

سال یازدهم، شماره های ۴۱ و ۴۲، پاییز و زمستان ۱۳۹۸

صاحب امتیاز: انجمن فناوری اطلاعات و ارتباطات ایران

مدیر مسؤول: دکتر مسعود شفیعی

سر دبیر: دکتر محمدشهرام معین

شمارگان: ۱۰۰۰

حروفچینی و صفحه آرایی: صابری

• اعضای اصلی:

دکتر مسعود شفیعی (رئیس انجمن)

دانشگاه صنعتی امیرکبیر

دکتر عباس محمدی

دانشگاه صنعتی امیرکبیر

مهندس ناصر یوسف پور

شرکت مخابرات همراه اول

دکتر سید علیرضا هاشمی گلپایگانی

دانشگاه صنعتی امیرکبیر

مهندس فرهاد رضایی

شرکت کوشا تجارت نوپدید

دکتر علیرضا اهری لاحق

شرکت داده پردازی ایران

دکتر محمد جوادیان

دانشگاه صنعتی کرمانشاه

مهندس سیدعلی صموتی

دانشگاه فنی حرفه ای شهید شمسی پور

دکتر علیرضا یاری

پژوهشگاه فناوری اطلاعات و ارتباطات

دکتر اکبر کاری دولت آبادی

دانشگاه شهید ستاری

دکتر جمشید شنبه زاده

دانشگاه خوارزمی

• اعضای علی البدل:

دکتر مهدی شجری (دانشگاه صنعتی امیرکبیر)

مهندس امیرارسلان داداشی جوردھی (شرکت ارتباطات سیار ایران)

بازرسان:

• بازرس اصلی:

دکتر محمود کمره (دانشگاه تهران)

• بازرس علی البدل:

دکتر امیر شهاب شاهمیری (دانشگاه جامع علمی کاربردی)

هیئت تحریریه:

فرخ حجت کاشانی (استاد)، دانشگاه علم و صنعت ایران

سید احمدرضا شرافت (استاد)، دانشگاه تربیت مدرس

مسعود شفیعی (استاد)، دانشگاه صنعتی امیرکبیر

محمدرضا عارف (استاد)، دانشگاه صنعتی شریف

کریم فائز (استاد)، دانشگاه صنعتی امیرکبیر

رضا فرجی دانا (استاد)، دانشگاه تهران

کمال محامدپور (استاد)، دانشگاه خواجه نصیرالدین طوسی

فرخ مروستی (استاد)، دانشگاه صنعتی شریف

سیداحمد معتمدی (استاد)، دانشگاه صنعتی امیرکبیر

روزبه معینی (استاد)، دانشگاه صنعتی امیرکبیر

کاظم اکبری (دانشیار)، دانشگاه صنعتی امیرکبیر

کامبیز بدیع (دانشیار)، مرکز تحقیقات مخابرات ایران

علی اکبر جلالی (دانشیار)، دانشگاه علم و صنعت ایران

حمیدرضا ربیعی (دانشیار)، دانشگاه صنعتی شریف

امیر البدوی (دانشیار)، دانشگاه تربیت مدرس

محمد تشنه لب (دانشیار)، دانشگاه خواجه نصیر طوسی

براساس نامه شماره ۳/۴۸۱۷ مورخ ۱۳۸۶/۶/۱۱ دفتر امور پژوهشی وزارت علوم، تحقیقات و فناوری و باینیه رأی کمیسیون بررسی نشریات علمی کشور در تاریخ ۱۳۸۶/۴/۲۳، درجه علمی - پژوهشی به این فصلنامه اعطا شده است

فهرست

فصلنامه علمی - پژوهشی فناوری اطلاعات و ارتباطات ایران
سال یازدهم، شماره‌های ۴۱ و ۴۲، پاییز و زمستان ۱۳۹۸

- ۱ ▪ روشی نوین برای کاهش تغییر روشی در تصاویر غیراخلاقی بر پایه شبکه عصبی عمیق فازی
ساسان کرمی‌زاده، ابوذر عرب‌سرخی
- ۱۳ ▪ ارائه مدل پذیرش فناوری اطلاعات در پروژه‌های تحقیقاتی دانشگاهی حوزه ICT بر اساس مدل پذیرش فناوری ITAIM
شهرام علی یاری، مسعود موحدی، سیروس کاظمیان
- ۳۳ ▪ ترکیب دوگانه سیستم استنتاج فازی با الگوریتم بهینه‌سازی ازدحام ذرات در پیش‌بینی قیمت سهام بورس اوراق بهادار تهران و مقایسه آن با مدل یادگیری عمیق
مجید عبدالرزاق نژاد، مهدی خرد
- ۵۷ ▪ ارائه روشی با استفاده از الگوریتم ژنتیک در تشخیص موضع افراد جامعه در رسانه‌های خبری و اجتماعی
مهدی سالخورده حقیقی، سیدمحمد ابراهیمی
- ۷۵ ▪ سیستم توصیه‌گر زمینه آگاه سفر با بهره‌گیری از اطلاعات عکس‌های برچسب‌دار جغرافیایی
رضوان محمدرضایی لرکی، رضا روانمهر، میلاد امرالهی
- ۹۷ ▪ نمودار قطبی نقاط با قطب متحرک
بهرام صادقی بی غم، فاطمه ربانی

روشی نوین برای کاهش تغییر روشنایی در تصاویر غیراخلاقی بر پایه شبکه عصبی عمیق فازی

*ساسان کرمی‌زاده **ابوذر عرب‌سرخی

*پژوهشگاه ارتباطات و فناوری اطلاعات (مرکز تحقیقات مخابرات ایران) تهران، ایران
**پژوهشگاه ارتباطات و فناوری اطلاعات (مرکز تحقیقات مخابرات ایران) تهران، ایران
تاریخ دریافت: ۱۳۹۹/۰۲/۰۲ تاریخ پذیرش: ۱۳۹۹/۰۶/۲۹

چکیده

در فضای اینترنت شناسایی عکس‌های غیراخلاقی امری ضروری در جهت حفاظت فیزیکی و ذهنی کودکان محسوب می‌شود؛ یکی از چالش‌های اصلی در تشخیص تصاویر غیر اخلاقی تغییرات روشنایی و رنگ پوست بدن است. به همین منظور در این مقاله روشی برای حل تغییرات روشنایی و بهبود تشخیص تصاویر غیراخلاقی ارائه شده است. در این مقاله از شبکه عصبی عمیق فازی برای بهبود روشنایی تصاویر غیراخلاقی استفاده شده است. در روش پیشنهادی از مدل یادگیری عمیق xception جهت تقسیم تصویر براساس شدت روشنایی به بخش‌های مختلف بهره گرفته شده است. تقسیم کردن تصویر به قسمت‌های مختلف باعث بهبود تغییرات روشنایی با حفظ جزئیات تصویر و نهایتاً شناسایی بهتر تصاویر غیراخلاقی شده است. به علاوه برای طبقه‌بندی رنگ پوست از ترکیب الگوریتم مبتنی بر Gaussian-KNN بهره گرفته شده است که روشی غیرپارامتری برای طبقه‌بندی‌ها و رگرسیون‌ها است؛ و در انتها از الگوریتم ماشین بردار پشتیبان برای طبقه‌بندی تصاویر استفاده شده است. به منظور پیاده‌سازی و ارزیابی روش پیشنهادی یک مجموعه شامل ۳۳۰۰۰ تصویر گردآوری شد، نتایج بدست آمده نشان می‌دهد که طرح پیشنهادی با دقت ۹۹/۷ درصد تصاویر غیراخلاقی را تشخیص می‌دهد.

کلید واژه: تغییرات روشنایی، طبقه‌بندی رنگ پوست، تغییر چهره، تصاویر اخلاقی و غیراخلاقی، Gaussian-KNN، الگوریتم ماشین بردار پشتیبان

۱- مقدمه

یکی از چالش‌های اصلی تشخیص تصاویر غیر اخلاقی مبتنی بر پردازش تصویر، تغییرات روشنایی و رنگ پوست بدن است. میزان روشنایی در تصاویر مختلف متفاوت است و تغییرات روشنایی یکی از مسائل حیاتی در تشخیص تصویر محسوب می‌شود. به عبارت دیگر شناسایی تصاویر در تغییرهای نور یک مسئله چالش‌انگیز است [۳].

امروزه حجم و تنوع اطلاعات در سطح اینترنت به شدت در حال افزایش است. گروه سنی کاربران این فضا نیز بسیار متنوع می‌باشد. دسترسی کودکان و گروه‌های سنی نوجوان به اطلاعات غیراخلاقی -اصطلاحاً تصاویر بزرگسالان- بسیار چالش‌انگیز است [۱]. تاکنون پژوهش‌های مختلفی برای شناسایی تصاویر غیراخلاقی مبتنی بر روش‌های پردازش تصویر و یادگیری ماشین ارائه شده است [۲].

طبقه‌بندی دودویی است که در طبقه‌بندی چند کلاس استفاده می‌شود. این روش دارای محاسباتی پیچیده است؛ در نتیجه اجرای آن به کندی انجام می‌شود [۱۲]. رویکردهای مبتنی بر قسمت بدن، چندین ویژگی معنایی مرتبط با پورنوگرافی همانند "سینه"، "شکم" و "پایین" را تعریف می‌کند، و برای این ویژگی‌ها ردیاب‌های مربوط به بخش بدن را آموزش می‌دهد [۱۳، ۱۴]. با توجه به تصویر از آشکارسازهای قسمت‌های بدن برای پوشش تصویر استفاده و نتایج بدست آمده جستجو می‌شود، آن‌گاه بوسیله یک بردار ویژگی معنایی طبقه‌بندی صورت می‌گیرد. با این حال، این آشکارسازهای قسمت بدن به دلیل پشتیبانی از وصله‌های کوچک و تغییرهای ظاهری زیاد در آموزش، دچار مشکل تشخیص می‌شوند. در این مقاله از ویژگی‌های تبدیل موجک مجتمع دوگانه برای بهبود استخراج ویژگی تصاویر استفاده شده است. در این تحقیق پیش از استفاده که یک تابع خطی است. این روش برای کاهش پارازیت در راستای حفظ مرز، بین اشکال استفاده می‌شود. باید به این موضوع توجه داشت که کاربرد توابع خطی در موقعیت‌های پیچیده مناسب نیست [۱۵، ۱۶]. یکی از عیب‌های این تحقیق، نرمال‌سازی نوردی بر روی کل تصویر است که در مقایسه با روش‌های عادی‌سازی دارای دقت کمتری است. در این تحقیق از نقشه خود سازماندهی جهت کاهش اثر تغییرات نور در تشخیص تصویر استفاده شده است [۱۰، ۱۷].

۳- روش پیشنهادی برای کاهش تغییرات روشنایی

روش پیشنهادی در این مقاله شامل دو بخش اصلی پیش‌پردازش و پردازش است. پیش‌پردازش شامل دو مرحله است. فرایند پیش‌پردازش مربوط به تغییر اندازه و عادی‌سازی نور است. تصویر ورودی با نگه داشتن نسبت طول و عرض به اندازه ۲۵۶ پیکسل اندازه‌گیری می‌شود، تا زمان محاسبات فرایند بعدی کاهش یابد. به منظور کاهش شدت نور در تصاویر، از یکسان‌سازی هیستوگرام استفاده شده است. پیکسل‌های پوست که دارای تغییرات نور زیادی هستند، نمی‌توانند در طبقه‌بندی به عنوان پوست عادی قرار گیرند. به عبارت دیگر، از هیستوگرام برای رسیدگی به تغییرات روشنایی تصاویر ورودی استفاده می‌شود. در بخش پیش‌پردازش برای متعادل‌سازی کنتراست از روش انحراف هیستوگرام

الگوریتم‌های تشخیص تصاویر بزرگسالان باید مقادیر قابل قبولی از تغییرهای نور بین تصاویر گالری و جستجو را در نظر بگیرند [۴، ۵]. تفاوت در میزان نور تصاویر به‌رغم شباهت‌های موجود در آنها به احتمال زیاد منجر به طبقه‌بندی اشتباه تصویرهای بزرگسالان خواهد شد [۶]. به منظور به حداقل رساندن تغییرهای نور در مجموعه داده‌های تصویر بزرگسالان از طرح عادی‌سازی روشنایی استفاده می‌شود. با این کار به دلیل وجود شباهت‌ها در ویژگی‌های تصویر با نورهای مختلف، احتمال طبقه‌بندی اشتباه کاهش می‌یابد [۷]. علاوه بر این، افراد با نژاد متفاوت می‌توانند برخلاف رنگ پوست‌های مختلف ویژگی‌های مشابهی داشته باشند. در نتیجه مقادیر عددی پیکسل‌ها نه تنها با تغییرات نور، بلکه به دلیل اختلاف رنگ‌ها نیز تغییر می‌کند. به دلیل این اختلاف مقادیر (رنگ/ شدت متغیر بین پیکسل)، برخی از محققان وابستگی به تغییرهای نور را اولویت‌بندی می‌کنند [۱]. هدف از این طبقه‌بندی بر روی بسیاری از اطلاعات بیومتریک نظیر پوست بدن، دست و پا، تعیین نوع پیکسل رنگ پوست است. طبقه‌بندی رنگ پوست باید به اندازه‌ای جامع باشد که رنگ‌های مختلف پوست همانند سفید، زرد و سیاه را پوشش دهد [۴]. فضاهای مختلف رنگی در تقسیم بندی پوست استفاده شده است. در برخی موارد، طبقه‌بندی رنگ پوست تنها با استفاده از رنگ‌آمیزی پیکسل انجام می‌شود، زیرا انتظار می‌رود در صورت از بین رفتن روشنایی پیکسل قابلیت شناسایی تغییرهای نور بیشتر شود [۸].

در این مقاله از روش ترکیبی فازی و شبکه عصبی کانولوشنی جهت بهبود شدت نور استفاده شده است. برای کاهش شدت تصویر غیراخلاقی از روش شبکه عصبی فازی برای تقسیم‌بندی عکس استفاده کرده‌ایم و بعد از تقسیم‌بندی عکس برای بهبود شدت نور از روش‌های هیستوگرام و کنتراست بر روی هر قسمت از عکس استفاده شده است.

۲- کارهای مرتبط

تاکنون مطالعات ارزشمندی در حوزه تشخیص برهنگی و تصاویر غیراخلاقی مبتنی بر روش‌های پردازش تصاویر انجام شده است. کارهای اولیه عمدتاً بر یافتن افراد برهنه در تصاویر براساس مدل ساختار انسانی تمرکز داشتند [۹-۱۱]. در این تحقیق از ماشین‌های بردار پشتیبان جهت طبقه‌بندی استفاده شده است. این الگوریتم معرف یک

۱-۳-تنظیم گر خودکار شدت نور

در این مرحله از طرح پیشنهادی، تلاش می‌شود که با استفاده از کوتاه کردن و کشش هیستوگرام کانال‌های رنگی تنظیم شوند، که شامل توزیع مجدد شدت در مقیاس مقادیر موجود است؛ به علاوه ممکن است کشش هیستوگرام در کل باند انجام شود به طوری که مقدار کمترین شدت صفر و بیشترین مقدار حداکثر مقدار باشد. در نتیجه، پیکسل‌های روشن به نورهایی روشن‌تر و پیکسل‌های تاریک به نورهای تاریک نزدیک‌تر می‌شوند که توزیع بهتری را تشکیل می‌دهد. معادلات هیستوگرام به یک فرایند غیرخطی اشاره دارد. تقسیم کانال‌ها و تنظیم هر کانال به‌طور جداگانه کاری نادرست است. علاوه بر این، متعادل کردن شدت رنگ تصویر به عنوان جزئی از مولفه‌های رنگ مطرح نیست. بنابراین برای یک تصویر رنگی ساده RGB، مقادیر هیستوگرام را نمی‌توان به طور مستقیم در کانال‌ها اعمال نمود. لازم است متعادل‌سازی شدت نور به نحوی انجام شود که بدون برهم زدن توازن رنگ تصویر بتوان این کار را عملیاتی نمود. بنابراین، اولین قدم این است که فضای رنگی تصویر را از RGB به یک فضای رنگی $YCbCr$ تبدیل کنیم که مقادیر شدت را از اجزای رنگ جدا می‌نماید [۱۳]. فضای رنگی $YCbCr$ به طور گسترده‌ای در پردازش تصویر استفاده می‌شود. در این قالب، اطلاعات توسط یک مؤلفه واحد ارائه می‌شود. اطلاعات Y و رنگ به عنوان دو جزء تفاوت رنگ، Cb و Cr ذخیره می‌شوند. جزء Cb تفاوت بین جزء آبی و یک مقدار مرجع است، و جزء Cr تفاوت بین جزء قرمز و یک مقدار مرجع است.

۲-۳-کاهش پارازیت

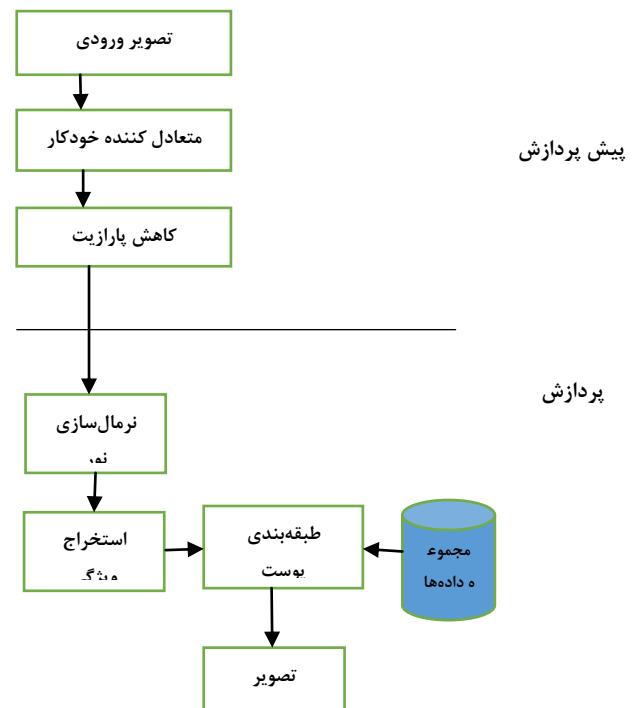
کاهش پارازیت^۲ یکی از مهمترین قسمت‌های پیش پردازش است. در این بخش، ضمن حفظ پارامترهای مهم، پارازیت حذف می‌شود. در این مقاله از سه نوع فیلتر وینر^۳، متوسط^۴ و میانگین^۵ جهت کاهش پارازیت استفاده شده است.

بعد از انجام تست و گرفتن نتایج مشخص شده که فیلتر وینر بهترین عملکرد را دارد. به این صورت که فرض اولیه آن

استفاده شده است. پس از متعادل کردن خودکار کنتراست، نیاز به کاهش پارازیت موجود در تصویر ورودی -همانند لکه‌های سیاه و سفید- است. برای این مرحله از فیلتر وینر استفاده شده است.

بعد از انجام پیش پردازش، تصویر به بخش پردازش ارسال می‌شود، که گام نخست در این بخش، نرمال‌سازی نور است. نرمال سازی نور از طریق شبکه عصبی عمیق فازی مبتنی بر تکنیک تقسیم‌بندی انجام می‌شود. هدف از تقسیم‌بندی تصویر حذف بهتر تغییرات نور و حفظ ویژگی‌های اصلی تصویر است. به عبارتی دیگر، این فعالیت نورهای زائد را حذف می‌کند. مرحله بعدی در طرح پیشنهادی، طبقه‌بندی رنگ پوست است که برای تشخیص تصویرهای متعدد در یک فریم تصویر استفاده می‌شود. به منظور استخراج ویژگی‌های عکس، از تکنیک استخراج ویژگی معادله تجزیه مقادیر منفرد^۱ بهره گرفته شده است. در نهایت ویژگی‌های استخراج شده برای طبقه‌بندی به الگوریتم ماشین بردار پشتیبان اعمال می‌شوند.

شکل ۱ بلوک دیاگرام روش پیشنهادی را به تصویر کشیده است. در ادامه جزئیات هر یک از بخش‌های روش پیشنهادی شرح داده می‌شود.



شکل ۱: بلوک دیاگرام روش پیشنهادی

¹ Singular Value Decomposition

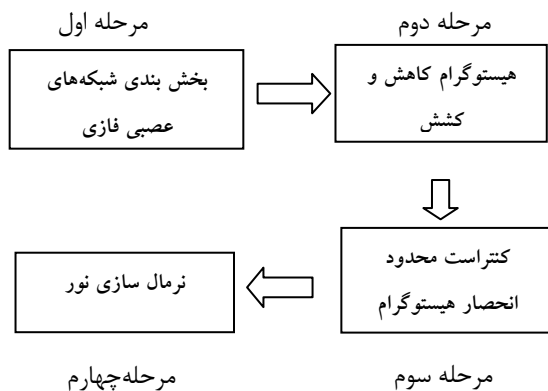
² Noise reduction

³ Wiener

⁴ mean

⁵ Meadin

مقاله، روش‌های نرمال‌سازی زیر مورد بررسی قرار می‌گیرند:
 (۱) تبدیل کسینوس گسسته، (۲) چندین مقیاس رتینکس^۷
 (۳) نرمال‌سازی مبتنی بر موجک، (۴) مقیاس تک رتینکس^۸
 و (۵) تفاوت گاوسی^۹. نرمال‌سازی نور تصویر بر اساس تکنیک تقسیم‌بندی شبکه فازی را می‌توان به چهار مرحله تقسیم نمود، در شکل ۲ نمای از مراحل نرمال‌سازی نور ارائه شده است.

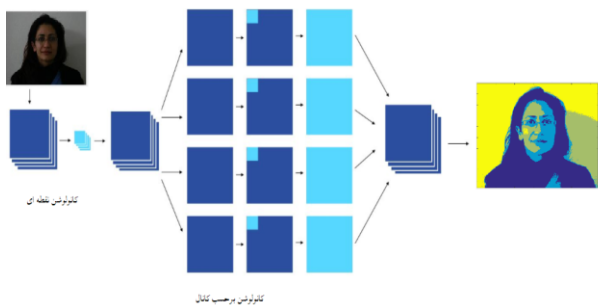


شکل ۲: مراحل نرمال سازی نور

۳-۳-۱- بخش‌بندی شبکه‌های عصبی عمیق فازی با

هدف بخش‌بندی تصویر

شبکه عصبی فازی عمیق در قالب شکل ۳ به تصویر کشیده شده‌است. همان‌طور که در شکل مشاهده می‌شود، این شبکه از پنج بخش تشکیل شده‌است. بخش اول تصویر ورودی است. در بخش دوم تابع عضویت فازی پردازش می‌شود. بخش سوم لایه فازی عصبی است که به وسیله شبکه عصبی عمیق xception ساخته شده است.



شکل ۳- مراحل تقسیم شبکه عصبی عمیق فازی

برای همه تصاویر، وجود لکه‌های کوچک پارازیت است. این الگوریتم برای از بین بردن پارازیت‌های ظریف بسیار کارآمد است. فیلتر وینر با استفاده از روش وینر سازگار با پیکسل و براساس آمار تخمین زده شده از اطراف هر پیکسل و با استفاده از تابع توزیع نرمال عمل می‌کند. معادله (۱) میانگین و واریانس اطراف هر پیکسل را برآورد می‌نماید.

$$G(u, v) = \frac{H^*(u, v)}{|H(u, v)|^2 + \frac{P_n(u, v)}{P_s(u, v)}} \quad (1)$$

در این رابطه،

$H(u, v)$: تابع کاهش

$H^*(u, v)$: تابع پیچیدگی تابع کاهش

$P_n(u, v)$: قدرت تراکم طیفی از پارازیت

$P_s(u, v)$: قدرت عددی طیفی تصویر ناسازگار.

اصطلاح P_n / P_s می‌تواند به عنوان نسبت سیگنال به پارازیت تفسیر شود. فیلترهای وینر توسط موارد زیر مشخص می‌شوند:

فرض: سیگنال‌های پارازیت معرف فرآیند خطی ثابت با ویژگی‌های طیفی شناخته شده هستند.

الزامات: فیلتر باید بطور فیزیکی قابل درک باشد.

معیارهای عملکرد: حداقل میانگین خطای مربع.

وینر به عنوان فیلتری برای برآورد یک فرآیند تصادفی مورد نظر یا هدف مطرح است و با فیلتر نمودن غیرخطی یک روند پارازیت اجرایی می‌شود. فیلتر وینر میانگین خطای مربع ما بین فرآیند تصادفی برآورد شده و روند مورد نظر را به حداقل می‌رساند.

۳-۳-۲- نرمال سازی نور

گام بعدی این فرآیند، نرمال‌سازی نور^۶ است. تنوع تابش می‌تواند باعث کاهش عملکرد سیستم‌های تشخیص تصویر شوند. بنابراین برای حل این مسئله، نرمال‌سازی نور تصویر را براساس روش تقسیم‌بندی شبکه عصبی فازی انجام می‌دهیم. تکنیک‌های عادی‌سازی نور پردازشی، برخی تغییرها را در تصویر ایجاد می‌کنند. این تغییرها سبب ایجاد یک فرم استاندارد بدون هیچ‌گونه تغییر در نور و حذف ضریب‌های تبدیل کسینوس گسسته در منطق لگاریتم می‌شوند. این کار معادل جبران تغییرات نور است. در این

⁷ Multi Scale Retinex

⁸ Single Scale Retinex

⁹ Difference of Gaussian

⁶ Illumination Normalization

کوچک از حجم ورودی استفاده می‌شود. در طول هر گام، یک فیلتر در عرض و ارتفاع عکس ورودی، نقطه‌ها بین ورودی‌های فیلتر و نقشه فعال ۲ بعدی ایجاد شده و از آن فیلتر را محاسبه می‌کند. در نتیجه، شبکه فیلتر را یاد می‌گیرد که وقتی یک نوع خاص از ویژگی‌ها را در یک موقعیت مکانی از ورودی مشاهده می‌کند، فعال می‌شود. قسمت پنجم - لایه پیش‌بینی V : در این مرحله تابع سافت مکس^{۱۰} اعمال می‌شود که برای طبقه‌بندی داده‌ها به کلاس‌های مختلف کاربرد دارد. این تابع به عنوان ورودی و برچسب مربوط به آن به معنی تبدیل تغذیه شبکه کاملاً متقارن فازی^{۱۱} از لایه ورودی به آخرین لایه کار اعمال می‌نماید. در ادامه، تابع سافت‌مکس می‌تواند به‌عنوان لایه خروجی با ورودی محاسبه شده و توسط معادله (۵) انجام شود.

$$y_{ic} = p(y_i | f_i) = \frac{e^{w_c \pi \Theta(f_i) + b_c}}{\sum_c e^{w_c \pi \Theta(f_i) + b_c}} \quad (5)$$

در این رابطه w_c و b_c نشان دهنده ضریب رگرسیون و C^{th} معرف کلاس $[y_{i1}, \dots, y_{ik}]$ است، که برچسب‌های پیش‌بینی شده شبکه عصبی با کلاس k است. آن‌گاه، از بین رفتن میانگین مربع شبکه عصبی با اتصال کامل می‌تواند بر روی بیش از m نمونه‌های آموزش تعریف شود که این موضوع در قالب فرمول (۶) نمایش داده شده‌است.

$$C = \frac{1}{m} \sum_i^m \|y_i - y_i\|_2^2 \quad (6)$$

۳-۳-۲ - هیستوگرام کاهش و کشش

مرحله دوم شامل کاهش و کشش هیستوگرام است. برش و کشش هیستوگرام یک تکنیک تکرار هیستوگرام محسوب می‌شود. این روش شامل توزیع شدت در مقیاس مقادیر موجود است. این مقدار برای کشش هیستوگرام روی تمام تصویر است، به طوری که مقدار کمترین شدت صفر باشد و

قسمت اول - لایه تابع عضویت فازی: هر نورون در لایه ورودی با توابع عضو چندگانه در ارتباط است که برچسب‌هایی را برای هر متغیر ورودی اختصاص می‌دهد. در این‌جا، متغیر ورودی یک بعد از بردار ورودی است. محاسبه درجه عضویت یک نورون ورودی به یک مجموعه فازی خاص توسط تابع عضویت معادله ۲ انجام می‌شود. در این لایه، نورون فازی نشان دهنده ورودی به عنوان درجه فازی است که در معادله ۲ نشان داده شده است.

$$o_i^{(l)} = u_i(a_k^{(l)}) = e^{-(a_k^{(l)} - \mu_i)^2 / \sigma_i^2}, \forall_i \quad (2)$$

در این معادله (۲) تابع عضویت گوسی با میانگین μ و واریانس σ^2 مورد استفاده قرار می‌گیرد.

$$o_i^{(l)} = \prod_j o_j^{(l-1)}, \forall j \in \Omega_i \quad (3)$$

که در معادله (۳)، Ω_i مجموعه‌ای از نورن‌ها را در لایه $(l-1)^{th}$ را تعریف می‌نماید که به i متصل می‌شوند. در این مورد، لایه $(l-1)^{th}$ ورودی است. خروجی‌های این بخش درجه‌های فازی هستند.

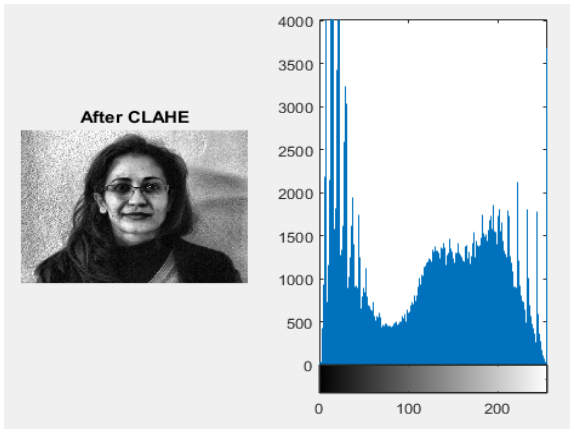
قسمت دوم - شبکه عصبی فازی: در این قسمت از مفهوم یادگیری عصبی استفاده می‌شود تا ورودی را به‌عنوان برخی از نمایندگی‌های سطح بالا تغییر دهد. در این قسمت لایه‌ها به طور کامل متصل می‌شوند، به این معنی که هر پیکسل در لایه (۱) به تمام پیکسل‌های لایه $(1-1)^{th}$ با پارامترهای $\theta(l) = \{w(l), b(l)\}$ متصل می‌شوند. این مثال در معادله (۴) ارائه شده‌است.

$$o_i^{(l)} = \frac{1}{1 + e^{-a_i^{(l)}}}, a_i^{(l)} = w_i^{(l)} o^{(l-1)} + b_i^{(l)} \quad (4)$$

قسمت سوم - لایه کانولوشنال: تصویر از طریق یک ستون از لایه‌های کانولوشن و پس از نمایش عصبی فازی عبور داده می‌شود. در این لایه از فیلتر 8×8 استفاده می‌شود. گام کانولوشن، ۱ پیکسل انتخاب شده‌است. این لایه همانند فیلتر به تصویر اعمال می‌شود. این فیلتر ویژگی‌هایی را از سطح پایین به سطح بالا می‌برد. توابع فعال‌سازی مختلف در لایه‌های متفاوت برای استخراج ویژگی‌های متنوعی استفاده می‌شوند. در لایه کانولوشن محاسبه ضرب نقطه‌ای بین وزن‌های نورون و یک ناحیه

¹⁰Soft-max

¹¹ fuzzy fully convolutional network



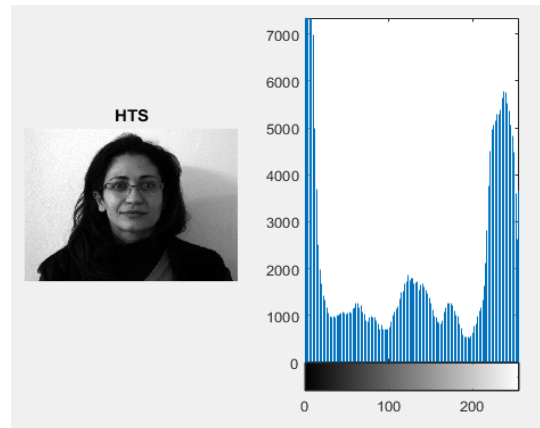
شکل ۵: بعد از هیستوگرام سازگار محدود کنتراست

۴-۳-۳- نرمال سازی

روش نرمال سازی تبدیل کسینوس گسسته جهت حذف اثرهای تنوع روشنایی به کار برده شده است. در این مقاله، تلاش بر این بود تا یک روش نرمال سازی نور را برای حذف تغییرات نور در حالی که خصیصه های اصلی تصویر بدون تغییر بماند را بررسی نماییم. ایده کلیدی رویکرد پیشنهادی این است که تغییرات نور را می توان با کاهش ضرایب تبدیل کسینوس تقریباً فرکانس پایین در دامنه لگاریتم تبدیل کسینوس گسسته کاهش داد. تکنیک نرمال سازی مبتنی بر تبدیل کسینوس گسسته یک مقدار ثابت از ضرایب تبدیل کسینوس گسسته را مطابق با فرکانس های پایین تعیین می کند، که با تغییر دادن غیرمستقیم روشنایی دنبال می شود. علاوه بر این، ضرایب تبدیل کسینوس گسسته با استفاده از الگوی زیگزاگ انتخاب می شوند.

تبدیل کسینوس گسسته به عنوان روشی قدرتمند در کاربردهای تشخیص تصویر و برنامه نویسی تصویر شناخته می شود. الگوریتم نرمال سازی نور تصویر براساس تکنیک تقسیم شبکه عصبی فازی عمل می نماید. تکنیک های نرمال سازی نورپردازی برخی تغییرهای در تصویر را برای ایجاد یک فرم استاندارد بدون هیچ گونه تغییرهای نورانی و حذف ضریب تبدیل کسینوس گسسته با فرکانس پایین در منطق لگاریتم انجام می دهند. پیاده سازی تبدیل کسینوس گسسته نیاز به دو حلقه تو در تو دارد. دستورهای رایانشی $M \times N$ بدان معنی است که افزایش M و N به طور قابل توجهی متناسب با زمان پردازش است. با توجه به این محدودیت ها، محققان تصویر ورودی را جهت غلبه بر این محدودیت ها تغییر می دهند. با این وجود، با استفاده از تبدیل کسینوس گسسته در قسمت های مختلف تصویر، زمان محاسبه کمتری را لازم دارد. تعداد زیادی از ضرایب

حداکثر مقدار آن بالاترین باشد. در نتیجه، پیکسل های نور حتی سبک تر و تاریک تر به سیاه نزدیک تر می شوند که توزیع بهتری را شکل می دهند. عمل کوتاه کردن، قبل از کشش انجام می شود تا قطر خاصی از انتهای پایین و بالای هیستوگرام را قطع کند. این کار اثر جعلی را از چندین پیکسل روشن یا تاریک بیش از بقیه حذف می نماید. به طور مستقیم کاهش و کشش هیستوگرام در هر بخش تقسیم شده موجب ایجاد یک مشکل خاص می شود. در این مورد، برخی از قطعه های محلی ممکن است به دهانه های مختلف تقسیم شوند و شدت این دهانه ها برای کشیدن کل تصویر گسترش یافته باشد. در نتیجه، این قطعات الگوهای منفرد فضایی را با شدت در جهت معکوس در مقایسه با تصویر اصلی نمایش می دهند که در شکل ۴ به تصویر کشیده شده است.



شکل ۴: تغییرات تصویر بعد از هیستوگرام کاهش و کشش

۳-۳-۳- هیستوگرام سازگار محدود کنتراست

گام دوم کنتراست محدود انحصار هیستوگرام سازگار^{۱۲} است. تکنیک های تقویت کنتراست در پردازش تصویر به طور گسترده ای مورد استفاده قرار می گیرد. انحصار هیستوگرام سازگار محدود کنتراست معرف یک نسخه بهبود یافته از انعطاف پذیری هیستوگرام سازگار است. این رویکرد، مناطق کوچک تصویر را با استفاده از خصیصه های هیستوگرام برای هر پیکسل به صورت جداگانه پردازش می کند. کنتراست، به ویژه در مناطق شدت همگن، می تواند برای جلوگیری از تقویت پارازیت محدود شود. که در شکل ۵ به تصویر کشیده شده است.

¹² Contrast Limited Adaptive Histogram Equalization

$$\delta(i, j) = \begin{cases} 1 & i = j \\ 0 & otherwise \end{cases} \quad (8)$$

بنابراین

$$class(x) = \begin{cases} \arg \max \sum_{i=1}^k f(x, NN_i(x)) & i = j \\ 0 & otherwise \end{cases} \quad (9)$$

توابع کرنل گاوسی به دو دلیل انتخاب شده‌است. اولین دلیل این است که، تابع گاوسی صاف است و از این رو تابع چگالی برآورد شده نیز به آرامی تغییر می‌کند. دومین دلیل، در فرم خاصی از توابع گاوسی که در آن تابع به طور شعاعی متقارن است، تابع می‌تواند به‌طور کامل توسط پارامتر واریانس مشخص شود که در معادله (۱۰) تابع گاوسی را نشان داده شده است.

$$f(x, \mu = y, \Sigma) = \frac{1}{\sqrt{(2\pi)^2 |\Sigma|}} \exp\left(-\frac{1}{2}(x - \mu)^T \Sigma^{-1}(x - \mu)\right) \quad (10)$$

که در آن X نمونه ورودی است و Y یک نمونه از ویژگی است؛ آن‌ها را بردارهای RGB می‌سازند، و μ مقدار میانگین است. طبقه‌بندی نزدیکتری همسایگی K پیشنهادی نسبت به طبقه اقلیت بسیار حساس است. ایده اصلی شناسایی نمونه‌های کلاس اقلیت در داده‌های آموزشی و سپس تعمیم آن‌ها به عملکرد گاوسی به‌عنوان مفهومی برای طبقه اقلیت عمل می‌نماید. این رویکرد مبتنی بر توسعه مرز تصمیم‌گیری برای طبقه اقلیت بنا نهاده شده است. معادله (۱۱) Gaussian-based K-NN weight را نشان می‌دهد:

$$KNN = class(x)$$

$$class(x) = \arg \max \sum_{i=1}^k f(x, NN_i(x)) \delta(class(NN_i(x)), j)$$

$$Gaussian = f(x, \mu = y, \Sigma)$$

$$f(x, \mu = y, \Sigma) = \frac{1}{\sqrt{(2\pi)^2 |\Sigma|}} \exp\left(-\frac{1}{2}(x - \mu)^T \Sigma^{-1}(x - \mu)\right)$$

$$Gaussian \text{ based K-NN} = K-NN + Gaussian$$

$$Class(x)' = KNN + Gaussian$$

$$= class(x) + f(x, \mu = y, \Sigma) \delta(class(NN_i(x)), j)$$

$$class(x)' =$$

$$\arg \max \sum_{i=1}^k \frac{1}{\sqrt{(2\pi)^2 |\Sigma|}} \exp\left(-\frac{1}{2}(x - \mu)^T \Sigma^{-1}(x - \mu)\right) \delta(class(NN_i(x)), j)$$

$$(11)$$

تبدیل کسینوس گسسته قرار داده شده در یک بردار، باید در مرحله پیاده‌سازی مرحله کمی‌سازی به صفر برسد. توالی زیگزاگ به‌عنوان بهترین روش در این فضا عمل می‌کند. تمام این اجزاء می‌توانند پس از ضرایب تبدیل کسینوس گسسته در شکل بردار از طریق توالی زیگزاگ به صفر برسند.

۴- طبقه‌بندی پوست

طبقه‌بندی پوست که یک ترکیب پیشرفته الگوریتم مبتنی بر Gaussian-KNN است، یک روش غیرپارامتری برای طبقه‌بندی‌ها و رگرسیون‌ها است. الگوریتم K-NN شامل برآورد شباهت ما بین نمونه ورودی و K نزدیکترین نمونه موجود در فضای برجسته است. هر ویژگی یک برجسته کلاس دارد؛ بنابراین، الگوریتم به شمارش تعداد مواردی می‌پردازد که متعلق به هر کلاس است. نتیجه طبقه‌بندی کلاس متاثر از حداکثر تعداد موارد اختصاص داده شده است. الگوریتم وزنی K-NN از یک روش مشابه که در فرمول (۷) نشان داده شده است برای تخمین ارزش هر رای برای بهبود عملکرد طبقه‌بندی به روش زیر استفاده می‌کند.

$$class(x) = \arg \max \sum_{i=1}^k f(x, NN_i(x)) \quad (7)$$

$$\delta(class(NN_i(x)), j)$$

در این رابطه موارد زیر را می‌توان مد نظر قرار داد:

- کلاس (X) : کلاس حقیقی نمونه X را باز می‌گرداند و یک کلاس پیش‌بینی را نشان می‌دهد.

- K: تعداد نزدیکترین همسایگان استفاده شده در-K-NN را نشان می‌دهد.

- NN_i(x): نزدیکترین همسایه ام با توجه به یک نمونه بدون برجسته X از 1, 2, ..., k را نشان می‌دهد.

- F(x,y): یک تابع وزنی برای رأی دهی است که میزان تاثیر نزدیکترین همسایه Y را به X تعیین می‌نماید.

- $\delta(class(x), class(y))$: به این معنی است که k-

NN وزن یکسانی را برای هر کدام از نزدیکترین همسایگان برای رأی‌گیری اکثریت تعیین می‌کند. بنابراین

f(x,y) مقدار هر کدام از ویژگی‌ها را تخمین می‌زند و تابع (j, δ i), تابع دلتا Kronecker است که به‌صورت زیر در

معادله‌های ۸ و ۹ تعریف می‌شود،

صعودی مرتب می‌کند، علاوه بر این K نزدیکترین فاصله را انتخاب می‌نماید. سیستم امتیازدهی مبتنی بر گاوسی، مقدار گاوسی را براساس فاصله نمونه از بردار محاسبه می‌کند. در واقع، این طرح همه امتیازها برای هر برچسب کلاس را به طور جداگانه محاسبه می‌نماید. برچسب کلاسی که با بالاترین نمره انتخاب می‌شود تحت عنوان برچسب نمونه شناخته می‌شود. نتیجه این مرحله تصویر طبقه‌بندی شده است که یکی از گروه‌های پوستی سفید، سیاه یا زرد را نشان می‌دهد.

۵- استخراج ویژگی‌ها

مرحله سوم پردازش تصویر به استخراج ویژگی اشاره دارد. در این مرحله از تجزیه مقادیر منفرد برای استخراج ویژگی‌هایی استفاده می‌شود که مبنای تشخیص طیف وسیعی از تصاویر هستند. این الگوریتم در تشخیص تصاویر غیراخلاقی مؤثرتر است، زیرا ویژگی ارزش واحد نسبت به تغییرات رنگ خاکستری در طی فرآیند شناخت حساس نیست و می‌تواند بر برخی از عوامل دیگر همانند تغییر نور غلبه کند. دو دیدگاه عمده پیرامون ویژگی‌های این الگوریتم وجود دارد. اولین دیدگاه این است که از تجزیه مقادیر منفرد می‌تواند به‌عنوان یک روش تبدیل متغیرهای هم‌بسته به مجموعه داده‌های اصلی استفاده نمود، و روابط مختلف را در میان اقلام داده‌های اصلی مؤثرتر نشان داد. دومین دیدگاه این است که از این روش برای شناسایی و ترتیب ابعاد داده‌ها که در آن نقاط داده‌ها بیشترین تغییر را نشان می‌دهند استفاده شود. استفاده از این سازوکار برای تعیین بهترین تقریب نقاط داده‌های اصلی با استفاده از ابعاد کم‌تر امکان‌پذیر است. از این‌رو تجزیه مقادیر منفرد به‌عنوان روشی برای کاهش داده‌ها مطرح است.

معادله تجزیه مقادیر منفرد برای یک $(m \times n)$ ماتریس منفرد $A = USV^T$ است، در جایی که U یک ماتریس قائم $(m \times m)$ است، V یک $(n \times n)$ است. ماتریس قائم S یک ماتریس مورب $(m \times n)$ است، که حاوی مقادیر منفرد A است که به ترتیب مرتبه کاهش می‌یابد. یک بردار در یک ماتریس متعامد می‌تواند به عنوان ترکیبی خطی از بردارهای دیگر بیان شود. بردارهای موجود در این فضا نیز به طور متقابل مستقل هستند و اکنون می‌توان یک راه حل برای X محاسبه کرد.

یک آستانه برای شناسایی داده‌های پوست در سیستم رنگ YC_bC_r استفاده می‌شود [۱۳]. آستانه ای در آن بیشتر افراد با رنگ پوست مختلف در تصویر می‌توانند با استفاده از تکنیک‌های هیستوگرام تقسیم شوند. بنابراین نیاز است که تصویر را به این سیستم تغییر داده و آن‌گاه تمام پیکسل‌هایی را که از یک محدوده داده از پیش تعیین شده خارج شده‌اند را حذف نمود. زمانی که میانگین پیکسل‌های استخراج شده از بردار R, G, B محاسبه می‌شود؛ که در معادله‌های ۱۲، ۱۳، ۱۴ بیان شده است.

$$Y = (0.299 \times R) + (0.587 \times G) + (0.114 \times B) \quad (12)$$

$$C_b = (B - Y) \times 0.564 + 135.0 \quad (13)$$

$$C_r = (R - Y) \times 0.713 + 128.0 \quad (14)$$

تصویر تبدیل شده $YC_b C_r$ به مرحله بعدی منتقل می‌شود، یعنی استخراج پیکسل‌های پوستی، پیکسل‌های پوست با استفاده از معادله (۱۵) زیر استخراج می‌شوند.

$$[85 < C_b < 135]$$

$$[135 < C_r < 180] \quad (15)$$

هر منطقه از تغییرات نور تصاویر رنگی بر اساس تغییرات برآورد شده توسط یک الگوریتم ثبات رنگ مبتنی بر لبه محلی^{۱۳} اصلاح می‌شود. سپس تصاویر تصحیح شده به یک فضای رنگی YC_bC_r تبدیل می‌شوند، و قسمت رنگ پوست و قسمت‌های رنگی غیر پوست از هم جدا شده‌اند. علاوه بر این، مورفولوژی^{۱۴} نیز در این فرآیند گنجانده شده است تا طبقه‌بندی پوست مثبت کاذب را حذف کند. ثابت شده است که این مدل عملکرد بهتری در شناخت تصاویر غیر اخلاقی ارائه می‌دهد [۱۷]. در این تبدیل، تطبیق هیستوگرام در سطح شدت Y انجام می‌شود و در نهایت می‌توان تصویر $YC_b C_r$ حاصل را به RGB برگرداند بردار رنگ به طبقه‌بندی K -NN منتقل خواهد شد. با توجه به اینکه K -NN نیاز به آموزش ندارد، وزن K -NN به‌عنوان طبقه‌بندی بسیار مناسب برای مجموعه داده‌های جزئی مورد توجه قرار می‌گیرد. طبقه‌بندی‌کننده K -NN فاصله اقلیدسی بردار رنگ و تمام نمونه‌های مجموعه داده را $k = 50$ محاسبه می‌کند. در ادامه فاصله‌ها را به‌صورت

¹³ Local edge-based

¹⁴ Morphological

و (۱۸) برای شاخص‌های عملکرد استفاده شد. در این مقاله ۷۰ درصد از مجموعه داده‌ها را برای آموزش داده‌ها و ۳۰ درصد از داده‌ها را برای تست استفاده شده است. به علاوه داشتن یک مجموعه داده جامع، دستیابی به سطح بالاتری از آموزش را فراهم می‌کند یکی از عوامل موفقیت است. FCN-AlexNet که از قبل روی مجموعه داده‌های SYNTHIA آموزش داده شده است که حاوی بیش از ۲۰۰،۰۰۰ تصویر است. پس از آن، یادگیری انتقالی در تقسیم‌بندی متراکم از نور با استفاده از تقسیم داده‌ها و طرح اولیه وزن MSRA انجام شده است. با توجه به تغییرات روشی در تصاویر غیراخلاقی که دارای پوست تیره بوده اند خطای کاذب بیشتر بوده است.

$$s = \begin{bmatrix} s_1 & 0 \\ 0 & s_m \end{bmatrix}$$

عناصر مورب ماتریس S به ترتیب نزولی $s_1 \geq s_2 \geq \dots \geq s_m \geq 0$ ذکر شده و مقادیر منفرد A نامیده می‌شوند.

۶- تشخیص تصویر

در این قسمت از الگوریتم ماشین بردار پشتیبان برای طبقه‌بندی عکس‌ها استفاده شده است. کلاس‌بندی ماشین بردار پشتیبان به وسیله فرمول (۱۶) بدست می‌آید.

$$g(x) = \sum_{i=1}^{L_S} aidik(xi, x)a0 \quad (16)$$

که در آن K تابع کرنل و xi معرف بردار پشتیبان، Ls معرف تعداد بردارهای پشتیبان و di معرف شماره کلاس Xi مربوطه است و در نهایت ai ها اعداد ثابتی هستند که در مرحله آموزشی به دست می‌آیند. با توجه به تعریف، بردارهای پشتیبان المان‌هایی از داده‌های آموزشی هستند که دقیقاً بر روی و یا داخل مرزهای تصمیم‌گیری طبقه‌بندی قرار می‌گیرند. به عبارتی دیگر این بردارها شامل نمونه‌هایی هستند که دسته‌بندی آن‌ها از بقیه نمونه‌ها مشکل‌تر است. ویژگی اصلی این روش توانایی بالا در استفاده از نمونه‌های تعلیمی کمتر و رسیدن به دقت بالاتر است.

به عبارتی دیگر این بردارها شامل نمونه‌هایی هستند که دسته‌بندی آن‌ها از بقیه نمونه‌ها مشکل‌تر است. باید به این نکته توجه شود که در یک مسئله ماشین بردار پشتیبان تنها کسر کوچکی از داده‌های آموزشی به عنوان بردارهای پشتیبان تعیین می‌گردند. توابع کرنل نقش مرکزی و اصلی را در نگاشت کردن منحنی بردار ورودی به فضای ویژگی با ابعاد بالاتر، برای به دست آوردن تفکیک‌پذیری بیشتر بازی می‌کنند، در این مقاله از کرنل چند جمله‌ای استفاده شده است.

۷- نتایج بدست آمده

به منظور پیاده سازی و ارزیابی روش پیشنهادی یک پایگاه داده از طریق جستجو در اینترنت مشتمل بر ۳۳۰۰۰ تصاویر اخلاقی و غیر اخلاقی ایجاد شد. دقت یک هدف مناسب در هر سناریوی یادگیری ماشینی محسوب می‌شود؛ زمانی که داده‌های یک کلاس در مقادیر بسیار بیشتری از داده‌های کلاس دیگر وجود دارند. معادله‌های (۱۶) و (۱۷)

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (16)$$

$$FNR = \frac{FN}{FN + TP} \quad (17)$$

$$FPR = \frac{FP}{FP + TN} \quad (18)$$

۷-۱- مقایسه مدل های یادگیری عمیق

در این قسمت عملکرد مدل‌های یادگیری عمیق به کار گرفته شده، مورد ارزیابی قرار گرفته است. عملکرد این مدل‌ها بر حسب پارامترهای ارزیابی ارائه شده است.

جدول شماره ۳: مقایسه مدل‌های یادگیری عمیق

precision	recall	زمان	حجم	معماری	کلاس
۰,۹۳	۰,۹۱	۰,۹۴	۳۰ms	۱۳۸ mi	VGG-16
۰,۹۴	۰,۹۴	۰,۹۹۰	۱۵ms	۲۵ mi	Resnet50v2
۰,۹۶	۰,۹۶	۰,۹۹۷	۲۴ms	۲۲ mi	xception

۳×۳ هست که پشت سر هم قرار گرفته‌اند. سپس، یک لایه ماکس پولینگ ۲×۲ با پرش به اندازه ۲ قرار گرفته است. این لایه ماکس پولینگ علاوه بر نمونه برداری، وظیفه کاهش بعد ویژگی‌ها به نصف را هم دارد. در ادامه، دو لایه کانولوشنی دیگر با ۱۲۸ فیلتر ۳×۳ و یک لایه ماکس پولینگ ۲×۲ و پرش ۲ قرار گرفته‌اند. به طور مشابه، سه لایه کانولوشنی با ۲۵۶ فیلتر ۳×۳ و یک لایه ماکس پولینگ ۲×۲ با پرش ۲ قرار گرفته‌اند ۳ لایه کانولوشنی با ۵۱۲ فیلتر ۳×۳ و یک لایه ماکس پولینگ ادامه این شبکه هست که البته دو بار تکرار می‌شود. در نهایت، ویژگی‌ها تبدیل به یک بردار ویژگی می‌شوند تا در اختیار لایه‌های نورونی تمام اتصال قرار گیرند. دو لایه نورونی به ابعاد ۴۰۹۶ پشت سر هم قرار گرفته‌اند. در نهایت، یک لایه نورونی به ابعاد ۱۰۰۰ که متناظر با تعداد کلاس‌های کاربرد ما هست، در نظر گرفته شده است لایه خروجی شامل ۱۰۰۰ نورون است. در تمامی لایه‌های کانولوشنی و لایه‌های نورونی از تابع فعال‌ساز بنام RELU استفاده شده است.

۴-۷- مقایسه روش انجام شده

در این بخش عملکرد روش پیشنهادی با روش سنتی شناسایی تصاویر غیراخلاقی و روش‌های نوین مبتنی بر تکنیک‌های یادگیری عمیق مقایسه شده است. مقایسه نتایج بین سه روش تشخیص عکس‌های غیراخلاقی با روش پیشنهادی پیرامون محاسبه دقت در قالب جدول ۴ ارائه شده است.

جدول ۴: مقایسه بین روش‌های موجود با روش

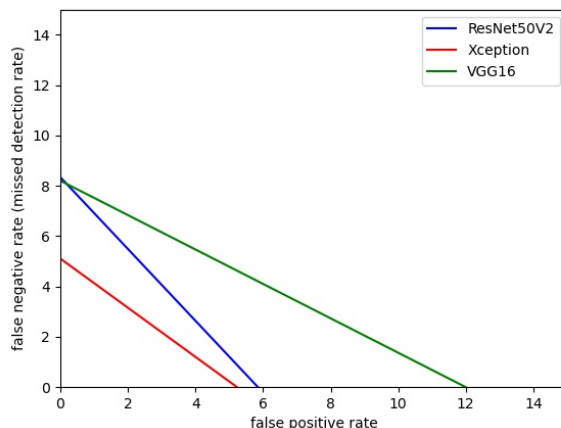
پیشنهادی	نتیجه غیراخلاقی	نتیجه اخلاقی	نتایج مجموع
Region-of-Interest Method[18]	۶۸/۳۰	۶۱/۴۱	۶۴/۸۵
Bag-of-Feature method[19]	۷۹/۷۹	۷۱/۸۷	۷۵/۸۳
Deep Fused CNN[20]	۹۲/۶۷	۸۹/۲۲	۹۰/۹۴
Pornographic Image Recognition Via Weighted Multiple Instance Learning[21]	۹۸/۱۸	۹۸/۵۱	۹۸/۳۵
روش ارایه شده	۹۹/۶۰	۹۹/۸۵	۹۹/۷۰

همانطور که در جدول شماره ۳ نشان داده شده است مدل xception بهترین دقت را نسبت به دو مدل دیگر ارائه نموده است، ولی از نظر سرعت مدل resnet50v2 بهترین نتیجه نسبت به دو مدل xception و VGG-16 به دست آورده است و چون شاخص دقت مهمتر بود مدل xception انتخاب شد.

۷-۲- منحنی خطای تشخیص

در منحنی خطای تشخیص خطا^{۱۵} در حالت ایده آل، احتمال بروز هر دو نوع خطا صفر خواهد بود. در عمل، این دو قابل معامله هستند: احتمال پذیرش کاذب می‌تواند با هزینه افزایش رد کاذب کاهش یابد.

منحنی‌های خطای تشخیص نشان می‌دهد چگونه ممکن است نقطه عملیاتی برای تغییر مبادلات بین FA و FR تغییر کند، این منحنی در شکل ۶ نشان داده شده است. بهترین حالت این است که هر دو خطای FA و FR در هر نقطه عملیاتی کم باشند. به عبارت دیگر، منحنی خطای تشخیص باید نزدیک یا در محورهای FA / FR باشد.



شکل ۶: منحنی خطای تشخیص خطا

منحنی خطای تشخیص برای سه مدل مختلف یادگیری عمیق نشان داده شده است. طبقه بندی بهتر دارای منحنی پایین تر است که به محورها نزدیک تر است که توسط خط مورب نشان داده شده است. که در این سه مدل، مدل xception بهترین نتیجه را ارائه نموده است.

۷-۳ معماری VGG-16

شبکه VGG-16 شامل ۱۶ لایه کانولوشنی یا ۱۶ لایه پارامتری است. شبکه VGG-16، شامل دو لایه کانولوشنی با ۶۴ فیلتر

¹⁵ Detection error tradeoff (DET)

خودکار شدت نور و کاهش پارزیت است. در قسمت پردازش برای نرمال‌سازی شدت نور از معماری exception استفاده شده است. در ادامه از طبقه‌بندی پوست استفاده شده تا به تشخیص تصویر کمک کند در این مرحله از الگوریتم Gaussian-KNN استفاده شده است. به علاوه در این مقاله از ۳۳۰۰۰ تصاویر اخلاقی و غیر اخلاقی استفاده شده که تصاویر از اینترنت جمع آوری شده‌است. نتایج بدست آمده نشان می‌دهد که روش پیشنهادی ۱/۳ درصد بهتر از روش‌های موجود در این حوزه عمل نموده‌است.

بررسی چالش‌های تصاویر غیراخلاقی به ویژه تصاویر سوء استفاده از کودکان به عنوان کارهای آینده پیشنهاد می‌شود؛ می‌توان با جمع آوری تصاویر سو استفاده از کودکان و با استفاده از مدل مختلف یادگیری عمیق ویژگی‌هایی بهتری رو استخراج نمود که باعث بهبود دقت در تشخیص این تصاویر شود.

۹- تشکر و قدردانی

این تحقیق به وسیله مرکز تحقیقات مخابرات ایران حمایت شده‌است. این روش به عنوان بخشی از فعالیت‌های تحقیقاتی طرح صیانت (فرهنگی - اجتماعی) با بهره‌گیری از سامانه‌های هوشمند در پژوهشگاه ارتباطات و فناوری اطلاعات طراحی و پیاده‌سازی شده‌است.

روش سنتی یک روش بازیابی شده است که با استفاده از تشخیص پوست، رنگ، شکل و ویژگی‌های پوست را بر روی ROI استخراج می‌کند [۱۸]. ویژگی BoF رویکرد مبتنی بر توصیف Hue-SIFT و نسخه بهبود یافته آن بر اساس مدل توجه بصری است [۲۱]. در هر دو روش عملکرد خوبی برای مجموعه داده‌های کوچک است [۱۹]. روش عمیق مبتنی بر CNN برای تشخیص تصویر غیراخلاقی وجود دارد [۲۰]. برای انجام یک مقایسه جامع با مقدمات عمیق مبتنی بر CNN، الگوریتم‌های توصیف شده در [۲۰]، و سه تغییر مبتنی بر CNN را با داده‌های آموزش پیاده‌سازی کردیم.

روش ارائه شده در این مقاله بر روی انواع تصاویر با کیفیت‌های مختلف کار می‌کند، به طور خاص، روش ارائه شده نسبت به سایر روش‌ها در تصاویر غیراخلاقی که بصورت غیرحرفه‌ای گرفته شده‌اند، مزایای بارزی را نشان می‌دهد. این بخش از تصاویر غیراخلاقی عمدتاً از تصاویر غیرحرفه‌ای در زندگی روزمره ناشی می‌شود که تغییرات زیادی در زمینه، مقیاس، روشنایی و افراد در معرض نمایش قرار می‌گیرند.

۸- نتیجه‌گیری

طرح پیشنهادی در این مقاله شامل دو بخش پیش پردازش و پردازش است. بخش پیش‌پردازش شامل تنظیم‌گر

۸- منابع

Generations' (Springer, 2016), pp. 1041-1050
 4.Nian, F., Li, T., Wang, Y., Xu, M., and Wu, J.: 'Pornographic image detection utilizing deep convolutional neural networks', *Neurocomputing*, 2016, 210, pp. 283-293
 5.Karamizadeh, S., Abdullah, S.M., Zamani, M., Shayan, J., and Nooralishahi, P.: 'Face recognition via taxonomy of illumination normalization': 'Multimedia Forensics and Security' (Springer, 2017), pp. 139-160
 6.Brancati, N., De Pietro, G., Frucci, M., and Gallo, L.: 'Dynamic Colour Clustering for Skin Detection Under Different Lighting Conditions', in Editor (Ed.)^(Eds.): 'Book Dynamic Colour

1.Wang, H., and Fan, A.: 'Pornographic information of Internet views detection method based on the connected areas', in Editor (Ed.)^(Eds.): 'Book Pornographic information of Internet views detection method based on the connected areas' (International Society for Optics and Photonics, 2017, edn.), pp. 1032228
 2.Wang, Y., Jin, X., and Tan, X.: 'Pornographic image recognition by strongly-supervised deep multiple instance learning', in Editor (Ed.)^(Eds.): 'Book Pornographic image recognition by strongly-supervised deep multiple instance learning' (IEEE, 2016, edn.), pp. 4418-4422
 3.Adnan, A., and Nawaz, M.: 'RGB and hue color in pornography detection': 'Information Technology: New

Applications of Mathematics and Computer Engineering, 2011, pp. 123-128

14.Karamizadeh, S., and Arabsorkhi, A.: 'Methods of pornography detection', in Editor (Ed.)^(Eds.): 'Book Methods of pornography detection' (2018, edn.), pp. 33-38

15.Sufyanu, Z., Mohamad, F.S., Yusuf, A.A., and Mamat, M.B.: 'Enhanced Face Recognition Using Discrete Cosine Transform', Engineering Letters, 2016, 24, (1)

16.Liu, Z., Zhao, H., Pu, J., and Wang, H.: 'Face recognition under varying illumination', Neural Computing and Applications, 2013, 23, (1), pp. 133-139

17.Anagha, K., and Ram, A.R.: 'Pose Tolerant Face Recognition: A Review', in Editor (Ed.)^(Eds.): 'Book Pose Tolerant Face Recognition: A Review' (IEEE, 2020, edn.), pp. 0147-0152

18.Jin, X., Wang, Y., and Tan, X.: 'Pornographic Image Recognition via Weighted Multiple Instance Learning', IEEE transactions on cybernetics, 2018, 49, (12), pp. 4412-4420

19.Lopes, A.P., de Avila, S.E., Peixoto, A.N., Oliveira, R.S., and Araújo, A.d.A.: 'A bag-of-features approach based on hue-sift descriptor for nude detection', in Editor (Ed.)^(Eds.): 'Book A bag-of-features approach based on hue-sift descriptor for nude detection' (IEEE, 2009, edn.), pp. 1552-1556

20.Moustafa, M.: 'Applying deep learning to classify pornographic images and videos', arXiv preprint arXiv:1511.08899, 2015

21.Ding, X., Li, B., Li, Y., Guo, W., Liu, Y., Xiong, W., and Hu, W.: 'Web Objectionable Video Recognition Based on Deep Multi Instance Learning with Representative Prototypes Selection', IEEE Transactions on Circuits and Systems for Video Technology, 2020

Clustering for Skin Detection Under Different Lighting Conditions', in Editor (Ed.)^(Eds.): 'Book Dynamic Colour Clustering for Skin Detection Under Different Lighting Conditions' (Springer, 2016, edn.), pp. 27-35

7.Karamizadeha, S., Mabduallah, S., Randjbaranc, E., and Rajabid, M.J.: 'A review on techniques of illumination in face recognition', Technology, 2015, 3, (02), pp. 79-83

8.Surinta, O., and Khamket, T.: 'Recognizing pornographic images using deep convolutional neural networks', in Editor (Ed.)^(Eds.): 'Book Recognizing pornographic images using deep convolutional neural networks' (IEEE, 2019, edn.), pp. 150-154

9.Noh, Y., Koo, D., Kang, Y.-M., Park, D., and Lee, D.: 'Automatic crack detection on concrete images using segmentation via fuzzy C-means clustering', in Editor (Ed.)^(Eds.): 'Book Automatic crack detection on concrete images using segmentation via fuzzy C-means clustering' (IEEE, 2017, edn.), pp. 877-880

10.Md Noor, S.S., Ren, J., Marshall, S., and Michael, K.: 'Hyperspectral image enhancement and mixture deep-learning classification of corneal epithelium injuries', Sensors, 2017, 17, (11), pp. 2644

11.Gross, R., Baker, S., Matthews, I., and Kanade, T.: 'Face recognition across pose and illumination': 'Handbook of face recognition' (Springer, 2005), pp. 193-216

12.Chen, Z., Liu, C., Chang, F., Han, X., and Wang, K.: 'Illumination processing in face recognition', International Journal of Pattern Recognition and Artificial Intelligence, 2014, 28, (05), pp. 1456011

13.Basilio, J.A.M., Torres, G.A., Pérez, G.S., Medina, L.K.T., and Meana, H.M.P.: 'Explicit image detection using YCbCr space color model as skin detection',

ارائه مدل پذیرش فناوری اطلاعات در پروژه‌های تحقیقاتی دانشگاهی حوزه ICT بر اساس مدل پذیرش فناوری ITAIM

*شهرام علی یاری *مسعود موحدی ***سیروس کاظمیان

*استادیار گروه مهندسی صنایع دانشگاه جامع امام حسین (ع)

** دانشیار گروه مهندسی صنایع دانشگاه جامع امام حسین (ع)

*** کارشناسی ارشد مهندسی صنایع دانشگاه جامع امام حسین (ع)

تاریخ پذیرش: ۱۳۹۹/۰۵/۰۸

تاریخ دریافت: ۱۳۹۸/۱۰/۳۰

چکیده

امروزه ظهور و گسترش فناوری‌هایی که امکان ارتباط هر چه گسترده‌تر را فراهم کرده‌اند، تغییرات قابل‌ملاحظه‌ای را در هم زندگی خصوصی و هم زندگی حرفه‌ای افراد به وجود آورده است. اجرای صحیح فناوری اطلاعات منشأ توسعه اقتصادی و فرهنگی و ارتقای کیفیت زندگی از طریق تبادل اطلاعات و ارائه خدمات عمومی و خصوصی است. هدف از این پژوهش ارائه مدل پذیرش فناوری اطلاعات در پروژه‌های فناوری اطلاعات مراکز تحقیقاتی ICT دانشگاهی ایران می‌باشد. در این مقاله به بررسی عوامل و پارامترهای مؤثر بر پذیرش فناوری اطلاعات در پروژه‌های ICT مراکز تحقیقاتی دانشگاهی حوزه فناوری اطلاعات پرداخته شده است. برای جمع‌آوری داده‌ها و اطلاعات موردنیاز از پرسشنامه استاندارد مرتبط استفاده شده است. جامعه آماری مورد نظر خبرگان شاغل در پروژه حوزه ICT مرکز تحقیقاتی دانشگاه بوده است که به دلیل انتخاب پروژه خاص در حوزه مذکور، نمونه برداری غیراحتمالی در دسترس (۳۰ نفر) انجام شده است. تجزیه و تحلیل داده نیز با استفاده از نرم‌افزارهای Smart pls3 و Spss 22 انجام شد. باتوجه به محاسبات انجام‌شده، می‌توان عوامل مؤثر بر پذیرش فناوری اطلاعات در مراکز تحقیقاتی دانشگاهی را به چهار دسته شامل عوامل مرتبط با فناوری اطلاعات، عوامل سازمانی، عوامل مرتبط با مدیر عالی اجرایی و عوامل فردی تقسیم نمود. نتایج تحقیق نشان می‌دهد، به ترتیب عوامل مدیریتی (به میزان ۰/۴۹۷)، عوامل مرتبط با فناوری اطلاعات (به میزان ۰/۴۶۰) و عوامل فردی (به میزان ۰/۴۵۷) بر پذیرش فردی فناوری اطلاعات و عوامل سازمانی (به میزان ۰/۴۶۹) و عوامل فردی (به میزان ۰/۴۵۹) تاثیرگذاری قابل توجهی بر پذیرش سازمانی فناوری اطلاعات دارند.

واژه‌های کلیدی: فناوری اطلاعات، پذیرش فناوری اطلاعات، پروژه‌های تحقیقاتی، مدل ITAIM

۱- مقدمه

امروزه ظهور و گسترش فناوری‌هایی که امکان ارتباط هر چه گسترده‌تر را فراهم کرده‌اند، تغییرات قابل‌ملاحظه‌ای را در هم‌زندگی خصوصی و هم‌زندگی حرفه‌ای افراد به وجود آورده است [۴]. در واقع فناوری اطلاعات و ارتباطات با تسهیل و گسترش مبادله اطلاعات و کاهش هزینه‌های دادوستد به‌عنوان سازوکاری در جهت افزایش بهره‌وری، کارایی، رقابت‌انگیزی و رشد در همه حیطه‌های فعالیت بشری مطرح است. اجرای صحیح فناوری اطلاعات منشأ توسعه اقتصادی و فرهنگی و ارتقای کیفیت زندگی از طریق تبادل اطلاعات و ارائه خدمات عمومی و خصوصی است [۶]. یکی از مواردی که جوامع از دیرباز با آن مواجه بوده‌اند ظهور پدیده‌های جدید و بحث تغییر و تحولاتی که وقوع آن تأثیر بسزایی در عملکرد آن جامعه و سازمان داشته و توجه و عدم توجه و واکنش مناسب نسبت به این تغییرات، تأثیر بسیار مهمی در موفقیت و شکست این پدیده‌ها داشته است [۷]. مزایای بالقوه این فناوری زمانی قابل حصول است که در بین تمام بخش‌های جامعه توسعه و گسترش یابد. از زمان ظهور فناوری اطلاعات محققان از دیدگاه‌های متفاوتی به بررسی آثار و تبعات استفاده از این فناوری پرداخته‌اند. با افزایش سرمایه‌گذاری در فناوری‌های جدید مطالعه پذیرش این فناوری‌ها بسیار مورد توجه قرار گرفته است [۱۱]. در عصر حاضر فناوری اطلاعات راهبرد جدیدی است که تمام ابعاد سازمان‌ها را تحت تأثیر قرار داده است به‌گونه‌ای که تصور سازمان بدون فناوری اطلاعات و همگام بودن با فناوری‌های نوین غیرممکن به نظر می‌رسد [۲۱]. تحقیقات نشان می‌دهد که یکی از مهم‌ترین دلایل تمایز سازمان‌ها از یکدیگر، درجه به‌کارگیری فناوری اطلاعات توسط آن‌ها در فعالیتهای سازمانی است. به‌این ترتیب سازمان‌هایی که از فناوری اطلاعات به‌طور^۱ است که برای موضوع تحقیق بومی سازی شده است. لازم به ذکر است مدل ITAIM یک مدل جامع فردی و سازمانی برای پذیرش فناوری اطلاعات در سازمان‌های دولتی ایران می باشد [۱۴] ابعاد مدل مناسب برای پذیرش فناوری

گسترده و بهینه استفاده کرده‌اند، دارای مزیت رقابتی پایدار هستند. باین وجود پروژه‌های فناوری اطلاعات در بیشتر موارد موفق نبوده‌اند، تقریباً نزدیک به ۷۰٪ از پروژه‌های پیاده‌سازی فناوری اطلاعات با شکست مواجه شده‌اند [۲۰]. آمارها نشان می‌دهد که درصد کمی از پروژه‌های فناوری اطلاعات با موفقیت اجرا می‌شوند. مطالعات گروه استندیش در سال ۲۰۱۴، نشان می‌دهد که حدود ۳۱٫۱٪ از پروژه‌ها قبل از به اتمام رسیدن به‌طور کامل منحل می‌شوند و ۵۲٫۷٪ از آن‌ها در گروه پروژه‌های چالشی قرار می‌گیرند و از محدوده زمانی و بودجه تعیین‌شده، فراتر می‌روند و هزینه‌ای برابر با ۱۸۹٪ بیشتر از برآورد اصلی خود خواهند داشت و اما در مقابل، به‌طور متوسط تنها حدود ۱۶٫۲٪ از پروژه‌های نرم‌افزاری در زمان و بودجه از پیش تعیین‌شده به اتمام رسیده‌اند [۴۰]. تحقیقات اخیر نیز کماکان نشان می‌دهد که شرکت‌های فعال در حوزه IT سالانه در حدود ۵۰ درصد مواقع از منظر هزینه کیفیت زمان اجرا و اثربخشی، با شکست مواجه شده‌اند در تحقیقی دیگر نیز از ۱۲۶ پروژه بررسی شده ۵۵ درصد آنها موفق نبوده‌اند در صورتیکه در سنوات قبل این عدد به حدود ۳۵ درصد بوده است (گزارش خطوط راهنما مدیریت پروژه ۲۰۱۸) شایان ذکر است عوامل مختلفی در موفقیت و شکست پروژه‌های حوزه ICT تأثیرگذار هستند [۳۵]. عدم توجه به عواملی نظیر عوامل فنی، انسانی، اقتصادی، فرهنگی و مدیریتی باعث شکست یا توقف پروژه‌ها و تحمل هزینه‌های هنگفت می‌شود [۲۴]. هدف این تحقیق ارائه مدلی برای پذیرش فناوری اطلاعات با محوریت پروژه‌های فناوری اطلاعات مراکز تحقیقاتی ICT دانشگاهی (مطالعه موردی یک مرکز تحقیقاتی دانشگاهی در حوزه مذکور) و همچنین شناسایی عوامل موثر بر پذیرش فناوری اطلاعات در این مراکز می‌باشد. مدل پایه این تحقیق مدل ITAIM اطلاعات در مراکز تحقیقاتی حوزه فناوری اطلاعات دانشگاهی چیست؟

۲- تعاریف فناوری اطلاعات

انجمن فناوری امریکا^۲، فناوری اطلاعات را مطالعه، طراحی، گسترش و توسعه، اجرا، پشتیبانی و مدیریت دستگاه‌های

^۱.Information technology Adoption Integrated Modeling

^۲.ITAA

معنی عمل موافقت یا اثبات و تأیید ایده، طرح و ... است [۱۴].

علی یاری در تعریفی جدید اذعان می‌دارد پذیرش فناوری اطلاعات در سازمان‌ها فرآیندی چند مرحله‌ای است که از قبل پیاده‌سازی فناوری اطلاعات شروع می‌شود و پس از تصمیم پذیرش و انتشار در سازمان، تا پس از پیاده‌سازی آن ادامه دارد و در نهایت بعد از پیاده‌سازی و نفوذ در سازمان، مورد پذیرش واقعی و استفاده کامل کاربران به عنوان ابزاری برای انجام فعالیت‌های جاری سازمان قرار می‌گیرد. کارتر در سال ۱۹۹۹ در تعریف پذیرش فناوری بیان می‌دارد که وقتی یک محصول یا سرویس به‌وسیله‌ی یک شرکت عرضه می‌شود لازم است که مشتری به خرید آن محصول متقاعد گردد و این تحریک نمودن مشتریان به‌وسیله تولیدکنندگان یا فراهم‌آوردگان سرویس، توسط فرآیند پیچیده‌ای که وقتی از آن بر خصوصیات و رفتارهای مشتریان صورت می‌پذیرد، این‌یک جریان ذهنی در رابطه با یک فرد است که از اولین شنود فرد درباره نوآوری توسط فرد پذیرش نام‌گرفته است. کارتر در سال ۲۰۰۳ تعریف جدیدی ارائه داده که پذیرش یک تصمیم منحصر به فرد است که یک مصرف‌کننده عادی در رابطه با استفاده از محصول می‌گیرد. راجرز و شومیکر تصمیم در به‌کارگیری ایده‌ی جدیدی همچون فناوری اطلاعات را به‌عنوان بهترین اقدام یا عمل پذیرش می‌گویند، پذیرش زمانی اتفاق می‌افتد که یک فرد ایده‌ی جدیدی را بپذیرد و در عمل به کار گیرد [۲۸]. پذیرش فناوری اطلاعات به مفهوم قبول و به‌کارگیری یا استفاده از فناوری اطلاعات می‌باشد. به‌عبارت‌دیگر درک عواملی که موجب پذیرش یک فناوری (فناوری اطلاعات) می‌شود و یا ایجاد شرایطی که تحت آن فناوری اطلاعات، مورد پذیرش قرار می‌گیرد، پذیرش فناوری اطلاعات اطلاق می‌شود [۴۱] پذیرش فناوری اطلاعات عبارت است از رضایت قابل‌شرح و اثبات به کار بردن سامانه‌های ارتباطی از نظر وظایفی که برای پشتیبانی آن‌ها طراحی شده‌اند [۴۳]. پذیرش پدیده‌ای چندبعدی است و مجموعه‌ای وسیع از متغیرهای مهم مانند ادراک‌ها، اعتقادات، نگرش‌ها و ویژگی‌های افراد و نیز میزان درگیری آنان با فناوری اطلاعات را شامل می‌شود [۲۸]، به‌بیان‌دیگر پذیرش فناوری میزان احساسی است که افراد به‌صورت اختیاری در قصد استفاده از

اطلاعاتی رایانه‌ای (کاربردهای سخت‌افزاری و نرم‌افزاری) تعریف نموده است. مطابق این تعریف فناوری اطلاعات به جنبه‌های مرتبط با فناوری یک سیستم اطلاعاتی اشاره دارد و شامل سخت‌افزار، بانک اطلاعاتی، شبکه‌های نرم‌افزاری و دیگر ابزارها می‌باشد. در این تعریف فناوری اطلاعات می‌تواند به‌عنوان یک زیرسیستم اطلاعاتی در نظر گرفته شود که قابلیت مطالعه، طراحی، گسترش و توسعه را خواهد داشت [۱۶]. محور فناوری اطلاعات رایانه و ارتباط از راه دور است. رایانه اساس کار پردازش و انباشت اطلاعات را انجام می‌دهد و ارتباط راه دور امکان پخش و توزیع اطلاعات را در سطح وسیعی فراهم می‌آورد [۸]. همچنین فناوری اطلاعات شامل مطالعه، طراحی و توسعه، اجرا، حمایت یا مدیریت نظام‌های اطلاعاتی مبتنی بر رایانه و به‌طور خاص کاربردهای نرم‌افزاری و سخت‌افزاری رایانه است [۱۰]. تعاریف متعددی از فناوری اطلاعات توسط خبرگان و محققان این حوزه ارائه شده است که برخی از این تعاریف به شرح زیر می‌باشند:

فناوری اطلاعات عبارت است از گردآوری، سازمان‌دهی، ذخیره و نشر اطلاعات اعم از صوت، تصویر، متن یا عدد که با استفاده از ابزار رایانه و مخابرات صورت می‌پذیرد [۵]. فناوری اطلاعات به ابزارها و روش‌هایی گفته می‌شود که اطلاعات را در شکل‌های مختلف جمع‌آوری، ذخیره، بازیابی، پردازش، تحلیل و توزیع می‌کند [۲۱].

فناوری اطلاعات معادل عناصر فناورانه از سامانه‌های اطلاعاتی به‌کارگیری شده برای جمع‌آوری، پردازش، تحلیل و انتشار اطلاعات برای هدف خاص تلقی می‌شود. فناوری اطلاعات وسیله‌ای برای بهبود کارایی و اثربخشی عملیاتی بوده و فرآیندهای راهبردی و راه‌کنشی در سازمان‌ها را دربرمی‌گیرد [۱۳].

فناوری اطلاعات واژه‌ای کلی است از سخت‌افزار، نرم‌افزار و فکر افزار که گردش و بهره‌برداری از اطلاعات را امکان‌پذیر می‌سازد [۲۸].

۳- پذیرش فناوری اطلاعات

در ادبیات موضوع، پذیرش فناوری اطلاعات با دو کلمه "Adoption" و "Acceptance" آمده است. در فرهنگ لغت انگلیسی "Adoption" به معنی تصمیم برای شروع استفاده از ایده، طرح یا یک اسم و "Acceptance" به

مدل‌های پذیرش فناوری اطلاعات

در دو دهه گذشته، نظریه‌ها و مدل‌های مختلفی در زمینه پذیرش فناوری، پیشنهاد، آزمون و اصلاح شده‌اند. این مدل‌ها به شناخت ما از عوامل مؤثر در پذیرش فناوری از سوی کاربران و روابط بین آنها کمک می‌کنند [۴۱]. در جدول (۱) مدل‌های پذیرش فناوری اطلاعات و عوامل مؤثر بر پذیرش فناوری اطلاعات بر اساس این مدل‌ها شرح داده شده است.

فناوری‌های خاص را دارند، از این‌رو پذیرش یا عدم پذیرش عاملی ضروری و تعیین‌کننده است و درک، شناسایی و ارزیابی عوامل مؤثر بر آن در حکم هدف اساسی باید در پژوهش‌ها مورد توجه قرار گیرد [۳]. نکته حایز اهمیت در بررسی ادبیات پذیرش فناوری اطلاعات و دقت در نحوه به کارگیری کلمات Adoption و Acceptance این است که کلمه Adoption در پذیرش فناوری اطلاعات سازمانی و Acceptance در پذیرش فناوری اطلاعات فردی مورد استفاده قرار می‌گیرد [۱۳].

۴- مبانی نظری

جدول ۱: مدل‌های پذیرش فناوری اطلاعات

عوامل	محقق / سال انتشار	عوامل مؤثر بر پذیرش فناوری اطلاعات
نظریه عمل مستدل	فیشبن و اجزن (۱۹۷۵)	عقاید و سنجش‌ها، نگرش‌ها، عقاید هنجاری و پیروی از انگیزش، هنجارهای ذهنی
اشاعه نوآوری	راجرز (۱۹۸۳)	مزیت نسبی، پیچیدگی، آزمون‌پذیری، قابلیت رؤیت پذیری، سازگاری
نظریه رفتار برنامه‌ریزی شده	اجزن (۱۹۸۵)	باورهای رفتاری و ارزیابی نتایج، باورهای رفتاری و انگیزه تبعیت، باورهای کنترلی، نگرش، هنجارهای ذهنی، کنترل رفتاری درک شده
نظریه تناسب وظیفه - فناوری	گودهیو (۱۹۸۸)	تجربه کار با فناوری، خصایص وظیفه، کاربردهای فناوری، ارزشیابی کاربر، عملکرد
مدل پذیرش فناوری	دیویس (۱۹۸۹)	سودمندی ادراکی و سهولت ادراکی
نظریه استفاده از رایانه شخصی	تامپسون و همکاران (۱۹۹۱)	عادت، تمایلات، شرایط تسهیل گر، احساس، عوامل اجتماعی، نتایج قابل انتظار
مدل انگیزشی	دیویس (۱۹۹۲)	انگیزه درونی، انگیزه بیرونی، سبک احساس
نظریه رفتار برنامه‌ریزی شده توزیع شده	تیلور و تاد (۱۹۹۵)	سازگاری، ادراک سودمندی، سهولت استفاده، نفوذ همکاران، نفوذ مافوقی‌ها، خودکارآمدی، منبع تسهیل گر شرایط، فناوری تسهیل گر شرایط
مدل تلفیقی پذیرش فناوری و نظریه تناسب وظیفه-فناوری	دیشاو و استرانگ (۱۹۹۹)	خصایص وظیفه، تناسب وظیفه فناوری، تجربه کار با فناوری، کارکردهای فناوری، درک سودمندی، درک سهولت استفاده،
مدل TAM 2	ونکاتش و دیویس (۲۰۰۰)	اختیاری بودن، تجربه، هنجار ذهنی، تصویر ذهنی، ارتباط شغلی، کیفیت خروجی، قابلیت رؤیت پذیری
مدل تلفیقی پذیرش و کاربرد فناوری	وانکاتش و همکاران	جنسیت، سن، تجربه، اختیاری بودن، تسهیل شرایط، نفوذ اجتماعی، پیش‌بینی تلاش، پیش‌بینی عملکرد
مدل TAM 3	ونکاتش و همکاران (۲۰۰۸)	هنجار ذهنی، تصویر ذهنی، ارتباط شغلی، کیفیت خروجی، نتایج قابل مشاهده، خودکارآمدی کامپیوتر، درک از کنترل بیرونی، اضطراب ناشی از کامپیوتر، تجربه، داوطلب بودن
مدل تلفیقی پذیرش و کاربرد فناوری ۲	ونکاتش و همکاران (۲۰۱۲)	جنسیت، سن، تجربه، عملکرد مورد انتظار، تلاش مورد نظر، اثر اجتماعی، شرایط تسهیل‌کننده، انگیزه لذت‌بخش، هزینه، عادت
مدل پذیرش فناوری برای سالمندان STAM	چن (۲۰۱۴)	خودکارآمدی تکنولوژی، اضطراب تکنولوژی، شرایط تسهیل گر، شرایط سلامت شخصی، توانایی شناختی، روابط اجتماعی، نگرش به زندگی و رضایت، عملکرد فیزیکی
مدل ITAIM	علی یاری (۱۳۹۳)	عوامل مرتبط با فناوری اطلاعات، عوامل سازمانی، عوامل محیطی، عوامل مرتبط با مدیر

عالی اجرایی و عوامل سازمانی

سازمانی، محیطی و فردی. [۴۲] عوامل مذکور را به صورت داخلی (سازمانی)، خارجی (محیطی) و فناوری دسته‌بندی کرده‌اند. در میان ویژگی‌های افراد، ویژگی‌های رهبر سازمان یا مدیر عالی اجرایی به عنوان یکی از پیش‌بینی‌کننده‌ها در فرآیند نوآوری در سازمان پیشنهاد شده است [۳۸]. بر اساس بررسی مدل‌های پذیرش فناوری (جدول ۱) و بررسی پژوهش‌های مرتبط با پذیرش فناوری اطلاعات می‌توان عوامل مؤثر بر پذیرش فناوری اطلاعات را به پنج دسته تقسیم نمود که در جدول (۲) نمایش داده شده‌اند.

جدول ۲: عوامل مؤثر بر پذیرش فناوری اطلاعات

عوامل	سازه‌ها
عوامل فناورانه	سازگاری، پیچیدگی، مزیت نسبی، آزمون‌پذیری، مشاهده‌پذیری، وسعت به‌کارگیری فناوری اطلاعات، خودکارآمدی رایانه
عوامل فردی	سن، تحصیلات، تجربه، مهارت، تلاش، هنجارهای ذهنی، تصویر ذهنی، توانایی شخصی، نگرش، باورها، برداشت ذهنی از آسانی استفاده، برداشت ذهنی از مفید بودن، عادت، تمایلات
عوامل سازمانی	اندازه سازمان، عمر سازمان، تمرکز، رسمیت، ساختار سازمانی، فرهنگ سازمانی، حمایت و پشتیبانی سازمانی، آموزش کارکنان، تخصیص مناسب منابع به فناوری اطلاعات، جوسازمانی، همکاری درون‌سازمانی و توانمندسازی کارکنان
عوامل مدیریتی	آگاهی مدیریت ارشد نسبت به فناوری اطلاعات، حمایت و پشتیبانی مدیریت، درگیر بودن مدیر در پروژه‌های فناوری اطلاعات، گرایش مدیران به نوآوری و ریسک‌پذیر بودن مدیران
عوامل محیطی	سیاست‌های دولت، زیرساخت‌های ملی فناوری، قوانین و مقررات

۵- پیشینه پژوهش

افزوده است. [۷] در پایان‌نامه خود با عنوان «ارائه یک مدل جدید برای پذیرش سرویس‌های دولت الکترونیک توسط کاربران» بیان می‌کند که یکی از بزرگ‌ترین چالش‌ها در این خصوص پذیرش سرویس‌های دولت الکترونیک توسط کاربران است. عوامل مطرح‌شده در این تحقیق به سه گروه فردی، سازمانی و اجتماعی تقسیم‌شده‌اند. [۱۲] در پژوهش خود تحت عنوان «پذیرش فناوری اطلاعات و ارتباطات: مطالعه موردی شرکت‌های کوچک و متوسط تهران»، عوامل مؤثر بر پذیرش از سوی این شرکت‌ها را در دسته‌های عوامل داخلی، عوامل خارجی و مشخصات فناوری اطلاعات و ارتباطات معرفی می‌کند. در همین راستا موانع پذیرش نیز مشتمل بر حمایت مدیریتی ضعیف، فقدان مهارت‌های لازم، چشم‌انداز

تحقیقات گذشته عوامل مؤثر متفاوتی را در پذیرش فناوری اطلاعات ارائه نموده‌اند [۳۴]. این عوامل در دسته‌بندی‌های مختلفی از ادبیات سیستم‌های اطلاعاتی گروه‌بندی شده‌اند. برای مثال [۴۲] پنج دسته‌بندی تأثیرگذار در انتشار IT شامل فناوری، سازمان، محیط، وظیفه و ویژگی کاربر را معرفی کرده‌اند. دسته‌بندی [۴۶] نیز شامل ویژگی‌های نوآوری، سازمان، محیط و ویژگی‌های مدیر عالی اجرایی بوده است. در ادبیات فناوری اطلاعات چهار دسته‌بندی عمده برای مولفه‌های تأثیرگذار معرفی شده‌اند که عبارتند از فناوری،

یقیناً محققان مختلفی موضوع پذیرش فناوری اطلاعات را در محیط‌های متفاوت با استفاده از مدل‌های مربوط به این حوزه موردبررسی قرار داده‌اند که از جمله این تحقیقات می‌توان به این موارد اشاره نمود؛ [۱] در پژوهش خود با عنوان «پذیرش بانکداری اینترنتی در ایران: بسط مدل پذیرش» بیان می‌کند که درک سودمندی و سهولت استفاده از مهم‌ترین عوامل در پذیرش فناوری‌های جدید در حوزه سیستم‌های اطلاعاتی است. با این حال این دو عامل به‌تنهایی نمی‌توانند رفتار کاربران را در محیط‌های نوظهور مانند بانکداری اینترنتی توصیف کنند، وی در پژوهش خود دو سازه برداشت ذهنی از توانایی‌های شخصی و اعتماد را به مدل پذیرش فناوری

استراتژیک ضعیف با توجه به فناوری اطلاعات، محیط‌زیست نوآورانه ضعیف و منابع ناکافی مالی (عوامل داخلی) و همچنین حمایت ضعیف دولتی، پذیرش فرهنگی کم و در نهایت درک پایین سودمندی و سازگاری کم (عوامل خارجی) از جمله عواملی هستند که مانع از پذیرش فناوری اطلاعات و ارتباط در این شرکت‌ها می‌باشند. [۱۲] نقش عوامل مؤثر بر پذیرش و به‌کارگیری فناوری اطلاعات و ارتباطات در بین جوانان روستایی بر اساس نظریه نشر نوآوری راجرز مورد بررسی قرار داده است و نتایج به‌دست‌آمده بیانگر آن است که متغیرهای مفید بودن، سهولت استفاده و نوع شغل در میزان پذیرش فناوری‌های اطلاعاتی تأثیر داشته و سازگاری فناوری اطلاعاتی و ارتباطات و میزان تحصیلات در پذیرش آن تأثیر ندارد. همچنین بین متغیرهای میزان آموزش، بهره‌مندی از کانال‌های ارتباطی و به‌کارگیری فناوری اطلاعات و ارتباطات، ارتباط معناداری ندارد. [۱۸] عوامل مؤثر بر پذیرش فناوری اطلاعات در محیط بالینی از دیدگاه پرستاران را مورد بررسی قرار داده و بیان می‌کند که بیشتر پرستاران سیستم اطلاعاتی بالینی را در کار روزانه خود پذیرفته‌اند و ملاحظات متعددی در حیطه‌های فنی و فردی را مدنظر قرار می‌دهند. [۲۲]، در تحقیقی با عنوان «بررسی عوامل مؤثر بر پذیرش فناوری‌های اطلاعاتی در سازمانها» به تاثیر عواملی نظیر ادراک مدیران از منافع و چالش‌های فناوری اطلاعات، اندازه سازمان، ثبات مدیریت و آمادگی تامین کنندگان سازمان بر پذیرش فناوری اطلاعات در سازمان اشاره نموده اند. [۱۴] در پژوهش خود با عنوان «مطالعه عوامل مؤثر بر پذیرش فناوری اطلاعات در صنایع کوچک و بزرگ» بیان می‌کند که متغیرهای مدل به‌غیر از رابطه میان سهولت درک شده بر نگرش به استفاده از فناوری از اطلاعات در صنایع کوچک و متوسط دارای تأثیر مثبت و معنادار می‌باشد. [۱۳] در پژوهش خود با عنوان «عوامل مؤثر بر پذیرش فناوری اطلاعات در سازمان‌های دولتی» عوامل مؤثر بر پذیرش فناوری اطلاعات در سازمان‌های دولتی ایران را در نظر گرفتن ویژگی‌های سازمان‌های دولتی ایران را مورد بررسی قرار داده است. بر اساس یافته‌های این پژوهش پنج دسته عوامل شامل عوامل مرتبط با فناوری اطلاعات، عوامل سازمانی، عوامل محیطی، عوامل مرتبط با مدیر عالی

اجرایی و عوامل فردی را بر پذیرش فناوری در سازمان‌های دولتی ایران را مؤثر می‌داند. [۱۷] نیز در تحقیق خود با عنوان «بررسی عوامل مؤثر بر پذیرش فن آوری رایانه توسط معلمان» ضمن بررسی موضوع با استفاده از نقشه‌های شناختی فازی به این نتیجه می‌رسد که استفاده معلمان از فن آوری رایانه بیشترین مرکزیت دارد. همچنین آموزش معلمان در زمینه رایانه، مهارت رایانه معلمان، استفاده از رایانه در تدریس و سودمندی رایانه در تدریس، در امر تدریس بیشترین تاثیرگذاری داشته است.

پژوهش‌ها و مدل‌های فراوانی در زمینه پذیرش و استفاده از فناوری اطلاعات در خارج از کشور انجام شده است. اما یکی از پرکاربردترین مدل‌ها در حوزه پذیرش فناوری‌های جدید، مدل پذیرش فناوری (TAM) است که توسط دیویس است که در سال ۱۹۸۵ بر مبنای نظریه عمل مستدل در رساله دکتری خود در دانشگاه «ام ای تی» معرفی نموده است. این مدل هم برای پیش‌بینی و هم برای توضیح رفتار کاربران مفید بوده و مدلی در سطح عوامل فردی است. لویس شاپر (۲۰۰۷) در پژوهشی تحت عنوان "مدل پذیرش فناوری اطلاعات و ارتباطات و استفاده‌ی از آن توسط کادر درمانی" که در بیمارستان‌های استرالیا انجام شده است نشان داده که عوامل مرتبط با زمینه‌های فناوری اطلاعات (انتظار عملکرد، انتظار تلاش) و زمینه‌های فردی (اضطراب کامپیوتر، خودکارآمدی، نگرش) و زمینه‌های اجرایی (نفوذ اجتماعی، سازگاری، شرایط تسهیل‌گر سازمانی) بر پذیرش و استفاده از فناوری اطلاعات و ارتباطات تأثیر بسزایی دارند. [۳۴] از مدل TAM برای پذیرش یک فناوری اطلاعات در حوزه سلامت در ایران استفاده کرده اند. همچنین [۳۲]، مدل پذیرش فناوری را به منظور بررسی پذیرش بسترهای اینترنتی بیمارستانی در بهبود سلامت افراد بزرگسال مورد استفاده قرار داده اند. مشارکت کنندگان در استفاده از پورتال طراحی شده، به کاهش وقت و هزینه در صورت پذیرش و استفاده از آن اشاره کرده اند. [۲۹] موضوع پذیرش فناوری‌های سیستم‌های مبتنی بر اینترنت را در سازمان‌های کارمندیابی امریکا مورد مطالعه قرار داده است و سه عامل

می‌کند که برای سالمندان موانع بسیاری برای پذیرش و استفاده از فناوری اطلاعات سلامت از جمله مسائل مربوط به آشنایی، تمایل به درخواست کمک، اعتماد فناوری، حریم خصوصی و چالش‌های طراحی مواجه هستند.

[۲۷] در پژوهش خود با عنوان «تجزیه و تحلیل رفتار کاربران مسن، پذیرش سیستم‌های فناوری سلامت خانگی (HTS) با استفاده از توسعه مدل UTAUT» نشان می‌دهد که انتظار عملکرد، انتظار پیشرفت، شرایط تسهیل‌کننده، سودآوری درک شده، و امنیت درک شده تأثیر مستقیم بر روی قصد رفتاری برای استفاده از HTS داشته است. علاوه بر این تأثیر نظرات پزشکان و نفوذ اجتماعی نقش مهمی در بین افراد سالمند برای پذیرش این فناوری داشته است. [۲۷] در پژوهشی تحت عنوان «مدل پذیرش یک سیستم اطلاعاتی بیمارستانی» بیان می‌کند که عوامل غیر فناوری مانند ویژگی‌های انسانی (یعنی سازگاری، امید به امنیت اطلاعات و خودکارآمدی) و ویژگی‌های سازمانی (یعنی حمایت مدیریت، تسهیل شرایط و مشارکت کاربران) و سهولت استفاده و مزایا فناوری نقش قابل توجهی در پذیرش و استفاده این فناوری را دارند.

[۳۷] در پژوهشی تحت عنوان «درک عوامل مؤثر بر پذیرش mhealth (موبایل سلامت) در بین سالمندان: توسعه مدل UTAUT» بیان می‌کند که انتظار عملکرد، امید به تلاش، نفوذ اجتماعی، اضطراب تکنولوژی و مقاومت در برابر تغییر تأثیر قابل توجهی در قصد رفتاری کاربران برای اتخاذ خدمات بهداشتی و درمانی دارند. [۲۶] در پژوهش خود با عنوان «توسعه مدل سه‌گانه پذیرش فناوری: پزشکان بیمارستان، چشم‌انداز مبتنی بر EHR» نشان می‌دهد که قصد استفاده از تکنولوژی جدید، سودمندی جمعی پیش‌بینی شده، نوآوری شخصی، سودمندی درک شده، سهولت استفاده درک شده، نفوذ بیمار، تغییر سازمانی، روابط حرفه‌ای، نظارت اداری، پشتیبانی سازمانی و مقاومت در برابر تغییر و اضطراب رایانه‌ای عوامل مهم بر پذیرش فناوری اطلاعات از سوی کاربران هستند. [۳۴] به پذیرش فناوری در بازاریابی دیجیتال پرداخته اند. این تحقیق به منظور بررسی پذیرش وب سایت و رسانه‌های اجتماعی در شرکت‌های متوسط و کوچک مالی انجام شده است. این تحقیق نشان می‌دهد که عملکرد

درک شخصی افراد از فناوری، محرک‌ها و موانع فناوری را بر پذیرش فناوری مؤثر دانسته است. یافته‌های مطالعه [۲۹] برای بررسی پذیرش فناوری سیستم‌های پشتیبان تصمیم توسط پزشکان با استفاده از نظریه تلفیق پذیرش و کاربرد فناوری، نشان دادند که هر دو متغیر انتظارات مربوط به عملکرد و انتظارات مربوط به تلاش بر قصد استفاده از فناوری تأثیر معناداری دارند. همچنین رابطه اثر اجتماعی با قصد استفاده و رابطه شرایط تسهیل‌کننده با رفتار استفاده تأیید شده است.

[۳۳] در پژوهش خود با عنوان «نظریه عمل منطقی برنامه برای پذیرش فناوری اطلاعات سبز» نشان می‌دهد که عوامل خارجی مانند باورهای فرد، پاسخ‌دهندگان و سطح آگاهی از تأثیر قابل توجهی در نگرش نسبت به فناوری اطلاعات سبز برخوردار است. [۳۰] در پژوهش خود با عنوان «پذیرش فناوری توسط سالمندان هنگ‌کنگ: مدل پذیرش فناوری سالمندان STAM» بیان می‌کند برای افراد سالمند هنگ‌کنگ ویژگی‌های فردی مانند سن، جنسیت، آموزش، خودکارآمدی و اضطراب، سلامت و توانایی و همچنین شرایط تسهیل‌گر به‌طور مستقیم بر شرایط تسهیل‌گر تأثیر دارند و به‌طور کلی نتایج به‌دست‌آمده نشان داده که مدل STAM از طرف سالمندان هنگ‌کنگ به‌شدت حمایت می‌شود. [۳۳] در پژوهشی با عنوان «عوامل تعیین‌کننده برای کادر درمانی در "پذیرش فناوری‌های جدید برای توان‌بخشی" استفاده از مدل یکپارچه پذیرش و استفاده از فناوری UTAUT» نشان می‌دهد که انتظار تلاش و سازه نفوذ اجتماعی از اهمیت چندانی برخوردار نبودند و درعین حال قصد رفتاری و شرایط تسهیل‌کننده و یا حمایت نهادی با استفاده صحیح از فناوری‌های نوین در توان‌بخشی مرتبط است. [۴۰] در پژوهش خود با عنوان «پزشکان، پذیرش تبادل الکترونیکی پزشکی: توسعه مدل TPB با اعتماد نهادی و ریسک درک شده» نشان می‌دهد که پنج عامل زیر نقش بسیار مهمی بر قصد استفاده پزشکان تایوانی برای استفاده از سیستم‌های تبادل اطلاعات (EMR) دارند: نگرش، محرک‌های ذهنی، کنترل رفتاری درک شده، اعتماد نهادی و خطر درک شده. [۳۳] در پژوهشی با عنوان «پذیرش و استفاده از فناوری‌های اطلاعاتی حوزه سلامت به‌وسیله جامعه بزرگ سالمندان» بیان

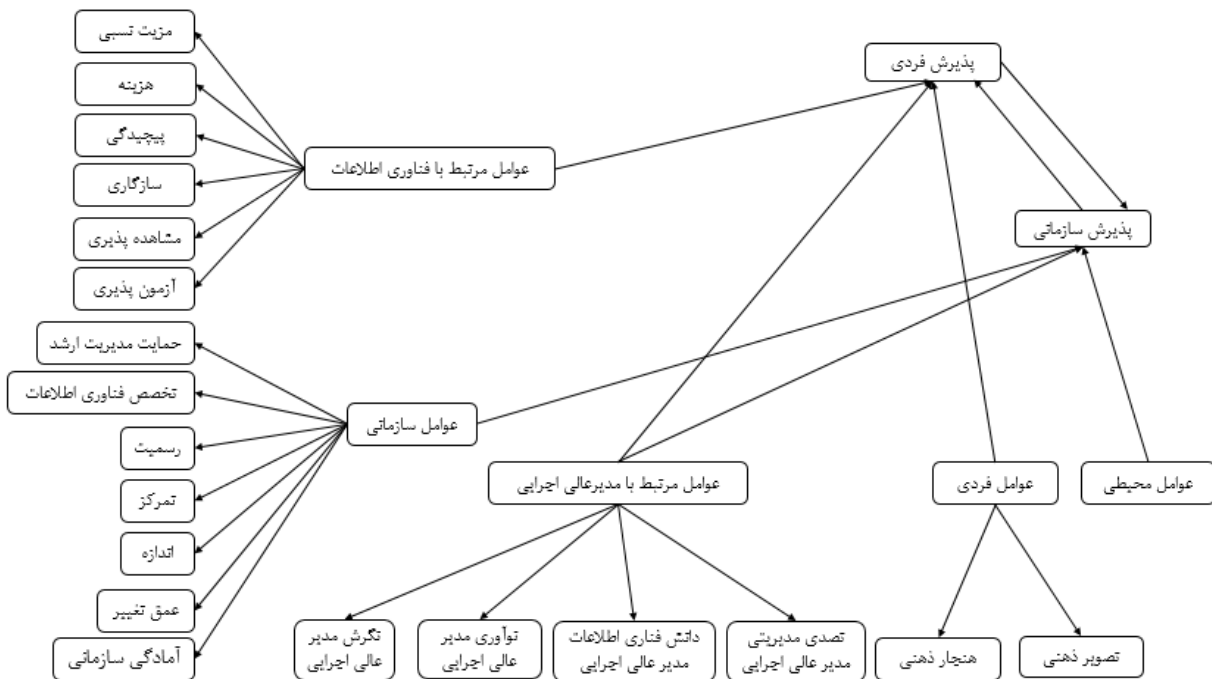
سازمانی، عوامل مدیر عالی اجرایی، عوامل محیطی و عوامل فردی) که به‌عنوان متغیرهای مکنون نیز نام‌برده می‌شوند و متغیرهای وابسته شامل پذیرش فردی و سازمانی می‌باشد. بر اساس مدل ITAIM، پذیرش فناوری اطلاعات در سازمان‌ها فرآیندی چندمرحله‌ای است که قبل از پیاده‌سازی فناوری اطلاعات شروع می‌شود و پس از تصمیم پذیرش و استفاده در سازمان تا پس از پیاده‌سازی آن ادامه دارد و در نهایت بعد از پیاده‌سازی و نفوذ در سازمان مورد پذیرش واقعی و استفاده‌ی کامل کاربران به‌عنوان ابزاری برای انجام فعالیت‌های جاری سازمان مورد توجه قرار می‌گیرد.

همان‌طور که بیان شد مدل ITAIM (شکل ۱) شامل پنج متغیر مکنون مستقل (عوامل مرتبط با فناوری اطلاعات، عوامل سازمانی، عوامل محیطی، عوامل مرتبط با مدیر عالی اجرایی و عوامل فردی) و دو متغیر وابسته (پذیرش فردی و پذیرش سازمانی) می‌باشد.

بازاریابان و تجارت شرکت‌های کوچک و متوسط پس از آشنایی با کانالهای بازاریابی و تبلیغاتی دیجیتال، بهبود قابل توجهی یافته است. [۳۲] در مقاله خود به بررسی ادبیات موضوع پذیرش مدل‌های فناوری اطلاعات و پذیرش نوآوری می‌پردازد. این مقاله ضمن بروز کردن مدل‌های پذیرش فناوری اطلاعات به برجسته کردن فاکتورهای موثر در پذیرش فناوریهای جدید پرداخته است.

۶- مدل ITAIM [۱۴]

مدل ITAIM، ترکیبی تعاملی از مدل‌های پذیرش سطح فردی و مدل‌های پذیرش در سطح سازمانی می‌باشد. این مدل از ویژگی‌های مدل‌های پایه مانند TOE, DOE, TPB, TRA, TAM بهره گرفته و با توجه به ویژگی‌های سازمان‌های دولتی ایران و در نظر گرفتن سطوح پذیرش فردی و سازمانی ارائه گردیده است. مدل ITAIM شامل پنج سازه (عوامل فناوری اطلاعات، عوامل



شکل ۱: مدل ITAIM

مشاهده‌پذیری و آزمون‌پذیری که از طریق ۲۲ شاخص موردسنجش واقع شده‌اند.

سازه‌ها و فاکتورهای مدل ITAIM به شرح زیر می‌باشند:

۱. عوامل مرتبط با فناوری اطلاعات (IT)؛ دارای ۶ متغیر آشکار مزیت نسبی، هزینه، پیچیدگی، سازگاری،

رسمیت (FORM): حدی است که فعالیت‌ها و مشاغل سازمان استاندارد شده‌اند، به عبارت دیگر درجه‌ای است که قوانین و فرآیندهای یک سازمان روی عملکرد اعضا قابل پیگیری است. رسمیت فناوری اطلاعات به عنوان ساختار سیستم اطلاعاتی یا استراتژی فناوری اطلاعات در سازمان در نظر گرفته می‌شود. در سازمان‌های که فرایندهای رسمی ایجاد شده است، برای پذیرش فناوری اطلاعات آمادگی بیشتری دارند.

تمرکز (CEN): میزانی است که تصمیم‌گیری در یک نقطه واحد سازمانی متمرکز شده است و به عبارت دیگر درجه‌ای است که قدرت و کنترل در دست تعداد اندکی از افراد سازمان است. اغلب مطالعات نشان می‌دهند که تمرکز وابستگی منفی با فناوری اطلاعات دارد.

اندازه (SIZE): به تعداد خدمات ارائه شده، تعداد کارکنان، فعالیت‌ها و گستره جامعه‌ای که خدمات دریافت می‌کنند اشاره دارد. این فاکتور در اکثر مطالعات پذیرش فناوری در سازمان‌ها مورد بررسی قرار گرفته است.

عمق تغییر (CH D): حدی است که فرآیندها و ساختار سازمان با پذیرش و پیاده‌سازی فناوری اطلاعات جدید دستخوش تغییر می‌شود.

آمادگی سازمانی (RED ORG): به مجموعه عواملی نظیر زیرساخت‌ها، منابع مالی و انسانی، آموزش، فرهنگ سازمان، واحد سازمانی فناوری اطلاعات، دانش سازمان و مدیران که روی پذیرش و پیاده‌سازی فناوری اطلاعات تأثیر دارند، اشاره دارد.

۳. عوامل محیطی (ENV): حمایت دولت نقش مهمی در ترویج فناوری جدید در سازمان دارد. دولت‌ها با دستورالعمل پیاده‌سازی و کمک‌های مالی و استراتژی بازار می‌توانند پذیرش فناوری اطلاعات را در سازمان‌ها تسهیل نمایند. دولت‌ها همچنین می‌توانند از طریق مشوق‌هایی سازمان‌ها را به پذیرش فناوری اطلاعات تشویق نمایند.

۴. عوامل مرتبط با مدیر عالی اجرایی (CEO): دارای ۴ متغیر آشکار (نگرش، نوآوری، دانش فناوری اطلاعات، تصدی مدیریت) است که از طریق ۱۴ شاخص موردسنجش واقع شده است.

مزیت نسبی (RA): این عامل به میزان درک فرد از بهتر بودن فناوری اطلاعات جدید نسبت به ایده‌های قبلی اشاره دارد.

هزینه (CO): هزینه مترتب شده در پذیرش، پیاده‌سازی و به‌کارگیری فناوری اطلاعات یک عامل مهم است. از هزینه به‌عنوان یک مانع در پذیرش فناوری اطلاعات یاد می‌شود که کاهش آن موجب پذیرش و استفاده از فناوری اطلاعات می‌شود.

پیچیدگی (COMP): به میزان درک فرد از دشواری یادگیری به کار بردن فناوری اطلاعات و آسانی استفاده از آن اشاره دارد.

سازگاری (COMPAT): میزان برداشت از هماهنگی فناوری با ارزش‌های موجود، تجربه‌های گذشته و نیازهای اوست. اکثر ناسازگاری‌ها با فرآیندها و سیستم‌های موجود سازمان است.

قابل‌رؤیت بودن (OBS): میزان برداشت از ملموس بودن نتایج حاصل از بکارگیری فناوری اطلاعات است.

آزمون‌پذیری (TRA): حدی است که یک فناوری اطلاعات را پیش از بکارگیری می‌توان به‌صورت آزمایشی و رایگان به کار گرفت. آزمون‌پذیر ارتباط مثبتی با پذیرش فناوری اطلاعات دارد.

۲. عوامل انسانی (ORG): دارای هفت متغیر آشکار حمایت مدیریت عالی سازمان، تخصص فناوری اطلاعات، رسمیت، تمرکز، اندازه، عمق تغییر و آمادگی سازمانی می‌باشد و از طریق ۲۹ شاخص موردسنجش واقع شده‌اند.

حمایت مدیریت عالی سازمان (TMS): اگر مدیریت منافع فناوری اطلاعات را درک کند، منابع موردنیاز برای پیاده‌سازی فناوری اطلاعات جدید را تخصیص خواهد داد. نقش مدیر عالی اجرایی در تخصیص منابع موردنیاز و فراهم آوردن یک جوسازمانی پشتیبانی‌کننده برای پذیرش کاربران بسیار مهم است.

تخصص فناوری اطلاعات (IT EX): عبارت است از تجربه گذشته سازمان و تنوع دانش فناوری اطلاعات آن. در سازمان‌ها دانش فناوری اطلاعات یک فاکتور اصلی در پذیرش فناوری‌های جدید است.

منزلت او در سیستم اجتماعی می‌شود، تعریف کرده‌اند. برخی از عوامل دیگر فردی مانند اجباری بودن بر اثر طی زمان کاهش می‌یابند، هنجارهای ذهنی نیز چنین است اما تأثیر تصویر ذهنی بر عواملی نظری درک سودمندی در طول زمان کاهش نمی‌یابد.

و در نهایت مدل شامل دو متغیر وابسته پذیرش فردی (IND ACC) و پذیرش سازمانی (ORG ADP) می‌باشد (علی یاری، ۱۳۹۳). مقاله حاضر به منظور بومی سازی مدل مذکور در پروژه‌های حوزه ICT مراکز تحقیقاتی دانشگاهی ارائه شده است.

۷- روش‌شناسی تحقیق

این تحقیق از جنبه هدف، کاربردی و توسعه‌ای است و از جنبه روش توصیفی و پیمایشی می‌باشد. برای دستیابی به نتایج مطلوب، مطالعاتی به صورت میدانی انجام می‌شود و با استفاده از نظر خبرگان و متخصصان، پارامترها و شاخص‌های مؤثر بر پذیرش فناوری اطلاعات در مراکز تحقیقاتی ICT دانشگاهی بر اساس مدل ITAIM شناسایی شده‌اند و سپس رابطه، شدت و نوع رابطه بین پارامترها شناسایی شده، بررسی شده است. در رابطه با محیط شناختی تحقیق می‌توان بیان نمود که به منظور بومی سازی مدل ITAIM در پروژه‌های حوزه ICT، یکی از مراکز دانشگاهی که صرفاً پروژه‌های ICT انجام می‌دهد، انتخاب شده است. لازم به ذکر است که به دلیل دانشگاهی بودن مرکز، کارکنان و سایر عوامل نمی‌توان به صرف دانشگاهی بودن و تصور سواد فناوری اطلاعات بالا، تصور نمود که پذیرش فناوری اطلاعات در این محیط حتماً به خوبی شکل می‌گیرد. بنابراین می‌بایست با دلیل علمی موضوع را بررسی نمود. در قسمت نتیجه‌گیری با استفاده از یافته‌های تحقیق، این موضوع بیشتر تشریح می‌گردد. برای جمع‌آوری اطلاعات و داده‌های مورد نیاز در این پژوهش از روش پرسشنامه با توزیع ۳۰ نمونه‌ای به روش هدفمند و سرشماری، مورد استفاده قرار گرفته است. لازم به ذکر است که این پرسشنامه برای بررسی دیدگاه‌ها و نظرات افراد در مورد شاخص‌ها و متغیرهای پژوهش طراحی شده است که شامل هشت بخش و ۱۱۶ پرسش می‌باشد و داده‌ها و اطلاعات حاصل از پرسشنامه‌ها توسط نرم افزارهای spss22 و Smart PLS3 مورد پردازش قرار گرفته‌اند. اولین

نگرش (AT CEO): برداشت مدیر عالی اجرایی از فناوری اطلاعات جدید نقش مهمی در پذیرش فناوری اطلاعات دارد. به طوری که نگرش مطلوب و نوآور مدیر عالی اجرایی نسبت به فناوری تأثیر مثبتی در پذیرش فناوری اطلاعات در سازمان دارد.

نوآوری (INN CEO): به میزان خلاقیت و نوآوری مدیر عالی اجرایی در مسائل سازمان اشاره دارد. مدیر عالی اجرایی با خصوصیت نوآور بودن و علاقه‌مندی به تغییر می‌تواند به پذیرش فناوری اطلاعات کمک نماید. مدیر عالی نوآور تمایل به انجام ریسک و ارائه راه‌حل‌های جدید است که قبل از آن سعی نشده بود.

دانش فناوری اطلاعات (IT K): مشخصه‌های فردی مدیر عالی اجرایی نقش اساسی در پذیرش فناوری اطلاعات در سازمان را دارد. دانش فناوری اطلاعات مدیر عالی اجرایی همبستگی قوی‌ای با پذیرش فناوری اطلاعات دارد [۴۱]. یک مدیر عالی اجرایی با دانش فناوری اطلاعات بالا می‌تواند فناوری مورد نظر را ارزیابی کند و احتمال پذیرش و پیاده‌سازی آن را افزایش دهد.

تصدی مدیریتی (TEN MAN): این مقوله به تجارب، قدمت و ثبات مدیریتی اشاره دارد. به عبارت دیگر این عامل به طول زمانی که مدیر در پست و شغل جاری‌اش مشغول فعالیت است، اشاره دارد. مدیران باتجربه می‌توانند فرآیند پذیرش را تسهیل نمایند و از اختیاراتشان برای ایجاد فضای مناسبی جهت پیاده‌سازی موفق فناوری اطلاعات در سازمان استفاده نمایند.

۵. عوامل فردی (INDU): دارای دو متغیر آشکار (هنجار ذهنی و تصویر ذهنی) می‌باشد که از طریق ۸ شاخص مورد سنجش واقع شده است.

هنجاری ذهنی (SN): در کنار باور و برداشت فردی، اجتماع نیز ممکن است بر رفتار فرد تأثیرگذار باشد. افراد غالباً بر مبنای ادراکاتشان از آنچه دیگران فکر می‌کنند، باید انجام دهند عمل می‌کنند و قصد آن‌ها جهت پذیرش رفتار به صورت بالقوه متأثر از افرادی است که ارتباط نزدیکی با آن‌ها دارد.

تصویر ذهنی (IM): مور و بن باسات تصویر ذهنی را می‌زانی که فرد فکر می‌کند استفاده از یک فناوری باعث ترفیع مقام و

پایایی ترکیبی محاسبه شده است. مقدار بحرانی بار عاملی در روش PLS برابر ۰/۴ می‌باشد. با توجه به محاسبات انجام شده در بخش سازه‌های مربوط به عوامل مرتبط با فناوری اطلاعات سازه مشاهده پذیری با توجه به حدنصاب نرسیدن مقدار بار عاملی این سازه و در بخش مربوط به عوامل سازمانی دو سازه تمرکز و عمق تغییر با توجه به حدنصاب نرسیدن بارهای عاملی جهت مناسب شدن پایایی مدل، از مدل تحقیق حذف شده‌اند. سایر سازه‌ها با توجه مقادیر محاسبه شده در جدول (۳) از مقدار بار عاملی مناسبی برخوردار هستند. مقدار بحرانی برای ضریب آلفای کرونباخ ۰/۷ می‌باشد و بر اساس مقادیر محاسبه شده در جدول (۳) همه سازه‌ها از مقدار مناسبی برخوردار هستند که نشان از پایداری درونی مناسب مدل می‌باشد. مقدار ملاک برای پایایی ترکیبی ۰/۷ می‌باشد و در صورتی که این مقدار کمتر از ۰/۶ باشد عدم پایایی مدل را نشان می‌دهد. بر اساس محاسبات انجام شده که در جدول (۳) نشان داده شده‌اند، همه سازه‌های مدل از مقادیر مناسبی برای پایایی ترکیبی برخوردار هستند که نشان دهنده پایداری درونی مناسب مدل می‌باشد. مقدار ملاک برای روایی همگرا ۰/۵ می‌باشد که همه مقادیر محاسبه شده برای سازه‌های مدل بیشتر از ۰/۵ می‌باشد و بیانگر همبستگی مناسب سازه‌ها با شاخص‌های خود می‌باشد.

مرحله تحلیل داده‌ها، توصیف یا تلخیص داده‌ها با استفاده از آمارهای توصیفی با استفاده از نرم‌افزار Spss22 است. در این تحقیق از مدل‌سازی معادلات ساختاری با رویکرد حداقل مربعات جزئی (PLS-SEM) استفاده گردید. علت استفاده از این رویکرد به دلیل عدم وابستگی روش به تعداد پرسشنامه توزیع شده یا نمونه انتخابی، همچنین عدم حساسیت به نرمال یا غیر نرمال بودن داده‌ها می‌باشد. بررسی مقالات چاپ شده دهه اخیر نشان از استفاده وسیع پژوهشگران از این روش برای تجزیه و تحلیل داده‌های پژوهش دارد. محققین دلایل زیادی را برای استفاده از این روش ذکر کرده‌اند که برخی از این دلایل به این شرح می‌باشند؛ حجم کم نمونه، داده‌های غیر نرمال، مدل‌های اندازه‌گیری از نوع سازنده، قدرت پیش‌بینی مناسب، پیچیدگی مدل، تحقیق اکتشافی، توسعه تئوری و نظریه، استفاده از متغیرهای طبقه‌بندی شده، بررسی همگرایی، آزمودن تئوری و آزمودن فرضیات شامل متغیرهای تعدیل‌گر و حجم نمونه اندک بهترین دلیل برای استفاده از PLS است. مراحل تجزیه و تحلیل داده‌ها با استفاده از مدل‌سازی معادلات ساختاری با استفاده از نرم‌افزار PLS در مطالب ارائه شده توسط داوری [۹]، ورتس و همکاران (۱۹۷۴)، چین (۱۹۹۸)، استون و گیزر (۱۹۷۵) و هنسler و همکاران ۳ از بعد مدل ساختاری و آزمون‌های مرتبط و مدل اندازه‌گیری و آزمون‌های مرتبط قابل مشاهده است. از اینرو در این مجال به آنها پرداخته نمی‌شود.

۸- یافته‌های پژوهش

در روش PLS پایایی مدل از طریق پایایی و روایی در دو بخش مدل‌های اندازه‌گیری و مدل‌های ساختاری بررسی می‌شود. برای بررسی برازش بخش اول سه مورد پایایی شاخص و روایی همگرا استفاده شده است که پایایی شاخص خود از طریق سه معیار بار عاملی، ضریب آلفای کرونباخ و

جدول ۳: ضرایب بارهای عاملی، آلفای کرونباخ، پایایی ترکیبی و روایی همگرا

عوامل	سازه‌ها	بار عاملی	ضریب آلفای کرونباخ	پایایی ترکیبی	روایی همگرا
IT	RA	۰/۷۱۵	۰/۸۹۸	۰/۹۲۲	۰/۶۵۵
	CO	۰/۶۱۵	۰/۷۷۱	۰/۸۶۸	۰/۶۷۸
	COMP	۰/۷۲۳	۰/۸۵۵	۰/۹۱۰	۰/۷۲۲
	COMPAT	۰/۸۲۹	۰/۹۰۱	۰/۹۳۳	۰/۷۷۷
	TRA	۰/۷۹۱	۰/۹۰۴	۰/۹۳۳	۰/۷۷۷
ORG	TMS	۰/۷۸۷	۰/۸۸۶	۰/۸۴۵	۰/۵۶۲
	IT EX	۰/۸۸۸	۰/۹۲۲	۰/۸۸۰	۰/۷۵۲
	FORM	۰/۸۴۱	۰/۹۰۵	۰/۸۶۱	۰/۷۰۵
	SIZE	۰/۶۱۲	۰/۹۳۳	۰/۸۹۷	۰/۸۲۴
	RDD ORG	۰/۹۴۵	۰/۹۵۱	۰/۹۴۴	۰/۵۷۰
ENV	ENV	۰/۹۳۴	۰/۹۰۲	۰/۸۵۵	۰/۷۰۲
CEO	AT CEO	۰/۹۳۱	۰/۸۸۶	۰/۹۱۷	۰/۶۸۹
	INN CEO	۰/۹۳۷	۰/۹۱۰	۰/۹۳۴	۰/۷۳۹
	IT K	۰/۷۸۷	۰/۹۴۳	۰/۹۶۳	۰/۸۹۷
	TEN MAN	۰/۶۴۶	۰/۹۶۴	۰/۹۶۷	۰/۹۷۵
IND	SN	۰/۸۹۸	۰/۹۵۹	۰/۹۶۹	۰/۵۹۳
	IM	۰/۸۷۰	۰/۸۷۱	۰/۹۲۰	۰/۶۹۵

باشد نشان از صحت رابطه بین سازه‌ها و در نتیجه تأیید فرضیه‌های پژوهش در سطح اطمینان ۹۵٪ است. با توجه به مقادیر محاسبه‌شده برای ضرایب معناداری Z که در جدول (۴) نشان داده شده است، می‌توان صحت رابطه بین تمام سازه‌ها و شاخص‌هایشان را در سطح اطمینان ۹۵٪ تأیید نمود. بر اساس محاسبات انجام‌شده در جدول (۴) تمام مسیرهای مدل به جز ENV به ORG ADP و CEO به ORG ADP را می‌توان تأیید نمود. بنابراین بر اساس مقادیر محاسبه‌شده برای ضرایب معناداری Z بایستی این دو مسیر را از مدل نهایی حذف نمود. بنابراین عوامل محیطی از مدل اصلی تحقیق حذف شده‌اند.

بنابراین با توجه به مقادیر محاسبه‌شده برای بارهای عاملی، ضرایب آلفای کرونباخ، پایایی ترکیبی و روایی همگرا می‌توان ادعا نمود که مدل تحقیق از پایایی مناسبی برخوردار است. برای برآزش بخش ساختاری مدل از معیارهای ضریب معناداری Z (t-values)، معیار R2، معیار Q2، معیار Redundancy استفاده شده است و در نهایت برای بررسی برآزش کلی مدل تحقیق از معیار GOF استفاده شده است. در جدول (۴) مقادیر ضریب معناداری Z (t-values) محاسبه شده است. مقدار ملاک برای ضرایب معناداری Z کمتر از ۱/۹۶ می‌باشد. در صورتی که این مقدار بیشتر ۱/۹۶

جدول ۴: مقادیر محاسبه شده برای ضرایب معناداری Z (t-values)

عوامل	سازه‌ها	مقادیر t-values	مسیر	مقادیر t-values	نتیجه
IT	RA	۴/۰۶۴	← IT	IND ACC	تأیید
	CO	۵/۱۸۶	← ORG	ORG ADP	تأیید
	COMP	۹/۱۵۱	← CEO	IND ACC	تأیید
	COMPAT	۲۰/۳۶۴	← CEO	ORG ADP	رد
	TRA	۱۰/۳۲۰	← IND	IND ACC	تأیید
ORG	TMS	۱۰/۰۱۰	← ENV	ORG ADP	رد
	IT EX	۲۸/۱۰۲	← IND ACC	ORG ADP	تأیید
	FORM	۱۵/۰۲۳			
	SIZE	۶/۴۲۳			
CEO	RDD ORG	۵۲/۸۲۶			
	AT CEO	۶۱/۴۳۵			
	INN CEO	۴۴/۰۲۴			
	IT K	۱۱/۲۷۱			
IND	TEN MAN	۱۰/۱۱۹			
	SN	۲۳/۱۱۱			
	IM	۲۸/۷۹۰			

این رابطه را با اطمینان ۹۵٪ می‌توان تأیید نمود. ضریب معناداری Z برای مسیر مربوط به عوامل مرتبط با مدیر عالی اجرایی با پذیرش سازمانی برابر ۱/۰۲۴ می‌باشد که نمی‌توان صحت این رابطه را مورد تأیید قرارداد. ضریب معناداری Z مربوط به مسیر عوامل فردی با پذیرش فردی برابر ۲/۷۶۱ می‌باشد. بنابراین صحت این رابطه را با اطمینان ۹۵٪ مورد تأیید قرارداد. ضریب معناداری Z مسیر مربوط به پذیرش فردی به پذیرش سازمانی برابر با ۲/۲۱۵ می‌باشد. لذا می‌توان صحت این رابطه را با اطمینان ۹۵٪ مورد تأیید قرارداد. گام بعدی برای بررسی برازش مدل ساختاری محاسبه معیار R2، معیار Q2، معیار Redundancy است. در جدول (۵) مقادیر معیارهای R2، Q2 و Redundancy محاسبه شده است.

ضرایب معناداری Z را برای مدل کلی نشان می‌دهد. ضریب مربوط به مسیر عوامل مرتبط با فناوری اطلاعات و پذیرش فردی برابر ۲/۴۸۰ می‌باشد. بنابراین معنادار بودن رابطه بین سازه‌های فناوری اطلاعات و پذیرش فردی با ضریب اطمینان ۹۵٪ مورد تأیید می‌باشد. ضریب مربوط به مسیر عوامل سازمانی و پذیرش سازمانی برابر ۳/۷۲۶ می‌باشد که در نتیجه با ضریب اطمینان ۹۵٪ می‌توان معنادار بودن این رابطه را تأیید کرد. ضریب مربوط به مسیر عوامل محیطی و پذیرش سازمانی برابر ۱/۰۸۴ می‌باشد لذا صحت این رابطه را نمی‌توان تأیید نمود. ضریب مربوط به مسیر عوامل مرتبط با مدیر عالی اجرای با پذیرش فردی برابر ۲/۴۸۸ می‌باشد. بنابراین صحت

جدول ۵: محاسبه مقادیر معیار های Q^2 , R^2 , Redundancy

عوامل	سازه‌ها	معیار R^2	معیار Q^2	معیار Redundancy
IT	RA	۰/۴۹۹	۰/۳۲۳	۰/۳۳۲
	CO	۰/۳۸۵	۰/۲۲۰	۰/۲۶۶
	COMP	۰/۵۲۱	۰/۳۷۵	۰/۳۷۶
	COMPAT	۰/۶۹۱	۰/۴۹۱	۰/۵۳۷
	TRA	۰/۶۳۱	۰/۴۶۸	۰/۴۹۰
ORG	TMS	۰/۶۱۶	۰/۲۶۱	۰/۳۴۸
	IT EX	۰/۷۹۲	۰/۵۷۵	۰/۵۹۵
	FORM	۰/۷۰۶	۰/۴۶۵	۰/۴۹۷
	SIZE	۰/۳۸۲	۰/۲۷۶	۰/۳۱۴
	RDD ORG	۰/۹۰۸	۰/۴۹۳	۰/۵۱۷
CEO	AT CEO	۰/۹۳۰	۰/۵۷۸	۰/۶۴۰
	INN CEO	۰/۹۳۸	۰/۶۴۲	۰/۶۵۱
	IT K	۰/۷۵۳	۰/۴۸۶	۰/۵۰۸
	TEN MAN	۰/۷۴۱	۰/۵۲۶	۰/۵۳۵
IND	SN	۰/۸۰۱	۰/۲۴۸	۰/۳۱۶
	IM	۰/۸۷۳	۰/۴۱۹	۰/۴۲۵
IND ACC	-	۰/۵۶۸	۰/۱۱۰	۰/۱۸۵
ORG ADP	-	۰/۸۲۲	۰/۱۸۴	۰/۲۴۰

هر چه مقدار R^2 مربوط به سازه‌های درونی یک مدل بیشتر باشد، نشان از برازش بهتر مدل است. چین^۱ (۱۹۹۸) سه مقدار ۰/۱۹، ۰/۳۳ و ۰/۶۷ را به‌عنوان ملاک برای مقادیر ضعیف، متوسط و قوی R^2 معرفی می‌کند. در جدول (۵) مقادیر R^2 برای سازه‌های درون‌زای مدل محاسبه شده‌اند. همان‌طور که مشاهده می‌شود کلیه مقادیر محاسبه شده برای سازه‌های دور نزا با توجه به مقدارهای ملاک، مقادیر مناسبی بوده و لذا می‌توان مناسب بودن برازش مدل ساختاری را تأیید نمود. معیار Q^2 قدرت پیش‌بینی مدل را مشخص می‌کند. هنسeler و همکاران^۲ (۲۰۰۹) در مورد شدت قدرت پیش‌بینی مدل در مورد سازه‌های درون‌زا سه مقدار ۰/۰۲، ۰/۱۵ و ۰/۳۵ را تعیین نموده‌اند. به اعتقاد آن‌ها اگر مقدار Q^2 در مورد یک سازه درون‌زا نزدیک به ۰/۰۲ باشد، نشان از آن دارد که مدل قدرت پیش‌بینی ضعیفی در قبال شاخص‌های آن سازه دارد. با توجه به مقادیر محاسبه شده در جدول (۵) می‌توان ادعا نمود که این مدل از قدرت پیش‌بینی

نسبتاً بالایی برخوردار است. در نتیجه می‌توان ادعا نمود که مدل از برازش مناسبی برخوردار است. معیار Redundancy حاصل ضرب مقادیر اشتراکی (Communality) سازه‌ها در مقادیر R^2 مربوط به آن‌ها به دست می‌آید و نشانگر مقدار تغییرپذیری شاخص‌های یک سازه درون‌زا است که از یک یا چندسازه برون‌زا تأثیر می‌پذیرد.

$$Red = Communality \times R^2 \quad (1)$$

با توجه به فرمول فوق مقادیر معیار Redundancy را برای سازه‌های درون‌زا مدل محاسبه شده است و در جدول (۵) نشان داده شده است و در نهایت مقدار \overline{Red} که برابر با مقدار میانگین معیارهای Redundancy محاسبه شده برای سازه‌های درون‌زا است، محاسبه شده است. هر چه این مقدار بیشتر باشد مدل از برازش مناسب‌تری برخوردار است. مقدار \overline{Red} برای مدل برابر است با:

$$\overline{Red} = ۰/۴۳۳$$

1.Chin
Henseler et al.¹

البته از منظر سایر مولفه های سازه فناوری اطلاعات با [۱۴] مطابقت دارد. سازه‌های عوامل سازمانی عبارت‌اند از حمایت مدیریت ارشد، تخصص فناوری اطلاعات، رسمیت، تمرکز، اندازه، عمق تغییر و آمادگی سازمانی؛ که سازه‌های تمرکز و عمق تغییر از مدل حذف‌شده‌اند. سازه عمق تغییر برای اولین بار در مدل ITAIM توسط علی یاری ارایه گردید که در سازمان های دولتی ایران پس از ارزیابی مورد پذیرش قرار گرفته بود. عمق تغییر حدی است که فرآیندها و ساختار سازمان با پذیرش و پیاده‌سازی فناوری اطلاعات دستخوش تغییر می‌شود، عمق تغییر بر پذیرش فناوری اطلاعات تأثیر منفی دارد. بر اساس یافته های تحقیق حاضر عمق تغییر به دلیل وجود بسترهای لازم در مراکز تحقیقاتی دانشگاهی تاثیر قابل ملاحظه ای در پذیرش ندارد. بنابراین از مدل نهایی حذف گردید. تمرکز نیز میزانی است که تصمیم‌گیری در یک نقطه واحد سازمانی متمرکز شده است. نتایج بدست آمده این تحقیق در رابطه با عوامل سازمانی با نتایج پژوهش [۴۳]، [۲۹]، [۴۴]، [۱۹] و [۷] مطابقت دارد. در این تحقیق تمرکز تصمیم‌گیری در یک نقطه تأثیر قابل توجهی بر پذیرش فناوری اطلاعات از دیدگاه کارکنان مراکز ندارد. می‌توان چنین استنباط نمود که تغییرات به وجود آمده در سازمان از دیدگاه کارکنان تأثیر قابل توجهی بر پذیرش و به‌کارگیری فناوری اطلاعات ندارد. ضریب معناداری Z برای سازه‌ی عوامل محیطی کمتر از مقدار بحرانی بوده و این نشان‌دهنده عدم تأیید مسیر بین عوامل محیطی و پذیرش سازمانی است. بنابراین می‌توان نتیجه‌گیری نمود که عوامل محیطی یا بیرونی بر پذیرش فناوری اطلاعات در مراکز تحقیقاتی دانشگاهی تأثیر چندانی ندارد. سازه عوامل محیطی از مدل نهایی تحقیق حذف‌شده است این مطلب نشان‌دهنده تأثیر اندک عوامل بیرونی از جمله سیاست و حمایت دولت در جهت به‌کارگیری فناوری اطلاعات می‌باشد. دسته عوامل مرتبط با مدیر عالی اجرایی شامل سازه‌های نگرش مدیر عالی اجرایی، نوآوری مدیر عالی اجرایی، دانش فناوری اطلاعات مدیر عالی اجرایی و تصدی مدیریتی مدیر عالی اجرایی همگی در مدل نهایی تحقیق بیان شده‌اند. یک مدیر باتجربه، ریسک‌پذیر، نوآور و دارای اطلاعات کافی در مورد فناوری اطلاعات در رأس یک مرکز دارای ساختار سلسله مراتب، می‌تواند نقش

و در نهایت معیار GOF مربوط به بخش کلی مدل‌های معادلات ساختاری است. بدین معنی که توسط این معیار، محقق می‌تواند پس از بررسی برازش بخش اندازه‌گیری و بخش ساختاری مدل کلی پژوهش خود، برازش بخش کلی را نیز کنترل کند. معیار GOF توسط تننهاوس و همکاران (۲۰۰۴) ابداع گردید و طبق فرمول زیر محاسبه می‌شود:

$$GOF = \sqrt{\overline{Communalities}} * R \text{ Sequare} \quad (2)$$

به‌طوری‌که

$\overline{Communalities}$ نشانه میانگین مقادیر اشتراکی هر سازه می‌باشد و $R \text{ Sequare}$ نیز مقدار میانگین مقادیر R^2 سازه‌های درون‌زای مدل است. که با توجه به محاسبات انجام‌شده مقدار $\overline{Communalities}$ برابر 0.663 و مقدار $R \text{ Sequare}$ برابر 0.624 می‌باشد. در نتیجه:

$$0.1643 = \sqrt{0.663 \times 0.624}$$

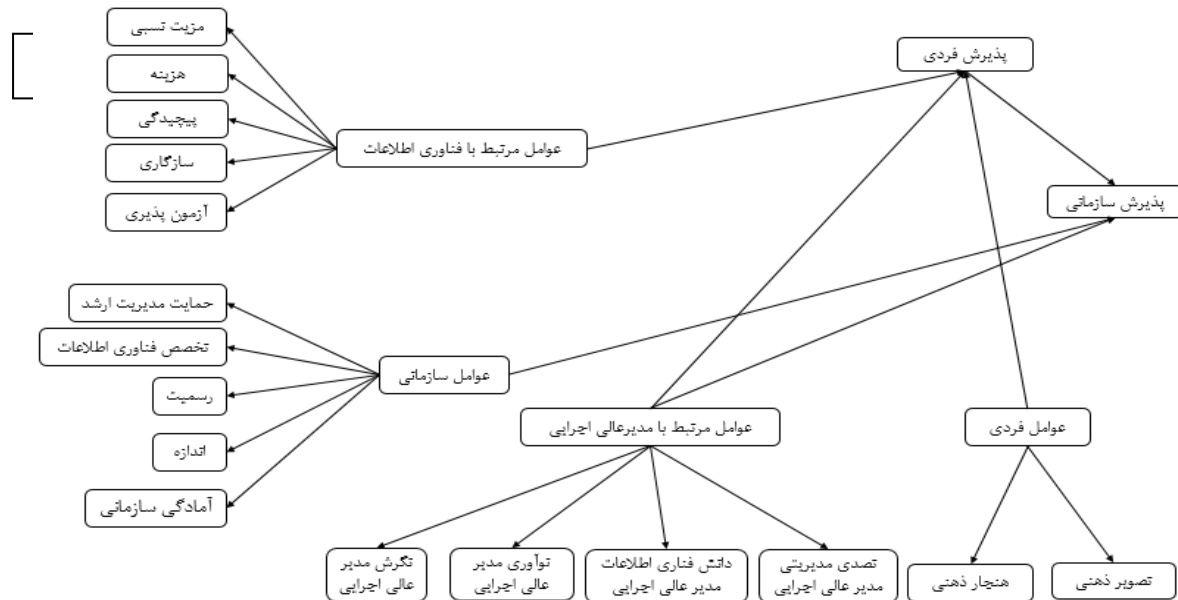
و با توجه به سه مقدار 0.1/0.1، 0.1/25 و 0.1/36 که به‌عنوان مقادیر ضعیف، متوسط و قوی برای معیار GOF معرفی شده است و حصول مقدار 0.1643، می‌توان برازش بسیار مناسب مدل کلی را تأیید نمود.

۹- نتیجه‌گیری و پیشنهادها

مدل اصلی تحقیق شامل پنج دسته عوامل مستقل که عبارتند از عوامل مرتبط با فناوری اطلاعات، عوامل سازمانی، عوامل محیطی، عوامل مرتبط با مدیریت عالی اجرایی و عوامل فردی و شامل دو متغیر وابسته پذیرش فردی و پذیرش سازمانی می‌باشد. سازه‌های مرتبط با فناوری اطلاعات عبارت‌اند از مزیت نسبی، هزینه، پیچیدگی، سازگاری، مشاهده‌پذیری و آزمون‌پذیری؛ که سازه مشاهده‌پذیری با توجه به حدنصاب نرسیدن ضریب بار عاملی جهت مناسب شدن پایایی مدل از مدل نهایی حذف‌شده است. این مطلب نشان‌دهنده این است که ملموس بودن نتایج حاصل از به‌کارگیری فناوری اطلاعات تأثیر قابل توجهی در پروژه های ICT مراکز تحقیقاتی دانشگاهی بر پذیرش فناوری اطلاعات ندارد. نتایج این بخش از تحقیق با نتایج پژوهش باکاندا و [۴۴]، [۴۵]، [۲۱] مطابقت دارد و با نتایج پژوهش [۱۴] صرفاً با توجه به حذف مولفه مشاهده‌پذیری مطابقت ندارد.

گردیده‌اند. نتایج به‌دست‌آمده نشان می‌دهد که عوامل محیطی از جمله حمایت دولت از بکار گیری فناوری اطلاعات در مراکز سازمانی و نظامی با توجه به ویژگی‌های این مراکز تأثیر معناداری بر روی پذیرش فناوری اطلاعات نداشته و همچنین مدیریت عالی مراکز سازمانی به‌صورت مستقیم اثرگذاری قابل توجهی بر پذیرش فناوری اطلاعات در سازمان مطبوع خویش ندارد. مدل نهایی این پژوهش جهت پذیرش فناوری اطلاعات در مراکز تحقیقاتی فناوری اطلاعات دانشگاهی بر اساس مدل پذیرش فناوری اطلاعات ITAIM بعد از انجام اصلاحات و تغییرات لازم در شکل (۲) نمایش داده شده است.

بسیاری در پذیرش و به‌کارگیری فناوری اطلاعات و پیشبرد اهداف سازمان داشته باشد. نتایج به دست آمده از این بخش تحقیق حاضر با نتایج پژوهش [۳۱]، [۱۹]، [۱۴] و [۱۵] همخوانی دارد. عوامل فردی شامل دو سازه هنجار ذهنی و تصویر ذهنی می‌باشند. این نشان‌دهنده این است که افراد سازمان از یکدیگر تأثیر پذیرفته و بر این باورند که فناوری اطلاعات منجر به ترفیع جایگاه و ارتقای منزلت آن‌ها در سازمان می‌شود. نتایج به دست آمده این تحقیق در مورد سازه عوامل فردی، با نتایج پژوهش [۲۸]، [۲۰]، [۵] و [۲۳] مطابقت دارد. با توجه به تجزیه و تحلیل‌های انجام‌شده و نتایج به‌دست‌آمده، برخی از مسیرهایی که معناداری آن‌ها مورد تأیید واقع نشده است و از مدل نهایی پژوهش حذف



شکل ۲: مدل نهایی تحقیق

سازمانی (ORG) شامل متغیرهای وابسته حمایت مدیریت عالی (TMS)، تخصص فناوری اطلاعات (IT EX)، رسمیت (FORM)، اندازه (SIZE) و آمادگی سازمانی (RED) ORG هست. بنابراین این سازه‌ها بر پذیرش سازمانی تأثیر مثبت و معناداری دارند. عوامل مرتبط با مدیر عالی اجرایی (CEO) شامل متغیرهای وابسته نگرش مدیریت عالی اجرایی (AT CEO)، نوآوری مدیر عالی اجرایی (INN)

این مدل شامل چهار دسته متغیرهای مستقل عوامل مرتبط با فناوری اطلاعات (IT)، عوامل سازمانی (ORG)، عوامل مرتبط با مدیر عالی اجرایی (CEO) و عوامل فردی (IND) و متغیرهای وابسته پذیرش فردی (IND ACC) و پذیرش سازمانی (ORG ADP) می‌باشد. عوامل مرتبط با فناوری اطلاعات شامل متغیرهای وابسته مزیت نسبی (RA)، هزینه (CO)، پیچیدگی (COMP)، سازگاری (COMPAT) و آزمون‌پذیری (TRA) می‌باشد. به این معنا که سازه‌ها بر پذیرش فردی فناوری اطلاعات اثرگذار هستند. عوامل

۸. داور پناه، محمدرضا، (۱۳۸۱)، موانع زیر ساختی بهره‌گیری از فناوری اطلاعات در کتابخانه‌های دانشگاهی ایران، کتابداری و اطلاع رسانی، دوره ۲، شماره ۵، صفحه ۱ تا ۲۳

۹. داوری، علی، (۱۳۹۲)، مدل‌سازی معادلات ساختاری با نرم‌افزار PLS، تهران: سازمان انتشارات جهاد دانشگاهی

۱۰. درودی، فریبرز، (۱۳۸۹)، برنامه‌ریزی راهبردی فناوری اطلاعات، کتاب ماه علوم و فنون، اردیبهشت ۱۳۸۹، شماره ۱۲۵ از ۲۰ تا ۳۵

۱۱. رضایی، مسعود، (۱۳۸۸)، نظریه‌های رایج درباره‌ی پذیرش فناوری اطلاعات و ارتباطات، پژوهش‌های ارتباطی (پژوهش و سنجش)، دوره ۱۶، شماره ۴، زمستان ۱۳۸۸

۱۲. رستمی، مسلم، (۱۳۹۱)، نقش عوامل موثر بر پذیرش و بکارگیری فناوری اطلاعات و ارتباطات در بین جوانان روستایی بر اساس نظریه نشر نوآوری راجرز- مطالعه موردی: روستاهای اورامانات استان کرمانشاه، کتابداری و اطلاع رسانی، دوره ۱۶، شماره ۲، از صفحه ۱۱۹ تا ۱۳۸

۱۳. علی یاری، شهرام، (۱۳۹۴)، نقش عوامل فردی و سازمانی و مدیریتی مؤثر بر پذیرش فناوری اطلاعات در سازمان‌های دولتی ایران، فصل‌نامه پژوهش‌های مدیریت منابع انسانی، سال هفتم، شماره ۳۰

۱۴. علی یاری، شهرام، (۱۳۹۳)، ارائه مدل پذیرش فناوری اطلاعات در سازمان‌های دولتی ایران، رساله دکترا، گروه مهندسی صنایع دانشکده فنی و مهندسی دانشگاه جامع امام حسین (ع)، ۱۳۹۳

۱۵. فهامی، ریحانه، زارع، حسین، (۱۳۹۲)، عوامل مؤثر بر پذیرش فناوری های جدید در آموزش از راه دور با استفاده از مدل پذیرش فناوری (مطالعه مورد دانشگاه پیام نور اصفهان)، رهیافتی نو در مدیریت، دوره ۴، شماره ۱، از ۶۷ تا ۷۹

۱۶. قلی پور، رحمت‌الله، (۱۳۸۳)، تأثیر فناوری اطلاعات بر ساختار سازمانی و نیروی کار، مدیریت فرهنگ سازمانی، شماره ۷ از ۱۲۷ تا ۱۲۵

۱۷. کیهان، جواد، محمدمرادی نقده، سارا، (۱۳۷۹)، بررسی عوامل مؤثر بر استفاده و پذیرش فناوری رایانه توسط معلمان در امر تدریس با استفاده از نقشه‌های شناختی فازی، پژوهش در نظام‌های آموزشی، شماره ۴۳ از ۲۳۳ تا ۲۴۹

CEO). دانش فناوری اطلاعات مدیر عالی اجرایی (IT K)

و تصدی

مدیریتی مدیر عالی اجرایی (TEN MAN) می‌باشد و در نهایت عوامل فردی (IND) شامل دو متغیر وابسته هنجار ذهنی (SN) و تصویر ذهنی (IM) می‌باشد

۱۰- منابع

۱. باقری، محمد، (۱۳۸۸)، پذیرش بانکداری اینترنتی در ایران: بسط مدل پذیرش فناوری، فصل‌نامه علوم فناوری اطلاعات، دوره ۲۴، شماره ۱۳، بهار ۱۳۸۸، ص ۳۴-۵

۲. بدیع، علی، دستجردی، علی (۱۳۹۵)، ارزیابی یک مدل مفهومی پذیرش فناوری در بانکداری الکترونیکی بر اساس مدل‌های TRA, TAM, TAMII, TPM همایش بین‌المللی تجارت و اقتصاد الکترونیکی

۳. جهانگیر، غلام؛ دیانی، محمدحسین؛ نو کاریزی، محسن؛ توسعه مدل پذیرش فناوری اطلاعات دیویس (TAM) از طریق سنجش تأثیر باورهای خودکار آمد اعضای خودکارآمدی دانشگاه علوم پزشکی مشهد بر پذیرش سامانه‌های اطلاعاتی، پژوهش‌نامه کتابداری و اطلاع‌رسانی، ۱۳۹۴، (۲)۵، ۳۱۹-۳۳۹

۴. حسینی، نگین (۱۳۹۲)، پذیرش و استفاده از پایگاه‌های اطلاعاتی پیوسته لاتین توسط اعضای هیئت‌علمی دانشگاه‌های شهر کرمانشاه بر اساس نظریه یکپارچه پذیرش و استفاده از فناوری UTAUT، فصل‌نامه دانش سیاسی، سال هفتم، شماره ۲۷

۵. حسین پور، جعفر، (۱۳۸۶)، نقش فناوری اطلاعات در تحول ساختار سازمان‌ها، اطلاعات سیاسی اقتصادی، شماره ۲۳۷ از ۱۸۲ تا ۱۹۵، خرداد و تیر ۱۳۸۶

۶. حقی‌شناس، اصغر، دلوی، محمدرضا، شفیعی، مسعود (۱۳۸۶)؛ نقش سرمایه‌های اجتماعی در توسعه، تدبیر، دی ماه، شماره ۱۸۸

۷. خلعتبری، احمد، (۱۳۹۰)، ارائه یک مدل جدید برای پذیرش سرویس‌های دولت الکترونیک توسط کاربران، پایان‌نامه دانشگاه خواجه‌نصیرالدین طوسی، دانشکده مهندسی صنایع، ۱۳۹۰

- model. *International journal of medical informatics*, 90, 22-31.
- 28.Chang, K.M & Cheung, W. (2001). Determinants of intention to use internet/ www.at work: a confirmatory study. *Information and management*, 39(1), 1-14
- 29.Chen,C.D; Fan, Y.W &Farn, C.K (2007) Predicting Electronic toll collection service Adoption: An Integration of the technology Acceptance model and theory of planned Behavior. *Transportation Research, part c*, 15, 300-311
- 30.Chen, K., & Chan, A. H. S. (2014). Gerontechnology acceptance by elderly Hong Kong Chinese: a senior technology acceptance model (STAM). *Ergonomics*, 57(5), 635-652.
- 31.Damanpour, F. (1991). Organizational Innovation: A Meta-analysis of Effects of Determinants and Moderators. *Academy of Management Journal*, 34 (3), pp. 555-590
- 32.Dickman Portz, J. et al . (2019) ,Using the Technology Acceptance Model to Explore User Experience, Intent to Use, and Use Behavior of a Patient Portal Among Older Adults With Multiple Chronic Conditions: Descriptive Qualitative Study. *Journal of Medical Internet Research*. Vol 21. , 1-12
- 33.Fischer, S. H., David, D., Crotty, B. H., Dierks, M., & Safran, C. (2014). Acceptance and use of health information technology by community-dwelling elders. *International journal of medical informatics*, 83(9), 624-635.
- 34.Farzandipour,M. Nadi-Ravandi,S. Gilasi,H. Soleimani.N. (2019). Iranian Health Information Technology Acceptance Model (IHITAM) from Users' Views. *ACTA Inform MED*. 27(4): 245-252
- 35.Heltzol, pual. (2019). IT manager's survival guide: 11 ways to thrive in the years ahead. CIO publisher.
- 36.Handayani, P. W., Hidayanto, A. N., Pinem, A. A., Hapsari, I. C., Sandhyaduhita, P. I., & Budi, I. (2017). Acceptance model of a Hospital Information System. *International journal of medical informatics*, 99, 11-28.
۱۸. کاهویی، مهدی (۱۳۹۲)، عوامل موثر بر پذیرش فناوری اطلاعات در محیط بالینی از دیدگاه پرستاران، پی‌اورد سلامت، دوره ۷، شماره ۴، از صفحه ۲۶۲ تا صفحه ۲۷۷
۱۹. مشایخی، علینقی و همکاران، بررسی عوامل کلیدی موثر بر کاربرد فناوری اطلاعات در سازمان های دولتی ایران: کاربرد روش دلفی، مجله مدرس علوم انسانی، دوره ۹، شماره ۳، پاییز ۱۳۸۴
۲۰. مانیان، امیر و همکاران، (۱۳۸۶)، استفاده از مدل معادلات ساختاری در ارائه مدلی برای موفقیت برنامه‌ریزی استراتژیک فناوری اطلاعات، فصلنامه دانش مدیریت دانشگاه تهران، شماره ۷۴، ۱۳۸۶، ص ۱۱۷ تا ۱۳۸
۲۱. محمدی، علی (۱۳۹۲)، شناسایی و تبیین عوامل مؤثر بر پذیرش نوآوری فناوری اطلاعات در سازمان‌های دولتی با رویکرد معادلات ساختاری، مدیریت فناوری اطلاعات، دوره ۵، شماره ۴، از صفحه ۱۹۵ تا ۲۱۸
۲۲. محترمی، امیر، خداداد حسینی، سید حمید، الهی، شعبان (۱۳۹۲)، بررسی عوامل مؤثر بر فناوری های اطلاعاتی در سازمان ها، فصلنامه مدیریت توسعه فناوری، شماره ۳، ص ۹۷ تا ص ۱۲۲
۲۳. نادری، بنی، محمود و همکاران (۱۳۹۴)، عوامل مؤثر بر پذیرش فناوری اطلاعات در هتل های شهر شیراز، مطالعات مدیریت گردشگری، سال نهم، شماره ۲۹
- 24.Anderson J. (2007). Social, ethical and legal barriers to e-health. *International Journal of Medical Informatics*; 76(5): 480-3
- 25.Akbari, M., & Alipour Pijani, A. (2013). ICT Adoption: A Case Study of SMEs in Tehran (Iran). *The International Journal of Humanities*, 20(3), 93-121.
- 26.Beglaryan, M., Petrosyan, V., & Bunker, E. (2017). Development of a tripolar model of technology acceptance: Hospital-based physicians' perspective on EHR. *International Journal of Medical Informatics*, 102, 50-61.
- 27.Cimperman, M., Brenčič, M. M., & Trkman, P. (2016). Analyzing older users' home telehealth services acceptance behavior—applying an Extended UTAUT

- for green information technology acceptance. *Computers in human behavior*, 36, 29-40.
46. Oreg S. Personality, (2006) context ,and resistance to organizational change. *European Journal of Work and Organizational Psychology*; 15(1): 73-101.
47. Pan, C.C., & Brophy, J. (2003). Students attitude equation modeling inquiry. *Jornal of Educational media and Library* 41(2), 181-194
48. Rogers, Everetm. ; Chafee, Steven, H. (1983). Communication as an Academic Discipline: A Dialogue. *Journal of Communication* Vol 33 no 3, pp: 18-30
49. Sun, H. & Zhang, P. (2006). The role of moderating factors in user technology acceptance. *Human- Computer studies*, 64: 53-78.
50. Schaper, Louise K., and Graham P. Pervan. (2007). " ICT and OTC: A model of Information and Communication technology acceptance and utilization by occupational therapists." *International journal of medical informatics* 76: s 212-s221.
51. Sindhu Yoga, I. Dyah Permatha Korry. N, Dhian Rani Yulianti. N. (2019). "Information Technology Adoption on Digital Marketing Communication Channel" *International Journal of Social Sciences and Humanities* Vol. 3 No. 2. pages: 95~104
52. Tsiknakis M, Kouroubali A. (2009) Organizational factors affecting successful adoption of innovative Health services: A case study employing the FITT framework. *International Journal of Medical Informatics*; 78(1): 39-52.
53. Turban .E. , Linder .D., Mclen E., Wetherbe. J., (2002) "Information technology for management " 6 THED , 720 pages
54. Thong, J. Y. L., and Yap, C.S., (1995). " CEO characteristics, organizational characteristics and information technology adoption in the small business. , *omega* 23(4), 429-442
55. Tung, F.C. Chang S.C. & Chou, C.M. (2008). An Extension of Trust and TAM
37. Hoque, R., & Sorwar, G. (2017). Understanding factors influencing the adoption of mHealth by the elderly: An extension of the UTAUT model. *International Journal of Medical Informatics*, 101, 75-84.
38. Hsieh, P. J. (2015). Physicians' acceptance of electronic medical records exchange: An extension of the decomposed TPB model with institutional trust and perceived risk. *International journal of medical informatics*, 84(1), 1-14.
39. IT Project management and reporting Guidelines, October, (2018). Information technology services university of California office of the president.
40. Jack T. Marchewka. (2015). *Information Technology Project Management: Providing Measurable Organizational Value*. 5th Edition: Wiley Publishing
41. King, W.R., He, J (2006). " A meta-analysis of the Teachnology acceptance model". *Information & management* vol, 43, pp : 740-755
42. Kwon, T.H. & Zumad, R.W. (1987) .Unifying the Fragmented Models of Information systems Implementation. In Boland, J.R. and Hirschheim R. (edition). *Critical Issues in information systems Research*. New York, Wiley.
43. Koh. Ch. Prybutok, V.R., Ryan. Sh & Ibragimova, B .(2006). The impotence of strategic in an emerging e-government environment. *Business Process Management Jornal*, 12(1):22-33
44. Kanda. M.J & Ogollah Kennedy, (2013), " Factors influencing adoption and use of information and communication technology at the ethics and anticorruption of Kenya". *Journal of Business Administration and Management Sciences Research* Vol. 2(11), pp. 224-309
45. Lee, J.D. & Jongsu Lee, J. (2009). E-government adoption in ASEAN: the case of Cambodia Sinawong Sang. *Internet Research*, 19(5): 517-534
46. Mishra, D., Akman, I., & Mishra, A. (2014). Theory of reasoned action application

studies. Management science, 46(2), 186-204

57.Venkatesh, V. Bala. H., (2008), " technology acceptance model3 and research agenda on interventions". Decision Science Volume 39, Number2

Model with IDT in the Adoption of the electronic logistics information system in HIS in the medical Industry. In.J. Medical informatics, 77. 324-335

56.Venketash, V. and Davis, F.D. (2006). A theoretical extension of the technology acceptance model: four longitudinal field

ترکیب دوگانه سیستم استنتاج فازی با الگوریتم بهینه‌سازی ازدحام ذرات در پیش‌بینی قیمت سهام بورس اوراق بهادار تهران و مقایسه آن با مدل یادگیری عمیق

*مجید عبدالرزاق نژاد **مهدی خرد

*استادیار، گروه مهندسی کامپیوتر، دانشکده مهندسی - دانشگاه بزرگمهر قائنات - قائن - ایران

**کارشناس ارشد، دانشکده مهندسی برق و کامپیوتر - دانشگاه بیرجند - بیرجند - ایران

تاریخ پذیرش: ۱۳۹۹/۰۵/۰۷

تاریخ دریافت: ۱۳۹۸/۱۱/۱۵

چکیده

پیش‌بینی قیمت سهام توسط تحلیلگران داده یک فرصت تجاری بزرگ را برای طیف گسترده سرمایه‌گذاران در بازار سهام ایجاد کرده است. اما این مهم به دلیل ماهیت بی‌ثبات و پویایی بیش از حد عوامل متعدد اقتصادی تاثیرگذار بر بازار سهام، امری دشوار است. در این پژوهش به منظور شناسایی ارتباط پیچیده ۱۰ متغیر اقتصادی بر قیمت سهام شرکت‌های فعال در بازار سهام تهران، دو مدل طراحی و پیاده‌سازی شده است. نخست یک سیستم استنتاج فازی ممدانی که مجموعه قوانین موتور استنتاج خود را توسط الگوریتم بهینه‌سازی ازدحام ذرات بدست می‌آورد طراحی می‌شود. سپس مدل یادگیری عمیق مشتمل بر ۲۶ نرون در ۵ لایه پنهان طراحی شده است. مدل‌های طراحی شده به منظور پیش‌بینی قیمت سهام نه شرکت فعال در بورس اوراق بهادار تهران پیاده‌سازی و نتایج بدست آمده حاکی از عملکرد بهتر مدل یادگیری عمیق بر مدل ترکیب دوگانه استنتاج فازی-ازدحام ذرات و نیز مدل رایج شبکه عصبی دارد. اما قدرت تفسیرپذیری الگوی بدست آمده، رفتار همسانتر و با واریانس به مراتب کمتر و نیز سرعت همگرایی بیشتری نسبت به سایر مدل‌ها را می‌توان از مزایای رقابتی قابل توجه مدل ترکیب دوگانه استنتاج فازی-ازدحام ذرات نام برد.

واژه‌های کلیدی: پیش‌بینی قیمت سهام، سیستم استنتاج فازی، یادگیری عمیق، شبکه عصبی و الگوریتم بهینه‌سازی ازدحام ذرات.

۱- مقدمه

باتوجه به حجم سرمایه در گردش از طریق بازارهای بورس سراسر جهان، پیش‌بینی قیمت سهام بورس موضوعی بوده است که هم تحلیلگران و هم محققان از مدت‌ها قبل به آن علاقه داشتند [۱]. از آنجا که طیف گسترده‌ای از عوامل مؤثر بر شاخص بورس اوراق بهادار در دسترس هستند و این عوامل دارای ویژگی‌های پویایی، پیچیدگی، غیرخطی و ناپارامتریک بودن می‌باشند، یافتن رابطه بین این متغیرهای کلان اقتصادی با قیمت سهام شرکت و پیش‌بینی نتایج دقیق ضروری است. از طرفی انجام این مهم

ترکیب دوگانه سیستم استنتاج فازی با الگوریتم بهینه‌سازی ازدحام ذرات در پیش‌بینی قیمت سهام بورس اوراق بهادار تهران و مقایسه آن با مدل یادگیری عمیق

الگوریتم بهینه‌سازی ازدحام ذرات^{۱۳} (PSO) جستجو و استخراج می‌شود. همچنین یک شبکه عصبی عمیق بعنوان رویکرد دوم، شامل ۵ لایه پنهان^{۱۴} و ۳۷ نرون به منظور شناسایی ویژگی‌های عمیق^{۱۵} و پنهان ۱۰ متغیر اقتصادی یادشده بر قیمت سهام طراحی و اجرا می‌شود. دو مدل ترکیب دوگانه طراحی شده و یادگیری عمیق برای پیش‌بینی قیمت سهام نه شرکت فعال در بورس اوراق بهادار تهران در دو بازه زمانی ۱۳۸۹ تا پایان ۱۳۹۴ برای شرکت‌های نفت پارس، غدیر و موتوزن و ۱۳۹۴ تا پایان ۱۳۹۸ برای شرکت‌های مس ایران، ایران خودرو، پالایشگاه اصفهان، پتروشیمی خلیج فارس، لوتوس و بانک تجارت تست و نتایج بدست آمده با یکدیگر مقایسه می‌گردند.

در ادامه مقاله و در بخش دوم نحوه پژوهش‌های پیشین تشریح و چالش‌های آنها مورد بررسی قرار می‌گیرند. در بخش سوم جزئیات دو مدل پیشنهادی ارائه شده است. در بخش چهارم نیز مدل‌های پیشنهادی، ماشین بردار پشتیبان و نیز یک شبکه عصبی پرسپترون پیاده‌سازی شده است و سپس نتایج بدست با یکدیگر مقایسه می‌شوند. در نهایت بخش آخر به نتیجه‌گیری و توصیف مسیر آتی تحقیق می‌پردازد.

۲- پژوهش‌های پیشین

در این بخش به بررسی و تحلیل معتبرترین پژوهش‌های انجام شده به منظور شناسایی قیمت سهام براساس ۹ رویکرد شبکه‌های عصبی، یادگیری عمیق، ماشین بردار پشتیبان، درخت تصمیم، جنگل تصادفی، k نزدیکترین همسایه، رگرسیون، بیزین و استنتاج فازی پرداخته می‌شود. شایان ذکر است که اگرچه یادگیری عمیق زیرمجموعه‌ای از شبکه‌های عصبی محسوب می‌شود، اما به دلیل استفاده جدید از یادگیری عمیق در پیش‌بینی قیمت سهام (در منابع معتبر داخلی یافت نشد ولی به تعداد بسیار محدود طی دو سال اخیر در منابع معتبر خارجی یافت شد [۱۹-۲۱]) و نیز باتوجه به کثرت پژوهش‌های انجام شده داخلی [۴-۶، ۸، ۱۸] و خارجی [۹-۱۷] براساس شبکه‌های عصبی و تنوع رویکردهای ترکیبی آنها، این دو مدل بصورت مستقل در این بخش طبقه‌بندی شده‌اند.

با هدف پیش‌بینی قیمت نفت وست تگزاس اینترمدیت^{۱۶} ترکیب دوگانه شبکه عصبی و الگوریتم ژنتیک^{۱۷} [۴] ارائه

بر اساس داده‌های تاریخی^۱ اثبات شده است که ناکافی می‌باشد [۲، ۳].

ورود تکنیک‌های هوش مصنوعی و یادگیری ماشین به منظور شناسایی رابطه بین متغیرهای اقتصادی^۲ بر قیمت سهام امری است که در سال‌های اخیر در بازارهای سهام ایران [۴-۸] و سراسر دنیا [۲، ۹] صورت پذیرفته است. از مهم‌ترین و کارآمدترین این تکنیک‌ها می‌توان به شبکه‌های عصبی^۳ [۴، ۵، ۹-۱۸]، یادگیری عمیق^۴ [۱۹-۲۱]، ماشین بردار پشتیبان^۵ [۲۲-۲۴]، درخت تصمیم^۶ [۲۵-۲۷]، جنگل تصادفی^۷ [۲۸-۳۰]، k نزدیکترین همسایه^۸ [۲۴، ۳۱]، رگرسیون^۹ [۳۲، ۳۳]، شبکه بیزین^{۱۰} [۳۴-۳۶] و روش‌های مبتنی بر استنتاج فازی^{۱۱} [۳۷-۴۱] نام برد. چالش‌های روش‌های پیش‌بینی قیمت سهام اشاره شده در فوق که در مجلات معتبر داخلی و خارجی به چاپ رسیده‌اند در بخش بعد مورد بررسی قرار خواهند گرفت. اما بطور وضوح مشاهده می‌شود که اولاً توانایی تفسیرپذیری الگوهای بدست آمده به منظور پیش‌بینی قیمت سهام برای تحلیلگران بازار سرمایه بسیار حائز اهمیت است، ثانیاً چالش‌های روش‌های موجود باعث شده ترکیب‌های دوگانه^{۱۲} آنها به منظور افزایش توانایی شناسایی ارتباط پیچیده بین متغیرهای تاثیرگذار اقتصادی بر قیمت سهام در دستور کار تحلیلگران داده قرار گیرد.

در این پژوهش نیز برای اولین بار یک مدل استنتاج فازی ممدانی (FIS) جهت شناسایی ارتباط ۱۰ متغیر اقتصادی کمترین قیمت سهام (LowP)، بالاترین قیمت سهام (HighP)، قیمت اولیه سهام (FirstP)، حجم معاملات (VolT)، ارزش معاملات (VlaueT)، شاخص بازار اول تالار بورس (FIndex)، قیمت شاخص کل بازار (PIndex)، نرخ ارز دلار (USD)، قیمت جهانی هر اونس طلا (GoldP)، قیمت جهانی نفت (برنت شمال) (OilP) بر قیمت سهام بازار بورس اوراق بهادار تهران طراحی شود. مجموعه قوانین فازی مدل استنتاج فازی ممدانی توسط

¹ Historical Data

² Economic Factors

³ Neural Networks

⁴ Deep Learning

⁵ Support Vector Machine

⁶ Decision Tree

⁷ Random Forest

⁸ K-Nearest Neighbor

⁹ Regression

¹⁰ Bayesian Network

¹¹ Fuzzy Inference-base Methods

¹² Hybrid

¹³ Particle Swarm Optimization Algorithm

¹⁴ Hidden Layer

¹⁵ Deep Feature

¹⁶ West Texas Intermediate Oil

MBO) به عنوان الگوریتم یادگیری شبکه عصبی کانولوشن عمیق، جدیدترین روش پیش‌بینی قیمت سهام می‌باشد که از پنج متغیر اقتصادی در طراحی این یادگیری عمیق بهبود یافته استفاده شده است [۲۱]. به اعتراف کلیه محقق این حوزه، الگوهای بدست آمده توسط نسخه‌های مختلف شبکه‌های عصبی و یادگیری عمیق مانند یک جعبه سیاه عمل کرده و امکان تفسیری این الگوها وجود ندارد. ماشین بردار پشتیبان (SVM) بعنوان یکی از معروف‌ترین روش‌های داده کاوی نیز استفاده قابل توجهی در این حوزه تحقیقاتی دارد، اگرچه با چالش پیچیدگی بالای زمانی در فاز یادگیری خود و شناسایی تابع کرنل مناسب بر روی داده‌ها با توزیع چندکلاسه غیرخطی روبرو بوده است. لذا تلاش شده بصورت ترکیب دوگانه SVM برای رفع چالش‌های یادشده استفاده شود. یک نسخه دوگانه از SVM بر روی ۱۲ متغیر اقتصادی بازار سهام چین با استفاده از میانگین حرکت رگرسیون خودکار^{۲۷} (ARMA) و کمترین مربعات ماشین بردار پشتیبان^{۲۸} (LS-SVM) اخیراً طراحی شده است [۲۲]. در این پژوهش با بهره‌گیری از تئوری مجموعه راف^{۲۹} تعداد متغیرهای اقتصادی اولیه مفروض به منظور رفع چالش‌های SVM کاهش یافت. رویکرد کاهش تعداد متغیرهای اقتصادی اولیه در [۲۳] براساس یک نسخه بهبود یافته الگوریتم جستجوی مستقیم ترکیبی^{۳۰} (F-SSFS) برای کاهش ۲۹ متغیر اقتصادی بورس نزدک^{۳۱} نیز انجام گرفت. همچنین محققین در [۲۴]، ابتدا با استفاده از معیار بهره اطلاعات^{۳۲} اقدام به وزن‌دهی متغیرهای اقتصادی اولیه بازار سهام چین (بازارهای شانگهای و شانژین) کردند. انتخاب متغیرهای اقتصادی اثرگذار جهت استفاده در SVM از بین ۹ متغیر اقتصادی وزن‌دار شده داده‌های تاریخی توسط الگوریتم k نزدیک‌ترین همسایه صورت پذیرفت. عددی شدن ضرایب ابرصفحات و انتقال ابعاد فضای اولیه به ابعاد بالاتر در استفاده از تابع کرنل SVM غیرخطی باعث می‌شود تفسیر پذیری الگوهای شناسایی شده در پیش‌بینی قیمت سهام توسط نسخه‌های مختلف ماشین بردار پشتیبان را بسیار دشوار و حتی ناممکن سازند.

گردید. در این ترکیب دوگانه، الگوریتم ژنتیک وظیفه یافتن ساختار بهینه اتصالات نرون‌ها در لایه‌های مختلف به یکدیگر و نیز تعیین مقادیر اولیه مناسب برای وزن و بایاس هر نرون را برعهده داشت. ترکیب دوگانه شبکه عصبی فازی و الگوریتم ژنتیک [۷، ۱۰، ۱۲]، ترکیب دوگانه شبکه عصبی بازگشتی^{۱۸} و الگوریتم کلونی زنبور عسل^{۱۹} [۸]، ترکیب دوگانه شبکه عصبی فازی و الگوریتم بهینه‌سازی ازدحام ذرات [۱۳، ۱۴]، ترکیب دوگانه شبکه عصبی و الگوریتم جستجوی فاخته^{۲۰} [۱۶] و ترکیب دوگانه شبکه عصبی و الگوریتم کرم شب تاب^{۲۱} [۱۸] با هدف بهینه‌سازی عملکرد شبکه عصبی در پیش‌بینی قیمت سهام ارائه شده‌اند. اگرچه در طراحی این مدل‌های دوگانه متغیرهای اقتصادی متفاوت برای پیش‌بینی قیمت سهام در بازارهای مختلفی در نظر گرفته شدند ولی نکته مشترک این پژوهش‌ها بر حساسیت شبکه‌های عصبی به پارامترهای ساختاری و نیز ضعف الگوریتم‌های یادگیری کلاسیک آن تاکید دارند. این چالش مهم را می‌توان در رویکرد اخذ شده توسط [۱۷] که با در نظر گرفتن ۳ معماری مختلف برای شبکه عصبی به پیش‌بینی قیمت سهام وارسا^{۲۲} پرداخته، نیز مشاهده نمود.

در [۱۹] یک مدل شبکه عصبی بازگشتی تحت عنوان شبکه حافظه طولانی کوتاه‌مدت^{۲۳} (LSTM) با ۲۵ نرون در تک لایه پنهان آن و یک شبکه عصبی عمیق با ۳ لایه پنهان به منظور پیش‌بینی سری‌های زمانی مالی در بازارهای مالی طراحی و نتایج بدست آمده از آنها را با روش‌های جنگل تصادفی و رگرسیون مقایسه شد. استفاده تجمیعی از ساختار نرون‌های کانولوشن و بازگشتی^{۲۴} در طراحی شبکه عصبی چند فیلتری عمیق^{۲۵} (MFNN) [۲۰] دومین تلاش در استفاده از یادگیری عمیق در شناسایی قیمت سهام است که بر روی شش متغیر تاثیر گذار بر شاخص بازار سهام چین انجام گرفت و نتایجی بهتر از LSTM، شبکه عصبی کانولوشن و بازگشتی داشته است. استفاده از یک نسخه ترکیبی جدید از الگوریتم بهینه‌سازی پروانه^{۲۶} (Rider-

¹⁷ Genetic Algorithm

¹⁸ Recurrent Neural Network

¹⁹ Bee Colony Optimization

²⁰ Cuckoo Search Algorithm

²¹ Firefly Algorithm

²² Warsaw

²³ Long Short-Term Memory Networks

²⁴ Convolutional and Recurrent Neurons

²⁵ Multi-Filters Neural Network

²⁶ Rider-based Monarch Butterfly Optimization

²⁷ Auto-Regressive Moving Average

²⁸ Least Squares Support Vector Machine

²⁹ Rough Set Theory

³⁰ F-Score and Supported Sequential Forward Search

³¹ NASDAQ Stock

³² Information Gain

ترکیب دوگانه سیستم استنتاج فازی با الگوریتم بهینه‌سازی ازدحام ذرات در پیش‌بینی قیمت سهام بورس اوراق بهادار تهران و مقایسه آن با مدل یادگیری عمیق

ناگهانی و قابل توجه متغیرهای اقتصادی) و تعداد متغیرهای اقتصادی بسیار حساس هستند. لذا اگرچه انتخاب ۱۰ متغیر اقتصادی چالش جدی برای مدل‌های طراحی شده ایجاد نخواهد کرد ولی وجود مقادیر نویز در داده‌های آموزش قطعاً چالش قابل توجهی برای آنها ایجاد می‌نماید. مدل گرافی-احتمالی شبکه بیزین [۳۴] به منظور پیش‌بینی قیمت سهام دو شرکت نیکی و تیویوتا موتور براساس تغییرات قیمت در دوره‌های ۲، ۵، ۷، ۹ و ۱۰ روزه طراحی شد. برای رفع چالش یادگیری شبکه بیزین در این پژوهش، ابتدا براساس قیمت روزانه سهام شبکه تعیین می‌شود و سپس به منظور پیش‌بینی قیمت در دوره‌های زمانی تعیین شده به کار گرفته شده که خود باعث تشدید پیچیدگی زمان محاسباتی مدل طراحی شده می‌شود. ترکیب دوگانه روش تنظیم بیزین^{۳۶} و شبکه عصبی پس انتشار^{۳۷} [۳۵] و گراف عامل بیزین پویا^{۳۸} بعنوان یک مدل از شبکه بیزین [۳۶] برای پیش‌بینی قیمت سهام شانگهای براساس ۹ متغیر اقتصادی طراحی و توسعه یافت.

باتوجه به ماهیت عدم قطعیت ارتباطات بین متغیرهای اقتصادی بر تغییرات قیمت سهام، استفاده از منطق فازی در این حوزه تحقیقاتی و در قالب‌های مجموعه قوانین فازی^{۳۹}، سیستم‌های استنتاج فازی تطبیقی^{۴۰} و سیستم عصبی-فازی خودسازگار^{۴۱} از جایگاه ویژه‌ی برخوردار است. یک سیستم استنتاج فازی تطبیقی براساس پنج متغیر اقتصادی به منظور پیش‌بینی قیمت سهام در هفت بازار سهام از جمله بورس‌های داوجونز، نزدک، انگلستان، آلمان و استرالیا طراحی شد [۳۷]. در این سیستم، مقادیر احتمال ورودی^{۴۲} داده‌های آموزشی ورودی جدید به سیستم توسط زنجیره مخفی مارکوف^{۴۳} محاسبه و قوانین فازی جدید باتوجه به این مقادیر تولید می‌شوند. یک سیستم استنتاج فازی تاکاگی-سوگنو-کنگ^{۴۴} (TSK) براساس هشت متغیر اقتصادی تاثیرگذار بر بازار سهام تایوان به منظور پیش‌بینی قیمت سهام طراحی شد [۳۸]. در این پژوهش از خوشه‌بندی k-means و

استفاده از درخت تصمیم بعنوان ابزار پیش‌بینی قیمت سهام اگرچه با حداقل پارامترهای الگوریتمیک را داشته و توانایی تفسیرپذیری الگوهای بدست آمده توسط آن بالاست ولی بسیار حساس به تعداد ویژگی‌ها و مقادیر نویز می‌باشد. لذا از تئوری مجموعه راف به منظور کاهش تعداد ۲۱ متغیر اقتصادی تاثیرگذار بر شاخص سهام بمبئی استفاده و پیش‌بینی قیمت سهام در این بازار بورس توسط الگوریتم C4.5 درخت تصمیم صورت گرفت [۲۵]. بهره‌گیری از معیار بهره اطلاعات به منظور انتخاب متغیرهای اقتصادی بازار سهام شانگهای و پیش‌بینی قیمت‌ها به کمک درخت‌های تصمیم فازی در [۲۶] و پیش‌بینی تنها افزایش یا کاهش قیمت سهام ۱۰ شرکت توسط درختان تصمیم گرادپان تقویتی^{۳۳} در [۲۷] انجام شدند. جنگل تصادفی [۲۸] نیز به منظور شناسایی میزان تقاضا در خرید یا فروش ۲ شرکت از بورس نزدک براساس درخت‌های چندگانه باینری روی شش متغیر اقتصادی طراحی شد. همچنین مدل ترکیبی جنگل تصادفی و حداقل مربعات تقویتی^{۳۴} برای پیش‌بینی قیمت سهام در بورس بمبئی براساس ۳ متغیر اقتصادی در [۲۹] مورد استفاده قرار گرفت. در [۳۰] نیز براساس ۶ متغیر اقتصادی به منظور پیش‌بینی روند قیمت سهام در بورس اوراق بهادار تهران از جنگ تصادفی بهره گرفته شد. در نظر گرفتن تعداد محدود متغیر اقتصادی و طراحی چندگانه درختان در این سه پژوهش را می‌توان در راستای رفع چالش‌های یاد شده برای درخت تصمیم ارزیابی نمود.

به منظور پیش‌بینی قیمت سهام شش شرکت فعال در بورس اردن براساس سه متغیر اقتصادی از الگوریتم k نزدیکترین همسایه استفاده شد [۳۱]. این الگوریتم اگرچه با چالش‌های SVM، درخت تصمیم و جنگل تصادفی روبرو نیست ولی تعیین بهینه مقدار k و حساسیت این الگوریتم به این مقدار چالش جدی آن می‌باشد. دو مدل خود رگرسیون با وقفه توزیعی [۳۲] ترکیب ادغام رگرسیون معمولی و رگرسیون فازی به همراه بهینه‌سازی و نافازی سازی^{۳۵} پارامترها با الگوریتم ژنتیک [۳۳] را می‌توان بعنوان پژوهش‌های صورت گرفته به منظور پیش‌بینی قیمت سهام بورس تهران در گروه صنایع کانی غیرفلزی براساس ۱۰ متغیر اقتصادی نامبرد. مدل‌های مبتنی بر رگرسیون نسبت به مقادیر نویز (تغییرات

³⁶ Bayesian Regularization Method

³⁷ Back Propagation Neural Network

³⁸ Dynamical Bayesian Factor Graph

³⁹ Fuzzy Rules Set

⁴⁰ Adaptive Fuzzy Inference System

⁴¹ Adaptive Neuro Fuzzy Inference System

⁴² Log-Likelihood Value

⁴³ Hidden Markov Model

⁴⁴ Takagi-Sugeno-Kang

³³ Gradient Boosted Decision Tree

³⁴ Least Square Boost

³⁵ Defuzzification

الگوریتم انجماد تدریجی^{۴۵} به منظور تولید و آموزش مجموعه قوانین فازی استفاده شده است. همچنین یک سیستم استنتاج فازی سه گانه TSK برای پیش‌بینی میزان خرید، نگهداری و فروش روزانه سهام در بازار نزدیک براساس ۴ متغیر اقتصادی ارائه گردید [۳۹]. در این پژوهش هر سیستم استنتاج شامل ۳ قانون فازی بود که مقادیر تابع عضویت تخصیص یافته به هر متغیر اقتصادی براساس میانگین و انحراف معیار لگاریتم تغییرات قیمت‌ها محاسبه شده است. لذا با توجه به عدم تضمین پیروی تغییرات قیمت‌ها از توزیع نرمال در بازه‌های زمانی مختلف، این نوع تولید قوانین فازی با چالش جدی روبرو می‌شود. لزوم تعیین بهینه مجموعه قوانین فازی مورد استفاده در سیستم‌های استنتاج فازی باعث شده است که در فرایند آموزش سیستم استنتاج فازی-عصبی خودسازگار از الگوریتم‌هایی همچون تبدیل موجک^{۴۶} [۴۰]، رگرسیون فازی [۴۱]، الگوریتم کرم شب تاب [۴۲]، الگوریتم جستجوی هارمونی^{۴۷} [۴۳]، الگوریتم ژنتیک [۴۴] و الگوریتم بهینه‌سازی ازدحام ذرات [۴۵] به منظور پیش‌بینی قیمت سهام و بازارهای مالی کمک گرفته شود.

۳- مدل‌های پیشنهادی

در این تحقیق به منظور پیش‌بینی قیمت سهام با استفاده از روش‌های هوشمند، به تشریح دو مدل ترکیب دوگانه سیستم استنتاج فازی ممدانی و الگوریتم بهینه‌سازی ذرات (MFIS-PSO) و مدل یادگیری عمیق براساس ۱۰ متغیر اقتصادی تاثیرگذار بر بازار سهام ایران می‌پردازیم. در ادامه این بخش، نخست در نحوه انتخاب و نرمال‌سازی داده‌های ورودی مطرح می‌شود. در ۲-۳ جزئیات مدل ترکیبی و در ۳-۳ جزئیات مدل یادگیری عمیق تشریح خواهند شد.

۳-۱- انتخاب و نرمال‌سازی داده‌های ورودی

گروه دوم: متغیرهای اقتصادی شامل نرخ ارز دلار (USD)، قیمت جهانی هر اونس طلا (GoldP)، قیمت جهانی نفت (برنت شمال) (OilP).

در این مرحله داده‌ها در بازه [0,1] با روش min-max [۴۶] نرمال می‌شوند تا در محدوده یکسانی قرار بگیرند. اجرای نرمال‌سازی باعث می‌شود تا تاثیر متغیرها با مقادیر بزرگتر بر متغیرهایی با مقادیر کوچکتر خنثی شود. نرمال‌سازی min-max یک تبدیل خطی را بر روی داده‌ها اجرا می‌کند و یک مقدار از مجموعه A مانند v_i به مقداری مانند v'_i در محدوده [0,1] با فرمول زیر نگاشت می‌شود:

$$v'_i = \frac{v_i - \min_A}{\max_A - \min_A} \quad (1)$$

در این فرمول \min_A و \max_A به ترتیب حداقل و حداکثر مقدار مجموعه A می‌باشند.

۳-۲- مدل ترکیب دوگانه فازی-ازدحام ذرات

این مدل شامل دو فاز است. در فاز اول ساختار سیستم استنتاج فازی ممدانی [۴۷] شامل تعداد ورودی‌ها، تعداد خروجی‌ها، تعداد و نوع توابع عضویت و نیز تعداد قوانین فازی تعریف می‌شوند. در فاز دوم با استفاده از الگوریتم بهینه‌سازی ازدحام ذرات جزئیات بهترین قوانین فازی یافت می‌شوند.

ساختار سیستم استنتاج فازی پیشنهادی بر اساس مجموعه داده‌ها مفروض برای پیش‌بینی قیمت سهام و

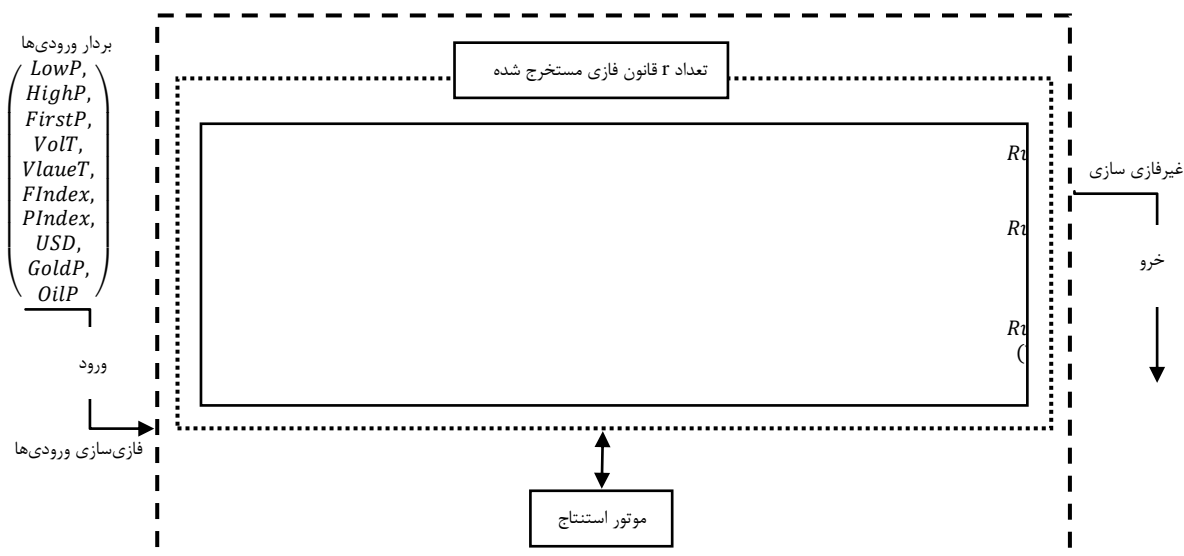
⁴⁵ Simulated Annealing Algorithm

⁴⁶ Wavelet Transform

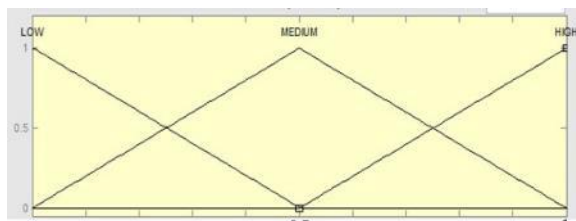
⁴⁷ Harmony Search Algorithm

⁴⁸ U.S Energy Information Administration

ترکیب دوگانه سیستم استنتاج فازی با الگوریتم بهینه‌سازی ازدحام ذرات در پیش‌بینی قیمت سهام بورس اوراق بهادار تهران و مقایسه آن با مدل یادگیری عمیق



شکل ۱: ساختار کلی سیستم استنتاج فازی مددانی براساس ۲ قانون فازی، ۱۰ متغیر ورودی و یک متغیر خروجی



شکل ۲: تابع عضویت متغیرهای سیستم فازی پیشنهادی

همانگونه که از مطالعه پژوهش‌های گذشته مشاهده شد، عدم توانایی آموزش سیستم‌های استنتاج فازی و حساسیت شدید این سیستم‌ها به مجموعه قوانین فازی موتور استنتاج باعث می‌شود از الگوریتم‌های کمکی به منظور شناسایی و تولید بهترین مجموعه قوانین استفاده شود. در این پژوهش، یک فرآیند آموزش تکاملی باعث شناسایی و بهبود کیفیت مجموعه قوانین موتور استنتاج می‌گردد. در این فرآیند آموزش تکاملی فرض بر این است که توابع عضویت از پیش تعریف شده‌ای در پایگاه داده وجود دارد و هدف یافتن بهترین مجموعه قوانین فازی توسط الگوریتم PSO می‌باشد. الگوریتم بهینه‌سازی ازدحام ذرات (PSO) [۴۸] یک الگوریتم جمعیت محور و حافظه‌مند است که از رفتار دسته جمعی ماهی‌ها و پرندگان برای یافتن غذا الهام گرفته شده است. در این الگوریتم، یک جواب در فضای مسئله را ذره می‌نامید. تغییر در موقعیت ذرات در فضای جستجو مبتنی بر رفتار اجتماعی موجودات در تقلید از موفقیت سایر موجودات است. تغییرات یک ذره در ازدحام، تحت تاثیر تجربیات خود و یا دانش بهترین ذره می‌باشد [۴۹]. هر ذره

متغیرهای تاثیرگذار آن در شکل ۱ نمایش داده شده است. در این سیستم از ۱۰ متغیر کمترین قیمت سهام (LowP)، بالاترین قیمت سهام (HighP)، قیمت اولیه سهام (FirstP)، حجم معاملات (Volt)، ارزش معاملات (VlaueT)، شاخص بازار اول تالار بورس (FIndex)، قیمت شاخص کل بازار (PIndex)، نرخ ارز دلار (USD)، قیمت جهانی هر اونس طلا (GoldP)، قیمت جهانی نفت (برنت شمال) (OilP) بعنوان ورودی استفاده شده است. این سیستم فازی نیز دارای یک خروجی است که قیمت سهام پیش‌بینی شده می‌باشد که با نام PPrice مشخص می‌شود.

برای ورودی‌ها و خروجی‌های یک سیستم استنتاج فازی می‌توان توابع عضویت فازی مختلفی را در نظر گرفت که با تغییر نوع و تعداد آنها می‌توان به پاسخ‌های متفاوتی از عملکرد سیستم دست یافت. در این پژوهش از متداول‌ترین تابع عضویت، یعنی تابع عضویت مثلثی برای تولید تابع عضویت هر متغیر اقتصادی ورودی و خروجی بهره گرفته شده است. هر کدام از این متغیرها سه متغیر زبانشناختی^{۴۹} High, Medium, Low (شکل ۲) تشکیل می‌شوند که به ترتیب نشان دهنده مقادیر کم، متوسط و زیاد متغیرها هستند.

⁴⁹ Linguistic Variable

ورودی و خروجی دارای سه مفهوم زبانشناختی low, medium, high است. لذا تعداد کل قوانین فازی ممکن برابر 3×4^{10} یعنی ۳۱۴۵۷۲۸ قانون خواهد بود. در مدل فازی پیشنهادی تعداد قواعد سیستم فازی به صورت ثابت r تعیین می‌شود و تعداد متغیرهای هر ذره برابر $12 \times r$ است که هر ۱۲ متغیر به صورت جداگانه مربوط به یک موتور استنتاج این سیستم فازی است به طوری که متغیرهای شماره ۱ تا ۱۲ مربوط به قانون فازی ۱، متغیرهای شماره ۱۳ تا ۲۴ مربوط به قانون فازی ۲، متغیرهای شماره ۲۵ تا ۳۶ مربوط به قانون فازی ۳ و به همین ترتیب متغیرهای شماره $11 - i \times 12$ تا $i \times 12$ مربوط به قانون i ام است تا زمانیکه $r = i$. برای هر قانون i متغیرهای ۱ تا ۱۰ مقدار ۰ تا ۳ را می‌گیرد که مقدار متغیر زبانشناختی مربوط به ۱۰ ورودی فازی را تعیین می‌کند. مقدار ۰ به معنی عدم حضور متغیر در قانون فازی، مقدار ۱ مفهوم زبانشناختی low، مقدار ۲ مفهوم زبانشناختی medium و مقدار ۳ مفهوم زبانشناختی High را برای متغیر مربوطه در قانون فازی مشخص می‌کند. متغیر شماره ۱۱ هر قانون فازی مقداری بین ۱ تا ۳ را می‌گیرند که مقدار متغیر زبانشناختی مربوط به خروجی فازی را نشان می‌دهد و متغیر شماره ۱۲ هر قاعده فازی نیز وزن قاعده را که بین ۰ و ۱ است نشان می‌دهد. نحوه نمایش ذرات از نوع عددی حقیقی است. در شکل ۳ ساختار یک ذره در روش پیشنهادی نشان داده شده است.

1	2	3	4	5	6	7	8	9	10	11	12	...	۳۱۲
360	360	360	360	360	360	360	361	361	160	360	360		160

شکل ۳: ساختار یک ذره در PSO برای مدل فازی ترکیبی پیشنهادی

تابع برازش، نحوه محاسبه کیفیت جواب‌ها (ذرات) به منظور ایجاد مقایسه کیفی میان آنها است. هدف الگوریتم بهینه‌سازی ذرات تعیین قوانین فازی موتور استنتاج برای سیستم فازی پیشنهادی به گونه‌ای است که خروجی آن کمترین خطا نسبت به مقادیر قیمت سهام واقعی را داشته باشد. لذا موتور استنتاج ارائه شده توسط هر ذره به ازای مقادیر ۱۰ متغیر اقتصادی ورودی در هر روز یک قیمت سهام را برای آن روز محاسبه می‌نماید. بنابراین میانگین مربع خطای^{۵۰} قیمت سهام پیش‌بینی شده نسبت به قیمت سهام

دارای یک مقدار شایستگی می‌باشد که توسط تابع شایستگی مسئله محاسبه می‌شود. اگر $X_i(t) = (x_{i1}(t), \dots, x_{in}(t))$ موقعیت ذره i در فضای جستجو n بعدی در گام زمانی t باشد موقعیت ذره با اضافه کردن سرعت $V_i(t) = (v_{i1}(t), \dots, v_{in}(t))$ به موقعیت جدید تغییر می‌کند:

$$X \quad (2)$$

به صورتی که $x_{ij}(0) \sim U(x_{min}, x_{max})$ سرعت ذره i به صورت زیر محاسبه می‌شود:

$$v_{ij}(t+1) = v_{ij}(t) + c_1 r_{1j}(t)[y_{ij}(t) - x_{ij}(t)] + c_2 r_{2j}(t)[\hat{y}_j(t) - x_{ij}(t)] \quad (3)$$

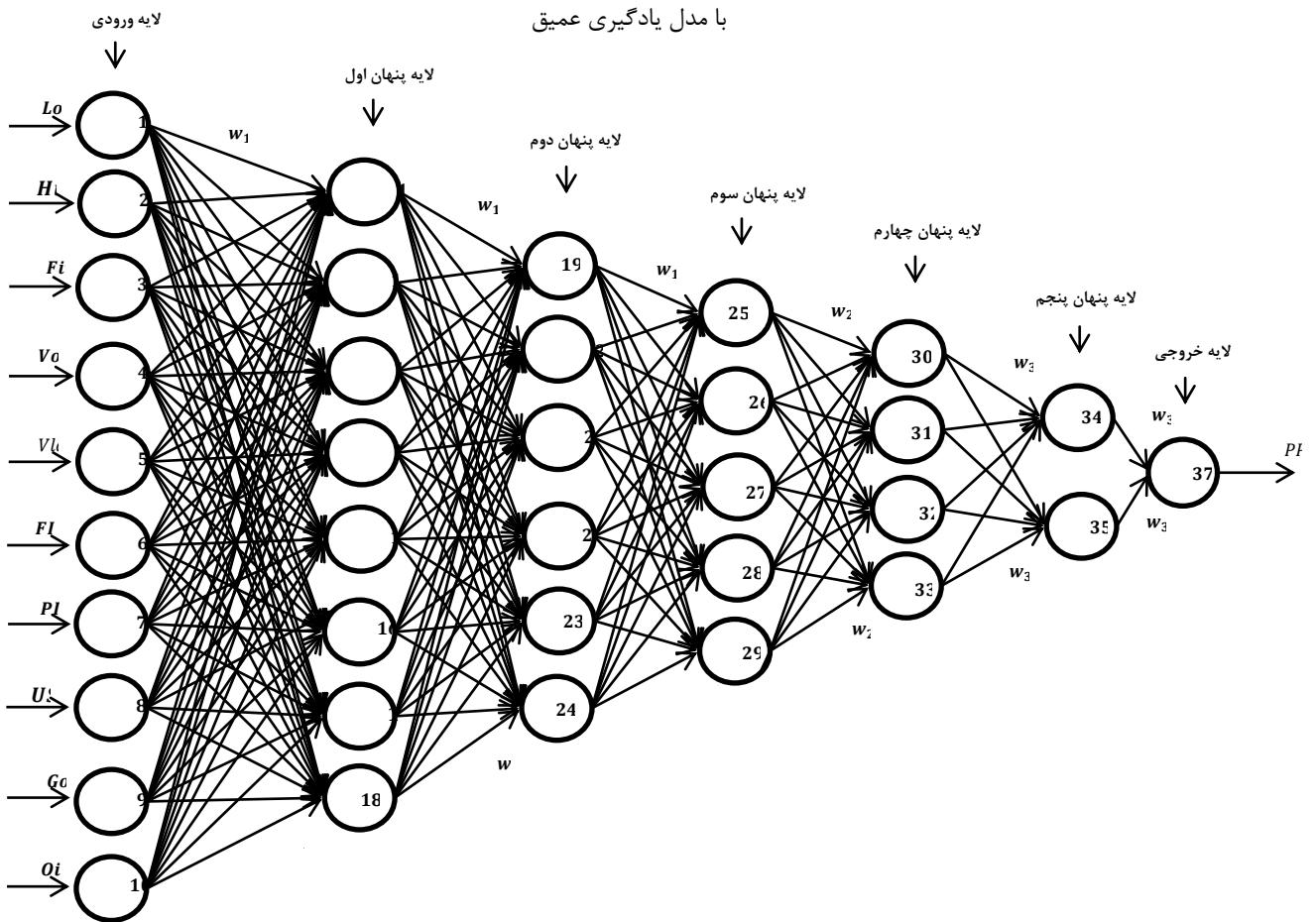
که $v_{ij}(t)$ سرعت ذره i در بعد j ام آن در گام زمانی t می‌باشد که $j=1, \dots, n$ است. $x_{ij}(t)$ موقعیت یا مکان ذره i در بعد j در گام زمانی t می‌باشد. c_1 و c_2 ثابت‌های مثبتی با نام ضرایب شتاب هستند که برای تنظیم میزان تاثیر پذیری مولفه‌های شناختی و اجتماعی بکار می‌روند. $r_{1j}(t)$ و $r_{2j}(t)$ مقادیر تصادفی در دامنه $[0, 1]$ می‌باشند که از توزیع یکنواخت نمونه برداری می‌شوند. بهترین موقعیت شخصی ذره i یا همان $Y_i(t) = (y_{i1}(t), \dots, y_{in}(t))$ ، بهترین موقعیت دیده شده توسط ذره i از اولین گام تا زمان t می‌باشد. $\hat{Y}(t) = (\hat{y}_1(t), \dots, \hat{y}_n(t))$ بهترین موقعیت سراسری در گام زمانی t تعریف می‌شود [۴۹]. مراحل الگوریتم PSO را می‌توان به صورت زیر تشریح نمود:

۱. تولید ذرات اولیه و محاسبه برازش آنها
۲. مقداردهی اولیه به پارامترهای سرعت و مکان ذرات در ازدحام اولیه
۳. تکرار مرحله ۴ برای همه ذرات
۴. بهنگام سازی بهترین موقعیت محلی ذره i ام
۵. بهنگام سازی بهترین موقعیت سراسری همه ذرات
۶. تکرار مراحل ۷ و ۸ برای همه ذرات
۷. محاسبه سرعت جدید ذره i ام براساس رابطه ۳
۸. محاسبه مکان جدید ذره i ام براساس رابطه ۲
۹. تکرار مراحل ۳ تا ۸ تا ارضا شرط خروج

از آنجا که در مدل فازی ترکیبی، نقش الگوریتم PSO تعیین قواعد فازی موتور استنتاج است، هر ذره حاوی اطلاعات تمامی قواعد موجود در پایگاه قوانین فازی بوده که متشکل از ۱۰ متغیر ورودی USD, PIndex, FIndex, LowP, OilP, VlaueT, VolT, FirstP, HighP و یک متغیر خروجی PredictP است. همانطور که اشاره شد، هر متغیر

⁵⁰ Mean Square Error

ترکیب دوگانه سیستم استنتاج فازی با الگوریتم بهینه‌سازی ازدحام ذرات در پیش‌بینی قیمت سهام بورس اوراق بهادار تهران و مقایسه آن با مدل یادگیری عمیق



شکل ۴: معماری شبکه عصبی پیش‌خور عمیق شامل ۵ لایه پنهان، لایه‌های ورودی و خروجی

که I_j ورودی خالص نرون j ($j = 11, 12, \dots, 37$) بوده و برابر با مجموعه وزندار شده خروجی نرون‌های لایه ماقبل ($\sum_i w_{ij} O_i$) بعلاوه بایاس نرون j (θ_j) می‌باشد. O_j خروجی نرون j است. کل بازه زمانی رصد شده برای ایجاد مدل پیش‌بینی قیمت سهام را به دو بازه روزهای آموزش و روزهای آزمون تقسیم می‌شوند. الگوریتم یادگیری این شبکه پیش‌خور عمیق، پس انتشار خطا با روش گرادیان نزولی Adam با نرخ یادگیری $l = 0.001$ می‌باشد که بر روی داده‌های روزهای مربوط به آموزش به شرح ذیل اجرا می‌شود.

۱. تمامی وزن‌ها و بایاس آنها را برای معماری شکل ۴ بصورت تصادفی در بازه $[-0.5, 0.5]$
۲. اگر شرط توقف برقرار است برو به ۱۱
۳. برای هر X (مقادیر ۱۰ متغیر اقتصادی مفروض در هر روز) اقدامات ذیل را انجام بده
۴. برای هر نرون $j = 1, \dots, 10$ از لایه ورودی $O_j = I_j$
۵. برای هر نرون $j = 11, \dots, 37$ ورودی را به صورت زیر و خروجی را براساس رابطه ۵ محاسبه کن

واقعی را در این پژوهش به عنوان برازش در نظر گرفته شده است. تابع برازش الگوریتم PSO برای تنظیم قوانین سیستم فازی پیشنهادی به منظور پیش‌بینی قیمت سهام در فرمول زیر آورده شده است.

$$f(X) = \frac{\sum_{i=1}^N (RP_i - PP_i)^2}{N} \quad (4)$$

که در آن $f(X)$ تابع برازش ذره X است، N تعداد روزهای بررسی قیمت سهام، RP_i مقدار قیمت سهام واقعی در روز i ام و PP_i مقدار قیمت سهام خروجی (پیش‌بینی شده) سیستم فازی با موتور استنتاج ذره X است.

۳-۳-۳ مدل یادگیری عمیق

برای این منظور شبکه عصبی پیش‌خور^{۵۱} هفت لایه مطابق با شکل ۴ طراحی شده است. در لایه ورودی باتوجه به وجود ۱۰ متغیر اقتصادی، ۱۰ نرون فاقد بایاس و تابع فعال‌سازی تعریف می‌شوند. برای هر نرون لایه‌های پنهان و لایه خروجی مقدار بایاس θ_j که $j = 11, 12, \dots, 37$ و تابع فعال‌سازی سیگموئیدی به صورت زیر تعریف می‌شوند:

$$O_j = \frac{1}{1 + e^{-I_j}} \quad (5)$$

⁵¹ Feed-Forward Neural Network

برنامه نویسی Python3.7 در محیط spyder بر روی یک کامپیوتر با حافظه‌ی اصلی ۸ گیگابایت و پردازنده پنج هسته‌ای ۲،۲ گیگا هرتز استفاده شده است. با توجه به محبوبیت و کثرت استفاده از شبکه‌های عصبی و ماشین بردار پشتیبان به منظور پیش‌بینی قیمت سهام در پژوهش‌های پیشین، ضمن مقایسه کیفیت عملکرد دو مدل پیشنهادی، نتایج بدست آمده از آنها با نتایج بدست آمده از شبکه عصبی پیشنهادی در [۱۷] و ماشین بردار پشتیبان [۲۳] مقایسه شده است. جزئیات مراحل پیاده‌سازی و اجرای مدل‌های پیشنهادی در زیر بیان شده است:

• بارگذاری و نرمال‌سازی داده‌ها: ابتدا باید ماتریس هر

شرکت بارگذاری و توسط رابطه ۱ نرمال‌سازی شوند تا مقادیر همه داده‌ها در بازه صفر تا یک قرار بگیرند. برای ذخیره‌سازی داده‌های هر یک از شرکت‌ها از یک ماتریس استفاده شده است. برای هر بایاس θ_j در شکل ۴ به صورت زیر به روز این ماتریس‌ها دارای ۱۱ ستون و ۲۰۰۰ سطر برای شرکت‌های گروه اول و ۱۸۰۰ سطر برای شرکت‌های گروه دوم هستند که ده ستون اول این ماتریس‌ها ۱۰ متغیر HighP.LowP، VlaueT، Volt.FirstP،

OilP.GoldP،USD.PIndex.FIndex، برای شرکت مورد نظر و ستون آخر قیمت سهام PPrice است. ورودی‌های هر کدام از شرکت‌های گروه اول شامل ۲۰۰۰ داده (روز) هستند که ۷۰٪ آنها (اطلاعات مربوط به ۱۴۰۰ روز) به صورت تصادفی به عنوان داده‌های آموزشی و مابقی داده‌ها (اطلاعات مربوط به ۶۰۰ روز)، داده‌های آزمایشی در نظر گرفته شده‌اند. همچنین ورودی‌های هر کدام از شرکت‌های گروه دوم شامل ۱۸۰۰ داده (روز) هستند که ۷۰٪ آنها (اطلاعات مربوط به ۱۲۶۰ روز) به صورت تصادفی به عنوان داده‌های آموزشی و مابقی داده‌ها (اطلاعات مربوط به ۵۴۰ روز)، داده‌های آزمایشی در نظر گرفته شده‌اند.

• طراحی سیستم استنتاج فازی ممدانی (شکل ۱): این سیستم دارای ۱۰ متغیر ورودی (HighP، LowP، USD، PIndex، Findex، VlaueT، Volt، FirstP، OilP، GoldP) و یک متغیر خروجی (قیمت سهام PPrice) است که با استفاده از جعبه ابزار Fuzzy نرم افزار Matlab، نوع طراحی می‌شود. در این سیستم فازی، نوع استنتاج از نوع ممدانی، روش غیر فازی‌ساز از نوع مرکز ثقل و برای عملگرهای AND، OR از مینیمم و ماکسیمم و برای توابع تجمیع و استدلال نیز به ترتیب از مینیمم و ماکسیمم استفاده شده است. همانطور که بیان شد هر یک از

$$I_j = \sum_i w_{ij} O_j + \theta_j \quad (6)$$

۶. خطای نرون ۳۷ را بصورت ذیل محاسبه کن

$$Err_{37} = O_{37}(1 - O_{37})(T_{37} - O_{37}) \quad (7)$$

که در آن T_{37} قیمت واقعی سهام در روز مربوط به داده X است.

۷. خطای هر نرون $j = 36, \dots, 11$ از لایه پنهان پنجم به لایه پنهان اول به صورت زیر محاسبه کن

$$Err_j = O_j(1 - O_j) \sum_k Err_k w_{jk} \quad (8)$$

که در آن k نرون‌های لایه ماقبل نرون j هستند.

۸. برای هر وزن w_{ij} در شکل ۴ به صورت زیر به روز رسانی کن

$$\Delta w_{ij} = (l)Err_j O_i \quad (9)$$

$$w_{ij} = w_{ij} + \Delta w_{ij} \quad (10)$$

که در آن l نرخ یادگیری با مقدار ۰،۰۰۱ می‌باشد.

۹. برای هر بایاس θ_j در شکل ۴ به صورت زیر به روز این ماتریس‌ها دارای ۱۱ ستون و ۲۰۰۰ سطر برای شرکت‌های گروه اول و ۱۸۰۰ سطر برای شرکت‌های گروه دوم هستند که ده ستون اول این ماتریس‌ها ۱۰ متغیر HighP.LowP، VlaueT، Volt.FirstP،

$$\Delta \theta_j = (l)Err_j \quad (11)$$

$$\theta_j = \theta_j + \Delta \theta_j \quad (12)$$

۱۰. اگر داده X از روزهای آموزش باقیمانده برو به ۳

۱۱. پایان

شروط توقف مختلفی را می‌توان برای خط ۲ الگوریتم یادگیری مطرح نمود که در این پژوهش رسیدن به حداکثر تعداد تکرار $MaxEpo$ می‌باشد. با پایان فرایند یادگیری، میانگین مربع خطای قیمت سهام پیش‌بینی شده نسبت به قیمت سهام واقعی براساس رابطه ۴ برای داده‌های آموزشی و آزمون محاسبه می‌گردد.

۴- پیاده‌سازی و ارزیابی

برای بررسی کارایی مدل ترکیب فازی-ازدحام ذرات (MFIS-PSO) و مدل یادگیری عمیق به عنوان مطالعه موردی از داده‌های نه شرکت فعال در بورس اوراق بهادار تهران (نفت پارس، غدیر، موتوژن، مس ایران، ایران خودرو، پالایشگاه اصفهان، پتروشیمی خلیج فارس، لوتوس و بانک تجارت) در بازه‌های ۱۳۸۹ تا پایان ۱۳۹۴ و از ۱۳۹۴ تا پایان ۱۳۹۸ جمع آوری شده که از ۱۰ متغیر تاثیرگذار بر قیمت سهام در این بازه استفاده شده است. برای پیاده‌سازی و اجرای مدل ترکیب فازی-ازدحام ذرات از نرم افزار Matlab R2017a بر روی یک کامپیوتر با حافظه‌ی اصلی ۸ گیگابایت و پردازنده پنج هسته‌ای ۲،۴ گیگا هرتز استفاده شده و برای پیاده‌سازی و اجرای مدل یادگیری عمیق از زبان

ترکیب دوگانه سیستم استنتاج فازی با الگوریتم بهینه‌سازی ازدحام ذرات در پیش‌بینی قیمت سهام بورس اوراق بهادار تهران و مقایسه آن با مدل یادگیری عمیق

متغیرهای ورودی و خروجی دارای سه مفهوم زبانشناختی پشتیبان استفاده شده و به منظور تنظیم کرنل مناسب از تابع low,medium,high از نوع تابع عضویت مثلثی هستند. هسته گاوسی برای تولید بردارهای پشتیبان مناسب بهره گرفته از آنجایی که مقدار داده‌ها در بازه ۰ تا ۱ نرمال شده است، شده است.

محور افقی توابع عضویت مربوط به متغیرهای سیستم فازی در محدوده ۰ تا ۱ تعریف شده است و پارامترهای مربوط به تابع عضویت مثلثی مفهوم زبانشناختی low برابر $[-0.4, 0, 0.4]$ ، مفهوم زبانشناختی medium برابر $[0.1, 0.5, 0.9]$ و مفهوم زبانشناختی high برابر $[0.6, 1, 1.4]$ است.

• **اجرای الگوریتم POS:** هر ذره در این الگوریتم در واقع ۲۰ قانون فازی را نشان می‌دهد و در نتیجه طول هر ذره برابر ۲۴۰ است. کیفیت هر ذره توسط رابطه ۴ محاسبه و هرچه قیمت سهام تخمین‌زده شده توسط مدل ترکیبی فازی-ازدحام ذرات به قیمت واقعی سهام نزدیک‌تر باشد، قوانین فازی بیان شده توسط ذره مورد نظر بهتر بوده و در نتیجه مقدار خروجی تابع برازش کمتر خواهد بود. نحوه تنظیم پارامترهای آن نیز در جدول ۲ آمده است.

جدول ۲: پارامترهای الگوریتم بهینه‌یابی ذرات

جمعیت اولیه	تعداد تکرار	C1	C2	W
۵۰۰	۲۰	۰٫۷	۱٫۵	۰٫۹

۴-۱- تحلیل نتایج داده‌های گروه اول داده‌های گروه اول شامل اطلاعات سه شرکت نفت پارس، موتوژن و غدیر طی دوره زمانی ۱۳۸۹ تا پایان ۱۳۹۴ است. دو مدل پیشنهادی و شبکه عصبی [۱۷]، هر کدام ۲۰ بار اجرا شده و نتایج بدست آمده برای داده‌های آموزش و آزمون آنها در جداول ۳، ۴ و ۵ نمایش داده شده است. از مقایسه کلی نتایج بدست آمده در این جداول مشاهده می‌شود که اگرچه مدل ترکیبی فازی-ازدحام ذرات موفق شده به میزان ۶۰٪ نتایج بهتری از شبکه عصبی داشته باشد ولی در مقایسه با یادگیری عمیق عملکرد ضعیف‌تری داشته و مدل یادگیری عمیق به میزان ۷۹٪ موفق‌تر از مدل ترکیبی فازی-ازدحام ذرات و ۹۲٪ موفق‌تر از شبکه عصبی در پیش‌بینی قیمت سهام عمل کرده است. همچنین از مقایسه نتایج جدول ۳ مشاهده می‌شود که مدل ترکیبی فازی-ازدحام ذرات در پیش‌بینی قیمت سهام موتوژن به میزان ۶٪ بهتر از پیش‌بینی قیمت سهام غدیر و ۱۷٪ بهتر از نفت پارس عمل کرده است. نتایج جدول ۴ نشان می‌دهند که مدل یادگیری عمیق در پیش‌بینی قیمت سهام غدیر بهترین عملکرد را داشته بطوریکه این نتایج ۶٪ دقیق‌تر از پیش‌بینی قیمت سهام نفت توسط این مدل و ۶۸٪ نتایج دقیق‌تری در پیش‌بینی قیمت سهام موتوژن داشته است. تحلیل آماری نتایج مدل شبکه عصبی در جدول ۵ نشان از موفقیت این مدل در پیش‌بینی قیمت سهام نفت پارس را دارد. بگونه‌ای که خطای پیش‌بینی قیمت سهام شبکه عصبی برای نفت پارس ۰٫۳٪ بهتر از خطای پیش‌بینی سهام موتوژن و ۳۲٪ بهتر از خطای پیش‌بینی غدیر بوده است.

نکته قابل توجه دیگر در تحلیل آماری نتایج جداول ۳، ۴ و ۵ شکاف عملکردی مدل‌های یادگیری عمیق و شبکه عصبی بر روی سه شرکت نفت پارس، غدیر و موتوژن می‌باشد در حالی که مدل ترکیبی فازی-ازدحام ذرات دارای عملکردی تقریباً مشابه برای این سه شرکت بوده است.

• **اجرای یادگیری عمیق:** شبکه پیشخور عمیق شکل ۴ با حداکثر تکرار ۱۰۰، روش گرادیان نزولی Adam، نرخ یادگیری ۰٫۰۰۱، تابع فعالسازی سیگموئیدی و استفاده از کتابخانه Keras پیاده‌سازی و اجرا شده است. کیفیت پیش‌بینی قیمت سهام شبکه عمیق یادگیری شده توسط رابطه ۴ برای داده‌های آموزش و آزمون محاسبه می‌شود.

• **اجرای شبکه عصبی مصنوعی [۱۷]:** این شبکه از نوع پیش‌خور با یک لایه ورودی شامل ۱۰ نرون، یک لایه پنهان شامل ۸ نرون و یک لایه خروجی شامل یک نرون است. برای آموزش آن نیز از روش پس انتشار خطا [۵۰] استفاده شده است و تابع فعال‌سازی نرون‌ها نیز تابع حلقوی [۴۹] تعیین شده است. شرط توقف برابری خطای آموزش و اعتبارسنجی است.

• **اجرای ماشین بردار پشتیبان [۲۳]:** باتوجه به ماهیت غیرخطی داده‌های مالی از نسخه غیرخطی ماشین بردار

جدول ۳: جزئیات نتایج ۲۰ اجرای مدل ترکیبی فازی - ازدحام ذرات

داده	نفت پارس		موتوزن		غدیر	
	آموزش	آزمون	آموزش	آزمون	آموزش	آزمون
۱	۰.۰۰۸۲	۰.۰۰۸۴	۰.۰۰۶۷	۰.۰۰۶۹	۰.۰۰۶۳	۰.۰۰۷۲
۲	۰.۰۰۷۱	۰.۰۰۷۹	۰.۰۰۶۲	۰.۰۰۷۳	۰.۰۰۶۱	۰.۰۰۷۲
۳	۰.۰۰۸۶	۰.۰۰۸۷	۰.۰۰۶۶	۰.۰۰۷۰	۰.۰۰۶۵	۰.۰۰۷۲
۴	۰.۰۰۸۸	۰.۰۰۸۹	۰.۰۰۶۸	۰.۰۰۶۹	۰.۰۰۶۷	۰.۰۰۷۶
۵	۰.۰۰۷۷	۰.۰۰۷۸	۰.۰۰۷۱	۰.۰۰۷۴	۰.۰۰۶۳	۰.۰۰۷۶
۶	۰.۰۰۷۴	۰.۰۰۸۱	۰.۰۰۶۶	۰.۰۰۷۱	۰.۰۰۶۸	۰.۰۰۷۶
۷	۰.۰۰۷۴	۰.۰۰۸۱	۰.۰۰۶۶	۰.۰۰۷۲	۰.۰۰۶۴	۰.۰۰۷۲
۸	۰.۰۰۷۷	۰.۰۰۸۴	۰.۰۰۶۱	۰.۰۰۷۲	۰.۰۰۶۹	۰.۰۰۷۴
۹	۰.۰۰۸۲	۰.۰۰۸۵	۰.۰۰۶۱	۰.۰۰۶۷	۰.۰۰۶۹	۰.۰۰۷۳
۱۰	۰.۰۰۷۳	۰.۰۰۸۶	۰.۰۰۶۲	۰.۰۰۶۹	۰.۰۰۷۰	۰.۰۰۷۶
۱۱	۰.۰۰۷۸	۰.۰۰۸۰	۰.۰۰۵۹	۰.۰۰۶۲	۰.۰۰۶۷	۰.۰۰۷۴
۱۲	۰.۰۰۷۱	۰.۰۰۸۵	۰.۰۰۶۲	۰.۰۰۶۳	۰.۰۰۶۲	۰.۰۰۷۷
۱۳	۰.۰۰۷۵	۰.۰۰۸۴	۰.۰۰۶۱	۰.۰۰۶۹	۰.۰۰۶۴	۰.۰۰۷۳
۱۴	۰.۰۰۷۱	۰.۰۰۷۶	۰.۰۰۶۲	۰.۰۰۷۲	۰.۰۰۷۲	۰.۰۰۷۳
۱۵	۰.۰۰۷۲	۰.۰۰۷۵	۰.۰۰۷۱	۰.۰۰۷۴	۰.۰۰۶۳	۰.۰۰۷۲
۱۶	۰.۰۰۷۰	۰.۰۰۷۳	۰.۰۰۶۴	۰.۰۰۶۵	۰.۰۰۷۱	۰.۰۰۷۲
۱۷	۰.۰۰۸۳	۰.۰۰۸۹	۰.۰۰۶۲	۰.۰۰۶۹	۰.۰۰۶۸	۰.۰۰۷۶
۱۸	۰.۰۰۷۶	۰.۰۰۷۹	۰.۰۰۶۲	۰.۰۰۶۸	۰.۰۰۷۳	۰.۰۰۷۵
۱۹	۰.۰۰۸۸	۰.۰۰۸۹	۰.۰۰۶۰	۰.۰۰۶۲	۰.۰۰۶۲	۰.۰۰۷۵
۲۰	۰.۰۰۷۱	۰.۰۰۷۷	۰.۰۰۶۰	۰.۰۰۶۴	۰.۰۰۶۶	۰.۰۰۷۲
بهترین	۰.۰۰۷۰	۰.۰۰۷۳	۰.۰۰۵۹	۰.۰۰۶۲	۰.۰۰۶۱	۰.۰۰۷۲
میانگین	۰.۰۰۷۷	۰.۰۰۸۲	۰.۰۰۶۴	۰.۰۰۶۹	۰.۰۰۶۶	۰.۰۰۷۴

جدول ۵: جزئیات نتایج ۲۰ اجرای شبکه عصبی مصنوعی [۱۷]

داده	نفت پارس		موتوزن		غدیر	
	آموزش	آزمون	آموزش	آزمون	آموزش	آزمون
۱	۰.۰۱۵۴	۰.۰۱۶۴	۰.۰۱۵۱	۰.۰۱۶۲	۰.۰۲۱۱	۰.۰۲۴۵
۲	۰.۰۱۵۱	۰.۰۱۶۱	۰.۰۱۴۹	۰.۰۱۶۱	۰.۰۲۰۰	۰.۰۲۴۴
۳	۰.۰۱۴۹	۰.۰۱۶۰	۰.۰۱۵۲	۰.۰۱۶۲	۰.۰۲۱۵	۰.۰۲۴۹
۴	۰.۰۱۵۰	۰.۰۱۶۱	۰.۰۱۴۸	۰.۰۱۶۲	۰.۰۲۰۶	۰.۰۲۴۴
۵	۰.۰۱۴۹	۰.۰۱۶۰	۰.۰۱۵۲	۰.۰۱۶۱	۰.۰۲۳۲	۰.۰۲۵۰
۶	۰.۰۱۴۶	۰.۰۱۶۰	۰.۰۱۴۸	۰.۰۱۶۴	۰.۰۲۳۴	۰.۰۲۴۸
۷	۰.۰۱۴۷	۰.۰۱۶۵	۰.۰۱۵۰	۰.۰۱۶۲	۰.۰۲۱۵	۰.۰۲۴۷
۸	۰.۰۱۴۶	۰.۰۱۶۵	۰.۰۱۵۱	۰.۰۱۶۲	۰.۰۲۰۴	۰.۰۲۴۷
۹	۰.۰۱۴۷	۰.۰۱۶۲	۰.۰۱۵۲	۰.۰۱۶۲	۰.۰۲۰۹	۰.۰۲۴۵
۱۰	۰.۰۱۴۷	۰.۰۱۵۹	۰.۰۱۴۸	۰.۰۱۶۲	۰.۰۲۱۴	۰.۰۲۴۷
۱۱	۰.۰۱۴۹	۰.۰۱۶۰	۰.۰۱۵۴	۰.۰۱۶۰	۰.۰۲۳۱	۰.۰۲۵۰
۱۲	۰.۰۱۴۵	۰.۰۱۵۹	۰.۰۱۵۲	۰.۰۱۶۱	۰.۰۲۰۹	۰.۰۲۴۷
۱۳	۰.۰۱۵۴	۰.۰۱۶۴	۰.۰۱۵۰	۰.۰۱۶۱	۰.۰۲۲۱	۰.۰۲۴۷
۱۴	۰.۰۱۵۴	۰.۰۱۵۹	۰.۰۱۵۰	۰.۰۱۶۰	۰.۰۲۲۵	۰.۰۲۴۵
۱۵	۰.۰۱۵۰	۰.۰۱۵۹	۰.۰۱۵۰	۰.۰۱۶۳	۰.۰۲۰۸	۰.۰۲۴۷
۱۶	۰.۰۱۵۰	۰.۰۱۶۰	۰.۰۱۴۹	۰.۰۱۶۰	۰.۰۲۰۴	۰.۰۲۴۸
۱۷	۰.۰۱۴۸	۰.۰۱۶۳	۰.۰۱۵۱	۰.۰۱۶۰	۰.۰۲۱۰	۰.۰۲۴۸
۱۸	۰.۰۱۵۴	۰.۰۱۶۳	۰.۰۱۵۱	۰.۰۱۶۰	۰.۰۲۱۱	۰.۰۲۴۶
۱۹	۰.۰۱۴۹	۰.۰۱۶۳	۰.۰۱۵۳	۰.۰۱۶۰	۰.۰۲۱۵	۰.۰۲۴۶
۲۰	۰.۰۱۴۶	۰.۰۱۶۲	۰.۰۱۴۷	۰.۰۱۵۹	۰.۰۲۱۸	۰.۰۲۵۰
بهترین	۰.۰۱۴۵	۰.۰۱۵۹	۰.۰۱۴۷	۰.۰۱۵۹	۰.۰۲۰۰	۰.۰۲۴۴
میانگین	۰.۰۱۴۹	۰.۰۱۶۱	۰.۰۱۵۰	۰.۰۱۶۱	۰.۰۲۱۴	۰.۰۲۴۷

جدول ۴: جزئیات نتایج ۲۰ اجرای مدل یادگیری عمیق

داده	نفت پارس		موتوزن		غدیر	
	آموزش	آزمون	آموزش	آزمون	آموزش	آزمون
۱	۰.۰۰۰۴	۰.۰۰۱۸	۰.۰۰۲۶	۰.۰۰۱۴	۰.۰۰۰۳	۰.۰۰۰۵
۲	۰.۰۰۰۶	۰.۰۰۰۱	۰.۰۰۲۵	۰.۰۰۳۴	۰.۰۰۰۳	۰.۰۰۰۴
۳	۰.۰۰۰۵	۰.۰۰۱۲	۰.۰۰۲۱	۰.۰۰۳۷	۰.۰۰۰۵	۰.۰۰۱۳
۴	۰.۰۰۰۵	۰.۰۰۱۱	۰.۰۰۲۲	۰.۰۰۲۷	۰.۰۰۰۵	۰.۰۰۰۹
۵	۰.۰۰۰۵	۰.۰۰۱۴	۰.۰۰۱۶	۰.۰۱۶۴	۰.۰۰۰۵	۰.۰۰۰۸
۶	۰.۰۰۰۵	۰.۰۰۱۲	۰.۰۰۲۳	۰.۰۰۰۸	۰.۰۰۰۶	۰.۰۰۰۶
۷	۰.۰۰۰۵	۰.۰۰۰۵	۰.۰۰۲۵	۰.۰۰۲۴	۰.۰۰۰۴	۰.۰۰۰۷
۸	۰.۰۰۰۸	۰.۰۰۱۸	۰.۰۰۲۲	۰.۰۰۲۱	۰.۰۰۰۳	۰.۰۰۰۶
۹	۰.۰۰۰۵	۰.۰۰۰۱	۰.۰۰۲۱	۰.۰۰۱۸	۰.۰۰۰۳	۰.۰۰۰۱
۱۰	۰.۰۰۰۴	۰.۰۰۱۳	۰.۰۰۲۳	۰.۰۰۲۱	۰.۰۰۰۷	۰.۰۰۰۷
۱۱	۰.۰۰۰۵	۰.۰۰۲۳	۰.۰۰۲۲	۰.۰۰۲۴	۰.۰۰۰۴	۰.۰۰۰۵
۱۲	۰.۰۰۰۷	۰.۰۰۰۱	۰.۰۰۲۲	۰.۰۰۳۴	۰.۰۰۰۳	۰.۰۰۰۵
۱۳	۰.۰۰۰۷	۰.۰۰۱۳	۰.۰۰۱۶	۰.۰۰۸۳	۰.۰۰۰۴	۰.۰۰۰۶
۱۴	۰.۰۰۰۴	۰.۰۰۰۴	۰.۰۰۲۳	۰.۰۰۱۲	۰.۰۰۰۴	۰.۰۰۰۸
۱۵	۰.۰۰۰۷	۰.۰۰۱۳	۰.۰۰۲۳	۰.۰۰۱۳	۰.۰۰۰۴	۰.۰۰۰۴
۱۶	۰.۰۰۰۴	۰.۰۰۱۲	۰.۰۰۲۲	۰.۰۰۲۷	۰.۰۰۰۵	۰.۰۰۰۴
۱۷	۰.۰۰۰۵	۰.۰۰۱۱	۰.۰۰۲۱	۰.۰۰۲۲	۰.۰۰۰۳	۰.۰۰۱۳
۱۸	۰.۰۰۰۵	۰.۰۰۱۸	۰.۰۰۲۱	۰.۰۰۲۳	۰.۰۰۰۳	۰.۰۰۰۴
۱۹	۰.۰۰۰۹	۰.۰۰۱۴	۰.۰۰۲۲	۰.۰۰۱۸	۰.۰۰۰۳	۰.۰۰۰۶

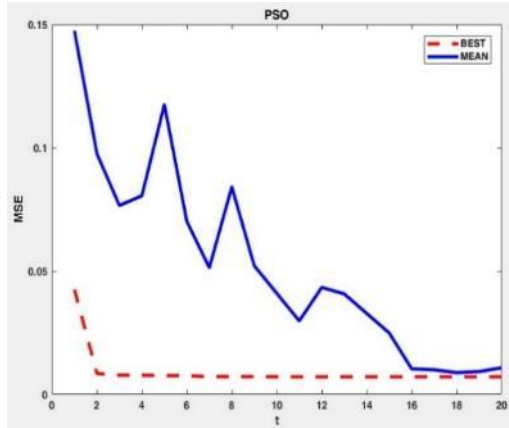
جدول ۶: مقایسه بهترین نتایج بدست آمده برای دو مدل پیشنهادی و شبکه عصبی برای داده‌های گروه اول

داده	نفت پارس		موتوزن		غدیر	
	آموزش	آزمون	آموزش	آزمون	آموزش	آزمون
مدل ترکیبی فازی - ازدحام ذرات	۰.۰۰۷۰	۰.۰۰۷۳	۰.۰۰۵۰	۰.۰۰۶۲	۰.۰۰۶۱	۰.۰۰۷۲
	کل	۰.۰۰۷۱	کل	۰.۰۰۵۴	کل	۰.۰۰۶۴
مدل یادگیری عمیق	۰.۰۰۰۴	۰.۰۰۰۴	۰.۰۰۱۶	۰.۰۰۰۳	۰.۰۰۰۴	۰.۰۰۰۴
	کل	۰.۰۰۰۴	کل	۰.۰۰۳۶	کل	۰.۰۰۰۳
شبکه عصبی	۰.۰۱۴۵	۰.۰۱۴۷	۰.۰۱۴۷	۰.۰۱۴۷	۰.۰۲۰۰	۰.۰۲۴۴
	کل	۰.۰۱۴۹	کل	۰.۰۱۵۰	کل	۰.۰۲۱۳

ترکیب دوگانه سیستم استنتاج فازی با الگوریتم بهینه‌سازی ازدحام ذرات در پیش‌بینی قیمت سهام بورس اوراق بهادار تهران و مقایسه آن با مدل یادگیری عمیق

این واقعیت در مقایسه تفاوت عملکرد پیش‌بینی مدل‌های یادگیری عمیق و شبکه عصبی بر روی سه شرکت که بدان اشاره شد کاملاً مشهود بوده و واریانس‌های نتایج بدست آمده برای این سه مدل نیز تایید کننده این واقعیت می‌باشند. مدل ترکیبی فازی-ازدحام ذرات با واریانس $5.69031E-07$ ، مدل یادگیری عمیق با واریانس $3.58E-06$ و مدل شبکه عصبی با واریانس $1.39E-05$ به ترتیب کمترین انحراف نتایج بدست آمده از شاخص متوسط خطای پیش‌بینی قیمت سهام برای سه شرکت را دارند.

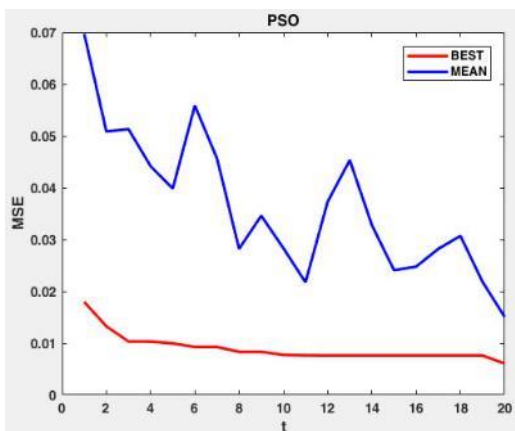
در جدول ۶ بهترین مقدار میانگین مجذور خطا (MSE) آموزش، آزمون و کل بدست آمده توسط دو مدل پیشنهادی و مدل شبکه عصبی برای هر یک از سه شرکت نفت پارس، موتوژن و غدیر آورده شده است. منظور از خطای کل، میانگین وزنی خطاهای آموزش و آزمون می‌باشد. بطور آشکار مشاهده می‌شود که مدل یادگیری عمیق نتایج بهتری در پیش‌بینی سهام شرکت‌های یاد شده کسب کرده است. اگرچه شکاف عملکردی مدل ترکیبی فازی-ازدحام ذرات از مدل یادگیری عمیق به مراتب کمتر بوده است.



شکل ۸: نمودار همگرایی PSO برای شرکت نفت پارس

	Low	P	High	P	First	P	VoIT	V	laue	T	F	Index	P	Index	USD	Gold	P	OilP	P	Price	Weight
Rule 1	N	H	H	N	H	M	L	H	H	N	L										0
Rule 2	H	H	N	H	H	M	N	H	H	H	H										0.9996
Rule 3	L	N	L	N	N	N	N	N	N	N	N										1
Rule 4	H	H	H	N	N	N	N	N	H	N	H										0.9982
Rule 5	N	H	H	N	H	H	H	H	N	H	N										0
Rule 6	N	N	L	L	H	N	H	L	N	H	L										0.9910
Rule 7	N	N	H	N	N	M	H	N	N	N	H										0.4079
Rule 8	H	N	N	H	N	N	H	H	H	N	N										0
Rule 9	H	N	H	N	N	H	N	N	H	H	H										1
Rule 10	N	N	N	H	H	H	N	N	N	H	H										0.0018
Rule 11	N	N	N	N	M	H	N	H	N	N	N										0.0732
Rule 12	H	H	H	H	N	N	N	N	L	H	N										1
Rule 13	N	N	M	N	H	H	N	N	N	N	N										0
Rule 14	N	N	N	N	N	N	L	N	N	N	L										1
Rule 15	H	N	M	H	H	L	H	H	H	N	H										0
Rule 16	N	H	N	N	N	H	H	N	N	N	H										0.7690
Rule 17	N	N	N	N	N	H	M	M	H	N	L										0
Rule 18	H	L	L	N	H	H	H	H	N	H	L										0
Rule 19	H	N	N	H	H	H	M	L	N	H	M										1
Rule 20	N	N	H	H	H	N	N	N	H	N	H										0.0401

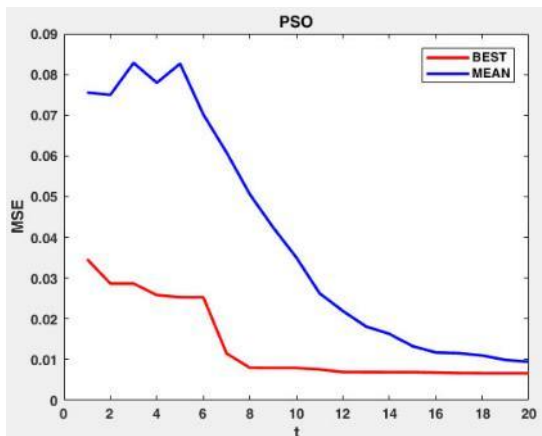
شکل ۵: قواعد فازی بدست آمده برای شرکت نفت پارس



شکل ۹: نمودار همگرایی PSO برای شرکت موتوژن

	Low	P	High	P	First	P	VoIT	V	laue	T	F	Index	P	Index	USD	Gold	P	OilP	P	Price	Weight
Rule 1	N	N	N	M	H	N	N	N	H	N	L										1
Rule 2	H	H	N	N	N	N	M	H	N	N	H										0.08...
Rule 3	N	N	N	H	H	N	H	H	N	L	M										0.99...
Rule 4	H	N	N	N	N	N	H	N	N	N	L										0.00...
Rule 5	N	N	H	H	H	M	H	H	H	N	L										8.58...
Rule 6	H	H	N	N	N	N	N	N	H	H	H										0
Rule 7	H	N	H	N	N	N	H	H	M	H	H										0.01...
Rule 8	H	N	M	N	N	H	H	N	H	N	H										1
Rule 9	L	N	L	N	L	N	N	N	N	N	L										1
Rule 10	L	N	H	H	N	H	N	N	N	N	M										0
Rule 11	N	H	H	N	H	H	N	H	N	H	H										0
Rule 12	H	H	N	H	H	N	N	N	L	H	L										1
Rule 13	N	H	L	H	H	N	H	H	N	H	H										0.91...
Rule 14	M	H	L	N	H	L	H	N	H	H	H										0
Rule 15	H	H	N	H	N	N	N	H	N	H	H										0
Rule 16	M	N	H	M	N	H	N	L	N	N	L										0
Rule 17	N	N	H	N	L	H	H	H	H	H	L										0.98...
Rule 18	M	L	L	L	L	N	M	L	N	H	H										0.17...
Rule 19	H	N	N	L	H	N	N	H	H	H	H										0
Rule 20	H	N	N	H	N	N	M	N	N	N	H										1

شکل ۶: قواعد فازی بدست آمده برای شرکت موتوژن



شکل ۱۰: نمودار همگرایی PSO برای شرکت غدیر

	Low	P	High	P	First	P	VoIT	V	laue	T	F	Index	P	Index	USD	Gold	P	OilP	P	Price	Weight
Rule 1	L	N	H	H	M	H	H	H	N	N	L										0.1551
Rule 2	N	H	L	N	N	H	H	N	N	N	M										1
Rule 3	M	N	N	M	N	L	L	H	N	N	M										1
Rule 4	M	N	N	H	N	L	N	H	L	L	M										1
Rule 5	L	N	N	H	H	N	H	H	H	H	L										0.8738
Rule 6	H	H	H	N	H	N	H	N	N	N	L										0
Rule 7	H	H	H	L	H	N	H	H	N	N	L										1
Rule 8	M	N	N	N	H	N	H	N	N	L	L										1
Rule 9	H	H	N	N	H	H	L	N	H	N	L										1
Rule 10	H	M	H	H	M	H	N	N	H	N	L										0.0530
Rule 11	N	L	N	N	N	N	M	H	H	H	H										0.9895
Rule 12	N	N	N	H	H	N	N	H	N	H	H										0.8866
Rule 13	H	N	N	N	N	N	H	M	N	H	H										1
Rule 14	N	N	N	H	H	N	N	H	H	H	H										0.0348
Rule 15	N	N	N	L	N	N	N	N	N	N	M										0.4046
Rule 16	L	N	N	N	N	N	N	N	N	N	L										1
Rule 17	N	N	H	N	H	N	N	N	N	N	L										0
Rule 18	N	H	H	H	N	H	N	N	M	H	H										0.4422
Rule 19	M	N	H	L	N	N	H	H	M	N	M										0.0713
Rule 20	H	N	N	H	H	H	H	N	H	H	L										0.0394

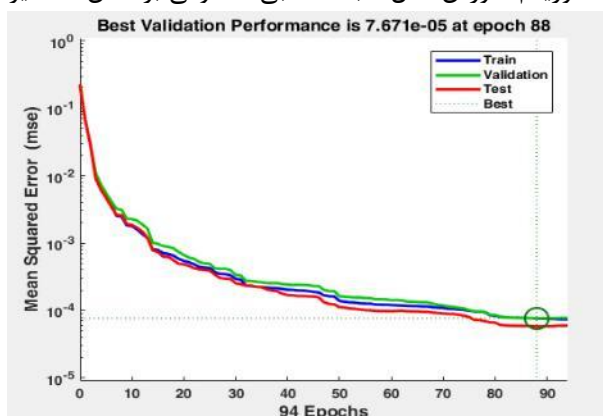
شکل ۷: قواعد فازی بدست آمده برای شرکت غدیر

معنی که متغیر مورد نظر در آن قانون فازی شرکت ندارد. هر ستون نیز متغیرهای ورودی و خروجی فازی و ستون آخر وزن قانون فازی در سیستم استنتاج فازی ممدانی مفروض را نشان می‌دهند. در تمامی این قوانین از عملگر AND به عنوان ترکیب کننده استفاده می‌شود. یکی از برتری‌های مدل

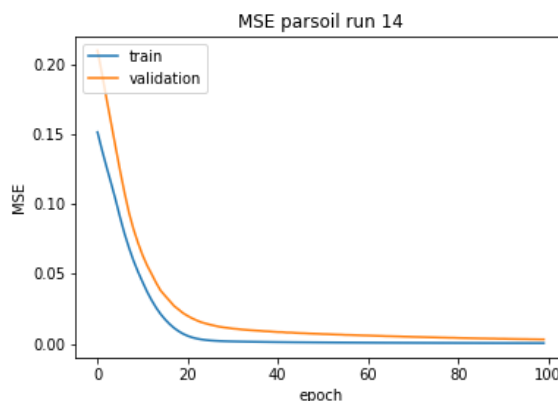
شکل‌های ۵ تا ۷ قوانین فازی بدست آمده توسط الگوریتم PSO برای موتور استنتاج شکل ۱ بر روی سه شرکت مفروض را نمایش می‌دهند. در این شکل‌ها هر ردیف یک قانون فازی را نشان می‌دهد که در آن L نشان دهنده متغیر فازی Low ، M نشان دهنده متغیر فازی Medium ، H نشان دهنده متغیر فازی High و N نشان دهنده Nounه است به این

ترکیب دوگانه سیستم استنتاج فازی با الگوریتم بهینه‌سازی ازدحام ذرات در پیش‌بینی قیمت سهام بورس اوراق بهادار تهران و مقایسه آن با مدل یادگیری عمیق

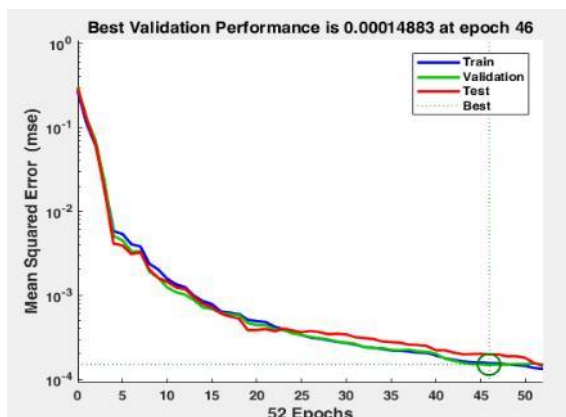
ترکیبی فازی-ازدحام ذرات، تفسیر راحت نتایج توسط کاربران است. این مهم توسط قوانین فازی بدست آمده محقق می‌گردد. اما در مورد مدل یادگیری عمیق و مدل شبکه عصبی این گونه تفسیری در دسترس نیست.



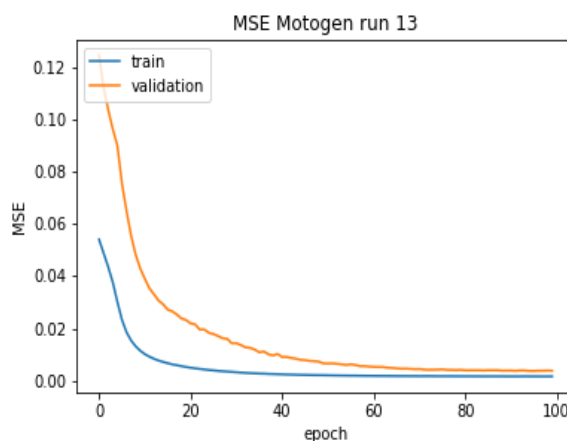
شکل ۱۴: نمودار همگرایی شبکه عصبی برای نفت پارس



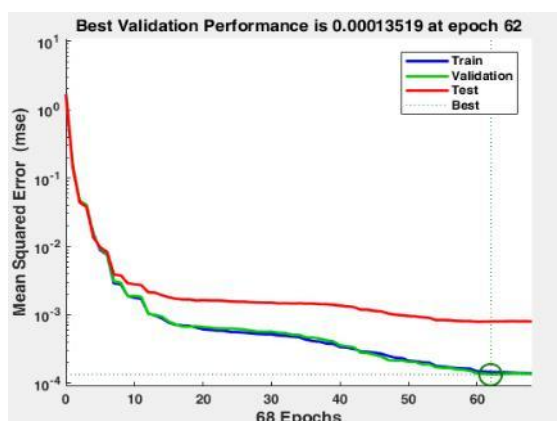
شکل ۱۱: نمودار همگرایی مدل یادگیر عمیق برای نفت پارس



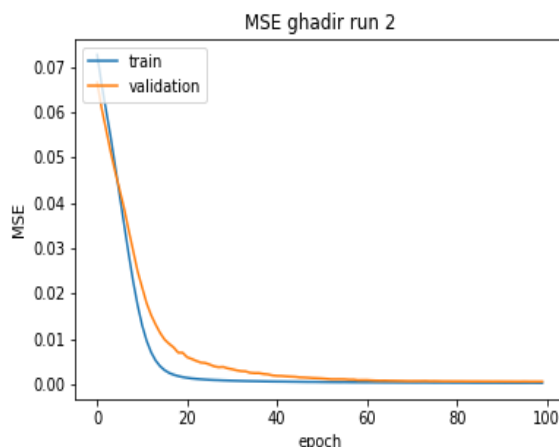
شکل ۱۵: نمودار همگرایی شبکه عصبی برای شرکت موتوزن



شکل ۱۲: نمودار همگرایی مدل یادگیری برای شرکت موتوزن



شکل ۱۶: نمودار همگرایی شبکه عصبی برای شرکت غدیر



شکل ۱۳: نمودار همگرایی مدل یادگیری برای شرکت غدیر

در شکل‌های ۸ الی ۱۰ نمودار همگرایی اجرای PSO براساس بهترین و متوسط کیفیت تابع برازش بدست آمده در هر تکرار برای داده‌های هر یک از شرکت‌ها نمایش داده شده

آموزش^{۵۲}، آزمایش^{۵۳} و اعتبارسنجی^{۵۴} نمایش داده شده است. در این نمودارها محور افقی تعداد تکرار الگوریتم و محور عمودی مقدار میانگین مجذور خطا را نشان می‌دهند.

⁵² Train

⁵³ Test

⁵⁴ Validation

ترکیب دوگانه سیستم استنتاج فازی با الگوریتم بهینه‌سازی ازدحام ذرات در پیش‌بینی قیمت سهام بورس اوراق بهادار تهران و مقایسه آن با مدل یادگیری عمیق

جدول ۸: جزئیات نتایج ۲۰ اجرا مدل ترکیبی فازی-ازدحام ذرات برای داده‌های گروه دوم

داده	مس ایران		ایران خودرو		پالایشگاه اصفهان		پتروشیمی خلیج فارس		لوتوس		بانک تجارت	
	آموزش	آزمون	آموزش	آزمون	آموزش	آزمون	آموزش	آزمون	آموزش	آزمون	آموزش	آزمون
۱	۰.۰۰۰۱۲	۰.۰۰۰۱۶	۰.۰۰۰۰۸	۰.۰۰۰۰۵	۰.۰۰۰۰۷	۰.۰۰۰۰۶	۰.۰۰۰۰۱	۰.۰۰۰۰۱	۰.۰۰۰۰۲۸	۰.۰۰۰۰۳۲	۰.۰۰۰۰۰	۰.۰۰۰۰۰
۲	۰.۰۰۰۰۱۴	۰.۰۰۰۰۰۶	۰.۰۰۰۰۰۴	۰.۰۰۰۰۰۳	۰.۰۰۰۰۰۴	۰.۰۰۰۰۰۳	۰.۰۰۰۰۰۳	۰.۰۰۰۰۰۲۵	۰.۰۰۰۰۰۴۴	۰.۰۰۰۰۰۵۷	۰.۰۰۰۰۰	۰.۰۰۰۰۰
۳	۰.۰۰۰۰۱۱	۰.۰۰۰۰۰۸	۰.۰۰۰۰۱۱	۰.۰۰۰۰۱۲	۰.۰۰۰۰۰۶	۰.۰۰۰۰۰۵	۰.۰۰۰۰۰۲	۰.۰۰۰۰۰۱۷	۰.۰۰۰۰۰۲۶	۰.۰۰۰۰۰۲۷	۰.۰۰۰۰۰	۰.۰۰۰۰۰
۴	۰.۰۰۰۰۰۹	۰.۰۰۰۰۱۰	۰.۰۰۰۰۰۸	۰.۰۰۰۰۰۸	۰.۰۰۰۰۱۶	۰.۰۰۰۰۰۳	۰.۰۰۰۰۰۲	۰.۰۰۰۰۰۲۲	۰.۰۰۰۰۰۱۶	۰.۰۰۰۰۰۲۲	۰.۰۰۰۰۰۳	۰.۰۰۰۰۰۲
۵	۰.۰۰۰۰۱۶	۰.۰۰۰۰۰۲۲	۰.۰۰۰۰۱۰	۰.۰۰۰۰۰۸	۰.۰۰۰۰۰۲۶	۰.۰۰۰۰۰۲	۰.۰۰۰۰۰۱	۰.۰۰۰۰۰۱۲	۰.۰۰۰۰۰۱۱	۰.۰۰۰۰۰۱۲	۰.۰۰۰۰۰	۰.۰۰۰۰۰
۶	۰.۰۰۰۰۰۶	۰.۰۰۰۰۰۹	۰.۰۰۰۰۱۲	۰.۰۰۰۰۰۹	۰.۰۰۰۰۰۶	۰.۰۰۰۰۰۶	۰.۰۰۰۰۰۱	۰.۰۰۰۰۰۱۳	۰.۰۰۰۰۰۱۳	۰.۰۰۰۰۰۱۶	۰.۰۰۰۰۰	۰.۰۰۰۰۰
۷	۰.۰۰۰۰۰۷	۰.۰۰۰۰۰۵	۰.۰۰۰۰۱۲	۰.۰۰۰۰۱۱	۰.۰۰۰۰۱۰	۰.۰۰۰۰۰۱	۰.۰۰۰۰۰۳	۰.۰۰۰۰۰۲۵	۰.۰۰۰۰۰۱۶	۰.۰۰۰۰۰۱۳	۰.۰۰۰۰۰	۰.۰۰۰۰۰
۸	۰.۰۰۰۰۰۸	۰.۰۰۰۰۰۵	۰.۰۰۰۰۰۷	۰.۰۰۰۰۱۶	۰.۰۰۰۰۰۹	۰.۰۰۰۰۰۵	۰.۰۰۰۰۰۱	۰.۰۰۰۰۰۱۶	۰.۰۰۰۰۰۴۵	۰.۰۰۰۰۰۴۶	۰.۰۰۰۰۰	۰.۰۰۰۰۰
۹	۰.۰۰۰۰۱۰	۰.۰۰۰۰۰۵	۰.۰۰۰۰۰۶	۰.۰۰۰۰۱۳	۰.۰۰۰۰۰۸	۰.۰۰۰۰۰۴	۰.۰۰۰۰۰۳	۰.۰۰۰۰۰۴۶	۰.۰۰۰۰۰۱۲	۰.۰۰۰۰۰۰۴	۰.۰۰۰۰۰	۰.۰۰۰۰۰
۱۰	۰.۰۰۰۰۰۶	۰.۰۰۰۰۰۳	۰.۰۰۰۰۰۴	۰.۰۰۰۰۰۴	۰.۰۰۰۰۱۰	۰.۰۰۰۰۰۲	۰.۰۰۰۰۰۲	۰.۰۰۰۰۰۲۴	۰.۰۰۰۰۰۱۶	۰.۰۰۰۰۰۱۸	۰.۰۰۰۰۰	۰.۰۰۰۰۰
۱۱	۰.۰۰۰۰۱۰	۰.۰۰۰۰۰۶	۰.۰۰۰۰۰۵۷	۰.۰۰۰۰۵۳	۰.۰۰۰۰۰۴۵	۰.۰۰۰۰۰۲	۰.۰۰۰۰۰۲	۰.۰۰۰۰۰۲۶	۰.۰۰۰۰۰۲۶	۰.۰۰۰۰۰۲۹	۰.۰۰۰۰۰۲	۰.۰۰۰۰۰۲
۱۲	۰.۰۰۰۰۱۸	۰.۰۰۰۰۱۲	۰.۰۰۰۰۱۸	۰.۰۰۰۰۱۹	۰.۰۰۰۰۰۶	۰.۰۰۰۰۰۶	۰.۰۰۰۰۰۳	۰.۰۰۰۰۰۳۰	۰.۰۰۰۰۰۱۶	۰.۰۰۰۰۰۰۸	۰.۰۰۰۰۰	۰.۰۰۰۰۰
۱۳	۰.۰۰۰۰۰۷	۰.۰۰۰۰۰۶	۰.۰۰۰۰۱۳	۰.۰۰۰۰۱۲	۰.۰۰۰۰۰۶	۰.۰۰۰۰۰۴	۰.۰۰۰۰۰۱	۰.۰۰۰۰۰۱۳	۰.۰۰۰۰۰۱۵	۰.۰۰۰۰۰۰۹	۰.۰۰۰۰۰	۰.۰۰۰۰۰
۱۴	۰.۰۰۰۰۰۵	۰.۰۰۰۰۰۴	۰.۰۰۰۰۰۷	۰.۰۰۰۰۱۰	۰.۰۰۰۰۰۱۷	۰.۰۰۰۰۰۱	۰.۰۰۰۰۰۱	۰.۰۰۰۰۰۱۵	۰.۰۰۰۰۰۳۱	۰.۰۰۰۰۰۵۱	۰.۰۰۰۰۰	۰.۰۰۰۰۰
۱۵	۰.۰۰۰۰۰۲۱	۰.۰۰۰۰۰۱۶	۰.۰۰۰۰۰۶	۰.۰۰۰۰۰۶	۰.۰۰۰۰۰۳۰	۰.۰۰۰۰۰۲	۰.۰۰۰۰۰۲	۰.۰۰۰۰۰۲۱	۰.۰۰۰۰۰۲۱	۰.۰۰۰۰۰۲۸	۰.۰۰۰۰۰	۰.۰۰۰۰۰
۱۶	۰.۰۰۰۰۰۸	۰.۰۰۰۰۰۶	۰.۰۰۰۰۱۳	۰.۰۰۰۰۰۹	۰.۰۰۰۰۰۱۸	۰.۰۰۰۰۰۷	۰.۰۰۰۰۰۱	۰.۰۰۰۰۰۱۳	۰.۰۰۰۰۰۱۵	۰.۰۰۰۰۰۱۲	۰.۰۰۰۰۰	۰.۰۰۰۰۰
۱۷	۰.۰۰۰۰۱۲	۰.۰۰۰۰۰۹	۰.۰۰۰۰۰۶	۰.۰۰۰۰۰۶	۰.۰۰۰۰۰۶	۰.۰۰۰۰۰۳	۰.۰۰۰۰۰۰	۰.۰۰۰۰۰۰۸	۰.۰۰۰۰۰۲۰	۰.۰۰۰۰۰۱۷	۰.۰۰۰۰۰	۰.۰۰۰۰۰
۱۸	۰.۰۰۰۰۱۵	۰.۰۰۰۰۰۸	۰.۰۰۰۰۰۴	۰.۰۰۰۰۰۳	۰.۰۰۰۰۰۳۲	۰.۰۰۰۰۰۲	۰.۰۰۰۰۰۲	۰.۰۰۰۰۰۱۸	۰.۰۰۰۰۰۱۴	۰.۰۰۰۰۰۲۴	۰.۰۰۰۰۰	۰.۰۰۰۰۰
۱۹	۰.۰۰۰۰۰۲۷	۰.۰۰۰۰۰۸	۰.۰۰۰۰۰۵	۰.۰۰۰۰۰۴	۰.۰۰۰۰۰۱۲	۰.۰۰۰۰۰۳	۰.۰۰۰۰۰۱	۰.۰۰۰۰۰۱۸	۰.۰۰۰۰۰۱۵	۰.۰۰۰۰۰۵۸	۰.۰۰۰۰۰	۰.۰۰۰۰۰
۲۰	۰.۰۰۰۰۰۳	۰.۰۰۰۰۰۳	۰.۰۰۰۰۰۶	۰.۰۰۰۰۰۶	۰.۰۰۰۰۰۵	۰.۰۰۰۰۰۶	۰.۰۰۰۰۰۰	۰.۰۰۰۰۰۱۱	۰.۰۰۰۰۰۱۳	۰.۰۰۰۰۰۱۲	۰.۰۰۰۰۰	۰.۰۰۰۰۰
بهترین	۰.۰۰۰۰۰۳	۰.۰۰۰۰۰۳	۰.۰۰۰۰۰۴	۰.۰۰۰۰۰۳	۰.۰۰۰۰۰۴	۰.۰۰۰۰۰۳	۰.۰۰۰۰۰۰	۰.۰۰۰۰۰۰۸	۰.۰۰۰۰۰۱۱	۰.۰۰۰۰۰۰۴	۰.۰۰۰۰۰	۰.۰۰۰۰۰
میانگین	۰.۰۰۰۰۱۱	۰.۰۰۰۰۰۸	۰.۰۰۰۰۱۰	۰.۰۰۰۰۱۰	۰.۰۰۰۰۱۳	۰.۰۰۰۰۰۱	۰.۰۰۰۰۰۱	۰.۰۰۰۰۰۱۹	۰.۰۰۰۰۰۲۰	۰.۰۰۰۰۰۲۴	۰.۰۰۰۰۰	۰.۰۰۰۰۰
ن	۲	۳	۸	۸	۹	۱	۳	۸	۶	۷	۴	۴

جدول ۹: جزئیات نتایج ۲۰ اجرا مدل یادگیری عمیق برای داده‌های گروه دوم

داده	مس ایران		ایران خودرو		پالایشگاه اصفهان		پتروشیمی خلیج فارس		لوتوس		بانک تجارت	
	آموزش	آزمون	آموزش	آزمون	آموزش	آزمون	آموزش	آزمون	آموزش	آزمون	آموزش	آزمون
۱	۰.۰۰۰۰۱۰	۰.۰۰۰۰۱۰	۰.۰۰۰۰۰۲	۰.۰۰۰۰۰۲	۰.۰۰۰۰۰۲	۰.۰۰۰۰۰۲	۰.۰۰۰۰۰۱	۰.۰۰۰۰۰۱۴	۰.۰۰۰۰۰۴۱	۰.۰۰۰۰۰۳۲	۰.۰۰۰۰۰۳	۰.۰۰۰۰۰۲
۲	۰.۰۰۰۰۰۹	۰.۰۰۰۰۰۸	۰.۰۰۰۰۰۲	۰.۰۰۰۰۰۱	۰.۰۰۰۰۰۲	۰.۰۰۰۰۰۲	۰.۰۰۰۰۰۱	۰.۰۰۰۰۰۱۳	۰.۰۰۰۰۰۴۱	۰.۰۰۰۰۰۳۹	۰.۰۰۰۰۰۴	۰.۰۰۰۰۰۳
۳	۰.۰۰۰۰۰۶	۰.۰۰۰۰۰۷	۰.۰۰۰۰۰۰	۰.۰۰۰۰۰۰	۰.۰۰۰۰۰۰	۰.۰۰۰۰۰۰	۰.۰۰۰۰۰۱	۰.۰۰۰۰۰۱۳	۰.۰۰۰۰۰۴۱	۰.۰۰۰۰۰۲۸	۰.۰۰۰۰۰۴	۰.۰۰۰۰۰۴

				۶		۲	۲	۲	۲			
۰,۰۰۰۰۴	۰,۰۰۰۰۳	۰,۰۰۰۰۳۶	۰,۰۰۰۰۴۱	۳	۰,۰۰۰۰۱۰	۱	۲	۱	۲	۰,۰۰۰۰۰۹	۰,۰۰۰۰۰۹	۴
۰,۰۰۰۰۳	۰,۰۰۰۰۳	۰,۰۰۰۰۴۲	۰,۰۰۰۰۴۱	۳	۰,۰۰۰۰۱۴	۳	۲	۲	۲	۰,۰۰۰۰۰۸	۰,۰۰۰۰۰۹	۵
۰,۰۰۰۰۳	۰,۰۰۰۰۳	۰,۰۰۰۰۳۷	۰,۰۰۰۰۴۱	۳	۰,۰۰۰۰۱۴	۱	۲	۳	۲	۰,۰۰۰۰۲۲	۰,۰۰۰۰۰۷	۶
۰,۰۰۰۰۳	۰,۰۰۰۰۲	۰,۰۰۰۰۳۳	۰,۰۰۰۰۴۱	۳	۰,۰۰۰۰۱۴	۲	۲	۳	۲	۰,۰۰۰۰۱۰	۰,۰۰۰۰۱۰	۷
۰,۰۰۰۰۲	۰,۰۰۰۰۳	۰,۰۰۰۰۵۲	۰,۰۰۰۰۴۴	۴	۰,۰۰۰۰۱۷	۲	۲	۶	۲	۰,۰۰۰۰۱۰	۰,۰۰۰۰۱۰	۸
۰,۰۰۰۰۳	۰,۰۰۰۰۳	۰,۰۰۰۰۴۲	۰,۰۰۰۰۴۱	۶	۰,۰۰۰۰۱۵	۱	۲	۲	۲	۰,۰۰۰۰۱۰	۰,۰۰۰۰۰۹	۹
۰,۰۰۰۰۳	۰,۰۰۰۰۳	۰,۰۰۰۰۴۲	۰,۰۰۰۰۴۰	۴	۰,۰۰۰۰۱۴	۲	۲	۱	۲	۰,۰۰۰۰۱۰	۰,۰۰۰۰۱۰	۱۰
۰,۰۰۰۰۳	۰,۰۰۰۰۳	۰,۰۰۰۰۴۶	۰,۰۰۰۰۴۰	۳	۰,۰۰۰۰۱۴	۱	۲	۲	۲	۰,۰۰۰۰۱۰	۰,۰۰۰۰۱۰	۱۱
۰,۰۰۰۰۳	۰,۰۰۰۰۳	۰,۰۰۰۰۲۸	۰,۰۰۰۰۴۱	۶	۰,۰۰۰۰۱۵	۳	۲	۲	۲	۰,۰۰۰۰۰۹	۰,۰۰۰۰۱۰	۱۲
۰,۰۰۰۰۳	۰,۰۰۰۰۳	۰,۰۰۰۰۳۶	۰,۰۰۰۰۴۱	۴	۰,۰۰۰۰۱۴	۱	۲	۳	۲	۰,۰۰۰۰۰۸	۰,۰۰۰۰۰۸	۱۳
۰,۰۰۰۰۴	۰,۰۰۰۰۴	۰,۰۰۰۰۴۳	۰,۰۰۰۰۴۰	۲	۰,۰۰۰۰۱۳	۴	۲	۲	۲	۰,۰۰۰۰۱۰	۰,۰۰۰۰۱۰	۱۴
۰,۰۰۰۰۳	۰,۰۰۰۰۴	۰,۰۰۰۰۵۲	۰,۰۰۰۰۴۰	۵	۰,۰۰۰۰۱۴	۲	۲	۱	۲	۰,۰۰۰۰۱۰	۰,۰۰۰۰۱۰	۱۵
۰,۰۰۰۰۵	۰,۰۰۰۰۴	۰,۰۰۰۰۴۵	۰,۰۰۰۰۴۰	۵	۰,۰۰۰۰۱۳	۴	۲	۱	۲	۰,۰۰۰۰۰۹	۰,۰۰۰۰۱۰	۱۶
۰,۰۰۰۰۵	۰,۰۰۰۰۴	۰,۰۰۰۰۵۱	۰,۰۰۰۰۴۰	۱	۰,۰۰۰۰۱۴	۲	۲	۴	۲	۰,۰۰۰۰۰۸	۰,۰۰۰۰۱۰	۱۷
۰,۰۰۰۰۶	۰,۰۰۰۰۴	۰,۰۰۰۰۵۶	۰,۰۰۰۰۴۴	۵	۰,۰۰۰۰۱۴	۱	۲	۲	۲	۰,۰۰۰۰۱۰	۰,۰۰۰۰۱۰	۱۸
۰,۰۰۰۰۳	۰,۰۰۰۰۴	۰,۰۰۰۰۳۹	۰,۰۰۰۰۴۱	۴	۰,۰۰۰۰۴۱	۲	۲	۲	۲	۰,۰۰۰۰۰۹	۰,۰۰۰۰۱۰	۱۹
۰,۰۰۰۰۵	۰,۰۰۰۰۴	۰,۰۰۰۰۶۲	۰,۰۰۰۰۴۲	۵	۰,۰۰۰۰۱۳	۲	۲	۱	۲	۰,۰۰۰۰۱۰	۰,۰۰۰۰۱۰	۲۰
۰,۰۰۰۰۲	۰,۰۰۰۰۲	۰,۰۰۰۰۲۸	۰,۰۰۰۰۰۴	۱	۰,۰۰۰۰۰۱	۱	۲	۱	۲	۰,۰۰۰۰۰۷	۰,۰۰۰۰۰۶	بهترین
۰,۰۰۰۰۳	۰,۰۰۰۰۳	۰,۰۰۰۰۴۲	۰,۰۰۰۰۴۱	۰,۰۰۰۰۰۱	۰,۰۰۰۰۱۵	۰,۰۰۰۰۰۰	۰,۰۰۰۰۰۰	۰,۰۰۰۰۰۰	۰,۰۰۰۰۰۰	۰,۰۰۰۰۰۹	۰,۰۰۰۰۰۹	میانگین
۵	۴	۰	۰	۴	۱	۲	۲	۲	۲	۸	۳	ن

جدول ۱۰: جزئیات نتایج ۲۰ اجرا مدل شبکه عصبی برای داده‌های گروه دوم

بانک تجارت		لوتوس		پتروشیمی خلیج فارس		پالایشگاه اصفهان		ایران خودرو		مس ایران		داده
آزمون	آموزش	آزمون	آموزش	آزمون	آموزش	آزمون	آموزش	آزمون	آموزش	آزمون	آموزش	خطا
۰,۰۰۰۹۳	۰,۰۰۰۹۳	۰,۰۰۱۱۸	۰,۰۰۱۱۰	۰,۰۰۰۷۱	۰,۰۰۰۶۹	۰,۰۰۱۴۷	۰,۰۰۱۴۳	۰,۰۰۰۶۶	۰,۰۰۰۶۸	۰,۰۰۱۰۵	۰,۰۰۱۰۶	۱
۰,۰۰۰۹۱	۰,۰۰۰۹۱	۰,۰۰۱۱۰	۰,۰۰۱۰۲	۰,۰۰۰۶۹	۰,۰۰۰۶۷	۰,۰۰۱۴۰	۰,۰۰۱۳۷	۰,۰۰۰۶۳	۰,۰۰۰۶۶	۰,۰۰۰۹۸	۰,۰۰۰۹۸	۲
۰,۰۰۰۸۷	۰,۰۰۰۸۷	۰,۰۰۱۰۶	۰,۰۰۰۹۹	۰,۰۰۰۶۴	۰,۰۰۰۶۲	۰,۰۰۱۲۷	۰,۰۰۱۲۵	۰,۰۰۰۶۰	۰,۰۰۰۹۳	۰,۰۰۰۹۵	۰,۰۰۰۹۶	۳
۰,۰۰۰۸۱	۰,۰۰۰۸۰	۰,۰۰۰۹۸	۰,۰۰۰۹۲	۰,۰۰۰۵۹	۰,۰۰۰۵۸	۰,۰۰۱۲۳	۰,۰۰۱۲۰	۰,۰۰۰۵۹	۰,۰۰۰۶۱	۰,۰۰۰۸۸	۰,۰۰۰۸۹	۴
۰,۰۰۰۷۵	۰,۰۰۰۷۵	۰,۰۰۰۹۷	۰,۰۰۰۹۱	۰,۰۰۰۵۶	۰,۰۰۰۵۴	۰,۰۰۱۱۳	۰,۰۰۱۱۰	۰,۰۰۰۵۵	۰,۰۰۰۵۷	۰,۰۰۰۹۷	۰,۰۰۰۹۷	۵
۰,۰۰۰۷۴	۰,۰۰۰۷۳	۰,۰۰۰۸۸	۰,۰۰۰۸۲	۰,۰۰۰۵۲	۰,۰۰۰۵۰	۰,۰۰۱۰۶	۰,۰۰۱۰۴	۰,۰۰۰۵۱	۰,۰۰۰۵۱	۰,۰۰۰۸۹	۰,۰۰۰۹۰	۶
۰,۰۰۰۷۱	۰,۰۰۰۷۱	۰,۰۰۰۸۴	۰,۰۰۰۷۸	۰,۰۰۰۵۱	۰,۰۰۰۴۹	۰,۰۰۱۰۰	۰,۰۰۰۹۸	۰,۰۰۰۴۸	۰,۰۰۰۵۰	۰,۰۰۰۸۷	۰,۰۰۰۸۸	۷
۰,۰۰۰۶۹	۰,۰۰۰۶۹	۰,۰۰۰۷۷	۰,۰۰۰۷۲	۰,۰۰۰۴۷	۰,۰۰۰۴۶	۰,۰۰۰۹۵	۰,۰۰۰۹۲	۰,۰۰۰۴۶	۰,۰۰۰۴۸	۰,۰۰۰۸۱	۰,۰۰۰۸۱	۸
۰,۰۰۰۶۵	۰,۰۰۰۶۵	۰,۰۰۰۷۰	۰,۰۰۰۶۵	۰,۰۰۰۴۶	۰,۰۰۰۴۴	۰,۰۰۰۸۸	۰,۰۰۰۸۶	۰,۰۰۰۴۵	۰,۰۰۰۴۷	۰,۰۰۰۹۹	۰,۰۰۱۰۰	۹
۰,۰۰۰۶۰	۰,۰۰۰۶۰	۰,۰۰۰۶۸	۰,۰۰۰۶۳	۰,۰۰۰۴۳	۰,۰۰۰۴۲	۰,۰۰۰۸۱	۰,۰۰۰۷۹	۰,۰۰۰۴۴	۰,۰۰۰۴۶	۰,۰۰۰۹۲	۰,۰۰۰۹۲	۱۰
۰,۰۰۰۶۰	۰,۰۰۰۵۹	۰,۰۰۰۶۴	۰,۰۰۰۶۰	۰,۰۰۰۴۲	۰,۰۰۰۴۱	۰,۰۰۰۷۹	۰,۰۰۰۷۷	۰,۰۰۰۴۰	۰,۰۰۰۴۲	۰,۰۰۰۹۰	۰,۰۰۰۹۱	۱۱
۰,۰۰۰۵۷	۰,۰۰۰۵۷	۰,۰۰۰۶۱	۰,۰۰۰۵۶	۰,۰۰۰۳۹	۰,۰۰۰۳۸	۰,۰۰۰۷۳	۰,۰۰۰۷۱	۰,۰۰۰۳۹	۰,۰۰۰۴۰	۰,۰۰۰۹۳	۰,۰۰۰۸۴	۱۲

ترکیب دوگانه سیستم استنتاج فازی با الگوریتم بهینه‌سازی ازدحام ذرات در پیش‌بینی قیمت سهام بورس اوراق بهادار تهران و مقایسه آن با مدل یادگیری عمیق

۰.۰۰۵۵	۰.۰۰۵۴	۰.۰۰۵۵	۰.۰۰۵۱	۰.۰۰۳۸	۰.۰۰۳۷	۰.۰۰۶۹	۰.۰۰۶۷	۰.۰۰۳۵	۰.۰۰۳۷	۰.۰۰۹۲	۰.۰۰۹۲	۱۳
۰.۰۰۴۹	۰.۰۰۵۰	۰.۰۰۵۰	۰.۰۰۴۷	۰.۰۰۳۶	۰.۰۰۳۵	۰.۰۰۶۵	۰.۰۰۶۴	۰.۰۰۳۳	۰.۰۰۳۵	۰.۰۰۸۵	۰.۰۰۸۶	۱۴
۰.۰۰۴۸	۰.۰۰۴۸	۰.۰۰۴۹	۰.۰۰۴۶	۰.۰۰۳۴	۰.۰۰۳۳	۰.۰۰۶۱	۰.۰۰۵۹	۰.۰۰۳۳	۰.۰۰۳۴	۰.۰۰۸۳	۰.۰۰۸۴	۱۵
۰.۰۰۴۵	۰.۰۰۴۴	۰.۰۰۴۶	۰.۰۰۴۳	۰.۰۰۳۲	۰.۰۰۳۱	۰.۰۰۵۷	۰.۰۰۵۶	۰.۰۰۳۱	۰.۰۰۳۲	۰.۰۰۷۷	۰.۰۰۷۸	۱۶
۰.۰۰۴۲	۰.۰۰۴۱	۰.۰۰۴۱	۰.۰۰۳۹	۰.۰۰۲۹	۰.۰۰۲۹	۰.۰۰۵۵	۰.۰۰۵۴	۰.۰۰۲۹	۰.۰۰۳۰	۰.۰۱۰۱	۰.۰۱۰۱	۱۷
۰.۰۰۴۰	۰.۰۰۳۹	۰.۰۰۴۰	۰.۰۰۳۷	۰.۰۰۲۹	۰.۰۰۲۸	۰.۰۰۵۱	۰.۰۰۵۰	۰.۰۰۲۶	۰.۰۰۲۷	۰.۰۰۹۳	۰.۰۰۹۴	۱۸
۰.۰۰۳۹	۰.۰۰۳۹	۰.۰۰۳۹	۰.۰۰۳۶	۰.۰۰۲۷	۰.۰۰۲۶	۰.۰۰۴۷	۰.۰۰۴۶	۰.۰۰۲۶	۰.۰۰۲۷	۰.۰۰۹۱	۰.۰۰۹۲	۱۹
۰.۰۰۳۸	۰.۰۰۳۷	۰.۰۰۳۷	۰.۰۰۳۵	۰.۰۰۲۴	۰.۰۰۲۳	۰.۰۰۴۴	۰.۰۰۴۳	۰.۰۰۲۶	۰.۰۰۲۷	۰.۰۰۸۴	۰.۰۰۸۵	۲۰
۰.۰۰۳۸	۰.۰۰۳۷	۰.۰۰۳۷	۰.۰۰۳۵	۰.۰۰۲۴	۰.۰۰۲۳	۰.۰۰۴۴	۰.۰۰۴۳	۰.۰۰۲۶	۰.۰۰۲۷	۰.۰۰۷۷	۰.۰۰۷۸	بهترین
۰.۰۰۶۱۹۵	۰.۰۰۶۱۶	۰.۰۰۶۹۹	۰.۰۰۶۵۲	۰.۰۰۴۴۴	۰.۰۰۴۳۱	۰.۰۰۸۶۰۵	۰.۰۰۸۴۰۵	۰.۰۰۴۲۷۵	۰.۰۰۴۵۹	۰.۰۰۹۱	۰.۰۰۹۱۲	میانگین

جدول ۱۱: جزئیات نتایج ۲۰ اجرای مدل ماشین بردار پشتیبان برای داده‌های گروه دوم

داده	مس ایران		ایران خودرو		پالایشگاه اصفهان		پتروشیمی خلیج فارس		لوتوس		بانک تجارت	
	آموزش	آزمون	آموزش	آزمون	آموزش	آزمون	آموزش	آزمون	آموزش	آزمون	آموزش	آزمون
۱	۰.۰۰۰۱	۰.۰۰۰۴	۰.۰۰۰۱	۰.۰۰۱۳	۰.۰۰۰۲	۰.۰۰۰۶	۰.۰۰۰۲	۰.۰۰۱۳	۰.۰۰۰۴	۰.۰۰۷۱	۰.۰۰۱۶	۰.۰۰۲۱
۲	۰.۰۰۰۲	۰.۰۰۰۳	۰.۰۰۰۱	۰.۰۰۰۱	۰.۰۰۰۲	۰.۰۰۰۷	۰.۰۰۰۱	۰.۰۰۵۴	۰.۰۰۹۵	۰.۰۲۶۱	۰.۰۰۰۲	۰.۰۰۰۱
۳	۰.۰۰۰۱	۰.۰۰۰۱	۰.۰۰۰۱	۰.۰۰۰۸	۰.۰۰۰۴	۰.۰۰۰۹	۰.۰۰۰۲	۰.۰۰۰۴	۰.۰۰۰۵	۰.۰۰۳۵	۰.۰۰۰۵	۰.۰۰۱۱
۴	۰.۰۰۰۱	۰.۰۰۰۱	۰.۰۰۰۱	۰.۰۰۰۱	۰.۰۰۰۳	۰.۰۰۱۱	۰.۰۰۰۵	۰.۰۰۰۵	۰.۰۰۰۶	۰.۰۰۲۸	۰.۰۰۱۵	۰.۰۰۱۲
۵	۰.۰۰۰۲	۰.۰۰۰۵	۰.۰۰۰۱	۰.۰۰۰۳	۰.۰۰۰۲	۰.۰۰۲۵	۰.۰۰۰۲	۰.۰۰۳۳	۰.۰۰۰۱	۰.۰۱۰۳	۰.۰۰۰۴	۰.۰۰۰۶
۶	۰.۰۰۰۶	۰.۰۰۰۵	۰.۰۰۰۳	۰.۰۰۰۶	۰.۰۰۰۲	۰.۰۰۰۶	۰.۰۰۰۳	۰.۰۰۰۷	۰.۰۰۰۶	۰.۰۰۳۹	۰.۰۰۱۵	۰.۰۰۰۸
۷	۰.۰۰۰۲	۰.۰۰۱۳	۰.۰۰۰۱	۰.۰۰۰۱	۰.۰۰۰۱	۰.۰۰۰۲	۰.۰۰۰۲	۰.۰۰۴۸	۰.۰۰۰۶	۰.۰۰۰۴	۰.۰۰۰۲	۰.۰۰۰۲
۸	۰.۰۰۰۱	۰.۰۰۰۳	۰.۰۰۰۷	۰.۰۰۱۳	۰.۰۰۰۳	۰.۰۰۰۷	۰.۰۰۰۲	۰.۰۰۴۳	۰.۰۰۰۴	۰.۰۰۰۲	۰.۰۰۵۵	۰.۰۰۲۸
۹	۰.۰۰۰۶	۰.۰۰۰۸	۰.۰۰۰۱	۰.۰۰۰۱	۰.۰۰۰۵	۰.۰۰۰۱	۰.۰۰۰۴	۰.۰۰۶۵	۰.۰۰۰۷	۰.۰۰۰۲	۰.۰۰۰۴	۰.۰۰۰۳
۱۰	۰.۰۰۰۱	۰.۰۰۰۱	۰.۰۰۰۱	۰.۰۰۰۱	۰.۰۰۰۴	۰.۰۰۰۸	۰.۰۰۰۲	۰.۰۰۲۲	۰.۰۰۰۹	۰.۰۰۷۵	۰.۰۰۰۲	۰.۰۰۰۲
۱۱	۰.۰۰۰۴	۰.۰۰۰۷	۰.۰۰۰۱	۰.۰۰۰۲	۰.۰۰۰۲	۰.۰۰۱۳	۰.۰۰۰۷	۰.۰۰۰۷	۰.۰۱۰۶	۰.۰۲۰۱	۰.۰۰۰۳	۰.۰۰۰۸
۱۲	۰.۰۰۰۳	۰.۰۰۰۶	۰.۰۰۰۲	۰.۰۰۰۴	۰.۰۰۰۳	۰.۰۰۰۳	۰.۰۰۰۷	۰.۰۰۰۷	۰.۰۰۰۶	۰.۰۰۴۹	۰.۰۰۰۲	۰.۰۰۰۳
۱۳	۰.۰۰۰۴	۰.۰۰۰۶	۰.۰۰۰۱	۰.۰۰۰۲	۰.۰۰۰۳	۰.۰۰۰۸	۰.۰۰۰۳	۰.۰۰۳۵	۰.۰۰۰۹	۰.۰۱۰۳	۰.۰۰۰۳	۰.۰۰۰۳
۱۴	۰.۰۰۰۲	۰.۰۰۰۲	۰.۰۰۰۱	۰.۰۰۰۲	۰.۰۰۰۲	۰.۰۰۰۶	۰.۰۰۰۳	۰.۰۰۳۸	۰.۰۰۰۶	۰.۰۰۰۳	۰.۰۰۰۲	۰.۰۰۰۷
۱۵	۰.۰۰۰۲	۰.۰۰۰۵	۰.۰۰۰۱	۰.۰۰۰۲	۰.۰۰۰۲	۰.۰۰۰۴	۰.۰۰۰۲	۰.۰۰۰۹	۰.۰۰۰۶	۰.۰۰۲۲	۰.۰۰۰۱	۰.۰۰۰۳
۱۶	۰.۰۰۰۲	۰.۰۰۰۳	۰.۰۰۰۱	۰.۰۰۰۲	۰.۰۰۰۳	۰.۰۰۰۳	۰.۰۰۰۴	۰.۰۰۲۶	۰.۰۰۰۴	۰.۰۰۲۲	۰.۰۰۰۵	۰.۰۰۰۶
۱۷	۰.۰۰۰۳	۰.۰۰۰۴	۰.۰۰۰۱	۰.۰۰۰۱	۰.۰۰۰۴	۰.۰۰۰۹	۰.۰۰۰۲	۰.۰۰۱۹	۰.۰۰۰۵	۰.۰۰۲۹	۰.۰۰۰۵	۰.۰۰۰۵
۱۸	۰.۰۰۰۳	۰.۰۰۰۴	۰.۰۰۰۲	۰.۰۰۲۵	۰.۰۰۰۵	۰.۰۰۰۸	۰.۰۰۰۲	۰.۰۰۲۱	۰.۰۰۰۷	۰.۰۰۲۳	۰.۰۰۰۲	۰.۰۰۰۳
۱۹	۰.۰۰۰۲	۰.۰۰۰۶	۰.۰۰۰۱	۰.۰۰۰۱	۰.۰۰۰۲	۰.۰۰۰۳	۰.۰۰۰۳	۰.۰۰۱۷	۰.۰۰۰۸	۰.۰۰۴۱	۰.۰۰۰۲	۰.۰۰۰۲
۲۰	۰.۰۰۰۱	۰.۰۰۰۲	۰.۰۰۰۱	۰.۰۰۰۱	۰.۰۰۰۲	۰.۰۰۰۵	۰.۰۰۰۳	۰.۰۰۰۱	۰.۰۰۰۵	۰.۰۰۲۳	۰.۰۰۰۳	۰.۰۰۰۵
بهترین	۰.۰۰۰۱	۰.۰۰۰۱	۰.۰۰۰۱	۰.۰۰۰۱	۰.۰۰۰۱	۰.۰۰۰۲	۰.۰۰۰۱	۰.۰۰۰۴	۰.۰۰۰۴	۰.۰۰۰۲	۰.۰۰۰۱	۰.۰۰۰۱
میانگین	۰.۰۰۰۲۴	۰.۰۰۰۴۴	۰.۰۰۰۱	۰.۰۰۰۴	۰.۰۰۰۳	۰.۰۰۰۷۶	۰.۰۰۰۳۰	۰.۰۰۰۲۷	۰.۰۰۰۱۵	۰.۰۰۶۳۲	۰.۰۰۰۰۷	۰.۰۰۰۰۷۴
ن	۵	۵	۵	۵	۷	۵	۵	۳	۷	۵	۴	۵

جدول ۱۲: مقایسه بهترین نتایج بدست آمده برای دو مدل پیشنهادی، ماشین بردار پشتیبان و شبکه عصبی برای

داده‌های گروه دوم

مدل ترکیبی فازی - ازدحام ذرات	مس ایران		ایران خودرو		پالایشگاه اصفهان		پتروشیمی خلیج فارس		لوتوس		بانک تجارت	
	آموزش	آزمون	آموزش	آزمون	آموزش	آزمون	آموزش	آزمون	آموزش	آزمون	آموزش	آزمون
مدل یادگیری عمیق	کل	۰.۰۰۰۰۳	۰.۰۰۰۰۳۷	۰.۰۰۰۰۳۷	۰.۰۰۰۰۳۷	۰.۰۰۰۰۳۷	۰.۰۰۰۰۶۵	۰.۰۰۰۰۱۱۳	۰.۰۰۰۰۴۰	۰.۰۰۰۰۲۷	۰.۰۰۰۰۲	۰.۰۰۰۰۲
	آموزش	۰.۰۰۰۰۶	۰.۰۰۰۰۲	۰.۰۰۰۰۲	۰.۰۰۰۰۲	۰.۰۰۰۰۲	۰.۰۰۰۰۱۰	۰.۰۰۰۰۴۰	۰.۰۰۰۰۱۰	۰.۰۰۰۰۴۰	۰.۰۰۰۰۲	۰.۰۰۰۰۲
مدل شبکه عصبی	کل	۰.۰۰۰۰۶۳	۰.۰۰۰۰۱۷	۰.۰۰۰۰۱۷	۰.۰۰۰۰۱۷	۰.۰۰۰۰۱۷	۰.۰۰۰۰۱۰۹	۰.۰۰۰۰۴۰۶	۰.۰۰۰۰۲۳	۰.۰۰۰۰۳۷	۰.۰۰۰۰۳	۰.۰۰۰۰۳
	آموزش	۰.۰۰۰۰۷۸	۰.۰۰۰۰۲۷	۰.۰۰۰۰۲۷	۰.۰۰۰۰۴۳	۰.۰۰۰۰۴۳	۰.۰۰۰۰۲۳	۰.۰۰۰۰۳۵	۰.۰۰۰۰۳۷	۰.۰۰۰۰۲۸	۰.۰۰۰۰۲۸	۰.۰۰۰۰۲۸

۰,۰۰۰۳۷۳	۰,۰۰۰۳۵۶	۰,۰۰۰۲۳۳	۰,۰۰۰۴۳۳	۰,۰۰۰۲۶۷	۰,۰۰۰۷۷۷	کل	مدل ماشین بردار پشتیبان
۰,۰۰۰۰۱	۰,۰۰۰۰۴	۰,۰۰۰۰۴	۰,۰۰۰۰۱	۰,۰۰۰۰۱	۰,۰۰۰۰۱	آموزش	
۰,۰۰۰۰۳	۰,۰۰۰۲۰	۰,۰۰۰۰۱	۰,۰۰۰۰۲	۰,۰۰۰۰۱	۰,۰۰۰۰۱	آزمون	
۰,۰۰۰۰۱۶	۰,۰۰۰۰۸۸	۰,۰۰۰۰۳۱	۰,۰۰۰۰۱۳	۰,۰۰۰۰۱	۰,۰۰۰۰۱	کل	

نسبت به دو مدل دیگر کندتر می‌باشد زیرا تعداد تکرارها مدل شبکه عصبی برای هر سه شرکت بیشتر از ۴۵ تکرار می‌باشد. اگرچه همگرایی دو مدل پیشنهاد شده در شکل‌های ۸ الی ۳ در کمتر از ۲۰ تکرار محقق شده است ولی می‌توان مشاهده نمود که سرعت همگرایی مدل ترکیبی فازی-عصبی از مدل یادگیری عمیق بیشتر است.

۴-۲- تحلیل نتایج داده‌های گروه دوم

داده‌های گروه دوم شامل اطلاعات شش شرکت مس ایران، ایران خودرو، پالایشگاه اصفهان، پتروشیمی خلیج فارس، لوتوس و بانک تجارت طی دوره زمانی ۱۳۹۴ تا پایان ۱۳۹۸ است. دو مدل پیشنهادی، شبکه عصبی [۱۷] و ماشین بردار پشتیبان [۲۳]، هر کدام ۲۰ بار اجرا شده و نتایج بدست آمده برای داده‌های آموزش و آزمون آنها در جداول ۸، ۹، ۱۰ و ۱۱ نمایش داده شده است. از مقایسه کلی نتایج بدست آمده در این جداول مشاهده می‌شود که اگرچه مدل ترکیبی فازی-ازدحام ذرات موفق شده به میزان ۹۸٪ نتایج بهتری از شبکه عصبی و به میزان ۸۸٪ نتایج بهتری از مدل ماشین بردار پشتیبان داشته باشند ولی در مقایسه با یادگیری عمیق عملکرد ضعیف‌تری داشته و مدل یادگیری عمیق به میزان ۱۷٪ موفق‌تر از مدل ترکیبی فازی-ازدحام ذرات، ۹۸٪ موفق‌تر از شبکه عصبی و ۹۰٪ موفق‌تر از مدل ماشین بردار پشتیبان در پیش‌بینی قیمت سهام عمل کرده است. برتری ۱۷٪ مدل یادگیری عمیق در مقابل مدل ترکیبی فازی-ازدحام ذرات در حالی است که مدل ترکیبی فازی-ازدحام ذرات برای شرکت‌های مس ایران، پتروشیمی خلیج فارس و لوتوس توانسته پیش‌بینی دقیق‌تری را ارائه دهد.

همچنین از مقایسه نتایج جدول ۸ مشاهده می‌شود که بهترین عملکرد مدل ترکیبی فازی-ازدحام ذرات در پیش‌بینی قیمت سهام مس ایران بوده بگونه‌ای که عملکرد این مدل بر روی داده شرکت مس به میزان ۱۰٪ بهتر عملکرد این مدل بر روی داده شرکت ایران خودرو، به میزان ۲۲٪ بهتر از شرکت پالایشگاه اصفهان، به میزان ۴۸٪ بهتر از شرکت پتروشیمی خلیج فارس، به میزان ۵۷٪ بهتر از شرکت لوتوس و ۳۰٪ بهتر از بانک تجارت بوده است.

با مقایسه نمودارهای همگرایی PSO از مدل ترکیبی فازی-ازدحام ذرات، همگرایی مدل یادگیری عمیق و شبکه عصبی مشاهده می‌شود که همگرایی مدل شبکه عصبی

	LowP	HighP	FirstP	VolT	VlaueT	FIndex	PIndex	USD	GoldP	OilP	PPrice	Weight
Rule1	L	N	M	M	N	L	M	H	M	N	H	1
Rule2	H	M	L	M	N	M	H	N	L	L	M	0
Rule3	H	L	N	H	H	N	H	H	M	M	M	0.6432
Rule4	H	H	M	L	N	M	L	H	L	M	H	0.0625
Rule5	N	M	H	L	L	H	N	N	N	N	M	1
Rule6	H	N	L	N	N	H	M	L	L	M	H	0.7789
Rule7	H	H	M	N	L	H	H	L	N	H	H	0.0252
Rule8	H	L	H	L	N	M	M	L	L	M	M	0
Rule9	N	N	N	H	L	H	M	L	L	N	L	0.9827
Rule10	N	H	H	L	N	L	M	H	L	N	L	0.0285
Rule11	H	H	L	H	M	N	M	N	L	N	H	0.4901
Rule12	H	M	H	H	N	N	M	H	N	H	1	
Rule13	L	N	H	H	L	M	H	N	H	M	H	0
Rule14	L	M	H	N	L	N	N	H	H	N	M	0.6122
Rule15	H	M	L	H	N	L	L	N	L	L	1	
Rule16	M	N	L	M	N	M	N	N	M	N	L	0.0558
Rule17	N	N	M	L	L	H	H	N	M	H	M	0.5621
Rule18	N	M	H	M	H	H	N	M	H	L	H	1
Rule19	N	N	L	N	M	M	H	M	M	N	L	0.9414
Rule20	N	H	M	N	N	L	H	N	M	M	H	0.1053

شکل ۲۰: قواعد فازی بدست آمده برای شرکت لوتوس

	LowP	HighP	FirstP	VolT	VlaueT	FIndex	PIndex	USD	GoldP	OilP	PPrice	Weight
Rule1	H	H	N	N	N	N	N	N	N	N	L	0
Rule2	H	L	N	M	N	H	N	N	N	M	M	0.0057
Rule3	N	N	H	H	N	H	H	M	H	H	L	1
Rule4	N	H	M	L	N	N	N	H	H	L	M	0.1200
Rule5	L	M	H	H	H	L	N	H	L	H	L	0.8932
Rule6	H	N	N	M	N	H	H	M	H	H	H	1
Rule7	N	H	L	L	H	M	N	M	H	H	L	0.8499
Rule8	N	H	L	N	H	N	M	H	H	N	H	1
Rule9	H	N	M	N	N	L	H	H	N	L	H	0
Rule10	L	H	H	L	H	L	H	N	N	N	L	0.4118
Rule11	N	M	L	H	L	N	N	N	L	N	H	0.3077
Rule12	M	L	L	M	N	N	M	N	M	M	L	1
Rule13	H	N	M	N	L	N	N	H	N	H	L	0.5695
Rule14	H	N	H	M	N	H	M	N	N	L	H	0.5956
Rule15	N	M	H	H	H	L	H	M	M	H	1	
Rule16	N	L	M	H	H	N	N	L	L	N	H	1.9160e...
Rule17	H	M	M	H	H	L	N	L	M	H	L	0.3986
Rule18	L	M	M	L	M	N	M	M	H	L	L	0
Rule19	H	M	N	H	N	L	M	L	M	L	L	0.8255
Rule20	H	N	N	L	M	N	H	L	H	H	1	

شکل ۲۱: قواعد فازی بدست آمده برای شرکت بانک تجارت

	LowP	HighP	FirstP	VolT	VlaueT	FIndex	PIndex	USD	GoldP	OilP	PPrice	Weight
Rule1	L	L	N	M	H	N	N	H	N	H	H	0.9110
Rule2	N	N	H	H	L	H	N	N	N	N	L	0.7187
Rule3	H	L	L	N	H	H	N	H	H	H	H	0.6127
Rule4	N	N	N	N	H	L	H	N	N	H	H	1
Rule5	H	H	N	M	N	H	N	N	L	M	H	0.4496
Rule6	H	N	M	N	N	H	N	L	H	N	L	0
Rule7	H	L	N	N	N	L	H	H	N	H	L	0.7178
Rule8	N	N	M	H	L	N	H	L	H	H	H	0.5136
Rule9	H	H	N	H	M	M	N	N	H	H	H	0.4477
Rule10	H	M	H	L	N	H	N	L	L	N	H	0.8178
Rule11	H	H	H	M	N	H	N	N	N	N	H	1
Rule12	N	N	H	H	H	N	N	L	L	H	0.0738	
Rule13	L	N	H	N	N	M	H	M	N	L	L	1
Rule14	N	L	N	N	M	N	N	M	M	L	1	
Rule15	M	H	N	N	M	H	N	H	N	L	0.0181	
Rule16	H	N	H	N	H	H	M	H	H	H	0.1092	
Rule17	N	H	N	N	M	N	H	H	M	L	1	
Rule18	H	M	H	N	L	H	M	L	M	H	0.4086	
Rule19	H	H	N	H	N	N	N	H	N	H	0.0568	
Rule20	M	L	H	N	H	N	N	H	L	M	L	0.0395

شکل ۲۲: قواعد فازی بدست آمده برای شرکت

ترکیب دوگانه سیستم استنتاج فازی با الگوریتم بهینه‌سازی ازدحام ذرات در پیش‌بینی قیمت سهام بورس اوراق بهادار تهران و مقایسه آن با مدل یادگیری عمیق

۲٪، پتروشیمی خلیج فارس ۸۶٪، لوتوس ۹۵٪ و بانک تجارت ۴۲٪ می‌باشد. نتایج جدول ۱۰ نشان می‌دهند که مدل شبکه عصبی در پیش‌بینی قیمت سهام شرکت پتروشیمی خلیج فارس نسبت به سایر شرکت‌ها عملکرد بهتری داشته است. این مهم در مقایسه خطای پیش‌بینی قیمت سهام شرکت پتروشیمی خلیج فارس نسبت به مس ایران به میزان ۵۲٪، نسبت به ایران خودرو ۱٪، نسبت به پالایشگاه اصفهان ۴۹٪، نسبت به لوتوس ۳۵٪ و نسبت به بانک تجارت ۲۹٪ می‌باشد. تحلیل آماری نتایج مدل ماشین بردار پشتیبان در جدول ۱۱ نشان از موفقیت این مدل در پیش‌بینی قیمت سهام ایران خودرو را دارد. بگونه ای که خطای پیش‌بینی قیمت سهام مدل ماشین بردار پشتیبان برای ایران خودرو ۱۳٪ بهتر از خطای پیش‌بینی سهام مس ایران، ۴۷٪ بهتر از خطای پیش‌بینی سهام پالایشگاه اصفهان، ۸۰٪ بهتر از پتروشیمی خلیج فارس، ۹۲٪ بهتر از لوتوس و ۶۰٪ بهتر از خطای پیش‌بینی سهام بانک تجارت بوده است.

جدول ۱۲ به منظور تسهیل در مقایسه کلی عملکرد مدل ترکیبی فازی-ازدحام ذرات و مدل یادگیری عمیق به دو مدل شبکه عصبی مصنوعی و ماشین بردار پشتیبان طراحی شده است که حاوی بهترین نتایج بدست آمده توسط این چهار مدل بر روی شش داده شرکت‌های گروه دوم می‌باشند. شکل‌های ۱۷ تا ۲۲ قوانین فازی بدست آمده توسط الگوریتم PSO برای موتور استنتاج شکل ۱ بر روی شش شرکت داده‌های گروه دوم را نمایش می‌دهند. همانطور که اشاره شد یکی از برتری‌های مدل ترکیبی فازی-ازدحام ذرات، تفسیری بالای مجموعه قوانین بدست آمده موتور استنتاج به عنوان الگوی خروجی این مدل توسط تحلیل گران بازار سرمایه است.

۵- نتیجه‌گیری

در این مقاله، یک مدل ترکیبی براساس استنتاج فازی ممدانی و الگوریتم بهینه‌سازی ازدحام (شکل ۱) و یک مدل یادگیری عمیق (شکل ۲) به منظور پیش‌بینی قیمت سهام در بورس اوراق بهادار ارائه گردید. ورودی مدل‌های پیشنهادی دارای ۱۰ متغیرها اقتصادی شامل کمترین قیمت سهام، بالاترین قیمت سهام، قیمت اولیه سهام، حجم معاملات، ارزش معاملات، شاخص بازار اول تالار بورس، قیمت شاخص کل بازار، نرخ ارز دلار، قیمت جهانی هر اونس طلا، قیمت جهانی نفت بوده و خروجی آنها نیز

	LowP	HighP	FirstP	VolIT	VlaueT	FIndex	PIndex	USD	GoldP	OilP	PPrice	Weight
Rule 1	M	L	H	M	M	N	L	H	N	M	M	0.5275
Rule 2	M	N	N	N	L	M	H	L	H	H	M	0.7380
Rule 3	N	L	N	H	N	N	L	M	N	N	H	1
Rule 4	H	N	H	N	M	N	N	H	N	H	H	0.6339
Rule 5	M	L	L	M	M	N	N	M	N	M	L	0.7691
Rule 6	L	M	L	N	N	H	L	M	L	M	M	0.1687
Rule 7	N	N	N	M	M	N	N	H	H	L	L	0
Rule 8	L	N	N	H	H	M	L	H	H	N	H	0
Rule 9	N	H	M	N	H	N	N	M	H	L	M	0.7626
Rule 10	L	H	L	N	M	H	M	N	L	M	L	0.4492
Rule 11	M	M	M	L	M	L	M	N	L	H	L	0.7315
Rule 12	N	M	H	H	M	H	N	N	N	M	M	0.0560
Rule 13	M	H	M	H	H	N	L	L	L	H	L	0.9276
Rule 14	N	H	L	H	L	N	M	L	H	N	M	1
Rule 15	N	N	H	N	N	L	N	L	H	L	L	0.0185
Rule 16	H	M	N	M	H	N	N	M	M	M	H	0.9960
Rule 17	N	M	M	N	N	L	H	H	M	H	M	0.8751
Rule 18	N	L	H	N	N	L	H	L	M	L	L	0.2739
Rule 19	M	L	L	H	H	H	L	M	N	N	M	0.1414
Rule 20	M	M	M	M	M	N	L	M	M	H	L	0

شکل ۱۷: قواعد فازی بدست آمده برای شرکت مس ایران

	LowP	HighP	FirstP	VolIT	VlaueT	FIndex	PIndex	USD	GoldP	OilP	PPrice	Weight
Rule 1	N	N	L	N	L	H	M	H	N	H	H	0
Rule 2	N	L	L	N	L	M	M	N	H	H	L	1
Rule 3	H	H	H	H	H	M	H	L	N	N	M	1
Rule 4	H	M	H	L	N	H	N	L	N	N	L	0.9038
Rule 5	N	H	N	L	M	N	M	N	M	N	H	0.0583
Rule 6	L	L	M	N	N	M	N	M	N	N	L	0.4879
Rule 7	H	H	N	M	N	L	M	H	N	M	H	0.9280
Rule 8	H	N	N	N	H	N	H	N	L	N	L	1
Rule 9	L	N	N	N	L	L	L	M	H	N	H	0.9550
Rule 10	L	N	N	N	L	H	H	N	L	L	H	0
Rule 11	L	H	N	L	L	N	N	H	H	H	H	1
Rule 12	N	H	H	L	H	N	N	M	N	H	H	0.7329
Rule 13	L	M	M	H	H	L	N	L	H	H	L	0.0287
Rule 14	L	N	N	N	H	H	L	L	L	N	H	0
Rule 15	N	L	L	M	N	M	H	N	N	N	H	0
Rule 16	H	M	N	M	H	N	M	M	L	H	H	0.8016
Rule 17	N	M	N	H	H	H	L	M	M	M	H	0
Rule 18	H	L	L	N	H	N	N	H	N	M	M	0.4795
Rule 19	N	H	M	N	M	H	H	N	N	L	L	0.9650
Rule 20	H	N	N	N	L	N	H	N	N	L	H	0.6277

شکل ۱۸: قواعد فازی بدست آمده برای شرکت ایران خودرو

	LowP	HighP	FirstP	VolIT	VlaueT	FIndex	PIndex	USD	GoldP	OilP	PPrice	Weight
Rule 1	L	H	L	M	N	H	M	L	L	M	L	0.1906
Rule 2	M	N	H	N	N	N	H	N	H	M	H	1
Rule 3	H	N	H	L	H	N	N	H	M	L	M	0.2987
Rule 4	L	H	N	H	N	M	H	M	N	H	M	0.9957
Rule 5	L	N	N	N	N	N	N	N	N	M	L	1
Rule 6	L	M	N	N	L	M	N	M	N	L	L	0.9808
Rule 7	L	M	N	H	N	M	H	M	H	L	H	0
Rule 8	N	N	M	H	H	N	N	N	N	L	H	0.0916
Rule 9	N	H	N	N	H	M	M	N	H	M	H	0.1683
Rule 10	H	M	N	H	M	N	N	N	N	N	L	0.1979
Rule 11	L	H	N	H	M	N	N	H	N	L	H	0.0642
Rule 12	L	N	L	H	H	H	H	M	L	H	L	0.9972
Rule 13	M	N	H	L	M	H	L	L	L	N	H	0
Rule 14	N	M	N	L	N	L	H	N	L	L	L	0.9854
Rule 15	N	M	M	N	M	L	H	H	H	L	L	0.0843
Rule 16	H	M	N	N	L	N	N	N	H	L	L	0.6771
Rule 17	H	L	H	M	M	N	N	N	M	H	H	0.0766
Rule 18	N	L	M	H	H	N	L	M	N	L	L	1
Rule 19	L	N	H	L	L	L	H	H	H	M	M	0.9522
Rule 20	M	H	H	H	N	H	N	L	N	L	L	0.7490

شکل ۱۹: قواعد فازی بدست آمده برای شرکت

پالایشگاه اصفهان

تحلیل آماری نتایج جدول ۹ نیز بیانگر عملکرد بهتر مدل یادگیری عمیق بر روی داده شرکت پالایشگاه اصفهان نسبت به سایر شرکت‌ها می‌باشد بگونه ای که این برتری نسبت به شرکت مس ایران به میزان ۷۹٪، ایران خودرو

۴. جهانتیغ، ف.، د.پ. تلگردویی، and صفورا، وقفه های زمانی بهینه در پیش بینی قیمت نفت توسط شبکه عصبی پویا اصلاح شده با الگوریتم ژنتیک. فصلنامه مطالعات اقتصاد انرژی، ۲۰۱۸. ۱۴(۵۶): ۱۱۵-۱۴۳.

۵. قربانی، et al.، پیش بینی سیگنال معاملات سهام با استفاده از شبکه های پتری رنگی و الگوریتم ژنتیک (مطالعه موردی: بازار بورس تهران). پژوهشنامه مدیریت اجرایی، ۲۰۱۹. ۱۱(۲۱): ۲۰۵-۲۲۷.

۶. نژاد، ف. and مینایی، پیش بینی رفتار بازار سهام بر اساس شبکه های عصبی مصنوعی با رویکرد یادگیری جمعی هوشمند. مدیریت صنعتی، ۲۰۱۸. ۱۰(۲): ۳۱۵-۳۳۴.

۷. رضانی and عاملی، پیش بینی قیمت سهام با استفاده از شبکه عصبی فازی مبتنی بر الگوریتم ژنتیک و مقایسه با شبکه عصبی فازی. تحقیقات مدل سازی اقتصادی، ۲۰۱۶. ۶(۲۲): ۶۱-۹۱.

۸. باباجانی، et al.، پیش بینی قیمت سهام در بورس تهران با استفاده از شبکه عصبی بازگشتی بهینه شده با الگوریتم کلونی زنبور عسل مصنوعی. راهبرد مدیریت مالی، ۲۰۱۹. ۷(۲): ۱۹۵-۲۲۸.

9. Kim, T. and H.Y. Kim, Forecasting stock prices with a feature fusion LSTM-CNN model using different representations of the same data. PloS one, 2019. 14(۲)

10. monajemi, abzari, and rayati, Stock price prediction in stock exchange stock exchange using fuzzy neural network and genetic algorithm and comparing it with artificial neural network. Quarterly Journal of Economics, 2010. 3(6): p. 1-26.

11. Hájek, P., V. Olej, and R. Myskova, Forecasting stock prices using sentiment information in annual reports: A neural network and support vector regression approach. WSEAS Transactions on Business and Economics, 2013. 10(4): p. 293-305.

12. Hadavandi, E., H. Shavandi, and A. Ghanbari, Integration of genetic fuzzy systems and artificial neural networks for stock price forecasting. Knowledge-Based Systems, 2010. 23(8): p. 800-808.

قیمت سهام است. برای یافتن بهترین موتور استنتاج، از PSO برای جستجو در بین ترکیب های ممکن ۲۰ تایی از ۳۱۴۵۷۲۸ قانون فازی استفاده شده است. به منظور ارزیابی عملکرد دو مدل پیشنهادی و مقایسه آنها با مدل مرسوم شبکه عصبی مصنوعی و ماشین بردار پشتیبان، دو گروه داده شرکت ها در نظر گرفته شده. در گروه اول، داده های سه شرکت فعال در بورس اوراق بهادار تهران همچون نفت پارس، موتوژن و غدیر در بازه ۱۳۸۹ تا پایان ۱۳۹۴ مستخرج و مدل ها بر روی آنها اجرا شدند. در گروه دوم، داده های شش شرکت مس ایران، ایران خودرو، پالایشگاه اصفهان، پتروشیمی خلیج فارس، لوتوس و بانک تجارت در بازه ۱۳۹۴ تا پایان ۱۳۹۸ مستخرج و مدل ها بر روی آنها پیاده سازی و اجرا شدند.

با تحلیل و مقایسه نتایج بدست آمده مشاهده شد که مدل یادگیری عمیق دارای خطای پیش بینی قیمت سهام کمتری نسبت به دو مدل ترکیبی فازی-ازدحام ذرات و شبکه عصبی بوده ولی رفتار مدل ترکیبی در پیش بینی قیمت سهام سه شرکت همسانتر و با واریانس به مراتب کمتر بود. همچنین سرعت همگرایی مدل ترکیبی فازی-ازدحام ذرات از سرعت همگرایی مدل یادگیری عمیق بیشتر بود.

منابع

1. Farmer, J.D. and A.W. Lo, *Frontiers of finance: Evolution and efficient markets*. Proceedings of the National Academy of Sciences, 1999. 96(18): p. 9991-9992.
2. Vachhani, H., et al. *Machine learning based stock market analysis: A short survey*. in *International Conference on Innovative Data Communication Technologies and Application*. 2019. Springer.
3. Jain, V.R., M. Gupta, and R.M. Singh, *Analysis and Prediction of Individual*

- method to the stock trend prediction. *Expert Systems with Applications*, 2009. 36(8): p. 10896-10904.
- 26.Chen, Y. and Y. Hao, A feature weighted support vector machine and K-nearest neighbor algorithm for stock market indices prediction. *Expert Systems with Applications*, 2010. 37(3): p. 340-355.
- 27.Nair, B.B., V. Mohandas, and N. Sakthivel, A decision tree—rough set hybrid system for stock market trend prediction. *International Journal of Computer Applications*, 2010. 6(9): p. 1-6.
- 28.Qiu, W., X. Liu, and L. Wang, Forecasting shanghai composite index based on fuzzy time series and improved C-fuzzy decision trees. *Expert Systems with Applications*, 2012. 39(9): p. 7680-7689.
- 29.Basak, S., et al., Predicting the direction of stock market prices using tree-based classifiers. *The North American Journal of Economics and Finance*, 2019. 47: p. 552-567.
- 30.Khaidem, L., S. Saha, and S.R. Dey, Predicting the direction of stock market prices using random forest. *arXiv preprint arXiv:1605.00003*, 2016.
- 31.Sharma, N. and A. Juneja. Combining of random forest estimates using LSboost for stock market index prediction. in *2017 2nd International Conference for Convergence in Technology (I2CT)*. 2017. IEEE.
۳۲. الهام، غ. and د. سیدمحمد رضا، پیش‌بینی روند قیمت در بازار سهام با استفاده از الگوریتم جنگل تصادفی. فصلنامه مهندسی مالی و مدیریت اوراق بهادار ۱۳۹۷. ۹(۳۵): p. 301-322.
- 33.Alkhatib, K., et al., Stock price prediction using k-nearest neighbor (kNN) algorithm. *International Journal of Business, Humanities and Technology*, 2013. 3(3): p. 32-44.
۳۴. زاده، et al. پیش‌بینی قیمت سهام با استفاده از روش خود رگرسیون با وقفه توزیعی (ARDL). *تحقیقات مالی*. ۱۳۸۶. ۹(۲۳): p. 49-60.
- 13.Chen, Y ,et al., Hybrid methods for stock index modeling. *Fuzzy Systems and Knowledge Discovery*, 2005: p. 490-490.
- 14.Wang, S., et al., Stock price prediction based on chaotic hybrid particle swarm optimisation-RBF neural network. *International Journal of Applied Decision Sciences*, 2017. 10(2): p. 89-100.
- 15.Khuat, T.T. and M.H. Le, An Application of Artificial Neural Networks *Stock Prices of Financial Sector Companies in NIFTY50*. *International Journal of Information Engineering and Electronic Business*, 2018. 11(2): p. 33.
- 16.Ghasemiyeh, R., R. Moghdani, and S.S. Sana, A Hybrid Artificial Neural Network with Metaheuristic Algorithms for Predicting Stock Price. *Cybernetics and Systems*, 2017. 48(4): p. 365-392.
- 18.Rajihy, Y., K. Nermend, and A. Alsakaa, Back-propagation artificial neural networks in stock market forecasting. An application to the Warsaw Stock Exchange WIG20. *Aestimatio*, 2017(15): p. 88.
۲۰. موسوی، س. علیرضا، and غلامی، استفاده از الگوریتم ترکیبی عصبی کرم شب‌تاب و روش رگولاسیون بی‌زین جهت پیش‌بینی قیمت سهام. مهندسی مالی و مدیریت اوراق بهادار، ۱۳۹۷. ۹(۳۶): p. 295-321.
- 21.Fischer, T. and C. Krauss, Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, 2018. 270(2): p. 654-669.
- 22.Long, W., Z. Lu, and L. Cui, Deep learning-based feature engineering for stock price movement prediction. *Knowledge-Based Systems*, 2019. 164: p. 163-173.
- 23.Kelotra, A. and P. Pandey, Stock market prediction using optimized deep-convlstm model. *Big Data*, 2020. 8 (۱): p. 5-24.
- 24.Xiao, C., W. Xia, and J. Jiang, Stock price forecast based on combined model of ARI-MA-LS-SVM. *Neural Computing and Applications*, 2020: p. 1-10.
- 25.Lee, M.-C., Using support vector machine with a hybrid feature selection

44. Nhu, H.N., S. Nitsuwat, and M. Sodanil. Prediction of stock price using an adaptive Neuro-Fuzzy Inference System trained by Firefly Algorithm. in 2013 International Computer Science and Engineering Conference (ICSEC). 2013. IEEE.
45. Dash, R. and P. Dash, Efficient stock price prediction using a self evolving recurrent neuro-fuzzy inference system optimized through a modified differential harmony search technique. *Expert Systems with Applications*, 2016. 52: p. 75-90.
46. Wei, L.-Y., A hybrid model based on ANFIS and adaptive expectation genetic algorithm to forecast TAIEX. *Economic Modelling*, 2013. 33: p. 893-899.
47. Bagheri, A., H.M. Peyhani, and M. Akbari, Financial forecasting using ANFIS networks with quantum-behaved particle swarm optimization. *Expert Systems with Applications*, 2014. 41(14): p. 6235-6250.
48. Han, J., J. Pei, and M. Kamber, *Data mining :concepts and techniques*. 2011: Elsevier.
49. Bova, S., et al. A logical analysis of Mamdani-type fuzzy inference, I theoretical bases. in *International Conference on Fuzzy Systems*. 2010. IEEE.
50. Kennedy, J. and R. Eberhart, *Particle Swarm Optimization*. IEEE, 1995: p. 1942-1948.
51. Engelbrecht, A.P., *Computational intelligence: an introduction*. 2 ed. 2007, England: John Wiley & Sons. 597.
52. Werbos, P.J., *Beyond Regression: New Tools for Prediction and Analysis in the Behavioural Sciences*. 1974, Harvard University: Boston, USA.
۳۵. قلی‌زاده، م.ح.، et al.، پیش‌بینی قیمت سهام با روش رگرسیون فازی. *پژوهشنامه اقتصاد کلان*، ۱۳۹۰، ۶(۱۲): p. 107-128
36. Kita, E., M. Harada, and T. Mizuno, Application of Bayesian Network to stock price prediction. *Artif. Intell. Research*, 2012. 1(2): p. 171-184.
37. Sun, Q., W.-G. Che, and H.-L. Wang, Bayesian regularization BP neural network model for the stock price prediction, in *Foundations and applications of intelligent systems*. 2014, Springer. p. 521-531.
38. Wang, L., et al., Stock market trend prediction using dynamical Bayesian factor graph. *Expert Systems with Applications*, 2015. 42(15-16): p. 6267-6275.
39. Hassan, M.R., et al., A HMM-based adaptive fuzzy inference system for stock market forecasting. *Neurocomputing*, 2013. 104: p. 10-25.
40. Chang, P.-C. and C.-H. Liu, A TSK type fuzzy rule based system for stock price prediction. *Expert Systems with applications*, 2 (۱): ۳۴ . . ۱۳۵-۱۴۴.
41. Lincy, G.R.M. and C.J. John, A multiple fuzzy inference systems framework for daily stock trading with application to NASDAQ stock exchange. *Expert Systems with Applications: An International Journal*, 2016. 44(C): p. 13-21.
42. Chandar, S.K., Fusion model of wavelet transform and adaptive neuro fuzzy inference system for stock market prediction. *Journal of Ambient Intelligence and Humanized Computing*, 2019: p. 1-9.
43. Feylizadeh, M.R., M.H. Keshavarz, and A. Hendalianpour, Presenting a model for predicting the Tehran Stock Exchange Index using ANFIS and fuzzy regression. *Journal of New Researches in Mathematics*, 2019.

ارائه روشی با استفاده از الگوریتم ژنتیک در تشخیص موضع افراد جامعه در رسانه‌های خبری و اجتماعی

*مهدی سالخورده حقیقی **سیدمحمد ابراهیمی

*استادیار، دانشکده مهندسی کامپیوتر و فناوری اطلاعات - دانشگاه صنعتی سجاد - مشهد - ایران

**دانشکده مهندسی کامپیوتر و فناوری اطلاعات - دانشگاه صنعتی سجاد - مشهد - ایران

تاریخ پذیرش: ۱۳۹۹/۰۵/۰۲

تاریخ دریافت: ۱۳۹۸/۰۲/۲۹

چکیده

گزارش‌های خبری ارائه شده در رسانه‌های اجتماعی و خبری با انواع اسناد و مدارک ارائه می‌شوند و شامل موضوعاتی هستند که جوامع و نظرات مختلف را در برمی‌گیرند. آگاهی از رابطه میان افراد در اسناد می‌تواند به خوانندگان کمک کند تا یک دانش اولیه در خصوص موضوع و هدف در اسناد مختلف به دست آورند. در این مقاله، روشهای تشخیص جوامع بررسی شده و تکنیک‌های مختلف خوشه‌بندی افرادی که نام آن‌ها در مجموعه‌ای از اسناد خبری آورده شده است نیز مورد بررسی قرار می‌گیرد. این افراد در جوامعی خوشه‌بندی می‌شوند که مواضع مرتبط با یکدیگر دارند. در این مقاله یک روش تشخیص موضع افراد جامعه مبتنی بر یک شبکه دوستی به عنوان مکانیزم پایه معرفی شده و مکانیزم تشخیص جوامع بهبود یافته‌ای بر مبنای آن ارائه گردیده است. در روش پیشنهادی از ساختار الگوریتم ژنتیک جهت بهبود نرخ تشخیص جوامع بهبود یافته‌ای بر مبنای آن ارائه گردیده است. در روش در نظر گرفته شده است که برای رسیدن به این هدف شاخص رند نیز استفاده گردیده است. نتایج حاصل از آزمایش‌ها که بر مبنای پایگاه‌های داده‌ی واقعی اسناد انتشار یافته در رسانه خبری گوگل نیوز در رابطه با یک موضوع خاص به دست آمده‌اند، حاکی از کارآمدی و بهره‌وری مطلوب روش پیشنهادی است.

واژه‌های کلیدی: رسانه‌های اجتماعی، رسانه‌های خبری، تشخیص جوامع، تشخیص موضع افراد

۱- مقدمه

با توجه به گستردگی استفاده از این شبکه‌ها، تبادل نظرها بین کاربران این شبکه‌ها نیز به سرعت و با حجم زیادی صورت می‌گیرد. از این رو کاربران این شبکه‌ها به دنبال افراد و گروه‌هایی هستند که نظرات مشابه یا متضادی دارند تا به ارزیابی این

امروزه شبکه‌های اجتماعی در بسیاری از زمینه‌های اجتماعی، سیاسی، اقتصادی، فرهنگی، هنری، ورزشی، آموزشی و زمینه‌های دیگر بطور گسترده‌ای مورد استفاده قرار می‌گیرند. کاربران این شبکه‌ها نیز طیف‌های وسیعی از افراد با سطوح دانش متفاوتی از این شبکه‌ها را تشکیل می‌دهند.

گسترده هستند و قابلیت ارتباط آن‌ها را به روش‌های مختلف فراهم می‌کند. آن‌ها همچنین تعامل و اشتراک اطلاعات را با افراد مختلف از جمله بستگان، همکاران، خانواده، دوستان، طرفداران و دیگران تسهیل می‌نمایند [3]. علاوه بر تسهیل ارتباطات، رسانه‌های مذکور، به‌روزرسانی، دوست داشتن، دوست نداشتن، ایجاد پروفایل‌ها و به اشتراک‌گذاری اطلاعات شخصی و عمومی را فراهم می‌کنند [4].

یکی از چالش‌های اساسی در مورد تجزیه و تحلیل رسانه-ها/شبکه‌های اجتماعی، کشف خودکار جوامع است [5]. جوامع به‌عنوان گروه‌ها، خوشه‌ها، یا زیرگروه‌ها در موقعیت‌های مختلف می‌باشند و کشف جامعه در یک رسانه اجتماعی و خبری به معنی شناسایی مجموعه‌ای از گره‌ها است که با یکدیگر ارتباطات بیشتری نسبت به سایرین داشته باشند.

اکثر محققان بر این باورند که تشخیص جوامع در شبکه-های اجتماعی با مسائل خوشه‌بندی در داده‌کاوی قابل مقایسه است. خوشه‌بندی در داده‌کاوی نوعی یادگیری بدون نظارت است که هدف آن تقسیم یک مجموعه داده بزرگ به گروه‌های همگن (خوشه) می‌باشد [6]. در حقیقت، کشف یک جامعه می‌تواند داده‌کاوی بر روی گراف‌ها باشد. علاوه بر این، تشخیص جوامع، بزرگ‌ترین حوزه مطالعه از کاربردهای داده‌کاوی در رسانه‌های اجتماعی و خبری است. کاربردهای دیگر مانند گراف‌کاوی در مراحل اولیه توسعه قرار دارند [7]. روش‌های تحلیل گراف می‌توانند درک عمیق‌تری از ساختار شبکه‌های اجتماعی ارائه نمایند. البته روش‌های مبتنی بر گراف هزینه محاسباتی بالایی از نظر زمان و فضا دارند که استفاده از آن‌ها را در این رابطه محدود می‌نماید [8].

در این پژوهش، هدف اصلی تشخیص جوامع بر اساس موضع افراد در مجموعه‌ای از اسنادی است که در رسانه‌های اجتماعی و خبری منتشر می‌شوند به‌طوری‌که در نهایت افراد در گروه‌های مختلف با مواضع مرتبط خوشه‌بندی می‌شوند. در همین راستا، در این پژوهش شکل بهبود یافته یک روش خاص تشخیص جوامع به نام شناسایی جامعه موضع افراد براساس تحلیل شبکه دوستی ارائه شده است [9]. این بهینه‌سازی در جهت بهبود تشخیص جوامع و افزایش دقت خوشه‌بندی و همچنین کار بر روی چالش مقداردهی اولیه و محلی بودن در روش

بپردازند. در نتیجه یافتن اطلاعات مورد نظر کاربران در میان انبوه اطلاعاتی که بطور پیوسته در این شبکه‌ها مبادله می‌شود یکی از چالش‌هایی است که کاربرانی این شبکه‌ها با آنها مواجه هستند.

علاوه بر موارد ذکرشده، تحلیل اطلاعات مبادله شده بین کاربران این شبکه‌ها می‌تواند راهی برای تحلیل رفتار آنها باشد تا بتوان از انجام برخی رفتارهای ناهنجار در این شبکه‌ها اطلاع حاصل نمود و از آنها جلوگیری کرد. لذا تشخیص وابستگی افراد به جوامع و گروه‌های مختلف از جمله زمینه‌های تحقیق جذاب در این شبکه‌ها می‌باشد.

روش پیشنهادی در این پژوهش یک روش بهبودیافته بر مبنای تشخیص جوامع بر اساس یک شبکه دوستی (گراف اولیه ارتباط بین گره‌ها) است. به عنوان نمونه، این روش گروه‌های مختلف افراد با گرایش‌ها و تمایلات یکسان را تشخیص می‌دهد. جوامع تشخیص داده شده برای آشنایی با پدیده‌های اجتماعی و تعاملات انسانی کاربرد خواهند داشت. اساساً بر طبق این روش، یک شبکه دوستی براساس نام افراد در ساختار الگوریتم ژنتیک تعریف شده و براساس یک تابع برازندگی، افراد در چند خوشه دسته بندی می‌شوند. به‌طور خلاصه، این پژوهش به بررسی مسئله‌ی تشخیص جامعه یا گروه افراد هدف (بر مبنای موضع آن‌ها) می‌پردازد تا در نهایت بتواند افراد هدف در جوامع با مواضع یکسان را به‌صورت بهبودیافته خوشه‌بندی نماید.

رسانه‌های اجتماعی به عنوان گروهی از برنامه‌های مبتنی بر اینترنت تعریف شده‌اند و امکان ایجاد و تبادل محتوای تولید شده توسط کاربران را می‌دهند. رسانه‌های اجتماعی متفاوت از یکدیگرند؛ انواع پایگاه‌های رسانه‌های اجتماعی شامل رسانه‌های سنتی مانند روزنامه، رادیو و تلویزیون و رسانه‌های غیر سنتی مانند فیس‌بوک، توئیتر و غیره می‌باشند [1]. به‌طور جزئی‌تر، از رسانه‌های اجتماعی با عنوان شبکه‌های اجتماعی نیز یاد می‌شود. رسانه‌های خبری، نوعی از مجموعه رسانه‌های اجتماعی هستند که هدف آن‌ها ارائه اخبار به عموم مردم یا جامعه‌ی هدفی از مردم است. این موارد شامل رسانه‌های چاپی (روزنامه، مجله‌های خبری)، پخش اخبار (رادیو و تلویزیون و اخیراً اینترنت)، روزنامه‌های برخط، وبلاگ‌های خبری، و غیره است [2].

رسانه‌های اجتماعی و خبری توان بالقوه‌ی قابل‌توجهی در ارتباطات و تعامل بین افرادی دارند که از لحاظ جغرافیایی

استخراج شده و برای استفاده بعدی به یک ساختار قابل درک تبدیل می‌گردد. نویسندگان از هفت ویژگی کاربر (مانند سن، جنس و غیره) برای خوشه‌بندی و تجزیه و تحلیل استفاده کردند. در نتیجه، مشخص گردید ویژگی‌های فیس بوکی انتخاب شده، تعیین کننده خوشه‌های مهم و معنا داری بودند.

تحقیق‌های مختلفی در زمینه تشخیص جوامع در شبکه‌های اجتماعی صورت گرفته است که در ادامه برخی از آنها بطور خلاصه مورد بررسی قرار می‌گیرند. در تحقیق‌های جداگانه‌ای، چن و همکارانش [15] [16] به مسئله تشخیص جوامع همت گمارده‌اند. آن‌ها در این راستا، استفاده از روش تحلیل مؤلفه‌های اصلی [17] را معرفی و تشریح کرده‌اند. به بیان مشخص، آن‌ها به بررسی نشانه‌های ورودی‌های بردار ویژه‌ای که با بزرگ‌ترین مقدار ویژه در ارتباط است، پرداخته‌اند تا بتوانند گروه‌های مربوط به موضع افراد را شناسایی نمایند. این روش می‌تواند تنها مباحث و موضوعاتی که دو موضع در آن‌ها مطرح است را تحت پوشش قرار دهد؛ اما در عمل، در بسیاری از موارد، بیش از دو موضع در مباحث و موضوعات مطرح می‌شود.

چن و همکارانش [9] روشی برای شناسایی گروه یا اجتماعات ارائه نمودند که مبتنی بر موضع و جایگاه آنها می‌باشد. به موجب آن، به‌طور خودکار و با استناد به سندها و مدرک‌های مورد نظر، یک شبکه‌ی دوستی برای افراد هدف تهیه می‌شود. روش‌های توسعه جامعه و اصلاح جامعه با هدف تعریف جوامعی با مواضع و جایگاه‌های مشخص از افراد هدف در شبکه‌ی دوستی طراحی شده‌اند. همچنین شناسایی افرادی که موضع آن‌ها با موضع هدف نامرتبط است نیز با این روش‌ها قابل انجام است. نتایج حاصل از آزمایش‌ها و تجربیاتی که بر مبنای پایگاه‌های داده‌ی واقعی و حقیقی به عمل آمده است، حاکی از کارآمدی و بهره‌وری مطلوب این روش دارند و نشان می‌دهند که این روش، کارایی بهتری از بسیاری از رویکردهای معروف تشخیص جوامع و الگوریتم‌های خوشه‌بندی دارد.

یادهوتموجو و همکارانش [18] یک روش تشخیص جامعه در یک شبکه استنادی با استفاده از الگوریتم رتبه پیوند ارائه نمودند. مقاله مذکور کاربرد یک الگوریتم تشخیص جامعه در یک شبکه استنادی را توصیف می‌نماید. هدف، تشخیص جوامع بر اساس روابط استنادی و تجزیه و تحلیل شباهت‌های موضوعات در هر جامعه است. نتایج آزمایش‌ها

مذکور است. در نهایت نیز روش پیشنهادی با چند روش دیگر خوشه‌بندی مورد مقایسه قرار گرفته که نتایج آزمایش‌ها نشان دهنده برتری روش پیشنهادی نسبت به سایر روش‌ها است.

۲- پیشینه تحقیق

اطلاعات رسانه‌های اجتماعی و خبری در قالب داده‌های متنی بدون ساختار مانند پیام‌های چت و اسناد متنی است. استخراج دستی اطلاعات از منابع متنی، که می‌تواند به یک پایگاه داده‌ی ساخت یافته و مناسب برای تجزیه و تحلیل بیشتر تبدیل شود، به‌ویژه هنگامی که داده‌ها حجیم باشند، قطعاً ناکارآمد و گاهی غیر ممکن است. بنابراین، استخراج خودکار اطلاعات یک رویکرد کارآمدتر در داده‌کاوی رسانه‌های اجتماعی برای تشخیص فعالیت‌های مجرمانه و همچنین روابط بالقوه میان شبکه است [10].

یکی از چالش‌های اساسی در مورد تجزیه و تحلیل رسانه‌های اجتماعی و خبری، کشف خودکار جوامع است [5]. جوامع به‌عنوان گروه‌ها، خوشه‌ها یا زیرگروه‌ها در مناطق مختلف دیده می‌شوند و کشف جامعه در یک رسانه اجتماعی به معنی شناسایی مجموعه‌ای از گره‌ها است که با یکدیگر ارتباطات بیشتری نسبت به سایرین داشته باشند. تشخیص جوامع در زمینه‌های مختلف در رسانه‌های اجتماعی مانند پیشنهاد‌های دوست، تقسیم مشتری، استخراج پیوند، رؤس برچسب‌گذاری و تحلیل نفوذ اجتماعی مفید است [11].

تعدادی از آثار به تشخیص جوامع با یک فرض بسیار محکم به نام جامعه اشاره می‌کنند. در یک جامعه، گروهی از رأس‌ها باید از یک ویژگی ساختاری بسیار محکم پیروی کنند. این کار شبیه مسئله معروف داده‌کاوی در تجزیه و تحلیل شبکه است (گراف کاوی). بعضی از نمونه‌های الگوریتم گراف کاوی در تعدادی از مقالات ارائه شده است [12] [6] [7] [13]. در تشخیص جامعه، تنها یک ساختار مهم وجود دارد و نتیجه دلخواه فهرستی از تمام گره‌های رأس است که ساختاری را در شبکه تشکیل می‌دهند.

تروساس و همکارانش [14] تشخیص جامعه را برای روابط بین خوشه‌های کاربر در فیس بوک برای آموزش زبان‌ها توصیف کردند. الگوریتم خوشه‌بندی K-means برای تعیین گروهی از کاربران با قابلیت‌ها و سبک‌های یادگیری یکسان، استفاده شد. اطلاعات از داده‌های کاربر

دو مرحله اطلاعات پیوندهای شبکه و اطلاعات گره‌ها را استفاده می‌کند. اطلاعات پیوندهای شبکه به جزءبندی ساختار شبکه به جوامع مختلف کمک می‌کند. اطلاعات گره‌ها، به تشخیص ناهنجاری‌های انفرادی کمک می‌کند. گائو و همکارانش [21] یک بهبود در روش بالا به‌وسیله یکپارچه‌سازی شبکه و اطلاعات گره‌ها برای تشخیص ناهنجاری‌های جوامع، ارائه کردند. روش پیشنهاد شده برای الگوریتم تشخیص ناهنجاری جامعه، از مدل ترکیب احتمالی طراحی شده برای شی داده چند متغیره (شیء با ویژگی‌های چندگانه) استفاده می‌نماید.

چن و همکارانش [22] یک روش تشخیص جامعه در شبکه‌های پیچیده با استفاده از حذف یال ارائه نموده‌اند. به منظور تشخیص ساختار جامعه در شبکه‌های پیچیده‌تر، الگوریتم جدیدی که به طور تکرار شونده یال‌ها را در جهت محدودیت‌ها حذف می‌کند، در مقاله مذکور ارائه شده است. این الگوریتم ابتدا از قدرت اتصال بین رأس‌ها برای تقسیم شبکه اصلی به برخی از جوامع استفاده می‌کند. همچنین قدرت اتصال بالا با قابلیت پیمانه‌ای بهینه با بهبود فرایند حذف یال را به کار می‌گیرد و در نهایت رئوس جدا شده را به جوامع اولیه برای بهینه‌سازی ساختار جامعه متصل می‌نماید.

یکی از چالش‌ها در تشخیص جوامع، وجود جوامع هم-پوشان است. در مقاله ارائه شده توسط رضوانی و همکارانش [23]، آنها ابتدا کیفیت جوامع هم‌پوشان را بررسی نموده و اثرات جداسازی آنها را نیز مورد توجه قرار می‌دهند. در تحقیق دیگری در رابطه با تشخیص جوامع هم‌پوشان، لیو و همکارانش [24] الگوریتمی دو مرحله‌ای را ارائه نمودند. در مرحله اول جوامع شناسایی شده و در مرحله دوم بر مبنای یک معیار شباهت، گره‌های پرت شناسایی می‌شوند. استفاده از توابع عضویت فازی نیز می‌تواند برای تشخیص این جوامع استفاده شود [25].

مسعودی و همکارانش [26] الگوریتم خفاش چند هدفه را برای تشخیص جامعه در شبکه‌های اجتماعی پویا به کار برده‌اند. بسیاری از الگوریتم‌های تکاملی برای حل مشکل تشخیص جامعه در شبکه‌های اجتماعی پویا پیشنهاد شده‌اند. بعضی از الگوریتم‌ها نیاز به پارامترهای پیش فرض دارند. دیگران از یک فرآیند تصادفی برای تولید جمعیت اولیه و برای اعمال اپراتورهای الگوریتم استفاده می‌کنند. این روشها فضای جستجو را افزایش می‌دهند و

نشان می‌دهند که الگوریتم ارائه شده قادر به شناسایی ۱۰۴۴۲ جامعه از ۱۸۸۵۱۴ گره است. پس از شناسایی جوامع، سه گروه برتر انتخاب شدند (کسانی که بیشترین تعداد اعضا را داشتند) و ۱۰ گره بالاترین امتیاز رتبه صفحه را در هر یک از این جوامع گرفتند. نمونه‌ها نشان می‌دهند که اکثر گره‌ها موضوع مشابهی دارند، اما هنوز گره‌هایی با موضوعات مختلف در یک جامعه مشابه وجود دارد. بررسی‌ها نشان داد که هفتاد درصد از گره‌ها موضوع مشابهی دارند، در حالی که سی درصد دیگر موضوعات متفاوتی دارند. با این وجود، مطالعه مذکور تأیید می‌کند که الگوریتم رتبه پیوند را می‌توان برای تشخیص جامعه در شبکه‌های جهت‌دار مورد استفاده قرار داد.

یک روش تشخیص ناهنجاری آماری بدون ناظر توسط ویسوانات و همکارانش [19] ارائه شده است که برای تشخیص رفتار غیرعادی در افراد کاربرد دارد. مجموعه داده‌های بدون برچسب فیس بوک در این تحقیق استفاده شده است و تعدادی از کاربران جعلی شناسایی شده‌اند. معیار استفاده شده برای جدا نمودن ناهنجاری‌ها و نرمال‌ها مقدار "پسندیدن"^۱ کاربران است.

در بسیاری از روش‌ها یک گراف ایجاد شده برای یک رسانه اجتماعی، به گروه‌ها یا جوامع مختلف تقسیم‌بندی می‌شود. این تقسیم‌بندی به‌وسیله حذف پیوندها از گره‌های مختلف یا به کاربردن الگوریتم‌های خاص خوشه‌بندی یا طبقه‌بندی، انجام می‌شود. برای مثال، ساختارهای جوامع به‌وسیله گیون و نیومن [20] بررسی شده‌اند. جوامع در قالب گروه‌های دوستی مختلفی هستند که در آنها قدرت پیوندها در بین گره‌ها در یک جامعه یا گروه دوستانه متراکم است؛ در حالی که در میان گره‌های دیگر خلوت است.

گائو و همکارانش [21] در زمینه‌ی تشخیص پرت‌های جوامع محلی و عمومی بسیار خوب کار کردند. یک روش ساده برای تشخیص پرت‌های جوامع روشی است که برای پرت‌های محلی و سراسری به کار می‌رود. برای مثال الگوریتم تشخیص پرت همسایه مستقیم، برای ناهنجاری‌های محلی، و الگوریتم تشخیص پرت سراسری برای ناهنجاری‌های سراسری به کار رفته‌اند. برای اینکه ناهنجاری‌های جوامع تشخیص داده شوند از اطلاعات موجود و هر گره جاری و همسایه‌هایش استفاده می‌گردد که این رویکرد، الگوریتم همسایه جامعه نامیده می‌شود که

منظور ابتدا یکی از روش‌های تشخیص جوامع معرفی می‌شود. روش مذکور که توسط چن و همکاران ارائه شده است [9]، تشخیص جامعه موضع افراد هدف با استفاده از تحلیل شبکه‌ی دوستی نام دارد که در این مقاله از آن برای ایجاد شبکه دوستی اولیه استفاده شده است تا در مراحل بعدی بهینه‌سازی بر روی آن انجام شود. در این روش، افرادی که نام آن‌ها در مجموعه‌ای از اسناد هدف آورده شده، در جوامعی با مواضع مرتبط خوشه‌بندی می‌شوند. این روش به این دلیل انتخاب شده است که خوشه‌بندی افراد براساس موضع آن‌ها در یک مجموعه از اسناد را به شکل مناسبی انجام می‌دهد بطوری که قابلیت بهبود بیشتر و بهینه‌سازی آن وجود دارد. اگرچه این روش انتخابی برای ایجاد شبکه دوستی اولیه بهترین روش نمی‌باشد، ولی انتخاب آن برای استفاده در روش پیشنهادی در این تحقیق با هدف نشان دادن قابلیت روش پیشنهادی برای بهینه‌سازی در چنین مسائلی است. نتایج حاصل از آزمایش‌ها که بر مبنای پایگاه‌های داده‌ی واقعی به دست آمده‌اند، حاکی از کارآمدی و بهره‌وری مطلوب روش پیشنهادی نسبت به روش پایه و دیگر روش‌های تشخیص جوامع است و نشان می‌دهد که این روش، کارایی بهتری نسبت به بسیاری از رویکردهای معروف تشخیص جوامع و الگوریتم‌های خوشه‌بندی دارد.

۳-۱- معرفی روش تشخیص جوامع

SCIFNET

در این روش تشخیص موضع افراد جامعه، آنهایی که نام آن‌ها در اسناد هدف آورده شده است، در جوامعی با مواضع یکسان طبقه‌بندی می‌شوند. شکل ۱ معماری این روش را نشان می‌دهد که از سه بخش تشکیل شده است: ساختار شبکه‌ی دوستی، توسعه جامعه موضع و اصلاح جامعه موضع. به بیان مشخص، روش مذکور ضمن استناد به یک مجموعه از اسناد و اخبار که گزارش‌هایی پیرامون یک مبحث با جوامع موضع خاص در آن‌ها آمده است، افراد هدف نامبرده در اسناد و اخبار را تشخیص می‌دهد. متعاقباً، یک شبکه‌ی دوستی برای افراد هدف ساخته می‌شود که مبتنی بر وقوع هم‌زمان نام افراد در اسناد و گرایش مواضع آن‌ها است. سپس، فرآیند توسعه جامعه موضع روی می‌دهد که به موجب آن، شناسایی جامعه موضع افراد هدف مانند یک فرآیند تشخیص جامعه تلقی می‌شود و گروه‌های موضع مشخص به‌طور مکرر در شبکه‌ی دوستی تکرار

باعث پیچیدگی فضایی می‌شوند. برای غلبه بر این نقاط ضعف، الگوریتم جدید چند هدفه‌ای با استفاده از الگوریتم انتقال میانگین برای تولید جمعیت اولیه برای به دست آوردن راه حل‌های با کیفیت بالا پیشنهاد شده است.

در شبکه‌های بزرگ و پیچیده، دستیابی به اطلاعات سراسری شبکه به منظور تشخیص جوامع بسیار دشوار یا حتی در بعضی موارد غیر ممکن است. در نتیجه از روش‌های تشخیص جوامع محلی برای این منظور استفاده می‌شود. در تازه‌ترین تحقیق انجام شده توسط لو و همکارانش [27] یک روش تشخیص جوامع محلی با استفاده از اطلاعات نزدیک‌ترین گره‌ها با مرکزیت بالا معرفی شده است. در زمینه تشخیص جوامع محلی، لو و همکارانش در ضمن تشکیل جامعه، به صورت پویا بر آن نظارت کرده و تابع عضویت مناسب را پیشنهاد می‌دهند [28].

روش‌های مبتنی بر جمعیت نیز برای تشخیص جوامع مورد استفاده قرار می‌گیرند. در تحقیقی که اخیراً توسط ون و همکارانش صورت گرفته [29]، چالش هم‌پوشانی جوامع در روش‌های تشخیص جوامع مورد بررسی قرار گرفته است. آنها از چند معیار و تابع چند هدفه برای تشخیص هم‌پوشانی در یک الگوریتم تکاملی مبتنی بر جمعیت استفاده نمودند. در بین الگوریتم‌های تکاملی تشخیص جامعه، استفاده از الگوریتم وال است که توسط زانگ و همکارانش اخیراً ارائه گردیده است [30].

در مقاله ارائه شده توسط پن و همکارانش [31]، روشی برای تشخیص جوامع ارائه شده است که از یک معیار شباهت محلی برای شناسایی نزدیک‌ترین همسایه‌های هر گره استفاده می‌نماید. سپس با ترکیب زوج گره‌ها زیرگراف‌های مرتبطی را ایجاد می‌کند و در نهایت جوامع را ایجاد می‌کند.

پژوهش حاضر، به بررسی مسئله‌ی تشخیص جامعه یا گروه افراد هدف (بر مبنای موضع آن‌ها) می‌پردازد تا در نهایت بتواند افراد هدف را در جوامع با مواضع یکسان، به صورت بهینه‌سازی شده خوشه‌بندی نماید.

۳- روش پیشنهادی

در این بخش روش پیشنهادی تشریح خواهد شد، که به موجب آن افراد هدفی که نام آن‌ها در مجموعه‌ای از اسناد آورده شده است، براساس یک تابع برازندگی و ساختار الگوریتم ژنتیک خوشه‌بندی خواهند شد. به این

حد واقعی نشان داده می‌شود. افراد هدفی که نام آن‌ها مکرراً در اسناد و اخبار مربوط به روابط موافق (مخالف) در کنار یکدیگر ظاهر می‌شوند، احتمالاً یک رابطه‌ی دوستانه (مخالفانه) دارند. روش پایه انتخابی، برای تعیین کمیت گرایش موضع در یک متن یا سند هدف، از روش ترنی و لیست‌من^۵ [35] استفاده نموده و ارزش (وزن) موضع، براساس رابطه (۱) محاسبه می‌شود:

(۱)

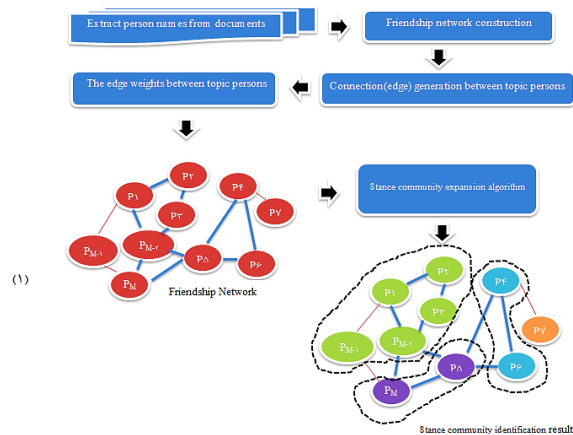
در این عبارت، SW_d نشان‌دهنده وزن یا ارزش موضع در سند D است؛ $Fwords$ و $Owords$ به ترتیب مجموعه‌های لغات با مفاهیم دوستانه و مخالف را نشان می‌دهند که توسط متخصصان زبان‌شناسی شناسایی شده‌اند. تابع $count(word_i, word_j)$ تعداد اسناد و اخباری که در آن‌ها، $word_i$ و $word_j$ به‌طور هم‌زمان ظاهر شده‌اند را محاسبه می‌کند. اساساً، در این معادله از اطلاعات دوسویه‌ی نقطه‌به‌نقطه^۶ (PMI) برای محاسبه وزن موضع یک سند استفاده شده است. وزن موضع SW_d مثبت خواهد بود اگر محتوای D ارتباط قوی و فراوانی با $Fwords$ داشته باشد. چن و همکاران [9]، ضریب همبستگی موضع محور (SOCOR) زیر را که مقدار/ ارزش یا وزن موضع را در ضریب همبستگی تلفیق می‌کند، طراحی کرده‌اند (رابطه ۲):

$$A = \left[\sum_{d \in D_{friendly}} SW_d * (p_{i,d} - \bar{p}_{i,friendly}) * (p_{j,d} - \bar{p}_{j,friendly}) + \sum_{d \in D_{opposing}} SW_d * (p_{i,d} - \bar{p}_{i,opposing}) * (p_{j,d} - \bar{p}_{j,opposing}) \right]$$

$$B = \sqrt{\sum_{d \in D_{friendly}} [\sqrt{SW_d} * (p_{i,d} - \bar{p}_{i,friendly})]^2 + \sum_{d \in D_{opposing}} [\sqrt{|SW_d|} * (p_{i,d} - \bar{p}_{i,opposing})]^2}$$

$$C = \sqrt{\sum_{d \in D_{friendly}} [\sqrt{SW_d} * (p_{j,d} - \bar{p}_{j,friendly})]^2 + \sum_{d \in D_{opposing}} [\sqrt{|SW_d|} * (p_{j,d} - \bar{p}_{j,opposing})]^2}$$

می‌گردند. در مرحله‌ی آخر، الگوریتم اصلاح جامعه شروع به کار کرده و نتیجه‌ی شناسایی را بر طبق یک تابع هدف، بهبود می‌بخشد.



شکل ۱: ساختار روش پایه تشخیص جامعه [9]

۱-۱-۳- ساخت یک شبکه‌ی دوستی

فرض کنید که $D = \{d_1, d_2, \dots, d_3\}$ مجموعه‌ای از اسناد و اخبار هدف است و $P = \{p_1, p_2, \dots, p_M\}$ یک مجموعه از افراد هدف می‌باشد که نام آن‌ها در مجموعه‌ی D آمده است. طی فرآیند ساخت شبکه‌ی دوستی، یک شبکه‌ی دوستی با عنوان $G = \{P, E\}$ ساخته می‌شود که در آن، افراد هدف موجود در مجموعه P ، گره‌های شبکه را تشکیل می‌دهند؛ و $E = (p_i, p_j)$ مجموعه‌ای از یال‌ها است که گرایش‌های دوستی افراد هدف را نشان می‌دهد (یعنی اینکه آیا رابطه‌ی میان افراد هدف، دوستانه است یا مخالفانه). به‌طور معمول، کشف گرایش‌های دوستی از روی متون دشوار است. باین‌حال، هریس^۲ [32] به این نتیجه رسیده است که واحدهای متنی با معانی متناقض و مغایر به‌ندرت در یک بافت مشابه ظاهر می‌شوند. علاوه بر آن، کانایاما و ناسوکاوا^۳ [33] نیز ثابت کرده‌اند که واحدهای متنی با گرایش‌های مشابه، در کنار یکدیگر ظاهر می‌شوند تا بافت را منسجم سازند. براین اساس، ضریب همبستگی^۴ [34] که وقوع هم‌زمان درجه‌ی افراد هدف در D را ارزیابی می‌کند، احتمالاً یک معیار یا شاخص مطلوب برای اکتشاف گرایش دوستی میان افراد هدف خواهد بود. باین‌حال، اسناد هدف گاهی مسائل جنجال‌برانگیز را تحت پوشش قرار می‌دهند. در این اسناد و اخبار، افراد با مواضع مختلف به‌شکلی یکدیگر را مورد انتقاد قرار می‌دهند. بنابراین تنها ضمن لحاظ کردن وقوع هم‌زمان درجه‌ی افراد هدف در D ، میزان دوستی رقبا بیش از حد واقعی و عملکرد تشخیص جامعه موضع، کمتر از

یا تضاد همسایگی) باعث افزایش قدرت دوستی می‌شود. و برعکس در حالت عدم دوستی یا تضاد یعنی $SOCOR(p_i, p_j) < -\theta$ قدرت دوستی منفی خواهد بود که نشانه قدرت عدم دوستی یا تضاد است.

۲-۱-۳- توسعه جامعه موضوع

شکل ۲ یک نمونه از توسعه جامعه موضوع و شکل ۳، نمودار الگوریتم توسعه جامعه را نشان می‌دهد. در نمودار شکل ۳، نماد $P_{unlabeled}$ نشان‌دهنده یک مجموعه از گره‌های بی‌نشان است (یعنی افراد هدف). در ابتدا داریم: $P_{unlabeled} = P$ ، یعنی تمام گره‌ها فاقد نشان هستند. این الگوریتم، گره‌های K را به‌طور تصادفی به‌عنوان دانه‌های جوامع موضوع انتخاب کرده و به‌صورت تکرار شونده، جوامع را ضمن ادغام گره‌های بی‌نشان توسعه می‌بخشد. در هر بار تکرار، یک مجموعه از گره‌های بی‌نشان U که مستقیماً به یک گروه متصل هستند، شناسایی می‌شود: $(U = \{p_i \in P_{unlabeled} | (p_i, p_j) \in E, p_j \in c_k, 1 \leq k \leq k\})$ هر گرهی p_i در U متعاقباً باهدف شناسایی یک نشان مناسب برای گروه، مورد بررسی قرار می‌گیرد. فرض کنید Z_i نشان‌دهنده مجموعه گروه‌هایی باشد که گرهی بی‌نشان p_i به‌طور مستقیم با آن‌ها در ارتباط است؛ این بدان معناست که خواهیم داشت:

$$Z_i = \{c_k | (p_i, p_j) \in E, p_j \in c_k, 1 \leq k \leq k\}$$

به‌عنوان مثال Z_4 که در شکل ۲ نشان داده شده است، گروه‌های c_1 و c_2 را تشکیل می‌دهد. چن و همکاران [16]، امتیاز ادغام برای هر گروه c_k در Z_i را به‌صورت رابطه ۴ محاسبه می‌نمایند:

$$ms_{i,k} = \sum_{p_j \in c_k, (p_i, p_j) \in E} \delta(p_i, p_j) \quad (۴)$$

در این رابطه، $ms_{i,k}$ نشان‌دهنده امتیاز ادغام p_i با c_k است. اساساً امتیاز ادغام یا همجوشی حاصل جمع وزن یال‌های مرتبط با p_i با جامعه موضوع c_k است. ادغام p_i با یک گروه که امتیاز ادغام مثبت دارد، باید یک گروه با مواضع یکپارچه و منسجم به دست دهد. زمانی که بیش از یک گروه دارای امتیاز ادغام مثبت باشند، الگوریتم اقدام به ادغام p_i با آن گروهی می‌کند که حداکثر امتیاز ادغام را به خود اختصاص داده است. این مورد قابل اثبات خواهد بود که این عمل بیشترین بازدهی را برای تابع هدف به همراه دارد. به این نکته توجه داشته باشید که اگر اکثر گره‌ها در c_k یک رابطه مخالف با p_i داشته باشند، امتیاز ادغام منفی

$$SOCOR(p_i, p_j) = \frac{A}{B * C} \quad (۲)$$

در این عبارت $D_{friendly} \subseteq D$ نشان‌دهنده‌ی مجموعه‌ای از اسناد هدف است که مقدار یا ارزش موضع آن‌ها مثبت است؛ $D_{opposing} \subseteq D$ مجموعه‌ای از اسناد هدف را نشان می‌دهد که مقدار یا ارزش موضع آن‌ها منفی است؛ و $\bar{p}_i.friendly$ و $\bar{p}_i.opposing$ بسامدهای میانگین P_i را نشان می‌دهند که به ترتیب در $D_{friendly}$ و $D_{opposing}$ ظاهر شده‌اند. بازه‌ی $SOCOR(p_i, p_j)$ نیز همانند ضریب همبستگی، معادل $[-1, 1]$ است. اگر وقوع هم‌زمان p_i و p_j در D مستقل از یکدیگر باشد، مقدار برابر با صفر لحاظ می‌شود. اما اگر p_i و p_j تمایل داشته باشند و به‌طور هم‌زمان در اسناد حاکی از روابط دوستانه (یا مخالف) حضور یابند، آنگاه $SOCOR(p_i, p_j)$ مثبت (یا منفی) خواهد بود. چن و همکاران [9] سپس، گرایش دوستی را با توجه به ضریب همبستگی موضع محور بررسی و تعریف نموده‌اند.

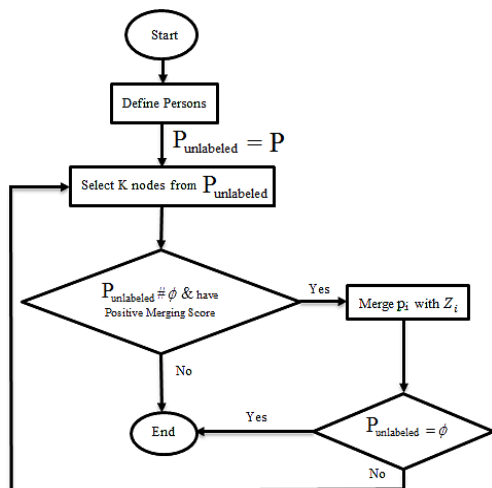
طبق تعریف اگر $p_i \in P$ باشد، همسایه‌های دوست p_i با نماد $\Gamma_{friendly}(p_i)$ نشان داده می‌شوند. همچنین همسایه‌های مخالف p_i با نماد $\Gamma_{opposing}(p_i)$ نشان داده می‌شوند.

استحکام و قدرت همسایگی مشترک دوستانه میان p_i و p_j با نماد $\gamma(p_i, p_j)$ نشان داده می‌شود و با استفاده از ضریب ژاکارد^۷ بدست می‌آید. قدرت و استحکام همسایگی مشترک مخالف میان p_i و p_j توسط $\omega(p_i, p_j)$ نشان داده می‌شود که این نیز با استفاده از ضریب ژاکارد محاسبه می‌شود.

قدرت و استحکام دوستی که با نماد $\delta(p_i, p_j)$ نشان داده می‌شود، نشان‌دهنده وزن یال (p_i, p_j) می‌باشد و براساس رابطه ۳ محاسبه می‌شود:

$$\delta(p_i, p_j) = \begin{cases} (SOCOR(p_i, p_j) + 1)^{\frac{\gamma(p_i, p_j) + \omega(p_i, p_j)}{2} + \beta} & .if \ SOCOR(p_i, p_j) > \theta \\ -(|SOCOR(p_i, p_j)| + 1)^{\left(1 - \frac{\gamma(p_i, p_j) + \omega(p_i, p_j)}{2}\right) + \beta} & .if \ SOCOR(p_i, p_j) < -\theta \end{cases} \quad (۳)$$

در رابطه (۳) اگر رابطه بین p_i و p_j دوستانه باشد یعنی $SOCOR(p_i, p_j) > \theta$ در این صورت افزایش مقدار توان γ قدرت دوستی همسایگی و ω قدرت عدم دوستی



شکل ۳: نمودار الگوریتم توسعه جامعه موضع [16]

در الگوریتم پیشنهادی در هر بار تکرار، یک مجموعه از گره‌های مرزی با عنوان $P_{boundary} \subseteq P$ تعریف می‌شوند. هر گره از $P_{boundary}$ به یک گروه از مواضع تعلق داشته و به تعدادی گروه دیگر متصل است. به عبارت دیگر:

$$P_{boundary} = \{p_i | (p_i, p_j) \in E, p_i \in C_m, p_j \in C_n, m \neq n\}$$

صادق است. این الگوریتم، هر گره مرزی را مجدداً در آن جامعه موضع که بیشترین امتیاز ادغام را دارد، ادغام می‌کند. بدین طریق تازمانی که $P_{boundary}$ خالی شده یا نتیجه‌ی شناسایی به ثبات برسد، گره‌های مرزی، شناسایی و خوشه‌بندی خواهند شد. اساساً، الگوریتم اصلاح جامعه پیشنهادی یک الگوریتم بالارونده است؛ چراکه نتیجه‌ی شناسایی را به‌طور تکرار شونده بهبود می‌بخشد. در الگوریتم پیشنهادی، با ایجاد هر جمعیت جدید از روی جمعیت قبلی، اعضای که مقدار تابع هدف آنها بیشترین مقدار است عیناً به جمعیت بعدی منتقل می‌شوند. به این ترتیب احتمال از دست دادن اعضای از جمعیت فعلی که ممکن است بتوانند مقادیر مطلوب‌تری برای تابع هدف در نسل‌های بعدی تولید کنند وجود نخواهد داشت. بنابراین انتظار می‌رود در جمعیت‌های بعدی تابع هدف غیر نزولی باشد.

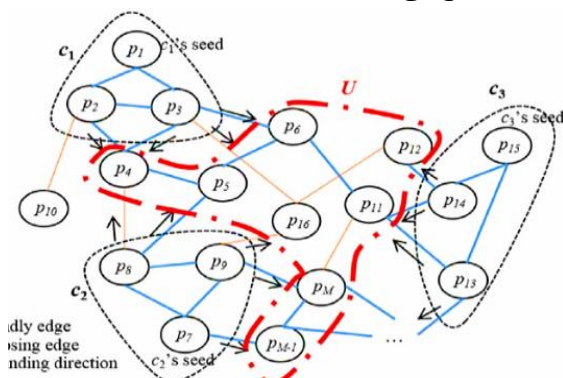
۲-۳- چارچوب الگوریتم پیشنهادی

بر طبق روش پیشنهادی، افرادی که نام آن‌ها در اسناد هدف آورده شده است، در جوامعی با مواضع یکسان طبقه‌بندی می‌شوند. شکل ۴، معماری روش پیشنهادی را نشان می‌دهد که از سه بخش تشکیل شده است: ساختار شبکه‌ی دوستی، ساختار الگوریتم ژنتیک (مقداردهی اولیه، ادغام، جهش و تابع برازندگی) و جامعه افرادی که تشخیص

خواهد بود. از آنجایی که ادغام p_i با یک گروه دارای موضع مخالف چندان مناسب نیست، این الگوریتم عملیات ادغام را مجدداً تکرار می‌کند؛ به شرطی که حداکثر امتیاز ادغام، منفی باشد. الگوریتم مذکور، گروه را به‌طور مکرر توسعه می‌بخشد؛ تا زمانی که تمام گره‌های بی‌نشان در شبکه‌ی دوستی، ادغام شده باشند یا هیچ گره‌ی بی‌نشانی دارای امتیاز ادغام مثبت با هیچ گروهی نباشد. سپس، یک نتیجه‌ی شناسایی گروه به دست خواهد آمد که با استفاده از الگوریتم اصلاح جامعه موضع (شکل ۳)، اصلاح می‌گردد.

۳-۱-۳- اصلاح جامعه موضع

الگوریتم توسعه جامعه موضع به‌طور تکرار شونده، گروه‌ها را از گره‌های دانه، توسعه و گسترش می‌دهد. در برخی از موارد، یک گره با یک گروه ادغام می‌شود، صرفاً به این دلیل که به دانه‌ای از آن گروه نزدیک باشد، حال آنکه بهتر است که با یک گروه دیگر ادغام شود. علاوه بر آن، نتیجه‌ی توسعه و گسترش، وابستگی بسیار زیادی به کیفیت دانه‌ها دارد. به‌منظور کمینه‌سازی اثر مسئله‌ی ادغام زودرس و کاهش تأثیر مقداردهی اولیه به دانه، الگوریتم پیشنهادی در این مقاله برای اصلاح یا پالایش جوامع موضع طراحی شده است. این الگوریتم، گروه‌ها را به‌طور تکرار شونده اصلاح می‌کند.

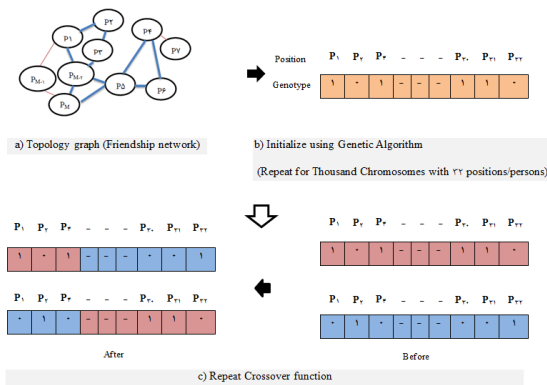


شکل ۴: نمونه از توسعه جامعه موضع [16]

$$\text{Fitness function} = \sum_{i=1}^m \sum_{j=1}^m \delta(i,j) + \sum_{i=1}^m \sum_{j=1}^m \delta(i,j) - 2 \quad (5)$$

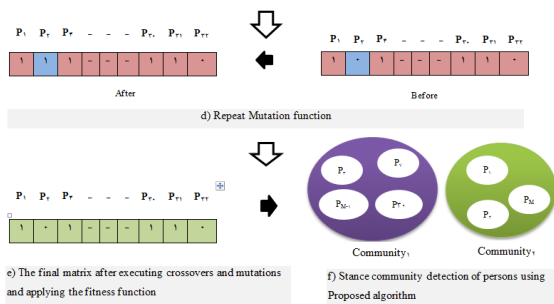
$$* \sum_{i=1}^m \sum_{j=1}^m \delta(i,j)$$

در این رابطه m تعداد افراد کلاس اول و n تعداد افراد کلاس دوم است. برای هر کروموزوم در هر نسل این رابطه تشکیل و خوشه‌بندی انجام می‌شود. لازم به ذکر است رابطه (۵) برای مسائلی با دو خوشه طراحی شده است. صورت نیاز می‌توان این رابطه را برای مسائلی با چند خوشه توسعه داد. با توجه به اینکه آزمایش‌ها در



شکل ۴: معماری الگوریتم پیشنهادی جهت تشخیص جامعه موضع افراد

در این رابطه اشخاصی که در کلاس اول هستند و با یکدیگر رابطه دارند (با موضع یکسان هستند) با سایر اشخاص همان کلاس مقدارشان جمع می‌شود. در عبارت دوم تابع مذکور نیز همین روال برای افراد کلاس دوم اتفاق



می‌افتد. در عبارت سوم چون هدف بهینه‌سازی است باید افرادی که در دو کلاس مختلف هستند مقدارشان براساس رابطه دلتا کاهش یابد پس یک مقدار منفی قبل از عبارت سوم درج می‌گردد و به دلیل اینکه دو کلاس وجود دارد در یک ضرب ۲ نیز ضرب می‌شوند.

داده شده‌اند. به بیان مشخص، روش مذکور ضمن استناد به یک مجموعه از اسناد و اخبار که گزارش‌هایی پیرامون یک مبحث در آن‌ها آمده است، افراد هدف نامبرده در اسناد و اخبار را تشخیص می‌دهد. یک شبکه‌ی دوستی برای افراد هدف ساخته می‌شود که مبتنی بر وقوع هم‌زمان نام افراد در اسناد و گرایش مواضع آن‌ها است. شبکه دوستی به‌عنوان جزئی از ساختار الگوریتم ژنتیک تعریف می‌شود. به بیان واضح‌تر، رابطه دلتا (رابطه ۳)، که نشان‌دهنده قدرت و استحکام دوستی است و به‌نوعی شبکه دوستی را ایجاد می‌نماید به‌عنوان جزئی از تابع برازندگی در ساختار الگوریتم ژنتیک مورد استفاده قرار می‌گیرد. سپس، ماتریس نام افراد به‌عنوان جمعیت اولیه در ساختار الگوریتم ژنتیک ایجاد و پس از اجرای توابع ادغام و جهش و ایجاد جمعیت‌های بعدی، این جمعیت‌ها توسط تابع برازندگی مورد ارزیابی قرار گرفته و بهترین ساختار کروموزوم (بردار)، به‌عنوان ماتریس نهایی انتخاب‌شده که در آن مشخص می‌شود افراد متعلق به چه جامعه‌ای هستند.

شکل ۴ روش ایجاد اعضای جمعیت و ایجاد جمعیت بعدی را نمایش می‌دهد و شکل ۵ الگوریتم روش پیشنهادی را نشان می‌دهد که بر طبق آن، افرادی که نام آن‌ها در اسناد هدف آورده شده است، در جوامعی با مواضع یکسان طبقه‌بندی می‌شوند. در این الگوریتم، ۳۲ شخص به‌عنوان یک ماتریس ۳۲ عضوی وارد ساختار الگوریتم ژنتیک می‌شوند. سپس تعداد بردارهای کروموزوم مشخص می‌گردد. اندازه جمعیت ۱۰۰۰ کروموزوم در نظر گرفته شده است که در تمام آزمایش‌ها ثابت است. هر کروموزوم ۳۲ عضو دارد (به تعداد اشخاص). در ادامه درصد ادغام‌ها و جهش‌ها مشخص می‌گردد. در الگوریتم شکل ۴ تعداد ۳۰۰ نسل تولید می‌شود که این تعداد در آزمایش‌ها ثابت در نظر گرفته شده است.

در مرحله بعد الگوریتم، رابطه بهینه‌سازی براساس رابطه دلتا تشکیل می‌گردد که همان تابع برازندگی است. رابطه (۵) که همان تابع برازندگی الگوریتم پیشنهادی است به جهت تشخیص یک جامعه منسجم طراحی گردیده است.

شکل ۵: الگوریتم پیشنهادی جهت تشخیص جامعه موضع افراد

۴- آزمایش‌ها

در این بخش، مجموعه داده‌های مورد استفاده در آزمایش‌ها معرفی خواهد شد؛ عملکرد روش پیشنهادی با روش SCIFNET و برخی دیگر از روش‌های خوشه‌بندی در رابطه با تشخیص جوامع مقایسه خواهد شد. متعاقباً، یکی از نتایج تشخیص جامعه که توسط روش پیشنهادی تشخیص داده شده است، معرفی می‌شود.

۴-۱- مجموعه داده تشخیص جوامع موضع افراد، یک حوزه پژوهشی نسبتاً جدید به شمار می‌آید. در این پژوهش از یک مجموعه داده با عنوان Kroa که شامل ۷۴ سند و اخبار مرتبط با آن است که همه‌ی آن‌ها از رسانه خبری Google News دانلود شده‌اند، استفاده گردید. مباحث و موضوعات جمع‌آوری شده، حوزه‌ی مسائل سیاسی را پوشش می‌دهند. بعلاوه، از برخی متخصصان انسانی نیز درخواست گردیده است تا اسناد غیر مرتبط را به صورت دستی از مجموعه جدا کنند تا اطمینان حاصل گردد که اسناد تحت بررسی، کاملاً با موضوعات تناسب دارند. برای استخراج نام افراد هدف که در اسناد هدف آمده‌اند، قطعه کدی نوشته شده است. فراوانی نام افراد هدف (λ)، به میزان ۶۰ درصد از فراوانی کل اسامی هدف استخراج شده مشخص گردید. میانگین تعداد اسامی بررسی شده تحت هر وضعیت از (λ) در جدول ۱ نشان داده شده است.

برای ارزیابی عملکرد، از شاخص رند^۸ که یک معیار ارزیابی خوشه‌بندی بسیار مهم است استفاده شده است [17]؛ زیرا در این روش، گروه‌های مختلف افراد هدف در خوشه‌های مختلف دسته‌بندی می‌شوند. شاخص رند، درصد تمام جفت‌های افراد را که به طور صحیح خوشه‌بندی شده‌اند نشان می‌دهد (یعنی اگر دو فرد با موضع یکسان در یک گروه واقع شده باشند یا دو فرد با مواضع مختلف در دو گروه مختلف قرار گرفته باشند). رابطه ۶ روش محاسبه این شاخص را نشان می‌دهد [17].

$$RI = \frac{TP + TN}{TP + TN + FP + FN} \quad (6)$$

در ادامه تابع ادغام انجام می‌شود که در آن، به طور تکرار شونده، دو کروموزوم (احتمال انتخاب کروموزوم‌ها با برابری بالاتر بیشتر است) انتخاب و ترکیب می‌شوند. در این تابع نوع ادغام، از نوع ادغام با یک نقطه قطع است و روش انتخاب کروموزوم‌ها به روش چرخ رولت می‌باشد. در ادامه الگوریتم پیشنهادی، جهش انجام می‌شود. در جهش به طور تکرار شونده، یک کروموزوم به صورت تصادفی انتخاب می‌شود و دو فیلد از آن به صورت تصادفی تعویض می‌شود. پس از انجام توابع ادغام و جهش و تولید نسل‌های جدید، بهترین کروموزوم بر اساس تابع برابری انتخاب می‌شود. در نهایت خروجی یک بردار است که به هر خانه مقدار ۰ یا ۱ نسبت داده که مشخص می‌کند هر شخص به کدام کلاس مربوط می‌شود (کلاس اول یا کلاس دوم).

جدول ۱: آمار نوشته‌های ارزیابی شده

topic	Number
# of topic documents	74
# of extracted topic persons	54
# of evaluated topic persons	32

Proposed algorithm

$$P = P_{persons}$$

The input of the genetic algorithm is the number of persons based on their position $P_{persons} = \gamma$

Maxiter = γ // Maximum number of replicas of the genetic algorithm

// Create a genetic algorithm structure

While (Maxiter) do

// According to γ Persons, the chromosome vectors (popsize) are γ -membered (nvar: $P_{persons}$)

// Determine percentage of crossover (pc) and mutations (pm) for the genetic algorithm

$$popsize = 1000 \quad nvar = P \quad pc = 0.7$$

$$ncross = \gamma * \left(\frac{popsize * pc}{\gamma} \right)$$

$$pm = 1 - nvar$$

$$nmult = popsize * pm$$

// Optimization relationship: Based on the delta matrix, the relation is as follows:

$$Fitness\ function = \sum_{i=1}^m \sum_{j=1}^m \delta(i,j) + \sum_{i=1}^n \sum_{j=1}^n \delta(i,j) - \gamma * \sum_{i=1}^m \sum_{j=1}^n \delta(i,j)$$

// m: is the number of first-class people and n: is the number of second-class people. For each

// chromosome in each generation, this relationship is formed and clustering is done.

// Carry out the crossover: the probable selection of γ chromosomes (the probability

// of chromosomes selection with higher fitness is higher) and their composition

// Carry out the mutation: Choosing a completely random chromosome

// and replacing two houses together

// Select the best chromosomes for the next generation and perform

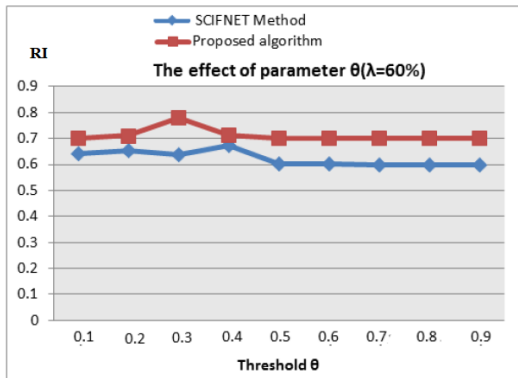
// crossovers and mutations to the specified number of repetitions

End while

Return

$$C_1 = \{P_1, P_2, \dots, P_m\}, C_2 = \{P_1, P_2, \dots, P_n\}$$

// The output is a vector that assigns a value of 0 or 1 to each position, which specifies which Persons are related to which class.



شکل ۶: دقت الگوریتم پیشنهادی و الگوریتم SCIFNET

در روش پیشنهادی، امتیاز اندازه‌ها نیز به صورت یک‌روال خطی پیش می‌رود و فقط در $\theta = 0.3$ افزایش می‌یابد که احتمالاً به دلیل ماهیت تصادفی بودن الگوریتم ژنتیک است. در روش پایه مقدار پایین λ نشان‌دهنده آن است که روال تشخیص جوامع موضع، دشوار است چرا که در این شرایط، فراوانی اسامی هدف در فرآیند شناسایی بسیار اندک است [16]. از آنجایی که روال ساخت یک شبکه دوستی با استناد به وقوع نام افراد در اسناد و اخبار هدف طی می‌شود، شمول نام افراد با فراوانی اندک می‌تواند کیفیت شبکه را کاهش داده و در نتیجه بر عملکرد تشخیص اثر بگذارد. اساساً، مقدار دو اندازه مذکور با افزایش مقدار θ افزایش پیدا می‌کند زیرا θ با مقدار بالا می‌تواند روابط نه‌چندان مهم میان افراد را پالایش و جداسازی کند تا کیفیت شبکه ارتقا یابد. زمانی که θ بالاتر از ۰,۴ باشد، امتیاز اندازه‌ها به تدریج افت پیدا می‌کند. در این شرایط میان گره‌ها نیز ارتباط و پیوندی شکل نمی‌گیرد. در نتیجه شبکه‌ی دوستی، به قدری کم تراکم خواهد بود که ارتباطات اطلاعاتی میان افراد، بازنمایی نمی‌شود و عملکرد شناسایی گروه چندان با کیفیت نخواهد بود.

با توجه به اهمیت تعداد تکرار (شرط پایان) الگوریتم ژنتیک در روش پیشنهادی، شکل ۷ مقادیر شاخص رند را تحت شرایط $\theta = 0.3$, $\beta = 1$ با تعداد تکرارهای مختلف از الگوریتم ژنتیک نشان می‌دهد. با توجه به شکل مذکور و نتایج آزمایش‌ها تعداد تکرار سیصد مرتبه بهترین نتیجه را در بر خواهد داشت.

در این رابطه TP تعداد افراد با مواضع یکسان که به درستی خوشه بندی شده‌اند، TN تعداد افراد با مواضع متفاوت که به درستی در خوشه‌های مختلف قرار داده شده‌اند، FP تعداد افراد با مواضع یکسان که به اشتباه خوشه بندی شده‌اند و FN تعداد افراد با مواضع متفاوت که به اشتباه خوشه بندی شده‌اند را نشان می‌دهد. هرچه امتیاز شاخص رند نیز بیشتر باشد، کیفیت خوشه‌بندی بالاتر است. لازم به ذکر است رابطه (۶) برای مسائلی با چند کلاستر نیز قابل استفاده است. توضیح روش به کار گیری آن برای چنین مسائلی در این مرجع آمده است [17]. از آنجایی که الگوریتم توسعه جامعه مواضع، وابستگی بسیار زیادی به مقدار اولیه‌ی دانه‌ی گروه دارد و نیز روش پیشنهادی پژوهش وابسته به معیارهای تصادفی است، این روش با مقادیر تصادفی ده بار تکرار می‌شود. مقادیر شاخص رند برای موضوع Korea تحت بررسی، میانگین گیری می‌شوند تا عملکرد کلی روال تشخیص، مشخص گردد.

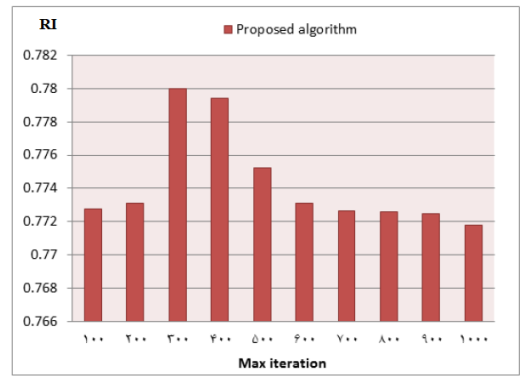
۲-۴- ارزیابی نتایج

در ابتدا، پارامتر θ که نشان‌دهنده‌ی آستانه‌ی گرایش دوستی است و برای تعیین یال‌ها در یک شبکه‌ی دوستی کاربرد دارد در نظر گرفته می‌شود. در این آزمایش، θ بین ۰,۱ تا ۰,۹ لحاظ می‌شود. جدول ۲، فهرست‌های Fwords و Owords را که توسط متخصص زبان‌شناسی تألیف و گردآوری شده‌اند، نشان می‌دهد [16] (یادآوری: Fwords و Owords به ترتیب مجموعه‌های لغات با مفاهیم دوستانه و مخالف را نشان می‌دهند که در رابطه‌ی ۱ به کار برده شده‌اند). از جدول مذکور در پیاده‌سازی‌ها و آزمایش‌های روش پیشنهادی و روش پایه استفاده شده است. فهرست لغات نشان‌دهنده مواضع، مورد استفاده‌ی ضریب همبستگی موضع محور بوده و برای محاسبه وزن یا ارزش موضع در یک سند هدف کاربرد دارد. پارامتر β که برای محاسبه قدرت و استحکام دوستی به کار می‌رود نیز برابر با ۱ در نظر گرفته می‌شود. شکل ۶، مقادیر شاخص رند را تحت شرایط مختلف θ و λ نشان می‌دهد.

تمام این روش‌ها به یک شبکه ورودی احتیاج دارند. بعلاوه باید توجه داشت که SM و FEC برای شبکه‌های نشان‌دار^۹ طراحی شده‌اند. روش‌های FastModularity، CODA، SCAN، بر آن فرض استوار شده‌اند که شبکه‌های تحت بررسی، بدون نشان هستند و به همین علت، ساختارهای پیوند را به جهت شناسایی گروه‌ها بررسی می‌کنند. الگوریتم‌های K-means و HAC بازنمایانگر یک فرد هدف هستند و آن را در قالب یک بردار فراوانی N بعدی نشان می‌دهند که در آن، یک ورودی، فراوانی وقوع اسم افراد هدف در یک سند هدف را نمایش می‌دهد و برای ارزیابی رابطه میان افراد هدف، از تشابه کسینوسی استفاده می‌شود که برای تعیین تشابه بردارهای بسامد یا فراوانی کاربرد دارد. خوشه‌بندی در روش‌های CODA و K-means تا حد زیادی به مقاردهی‌های اولیه آن‌ها وابسته است.

همان‌طور که در جدول ۳ ذکر گردیده است، روش پیشنهادی نسبت به روش پایه و سایر روش‌های ذکر شده در جدول، بهترین عملکرد را از خود نشان داده است. نتایج نشان می‌دهند که HAC و K-means افراد هدف مشهور را در کنار یکدیگر خوشه‌بندی می‌کنند. این بدان علت است که در تشابه کسینوسی، ضرب داخلی دو بردار فراوانی نرمال‌سازی شده است و اگر بردارهای محاسبه شده حاوی تعداد زیادی ورودی غیر صفر باشند، یک امتیاز تشابه بالا به دست می‌آید [16].

از آنجایی که اسامی افراد مشهور در بسیاری از اسناد و اخبار قید شده است، بردارهای فراوانی نرمال‌سازی شده حاوی تعداد زیادی ورودی غیر صفر خواهند بود. بنابراین، روش‌های خوشه‌بندی، رابطه میان افراد هدف مشهور را بالاتر از آنچه که هست برآورد می‌کنند و افراد مشهور دارای مواضع مختلف را در کنار یکدیگر خوشه‌بندی می‌نمایند. در نتیجه، عملکرد روش تضعیف می‌شود. عملکرد نه‌چندان قوی استراتژی پیوند منفرد در روش HAC، متأثر و ناشی از نقصان فوق است؛ چراکه طبق این استراتژی، تشابه دو خوشه از طریق بررسی شبیه‌ترین جفت افراد حاضر در خوشه انجام می‌شود. در نتیجه، این استراتژی خوشه‌های حاوی افراد مشهور را با یکدیگر ادغام می‌کند، حال آن‌که ممکن است این خوشه‌ها نمایانگر مواضع متفاوت باشند.



شکل ۷: دقت الگوریتم پیشنهادی تحت تاثیر تکرار الگوریتم ژنتیک

در ادامه، روش پیشنهادی با روش پایه و با شش روش دیگر تشخیص جامعه به نام روش‌های FastModularity [36]، [37]SCAN، [38]FEC، [21]CODA، [39]SM و دو روش شناخته شده خوشه‌بندی K-means و HAC [40] مقایسه خواهد شد. نتایج این مقایسه‌ها در جدول ۳ ذکر گردیده است.

جدول ۳: کارایی Rand index براساس $\lambda=60\%$

Method	(Rand index)%
proposed method	77.99
SCIFNET	75.89
Fast Modularity	64.42
SCAN	65.23
CODA	68.87
SM	69.19
FEC	68.05
HAC (Single-Link)	53.23
K-means	75.1

	Good Team Work Partner Advocate Friend
Stance-opposing words (Owords)	Campaign Opposite Rival Fraud Accusation Contest Lost Beat Debate

آن، تابع هدف به‌سادگی می‌تواند مجموع وزن‌های یال‌ها در هر گروه را به حداکثر رسانده و رابطه میان گروه‌ها را نادیده انگارد. به همین علت، تعداد زیادی فرد مشهور با مواضع مختلف در یک گروه قرار می‌گیرند.

تابع هدف در روش پیشنهادی (که تابع برازندگی نامیده می‌شود) و همچنین روش پایه علاوه بر بیشینه‌سازی رابطه میان گره‌های موجود در گروه‌ها، رابطه بین گروه‌ها را به حداقل می‌رساند. بنابراین، روش پیشنهادی و روش پایه یک برتری در شناسایی گروه‌ها پیدا می‌کنند. روش FEC به‌طور معمول عملکرد بهتری نسبت به دیگر روش‌های تحت مقایسه دارد؛ این بدان علت است که روش مذکور برای شبکه‌های نشان‌دار طراحی شده و پیوندهای منفی در شبکه‌های دوستی ایجاد شده وجود ندارند. باین‌حال عملکرد روش پایه همچنان از FEC بهتر است. روش SM نیز برای شبکه‌های نشان‌دار طراحی شده است. اما SM گاهی نمی‌تواند برای یک مبحث تحت ارزیابی، K گروه بسازد؛ زیرا نشانه‌های ورودی‌ها در بردارهای ویژه‌ی اصلی همگی مثبت هستند. به همین علت، افرادی با مواضع متفاوت در یک گروه کنار هم قرار می‌گیرند. در نتیجه روش پایه عملکرد بهتری نسبت به روش SM خواهد داشت.

روش پیشنهادی عملکرد بهتری را نسبت به روش پایه و سایر روش‌های ذکر شده در جدول ۵ از خود نشان داده است. به‌طور خاص مزیت‌های روش پیشنهادی نسبت به روش پایه عبارت‌اند از: ۱- دقت خوشه‌بندی افزایش می‌یابد، ۲- عملیات خوشه‌بندی سراسری انجام می‌شود. در روش پایه عملیات خوشه‌بندی به‌صورت محلی انجام می‌گیرد، یعنی الگوریتم از یک نقطه شروع نموده و مقادری اولیه انجام می‌دهد (که ممکن است غلط باشد) سپس عضو جدید را با اعضای خوشه مقایسه می‌کند و بر اساس این‌که قدرت و استحکام دوستی (رابطه $\delta(p_i, p_j)$) عضو جدید با کدام خوشه بیشتر است، عضو جدید به آن خوشه اختصاص می‌یابد و به‌تدریج خوشه‌ها تشکیل می‌شوند. در این حالت ممکن است به دلیل عدم شکل‌گیری کامل خوشه‌ها در این مرحله، یک گره (فرد) جدید به‌اشتباه به خوشه‌ای اختصاص یابد. اما در روش پیشنهادی از همان ابتدا با استفاده از روش ارائه‌شده، در ساختار الگوریتم ژنتیک همه حالت‌ها به‌صورت سراسری بررسی می‌شود تا به برازندگی بهتری برسد. ۳- در روش

همچنین روش پیشنهادی رابطه میان افراد هدف را بر اساس ساختار الگوریتم ژنتیک و یک تابع برازندگی تعیین می‌نماید. روش پایه رابطه میان افراد مشهور را برحسب ضریب همبستگی موضع محور و قدرت همسایگی مشترک ارزیابی می‌کند. ضریب همبستگی موضع محور برخلاف تشابه کسینوسی، تفاوت در وقوع اسامی دو فرد هدف در کنار هم در یک مجموعه از اسناد مرتبط با یک گرایش را بررسی می‌کند؛ و در نتیجه قادر است رابطه‌ی میان افراد هدف مشهور را به‌درستی برآورد نماید.

الگوریتم FastModularity، گره‌ها را در گروه‌ها و با توجه به معیار مدولاریته^۱ ادغام می‌کند؛ در این شرایط گروه‌هایی که با چندین یال در ارتباط هستند، ادغام می‌شوند. باین‌حال، این معیار وزن‌های یال در گره‌ها را نادیده می‌گیرد. بسیاری از یال‌های متصل، وزن‌های اندکی دارند که بر انسجام گروه ادغام‌شده اثر گذاشته و عملکرد الگوریتم را با اختلال مواجه می‌سازند. روش پیشنهادی گروه‌های اولیه (ساختار کروموزوم‌ها) را بر اساس جهش‌ها و تقاطع‌ها می‌سازد و بر اساس یک تابع برازندگی انتخاب می‌نماید. روش پایه، گروه‌ها را بر مبنای امتیاز ادغام آن‌ها در هم ادغام و ترکیب می‌کند (رابطه ۶). از آنجایی که این امتیاز برحسب وزن‌های یال‌ها (قدرت دوستی) تعیین می‌شود، گره‌ها در یک گروه، ارتباطی قوی و مستحکم با یکدیگر خواهند داشت.

متعاقباً، نتیجه تشخیص جوامع در روش پیشنهادی نسبت به نتیجه حاصل از الگوریتم پایه و روش FastModularity بهتر خواهد بود. در روش SCAN از یک تشابه شبه ژاکارد برای ارزیابی قدرت همسایگی مشترک میان گره‌ها استفاده شده و یک گره با یک گروه ترکیب می‌شود اگر قدرت همسایگی مشترک آن‌ها زیاد باشد. باین‌حال روش SCAN نیز همانند روش FastModularity، وزن‌های یال‌ها را نادیده می‌انگارد و عملکرد را دچار اختلال می‌سازد. علاوه بر قدرت همسایگی مشترک، قدرت دوستی در روش پایه نیز بر وقوع هم‌زمان گره‌ها در اسناد هدف با مواضع یکسان متمرکز است. در نهایت نتیجه گرفته می‌شود که روش پایه عملکرد بهتری نسبت به روش SCAN دارد.

به‌رغم آنکه در روش CODA، وزن‌های یال‌ها در تابع هدف خوشه‌بندی آن تلفیق می‌شود، اما این وزن‌ها مبتنی بر تشابه کسینوسی بردارهای فراوانی حاصل شده‌اند. علاوه بر

(همانطور که در شکل ۸ نشان داده شده است).
بازده نهایی، معادل با نتیجه‌ی مقایسه‌ای است که در بخش
پیش ارائه شد

شکل ۸: نتیجه خوشه‌بندی اسامی اشخاص بر اساس
الگوریتم پیشنهادی

۵- نتیجه‌گیری

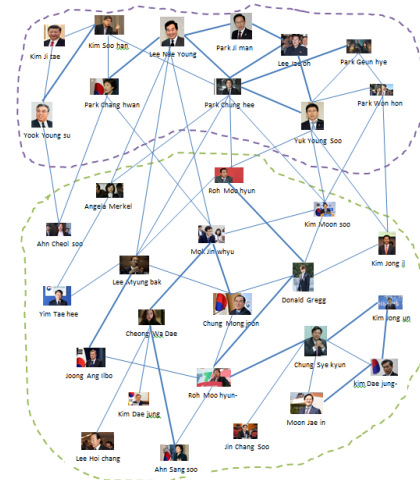
رسانه‌های اجتماعی و خبری در بستر اینترنت تبدیل به
یک ابزار مهم و فراوان برای انتشار و اکتساب آخرین
اطلاعات پیرامون موضوعات مختلف شده است. با این حال،
کاربران اینترنت در غالب اوقات با تعداد زیادی از اخبار
مختلف و غیر مرتبط با موضوع مورد نظرشان احاطه
می‌شوند. اساساً زمان‌ها، مکان‌ها و افراد از جمله عناصر
کلیدی موضوعات خبری به شمار می‌آیند. آگاهی ارتباط
میان افراد می‌تواند کمک شایانی به خواننده نموده تا بتواند
یک دانش پیش‌زمینه‌ای در خصوص مبحث به دست آورده
و اسناد و اخبار مربوط به آن را به سرعت شناسایی نماید.
یکی از چالش‌های اساسی در مورد تجزیه و تحلیل رسانه‌های
اجتماعی و خبری، کشف خودکار جوامع است. با توجه به
محبوبیت رسانه‌های اجتماعی و خبری، تعداد کاربران
شبکه به‌طور چشمگیری افزایش یافته است. کشف الگوهای
ناشناخته در چنین رسانه‌هایی در بسیاری از برنامه‌های
کاربردی همچنان کار دشواری است که می‌تواند
مفید باشد و در همین رابطه تشخیص جامعه می‌تواند در
زمینه‌های مختلف استفاده شود.

در پژوهش حاضر، به بررسی مسئله شناسایی اجتماع یا
گروه افراد هدف در رسانه‌های اجتماعی و خبری پرداخته
شد. یکی از روش‌های مهمی که تاکنون در این حوزه ارائه
گردیده به نام SCIFNET مورد بررسی دقیق قرار گرفته
و پیاده‌سازی گردید. در نهایت روش پیشنهادی جهت
بهبود روش مذکور که روش پایه نامیده شد ارائه شد. در
روش پیشنهادی با استفاده از ساختار الگوریتم ژنتیک و نیز
ساخت کروموزوم‌ها بر اساس جهش و تقاطع آن‌ها و نیز
انتخاب بهترین‌ها بر اساس یک تابع برازندگی، بهترین
گروه‌ها بر اساس نام افرادی که در اسناد آورده شده‌اند،
شناسایی گردیدند. نتایج حاصل از آزمایش‌ها در دنیای
واقعی، حاکی از بهره‌وری مطلوب روش پیشنهادی نسبت
به روش پایه و سایر روش‌ها بوده و نشان می‌دهد که این
روش عملکرد بهتری نسبت به بسیاری از دیگر روش‌های
شناسایی و خوشه‌بندی دارد.

پایه یکی از مسائل، مقداردهی اولیه بود به‌طوری‌که از
همان ابتدا به صورت تصادفی چند گره به‌عنوان خوشه‌های
اولیه انتخاب می‌شد و این باعث تأثیر منفی و کاهش دقت
در خوشه‌بندی در روش پایه می‌گردید اما در روش
پیشنهادی به دلیل استفاده از ساختار کروموزوم‌ها در
الگوریتم ژنتیک و عدم نیاز به گره اولیه این مشکل نیز رفع
گردید.

آزمایش‌های مذکور، عملکرد روش پیشنهادی را به لحاظ
کمی و کیفی مورد بررسی و ارزیابی قرار داده‌اند. همان‌طور
که در بخش‌های قبلی بیان شد برای آزمایش‌ها از مجموعه
داده‌ای شامل ۷۴ سند مربوط به گوگل نیوز، استفاده شده
است. مجموعه داده مذکور مربوط به یک مبحث سیاسی با
عنوان Korea است، لذا از این مجموعه داده استفاده
می‌شود تا نتایج تشخیص جامعه در روش پیشنهادی ارزیابی
شود. این مبحث، پیرامون نامزدهای انتخابات ریاست
جمهوری کره جنوبی و وابستگان آن‌ها است. شکل ۸،
نتیجه تشخیص جامعه افراد (خوشه‌بندی اسامی اشخاص)
بر اساس الگوریتم پیشنهادی را نشان می‌دهد (با در نظر
گرفتن $\theta = 0.3$ و $\beta = 1$). تعداد زیادی از اسناد و
اخبار هدف، حاوی گزارش‌هایی در خصوص برگزاری
انتخابات هستند.

نواحی رنگی موجود در شکل ۸، نتایج تشخیص جامعه را
نشان می‌دهند. در این مثال، بسیاری از افراد هدف
به‌درستی در گروه‌های مختلف تقسیم‌بندی شده‌اند. به طور
مثال پارک گون هیه^{۱۱} دختر پارک چانگ هی^{۱۲} است که
نام آنها در بسیاری از اسناد مربوط به مجموعه داده به طور
همزمان تکرار شده است. در نتیجه با اعمال الگوریتم
پیشنهادی این دو فرد در یک جامعه قرار گرفته‌اند



محل انجام می‌گردید یعنی الگوریتم از یک نقطه شروع می‌نمود و مقداردهی انجام می‌شد (که امکان اشتباه وجود داشت).

سپس اعضای جدید را با گروه مقایسه می‌نمود و بر اساس این که عضو جدید با کدام گروه رابطه دارد به آن گروهی که مقدار $ms_{i,k}$ بیشتر است اختصاص می‌یافت (یادآوری: رابطه ۴). (در این عبارت، $ms_{i,k}$ نشان‌دهنده امتیاز ادغام p_i با C_k است). اساساً امتیاز ادغام یا همجوشی حاصل جمع اوزان یال مرتبط با p_i با جامعه موضع C_k است. ادغام p_i با یک گروه که امتیاز ادغام مثبت دارد، باید یک گروه با مواضع یکپارچه و منسجم به دست دهد. زمانی که بیش از یک گروه دارای امتیاز ادغام مثبت باشند، الگوریتم اقدام به ادغام p_i با آن گروهی می‌کند که حداکثر امتیاز ادغام را به خود اختصاص داده است. به تدریج که گروه‌ها تشکیل می‌شوند، در این روش ممکن است چون گروه‌ها تکمیل نگردیده‌اند، یک گره جدید (شخص جدید) به یک گروه اختصاص یابد، که درست نباشد (به‌طور منطقی باید گروه‌ها کامل تشکیل شوند تا وضعیت مشخص شود). لذا در روش پیشنهادی در این پژوهش از همان ابتدا با استفاده از روش ارائه‌شده، با ساختار الگوریتم ژنتیک و کروموزوم‌ها همه حالت‌ها بررسی می‌شوند تا به برازندگی بهتری برسد. یعنی تمامی حالت‌ها به‌صورت سراسری بررسی می‌شود. همچنین در روش پیشنهادی، چالش مقداردهی اولیه موجود روش پایه با توجه به الگوریتم پیشنهادی و ساختار الگوریتم ژنتیک مرتفع گردید. به بیان واضح‌تر، در روش پایه گروه‌های اولیه به صورت تصادفی انتخاب می‌شدند. که باعث تغییرات زیادی در تشخیص جوامع می‌گردید.

در این پژوهش از یک مجموعه داده برای ارزیابی استفاده شد. این پیکره با عنوان Kroa و شامل ۷۴ سند و اخبار مرتبط با آن است که همه‌ی آن‌ها از رسانه خبری Google News بارگذاری شده‌اند. مباحث و موضوعات جمع‌آوری شده، حوزه‌ی مسائل سیاسی را پوشش می‌دهند؛ بعلاوه، از برخی متخصصان انسانی نیز درخواست گردیده است تا اسناد غیر مرتبط را به‌صورت دستی از مجموعه جدا کنند تا اطمینان حاصل گردد که اسناد تحت بررسی، کاملاً با موضوعات تناسب دارند.

در طی آزمایش‌های این پژوهش، حوزه‌های پژوهشی جالب توجهی برای آثار آتی شناسایی گردید. به‌عنوان مثال علیرغم آنکه ضریب همبستگی کاربرد فراوانی در شناسایی گرایش ارتباط میان افراد هدف دارد، اما توسط مسئله‌ی فراوانی کم نام افراد تحت تأثیر قرار می‌گیرد. از آنجایی که نام افرادی که فراوانی تکرار آن‌ها کمتر است در بسیاری از متون و اخبار غایب می‌باشد، ضریب همبستگی مبتنی بر موضع احتمالاً استحکام رابطه را بیش از حد واقعی برآورد می‌کند. کاهش وزن اسناد و متون در زمانی که نام افراد با فراوانی اندک از متون غایب هستند، می‌تواند مسئله‌ی برآورد فراتر از انتظار را حل نماید. علاوه بر این، لحاظ کردن متون و اخبار غیر مرتبط با مبحث موردنظر می‌تواند موجب وجود نام افراد غیر مرتبط در فرآیند شناسایی شود و عملکرد سیستم را با مشکل مواجه سازد. بنابراین، رویکردهای مبتنی بر حذف مباحث غیر مرتبط باید در جهت توسعه‌ی عملکرد شناسایی مورد استفاده قرار بگیرند. همچنین، اسناد و اخبار ورودی به‌صورت دستی جمع‌آوری شده بودند. برای آنکه به کاربران اینترنتی کمک شود مباحث و موضوعات نوظهور را به‌خوبی درک کنند، روش پایه می‌تواند با فن‌های مختلف شناسایی و مسیریابی موضوع تلفیق شود؛ بدین صورت اسناد و اخبار پیرامون یک مبحث به‌طور خودکار و از منابع اطلاعاتی مختلف (مثلاً آژانس‌های خبری)، شناسایی و رهگیری می‌شوند. در همین رابطه با توجه به روش بهینه ارائه شده با الهام از ساختار الگوریتم ژنتیک، در این پژوهش مشکل مقداردهی اولیه و نیز محلی بودن روش پایه حل گردید و باعث افزایش دقت خوشه‌بندی نیز گردید. لذا به دلیل ماهیت تصادفی بودن الگوریتم ژنتیک می‌توان این استنباط را داشت که بهبود در روش پایه قبل از بهره‌گیری از الگوریتم ژنتیک نیز می‌تواند مفید واقع گردد.

در روش پیشنهادی و در مقایسه با روش پایه، نتیجه‌های ذیل به دست آمد. اول در روش پیشنهادی دقت خوشه‌بندی نسبت به روش پایه و سایر روش‌های تشخیص جوامع، افزایش یافت. همچنین در روش پیشنهادی مسئله (چالش) حساس بودن به مقداردهی اولیه مرتفع گردید و در نهایت مسئله خوشه‌بندی محلی در روش پایه نیز حل شده به‌طوری که با روش ارائه شده، خوشه‌بندی (تشخیص جوامع) به‌صورت سراسری صورت گرفت. لازم به توضیح است که در روش پایه عمل خوشه‌بندی به‌صورت

منابع

- Media:Architecture, Tools, and Approaches to Detect Criminal Activity,” در *Application of Big Data for Application of Big Data for National Security*, Elsevier, 2015, pp. 155-172.
- 11.G.-J. Qi , C. Aggarwal و T. Huang, “Community detection with edge content in social media networks,” در *Paper presented at the 2012 IEEE 28th international conference on data engineering* .۲۰۱۲ ,
- 12.S. Borgatti و M. Everett , “Graph colorings and power in experimental exchange networks,” *Soc Netw*, شماره ۱۴, جلد ۳, pp. 287–308, 1992.
- 13.S. Nijssen و J. Kok , “A quickstart in frequent structure mining can make a difference,” در *Paper presented at the Proceedings of the tenthACMSIGKDD international conference on knowledge discovery and data mining* .۲۰۰۴ ,
- 14.C. Troussas , M. Virvou, J. Caro و K. Espinosa , “Mining relationships among user clusters in Facebook for language learning,” در *Paper presented at the international conference on computer, information and telecommunication systems (CITS)* .۲۰۱۳ ,
- 15.C. Chen, Z.-Y. Chen و C.-Y. Wu , “An unsupervised approach for person name bipolarization using principal component analysis,” *IEEE Trans. Knowl. Data Eng.*, جلد ۲۴, pp. 1963-1976, 2012.
- 16.C. Chen و C.-Y. Wu , “Bipolar person name identification of topic documents using principal component analysis,” در *Proceeding of the 23rd International Conference on Computational Linguistics* , .۲۰۱۰
- 17.C. D. Manning, P. Raghavan and H. Schütze, "An Introduction to Information Retrieval," in *Cambridge University Press*, New York, 2009.
- 1.A. M. Kaplan و M. Haenlein, “Users of the world, unite! The challenges and opportunities of social media,” *Business Horizons*, شماره ۱, جلد ۵۳, pp. 59-68, 2010.
- ۲.۲۰۱۴۸. [درون خطی.
- 3.A. Mislove , “Online social networks: measurement, analysis, and applications to distributed information systems,” 2009.
- 4.M. Sachan , D. Contractor, T. Faruque و L. Subramaniam , “Using content and interactions for discovering communities in social networks,” در *Paper presented at the proceedings of the 21st international conference on world wide web* .۲۰۱۲ ,
- 5.D. Ganley و C. Lampe, “The ties that bind: social network principles in online communities,” *Decision Support Systems*, شماره ۳, جلد ۴۷, pp. 266-274, 2009.
- 6.M. Kuramochi و G. Karypis , “Finding frequent patterns in a large sparse graph,” *Data Min Knowl Discov* , شماره ۳, جلد ۱۱, p. 243–271.
- 7.X. Yan و J. Han, “gspan: graph-based substructure pattern mining,” در *Paper presented at the Proceedings of the IEEE international conference on data mining (ICDM 2002)* .۲۰۰۲ ,
- 8.H. Cai, V. W. Zheng و K. C.-C. Chang, “A Comprehensive Survey of Graph Embedding: Problems, Techniques and Applications,” *IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING*, شماره ۹, جلد ۳۰, .۲۰۱۷
- 9.Z.-Y. Chen و C. C. Chen, “SCIFNET: Stance community identification of topic persons using friendship network analysis,” *Knowledge-Based Systems*, شماره ۱۱۰, p. 30–48, 2016.
- 10.B. Akhgar, G. B. Saathoff, H. R. Arabnia, R. Hill, A. Staniforth و P. Saskia Bayerl, “Mining Social کندگان,”

- 26.I. Messaoudi و N. Kamel, "A multi-objective bat algorithm for community detection on dynamic social networks," *Applied Intelligence*, pp. 1-18, 1 2019.
- 27.W. Luo, N. Lu, L. Ni, W. Zhu و W. Ding, "Local community detection by the nearest nodes with greater centrality," *Information Sciences*, جلد ۵۱۷، pp. ۳۹۲-۳۷۷، ۲۰۲۰.
- 28.W. Luo, . D. Zhang , . H. Jiang , L. Ni . و Y. Hu , "Local Community Detection With the Dynamic Membership Function," *IEEE Transactions on Fuzzy Systems*, جلد ۲۶، شماره ۵، ۲۰۱۸، ۵.
- 29.X. Wan, X. Zuo و F. Song, "Solving dynamic overlapping community detection problem by a multiobjective evolutionary algorithm based on decomposition," *Swarm and Evolutionary Computation*, ۲۰۲۰.
- 30.Y. Zhang, Y. Liu, J. Li, J. Zhu, C. Yang, W. Yang و C. Wenc, "WOCDA: A whale optimization based," *Physica A: Statistical Mechanics and its Applications*, جلد ۵۳۹، ۲۰۲۰.
- 31.X. Pan, . G. Xu , . B. Wang . و T. Zhang , "A Novel Community Detection Algorithm Based on Local Similarity of Clustering Coefficient in Social Networks," *IEEE Access* , جلد ۷، ۲۰۱۹، ۷.
- 32.Z. Harris, "Distributional structure," *Word* , جلد ۱۰، p. 146-162, 1954.
- 33.H. Kanayama و T. Nasukawa, "Fully automatic lexicon expansion for domain-oriented sentiment analysis," در *Proceedings of the Conference on Empirical Methods in Natural Language Processing* ۲۰۰۶،
- 34.G. Keller , "Statistics for Management and Economics," *Cengage Learning*, ۲۰۰۸
- 35.P. Turney و M. Littman , "Measuring praise and criticism: inference of semantic orientation from association," *ACM Trans.*
- 18.S. B. Yudhoatmojo و M. A. Samuar, "Community Detection On Citation Network Of DBLP Data Sample Set Using LinkRank Algorithm," *۴th Information Systems International Conference 2017, ISICO 2017* ,Bali, Indonesia, 2017.
- 19.B. Viswanath, M. Bashir , M. Crovella , S. Guha, K. Gummadi, B. Krishnamurthy و A. Mislove, "Towards detecting anomalous user behavior in online social networks.," در *Proceedings of the 23rd USENIX security symposium (USENIX Security ۲۰۱۴)* ,
- 20.M. Girvan و M. Newman, "Community structure in social and biological networks," *Proc Natl Acad Sci* , جلد ۹۹(۱۲)، p. 7821-6, 2002.
- 21.J. Gao, F. Liang , W. Fan, C. Wang , Y. Sun و J. Han, "On community outliers and their efficient detection in information networks," در *Proceedings of the 16th ACM SIGKDD international conference Proceedings of the 16th ACM SIGKDD international conference ۲۰۱۰* ,
- 22.X. Chen and J. Li, "Community detection in complex networks using edge-deleting with restrictions," *Physica A: Statistical Mechanics and its Applications*, vol. 519, pp. 181-194, 4 2019.
- 23.M. Rezvani, W. Liang, . C. Liu . و J. Xu Yu , "Efficient Detection of Overlapping Communities Using Asymmetric Triangle Cuts," *IEEE Transactions on Knowledge and Data Engineering*, شماره ۱۱، جلد ۳۰، ۲۰۱۸.
- 24.Z. Liu, B. Xiang, W. Guo , Y. Chen, K. Guo . و J. Zheng , "Overlapping Community Detection Algorithm Based on Coarsening and Local Overlapping Modularity," *IEEE Access*, جلد ۷، ۲۰۱۹، ۷.
25. T. Meng , L. Cai , . T. He , L. Chen . و Z. Deng, "Local Higher-Order Community Detection Based on Fuzzy Membership Functions," *IEEE Access*, جلد ۷، ۲۰۱۹، ۷.

- 38.B. Yang, W. Cheung و J. Liu, "Community mining from signed social networks," *IEEE Trans. Knowl. Data Eng.*, جلد ۱۹, p. 1333–1348, 2007.
- 39.P. Anchuri و M. Magdon-Ismai, "Communities and balance in signed networks: a spectral approach," در *Proceedings of the International Conference on Advances in Social Networks Analysis and Mining*, ۲۰۱۲ ,
- 40.T. Mitchel, *Machine Learning*, McGrawHill, 1997.
- Inf. Syst.*, جلد ۲۱, p. 315–346, 2003.
- 36.M. Newman, "Fast algorithm for detecting community structure in networks," *Phys. Rev.*, جلد ۷۴, ۲۰۰۴ ,
- 37.X. Xu, N. Yuruk, Z. Feng و T. Schweiger, "SCAN: a structural clustering algorithm for networks," در *Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ۲۰۰۷ ,

سیستم توصیه‌گر زمینه آگاه سفر با بهره‌گیری از اطلاعات عکس‌های برچسب‌دار جغرافیایی

*رضوان محمدرضایی لرکی **رضا روانمهر ***میلاد امرالهی

*دانشجوی دکتری، گروه مهندسی کامپیوتر، واحد تهران مرکزی، دانشگاه آزاد اسلامی، تهران
**استادیار، گروه مهندسی کامپیوتر، واحد تهران مرکزی، دانشگاه آزاد اسلامی، تهران
***کارشناسی ارشد، گروه مهندسی کامپیوتر، واحد تهران مرکزی، دانشگاه آزاد اسلامی، تهران
تاریخ دریافت: ۱۳۹۸/۱۲/۲۲ تاریخ پذیرش: ۱۳۹۹/۰۶/۲۹

چکیده

سیستم‌های توصیه‌گر بر اساس اطلاعات کاربران و آیتم‌ها، سلايق و علاقه‌مندی‌های مورد نظر کاربران را پیش‌بینی کرده و بر اساس آنها آیتم‌های مناسب را به کاربران پیشنهاد می‌دهند. اکثر رویکردهای موجود برای سیستم‌های توصیه‌گر بر روی پیشنهاد دادن مرتبط‌ترین آیتم‌ها تمرکز می‌کنند و اطلاعات زمینه‌ای مانند زمان یا مکان را در نظر نمی‌گیرند. در این مقاله یک روش برای سیستم توصیه‌گر زمینه آگاه سفر با بهره‌گیری از اطلاعات عکس‌های برچسب‌دار جغرافیایی ارائه شده است که از دقت بالایی برخوردار است. این روش نسبت به مقالات مشابه، تعداد زمینه‌های بیشتری مانند (وضعیت آب و هوا، وضعیت روحی کاربر، میزان ترافیک و ...) را در نظر می‌گیرد. با این کار باعث نزدیک شدن کاربران یک خوشه و در نتیجه افزایش دقت می‌شود. روش پیشنهادی به جای استفاده از الگوریتم خوشه‌بندی کلاسیک از ادغام دو الگوریتم رقابت استعماری و خوشه‌بندی C-Means فازی استفاده می‌کند. ارزیابی روش ارائه شده بر روی مجموعه داده فلیکر انجام شده و نتایج ارزیابی حاکی از آن است که روش پیشنهادی قادر به فراهم کردن توصیه‌های موثر و دقیق مطابق با علاقه‌مندی‌های کاربر و موقعیت فعلی بازدید او می‌باشد. رویکرد پیشنهادی در معیارهای دقت، متوسط مطلق دقت و میانگین قدر مطلق خطا به ترتیب به ۰,۶۹ و ۰,۵۳ و ۰,۳۱ دست یافته است.

کلمات کلیدی: سیستم توصیه‌گر زمینه آگاه، سیستم توصیه‌گر سفر، زمینه، الگوریتم رقابت استعماری، خوشه‌بندی C-Means فازی.

۱- مقدمه

شاید تا به حال برای شما این سؤال پیش آمده باشد که شرکت گوگل با چه الگوریتمی تبلیغات هوشمند را به کاربران نمایش می‌دهد؟ یا فیس بوک چگونه هنگام آپلود عکس‌ها و برچسب‌زنی آن‌ها، نام دوستانمان را به ما پیشنهاد می‌دهد؟

پاسخ این سوالات در استفاده این سایت‌ها از سامانه‌های توصیه‌گر است. سیستم‌های توصیه‌گر با توجه به رشد روزافزون اطلاعات در دنیای حاضر، افزایش چشمگیر کاربران آنلاین و همچنین ضرورت تحویل سریع اطلاعات درست به کاربران، اهمیتی حیاتی دارند و با هدف حذف سربارهای اطلاعاتی، یکی از ابزارهای مهم تجارت الکترونیک شده‌اند [۱]. مهم‌ترین هدف در این سیستم‌ها، جذب مشتریان و جلب اعتماد آن‌ها از طریق ارائه بهترین و مناسب‌ترین پیشنهاد خرید محصولات، با توجه به علایق و سلیقه آن‌ها می‌باشد [۲]. تحقیقات در حوزه الگوریتم‌های توصیه‌گر به صورت روز افزون در حال افزایش هستند.

شرکت‌ها با ایجاد سامانه‌های هوشمندی که رفتار گذشته کاربر و نیز رفتار کاربران مشابه را بررسی می‌کنند، پیشنهادات و توصیه‌های مناسب و مورد توجه کاربر را به او ارائه می‌دهند [۱]. مواردی از قبیل فرصت‌های شغلی مناسب کاربر، فیلم‌های مورد علاقه کاربر، ویدئوهای پیشنهادی، دوستان فیس بوکی که احتمالاً آن‌ها را می‌شناسید و ... از جمله موارد پیشنهادی به کاربران می‌تواند باشد.

با توجه به جایگاه فناوری اطلاعات در صنعت گردشگری و پیدایش گردشگری الکترونیک، نقطه‌ی عطفی میان این صنعت و فناوری اطلاعات به وجود آمده است. توجه به گردشگری و استفاده از فناوری‌های نوین در حال گسترش آن، زمانی روشن می‌شود که بدانیم درآمد حاصل از آن، هم ردیف درآمدهای نفتی می‌باشد [۳]. بنابراین یکی از مهم‌ترین کاربردهای فناوری اطلاعات و ارتباطات، صنعت گردشگری است که نقش چشم‌گیری در جهت توسعه‌ی منافع ملی ایفا می‌کند.

از نکات قابل توجه در این زمینه تأمین نیازهای کاربران و ارائه‌ی خدمات مطلوب‌تر به آن‌ها است. در سال‌های اخیر الگوریتم‌های مختلفی در جهت بهبود این سیستم‌ها ابداع شده‌اند که در اکثر کاربردها و زمینه‌های مبتنی بر وب، شخصی سازی نقش مهمی را ایفا می‌کند [۴]. شخصی سازی به فرایندی گفته می‌شود که در طی آن رفتار سامانه

بر اساس دانش استخراج شده از رفتارهای کاربران، تغییر می‌کند [۵]. این مقاله سعی دارد با توجه به فرآیند شخصی سازی، رویکردهای مرتبط با سیستم‌های توصیه‌گر و چگونگی به‌کارگیری و کارایی این الگوریتم‌ها را در محیط‌های گردشگری الکترونیک بررسی نماید.

با توجه به این که کیفیت توصیه‌ها در سیستم‌های توصیه‌گر گردشگری بسیار حائز اهمیت است، اگر توصیه‌ای ارائه شود که کاربر قبلاً آن را دریافت کرده باشد و یا به آن نیازی نداشته باشد به کاهش نگرش مثبت وی به توصیه‌ها و بی‌توجهی به آن‌ها می‌انجامد. به همین دلیل لازم است سیستم‌های توصیه‌گر در ارائه توصیه به گردشگران از دقت بالایی برخوردار باشند که بکارگیری منطق فازی در کاهش عدم قطعیت و ارائه توصیه‌های موثر، می‌تواند مفید واقع شود [۶]. در این پژوهش ترکیب دو الگوریتم رقابت استعماری و خوشه‌بندی فازی به منظور پردازش بیشتر و قوی‌تر در مقابل استفاده از خوشه‌بندی کلاسیک، موجب افزایش دقت پیشنهادهای ارائه شده توسط سیستم است.

آنچه موجب متمایز شدن روش پیشنهادی این مقاله شده است افزودن تعداد زمینه‌های 'بیشتر' (وضعیت آب و هوا، وضعیت روحی کاربر، میزان ترافیک و ...) به نسبت مقالات مشابه است که فقط از زمینه‌های زمانی و مکانی استفاده کرده‌اند. این نکته موجب نزدیک تر شدن کاربران درون یک خوشه و به موجب آن افزایش دقت گردید.

در روش ارائه شده به منظور پارامترهای کیفیت سرویس دقت، متوسط مطلق دقت و میانگین قدر مطلق خطا اندازه‌گیری و ارزیابی می‌شوند. در این تحقیق سعی می‌شود پارامترهای ذکر شده بهبود یابند. به منظور ارزیابی روش ارائه شده از دیتاست فلیکر، یکی از بزرگ‌ترین سایت‌های اشتراک‌گذاری تصویر و ویدئو آنلاین استفاده شده است. به طور خلاصه نوآوری‌های این پژوهش به ترتیب زیر عبارتند از:

- ارائه یک روش مناسب برای سیستم زمینه آگاه سفر بر اساس ترکیب دو الگوریتم رقابت استعماری و خوشه‌بندی C-Means فازی
- بهره‌گیری از اطلاعات عکس‌های برچسب‌دار جغرافیایی جهت خوشه‌بندی و مدلسازی زمینه‌های داده‌ها

علاوه بر اطلاعات ضمنی و صریح برخی از سیستم‌ها نیز هستند که از اطلاعات شخصی کاربران استفاده می‌کنند. به عنوان مثال سن، جنسیت و ملیت کاربران می‌توانند منابع خوبی برای شناخت کاربر و ارائه پیشنهاد به وی باشند. به این گونه از اطلاعات، اطلاعات آماری گفته می‌شود که گروهی از سیستم‌های پیشنهاددهنده مبتنی بر همین اطلاعات بنا شده‌اند. با ظهور وب ۲ و گسترش شبکه‌های اجتماعی در سال‌های اخیر محققین به منبع اطلاعاتی دیگری برای بهبود کیفیت پیشنهادها پی بردند که همان اطلاعات موجود در شبکه‌های اجتماعی بود و بر همین اساس کارهای تحقیقاتی زیادی در این حوزه انجام شد [۱].

۲-۲- داده‌های زمینه‌ای و انواع آن‌ها

آدوماویکوییس و توژیلین [۸] بر این واقعیت تاکید دارند که با توجه به این که مفهوم زمینه بسیار گسترده است، تمرکز زمینه باید به طور مستقیم بر حوزه‌های مرتبط با سیستم‌های توصیه‌گر نظیر داده کاوی، شخصی‌سازی تجارت الکترونیک، پایگاه‌های داده، بازیابی اطلاعات، سیستم‌های فراگیر آگاه از زمینه، بازاریابی و مدیریت باشد. نیاز به تعریف و مدل‌سازی یک زمینه در سیستم‌های توصیه‌گر از طریق رویه‌ای سازگار، توسط پژوهشگران مختلف شناسایی شده است. بنا به نظر وربرت و همکارانش [۹] تعریف و مدل‌سازی دقیق زمینه، شناسایی آن چه که زمینه را می‌سازد آسان نموده و استفاده مجدد و تبادل داده‌های زمینه‌ای در برنامه‌های کاربردی را امکان‌پذیر می‌کند. توصیف رده‌های مختلف زمینه و انواع قابل به‌کارگیری آن در سیستم‌های توصیه‌گر در زیر ارائه شده است:

✓ فیزیکی: زمینه فیزیکی، موقعیت‌های محیطی کاربر یا سیستم در یک مکان خاص را نشان می‌دهد. زمینه فیزیکی شامل ویژگی‌هایی مثل نور، گرما و صدا است. بر اساس نظر وربرت و همکارانش [۹]، زمینه فیزیکی به طور گسترده در تحقیقات مربوط به خودکارسازی منازل مورد بررسی قرار گرفته است.

✓ محاسبه: زمینه محاسبه به طور گسترده توسط ریکی [۱۰] و انجمن تحقیقاتی یادگیری فراگیر [۱۱] مورد بررسی قرار گرفته است. کسب زمینه محاسباتی لازم برای حمایت از واسط‌های هوشمند

• استفاده از اطلاعات زمینه‌ای مکان، تعداد بازدید (میزان محبوبیت)، وضعیت آب و هوا، موقعیت زمانی در روز، حالت روحی کاربر، میزان ترافیک، فصل و دما

در ادامه مقاله، در بخش ۲ به مبانی تحقیق و معرفی سیستم‌های پیشنهاددهنده، دسته‌ها و انواع زمینه و مسائل باز در این حوزه خواهیم پرداخت و در بخش ۳ کارهای مرتبط در زمینه سیستم‌های توصیه‌گر زمینه آگاه بررسی می‌شود. در بخش ۴ روش پیشنهادی مطرح شده، نتایج ارزیابی روش پیشنهادی به تفصیل در بخش ۵ بیان می‌شود و در پایان نتیجه‌گیری و کارهای آینده بیان شده است.

۲- مبانی تحقیق

سیستم توصیه‌گر سفر زمینه آگاهی که در این مقاله معرفی شده است، مبتنی بر الگوریتم خوشه‌بندی فازی و الگوریتم بهینه‌سازی رقابت استعماری است. در این بخش ابتدا مفاهیم سیستم‌های توصیه‌گر و داده‌های زمینه‌ای و سپس الگوریتم‌های خوشه‌بندی C-Means فازی و الگوریتم بهینه‌سازی رقابت استعماری به عنوان مبانی و بلاک‌های سازنده روش پیشنهادی مورد بررسی قرار می‌گیرد.

۲-۱- سیستم‌های توصیه‌گر

سیستم‌های توصیه‌گر سیستم‌هایی هستند که در پیدا کردن و انتخاب نمودن آیتم‌های موردنظر کاربران به آن‌ها کمک می‌کنند. طبیعی است که این سیستم‌ها بدون در اختیار داشتن اطلاعات کافی و صحیح در مورد کاربران و آیتم‌های موردنظر آن‌ها (به عنوان مثال فیلم، موزیک، کتاب و...) قادر به تولید پیشنهاد مناسب نمی‌باشند. بنابراین یکی از اساسی‌ترین اهداف آن‌ها جمع‌آوری اطلاعات گوناگون در رابطه با سلیق کاربران و آیتم‌های موجود در سیستم است [۱]. منابع و روش‌های گوناگونی برای جمع‌آوری چنین اطلاعاتی وجود دارد. یکی از روش‌های جمع‌آوری اطلاعات به صورت صریح است که در آن کاربر با صراحت اعلام می‌کند که به چه آیتم‌هایی علاقه دارد (به عنوان نمونه با امتیاز دادن به یک موزیک) و روش دیگر روش ضمنی است که کمی دشوارتر است [۷] و در آن سیستم باید سلیق کاربر را با بررسی و دنبال کردن رفتارها و فعالیت‌های او بیابد. (به عنوان مثال باید ببینید کاربر به چه موزیک‌هایی بیشتر گوش می‌دهد، چه صفحاتی را بازدید می‌کند، با چه کسانی در ارتباط است و ...).

² Demographic

خوشه‌بندی با فراهم کردن قابلیت ورود به فضای داده و تشخیص ساختار آن‌ها، یکی از ایده‌آل‌ترین روش‌ها برای کار با دنیای عظیم داده‌ها است [۱۶]. خوشه‌بندی یکی از شاخه‌های یادگیری بدون نظارت است و فرایند خودکاری است که طی آن، نمونه‌ها به دسته‌هایی که اعضای آن مشابه یکدیگر می‌باشند تقسیم می‌شوند که به این دسته‌ها خوشه گفته می‌شود. بنابراین خوشه مجموعه‌ای از اشیاء است که در آن اشیاء با یکدیگر مشابه بوده و با اشیاء موجود در خوشه‌های دیگر نامشابه می‌باشند [۱۷]. برای تشابه می‌توان معیارهای مختلفی را در نظر گرفت مثلاً می‌توان معیار فاصله را برای خوشه‌بندی مورد استفاده قرارداد و اشیائی که به یکدیگر نزدیک تر هستند را به‌عنوان یک خوشه در نظر گرفت که به این نوع خوشه‌بندی، خوشه‌بندی مبتنی بر فاصله نیز گفته می‌شود. روش‌های موجود خوشه‌بندی را می‌توان در پنج دسته، روش‌های سلسله مراتبی، جزءبندی، مبتنی بر تراکم، مبتنی بر شبکه و مبتنی بر مدل دسته‌بندی کرد [۱۸].

در میان روش‌های مختلف در خوشه‌بندی داده‌ها، روش خوشه‌بندی K-Means و C-Means فازی به‌طور گسترده در زمینه‌های مختلف مورد استفاده قرار می‌گیرد. الگوریتم K-Means یک از معروف‌ترین روش‌ها برای خوشه‌بندی داده‌ها به شمار می‌رود. در این الگوریتم، داده‌ها پس از تعدادی تکرار به k خوشه مختلف دسته‌بندی می‌شوند. یکی از مشکلات الگوریتم K-Means و اکثر الگوریتم‌های سنتی خوشه‌بندی این است که تعلق داده به هر خوشه با عدد صفر یا یک مشخص می‌شود. به عبارت دیگر در خوشه‌بندی کلاسیک، هر نمونه ورودی متعلق به یک و فقط یک خوشه است و نمی‌تواند عضو دو خوشه و یا بیشتر باشد. در واقع، خوشه‌ها همپوشانی ندارند.

برای حل چالش‌های روش خوشه‌بندی k-means، الگوریتم خوشه‌بندی فازی ارائه شده است [۱۹]. در خوشه‌بندی فازی یک نمونه می‌تواند متعلق به بیش از یک خوشه باشد. در الگوریتم خوشه‌بندی فازی تعلق هر داده به یک خوشه خاص با یک عدد حقیقی بین صفر و یک مشخص می‌شود. ایده بنیادین در خوشه‌بندی فازی به این ترتیب است که فرض کنیم هر خوشه مجموعه‌ای از عناصر است، سپس با تغییر در تعریف عضویت عناصر در این مجموعه از حالتی که یک عنصر فقط بتواند عضو یک خوشه باشد (حالت افزایی)، به حالتی که هر عنصر می‌تواند با درجه

که بتوانند منابع مناسب سیستم توصیه‌گر را انتخاب و توصیه کنند مهم است.

✓ مکان: زمینه مکان در تحقیقات زمینه آگاه بسیار فراگیر شده است [۱۲, ۱۳]. به دلیل افزایش استفاده از ابزارهای متحرک، سیستم‌های توصیه‌گر مبتنی بر مکان به طور گسترده مورد استفاده قرار گرفته‌اند [۹].

✓ زمان: زمینه زمان از اطلاعات تاریخ و زمان که معمولاً شامل دقیقه، ساعت، هفته، ماه، ترم، فصل و غیره است تشکیل می‌شود. زمان معمولاً مرکب بوده و در رابطه با دسته‌های دیگری مانند مکان به صورت مدت زمان یا مهر زمانی مورد استفاده قرار می‌گیرد. زمینه زمان، یک نمونه یا دوره را در مدتی که اطلاعات زمینه‌ای مورد نیاز است یا مربوط به کاربر سیستم توصیه‌گر است نشان می‌دهد [۹, ۱۴].

✓ کاربر: شخصی است که از سیستم توصیه‌گر استفاده می‌کند یا آیتم‌های توصیه برای وی تولید می‌شود. برای مثال کاربران در آموزش الکترونیک، یادگیرنده‌های الکترونیکی هستند و یک مدل آموزشی را برای تولید توصیه‌های دقیق و ارزشمند یاد می‌گیرند. مدل‌های یادگیر به طور گسترده به مدل‌سازی آموزشی کاربر و ابررسانه تطبیقی آموزشی دسترسی دارند [۹]. زمینه کاربر در چارچوب سیستم توصیه‌گر می‌تواند شامل اطلاعات شخصی، دانش/عملکرد، علائق، روش‌های یادگیری و شناختی و پس زمینه کاربر باشد [۱۵]

✓ روابط اجتماعی^۳: روابط اجتماعی، حلقه‌های اجتماعی، انجمن‌ها، ارتباطات و وابستگی بین دو یا چند شخص را توصیف می‌کنند. برای مثال، روابط اجتماعی می‌تواند اطلاعاتی در مورد دوستان، اشخاص، همسایگان، همکاران، دشمن‌ها و فامیل باشد. بعضی از محققین روابط اجتماعی را به عنوان بعد زمینه‌ای مهم شناسایی نموده‌اند [۹].

۲-۳- خوشه بندی C-Means فازی

3 Time Stamp

4 Educational Adaptive Hypermedia

5 Social relations

و بر مبنای آن، کشورهای مستعمره، بین استعمارگران تقسیم می‌شوند:

$$P_n = \left| \frac{C_n}{\sum_{i=1}^{N_{imp}} C_i} \right| \quad (2)$$

بنابراین تعداد اولیه مستعمرات یک استعمارگر برابر خواهد بود با:

$$N.C_n = \text{round}\{P_n . (N_{col})\} \quad (3)$$

که در آن $N.C_n$ ، تعداد اولیه مستعمرات یک امپراطوری و N_{col} نیز تعداد کل کشورهای مستعمره موجود در جمعیت کشورهای اولیه است. با داشتن حالت اولیه تمام امپراطوری-ها، الگوریتم رقابت استعماری شروع می‌شود. سیاست کلی در این الگوریتم به این صورت است که کشورهای استعمارگر، کشورهای مستعمره را به طرف خود می‌کشند. همان طور که در رابطه ۴ نشان داده شده است، هزینه کل هر امپراطوری، علاوه بر هزینه کشور استعمارگر از بخشی از میانگین هزینه مستعمره‌ها نیز تشکیل می‌شود.

$$C_n = \text{cost}(\text{imperialist}) + \xi \text{mean}(\text{cost}(\text{colonies of impire}_n)) \quad (4)$$

کشور مستعمره، به اندازه X واحد در جهت استعمارگر حرکت کرده و به موقعیت جدیدی کشانده می‌شود. اگر d را فاصله میان استعمارگر و مستعمره و X نیز عددی تصادفی با توزیع یکنواخت باشد؛ برای X داریم:

$$x \sim U(0, \beta \times d) \quad (5)$$

که در آن β عددی بزرگ‌تر از یک و نزدیک به ۲ است. مقدار $\beta=2$ می‌تواند یک انتخاب مناسب باشد. وجود ضریب $\beta > 1$ باعث می‌شود تا کشور مستعمره در حین حرکت به سمت کشور استعمارگر، از جهت‌های مختلف به آن نزدیک شود. برای افزایش ناحیه جستجوی اطراف استعمارگر، یک انحراف زاویه ای برابر θ که از یک توزیع یکنواخت تصادفی تبعیت می‌نماید، به بردار اصلی اضافه می‌شود:

$$\theta \sim U(-\gamma, \gamma) \quad (6)$$

γ پارامتری است که محدوده انحراف زاویه ای را کنترل می‌کند. گام بعدی، گام رقابت استعماری می‌باشد. در این مرحله ضعیف‌ترین مستعمره از ضعی ترین امپراطوری انتخاب شده و به یک امپراطوری قوی (نه لزوماً قوی‌ترین امپراطوری) داده می‌شود. البته نحوه انتخاب امپراطوری موردنظر به این صورت است که هرچه یک امپراطوری قوی‌تر باشد، احتمال

عضویت‌های مختلف به چندین خوشه مختلف تعلق داشته باشد، خوشه‌بندی‌هایی که انطباق بیشتری با واقعیت دارند ارائه کنیم [۱۹]. در سال‌های اخیر نسخه‌های بهبودیافته‌ای از این الگوریتم نیز ارائه شده است. برای حل مشکلات ناشی از مقاردهی اولیه خوشه‌ها، استتکو و همکارانش یک طرح جدید در مقاردهی اولیه خوشه‌ها در خوشه‌بندی C-Means روش فازی ارائه دادند [۲۰]. روش فازی C-Means به‌طور گسترده در زمینه‌های مختلف مانند سنجش از راه دور، خوشه‌بندی سری‌های زمانی و قطعه‌بندی تصاویر رنگی و ... مورد استفاده قرار می‌گیرد.

یکی دیگر از مشکلات خوشه‌بندی، تعیین تعداد بهینه خوشه‌ها است. در این مقاله برای حل این مشکل و اجرای حدس اولیه تعیین تعداد خوشه‌ها پیش از اجرای الگوریتم، از الگوریتم بهینه‌سازی رقابت استعماری استفاده شده است که در بخش ۴-۲ توضیح داده شده است.

۲-۴- الگوریتم بهینه سازی رقابت استعماری

الگوریتم بهینه‌سازی رقابت استعماری، یک الگوریتم تکاملی است که با هدف بهینه‌سازی مسائل مختلف از رقابت استعمارگران الهام گرفته شده است. همانند دیگر الگوریتم-های تکاملی، این الگوریتم، نیز با تعدادی جمعیت اولیه تصادفی که هر کدام از آنها یک کشور نامیده می‌شوند، شروع می‌شود. تعدادی از بهترین عناصر به منزله‌ی استعمارگر انتخاب و باقیمانده جمعیت نیز به منزله‌ی مستعمره در نظر گرفته می‌شوند.

در این الگوریتم، برای حل مساله بهینه‌سازی مفروض، N کشور در نظر گرفته می‌شود که هر کدام با یک بردار نمایش داده شده و نشان دهنده یک نقطه در فضای n بعدی می‌باشد. از بین این نقاط، نقاطی که کمترین هزینه را طبق تابع بهینه‌سازی داشته باشند، به عنوان استعمارگر و بقیه نقاط به عنوان مستعمره در نظر گرفته می‌شوند [۲۱]. در ابتدا برای هر استعمارگر هزینه نرمال شده به صورت زیر محاسبه می‌شود:

$$C_n = c_n - \max_i \{c_i\} \quad (1)$$

در رابطه (۱) C_n ، هزینه نرمال شده استعمارگر n ام، $\max\{c_i\}$ بیشترین هزینه میان استعمارگران و c_n هزینه استعمارگر n ام می‌باشد. با داشتن هزینه نرمال شده، قدرت نسبی نرمال شده هر استعمارگر، به صورت زیر محاسبه شده

✓ مدلسازی زمینه ای^۸ از اطلاعات زمینه‌ای به طور مستقیم در مدل تولید پیشنهادات استفاده می‌شود. تحقیقات بسیاری در ارتباط با سیستم‌های توصیه‌گر زمینه آگاه و کاربردهای آن‌ها در زمینه سفر و گردشگری انجام شده است که در این بخش تعدادی از آن‌ها بر اساس نوع الگو، نوع اطلاعات زمینه‌ای مورد استفاده، زمینه کاربردی و روش به کار برده شده برای تولید پیشنهادات، در جدول شماره ۱ ارائه داده شده‌اند. در مقایسه با اکثر روش‌های ارائه شده در این زمینه، داده‌های زمینه‌ای روش پیشنهادی در این مقاله علاوه بر داده‌های مکانی که شامل مختصات جغرافیایی مکان بازدید می‌شود، از داده‌های زمانی مانند موقعیت زمانی بازدید در روز، فصل بازدید و همچنین داده‌های مربوط به تعداد بازدید (میزان محبوبیت)، وضعیت آب و هوا، حالت روحی کاربر، میزان ترافیک، و دما نیز جهت افزایش دقت پیشنهادات استفاده می‌کند. روش پیشنهادی با هدف مدلسازی داده‌ها بر اساس داده‌های زمینه‌ای و با استفاده از ادغام دو الگوریتم رقابت استعماری و روش C-Means فازی، خوشه‌بندی انجام می‌دهد و مبتنی بر نتایج به دست آمده، نقاط بازدید مناسب را به کاربران پیشنهاد می‌دهد. روش پیشنهاد داده شده در این مقاله مبتنی بر الگوهای مدلسازی زمینه‌ای و پس پالایش زمینه‌ای است.

انتخاب آن بالاتر می‌رود. در نهایت وقتی یک امپراطوری همه مستعمره‌های خود را از دست داد، آن امپراطوری از لیست امپراطوری‌ها خارج شده و در یک رقابت استعماری به عنوان مستعمره به امپراطوری‌های دیگر داده می‌شود. روند تکامل در یک حلقه قرار دارد که تا برآورده شدن شرط توقف ادامه می‌یابد [۲۱].

الگوریتم رقابت استعماری در روش پیشنهادی این مقاله، برای به دست آوردن تعداد خوشه‌ها و مراکز آن‌ها به کار برده شده است. در واقع مقادیر به دست آمده از این الگوریتم، تعداد خوشه‌ها و مراکز آن‌ها را برای اجرای خوشه-بندی فازی فراهم می‌کنند.

۳- مروری بر تحقیقات پیشین

در گذشته، سیستم‌های توصیه‌گر، پیشنهادات را فقط مبتنی بر اطلاعات کاربران و آیتم‌ها تولید می‌کردند اما سیستم‌های توصیه‌گر مبتنی بر زمینه، برای ارائه پیشنهادات، علاوه بر اطلاعات کاربر و آیتم‌ها از اطلاعات مختلف زمینه‌ای نیز استفاده می‌کنند. بر اساس شیوه و زمان استفاده از اطلاعات زمینه‌ای در سیستم‌های توصیه‌گر، سه نوع الگوی مختلف وجود دارد که عبارتند از [۲۲]:

✓ پیش پالایش زمینه ای؟ اطلاعات زمینه‌ای پیش از اجرای فرایند پیشنهاد دهنده، به عنوان یک روش فیلتر-کننده بر روی داده‌ها اعمال می‌شوند.

✓ پس پالایش زمینه ای؟ اطلاعات زمینه‌ای در ابتدای فرایند نادیده گرفته می‌شوند و پس از آن که پیشنهادات بر اساس مدل توصیه‌گر ایجاد شد، از اطلاعات زمینه‌ای برای فیلتر کردن نتایج و ارائه پیشنهادات مطابق با علاقه مندی-های کاربر، استفاده می‌شود.

⁶ Contextual Pre-filtering

⁷ Contextual post-filtering

⁸ Contextual modeling

جدول ۱: مقایسه تحقیقات پیشین بر اساس تکنیک پیشنهادی، نوع الگو، نوع اطلاعات زمینه‌ای، کاربرد و رویکرد

مقاله	تکنیک سیستم توصیه‌گر	نوع الگو	نوع اطلاعات زمینه‌ای	کاربرد	رویکرد مقاله نسبت به اطلاعات زمینه‌ای
[۲۳]	مبتنی بر مدل، رگرسیون خطی ^۹	مدلسازی زمینه‌ای	شبکه‌های اجتماعی، زمان	هتل، گردشگری	بر اساس مدل رگرسیون خطی و داده‌هایی که ارائه دهنده زمینه مورد علاقه کاربر هستند، علاقه مندی‌های کاربر پیش بینی می‌شود. این داده می‌تواند بر اساس سه روش احتمالی محاسبه شوند: روش مبتنی بر اطلاعات انحصاری، روش مبتنی بر تقویت اطلاعات، روش مبتنی بر احتمال خی-۲
[۲۴]	روش‌های مکاشفه‌ای	پس پالایش زمینه‌ای	مکان، زمان، آب و هوا	هتل و گردشگری	ویژگی‌های زمینه‌ای مسافرت‌های گذشته کاربر بر اساس هر مکان بررسی می‌شود. پیشنهادات مبتنی بر زمینه از طریق یافتن کاربرانی که شباهت بیشتری به یکدیگر دارند استخراج شده و برای هر مکان امتیازی به آن تعلق می‌گیرد و پیشنهاداتی که مکان آنها شرایط زمینه‌ای را ندارند فیلتر و حذف می‌شوند.
[۲۵]	مدلسازی موضوعی، شبکه عصبی مصنوعی، ماشین بردار پشتیبان	پیش پالایش و مدلسازی زمینه‌ای	مکان، شرایط آب و هوا، داده‌های بلاگ‌های گردشگری	مکان‌های گردشگری	روش تخصیص دریکله پنهان، موضوعات نظرات گردشگران را استخراج کرده و روش ماشین بردار پشتیبان برای آن‌ها تحلیل احساسات انجام می‌دهد. سپس این داده‌ها به همراه امتیازها، رتبه‌های مکان‌ها و اطلاعات تاریخچه‌ای کاربران و مکان‌ها با استفاده از یک شبکه عصبی مصنوعی پردازش شده و مکان‌های موردعلاقه کاربر را که کمتر موردتوجه بوده‌اند، به کاربر پیشنهاد داده می‌شود.
[۲۶]	کاوش داده‌های استخراج شده از وب	مدلسازی زمینه‌ای	زمان، مکان	شهرهای گردشگری	در این سیستم پیشنهاد دهنده گردشگری چند سطحی، علاوه بر علاقه مندی‌های کاربر، محدودیت‌های او نیز در نظر گرفته می‌شود. ابتدا بر اساس تاریخ‌های سفر و بودجه گردشگران، تعدادی از شهرها انتخاب می‌شوند. سپس برای هر شهر یک امتیاز در نظر گرفته می‌شود که بر اساس محدودیت‌ها و علاقه مندی‌های کاربران محاسبه می‌شود و بر اساس آن‌ها شهرهای مناسبی به کاربر پیشنهاد داده می‌شود.
[۲۷]	تکنولوژی‌های اینترنت اشیا، آنتولوژی، وب معنایی	پیش پالایش و مدلسازی زمینه‌ای	زمان، مکان، شبکه‌های اجتماعی، داده‌های سلامت کاربران، سنسور ابزارها	گردشگری	یک سیستم گردشگری هوشمند مبتنی بر آنتولوژی و زمینه آگاه که با استفاده از تکنولوژی اینترنت اشیا و سرویس‌های وب معنایی به گردشگران سرویس‌های متناسبی را پیشنهاد می‌دهد
[۲۸]	ماشین بردار پشتیبان	پیش پالایش و مدلسازی زمینه‌ای	داده‌های جمع‌آوری شده از سایت‌های گردشگری	گردشگری	سیستم توصیه‌گر پروفایل‌های کاربران را از طریق TripAdvisor استخراج کرده و مجموعه داده به دست آمده را با روش ماشین بردار پشتیبان پردازش می‌کند. هدف این سیستم، ایجاد سهولت و بالا بردن دقت در پیشنهادات ارائه شده به گردشگران بر اساس داده‌های باز الت.
[۲۹]	پیش بینی لینک نظارت شده و مبتنی بر همسایه	مدلسازی زمینه‌ای	مکان	هتل	در این سیستم یک شبکه دو بخشی تعاملات مسافر و هتل ساخته می‌شود که از روش پیش‌بینی لینک مبتنی بر همسایگی برای پیش‌بینی ارتباطاتی که بر روی شبکه وجود ندارند استفاده می‌شود.
[۳۰]	تئوری گراف، متدهای تحلیل شبکه، فیلترکننده ترکیبی	مدلسازی زمینه‌ای	جنسیت، زبان و کشور گردشگر، مکان، امتیاز و نوع هر مکان توریستی	جذابیت‌های گردشگری	این سیستم توصیه‌گر ترکیبی شخصی‌سازی شده مبتنی بر داده‌های اجتماعی، با استفاده از میزان شباهت، اعتماد و روابط اجتماعی میان گردشگران، جذابیت‌های گردشگری موردعلاقه گردشگران را به آن‌ها پیشنهاد می‌دهد.
[۳۱]	فیلترکننده مشارکتی، تجزیه ماتریس، کاهش گرادیان تصادفی	پس پالایش زمینه‌ای	صرفه اقتصادی، هتل، مکان هتل، نظرات و احساسات مهمان-های پیشین	هتل	اطلاعات انبوه سپاری هتل‌ها شامل اطلاعات رسمی هتل، امتیازات چند معیاری و نظرات متنی مسافران، استخراج می‌شود تا با استفاده از فیلترکننده مشارکتی برای هتل‌ها و مسافران پروفایل ایجاد شود. سپس بر اساس روش تجزیه ماتریس کاهش گرادیان تصادفی پیشنهادات اولیه پیش‌بینی شده و بعد از اعمال اطلاعات زمینه‌ای بر روی آن‌ها، پیشنهادات نهایی ارائه می‌شود.

9-linear regression

1 Open data

سیستم توصیه‌گر زمینه آگاه سفر با بهره‌گیری از اطلاعات عکس‌های برچسب‌دار جغرافیایی

[۳۲]	خوشه‌بندی P-DBSCAN , تئوری بیز	پس‌پالایش و مدل‌سازی زمینه‌ای	مکان، جنسیت کاربر	گردشگری	از معیار تحرک مبتنی بر آنتروپی برای دسته‌بندی عکس‌هایی که برچسب جغرافیایی دارند استفاده شده و با تشخیص چهره در عکس‌هایی که مربوط به تورهای گردشگری هستند، تعیین جنسیت انجام می‌شود. سپس پروفایلی از کاربران و مکان‌های گردشگری مبتنی بر جنسیت ایجاد شده و با بررسی جنسیت کاربران و شباهت آن‌ها، مکان‌های گردشگری پیشنهاد داده می‌شود.
[۳۳]	فیلترکننده مشارکتی	پس‌پالایش و مدل‌سازی زمینه‌ای	شرایط آب و هوایی، زمان، مکان	مکان گردشگری	با بهره‌گیری از روش فیلترکننده مشارکتی، از تاریخچه سفرهای گردشگران و اطلاعات زمینه‌ای مانند شرایط آب و هوایی، زمان و مکان آن‌ها استفاده می‌شود تا مکان‌هایی در شهرهای جدید به آن‌ها پیشنهاد داده شود که مشابه با مکان‌های قبلی مورد علاقه گردشگران باشد.
[۳۴]	رتبه‌بندی بر اساس فیلترکننده مشارکتی	پیش‌پالایش و مدل‌سازی زمینه‌ای	شرایط آب و هوایی، زمان، مکان	مکان گردشگری	ارایه مکان‌های گردشگری پیشنهادی مبتنی بر معنا، با بهره‌گیری از اطلاعات زمینه‌ای صورت می‌گیرد. اساس این روش بر الگوریتم فیلترکننده مشارکتی و رتبه‌بندی اطلاعات زمینه‌ای است.
[۳۵]	مدل کردن موضوعی پویا، تجزیه ماتریس	مدل‌سازی و پس‌پالایش زمینه‌ای	داده شخصی گردشگران، زمان، مکان، اطلاعات ورود و خروج، اطلاعات هواشناسی	مکان گردشگری	مدل‌سازی موضوعی پویا برای بدست آوردن توزیع مبتنی بر زمان و اطلاعات ضمنی کاربر به کار گرفته شده و علاوه بر آن حجم زیادی از داده‌های صریح بر اساس اطلاعات ورود و خروج، محتویات بصری و دسته‌بندی‌های مکان‌های مورد علاقه گردشگران استخراج می‌شوند. سپس این اطلاعات برای تشخیص شباهت‌های کاربر-کاربر و مکان-مکان به کار گرفته می‌شود و تجزیه ماتریس بر اساس آن‌ها صورت می‌گیرد.
[۳۶]	ترکیب پالایش مشارکتی و مبتنی بر محتوی	پس‌پالایش زمینه‌ای	اطلاعات اینترنت اشیا و شبکه‌های اجتماعی، اطلاعات زمینه‌ای کاربر	مکان رویداد	این سیستم توصیه‌گر مبتنی بر مکان و گرایش آگاه که به صورت ترکیبی طراحی شده است، برای پیشنهاد رویدادها علاوه بر اطلاعات زمینه‌ای کاربران، از اطلاعات زمینه‌ای دستگاه‌های هوشمند مربوط به کاربران با استفاده از اینترنت اشیا نیز بهره می‌برد. همچنین از اطلاعات مربوط به اجتماعات مجازی در شبکه‌های اجتماعی برای زمانی که علاقه مندی‌های کاربر بر اساس داده‌های اینترنت اشیا قابل دستیابی نیست، استفاده می‌کند.
[۳۷]	تجزیه ماتریس احتمالاتی، تحلیل احساسات	پیش‌پالایش و مدل‌سازی زمینه‌ای	داده زمینه‌ای احساسات، اطلاعات جغرافیایی مکان	نقطه مورد علاقه	خصوصیات احساسی کاربران در مکان‌های مختلف مشخص می‌شود و برای کاوش مکان مورد علاقه باید اطلاعات زمینه‌ای مکانی و احساسی را با هم ترکیب کند. سپس با مشارکت فاکتورهای فاصله‌های جغرافیایی و شباهت احساسی، مکان مورد علاقه کاربر را پیشنهاد می‌دهد.
[۳۸]	TF-IDF، دسته‌بندی متون، FastText	مدل‌سازی زمینه‌ای	برچسب جغرافیایی عکس‌ها، اطلاعات جغرافیایی مکان	پیشنهاد مکان	این سیستم توصیه‌گر مبتنی بر مدل، برای هر مکان یک پروفایل ایجاد می‌کند تا بر اساس آن‌ها و علاقه‌مندی‌های کاربران، مکان‌هایی را به آن‌ها پیشنهاد دهد. در این روش ابتدا با استفاده از TF-IDF، برچسب‌های مکان‌های استخراج شده از توضیحات عکس‌ها در اینستاگرام، رتبه‌بندی می‌شوند و سپس با روش FastText برچسب‌های توضیحات عکس‌ها دسته‌بندی می‌شوند.
[۳۹]	تجزیه ماتریس، K-means	مدل‌سازی زمینه‌ای، پیش‌پالایش	داده‌های زمینه‌ای مکانی مربوط به مسیر و مقصد سفر	مکان سفر، مسیر سفر	ابتدا یک روش داده محور برای کاوش سفرها در شبکه‌های اجتماعی مبتنی بر مکان ارایه می‌شود که نحوه سفر کردن کاربران را بررسی می‌کند و الگوهای حرکتی آن‌ها را به دست می‌آورد. از اطلاعات به دست آمده برای خوشه‌بندی مسافران و همچنین خوشه‌بندی فاصله‌های مقصد‌های سفرها استفاده می‌کند.
[۴۰]	تجزیه ماتریس، تخصیص دریکله پنهان (LDA)	مدل‌سازی زمینه‌ای، پیش‌پالایش	برچسب جغرافیایی عکس‌ها، اطلاعات متنی، ترتیبی و تصویری	مکان سفر	با استفاده از عکس‌هایی دارای برچسب جغرافیایی، یک سیستم توصیه‌گر سفر شخصی‌سازی شده ارایه شده که مبتنی بر تجزیه ماتریس محدود شده برای اطلاعات چندگانه وزن دار است. با استفاده از اطلاعات چندگانه ابتدا برای کاربران و مکان‌های سفر پروفایل ایجاد شده و برای پیشنهادات مکان سفر، به داده‌های متنی، تصویری و فاصله‌ای، وزن‌های مختلفی بر پایه شباهت آن‌ها اختصاص داده می‌شود و از این وزن‌ها برای تجزیه ماتریس کاربر-مکان استفاده می‌شود.

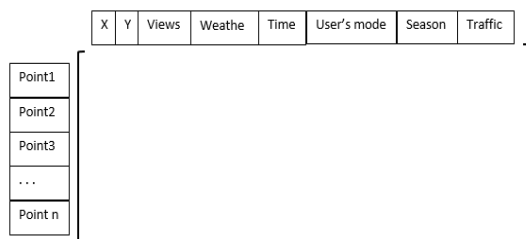


شکل ۱: فرآیند اجرای روش پیشنهادی

وضعیت آب و هوا، موقعیت زمانی در روز، حالت روحی کاربر، میزان ترافیک، فصل و دما. بنابراین داده‌های زمینه‌ای مکانی شامل طول و عرض جغرافیایی و داده‌های زمینه‌ای زمانی شامل فصل و زمان بازدید هستند. ویژگی‌های تعداد بازدید از مکان موردنظر، وضعیت آب و هوا، میزان ترافیک و وضعیت روحی کاربر در جهت افزایش دقت روش پیشنهادی و ارائه پیشنهادات دقیق‌تر مورد استفاده قرار گرفته است. شکل ۲ ساختار (سطرها و ستون‌های) ماتریس مربوط به داده‌های زمینه‌ای را نشان می‌دهد.

جدول ۲: ویژگی‌های نقطه بازدید

نام ویژگی	شرح
X (Lat)	مختصات جغرافیایی X نقطه بازدید
Y(Lon)	مختصات جغرافیایی Y نقطه بازدید
Views	تعداد بازدید از مکان موردنظر (محبوبیت مکان)
Current Weather	وضعیت هوای در زمان گرفتن عکس
Time Of The Day	زمان بازدید در طول روز
User's Mode	وضعیت روحی کاربر در زمان بازدید
Season	فصل زمان بازدید
Traffic	میزان ترافیک در زمان ترافیک



شکل ۲: ساختار ماتریس ویژگی مبتنی بر اطلاعات زمینه‌ای

۴- روش پیشنهادی

سیستم توصیه‌گر سفر زمینه آگاهی که در این مقاله ارائه شده است، مبتنی بر الگوریتم خوشه‌بندی فازی و الگوریتم بهینه‌سازی رقابت استعماری است. شکل ۱ نشان دهنده فرآیند اجرای روش پیشنهادی است که هر یک از فازهای فرآیند روش پیشنهادی به تفکیک در این بخش توضیح داده شده است.

۴-۱- استخراج اطلاعات زمینه‌ای و ایجاد ماتریس

ویژگی

روش پیشنهادی بر اساس مجموعه داده فلیکر^۱ یکی از بزرگ‌ترین سایت‌های اشتراک‌گذاری تصویر و ویدیو برخط^{۱۲} و قابل دسترسی برای عموم، ارزیابی شده است. مقاله [۴۰] بر تاثیر اطلاعات و داده‌های جغرافیایی بر روی بهبود کارایی سیستم‌های توصیه‌گر تاکید کرده است. دلیل این نکته نزدیک بودن اغلب مکان‌های توریستی به یکدیگر است، به طور مثال ۷۶ درصد مکان‌های توریستی شهر پاریس کمتر از ۵ کیلومتر از هم فاصله دارند. اما باید این نکته را در نظر بگیریم که فقط ویژگی‌های جغرافیایی و فاصله‌های مکانی حائز اهمیت نیستند. به عنوان مثال اگر به کاربری که در مرکز شهر است و نیاز به استراحت دارد، فقط بر اساس طول و عرض جغرافیایی، مکانی نزدیک پیشنهاد داده شود، ممکن است باعث عدم رضایت کاربر شده و مکان پیشنهادی را انتخاب نکند. بنابراین علاوه بر برچسب‌های جغرافیایی باید ویژگی‌های دیگری نیز به عنوان تکمیل کننده اطلاعات و ارائه پیشنهادات دقیق‌تر در نظر گرفته شوند [۳۷].

برای استخراج اطلاعات زمینه‌ای موردنیاز برای روش پیشنهادی، برچسب‌های نشان داده شده در جدول ۲ به ازای هر مکان بازدید شده، از سایت استخراج شده‌اند. ویژگی‌های استخراج شده به ازای هر نقطه از بازدید عبارتند از: طول و عرض جغرافیایی مکان، تعداد بازدید (میزان محبوبیت)،

1 -www.flickr.com/services/api

1

1 -online

2

۴-۲- تعیین تعداد خوشه‌ها با استفاده از الگوریتم

رقابت استعماری

برای خوشه‌بندی مکان‌های مختلف، ابتدا با استفاده از الگوریتم رقابت استعماری، بر اساس ویژگی‌های زمینه‌ای مختلف مربوط به مکان‌ها، مکان‌های مشابه در یک خوشه قرار می‌گیرند. بنابراین بر اساس برچسب‌های مکان‌های بازدید و بدون نیاز به مشخص کردن تعداد خوشه‌ها از پیش، به‌طور خودکار مکان‌های بازدید خوشه بندی می‌شوند. همان‌طور که در بخش ۲-۳ شرح داده شده است، ابتدا به کشورها مقدردهی اولیه می‌شود و استعمارگران و مستعمره‌های آن‌ها تعیین می‌شود. سپس حرکت تصادفی کشورهای مستعمره به سمت استعمارگران آن‌ها انجام شده و در صورتی که قدرت مستعمره‌ها بیشتر از استعمارگران نظیر آن‌ها باشد، مستعمره‌ای که قدرت بیشتری دارد، به عنوان استعمارگر جدید شناخته شده و استعمارگر قدیم نیز مستعمره آن خواهد شد. سپس ضعیف‌ترین مستعمره از ضعیف‌ترین استعمارگر به طور تصادفی به یکی دیگر از استعمارگرها ملحق می‌شود و بررسی می‌شود که اگر استعمارگری بدون مستعمره وجود دارد حذف شود. این مراحل تا زمانی تکرار می‌شوند که شرط پایان اجرا برقرار شود. شبه کد الگوریتم رقابت استعماری مطابق الگوریتم ۱ است.

الگوریتم ۱. شبه کد الگوریتم رقابت استعماری

The pseudo-code of the Imperialist Competitive Algorithm

Input: Numpopulation, NumInitialImperialists

Output: Best Imperialists

1. initialize population randomly
2. **For** $i=1$ **to** Numpopulation
3. Compute the evaluation cost c_i
4. Sort the computed cost c_i in descending order
5. Select NumInitialImperialists out of Numpopulation
6. Normalized the cost of each Imperialist c_n
7. Compute the normalized power of each Imperialist p_n
8. Assign Numcolonies remined countries to the Imperialists
9. **EndFor**
10. **For** $j=1$ **to** NumInitialImperialists
11. Move the colony toward the relevant Imperialist (assimilation)
12. Compute the cost of assimilated countries
13. Perform revolution on new colony
14. **If** the cost of new colony is less than cost of Imperialist
15. **Then** exchange the position of colony and Imperialist
16. Pick the weakest colony from the weakest empire and assign it to the empire that has most likelihood to possess it
17. **EndFor**
18. **If** there is Imperialist with no colonies
19. **Then** eliminate the Imperialists
20. **Until** stopping condition is reached

خوشه‌بندی مکان‌ها بر اساس ویژگی‌های زمینه‌ای آن‌ها به کار برده شده است. در [۴۱] یک الگوریتم رقابت استعماری بهبود یافته ارائه شده است که نسبت به الگوریتم رقابت استعماری کلاسیک عملکرد بهتری در زمینه خوشه بندی دارد که در روش پیشنهادی این مقاله به کار برده شده است. میزان خوب بودن نتیجه یک خوشه‌بندی را می‌توان بر اساس معیارهای ارزیابی خوشه‌بندی که مطابق با توابع ریاضی-آماري هستند، نشان داد.

یک معیار ارزیابی خوشه‌بندی دو هدف را دنبال می‌کند: ۱. مشخص کردن تعداد خوشه‌ها ۲. به دست آوردن بهترین حالت خوشه‌بندی با توجه به تعداد خوشه‌ها. هر معیار ارزیابی خوشه‌بندی باید دو وجه خوشه‌بندی زیر را مدنظر قرار دهد: ۱. پیوستگی یا فشردگی: الگوهای موجود در یک خوشه باید تا حد امکان به یکدیگر شبیه باشند. واریانس یا پراکندگی الگوهای موجود در یک خوشه نمایان‌گر پیوستگی یا فشردگی الگوهای درون یک خوشه هستند. ۲. تفکیک: خوشه‌ها باید تا حد امکان از هم فاصله داشته باشند. فاصله بین مراکز خوشه‌ها (به عنوان مثال فاصله اقلیدسی) می‌تواند نمایانگر تفکیک خوشه‌ها باشد [۴۲].

معیارهای ارزیابی مختلفی برای ارزیابی خوشه‌بندی غیر فازی وجود دارند که از جمله آن‌ها می‌توان به معیار PBM^۱، DI^۲، DB [۴۳]، CS [۴۴] اشاره کرد. برای تمام این معیارها، مقدار بیشینه و یا کمینه آن‌ها نشان‌دهنده خوشه‌بندی بهینه مجموعه الگوها و یا داده‌هاست. از این‌رو می‌توان آن‌ها را در الگوریتم‌های بهینه‌سازی فرا ابتکاری استفاده نمود. پیچیدگی محاسباتی معیار DI با افزایش تعداد خوشه‌ها و اندازه داده‌ها به شدت افزایش می‌یابد و همچنین معیار PBM بیشتر در مواردی که تعداد خوشه‌ها کم و داده‌ها دارای پراکندگی پایین و تراکم بالا هستند، مورد استفاده قرار می‌گیرد. معیار DB علاوه بر اینکه در مجموعه داده‌های بزرگ و مجموعه داده‌های سلسله مراتبی، پیچیدگی محاسباتی قابل قبولی دارد، برای داده‌هایی با ابعاد متفاوت نیز می‌تواند نتایج خوبی ارائه دهد. معیار CS نیز برای خوشه‌هایی با چگالی‌های متفاوت به صورت موثر عمل می‌کند. بنابراین در این مقاله برای ارزیابی خوشه‌بندی از

¹ - Pakhira, Bandyopadhyay & Maulik (PBM)

¹ - Dunn index (DI) 4

¹ - Davies & Bouldin (DB) 5

¹ - Chou, Su, Lai (CS) 6

در روش پیشنهادی این مقاله، الگوریتم رقابت استعماری هدف به دست آوردن تعداد خوشه‌ها و مراکز آن‌ها، برای

صورت کسر کمینه و مخرج بیشینه و در کل مقدار CS کمینه شود.

در واقع در هر تکرار الگوریتم فاصله هر نمونه از مرکز خوشه مشخص می‌شود و نمونه‌ها به خوشه‌ای اختصاص داده می‌شوند که کمترین فاصله را تا مرکز خوشه داشته باشد. سپس الگوریتم رقابت استعماری بهبود یافته بر روی مراکز خوشه‌ها اعمال می‌شود و تا جایی این فرایند تکرار می‌شود که تعداد و مراکز بهینه خوشه‌ها مشخص شود.

از آنجا که در ابتدا تعداد خوشه‌ها توسط کاربر مشخص نشده است، نحوه نمایش جواب‌ها باید طوری باشد که بتوان تعداد خوشه‌ها را همزمان با انجام خوشه‌بندی داده‌ها در زمان اجرا به دست آورد. بنابراین با فرض اینکه تعداد بیشینه خوشه‌ها برابر با K و تعداد ابعاد یا ویژگی‌های مجموعه داده برابر با d باشد، آنگاه هر یک از جواب‌های ممکن یک ماتریس با ابعاد $(K \times (d+1))$ خواهد بود که برای بدست آوردن تعداد بهینه خوشه‌ها، یک متغیر تصمیم به متغیرهای تصمیم موجود برای حل مسئله خوشه‌بندی اضافه می‌شود. به عبارتی هر یک از مرکز خوشه‌ها یک متغیر تصمیم به نام متغیر تصمیم فعال‌سازی دارد که مشخص می‌کند خوشه مربوطه فعال باشد و یا غیرفعال باشد. مقادیر مجاز برای متغیر تصمیم فعال‌سازی اعداد حقیقی در بازه $(0, 1)$ است. مرکز یک خوشه در صورتی فعال می‌شود که مقدار متغیر تصمیم مربوط به آن بزرگتر از 0.5 باشد. در غیر این صورت مرکز خوشه غیرفعال است.

برای ارزیابی جواب‌ها در الگوریتم رقابت استعماری، هر یک از معیارهای ارزیابی خوشه‌بندی می‌توانند مورد استفاده قرار گیرند. در این مقاله دو معیار ارزیابی خوشه‌بندی DB و CS به عنوان تابع هدف طبق رابطه (۱۳) در نظر گرفته شدند که هدف آن‌ها کمینه‌سازی مقدار معیار CS و DB تعریف شده است.

$$f_1 = CS(K), f_2 = DB(K) \quad (13)$$

۳-۴- خوشه‌بندی فازی داده‌ها

در خوشه‌بندی C-Means فازی اگر مجموعه‌ی مکان‌های بازدید را $L = [l_1, l_2, \dots, l_n]$ و مجموعه خوشه‌ها را برابر با $C = [c_1, c_2, \dots, c_c]$ در نظر گرفته شود هر عضو مجموعه L می‌تواند با یک مقدار بین ۰ تا ۱ به عنوان درجه عضویت در بیش از یک خوشه قرار بگیرد. بنابراین، با توجه به این‌که هر مکان بازدید در بیش از یک خوشه قرار

معیارهای DB و CS استفاده شده است. نتایج ارزیابی این معیارها در فصل ۵-۴ آمده است.

✓ **معیار ارزیابی DB:** این معیار تابعی از نسبت مجموع پراکندگی درون خوشه‌ای به پراکندگی بین خوشه‌هاست. ابتدا پراکندگی خوشه i ام و سپس فاصله بین خوشه i و خوشه j طبق روابط (۷) و (۸) محاسبه می‌شوند [۴۳]:

$$S_{i,a} = \left[\frac{1}{x} \sum \| \vec{X} - \vec{m} \|_2^q \right]^{\frac{1}{q}} \quad (7)$$

$$D_{i,j,t} = \left\{ \sum_{p=1}^d |m_{i,p} - m_{j,p}|^t \right\}^{\frac{1}{t}} \| \vec{m}_i - \vec{m}_j \|_t \quad (8)$$

که m_i مرکز خوشه i ام، $t \geq 1$ ، q و t می‌توانند به طور مستقل مقداردهی شوند. N_i تعداد الگوهای متعلق به خوشه C_i است. $D_{i,j,t}$ نرم t برای مراکز خوشه (m_i) و C_j خوشه (m_j) را محاسبه می‌کند. مقدار معیار ارزیابی DB بر اساس رابطه (۹) و (۱۰) محاسبه می‌شود.

$$R_{i,qt} = i \in K, j \neq i \left\{ \frac{S_{i,q} + S_{j,q}}{D_{ij,t}} \right\} \quad (9)$$

$$DB(K) = \frac{1}{K} \sum_{i=1}^K R_{i,qt} \quad (10)$$

در رابطه (۱۰)، K نشان‌دهنده خوشه K ام است. کمترین مقدار به دست آمده برای معیار DB، نشان‌دهنده خوشه‌بندی بهینه می‌باشد.

✓ **معیار ارزیابی CS:** قبل از محاسبه معیار ارزیابی CS، مرکز هر خوشه به وسیله میانگین الگوهای آن خوشه طبق رابطه (۱۱) مشخص می‌شود [۴۴]:

$$\vec{m}_i = \frac{1}{N_i} \sum_{x_i \in C_i} \vec{x}_j \quad (11)$$

در رابطه (۱۱)، N_i تعداد الگوهای متعلق به خوشه‌ی C_i است. معیار فاصله بین دو نمونه \vec{X}_i و \vec{X}_j به صورت $d(\vec{X}_i, \vec{X}_j)$ نشان داده می‌شود. معیار CS به صورت رابطه (۱۲) تعریف می‌شود:

$$CS(K) = \frac{\sum_{i=1}^K \left[\frac{1}{N_i} \sum_{\vec{X}_i \in C_i} \vec{X}_i \in \vec{X}_q \right]}{\sum_{i=1}^K [i \in K, j \neq i]} \quad (12)$$

همانند معیار ارزیابی DB، معیار CS برابر با نسبت فاصله درون خوشه‌ای به فاصله بین خوشه‌ای است و بنابراین باید

را اختیار کنند اما مجموع مؤلفه‌های هر یک از ستون‌ها باید برابر ۱ باشد و داریم:

$$\sum_{i=1}^c u_{ik} = 1, \forall k = 1, \dots, n \quad (15)$$

این شرط به این معنا است که مجموع تعلق هر نمونه به C خوشه باید برابر ۱ باشد. با استفاده از این شرط و کمینه کردن تابع هدف خواهیم داشت:

(۱۶)

(۱۷)

۴-۴- پیشنهاد نقاط بازدید به کاربر

پس از مشخص شدن مراکز و تعداد خوشه‌ها توسط الگوریتم رقابت استعماری، مراکز به عنوان مراکز اولیه الگوریتم C -Means فازی در نظر گرفته می‌شود. سپس الگوریتم C -Means فازی اجرا می‌شود و در نهایت مکان‌های خوشه‌بندی شده به دست می‌آید. سپس برای هر کاربر خوشه‌هایی که متعلق به مکان‌های آن کاربر هستند مشخص می‌شود. مثلاً ممکن است مکان‌هایی که یک کاربر از آن‌ها دیدن کرده است متعلق به خوشه‌های ۱ و ۲ و ۴ باشد. خوشه‌ای که بیش‌ترین مکان بازدید شده توسط کاربر را داراست به عنوان خوشه موردعلاقه کاربر در نظر گرفته می‌شود. سپس برای پیش‌بینی مکان‌های بعدی به کاربر مورد نظر، از مکان‌های متعلق به آن خوشه استفاده می‌شود. برای این کار مکان‌های متعلق به خوشه بر اساس نزدیک‌ترین مکان به مرکز خوشه مرتب می‌شوند. سپس مکان‌هایی که توسط کاربر بازدید شده‌اند و به آن خوشه تعلق دارند حذف می‌شوند و در نهایت از بقیه مکان‌ها بر اساس تعداد مکان مورد نیاز، مکان‌هایی به کاربر پیشنهاد می‌شود.

۵-ارزیابی

در این بخش، مجموعه داده استفاده شده برای ارزیابی روش پیشنهادی معرفی می‌شود، سپس نحوه پیش‌پردازش داده‌ها و مقداردهی به پارامترها، نتایج حاصل از ارزیابی‌ها و در نهایت مقایسه نتایج حاصل با دیگر روش‌های مشابه ارائه می‌شوند. در نمودارها و جداول ارزیابی از برچسب TRS^{18} GP برای روش پیشنهادی استفاده می‌شود.

۵-۱- مجموعه داده

فلیکر یکی از بزرگ‌ترین سایت‌های اشتراک‌گذاری تصویر و ویدئو، خدمات وب و جوامع آنلاین است که توسط شرکت

می‌گیرد، یکی از چالش‌های سیستم‌های توصیه‌گر که خلوت بودن داده‌ها است تا حدود زیادی حل می‌شود و همسایگان بیشتری برای مکان‌های بازدید هر خوشه وجود خواهد داشت. شبه کد الگوریتم خوشه‌بندی فازی مطابق الگوریتم ۲ است:

الگوریتم ۲. شبه کد الگوریتم خوشه‌بندی فازی

<p>The pseudo-code of the Fuzzy C-Means (FCM) Input: L and C. Output: final Fuzzy C-Means clusters. 1: Select an initial fuzzy pseudo-partition, i.e., assign values to all u_{ij} 2: Repeat 3: Compute the centroid of each cluster using the fuzzy partition. 4: Update the fuzzy partition, i.e., the u_{ij}. 5: Until the centroids don't change.</p>

همان‌طور که در شبه کد الگوریتم ۲ مشاهده می‌شود، اگر L مجموعه مکان‌های بازدید و C مجموعه خوشه و I و J به ترتیب شاخص‌های هر یک از خوشه‌ها و مکان‌ها در نظر گرفته شود، ابتدا میزان تعلق هر مکان به خوشه‌ها که بر اساس $u_{i,j}$ به دست می‌آید، محاسبه شده و سپس مرکز هر خوشه مشخص می‌شود. سپس مجدداً خوشه‌بندی مکان‌ها با توجه به مراکز جدید انجام می‌شود. این کار تا زمانی که تغییرات مراکز خوشه‌ها از یک حد آستانه مشخص کوچکتر شود، تکرار شده و در نهایت خوشه‌های نهایی به دست می‌آیند.

در الگوریتم خوشه‌بندی فازی تعداد خوشه‌ها (C) باید از قبل مشخص شده باشند تابع هدفی که برای این الگوریتم تعریف شده است به صورت زیر می‌باشد [۱۹]:

$$J = \sum_{i=1}^c \sum_{k=1}^n u_{ik}^m \|x_k - v_i\|^2 \quad (14)$$

در رابطه (۱۴) m یک عدد حقیقی بزرگ‌تر از ۱ است که در اکثر موارد برای m عدد ۲ انتخاب می‌شود. X_k نمونه k ام است و V_i مرکز خوشه i ام و n تعداد مکان‌های بازدید است. U_{ik} میزان تعلق نمونه i ام به خوشه k ام را نشان می‌دهد. علامت $\|\cdot\|^2$ میزان تشابه (فاصله) نمونه با (از) مرکز خوشه می‌باشد که می‌توان از هر تابعی که بیانگر تشابه نمونه و مرکز خوشه باشد را استفاده کرد. از روی U_{ik} می‌توان یک ماتریس U تعریف کرد که دارای C سطر و n ستون می‌باشد و مؤلفه‌های آن هر مقداری بین ۰ تا ۱ را می‌توانند اختیار کنند. اگر تمامی مؤلفه‌های ماتریس U به صورت ۰ و یا ۱ باشند الگوریتم مشابه C -Means کلاسیک خواهد بود. با اینکه مؤلفه‌های ماتریس U می‌توانند هر مقداری بین ۰ تا ۱

¹ Travel Recommender System- Geotagged Photo

ترتیب فصول سال و میزان ترافیک به اعداد صحیح نگاشت شده است. پس از انجام پیش‌پردازش بر روی مجموعه داده اولیه، مجموعه داده جدید با ساختار موجود در شکل ۳ بدست می‌آید.

همان‌گونه که در بخش ۴-۲ به آن اشاره شد اجرای الگوریتم رقابت استعماری نیاز به تعیین یک سری پارامترهای اولیه دارد که باید مشخص شود نظیر پارامترهای β ، γ ، تعداد جمعیت (کشورها)، تعداد استعمارگران اولیه و تعداد تکرارها. این مقادیر در جدول ۳ نشان داده شده است.

جدول ۳: مقادیر پارامترهای الگوریتم رقابت استعماری

پارامتر	مقدار
تعداد کشورها	۱۰۰
تعداد استعمارگران اولیه	۸
β	۲
γ	۰٫۷
ξ	۰٫۰۵
تعداد تکرار	۲۰۰

Ludicorp در سال ۲۰۰۴ تاسیس شده و در سال ۲۰۰۵ توسط یاهو خریداری شد. این سایت به مرور به یک شبکه اجتماعی مبتنی بر عکس تبدیل شده است [۴۵]. در این مقاله برای ارزیابی روش پیشنهادی، از مجموعه داده‌ای استفاده شده است که یک خزنده^۱ در بازه زمانی سال ۲۰۱۵ تا ۲۰۱۹ مجموعه‌ای از مکان‌های بازدید شده در کشور آمریکا و عمدتاً در قاره اروپا را بازیابی کرده و برچسب‌های مختصات جغرافیایی مکان، تعداد بازدید (میزان محبوبیت)، وضعیت آب و هوا، موقعیت زمانی در روز، حالت روحی کاربر، میزان ترافیک، فصل و دما را برای بازدیدها جمع‌آوری کرده است. جدول ۲ برچسب‌های استخراج شده به ازای هر مکان بازدید شده، را نشان می‌دهد.

۵-۲- پیش‌پردازش داده‌ها و مقداردهی پارامترها

به منظور یکسان‌سازی ویژگی‌ها و عملیات پاک‌سازی داده‌ها از میان ۱۰۰۰۰ رکورد، ۲۵۰۰ رکورد که شامل تمام ویژگی‌ها بود، برای شبیه‌سازی بکار رفته است. در مجموعه داده مورد استفاده ۱۰۰ کاربر وجود دارد که برای هر کاربر ۲۰ مکان ملاقات شده مشخص شده است. از بین این ۲۰ مکان، ۱۵ مکان برای آموزش و ۵ مکان برای آزمایش مدل پیشنهادی در نظر گرفته شده است. به عبارت بهتر، با استفاده از ۱۵ مکان ملاقات شده، ۵ مکان برای هر کاربر پیش‌بینی می‌شود. ویژگی‌های مجموعه داده مورد استفاده از مقادیر عددی و غیر عددی تشکیل شده است. به منظور استفاده از مقادیر غیر عددی برخی ویژگی‌ها، نگاشتی به‌کار گرفته می‌شود تا ویژگی‌های غیر عددی با مقادیر عددی مقداردهی شوند که در رابطه (۱۸) مشخص شده است:

$$NormalizedValue = \frac{\max(field) - \min(field)}{number} \quad (18)$$

$field$ مقدار ویژگی مورد نظر، $number$ تعداد بازه‌هایی است که برای نگاشت مقادیر غیر عددی در نظر گرفته شده و $NormalizedValue$ مقدار صحیح و نرمال‌سازی شده ویژگی مورد نظر است. در واقع مقادیر تمامی ویژگی‌ها به منظور نرمال‌سازی به مقادیر صحیح تبدیل شده است. به عنوان مثال مقادیر وضعیت هوا (آفتابی، ابری و بارانی، برفی و سرد، و مرطوب و معتدل) به مقادیر ۱ تا ۴، زمان بازدید در طول روز (صبح، ظهر، عصر و شب) به مقادیر ۱ تا ۴، وضعیت روحی کاربر (نرمال و طبیعی، شاد، عصبی، ناراحت، هیجان‌زده، خسته و نگران) به مقادیر ۱ تا ۶، و به همین

1	Lat	Lon	views	date	current weather	time of the day	user's mode	season code	traffic	temperature	time range season
2	51.02815	4.48078	899661	41622	1	2	1	1	14	-15.00	1 winter
3	38.20551	-85.7028	534764	41217	4	2	2	3	4	-9.00	9 autumn
4	39.95653	-75.1722	1117325	42043	1	1	2	1	11	-37.00	4 winter
5	45.53467	10.23746	1534894	42048	3	2	3	1	14	-23.00	5 winter
6	43.17257	2.345066	1784617	40678	2	3	1	4	5	-5.00	6 Spring
7	-1.75856	36.05921	913359	40566	2	1	3	1	14	-60.00	2 winter
8	38.47316	-111.354	94683	42823	1	4	6	4	4	28.00	4 Spring
9	56.10125	-4.56877	92240	42775	4	1	6	1	15	-16.00	5 winter
10	35.57502	-100.963	88318	42858	4	3	5	4	1	17.00	8 Spring
11	51.76378	6.826575	49690	42859	2	2	6	4	4	-1.00	2 Spring
12	58.67722	6.368225	47634	42812	2	1	4	4	4	17.00	9 Spring
13	51.21709	4.421192	47032	41718	3	3	2	4	3	-4.00	5 Spring
14	37.07028	-91.7361	46734	42750	4	2	1	1	13	-25.00	3 winter
15	59.9075	10.75216	59666	42859	2	4	3	4	3	0.00	0 Spring
16	42.41321	-8.76705	86236	42859	3	3	3	4	4	25.00	2 Spring
17	40.29688	-4.66627	99668	40096	4	3	2	3	1	6.00	7 autumn
18	43.08798	-2.54714	100025	42823	4	4	6	4	1	1.00	2 Spring
19	44.36205	4.782507	80172	42860	1	4	3	4	5	23.00	2 Spring
20	47.836	12.89598	96978	42834	2	4	3	4	5	27.00	7 Spring

شکل ۳: بخشی از داده‌های نگاشت شده مجموعه داده فلیکر

۳-۵- معیارهای ارزیابی

مفید مورد بررسی قرار می‌دهد. در این مقاله، برای مکان‌های پیشنهاد داده شده، دو حالت بازدید شده یا بازدید نشده در نظر گرفته شده است. معیار میانگین قدر مطلق خطا بر اساس رابطه (۲۱) محاسبه می‌شود که در آن، مقدار Accuracy بر اساس تعداد پیشنهادات مورد علاقه کاربر نسبت به کل پیشنهادات به دست آمده است [۴۶]:

$$MAE = 1 - Accuracy \quad (21)$$

۴-۵- پیکربندی سخت افزار/نرم افزار

مشخصات پیکربندی نرم افزار و سخت افزار برای پیاده‌سازی روش پیشنهادی در جدول ۴ ارائه شده است.

جدول ۴: مشخصات پیکربندی سخت افزار و نرم افزار

۵-۵- نتایج ارزیابی TRS-GP

برای ارزیابی TRS-GP در ابتدا دقت فرایند خوشه‌بندی برای روش‌های C-Means فازی و K-Means بر اساس تعداد مختلف خوشه‌ها بررسی می‌شود. سپس معیارهای الگوریتم رقابت استعماری بر اساس فاکتورهای DB و CS بررسی می‌شود. نتایج به دست آمده برای ارزیابی دقت با در نظر گرفتن تعداد خوشه‌های مختلف برای الگوریتم‌های خوشه‌بندی K-Means در جدول ۵ آمده است.

جدول ۵: مقادیر دقت با در نظر گرفتن تعداد خوشه‌های مختلف

برای الگوریتم K-Means

تعداد خوشه	دقت
۳	۰,۶۶
۴	۰,۶۳
۵	۰,۶۲

برای ارزیابی روش پیشنهادی از معیارهای دقت و متوسط مطلق دقت و میانگین قدر مطلق خطا استفاده می‌شود. معیار دقت بر اساس نسبت تعداد پیش‌بینی‌های صحیح به تعداد کل پیش‌بینی‌ها به دست می‌آید که طبق رابطه (۱۹) محاسبه می‌شود [۱].

$$precision = \frac{\text{number of items that are relevant}}{\text{number of all recommended items}} \quad (19)$$

معیار متوسط مطلق دقت نیز بر اساس رابطه ۲۰ محاسبه می‌شود [۱]:

$$MAP = \frac{1}{Nq} \sum_{i=1}^{Nq} AP_i \quad (20)$$

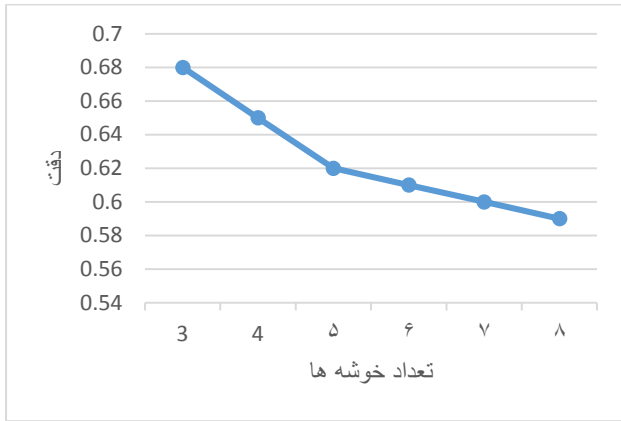
در این رابطه Nq تعداد پرس‌وجوی مورد آزمایش و AP_i متوسط دقت برای پرس‌وجوی i ام می‌باشد. AP_i بر اساس رابطه $\frac{1}{r}$ محاسبه می‌شود که r نشان‌دهنده رتبه آیتم مورد نظر در لیست مرتب شده آیتم‌ها است. در واقع معیار دقت و متوسط مطلق دقت نشان‌دهنده کیفیت پیشنهادهای ارائه شده هستند و هر چه مقدار آن‌ها بیشتر باشد، نشان‌دهنده توانایی بیشتر سیستم توصیه‌گر برای ارائه پیشنهادهای درست است.

یکی دیگر از روش‌های ارزیابی نتایج سیستم‌های پیشنهاددهنده، بررسی میزان خطای پیشنهادات ارائه شده به کاربران است. منابع مختلف، روابط متفاوتی را برای محاسبه میزان خطا ارائه داده‌اند که یکی از ساده‌ترین و پایه‌ای‌ترین رابطه‌ها که در [۴۶] مطرح شده است به جای در نظر گرفتن مقادیر عددی رتبه‌های پیش‌بینی‌شده و رتبه‌های واقعی، مقادیر آن‌ها را به صورت باینری در نظر گرفته و هر پیشنهاد را در دو حالت مفید و غیر

0 - Precision
1 - Mean Absolute Precision (MAP)

2 - Mean Absolute Error

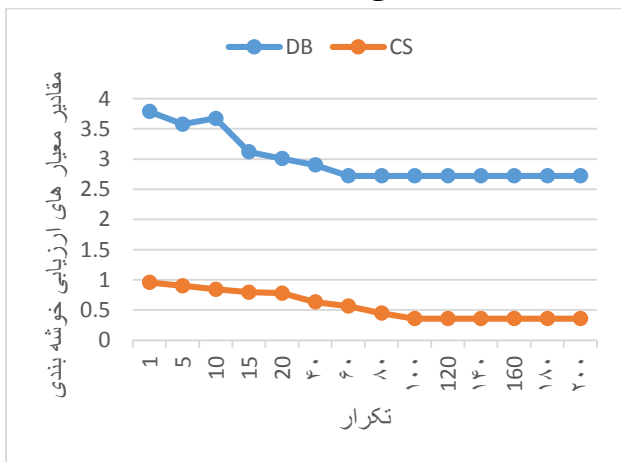
۰,۵۹	۸
------	---



شکل ۵: مقادیر معیار دقت الگوریتم خوشه بندی C-Means فازی را با در نظر گرفتن تعداد خوشه های مختلف

برای الگوریتم K-Means همانند الگوریتم C-Means فازی بالاترین میزان دقت با تعداد ۳ خوشه به دست می‌آید.

در مورد معیارهای موثر در الگوریتم رقابت استعماری نتایج ارزیابی نشان می‌دهند که معیار DB نسبت به معیار CS مناسب‌تر بوده و برای تعیین تعداد و مراکز خوشه‌ها بهتر عمل می‌کند. معیار CS نسبت به DB پیچیدگی محاسباتی نسبتاً بالاتری دارد و نتایج ارزیابی‌ها در [۴۷] نیز نشان می‌دهد که معیار DB نسبت به دیگر معیارهای ارزیابی خوشه بندی عملکرد بهینه‌تری را ارائه می‌دهد. شکل ۶ تغییرات این دو معیار را در تکرارهای مختلف نشان می‌دهد.



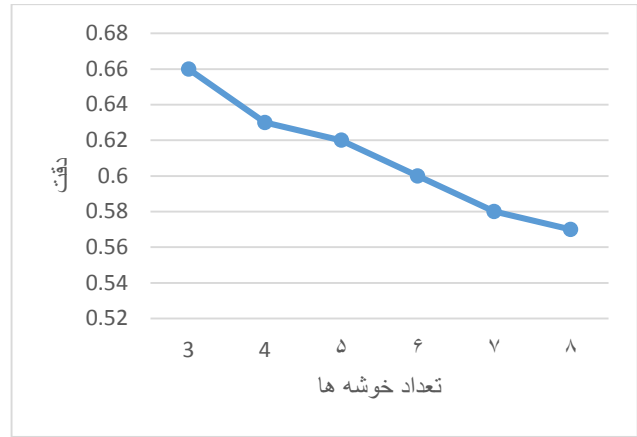
شکل ۶: مقایسه معیارهای ارزیابی خوشه‌بندی DB و CS

۵-۵- مقایسه TRS-GP با روش‌های موجود

در این بخش، TRS-GP را با برخی از روش‌های موجود در سیستم توصیه‌گر زمینه آگاه سفر مقایسه می‌کنیم تا جایگاه روش پیشنهادی در پژوهش‌های موجود در این حوزه مشخص شود.

۰,۶	۶
۰,۵۸	۷
۰,۵۷	۸

شکل ۴: مقادیر معیار دقت الگوریتم خوشه بندی K-Means را با در نظر گرفتن تعداد خوشه‌های مختلف نشان می‌دهد.



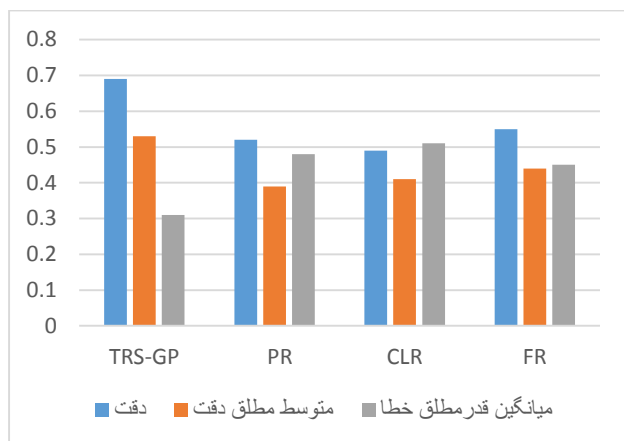
مشخصات	اجزا	نرم افزار	سخت افزار
MATLAB 9.4	زبان برنامه نویسی		
Windows 7	ویندوز		
Intel core(TM) i7-7700HQ CPU-2.80GHz, 4 Core(s), 8 Logical Processor(s)	CPU		
GetForce GTX 1050 Ti, NVIDIA compatible - 4GB RAM	GPU		
16 GB	RAM		

شکل ۴: مقادیر معیار دقت الگوریتم خوشه بندی K-Means را با در نظر گرفتن تعداد خوشه های مختلف

معیار دقت با در نظر گرفتن تعداد خوشه‌های مختلف برای الگوریتم خوشه‌بندی C-Means فازی در جدول ۶ و نمودار مقادیر آن در شکل ۵ نشان داده شده است.

جدول ۶: مقادیر دقت با در نظر گرفتن تعداد خوشه‌های مختلف برای الگوریتم FCM

تعداد خوشه	دقت
۳	۰,۶۸
۴	۰,۶۵
۵	۰,۶۲
۶	۰,۶۱
۷	۰,۶



شکل ۷: میزان دقت، متوسط مطلق دقت و میانگین قدرمطلق خطا

TRS-GP در مقایسه با روش‌های پایه‌ای

با هدف بررسی تاثیرگذاری روش پیشنهادی، نتایج آن نسبت به چهار روش‌های پایه که در این قسمت توضیح داده شده‌اند، مقایسه شده است.

- **Popularity Rank (PR):** در این روش مکان‌ها بر اساس میزان محبوبیت عمومی رتبه بندی می‌شوند. میزان محبوبیت هر مکان بر اساس تعداد بازدیدهای غیرتکراری از مکان موردنظر است [۴۸].
 - **Classic Rank (CLR):** این روش، همه کاربرها را یکسان در نظر نمی‌گیرد و بر اساس استنتاج مبتنی بر HITS^۳، جذابیت‌های توریستی مکان‌ها و تجربه‌های سفر کاربران را برحسب امتیاز اعتباری برای آن‌ها محاسبه می‌کند و همچنین امتیاز محوری را از داده‌های مربوط به ارتباطات کاربران و مکان‌ها استخراج می‌کند [۴۹].
 - **Frequent Rank (FR):** در این روش برای تعیین مکان‌های پرتکرار و الگوهای ترتیبی که فراوانی‌های آن‌ها کوچک‌تر از مقدار حد‌آستانه پشتیبان کمینه نباشد، از روش PrefixSpan استفاده می‌شود [۵۰].
- نتایج نشان می‌دهد که عملکرد TRS-GP نسبت به دیگر روش‌های پایه بهتر است. دلیل آن، استفاده از داده‌های زمینه‌ای مبتنی بر زمان، مکان، شرایط آب و هوا، دما، حالت روحی کاربر، میزان ترافیک و همچنین تحلیل و خوشه‌بندی دقیق داده‌ها

برای مقایسه روش پیشنهادی با سایر پژوهش‌های موجود در این زمینه از دو رویکرد زیر استفاده می‌شود تا نتایج به دست آمده در یک چهارچوب استاندارد صحت‌گذاری مناسبی گردد.

در ابتدا، TRS-GP با یکسری روش‌های پایه‌ای استاندارد که در بسیاری از مقالات برای سنجش دقت از آن‌ها استفاده می‌شود مقایسه می‌گردد، نظیر [48] Popularity Rank (PR)، [49] Classic Rank (CLR)، [50] Frequent Rank (FR). لازم به ذکر است که تمامی این روش‌های پایه‌ای با مجموعه داده ارائه شده در این مقاله پیاده‌سازی و ارزیابی شده‌اند. سپس، تعدادی از پژوهش‌های موجود در زمینه سیستم‌های توصیه‌گر سفر مبتنی بر عکس‌های برچسب‌دار جغرافیایی استخراج شده از فلیکر ([۲۴، ۳۲، ۳۳، ۳۴]) برای مقایسه با روش پیشنهادی ارزیابی می‌شود. معیارهای ارزیابی در تمام مقالات پارامترهای مرتبط با کیفیت توصیه سیستم نظیر دقت، متوسط مطلق دقت و میانگین قدرمطلق خطا است.

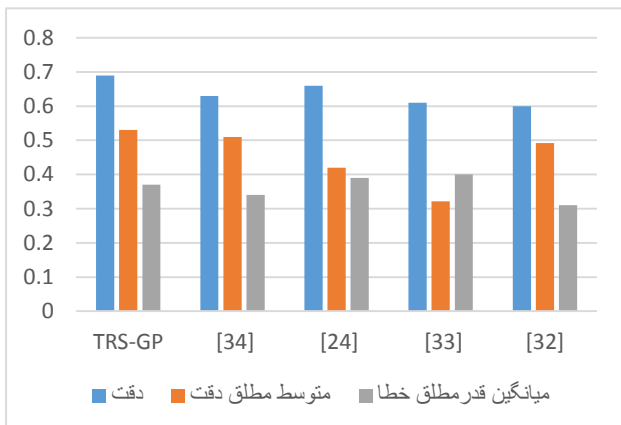
جدول ۷ مقادیر دقت، متوسط مطلق دقت و میانگین قدرمطلق خطا اندازه‌گیری شده TRS-GP را در مقایسه با روش‌های پایه‌ای دیگر نشان می‌دهد. برای ارزیابی تمام روش‌ها، از مجموعه داده استفاده شده در این پژوهش تعداد ۲۵۰۰ رکورد از دیتاست فلیکر برای ۱۰۰ کاربر استخراج شده است. شکل ۷ نیز این ارزیابی را نشان می‌دهد.

جدول ۷: میزان دقت، متوسط مطلق دقت و میانگین قدرمطلق خطا TRS-GP در مقایسه با روش‌های پایه‌ای

روش	دقت	متوسط مطلق دقت	میانگین قدرمطلق خطا
TRS-GP	۰,۶۹	۰,۵۳	۰,۳۱
PR	۰,۵۲	۰,۳۹	۰,۴۸
CLR	۰,۴۹	۰,۴۱	۰,۵۱
FR	۰,۵۵	۰,۴۴	۰,۴۵

² Hypertext Induced Topic Search (HITS)

ترافیک و ...) است. در حالی که در بسیاری از مقالات مشابه فقط از زمینه‌های زمانی و مکانی استفاده شده است.



شکل ۸: میزان دقت، متوسط مطلق دقت و میانگین قدرمطلق خطا TRS-GP در مقایسه با مقالات دیگر

۶- نتیجه گیری

این مقاله بر اساس تاریخچه سفرها و بازدیدهای کاربر از مکان‌های مختلف، بر اساس داده‌های زمینه‌ای، مکان‌های جدیدی را برای سفر به کاربر توصیه می‌کند. روش پیشنهادی، از تصاویر با برچسب‌های جغرافیایی در رسانه‌های اجتماعی، برای تولید توصیه‌های مکانی مطابق با علاقه‌مندی‌های کاربران و موقعیت‌های بازدید، استفاده می‌کند و در این راستا قادر به درک موقعیت، زمان، تاریخ، آب و هوا، دما و حالات روحی کاربر نیز می‌باشد. این روش، مبتنی بر ترکیب دو الگوریتم رقابت استعماری و خوشه‌بندی C-Means فازی است که از آن‌ها، برای شناسایی خوشه‌های مکان‌های توریستی استفاده می‌شود و سوابق سفرها از تصاویری با برچسب‌های جغرافیایی استخراج می‌شوند. روش پیشنهادی به بعضی از روش‌های توصیه مکان سفر با استفاده از مجموعه تصاویر فلیکر که در دسترس عموم قرار دارد ارزیابی و مقایسه شده است. نتایج ارزیابی بیانگر آن است که این روش قادر به فراهم کردن توصیه‌های مکانی مطابق با علاقه‌مندی‌های کاربر و موقعیت فعلی بازدید او می‌باشد و با

است که با بررسی تاریخچه مربوط به بازدیدهای گذشته کاربر، پیشنهادات دقیقی را به کاربر ارائه می‌دهد.

در جدول ۸ TRS-GP با مقالات مشابه بر اساس پارامترهای دقت، متوسط مطلق دقت و میانگین قدرمطلق خطا مقایسه شده است. همانطور که قبلاً نیز ذکر شد تمام روش‌ها از مجموعه داده‌های فلیکر که در مکان‌های مختلف جمع‌آوری شده‌اند استفاده کرده‌اند. شکل ۸ نیز این ارزیابی را نشان می‌دهد.

جدول ۸: میزان دقت، متوسط مطلق دقت و میانگین قدرمطلق خطا TRS-GP در مقایسه با مقالات دیگر

مقاله	دقت	متوسط مطلق دقت	میانگین قدرمطلق خطا
[۳۴]	۰,۶۳	۰,۵۱	۰,۳۷
[۲۴]	۰,۶۶	۰,۴۲	۰,۳۴
[۳۳]	۰,۶۱	۰,۳۲	۰,۳۹
[۳۲]	۰,۶۰	۰,۴۹	۰,۴۰
TRS-GP	۰,۶۹	۰,۵۳	۰,۳۱

همانطور که در جدول ۸ و شکل ۸ مشخص شده است میزان دقت و متوسط مطلق دقت TRS-GP نسبت به دیگر روش‌ها بالاتر و میانگین قدرمطلق خطای آن نسبت به روش‌های دیگر پایین‌تر است. استفاده از چندین زمینه مختلف و ترکیب دو الگوریتم خوشه‌بندی فازی و بهینه‌سازی رقابت استعماری قبل از ارائه توصیه نهایی در TRS-GP در مقایسه با [۲۴] که از تعداد زمینه‌های کمتری استفاده می‌کند و تنها مبتنی بر الگوریتم خوشه‌بندی K-Means است، دقت را افزایش و خطا را کاهش داده است. همچنین در مقایسه با روش مقالات [۳۴] و [۳۳] که بر اساس اولویت‌های سفر کاربران از سفرهای قبلی آن‌ها و همچنین رویکرد فیلترکننده مشارکتی پیشنهاد سفر می‌دهند TRS-GP از دقت بالاتری برخوردار است زیرا TRS-GP اطلاعات کامل‌تر و همه‌جانبه‌تری از یک بازدید را مورد پردازش قرار داده است و بر اساس آن‌ها خوشه‌بندی و ارائه پیشنهادات دقیق‌تری را انجام می‌دهد. در روش [۳۲] مکان‌ها بر اساس نوع جنسیت گردشگران دسته‌بندی می‌شوند و اطلاعات زمینه‌ای کمتری نسبت به TRS-GP به کار گرفته شده است. آنچه موجب تمایز اصلی میان TRS-GP در این مقاله و روش‌های ارائه شده در مقالات دیگر شده است افزودن تعداد زمینه‌های بیشتری (وضعیت آب و هوا، وضعیت روحی کاربر، میزان

۷- منابع

1. J. Bobadilla, F. Ortega, A. Hernando and A. Gutiérrez, "Recommender systems survey," *Know.-Based Syst.*, vol. 46, pp.109-132, 2013.
۲. بهشتی نژاد، راحله. سمیع، محمد ابراهیم. حمزه، علی. (۱۳۹۸). «بهبود سیستم‌های توصیه‌گر با کمک وب معنایی»، *نشریه فناوری اطلاعات و ارتباطات ایران* ۹، شماره ۳۱ (۱۳۹۸): ۴۵-۵۶.
3. M. Slehat, "Evaluation of potential tourism resources for developing different forms of tourism : case study of Iraq Al-Amir and its surrounding areas – Jordan," PhD thesis, Catholic University of Eichstätt-Ingolstadt. 2018.
۴. منتظر، غلامعلی. فتحی، وحید. (۱۳۹۴). «تخصیص بهینه درس پار به کمک الگوریتم بهینه‌سازی گروه ذرات»، *نشریه فناوری اطلاعات و ارتباطات ایران* ۶، شماره ۲۱ (۱۳۹۴): ۱۵-۲۶.
5. N. Henze, P. Dolog and N. Wolfgang, "Reasoning and Ontologies for Personalized Elearning in the Semantic Web," *Educational Technology & Society*, Vol. 7, 82-97.
۶. صابری، نفیسه. منتظر، غلامعلی. (۱۳۸۹). «شخصی‌سازی محیط یادگیری الکترونیکی به کمک توصیه‌گر فازی مبتنی بر تلفیق سبک یادگیری و سبک شناختی» *نشریه فناوری اطلاعات و ارتباطات ایران* ۲، شماره ۳ (1389): ۹۱-۱۰۹.
7. K. Choi, D. Yoo, G. Kim and Y. Suh, "A hybrid online-product recommendation system: Combining implicit Rating based collaborative filtering and sequential pattern analysis," *Electronic Commerce Research and Applications*, Vol.11, pp. 309-317, 2012.
8. G. Adomavicius and A. Tuzhilin, "Context-Aware Recommender Systems," In: *Recommender Systems Handbook*, F. Ricci, L. Rokach, B. Shapira, Ed. Springer, Boston, MA, Springer US, pp. 217-253. 2011.
9. K. Verbert, N. Manouselis, X. Ochoa, M. Wolpers, H. Drachsler, I. Bosnic and E. Duval, "Context-aware recommender systems for learning: a survey and future challenges," *IEEE Transactions on Learning Technologies*, Vol. 5, no. 4, pp. 318-335, 2012.

در نظر گرفتن شباهت‌های میان کاربران (مدل تشابه کاربر-کاربر) بر اساس سفرهایشان برای توصیه‌های سفر مناسب است.

بررسی نتایج ارزیابی‌های روش پیشنهادی نشان می‌دهند که بهره‌گیری از توان نفوذ اجتماعی و استفاده از داده‌های زمینه‌ای برای آیت‌های حساس به مکان موثر است. در واقع، روش پیشنهادی با ترکیب دو الگوریتم رقابت استعماری و خوشه‌بندی C-Means فازی موجب افزایش دقت پیشنهادها شده و استفاده از سرویس‌های شخصی‌سازی شده و همچنین داده‌های زمینه‌ای بیشتر (مکان، تعداد بازدید، وضعیت آب و هوا، موقعیت زمانی در روز، حالت روحی کاربر، میزان ترافیک، فصل و دما) باعث بهبود عملکرد سیستم توصیه‌گر و آرایه پیشنهادهای دقیق‌تر با خطای کمتر می‌شود.

برای بهبود کارایی سیستم توصیه‌گر مبتنی بر زمینه می‌توان پژوهش‌های زیر را برای کارهای آینده بیشتر مورد بررسی قرار داد:

- ✓ آرایه راهکارهایی برای حفظ محرمانگی و امنیت داده‌ها.
- ✓ توسعه واسط‌های کاربر با کیفیت بالا و پیچیدگی کم که تعامل بین کاربران نهایی و سیستم توصیه‌گر را ساده کنند.
- ✓ بررسی برخی محدودیت‌های زمانی- مکانی برای نتایج توصیه‌ای (مثلاً چه مدت زمان برای بازدید نیاز است و کاربر چقدر زمان دارد؟)
- ✓ شخصی‌سازی در ترکیب اطلاعات به دست آمده از موقعیت در سفرها (پیگیری مکان‌ها) برای توصیه‌های توریستی
- ✓ آرایه یک مجموعه داده منسجم و کاربردی
- ✓ استخراج موقعیت جغرافیایی (طول و عرض جغرافیایی) و همچنین امکان افزودن برچسب به صورت دستی با استفاده از سیستم‌های موبایل
- ✓ وزن‌دهی به اطلاعات زمینه‌ای مختلف

20. A. Stetco, X. Zeng and J. Keane, "Fuzzy C-means++: Fuzzy C-means with effective seeding initialization," *Expert Systems with Applications*, Vol. 42, no. 21, pp. 7541–7548, 2015.
21. A. Gargari and E. Lucas, "Imperialist competitive algorithm: an algorithm for optimization inspired by imperialistic competition," *IEEE Congress on Evolutionary Computation*, pp. 4661-4667, 2007.
22. N.M. Villegas, C. Sanchez, J. Dóaz-Cely and G. Tamura, "Characterizing Context - Aware Recommender Systems: A Systematic Literature Review," *Knowledge-Based Systems*. Vol. 140, pp. 173-200, 2017.
23. G. Chen and L. Chen, "Augmenting service recommender systems by incorporating contextual opinions from user reviews," *User Modeling and User-Adapted Interaction*, Vol. 25, no. 3, pp. 295–329, 2015.
24. Z. Xu, C. Ling and C. Gencai, "Topic based context-aware travel recommendation method exploiting geotagged photos," *Neurocomputing*, Vol. 155, pp. 99-107, 2015.
25. Shafaqat, W. and Byun, Y.C., 2020. A Recommendation Mechanism For Under-Emphasized Tourist Spots Using Topic Modeling And Sentiment Analysis. *Sustainability*, 12(1), p.320.
26. Alrasheed, H., Alzeer, A., Alhowimel, A. and Althyabi, A., 2020. A Multi-Level Tourism Destination Recommender System. *Procedia Computer Science*, 170, pp.333-340.
27. H. Khallouki, A. Abatal and M. Bahaj, "An ontology-based context awareness for smart tourism recommendation system," In: *Proceedings of the Inter-national Conference on Learning and Optimization Algorithms: Theory and Applications*, LOPAL 2018, Rabat, Morocco, May 2-5, 2018, pp43:1–43:5.
28. E. Pantano, C-V. Priporas, N. Stylos and C. Dennis, "Facilitating tourists' decision making through open data analyses: A novel recommender system," *Tourism Management Perspectives*, Vol. 31, pp. 323-331, 2019.
29. B. Kaya, "A hotel recommendation system based on customer location: a link prediction
10. F. Ricci, "Mobile recommender systems," *Information Technology & Tourism journal*, vol. 12, no. 3, pp.205-231., 2010.
11. R. Beale and P. Lonsdale, "Mobile context aware systems: The intelligence to support tasks and effectively utilise resources," In: *Mobile Human-Computer Interaction-Mobile HCI 2004*. S. Brewster, M. Dunlop, Ed. Springer Berlin Heidelberg. pp. 240-251. 2004.
12. D. Weib, M. Duchon, F. Fuchs and C. Linnhoff-Popien, "Context-aware personalization for mobile multimedia services," *Proceedings of the 6th International Conference on Advances in Mobile Computing and Multimedia*. ACM, 2008.
13. N. Manouselis, H. Drachsler, R. Vuorikari, H. Hummel and R. Koper, "Recommender systems in technology enhanced learning," *Recommender systems handbook*. Springer US. pp. 387-415. 2011.
14. A. K. Dey, A. Gregory and D. Salber, "A conceptual framework and a toolkit for supporting the rapid prototyping of context-aware applications," *International Journal of Human-Computer Studies*, Vol. 16, no. 2, pp. 97-166. 2001.
15. P. Brusilovsky and E. Millán, "User models for adaptive hypermedia and adaptive educational systems," *The adaptive web*. Springer-Verlag, 2007.
16. K. Xu, Clustering. In: Dubitzky W., Wolkenhauer O., Cho KH., Yokota H. (eds) *Encyclopedia of Systems Biology*. 2013, Springer, New York, NY.
17. S. Chen, Z. Xu and Y. Tang, "A Hybrid Clustering Algorithm Based on Fuzzy c-Means and Improved Particle Swarm Optimization," *Arabian Journal for Science and Engineering*, Vol. 39, pp.8875–8887, 2014.
18. KP. Soman, S. Diwakar and V. Ajay, "Data mining: theory and practice," PHI Learning Pvt. Ltd.; 2006.
19. J. Nayak, B. Naik and HS. Behera, "Fuzzy C-means (FCM) clustering algorithm: a decade review from 2000 to 2014," In *Computational intelligence in data mining-Vol. 2*, pp. 133-149. 2015, Springer, New Delhi.

- Geotagged Photos On Instagram,” 2020 6th International Conference on Web Research (ICWR), Tehran, Iran, 2020, pp. 203-208.
39. L.W. Dietz, A. Sen, R. Roy and W. Worndl, “Mining trips from location-based social networks for clustering travelers and destinations,” *Information Technology & Tourism*, Vol. 22, pp. 131–166, 2020.
40. D. Lyu, L. Chen, Z. Xu and S. Yu, “Weighted multi-information constrained matrix factorization for personalized travel location recommendation based on geo-tagged photos,” *Applied Intelligence*, Vol. 50, pp. 924–938, 2020.
41. A. Chaghari and M. Feizi-Derakhshi, “Automatic Clustering Using Improved Imperialist Competitive Algorithm,” *JSDP*. 2017; 14 (2) :159-169.
۴۲. امیری، مریم. ختن لو، حسن. (۱۳۹۲). «خوشه بندی اسناد، مبتنی بر آنتولوژی و رویکرد فازی» نشریه فناوری اطلاعات و ارتباطات ایران ۵، شماره ۱۷ (۱۳۹۲): ۹۶-۷۳.
43. D. L. Davies and D. W. Bouldin, “A cluster separation measure,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 1, no. 2, pp. 224–227, 1979
44. H. Chou, M. C. Su, and E. Lai, “A new cluster validity measure and its application to image compression,” *Pattern Analysis and Applications* vol. 7, no. 2, pp. 205–220, Jul. 2004.
45. “Flickr 10K Dataset” <https://www.kaggle.com>, [accessed: Sep 2020]
46. Del Olmo, F.H. and Gaudioso, E., 2008. Evaluation of recommender systems: A new approach. *Expert Systems with Applications*, 35(3), pp.790-804.
- approach,” *Multimedia Tools and Applications*, Vol. 79, pp. 1745–1758, 2020.
30. L. Esmaeili , S. Mardani , A. Hashemi Golpayegani and Z. Zanganeh Madar. “A Novel Tourism Recommender System in the Context of Social Commerce,” *Expert Systems With Applications*, Vol.149, 113301, July 2020.
31. BM. Veloso, F. Leal, B. Malheiro and JC. Burguillo, “On-line guest profiling and hotel recommendation,” *Electronic Commerce Research and Applications*, Vol. 34, 100832, 2019.
32. Z. Xu, L. Chen, H. Guo, M. Lv and G. Chen, “User similarity-based gender-aware travel location recommendation by mining geotagged photos,” *International Journal of Embedded Systems*, Vol. 10, no. 5, 356, 2018.
33. A. Majid, L. Chen, G. Chen, H. Mirza, I. Hussain and J. Woodward, “A context-aware personalized travel recommendation system based on Geotagged social media data mining,” *International Journal of Geographical Information Science*, Vol. 27, no. 4, pp. 662-684, 2013.
34. I. Memon, L. Chen, A. Majid, M. Lv, I. Hussain and G. Chen, “Travel recommendation using geo-tagged photos in social media for tourist,” *Wireless Personal Communications*. vol. 80, no. 4, pp. 1347-1362, 2015.
35. Z. Xu, L. Chen, Y. Dai and G. Chen. “A Dynamic Topic Model and Matrix Factorization based Travel Recommendation Method Exploiting Ubiquitous Data,” *IEEE Transactions on Multimedia*. Vol. 19, no. 8, pp. 1933-1945, 2017.
36. S. Ojagh, M. Malek, S. Saeedi, and S. Liag, “A location-based orientation-aware recommender system using IoT smart devices and Social Networks,” *Future Generation Computer Systems*, Vol. 108, pp. 97-118, 2020.
37. G. Zhao, P. Lou, X. Qian and X. Hou, “Personalized location recommendation by fusing sentimental and spatial context,” *Knowledge-Based Systems*, Vol. 196, 2020.
38. M. Memarzadeh and A. Kamandi, “Model-Based Location Recommender System Using

48. A. Majid, L. Chen, G. Chen, H. T. Mirza and I. Hussain, "GoThere: Travel suggestions using geotagged photos," In WWW 2012 companion, April 16–20, 2012. Lyon, France: ACM.
49. Y. Zheng, L. Zhang, X. Xie and W.Y. Ma, "Mining interesting locations and travel sequences from GPS trajectories," in: Proceedings of the 18th International Conference on World Wide Web, WWW '09, ACM, New York, NY, USA, 2009, pp. 791–800.
50. Z. Yin, L. Cao, J. Han, J. Luo and T.S. Huang, "Diversified trajectory pattern ranking in geo-tagged social media," in: Proceedings of SIAM International Conference on Data Mining, 2011, pp. 980–991.
47. E. Rendón, I. M. Abundez, C. Gutierrez, S. D. Zagal, A. Arizmendi, E. M. Quiroz and H. E. Arzate, "A comparison of internal and external cluster validation indexes," In Proceedings of the 2011 American conference on applied mathematics and the 5th WSEAS international conference on Computer engineering and applications (AMERICAN-MATH'11/CEA'11). World Scientific and Engineering Academy and Society (WSEAS), Stevens Point, Wisconsin, USA, 158–163, 2011.

نمودار قطبی نقاط با قطب متحرک

*بهرام صادقی بی غم **فاطمه ربانی

*دانشیار، دانشکده علوم کامپیوتر و فناوری اطلاعات، دانشگاه تحصیلات تکمیلی علوم پایه زنجان
**کارشناسی ارشد، دانشکده علوم کامپیوتر و فناوری اطلاعات، دانشگاه تحصیلات تکمیلی علوم پایه زنجان
تاریخ دریافت: ۱۳۹۸/۱۲/۱۰ تاریخ پذیرش: ۱۳۹۹/۰۶/۲۲

چکیده

مسئله نمودار قطبی یکی از تعمیم‌های نمودار ورونوی است که در آن به جای متر اقلیدسی از مقدار زاویه برای محاسبه فاصله استفاده می‌شود. این مسئله کاربردهای زیادی در پردازش تصویر، مخابرات و مباحث مربوط به آنتن، رؤیت‌پذیری و مسیریابی ربات دارد. در سالهای اخیر دو نوع نمودار قطبی مطرح شده و برای انواع سایت‌ها الگوریتم‌های مناسبی ارائه شده است. همچنین روی همین مسائل با داده‌های جنبشی و حالات پویا الگوریتم‌هایی ارائه شده است. در این مقاله قطب به عنوان ناظر متحرک در نظر گرفته شده و الگوریتمی ارائه می‌شود که مسئله بازسازی نمودار قطبی با قطب نزدیک را به صورت کارا و در زمان خطی حل می‌کند. در این حالت زمان پیش‌پردازش الگوریتم $O(n^4 \log_2 n)$ و زمان باز رسم نمودار در هر حرکت متوالی قطب برابر با $O(\log n + k)$ است که در آن k تعداد سایت‌های درون ناحیه T است که احتمال تغییر در آنها وجود دارد.

کلمات کلیدی: نمودار قطبی، نمودار ورونوی، مخابرات، آنتن، زاویه قطبی، رویت‌پذیری

۱- مقدمه

پردازش برای بازسازی نمودار است. هنگامی که قطب به عنوان یک ناظر در نظر گرفته شود، این مسئله کاربردهای زیادی در پردازش تصاویر خواهد داشت. افزایش صفحه بعنوان پیش‌پردازش، همواره باعث افزایش سرعت در اجرای برنامه‌های آنلاین می‌شود که سرعت در آنها از اهمیت خاصی برخوردار است. نمونه‌ای از این کاربردها را میتوان در بازی‌های کامپیوتری دید. هنگامی که شی متحرکی (در اینجا منبع نور بعنوان روشنایی) لابلای موانع ثابت در حرکت است صحنه‌ای که در اثر سایه‌ها تولید می‌شود در هر لحظه متفاوت است که نیاز به بازسازی دارد. بدون این پیش‌پردازش و با استفاده از الگوریتم‌های قبلی مانند الگوریتم نقاش، محاسبه‌ی ناحیه‌ها زمان زیادی صرف می‌کند. با استفاده از الگوریتم ارایه شده و پس از یک بار پیش‌پردازش میتوان تاثیرات نور و سایه در محیط را بسیار سریع تر محاسبه کرد. در این نوع کاربرد فرض کنید قطب نوری به

یکی از مسائل مهم و پر کاربرد در هندسه محاسباتی مسئله نمودار ورونوی است که از نظر نوع سایت‌ها، متر، محدودیت‌های صفحه و موانع دارای انواع تعمیم‌هاست. یکی از این تعمیم‌ها حالتی است که معیار فاصله بین نقاط و سایت‌ها زاویه است. این نوع نمودار کاربردهای مهمی در بینایی ماشین و مخابرات دارد. این مسئله در دو نوع کلی نمودار قطبی با قطب دور (کلاسیک) و قطب نزدیک معرفی شده است که در هر مورد به تعمیم‌هایی از آن نیز پرداخته شده است. همچنین به حالاتی از مسئله که سایت‌ها در حالت حذف یا اضافه شدن باشند نیز توجه شده است. این مقاله حالتی از مسئله نمودار قطبی با قطب نزدیک را بررسی می‌کند که سایت‌ها به صورت نقاطی ثابت در صفحه باشند و قطب بتواند در خارج از پوسته محدب سایت‌ها جابجا شود. هدف اصلی، ارائه روشی سریع و کارا پس از یک پیش

در نمودار قطبی با قطب نزدیک، قطب در نزدیک سایت‌ها قرار دارد. فرض می‌کنیم که قطب در سمت چپ همه سایت‌ها قرار داشته باشد. در واقع ورودی این مسئله مجموعه‌ای از سایت‌ها هستند که این مجموعه را با حرف S نشان می‌دهیم و یک نقطه P به عنوان قطب. خروجی مسئله یک افراز از صفحه می‌باشد که با استفاده از آن برای هر نقطه دلخواه X در صفحه می‌توان گفت که X متعلق به ناحیه قطبی کدام سایت S_i است. شکل ۱(b) نمونه‌ای از نمودار قطبی با قطب نزدیک می‌باشد. نمودار قطبی اولین بار در سال ۱۹۹۸ در چهاردهمین کارگاه اروپایی هندسه محاسباتی در اسپانیا توسط گریم^۱ و همکارانش معرفی شد [۱]. آن‌ها ضمن معرفی نمودار قطبی به ارائه الگوریتم رسم آن نیز پرداختند. همچنین نمودار قطبی را در مورد پاره‌خط‌ها، چندضلعی‌ها و دوایر گسترش دادند و به کاربردهای این نمودار در مسائل پوسته محدب^۲ و رؤیت‌پذیری^۳ اشاره کردند [۲].

در سال ۲۰۰۳، ارتگا^۴ و همکارانش در مورد کاربرد نمودار قطبی در مسائل برنامه‌ریزی حرکت^۵ و مسائل رؤیت‌پذیری بحث نمودند [۳] و الگوریتم‌هایی را برای حل این مسائل ارائه کردند که از یک پیش‌پردازش زاویه‌ای بر پایه نمودار قطبی استفاده می‌کرد.

دوگان نمودار قطبی توسط صادقی و همکارانش معرفی و الگوریتمی برای محاسبه آن ارائه شد [۴]. این الگوریتم دوگان را در زمان بهینه $\theta(n \log n)$ محاسبه می‌کند. در سال ۲۰۱۰ ارتگا و همکارانش این مسئله را به طور گسترده مورد بررسی قرار دادند و تحت عنوان یک راه حل برای مسئله مسیریابی با استفاده از نمودار قطبی به عنوان پیش‌پردازش، مقاله‌ای به چاپ رساندند [۵]. آنها در این مقاله الگوریتمی برای مسئله مسیریابی ارائه کردند که با استفاده از نمودار قطبی صفحه را به چندین ناحیه تقسیم بندی می‌کند. همچنین در سال ۲۰۱۳ ارتگا و همکارانش به کاربرد نمودار قطبی در مسائل رؤیت‌پذیری پرداختند [۶]. در الگوریتم بیان شده ابتدا چهار نمودار قطبی به عنوان پیش‌پردازش محاسبه و سپس در گام‌های الگوریتم از این چهار نمودار با

رنگ خاص سانس می‌کند و هر سایت نور و رنگ مخصوص به خود را دارند.

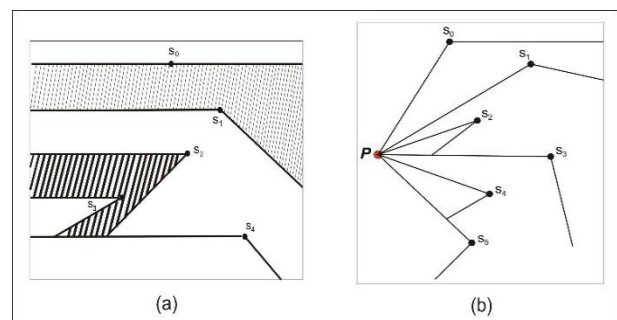
همچنین فرض کنید هر نقطه ی دلخواه از صفحه دارای رنگی است که از ترکیب رنگ سایت والد و قطب حاصل می‌شود. محاسبه ی رنگ هر پیکسل از صفحه با الگوریتم های رایج قبلی زمانی طولانی لازم دارد و با روش جدید و البته پس از یک بار پیش پردازش سرعت انجام اینکار زیر خطی است.

نمودار قطبی یکی از نمودارهای افراز صفحه مانند نمودار ورونوی است با این تفاوت که معیار اصلی در نمودار قطبی زاویه قطبی می باشد. زاویه قطبی سایت S_i نسبت به قطب P برابر است با زاویه ای که توسط نیم خط افقی رسم شده از P به سمت راست (جهت مثبت محور X ها) و پاره خط $S_i P$ تولید می شود. فرض می شود قطب در سمت چپ تمام سایتها قرار دارد و بنابراین مقدار زاویه قطبی در بازه ی $[-90, 90]$ قرار دارد.

ناحیه قطبی یک سایت دلخواه S_i که با $P_S(S_i)$ نشان می دهیم، برابر با مکان هندسی تمام نقاطی از صفحه است که زاویه قطبی یا آنها تا سایت S_i کوچکتر از زاویه قطبی آن نقاط تا هر سایت دیگری باشد. زاویه قطبی نقطه ی (x, y) نسبت به سایت S_i یعنی $\text{ang}(S_i)(x, y)$ زاویه ای است که بین نیمخط افقی با مبدا سایت S_i و به سمت منفی محور X و خط واصل بین سایت S_i و نقطه ی (x, y) است. ناحیه قطبی را می‌توان به صورت فرمول ۱ نشان داد:

$$P_S(S_i) = \{(x, y) \in E^2 \mid \text{ang}(S_i)(x, y) < \text{ang}(S_j)(x, y), \forall j \neq i\} \quad (1)$$

تاکنون دو نوع نمودار قطبی تعریف شده است. نمودار قطبی با قطب در بینهایت و نمودار قطبی باقطب نزدیک. در نمودار قطبی با قطب در بینهایت، قطب در دور از سایت‌ها و در بینهایت قرار دارد. شکل ۱ نمونه‌ای از نمودار قطبی با قطب در بینهایت را نشان می‌دهد.



شکل ۱ (a) نمودار قطبی با قطب دور، (b) نمودار قطبی با قطب نزدیک

Grima^۱
Convex hull^۲
Visibility^۳
Ortega^۴
Motion planning^۵

پژوهش با استفاده از ساختمان داده Y DCEL، مسئله با زمان پیش‌پردازش $O(n^4 \log_2 n)$ و نیز با بروز کردن نمودار در زمان $O(\log n + k)$ حل شده است. در ادامه به معرفی نمودار قطبی با قطب نزدیک پرداخته می‌شود و سپس حالتی را مورد بررسی قرار خواهیم داد که قطب میتواند به طور پیوسته و البته خارج از پوسته محدب سایت‌ها در صفحه حرکت کند. در بخش آخر نیز تحلیل نهایی از الگوریتم ارائه شده برای به روز رسانی نمودار قطبی با قطب نزدیک متحرک، صورت می‌گیرد.

۲- نمودار قطبی با قطب نزدیک متحرک برای نقاط

فرض کنید مجموعه S شامل n نقطه به عنوان سایت در صفحه و قطب P در نزدیک سایت‌ها داده شده است. همچنین فرض کنید که نمودار قطبی مجموعه S و قطب P با نام $NPPD(S)$ با استفاده از الگوریتم افزایشی بیان شده در [۸] محاسبه شده و سپس قطب P شروع به حرکت پیوسته در صفحه می‌کند. با حرکت قطب در صفحه، نمودار قطبی دچار تغییر خواهد شد. باید الگوریتمی ارائه شود که $NPPD(S)$ را در کمترین زمان ممکن بروزرسانی نماید.

ابتدا باید بررسی شود که با حرکت قطب چه رویدادهایی ممکن است برای نمودار قطبی رخ دهد و $NPPD(S)$ تحت چه شرایطی دست خوش تغییرات می‌گردد.

اگر به ازای هر سایت $S_i \in S$ که پیش از این در ناحیه قطبی سایت $S_k \in S$ بوده است (یا به عبارت دیگر $S_i \in P_s(S_k)$) همچنان سایت S_i در ناحیه قطبی سایت S_k بماند و عضو ناحیه قطبی سایت دیگری نشود، می‌گوییم توپولوژی نمودار $NPPD$ دچار تغییر نشده است. بنابراین هر گاه والد یکی از سایت‌ها تغییر کند، توپولوژی نمودار تغییر خواهد کرد و لازم است نمودار در زمان سریع و قابل قبول ترمیم شود که در واقع هدف اصلی مقاله می‌باشد.

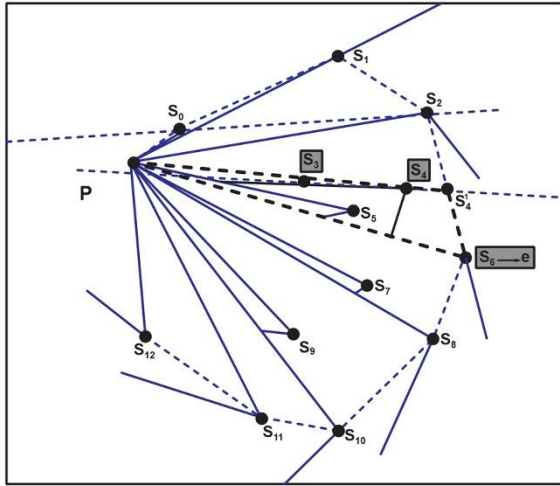
می‌توان گفت که در الگوریتم رسم نمودار $NPPD$ هر تغییر در لیست مرتب شده ی سایت‌ها بر اساس زاویه قطبی، باعث تغییر در توپولوژی نمودار $NPPD$ خواهد شد.

فرض کنید قطب به طور پیوسته و البته خارج از پوسته محدب سایت‌ها در حال حرکت در صفحه است، با توجه به نکته قبلی، کافی است که در هر لحظه لیست مرتب شده سایت‌ها بر اساس زاویه قطبی بررسی شود، اگر این لیست

توجه به جهت مشخص، استفاده شده است. در سال ۲۰۰۷ نوع دیگری از نمودار قطبی توسط صادقی و همکارانش معرفی شد که در آن قطب در نزدیک سایت‌ها قرار داشت [۷]. به این نوع از نمودار قطبی، نمودار قطبی با قطب نزدیک گفته می‌شود. این گروه علاوه بر معرفی نمودار قطبی با قطب نزدیک، در مورد دوگان این نوع نمودار قطبی نیز بحث نمودند. همچنین نمودار قطبی با قطب نزدیک را برای پاره‌خط‌ها و چندضلعی‌ها و دوائر گسترش دادند و به کاربردهای این نمودار در مسائل رؤیت پذیری و مسیریابی اشاره کردند [۸]. بعد از آن و در سال ۲۰۰۸، صادقی و همکارانش نمودار قطبی پویا را معرفی نمودند [۹]. در نمودار قطبی پویا دو موضوع مورد بررسی قرار گرفته است. یکی حالتی است که در نمودار قطبی یک سایت اضافه و یا حذف شود. نمودار قطبی موجود با توجه به این که کدام اتفاق رخ داده است، تغییر می‌کند و به روز می‌گردد. این تغییر از نمودار به گونه‌ای است که بخش زیادی از نمودار دست نخورده باقی می‌ماند و این موجب افزایش سرعت به روز رسانی نمودار می‌شود. این الگوریتم تغییرات را در زمان قابل قبولی در نمودار قطبی اعمال می‌کرد. سپس در سال ۲۰۱۰ احسانفر و همکارانش الگوریتمی ارائه کردند که اضافه یا حذف شدن سایت‌ها در نمودار قطبی پویا به صورت بهینه انجام گردد [۱۰]. نوع دیگری از نمودار قطبی پویا با عنوان نمودار قطبی اشیاء متحرک نیز توسط نوری بیگی و همکارش در سال ۲۰۰۸ مورد بررسی قرار گرفت [۱۱]. در این نوع از نمودار قطبی پویا سایت‌ها در صفحه به صورت پیوسته حرکت می‌کنند و هدف بروز رسانی بهینه نمودار قطبی مجموعه سایت‌ها می‌باشد. آنها از ساختمان داده جنبشی^۶ برای حل این مسئله استفاده کرده‌اند. همچنین الگوریتم بیان شده نمودار قطبی را در زمان بهینه بروز رسانی می‌نماید. در سال ۲۰۰۹ بزاز در پایان‌نامه کارشناسی ارشد خود در مورد نوع دیگری از نمودار قطبی بحث کرد که به کاربردهایی از آن اشاره دارد [۱۲]. نتایج بدست آمده در پایان‌نامه ی بزاز و سپس در پایان نامه ی فاطمه ربانی [۱۷] در خصوص نقاط در این پژوهش بررسی و به کار گرفته شده است در این نوع نمودار قطبی، قطب میتواند به طور پیوسته در صفحه و خارج از پوسته محدب سایت‌ها حرکت کند و نمودار قطبی با توجه به این حرکت به روز شود. در این

⁷ Doubly-Connected Edges List

⁶ Kinetic Data Structure



شکل ۳: عبور قطب از امتداد خط S_2S_3

از حافظه $O(n)$ استفاده می‌کند و هر پرس و جو از مکان‌یابی یک نقطه را در زمان $O(\log n)$ پاسخ می‌دهد [۱۴]. در روشی که برای افزایش صفحه در مسئله نمودار قطبی با قطب نزدیک متحرک بیان شد، تعداد خطوط از مرتبه $O(n^2)$ و تعداد یال‌های به وجود آمده $O(n^4)$ می‌باشد. بنابراین بررسی این که قطب در کدام ناحیه قرار گرفته است، در زمان $O(\log n)$ امکان پذیر می‌باشد [۱۵]. حال باید با استفاده از این ساختار داده و با توجه به تغییراتی که با حرکت قطب در نمودار NPPD رخ می‌دهد، زمان لازم برای بروز رسانی نمودار را تحلیل کنیم. در ادامه فرض کنید خط گذرنده از دو سایت S_i و S_j را با $L(s_i, s_j)$ نشان می‌دهیم و هر خط صفحه را به دو نیم‌صفحه تقسیم می‌کند. واضح است که هر گاه قطب از خط $L(s_i, s_j)$ عبور کند، ترتیب دو سایت S_i و S_j در لیست مرتب شده بر اساس زاویه قطبی، جابجا می‌شود و هر گاه قطب در درون یکی از ناحیه‌ها حرکت کند، توپولوژی نمودار NPPD تغییر نخواهد کرد و تنها زمانی نمودار NPPD دچار تغییر می‌شود که قطب از خط گذرنده بین دو سایت S_i و S_j عبور کند. در واقع با حرکت قطب در صفحه و عبور آن از یکی از خطوط مجاور رسم شده بین هر دو سایت دلخواه، تنها دو سایت متوالی در لیست مرتب شده‌ی سایت‌ها بر اساس زاویه قطبی، می‌توانند با یکدیگر جابجا شوند. ابتدا رویه محدب مجموعه سایت‌ها و قطب محاسبه می‌شود. این رویه محدب با CH_{S+p} نشان داده می‌شود و با حرکت قطب به روز رسانی می‌گردد. رسم CH_{S+p} برای اولین بار از مرتبه زمانی $\theta(n \log n)$ می‌باشد ولی بروز کردن آن مطابق با حرکت قطب، در زمان ثابت امکان پذیر خواهد بود. در واقع حرکت پیوسته قطب و

مرتب شده دچار تغییر شده باشد، به این معنا است که توپولوژی نمودار قطبی تغییر کرده است.

بررسی لیست مرتب شده در هر لحظه کاری پرهزینه است. لذا روشی دیگر پیشنهاد می‌شود که در آن نیازی نیست که در هر لحظه لیست مرتب شده بررسی گردد. در این روش ابتدا با یک پیش پردازش، صفحه به چندین ناحیه تقسیم‌بندی می‌شود. ویژگی هر ناحیه در این تقسیم‌بندی این است که با حرکت قطب در هر یک از این ناحیه‌ها، ترتیب سایتها در لیست مرتب شده تغییر نمی‌کند و لذا توپولوژی NPPD بدون تغییر باقی می‌ماند. ویژگی دیگر این نواحی این است که با عبور قطب از مرز هر یک از آنها، دقیقاً جای دو سایت در لیست مرتب شده جابجا خواهند شد. در ادامه به تشریح نحوه‌ی افزایش صفحه به این نواحی پرداخته می‌شود.

برای تقسیم‌بندی صفحه، خط گذرنده از هر دو سایت، در صفحه رسم می‌شود. واضح است که با این روش $O(n^2)$ خط در صفحه رسم خواهد شد. اثبات شده است که در یک چیدمان از خطوط که شامل n خط می‌باشد، $O(n^2)$ رأس، یال و وجه وجود دارد [۱۳]. بنابراین با وجود $O(n^2)$ خط در صفحه، $O(n^4)$ رأس، یال و وجه در صفحه وجود خواهد داشت که این همان پیچیدگی فضایی الگوریتم نیز است. در کلیه‌ی مراحل پیش پردازش و پرس و جو این مقدار فضا برای تمام ذخیره سازی‌ها کافی است. هر یک از ناحیه‌های بدست آمده، یک ناحیه محدب می‌باشد و این ناحیه‌ها یا کراندار هستند و یا یک ناحیه بدون مرز. برای ذخیره این افزایش از صفحه می‌توان از ساختار داده‌ای DCEL استفاده نمود. این ساختمان داده برای نگهداری تقسیم‌بندی‌هایی از صفحه به کار می‌رود که تمامی نواحی کراندار و محدود شده باشند. پس از معرفی ساختمان داده مناسب برای ذخیره سازی افزایش صفحه باید دید با حرکت قطب چه رویدادهایی برای نمودار قطبی رخ خواهد داد. قبل از این که قطب شروع به حرکت کند، لازم است که موقعیت قطب مشخص گردد. بدین معنی که ابتدا باید دید که قطب در کدام یک از وجه‌ها قرار دارد. بررسی این که قطب در کدام یک از این وجه‌ها قرار دارد، یکی از مسائل مکان‌یابی^۸ می‌باشد. میدانیم که اگر S یک افزایش از صفحه شامل n یال باشد، در زمان مورد انتظار $O(n \log n)$ می‌توان ساختمان داده‌ای ساخت که

^۸ Point location

همان‌طور که در تصویر نیز مشخص شده است، وقتی قطب از خط $L_{(S_2, S_3)}$ عبور کرد، سایت S_3 از رویه محدب حذف می‌گردد. حال اگر فرض کنیم که ابتدا قطب در مکان P قرار داشته باشد، با حرکت قطب از سمت چپ خط $L_{(S_2, S_3)}$ به سمت راست آن، سایت S_3 به رویه محدب اضافه می‌گردد.

در نمودار قطبی رسم شده، هر سایت که روی CH_{S+P} قرار دارد، در ناحیه قطبی سایتی قرار دارد که در CH_{S+P} در ترتیب ساعتگرد، قبل از آن قرار دارد. به عنوان مثال می‌توان در شکل ۱- (b) گفت که سایت‌های S_0, S_1, S_3, S_5 به ترتیب در ناحیه قطبی سایت قبلی خود قرار دارند.

همچنین در همان شکل ۱- (b) مشخص است که اگر هر سایت v (به طور مثال سایت S_3) روی CH_{S+P} قرار داشته باشد، این سایت مانع دید سایت‌های قبل از خود ش (S_0, S_1, S_2) نسبت به سایت‌های بعدی (S_4, S_5) است.

اکنون با توجه به مطالب بیان شده، ناحیه‌ای را تعیین می‌کنیم که با حرکت قطب، تنها نمودار قطبی سایت‌های درون این ناحیه می‌توانند دچار تغییر شوند:

فرض کنید قطب ابتدا در مکان P قرار داشته باشد. حال در صدد آن هستیم که با عبور قطب از خط $L_{(S_i, S_j)}$ به طوری که S_j در لیست مرتب شده قبل از S_i قرار بگیرد (برای مثال در شکل ۳ عبور قطب از خط $L_{(S_3, S_4)}$ ، ناحیه‌ای از صفحه مانند T را مشخص کنیم که فقط نمودار قطبی داخل آن امکان تغییر دارد. T ناحیه مثلثی است که یکی از رئوس آن در مکان قطب (P) قرار دارد و دو رأس دیگر آن به صورت زیر تعیین می‌شوند:

T ناحیه‌ای در CH_{S+P} است که بین دو خط $L_{(P, S_j)}$ و $L_{(P, e)}$ قرار گرفته است که در آن سایتی است که در دو شرط زیر به طور هم‌زمان صدق کند:

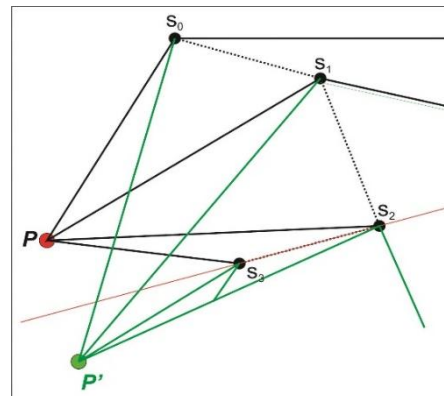
۱. در لیست مرتب شده ی سایتها، بعد از سایت‌های S_j و S_j قرار گرفته باشد. (لزومی ندارد دقیقاً سایت بعدی آنها در لیست مرتب شده باشد، بلکه کافی است که شیب Pe از شیب Ps_j و Ps_j کمتر باشد).

۲. مانند شکل ۳ روی CH_{S+P} قرار گرفته باشد.

با توجه به مطالب گفته شده، می‌توان قضیه زیر را جمع‌بندی نمود.

به روز رسانی نمودار قطبی به این صورت انجام می‌پذیرد که با عبور قطب از هر یک از خطوط مربوط به افزایش صفحه، تغییرات لازم در نمودار قطبی اعمال می‌گردد. همچنین به روز رسانی CH_{S+P} نیز هنگامی انجام می‌شود که قطب از یکی از خطوط افزایش صفحه عبور نموده است. به همین دلیل اگر قطب از خط $L_{(S_i, S_j)}$ عبور کرده باشد، تنها دو سایت S_i و S_j برای بروز رسانی CH_{S+P} مورد بررسی قرار خواهند گرفت. اگر خط $L_{(S_i, S_j)}$ یکی از اضلاع CH_{S+P} نباشد، آنگاه رویه محدب تغییری نخواهد کرد و اگر $L_{(S_i, S_j)}$ یکی از اضلاع رویه محدب باشد، با توجه به جهت حرکت قطب یکی از دو سایت S_i و S_j از رویه محدب حذف و یا به آن اضافه می‌شوند. این یک سایت که به CH_{S+P} اضافه می‌شود یا از آن حذف می‌شود، لزوماً یکی از اعضا رویه محدب مجموعه S می‌باشد که در زمان ثابت می‌توان آن را مشخص کرد.

فرض کنید قطب از خط $L_{(S_a, S_c)}$ عبور کرده باشد S_c و S_a از رئوس رویه محدب مجموعه S می‌باشند. اگر قطب از سمت چپ این خط به سمت راست آن حرکت کرده باشد، هرکدام از دو سایت S_c و S_a که تا کنون عضو CH_{S+P} نبوده‌اند، از این پس یکی از رئوس CH_{S+P} خواهند بود و اگر قطب از سمت راست این خط به سمت چپ آن حرکت کند، خط مماس از قطب به رویه محدب مجموعه S ، از یکی از دو سایت S_c و S_a خواهد گذشت که این سایت عضو CH_{S+P} باقی می‌ماند و سایت دیگر از CH_{S+P} حذف خواهد شد.



شکل ۲: عبور قطب از امتداد خط S_3S_4

برای مثال تصویر ۲ نشان می‌دهد که قطب از خط $L_{(S_2, S_3)}$ عبور کرده است. خط $L_{(S_2, S_3)}$ امتداد یکی از اضلاع رویه محدب مجموعه $\{S_0, S_1, S_2, S_3, P\}$ می‌باشد و قطب از سمت راست این خط به سمت چپ آن حرکت کرده است. با عبور قطب از سایت S_2 و S_3 مورد بررسی قرار می‌گیرند.

قضیه ۲,۱: فرض کنیم قطب از خط $L(s_i, s_j)$ عبور کرده باشد، T مکان هندسی ناحیه‌ای از نمودار است که در خارج از آن تغییر توپولوژی رخ نخواهد داد. اثبات: سایت‌هایی که در لیست مرتب شده براساس زاویه قطبی قبل از s_i و s_j هستند، تحت تأثیر ناحیه قطبی این دو سایت قرار ندارند. بنابراین مرز بالایی ناحیه‌ای که تغییر توپولوژی در خارج از آن اتفاق نمی‌افتد، خط $L(p, s_j)$ خواهد بود. همانطور که قبلاً گفته شد، سایت e مانع از دید سایت‌های قبل از خود نسبت به سایت‌هایی می‌شود که در لیست مرتب شده بعد از آن قرار دارند. بنابراین مرز پایینی مکان هندسی مورد نظر خط $L(p, e)$ است. □

شکل ۳ نمودار قطبی را پس از عبور قطب از خط $L(s_3, s_4)$ نشان می‌دهد. تغییرات توپولوژی فقط در مثلث S'_4PS_6 رخ داده است.

۲-۲- پیچیدگی زمانی به روز رسانی نمودار قطبی با

Algorithm 1
Input: S includes n Point as a set of sites, Point $P=(P_x, P_y)$ as a pole, $NPPD(S)$, V as set of vertexes in planar subdivision, F as set of faces in planar subdivision, E as set of edges in planner region of P
Output: Updated $NPPD(S)$ when pole moves to new point $P'=(P'_x, P'_y)$ that P' is in the adjacent region of P
Step1: Determine the polar region where the pole is located
Step2: Determine the face ($f_i \in F$) where the pole is located
Step3: Determine the edge of f_i that is crossed by moving the Pole (P), suppose $e_i=L(s_a, s_b)$
Step4: Specify region T and update polar region of the sites inside the T , suppose that number of sites inside the T be k

قطب نزدیک متحرک

قضیه ۲,۲ برای n سایت داده شده در صفحه و نمودار قطبی با قطب نزدیک و متحرک P ، زمان مورد نیاز برای پیش‌پردازش $O(n^4 \log_2 n)$ و زمان بازسازی نمودار در هر حرکت متوالی قطب برابر $O(\log n + k)$ است که در آن k تعداد سایت‌های داخل ناحیه T است.

۱-۲- الگوریتم به روز رسانی نمودار قطبی با قطب نزدیک متحرک

تصویر ۴ مراحل اجرای الگوریتم نمودار قطبی با قطب نزدیک متحرک را به صورت شبه کد نشان می‌دهد. برای محاسبه پیچیدگی زمانی بروز رسانی نمودار قطبی، ابتدا در هر گام از الگوریتم، پیچیدگی زمانی را تحلیل می‌کنیم. در گام اول مشخص می‌شود که قطب در ناحیه قطبی کدام سایت قرار گرفته است. با داشتن $NPPD(S)$ ، در زمان $O(\log n)$ می‌توان تعیین کرد که قطب در ناحیه قطبی کدام سایت است. فرض کنید قطب در ناحیه قطبی سایت s_i قرار داشته باشد.

در گام دوم تعیین می‌گردد که قطب در کدام وجه از تقسیم‌بندی صفحه قرار گرفته است. طبق قضیه ای که پیش‌تر بیان شده است، با $O(n^2)$ یال در تقسیم‌بندی صفحه، در زمان $O(\log n)$ می‌توان تعیین نمود که قطب در کدام وجه قرار دارد [۱۵].

در گام سوم تعیین می‌شود که قطب با حرکت پیوسته خود، از کدام یک از یال‌های وجه‌ای که در آن قرار گرفته است، عبور کرده است و این یال مربوط به خط واصل کدام دو سایت می‌باشد.

افراز صفحه به روش بیان شده، موجب می‌شود که وجه‌هایی به صورت چندضلعی‌های محدب (کراندار یا نامحدود) در صفحه ایجاد شوند. این چندضلعی‌ها در بدترین حالت ممکن است $O(n)$ ضلع داشته باشند. برای این که مشخص شود

قضیه ۲,۲ نتایج نظری به دست آمده را خلاصه سازی می کند. این نتایج نظری به راحتی می توانند در انواع کاربردها در زمینه های ردیابی اتومبیل های خودراننده [۱۸] هدایت اشیا زیر سطحی (نظیر زیر دریایی) [۱۹]. ارتباطات مغناطیسی [۲۰] و سایر کاربردها دیده شوند و بدیهی است استفاده از نتایج جدید در خصوص پیچیدگی زمانی، باعث ارتقا کیفیت و سرعت پردازش نرم افزارهای مرتبط خواهد شد.

۳- نتیجه گیری

مسئله اصلی که در این مقاله به عنوان پژوهش اصلی مورد توجه است، حل مسئله نمودارقطبی نقاط با قطب متحرک است که در آن قطب در نزدیکی سایتها و البته خارج پوسته محدب آنها، در حال حرکت است. الگوریتمی که در این مقاله ارائه شده است، ابتدا با یک پیش پردازش صفحه را افزایش می کند و ناحیه ای را که قطب در آن قرار گرفته است، مشخص می کند. سپس ناحیه ای را تعیین می کند که با حرکت قطب، دچار تغییرات شده است و نیاز است که این بخش از نمودار قطبی با قطب نزدیک به روز رسانی گردد. در این الگوریتم زمان پیش پردازش الگوریتم $O(n^4 \log_2 n)$ و زمان باز رسم نمودار در هر حرکت متوالی قطب برابر با $O(\log n + k)$ است که در آن k تعداد سایت های درون ناحیه T است که احتمال تغییر در آنها وجود دارد.

با توجه به کاربردهای بالقوه زیاد این مفهوم در بینایی ماشین و مخابرات پیشنهاد میشود علاقمندان در آینده همین مساله را برای اشیا مختلف حل کنند و کاربردهای دقیقی از آن بدهند.

برای محاسبه پیچیدگی زمان این الگوریتم دو نوع زمان مورد بررسی قرار می گیرد. یکی زمان پیش پردازش و دیگری زمان اجرا.

در زمان پیش پردازش، نمودار قطبی $NPPD(S)$ محاسبه می گردد و افزایش صفحه انجام می شود و همچنین یک پیش پردازش نیز برای مسئله مکان یابی نیز صورت می گیرد که مرتبه زمانی تمام این موارد برابر است با:

$$O(n \log n + n^2 + n^4 \log n) = O(n^4 \log n)$$

در زمان اجرا پس از عبور قطب از هر یک از خطوط افزایش صفحه، ابتدا باید مشخص نمود که قطب از کدام خطوط افزایش صفحه عبور کرده است که با توجه به مطالب گفته شده این کار در زمان $O(\log n)$ انجام می گیرد. سپس باید نمودار قطبی به روز رسانی گردد. که تنها کافیست نمودار قطبی در درون ناحیه T دوباره رسم شود. با فرض این که k سایت در این ناحیه وجود دارند که ترتیب آنها بر اساس زاویه قطبی مشخص است، میتوان در زمان $O(\log n + k)$ نمودار قطبی را بروز رسانی نمود. بنابراین پیچیدگی زمان اجرا برابر است با: $O(\log n + k)$

در ادامه و در قالب یک قضیه نتایج فوق در مورد الگوریتم و پیچیدگی های زمانی و فضایی را جمع بندی می کنیم.

قضیه ۲,۲: الگوریتم ۱ مساله ی محاسبه ی دیاگرام قطبی با قطب نزدیک و متحرک را برای سایت با صرف $O(n^4)$ فضا و در زمان پیش پردازش $O(n^4 \log_2 n)$ حل میکند. این زمان برای هر پرس و چو در مراحل بعدی برابر $O(\log n + k)$ است.

Sadeghi, and Madadi, Najmeh. An optimal solution for dynamic Polar diagram. in CCCG, PP. 51–54, 2010.e

11. Beygi, Mojtaba Nouri and Ghodsi, Mohammad. Polar diagram of moving objects. In 20th Annual Canadian Conference on Computational Geometry, P. 51. Citeseer, 2008.

۱۲. بزاز، زینب، کاربرد دیاگرام ورونوی حساس به زاویه در مسایل بینایی، پایان نامه کارشناسی ارشد، دانشگاه صنعتی امیرکبیر، ۱۳۸۹

13. Aronov, Boris, Edelsbrunner, Herbert, Guibas, Leonidas J., and Sharir, Micha. The number of edges of many faces in a line segment arrangement. *Combinatorica*, 12(3):261–274, 1992.

14. De Berg, Mark, Cheong, Otfried, van Kreveld, Marc, and Overmars, Mark. *Computational geometry*. Springer Berlin Heidelberg, Berlin, third ed., 2008.

15. KirkPatrick, David. Optimal search in planar subdivisions. *SIAM Journal on Computing*, 12(1):28–35, 1983.

16. Chazelle, Bernard and Dobkin, David P. Intersection of convex objects in two and three dimensions. *Journal of the ACM (JACM)*, 34(1):1–27, 1987.

۱۷. ربانی، فاطمه، دیاگرام قطبی با قطب متحرک. پایان نامه کارشناسی ارشد، دانشگاه تحصیلات تکمیلی علوم پایه زنجان، ۱۳۹۱

18. Sun, Qinbo, et al. "Tacking Control of an Autonomous Sailboat Based on Force Polar Diagram." 2018 13th World Congress on Intelligent Control and Automation (WCICA). IEEE, 2018.

19. The magnetotelluric (MT) method is commonly used to estimate the subsurface conductivity structure.

20. Pranata, Erick, Selvi Misnia Irawati, and Sintia Windhi Niasari. "Magnetotelluric Data Analysis using Swift Skew, Bahr Skew, Polar Diagram, and Phase Tensor: a Case Study in Yellowstone, US."

۴-منابع

1. Grima, CI, Márquez, A, and Ortega, L. A locus approach to angle Problems in computational geometry. In Proc. of 14th European Workshop in Computational Geometry, Barcelona, 1998.

2. Grima, CI, Márquez, A, and Ortega, L. Polar diagrams of geometric objects. In 15th European Workshop in Computational Geometry, 1999.

3. Grima, C, Márquez, A, and Ortega, L. Motion Planning and visibility Problems using Polar diagrams. In Annual conference of the European association for computer graphics, EG. Citeseer, 2003.

4. Bigham, B Sadeghi and Mohades, Ali. The dual of Polar diagrams and its extraction. In International Conference of Computational Methods in Sciences and Engineering ICCMSE, vol. 7, PP. 451–454, 2006.

5. Ortega, Lidia M, Rueda, Antonio J, and Feito, Francisco R. A solution to the Path Planning Problem using angle Pre-Processing. *Robotics and Autonomous Systems*, 58(1):27–36, 2010.

6. Ortega, Lidia and Robles-Ortega, M Dolores. Visibility resolution with Polar diagrams. *APPL. Math*, 7(5):1651–1669, 2013.

7. Bigham, B Sadeghi and Mohades, Ali. Polar diagram with respect to a near Pole. In 23rd European Workshop on Computational Geometry EWCG07, Austria, PP. 206–209. Citeseer, 2007.

8. Bigham, Bahram Sadeghi, Eskandari, Marzieh, and Tahmasbi, Maryam. Near-Pole Polar diagram of objects and duality. *Journal of Computational Science*, 3(3):127–131, 2012.

9. Sadeghi Bigham, Bahram, Mohades, Ali, and Ortega, Lidia. Dynamic Polar diagram. *Information Processing Letters*, 109(2):142–146, 2008.

10. Ehsanfar, Ebrahim, Bigham, Bahram

راهنمای نگارش

مقالات فصلنامه علمی - پژوهشی فناوری اطلاعات و ارتباطات ایران

ج - چکیده: متنی حاوی حداکثر ۲۰۰ واژه که بیانگر نکات اصلی مقاله باشد.

د - کلید واژگان: شامل حداکثر ۵ واژه اصلی که بتواند در نمایه‌گذاری مقاله، ویژگی‌های آن را توصیف کند.

ه - بدنه مقاله: شامل مقدمه، تبیین مسئله اصلی تحقیق، روش کاربردی برای حل مسئله و نتایج آن. این بخش باید چنان کامل باشد که خواننده با مطالعه آن به اقتناع علمی دست یابد از سویی لازم است نگارندگان از تطویل کلام بپرهیزند و بخشهای غیر ضروری را از مقاله حذف کنند.

و - مراجع: شامل همه کتابها، مقاله‌ها، گزارشها و صفحه‌های وب مورد استناد در مقاله.

یادآوری ۱: شکلها، جدولها و مراجع باید به ترتیبی که در متن آمده‌اند شماره‌گذاری شود. (عنوان شکل در زیر آن و عنوان جدول در بالای آن ذکر می‌شود).

یادآوری ۲: شیوه نگارش مراجع و ارجاع به آن در متن، مطابق استاندارد نشریات IEEE است.

یادآوری ۳: در مقاله‌های فارسی لازم است همه اصطلاحها، اسامی و اعداد کاربردی در متن مقاله، جدولها و شکلها به زبان فارسی نوشته شود و در صورت نیاز معادل فارسی آن در زیرنویس صفحه آورده شود. مبنای انتخاب معادل فارسی، مصوبت فرهنگستان زبان و ادب فارسی است.

یادآوری ۴: شیوه نگارش فارسی فصلنامه براساس «دستور خط فارسی» مصوب فرهنگستان زبان و ادب فارسی است و نگارندگان لازم است اصول نگارش آن را در مقاله رعایت کنند.

۵- مقاله‌های دریافتی به هر دو زبان فارسی و انگلیسی قابل پذیرش است. فصلنامه شماره‌های ویژه‌ای را به مقاله‌های انگلیسی زبان اختصاص خواهد داد. به هر حال همه مقاله‌ها (خواه فارسی، خواه انگلیسی) باید دارای چکیده دو زبانه باشند.

۶- لازم است نگارندگان مقاله خود را در محیط MS-Word به صورت دو ستونه، با فاصله خطوط ۱ (Single) و با قلم بی-لوتوس (B Lotus) ۱۱ (برای مقاله فارسی) و Times New Roman ۱۱ (برای مقاله انگلیسی) و با رعایت حاشیه بالا ۳ سانتی‌متر، پایین ۱ سانتی‌متر، چپ ۲٫۵ سانتی‌متر، راست ۲٫۵ سانتی‌متر و با توجه به نکات بند ۴، آماده و به دفتر نشریه

فصلنامه علمی-پژوهشی «فناوری اطلاعات و ارتباطات ایران»، نشریه‌ای علمی است که با هدف ایجاد ارتباط میان پژوهشگران و اندیشمندان عرصه فناوری اطلاعات و اعتلای سطح دانش نظری این حوزه علمی منتشر می‌شود. این فصلنامه در پی آن است که زمینه‌ای مناسب را برای ارائه آخرین یافته‌های علمی محققان در حوزه‌های مختلف فناوری اطلاعات در داخل و خارج از ایران پدید آورد.

فصلنامه پذیرای مقاله‌هایی است که حاصل پژوهشهای اصیل نظری، کاربردی و توسعه‌ای در زمینه‌های تخصصی فناوری اطلاعات و فناوری ارتباطات باشد، از این رو از همه پژوهشگران و متخصصان دعوت می‌شود تا نتایج تحقیقات خود را برای انتشار به فصلنامه ارسال کنند. در این خصوص اشاره به نکات زیر ضروری است.

۱- مقاله‌های ارسالی به فصلنامه باید دربرگیرنده نتایج نو در هر یک از حوزه‌های تخصصی فصلنامه باشد و نباید قبلاً در هیچ نشریه دیگری به چاپ رسیده یا به طور همزمان به سایر نشریه‌های داخلی و خارجی برای داوری ارسال شده باشد. شایان ذکر است انتشار مقاله کاملی که خلاصه نتایج آن پیشتر در مجامع علمی ارائه و چاپ شده باشد، بلامانع است.

۲- مقاله‌های ارسالی به فصلنامه باید حاوی مطالب جدید و واجد نتایج، روشها، مفاهیم، کاربردها یا ترکیبی از این موارد باشد.

۳- فصلنامه از مقاله‌های مروری، تحلیلی و توصیفی (با دعوت از صاحب‌نظران برجسته) و نیز نامه به سردبیر (در نقد، تحلیل و تشریح مقاله‌ای خاص) استقبال می‌کند.

۴- مقاله‌ها باید بتوانند اطلاعات کافی را به خواننده منتقل کنند، از این رو لازم است حاوی بخشهای زیر باشند:

الف - عنوان مقاله: گویای اصلی‌ترین نکته و یافته مقاله (طول عنوان نباید از ۱۵ واژه بیشتر باشد).

ب - مشخصات نویسندگان: شامل نام و نام خانوادگی، شغل، وابستگی سازمانی، نشانی پستی و الکترونیکی.

۸- هر مقاله پس از بررسی سردبیر به سه داور ارسال می‌شود. جمع‌بندی نظریات داوران برعهده سردبیر فصلنامه است. فصلنامه حق پذیرش، ویرایش یا رد مقاله‌ها را برای خود محفوظ می‌دارد.

۹- فصلنامه متعهد می‌شود در ازای چاپ هر مقاله، تعداد پنج نسخه از بازچاپ مقاله (یا پنج نسخه از مجله) را برای نویسنده عهده‌دار مکاتبات ارسال کند. ضمناً دسترسی به نسخه الکترونیکی مقاله برای همه نویسندگان مقاله امکان‌پذیر خواهد بود.

ارسال کنند. (در ضمن در صورت امکان، PDF آن را نیز ارسال شود).

یادآوری: ضروری است نشانی کامل پستی و الکترونیکی و نیز شماره تلفن نویسنده عهده‌دار مکاتبات (Corresponding Author) در مقاله ارسالی معلوم باشد تا امکان ایجاد ارتباط با نگارندگان فراهم آید.

۷- حق نشر مقاله‌ها برای فصلنامه محفوظ است؛ از این رو ضروری است نویسنده عهده‌دار مکاتبات به نمایندگی از همه نویسندگان «مجوز حق انتقال نشر» را امضا کند، در غیر این صورت فصلنامه از چاپ مقاله پذیرفته شده معذور است.

معرفی انجمن فناوری اطلاعات و ارتباطات ایران

اهداف

- انجمن فناوری اطلاعات و ارتباطات ایران در سال ۱۳۸۳ با اهداف زیر تشکیل گردید:
- ایجاد ارتباط با پژوهشگران و افراد خیره در حوزه‌های مرتبط با فناوری اطلاعات (IT)، فناوری ارتباطات (CT) و فراهم نمودن زمینه‌های شناسایی و همکاری مشترک بین آنها.
 - همکاری با نهادهای اجرائی، علمی و پژوهشی مرتبط با فناوری اطلاعات و ارتباطات، اقدام در جهت پیوند و همکاری دستگاههای اجرائی با مجموعه‌های علمی - تحقیقاتی.
 - توسعه فرهنگ استفاده مطلوب از توانمندیهای فناوری اطلاعات و ارتباطات.
 - ارزیابی و تحلیل عوامل رشد کشورهای مشابه دیگر و استفاده مطلوب از تجربیات دیگران در توسعه کشور در حوزه ارتباطات و فناوری اطلاعات.

برنامه‌ها

- انجمن به منظور تحقق اهداف خود، برنامه‌های زیر را در اولویت فعالیت‌هایش قرار داده است.
- ایجاد و تقویت ارتباط علمی، پژوهشی و آموزشی در سطوح ملی، منطقه‌ای و بین‌المللی.
 - برگزاری همایشهای علمی در سطح ملی، منطقه‌ای و بین‌المللی.
 - ایجاد تسهیلات برای نشر نتایج آخرین دستاوردهای علمی و پژوهشی مرتبط با نیازهای کشور در حوزه ارتباطات و فناوری اطلاعات (انتشار کتب و نشریات علمی).
 - برگزاری کارگاههای تخصصی در حوزه ارتباطات و فناوری اطلاعات.

برنامه‌ریزی فعالیتها

- انتشار خبرنامه انجمن
- تشکیل کمیته‌های تخصصی به منظور تحقق اهداف انجمن
- اقدامات وسیع برای معرفی انجمن در همایشهای علمی، سازمانهای اجرائی و دانشگاهها
- انتشار نشریه علمی - پژوهشی ICT
- همکاری در برگزاری همایشهای علمی از قبیل چهارمین کنفرانس فناوری اطلاعات و دانش (IKT)
- تشکیل جلسات با مسئولین مملکتی ذیربط جهت شناساندن اهداف انجمن و مشخص شدن نحوه همکاری

کمیته‌های تخصصی انجمن

- کمیته پذیرش و اطلاع‌رسانی
- کمیته آموزش
- کمیته پژوهش
- کمیته مالی و پشتیبانی
- کمیته ارتباط صنعت دانشگاه
- کمیته جامعه اطلاعاتی

از کلیه اندیشمندان، اساتید دانشگاه، پژوهشگران و سایر علاقه‌مندان در حوزه ارتباطات و فناوری اطلاعات دعوت می‌شود با عضویت در انجمن، ما را در تحقق اهداف والای آن یاری نمایند.

نشانی دبیرخانه: تهران - خیابان حافظ - دانشگاه صنعتی امیرکبیر - ساختمان ابوریحان - طبقه ششم - اتاق ۶۱۲

تلفن: ۰۲۱-۶۶۴۸۵۸۵۶، تلفکس: ۰۲۱-۶۶۴۹۵۴۳۳

اعضای حقوقی انجمن در سال ۹۲	
<ul style="list-style-type: none"> عصر ارتباطات و انتقال داده‌های سپاهان 	<ul style="list-style-type: none"> امیر صفوی شماره عضویت: ۹۲۴۸۶۵۱۸۴
<ul style="list-style-type: none"> موسسه آموزش عالی غیر انتفاعی نوین 	<ul style="list-style-type: none"> مجید میرزا محمدی شماره عضویت: ۹۲۴۴۵۳۱۸۵
<ul style="list-style-type: none"> جامع علمی و کاربردی خانه کارگر استان اردبیل 	<ul style="list-style-type: none"> ودود مبینی شماره عضویت: ۹۲۴۴۵۳۱۸۶
اعضای پیوسته انجمن در سال ۹۲	
انجمن فناوری اطلاعات و امیر شهاب شاهمیری - ارتباطات	شماره عضویت: ۸۹۱۲۱۳۵۴۰
سید مهدی حاتمیان - وزارت فناوری اطلاعات و ارتباطات ایران	شماره عضویت: ۸۹۱۲۱۲۵۴۳
شیرین گیلکی - جهاد دانشگاهی	شماره عضویت: ۹۰۳۲۱۵۵۸۹
فریدون عبدی - دانشگاه افسری امام علی (ع)	شماره عضویت: ۹۱۳۲۱۳۶۷۰
نرگس رضایی ملک - شرکت هوای فشرده ایمن	شماره عضویت: ۹۲۱۲۱۶۶۸۴
محمد نوری مطلق -	شماره عضویت: ۹۲۱۵۸۶۶۸۶
احمد استیری -	شماره عضویت: ۹۲۱۲۱۶۶۸۶
محمدآبادی - دانشگاه آزاد اسلامی واحد مشهدالهام میرزایی	شماره عضویت: ۹۲۱۵۱۶۶۸۶
ساسان عظیمی - دفتر مطالعه و نوآوری	شماره عضویت: ۹۲۱۲۱۳۷۰۵
زهرا عزیزی - شرکت ایزایران	شماره عضویت: ۹۲۱۲۱۳۷۰۶
غلامحسین قاسمی - دانشکده برق و کامپیوتر، دانشکده فنی	شماره عضویت: ۹۲۱۵۶۳۷۰۷
قائن	
رباب قاسم شربیانی - شرکت مهندسی مشاور نیروی	شماره عضویت: ۹۲۱۴۱۵۷۰۹
آذربایجان (منا)	
سجاد رضائیان - دانشگاه صنعتی شیراز	شماره عضویت: ۹۲۱۷۳۳۷۱۱
پانید علیپور اقدم -	شماره عضویت: ۹۲۱۲۱۵۷۱۲
پریسا پورحسن -	شماره عضویت: ۹۲۱۱۲۳۷۱۳
سامان سیادتی -	شماره عضویت: ۹۲۱۲۱۳۷۱۴
عرفانه غروی -	شماره عضویت: ۹۲۱۷۱۳۷۱۵
کامبیز رضایی - پردیس شهید رجایی فارس	شماره عضویت: ۹۲۱۷۱۳۷۱۶
ایمان قاصدیان - بانک تجارت استان فارس	شماره عضویت: ۹۲۱۷۱۶۷۱۷
شریفی - دانشگاه علوم پزشکی شیرازمژگان شیال	شماره عضویت: ۹۲۱۷۱۳۷۱۸
سید هادی موسوی - دانشگاه صنعتی بیرجند	شماره عضویت: ۹۲۱۵۶۳۷۱۹
نوبری - دانشکده تربیت بدنی و علوم ورزشی دانشگاه دل‌بابک زنده	
تهران	
شماره عضویت: ۹۲۱۲۱۳۷۲۰	

سهبیل افراز- دانشگاه آزاد اسلامی واحد اردبیل شماره عضویت : ۹۲۱۴۵۳۷۲۱	عبدالله مهدوی- دانشگاه علوم پزشکی اردبیل شماره عضویت : ۹۲۱۴۵۳۷۲۲
فاطمه حورعلی- شماره عضویت : ۹۲۱۵۷۳۷۳۳	علی موثق- شماره عضویت : ۹۲۱۲۱۳۷۲۴
محمد رضا کنعانی- اداره کل حفاظت محیط زیست استان مازندران شماره عضویت : ۹۲۱۱۵۳۷۲۵	لیلا انصاری فرد- شماره عضویت : ۹۲۱۴۵۳۷۲۶
اعضای دانشجویی انجمن در سال ۹۲	
محسن گلبند- شماره عضویت : ۹۲۳۲۱۳۶۳۳	فراهانی-دانشگاه صنعتی شریف الهام درمنکی شماره عضویت : ۹۲۳۲۱۳۶۳۴
الهه پوریان- شماره عضویت :	شاهرخ بحتوئی شماره عضویت : ۹۲۳۷۱۳۶۳۵
محمد کاظم حیدری- شماره عضویت : ۹۲۳۷۱۳۶۳۶	پیمان غلامی- دانشگاه آزاد اسلامی اراک شماره عضویت : ۹۲۳۸۶۳۶۳۷
نژاد سهیلا بارچی شماره عضویت : ۹۲۳۳۴۳۶۳۸	میثم افتخاری- دانشگاه آزاد اسلامی واحد اراک شماره عضویت : ۹۲۳۸۶۳۶۳۹
آیت قلندری- دانشگاه آزاد اسلامی واحد اراک شماره عضویت : ۹۲۳۶۶۳۶۴۰	موسوی- مدرسه راهنمایی فرزنانگان ۲ و جبهه لوح شماره عضویت : ۹۲۳۳۱۳۶۴۱
آبادی-دانشگاه قم ملکزهره گلی شماره عضویت : ۹۲۳۳۳۳۶۴۲	بیدگلی- پردیس فنی دانشگاه تهران علی بلالی شماره عضویت : ۹۲۳۲۱۳۶۴۳
محمد مهدی فیروزی- دانشگاه صنعتی امیرکبیر شماره عضویت : ۹۲۳۲۱۳۶۴۴	مسعود نجفی- دانشگاه آزاد اسلامی نجف آباد اصفهان شماره عضویت : ۹۲۳۷۷۳۶۴۵
سپهر قاسمی- پردیس دانشکده فنی دانشگاه تهران شماره عضویت : ۹۲۳۲۱۳۶۴۶	سالار محتاج- دانشگاه صنعتی امیرکبیر شماره عضویت : ۹۲۳۲۱۳۶۴۷
لیلا آذری- دانشگاه آزاد اسلامی واحد علوم و تحقیقات خوزستان شماره عضویت : ۹۲۳۲۱۳۶۴۸	زنجان- دانشگاه صنعتی امیرکبیر محمد دریایی شماره عضویت : ۹۲۳۶۱۳۶۴۹
محسن مرادی- دانشگاه صنعتی مالک اشتر شماره عضویت : ۹۲۳۲۱۳۶۵۰	امیر اسماعیلی- دانشگاه صنعتی امیرکبیر شماره عضویت : ۹۲۳۲۱۳۶۵۱
آتنا کاویان- دانشگاه صنعتی امیرکبیر شماره عضویت : ۹۲۳۲۱۳۶۵۲	آزاده پیرحیاتی- دانشگاه آزاد اسلامی همدان شماره عضویت : ۹۲۳۶۶۳۶۵۳
ناهید فتوحی- دانشگاه سلمان فارسی کازرون شماره عضویت : ۹۲۳۷۱۳۶۵۴	بابالی- وزارت فناوری اطلاعات و ارتباطات میراسماعیل میرنی مازندران شماره عضویت : ۹۲۳۱۱۳۶۵۵
ندا بینش شماره عضویت : ۹۲۳۲۳۳۶۵۶	محمد مهدی محصولی- دانشگاه صنعتی امیرکبیر شماره عضویت : ۹۲۳۲۱۳۶۵۷

خواه- دانشگاه علوم پزشکی شیراز مهدی شرف شماره عضویت : ۹۲۳۷۱۳۶۵۸	حمیدرضا زارعیان- بانک تات شماره عضویت : ۹۲۳۲۱۳۶۵۹
عباس خسروانی- دانشگاه صنعتی امیرکبیر شماره عضویت : ۹۲۳۲۱۳۶۶۰	بابایی فاطمه حاجی شماره عضویت : ۹۲۳۵۱۳۶۶۱
رامین ساروقی- دانشگاه سلمان فارسی کازرون شماره عضویت : ۹۲۳۷۱۳۶۶۲	جم- دانشگاه تبریز معصومه رضائی شماره عضویت : ۹۲۳۸۷۳۶۶۳
بهرام هدایتی شماره عضویت : ۹۲۳۷۳۳۶۶۵	اصغر ترکی- فولاد مبارکه اصفهان شماره عضویت : ۹۲۳۳۱۳۶۶۶
فاطمه قنبریان شماره عضویت : ۹۲۳۷۱۳۶۶۶	محمد ضیاءالدینی- دانشگاه باهنر شماره عضویت : ۹۲۳۳۴۳۶۶۷
خلجی فاطمه کیماسی شماره عضویت : ۹۲۳۳۷۳۶۶۸	وحید ظهیرپور- دانشگاه علم و صنعت ایران شماره عضویت : ۹۲۳۲۱۳۶۶۹
داریوش سرخه- دانشگاه آزاد اسلامی شوش شماره عضویت : ۹۲۳۶۴۳۶۷۰	محسن رضانی شماره عضویت : ۹۲۳۸۷۳۶۷۱
فرزانه کرمی- دانشگاه صنعتی امیرکبیر شماره عضویت : ۹۲۳۲۱۳۶۷۲	محسن یعقوبی سورکی شماره عضویت : ۹۲۳۱۵۳۶۷۳
امیرحسین خانیکی شماره عضویت : ۹۲۳۵۱۳۶۷۴	محمدی علیرضا ملک شماره عضویت : ۹۲۳۳۸۳۶۷۵
رضا سالارمهر- دانشگاه صنعتی امیرکبیر شماره عضویت : ۹۲۳۲۱۳۶۷۶	فاطمه مجیدی- دانشگاه جامع علمی کاربردی خانه کارگر شماره عضویت : ۹۲۳۴۵۳۶۷۷
میثم معمربور- موسسه غیرانتفاعی مقدس اردبیلی شماره عضویت : ۹۲۳۴۵۳۶۷۸	تپراقویوسف عطائی شماره عضویت : ۹۲۳۴۵۳۶۷۹
سلیمان نظیری- موسسه غیرانتفاعی مقدس اردبیلی شماره عضویت : ۹۲۳۴۵۳۶۸۰	خواه- موسسه غیرانتفاعی مقدس اردبیلی شبنم وطن شماره عضویت : ۹۲۳۴۵۳۶۸۱
امید تقوائی- موسسه غیرانتفاعی مقدس اردبیلی شماره عضویت : ۹۲۳۴۵۳۶۸۲	جمادی- موسسه غیرانتفاعی مقدس اردبیلی فریده نعمتی شماره عضویت : ۹۲۳۴۵۳۶۸۳
مجتبی لاله- شورای اسلامی شهر اردبیل شماره عضویت : ۹۲۳۴۵۳۶۸۴	هاجر کوثرزاده- دانشگاه آزاد اسلامی واحد اردبیل شماره عضویت : ۹۲۳۴۵۳۶۸۵
میلاد راستگو- موسسه غیرانتفاعی مقدس اردبیلی شماره عضویت : ۹۲۳۴۵۳۶۸۶	میلاد مددپور- دانشگاه آزاد اسلامی واحد اردبیل شماره عضویت : ۹۲۳۴۵۳۶۸۷
سحر ولی نژادروح بخش- موسسه غیرانتفاعی مقدس اردبیلی شماره عضویت : ۹۲۳۴۵۳۶۸۸	گلناز هاشمی- موسسه غیرانتفاعی مقدس اردبیلی شماره عضویت : ۹۲۳۴۵۳۶۸۹
رضا فرجی- موسسه غیرانتفاعی مقدس اردبیلی شماره عضویت : ۹۲۳۴۵۳۶۹۰	بهاءالدین کاوه پور- موسسه غیرانتفاعی مقدس اردبیلی شماره عضویت : ۹۲۳۴۵۳۶۹۱
مریم قاسمی- موسسه غیرانتفاعی مقدس اردبیلی	فاطمه جعفرزاده- موسسه غیرانتفاعی مقدس اردبیلی

۹۲۳۴۵۳۶۹۳ شماره عضویت :	۹۲۳۴۵۳۶۹۲ شماره عضویت :
توحید فرهادی- دانشگاه آزاد اسلامی اردبیل ۹۲۳۴۵۳۶۹۵ شماره عضویت :	زینب تیزپر- موسسه غیرانتفاعی مقدس اردبیلی ۹۲۳۴۵۳۶۹۴ شماره عضویت :
جواد مؤذنی- موسسه غیرانتفاعی مقدس اردبیلی ۹۲۳۴۵۳۶۹۷ شماره عضویت :	محمد وجدانی- دانشگاه آزاد اسلامی اردبیل ۹۲۳۴۵۳۶۹۶ شماره عضویت :
حجت همپایی- موسسه غیرانتفاعی مقدس اردبیلی ۹۲۳۴۵۳۶۹۹ شماره عضویت :	سید مسعود سجادی- موسسه غیرانتفاعی مقدس اردبیلی ۹۲۳۴۵۳۶۹۸ شماره عضویت :
سجاد آخربین- موسسه غیرانتفاعی مقدس اردبیلی ۹۲۳۴۵۳۷۰۱ شماره عضویت :	عباسی- موسسه غیرانتفاعی مقدس برحامد بیرون اردبیلی ۹۲۳۴۵۳۷۰۰ شماره عضویت :
زاده- موسسه غیرانتفاعی مقدس اردبیلی مینا عالی ۹۲۳۴۵۳۷۰۳ شماره عضویت :	زاده- موسسه غیرانتفاعی مقدس اردبیلی مهدی بیک ۹۲۳۴۵۳۷۰۲ شماره عضویت :
اعضای وابسته انجمن در سال ۹۲	
الهه قنبریان ۹۲۲۷۱۶۳۸۴ شماره عضویت :	سعید صفوی ۹۱۵۲۱۵۳۴۸ شماره عضویت :
نقی زندیان اجیرلو- موسسه غیرانتفاعی مقدس اردبیلی ۹۲۲۴۵۳۳۸۶ شماره عضویت :	علیرضا ابریشمی شیرازی ۹۲۲۷۱۶۳۸۵ شماره عضویت :
حمید نصیری ۹۲۲۴۵۳۳۸۸ شماره عضویت :	سید عباس پیغمبرزاده- هنرستان غیرانتفاعی ایران مهارت شماره عضویت: ۹۲۲۴۵۳۳۸۷
حسین رهبرفر ۹۲۲۴۵۳۳۹۰ شماره عضویت :	جعفر زاهدی اول ۹۲۲۵۷۳۳۸۹ شماره عضویت:
مهدی جانفشار- دانشگاه آزاد اسلامی واحد اردبیل ۹۲۲۴۵۳۳۹۲ شماره عضویت :	محمد رضا سید پیروی ۹۲۲۴۵۳۳۹۱ شماره عضویت:
سمیرا نجفی ۹۲۲۴۵۳۳۹۴ شماره عضویت :	توحید ابراهیمی- اداره کل راه و شهرسازی اردبیل ۹۲۲۴۵۳۳۹۳ شماره عضویت :

New Method to Improve Illumination Variations in Adult Images Based on Fuzzy Deep Neural Network

Sasan Karamizadeh

Arabsorkhi Abouzar

Abstract:

In the era of the Internet, recognition of adult images is important to children's physical and mental protection. It is a challenge to recognize adult images with changes in the illumination and skin color. In this paper, we proposed a new method for solving illumination normalization with skin color classification in the diagnosis of the adult image. In this paper, the deep fuzzy neural network method is utilized to improve the illumination normalization of adult images, which has improved the recognition of adult images is utilized. Using Xception to dividing the images and reduce the illumination variations in each part separately, which makes it possible to reduce the illumination variation in the whole image without losing details. In addition, the advanced color combination algorithm based on Gaussian-KNN algorithm is used for skin color classification, a non-parametric method is used for classifications and regressions. Finally, the SVM algorithm is utilized for image classification. In this paper, 33,000 different types of images are collected from the Internet. The results show that the proposed method of 1/3 has improved the accuracy of the recognition.

Keywords: ILLUMINATION VARIATION, SKIN CLASSIFICATION, ADULT IMAGE, SVM, GAUSSIAN-KNN

A model information technology adoption in academic research projects in the filed ICT based on Information Technology Adoption Integrated Modeling (ITAIM)

Shahram Aliyari

Masoud Movahedi

Sirous Kazemian

Abstract:

Today, the emergence and expansion of technologies that provide the widest possible connection have brought about significant changes in the private life and professional life of individuals. Correct implementation of information technology is the source of economic and cultural development and the promotion of quality of life through the exchange of information and the provision of public and private services. The purpose of this research is to present the model of information technology acceptance in Iranian ICT research centers. Employed experts worked on ICT projects in this research Statistical population. It was provided in one of the university searching centers. And so a convenience and purposeful Nonprobability sampling does (30 person). This paper examines the factors and parameters affecting the acceptance of information technology in ICT projects of the university research centers in the field of information technology. To collect the required data and information, a questionnaire was used & to analyze the data and information obtained from the questionnaires using the Spss 22 and Smart pls3 software. According to the calculations, the factors affecting the acceptance of information technology in university research centers can be divided into four categories: IT related factors, organizational factors, factors related to executive director and individual factors that are related to management (0.497), IT (0.460) and Individual factors (0.457) have an impact on the individual acceptance of information technology respectively and organizational factors (0.469) on the adoption of an IT organization

Keywords: Information Technology, Information Technology Acceptance, Research Projects, ITAIM Model

Introducing a genetic algorithm based Method for Community person's stance Detection in social media and news

Mehdi Salkhordeh haghghi

Seyyed Mohammad Ebrahimi

Abstract:

News reports in social media are presented with large volumes of different kinds of documents. The presented topics in these documents focus on different communities and person stances and opinions. Knowing the relationships among persons in the documents can help the readers to obtain a basic knowledge about the subject and the purpose of various documents. In the present paper, we introduce a method for detecting communities that includes the persons with the same stances and ideas. To do this, the persons referenced in different documents are clustered into communities that have related positions and stances. In the presented method. Community-based personalities are identified based on a friendship network as a base method. Then by using a genetic algorithm, the way that these communities are identified is improved. The criterion in the tests is rand index of detection of these communities. The experiments are designed based on real databases that published in Google News on a particular topic. The results indicate the efficiency and desirability of the proposed method

Keywords: Social media, news media, community recognition, individual position recognition

Provide a method using genetic algorithm to identify the position of people in society in news and social media

Mehdi Salkhordeh haghghi

Seyyed Mohammad Ebrahimi

Abstract:

News reports in social media are presented with large volumes of different kinds of documents. The presented topics in these documents focus on different communities and person stances and opinions. Knowing the relationships among persons in the documents can help the readers to obtain a basic knowledge about the subject and the purpose of various documents. In the present paper, we introduce a method for detecting communities that includes the persons with the same stances and ideas. To do this, the persons referenced in different documents are clustered into communities that have related positions and stances. In the presented method. Community-based personalities are identified based on a friendship network as a base method. Then by using a genetic algorithm, the way that these communities are identified is improved. The criterion in the tests is rand index of detection of these communities. The experiments are designed based on real databases that published in Google News on a particular topic. The results indicate the efficiency and desirability of the proposed method.

Keywords: Social network, social media, community detection, community persons stance

context-aware travel recommender system exploiting from Geo-tagged photos

Rezvan Mohamadrezai larki

Reza Ravanmehr

Milad Amrolahi

Abstract:

Recommender systems are the systems that help users find and select their target items. Most of the available events for recommender systems are focused on recommending the most relevant items to the users and do not include any context information such as time, location . This paper is presented by the use of geographically tagged photo information which is highly accurate. The distinction point between this thesis and other similar articles is that this paper includes more context (weather conditions, users' mental status, traffic level, etc.) than similar articles which include only time and location as context. This has brought the users close to each other in a cluster and has led to an increase in the accuracy. The proposed method merges the Colonial Competitive Algorithm and fuzzy clustering for a better and stronger processing against using merely the classic clustering and this has increased the accuracy of the recommendations. Flickr dataset is used to evaluate the presented method. Results of the evaluation indicate that the proposed method can provide location recommendations proportionate to the users' preferences and their current visiting location.

Keywords: context aware recommender system, travel recommender system, context, Colonial Competitive Algorithm, fuzzy clustering

Polar Diagram of Points with Moving Pole

Bahram Sadeghi Bigham

Fateme Rabani

Abstract:

Polar diagram is a generalization of Voronoi diagram in which the angle is used as the metric. This Problem has many applications in visibility, image Processing, telecommunication, antenna, and Path Planning Problems. In recent years two kinds of Polar diagram have been proposed and appropriate algorithm have been Presented for some types of sites. Also, some algorithms has presented for kinetic data and dynamic states. In this Paper, it is assumed that the Pole is moving and an algorithm is presented that updates near Pole Polar diagram of sites with moving pole efficiently and in a sub linear time. In this approach, the Preprocessing time is $O(n^4 \log n)$ and updating time for diagram with each successive movement is $O(\log(n+k))$ that k is the number of sites in region T which its site's regions may be changed

Keywords: Polar Diagram, Voronoi Diagram, Polar Angle, Telecommunication, Antenna, Visibility.

Contents

Iranian Journal of
Information Technology & Communication
No. 41-42, Vol.11, September-March 2020

- **New Method to Improve Illumination Variations in Adult Images Based on Fuzzy Deep Neural Network** 1
Sasan Karamizadeh, Arabsorkhi Abouzar
 - **A model information technology adoption in academic research projects in the filed ICT based on Information Technology Adoption Integrated Modeling (ITAIM)** 13
Shahram Aliyari, Masoud Movahedi, Sirous Kazemian
 - **Introducing a genetic algorithm based Method for Community person's stance Detection in social media and news** 33
Mehdi Salkhordeh haghghi , Seyyed Mohammad Ebrahimi
 - **Provide a method using genetic algorithm to identify the position of people in society in news and social media** 57
Mehdi Salkhordeh haghghi, Seyyed Mohammad Ebrahimi
 - **context-aware travel recommender system exploiting from Geo-tagged photos** 75
Rezvan Mohamadrezaei larki , Reza Ravanmehr , Milad Amrolahi
 - **Polar Diagram of Points with Moving Pole** 97
Bahram Sadeghi Bigham , Fateme Rabani
-