



دانشگاه صنعتی امیرکبیر
(پلی تکنیک تهران)
دانشکده مهندسی کامپیوتر و فناوری اطلاعات

عقیده کاوی و تحلیل احساسات

نگارش
زهرا یوسفی
تینا غلامی

استاد درس
دکتر احسان ناظر فرد

خرداد ماه ۱۳۹۹

چکیده

در سال‌های اخیر، افزایش گسترده‌ای در میزان داده‌های خام تولید شده توسط شرکت‌های آنلاین و به خصوص، کاربران آن‌ها وجود داشته است. با این حال، چنین داده‌هایی بدون وجود داشتن هیچ تحلیلی از آن‌ها، ارزشی برای صاحبان کسب‌وکارها ندارند. این اطلاعات ساختاریافته نیستند و چون برای مخاطب انسانی و نه ماشینی طراحی شده‌اند، توسط ماشین‌ها قابل پردازش نیستند اما با تجربه و تحلیل این اطلاعات، شرکت‌ها می‌توانند اطلاعات ارزشمندی در مورد نیازها، افکار، واکنش‌ها و ایده‌های کاربران و مشتریان خود نسبت به محصولات و خدماتی که ارائه می‌دهند به دست آورند. این موضوع منجر به پیدایش علوم عقیده‌کاوی^۱ و تحلیل احساسات^۲ شده است. در این علوم، تحقیقات بر تولید ابزارهای اتوماتیک برای جمع‌آوری نظرات و نقدهای فضای مجازی و تحلیل آن‌ها تمرکز دارد.

واژه‌های کلیدی:

عقیده‌کاوی، تحلیل احساسات، پردازش زبان طبیعی، قطبی

^۱ Opinion Mining

^۲ Sentimental Analysis

فهرست مطالب

فصل اول مقدمه.....	۴
فصل دوم پیش زمینه.....	۶
۱-۲- پردازش زبان طبیعی.....	۷
۱-۲-۱ تعبیه کلمه.....	۷
فصل سوم عقیده کاوی، تحلیل احساسات و نیازهایی که پاسخگوی آن هستند.....	۹
۱-۳- عقیده کاوی و تحلیل احساسات.....	۱۰
۲-۳- تحقیقات در زمینه ی عقیده کاوی.....	۱۱
فصل چهارم روش ها و چالش ها.....	۱۳
۱-۴- روش های کلی.....	۱۴
۲-۴- از کلیدواژه به مفاهیم و معانی.....	۱۶
۱-۲-۴ روش های مبتنی بر یافتن کلیدواژه ها.....	۱۷
۲-۲-۴ روش های مبتنی بر وابستگی کلمات.....	۱۸
۳-۲-۴ روش های آماری.....	۱۸
۴-۲-۴ روش های مبتنی بر مفهوم.....	۱۸
فصل پنجم جمع بندی و نتیجه گیری.....	۱۹
منابع و مراجع.....	۲۱

فهرست اشکال

۱۲ شکل ۱-۳ مدل Bag-of-Concepts

فصل اول

مقدمه

مقدمه

هنگامی که می‌خواهیم از میان گزینه‌های متعدد یکی را انتخاب کنیم، نظرات سایر افراد می‌توانند تاثیر به سزایی در این انتخاب داشته باشند. به ویژه اگر هزینه‌های زمانی، مالی و ... در این انتخاب دخیل باشند. تا چندی پیش، منبع اصلی اطلاعات برای اکثر تصمیم‌گیری‌های ما، دوستان و اطرافیان، مجله‌های اختصاصی هر موضوع و وبسایت‌های مختلف بودند. اما امروزه، اینترنت به ما امکان استفاده از ابزارهای جدید برای ایجاد و به اشتراک‌گذاری نظرات و عقاید خود با سایر افراد متصل به شبکه‌ی جهانی وب را می‌دهد. بلاگ‌ها، شبکه‌های اجتماعی و سرویس‌های به اشتراک‌گذاری محتوا به افراد مختلف امکان به اشتراک‌گذاری اطلاعات متفاوتی را می‌دهد. این اطلاعات ساختاریافته نیستند و چون برای مخاطب انسانی و نه ماشینی طراحی شده‌اند، توسط ماشین‌ها قابل پردازش نیستند.

در حال حاضر، توجه محققین و صنعت به جمع‌آوری نظر عموم مردم درباره‌ی رویدادهای اجتماعی، جنبش‌های سیاسی، تصمیمات شرکت‌های بزرگ صنعتی، کمپین‌های بازاریابی و میزان محبوبیت یک محصول جدید روز به روز در حال افزایش است. این توجه منجر به به وجود آمدن علوم **عقیده‌کاوی** و **تحلیل احساسات** شده است.

فصل دوم

پیش زمینه

پیش‌زمینه

خوب است ابتدا پیش از ورود به بحث اصلی، با علم پردازش زبان طبیعی بیشتر آشنا شویم. این شناخت به ما کمک می‌کند تا دلیل نیاز به این علم در عقیده‌کاوی و تحلیل احساسات و مفاهیم آن را بهتر درک کنیم.

در ادامه بیشتر با این علوم آشنا خواهید شد.

۲-۱- پردازش زبان طبیعی

پردازش زبان طبیعی یکی از زیرشاخه‌های مهم علوم کامپیوتر و هوش مصنوعی است که بر تعامل بین انسان‌ها و ماشین‌ها تمرکز دارد. چالش اصلی در این علم، برنامه‌ریزی ماشین به نحوی است که بتواند حجم بزرگی از اطلاعات که به زبان طبیعی انسان نوشته شده‌اند را پردازش و درک کند. از دیگر چالش‌های این علم، تولید زبان قابل فهم برای انسان (زبان طبیعی) توسط ماشین است.

کاربردهای پردازش زبان طبیعی به دو دسته‌ی کلی تقسیم می‌شوند: کاربردهای نوشتاری و کاربردهای گفتاری. از کاربردهای نوشتاری این حوزه می‌توان به استخراج اطلاعات خاص از یک متن و نیز ترجمه متن از یک زبان طبیعی به زبانی دیگر، اشاره کرد. کاربردهای گفتاری این حوزه نیز شامل سیستم‌های پرسش و پاسخ انسان با رایانه است. در سال‌های اخیر تحقیقات قابل ملاحظه‌ای در این زمینه صورت گرفته است.

۲-۱-۱ تعبیه کلمه

تعبیه کلمه^۳ نامی تجمعی است که به مجموعه‌ای از تکنیک‌های یادگیری ویژگی و مدل‌سازی زبان در پردازش زبان طبیعی اطلاق می‌شود. در این تکنیک‌ها، کلمات و عبارات از یک سند به بردارهای عددی نگاشت می‌شوند.

^۳ Word Embedding

به طور مفهومی، این تکنیک مستلزم تعبیه سازی ریاضی^۴ از فضایی با ابعاد زیاد به ازای هر کلمه به فضای برداری پیوسته با ابعاد بسیار کمتر است. روش‌هایی که جهت تولید این نگاشت‌ها مورد استفاده قرار می‌گیرند شامل شبکه‌های عصبی و مدل‌های احتمالاتی می‌شود. استفاده از روش تعبیه کلمه و عبارت به عنوان بازنمایی ورودی سبب افزایش کارایی در کاربردهای مبتنی بر پردازش زبان طبیعی همانند عقیده‌کاوی و تحلیل احساسات می‌گردد.

⁴ Mathematical Embedding

فصل سوم

عقیده کاوی، تحلیل احساسات و نیازهایی که پاسخگوی آن هستند

درباره‌ی عقیده‌کاوی و تحلیل احساسات

خوب است ابتدا کمی بیشتر با عقیده‌کاوی و تحلیل احساسات آشنا شویم.

۳-۱- عقیده‌کاوی و تحلیل احساسات

اگرچه عموماً از هر دو اصطلاح برای اشاره به یک مفهوم استفاده می‌شود، عقیده‌کاوی بر تشخیص قطبیت^۵ و تحلیل احساسات بر تشخیص احساساتی که گوینده در بیان یک متن دارد تمرکز دارد. در هر دو زمینه از علوم داده‌کاوی^۶ و پردازش زبان طبیعی^۷ برای تشخیص، بازیابی و استخراج اطلاعات گسترده‌ی شبکه‌ی جهانی وب استفاده می‌شود.

تحلیل عقاید و احساسات یک سند چالش بزرگی است. این کار نیاز به درک عمیق زبان آن سند دارد. این موضوع عقیده‌کاوی را به مسئله‌ای مهم در پردازش زبان طبیعی تبدیل می‌کند زیرا این مسئله به محققین فرصت پیشرفت در تمامی قسمت‌های پردازش زبان طبیعی را می‌دهد. اما متأسفانه در حال حاضر اکثر تحقیقات در زمینه‌ی تحلیل احساسات بر متون انگلیسی تمرکز دارند و به همین دلیل، اکثر منابع آموزشی و کاربردی این زمینه نیز به زبان انگلیسی هستند.

شرکت‌های زیادی از عقیده‌کاوی و تحلیل احساسات در تحقیقات خود استفاده می‌کنند. به عنوان مثال، شرکت‌ها از طریق سکوهایی مانند وبسایت تویتر^۸ به طور مداوم حجم بزرگی از اطلاعات شامل نقد محصولات و سرویس‌های ارائه شده توسط آن شرکت را جمع‌آوری کرده و با استفاده از عقیده‌کاوی و تحلیل احساسات، برای بهبود محصولات و خدمات خود، رابطه‌ی شرکت و مشتری و سیستم‌های توصیه‌گر

^۵ Polarity

^۶ Data Mining

^۷ Natural Language Processing

^۸ Twitter

^۹ تلاش می کنند. همچنین می توان از این روش ها برای تشخیص و حذف نظرات پرتنش با محتوای نامناسب استفاده کرد. [۱]

۳-۲- تحقیقات در زمینه ی عقیده کاوی

تحقیقات در این زمینه بر تولید ابزارهای اتوماتیک برای جمع آوری نظرات و نقدهای فضای مجازی و تحلیل آن ها تمرکز دارد. گروه های مختلف، ابزارهایی برای جمع آوری و تحلیل نظرات کاربران و خلاصه سازی و نمایش گرافیکی آن ها ارائه می دهند که از میان آن ها می توان به گروه سنتیکنت (SenticNet) اشاره کرد.

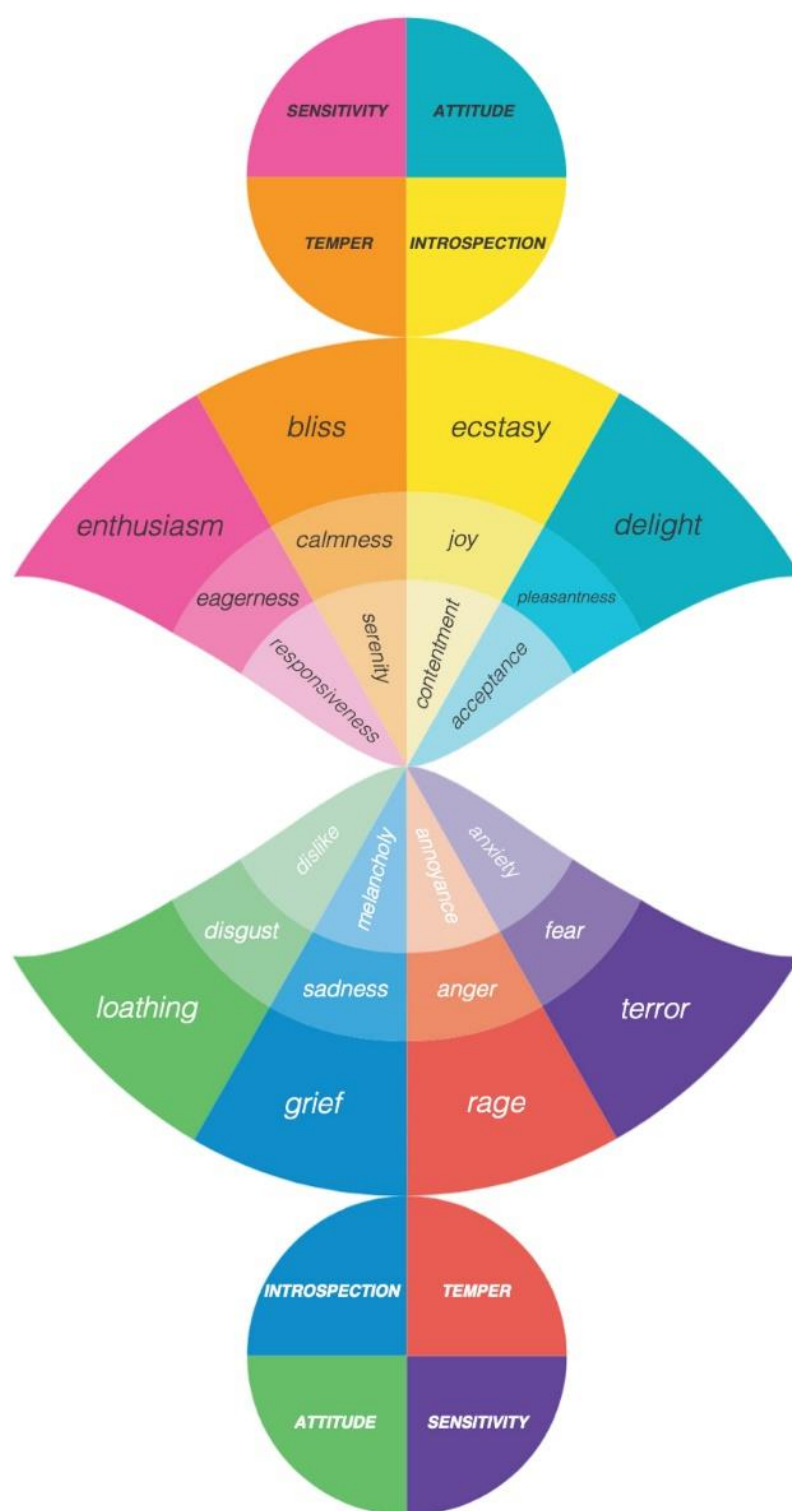
سنتیکنت در سال ۲۰۰۹ در آزمایشگاه رسانه ی دانشگاه ام. آی. تی.^{۱۰} با همکاری دانشگاه استرلینگ^{۱۱} به وجود آمد. تمرکز این گروه از ابتدای کار بر طراحی و توسعه ی برنامه های هوشمند آگاه بر احساسات در زمینه های متعدد بوده است. هدف اصلی سینتکنت مشابه پردازش زبان طبیعی، تبدیل زبان انسان به شکلی است که برای ماشین ها قابل درک باشد.

سنتیکنت برای این کار از مدل Bag-of-Concepts استفاده می کند. در این مدل، به جای شمارش ساده ی کلمات و بررسی فرکانس تکرار آن ها که از روش های رایج بازیابی اطلاعات یک سند هستند، با تکیه بر الگوهای زبان شناسی، جریان احساسات در یک سند بررسی می شود. [۲] در صفحه ی بعد شکل کلی این مدل را مشاهده می کنید:

^۹ Recommender Systems

^{۱۰} Massachusetts Institute of Technology (MIT)

^{۱۱} Stirling University



شکل ۳-۱: مدل Bag-of-Concepts

فصل چهارم

روش‌ها و چالش‌ها

در این قسمت به معرفی و بررسی روش‌های رایج عقیده‌کاوی و تحلیل احساسات و چالش‌های آن می‌پردازیم.

۴-۱- روش‌های کلی

وظیفه‌ی اصلی عقیده‌کاوی، دسته‌بندی با استفاده از قطبیت‌هاست. این اتفاق وقتی رخ می‌دهد که یک متن که بیان‌کننده‌ی نظری درباره‌ی یک موضوع خاص است، در یکی از دو دسته‌ی احساسی که برای طبقه‌بندی آن متن استفاده می‌شوند، قرار گیرد. نظردهی به اشکال thumbs up و thumbs down و یا like و dislike نمونه‌ای از دسته‌بندی با استفاده از قطبیت‌هاست. همچنین سیستم می‌تواند پس از تشخیص قطبیت متن به قطبیت آن امتیاز دهد. به عنوان مثال، اگر دو قطب دسته‌بندی ما قطب‌های "مثبت" و "منفی" باشند، سیستم درجه‌ی مثبت بودن یک متن را مشخص کند و به آن امتیاز دهد.

تشخیص توافق^{۱۲} نیز شکل دیگری از دسته‌بندی باینری عقاید و احساسات است. تشخیص توافق مشخص می‌کند که آیا دو متن متفاوت که در مورد یک موضوع خاص صحبت می‌کنند برچسب احساس مشابهی دارند یا خیر.

اگر متنی حاوی عقاید قوی و محکم درباره‌ی موضوع خود نباشد یا بیش از یک موضوع را پوشش دهد، با چالش‌های جدیدی روبرو خواهیم بود. در این حالت برای تشخیص قطبیت صحیح متن باید این موضوعات را تشخیص داده، نظرات مربوط به هر کدام را استخراج کرده و به صورت جداگانه بررسی کنیم. علاوه بر این، اگر متن حاوی نظرات مثبت و منفی نسبت به یک موضوع باشد، باید به طریق مشابه این نظرات را استخراج کرده و به صورت جداگانه بررسی کرد. برای این کار از روش‌های در سطح بخش^{۱۳} استفاده می‌شود. در یکی از تحقیقاتی که در این زمینه انجام شد، محتوای متن به جای سطح پاراگراف در سطح پایین‌تر جمله بررسی شد. در این روش با استفاده از بررسی حضور کلماتی که نشان‌دهنده‌ی افکار و نظرات خاصی هستند و ارتباط آن‌ها با یکدیگر (n-گرم^{۱۴}) تحلیل احساسات متن انجام شد. [۳]

¹² Agreement Detection

¹³ Segment-Level Opinion Analysis

¹⁴ N-gram

اما حتی گاهی این روش‌ها هم از تشخیص احساس مرتبط با یک شی^{۱۵} و یا صورت و جنبه‌ی آن ناتوانند. برای حل این مشکل، محققان از رویکردهای جنبه‌نگرانه^{۱۶} استفاده کردند. در این روش، هر نظر و ایده و مخاطب یا هدف آن مشخص شده و احساس آن نیز به دست می‌آید. برای مثال جمله‌ی "صفحه‌ی نمایش تلفن همراه آیفون عالی است اما قابلیت نگهداری شارژ باتری آن ضعیف است." دو جنبه را بررسی می‌کند: صفحه‌ی نمایش تلفن همراه و قابلیت نگهداری شارژ باتری. احساس گوینده‌ی جمله نسبت به جنبه‌ی اول مثبت و نسبت به دومی منفی است. حال، براساس این سطح از بررسی و کاوش، می‌توانیم نتایج سازمان‌دهی شده‌تر و یا آماری دقیق‌تر از آن شی و ویژگی‌های مختلف آن را ارائه دهیم. [۴-۶]

علاوه بر این چالش، یک متن می‌تواند بدون اینکه درباره‌ی موضوعی نظر دهد، قطبیت داشته باشد. به عنوان مثال، یک خبر می‌تواند بدون اینکه حاوی نظر یا عقیده خاصی باشد در دسته‌ی اخبار خوب یا بد قرار گیرد.

در حال حاضر، عقیده‌کاوی و تحلیل احساسات بر استخراج برداری^{۱۷} محتوای متن برای نمایش ویژگی‌های مهم آن تکیه دارد. ما می‌توانیم با استفاده از این بردار، ویژگی‌هایی همچون **فرکانس عبارت**^{۱۸} و **حضور عبارت**^{۱۹} برای کلمات موجود در متن را مشخص کنیم.

حضور عبارت، یک متغیر باینری است که مشخص می‌کند عبارتی خاص در متن حضور دارد (مقدار ۱) یا ندارد (مقدار ۰). این ویژگی در تشخیص قطبیت متن نقش مهمی دارد. فرکانس عبارت اشاره به تعداد دفعات تکرار یک عبارت در متن دارد.

اگرچه کلمات پرتکرار در متن اهمیت بالایی دارند اما تکرار یک کلمه می‌تواند در قطبیت و تحلیل احساسات نهایی متن بی‌تاثیر باشد. به عنوان مثال، تکرار نام یک شرکت در خبری درباره آن شرکت تاثیری در قطبیت نهایی متن ندارد. در چنین حالاتی می‌توان از معیار دیگری به نام n -قلم استفاده کرد. به

¹⁵ Entity

¹⁶ Aspect Level

¹⁷ Vector Extraction

¹⁸ Term Frequency

¹⁹ Term Presence

دنباله‌ای پیوسته از n -قلم در یک دنباله معین از متن یا کلام n -قلم می‌گوییم. اقلام می‌توانند واج، هجا، حرف، واژه و ... باشند. جایگاه یک کلمه در ساختار جمله (فعل، فاعل، مفعول و ...) نیز می‌تواند معیار مناسبی برای تشخیص قطبیت متن باشد. از دیگر معیارها می‌توان به فاصله‌ی بین عبارات اشاره کرد. از این معیار برای بررسی چگالی گروه کلمات استفاده می‌شود.

۴-۲- از کلیدواژه به مفاهیم و معانی

ما می‌توانیم سیر تحول حوزه‌ی عقیده‌کاوی و تحلیل احساسات را به وسیله‌ی بلوک‌های ساختاری متن^{۲۰} و اطلاعات مربوط به هر کدام بررسی کنیم. بر این اساس، روش‌های استفاده شده در مسائل عقیده‌کاوی و تحلیل احساسات به چهار گروه تقسیم می‌شوند:

- روش‌های مبتنی بر یافتن کلیدواژه‌ها^{۲۱}
- روش‌های مبتنی بر وابستگی کلمات^{۲۲}
- روش‌های آماری^{۲۳}
- روش‌های مبتنی بر مفهوم^{۲۴}

²⁰ Token

²¹ Keyword Spotting

²² Lexical Affinity

²³ Statistical Methods

²⁴ Concept-Based Techniques

۴-۲-۱ روش‌های مبتنی بر یافتن کلیدواژه‌ها

این روش علی‌رغم اینکه ساده‌ترین روش استفاده شده در مسائل عقیده‌کاوی است، اما به دلیل سادگی‌اش بسیار محبوب است. در این روش، دسته‌بندی متون با استفاده از لیستی از کلمات موثر احساسی^{۲۵} در متنی که حاوی لغات واضح و غیرمبهم احساسی مانند "شاد"، "ناراحت"، "ترس" و "خسته" است انجام می‌شود. برای مثال، روش Affective Reasoner که توسط کلارک الیوت^{۲۶} معرفی شد، در متن به جست‌وجوی ۱۹۸ کلیدواژه‌ی احساسی (مانند "نگران" و "عصبانی")، قیده‌ای شدت (مانند "خیلی زیاد"، "تقریباً" و "خیلی کم") و تعدادی عبارت کمکی (مانند "انجام داد" و "خواست") می‌پردازد.

اما این روش دو ضعف دارد:

- نمی‌تواند به طور موثر و قابل اطمینان کلمات موثر را شناسایی کند.
- غالباً به ویژگی‌های سطحی^{۲۷} متن می‌پردازد.

با وجود اینکه این روش می‌تواند به درستی جمله‌ی "امروز روز خوبی بود." را به عنوان یک جمله مثبت تشخیص دهد، اما احتمالاً در دسته‌بندی جمله‌ی "امروز اصلاً روز خوبی نبود." موفق نخواهد بود. همچنین، اگر یک جمله به جای استفاده‌ی واضح از کلمات موثر، احساس گوینده را در قالب معنایی که در خود نهفته است منتقل کند، این روش در دسته‌بندی آن موفق نخواهد بود. برای مثال جمله‌ی "همسرم تقاضای طلاق کرده است و قصد دارد حق حضانت بچه‌هایمان را از من بگیرد." احساسات قوی گوینده را منتقل می‌کند اما در آن از کلمات موثر استفاده نشده است. در نتیجه، این روش نمی‌تواند به درستی این نوع جملات را دسته‌بندی کند.

²⁵ Affect Words

²⁶ Clarck Elliott

²⁷ Surface Features

۲-۲-۴ روش‌های مبتنی بر وابستگی کلمات

این روش کمی پیچیده‌تر از روش قبلی است زیرا نه تنها لغات موثر را به درستی شناسایی می‌کند بلکه به هر لغت یک احتمال ارتباط با یک احساس خاص را نسبت می‌دهد. برای مثال، این تکنیک ممکن است به کلمه‌ی "تصادف" ۷۵٪ احتمال بیان یک حس منفی را بدهد. مانند عبارات "تصادف ماشین" و "کشته‌شده در تصادف". [۷]

با وجود اینکه این روش نسبت به روش قبلی برتری دارد اما نقطه ضعف‌هایی نیز دارد:

- نمی‌تواند به درستی احساس موجود در جمله‌ی "از وقوع تصادف جلوگیری کردم." را از جملات مشابه اما با معنای متفاوت مانند "دوستم را تصادفی ملاقات کردم." تفکیک کند.
- احتمالات نسبت داده شده به یک کلمه معمولاً به موضوع متن بستگی دارد که این، ساخت مدلی با قابلیت استفاده‌ی مجدد برای متون مختلف و مستقل از موضوع متن را مشکل‌تر می‌سازد.

۳-۲-۴ روش‌های آماری

این روش ترکیبی از دو روش قبلی است که این کار با train کردن یک الگوریتم یادگیری ماشین با استفاده از مجموعه‌ای بزرگ از متون نشانه‌گذاری شده انجام می‌شود. اما به طور کلی این دسته از روش‌ها ضعیف هستند، زیرا برای موثر بودن نیاز به دسته بزرگی از متون و اسناد برای train کردن الگوریتم خود دارند و در غیر این صورت تنها برای دسته‌بندی متون کوتاه و در سطح جمله یا پاراگراف کاربرد دارند.

۴-۲-۴ روش‌های مبتنی بر مفهوم

در این دسته از روش‌ها با استفاده از شبکه‌ی جهانی وب و شبکه‌های معنایی^{۲۸} به بررسی نحوی و معنایی یک متن می‌پردازند. این کار به سیستم کمک می‌کند تا اطلاعات مفهومی و احساسی مربوط به یک متن را در زبان طبیعی به دست بیاورد (به جای استفاده از کلیدواژه‌ها و احتمال حضور کلمات خاص در متن).

²⁸ Semantic Networks

فصل پنجم

جمع‌بندی و نتیجه‌گیری

جمع‌بندی و نتیجه‌گیری

عقیده‌کاوی و تحلیل احساسات به تدریج در حال تبدیل شدن به یک حوزه‌ی تحقیقاتی جداگانه است که بین حوزه‌های پردازش زبان طبیعی و فهم زبان طبیعی^{۲۹} قرار می‌گیرد. برخلاف حوزه‌ی پردازش زبان طبیعی که بر روی موضوعات مفهومی مانند خلاصه‌سازی^{۳۰} و دسته‌بندی اتوماتیک^{۳۱} متمرکز هستند، تحقیقات در این دو حوزه بیشتر بر روی بار معنایی و اطلاعات موثر دریافتی از زبان طبیعی تمرکز دارند و نیازی به فهم عمیق متن نیست.

از طرفی، شبکه‌ی جهانی وب باعث شده است که کاربران علاقه‌مند به ارتباط و به اشتراک‌گذاری اطلاعات از طریق شبکه‌های اجتماعی، گروه‌های آنلاین، بلاگ‌ها، ویکی و دیگر شبکه‌های مشترک روی بیاورند. دانش جمعی در سراسر وب گسترده شده است، به خصوص در حوزه‌های مربوط به زندگی روزمره مانند رفت و آمد، توریسم، آموزش و بهداشت و سلامت. اما علی‌رغم پیشرفت‌های صورت گرفته، تحلیل احساسات و عقیده‌کاوی همچنان در حال یافتن جایگاه خود به عنوان یک حوزه‌ی جدید بین رشته‌ای است.

²⁹ Natural Language Understanding

³⁰ Summarization

³¹ Auto-Categorization

منابع و مراجع

- [١] Liu, B. and Zhang, L., 2012. A survey of opinion mining and sentiment analysis. In *Mining text data* (pp. 415-463). Springer, Boston, MA.
- [٢] <https://sentic.net/about/>
- [٣] Riloff, E. and Wiebe, J., 2003. Learning extraction patterns for subjective expressions. In *Proceedings of the 2003 conference on Empirical methods in natural language processing* (pp. 105-112).
- [٤] Hu, M. and Liu, B., 2004, August. Mining and summarizing customer reviews. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 168-177).
- [٥] Lu, B., Ott, M., Cardie, C. and Tsou, B.K., 2011, December. Multi-aspect sentiment analysis with topic models. In *2011 IEEE 11th international conference on data mining workshops* (pp. 81-88). IEEE.
- [٦] Di Fabbri, G., Aker, A. and Gaizauskas, R., 2011, December. Starlet: multi-document summarization of service and product reviews with balanced rating distributions. In *2011 IEEE 11th International Conference on Data Mining Workshops* (pp. 67-74). IEEE.
- [٧] Stevenson, R.A., Mikels, J.A. and James, T.W., 2007. Characterization of the affective norms for English words by discrete emotional categories. *Behavior research methods*, 39(4), pp.1020-1024.