



پژوهش‌های نوین در تصمیم‌گیری

دوره ۶، شماره ۴، زمستان ۱۴۰۰، صص ۵۳-۳۵

نوع مقاله: پژوهشی

طراحی سیستم توصیه‌گر محصولات شوینده بر مبنای تحلیل احساسات فازی

مجید بهروان^۱، محمدرضا معتدل^{۲*}، عباس طلوعی اشلقی^۳، رضا رادفر^۴

۱. دانشجوی دکتری، گروه مدیریت فناوری، واحد تهران مرکزی، دانشگاه آزاد اسلامی، تهران، ایران.
۲. استادیار، گروه مدیریت صنعتی، واحد تهران مرکزی، دانشگاه آزاد اسلامی، تهران، ایران
۳. استاد، گروه مدیریت صنعتی، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران.
۴. استاد، گروه مدیریت صنعتی، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران.

تاریخ پذیرش: ۱۴۰۰/۰۶/۲۴

تاریخ دریافت: ۱۴۰۰/۰۳/۰۳

چکیده

امروزه شبکه‌های اجتماعی به‌عنوان محلی برای ارائه نظرات کاربران نسبت به موضوعات مختلف از جمله محصولات مصرفی تبدیل شده‌است و این نظرات به منبعی ارزشمند برای تحلیل احساسات و استخراج انتظارات مشتریان از محصولات تبدیل شده‌است. این موضوع فرصت مناسبی را برای شرکت‌ها جهت بازطراحی محصولات خود بر اساس نظرات کاربران فراهم کرده‌است. در این پژوهش برای ارائه توصیه به واحدهای طراحی محصولات شوینده، از تحلیل احساسات مشتریان و مصرف‌کنندگان این محصولات در شبکه‌های اجتماعی استفاده کرده‌ایم. بیش از ۴۲۰۰ مورد از نظرات کاربران شبکه اجتماعی توئیتر بر اساس موضوع پژوهش در سال ۲۰۱۹ استخراج و در مرحله پیش‌پردازش متون پالایش و تگ‌گذاری شدند و پس از طی این مرحله با بکارگیری منطق فازی و مدل‌یابی موضوعی به تحلیل احساسات پرداخته شده‌است. مدل‌یابی موضوعی را به‌منظور یافتن ویژگی‌های مورد اشاره در نظرات برای داشتن رویکردی بهتر در واحدهای طراحی محصولات به‌کار بردیم و از منطق فازی برای استحصال میزان قطبیت نظرات به ۵ دسته بسیار مثبت، مثبت، خنثی، منفی و بسیار منفی استفاده کرده‌ایم. نتایج به‌دست‌آمده از تحلیل احساسات در مدل پژوهش با استفاده از ماتریس آشفتگی مورد ارزیابی قرار گرفت و صحت ۸۶.۱۵٪ حاصل شد. در این پژوهش از زبان و کتابخانه‌های پایتون برای جمع‌آوری، پالایش و تحلیل اطلاعات استفاده گردید.

کلیدواژه‌ها: تحلیل احساسات، سیستم توصیه‌گر، مدل‌یابی موضوعی، هوش تجاری، منطق فازی



۱- مقدمه

تحلیل احساسات به‌طور گسترده در موضوعاتی مثل بازاریابی، تجارت‌آنلاین، بهداشت و سلامت، سیاست و آموزش مطرح شده است تا بتواند کیفیت محصولات و خدمات را بهبود بخشد [۱]. مطالعات اخیر نشان می‌دهد که الگوهای انتشار احساسات در توئیتر ارتباط نزدیک با قابلیت احساسات پیام‌های توئیتر دارند [۲] و استفاده از محتوا کاوی در شبکه‌های اجتماعی به‌عنوان رویکردی نوین در مشتری‌مداری و کسب نظر مشتریان راجع به محصولات و خدمات مورد استفاده قرار گرفته است تا با دستیابی مستقیم به نظرات صریح مشتریان، امکان افزایش تطبیق محصولات با نیاز بازار فراهم و سطح هوشمندی کسب‌وکار افزایش یابد.

غالباً مشتریان در شبکه‌های اجتماعی برندهای گوناگونی را دنبال می‌کنند تا اطلاعاتی راجع به محصول یا برند مورد علاقه خود به‌دست آورند که آن‌ها این روش ارتباطی را به نسبت فعالیت‌های ترفیعی سنتی، منبعی قابل اعتمادتر قلمداد می‌کنند [۳].

تجزیه و تحلیل احساسات فرآیند استخراج احساسات و یا نظرات از یک متن برای یک موضوع خاص که امکان درک نگرش‌ها، نظرات و احساسات در متن را فراهم می‌نماید [۴]. احساسات مشتری یکی از جنبه‌های اساسی برای تغییر رویکردهای تجاری برای مقابله با رقبا در بازار سریع امروز است. امروزه، با رشد فعالیت وبسایت‌های شبکه‌های اجتماعی در رونق انواع کسب‌وکارها و اعتباربخشی کانال‌های توزیع آن‌ها، محتوای تولید شده زیادی در مورد محصول، بازار، رقبا، تأمین‌کنندگان و توزیع‌کنندگان در رسانه‌های اجتماعی مورد تبادل قرار می‌گیرد [۵]. هر شرکت بزرگی هر روز با تعداد زیادی مشتری سروکار دارد و بنابراین دریافت بازخورد منفرد از هر یک از آن‌ها عملی نیست. از طریق سایت‌های شبکه‌های اجتماعی مانند توئیتر، فیس‌بوک و غیره می‌توان از احساسات مشتری اطلاعات کسب کرد و کسب‌وکارها می‌توانند به‌منظور دریافت بینش و تصمیم‌گیری در مورد بهبود کیفیت محصول خود مورد استفاده قرار دهند [۶].

با استخراج نظرات مشتریان درباره ویژگی‌های محصولات، امکان ارائه توصیه به واحد طراحی برای بازطراحی محصولات در راستای جلب رضایت مشتریان، موجب افزایش فروش محصولات در سراسر شبکه فروش شرکت خواهد شد.

در سال‌های اخیر با افزایش کاربران اینترنت در ایران فرصت مناسبی برای شرکت‌ها به وجود آمده است تا با بهره‌گیری از این روش نوین در کنار روش‌های سنتی بازاریابی، علاوه بر افزایش سهم خود از بازار، شناخت مشتریان از محصولات خود را افزایش داده و نام تجاری



شرکت را بهبود بخشند [۷].

احساسات در کلمات می‌توانند بنا بر جایگاه کلمه در جملات و ترکیب با سایر کلمات موجود در جمله بار معنایی متفاوتی داشته باشند، لذا تحلیل فازی احساسات امکان تعدیل و تصحیح بار معنایی کلمات را برای استخراج احساس هر کلمه در جمله فراهم می‌نماید.

در مطالعاتی که تاکنون در زمینه نظرات کاربران در خصوص رتبه‌بندی محصولات انجام گرفته است، معمولاً رابطه رتبه‌بندی محصول با میزان فروش و یا کمک به تولیدکننده در شناخت ویژگی‌های محصول مورد توجه بوده است [۸].

برای افزایش دقت پیشنهادات در سیستم توصیه‌گر نیاز به استخراج ویژگی‌های مرتبط با موضوع به نحوی است که احساسات راجع به ویژگی‌های مختلف یک موضوع مورد ارزیابی قرار گیرد و پیشنهاد ارائه شده به تفکیک ویژگی‌های موضوع ارائه شود تا امکان بازطراحی محصولات بر اساس ویژگی‌هایی که نیاز به اصلاح دارند، فراهم شود.

تاکنون استفاده از محتوا کاوی در شبکه‌های اجتماعی به‌عنوان رویکردی نوین در مشتری‌مداری و کسب نظر مشتریان راجع به محصولات و خدمات مورد استفاده قرار گرفته است تا با دستیابی مستقیم به نظرات صریح مشتریان، امکان افزایش تطبیق محصولات با نیاز بازار فراهم و سطح هوشمندی کسب‌وکار افزایش یابد [۹] اما پردازش متون فارسی کمتر مورد توجه پژوهشگران قرار گرفته است و ابزارهای لازم برای پردازش آن کمتر توسعه یافته‌اند؛ لذا توسعه مدل مبتنی بر پردازش زبان فارسی بخشی از این پژوهش خواهد بود که امکان تحلیل احساسات متون فارسی را فراهم نماییم.

در این پژوهش ما از تحلیل احساسات برای ارائه پیشنهاد به واحدهای تولید برای بهبود طراحی محصولات شوینده استفاده کردیم. ابتدا جمع‌آوری، پالایش و آماده‌سازی دیتا از توئیتر با فیلتر روی توئیتهایی که شامل کلمات مرتبط با مواد شوینده هستند، انجام گردید. نوآوری موجود در این پژوهش شامل ترکیب استفاده از الگوریتم مدل‌سازی موضوعی برای استخراج ویژگی‌های محصولات در متون و رویکرد فازی برای استحصال قطبیت احساسات کاربران شبکه‌های اجتماعی نسبت به آن ویژگی‌ها می‌باشد.

۲- پیشینه نظری پژوهش

کلمه "احساس" در فرهنگ لغت این‌گونه معنی شده است: "یک نگرش، عقیده یا قضاوت که منشأ آن احساسات است" و یا یک "منظر یا نظریه خاص: دیدگاه" و "عواطف". کلمه نظر



معمولاً بیانگر یک منظر، قضاوت یا تأیید است که برای مسئله به‌خصوصی در ذهن شکل می‌گیرد [۱]. این‌گونه توضیح می‌دهد: "دیدگاه‌ها معمولاً عبارات ذهنی هستند که احساسات، تأییدیات یا نگرش‌های مردم را نسبت به موجودیت‌ها، رویدادها و ویژگی‌های آن‌ها وصف می‌کند." اگرچه موضوع آنالیز احساسات و عقیده‌کاوی به تازگی توجه بسیاری از محققان و تاجران را کسب کرده است، اما همیشه حس عدم رضایت و مخالفتی مخفیانه برابر کاوش عقاید وجود داشته است. امروزه اغلب مطالعات در زمینه پردازش اطلاعات متنی بر استخراج و بازیابی اطلاعات حقیقی از جمله بازیابی اطلاعات، طبقه‌بندی متن یا خوشه‌بندی متن متمرکز شده‌اند [۱۰]. یکی از دلایل اصلی برای کمبود تحقیقات در زمینه تحلیل احساسات، این است که قبل از عصر تارنمای جهانی^۱، نظرات متنی زیادی موجود نبود. افراد در آن زمان قبل از تصمیم‌گیری نظر دوستان یا خانواده‌شان را جویا می‌شدند یا سازمان‌ها برای دستیابی به نظرات مردمی در مورد کالا یا خدماتش، معمولاً صندوق نظرسنجی یا گروه کانونی تشکیل می‌دادند. پس از اینکه به موجب پیشرفت تکنولوژی، اینترنت به‌طور گسترده مورد استفاده قرارگرفت، مردم می‌توانند عقاید و احساساتشان را برای خدمات و محصولات به‌صورت آنلاین ارائه دهند. در گفتگوهایی که بین مردم شکل می‌گیرد، یکدیگر را به استفاده یا عدم استفاده از یک کالا ترغیب می‌کنند. این رفتار از سوی مشتریان منابع جدید و قابل‌سنجشی از اطلاعات با کاربردهای عملی فراوان در دسترس ما قرار می‌دهد [۱۰].

۲-۱- روش‌های آنالیز احساسات

آنالیز احساسات به لطف ارزش و پتانسیل بالایی که برای کاربردهای مختلف دارد، مورد توجه مجامع علمی و تجاری قرار گرفته است [۱۱]. طبقه‌بندی احساسات می‌تواند به سه بخش تقسیم شود [۱۲]:

۱. تشخیص ذهنیت، یعنی اینکه مشخص شود آیا متن داده شده حاوی اطلاعات واقعی است یا اطلاعات ذهنی (نظرات شخصی).
 ۲. تشخیص جهت‌گیری یا قطبیت متن، یعنی مشخص شود که متن ذهنی مورد نظر، بیانگر نظرات مثبت است یا منفی.
 ۳. تشخیص شدت جهت‌گیری.
- در تحقیقاتی که بر روی آنالیز احساسات انجام می‌گیرد، دو روش اصلی مطرح می‌شود:



روش معنا محور و روش یادگیری ماشین [۱۳]. در برخی موارد، به روش معنا محور، روش مبتنی بر فرهنگ لغت هم می‌گویند. در این روش لغات و احساسات به عنوان شاخص‌های جهت‌گیری معنایی در نظر گرفته می‌شوند و قطبیت کلی متن، مجموع متوسط قطبیت شاخص‌ها است [۱۴]. روش دیگر بر پایه استفاده از تکنیک‌های یادگیری ماشین بنا شده است. یعنی هدف ما انتخاب صحیح الگوریتم‌های یادگیری ماشین و ویژگی‌های مناسب متن است تا به این صورت قطبیت متن را طبقه‌بندی کنیم. محققین و دانشمندان دیگر این دو روش را با نام یادگیری نظارت‌شده و یادگیری نظارت‌نشده می‌نامند. امروزه محققان سعی می‌کنند از مزایای هر دو روش بهره‌مند شوند. به علاوه، با ترکیب هر دو روش، سیستم‌های طبقه‌بندی هیبرید آنالیز احساسات ابداع شده است [۱۵].

هر روش مزایا و محدودیت‌های مخصوص به خود را دارد. تحقیقات قبلی هم از تکنیک یادگیری ماشین و هم از تکنیک‌های معنایی برای آنالیز احساسات در حوزه‌های مختلف استفاده کرده‌اند. در چند سال اخیر در بعضی آزمایش‌ها، روش‌های ترکیبی و خلاقانه‌ای پیشنهاد شده است تا کارایی و عملکرد آن روش افزایش یابد.

پژوهش‌های کمی در زمینه متون فارسی تاکنون انجام پذیرفته است و میزان صحت قطبیت نظر مربوط به ویژگی‌های محصولات در این پژوهش که حاصل ترکیب الگوریتم مدلیابی موضوعی و منطق فازی است، ۸۶.۱۵٪ حاصل گردید که صحت قابل قبولی نسبت به اکثر پژوهش‌های انجام شده در این حوزه را نشان می‌دهد.

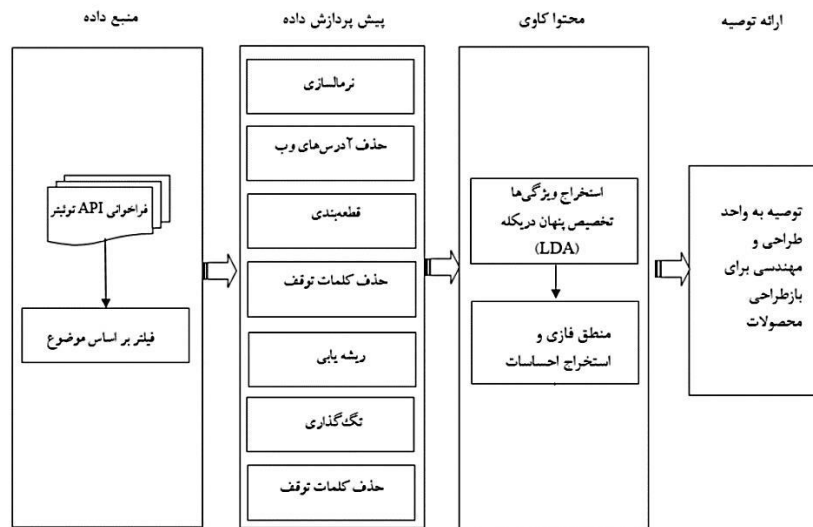


جدول ۱. خلاصه از تحقیقات گذشته

نویسنده	سال	روش	دیتاست	صحت
Montejo-Ráez et al [۱۶]	۲۰۱۲	SentiWordNet گام تصادفی	پیکره توئیتر گردآوری توسط خود شخص	۶۳٪
Saif et al [۱۷]	۲۰۱۲	SentiWordNet; SentiCircles	دیتاست توئیتر	۸۷/۵۰٪
Ortega et al [۱۸]	۲۰۱۳	SentiWordNet; برچسب‌گذاری مبتنی بر قاعده	SemEval ۲۰۱۳ Twitter&S MS	۵۰٪
Balage and Pardo [۱۹]	۲۰۱۳	روش لغت‌نامه‌ای یادگیری ماشین	داده‌های توئیتر	۶۵/۳۹٪
Jain and Pandey [۲۰]	۲۰۱۳	SentiWordNet; فقط صفت‌ها	نظرات مشتریان گردآوری توسط خود شخص	۶۹٪
Dhande and Patnaik [۲۱]	۲۰۱۴	یادگیری ماشین (Naïve Bayes Neural)	نظرات مربوط به فیلم‌ها توسط Pang et al. (۲۰۰۲)	۸۰/۶٪
Khan and Bashir [۲۲]	۲۰۱۶	SentiWordNet; SentiMi فقط برای صفت‌ها	نظرات مربوط به فیلم‌ها	۷۶٪
Li Yang et al [۲۳]	۲۰۲۰	SentiWordNet & Convolutional Neural Network	نظرات محصولات فروشگاه تجارت الکترونیک	۸۴.۳٪

۳- مدل پژوهش

نمای کلی مدل پیشنهادی برای سیستم توصیه‌گر با رویکرد تحلیل احساسات فازی در شکل ۱ نشان داده شده است.



شکل ۱. نمای کلی مدل پژوهش

۳-۱- استخراج داده‌های توئیتر

در این پژوهش از وب سرویس رسمی ارائه شده توسط توئیتر برای استخراج نظرات استفاده می‌شود. با ایجاد یک اکانت در <https://developer.twitter.com> امکان استخراج توئیتهای به صورت برخط وجود دارد.

همچنین برای فراخوانی سرویس فوق در این پژوهش از زبان برنامه‌نویسی پایتون استفاده شده است. برای این منظور، متن توئیتهای بر اساس کلمات مرتبط با موضوع پژوهش فیلتر شده و وارد مراحل پیش‌پردازش جهت پاک‌سازی و پالایش متون می‌شود.

۳-۲- پیش پردازش داده‌ها

پیش پردازش داده‌ها برای بهبود کیفیت داده‌ها و آماده‌سازی دیتاست جهت تحلیل و پردازش داده‌هاست. در این مرحله به جزئیات لازم برای دسته‌بندی و استخراج احساسات و آماده‌سازی توئیتهای برای استفاده در مدل فازی پرداخته می‌شود.

در این پژوهش با استفاده از جستجوی کلمات مرتبط با موضوع، فیلتر توئیتهای انجام می‌شود و داده‌های نامرتب و نویزی حذف می‌شود و این روش موجب کاهش حجم پردازش داده در



مراحل بعدی خواهد شد.

در این پژوهش با رویکرد منطق فازی به دنبال یافتن قطبیت توثیت‌ها در دسته‌های بسیار مثبت، بسیار منفی، منفی، مثبت و خنثی هستیم که نشان دهنده شدت نظرات و احساسات نسبت به محصولات مورد نظر می‌باشد. به کارگیری منطق فازی امکان درک احساس متفاوت نسبت به کلمات مختلف را فراهم می‌نماید، به عنوان مثال احساسی که از کلمات "فوق‌العاده"، "خوب" و "قابل قبول" (که هر سه دارای بار احساسی مثبت هستند) حاصل می‌شود قابل ارزیابی و محاسبه می‌شود.

۳-۲-۱- یکسان‌سازی^۱ متون

منظور از یکسان‌سازی، یکنواخت کردن کاراکترهای استفاده‌شده در متون است. با توجه به اینکه ممکن است اسناد منتشرشده، از کدگذاری‌های متفاوتی استفاده کنند لذا برای بهبود تحلیل‌های متنی نیازمند یکسان‌سازی این متون هستیم. در واقع حذف کد حروف غیرمناسب (مانند اعراب حروف) یا تبدیل کدهای مختلف حروف (مانند کد /ی/ عربی و فارسی) به یک کد واحد در این مرحله انجام می‌شود. همچنین کلماتی که دارای ساختار نوشتاری مختلف هستند، به یک کلمه تبدیل می‌شود. در اولین گام باید متون برای استفاده در گام‌های بعدی به شکلی استاندارد درآیند. از آنجایی که متون مختلف ممکن است بسیار به هم شبیه باشند اما به دلیل تفاوت‌های ساده ظاهری از نظر ماشین متفاوت باشند؛ به همین دلیل سعی شده است این تفاوت‌های ساده ظاهری برطرف گردد. برای رسیدن به این هدف، قبل از مقایسه متون، پیش‌پردازش‌هایی روی آن‌ها انجام می‌شود. طبیعتاً هر چه این پیش‌پردازش‌ها قوی‌تر باشد، نتایج حاصل از مقایسه متون قابل اطمینان‌تر خواهد بود. شایان ذکر است که از آنجایی که زبان فارسی جزء زبان‌های غیرساخت‌یافته است، با مشکلات بسیار بیشتری نسبت به سایر زبان‌ها مواجه خواهیم شد. متون غیرساخت‌یافته، متونی هستند که پیش‌فرض خاصی در مورد قالب آن‌ها نداریم و آن‌ها را به صورت مجموعه‌ای مرتب از جملات در نظر می‌گیریم.

۳-۲-۲- قطعه‌بندی^۲ متن توثیت‌ها

قطعه‌بندی ابزاری برای شکستن یک متن بر اساس واحدهای با معنی مانند کلمه، پاراگراف، نمادهای معنادار مانند فاصله می‌باشد. لازمی ایجاد این ابزار جمع‌آوری واحدهایی است که در



زبان فارسی به‌عنوان واحدهای مستقل معنایی شناخته می‌شوند. سپس بر اساس انتخاب هر کدام از این واحدها متن بر اساس آن شکسته خواهد شد. قطعه‌بند مرز کلمات را در متون تشخیص داده و متن را به دنباله‌ای از کلمات تبدیل می‌کند و آن را برای تحلیل‌های بعدی آماده می‌کند. در واقع این واحد ساز جزء مراحل پیش‌پردازش متن به حساب می‌آید. معمولاً بعد از این مرحله، نوبت به مرحله حذف کلمات توقف یا Stop Words می‌باشد.

۳-۲-۳- حذف کلمات توقف^۱

کلمات توقف، کلماتی پرتکرار هستند که در ساختار زبان مورد استفاده قرار می‌گیرند. در فارسی کلماتی مانند "از"، "تا"، "به" کلمات توقف به‌شمار می‌روند که در جملات به‌وفور به‌کار می‌روند و در ساختار جملات ضروری می‌باشند ولی در فرآیند تحلیل جملات کارایی ندارند. در این پژوهش، جملات به لیستی از کلمات تقسیم می‌شوند و سپس با لیستی از کلمات توقف مقایسه می‌شوند و در صورتی که در این لیست باشند، حذف خواهند شد.

۳-۲-۴- ریشه‌یابی^۲ کلمات

ریشه‌یابی عبارت است از حذف پس‌وندها و پیش‌وندهای کلمات و استخراج ریشه آن‌ها. در هر زبان، واژه‌ها با توجه به نقش معنایی و نحوی خود در جمله به شکل‌های ظاهری متفاوتی حضور می‌یابند، این شکل ظاهری متفاوت از جهتی نشان‌دهنده معنای متفاوت این واژه‌هاست، اما با توجه به این که تمامی آن‌ها از یک ریشه مشتق شده‌اند، از نظر معنا قرابت نسبتاً زیادی خواهند داشت. از همین رو در بسیاری از کاربردهای پردازش زبان طبیعی و بازیابی اطلاعات، نیاز داریم تا همه مشتقات یک واژه را به ریشه‌ی آن، که همان شکل ساده واژه می‌باشد، تبدیل نماییم. ریشه‌یابی کلمات در سیستم‌های بازیابی اطلاعات یک امر ناگزیر است.

۳-۲-۵- برچسب زدن با استفاده از نگه‌های POS

ویژگی‌های محصولات با استفاده از جایگاه کلمات در جملات استخراج می‌شود. در برچسب‌زنی کلمات براساس نقش آن‌ها در جمله هر کلمه یکی از نقش‌های صفت، اسم، قید و فعل را به خود اختصاص می‌دهد و تحلیل احساسات بر اساس برچسب هر کلمه امکان‌پذیر



می‌شود. در این پژوهش برای شناسایی نقش کلمات در جملات از کتابخانه هضم استفاده شده است.

جدول ۲. برچسب‌گذاری توئیت‌ها

متن نظر	POS
کاملاً موها رو نرم کرده و حالت ابریشمی به موها میده پیشنهاد می‌کنم حتماً استفاده کنید	['Ne', 'حالت'], ['AJ', 'نرم'], ('N', 'موها'), ['V', 'می‌ده'], ('N', 'موها'), ('AJe', 'ابریشمی'), ['N', 'استفاده'], ('ADV', 'حتماً'), ('N', 'پیشنهاد')]
خیلی کوچیک و گران. اصلاً ارزش خریدن نداره	['AJ', 'کوچک'], ('AJ', 'خیلی'), ('ADV', 'خیلی'), ['N', 'ارزش'], ('N', 'ارزش'), ('AJ', 'ارزش'), ['V', 'نداره'], ('N', 'نداره')]
به تازگی سفارش دادم، خیلی به صرفه و خیلی موهام براق و بانشاط شده	['V', 'دادم'], ('N', 'سفارش'), ('Ne', 'تازگی'), ['ADV', 'خیلی'], ('PUNC', ':'), ('ADV', 'صرفه'), ['Ne', 'بانشاط'], ('Ne', 'براق'), ('N', 'موهام'), ['AJ'], ('AJ'), ('AJ')]
به نظرم شامپوی بدی نبود. موها رو ضخیم می‌کنه ولی کمی خشک می‌کنه که من مجبور میشم ماسک مو بعدش استفاده کنم. بوی گیاه میده ولی خوشاینده	['AJ', 'بدی'], ('Ne', 'شامپوی'), ('N', 'نظرم'), ['AJ', 'ضخیم'], ('N', 'موهارو'), ('PUNC', ':'), ['V', 'می‌کنه'], ('AJ', 'خشک'), ('V', 'کنه می'), ['Ne', 'ماسک'], ('V', 'می‌شم'), ('AJ', 'مجبور'), ['N', 'استفاده'], ('ADV', 'بعدش'), ('N', 'مو'), ['PUNC', ':'], ('N', 'گیاه'), ('Ne', 'بوی'), ('PUNC', ':'), ['N', 'خوشاینده'], ('V', 'خوشاینده')]

۴- روش‌شناسی پژوهش

۴-۱- مدل‌سازی موضوعی

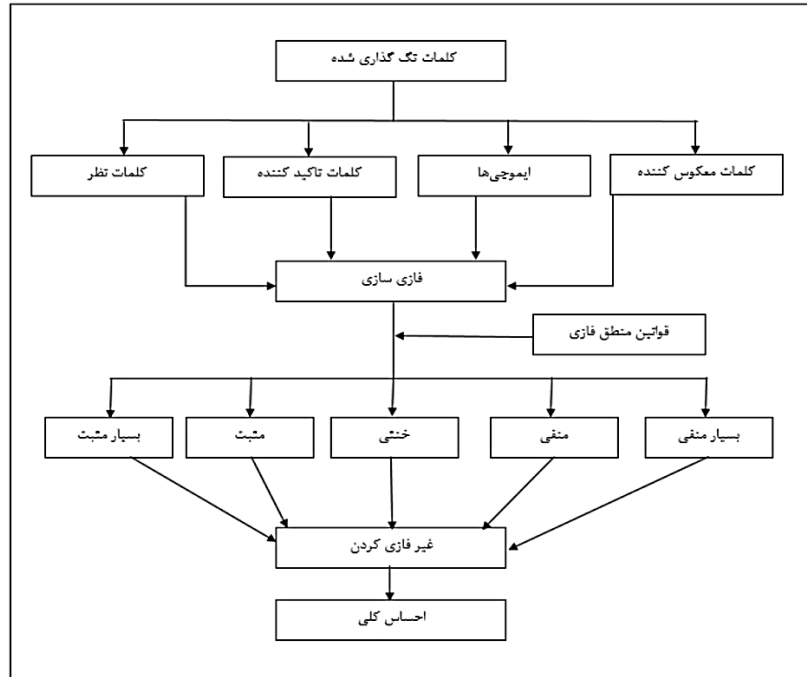
برای استخراج ابعاد موضوعات مطرح شده در متن توئیت‌ها از الگوریتم‌های مدل‌سازی موضوعی استفاده می‌کنیم که روش‌های آماری هستند که کلمات داخل یک متن را تحلیل کرده و از این طریق موضوعات داخل متون را استخراج می‌کنند. همچنین ارتباط این موضوعات با یکدیگر و نیز تغییر آن‌ها در طول زمان را مشخص می‌کنند. این الگوریتم‌ها نیازی به هیچ فرض اولیه‌ای در مورد موضوعات متون و یا برچسب‌گذاری متون ندارند. بلکه ورودی آن‌ها متن اصلی است. الگوریتم‌های مدل‌سازی موضوعی به ما این امکان را می‌دهند تا سازمان‌دهی و خلاصه‌سازی آرشیوهای الکترونیکی‌مان را در ابعادی که از عهده‌ی انسان بر نمی‌آید، انجام



دهیم. در این پژوهش از الگوریتم LDA^۱ استفاده کردیم. در این پژوهش موضوعات استخراج شده درباره مواد شوینده را در قالب ۳ دسته متفاوت توسط الگوریتم LDA تفکیک کردیم و متغیر آلفای ورودی الگوریتم را ۰.۵ در نظر گرفتیم و میزان ارتباط هر توئیت با هر دسته را محاسبه کردیم تا در مرحله فازی‌سازی از آن استفاده نماییم. به‌عنوان مثال در توئیت " کاملاً موها رو نرم کرده و حالت ابریشمی به موها میده پیشنهاد میکنم حتماً استفاده کنید" میزان ارتباط با موضع کارایی ۸۱٪ و با موضوع ویژگی‌ها ۱٪ و با موضوع ظاهر محصول ۱۹٪ مرتبط می‌باشد.

۲-۴- مدل دسته‌بندی فازی برای تحلیل احساسات

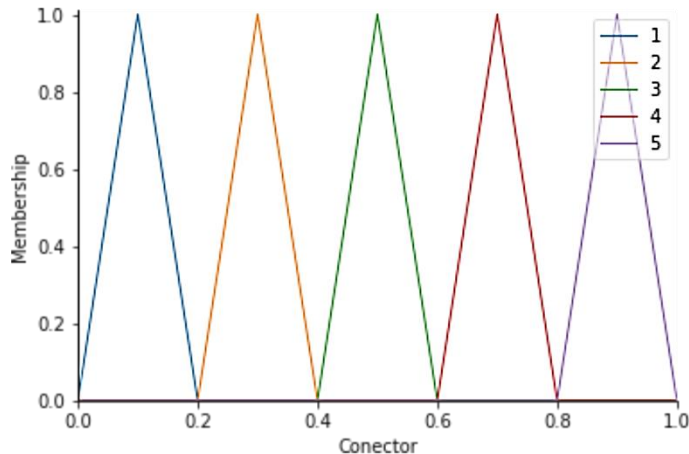
برای استفاده از منطق فازی در پردازش عبارات، بایستی از تکنیک‌های عددی کردن ویژگی‌ها استفاده کرد تا امکان محاسبات فازی فراهم شود. ۴ نوع تگ کلمات به عنوان ورودی به سیستم فازی وارد خواهند شد. این ۴ گروه شامل کلمات نظر، کلمات معکوس کننده، کلمات تأکید کننده و ایموجی‌ها می‌شود.



شکل ۲. مدل استخراج احساسات با استفاده از منطق فازی

۱-۲-۴- دسته‌بندی با استفاده از منطق فازی

اعمال منطق فازی، شامل فرآیندی است که در آن ورودی‌ها با استفاده از تابع عضویت به مجموعه‌های فازی نگاشت می‌شوند. در این پژوهش از تابع عضویت مثلثی برای تخصیص وزن متناسب به کلمات ورودی استفاده می‌شود. به‌عنوان مثال تابع عضویت کلمات تأکید کننده شامل کلمات "بسیار زیاد"، "زیاد"، "کاملاً"، "باور نکردنی" (به ترتب با کدهای ۱ تا ۵ در شکل ۳) و تابع عضویت نودنقه‌ای برای تخصیص دسته خروجی به یکی از ۵ دسته خیلی منفی، خیلی مثبت، منفی، مثبت و خنثی استفاده می‌شود.



شکل ۳. تابع عضویت کلمات تأکید کننده

۲-۲-۴- قوانین فازی

مجموعه از عبارات اگر-آنگاه برای شرح قوانین فازی و نگاشت یک یا چند ورودی به یک خروجی مورد استفاده قرار می‌گیرد. در این پژوهش این قوانین در زبان برنامه‌نویسی پایتون تعریف خواهند شد.

۵- نتایج

۵-۱- شاخص‌های ارزیابی

برای ارزیابی مدل این پژوهش، شاخص‌های صحت^۱، دقت^۲ و فراخوانی^۳ و اندازه‌گیری F براساس نتایج به‌دست آمده از مدل و به کمک ماتریس آشفتگی محاسبه می‌شود. استفاده از این معیارها در ارزیابی مدل‌های پردازش زبان طبیعی کاربرد فراوانی دارد که شامل زمینه تحلیل احساسات نیز می‌شود [۲۸]. با فرض اینکه هدف، بازیابی یا طبقه‌بندی مجموعه‌ای از اسناد مورد نظر (TP+FP) است، داده‌های انتخاب شده جهت ارزیابی (TP+FP) هستند. نماد TP برای پیش‌بینی‌های مثبت و صحیح، TN برای موارد منفی و صحیح توسط سیستم بازیابی یا طبقه‌بندی استفاده می‌شوند؛ در حالی که موارد طبقه‌بندی شده با نماد FP مربوط به موارد مثبت و نادرست و موارد موجود در FN شامل موارد منفی و نادرست هستند [۲۹].



جدول ۳. ماتریس آشفتگی

پیش‌بینی ناصحیح	پیش‌بینی صحیح	
FN	TP	موارد مثبت
TN	FP	موارد منفی

شاخص صحت، دقت، فراخوانی و شاخص F از طریق فرمول‌های زیر محاسبه می‌شوند:

$$accuracy = \frac{tp + tn}{tp + fp + fn + tn}$$

$$precision = \frac{tp}{tp + fp}$$

$$recall = \frac{tp}{tp + fn}$$

$$F = \frac{2 * precision * recall}{precision + recall}$$

شاخص صحت به‌عنوان شاخص کلی ارزیابی مدل مورد استفاده قرار می‌گیرد که نشانگر میزان پیش‌بینی‌های صحیح توسط مدل پژوهش است. شاخص دقت در صورتی که بالا باشد، نشانگر پیش‌بینی ناصحیح مثبت کمتر و دقت پایین به‌معنای صحیح مثبت بیشتر است. شاخص فراخوانی توسط شاخص دقت مورد استفاده قرار می‌گیرد. شاخص فراوانی نشانگر میزان حساسیت مدل است. ارتباط بین دقت و فراخوانی به‌نحوی است که عموماً با افزایش یکی از شاخص‌ها دیگری کاهش می‌یابد. از شاخص F برای ادغام شاخص‌های دقت و فراخوانی جهت ارزیابی عملکرد کلی مدل استفاده می‌شود. به‌دلیل ماهیت ویژه تحقیقات در حوزه پردازش زبان طبیعی می‌توان از یک و یا ترکیبی از این شاخص‌ها برای ارزیابی مدل استفاده کرد [۲۴]. در ارزیابی مدل این تحقیق از هر چهار معیار استفاده شده است.

۲-۵- تحلیل نتایج

ماتریس آشفتگی در محاسبه عملکرد مدل و صحت و دقت عملکرد مورد استفاده قرار می‌گیرد. بدین منظور تعدادی توئیت که به‌صورت دستی تگ‌های ۵ گانه احساسات را دارند، به‌عنوان داده



آزمایشی به مدل داده می‌شوند و نتایج به‌دست آمده در ماتریس آشفتگی ثبت و ارزیابی مدل صورت می‌پذیرد.

جدول ۴. احساسات استخراج شده از توئیت‌ها

احساس فازی	متن نظر
بسیار مثبت	کاملاً موها رو نرم کرده و حالت ابریشمی به موها میده پیشنهاد میکنم حتماً استفاده کنید
مثبت	به تازگی سفارش دادم، خیلی به‌صرفه و خیلی موهام براق و بانشاط شده
خنثی	به نظرم شامپوی بدی نبود. موها رو ضخیم میکنه ولی کمی خشک میکنه که من مجبور میشم ماسک مو بعدش استفاده کنم. بوی گیاه میده ولی خوشاینده
بسیار منفی	عملکردش خیلی ضعیفه. هیچ تأثیری هم نداشت. با چشم که چیز خاصی دیده نشد. بوی نامطبوعی هم داره.
بسیار مثبت	یه محصول با کیفیت با قیمت پایین که واقعاً برای تمیز کردن ماشین ظرفشویی کاربردی هستش و خیلی سریع و راحت جرم‌گیری میکنه
بسیار منفی	دوستان اصلاً نخزید. من ماشینم ال جی کره‌ای هست. این پودر را که استفاده کردم قسمت‌های استیل داخلی ماشینم زرد و قسمت‌های پلاستیکی هم همه خرد شد!

جدول ۵. ماتریس آشفتگی مدل دسته‌بندی فازی احساسات

	پیش‌بینی بسیار مثبت	پیش‌بینی مثبت	پیش‌بینی خنثی	پیش‌بینی منفی	پیش‌بینی خیلی منفی	
بسیار مثبت	۵۲	۳	۰	۳	۰	۵۸
مثبت	۴	۵۵	۴	۲	۰	۶۵
خنثی	۲	۳	۳۰	۷	۰	۴۲
منفی	۰	۴	۳	۶۰	۱	۶۸
بسیار منفی	۰	۰	۲	۵	۶۰	۶۷
	۵۸	۶۵	۳۹	۷۷	۶۱	۳۰۰

جدول ۶. ارزیابی عملکرد مدل

	بسیار مثبت	مثبت	خنثی	منفی	خیلی منفی	میانگین
دقت	۹۰.۶۳٪	۸۴.۶۲٪	۷۶.۹۲٪	۷۷.۹۲٪	۹۸.۳۶٪	۸۵.۶۹٪
فراخوانی	۸۹.۶۶٪	۸۴.۶۲٪	۷۱.۴۳٪	۸۸.۳۴٪	۸۹.۵۵٪	۸۴.۷۰٪
شاخص F	۸۹.۲۳٪	۸۴.۶۲٪	۷۴.۰۷٪	۸۲.۷۶٪	۹۳.۷۵٪	۸۴.۸۹٪
صحت	۸۶.۱۵٪					



هدف این پژوهش، ارائه چارچوبی برای یک سیستم توصیه‌گر براساس ارزیابی فازی احساسات مشتریان نسبت به محصولات شوینده بود. طراحی مدل براساس منطق فازی، امکان استخراج دقیق‌تر احساسات موجود در نظرات مشتریان محصولات تجاری در شبکه اجتماعی توئیتر را فراهم نمود. استخراج ویژگی‌های محصولات، کلمات نظر، ایموجی‌ها و کلمات معکوس‌کننده در ساختار فازی مدل امکان دسته‌بندی دقیق‌تر نظرات را فراهم آورد. دسته‌بندی نظرات با جزئیات بیشتر و در قالب بسیار مثبت، مثبت، خنثی، منفی و بسیار منفی ارائه توصیه در بخش توصیه‌گر پژوهش را بهبود داد. با ارزیابی مدل با استفاده از ماتریس آشفتگی میزان صحت مدل در بخش استخراج احساسات به‌صورت فازی معادل ۸۶.۱۵٪ به‌دست آمد که امکان استفاده از این مدل در بخش‌های تجاری نیز میسر می‌نماید. در نهایت براساس نتایج به‌دست آمده از تحلیلی فازی احساسات، سیستم توصیه‌گر ایجاد شده، به ارائه پیشنهاداتی برای تغییر در طراحی محصولات می‌پردازد.

۳-۵- پیشنهادات کاربردی و پیشنهاد برای تحقیقات آتی

پیشنهاد می‌شود در پژوهش‌های آتی درحوزه تحلیل احساسات و استخراج ویژگی‌های محصولات در بخش الگوریتم و روش، ترکیب یادگیری عمیق شامل شبکه‌های عصبی بازگشتی و منطق فازی مورد ارزیابی قرار گیرید و نتایج حاصله در بخشی مانند بازاریابی اینترنتی و فروش محصولات، مورد استفاده قرار گیرد.

۶- پی‌نوشت‌ها

۱. World Wide Web
۲. Normalizer
۳. Tokenize
۴. Stop Word
۵. Stemmer
۶. Latent Dirichlet Allocation
۷. Accuracy
۸. precision
۹. Recall



۷- منابع

- [۱] Khan, F. H., Qamar, U., & Bashir, S. (۲۰۱۶). 'SentiMI: Introducing point-wise mutual information with SentiWordNet to improve sentiment polarity detection'. *Applied Soft Computing*, ۳۹, pp. ۱۴۰-۱۵۳.
- [۲] Lei Wang, Jianwei Niu, Shui Yu. (۲۰۱۹). 'SentiDiff: Combining Textual Information and Sentiment Diffusion Patterns for Twitter Sentiment Analysis', *IEEE Transactions on Knowledge and Data Engineering*. Volume: ۳۲, Issue: ۱۰.
- [۳] Zohreh Dehdashti Shahrokh, Maryam Naeli. (۲۰۲۰) 'The Impact of Social Media Marketing Activities On Customer Equity of Luxury Brands A Study of Dorsa Brand'. *Journal of Management Research in Iran*.
- [۴] Gamon, M., Aue, A., Corston-Oliver, S., and Ringger, E. (۲۰۰۵). 'Pulse: Mining customer opinions from free text', In *Advances in Intelligent Data Analysis VI*, pp. ۱۲۱-۱۳۲. Springer Berlin Heidelberg.
- [۵] Azim Zarei, Davood Feiz, Ghazale Taheri. (۲۰۲۰). 'Providing Social Market Intelligence Framework based on web ۲.۰ Using Text-Mining Technique on Social Media Websites (Case Study: Competitive Analysis between Samsung and Emersun Brands)'. *Journal of Management Research in Iran*.
- [۶] N Srivats Athindran, S. Manikandaraj, R. Kamaleshwar. (۲۰۱۸). 'Comparative Analysis of Customer Sentiments on Competing Brands using Hybrid Model Approach', ۳rd International Conference on Inventive Computation Technologies (ICICT).
- [۷] Ali Mohaghar, Seyed Hojjat Bazazzadeh, Roya Eghbal. (۲۰۱۷) "Identification and Prioritization of Effective Factors on Online Advertising in Iran's Market by Use of Fuzzy MADM Technics (Case Study: Clothing Industry)"; *Journal of Modern Research in Decision Making*.
- [۸] Omid afsharizadeh jafari, Morteza Maleki MinBashRazgah, Azim Zarei, Mohsen Shafiei Nikabadi. (۲۰۲۱). "Designing a ranking system for purchased products based on the consumer's and expert's opinions using an aspect-based sentiment



- analysis approach”; Journal of Modern Research in Decision Making.
- [۹] Lin, Y., Zhang, J., Wang, X., and Zhou, A. (۲۰۱۲). ‘An information theoretic approach to sentiment polarity classification’, In Proceedings of the ۲nd Joint WICOW/AIRWeb Workshop on Web Quality, pp. ۳۵-۴۰. ACM.
- [۱۰] Liu, B. (۲۰۱۵). Sentiment analysis: Mining opinions, sentiments, and emotions. Cambridge University Press.
- [۱۱] Brooke, J. (۲۰۰۹). ‘A semantic approach to automated text sentiment analysis’, Doctoral dissertation, Simon Fraser University.
- [۱۲] Liu, B. (۲۰۱۲). ‘Sentiment analysis and opinion mining’, Synthesis lectures on human language technologies, ۵(۱), pp. ۱-۱۶۷.
- [۱۳] Pang, B. and Lee, L. (۲۰۰۸). ‘Opinion mining and sentiment analysis’, Foundations and trends in information retrieval, ۲(۱-۲), pp. ۱-۱۳۵.
- [۱۴] Esuli, A, and Sebastiani, F. (۲۰۰۵). ‘Determining the semantic orientation of terms through gloss classification’, In Proceedings of the ۱۴th ACM International Conference on Information and Knowledge Management. Bremen, Germany.
- [۱۵] Medhat, W., Hassan, A. and Korashy, H. (۲۰۱۴). ‘Sentiment analysis algorithms and applications: A survey’, Ain Shams Engineering Journal, ۵(۴), pp. ۱۰۹۳-۱۱۱۳.
- [۱۶] Montejo-Ráez, A., Martínez-Cámara, E., Martín-Valdivia, M. T. and Urena-Lopez, L. A. (۲۰۱۲). ‘Random walk weighting over sentiWordNet for sentiment polarity detection on twitter’. In Proceedings of the ۳rd Workshop in Computational Approaches to Subjectivity and Sentiment Analysis, pp. ۳-۱۰. Association for Computational Linguistics.
- [۱۷] Saif, H., He, Y. and Alani, H. (۲۰۱۲). ‘Semantic sentiment analysis of twitter’, In The Semantic Web–ISWC ۲۰۱۲, pp. ۵۰۸-۵۲۴. Springer Berlin Heidelberg.
- [۱۸] Ortega, R., Fonseca, A. and Montoyo, A. (۲۰۱۳). ‘SSA-UO: Unsupervised Twitter sentiment analysis’, In Second Joint Conference on Lexical and Computational Semantics (* SEM) Vol. ۲, pp. ۵۰۱-۵۰۷.
- [۱۹] Balage Filho, P. P. and Pardo, T. A. (۲۰۱۳). ‘Nilec usp: A hybrid system for



- sentiment analysis in twitter messages'.In Second Joint Conference on Lexical and Computational Semantics (* SEM),Vol. ۲, pp. ۵۶۸-۵۷۲.
- [۲۰] Jain, A. K., & Pandey, Y. (۲۰۱۳). 'Analysis and Implementation of Sentiment Classification Using Lexical POS Markers', International Journal, ۲(۱).
- [۲۱] Dhande, L., and Patnaik, G. (۲۰۱۴). 'Analyzing sentiment of movie review data using Naive Bayes neural classifier', Int J Emerg Trends Technol Comput Sci, ۳, pp.۳۱۳۳۲۰.
- [۲۲] Khan, F. H., Qamar, U., & Bashir, S. (۲۰۱۶). 'SentiMI: Introducing point-wise mutual information with SentiWordNet to improve sentiment polarity detection'. Applied Soft Computing, ۳۹, pp.۱۴۰-۱۵۳.
- [۲۳] Li Yang, Ying Li, Jin Wang, R. Simon Sherratt, (۲۰۲۰), 'Sentiment Analysis for E-Commerce Product Reviews in Chinese Based on Sentiment Lexicon and Deep Learning', IEEE Access (Volume: ۸).
- [۲۴] Prabowo, R. and Thelwall, M. (۲۰۰۹). 'Sentiment analysis: a combined approach', Journal of Informetrics, ۳, pp.۱۴۳-۱۵۷.