# مدیریت احساسات مشتریان در رسانه های اجتماعی جهت بهبود تبلیغات و افزایش خرید

لیلا خواجه‌وند[۱]، عباس طلوعی اشلقی[۲،*]، مرتضی موسی‌خانی[۳]

[۱]دانشجوی کارشناسی ارشد، گروه مدیریت فناوری اطلاعات، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران

[۲]استاد، گروه مدیریت صنعتی، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران (عهده‌دار مکاتبات)

[۳]استاد، گروه مدیریت دولتی، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران

تاریخ دریافت: اسفند ۱۴۰۱، اصلاحیه: فروردین ۱۴۰۲، پذیرش: تیر ۱۴۰۲

**چکیده :**

امروزه شبکه‌های اجتماعی توجه ویژه‌ای را به خود جلب نموده‌اند. در شبکه‌های اجتماعی گوناگون، کاربران دائما در حال ابراز نظرات عمومی و همچنین خصوصی خود درباره‌ی موضوعات مختلف هستند. توییتر یکی از این شبکه‌های اجتماعی است که در دهه اخیر محبوبیت بسیاری یافته است. این شبکه اجتماعی روشی سریع و موثر برای تحلیل احساسات، دیدگاه‌ها و انتقادات مشتریان برای موفقیت در بازار را به سازمان‌ها ارائه می‌دهد. تحلیل احساسات یا عقیده کاوی فرآیندی است که در آن نظرات، احساسات و نگرش افراد در ارتباط با موضوعی خاص استخراج می‌شود. پژوهش‌های زیادی در رابطه با تحلیل احساسات بر روی نظرات کاربران، مستندات و مقالات انجام شده است. تحلیل بر روی موارد بیان شده تفاوت عمده‌ای با داده های توییتر دارد، به این سبب که توییت‌های توییتر محدودیت ۲۸۰ کاراکتری دارند و کاربران را وادار به بیان احساسات خود به صورت فشرده و کوتاه می‌نمایند. بهترین نتایج به دست آمده در طبقه بندی احساسات از تکنیک‌های یادگیری ماشین مثل بیز ساده و ماشین بردار پشتیبان حاصل شده است.

در این پژوهش به ارائه روشی برای تحلیل احساسات در شبکه‌های اجتماعی پرداخته می‌شود. در این راستا سعی شده با تمرکز بر مراحل پیش پردازش داده‌ها و انتخاب ویژگی، طبقه‌بندی متن توسط روش بیز را تا حدودی بهبود بخشیم. به عبارتی، با تعریف این مسئله به صورت یک مسئله کلاس‌بندی باینری بر اساس خصیصه‌های پیشنهادی به تحلیل احساسات کاربران پرداخته می‌شود. مسئله کلاس‌بندی با استفاده از جدیدترین دستاوردهای حوزه یادگیری ماشین فرموله و حل شده است. برای ارزیابی روش پیشنهادی در این رساله از سناریو مجموعه دادگان توییتر می‌باشد. روش پیشنهادی با سایر روش‌های طبقه‌بندی مقایسه می‌شود. بهترین عملکرد را از خود نشان داده است.

**واژه‌های اصلی**: شبکه‌های اجتماعی، تحلیل احساس، محتوا ،یادگیری ماشین، کاربران.

## ۱- مقدمه

رسانه‌های اجتماعی با سرعتی تصاعدی در حال گسترش هستند و در حال حاضر آن ها فرصتی بی‌نظیر برای کسب وکارها ایجاد کرده‌اند. فیس‌بوک، به‌عنوان بزرگ‌ترین شبکه اجتماعی در حال حاضر بیش از ۹۰۰ میلیون کاربر فعال دارد بیش از ۷۰٪ این کاربر به‌صورت روزانه وارد این سایت شده و در حدود ۹۴۰ میلیارد دقیقه را صرف بازدید از این سایت می‌نمایند. همین وضعیت برای دیگر رسانه‌های اجتماعی نظیر توئیتر صادق است. با شیوع بیماری کرونا و همه‌گیر شدن آن، کاربران از صنایع مختلف انتظار دارند سرویس و محصولات خود را در بستر آنلاین به فروش برسانند و آن ها بتوانند در هر شرایطی به محصولات مورد نیاز خود دسترسی داشته باشند تحقیق کن ۲۰۲۰ بیش از ۹۸٪ کاربران آمریکایی رسانه‌های اجتماعی عقیده دارند که شرکت‌ها بهتر است خدمات خود را در قالب رسانه‌های اجتماعی ارائه نمایند و بیش از ۸۵٪ معتقدند که کارخانجات باید تعامل بهتری از طریق رسانه‌های اجتماعی با مشتریان خود داشته باشند[۱۳]. این آمار و ارقام نشان دهنده این موضوع است که چقدر رسانه های اجتماعی و ارائه سرویس برروی این سامانه ها برای مشتریان اهمیت دارد و کسب و کارها برای افزایش سهم بازار و پیشرو شدن در رقبا بر روی این پروژه ها سرمایه گذاری می کنند. استفاده از رسانه‌های اجتماعی برای فرآیندهای مرتبط با مشتریان، که به آن مدیریت ارتباط با مشتری اجتماعی گفته می‌شود، بخش بزرگ‌تری از استفاده صرف کسب‌و کارها از رسانه‌های اجتماعی را شامل می‌شود. در اکثر منابع، مدیریت ارتباط با مشتری به چهار شکل کاملاً متفاوت مورد توجه قرارگرفته است: استراتژیک، عملیاتی، تحلیلی و تعاملی. در نگاه استراتژیک به مدیریت ارتباط با مشتری، CRM یک استراتژی محوری و کلیدی در کسب‌و کار است و قرار است مشتریان سود ده را جذب کسب‌و کار کرده و آن‌ها را برای ما حفظ کند.

*edu.myresearch@gmail.com

در نگاه عملیاتی به مدیریت ارتباط با مشتری، هر فرایندی که به‌نوعی با مشتری مرتبط است. با استفاده از سامانه‌های نرم‌افزاری، به ابزارهای خودکار تجهیز می‌شود. بازاریابی، فروش و خدمات مشتریان، ازجمله این فرایندها هستند. در نگاه تحلیلی به مدیریت ارتباط با مشتری، CRM ابزاری برای تحلیل هوشمندانه داده‌ها و اطلاعات مربوط به مشتری باهدف‌های استراتژیک و یا عملیاتی است. در نگاه تعاملی به مدیریت ارتباط با مشتری، CRM از فناوری برای مدیریت مرزهای سازمان، چه در رابطه با مشتری و چه شرکای تجاری استفاده می‌شود و هر نوع داده و اطلاعاتی که از مرزهای سازمان عبور می‌کند توسط این سیستم مدیریت می‌شود. هدف این سیستم ایجاد ارزش برای مشتریان و همکاران سازمان است. با هر دیدگاهی که به CRM توجه شود رسانه‌های اجتماعی می‌تواند جایگاه مهمی داشته باشد. یکی از زیر بخش‌های مهم برای تحقق SCRM، تحلیل احساسات و علایق مشتریان می‌باشد. با توجه به وسعت و قابلیت‌های تعامل کاربران، این بخش نیازمند تحول بنیادین با توجه به نیازهای روزافزون فضای کسب‌و کار و روش‌های بازاریابی می‌باشد[۱].

تحلیل احساسات و علایق مشتریان قلب هر سامانه مدیریت ارتباط با مشتری محسوب می‌شود. مسئله مطرح در این پروژه، ارائه روشی برای تحلیل احساسات مشتریان در سامانه‌های مدیریت ارتباط با مشتری اجتماعی است. این روش قادر خواهد بود با بررسی پست‌ها و فعالیت‌های کاربران در رسانه‌های اجتماعی به شناسایی احساسات، اقدام به گروه‌بندی مشتریان بر حسب علایق و احساسات نماید. همین‌طور این روش قادر است گزارش‌گیری کارآمدی از تغییرات نظرات و احساسات مشتریان در بازه‌های زمانی مختلف را ارائه نماید. تحلیل احساس یکی از فعالترین حوزه های پژوهشی در پردازش زبان طبیعی است که به دلیل اهمیت آن در کسب وکار و جامعه به خارج از علم کامپیوتر مانند مدیریت و علوم اجتماعی نیز گسترش یافته است. اهمیت درحال رشد تحلیل احساسات با رشد رسانه‌های اجتماعی مانند بررسی‌ها، انجمن‌ها، بحث، وبلاگ‌ها، میکروبلاگ‌ها، توییتر و شبکه‌های اجتماعی همخوانی دارد. سیستم‌های سنجش در هر حوزهٔ تجاری و اجتماعی کاربرد دارند؛ زیرا نظرات در همه احساسات تقریبا فعالیت‌های انسانی متمرکز هستند و تأثیرات کلیدی رفتارهای ما به شمار می‌روند. کسب وکارها با تحلیل احساسات کاربران خود می‌توانند سرویس یا محصول خود را ارتقا دهند یا محصولات و سرویس‌های جدید را به بازار ارائه دهند. از طرفی تحلیل احساسات کاربران می‌تواند به تصمیمات مدیران سازمان جهت ارائه تبلیغات موثرتر، بهبود پرموشن‌ها و جذب کردن کاربران کمک کند. همچنین در صورتی که مشتریان نسبت به محصولات یا سرویس‌های فعلی احساس ناخوشایندی داشته باشند می‌توان با دخیل کردن نظرات آن‌ها این حس ناخوشایند را به احساس مثبت و افزایش رضایتمندی آن‌ها تبدیل کرد.

در بخش دوم به ارائه کارهای مرتبط پرداخته می‌شود. در بخش سوم کار پیشنهادی مطرح می‌شود و در نهایت در بخش چهارم به ارزیابی روش پیشنهادی پرداخته می‌شود.

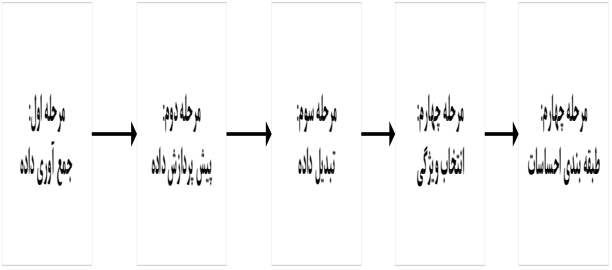
## ۲- پیشینه تحقیق

بیشتر مطالعات مرتبط با تحلیل احساس در گذشته بر اساس الگوریتم‌های یادگیری با نظارت انجام گرفته است که نیاز به تهیه داده برچسب خورده دارند. مدل بیز ساده، ساده‌ترین و پراستفاده‌ترین الگوریتم احتمالاتی برای دسته‌بندی است و بر مبنای قضیه بیز کار می‌کند. این مدل احتمالات پسین رویدادها را محاسبه کرده و برچسبی که بیشترین احتمال پسین را دارد به رویداد نسبت می‌دهد. دسته‌بندی پرکاربردآنتروپی بیشینه است که کار دسته‌بندی را می‌توان با آن انجام داد. این روش بر پایه مدل نمایی و اصل حداکثر آنتروپی است. استفاده از این روش تجربه‌های موفقی در کار پردازش زبان طبیعی از جمله در تحلیل احساس به ارمغان آورده است. این روش در اکثر (و نه در همه) مواقع نسبت به مدل بیز ساده برتری دارد.ماشین بردار پشتیبان کار دسته‌بندی اسناد بر مبنای موضوعات مشابه بسیار مفید است. روش SVM یک مدل یادگیری بانظارت است که کار آن دسته‌بندی کردن اشیا در کلاسهای مختلف با استفاده از ویژگی‌های استخراج شده است. این دسته‌بندی با ایجاد ابرصفحه‌ای میان نمونه‌های هر کلاس و حداکثرکردن فاصله نمونه‌ها از این صفحه صورت می‌گیرد. برتری این روش نسبت به دیگر روشهای مطرح یادگیری ماشین آن است که در مورد داده‌های ورودی پیش فرضی ندارد و به جای تکیه بر ارزش‌های احتمالاتی، سعی دارد تا بهینه‌ترین دسته‌بندی را با داده‌های موجود انجام دهد و نتایج به دست آمده از آن در تحلیل احساس برتری محسوسی به دیگر روشهای یادگیری ماشین در زبان انگلیسی دارد[۸]. در سالهای اخیر روشهای یادگیری عمیق به خصوص شبکه‌های (RNN) در تحلیل احساس برای زبان انگلیسی چینی و آلمانی در میان زبانهای مختلف، با استفاده از بردارهای مختلف نمایش کلمات کاربرد زیادی داشته است. آنها برای درک و کنترل ترکیب معنایی در کارهای پیچیده‌ای مانند تحلیل احساس مفید هستند. شبکه‌های RNN برای داده‌هایی با قابلیت تبدیل به مقادیر متوالی به کار می‌روند و با استفاده از ایده اشتراک‌گذاری پارامترها برای رسیدن به وزن‌های مطلوب، توانایی پردازش توالی‌هایی با طول‌های متفاوت را دارند. با وجود این که استفاده از آنها در تحلیل احساس برای زبان انگلیسی با نتایجی بهتر از روشهای یادگیری بانظارت همراه بوده است. با رشد ساختار شبکه‌های RNN ،ابعاد ماتریس‌ها در مرحله بازپخش به صورت توانی رشد می‌کنند و در عمل استفاده از آنها غیر ممکن می‌شود[۱۱]. شبکه‌های پیچشی که کولوبرت و دیگران در ابتدا برای کاربرد در بینایی رایانه‌ای ارائه کرده‌اند، اخیراً در بسیاری از کارهای پردازش زبان طبیعی مانند تجزیه نحوی، تجزیه سطحی، برچسب‌زنی نقش معنایی مورد استفاده قرار گرفته است. استفاده از

شبکه‌های پیچشی قطعه‌بندی در تحلیل احساس نیز برای زبان‌ها با منابع فراوان مورد استفاده قرار گرفته و باعث بهبود قابل توجه دقت و کاهش زمان مرحله آموزش نسبت به دیگر روشهای یادگیری عمیق شده است. پژوهش‌های حوزه تحلیل احساس در زبان فارسی معمولاً یا با استفاده از روشهای مبتنی بر قاعده هستند یا مبتنی بر پیکره. برای بهبود نتایج معمولاً از پیشپردازش نظرات و ویژگی‌های لغتنامه استفاده شده است. بصیری و همکاران[۶] یک چارچوب مبتنی بر لغتنامه ارائه کردند که به صورت بدون نظارت با استفاده از قواعد از پیش تعیین شده و لغتنامه تعریف شده جهت‌گیری متون محاوره را تشخیص می‌دهد. استفاده از SVM برای تحلیل احساس در زبان فارسی بر روی داده مربوط به نقد فیلم، منجر به نتایج بهتری نسبت به روشهای دیگر یادگیری ماشین شده است. بازدهی این روشها وابسته به کیفیت برچسبدهی در پیکره‌ها و شیوه گزینش ویژگی‌ها پیش از شروع کار دسته‌بندی است. روشنفکر و همکاران[۷] برای اولین بار از شبکه‌های عصبی LSTM برای تشخیص احساس متون فارسی استفاده کردند و توانستند نسبت به روشهای یادگیری سنتی نتایج بهتری داشته باشند، اما این نوع شبکه ها برای آموزش نیاز به داده‌های خیلی زیادی دارند. همچنین آنها در کار خود فقط دو سطح از احساس را در نظر گرفتند و از جاسازی ساده کلمات استفاده کردند. اینترنت از قابلیت‌ها و امکانات زیادی برای ایفای کارکرد حوزه عمومی برخوردار است. شبکه اینترنت امکاناتی در اختیار مردم جوامع می‌گذارد تا در فضایی مناسب به گفتگوی آزاد و برابر با هم بپردازند و در نتیجه فرآیندهای گفتگو و مباحثه، به نقطه‌نظرهای مشترکی درباره مسائل سیاسی و اجتماعی دست یابند و به افکار عمومی شکل دهند. (میناوند، محمدقلی) مجموعه داده‌های نشات گرفته از تحقیق بر روی شبکه‌های اجتماعی در زمینه‌های بسیاری مانند جامعه شناسی و روانشناسی باارزش هستند. اما حمایت از دیدگاه فنی به اندازه کافی دور است، و به روشهای خاص فوری نیاز دارند. (اکسیو وانگ و همکار) تحقیق روانشناسی برای تشخیص کاربران افسرده شبکه‌های اجتماعی[۱۶] را با استفاده از داده‌کاوی انجام داده است. اخیرا ایده استفاده از تجزیه و تحلیل احساسات کاربران شبکه‌های اجتماعی برای بهبود عملکرد برنامه‌های کاربردی در وب سایت‌های خرید آنلاین توجه پژوهشگران را به خود جلب کرده است. وب سایت های خرید آنلاین بطور گسترده‌ای بررسی در مورد یک محصول ارائه می‌کنند و مشتریان می‌توانند استفاده کنند[۴]. در اینجا تجزیه و تحلیل احساسات مشتریان در باره هر محصول انجام می‌شود. در بررسی از سایت‌های شبکه اجتماعی قوانین برای احساسات مثبت یا منفی بسته به نمره کلی آن، با کمک SentiWordNet محاسبه می‌شود .تجزیه و تحلیل تمایلات شامل تشخیص ذهنیت و احساسات موجود در نظرات است. نظرات عبارات توصیف عواطف و احساسات مردم در مورد یک موضوع، نهاد و یا رویداد است با استفاده از تکنیک‌های زبان طبیعی انجام توصیف می‌کنند.[۱۴] تقاضا برای اطلاعات فرابری شده از منابع متنی به طور فزاینده‌ای در حال افزایش است. ذات غیرساخت یافته‌ی این متون، اعمال همان روشهایی را که ما در مورد دیتابیس‌ها بکار می‌بریم، غیر ممکن می‌سازد. کاربردهای مهمی را که از پردازش متون مورد انتظار است، بررسی می‌کنیم. به اینگونه پردازش‌ها که روی متون اعمال می‌شود، متن‌کاوی می‌گوییم[۵]. وظیفه‌ی اصلی عقیده‌کاوی طبقه‌بندی قطبیت است. طبقه‌بندی قطبیت وقتی اتفاق می‌افتد که یک تکه متن که یک عقیده در مورد یک موضوع را بیان کند به یکی از دو احساس متضاد تقسیم شود. نظراتی مثل «موافق» در مقابل «مخالف»، «دوست داشتن» در مقابل «دوست نداشتن» مثال‌هایی از طبقه‌بندی عقاید هستند. طبقه‌بندی قطبیت بیانات موافق و مخالف را تشخیص می‌دهد و به تولید ارزیابی‌های معتمدتر کمک می‌کند[۲]. بسیاری از شرکت‌ها از عقیده کاوی و تحلیل احساسات به عنوان جزئی از تحقیقاتشان استفاده می‌کنند. مثلاً شرکت‌ها از عقیده کاوی برای ساخت و نگه‌داری نظرات استفاده می‌کنند. سیستم‌های آن‌ها به طور مداوم اطلاعات را از وب مثل نظرات محصولات، دریافت برند و مسائل سیاسی جمع‌آوری می کند. دیگر سیستم‌ها نیز ممکن است از عقیده کاوی و تحلیل احساسات به عنوان یک فناوری زیر مؤلفه برای بهبود مدیریت روابط مشتری و سیستم توصیه‌گر از طریق بازخوردهای مثبت و مشتریان استفاده کنند. به طور مشابه، عقیده کاوی و تحلیل احساسات ممکن است شعله‌ها (زبان خصومت‌آمیز و گرمای اضافی) را در روابط اجتماعی شناسایی و حذف کنند[۱۲] .طبقه‌بندی بیان احساسات بر اساس معنای آن‌ها و دانش قبلی تمایلات معنایی نامیده می‌شود. با وجود اینکه تحلیل نحوی نقشی کلیدی در طبقه‌بندی اسناد بازی می‌کند اما این برای استخراج مفاهیم از متن فقط از طریق نحو کافی نیست. معیارهای تئوری اطلاعات و دانش معنایی یک سلسه مراتب را با استفاده از WordNet ترکیب کردند تا مفاهیم را به طور اتوماتیک از متن استخراج کنند[۱۵] دنکه نقش مدل بر پایه قوانین و یادگیری ماشین در یک دامنه چندگانه، بر روی سناریو طبقه‌بندی تست کرده است نتایج آنها نشان می‌دهد که رویکرد مبتنی بر واژگان، که با استفاده از SentiWordNet ساخته شده است، دقت آن در مقایسه با روش‌های یادگیری ماشین محدودتر است[۱۰].

## ۳- روش پیشنهادی

در این قسمت به ارائه روش پیشنهادی برای تحلیل احساسات مشتریان در رسانه اجتماعی خواهیم پرداخت. این روش شامل پنج مرحله به صورت شکل می‌باشد. این مراحل شامل جمع‌آوری داده، پیش پردازش داده، آماده‌سازی داده، برچسب‌گذاری کلمات، شناسایی خصیصه‌های مرتبط با دامنه و خوشه‌بندی کلمات مرتبط با احساسات می‌باشد. این مراحل به صورت تفصیلی در ادامه بخش تشریح خواهد شد.




**شکل (۱) : روش پیشنهادی**

*مرحله اول ،ورودی نظرات کاربران*

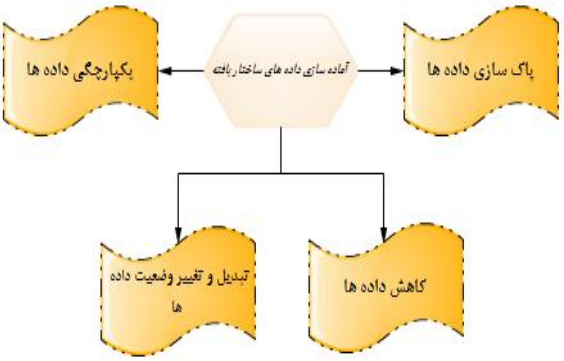
کاربران با ثبت‌نام بر روی شبکه‌های اجتماعی، فعالیت خود را بر بستر این شبکه‌ها آغاز می‌کنند. این فعالیت‌ها شامل تولید محتوا توسط کاربران، پیگیری علاقمندی هایشان، عضو شدن در گروه‌های متنوع و غیره می‌باشد. تمامی فعالیت‌های انجام شده توسط این کاربران در پایگاه داده ذخیره می‌شود، بنابراین انواع مختلفی از داده‌های متنی یا غیر متنی، داده‌هایی با ساختارهای متنوع، داده‌های غیر دقیق و نادرست وجود دارد. در مراحل بعدی سعی در تمیز کردن داده‌ها می‌شود. در مرحله نخست به جمع‌آوری داده پرداخته می‌شود. داده‌هایی می‌توانند از رسانه‌های مختلف جمع‌آوری شود. در این پایان‌نامه داده توییتر انتخاب شده است. برای جمع‌آوری داده‌های توئیتر چندین ابزار کاربردی متفاوت، وجود دارد. ۱) برنامه کاربردی جستجوی توئیتر، ۲) برنامه کاربردی جریان توئیتر، ۳) ابزار آتش‌نشانی توئیتر.

*مرحله دوم، پیش پردازش داده‌ها: جداسازی جملات*

مشتریان با ثبت نام بر روی رسانه‌های اجتماعی، فعالیت‌های خود را بر بستر این شبکه‌ها آغاز می‌کنند. این فعالیت‌ها شامل تولید محتوا توسط کاربران، پیگیری علاقمندی‌هایشان، عضو شدن در گروه‌های متنوع، ارائه نظرات مثبت و منفی خود نسبت به یک محصول یا سرویس خاص می‌باشد. تمامی فعالیت‌های انجام شده توسط این مشتریان در پایگاه داده ذخیره می‌شود، بنابراین انواع مختلفی از داده‌های متنی یا غیر متنی، داده‌هایی با ساختارهای متنوع، داده‌های غیر دقیق و نادرست وجود دارد. در این مرحله برای از بین بردن عدم این ناسازگاری‌ها، کمبودهای یکپارچگی و بهبود تحلیل نتایج به آماده سازی داده‌ها پرداخته می‌شود، برای این کار داده‌ها، به سه دسته ساختاریافته، نیمه ساختاریافته و غیر ساختاریافته تقسیم می‌شود.

*-آماده‌سازی داده‌های ساختاریافته*

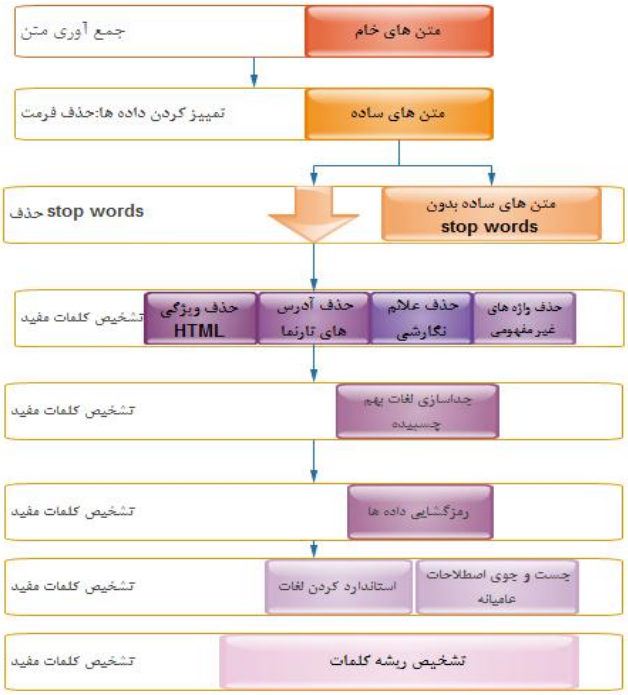
داده های ساختاریافته از درجه بالای سازمان‌دهی برخوردار هستند. برای آماده‌سازی داده‌های ساختار یافته چهار گام اصلی که در شکل ۲ مشاهده می‌شود، طی خواهد شد.




**شکل (۲): آماده‌سازی داده‌های ساختار یافته**

*- آماده‌سازی داده‌های نیمه ساختاریافته و غیر ساختار یافته*

داده‌های نیمه‌ساختاریافته شکلی از داده‌های ساختاریافته‌ای هستند که از ساختار رسمی از جداول و مدل‌های داده‌ای وابسته به پایگاهِ داده‌های رابطه‌ای مطابقت نمی‌کنند، اما با این وجود شامل برچسب‌ها یا علامت‌ها و شاخص‌هایی هستند که عناصر معنایی را از یکدیگر جدا می‌کنند و سلسله مراتبی از رکوردها و فیلدها را بین داده‌ها ایجاد می‌کنند. داده‌های غیر ساختار یافته به داده‌هایی گفته می‌شود که از هیچ مدل از قبل تعریف شده ای تبعیت نمی‌کنند مثالی از داده‌های ساختاریافته متن‌های سنگین می‌باشد.




**شکل (۳) : آماده‌سازی داده‌های نیمه ساختار یافته و غیر ساختار یافته**

*مرحله سوم: تبدیل داده*

در این بخش داده ها را به فرمتی مناسب تبدیل کرده و آن‌ها برای مرحله بعدی آماده می‌شوند.

گاهی داده‌های خامی که برای تحلیل داریم مناسب گروهی از آزمونهای آماری نیستند و برای این‌که بتوانیم از این دسته آزمون‌های آماری استفاده کنیم و همچنین دقت تحلیل را بالا ببریم باید در داده‌های خام تغییراتی ایجاد کنیم. یکی از این تغییرات، تبدیل داده ها نام دارد. تبدیل داده‌ها، روش‌هایی ریاضی است که برای تعدیل متغیرهایی به‌کار می‌رود که از مفروضه‌های آماری نرمال بودن، خطّی بودن و یکسانی پراکندگی پیروی نمی‌کنند یا الگوهایی با داده‌های پرت غیرمعمول دارند

در مجموع زمانی که پیش شرط‌های آزمون‌های چندمتغیره برقرار نباشد، باید داده‌های به دست آمده را تبدیل کنیم تا امکان استفاده از برخی آزمون‌های مدنظر (عموما پارامتریک) فراهم شود.

در ابتدا باید میزان تخطّی و تفاوت داده‌ها از پیش‌فرض‌های ذکر شده را تعیین کرد و در صورتی که پیش‌فرض‌ها یا پیش شرط های آماری به دست آمده دارای تفاوت قابل اعتنایی با مقدار معیار باشند از روش تبدیل داده ها استفاده کرد. تبدیل داده‌ها با هدف تعدیل متغیرها از جنبه علمی روشی پذیرفته شده است. البته زمانی که اختلاف داده‌ها با پیش‌فرض‌های آماری اندک باشد و به طور تقریبی مفروضات آماری برقرار باشد می‌توان از تبدیل داده‌ها صرف نظر کرد.

باید توجه داشت که تبدیل داده‌ها تا اندازه‌ای مانند شمشیر دولبه است. حسن این روش این است که می تواند دقت معنی‌داری تحلیل‌های آماری را افزایش دهد و عیب آن این است که ممکن است تفسیر داده‌ها را دشوارتر کند. در نتیجه باید از روش تبدیل داده‌ها به شیوه‌ای مدبرانه استفاده کرد.

دشوارکردن تفسیر داده‌ها بدین معناست که وقتی داده‌ها را تبدیل می‌کنیم، مقدار حداقل و حداکثر و شیوه توزیع متغیر و تمامی شاخص‌های میانگین و انحراف استاندارد تغییر می‌کند و با حالت معمول و عادی تفاوت پیدا می‌کند. مثلا اگر سن افراد که به صورت کمّی (نسبی) سنجیده شده است را به توان دو برسانیم شاخص های آماری سن افراد تغییر می‌کند و با سن‌های غیر عادی مثل ۲۵۰ ، ۳۰۰ و غیره مواجه می‌شویم. یا وقتی متغیری مانند اعتماد اجتماعی داریم و با ۱۰ سوال این متغیر را سنجیدیم و دامنه میانگین این متغیر بین ۱ تا ۵ باشد، لگاریتم گرفتن از این متغیر دامنه نمرات را تغییر می‌دهد و توضیح و تفسیر متغیر را با مشکل مواجه می‌کند. یکی از راه‌های رفع این مشکل این است که هنگام گزارش یافته‌های توصیفی و شاخص‌های آماری (مانند میانگین، انحراف استاندارد و مقدار حداقل و حداکثر)؛ یافته‌ها و شاخص‌های آماری را هم به صورت عادی (قبل از تبدیل داده‌ها) و هم بعد از تبدیل داده‌ها گزارش کنیم.

*مرحله چهارم: انتخاب ویژگی‌ها*

در این مرحله به کلمات وزن داده می‌شود. کلمات با توجه به وزنی که در این مرحله و با استفاده از معیارهای تعریف شده در مرحله بعدی دریافت می‌کنند، مشخص می‌شوند. نام دیگر این مرحله استخراج ویژگی است. یعنی در این مرحله ویژگی‌های موردنظر را مشخص می‌کنیم تا در مرحله بعدی انتخاب شوند.

در مرحله انتخاب ویژگی، ویژگی‌هایی که معیارهای تعریف شده را ماکزیمم می‌کنند، انتخاب خواهند شد. این روش با کاهش نمونه‌ها سعی در ایجاد یک دسته‌بندی مناسب دارد.

*-تشخیص توییت‌های مورد علاقه کاربر*

اولا بایستی مرتبط بودن توئیت را با علایق کاربر پیدا کرد. به منظور تعیین میزان ارتباط یک توئیت با موضوعات مورد علاقه کاربر در اغلب کارهای انجام شده از روش TF-IDF و کسینوس زاویه بین بردار کلمات $\boldsymbol{tw_i}$ و $\boldsymbol{p_{u_i}}$ استفاده می‌گردد DF. این معیار تعداد تکرار ویژگی‌ها را در متون بررسی می‌کند. در این معیار یک میزان آستانه تعریف می‌شود. آن ویژگی‌هایی که تعداد تکرارشان بیشتر یا کمتر از میزان آستانه است، حذف می‌شوند. تکرار زیاد یک ویژگی در اینجا احتمال انتخاب آن ویژگی را بیشتر می‌کند. ازجمله ویژگی‌های این روش، مقیاس‌پذیری، سادگی و تأثیر آن است. معیار دیگر مورد استفاده "بهره اطلاعاتی" می‌باشد. این روش با استفاده از آنتروپی قابل محاسبه است. و آن ویژگی‌هایی که میزان gain بیشتری دارند را انتخاب می‌کند. برای افزایش دقت روش پیشنهادی از آنتولوژی استفاده می‌گردد. آنتولوژی برای مدل‌سازی شرایط در یک دامنه مورد علاقه و همچنین روابط میان این شرایط استفاده می‌شود. مهم‌ترین بخش آنتولوژی نقش کلیدی آن در توسعه وب معنایی است. تحلیل احساسات با استفاده از آنتولوژی به این صورت است که از آنتولوژی جهت استخراج مفاهیم مرتبط استفاده می‌شود. این بخش به بخش استخراج ویژگی اعمال شده است (در شرح مسئله، بخش استخراج ویژگی‌ها توضیح داده‌شده است). در واقع آنتولوژی یک نوع معناشناسی انجام می‌دهد.کلمات یکسان، ممکن است معانی مختلف و کلمات مختلف، ممکن است معانی یکسان داشته باشند. آنتولوژی کلماتی را که از نظر مفهوم به آن ویژگی‌ها نزدیک‌ترند، مشخص می‌کند.

*-بررسی تعداد لایک ، تعداد کامنت‌ها و تعداد ذکر شدن‌ها:*

لایک شدن یک پست توسط کاربر نشان دهنده آن است که کاربر نسبت به آن موضوع حساسیت بیشتری دارد. بنابراین تعداد لایک می‌تواند پارامتر مهمی در تشخیص احساس کاربر باشد. از طرفی اگر تعداد منشن‌ها زیاد باشد یعنی کاربران تمایل دارند موضوع را با سایر دوستان خود به اشتراک بگذارند. و همین‌طور تعداد کامنت ها نشان دهنده آن است که موضوع برای کاربر جذاب بوده و به ارائه ایده خود پرداخته است.

*مرحله پنجم: کلاس‌بندی*

شبکه‌های عصبی پیچشی رده‌ای از شبکه‌های عصبی عمیق هستند که معمولاً برای انجام تحلیل‌های تصویری یا گفتاری در یادگیری ماشین استفاده می‌شوند. این شبکه یک الگوریتم یادگیری عمیق است که تصویر ورودی را دریافت می‌کند و به هر یک از اشیا / جنبه‌های موجود در تصویر میزان اهمیت (وزن‌های قابل یادگیری و بایاس) تخصیص می‌دهد و قادر به متمایزسازی آن‌ها از یکدیگر است. در الگوریتم ConvNet در مقایسه با دیگر الگوریتم‌های دسته‌بندی به پیش‌پردازش کمتری نیاز است. در حالی‌که فیلترهای روش‌های اولیه به صورت دستی مهندسی شده‌اند، شبکه عصبی پیچشی، با آموزش دیدن به اندازه کافی، توانایی فراگیری این فیلترها / مشخصات را کسب می‌کند. معماری ConvNet مشابه با الگوی اتصال نورون‌ها در مغز انسان است و از سازمان‌دهی قشر بصری در مغز الهام گرفته شده است. هر نورون به محرک‌ها تنها در منطقه محدودی از میدان بصری که تحت عنوان میدان تاثیر شناخته شده است پاسخ می‌دهد. یک مجموعه از این میدان‌ها برای پوشش دادن کل ناحیه بصری با یکدیگر هم‌پوشانی دارند.ConvNet قادر است به طور موفقی وابستگی‌های زمانی و فضایی را در یک تصویر با استفاده از فیلترهای مرتبط ثبت کند و همچنین، معماری فیلترگذاری بهتری را روی مجموعه داده تصویر به دلیل کاهش تعداد پارامترهای درگیر و استفاده مجدد از وزن‌ها انجام می‌دهد.

**۴- ارزیابی**

در این بخش به ارزیابی روش پیشنهادی که برحسب مجموعه دادگان تویتر است می پردازیم.

۴-۱ مجموعه دادگان

برای ارزیابی روش پیشنهادی بر طبق مراحل زیر عمل کرده و خروجی روش پیشنهادی با سایر الگوریتم های یادگیری ماشین مقایسه خواهد شد.برای جمع آوری داده در توییتر از روش جریان توییتر استفاده شده است. بر طبق این ابزار توییت های مورد نظر بر حسب موضوعوات محتلف جمع آوری شده است. برای این مقاله از سه موضوع مختلف ، کوید ۱۹ ، آتشسوزی در جنگل های استرالیا و بهار عرب استفاده شده است.

**جدول (۱) : جزئیات مجموعه داده**

| شماره | مجموعه دادگان | تعداد پست ها | تعداد کاربران | تعداد بازنشرها |
|---|---|---|---|---|
| مجموعه دادگان یک | کوید ۱۹ | ۲۳۴۵۷ | ۶۵۸۹۰ | ۶۵۴۷۸۰ |
| مجموعه دادگان دو | آتشسوزی در جنگل های استرالیا | ۴۵۰۰ | ۷۸۹۰۴ | ۵۶۷۸۹۰ |
| مجموعه دادگان سه | بهار عرب | ۵۴۹۸ | ۴۵۶۷۸۹ | ۴۷۸۹۰ |

**جدول (۲) : معیار ارزیابی**

| | نتایج پیش بینی | |
|---|---|---|
| نتایج واقعی | غلط negative | درست positive |
| درست positive | True Negative | True Positive |
| غلط negative | False Negative | False Positive |

**جدول (۳) : روابط روش‌های مختلف ارزیابی**

| | |
|---|---|
| رابطه ۱ | $\text{Precision} = \frac{tp}{tp + fp}$ |
| | |
| رابطه ۲ | $Recall = \frac{tp}{tp + fn}$ |
| رابطه ۳ | $F = ۲ \cdot \frac{precision \cdot recall}{precision + recall}$ |
| رابطه ۴ | $Accuracy = \frac{tp + tn}{tp + tn + fp + fn}$ |

دانشی که در مرحله یادگیری مدل تولید می‌شود، می‌بایست در مرحله ارزیابی مورد تحلیل قرار گیرد تا بتوان ارزش آن را تعیین نمود و در پی آن کارائی الگوریتم یادگیرنده مدل را نیز مشخص کرد. این معیارها را می‌توان هم برای مجموعه داده‌های آموزشی در مرحله یادگیری و هم برای مجموعه رکوردهای آزمایشی در مرحله ارزیابی محاسبه نمود.

در جدول ۴ مقادیر دقیق حساسیت روش‌های مختلف ارزیابی شده را نشان می‌دهد. در بررسی نتایج، نکته حائز اهمیت دقت بالای روش SASMپیشنهاد شده در مقایسه با دیگر روش‌هاست. SASM در برابر روشی مانند KNN بسیار کند است، با این وجود، دقت بالای این روش علی‌رغم کندی آن می‌تواند قابلیت‌های این روش را نشان دهد. دقت بالای روش SASM به ماهیت این روش برمی‌گردد. در میان روش‌های دیگر، روش SVM عملکرد مناسبی داشته است. علت این برتری در قدرت طبقه‌بندی این روش می‌باشد.

**جدول (۴) : مقایسه نرخ حساسیت روش‌های پیشنهادی با بقیه روش‌ها**

| معیار حساسیت | میانگین | مجموعه دادگان اول | مجموعه دادگان دوم | مجموعه دادگان سوم |
|---|---|---|---|---|
| **SASM** | ۰/۸۵۴۸۱۵۷۷۱ | ۰/۸۹۶۶۶۶۷۰۰ | ۰/۸۸۱۲۵۰۰۰ | ۰/۷۸۶۵۳۰۶۱۲ |
| **KNN** | ۰/۴۵۲۴۰۲۴۲۸۰ | ۰/۴۱۶۳۶۳۶۰۰ | ۰/۴۵۶۷۸۵۷۱۴ | ۰/۴۸۴۰۵۷۹۷۱ |
| نیوبیزین | ۰/۵۳۳۲۷۰۹۷۰ | ۰/۵۱۶۴۸۳۵۱۶ | ۰/۵۴۴۲۵۵۳۱۹ | ۰/۵۳۹۰۷۴۰۷۴ |
| **SVM** | ۰/۷۲۶۳۳۰۹۹۰ | ۰/۶۹۶۳۶۳۶۰۰ | ۰/۷۷۸۵۷۱۴۰۰ | ۰/۷۰۴۵۰۵۷۹۷۱ |

در جدول ۵، روش‌های پیشنهادی با سه روش دیگر از نظر معیار تشخیص مقایسه شده است. در این ارزیابی، میزان پیش‌بینی درست تصمیمات عدم بازنشر، در معیار تشخیص بسیار موثر است. همان‌طور که در جدول ۵ ، مشاهده می‌شود، SASM بهترین عملکرد را داشته است. در معیار تشخیص نیز، روش SVM به صورت میانگین، بهترین عملکرد را بعد از SASMبه خود اختصاص داده است. روش SASM توانسته است، موارد عدم بازنشر را بهتر از روش‌های دیگر، تشخیص دهد. در این بخش، روش KNN بدترین عملکرد را به خود اختصاص داده است.

**جدول (۵): مقایسه معیار تشخیص روش‌های پیشنهادی با بقیه روش‌ها**

| **معیار تشخیص** | میانگین | **مجموعه دادگان اول** | **مجموعه دادگان دوم** | **مجموعه دادگان سوم** |
|---|---|---|---|---|
| SASM | ۰/۷۳۸۱۶۸۷۹۰ | ۰/۶۶۰۷۳۱۷۰ | ۰/۷۸۹۶۹۷۰۰ | **۰/۷۶۴۰۷۷۶۷** |
| KNN | ۰/۳۹۴۱۴۸۵۲۷ | ۰/۳۵۹۳۴۰۷۰ | ۰/۳۹۵۳۲۷۱۰ | **۰/۴۲۷۷۷۷۷۸** |
| نیوبیزین | ۰/۵۴۹۶۹۱۵۴۳ | ۰/۵۱۰۳۱۲۵۰ | ۰/۵۶۵۵۹۱۳۹۸ | **۰/۵۷۳۱۷۰۷۳** |
| SVM | ۰/۵۹۱۶۶۸۵۵۷ | ۰/۴۵۹۱۸۳۷۰ | ۰/۵۸۵۷۱۴۳۰ | **۰/۷۳۰۱۵۸۷۳** |

در ادامه ارزیابی روش‌های پیشنهادی برای پیش‌بینی تصمیم بازنشر کاربر در مواجهه با یک پست، به بررسی معیار نرخ خطای روش‌های پیشنهادی در مقایسه با روش‌های دیگر می‌پردازیم. معیار دقت، یکی از پرکاربردترین معیارها برای ارزیابی روش‌های پیش‌بینی و کلاس‌بندی است. هر چه نسبت میزان پیش‌بینی درست تصمیم بازنشر، به پیش‌بینی‌های نادرست، بیشتر باشد، دقت روش مربوطه بالاتر خواهد بود. جزئیاتی بیشتری از نتایج مقایسه روش‌های پیشنهادی با دیگر روش‌ها از نظر معیار دقت، در جدول ۶ آورده شده است.

**جدول (۶): مقایسه نرخ دقت روش‌های پیشنهادی با بقیه روش‌ها**

| معیار دقت | میانگین | مجموعه دادگان اول | مجموعه دادگان دوم | مجموعه دادگان سوم |
|---|---|---|---|---|
| SASM | ۰/۷۶۶۶۷۵۷۲۶ | ۰/۶۸۵۸۵۷۱۰ | ۰/۸۶۱۵۳۸۵ | ۰/۷۵۲۶۳۱۵۷۹ |
| KNN | ۰/۴۰۶۳۶۶۱۲۷ | ۰/۳۴۳۶۷۸۲۰ | ۰/۴۸۵۷۱۴۳۰ | ۰/۳۸۹۷۰۵۸۸۲ |
| نیوبیزین | ۰/۶۴۵۳۸۲۱۷۳ | ۰/۶۸۲۱۲۱۲۱۲ | ۰/۶۹۱۷۵۲۵۷۷ | ۰/۵۶۲۲۷۲۷۳۰ |
| SVM | ۰/۷۰۷۷۳۰۴۲۶ | ۰/۵۹۵۲۳۸۱۰ | ۰/۷۸۸۷۳۲۴۰ | ۰/۷۳۹۲۲۰۷۷۹ |

همان طور که در جدول ۶ مشاهده می‌شود، روش SASM در این معیار، بهترین عملکرد را داشته است. بعد از این روش ماشین بردار پشتیان بهترین دقت را در بین روش‌های دیگر داشته است. یکی از دلایل برتری روش‌های پیشنهادی بر اساس معیار دقت در ماهیت این روش‌ها می‌باشد. در اجتماعات برخط ممکن است کاربران دچار تغییر سلایق شوند. این برتری روش پیشنهادی در دقت پیش‌بینی‌های انجام شده بسیار موثر است.

در جدول ۷، معیار F-measure روش‌های مختلف مقایسه شده است. همان‌طور که پیش‌تر اشاره شد، معیار F-measure بکارگرفته شده در این رساله، میانگین موزون دقت و بازیابی (حساسیت) می‌باشد ($\beta$=۱). همان‌طور که اشاره شد، میزان این معیار، ارتباط مستقیمی به دقت و بازیابی روش‌ها دارد. هر دو روش پیشنهادی، بهترین عملکرد را داشته‌اند. نکته جالب در این جدول، پایین بودن مقدار F- measure برای روش KNN است. با وجود دقت نسبتاً خوب این روش، با توجه به ضعف شدید این روش در بازیابی، بدترین عملکرد را در این معیار دارد.

**جدول (۷): مقایسه معیار F measure روش‌های پیشنهادی با بقیه روش‌ها**

| معیار F measure | میانگین | مجموعه دادگان اول | مجموعه دادگان دوم | مجموعه دادگان سوم |
|---|---|---|---|---|
| SASM | ۲/۳۸۷۷۸۰۶۱۲ | ۲/۳ | ۲/۲۸۳۷۷۵ | ۲/۵۷۹۵۹۱۸۳۷ |
| KNN | ۱/۵۱۳۳۸۴۱۹۷ | ۱/۶۹۹۰۹۰۹ | ۱/۱۸۵۷۱۴۳ | ۱/۶۵۵۳۴۷۳۹۱ |
| نیوبیزین | ۱/۵۳۱۶۱۲۱۵۲ | ۱/۵۹۷۰۱۱۲۹۹ | ۱/۴۴۱۱۲۴۱۲۷ | ۱/۵۵۶۷۰۱۰۳۰ |
| SVM | ۱/۷۸۵۸۹۴۵۱۷ | ۱/۷۱۴۲۸۵۷ | ۱/۴۴۸۲۷۵۹ | ۲/۱۹۵۱۲۱۹۵۱ |

## ۵-نتیجه‌گیری

تحلیل انتشار اطلاعات و نفوذ اجتماعی در شبکه‌های اجتماعی دارای کاربردهای بسیار زیادی در جهان واقعی دارد. یکی از مثال‌های کاربردی آن بیشنه‌سازی نفوذ در بازاریابی ویروسی می‌باشد. تعیین کاربران تأثیر گذار به عنوان یکی از اصلی‌ترین موضوعات موجود در شبکه‌های اجتماعی می‌باشدکه اهمیت فراوانی دارد. چنانچه این کاربران به صورت دقیق‌تری شناسایی شوند، عملیاتی که بر مبنای این کاربران انجام می‌شود با نفوذتر خواهد بود. هدف از انجام این پروژه ارائه روشی برای تحلیل احساسات مشتری در رسانه‌های اجتماعی جهت استفاده در سامانه‌های تبلیغات است.برای این امر از API جست و جوی توییتر استفاده شده است. این مجموعه داده شامل اطلاعات مرتبط با کاربران و فعالیت‌هایشان در بستر شبکه اجتماعی توییتر بوده است که در پایگاه داده ذخیره شده است. این داده‌های ذخیره شده برای آماده سازی به سه دسته ساختار یافته، نیمه ساختار یافته و غیر ساختار یافته تقسیم شده است و برای هر کدام از ساختارها مراحلی برای آماده سازی داده انجام شده است. سپس داده‌ها به فرمت مناسب تبدیل شده است و ویژگی های مورد نظر استخراج می‌شود و در نهایت به طبقه‌بندی کاربران براساس احساسات مثبت و منفی با استفاده از روش شبکه عصبی چرخشی خواهیم پرداخت. این روش با سه روش دیگر نیز برای ارزیابی مقایسه شد و از سه روش دیگر عملکرد بهتری داشت.

رفتارشناسی کاربران در شبکه‌های اجتماعی یکی از جذاب‌ترین بحث‌های حوزه فناوری اطلاعات در دهه اخیر می‌باشد. رفتارشناسی کاربران این امکان را در اختیار توسعه‌دهندگان فناوری اطلاعات فراهم می‌آورد که با استفاده از نیازسنجی تعاملات و برهم‌کنش علایق و خصوصیات کاربران، به ارائه سرویس بپردازند. این سرویس‌ها ممکن است در قالب ارائه آگهی تبلیغاتی، یا یک فرآیند مدیریت دانش و یا حتی به صورت یک توصیه صورت پذیرد. به عنوان پیشنهاد برای کارهای آینده، می‌توان روی رفتارشناسی انتشار در شبکه‌های اجتماعی مطالعات تکمیلی صورت پذیرد. کاربران تاثیرکذار بر روی سایر کاربران اکتشاف شوند، این کاربران به گونه‌ایی هستند که بیشترین احساس مثبت در جذب کاربران برای خرید یک محصول به سایر کاربان می‌دهند و همینطور می‌توانند بیشترین احساس منفی را برای دفع کاربران برای خرید یک محصول را در شبکه های اجتماعی داشته باشند. پست‌های تاثیرگذاری که بیشترین احساس مثبت و منفی را برای جذب یا دفع به خرید یک محصول می‌انجامد را اکتشاف کرد. با اتخاذ روش‌های دیگر برای مدل پیشنهادی

می‌توان آن را مقیاس پذیرتر نمود. با دخیل کردن ویژگی‌هایی همچون فرهنگ، نژاد، قومیت به مدل پیشنهادی می‌توان نتایج گسترده‌تری را کسب نمود. با ترکیب مدل‌های طبقه‌بندی و قراردادن وزن‌های متفاوت به ویژگی‌های تأثیرگذار، عملکرد روش پیشنهادی را افزایش داد. از طرفی علایق کاربران با گذشت زمان تغییر می‌کند، این تغییرات یا بستگی به برهه‌های خاص زمانی دارد، مانند اوایل سال جدید و یا بر اثر تغییر طبع کاربر با گذر زمان ایجاد می‌شود، با در نظر گرفتن این پویایی در شبکه‌های اجتماعی می‌توان روش پیشنهادی را انعطاف‌پذیرتر نمود.

## منابع و ماخذ


[۱] Arora, L., Singh, P., Bhatt, V., Sharma, B. (۲۰۲۱). **Understanding and Managing Customer Engagement through Social Customer Relationship Management.** Journal of Decision Systems, ۱-۲۱

[۲] Bagheri A., Saraee, M. (۲۰۱٤). **Persian Sentiment Analyzer: a Framework Based on a Novel Feature Selection Method**. International Journal of Artificial Intelligence™, Vol. ۱۲, No. ۲, pp. ۱۱۵-۱۲۹.

[۳] Chen, Z S., Zhang, X., Govindan, K., Wang, X. J., Chin, K. S. (۲۰۲۱). **Third-Party Reverses Logistics Provider Selection: A Computational Semantic Analysis-Based Multi-Perspective Multi-Attribute Decision-Making Approach**. Expert Systems with Applications, ۱٦٦, ۱۱٤۰۵۱.

[٤] Cieliebak, M., Deriu, J., Egger, D., Uzdilli, F. (۲۰۱۷). **A Twitter Corpus and Benchmark Resources for German Sentiment Analysis.** in Proc of the ۵th Ine**,** Workshop on Natural Language Processing for Social Media, SocialNLP, pp. ٤۵-۵۱, Boston, USA.

[۵] Collobert, R., Weston, J., Bottou, L., Karlen, M., Kavukcuoglu, K., Kuksa, P. (۲۰۱۱). **Natural Language Processing (Almost) from Scratch.** Journal of Machine Learning Research, vol. ۱۲, No. ۷٦, pp. ۲٤۹۳-۲۵۳۷.

[٦] Cortes C., Vapnik, V. (۱۹۹۵). **Support-Vector Networks.** Machine Learning, Vol. ۲۰, No. ۳, pp. ۲۷۳-۲۹۷.

[۷] Dos Santos C.N., Gatti, M. (۲۰۱٤). **Deep Convolutional Neural Networks for Sentiment Analysis of Short Texts**. in Proc of the ۲۵th International Conf. on Computational Linguistics, COLING'۱٤, pp. ٦۹-۷۸, Dublin, Ireland, ۲۵–۲۹.

[۸] Jaynes, E.T. (۱۹۵۷). **Information Theory and Statistical Mechanics.** Physical Review, Vol. ۱۰٦, No. ٤, pp. ٦۲۰.

[۹] Maulud, D. H. (۲۰۲۱). **State of Art for Semantic Analysis of Natural Language Processing.** Qubahan Academic Journal ۱,۲, ۲۱-۲۸.

[۱۰] Mikolov, T., I. Sutskever, I., Chen, K., Corrado, G.S., Dean, J. (۲۰۱۳). **Distributed Representations of Words and Phrases and their Compositionality**. in Proc Advances in Neural Information Processing Systems, NIPS'۱۳, pp. ۳۱۱۱-۳۱۱۹, Lake Tahoe, CA, USA, ۵-۱۰.

[۱۱] Neethu M.S., Rajasree, R. (۲۰۱۳). **Sentiment Analysis in Twitter using Machine Learning Techniques.** in Proc IEEE ٤th Int. Conf. on, Computing, Communications and Networking Technologies, ICCCNT'۱۳, ۵ pp., Tiruchengode, India, ٤-٦.

[۱۲] Roshanfekr, B., Khadivi, S., Rahmati, M. (۲۰۱۷). **Sentiment Analysis using Deep Learning on Persian Texts.** in Iranian Conf, on Electrical Engineering, ICEE'۱۷, pp. ۱۵۰۳-۱۵۰۸, Tehran, Iran, ۲-٤.

[۱۳] Shearer, E., Amy, M. (۲۰۲۱). **News Use across Social Media Platforms in ۲۰۲۰**.

[۱٤] Socher, R., Perelygin, A., Wu, J.Y., Chuang, J., Manning, C.D., Ng, A.Y., Potts, C. (۲۰۱۳). Recursive Deep Models for Semantic Compositionality over a Sentiment Treebank. in Proc of the Conf.

[۱۵] Wang, K., Wang, X., Lin, L., Wang, M., Zuo, W. (۲۰۱٤). **۳D human Activity Recognition with Reconfigurable Convolutional Neural Networks.** in Proc of the ۲۲nd ACM International Conf. on Multimedia, pp. ۹۷-۱۰٦, Orland, FL, USA, ۱۸-۱۹.

[۱٦] Zhang, Y., Chen, M., Liu, L., Wang, Y. (۲۰۱۷). **An Effective Convolutional Neural Network Model for Chinese Sentiment Analysis.** in Proc AIP Conf. Proc., vol. ۱۸۳٦, pp. ۰۲۰۰۸٤, Rome, Italy, ۲۷-۲۹.

[۱۷] on Empirical Methods in Natural Language Processing, EMNLP'۱۳, vol. ۱٦۳۱, pp. ۱٦۳۱-۱٦٤۲, Seattle, WA, USA, ۱۸-۲۱.