

ارائه یک سیستم توصیه گر برای شناخت کاربران بد در شبکه های اجتماعی با استفاده از تکنیک داده کاوی (بررسی موردی شبکه پیام رسان تلگرام)

سیده منیر موسوی^۱

^۱کارشناس ارشد ادبیات نمایشی، S_m_mosavi@yahoo.com

چکیده

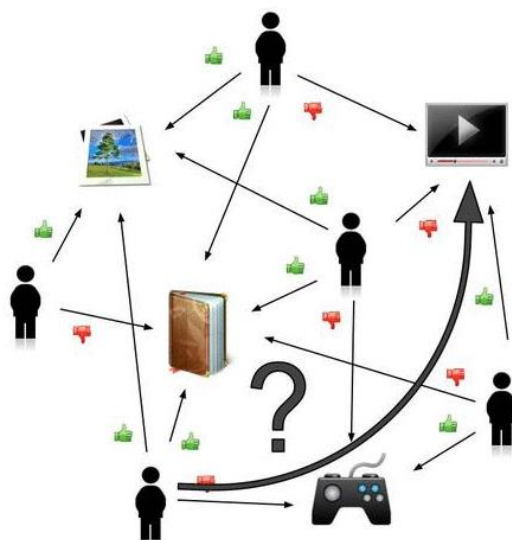
میزان تبادل اطلاعات و نزدیکی روابط در شبکه های اجتماعی برخط ضرورت کنترل و نظارت بر این شبکه ها رو دو چندان کرده است. در چنین شبکه هایی که هر یک برای پوشش هدفی خاص ایجاد شده اند همواره کاربران متنوعی وجود دارد که هریک اهداف خاصی را از عضویت دنبال می کنند. در چنین فضایی که فرزندان، شخصیت و هویت افراد به تبع حضور برخی از کاربران بد در معرض تهدید است طراحی سامانه ای توصیه گر برای تشخیص کاربران بد بسیار ضروری به نظر می رسد. بر همین اساس در این پژوهش روشی برای تشخیص کاربران بد و عادی در پیام رسان اجتماعی تلگرام پیشنهاد گردید. در روش پیشنهادی از طبقه بند ماشین بردار پشتیبان بهینه سازی شده با الگوریتم ژنتیک بهره گرفته شد. بررسی نتایج پژوهش نشان می دهد روش پیشنهادی توانسته ۶٪ روش پایه را ماشین بردار پشتیبان می باشد؛ بهبود بخشد. طبق یافته ها دقت روش پیشنهادی ۹۱٪ و روش پایه ۸۵٪ می باشد. جدول – مقایسه دقت روش پایه و پیشنهادی را نشان می دهد

کلمات کلیدی: سیستم توصیه گر، تلگرام، کاربران بد، شبکه اجتماعی.

۱- مقدمه و بیان مساله

انسان‌ها اجتماعی‌ترین موجوداتی هستند که در تعامل با دیگران یاد می‌گیرند و این تعامل باعث وابستگی متقابل بین آنها می‌شود. شبکه اجتماعی یک ساختار اجتماعی متشکل از گروه‌هایی است که از طریق یک یا چند نوع وابستگی معین نظیر ارزش‌ها، بینش‌ها، داد و ستد مالی، دوستی، خویشاوندی، تنفر، کشمکش از این دست وابستگی‌ها، با هم در ارتباط هستند. شبکه‌های اجتماعی برخط مانند فیس‌بوک، توئیتر و گوگل پلاس، به هر کاربر خود امکان می‌دهند که از خود یک سیمای مجازی تحت عنوان پروفایل شخصی بسازند و آن را به دیگر پروفایل‌های آن شبکه متصل کنند. برخلاف ارتباط رودررو در جهان واقعی، در شبکه‌های اجتماعی برخط، یک کاربر در چگونگی شخصی‌سازی پروفایل خود اختیار کامل دارد به‌گونه‌ای که می‌تواند آزادانه انتخاب کند که کدام تصویرش را به نمایش بگذارد، چه اطلاعاتی را از خود منتشر کند، و چه کسانی را به عنوان دوستان و آشنایان خود معرفی نماید. شبکه‌های اجتماعی برخط یک اصل را به نمایش می‌گذارند: به‌ازای هر دو نفر از جمعیت جهان، انواع روابط می‌تواند شکل بگیرد.

با گسترش روزافزون استفاده از شبکه‌های اجتماعی، روش‌های جست و جوی اطلاعات، به اشتراک گذاری دانش و تجربه و نحوه ارتباط افراد تحت تاثیر قرار گرفته و دچار تغییر شده اند. عامل مهمی که باعث شده است این تکنولوژی با این سرعت رشد پیدا کند، امکان تولید آسان محتوا توسط کاربر در آن می باشد. تقاضا برای جست و جوی دانش از مجموعه داده‌های موجود در این شبکه‌ها و یادگیری رفتار کاربران در آن باعث شده است نیاز به وجود روش‌هایی جهت حذف اطلاعات اضافی و غیر مفید و برجسته‌تر نمودن اطلاعات کارآمد و کاربردی بیش از پیش مورد توجه قرار بگیرد. سیستم‌های توصیه‌گر به عنوان راه حلی برای مواجهه با مشکل سربار اطلاعات و سردرگمی در یافتن اطلاعات مفید و کاربردی ظهور یافته‌اند و به صورت مشاوران مجازی هوشمند با ترکیب اطلاعات موجود در پروفایل کاربر، تصفیه اطلاعات و یادگیری در ارائه خدمات اطلاعاتی هوشمند و فعال و تولید پیشنهاداتی که به نیازها و سلیق کاربران نزدیک است، بسیار مفید و مثمر ثمر واقع شده‌اند. این سیستم‌ها از روش‌های گوناگونی جهت تصفیه اطلاعات و ارائه توصیه‌ها استفاده می‌نمایند [۱].



شکل ۱- نمایی از ارتباطات و سلیق کاربران در شبکه‌های اجتماعی

سیستم توصیه‌گر، علائق کاربران را نسبت به مجموعه‌ای از اقلام جمع‌آوری می‌کند. این اطلاعات به صورت صریح یا ضمنی بدست می‌آیند. سیستم‌های توصیه‌گر تلاش می‌کنند که به وسیله پیشنهاد اطلاعاتی که مورد علاقه کاربران باشد بر مشکل سرریز اطلاعات غلبه کنند. با رشد وب بهره‌گیری از اطلاعات شبکه‌های اجتماعی در سیستم‌های توصیه‌گر افزایش پیدا کرده است. سیستم‌های توصیه‌گر اجتماعی با بهره‌گیری از تعاملات اجتماعی نقاط ضعف سیستم‌های توصیه‌گر سنتی را مرتفع می‌کنند [۲]. در توصیه‌گرهای مبتنی بر روابط اجتماعی، ارتباط میان کاربران از شبکه اجتماعی استنتاج می‌شود و با

استفاده از روابطی که به صورت صریح میان کاربران و

از طرف خود آنها اعلام شده است، موارد مشابه محاسبه می‌شوند و پیشنهادهایی برای ارائه به یک کاربر خاص آماده می‌شوند. معلمان می‌توانند از طریق شبکه های اجتماعی به آسانی به دانش آموزان خود دسترسی پیدا کنند و یک محیط دوستانه برای مطالعه و تعامل دانش آموزان ارائه کنند. کارفرمایان می‌توانند از طریق شبکه های اجتماعی نسبت به شناسایی افراد با استعداد و علاقه مند اقدام کنند. این بررسی پس زمینه می‌تواند با استفاده از این شبکه های اجتماعی به درستی انجام شود. بسیاری از شبکه های اجتماعی رایگان هستند اما برخی از آنها برای عضویت، هزینه ای را در نظر گرفته اند و این هزینه برای ارائه خدمات بیشتر است [۳]. در عصر سایبری امروز، شبکه های اجتماعی باعث شده اند کل جهان به هم نزدیک شوند و در عین حال مشکلات زیادی مانند هویت های جعلی، شبیه سازی پروفایل، حملات هرزنامه ها، نشر اکاذیب، نشر متون و تصاویر مستحجن و غیره ایجاد کرده اند. پژوهش پیشنهادی به دنبال این خواهد بود تا با بهره گیری از ترکیب الگوریتم تکاملی جدید میگو و شبکه های عصبی سامانه ای توصیه گر مبتنی بر محتوا برای شناسایی کاربران بد در شبکه اجتماعی ارائه دهد. در این مقاله پس از تبیین موضوع کارهای مرتبط بررسی خواهد شد. در ادامه روش پیشنهادی شرح داده شده و نتایج حاصل تحلیل می‌شود. در خاتمه یک جمع بندی از مطالب ذکر شده ارائه می‌گردد.

۲- کارهای پیشین

در [۴۶] یک سیستم توصیه گر خدمات گردشگری معرفی گردیده است که با ترکیب داده های اولیه ی موجود و داده های ثانویه ی حاصل از تعامل مستقیم و غیر مستقیم با کاربران، توصیه هایی هدفمند و با قابلیت بهبود مداوم دقت و تناسب آنها با هر کاربر را ارائه می‌نماید. در این سیستم، بر اساس پروفایل کاربر و تعاملات وی درون سیستم، میزان تعلق کاربر با گروه های گردشگری استاندارد، پیش بینی می‌شود که در هر گروه نیز، وزن علاقمندی ها و جاذبه های گردشگری، متفاوت و متناسب با حداکثر کاربران آن گروه می‌باشد. در ادامه، بر اساس موقعیت جغرافیایی کاربر، از طریق داده های جی پی اس، متناسب ترین توصیه ها بر اساس زمان، فاصله، هزینه و ... ارائه می‌گردد. در هر مرحله از فرآیند جمع آوری اطلاعات با هدف کاهش شروع سرد، گسترش پروفایل کاربر و دریافت بازخورد از کاربر، وزن دهی متغیرهای مختلف، از جمله وزن کاربر در گروه و وزن جاذبه ها و علاقمندی ها برای کاربر، مبتنی بر مفاهیم هستان شناسی، اصلاح می‌گردد.

در [۴۷] یک سیستم توصیه گر تحت موبایل برای گردشگران گروهی و انفرادی جهت توصیه جاذبه، رستوران و اقامتگاه ارائه شده است. به منظور توصیه صحیح و دقیق، الگوریتمی مرکب از تکنیکهای دموگرافی، فیلتر همبستگی و مبتنی بر محتوا با در نظر گرفتن موقعیت جغرافیایی کاربر و ساعت کاری جاذبه ها معرفی می‌شود و همچنین جهت استخراج گزینههای مناسب برای گروه، از روشهای تجمیع و تقاطع استفاده میشود. نتایج حاصل از ارزیابی این پژوهش روشن نمود که الگوریتم ترکیبی به همراه روش تقاطع نتایج خوبی ارائه می‌دهد.

در [۴۸] روشی ارائه شده که همه اطلاعات موجود در یک مقاله را دریافت کرده و از طریق تولید پرس و جو های کوتاه و طولانی، برای هر کاربر پروفایلی تشکیل می‌دهد، که این پروفایل برای هر فرد شخصی سازی شده است و سپس از طریق پروفایل کاربر به وی مرتبط ترین مقاله را توصیه می‌کند. نتایج حاصله نشان می‌دهد که روش پیشنهادی نسبت به روش های مشابه بهتر عمل می‌کند.

در [۴۹] یک روش جدید به منظور بهبود کارایی سیستم های توصیه گر مبتنی بر اعتماد معرفی شده است. در واقع با استفاده از اطلاعات اجتماعی، علاوه بر رفع ضعف های مربوط به سیستم های پالایش گروهی، میزان دقت و پوشش بهبود پیدا کرده است.

در [۲] سیستم های توصیه گر اجتماعی و نقش شبکه اجتماعی در بهبود سیستم های توصیه گر سنتی مورد بررسی قرار گرفته است. سیستم های توصیه گر اجتماعی در این مقاله به دو دسته مبتنی بر حافظه و مبتنی بر مدل تقسیم می‌شوند و مورد مطالعه و نقد قرار می‌گیرند

در [۵۰] سیستمی معرفی می‌شود که از موضوع های پنهان اطلاعات غیر سازمان یافته موجود در بازار های الکترونیکی جهت رفع مشکلات ذکر شده بهره می‌برد. در این سیستم، از یک مد موضوعی احتمالاتی برای یافتن موضوع های پنهان، استفاده می‌شود. موضوع های پنهان، اساس محاسبه شباهت بین واژه ها و کالا ها و فراهم آوردن توصیه های مناسب

هستند. در ادامه، ساختار سیستم پیشنهادی معرفی

می شود و کیفیت موضوع های استخراج شده مورد ارزیابی قرار می گیرد. نتایج ارزیابی بیانگر عملکرد بهتر این سیستم می باشد

در [۱] با بررسی تعدادی از مدل های اعتماد مورد استفاده و تحلیل عملکرد آن ها، یک مدل استخراج اعتماد ارائه گردیده است که بر اساس تاریخچه فعالیت کاربران در شبکه اجتماعی، مجموعه اعتماد آن ها را شناسایی می کند. برای ارزیابی این مدل نیز از داده های شبکه اجتماعی *epinions* استفاده شده و نتایج آزمایش با مجموعه های اعتمادی که در حال حاضر در این شبکه موجود هستند مقایسه شده است.

در [۵۱] پس از بیان توضیحات و مروری جامع بر سیستم های توصیه گر، روشی پیشنهادی مبتنی بر روش فیلترینگ جمعی همراه با پیاده سازی و نتایج حاصل از آن در مقایسه با دیگر روش های موجود ارائه گردیده است.

در [۵۲] پژوهشگران سعی بر ایجاد و معرفی معیار جدیدی برای ارزیابی شباهت بین کاربران کرده اند. آنها معیار جدید را «آلفا» نام گذاری کرده اند. ارزیابی بر اساس داده های واقعی، نشان از میانگین خطای قدر مطلق بهتر معیار آلفا دارد.

۳- روش پژوهش

روش پیشنهادی این مقاله شامل سه مرحله کلی پیش پردازش، پردازش و تشخیص می باشد. در ادامه توضیحات هر یک از این مراحل آمده است.

۳-۱- پیش پردازش:

مرحله ای از فرآیند پردازش متن است که طی آن متن ورودی به ساختاری قابل پردازش برای مراحل بعد تبدیل می شود. این ساختار بستگی به ویژگی هایی دارد که برای استفاده در بخش های بعدی مورد نیاز است. مهم ترین فرآیندهای پیش پردازشی صورت گرفته روی متن ورودی، در این پایان نامه به شرح زیر است.

یک دست سازی پیکره های متنی: گاهی نویسه های بکار رفته در دو واژه ی یکسان، با هم متفاوت هستند؛ این باعث می شود که هنگام شمردن واژه ها، دو واژه یکسان با املاهای متفاوت به عنوان دو واژه مختلف در نظر گرفته شوند. برای جلوگیری از بروز این مشکل، به ابزاری برای یک دست سازی پیکره متنی نیاز داریم. به عنوان مثال، پیشوند «می» و پسوند «ها» در ابتدای و انتهای واژه ها، ممکن است به سه صورت مختلف دیده شوند.

چسباندن	جدا با فاصله	جدا با نیم فاصله
کتابها	کتاب ها	کتاب‌ها
می‌رود	می رود	می‌رود

و یا واژه های «مسئول»، «مجموعه»، و «پاییز» به صورت های زیر در پیکره ی متنی دیده شوند

مسئول	مسؤول	مسوول
مجموعه ی	مجموعه	مجموعه
پاییز	پائیز	

به عنوان مثال در مورد پسوند «ها» می توان شکل های «جدا با فاصله» و «جدا بدون فاصله» را ناندیده گرفت و با مشاهده هر یک از این دو، آن را به شکل چسباندن درآورد. در ادامه یک مجموعه تبدیل های نسبتاً جامع و مؤثر برای یک دست سازی متون فارسی ارائه شده است. تجربه نشان می دهد که با استفاده از تبدیل های زیر، می توان پیکره متنی را به خوبی یک دست سازی کرد:

تبدیل نویسه های «ی» و «ک» عربی به نوع فارسی آن.

تبدیل نویسه های «ؤ» به «و»، «ئ» به «ی» و «أ» به «ا».

تبدیل نویسه های «ة» و «ه» به «ه» در آخر

واژه ها.

حذف «ی» از آخر واژه هایی مانند «خانه ی».

حذف فتحه، کسره و ضمه (نویسه های َ ، ِ ، ُ) از واژه ها.

حذف تنوین (نویسه های ً ، ٍ ، ٌ) از انتهای واژه ها.

حذف تشدید (شناسه ّ) از واژه ها.

حذف شناسه «ء» در آخر بعضی واژه ها مانند «شهداء».

چسباندن پیشوندهای «می»، «درمی»، «برمی»، «نمی»، و «بی» به ابتدای واژه ها.

چسباندن پیشوند «هم» به واژه های مانند «هم چنین» به ابتدای واژه ها.

چسباندن پیشوند «به» به واژه هایی مانند «به ندرت» به ابتدای واژه ها.

چسباندن پسوندهای «ها»، «های»، «هایی»، «هایم»، «هایت»، «هایش»، «هایمان»، «هایتان»، «هایشان» به انتهای

واژه ها.

چسباندن پسوندهای «تر» و «ترین» به آخر واژه ها.

حذف فاصله بعد از پیشوند «بر» در واژه هایی مانند «بر می گردد».

تبدیل واژه هایی مانند «مسیول» و «مسئله» به «مسئول» و «مسئله». واژه های «مستول» و «مسئله»، با تبدیل دو به

«مسیول» و «مسئله» تبدیل می شوند و تبدیل ۱۵ این شکل های میانی را شکل به نهایی تبدیل می کند.

حذف نویسه «-» که برای کشش نویسه های چسبان مورد استفاده قرار می گیرد. مانند تبدیل «بر» و «بر» به «بر». معمولاً از این

نویسه برای تراز کردن طول خطها استفاده می شود.

تعیین مرز واژه ها: تمامی سامانه های پردازشی متن نیازمند شناسایی مرز واژه ها برای تفکیک آن ها از هم هستند.

شناسایی واژه ها نیز از طریق بررسی علائم قابل انجام است. این علائم عبارتند از: فضای خالی، علامت پرش، علامت خط جدید، «،»، «>»، «<»، «[»، «>»، «-»، «_» و «/». در اینجا هم بررسی علائم به تنهایی کافی نیست و معمولاً برای بهبود

کارایی باید از روش های پیچیده تری استفاده نمود. به عنوان مثال، تعیین مرز بعضی فعل ها (مانند «جارو زدن»، «زده

بوده باشد») و واژه های مرکب (مانند «ساده زیست») به این روش امکان پذیر نیست.

حذف واژه های غیر مهم: واژه های غیر مهم، واژه هایی هستند که پرتکرارند، اما مهم نیستند. در صورتی که فراوانی

واژه ها را گواهی بر اهمیت آن ها بدانیم، ضروری است که غیر مهم بودن آن را مورد بررسی قرار دهیم.

ریشه یابی: ریشه یابی به فرآیند کاهش دادن لغات به ریشه های آنها اطلاق می گردد. بنابراین «computer» و

«compute» و «computing» به «compute» که ریشه اصلی است کاهش می یابند. لازم به ذکر است که منظور از ریشه

در این بخش، دقیقاً ریشه کلمات که در زبان شناسی استفاده می باشد نمی باشد. بلکه منظور از ریشه یک نماینده برای کلماتی

است که از لحاظ معنایی و نحوی در یک حوزه قرار می گیرند. تمامی سیستم های بازیابی اطلاعات نوع یکسانی از «ریشه یاب» را

مورد استفاده قرار نمی دهند. در انگلیسی معروف ترین ریشه یاب، الگوریتم ریشه یاب «مارتین پورتر» است [5]. در روش

پیشنهادی از ابزار ریشه یابی دانشگاه فردوسی استفاده شده است.

۳-۲- مرحله پردازش:

در این پژوهش متون کاربران در شبکه اجتماعی به صورت برداری از ویژگی ها مدل می شود که در آن هر بردار شامل ۵ ویژگی

متنی است. روش پیشنهادی با استفاده از حاصل جمع وزن ویژگی ها، برای متون کاربران یک ارزش کل اختصاص می دهد. ویژگی

های مورد استفاده در سامانه پیشنهادی به شرح ذیل می باشند. در ادامه نحوه محاسبه این ویژگی ها آمده است.

جدول ۱- معرفی ویژگی های متن

دسته ویژگی ها	عنوان ویژگی
---------------	-------------

ویژگی‌های مکانی	طول نسبی متن
	تأثیر کلمات زاید
ویژگی‌های آماری و کلاسیک	میانگین فراوانی واژه‌ها
	شباهت با بردار کلمات اختصاصی
	شباهت با بردار زمینه

با توجه به اینکه مقادیر به دست آمده برای هر یک از ویژگی‌ها در یک دامنه نخواهند بود؛ ضرورت دارد تمامی مقادیر به دست آمده برای متون موجود در دیتاست استاندارد سازی شوند. برای این منظور در روش پیشنهادی از رابطه (۱) استفاده شده است. با استفاده از این رابطه تمامی مقادیر به صورت استاندارد به بازه (۱ و ۰) منتقل می‌شوند.

(۱)

$$x_{in} = \frac{x_i - x_{min}}{x_{max} - x_{min}}$$

تأثیر کلمات ایست: متونی که دارای تعداد زیادی کلمات ایست هستند؛ بار معنایی کم دارند. هر چقدر متنی با استفاده از کلمات کمتری بتواند معنای بیشتری را منتقل نماید، اهمیت آن بیشتر خواهد بود. برای تعیین میزان اطلاع رسانی جملات از رابطه (۲) استفاده می‌شود.

(۲)

$$x_i = \frac{|S_{NS}^i|}{|S^i|}$$

که در آن:

$|S_{NS}^i|$ تعداد کلمات غیر ایست متن i -ام و $|S^i|$ طول متن برحسب تعداد کل کلمات.

طول متن: برای محاسبه این ویژگی، ابتدا طول متن پس از نرمالسازی، حذف کلمات توقف و ریشه یابی کلمات به طول واقعی خود کاهش می‌یابد. سپس طول تمامی متون موجود در دیتاست محاسبه می‌شود. قدرمطلق طول واقعی هر متن با میانگین به دست آمده، بیانگر ارزش طول متن خواهد بود. رابطه (۳) چگونگی محاسبه این ویژگی را نشان می‌دهد.

(۳)

$$L = w_{avg} - |w_{avg} - w_s|$$

فراوانی واژه‌های با اهمیت: فراوانی واژه‌ها در متن بعنوان معیاری برای تشخیص اهمیت یک متن می‌باشد. یعنی اگر متنی شامل واژه‌هایی باشد که فراوانی بالایی را به خود اختصاص داده اند پس آن متن اهمیت بیشتری خواهد داشت. در شبکه‌های اجتماعی وقتی این وزن برای متن اشاعه شده توسط کاربری بیشتر باشد بیانگر این امر خواهد بود که کاربر دیدگاه‌های خود را هم راستا با سایر کاربران منتشر می‌کند.

(۴)

$$Score(w) = f(w) * \log(N)/S(W)$$

که $f(w)$ میزان تکرار کلمه (کاندید) در یک متن، N تعداد کلمات موجود در یک متن و $S(W)$ تعداد متونی از دیتاست است که حاوی آن کلمه است. مجموع وزن کلمات (کاندید) نشان دهنده اهمیت آن جمله را نشان می‌دهد.

شباهت با بردار کلمات تخصصی: در سامانه پیشنهادی متن هر کاربر پس از عملیات پیش پردازشی به برداری از کلمات تبدیل می‌شود تا شباهت بردار تشکیل شده با بردار کلمات تخصصی (در این پایان نامه کلماتی می‌باشد که معرف یک کاربر بد است) به دست آید. این میزان هرچه قدر برای کاربری بیشتر باشد بیانگر بد بودن آن کاربر است. رابطه (۵) نحوه محاسبه این شباهت را معرفی می‌کند.

(۵)

$$SimK = \frac{\text{keywords in } UT \cap \text{keywords in } ST}{\text{keywords in } UT \cup \text{keywords in } ST}$$

در رابطه بالا:

سایت متلب کار - مرجع پروژه‌های آماده و شبیه سازی با متلب

Keywords in UT: کلمات موجود در بردار متن کاربر می

باشد.

Keywords in ST: کلمات موجود در بانک کلمات تخصصی می باشد.

شباهت با زمینه: در روش پیشنهادی پس از پیش پردازش کلیه متون منتشر شده توسط کاربر (کلیه متون موجود در دیتاست) و حذف کلمات مشابه به یک مجموعه از کلمات خواهیم رسید که این مجموعه به صورت برداری از کلمات بیانگر زمینه اصلی دیتاست خواهد بود. هر چه قدر شباهت متن منتشر شده توسط کاربری به زمینه بیشتر باشد بیانگر این امر خواهد بود که کاربر پیرامون زمینه در شبکه اجتماعی محاوره می کند. این میزان شباهت طبق رابطه (۶) محاسبه خواهد شد.

(۶)

$$SimB = \frac{keywords\ in\ UT \cap keywords\ in\ BT}{keywords\ in\ UT \cup keywords\ in\ BT}$$

در رابطه بالا:

Keywords in UT: کلمات موجود در بردار متن کاربر می باشد.

Keywords in BT: کلمات موجود در بردار زمینه می باشد.

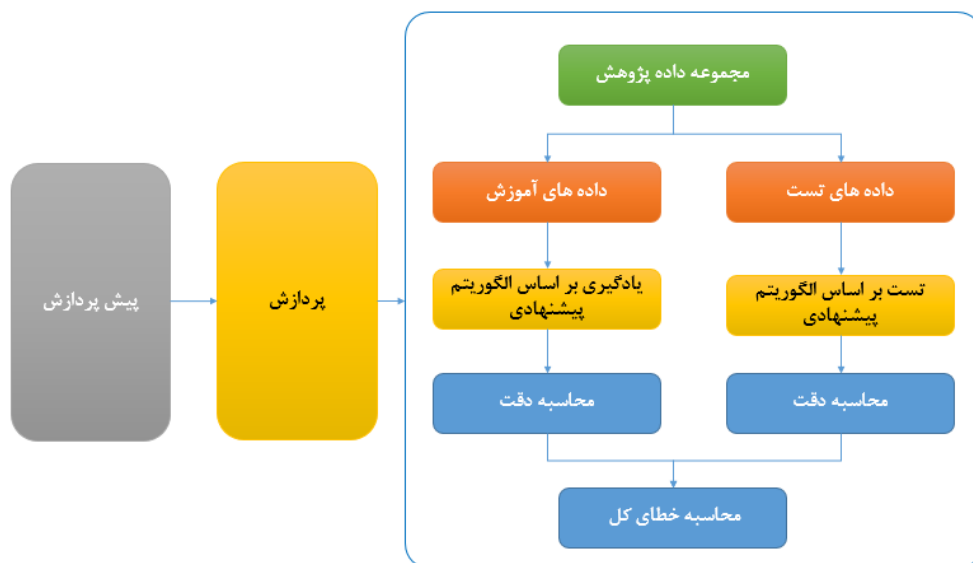
۳-۳- تشخیص:

پس از استخراج ویژگی های متنی برای پیام های منتشر شده توسط کاربران در شبکه اجتماعی، مقدار هر از ویژگی ها برای هر متن در مجموعه داده پژوهش درج می شود تا مجموعه داده نهایی پژوهش برای تشخیص کاربران عادی و بد شکل بگیرد. در مجموعه داده پژوهش هر یک از متون توسط گروه های انسانی برچسب گذاری شده است به این صورت که برچسب ۱ بیانگر کاربر بد و برچسب صفر بیانگر کاربر عادی می باشد. ساختار نهایی مجموعه داده پژوهش به صورت جدول زیر خواهد بود.

جدول ۲- ساختار فرضی دیتاست نهایی روش پیشنهادی

ردیف	متن	برچسب	کاربر	طول متن کاربر	تاثیر کلمات زائد	فراوانی واژه های متن	شباهت کلمات تخصصی	با شباهت با زمینه
۱	فردا نمرات درس هوش مصنوعی پیشرفته در سایت دانشگاه قرار می گیرد.	۱	U1	۰/۶	۰/۸	۰/۵	۰/۷	۰/۸
۲	امیدوارم همه دوستان در درس هوش مصنوعی پیشرفته قبول بشوند.	۱	U2	۰/۵	۰/۷	۰/۵	۰/۸	۰/۷
۳	من فردا با حسین میخوام برم کوهنوردی برای خوش گذارنی.	۰	U3	۰/۶	۰/۶	۰/۲	۰/۱	۰/۱

پس از مراحل پیش پردازش و پردازش، مجموعه داده نهایی پژوهش به دو مجموعه یادگیری و آزمون تقسیم می شود. روش پیشنهادی در مرحله یادگیری بر اساس الگوهای برچسب گذاری شده به یک رابطه منطقی برای تشخیص کاربران بد و عادی در شبکه اجتماعی دست پیدا می کند. سپس دقت روش پیشنهادی بر اساس معیار ارزیابی پژوهش بررسی می شود. شکل ۱ چگونگی این فرآیند را نشان می دهد.

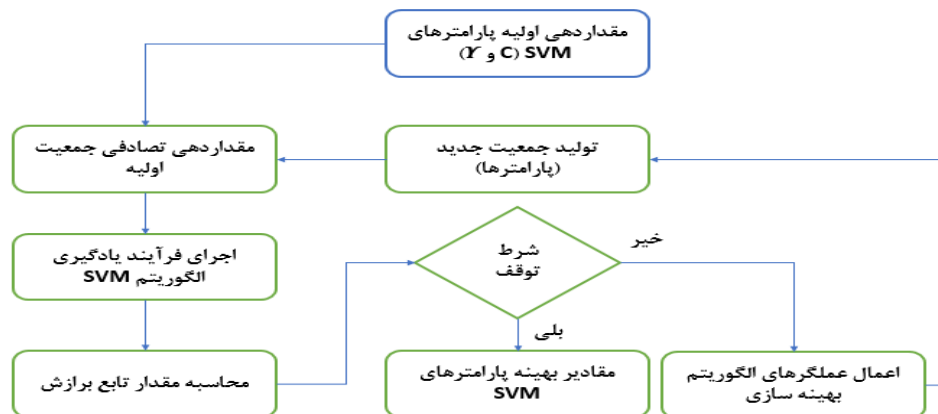


شکل ۲- بلوک دیاگرام کلی روش پیشنهادی

ماشین‌های بردار پشتیبان (SVM): SVM یک نوع سیستم یادگیری است که هم برای دسته‌بندی داده‌های ورودی و هم برای تخمین و برآورد تابع برازش داده‌ها به کار می‌رود، به طوری که کمترین خطا در دسته‌بندی داده‌ها و تابع برازش رخ دهد. داده‌ها کلاً به سه دسته آموزشی، صحت‌سنجی و آزمون تقسیم می‌کنیم به طوری که داده‌های آموزشی باعث آموزش ماشین بردار پشتیبان می‌شوند، داده‌های صحت‌سنجی به واسنجی پارامترهای ماشین می‌پردازد و در نهایت از این ماشین برای طبقه‌بندی یا برآورد داده‌های آزمون استفاده می‌شود. این روش بر مبنای تئوری بهینه‌سازی مقید است که از اصل کمینه‌سازی خطای ساختاری استفاده کرده و منجر به یک جواب بهینه کلی می‌گردد.

بهبود عملکرد ماشین بردار پشتیبان: برای ساخت مدل SVM موثر، پارامترهای آن باید با دقت تنظیم شود. در الگوریتم ماشین بردار پشتیبان پارامترها به صورت پیش فرض تعریف شده اند و هرگز نمی توان ادعا کرد الگوریتم طراحی شده همواره در بهترین وضعیت خود قرار داشته و بهترین جواب ها را تولید خواهد کرد. لذا تعیین پارامترها را می توان به عنوان یک مسئله بهینه سازی در نظر گرفت. در این مقاله استفاده از الگوریتم های بهینه سازی برای تعیین مقدار بهینه پارامترها پیشنهاد شده است. در کنار انتخاب بهینه پارامترها، انتخاب صحیح هسته ماشین بردار پشتیبان نیز در روند پیش بینی موثر خواهد بود. با توجه به ماهیت دیتاست مقاله و نظر به اینکه کرنل (هسته) خطی می تواند جواب های مناسب و مورد قبولی را تولید کند در ماشین بردار پشتیبان پیشنهادی پژوهش حاضر از کرنل خطی استفاده شد.

در ماشین بردار پشتیبان دو پارامتر اصلی به نام های پارامتر جریمه خطا (C) و پارامتر تابع هسته γ وجود دارد که به منظور دستیابی به نتایج مطلوب ضرورت دارد مقادیر بهینه ای برای آنها انتخاب شود. روند کلی بهینه سازی پارامترهای ماشین بردار پشتیبان با الگوریتم های بهینه سازی در بلوک دیاگرام شکل (۲) آمده است.

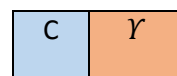


شکل ۳- روند کلی بهینه سازی پارامترهای ماشین بردار پشتیبان با الگوریتم های بهینه سازی

۳-۱- گام های بهینه سازی پارامترهای رگرسیون بردار پشتیبان

باتوجه به شکل (۴-۱) در ادامه هر یک از گام های روند کلی بهینه سازی پارامترهای رگرسیون بردار پشتیبان با الگوریتم های بهینه سازی به اختصار بیان می شود.

مقداردهی اولیه پارامترها (نمایش برداری پارامترها): هدف از بهبود رگرسیون بردار پشتیبان در این پایان نامه، انتخاب مقادیر بهینه برای پارامترهای C و γ در الگوریتم SVM می باشد. لذا در هر عضو از جمعیت اولیه به صورت برداری از دو ویژگی C و γ مطابق با شکل (۳) تعریف می شود.



شکل ۴- آرایه ویژگی ها

نرمالسازی ویژگی ها دیتاست: به منظور دستیابی به نتایج بهتر و غلبه بر مشکلات ناشی از عدم تناسب ویژگی ها، تغییر مقیاس (نرمال سازی) روی مقادیر ویژگی بردارهای ورودی انجام می شود. بدین ترتیب محدوده مقادیر ویژگی های نرمال شده بصورت خطی به بازه [۱-۱] تغییر مقیاس می یابند.

مقداردهی تصادفی پارامترها: در این مرحله جمعیت اولیه ای از مقادیر سه پارامتر اصلی رگرسیون بردار پشتیبان (C و γ) مقداردهی و تولید می شود.

یادگیری ماشین بردار پشتیبان: در این مرحله بخشی از مجموعه داده پژوهش، به عنوان مجموعه داده مرحله یادگیری SVM مورد استفاده قرار می گیرد. سپس برحسب مقادیر برجسب گذاری شده و رویکردهای متفاوت یادگیری ماشین، SVM به یک رابطه مناسبی برای دسته بندی کاربران در شبکه های اجتماعی دست پیدا می کند. به عبارتی در این مرحله پس از تولید جمعیت اولیه، مقدار تابع هزینه برای هر نمونه از جمعیت اولیه مورد بررسی قرار می گیرد. مقدار تابع هزینه هر نمونه در این تحقیق، میانگین درصد خطای مطلق می باشد.

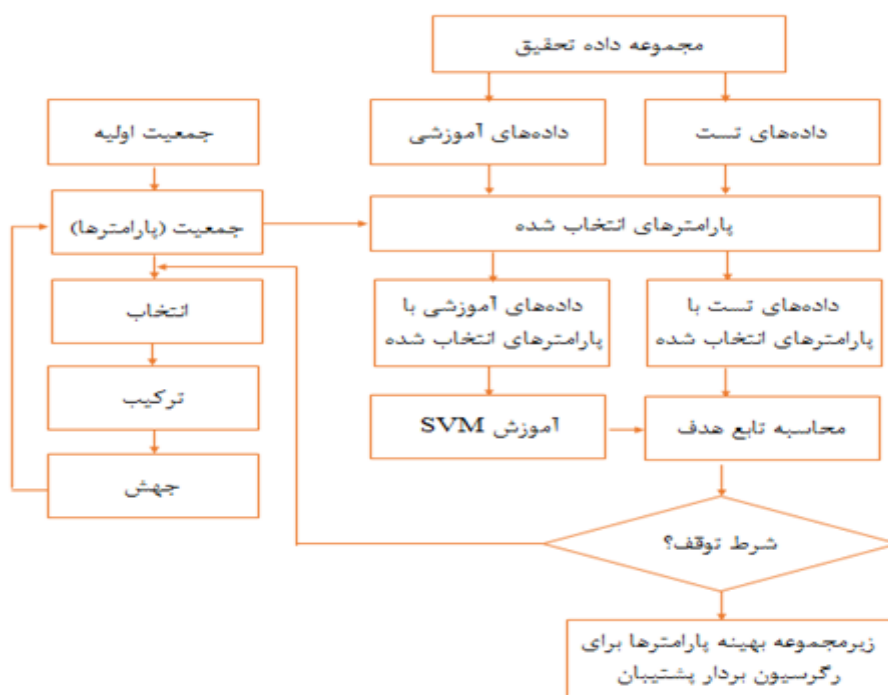
محاسبه تابع هدف: تابع هدف بهینه سازی پارامترهای ماشین بردار پشتیبان در این مقاله دقت SVM در تشخیص کاربران بد خواهد بود.

شرط توقف: پس از بدست آوردن بردار ویژگی بهینه، مرحله بهینه سازی پارامترهای SVM خاتمه می یابد.

۳-۲- انتخاب پارامترهای بهینه با استفاده از الگوریتم ژنتیک

چنانچه پارامتربررسی شده شرط توقف الگوریتم را ارضا کند، آن دسته از پارامترهای (C و γ) به عنوان پارامترهای بهینه الگوریتم ماشین بردار پشتیبان در نظر گرفته می شود. حال اگر شرط توقف ارضا نشود، مجدد عملگرهای الگوریتم بهینه ساز بکارگرفته می شوند تا نمونه های جدیدتری تولید شود. این فرآیند آنقدر ادامه پیدا می کند تا شرط توقف بوسیله یک نمونه ارضا شود.

در این رویکرد پس از تولید جمعیت اولیه (کروموزوم ها که هر ژن آن بیانگر یکی از پارامترهای ماشین بردار پشتیبان است) با استفاده از عملگر های ترکیب و جهش فرزندان جدیدی تولید می شود. سپس به وسیله تابع برازش، برازندگی هر کروموزوم محاسبه می شود (تابع برازش در این مسئله کمینه سازی خطای پیش بینی مصرف آب مشترکین است). در ادامه توسط عملگر انتخاب جمعیت نسل بعدی انتخاب می شود و بهترین کروموزوم ذخیره می شود. این فرآیند تا زمانی ادامه پیدا می کند که برای نسل های متوالی کروموزوم بهتری تولید نشود. بلوک دیاگرام شکل (۴) چگونگی این فرآیند را نشان می دهد.



شکل ۵- چارت روند بهینه سازی پارامترهای رگرسیون بردار پشتیبان با الگوریتم ژنتیک

۴- معیار ارزیابی

معیار ارزیابی این پژوهش دقت^۱ خواهد بود. این معیار دقت کل یک دسته بند را محاسبه می کند. در واقع این معیار مشهورترین و عمومی ترین معیار محاسبه کارایی الگوریتم های دسته بندی است که نشان می دهد، دسته بند طراحی شده چند درصد از کل مجموعه رکوردهای آزمایشی را بدرستی دسته بندی کرده است. دقت دسته بندی با استفاده از رابطه زیر بدست می آید که بیان می کند دو مقدار TP و TN مهمترین مقادیری هستند که در یک مسئله دودسته ای باید بیشینه شوند. (۷)

¹ Accuracy

$$AC = \frac{TN + TP}{TN + FN + TP + FP}$$

جدول ۳- معرفی ماتریس ارزیابی مقاله

کل	مثبت شرایط	منفی شرایط
شرایط پیش‌بینی مثبت	مثبت حقیقی	مثبت کاذب
شرایط پیش‌بینی منفی	منفی کاذب	منفی واقعی

در ماتریس فوق مفهوم عبارت‌ها به شرح ذیل است:

- **منفی واقعی:** بیانگر تعداد کاربرانی است که دسته‌ی واقعی آن‌ها بد بوده و الگوریتم نیز آن‌ها را به‌درستی بد تشخیص داده است.
- **مثبت حقیقی:** بیانگر تعداد کاربرانی است که دسته‌ی واقعی آن‌ها عادی بوده و الگوریتم نیز دسته‌ی آن‌ها را به‌درستی عادی تشخیص داده است.
- **مثبت کاذب:** بیانگر تعداد کاربرانی است که دسته‌ی واقعی آن‌ها بد بوده و الگوریتم آن‌ها را به‌اشتباه عادی تشخیص داده است.
- **منفی کاذب:** بیانگر تعداد کاربرانی است که دسته‌ی واقعی آن‌ها عادی بوده و الگوریتم آن‌ها را به‌اشتباه بد تشخیص داده است.

۵- یافته‌های پژوهش

همانطور که در بخش‌های قبلی قاله نیز اشاره شد برای ارزیابی روش پیشنهادی از معیار دقت استفاده شده است. برای این منظور الگوریتم SVM و SVMGA بر روی مجموعه داده پژوهش اعمال گردید. در ادامه به بررسی هر یک از روش‌های موصوف پرداخته شده است.

تشخیص کاربران بد در شبکه اجتماعی تلگرام با SVM: جدول ۴ یافته پژوهش را به وسیله الگوریتم طبقه بندی کننده SVM برای تشخیص کاربران بد در شبکه اجتماعی تلگرام نمایش می دهد. یافته‌ها بیانگر این امر می باشد که ماشین بردار پشتیبان توانسته است با دقت ۸۵٪ کاربران بد را در پیام رسان تلگرام تشخیص دهد.

جدول ۴- ماتریس درهم ریختگی دسته معاشرت انسانی

	پیش بینی مثبت	پیش بینی منفی
تشخیص دسته	۱۰۴	۷
عدم تشخیص دسته	۱۳	۱۲۳

تشخیص کاربران بد در شبکه اجتماعی تلگرام با SVMGA: جدول ۵ یافته پژوهش را به وسیله الگوریتم طبقه بندی کننده SVM برای تشخیص کاربران بد در شبکه اجتماعی تلگرام نمایش می دهد. یافته‌ها بیانگر این امر می باشد که ماشین بردار پشتیبان توانسته است با دقت ۹۱٪ کاربران بد را در پیام رسان تلگرام تشخیص دهد.

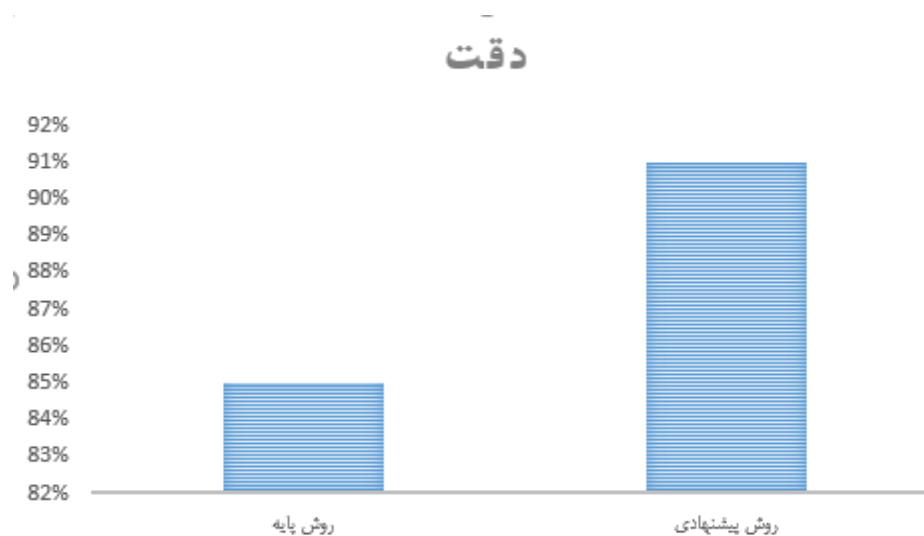
جدول ۵-ماتریس درهم ریختگی دسته معاشرت انسانی

	پیش بینی مثبت	پیش بینی منفی
تشخیص دسته	۶	۸
عدم تشخیص دسته	۴	۳۱

۶- نتیجه گیری

سیستم های توصیه گر در شبکه های اجتماعی به منظور مختلفی طراحی می شوند؛ فیلترینگ، شناسایی لینک های مشکوک، موجودیت های مخرب و ... از جمله مواردی هست که نظر پژوهشگران را به سمت خود سوق داده است. میزان تبادل اطلاعات و نزدیکی روابط در شبکه های اجتماعی برخط ضرورت کنترل و نظارت بر این شبکه ها رو دو چندان کرده است. در چنین شبکه هایی که هر یک برای پوشش هدفی خاص ایجاد شده اند همواره کاربران متنوعی وجود دارد که هریک اهداف خاصی را از عضویت دنبال می کند. در چنین فضایی که فرزندان، شخصیت و هویت افراد به تبع حضور برخی از کاربران بد در معرض تهدید است طراحی سامانه ای توصیه گر برای تشخیص کاربران بد بسیار ضروری به نظر می رسد. امری که به جرات می توان گفت مطالعات کمی روی آن انجام شده است. بر همین اساس در این پژوهش روشی برای تشخیص کاربران بد و عادی در پیام رسان اجتماعی تلگرام پیشنهاد گردید. در روش پیشنهادی از طبقه بند ماشین بردار پشتیبان بهینه سازی شده با الگوریتم ژنتیک بهره گرفته شد.

بررسی نتایج پژوهش نشان می دهد روش پیشنهادی توانسته ۶٪ روش پایه را ماشین بردار پشتیبان می باشد؛ بهبود بخشد. طبق یافته ها دقت روش پیشنهادی ۹۱٪ و روش پایه ۸۵٪ می باشد. جدول - مقایسه دقت روش پایه و پیشنهادی را نشان می دهد.



شکل ۶- مقایسه دقت روش پیشنهادی و روش پایه

مراجع:

۱. تیر، ش.م.م.ش.ک.، تاثیر اعتماد در بهبود عملکرد سیستم های توصیه گر در شبکه های اجتماعی آنلاین. ۱۳۹۳، وزارت علوم، تحقیقات و فناوری - دانشگاه صنعتی خواجه نصیرالدین طوسی - دانشکده صنایع.
۲. محمد، ک. and ف. محمد، بررسی سیستم های توصیه گر اجتماعی، in کنفرانس ملی مدیریت و فناوری اطلاعات و ارتباطات. ۱۳۹۳.
- ۳ Sarode, A.J. and A. Mishra. *Audit and Analysis of Impostors: An experimental approach to detect fake profile in online social network*. in *Proceedings of the Sixth International Conference on Computer and Communication Technology* ۲۰۱۵. ۲۰۱۵. ACM.
- ۴ Rich, E., *User modeling via stereotypes*. Cognitive Science, ۱۹۷۹. ۳(۴): p. ۳۵۴-۳۲۹
- ۵ Powell, J.D., *Approximation Theory and Methods*. ۱۹۸۱: Cambridge University Press.
- ۶ Salton, G., *Automatic text processing: the transformation, analysis, and retrieval of information by computer*. ۱۹۸۹: Addison-Wesley.
- ۷ Armstrong, J.S., *Principles of Forecasting: A Handbook for Researchers and Practitioners*. ۲۰۰۱: Springer.
- ۸ Murthi, B.P.S. and S. Sarkar, *The Role of the Management Sciences in Research on Personalization*. Manage. Sci., ۲۰۰۳. ۴۹(۱۰): p. ۱۳۶۲-۱۳۴۴
- ۹ Lilien, G.L., P. Kotler, and K.S. Moorthy, *Marketing models*. ۲۰۰۳: Prentice-Hall of India.
- ۱۰ Guy, I. and D. Carmel, *Social recommender systems*, in *Proceedings of the ۲۰th international conference companion on World wide web*. ۲۰۱۱, ACM: Hyderabad, India. p. ۲۸۴-۲۸۳
- ۱۱ Guy, I., et al., *Social media recommendation based on people and tags*, in *Proceedings of the ۳۳rd international ACM SIGIR conference on Research and development in information retrieval*. ۲۰۱۰, ACM: Geneva, Switzerland. p. ۲۰۱-۱۹۴
- ۱۲ Sigurbjörnsson, B. and R.v. Zwol, *Flickr tag recommendation based on collective knowledge*, in *Proceedings of the ۱۷th international conference on World Wide Web*. ۲۰۰۸ ACM: Beijing, China. p. ۳۳۶-۳۲۷
- ۱۳ Guy, I., I. Ronen, and E. Wilcox, *Do you know?: recommending people to invite into your social network*, in *Proceedings of the ۱۴th international conference on Intelligent user interfaces*. ۲۰۰۹, ACM: Sanibel Island, Florida, USA. p. ۸۶-۷۷
- ۱۴ Chen, W.-Y., et al., *Collaborative filtering for orkut communities: discovery of user latent behavior*, in *Proceedings of the ۱۸th international conference on World wide web*. ۲۰۰۹, ACM: Madrid, Spain. p. ۶۹۰-۶۸۱
- ۱۵ Jannach, D., et al., *Recommender Systems: An Introduction*. ۲۰۱۰: Cambridge Univ Pr.
- ۱۶ Singla, P. and M. Richardson, *Yes, there is a correlation:-from social networks to personal behavior on the web*, in *Proceedings of the ۱۷th international conference on World Wide Web*. ۲۰۰۸, ACM. p. ۶۶۴-۶۵۵
- ۱۷ O'Donovan, J. and B. Smyth, *Trust in recommender systems*, in *Proceedings of the ۱۰th international conference on Intelligent user interfaces*. ۲۰۰۵, ACM: San Diego, California, USA. p. ۱۷۴-۱۶۷
- ۱۸ Jamali, M. and M. Ester, *TrustWalker: a random walk model for combining trust-based and item-based recommendation*, in *Proceedings of the ۱۵th ACM SIGKDD international conference on Knowledge discovery and data mining*. ۲۰۰۹, ACM. p. ۴۰۶-۳۹۷

- Golbeck, J.A., *Computing and applying trust in web-based social networks*. ۲۰۰۵, University of Maryland at College Park. p. ۱۹۹
- ۲۰ Massa, P. and P. Avesani, *Trust-aware recommender systems*, in *Proceedings of the ۲۰۰۷ ACM conference on Recommender systems*. ۲۰۰۷, ACM. p. ۱۷-۲۴
- ۲۱ Seth, A. J. Zhang, and R. Cohen, *Bayesian credibility modeling for personalized recommendation in participatory media*. User Modeling, Adaptation, and Personalization, ۲۰۱۰: p. ۲۷۹-۲۹۰
- ۲۲ Chowdhury, M., A. Thomo, and B. Wadge, *Trust-based infinitesimals for enhanced collaborative filtering*, in *Proceedings of the ۱۵th International Conference on Management of Data (COMAD)*. ۲۰۰۹
- ۲۳ Ray, S. and A. Mahanti, *Improving prediction accuracy in trust-aware recommender systems*, in *System Sciences (HICSS), ۴۳rd Hawaii International Conference on*. ۲۰۱۰, IEEE. p. ۹-۱
- ۲۴ Salton, G. and C. Buckley, *Term-weighting approaches in automatic text retrieval*. Information Processing & Management, ۱۹۸۸. ۲۴(۵): p. ۵۱۳-۵۲۳
- ۲۵ Lang, K., *Newsweeder: Learning to filter netnews*, in *In Proceedings of the Twelfth International Conference on Machine Learning*. ۱۹۹۵, Citeseer.
- ۲۶ Mooney, R.J. and L. Roy, *Content-based book recommending using learning for text categorization*, in *Proceedings of the fifth ACM conference on Digital libraries*. ۲۰۰۰, ACM: San Antonio, Texas, United States. p. ۱۹۵-۲۰۴
- ۲۷ Pazzani, M. and D. Billsus, *Learning and Revising User Profiles: The Identification of Interesting Web Sites*. Machine Learning, ۱۹۹۷. ۲۷(۳): p. ۳۱۳-۳۳۱
- ۲۸ Adomavicius, G. and A. Tuzhilin, *Recommendation Technologies: Survey of Current Methods and Possible Extensions*. ۲۰۰۳
- ۲۹ Baeza Yates, R. and B. Neto, *Modern Information Retrieval*. ۱۹۹۹: Addison-Wesley Longman Publishing Co., Inc.
- ۳۰ Miller, B.N., et al., *MovieLens unplugged: experiences with an occasionally connected recommender system*, in *Proceedings of the ۱۱th international conference on Intelligent user interfaces*. ۲۰۰۳, ACM: Miami, Florida, USA. p. ۲۶۳-۲۶۶
- ۳۱ Shardanand, U. and P. Maes, *Social information filtering: algorithms for automating "word of mouth"*, in *Proceedings of the SIGCHI conference on Human factors in computing systems*. ۱۹۹۵, ACM Press/Addison-Wesley Publishing Co. p. ۲۱۰-۲۱۷
- ۳۲ Konstan, J.A., et al., *GroupLens: applying collaborative filtering to Usenet news*. Commun. ACM, ۱۹۹۷. ۴۰(۳): p. ۷۷-۸۷
- ۳۳ Sarwar, B., et al., *Item-based collaborative filtering recommendation algorithms*, in *Proceedings of the ۱۰th international conference on World Wide Web*. ۲۰۰۱, ACM: Hong Kong, Hong Kong. p. ۲۸۵-۲۹۵
- ۳۴ Breese, J.S., D. Heckerman, and C. Kadie, *Empirical analysis of predictive algorithms for collaborative filtering*, in *Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence*. ۱۹۹۸, Morgan Kaufmann Publishers Inc. p. ۴۳-۵۲
- ۳۵ Claypool, M. et al. *Combining Content-Based and Collaborative Filters in an Online Newspaper*. ۱۹۹۹; Available from: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=۱۰.۱.۱.۱۴۵.۹۷۹۴>.
- ۳۶ Cotter, P. and B. Smyth, *PTV: Intelligent Personalised TV Guides*, in *Proceedings of the*

Seventeenth National Conference on Artificial Intelligence and Twelfth Conference on Innovative Applications of Artificial Intelligence. ۲۰۰۰, AAAI Press. p. ۹۶۴-۹۵۷

۳۷ Basu, C., H. Hirsh, and W. Cohen, *Recommendation as classification: using social and content-based information in recommendation*, in *Proceedings of the fifteenth national/tenth conference on Artificial intelligence/Innovative applications of artificial intelligence*. ۱۹۹۸, American Association for Artificial Intelligence: Madison, Wisconsin, United States. p. ۷۲۰-۷۱۴

۳۸ Burke, R., *Hybrid Recommender Systems: Survey and Experiments*. User Modeling and User-Adapted Interaction, ۲۰۰۲. ۱۲(۴): p. ۳۷۰-۳۳۱

۳۹ Balabanović, M., *Exploring Versus Exploiting when Learning User Models for Text Recommendation*. User Modeling and User-Adapted Interaction, ۱۹۹۸. ۸(۲-۱): p. ۱۰۲-۷۱

۴۰ Rocchio, J.J., *Relevance feedback in information retrieval*, in *The SMART Retrieval System: Experiments in Automatic Document Processing*, G. Salton, Editor. ۱۹۷۱ Prentice-Hall, Englewood Cliffs NJ. p. ۳۲۳-۳۱۳

۴۱ Pazzani, M.J., *A Framework for Collaborative, Content-Based and Demographic Filtering*. Artif. Intell. Rev., ۱۹۹۹. ۱۳(۶-۵): p. ۴۰۸-۳۹۳

۴۲ Littlestone, N. and M.K. Warmuth, *The Weighted Majority Algorithm*. Information and Computation, ۱۹۹۴. ۱۰۸(۲): p. ۲۶۱-۲۱۲

۴۳ Condiliff, M., et al., *Bayesian Mixed-Effects Models for Recommender Systems*, in *Proceedings of the ACM SIGIR Workshop on Recommender Systems: Algorithms and Evaluation*. ۱۹۹۹

۴۴ Schwab, I., A. Kobsa, and I. Koychev. *Learning User Interests through Positive Examples using Content Analysis and Collaborative Filtering*. ۲۰۰۱; Available from: citeseer.ist.psu.edu/schwab/learning.html.

۴۵ Shinde, S.K. and U. Kulkarni, *Hybrid personalized recommender system using centering-bunching based clustering algorithm*. Expert Systems with Applications, ۲۰۱۲. ۳۹(۱): p. ۱۳۸۷-۱۳۸۱

۴۶ عاطفه، ق. and س. میرعلی، سیستم توصیه گر خدمات گردشگری بر اساس تحلیل مکان، in دومین کنفرانس ملی رویکردهای نو در مهندسی برق و کامپیوتر. ۱۳۹۵.

۴۷ وحید، س. and ک. محمدجواد، یک روش توصیه گر ترکیبی مبتنی بر موبایل برای گردشگران انفرادی و گروهی، in کنفرانس بین المللی توسعه با محوریت کشاورزی، محیط زیست و گردشگری. ۱۳۹۴.

۴۸ شیماء، آ. and ه. علیرضا، بهبود سیستم توصیه گر مقالات علمی با استفاده از استخراج پروفایل کاربر در داده های کلان، in کنفرانس بین المللی مهندسی و علوم کاربردی. ۱۳۹۴.

۴۹ الهام نیکبخت، م.، استفاده از تکنیک های مبتنی بر تیوری گراف برای افزایش عملکرد سیستم های توصیه گر، in اولین همایش ملی کاربردهای سیستم های میکاترونیکی و رباتیکی. ۱۳۹۵.

۵۰ فهیمه محمدی، ه. and م. پروانه، یک سیستم توصیه گر برای مزایده الکترونیکی کالا با استفاده از رویکرد مدل سازی موضوعی، in دومین همایش ملی پژوهش های کاربردی در علوم کامپیوتر و فناوری اطلاعات. ۱۳۹۳.

۵۱ شاهچراغی، م.ذ.ج.ن.، طراحی یک سیستم توصیه گر با استفاده همزمان شباهت کاربران و اقلام. ۱۳۹۱، وزارت علوم، تحقیقات و فناوری - دانشگاه شیراز - مرکز آموزش الکترونیکی.

۵۲ علیزاده، م.و.ج.ه.، ارائه راهکاری جدید مبتنی بر اعتماد و معنا در سیستم های توصیه گر. ۱۳۹۲، دانشگاه امام رضا علیه اسلام - دانشکده مهندسی کامپیوتر.