

ارائه یک مدل مبتنی بر زمینه برای رفع ابهام از مفاهیم حسی با کمک دانش عرفی

زینب رجیبی^۱، مریم حورعلی^۲، محمدرضا ولوی^۳

تاریخ دریافت: ۱۳۹۷/۰۳/۳۰

تاریخ پذیرش: ۱۳۹۷/۰۷/۲۴

چکیده

با گسترده شدن روز افزون شبکه‌های اجتماعی، حجم زیادی از داده و محتوا توسط کاربران تولید می‌شود. تحلیل محتواهای تولید شده توسط کاربران می‌تواند برای نیل به اهداف اقتصادی، سیاسی، فرهنگی، اجتماعی، دفاعی و امنیتی در جهت افزایش کارایی و بهره‌وری بسیار موثر می‌باشد. امروزه با نفوذ شبکه‌های اجتماعی در بستر اینترنت، بسیاری از جنبه‌های زندگی بشر تحت تأثیر قرار گرفته است. این شبکه‌ها علاوه بر به وجود آوردن فرصت‌هایی که قبلاً وجود نداشته‌اند، تهدیدات و مشکلات فراوانی را بر علیه کاربران خود نیز به وجود آورده‌اند؛ بنابراین تحلیل احساسات در شبکه‌های اجتماعی از منظر امنیتی اهمیت قابل توجهی دارد. در این مقاله مدلی جدید برای تعیین قطبیت نظرات کاربران با رویکرد کیسه مفاهیم و بر مبنای دانش عرفی ارائه شده است. ابتدا روشی برای شناخت مفاهیم حسی مبهم پیشنهاد شده است، سپس روشی برای رفع ابهام از مفاهیم حسی با کمک کیسه مفاهیم و پیوندهای آن با مفاهیم دیگر ارائه شده است. قطبیت مفاهیم مبهم، با توجه به بردارهای زمینه مثبت و منفی شناخته شده و ارتباط با مفاهیم پایگاه دانش معنایی، دقیق‌تر تعیین می‌شود. در نهایت مدل پیشنهادی با بکارگیری برای پیکره مورد نظر، مورد ارزیابی قرار گرفته و نتایج تجربی نشان‌دهنده اثر بخش بودن مدل پیشنهادی می‌باشد.

کلیدواژه‌ها:

تحلیل احساسات، نظرکاوی، مفاهیم حسی، رفع ابهام، زمینه سازی، دانش عرفی، کیسه مفاهیم

^۱دانشکده برق و کامپیوتر، دانشگاه صنعتی مالک اشتر، ایمیل: rajabi.ze@gmail.com

^۲هیئت علمی دانشکده برق و کامپیوتر، دانشگاه صنعتی مالک اشتر، ایمیل: mhourali@mut.ac.ir

^۳هیئت علمی دانشکده برق و کامپیوتر، دانشگاه صنعتی مالک اشتر، نویسنده مسئول، ایمیل: valavi@mut.ac.ir

مقدمه

با نفوذ شبکه‌های اجتماعی مبتنی بر اینترنت در زندگی انسان‌ها، بسیاری از جنبه‌های زندگی بشر تحت تأثیر قرار گرفته است. این شبکه‌ها علاوه بر به وجود آوردن فرصت‌هایی که قبلاً وجود نداشته‌اند، تهدیدات و مشکلات فراوانی را بر علیه کاربران خود نیز به وجود آورده‌اند؛ به گونه‌ای که این تهدیدات از طرفی توسط سایر کاربران این شبکه‌ها و از طرف دیگر، از جانب دولت‌ها و نهادهای وابسته به دولت‌های متخاصم بر کاربران تحمیل می‌شوند. بدین ترتیب، گستردگی کاربری از شبکه‌های اجتماعی مبتنی بر فضای مجازی نباید ما را از بررسی تهدیدات بالقوه‌ی این‌گونه فضاها غافل گرداند. امروزه تحقیقات گسترده‌ای در این زمینه در ابعاد کاربردی متنوع از کسب و کار، تعیین روابط اجتماعی و نیز تحلیل‌های حوزه امنیتی انجام می‌شود. اگرچه این تحقیقات جدید و شرایط اولیه خود را سپری می‌کند لکن به شدت مورد توجه سازمان‌های امنیتی قرار گرفته و به عنوان ابزار کارآمد در تحلیل رفتار و علائق اجتماعی و موارد دیگر استفاده می‌گردد. تحلیل داده‌های تولید شده توسط کاربران، می‌تواند به بهره‌برداری اطلاعات مستتر از انبوه داده‌ها با کارکردهای امنیتی، ادراک و ذائقه سنجی جامعه کمک نماید. بنابراین پژوهش حاضر سعی در بهبود دقت روش‌های تحلیل احساس^۴ که یکی از تحلیل‌هایی است که روی کلان‌داده‌هایی^۵ که در فضای مجازی ایجاد می‌شود، داشته و از منظر امنیتی اهمیت قابل توجهی دارد. تحلیل احساس یا نظرکاوی^۶، حوزه مطالعاتی است که نظر، احساس، ارزیابی‌ها، گرایش‌ها و تمایلات افراد را نسبت به موجودیت‌هایی مانند محصولات، سرویس‌ها، سازمان‌ها، افراد، قضیه‌ها، موضوعات و ویژگی‌ها

های هرکدام از آنها را مورد تحلیل قرار می‌دهد [۱]. اصلی‌ترین وظیفه تحلیل احساس تعیین قطبیت یا تعیین مثبت یا منفی بودن یک نظر هست که در سال‌های اخیر تلاش‌های فراوانی را به خود معطوف کرده است.

روش‌های تحلیل احساس در آخرین پیشرفت‌های علمی به دقت‌های چشمگیری دست یافته‌اند، ولی هنوز فاصله زیادی با توانایی انسان در درک نظر نویسنده دارند. اگر چه در حال حاضر روش‌های بسیار دقیق و زیادی برای تجزیه و تحلیل و استخراج دانش مربوط به داده‌های ساخت یافته وجود دارد، وظیفه استخراج اطلاعات مفید از داده‌های بدون ساختار هنوز با چالش‌های مهمی درگیر است [۲]. تحلیل احساس بیشتر به عنوان یک مسئله طبقه‌بندی در نظر گرفته شده است و سعی می‌کند متون یا جملات را به دو دسته مثبت و منفی طبقه‌بندی کند. با نگاه عمیق‌تر به این مسئله، روش‌هایی که دانش زمینه را لحاظ نکنند فاقد ارزش کافی می‌باشند [۱].

روش‌های آماری تحلیل احساس از کلید واژه‌ها به طور کورکورانه استفاده می‌نمایند و تعداد دفعات وقوع همزمان آنها را می‌شمارند. این روش‌ها کمتر روی معنا و ویژگی‌های ضمنی مرتبط با مفاهیم زبان طبیعی تکیه دارند [۳]. رویکردهای مبتنی بر مفهوم بهتر از تکنیک‌های نحوی صرف و به طور موشکافانه‌تری نظرات را تشخیص می‌دهند. این رویکردها عبارت‌های چند کلمه‌ای که صریحاً قطبیت را نمی‌رسانند، بلکه مرتبط با مفاهیم هستند را می‌توانند تحلیل کنند. یک سیستم تحلیل احساس، بدون منبع فراگیر دانش انسانی به سختی می‌تواند معنای یک متن زبان طبیعی را درک نماید [۴]. این پژوهش روی استفاده از پایگاه دانش انسانی تکیه داشته و به دنبال بهبود عملکرد چنین رویکردهایی می‌باشد. توجه به پایگاه دانش معنایی در

^۴ Sentiment analysis

^۵ Big data

^۶ Opinion mining

فرآیند تحلیل احساس حائز اهمیت می‌باشد. در این پژوهش یک مدل جدید در تحلیل احساس به کمک بازنمایی دانش عرفی برای تحلیل نظرات نقد و بررسی محصولات ارائه خواهد شد. بکارگیری دانش معنایی و دانش عرفی نهفته در منابع خاص مانند ConceptNet مدلی غنی‌تر و انتزاعی‌تر از زمینه را فراهم می‌کند و دقت و فراخوانی بیشتر الگوریتم‌های تحلیل احساس در تعیین قطبیت را به همراه دارد.

ترسیم زمینه گفتگوهای شبکه اجتماعی و موثر قراردادن زمینه در تحلیل نظرات کاربران در شبکه‌های اجتماعی در دقت تحلیل‌ها تاثیر زیادی خواهد داشت. زمینه از چندین جهت قابل بررسی است: ۱- زمینه ذهنی افراد، ارتباطاتی که بین مفاهیم موجود در متن و مفاهیمی که در ذهن افراد وجود دارد، که در واقع دانش زمینه نامیده می‌شود و از دو بخش تشکیل شده است: دانش عرفی و دانش دامنه ۲- دامنه و حوزه مورد بحث مانند حوزه نقد و بررسی محصولات الکترونیک یا حوزه نقد و بررسی هتل ۳- مفاهیمی که در اطراف مفاهیم حسی رخ می‌دهد و با آنها هم‌رخداد است.

مهمترین شاخص در تحلیل احساس، واژگان حسی^۷ هستند، این کلمات برای بیان احساسات مثبت یا منفی مورد استفاده قرار می‌گیرد [۱]. در مسئله تحلیل احساس کلاسیک، قطبیت هر عبارت از سند مستقل از دامنه‌ای که سند متعلق به آن است محاسبه می‌شود در حالیکه دامنه سخن و حوزه گفتگو تاثیر بسزایی روی تعیین قطبیت دارد. همچنین کلمات در حوزه‌های مختلف دارای معانی متفاوت می‌باشند و بی‌توجهی الگوریتم‌ها در تعیین معانی آنها می‌تواند نتایجی ضعیف را به همراه داشته باشد.

اگرچه پژوهش‌های زیادی در تعیین قطبیت وابسته به حوزه خاص تلاش نمودند [۵] و [۶] اما مسئله مهم‌تر آن

است که قطبیت واژگان در یک حوزه نیز ثابت نیست. قطبیت نه تنها به یک حوزه وابسته است بلکه به زمینه نیز بستگی دارد. منظور از زمینه، محلی است که واژه در آن رخ می‌دهد و در نظر گرفتن مفاهیمی که با آن هم‌رخداد است. بنابراین چالش موجود در تحلیل احساس عبارت است از واژگان حسی با توجه به زمینه‌ای که در آن قرار می‌گیرد می‌تواند دارای قطبیت مثبت یا منفی باشد. قطبیت یک واژه با توجه به زمینه تغییر می‌کند. این واژه‌ها، مفاهیم مبهم نامیده می‌شوند. تعداد معدودی از پژوهش‌ها به این مسئله پرداخته‌اند [۷] و [۸]. به عنوان مثال در دامنه دارو، "افزایش وزن" یا "کاهش اشتها" مفاهیمی هستند که می‌توانند با توجه به زمینه، قطبیت مثبت یا منفی داشته باشند. نشانه‌های زمینه، نقش اصلی را تعیین قطبیت مفاهیم مبهم دارند. به منظور تعیین قطبیت دقیق‌تر با توجه به زمینه، به سؤالاتی که در ادامه آمده است پاسخ می‌دهیم: مفاهیم مبهم کدام مفاهیم هستند؟ چگونه می‌توان با کمک منابع ساختاریافته دانش معنایی و دانش عرفی خارجی مفاهیم حسی مبهم را غنی‌تر و رفع ابهام نمود؟

ادامه مقاله به صورت بدین صورت سازماندهی شده است، بخش ۲ کارهای مرتبط، بخش ۳ مدل پیشنهادی، بخش ۴ ارزیابی و بخش ۵ نتیجه‌گیری آورده شده است.

۲- کارهای مرتبط

پژوهش‌هایی که از دانش زمینه در فرآیند تحلیل احساس استفاده می‌کنند به دو دسته تقسیم می‌شوند پژوهش‌هایی که از دانش خاص دامنه کمک می‌گیرند مانند [۹] و [۱۰] و پژوهش‌هایی که از دانش عرفی استفاده می‌کنند [۷] و [۸]. از طرفی دیگر تعدادی تلاش می‌کنند این دانش را از روی پیکره‌ها با کمک هستان‌شناسی‌ها به طور دستی [۹] و [۱۰] و یا نیمه خودکار [۱۱] ایجاد نمایند و تعدادی از پایگاه دانش عرفی [۱۲] و [۱۳] و [۷] و [۸] که

^۷ Sentiment Lexicon

ویژگی‌ها استفاده نموده‌اند. تعداد معدودی از پژوهش‌های [۷] و [۸] از ConceptNet برای غنی‌سازی و رفع ابهام واژگان حسی استفاده نموده‌اند که تمرکز این دو پژوهش روی واژگان بوده، و توجهی به چارچوب CLSA و تحلیل احساس مبتنی بر مفهوم نداشته است.

ستیلو [۱۵-۱۷] با کشف ارتباط بین عناصر جمله، مفاهیم اصلی و ارتباطات آنها را شناسایی و در گراف rdf نشان می‌دهد. OntoSentilo یک هستان‌شناسی برای جملات حاوی نظر می‌باشد و SentiloNet یک منبع واژگانی است که امکان ارزیابی نظرات را از طریق رویدادها و موقعیت‌ها فراهم می‌کند.

پژوهش‌های زیادی با روش‌های متعددی مانند الگوریتم قدم زدن تصادفی [۱۸]، تکنیک رفع ابهام کلمات [۱۹]، جفت واژه حسی و ویژگی [۲۰]، انتشار برچسب‌های مقید [۲۱]، یادگیری بدون ناظر با کمک گراف جهت‌دار بدون دور از جفت واژه نظر و جنبه [۲۲]، مدل‌سازی قطبیت با منطق فازی [۵] سعی در تعیین قطبیت واژگان حسی با توجه به دامنه کرده‌اند. این پژوهش‌ها به معنانشناسی مبتنی بر زمینه و اثر نشانه‌های زمینه در تعیین قطبیت کمتر توجه کرده‌اند. اینکه قطبیت بسیاری از کلمات حتی در دامنه خاص باتوجه به زمینه می‌تواند متفاوت باشد. در بین پژوهش تعدادی از آنها [۸] و [۷] و [۲۳] و [۲۴] اشاره به این مسئله کرده‌اند که قطبیت علاوه بر دامنه، به زمینه نیز وابسته است. سیف در پژوهش [۲۳] و [۲۴] یک رویکرد مبتنی بر واژگان جدید با استفاده از یک نمایش زمینه‌ای از کلمات به نام SentiCircles که قادر به ضبط معانی پنهان کلمات از الگوهای هم‌رخداد خود است و براساس آن، جهت‌گیری‌های احساسی خود را بروز می‌کند، را ارائه می‌دهد. بدین شیوه تلاش می‌کنند زمینه را با کمک الگوهای هم‌رخداد در سیستم تحلیل احساس وارد نمایند. قطبیت بسیاری از واژه‌ها فقط وابسته به دامنه نیست، بلکه وابسته به زمینه هم هست. زمینه علاوه بر دانش پس‌زمینه، مفاهیم

به طور عمومی ایجاد شده‌اند، استفاده می‌کنند. پایگاه دانش عرفی، از این جهت که قابل دسترس می‌باشند و از پیش آماده شده‌اند گزینه مناسب‌تری هستند ولی مشکل تنگی داده‌ها دارند و دانش دامنه و دانش تخصصی کمتر در آنها وجود دارد. از طرفی طراحی پایگاه دانش در حوزه خاص به صورت تخصصی با کمک هستان‌شناسی‌ها، کار زمانگیر و دشواری است. پژوهش‌هایی نیز سعی داشته‌اند برای حوزه خاص یک پایگاه دانش ایجاد کنند. این پژوهش‌ها دانش دامنه را در نظر می‌گیرند بیشتر خود به تشکیل هستان‌شناسی و ایجاد پایگاه دانش از داده‌ها می‌پردازند و در ادامه کار از آنها استفاده می‌کنند. نوفرستی [۱۱، ۱۴] یک پایگاه دانش برای دامنه دارو با کمک داده پیوندی به روش نیمه خودکار ایجاد نموده است و برای تحلیل احساس نظرات غیر مستقیم در دامنه دارو استفاده نموده است. هرچه پایگاه دانش غنی‌تری مبتنی بر هستان‌شناسی‌ها در اختیار داشته باشیم، می‌توانیم تحلیل احساس را با دقت بیشتر انجام دهیم و نیاز ما به پیکره‌های بزرگ برای انجام تحلیل احساس کمتر می‌شود.

پژوهش‌هایی که از دانش زمینه و هستان‌شناسی‌ها استفاده نموده‌اند بیشتر برای استخراج ویژگی از این دانش استفاده نموده [۹] و [۱۲] و [۱۳] و [۱۰] و تمرکز آنها بر تحلیل احساس در سطح ویژگی/جنبه می‌باشد و از هستان‌شناسی برای کشف ارتباطات بین ویژگی‌ها استفاده نموده‌اند. به عنوان نمونه [۱۰] هستان‌شناسی دامنه را با کمک دو روش FCA و یادگیری هستان‌شناسی ایجاد نموده و از هستان‌شناسی دامنه برای استخراج ویژگی‌ها استفاده نموده و تحلیل احساس را برای مجموعه داده‌های مربوط به گفتگوهای کوتاه کاربران در سطح ویژگی انجام داده است. علاوه بر این، پژوهش‌هایی که به طور خاص روی داده‌های نقد و بررسی محصولات انجام شده است، بیشتر تمرکز روی استخراج ویژگی و تحلیل احساس در سطح جنبه/ویژگی دارند. آنها از هستان‌شناسی برای کشف ارتباطات بین

سطح کلمه را مورد توجه قرار داده است. ب) همه کارهای مرتبط با فرآیند پردازش زبان طبیعی برای استخراج اطلاعات نظر از متن ضروری دانسته است.

مدل CLSA بر روی پایه‌های محاسباتی پژوهش‌های تحلیل احساس برای تعیین ۸ بخش پردازش زبان طبیعی کلیدی خود تمرکز دارد. ۸ گام کلی را ارائه داده که جالب توجه می‌باشد: تحلیل میکرومتن^۸، تجزیه معنایی^۹، تشخیص ذهنیت^{۱۰}، تشخیص ارجاعات^{۱۱}، تشخیص کنایه‌ها^{۱۲}، شناسایی عنوان^{۱۳}، استخراج جنبه‌ها^{۱۴}، تشخیص قطبیت^{۱۵}. مدل پیشنهادی در این مقاله بر پایه چارچوب CLSA می‌باشد. تجزیه معنایی یا تجزیه مفهومی گام اساسی این چارچوب هست، که در این مدل نیز از آن بهره گرفته شده است.

تفاوت اصلی رویکردهای سنتی تحلیل احساس و رویکردهای مبتنی بر مفهوم این است که رویکردهای سنتی تکیه به قسمت‌هایی دارند که نظر را به صراحت بیان می‌کند و واژگان حسی در آنها وجود دارد ولی در رویکرد مبتنی بر مفهوم روابط بین موجودیت‌های جمله و ارتباط آنها با دانش پس زمینه بسیار مهم است. بنابراین این رویکردها به شناسایی مفاهیم موجود در متن و روابط آنها و روابط متقابل آنها با دانش پس‌زمینه می‌پردازند و تلاش می‌کنند با کمک این روابط، نظراتی که به طور صریح و ضمنی احساسات را می‌رسانند، مورد بررسی قرار دهند.

۲-۲- تجزیه مفهومی

تجزیه متن نوشته شده با زبان طبیعی به مفاهیم را تجزیه مفهومی می‌گویند [۲۸]. یک مفهوم یک موجودیت در متن می‌باشد که می‌تواند اسم، فعل، صفت و قید باشد و همچنین می‌تواند یک مفهوم رویداد^{۱۶} و مفهوم پیچیده باشد. در [۲۹]

هم‌رخداد نیز می‌تواند باشد، توجه به مفاهیم هم‌رخداد کمک قابل توجهی به تعیین قطبیت نظرات ضمنی دارد.

۱-۲- تحلیل احساس در سطح مفهوم

رویکردهای تحلیل احساس به دو دسته کلی تقسیم می‌شوند [۲۵] و [۲۶]: رویکردهای مبتنی بر واژه [۲۷] و رویکردهای یادگیری ماشین که خود به دو دسته با ناظر و بدون ناظر تقسیم می‌شوند. کمبریا در [۴] دسته دیگری با عنوان تحلیل احساس در سطح مفهوم ارائه کرده است. تحلیل احساس در سطح مفهوم تاکید بر تحلیل معنایی متن از طریق هستان‌شناسی و شبکه معنایی دارد که فرصت جمع اطلاعات احساسی و مفهومی مرتبط با نظرات را می‌دهد. با تاکید بر پایگاه دانش معنایی بزرگ، چنین روش‌هایی از استفاده کورکورانه از کلمات کلیدی و شمارش تعداد دفعات وقوع همزمان آنها دوری گزیده و به جای آن بر روی ویژگی‌های ضمنی مرتبط با مفاهیم زبانی تکیه دارند. برخلاف تکنیک‌های نحوی محض، رویکردهای مبتنی بر مفهوم قادر به تشخیص قطبیت که به شیوه‌ای ظریف بیان شده هستند. به عنوان نمونه می‌توان به قطبیت-هایی که به طور صریح بیان نشده‌اند ولی از طریق تحلیل مفاهیم و با پیوند به مفاهیم به طور ضمنی کشف می‌شوند، تشخیص داد.

کمبریا و همکارانش پیشرو در ارائه چنین روش‌هایی هستند و چارچوب CLSA (Concept-Level sentiment analysis) [۲۸] را معرفی کرده‌اند. این چارچوب به عنوان یک مرجع برای پژوهشگران که مایل هستند رویکرد کل نگرانه و آگاه از معنی نسبت تحلیل احساس داشته باشند، می‌باشد. این چارچوب دارای ویژگی‌های زیر می‌باشد: الف) تحلیل متن حاوی نظر در سطح مفهوم و به جای

¹³ Topic Spotting

¹⁴ Aspect extraction

¹⁵ Polarity detection

¹⁶ Event concept

⁸ microtext analysis

⁹ semantic parsing

¹⁰ Subjectivity detection

¹¹ anaphora resolution

¹² sarcasm detection

عاطفی مجموعه مفاهیم^{۱۹} Open Mind استفاده می‌کند و آنها را در قالب قابل دسترس برای ماشین و قابل پردازش نمایش می‌دهد [۳۵]. این منبع شامل بیش از ۵۰۰۰۰ مفهوم می‌باشد. این مفاهیم ترکیبی از مفاهیم یک کلمه‌ای و یا مفاهیم چند کلمه می‌باشد، ۳۰۰۰۰ مفهوم چند کلمه‌ای دارد. به عنوان نمونه یک پرس‌وجو برای مفهوم "good_customer", اطلاعات حسی و مقدار قطبیت مثبت را نشان می‌دهد که در شکل ۱ نمایش داده شده است.

۲-۴- شبکه معنایی ConceptNet

ConceptNet یک پایگاه دانش عرفی چندزبانه، که کلمات و عباراتی که افراد استفاده می‌کنند و روابط عرفی بین آنها را نمایش می‌دهد [۳۶] و [۳۷] و [۳۸]. دانش عرفی موجود در آن با مشارکت مردم عادی در اینترنت گردآوری شده است. این شبکه معنایی بزرگ متشکل از تعداد زیادی از مفاهیم عرفی است. یک مفهوم، یک موجودیت است که دانش عرفی را تعریف می‌کند [۲۲].

مفاهیم ساده را با روش مبتنی بر گراف استخراج کرده است و پژوهش [۳۰] و [۳۱] مجموعه‌ای از قوانین پیشنهاد داده است که مفاهیم پیچیده را با کمک تجزیه وابستگی کشف نموده است. بنابراین یک مفهوم به صورت یک کلمه‌ای، دو کلمه یا چند کلمه‌ای استخراج می‌شود. بعد از استخراج مفاهیم، کیسه مفاهیم^{۱۷} به جای کیسه کلمات^{۱۸} برای ادامه پردازش‌ها در دسترس می‌باشد [۲۹] و می‌تواند به عنوان خوراک یک الگوریتم استدلال عرفی بکارگرفته شود. کیسه‌ای از مفاهیم خیلی بهتر از کیسه‌ای از کلمات است. به عنوان نمونه جمله "I contact to phone services to help me for discharging my device" مفاهیم phone_service, contact_to_phone_service, discharge_device تجزیه می‌شود. استخراج مفهوم یکی از مراحل کلیدی تحلیل احساس در سطح مفهوم می‌باشد [۲۸] و [۳۰]. رویکردهای موجود برای بدست آوردن مفاهیم از متن می‌توانند به دو دسته طبقه‌بندی بشوند: رویکردهای قواعد زبانی [۲۹] و [۳۱] و [۳۰] و رویکردهای آماری [۳۲]. ترکیب رویکردهای قواعد زبانی و آماری فرآیند استخراج مفهوم را بهبود می‌دهد.

۲-۳- SenticNet

SenticNet [۳۳] یک منبع برای تحلیل نظر در سطح مفهوم می‌باشد. منبع واژگانی که بوسیله خوشه‌بندی مدل فضای بردار دانش عرفی عاطفی استخراج شده از ConceptNet، ساخته شده است [۳۴]. این منبع لیستی از مفاهیم به همراه قطبیت آنها تولید نموده است و مقدار قطبیت در محدوده ۱- تا ۱+ می‌باشد. مقدار قطبیت مفاهیم ارائه شده در این فرهنگ لغت با بهره‌گیری از هوش مصنوعی و تکنیک‌های وب معنایی محاسبه می‌شود. به طور ویژه، SenticNet از کاهش ابعاد برای محاسبه ارزش

^{۱۹} <http://commons.media.mit.edu/en/>

^{۱۷} Bag-of-concepts

^{۱۸} Bag-of-words

```

- <rdf:RDF>
- <rdf:Description rdf:about="http://sentic.net/api/en/concept/good_customer">
  <rdf:type rdf:resource="http://sentic.net/api/concept"/>
  <text>good customer</text>
- <semantics>
  <concept rdf:resource="http://sentic.net/api/en/concept/save_money"/>
  <concept rdf:resource="http://sentic.net/api/en/concept/buy_cheaper"/>
  <concept rdf:resource="http://sentic.net/api/en/concept/better_price"/>
  <concept rdf:resource="http://sentic.net/api/en/concept/sale"/>
  <concept rdf:resource="http://sentic.net/api/en/concept/on_sale"/>
</semantics>
- <sentic>
  <pleasantness rdf:datatype="http://w3.org/2001/XMLSchema#float">0.799</pleasantness>
  <attention rdf:datatype="http://w3.org/2001/XMLSchema#float">-0.92</attention>
  <sensitivity rdf:datatype="http://w3.org/2001/XMLSchema#float">0</sensitivity>
  <aptitude rdf:datatype="http://w3.org/2001/XMLSchema#float">0.896</aptitude>
</sentic>
- <moodtags>
  <concept rdf:resource="http://sentic.net/api/en/concept/surprise"/>
  <concept rdf:resource="http://sentic.net/api/en/concept/admiration"/>
</moodtags>
- <polarity>
  <value>positive</value>
  <intensity rdf:datatype="http://w3.org/2001/XMLSchema#float">0.874</intensity>
</polarity>
</rdf:Description>
</rdf:RDF>

```

شکل ۱ مفهوم "good_customer" در SenticNet [۳۳]

روابط در شبکه معنایی ConceptNet با توجه متونی کاربران عادی نوشته‌اند تعیین شده است ولی روابط در WordNet توسط متخصصین تعریف شده است. بیشتر تمرکز آن روی روابط عرفی بین مفاهیم می‌باشد و هدف از تشکیل ConceptNet استدلال دانش عرفی^{۲۰} می‌باشد. ConceptNet نسبت گره‌های شبکه معنایی WordNet که از اقلام صرفاً واژگانی (کلمات و عبارات ساده با معنای اتمیک) گسترش می‌یابد و شامل مفاهیم مرکب مرتبه بالاتر مانند "talking to someone" و "chatting with friend" برای نشان دادن دانش طیف وسیعی از مفاهیم موجود در زندگی روزمره می‌باشد. علاوه بر این، روابط معناشناسی از سه گانه مترادف^{۲۱}، IsA و PartOf، که در WordNet وجود دارد به بیش از بیست رابطه معنایی از جمله EffectOf (علیت^{۲۲})، SubeventOf

بزرگترین منبع عرفی قابل استفاده برای ماشین که شامل بیش از ۲۵۰۰۰۰ رابطه است. ConceptNet در دسترس عموم است و می‌تواند برای استخراج‌های مختلف کاوش در متن استفاده شود. از گره‌هایی (مفاهیم) که با یال‌ها (روابط بین مفاهیم) وصل می‌شوند، تشکیل شده است.

گراف معنایی ConceptNet نشان‌دهنده اطلاعات از پیکره OpenMind می‌باشد که به شکل یک گراف جهت‌دار می‌باشد، به گونه‌ای که گره‌های آن مفاهیم و یال‌های برچسب زده شده اظهارات عرفی که مفاهیم را بهم وصل می‌کنند، می‌باشد. به عنوان مثال در ConceptNet، phone رابطه UsedFor با making a phone call دارد و cellphone رابطه Isa type of با phone دارد و غیره.

²² causality

²⁰ Common sense reasoning

²¹ synonym

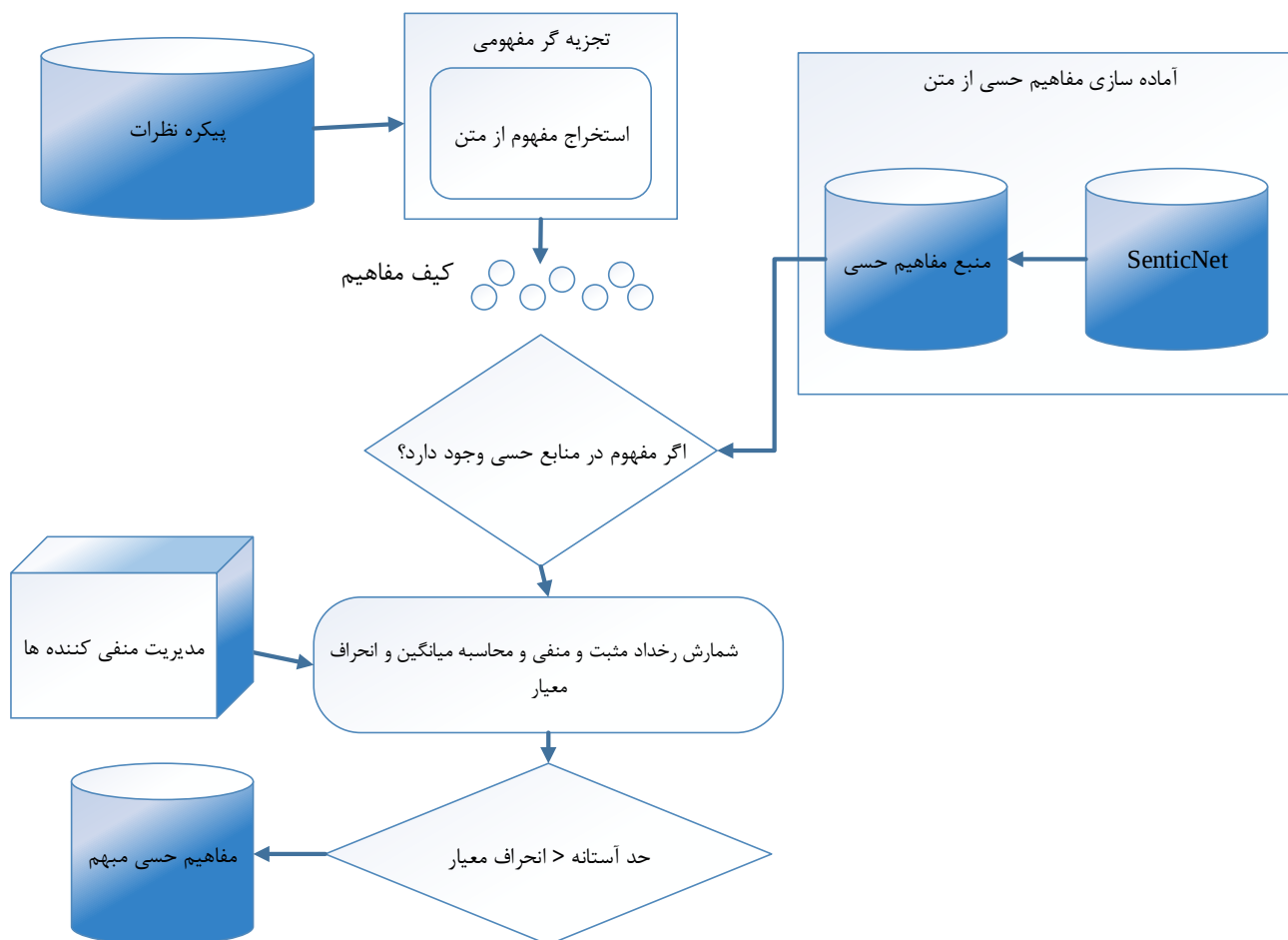
اشاره‌ای نکرده‌اند. در تحلیل احساس معنایی و مبتنی بر مفهوم، مفاهیم به جای کلمات از اهمیت برخوردار هستند. مفاهیم حسی مبهم در یک حوزه قطبیت ثابتی ندارند و با توجه به زمینه، قطبیت آنها تغییر می‌کند. از آنجایی که قطبیت این دسته از مفاهیم وابسته به زمینه می‌باشد، برای اینکه تعیین قطبیت بهتر انجام شود، باید به زمینه توجه نمود. به منظور یافتن چنین مفاهیمی از متن، ۴ گام باید انجام شود. شکل ۲ مدل پیشنهادی برای تشخیص مفاهیم حسی مبهم را نشان می‌دهد.

(سلسله مراتب رویداد^{۲۳}) CapableOf (توانایی عامل^{۲۴})، MotivationOf (تاثیر)، LocationOf، PropertyOf، EffectOf گسترش پیدا کرده است [۳۸]. همچنین دانش ConceptNet ماهیت غیر رسمی، قابل انکار و از لحاظ عملی ارزشمند دارد.

۳- مدل پیشنهادی

۳-۱- یافتن مفاهیم حسی مبهم

پژوهش‌های [۱۳] و [۸] و [۷] روش‌هایی برای شناخت واژگان حسی مبهم ارائه داده‌اند ولی به مفاهیم حسی مبهم



شکل ۲ مدل پیشنهادی برای تشخیص مفاهیم حسی مبهم

NCnt تعداد رخدادهای منفی، NScore میزان امتیاز منفی جمله

می باشد.

$$\mu_i = \frac{PCnt(c_i) * PScore(c_i) - NCnt(c_i) * NScore(c_i)}{PCnt(c_i) + NCnt(c_i)} \quad 1$$

$$\delta i = \sqrt{\frac{\left(\left(\mu_i - PScore(c_i) \right)^2 * PCnt(c_i) + \left(\mu_i - NScore(c_i) \right)^2 * NCnt(c_i) \right)}{(PCnt(c_i) + NCnt(c_i))}} \quad 2$$

۳-۱-۳- جداکردن مفاهیم حسی مبهم از غیر مبهم

در صورتی که انحراف معیار مجموعه رخدادهای مفهوم از حد آستانه بالاتر باشد، بدین معنی هست که قطبیت مفهوم دارای پراکندگی است. در نتیجه مفهوم مبهم می باشد و قطبیت آن باید با توجه به زمینه اصلاح گردد. حد آستانه h به طور تجربی با توجه به پیکره بدست می آید که مقدار ۰٫۸۵ برای داده های این پژوهش بدست آمده است.

$$\delta i \geq h$$

بدین ترتیب منبعی از مفاهیم حسی که با توجه به زمینه مبهم هستند در این بخش تعیین و در بخش بعدی از آنها رفع ابهام می شود.

۳-۲- رفع ابهام از مفاهیم حسی مبهم

در این بخش روشی تشریح می گردد که مفاهیم با توجه به زمینه تعیین قطبیت می شوند و مسئله این پژوهش رفع ابهام معنایی کلمات که وظیفه استخراج معنی صحیح واژه های که چندمعنا دارد، نمی باشد. به منظور رفع ابهام از مفاهیم حسی مبهم با محوریت زمینه توجه به نشانه های زمینه ضروری می باشد. نشانه های زمینه، به تعیین قطبیت مفهوم مبهم کمک می کنند. این نشانه می تواند نشانه های متنی یا غیرمتنی باشند. در این پژوهش از نشانه های متنی استفاده می شود. نشانه ها مفاهیمی هستند که در اطراف مفهوم مبهم رخ می دهند و به

۳-۱-۱- استخراج لیست مفاهیم موجود در متن

در این بخش، مفاهیم موجود از متن پیکره نقد و بررسی محصولات با کمک تجزیه گر مفهومی استخراج می شود. یک مفهوم یک موجودیت در متن می باشد که می تواند اسم، فعل، صفت و قید باشد و همچنین می تواند یک مفهوم رویداد^{۲۵} و یا یک مفهوم پیچیده باشد. بنابراین یک مفهوم به صورت یک کلمه ای، دو کلمه یا سه کلمه ای با روش های [۲۹]، [۳۰] و [۳۱] استخراج می شود. بعد از استخراج مفاهیم، کیسه مفاهیم به جای کیسه کلمات برای ادامه پردازش ها در دسترس می باشد.

۳-۱-۲- آماده سازی منبع مفاهیم حسی

مفاهیم حسی از روی منبع واژگان حسی SenticNet گردآوری می شود. این منبع به دلیل مبتنی بر مفهوم بودن انتخاب گردیده است. علاوه بر آن معادل های معنایی تعریف شده در این منبع نیز اضافه شده است تا پوشش خوبی نسبت به داده ها بدهد.

۳-۱-۴- شمارش رخدادهای مثبت و منفی و محاسبه میانگین و انحراف معیار

هر مفهوم حسی ممکن است در یک جمله با قطبیت مثبت و یا در جمله با قطبیت منفی رخ بدهد. توزیع هر مفهوم حسی محاسبه می شود. دو پارامتر آماری که کمک به تعیین ابهام می کند، میانگین و انحراف معیار می باشد. انحراف معیار بالا نشان دهنده این است که مفهوم، مبهم است. در این بخش تعداد رخدادهای مفاهیم حسی در جملات منفی و مثبت می شماریم و در ادامه میانگین و انحراف معیار را برای این مفهوم محاسبه می نماییم. رابطه ۱ و ۲ نحوه محاسبه میانگین و انحراف معیار را نشان می دهد. در این رابطه c_i مفهوم، PCnt تعداد رخدادهای مثبت، PScore میزان امتیاز مثبت جمله،

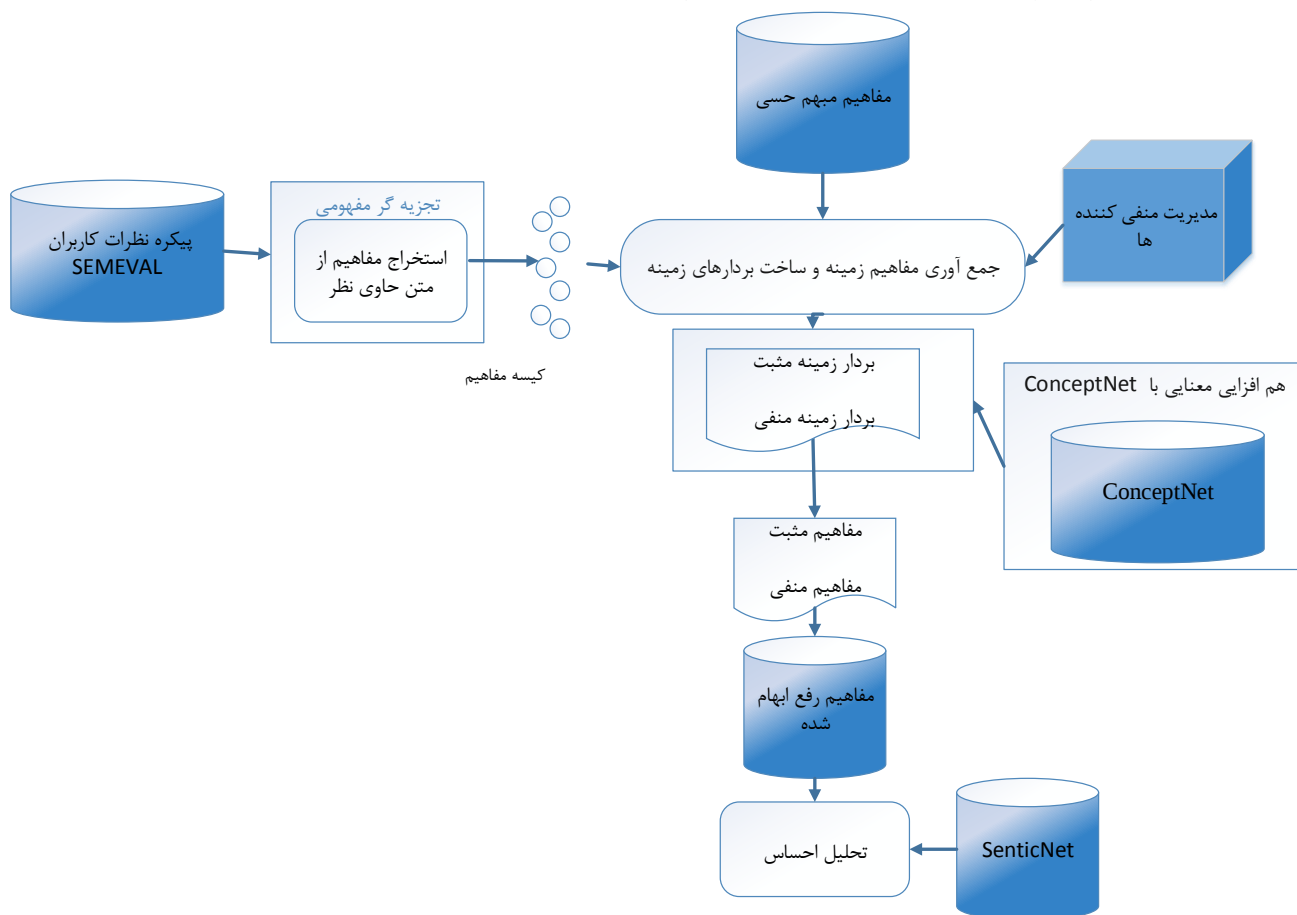
²⁵ Event concept

مبهم رخداد پربسامدی داشته باشند، به عنوان مفاهیم زمینه به حساب آورده می‌شوند. برای هر مفهوم مبهم، لیست مفاهیم که در جملات مثبت هم وقوع هستند بردار زمینه مثبت و لیست مفاهیم که در جملات منفی هم وقوع هستند بردار زمینه منفی را تشکیل می‌شود. بردار زمینه مفاهیم اصلی اطراف مفهوم مبهم را نشان می‌دهد.

صورت بردار زمینه در نظر گرفته می‌شود. علاوه بر آن اطلاعات معنایی نهفته در دانش عرفی به منظور غنی‌تر کردن نشانه‌های زمینه مورد استفاده قرار می‌گیرد و به همین منظور ConceptNet بکار گرفته می‌شود. شکل ۳ نمای کلی مدل پیشنهادی برای رفع ابهام مفاهیم حسی مبتنی بر زمینه را نشان می‌دهد.

۳-۲-۱- جمع‌آوری مفاهیم زمینه و ایجاد بردار زمینه

مفاهیم زمینه، مفاهیمی هستند که فرکانس رخداد همزمان آنها در پیکره با مفهوم مبهم بالا می‌باشد. مفاهیمی که با مفهوم



شکل ۳ مدل پیشنهادی برای رفع ابهام مفاهیم حسی مبتنی بر زمینه

مورد انتخاب قرار می‌گیرد. با بکارگیری دانش معنایی و دانش عرفی نهفته در ConceptNet مدل‌سازی زمینه به طور غنی‌تر و انتزاعی‌تر انجام می‌شود و دقت و فراخوانی بیشتر در روش-های تحلیل احساس در تعیین قطبیت بدست می‌آید. بدین ترتیب هریک از مفاهیم یک کلمه‌ای و چند کلمه‌ای استخراج شده از جملات گسترش پیدا می‌کند. روابط ConceptNet به صورت دسته‌بندی شده در جدول ۱ آورده شده است.

جدول ۱ دسته بندی روابط پر اهمیت در ConceptNet

Things: IsA, part of, MemberOf, HasA, HasProperty, Synonym, Antonym, DerivedFrom, DefinedAs, TranslationOf
Functional: UsedFor, CapableOf
Spatial: AtLocation
Event: HasSubevent, HasFirstSubevent, HasLastSubevent, HasPrerequisite,
Causal: Causes, Desires
Motivation: MotivatedByGoal, ObstructedBy
Other: RelatedTo, CreatedBy

۳-۴- مدیریت منفی کننده ها

منفی‌کننده‌ها نقش کلیدی در تعیین قطبیت دارند و می‌توانند جهت قطبیت را تغییر بدهند، بدین معنی که مثبت را به منفی و منفی را به مثبت تبدیل کنند. بنابراین تشخیص نفی و اعمال آن در سیستم‌های تحلیل احساس پر اهمیت می‌باشد. در رویکرد تحلیل احساس در سطح مفهوم، منفی‌کننده باید شناخته شود. در این مقاله با روش لیست کلمات کلیدی رایج [۲۱] و [۳۹]، نفی موجود در جملات شناسایی و در فرآیند رفع ابهام از آن استفاده می‌کنیم.

۴- ارزیابی

به منظور ارزیابی مدل پیشنهادی از پیکره استاندارد نقد و بررسی محصولات الکترونیکی با نام Semeval از آدرس <https://github.com/davidsbatista/Aspect-Based-Sentiment-Analysis/tree/master/datasets/ABSA-SemEval2015> استفاده شده است. این پیکره حاوی ۲۵۰۰ جمله نقدوبررسی کاربران در

مفاهیم هم‌رخداد در جملات مثبت، بردار ویژگی مثبت و مفاهیم هم‌رخداد در جملات منفی بردار ویژگی منفی را تشکیل می‌دهد. مفاهیم هم‌رخداد در واقع نشانه‌های زمینه مفهوم مبهم می‌باشند. از تکنیک بیزین ساده برای تعیین قطبیت دقیق با توجه مفاهیم زمینه استفاده می‌شود. رابطه ۳ نحوه تعیین قطبیت مفهوم مبهم را نشان می‌دهد.

$$c = \{c1, \dots, cn\}$$

$$p(C - /c) = \frac{p(C -) \prod_{i=1}^n p(ci / C -)}{\prod_{i=1}^n p(ci)}$$

$$p(C + /c) = \frac{p(C +) \prod_{i=1}^n p(ci / C +)}{\prod_{i=1}^n p(ci)}$$

در واقع برای هریک از مفاهیم مبهم، مفاهیم زمینه‌ای که به تعیین قطبیت مفاهیم حسی مبهم کمک می‌کند، تعیین می‌شود. بسیاری از مفاهیم مبهم در منبع مفاهیم حسی SenticNet دارای قطبیت ثابت می‌باشند، که در بسیاری از موارد قابل قبول نیست و قطبیت علاوه بر وابستگی به دامنه به زمینه نیز وابسته می‌باشد. مفاهیم زمینه و معانی و روابط معنایی آنها کمک به شناخت بهتر مفهوم مبهم و تعیین قطبیت دقیق‌تر آن با توجه به زمینه می‌باشد.

۳-۳- هم افزایی معنایی با کمک دانش عرفی

با کمک روابط در پایگاه دانش معنایی از منبع خارجی، مفاهیم موجود در زمینه هم‌افزایی می‌شود. به مجموعه مفاهیم استخراج شده از جملات، گروه‌های معنایی با کمک پایگاه دانش معنایی ConceptNet، اضافه می‌شود. به عنوان مثال مفهومی مانند "laptop" در ConceptNet دارای تعداد ۲۴ گروه مرتبط می‌باشد و مفهوم "going on the internet" رابطه "Things that require" با مفهوم "laptop" دارد. از آنجاکه تعداد یال‌های مرتبط زیاد می‌باشد، یال‌هایی که وزن آنها بالا می‌باشد

مورد محصولات الکترونیکی می‌باشد و قطبیت هر جمله با برچسب‌های positive، negative و neutral تعیین شده است. این پیکره به صورت استاندارد در شبکه اینترنت وجود دارد و توسط پژوهشگران برای ارزیابی پژوهش‌های تحلیل احساس مورد استفاده قرار می‌گیرد. علاوه بر این داده‌ها به دو دسته آموزش و آزمایش تقسیم شده است. ۹۰ درصد داده‌ها برای آموزش و ۱۰ درصد برای تست در نظر گرفته شده است. برای پیش پردازش داده‌ها کارهایی که انجام شده‌اند عبارتند از: نقطه‌گذاری صحیح برای تشخیص بهتر جملات، اصلاح کلمات و حذف حروف تکراری از واژه‌ها، اصلاح املای بعضی از کلمات، اصلاح علامت-های؟!، جایگذاری اختصارها، اصلاح پرانتزها، شکستن جملات پیچیده به جملات ساده. علاوه بر آن تشخیص موجودیت‌های نامدار و جانشینی نیز انجام شده است. بیشتر پردازش‌ها به کمک بسته NLTK صورت گرفته است. همچنین مفاهیم هر جمله با کمک تجزیه‌گر مفهومی^{۲۶} پیاده‌سازی شده توسط گروه سنتیک استخراج شده است. از منبع SenticNet و معادل‌های معنایی آن برای ایجاد منبع مفاهیم حسی استفاده گردید. پیوند بین مفاهیم با کمک ConceptNet با فرمت داده‌های پیوندی انجام شد.

۴-۱- تحلیل نتایج

همانطور که در بخش‌های قبل مورد بررسی قرار گرفت، روش‌های تحلیل احساس مبتنی بر واژه بر اساس منبعی از واژگان حسی هستند. منابع موجود ایستا هستند، بدین معنی که به هریک از واژه‌ها میزان قطبیتی ثابت اختصاص داده شده است. روش‌های مبتنی بر یادگیری ماشین نیز کمتر به مفاهیم معنایی جمله و پیوندهایی که در پس‌زمینه وجود دارد توجه کرده‌اند و به جای آن بیشتر به کلیدواژه‌ها و رخداد همزمان آنها پرداخته‌اند. مدل پیشنهادی ترکیبی از روش‌های یادگیری ماشین و مبتنی بر دانش می‌باشد.

در تعیین قطبیت برخی از مفاهیم در نظر گرفتن نشانه‌های زمینه ضروری است. در مدل پیشنهادی روشی ارائه شده است که به جای استفاده از کیسه کلمات و bigram و n-gram، از مفاهیم جمله و ارتباطاتی که بین عناصر جمله وجود دارد و همچنین هم‌فزایی معنایی آنها با بکارگیری منبع دانش خارجی، استفاده می‌کند. مفاهیم جمله به عنوان نشانه‌های زمینه در نظر گرفته شده است و با کمک نشانه‌های زمینه از مفاهیم حسی مبهم رفع ابهام می‌شود. مفاهیمی که در اطراف مفاهیم مبهم بیشترین فرکانس رخداد را دارد و نقش کلیدی در شناخت زمینه دارد، در نظر گرفته شده است. در نمونه‌هایی از جملات، مفاهیم جمله که توسط تجزیه‌گر مفهومی استخراج شده، مفهوم حسی مبهم شناسایی شده و مفاهیم زمینه را نمایش داده شده است. جمله Save your money and go for a better device. مشتمل بر مفاهیم go_money، go_for_device، better_device، save_device، save_money می‌باشد، که با توجه به حضور better_device به نظر می‌رسد قطبیت جمله باید مثبت باشد ولی better_device در مجموعه مفاهیم مبهم قرار گرفته و به دلیل حضور مفهوم save_money در زمینه قطبیت جمله منفی نتیجه‌گیری می‌شود.

پانگ، لی و وایتایانان در سال ۲۰۰۲ [۴۰] از کیسه کلمات (unigrams) به عنوان ویژگی در طبقه‌بند بیزین ساده استفاده کرده‌اند و به دقت ۸۱٫۰ برای طبقه‌بندی نقد و بررسی فیلم در دو کلاس مثبت و منفی رسیده‌اند. در این پژوهش برای ارزیابی مدل پیشنهادی، روش ارائه شده در مقاله پانگ، لی و وایتایانان به عنوان مبنا و برای داده‌های محصولات الکترونیک پیاده‌سازی و به دقت ۷۴٫۱۳ دست یافته شده است. برای ارزیابی کارایی از معیارهای دقت، بازخوانی و F-Measure استفاده می‌شود. یکبار بخشی از مدل پیشنهادی

²⁶ <http://sentic.net>

مثبتی روی مدل پیشنهادی می‌گذارد. علاوه بر آن دقت مدل پیشنهادی به شناخت نفی و مدیریت منفی‌کننده‌ها نیز وابسته است و هرچه نفی بهتر تشخیص داده شود، دقت مدل نیز بالاتر می‌شود. از طرفی هم پایگاه دانش معنایی قوی پشتیبانی بهتری از داده‌ها می‌کند و مشکل تنگی داده‌ها کمتر دیده می‌شود. تمرکز این مقاله بیشتر روی مفاهیم عرفی و رایج است که کاربران هنگام بیان نظرات خود هنگام خرید استفاده می‌کنند، می‌باشد.

برای پیکره مورد نظر با کیسه مفاهیم پیاده‌سازی شده و دقت ۷۹,۱۷ دریافت نموده‌ایم. بار دیگر کل مدل پیشنهادی با کیسه مفاهیم و پیوند با دانش عرفی برای پیکره مورد نظر اعمال و به دقت ۸۰,۰۵ دست یافته شد. هر کدام از دقت‌ها در مقایسه روش کیسه کلمات که برای مقایسه مدل پیشنهادی به عنوان روش پایه در نظر گرفته شده است به پیشرفت خوبی دست پیدا کرده‌اند.

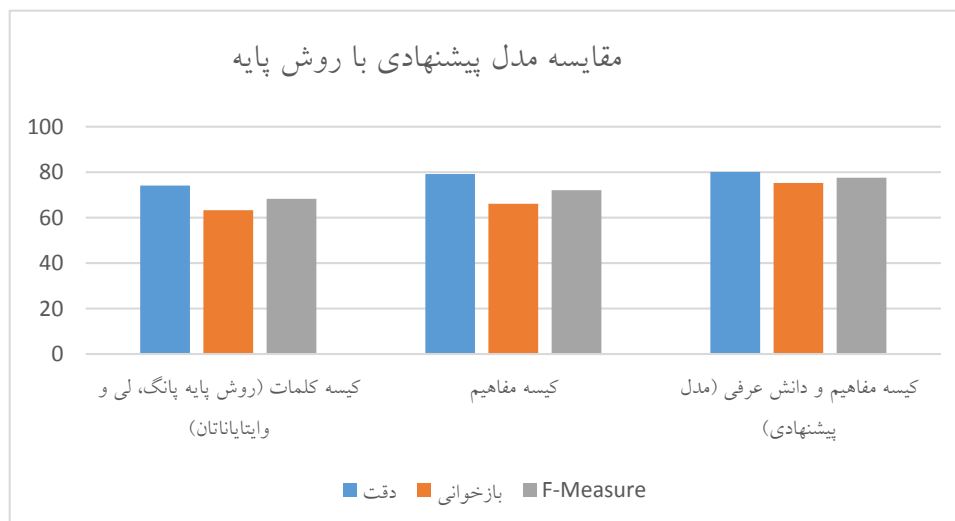
دقت روش پیشنهادی وابستگی زیادی به دقت تجزیه‌گر مفهومی دارد، هرچه دقت تجزیه‌گر مفهومی بیشتر باشد تاثیر

جدول ۲ نمونه‌هایی از رفع ابهام مفاهیم حسی مبهم

مفهوم زمینه	مفهوم حسی مبهم	کیسه مفاهیم	قطبیت جمله	جمله
worth (مفهوم زمینه منفی)	extra dollar	be_worth_dollar, few_dollar, extra_dollar, dollar, worth, be_worth	مثبت	It was worth the few extra dollars.
save_money (مفهوم زمینه منفی)	better_device	go_money, save_money, go_for_device, better_device, device, save	منفی	Save your money and go for a better device.

جدول ۳ مقایسه معیارهای ارزیابی در مدل پیشنهادی

کیسه کلمات (روش پایه پانگ، لی و وایتایاناتان)	کیسه مفاهیم	کیسه مفاهیم و دانش عرفی (مدل پیشنهادی)	
74.13	79.17	80.05	دقت
63.25	66.09	75.23	بازخوانی
68.25	72.04	77.56	F-Measure



شکل ۴ نمودار مقایسه دقت مدل پیشنهادی با روش پایه

۵- نتیجه گیری

در این مقاله مدلی برای رفع ابهام از قطبیت بر مبنای کیسه مفاهیم و دانش عرفی ارائه شد. این پژوهش تلاش کرده است در چارچوب مبتنی بر مفهوم، مفاهیم مبهم را با توجه به زمینه شناسایی و رفع ابهام نماید. کیسه مفاهیم با نگاه عمیق‌تری متن را مورد تجزیه قرار می‌دهد.

وارد نمودن دانش زمینه علی‌الخصوص دانش عرفی در فرآیند تحلیل احساس، باعث افزایش دقت سیستم تحلیل احساس می‌شود. به منظور بازنمایی دانش عرفی از پایگاه دانش معنایی ConceptNet استفاده شده است. استخراج مفاهیم و پیوند بین مفاهیم درک عمیق‌تری را نسبت به متن می‌دهد و مقدمه‌ای برای تحلیل نظرات ضمنی می‌باشد.

شناسایی مفاهیم مبهم با کمک پیکره آموزشی و روش‌های آماری انجام شده است و منبع مبتنی بر مفهوم SenticNet، به عنوان پایه در نظر گرفته شده است. زمینه‌سازی و رفع ابهام از مفاهیم حسی مبهم با روش بیزین ساده انجام می‌شود. بردارهای زمینه برای مفهوم متن و مفاهیم پیوند خورده با پایگاه دانش عرفی ایجاد می‌شود. نتایج آزمایش‌ها نشان می‌دهد که روش پیشنهاد شده به دقت نسبتاً خوبی رسیده است. می‌توانیم نتیجه بگیریم که مدل در طبقه‌بندی قطبیت نقد و بررسی محصولات به خوبی عمل کرده است. بنابراین در کارهای آینده پیشنهاد می‌شود که مدل را روی تحلیل نظرات ضمنی و غیرمستقیم به طور متمرکزتر بسط داده شود. علاوه بر این نیاز هست روی روش‌هایی که دقت تجزیه مفهومی را افزایش می‌دهد، کار شود.

منابع

15. Recupero, D.R., et al., Sentilo: frame-based sentiment analysis. *Cognitive Computation*, 2015. 7(2): p. 211-225.
16. Presutti, V., et al., Sentilo: Semantic Web-based Sentiment Analysis.
17. Recupero, D.R., et al. A semantic web based core engine to efficiently perform sentiment analysis. in *European Semantic Web Conference*. 2014. Springer.
18. Tan, S. and Q. Wu, A random walk algorithm for automatic construction of domain-oriented sentiment lexicon. *Expert Systems with Applications*, 2011. 38(10): p. 12094-12100.
19. Hung, C. and S.-J. Chen, Word sense disambiguation based sentiment lexicons for sentiment classification. *Knowledge-Based Systems*, 2016. 110: p. 224-232.
20. Kansal, H. and D. Toshniwal, Aspect based summarization of context dependent opinion words. *Procedia Computer Science*, 2014. 35: p. 166-175.
21. Huang, S., Z. Niu, and C. Shi, Automatic construction of domain-specific sentiment lexicon based on constrained label propagation. *Knowledge-Based Systems*, 2014. 56: p. 191-200.
22. Ofek, N., et al., Unsupervised Commonsense Knowledge Enrichment for Domain-Specific Sentiment Analysis. *Cognitive Computation*, 2016. 8(3): p. 467-477.
23. Saif, H., et al. Semantic patterns for sentiment analysis of Twitter. in *International Semantic Web Conference*. 2014. Springer.
24. Saif, H., et al., Contextual semantics for sentiment analysis of Twitter. *Information Processing & Management*, 2016. 52(1): p. 5-19.
25. Medhat, W., A. Hassan, and H. Korashy, Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, 2014. 5(4): p. 1093-1113.
26. Ravi, K. and V. Ravi, A survey on opinion mining and sentiment analysis: Tasks, approaches and applications. *Knowledge-Based Systems*, 2015. 89: p. 14-46.
27. Gilbert, C.H.E. Vader: A parsimonious rule-based model for sentiment analysis of social media text. in *Eighth International Conference on Weblogs and Social Media (ICWSM-14)*. Available at (20/04/16) <http://comp.social.gatech.edu/papers/icwsm14.vader.hutto.pdf>. 2014.
28. Cambria, E., et al. The CLSA model: a novel framework for concept-level sentiment analysis. in *International Conference on Intelligent Text Processing and Computational Linguistics*. 2015. Springer.
29. Rajagopal, D., et al. A graph-based approach to commonsense concept extraction and semantic similarity detection. in *Proceedings of the 22nd*
1. Liu, B., Sentiment analysis and opinion mining. *Synthesis lectures on human language technologies*, 2012. 5(1): p. 1-167.
2. Montoyo, A., P. Martínez-Barco, and A. Balahur, Subjectivity and sentiment analysis: An overview of the current state of the area and envisaged developments. *Decision Support Systems*, 2012. 53(4): p. 675-679.
3. Cambria, E., et al., New avenues in opinion mining and sentiment analysis. *IEEE Intelligent Systems*, 2013. 28(2): p. 15-21.
4. Cambria, E. An introduction to concept-level sentiment analysis. in *Mexican International Conference on Artificial Intelligence*. 2013. Springer.
5. Dragoni, M., A.G. Tettamanzi, and C. da Costa Pereira, Propagating and aggregating fuzzy polarities for concept-level sentiment analysis. *Cognitive Computation*, 2015. 7(2): p. 186-197.
6. Blitzer, J., M. Dredze, and F. Pereira. Biographies, bollywood, boom-boxes and blenders: Domain adaptation for sentiment classification. in *ACL*. 2007.
7. Weichselbraun, A., S. Gindl, and A. Scharl, Enriching semantic knowledge bases for opinion mining in big data applications. *Knowledge-based systems*, 2014. 69: p. 78-85.
8. Gindl, S., A. Weichselbraun, and A. Scharl, Extracting and Grounding Contextualized Sentiment Lexicons. 2013.
9. Shein, K.P.P. Ontology based combined approach for sentiment classification. in *Proceedings of the 3rd International Conference on Communications and information technology*. 2009. World Scientific and Engineering Academy and Society (WSEAS).
10. Kontopoulos, E., et al., Ontology-based sentiment analysis of twitter posts. *Expert systems with applications*, 2013. 40(10): p. 4065-4074.
11. Noferesti, S. and M. Shamsfard, Resource construction and evaluation for indirect opinion mining of drug reviews. *PloS one*, 2015. 10(5): p. e0124993.
12. Sureka, A., et al. Generating domain-specific ontology from common-sense semantic network for target specific sentiment analysis. in *Proceedings of the fifth international conference of the Global WordNet Association*. Mumbai, India. 2010.
13. Agarwal, B., et al., Sentiment analysis using common-sense and context information. *Computational intelligence and neuroscience*, 2015. 2015: p. 30.
14. Noferesti, S. and M. Shamsfard, Using Linked Data for polarity classification of patients' experiences. *Journal of biomedical informatics*, 2015. 57: p. 6-19.

- common sense knowledge. in Recent advances in natural language processing. 2007. Citeseer.
37. Speer, R. and C. Havasi, ConceptNet 5: A large semantic network for relational knowledge, in The People's Web Meets NLP. 2013, Springer. p. 161-176.
 38. Liu, H. and P. Singh, ConceptNet—a practical commonsense reasoning tool-kit. BT technology journal, 2004. 22(4): p. 211-226.
 39. Marrese-Taylor, E., J.D. Velásquez, and F. Bravo-Marquez, A novel deterministic approach for aspect-based opinion mining in tourism products reviews. Expert Systems with Applications, 2014. 41(17): p. 7764-7775.
 40. Pang, B., L. Lee, and S. Vaithyanathan. Thumbs up?: sentiment classification using machine learning techniques. in Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10. 2002. Association for Computational Linguistics.
 - International Conference on World Wide Web. 2013. ACM.
 30. Agarwal, B., et al., Concept-level sentiment analysis with dependency-based semantic parsing: a novel approach. Cognitive Computation, 2015. 7(4): p. 487-499.
 31. Poria, S., et al. Dependency-based semantic parsing for concept-level text analysis. in International Conference on Intelligent Text Processing and Computational Linguistics. 2014. Springer.
 32. Kim, H.K., H. Kim, and S. Cho, Bag-of-concepts: Comprehending document representation through clustering words in distributed representation. Neurocomputing, 2017. 266: p. 336-352.
 33. SenticNet. Available from: <http://sentic.net/>.
 34. Poria, S., et al., Sentic patterns: Dependency-based rules for concept-level sentiment analysis. Knowledge-Based Systems, 2014. 69: p. 45-63.
 35. Cambria, E. and A. Hussain, Sentic computing: a common-sense-based framework for concept-level sentiment analysis. Vol. 1. 2015: Springer.
 36. Havasi, C., R. Speer, and J. Alonso. ConceptNet 3: a flexible, multilingual semantic network for