بسم الله الرحمن الرحیم

دانشکده پردیس خوارزمی
رشته مهندسی کامپیوتر،گرایش هوش مصنوعی

رساله دکتری

طراحی یک واژه‌نامه‌ی احساسی انگلیسی با توانایی کاربست کلمات تقویت‌کننده و منفی‌کننده‌ی متوالی

نگارنده:

حامد زرگری

استاد راهنما:

دکتر مرتضی زاهدی

استاد مشاور:

دکتر مرضیه رحیمی

شهریور ۱۴۰۰

دانشکده: پردیس خوارزمی

گروه: هوش مصنوعی

رساله دکتری آقای حامد زرگری

تحت عنوان: طراحی یک واژه‌نامه‌ی احساسی انگلیسی با توانایی کاربست کلمات تقویت‌کننده و منفی‌کننده‌ی متوالی

شماره: ۳۱۰۰/۲۴۳۵۱/الف

تاریخ: ۱۴۰۰/۰۹/۱

ویرایش:

باسمه تعالی

مدیریت تحصیلات تکمیلی

پیوست شماره ۲

دانشکده :

گروه :

پایان نامه کارشناسی ارشد / رساله دکتری آقای حامد زرگری

تحت عنوان: طراحی یک واژه‌نامه‌ی احساسی انگلیسی با توانایی کاربست کلمات تقویت‌کننده و منفی‌کننده‌ی متوالی

در تاریخ ۱۴۰۰/۰۶/۱۵ توسط کمیته تخصصی زیر جهت اخذ مدرک کارشناسی ارشد/ رساله دکتری ارزیابی گردید و با درجه ..... مورد پذیرش قرار گرفت.

(هیجده و پنجاه و هشت صدم) ۱۸,۵۸

اساتید راهنما	امضاء	اساتید مشاور	امضاء
نام و نام خانوادگی: مرتضی زاهدی		نام و نام خانوادگی: مرضیه رحیمی	
نام و نام خانوادگی:		نام و نام خانوادگی:	

اساتید داور	امضاء	نماینده تحصیلات تکمیلی	امضاء
نام و نام خانوادگی: محسن رضوانی		نام و نام خانوادگی: دکتر سید ایمان آقایان	
نام و نام خانوادگی: هدی مشایخی			
نام و نام خانوادگی: محمد قاسم زاده			

تقدیم به آن غایب از نظری که همیشه و همه جا حاضر است

تقدیر و تشکر

سپاس و ستایش مر خدای را جل و جلاله که آثار قدرت او بر چهره روز روشن، تابان است و انوار حکمت او در دل شب تار، درفشان.
آفریدگاری که خویشتن را به ما شناساند و درهای علم را بر ما گشود و عمری و فرصتی عطا فرمود تا بدان، بنده ضعیف خویش را در طریق علم و معرفت بیازماید.

در آغاز از استاد بزرگ و دانشمند جناب آقای دکتر مرتضی زاهدی که راهنمایی این پایانامه را به عهده داشته اند کمال تشکر را دارم.

از سرکار خانم مرضیه رحیمی که استاد مشاور این پایانامه بوده اند نیز قدردانی می نمایم.

تقدیم با بوسه بر دستان پدرم:

به او که نمی دانم از بزرگی اش بگویم یا مردانگی سخاوت، سکوت، مهربانی و .....

تقدیم به مادر عزیزتر از جانم:

مادرم هستی من ز هستی توست تا هستم و هستی دارمت دوست.

و در نهایت تقدیم به همسرم

به پاس قدردانی از قلبی آکنده از عشق و معرفت که محیطی سرشار از سلامت و امنیت و آرامش و آسایش برای من فراهم آورده است

و فرزند دلبندم سجان که با دلی پاک و دستانی کوچک برای موفقیت من دعا می کرد

تعهد نامه

اینجانب **حامد زرگری** دانشجوی دوره دکترای رشته **کامپیوتر** گرایش **هوش مصنوعی** دانشگاه **صنعتی شاهرود** نویسنده پایان‌نامه" **طراحی یک واژه‌نامه‌ی احساسی انگلیسی با توانایی کاربست کلمات تقویت‌کننده و منفی‌کننده‌ی متوالی** " تحت راهنمایی **دکتر مرتضی زاهدی** متعهد می‌شوم.

- تحقیقات در این پایان‌نامه توسط اینجانب انجام شده است و صحت و اصالت برخوردار است.
- در استفاده از نتایج پژوهش‌های محققان دیگر مرجع مورداستفاده استناد شده است.
- مطالب مندرج در پایان‌نامه تاکنون توسط خود یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ جا ارائه نشده است.
- کلیه حقوق معنوی این اثر متعلق به دانشگاه صنعتی شاهرود می‌باشد و مقالات مستخرج با نام دانشگاه صنعتی شاهرود و یا «Shahrood University of Technology» به چاپ خواهد رسید.
- حقوق معنوی تمام افرادی که در به‌دست‌آمدن نتایج اصلی پایان‌نامه تاثیرگذار بوده‌اند در مقاله مستخرج از پایان‌نامه رعایت می‌گردد.
- در کلیه مراحل انجام این پایان‌نامه در مواردی که از موجود زنده یا بافت‌های آنها استفاده شده است و اصول اخلاقی رعایت شده است.
- در کلیه مراحل انجام این پایان‌نامه در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته یا استفاده شده است اصل رازداری ضوابط و اصول اخلاق انسانی رعایت شده است.

تاریخ:**۱۴۰۰/۰۶/۱۵**

امضای دانشجو

مالکیت نتایج و حق نشر

کلیه حقوق معنوی این اثر و محصولات آن (مقالات مستخرج از کتاب برنامه‌های رایانه‌ای نرم افزارها و تجهیزات ساخته شده است) متعلق به دانشگاه صنعتی‌شاهرود می‌باشد این مطلب باید به نحو مقتضی در تولیدات علمی مربوطه ذکر شود استفاده از اطلاعات و نتایج موجود در پایان نامه بدون ذکر مرجع مجاز نمی باشد

چکیده

در تمامی روش‌های تشخیص قطبیت متون، یکی از عوامل اصلی کلمات یا عبارات احساسی موجود در آن متن هستند. بیشتر تحقیقات موجود مبتنی بر دیکشنری، معطوف به استخراج کلمات احساسی، تعیین قطبیت و در نهایت تخصیص شدت احساسی به این کلمات بوده است. از سوی دیگر تقریباً تمامی محققان حوزه تجزیه‌وتحلیل احساسات بر این نکته تاکید دارند که یک متن می‌تواند شامل پدیده‌های زبانی باشد که باوجوداینکه حاوی هیچ‌گونه بار احساسی و قطبیتی نیستند اما قادرند بار احساسی متن را به‌شدت تحت تأثیر خود قرار دهند. منفی‌کننده‌ها، و تشدیدکننده‌ها دو نمونه مهم از پدیده‌های زبانی را تشکیل می‌دهند که توجه به آنها باعث افزایش میزان دقت و کارایی فرایند تجزیه‌وتحلیل احساسات می‌شود. غالب دیکشنری‌های احساسی موجود تنها حاوی یونیگرم‌ها به همراه نمرات احساسی تخصیص‌داده‌شده به آنها هستند و تنها تعداد بسیار اندکی از آنها به بررسی نقش تشدیدکننده‌ها در محاسبه قطبیت کلمات احساسی پرداخته‌اند و تقریباً در هیچ‌کدام از این دیکشنری‌ها تأثیر موقعیت مکانی تشدیدکننده بر شدت قطبیت کلمه احساسی و همچنین اثر متقابل قطبیت کلمه احساسی بر ضریب تأثیر تشدیدکننده موردتوجه واقع نشده است.

باتوجه‌به موارد فوق این پژوهش به لحاظ‌کردن موقعیت مکانی تشدیدکننده‌ها در یک عبارت احساسی بر شدت قطبیت آن عبارت و همچنین محاسبه اثر قطبیت کلمه احساسی بر این تشدیدکننده‌ها و ضرایب تأثیر آنها در یک عبارت احساسی پرداخته است و برای هر تشدیدکننده باتوجه‌به موقعیت قرارگیری و قطبیت کلمه احساسی چهار ضریب اثر مختلف را محاسبه نموده است. در روش پیشنهادی با استفاده از روش‌های نیمه نظارتی یک دیکشنری احساسی که توانایی پوشش غالب ترکیبات احساسی حاصل از کلمات احساسی و تشدیدکننده‌ها و منفی‌کننده‌ها را دارد ساخته شده است. در روش پیشنهادی چهار ضریب تأثیر به هر تشدیدکننده با استفاده از یک پیکره از نظرات برچسب خورده کاربران استخراج شده از سایت آمازون و روش‌های ترکیبی فازی و آماری اختصاص‌داده‌شده است. امتیازات کلمات احساسی از فرهنگ لغت دستی VADER استخراج شده است. نتایج شبیه‌سازی نشان می‌دهد فرهنگ لغت پیشنهادی حداقل ۹٪ در مقایسه با آخرین فرهنگ لغات احساسات در پایگاه‌داده‌های استاندارد بهبود داشته و در حوزه تشخیص اسناد منفی نیز به‌صورت نسبی ۴۹.۶٪ بهبود را از خود نشان داده است.

کلمات کلیدی:

آنالیز احساسات، دیکشنری احساسی، تشدیدکننده‌ها، موقعیت مکانی تشدیدکننده، تشدیدکننده‌های متوالی، اثر قطبیت کلمه احساسی و تشدیدکننده

لیست مقالات ژورنالی مستخرج از پایان نامه

Zargari, Hamed, Morteza Zahedi, and Marziea Rahimi. "GINS: A Global intensifier-based N-Gram sentiment dictionary", Journal of Intelligent & Fuzzy Systems, vol. 40, no. 6, pp. 11763-11776, 2021

لیست مقالات کنفرانسی مستخرج از پایان نامه

حامد زرگری، مرتضی زاهدی،۱۳۹۸، افزایش دایره واژگان و بهبود عملکرد واژه‌نامه احساسی Senti-N-Gram با استفاده از پسوندها و پیشوندهای منفی‌ساز. پذیرفته شده برای ارائه در پنجمین کنفرانس پردازش سیگنال و سیستم‌های هوشمند ایران (ICSPIS)، شاهرود، دانشگاه صنعتی شاهرود

فهرست مطالب

عنوان **صفحه**

فهرست جدول‌ها ..... د
فهرست شکل‌ها ..... و

فصل ۱- مقدمه و کلیات تحقیق ۱
۱-۱ مقدمه ۲
۱-۲ تعریف مسئله و چالش‌های موجود ۳
۱-۳ دلایل دشواری مسئله ۵
۱-۴ انگیزه‌ی پژوهش ۶
۱-۵ برخی اصلاحات و مفاهیم موجود ۸
۱-۶ ساختار پایان‌نامه ۱۲

فصل ۲- مفاهیم مقدماتی ۱۳
۲-۱ مقدمه ۱۴
۲-۲ اصطلاح‌شناسی و مفاهیم پس‌زمینه ۱۵
۲-۲-۱ هیجان و احساسات از دیدگاه روان‌شناسی ۱۵
۲-۲-۲ مؤلفه‌های کلیدی یک نظر ۱۶
۲-۲-۳ انواع مختلف نظرات ۱۷
۲-۳ نظرکاوی و تحلیل احساسات ۱۹
۲-۳-۱ آنالیز احساسات مبتنی بر عملکرد ۱۹
۲-۳-۲ سطوح مختلف آنالیز احساسات ۲۲
۲-۳-۳ تکنیک‌های آنالیز احساسات ۲۴
۲-۴ جمع‌بندی ۲۶

فصل ۳- پیشینه تحقیق ۲۷
۳-۱ مقدمه ۲۸
۳-۲ روش‌های طبقه‌بندی احساسات ۲۸
۳-۲-۱ روش‌های مبتنی بر یادگیری ماشین ۲۸
۳-۲-۲ روش‌های مبتنی بر واژه‌نامه ۳۰
۳-۲-۳ روش‌های ترکیبی ۳۱
۳-۳ پدیده‌های زبانی ۳۱
۳-۳-۱ منفی‌کننده‌ها ۳۲

۳-۳-۲ تشدیدکننده‌ها ۳۲
۳-۴ روش‌های تولید واژه‌نامه ۳۴
۳-۴-۱ روش‌های دستی ۳۵
۳-۴-۲ Bootstrapping یک واژه‌نامه‌ی احساسی با استفاده از یک مجموعه از کلمات بذر ۳۶
۳-۴-۳ تطبیق یک واژه‌نامه از حوزهی دیگر با استفاده از نوعی از انتقال یادگیری ۳۷
۳-۴-۴ روش‌های مبتنی بر یادگیری ماشین ۳۷
۳-۵ معرفی برخی از واژه‌نامه‌های موجود ۳۸
۳-۵-۱ واژه‌نامه‌هایی با جهتگیری معنایی «مبتنی بر قطبیت لغات» ۳۸
۳-۵-۲ واژه‌نامه‌های با شدت احساسات «valance based» ۴۰
۳-۵-۳ واژه‌نامه‌های آگاه از زمینه «context- awareness» ۵۳
۳-۶ روش‌های پایه آنالیز احساسات با استفاده از Unigram و Ngram ۶۵
۳-۷ جمع‌بندی ۶۷

فصل ۴- بررسی اثر موقعیت مکانی تشدیدکننده‌ها نسبت به کلمه احساسی ۶۹

۴-۱ مقدمه ۷۰
۴-۲ آماده‌سازی داده‌ها ۷۱
۴-۳ پایگاه‌داده ۷۴
۴-۴ معیارهای مقایسه ۷۷
۴-۵ روش پیشنهادی ۷۸
۴-۵-۱ استخراج الگوی فازی عبارات احساسی ۷۹
۴-۵-۲ محاسبه ضریب تأثیر تشدیدکننده‌ها ۸۶
۴-۵-۳ الگوریتم‌های مورداستفاده برای ارزیابی کارایی روش‌های پیشنهادی ۹۰
۴-۶ نتایج ۹۲
۴-۷ الگوی فازی ۹۲
۴-۸ ارزیابی دیکشنری پیشنهادی ۹۴
۴-۹ جمع‌بندی و خاتمه ۹۹

فصل ۵- بررسی تأثیر موقعیت مکانی تشدیدکننده و اثر متقابل قطبیت کلمه احساسی بر تشدیدکننده‌ها ۱۰۱

۵-۱ مقدمه ۱۰۲
۵-۲ روش پیشنهادی ۱۰۳
۵-۲-۱ نمونه‌برداری تصادفی طبقه‌ای انتساب متناسب ۱۰۳
۵-۲-۲ استخراج الگوی فازی عبارات احساسی ۱۰۶
۵-۳ محاسبه ضریب تأثیر تشدیدکننده‌ها ۱۰۷
۵-۴ شبیه‌سازی و نتایج ۱۱۱

۵-۴-۱ ارزیابی روش پیشنهادی ۱۱۱
۵-۵ جمع‌بندی ۱۱۷
۵-۶ نتیجه‌گیری و کارهای آینده ۱۱۸
۵-۷ جمع‌بندی روش پیشنهادی در رساله ۱۱۹
۵-۸ کارهای آینده ۱۲۱
فهرست مراجع ۱۲۲

فهرست جدول‌ها

عنوان **صفحه**

جدول ۱. اثر مساوی تشدیدکننده‌ها بر روی کلمات احساسی در دیکشنری VADER ۴۵
جدول ۲. تعداد یک گرمی‌ها، دو گرمی‌ها و سه گرمی‌ها در واژه‌نامه‌ی SCL-OPP ۴۹
جدول ۳. چند نمونه از کلمات موجود در دیکشنری DAL و امتیازات آنها ۵۱
جدول ۴. تعداد کلمات در واژه‌نامه‌ی WKWSCI با انواع برچسب‌گذاری واژگان کلامی و مقدار شدت احساسات ۵۶
جدول ۵. بخشی از واژه‌نامه‌ی ذهنی MPQA ۵۷
جدول ۶. تعداد عناصر هرکدام از دسته‌بندی‌های موجود در واژه‌نامه‌ی NRC ۵۸
جدول ۷ . حوزه‌های احساسی واژه‌نامه‌یWAL ۵۹
جدول ۸. کلمات و Synsetهای احساسی دسته‌بندی شده بر اساس برچسب‌گذاری نحوی ۶۱
جدول ۹. نمونه ای از کلمات بهمراه مقادیر آنها در واژه‌نامه Sentiful ۶۲
جدول ۱۰. دسته‌بندی انواع پیوست‌های استفاده شده برای تولید کلمات جدید در SentiFul ۶۳
جدول ۱۱. لیست کلمات موجود در واژه‌نامه‌ی SentiFul به تفکیک منبع ۶۴
جدول ۱۲. روش‌های نیمه‌خودکار مبتنی بر واژگان برای اصلاح نمره Ngramها با افزودن تقویت‌کننده‌ها و منفی‌کنندها ۶۶
جدول ۱۳. اطلاعات آماری برخی از واژه‌نامه‌های احساسی ۶۷
جدول ۱۴. مقایسه واژه‌نامه‌های موجود بر مبنای توسعه واژگان و روش امتیازدهی ۶۸
جدول ۱۵ . محتوا و نظر هر بخش موضوعی از پایگاه‌داده آمازون ۷۵
جدول ۱۶. نمادهای مربوط به واژه‌نامه احساسی پیشنهادی اول ۸۵
جدول ۱۷. جدول حالات مختلف بررسی و تأیید امتیازات استخراج شده برای اصطلاحات Bigram ۸۷
جدول ۱۸. ضرایب محاسبه شده برای برخی از تشدیدکننده‌ها باتوجه‌به موقعیت آنها در عبارت احساسی در روش پیشنهادی اول ۹۰
جدول ۱۹. نتایج آماری و فازی یک نمونه Unigram, Bigram, Trigram ۹۳
جدول ۲۰. مقایسه پارامتر دقت بر مبنای دوروش طبقه‌بندی اسناد در سطح جمله با استفاده از واژه‌نامه‌های VADER، Senti-N-gram و واژه‌نامه پیشنهادی اول ۹۵
جدول ۲۱. مقایسه پارامتر دقت بر مبنای روش طبقه‌بندی اسناد در سطح جمله[153] «الگوریتم اول» با استفاده از واژه‌نامه‌های VADER،Senti-N-gram و واژه‌نامه پیشنهادی نخست در حوزه‌های مختلف ۹۵
جدول ۲۲. مقایسه پارامترهای Precision، Recall و F-measure برمبنای روش طبقه‌بندی اسناد در سطح جمله[153] «الگوریتم اول» با استفاده از واژه‌نامه‌های VADER،Senti-N-gram و واژه‌نامه پیشنهادی نخست در حوزه‌های مختلف ۹۶

جدول ۲۳. مقایسه پارامتر دقت بر مبنای روش طبقه‌بندی اسناد در سطح جمله[2] «الگوریتم دوم» با استفاده از واژه‌نامه‌های Senti-N-gram،VADER و واژه‌نامه پیشنهادی نخست در حوزه‌های مختلف.....۹۷
جدول ۲۴. مقایسه پارامترهای Precision، Recall و F-measure برمبنای روش طبقه‌بندی اسناد در سطح جمله[2] «الگوریتم دوم» با استفاده از واژه‌نامه‌های Senti-N-gram،VADER و واژه‌نامه پیشنهادی نخست در حوزه‌های مختلف.....۹۸
جدول ۲۵. نمادهای مربوط به الگوریتم پیشنهادی دوم ۱۰۵
جدول ۲۶. جدول حالات مختلف بررسی و تأیید امتیازات استخراج شده برای اصطلاحات Trigram..۱۰۸
جدول ۲۷. ضرایب محاسبه شده برای چند تشدیدکننده نمونه باتوجه‌به موقعیت آنها در عبارت احساسی و قطبیت کلمه احساسی در روش پیشنهادی دوم ۱۱۰
جدول ۲۸. مقایسه پارامتر دقت بر مبنای دوروش طبقه‌بندی اسناد در سطح جمله با استفاده از واژه‌نامه‌های VADER، Senti-N-gram و واژه‌نامه پیشنهادی دوم ۱۱۲
جدول ۲۹. مقایسه پارامتر دقت بر مبنای روش طبقه‌بندی اسناد در سطح جمله [153] «الگوریتم اول» با استفاده از واژه‌نامه‌های Senti-N-gram،VADER و واژه‌نامه پیشنهادی دوم در حوزه‌های مختلف.....۱۱۳
جدول ۳۰. مقایسه پارامترهای Precision، Recall و F-measure بر مبنای روش طبقه‌بندی اسناد در سطح جمله [153] «الگوریتم اول» با استفاده از واژه‌نامه‌های Senti-N-gram،VADER و واژه‌نامه پیشنهادی دوم در حوزه‌های مختلف ۱۱۳
جدول ۳۱. مقایسه پارامتر دقت بر مبنای روش طبقه‌بندی اسناد در سطح جمله [2] «الگوریتم دوم» با استفاده از واژه‌نامه‌های Senti-N-gram،VADER و واژه‌نامه پیشنهادی دوم در حوزه‌های مختلف.....۱۱۴
جدول ۳۲. مقایسه پارامترهایPrecision ، Recallو F-measure بر مبنای روش طبقه‌بندی اسناد در سطح جمله [2] «الگوریتم دوم» با استفاده از واژه‌نامه‌هایVADER ، Senti-N-gramو واژه‌نامه پیشنهادی دوم در حوزه‌های مختلف ۱۱۵

فهرست شکل‌ها

عنوان **صفحه**

شکل ۱ . مدل ابعادی رابرت پلاچیک ۱۶
شکل ۲. تحقیقات آنالیز احساسات از جنبه‌های چندگانه ۲۰
شکل ۳. حذف جملات عینی برای بهبود کلاس‌بندی قطبی ۲۱
شکل ۴. نمایی از بازی Tower of Babel ۳۵
شکل ۵. ساختار دوبعدی صفات (مترادف متضاد) در WordNet ۴۰
شکل ۶. اشکال گرافیکی آزمون SAM برای سه بعد خوشایندی (اشکال بالا) برانگیختگی (اشکال وسط) و تسلط (اشکال پایین) ۴۱
شکل ۷. نمونه ای از امتیازدهی کلمات در دیکشنری ANEW ۴۲
شکل ۸. مدل گرافیکی sentiWordNet برای یک Synset ۴۴
شکل ۹ .فلوچارت مراحل ایجاد واژه‌نامه TF_IDF ۴۶
شکل ۱۰. گراف جزئی از حوزه‌ها و سلسله‌مراتب WAL ۶۰
شکل ۱۱. نمونه ای از نظرات ثبت شده در پایگاه داده آمازون ۷۶
شکل ۱۲ . فلوچارت روش پیشنهادی اول ۸۴
شکل ۱۳. شبه کد مربوط به فرآیند بررسی اعتبار امتیازات Bigram ۸۸
شکل ۱۴. فرآیند تعیین قطبیت اسناد ۹۱
شکل ۱۵ . نمودار فازی برای سه عبارت good، Very good و Very very good ۹۲
شکل ۱۶.مقایسه پارامتر دقت بر مبنای روش طبقه‌بندی اسناد در سطح جمله **[153]** «الگوریتم اول» با استفاده از واژه‌نامه‌های VADER،Senti-N-gram و واژه‌نامه پیشنهادی نخست در حوزه‌های مختلف ۹۸
شکل ۱۷.مقایسه پارامتر دقت بر مبنای روش طبقه‌بندی اسناد در سطح جمله[2] «الگوریتم دوم» با استفاده از واژه‌نامه‌های VADER،Senti-N-gram و واژه‌نامه پیشنهادی نخست در حوزه‌های مختلف ۹۹
شکل ۱۸. فلوچارت پیاده سازی دیکشنری پیشنهادی دوم ۱۰۴
شکل ۱۹. مقایسه پارامتر دقت بر مبنای روش طبقه‌بندی اسناد در سطح جمله [153] «الگوریتم اول» با استفاده از واژه‌نامه‌های VADER،Senti-N-gram و واژه‌نامه پیشنهادی دوم در حوزه‌های مختلف ۱۱۶
شکل ۲۰. مقایسه پارامتر دقت بر مبنای روش طبقه‌بندی اسناد در سطح جمله[2] «الگوریتم دوم» با استفاده از واژه‌نامه‌های VADER،Senti-N-gram و واژه‌نامه پیشنهادی دوم در حوزه‌های مختلف ۱۱۷

فصل اول

مقدمه و کلیات تحقیق

فصل ۱– مقدمه و کلیات تحقیق

۱-۱ مقدمه

هدف اصلی از تجزیه‌وتحلیل احساسات یافتن روشی خودکار برای شناسایی احساسات مثبت، منفی یا خنثی در اسناد است[1]. باتوجه‌به نقش پایه‌ای واژگان احساسی در روش‌های مبتنی بر تجزیه‌وتحلیل احساسات، واژه‌نامه‌های احساسی را می‌توان به‌عنوان یکی از مولفه‌های کلیدی اغلب روش‌های تجزیه‌وتحلیل احساسات درنظر گرفت[2,3]. این واژه‌نامه‌ها همچون پلی مابین محتوای متنی و کلمات عمل می‌نمایند. در نتیجه تولید واژه‌نامه‌های احساسی که از روش‌های مناسب‌تری برای امتیازدهی به واژه‌های احساسی استفاده نموده باشند توجه محققان زیادی را به خود جلب نموده است. علی‌رغم هزینه بالای زمانی و مالی در ساخت واژه‌نامه‌های احساسی، مزایای فوق‌العاده این قبیل از واژه‌نامه‌ها همچون عدم نیاز به منابع آموزشی و پیکره‌های حاشیه‌نویسی شده عظیم[3,4]، قابلیت این واژه‌نامه‌ها در استخراج ویژگی‌های مرتبط با احساس در روش‌های مبتنی بر یادگیری ماشین[5]، تولید خودکار واژه‌نامه‌های احساسی بزرگتر با استفاده از تکنیکهاینیمه نظارتیی[6] را می‌توان از جمله دلایل اهمیت این واژه‌نامه‌ها برشمرد. قابلیت تعمیم‌پذیری و قابلیت تبدیل آنها به زبانهای مختلف و حوزه‌های متفاوت[7]، توانایی تعریف و بهره‌گیری از پدیده‌های زبانی همچون تشدیدکننده‌ها[8]، منفی‌کننده‌ها و افعال کمکی، امکان افزودن شکلک‌ها[1]، اختصارات و عبارات محاوره‌ای [2]، قابلیت افزودن ویژگی‌های وابسته به زبان[9] و همچنین قابلیت درک و اصلاح بهتر برای انسان[7] از جمله مزایایی است که محققان فراوانی را برای تولید این قبیل از واژه‌نامه‌ها تحریک می‌نماید.

از سوی دیگر، پدیده‌های زبانی یکی از مؤثرترین عوامل تغییر قطبیت اولیه کلمات و عبارات احساسی در حوزه تجزیه‌وتحلیل احساسات محسوب می‌شوند. از جمله این عوامل می‌توان به نقش منفی‌کننده‌ها و تشدیدکننده‌ها در یک عبارت احساسی اشاره نمود. عدم توجه دقیق به تأثیر منفی‌کننده‌ها و یا تشدیدکننده‌ها

[1] Emoticons
[2] Colloquial Expressions

و همچنین موقعیت این پدیده‌ها در یک عبارت احساسی از یک سو و عدم توجه به تأثیر کلمات احساسی بر این پدیده‌های زبانی باعث کاهش کارایی سیستم‌های تجزیه‌وتحلیل احساسات می‌شود. برطبق نتایج و مشاهدات موجود واژه‌نامه‌های احساسی کنونی غالباً تنها حاوی کلمات احساسی هستند که قطبیت اولیه آنها به‌صورت یک امتیاز به آنها نسبت‌داده‌شده است. هرچند تأثیر تشدیدکننده‌ها و منفی‌کننده‌های متوالی در بهبود تجزیه‌وتحلیل احساسات بدیهی است اما تعداد بسیار اندکی[2,10–13] از این واژه‌نامه‌ها به نقش این قبیل از پدیده‌های زبانی توجه نموده‌اند و واژه‌نامه‌های موردنظر نیز تمامی حالات مختلف را شامل نمی‌شوند. برطبق تحقیقات ما تقریباً هیچ‌کدام از آنها به موقعیت قرارگیری این پدیده‌های زبانی پیش از کلمه احساسی و همچنین اثر متقابل قطبیت کلمه احساسی بر روی این پدیده‌های زبانی توجهی نداشته‌اند.

در این رساله روشی برای تولید یک واژه‌نامه‌ی احساسی با قابلیت امتیازدهی به عبارات احساسی Ngram ارائه نموده‌ایم. روش پیشنهادی ما یک روش مبتنی بر قانون[1] بوده که برای تولید این واژه‌نامه به‌صورت خودکار از آن استفاده می‌نماییم. برای محاسبه‌ی امتیاز موردنیاز از تعدادی از نظرات کاربران که دارای رتبه‌بندی در مقیاس ۵ می‌باشند و به‌صورت تصادفی از پیکره نظرات کاربران سایت آمازون انتخاب شده‌اند استفاده خواهیم نمود.

۲-۱ تعریف مسئله و چالش‌های موجود

ازآنجایی‌که معمولاً فرایند تصمیم‌گیری افراد عموماً تحت تأثیر تفکرات رهبران فکری و یا نظرات عمومی دیگران است، مسئله‌ی نظرکاوی یکی از موضوعات بسیار مهم در طول تاریخ زندگی بشر محسوب می‌شود. از سوی دیگر افراد مختلف به دلایل اجتماعی همچون علاقه به دیده‌شدن، کسب احساس صاحب‌نظر بودن، بیان تجربیات شخصی و... علاقه‌مند به بیان نظراتشان در رسانه‌های اجتماعی هستند. یکی از مهم‌ترین مسائل در حوزه‌ی نظرکاوی، استخراج کلمات احساسی موجود در نظر و سپس تعیین قطبیت متن، جمله و یا دیگر سطوح موردنظر است. به‌طورکلی سه روش برای استخراج اتوماتیک احساسات عبارتند از روش‌های مبتنی بر

[1] Rule-Based Approach

واژه‌نامه [12,14]، روش‌های مبتنی بر یادگیری ماشین[15,16] و روش‌های ترکیبی[17,18] که از ترکیب دو روش فوق با یکدیگر صورت پذیرفته است.

روش‌های مبتنی بر یادگیری ماشین با استفاده از ابزارهای آماری و هوش مصنوعی و همچنین پردازش پایگاه‌داده‌های برچسب خورده موجود در حوزه زبان طبیعی اقدام به استخراج ویژگی‌های هدفمند نموده و سپس با طی یک مرحله آموزش بر مبنای نظرات برچسب‌گذاری شده مربوط به یک دامنه خاص، اقدام به طبقه‌بندی جملات یا متون وارد شده جدید می‌نمایند. از روش‌های مرسوم در یادگیری ماشین می‌توان به روش‌هایی مانند نایوبیز[19,20]، ماکزیمم آنتروپی [21,22]، ماشین بردار پشتیبان [23,24] اشاره نمود. در تمامی روشهای تجزیه‌وتحلیل احساسات واژگان احساسی نقشی اساسی دارند، بنابراین ساخت یک مجموعه واژگان یا واژه‌نامه احساسی با کیفیت بالا و امتیازدهی مناسب توجه محققان زیادی را به خود جلب کرده است. حتی در روش‌های مبتنی بر یادگیری ماشین نیز یک واژه‌نامه احساسی می‌تواند برای استخراج ویژگی‌های مرتبط با احساس که هدفمندی و کارایی مناسب‌تری در تشخیص احساسات متن دارد استفاده شود [5].

در روش‌های مبتنی بر واژه‌نامه معمولاً لغات احساسی به‌صورت دستی و با استفاده از افراد بومی آن زبان استخراج و قطبیت و یا امتیازی به این کلمات اختصاص می‌یابد. مسلماً تولید این نوع از واژه‌نامه‌ها که غالباً عمومی بوده و به دامنه‌ی خاصی نیز وابسته نیستند به‌صورت دستی امری بسیار هزینه‌بر و وقت‌گیر خواهد بود. اما دراین‌بین آنچه غالباً مورد غفلت قرار گرفته است تأثیری است که کلمات تشدیدکننده و یا منفی‌کننده در جمله از خود باقی می‌گذارند. واضح و بدیهی است که در روش‌هایی که تنها بر مبنای کلمات احساسی عمل می‌نماید جمله‌ی "*شما انسان مهربانی هستید.*" در مقایسه با جمله‌ی "*شما انسان بسیار مهربانی هستید.*" دارای امتیاز احساسی یکسانی است حال اینکه، کلمه‌ی "بسیار" در اینجا نقش تشدیدکننده را داشته و جمله‌ی دوم از شدت احساسی بیشتری برخوردار است. این امر در مورد کلمات منفی‌کننده نیز صدق می‌نماید. مسئله زمانی پیچیده‌تر خواهد شد که ما به‌جای استفاده از یک تشدیدکننده به طور هم‌زمان از دو یا سه تشدیدکننده و منفی‌کننده‌ی متوالی، در یک جمله استفاده نماییم. به‌عنوان‌مثال "من معلمم را بسیار زیاد دوست دارم"

و یا "حال من خیلی بد نیست". بدیهی است که محاسبه امتیازات ترکیبات مختلف حاصل از عبارات احساسی و تشدیدکننده‌ها و منفی‌کننده‌ها امری غیرممکن و ناشدنی خواهد بود.

برطبق تحقیقات ما در واژه‌نامه‌های توسعه داده شده گذشته، توجه به پدیده‌های زبانی یا اصلاً صورت نگرفته یا بسیار اندک بوده است. برای نمونه در غالب واژه‌نامه‌های موجود، تأثیر تشدیدکننده‌ها مانند "very" در ترکیب با کلمات احساسی موجود یا اصلاً مدنظر قرار نگرفته و یا درصورتی‌که مدنظر قرار گرفته تأثیر آنها به‌صورت اضافه‌شدن یا کم‌شدن یک عدد ثابت به امتیاز کلمه احساسی و یا ضرب امتیاز کلمه احساسی در یک وزن محاسبه شده بوده است. ایراد عمده چنین رویکردهایی عدم توجه به تکرار، موقعیت مکانی پدیده زبانی و تفاوت‌های کلمات احساسی با یکدیگر هنگام ترکیب با پدیده‌های زبانی است که می‌تواند تأثیر متفاوتی بر شدت قطبیت و اهمیت ترکیب بگذارد و لذا در محاسبات مربوط به قطبیت متن توجه به پدیده‌های زبانی و موقعیت مکانی آنها نیز می‌تواند کارایی روش را تا حد زیادی بهبود دهد. از سوی دیگر برطبق نتایج آماری به‌دست‌آمده اثر یک تشدیدکننده و یا حتی یک منفی‌کننده هنگامی‌که پیش از کلمه‌ی احساسی مثبت قرار می‌گیرد با اثر همان تشدیدکننده و یا منفی‌کننده در هنگامی‌که پیش از کلمه‌ی منفی قرار می‌گیرد متفاوت است[6].

لذا باتوجه‌به دستاوردهای این پژوهش در راستای ارتقاء واژه‌نامه‌های احساسی عمومی موجود، تحقیق حاضر تلاشی برای ارائه‌ی یک واژه‌نامه‌ی احساسی عمومی Ngram است که امکان ارزیابی تأثیر تشدیدکننده‌ها و منفی‌کننده‌های متوالی را باتوجه‌به اثر قطبیت کلمه احساسی پایه و موقعیت مکانی تشدیدکننده نسبت به کلمه احساسی در یک جمله فراهم می‌نماید.

۳-۱ دلایل دشواری مسئله

غالباً تمامی روش‌های که تأثیر یک تشدیدکننده در جملات احساسی را مدنظر قرار می‌دهند از وزن‌دهی یکسان به تشدیدکننده‌ها استفاده نموده‌اند. معمولاً این وزن‌دهی به‌صورت دستی و با صرف وقت و هزینه‌ی بسیار بالایی صورت پذیرفته است. باتوجه‌به تحقیقات صورت پذیرفته بر روی منفی‌کننده‌ها - که بسیار عمومی

و کم تعداد نیز می‌باشند - تأثیر یک منفی‌کننده در هنگام قرارگرفتن پیش از یک کلمه‌ی احساسی مثبت با اثر همان منفی‌کننده هنگام قرارگیری پیش از یک کلمه‌ی منفی متفاوت است. لذا محاسبه‌ی دستی اثر تشدیدکننده‌های موجود، در هنگامی‌که قبل از هر کلمه‌ی احساسی خاص قرار می‌گیرند امری امکان‌ناپذیر خواهد بود. انجام این عمل به‌صورت خودکار نیز بنحویکه واژه‌نامه‌ی تولیدی قادر به محاسبه‌ی تأثیر دو یا سه تشدیدکننده و منفی‌کننده‌ی متوالی بوده و تمامی حالات مختلف موجود را تحت پوشش خود قرارداده باشد نیازمند حجم بسیار عظیمی از اطلاعات و محاسبات پیچیده است. تجزیه‌وتحلیل آماری این مسئله به‌صورت کامپیوتری نیز نیازمند حجم بسیار زیاد حافظه و پردازش فوق‌العاده بالایی خواهد بود. الگوریتم‌های طراحی شده برای پردازش نیز باید به‌گونه‌ای طراحی شده باشند که قادر به پالایش و نگهداری این اطلاعات در حافظه‌ی اصلی بوده و نتایج تولید شده توسط این واژه‌نامه را نسبت به واژه‌نامه‌های دیگر به طور محسوسی بهبود دهند.

۴-۱ انگیزه‌ی پژوهش

کاربردهای مختلف سیستم‌های هوشمند تشخیص و تحلیل احساسات که قادر به ارائه برخط نتایج هستند را می‌توان در طیف وسیعی از حوزه‌های متنوع مانند حوزه سینمایی و نقد فیلم، آموزشی همچون بررسی کیفیت تدریس، تجاری و بازخورد از محصولات، یادگیری الکترونیکی، هتلداری، توریسم و ... مشاهده نمود. به‌عنوان‌مثال در حوزه تجاری تشخیص قطبیت و جهت‌گیری نظرات ارائه شده توسط کاربران محصولات راجع به یک محصول در ارزیابی محصول و ارائه بازخورد به سازنده و تولیدکننده جهت ورود و یا ارتقاء محصولات موجود و یا ارائه بهتر خدمات کمک می‌نماید [25]. همچنین باعث ارائه تجزیه‌وتحلیل آماری کاملاً هوشمند و مشاوره‌های آنلاین به هر دو گروه مشتری و بازرگان [26] می‌گردد. در حوزه سیاسی تحلیل احساسات در فروم‌های سیاسی برای تحلیل جهت‌دهی و سوگیری مردم در قبال کارزار سیاسی [27,28] مورداستفاده قرار می‌گیرد. برای ارائه‌دهندگان آموزشهای الکترونیکی، تحلیل احساس توانایی تشخیص سریعتر مشکلاتی که ممکن است در حین آموزش برای یادگیرندگان رخ دهد را فراهم می‌نمایند[29].

تمایل بسیار فراوان افراد جامعه در به‌اشتراک‌گذاری نظرات خود در شبکه‌های اجتماعی و همچنین بهره‌گیری از نظرات افراد دیگر برای تصمیم‌گیری بهتر باعث توجه بسیار زیادی به این حوزه و درنتیجه تولید حجم عظیمی از این اطلاعات بسیار سودمند گردیده و عملاً امکان بررسی دستی این اطلاعات را ناممکن ساخته است. ازاین‌رو اهمیت مسئله تجزیه‌وتحلیل احساسات در مسائل اقتصادی، سیاسی، بهداشتی و تجاری باعث افزایش علاقه‌مندی پژوهشگران در حوزه‌ی علوم داده‌کاوی و پردازش زبان طبیعی گردیده است.

نیاز به دستگاه‌هایی که فارغ از حوزه‌ی مورد تجزیه‌وتحلیل، قادر به ارزیابی و کشف احساسات موجود در سطوح مختلفی همچون سند یا جمله بوده و توانایی پردازش حجم عظیمی از این نظرات غالباً برخط را با دقت و صحت بهتری داشته باشند بر هیچ‌کس پوشیده نیست. روش‌های یادگیری ماشینی به سبب نیاز به داده‌های آموزشی فراوان و عدم انعطاف‌پذیری در حوزه‌های مختلف و همچنین نیاز به بروزرسانی‌های متوالی قادر به پاسخگویی در این زمینه نیستند. واژه‌نامه‌های احساسی موجود نیز با درنظرگرفتن این موضوع که غالباً تنها حاوی کلمات احساسی بوده و توجه کمتری به نقش کلمات تقویت‌کننده، تضعیف‌کننده و منفی‌کننده داشته‌اند قادر به ارزیابی دقیقی نخواهند بود.

در این پایان‌نامه باتوجه‌به چالش‌های ارائه شده در بخش قبل و به جهت رفع کاستی‌های موجود در روش‌های پیشین مبتنی بر واژه‌نامه، قصد ارائه یک واژه‌نامه احساسی با توانایی کاربست کلمات تقویت‌کننده و منفی‌کننده‌ی متوالی و تخصیص امتیاز به عبارات احساسی چندکلمه‌ای به جهت افزایش میزان دقت و صحت تشخیص احساسات در میان متون غیر ساختاریافته و جملات احساسی را داریم. متدولوژی در نظر گرفته شده استفاده از ابزارهای داده‌کاوی و متن‌کاوی به همراه خودکار ساختن روش استخراج اطلاعات خواهد بود. الگوریتم پایه‌ی موردنیاز با استفاده از پایگاه‌داده‌ی مبتنی بر آمازون پیاده‌سازی و قابلیت و کارایی آن در دسته‌بندی و ایجاد واژه‌نامه‌های خودکار به طور کامل نشان داده است.

نتایج به‌دست‌آمده از روش‌های پیشنهادی در این پایان‌نامه حکایت از تأثیر قابل‌توجه این دیدگاه و روش‌های مبتنی بر آن در افزایش میزان دقت و صحت تشخیص جملات و متون احساسی داشته و نشان از درستی هدف‌گذاری تعریف شده در این پایان‌نامه است.

۵-۱ برخی اصلاحات و مفاهیم موجود

- **عقیده‌کاوی**: به استفاده از روش‌های پردازش زبان طبیعی، تحلیل متن، زبان‌شناسی رایانشی[1] و زیست‌سنجشی[2] برای شناسایی، استخراج، کیفیت‌یابی، مطالعه‌ی وضعیت، عواطف و اطلاعات موضوعی سیستماتیک گفته می‌شود.
- **نظر**[3]: به متنی احساسی که نظر کاربر را در مورد یک موجودیت مشخص ارائه می‌دهد، نظر گفته می‌شود.
- **موجودیت**[4]: یک موجودیت، یک شئی حقیقی و یا انتزاعی مانند محصول، خدمت، انسان، رویداد، سازمان و... است که در مورد آن نظر داده می‌شود.
- **منظر**: منظر یا وجه که به آن ویژگی گفته می‌شود، جزئی از موجودیتی است که در یک نظر موجود است. یک ویژگی به دو صورت صریح و ضمنی در یک نظر به کار می‌رود. چنانچه به‌صورت صریح در متن به یک ویژگی اشاره شود، به آن ویژگی صریح و در غیر اینصورت به آن ویژگی ضمنی گفته می‌شود.
- **کلمات احساسی:** به کلماتی گفته می‌شود که تمایل و جهت‌گیری نظر کاربر را در مورد یک موجودیت و یا مؤلفه‌ای از آن موجودیت را نشان می‌دهد.
- **واژه‌نامه‌ی احساسی**: مجموعه‌ای است از کلمات یا عبارات که معمولاً برای ابراز احساسات مورداستفاده قرار می‌گیرد.

[1] Computational Linguistics
[2] Biometrics
[3] Review
[4] Entity

- **کلمه‌ی بذر:** یک کلمه‌ی بذر کلمه‌ای با میزان قطبیت احساسی شناخته شده است. یک seed lexicon مجموعه‌ای از کلمات بذر است. کلمات بذر غالباً مجموعه‌ی کوچکی از کلمات هستند که از لحاظ قطبیت به‌شدت مثبت یا به‌شدت منفی هستند.

- **ابهام‌زدایی کلمه‌ی احساسی:** فرایند شناسایی احساس کلمه درون جمله زمانی که کلمه دارای چند معنی مختلف است.

- **سطح جنبه**[1]: تمرکز بر تمام حالات احساسات موجود در سندِ داده شده و جنبه‌ای که به آن اشاره می‌شود.

- **آزمون آزگود:** هر واژه را دارای دو سطح معنایی می‌داند: معنای صریح که همان معنای عینی و ارجاعی آن است و معنای ضمنی که متأثر از بازخوردها، نگرش‌ها و ارزیابی‌ها است. آزگود این جنبه‌ها را به چهار دسته‌ی ارزیابی، قدرت، فعالیت و خطر تقسیم‌بندی کرده است.

- **شُمول معنایی**[2]: در زبان‌شناسی، در ساخت معنایی, رابطه‌ای است که بین عناصر واژگان خاص و عام برقرار است. به حالتی اشاره دارد که معنای یک واژه‌ی (خاص) به طور کامل توسط یک واژه‌ی دیگر (عام) که معنای فراگیرتری دارد پوشش داده شود، برای نمونه معنای چای توسط معنای نوشیدنی پوشش داده می‌شود و معنای سرخ توسط معنای رنگ. در این نمونه‌ها واژه‌ی چای و سرخ واژه‌های ***زیرشمول***[3] و نوشیدنی و رنگ واژه‌های ***فراشمول***[4] نامیده می‌شوند. واژه‌هایی که تحت یک واژه‌ی کلی (بالاتر) قرار گرفته‌اند، واژه‌های ***هم‌شمول***[5] نامیده می‌شوند، بنابراین، کلمه‌های اسب و سگ که هر دو تحت معنای حیوان قرار دارند کلمه‌هایی هم‌شمول‌اند. واژه‌ی حیوان نیز در اینجا واژه‌ای فراشمول است.

[1] Aspect Level
[2] Hyponymy
[3] Hyponym
[4] Hypernym
[5] Cohyponym

- **مدل هیجانی PAD:** مدل هیجانی PAD یک مدل روان‌شناختی است که توسط آلبرت مهرابیان و جیمز راسل برای توصیف و اندازه‌گیری حالت‌های هیجانی توسعه‌یافته است. PAD از سه بعد عددی، خوشایندی، تحریک‌پذیری و سلطه برای نشان‌دادن تمام هیجان‌ها استفاده می‌کند.

مقیاس خوشایندی - ناخوشایندی: سنجشی از میزان مطبوع یا نامطبوع بودن چیزی است که احساس می‌شود به‌عنوان‌مثال خشم و ترس هر دو هیجان نامطبوع هستند و در کنار هیجان‌های نامطبوع دسته-بندی می‌شوند. بااین‌حال لذت یک هیجان مطبوع است.

مقیاس تحریک‌پذیری- تحریک‌ناپذیری: می‌سنجد که چه چیزی پرانرژی یا کرخت احساس می‌شود و به مسئله‌ی فعالیت فیزیکی و تغییرات روانی - فیزیولوژیک اشاره می‌کند. نکته‌ی مهم این است که این سنجش شدت هیجان نیست. اندوه و افسردگی می‌توانند دارای تحریکات کم و احساسات شدید باشند. درحالی‌که خشم و غضب هیجان نامطبوع هستند، غضب دارای شدت بالاتر یا یک حالت تحریک‌پذیری بالاتر است. بااین‌حال ملالت که یک حالت نامطبوع است دارای ارزش تحریک‌پذیری کم است.

- **شبکه‌ی واژگان:** شبکه‌ی واژگان حاوی لغات زبان طبیعی در قالب مجموعه‌های کلمات هم‌معنی[1] است که باتوجه‌به نقش نحوی همچون فعل، اسم، صفت و قید دسته‌بندی شده‌اند. گروه‌های هم‌معنی در هر بخش دستوری دارای روابط معنایی خاص خود هستند. به‌عنوان‌مثال کلمات در مقوله‌ی اسم باهم دارای روابط معنایی مانند شمول معنایی، تضاد معنایی[2]، رابطه شمول معنایی دربرداشتن[3]، روابط سلسله‌مراتبی[4] زیرشمول و فراشمول و... دارند. شبکه‌ی واژگان غالباً برای ابهام‌زدایی و تعیین شباهت معنایی در کاربردهای مختلف پردازش زبان طبیعی و بازیابی اطلاعات مانند ترجمه ماشینی، استخراج اطلاعات، خلاصه‌سازی و... استفاده می‌شود.

- **عبارات دووجهی:** عباراتی که شامل حداقل یک کلمه‌ی مثبت و حداقل یک کلمه‌ی منفی باشند.

[1] Synonymous Set
[2] Antonymy
[3] Meronymy
[4] Taxonomic

- **سیستم بین‌المللی تصاویر عاطفی**[1]: این تصاویر شامل بیش از ۱۰۰۰ تصویر است که هنجارهای عاطفی تصاویر آن (خوشایندی، انگیختگی و تسلط) است. این مجموعه تصاویر شامل تصاویر مختلفی مانند قطع عضو، مارها، حشرات، صحنه‌های مربوط به حمله، تصادفات، آلودگی، بیماری، فقر، کودکان، مناظر طبیعی، تصاویر عاشقانه و... می‌باشد که به نوبه‌ی خود باعث ایجاد هیجان‌های اساسی و جهان-شمول مانند خشم، نفرت، غم، شادی، لذت، آرامش و... می‌شوند.

- **آزمون خودارزیابی تصاویر آدمک:** یک نمایش تصویری از ابعاد PDA است که به‌وسیله لانگ به‌عنوان جایگزینی برای مقیاس‌های خودگزارشی ابداع شده است. نمایش‌های آن در هر بعد با یک ویژگی تصویری در یک مقیاس ۹ درجه‌ای است که پاسخ‌دهنده می‌تواند آنچه را که احساس می‌کند انتخاب کند. SAM از شکل‌های آدمک در یک مقیاس برای هر یک از ابعاد عاطفی استفاده می‌کند. لانگ SAM را به‌عنوان یک مقیاس کاربردی تصویری برای سنجش ابعاد خوشایندی، انگیختگی و تسلط ابداع کرد. چون SAM یک ابزار میان‌فرهنگی (نابسته به فرهنگ و نابسته به زبان است)، برای استفاده در کشورها و فرهنگ‌های مختلف مناسب است.

- **قطبیت اولیه**[2]: درنظرگرفتن قطبیت نرمال یک کلمه در تمامی زمینه‌ها به‌جای درنظرگرفتن قطبیت کلمه در زمینه‌ی خاصی که آن کلمه یافت شده است.

- **Synset**[3]: مجموعه‌ای از کلمات مترادف و هم معنا با یک تعریف.

- **محاسبات Sentic:** روشی برای نظرکاوی و تحلیل احساسات که از تکنیک‌های هوش مصنوعی و وب معنایی برای سازمان‌دهی بهتر، تفسیر و پردازش نظرات و احساسات در متون زبان طبیعی استفاده می‌نماید. این روش یک روش یادگیری آماری نبوده و از ابزارهای استدلال عقلی و هستان‌شناسی خاص دامنه استفاده می‌نماید. این روش برخلاف طبقه‌بندی‌های آماری که نیازمند داده‌های ورودی عظیمی

[1] International Affective Picture System
[2] Prior Polarity
[3] Sets of synonyms

برای ارزیابی مطلوب هستند قادر به تجزیه‌وتحلیل اسناد، نه‌تنها در سطح صفحه یا پاراگراف که حتی در سطح جمله نیز می‌باشد. به طور خلاصه محاسبات sentic شامل تکنیک‌های هوش مصنوعی و وب‌معنایی برای ارائه و استنتاج دانش است، ریاضیات برای انجام فرایند جستجوی گراف و کاهش چند بعدی، زبان‌شناسی برای تجزیه‌وتحلیل گفتمان، روان‌شناسی برای مدل‌سازی شناختی و عاطفی، جامعه‌شناسی برای درک پویای شبکه‌های اجتماعی و نفوذ اجتماعی آنها و درنهایت اخلاقیات برای موضوعات مرتبط با ماهیت ذهن است[30].

۱-۶ ساختار پایان‌نامه

مطالب عنوان شده در این رساله در غالب ۵ فصل ارائه شده است. در فصل نخست که به مقدمه‌ی پژوهش اختصاص دارد به بیان برخی از دلایل ضرورت انجام این پژوهش و مزایای استفاده از واژه‌نامه‌های Ngram احساسی پرداخته شده است. در فصل ۲ به بررسی مفهوم نظرکاوی یا تجزیه‌وتحلیل احساسات خواهیم پرداخت و هرآنچه برای درک مطالب و اصطلاحات بکار رفته در این رساله موردنیاز است را ارائه خواهیم نمود. سپس در فصل ۳ به معرفی و بررسی انواع روش‌های تولید واژه‌نامه‌های احساسی خواهیم پرداخت و تعدادی از مهم‌ترین واژه‌نامه‌های احساسی تولید شده تاکنون را از لحاظ مراحل، نحوه‌ی ساخت و عناصر تشکیل‌دهنده‌ی واژه‌نامه‌ی احساسی مورد بررسی قرار خواهیم داد. در ادامه و در فصل ۴ به بیان روش پیشنهادی اول و پایگاه‌داده‌های مورداستفاده در واژه‌نامه‌های احساسی پیشنهادی پرداخته و روش پیشنهادی نخست را توضیح داده و نتایج هریک را ارائه خواهیم داد. در فصل ۵ نیز روش پیشنهادی دوم و نتایج حاصل از مقایسه واژه‌نامه پیشنهادی با واژه‌نامه‌های دیگر ارائه شده است. در نهایت به جمع‌بندی مطالب پرداخته شده است و برخی از موضوعاتی که می‌تواند در آینده به‌منظور بهبود کارایی واژه‌نامه احساسی و الگوریتم‌های تشخیص احساس مورد بررسی قرار گیرد ارائه شده است.

فصل دوم

مفاهیم مقدماتی

فصل ۲–مفاهیم مقدماتی

۲-۱ مقدمه

ازآنجایی‌که معمولاً فرایند تصمیم‌گیری افراد عموماً تحت تأثیر تفکرات رهبران فکری و یا نظرات عمومی دیگران است، مسئله‌ی نظرکاوی یکی از موضوعات بسیار مهم در طول تاریخ زندگی بشر محسوب می‌شود. آنالیز احساسات و نظرکاوی یکی از مهم‌ترین کاربردهای پردازش زبان طبیعی است که با استخراج خودکار نظرات یا احساسات انسان‌ها نسبت به محصولات و سرویس‌ها از محتواهای تولید شده توسط کاربران، سروکار دارد [31]. به‌طورکلی هدف از تجزیه‌وتحلیل احساسات شناسایی اطلاعات ذهنی و تعیین جهت‌گیری احساسی در اسناد است [32].

با رشد روزافزون رسانه‌های فراگیر اجتماعی همچون توییتر، فیس‌بوک، ویچت، میکرو بلاگ‌ها، مقالات خبری و ... تعداد بسیار زیادی از گروه‌های سیاسی و اجتماعی، اقتصادی، بهداشتی، آموزشی و نظامی علاقه‌مند به کسب نظرات گروه‌های هدف برای تصحیح، تغییر یا بهبود عملکرد خود هستند. به‌عنوان‌مثال باتوجه‌به سرعت تغییر و افزایش پیچیدگی در بازارهای مالی اطلاعات حاصل از تجزیه‌وتحلیل احساسات به شرکت‌ها و کمپانی‌ها برای شناسایی خطرات و ریسک‌های مختلف مرتبط با سرمایه‌گذاری‌های جدید و بهبود محصولات و خدمات کمک می‌نماید [33]. سیاستمداران نیز مایلند تا از میزان رضایتمندی و یا نارضایتی رای‌دهندگانشان از سیاست‌های اعمالی و درنتیجه میزان محبوبیت خود مطلع گردند [34,35]. دولت‌ها نیز تمایل دارند تا با استفاده از رسانه‌های داخلی و خارجی از نتیجه اعمال سیاستهایشان مطلع گردند [36]، تجزیه‌وتحلیل احساسات منجر به تشخیص بهتر و سریعتر معایب و کاستی‌های موجود در حین آموزش‌های الکترونیکی گردیده و امکان برطرف نمودن سریعتر این معایب را برای ارائه‌دهندگان آموزش فراهم می‌نماید[29]. پزشکان با بررسی نظرات و احساسات مشترک بیماران قادر به ارزیابی و بهبود شیوه‌های درمانی خود می‌گردند[37]. برخی از محققان با استفاده از تجزیه‌وتحلیل احساسات جنبه‌های مهم محصولات را جهت سرمایه‌گذاری و تولید محصولات جدید شناسایی می‌نمایند[38,39].

تجزیه‌وتحلیل احساسات فرایندهای مختلفی از قبیل شناسایی جملات ذهنی[1][40,41]، دسته‌بندی قطبی اسناد[42,43]، طبقه‌بندی بر مبنای شدت قطبیت[2][44,45]، پاسخ به پرسشهای احساسی[45,46]، تشخیص نظرات اسپم[49–47]، تشخیص جرم [23,50]، خلاصه‌سازی نظرات [32,51] و بسیاری از فرآیندهای دیگر را در بر می‌گیرد.

۲-۲ اصطلاح‌شناسی و مفاهیم پس‌زمینه

«خِرَد»، با تعریف اصطلاحات آغاز می‌شود. «سقراط»

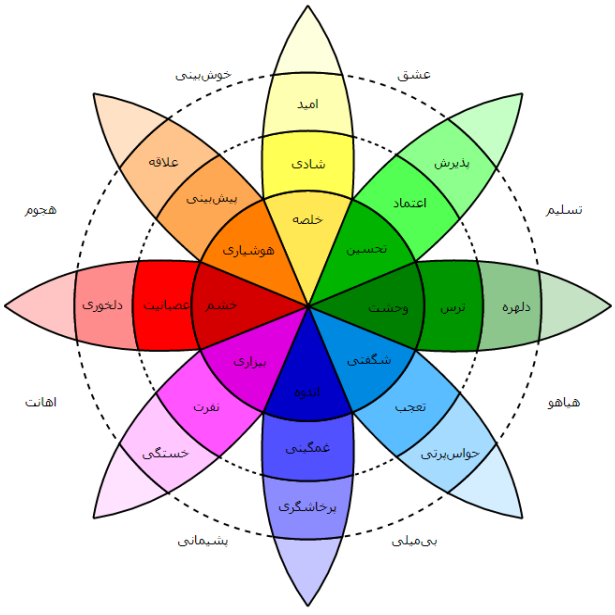
به آنالیز و استخراج خودکار جهت‌گیری احساسی، تحت عناوین مختلفی همچون، آنالیز احساسات[52]، ذهنیت[3][53,54]، نظرکاوی[52]، آنالیز حالت[4][55,56]، ارزیابی[5][57]، دیدگاه[6][58,59] و برخی از لغات دیگر اشاره شده است. در این بخش به توصیف مختصر اصطلاحات و مفاهیم پس‌زمینه در حوزه‌ی تجزیه‌وتحلیل احساسات می‌پردازیم همچنین تفاوت مابین هیجان و احساسات را از دیدگاه روان‌شناسی موردبحث قرارداده و تعریفی از تجزیه‌وتحلیل احساسات در رسانه‌های اجتماعی ارائه می‌نماییم و انواع مختلف نظرات را بیان می‌نماییم.

۲-۲-۱ هیجان[7] و احساسات[8] از دیدگاه روان‌شناسی

انجمن روان‌شناسی آمریکا هیجان را "به‌عنوانی الگوی پیچیده‌ای از عناصر تجربی، رفتاری و فیزیولوژیکی، که توسط آن فرد تلاش می‌نماید تا با یک موضوع یا یک رویداد مهم شخصی برخورد نماید[9]" تعریف کرده است. به‌طورکلی هیجان، احساس یا واکنشی است که یک شخص نسبت به یک رویداد خاص از خود بروز

[1] Subjectivity Detection
[2] Intensity Classification
[3] Subjectivity
[4] Analysis Of Stance
[5] Appraisal
[6] Point Of View
[7] Emotion
[8] Sentiment
[9] https://dictionary.apa.org/emotion

می‌دهد. پل‌اکتمن[1] شش احساس اولیه، یعنی شادی، غم و اندوه، خشم، ترس، تعجب و انزجار را به‌عنوان عناصر تشکیل‌دهنده هیجان پیشنهاد کرده است. بعدها روانشناسان دیگر با اضافه‌کردن احساساتی همچون غرور، هیجان، خجالت، تحقیر و شرم این لیست را غنی‌تر کرده‌اند. شکل ۱ مدل ابعادی پلاچیک را نشان می‌دهد که به‌صورت دوایر هم مرکز مرتب شده‌اند به نحویکه دوایر درونی‌تر بنیادی‌تر بوده اما دوایر بیرونی‌تر از ترکیب هیجان‌های مجاور به وجود آمده‌اند. توجه به این نکته مهم است که در دنیای کنونی دو کلمه هیجان



شکل ۱. مدل ابعادی رابرت پلاچیک

و احساسات معادل یکدیگر در نظر گرفته می‌شود اما احساسات عمومی‌تر بوده و قطبیت هیجان را شامل می‌شود (به‌عنوان‌مثال مثبت، منفی یا خنثی)[60].

۲-۲-۲ مؤلفه‌های کلیدی یک نظر[2]

بر مبنای[61]، یک نظر از پنج مؤلفه‌ی کلیدی تشکیل شده است. محققان یک نظر را به‌عنوان یک پنج‌گانه به‌صورت زیر تعریف می‌نمایند.

$$(e_i , a_{i\,j} , s_{i\,j\,kl} , h_k , t_l)$$

[1] Paul Eckman
[2] Opinion

جایی که e_i نام یک موجودیت یا شئی، a_{ij} یک جنبه از e_i، $S_{i\,j\,Kl}$ احساسات بر روی جنبه‌ای از $a_{i\,j}$ از موجودیت e_i، h_k دارنده‌ی نظر[1] و t_l زمانی است که نظر به‌وسیله‌ی h_k بیان شده است. احساسات $S_{i\,j\,Kl}$ می‌تواند مثبت، منفی یا خنثی بوده و یا با سطوح مختلفی از قدرت/شدت «به‌عنوان‌مثال ۱ تا ۵ ستاره» بیان گردد.

از سوی دیگر همان‌طور که کاربران مختلف از موقعیت‌های جغرافیایی متفاوت موضوعات محلی خود را به‌وسیله‌ی توییت‌های مختلف پست می‌نمایند، می‌توان از اطلاعات موقعیتی پایه برای تشخیص رویدادهای محلی در حال ظهور، حدس زدن محل کاربر و آشکار نمودن ویژگی‌های رفتاری کاربران محلی استفاده نمود. از دیدگاه لین و همکاران در[26] می‌توان مؤلفه‌ی دیگری به مجموعه پنج‌تایی فوق اضافه نمود و شش‌تایی زیر را تولید نمود.

$$(e_i\ ,\ a_{i\,j}\ ,\ s_{i\,j\,klt}\ ,\ h_k\ ,\ t_l\ ,\ p_t)$$

به‌نحوی‌که مؤلفه‌ی ششم یعنی p_t محلی است که احساسات توسط h_k بیان شده است. توجه داشته باشید که یک موجودیت می‌تواند مسائل، محصولات، سازمان‌ها، خدمات، حوادث، افراد، موضوعات و غیره را در بربگیرد[61].

۲-۲-۳ انواع مختلف نظرات

انواع مختلفی از نظرات در نظرکاوی معرفی و مقایسه شده‌اند که نظرات منظم[2] و تطبیقی[3]، نظرات صریح[4] و ضمنی[5]، ذهنیت[6] و هیجان[7] را شامل می‌شوند. در این بخش ما به طور خلاصه این موضوعات را مورد بررسی قرار می‌دهیم.

[1] Opinion Holder
[2] Regular Opinions
[3] Comparative Opinions
[4] Explicit
[5] Implicit
[6] Subjectivity
[7] Emotion

نظرات منظم[1]: این نظرات که غالباً در ادبیات با عنوان یک نظر استاندارد به آن اشاره می‌شود از دو زیرگروه اصلی تشکیل شده است

- *نظر مستقیم:* عقاید مستقیم بیان‌کننده‌ی نظرات بر روی یک شئی یا یک جنبه از شئی است و مستقیماً در مورد موجودیت وجود دارد (به‌عنوان‌مثال، "کیفیت تصاویر دوربین آیفون بسیار عالی است").
- *نظر غیرمستقیم:* این عقاید بر اساس تأثیرات آن بر روی موجودیت‌های دیگر به‌صورت غیرمستقیم ابراز می‌گردند (به‌عنوان‌مثال عبارت "بعد از پایان مسابقه تمام بازیکنان احساس خستگی کردند") بیانگر اثر نامطلوب مسابقه بر تمام بازیکنان است و به طور غیرمستقیم نظر یا احساس منفی دررابطه‌با مسابقه را بیان می‌نماید.

نظرات تطبیقی: این نظرات رابطه بین دو یا چند موجودیت را از لحاظ شباهت، اختلاف و یا اولویت بر اساس جنبه‌های مشترک آنها بیان می‌نمایند (به‌عنوان‌مثال "لینوکس عملکرد بهتری نسبت به ویندوز دارد").

نظرات صریح: یک نظر صریح یک نظر تطبیقی یا منظم است که به‌صورت گزاره‌ای ذهنی[2] ارائه می‌شود (به‌عنوان‌مثال روشنایی صفحه‌نمایش آیفون عالی است).

نظرات ضمنی: گزاره‌ای عینی[3] است که بر یک نظر منظم یا مقایسه‌ای دلالت داشته و معمولاً بیان‌کننده یک واقعیت است. (به‌عنوان‌مثال "نجات سرباز رایان بسیار خشن‌تر از "تنها بازمانده" است"). همان‌طور که مشخص است درک نظر پنهان در این مثال برای انسان دشوار است.

[1] Regular Opinions
[2] Subjective Statement
[3] Objective Statement

ذهنیت: مفهوم ذهنیت به جنبه‌هایی از زبان گفته می‌شود که برای بیان عقاید، احساسات، ارزیابی‌ها و گمانه‌زنی‌ها استفاده می‌شود. به‌عبارت دیگر یک جمله عینی[1] بیان‌کننده حقایق و اطلاعات بوده در حالی کهه یک جمله ذهنی بیانگر برخی از احساسات، دیدگاه‌ها یا عقاید شخص می‌باشد.

هیجان: هیجانات، احساسات و افکار ذهنی هستند که با احساسات مرتبط هستند. شدت احساسات یا یک نظر معمولاً با شدت ایجاد هیجان مرتبط است مثلاً "خشم و شادی".

۲-۳ نظرکاوی و تحلیل احساسات

در این بخش قصد داریم تا به‌صورت مختصر تمامی جنبه‌های مختلف نظرکاوی و تحلیل احساسات را توصیف نماییم. همان‌طور که در شکل ۲ مشاهده می‌نمایید سه حوزه‌ی مختلف در این زمینه عبارت‌اند از: آنالیز احساسات ازلحاظ عملکرد، آنالیز احساسات از لحاظ سطح مورد تجزیه‌وتحلیل و آنالیز احساسات از لحاظ روش مورداستفاده.

۲-۳-۱ آنالیز احساسات مبتنی بر عملکرد

همان‌طور که در شکل ۲ نشان‌داده‌شده است در این بخش به بررسی پنج عملکرد مختلف در تجزیه‌وتحلیل احساسات خواهیم پرداخت. این پنج عملکرد عبارت‌اند از طبقه‌بندی قطبی، شدت[2] یا انگیختگی[3] در مقیاس مشخص، طبقه‌بندی فراتر از قطبیت، شناسایی ذهنی/عینی[4] و تجزیه‌وتحلیل احساسات مبتنی بر ویژگی/جنبه.

۲-۳-۱-۱ طبقه‌بندی قطبی

در تجزیه‌وتحلیل احساسات، طبقه‌بندی قطبی فرایندی است که در طی آن قطبیت یک متن یا سند و یا جمله بر مبنای سه طبقه مثبت، منفی و خنثی دسته‌بندی می‌گردد. روش‌های متنوع و مختلفی برای تشخیص قطبیت مورداستفاده قرار می‌گیرد. این روش‌ها در سه گروه روش‌های مبتنی بر یادگیری ماشین[15,16]،

[1] Objective
[2] Valence
[3] Arousal
[4] Subjectivity/Objectivity Identification

روش‌های مبتنی بر دیکشنری [12,14] و روش‌های ترکیبی[17,18] دسته‌بندی می‌شوند. این روش‌ها در فصل بعد توصیف شده‌اند.

شکل ۲. تحقیقات آنالیز احساسات از جنبه‌های چندگانه

۲-۳-۱-۲ طبقه‌بندی بر مبنای سطوح شدت در یک مقیاس مشخص

به‌طورکلی طبقه‌بندی احساسات را می‌توان از دو جنبه موردتوجه قرارداد. جنبه نخست انتساب قطبیت یک سند است که تحت عنوان طبقه‌بندی قطبی شناخته می‌شود. جنبه دیگر این طبقه‌بندی انتساب میزان شدت احساسی یک سند بوده که میزان قطبیت یک سند را بر مبنای روش‌های مختلفی همچون دسته‌بندی

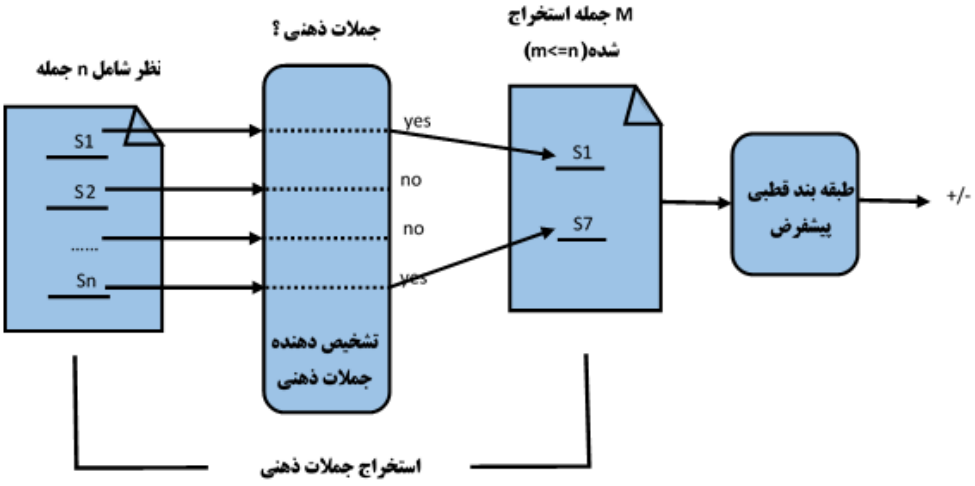
دو کلاسه (مثبت ومنفی) و یا طبقه‌بندی سه کلاسه (مثبت، منفی و خنثی) در مقیاس‌های عددی مشخص می‌نماید.

۲-۳-۱-۳ فراتر از طبقه‌بندی قطبی (نهاد، صاحب‌نظر، اطلاعات مکانی و تجزیه‌وتحلیل اطلاعات زمان)

علاوه بر طبقه‌بندی قطبی و همچنین مشخص نمودن میزان شدت قطبیت یک سند یا عبارت احساسی در یک دسته‌بندی ریزتر می‌توان شناسایی انواع موجودیت‌های احساسی از قبیل صاحب‌نظران[62][1] دیدگاه‌ها[2][63]، نظرسنجی‌ها[3][64]، جنبه‌های هدف[4][65] و منابع تفکر[5][66] موردتوجه قرارداد.

۲-۳-۱-۴ شناسایی ذهنیت یا عینیت

اطلاعات متنی را می‌توان به دو دسته عینی یا ذهنی طبقه‌بندی نمود. اظهارات عینی نشان‌دهنده حقایق و اطلاعات هستند درحالی‌که نظرات و دیدگاه‌های افراد در دسته اظهارات ذهنی قرار می‌گیرند. در برخی از مواقع تشخیص ذهنی یا عینی بودن مشکل‌تر از یک مسئله طبقه‌بندی قطبیت است. تشخیص این موضوع که کدام بخش از اسناد واقعیت و کدام بخش از اسناد نظر یا عقیده نویسنده است می‌تواند در طبقه‌بندی احساسی سند سودمند باشد.



شکل ۳. حذف جملات عینی برای بهبود کلاس‌بندی قطبی

[1] Opinion Holders
[2] Opinion Expressions
[3] Opinion Targets
[4] Aspects Of A Target
[5] Opinion Sources

همان‌طور که در شکل ۳ مشاهده می‌نمایید پنگ و همکاران در[67] نشان داده‌اند که حذف جملات عینی از یک سند پیش از طبقه‌بندی احساسات آن سند می‌تواند سودمند باشد. به‌طورکلی یک عبارت ذهنی هر کلمه یا عبارتی است که بیان‌کننده‌ی یک نظر، احساس، ارزیابی، حدس، یا گمانه‌زنی باشد[68]. در حالی کهه یک عبارت عینی حاوی واقعیت‌های موجود درارتباطبا یک موجودیت است.

از سوی دیگر ذکر این نکته نیز بسیار ضروری است که یک حقیقت برای تمامی حالات درست است درحالی‌که یک نظر مبتنی بر ذهنیت افراد بوده و لذا می‌تواند از فردی به فرد دیگر تغییر نماید و این تفاوت نظرات یکی از مسائل بسیار چالش‌برانگیز محسوب می‌گردد.

۲-۳-۱-۵ تحلیل احساسات مبتنی بر ویژگی/جنبه

طبقه‌بندی احساسات در این سطح یک طبقه‌بندی ریزدانه است و هدف آن استخراج ویژگی‌های معنایی متن است که با جنبه‌های خاصی مطابقت دارد و قطبیت احساسی را برای آن جنبه خاص پیش‌بینی می‌نماید[69]. یک جنبه، ویژگی، صفت یا مولفه‌ای از موجودیت است لذا در این سطح دو موضوع جنبه و قطبیت مطرح می‌شود بدین ترتیب که نخست جنبه‌های موجودیت شناسایی شده و سپس برای هر کدام از آنها آنالیز احساسات صورت می‌پذیرد[70].

۲-۳-۲ سطوح مختلف آنالیز احساسات

به‌طورکلی تجزیه‌وتحلیل احساسات عمدتاً در سه سطح سند، جمله و سطح کلمه انجام می‌گیرد که در ادامه به توضیح مختصری دررابطه‌با هر یک از این سطوح خواهیم پرداخت.

۲-۳-۲-۱ تجزیه‌وتحلیل احساسات در سطح سند

هدف از تجزیه‌وتحلیل احساسات در سطح سند، تعیین قطبیت کلی سند به‌عنوان مثبت، منفی و خنثی است. به‌طورکلی احساسات در سطح متن را می‌توان به شکل پنج‌تایی زیر در نظر گرفت.

(e,GENERAL,t,h,so)

در این پنج‌تایی، e موجودیت موردنظر، h نویسند یا گوینده‌ی متن، t زمان ارائه و بیان متن، GENERAL و so جهت‌گیری کلی متن از جنبه‌ی کلی آن است[61]. این نوع تجزیه‌وتحلیل با این فرض که هر سند به طور خاص دررابطه‌با یک موجودیت واحد تولید شده است کاربرد داشته و برای اسنادی که چندین موجودیت مختلف را مورد ارزیابی و مقایسه قرار می‌دهند قابل استفاده نیست[71].

۲-۳-۲-۲ تجزیه‌وتحلیل احساسات در سطح جمله

در این سطح به هر جمله موجود در سند یک احساس نسبت داده می‌شود. ازآنجایی‌که جملات را می‌توان به دو گروه جملات عینی و جملات ذهنی دسته‌بندی نمود و باتوجه‌به این موضوع که جملات عینی[1] بیان-کننده‌ی حقایق و اطلاعات هستند و جملات ذهنی[2] بیان‌کننده‌ی احساسات، دیدگاه‌ها و نظرات روشن، لذا بدیهی است که در حوزه‌ی تجزیه‌وتحلیل احساسات در سطح جمله، ما با جملات ذهنی سروکار خواهیم داشت. ازاین‌رو فرایند تجزیه‌وتحلیل احساسات در سطح جمله در طی دو مرحله صورت می‌پذیرد نخستین مرحله استخراج جملات ذهنی و سپس تعیین قطبیت این جملات است[70].

۲-۳-۲-۳ تجزیه‌وتحلیل احساسات در سطح کلمه

کلمه مبنای تجزیه‌وتحلیل احساسات متن است. از سوی دیگر، کلمات می‌توانند در حوزه‌های مختلف معانی احساسی مختلفی را از خود بروز دهند و احساسات ابراز شده در یک عبارت یا یک جمله صرفاً مجموعه‌ی ساده‌ای از احساسات انتقال‌داده‌شده توسط کلمات آن نیست[72]، با این وجود کلمات احساسی عنصر بسیار مهمی برای هر الگوریتم تشخیص احساساتی محسوب می‌شوند. روش‌های تجزیه‌وتحلیل احساسات در سطح کلمه غالباً از واژه‌نامه‌های احساسی که قطبیت کلمات در آنها از پیش تعریف شده است استفاده می‌نمایند.

مطالعات متعددی در مورد جهت‌گیری عاطفی کلمات بر مبنای استفاده از واژه‌نامه‌های احساسی وجود دارد. به‌طورکلی پردازش در سطح کلمه را می‌توان در سه دسته‌ی زیر خلاصه نمود:

[1] Objective
[2] Subjective

- تعیین نظری بودن کلمه در سند (آیا کلمه ذهنی است یا عینی)
- تعیین گرایش یا قطبیت یک کلمه (مثبت یا منفی بودن کلمه)
- تعیین شدت و میزان گرایش یک کلمه

۲-۳-۳ تکنیک‌های آنالیز احساسات

در هر کلمه‌ای احساسی نهفته است، احساس یک عبارت به طور قابل ملاحظه‌ای می‌تواند متفاوت با احساس کلمات تشکیل‌دهنده‌ی آن باشد. ترکیب احساسات[1]، تعیین احساس یک واحد زبانی چندکلمه‌ای همچون یک عبارت یا یک جمله مبتنی بر اجزای آن است[73]. روش‌های موجود جهت آنالیز احساست را می‌توان به سه گروه تقسیم نمود: یادگیری با ناظر، یادگیرینیمه نظارتیی و یادگیری بدون ناظر.

۲-۳-۳-۱ رویکرد یادگیری با ناظر

این روش‌ها به داده‌های آموزشی برچسب‌گذاری شده و یک مجموعه از نمونه‌های آموزشی متکی هستند. الگوریتم‌های یادگیری با ناظر داده‌های آموزشی را تجزیه‌وتحلیل نموده و عملکردی را استنباط می‌نمایند که ما برای نگاشت آنها به نمونه‌های جدید استفاده می‌نماییم. این الگوریتم‌ها نیازمند پیکره‌های بزرگ آموزشی هستند که باید با استفاده از برچسب‌های احساسی حاشیه‌نویسی شده باشند. فرایند حاشیه‌نویسی یکی از بزرگترین معایب این روش‌ها تلقی می‌شود بااین‌حال مهم‌ترین مزیت آنها تنظیم خودکار پارامترهای مدل در فرایند یادگیری است. الگوریتم‌های مبتنی‌برقانون، درخت تصمیم، ماشین بردار پشتیبان، الگوریتم نایوبیز، شبکه‌های عصبی مصنوعی، یادگیری عمیق و درون یابی نمونه‌هایی از این گروه از الگوریتم‌ها هستند. مشکل اساسی در این‌گونه از روش‌ها کمبود پیکره‌های حاشیه‌نویسی شده‌ی بزرگ است که باعث حرکت به سمت روش‌های یادگیری جایگزین همچون یادگیرینیمه نظارتیی و یادگیری بدون ناظر گردیده است.

[1] Sentiment Composition

۲-۳-۳-۲ روش یادگیری بدون ناظر

هدف الگوریتم‌های بدون ناظر یافتن ساختارهای پنهان در داده‌های بدون برچسب است تا بدین‌وسیله مدل‌هایی را برای طبقه‌بندی احساسات ایجاد نمایند. این روش‌ها از شباهت‌ها و قاعده‌های موجود در داده‌ها برای کشف گروه‌بندی طبیعی استفاده می‌نمایند. ازآنجایی‌که در این روش‌ها هیچ سابقه یا پیش‌زمینه داده‌ای در مورد الگوی یک پدید وجود ندارد لذا این روش‌ها چالش برانگیزتر هستند[74]. برخلاف طبقه‌بندی، یادگیری بدون ناظر هیچ‌گونه وابستگی به طبقه‌های از پیش تعریف شده یا داده‌های آموزشی برچسب‌گذاری شده ندارد. خوشه‌بندی مشهورترین روش بدون ناظر است که فضای ورودی را بر مبنای اندازه‌گیری فاصله به تعدادی ناحیه تقسیم می‌نماید. الگوریتم تقسیم[۱]، الگوریتم سلسله مراتبی[۲] و الگوریتم مبتنی بر تراکم[۳] از جمله الگوریتم‌های بدون ناظر محسوب می‌شوند.

۲-۳-۳-۳ رویکرد یادگیری نیمه نظارتی

مشکل برچسب‌گذاری دستی پیکره‌های آموزشی در روش‌های یادگیری با ناظر را می‌توان با استفاده از برچسب‌گذاری اتوماتیک حل نمود. این فرایند با نام یادگیری نیمه نظارتی شناخته می‌شود. این روش‌ها به این سبب که توازنی مابین یادگیری با ناظر و یادگیری بدون ناظر پدید می‌آورند ارزشمند هستند. تکنیک‌های یادگیرینیمه نظارتیی (SSL[4]) قادر به استفاده از داده‌های بدون برچسب در فرایند آموزش و بهبود طبقه‌بندی در کاربردهایی که داده‌های برچسب‌دار در آن‌ها کمیاب است، می‌باشند. به‌طورکلی چندین روش کلاسیک در یادگیرینیمه نظارتیی وجود دارند که عبارت‌اند از: روش‌های مبتنی بر گراف، روش‌های مبتنی بر بسته‌بندی[5] و روش‌های مبتنی بر موضوع[۶].

[1] Partitional Algorithm
[2] Hierarchical Algorithm
[3] Density-Based Algorithm
[4] Semi-Supervised Learning
[5] Wrapper-Based Method
[6] Topic-Based Method

۲-۴ جمع‌بندی

در این فصل با نگاهی سطح بالا سعی شد تا با مبانی اولیه نظرکاوی و تجزیه‌وتحلیل احساسات آشنا شویم و مفاهیم موردنیاز جهت مطالعه فصول آتی شرح داده شود. به طور مشخص پس از آنکه به تعریف دقیق هیجان و احساسات از دیدگاه روان‌شناسی پرداختیم، مؤلفه‌های کلیدی یک نظر را باتوجه‌به دو دیدگاه معروف موردتوجه قرار داده و آنگاه در ادامه تفاوت نظرات تطبیقی و منظم مورد بررسی قرار گرفته و تفاوت مابین نظرات صریح و ضمنی بیان گردید در ادامه مسئله نظرکاوی و تجزیه‌وتحلیل احساسات را از جنبه‌های مختلف و سطوح متفاوت و تکنیک‌های موجود مورد بررسی قراردادیم. در این قسمت ضمن توصیف جملات ذهنی و عینی، انواع روش‌های طبقه‌بندی احساسات مورد بررسی قرار گرفته و سپس سطوح مختلف تجزیه‌وتحلیل احساسات بیان گردید. در ادامه با ارائه روش‌های مختلف آنالیز احساسات در رویکردهای با ناظر و بدون ناظر و نیمه نظارتی، برخی از مفاهیم مشترک مورداستفاده در سراسر این پایان‌نامه و فصول مختلف مورد بررسی قرار گرفته و شرح داده شد.

فصل سوم

پیشینه تحقیق

فصل ۳– پیشینه تحقیق

۳-۱ مقدمه

در این فصل ابتدا به مروری کلی بر کارهای انجام شده که در این حوزه صورت‌گرفته پرداخته و سپس در یک جمع‌بندی دقیق‌تر مقالات مرتبط با حیطه انتخاب شده در انتهای فصل به‌صورت دسته‌بندی آورده شده‌اند. ابتدا به سراغ پیشینه عمومی روش‌های طبقه‌بندی احساسات می‌رویم.

۳-۲ روش‌های طبقه‌بندی احساسات

به‌طورکلی دو روش اصلی برای مسئله‌ی تجزیه‌وتحلیل احساسات معرفی شده است. نخست روش‌های مبتنی بر واژه‌نامه که شامل محاسبه‌ی جهت‌گیری احساسی یک سند بر اساس جهت‌گیری معنایی لغات یا عبارات تشکیل‌دهنده‌ی آن سند است [75] و روش‌های طبقه‌بندی متن با استفاده از دسته‌بندی کننده‌ها که از روی نمونه‌های برچسب‌خورده‌ی متن یا جملات [71] عمل می‌نمایند و تحت عنوان روش‌های آماری یا یادگیری ماشین شناخته می‌شوند، این روش‌ها وابسته به دامنه بوده [76] و با ورود داده‌های جدید نیازمند یادگیری مجدد هستند. برخی تلاش‌ها [17,18] نیز برای ترکیب دو روش فوق با یکدیگر صورت پذیرفته است. انتخاب هرکدام ازاین‌رویکردها برای استفاده امری بسیار پراهمیت است، زیرا بر دقت طبقه‌بندی احساسات تأثیر می‌گذارد و نیازمند تطبیق دقیق با نوع داده‌های مورد آنالیز است [77].

۳-۲-۱ روش‌های مبتنی بر یادگیری ماشین

روش‌های مبتنی بر یادگیری ماشین را می‌توان روش‌هایی در حوزه هوش مصنوعی دانست که برای ایجاد توانایی یادگیری از داده‌ها به رایانه‌ها متمرکز است [74]. یک روش رایج برای تجزیه‌وتحلیل احساسات، انتخاب یک الگوریتم یادگیری ماشین و روشی برای استخراج ویژگی‌ها از متن‌ها و سپس آموزش طبقه‌بندی کننده با یک پیکره‌ی کدگذاری شده توسط انسان در یک حوزه‌ی خاص است. ویژگی‌های مورداستفاده غالباً کلمات هستند اما این ویژگی‌ها می‌توانند ریشه‌ی کلمه یا برچسب‌گذاری نحوی کلمات و همچنین ترکیبی از کلمات

به‌صورت Bigram و Trigram هم باشند[52]. یافتن مقادیر عددی ویژگی‌ها با استفاده از مکانیسم‌هایی از قبیل رتبه‌بندی TF-IDF، آموزش طبقه‌بندی کننده‌هایی از قبیل نایوبیز، ماشین بردار پشتیبان، ماکزیمم بی‌نظمی و... صورت می‌پذیرد و درنهایت از طبقه‌بندی‌های آموزش داده شده برای تشخیص قطبیت احساسی استفاده می‌گردد.

اگرچه که روش‌های مبتنی بر یادگیری ماشین نشان داده‌اند که در بسیاری از حالات دارای کارایی بهتری هستند[21] اما از دو مشکل اساسی رنج می‌برند، نخست آنکه ایجاد الگوهای آموزشی از کاربران واقعی بسیار وقت‌گیر و گران است [21] و دوم انتخاب ویژگی‌های صحیح و مقداردهی عددی به آن‌ها است[78]. بسیاری از محققان برای حل مسئله‌ی نخست از رویکرد اعتبارسنجی بین دامنه‌ای حمایت می‌نمایند[78] که در آن الگوهای آموزش و آزمایش از دو مجموعه داده‌ی غیر مرتبط هستند. برای مقابله با مشکل دوم روش‌هایی همچون استفاده از امتیازدهی TF-IDF مرتبط با Unigramها، Bigramها یا یک Ngram عمومی و روش‌های مبتنی بر واژه‌نامه برای امتیازدهی ویژگی‌ها قابل‌استفاده است[79].

از سوی دیگر طبقه‌بندی خودکار احساسات بیان شده در رسانه‌های اجتماعی به چند دلیل به چالش کشیده شده است. نخست اینکه شناسایی نظرات و احساسات از متن نیازمند به کسب درک عمیقی از قوانین صریح و ضمنی، باقاعده و بی‌قاعده و قوانین نحوی و معنایی زبان طبیعی دارد [80]. علاوه بر این در تجزیه‌وتحلیل احساسات با استفاده از پردازش زبان طبیعی ما با متن‌های بدون ساختار مواجه هستیم که معمولاً این گونه متن‌های تولید شده توسط کاربران که در رسانه‌های اجتماعی تولید می‌گردد، متداول به حساب می‌آید. به‌این‌ترتیب، محتوای متون تولید شده توسط کاربران در رسانه‌های اجتماعی به‌طورمعمول ماهیت فوری و غیررسمی را نشان می‌دهد [81]. محتویات این متن‌ها از لحاظ لغات تشکیل‌دهنده و گرامر کاملاً متنوع است [82]. معمولاً شامل اختصارات، غلط‌های املایی، شکلک‌ها و ایموجی‌ها بوده و غالباً از نحو پیام‌های کوتاه پیروی می‌نمایند که متدهای تجزیه‌وتحلیل احساسات فعلی معمولاً از آن‌ها پشتیبانی نمی‌نمایند. علاوه بر این ویژگی‌های پلتفرم‌های خاص، همچون محدودیت ۱۴۰ کلمه‌ای پیام‌ها در توییتر، مانع از کارایی ابزارهای

تجزیه‌وتحلیل خودکار احساسات می‌شوند [6]. حجم زیاد مکالمات رسانه‌های اجتماعی که یک چالش بزرگ محسوب می‌گردد و درنهایت حوزه محور بودن روش‌های مبتنی بر یادگیری ماشین و وابستگی آنها به دامنه‌ای خاص [76] از جمله معایب این روش‌ها محسوب می‌شود. ذکر این نکته ضروری است که قدرت روش‌های مبتنی بر یادگیری ماشینی به داده‌های آموزشی آن وابسته است و تولید مجموعه داده‌های آموزشی عالی و باکیفیت بالا بسیار زمان‌بر و پرهزینه است و چنین مجموعه‌هایی غالباً به‌ندرت در دسترس هستند.

۳-۲-۲ روش‌های مبتنی بر واژه‌نامه

یک واژه‌نامه‌ی احساسی که کلمات قطبی[1] کلمات حاوی نظر[2] یا کلمات opinion-bearing نیز نامیده می‌شود[83] مجموعه‌ای است از کلمات یا عبارات که معمولاً برای ابراز احساسات مورداستفاده قرار می‌گیرد[52]. در این روش‌ها، از ابزارهای پردازش زبان طبیعی برای استخراج کلمات احساسی استفاده نموده و سپس برای یافتن قطبیت کلی از واژه‌نامه‌های احساسی همچونSentiWordNet, SenticNet, VADER استفاده می‌نمایند.

معمولاً هر موجودیت درون این واژه‌نامه‌ها با یک امتیازدهی احساسی مرتبط است. در بیشتر واژه‌نامه‌ها این امتیازدهی تنها بیانگر جهت‌گیری‌های احساسی، مثبت، منفی و یا خنثی هستند [84]. برخی از واژه‌نامه‌ها[2,10,12] نیز از مقیاس برای نمایش شدت احساسات استفاده نموده‌اند. این روش‌ها می‌توانند با اطلاعات دیگری مانند، لیست شکلک‌ها و قوانین معنایی همچون نحوه‌ی برخورد با منفی‌کننده‌ها [12]تکمیل گردند. در برخی از واژه‌نامه‌های احساسی [12,85]از برچسب‌گذاری واژگان کلامی[3] برای هر کلمه استفاده شده است. این برچسب‌ها امکان تفسیر[4] کلماتی را که دارای چندین برچسب هستند فراهم می‌نمایند[68]. در روش‌های مبتنی بر واژه‌نامه، شاخص‌های احساسی درون سند برای طبقه‌بندی بر مبنای واژه‌نامه‌ی مورد

[1] Polar Words
[2] Opinion Words
[3] Pos Tag
[4]. بعنوان مثال کلمه "Good" زمانیکه بعنوان اسم مورد استفاده قرار می‌گیرد حاوی هیچ احساسی نیست اما زمانیکه بعنوان صفت مورد استفاده قرار می‌گیرد حاوی احساس مثبت است

جستجو قرار می‌گیرد [61] و احساس کلی سند به‌وسیله قطبیت غالب (به‌عنوان‌مثال مثبت یا منفی) بین شاخص‌ها تعیین می‌شود. این روش‌ها، وابسته به واژگان احساسی گسترده و دقیق هستند. بااین‌حال تأثیر واژه‌نامه‌های موجود در هنگام استفاده از آن‌ها در مسائل جدید بسیار محدود است[86].

واژه‌نامه‌های احساسی عمومی از دو جنبه نارسایی[1] و عدم دقت[2] دارای ضعف هستند [87]. نارسایی به این موضوع اشاره می‌نماید که در این نوع واژه‌نامه‌های احساسی عمومی، واژه‌های احساسی خاص برای یک حوزه وجود ندارد و عدم دقت به این معناست که جهت‌گیری احساسی یک لغت می‌تواند باتوجه‌به تغییر دامنه تغییر نماید[87,88].

۳-۲-۳ روش‌های ترکیبی

برخی از تلاش‌ها برای ترکیب دو روش مبتنی بر واژه‌نامه و روش‌های یادگیری ماشین صورت پذیرفته است. ترکیب این دو متد روشی برای پرهیز از طبقه‌بندی دستی داده‌ها جهت فاز آموزش فراهم می‌نماید[89]. ترکیب ویژگی‌های همچون پایداری[3] و قابلیت خوانی یک واژه‌نامه دقیق طراحی شده و دقت بالا یک الگوریتم یادگیری ماشین قدرتمند از جمله مزایای روش‌های ترکیبی محسوب می‌شود[78].

۳-۳ پدیده‌های زبانی

درحالی‌که تحقیقات بسیار زیادی بر روی توسعه سیستم‌های باناظر و بدون ناظر در حوزه آنالیز احساسات صورت پذیرفته است اما توجه بسیار کمی در مورد پدیده‌های مختلف زبانی در این حوزه صورت‌گرفته است[90]. پدیده‌های زبانی یک از مؤثرترین عوامل تغییر قطبیت اولیه کلمات و عبارات احساسی حوزه تجزیه‌وتحلیل احساسات محسوب می‌شوند. از جمله این عوامل می‌توان به نقش منفی‌کننده‌ها و تشدیدکننده‌ها در یک

[1] Insufficiency
[2] Inaccuracy
[3] Stability

عبارت احساسی اشاره نمود. عدم توجه دقیق به تأثیر منفی‌کننده‌ها و یا تشدیدکننده‌ها بر روی کلمات و عبارت احساسی باعث کاهش کارایی سیستم‌های تجزیه‌وتحلیل احساسات می‌شود.

۳-۳-۱ منفی‌کننده‌ها

منفی‌کننده‌ها یکی از مهم‌ترین پدیده‌های زبانی مؤثر در تغییر قطبیت کلمات احساسی محسوب می‌شوند. تعداد بسیار زیادی از محققان [91,92] نشان داده‌اند که توجه به ساختارهای زبانی همچون منفی‌کننده‌ها باعث افزایش کارایی سیستم‌های تجزیه‌وتحلیل احساسات می‌گردند.

به‌طورکلی ۴ روش برای برخورد با منفی‌کننده‌ها وجود دارد

- روش‌هایی که باعث معکوس شدن امتیاز کلمه احساسی پس از منفی‌کننده می‌شوند[95–93].
- روش‌هایی که تنها باعث شیف ثابت امتیاز احساسی کلمه احساسی می‌شوند[12,96,97].
- روش‌هایی که برای هر کلمه احساسی اثر متفاوتی را در نظر می‌گیرند یا روش‌های اکتشافی[2,6,7].
- روش‌هایی مبتنی بر یادگیری ماشینی[98].

باتوجه‌به پیچیدگی‌های بسیار زیاد منفی‌کننده‌ها و کاربردهای متفاوت آنها در تغییر قطبیت، تاکید بر منفی بودن و یا تضعیف ادعا[1] [99] بدون داشتن دانش وسیع و عمیق از نوع و کاربرد مورداستفاده هر منفی‌کننده در هر موقعیت و عدم توجه به تأثیر متقابل منفی‌کننده‌ها وجنبه‌های مختلف زبان بر یکدیگر، باعث کاهش کارایی سیستم‌های تجزیه‌وتحلیل احساسات خواهد شد.

۳-۳-۲ تشدیدکننده‌ها

همان‌طور که پیش‌تر نیز بیان گردید هرچند که اهمیت پدیده‌های زبانی بر هیچ‌کس پوشیده نیست اما تاکنون به‌صورت جامع و وسیعی به این مقوله‌های زبانی پرداخته نشده است[90]. آزمایشات صورت پذیرفته

[1] Make Weaker Claims

توسط سیستم‌های تجزیه‌وتحلیل احساساتی که از پدیده‌های زبانی مانند تشدیدکننده‌ها استفاده نموده‌اند نشان از افزایش میزان دقت و کارایی فرایند آنالیز احساسات داشته‌اند[100].

هرچند که تشدیدکننده‌ها به طور ذاتی دارای قطبیت نیستند[96] اما باعث تغییر شدت کلمات احساسی می‌گردند[101]. به‌طورکلی تشدیدکننده‌ها را می‌توان به دو گروه تقویت‌کننده و تضعیف‌کننده تقسیم نمود. در دسته اول تشدیدکننده‌هایی همچون very, entirely قرار دارد که باعث افزیش تأثیر یک کلمه احساسی می‌گردند این گروه با نامهایی همچون تقویت‌کننده[1] یا reinforcers، overstatements نامگذاری شده‌اند در گروه دوم تشدیدکننده‌هایی همچون quite, slightly قرار دارند که باعث کاهش تأثیر یک کلمه احساسی می‌شوند این گروه از تشدیدکننده‌ها با نامهایی همچون downtoners یا attenuators، understatements نامگذاری شده‌اند[102]. آنچه بر پیچیدگی عملکرد تشدیدکننده‌ها افزوده است تأثیر متقابل تشدیدکننده و کلمه احساسی پس از تشدیدکننده است. به نحویکه اثر تشدیدکننده‌ها را فراتر و پیچیده‌تر از یک جمع و تفریق ساده نموده است [103]. سیستم‌های مختلف از روش‌های متنوع و گوناگونی جهت اعمال این تأثیرات استفاده نموده‌اند. در یک جمع‌بندی کلی این روش‌ها را می‌توان به چهار دسته تقسیم نمود :

- اضافه یا کم کردن یک مقدار ثابت[94,97,104].
- دسته‌بندی تشدیدکننده‌ها در گروه‌های مختلف و تخصیص ضریب ثابتی به هر گروه[7,41,105]. به‌عنوان‌مثال الگوریتم ارائه شده توسط Jurek و همکاران تشدیدکننده‌ها را به سه دسته downtoners، weak amplifiers و strong amplifiers تقسیم‌بندی نموده است به نحویکه به ترتیب دارای ضریب تأثیر ۵۰٪- ،۵۰٪ و ۱۰۰٪ هستند.
- اختصاص یک ضریب ثابت به هر تشدیدکننده [12,96,106]. در این روش‌ها پروسه افزایش یا کاهش امتیاز یک کلمه در هنگام ترکیب با یک تشدیدکننده معمولاً به‌صورت ضرب امتیاز کلمه احساسی در امتیاز تشدیدکننده صورت می‌پذیرد.

[1] Amplifiers

- اعمال ضرایب متفاوت به هر تشدیدکننده باتوجه‌به کلمه احساسی پس از آن[2].

باتوجه‌به افزایش دقت در تمامی روش‌های تجزیه‌وتحلیل احساسات در هنگام استفاده از تشدیدکننده‌ها لزوم توجه بیش از پیش به تولید دیکشنری‌های احساسی عمومی که از ترکیبات مختلف تشدیدکننده‌های متوالی حمایت نمایند آشکار می‌گردد.

بدیهی است که اضافه یا کم کردن مقدار ثابت به امتیاز یک کلمه احساسی بدون در نظر گرفتن شدت کلمه احساسی پس از تشدیدکننده قادر به اعمال مطلوب اثر یک تشدیدکننده نخواهد بود. با توجه به خواص متفاوت هر تشدیدکننده دسته بندی تشدیدکننده‌ها در گروههای مختلف و اعمال ضرایب ثابت به هر گروه نیز قادر به اعمال اثرات متفاوت هر تشدیدکننده بر کلمات احساسی مختلف نیست. از سوی دیگر تخصیص ضریب به هر تشدیدکننده بدون در نظر گرفتن کلمه احساسی پس از آن نیز، به درستی اثر یک تشدید کننده را اعمال نخواهد نمود . لذا بر طبق تحقیقات ما، تنها الگوریتمهایی که

- اثر تشدید کننده بر کلمه احساسی را مورد توجه قرار دهد .
- تاثیر کلمه احساسی بر تشدیدکننده را مورد توجه قرار داده باشد
- اثر موقعیت مکانی تشدیدکننده نسبت به کلمه احساسی را مورد توجه قرار داده باشد.
- تمامی ترکیبات مختلف ممکن را تحت پوشش قرار داده باشد.

می توانند با عث ایجاد نتایج مطلوب‌تر در حوزه تجزیه و تحلیل احساسات باشند.

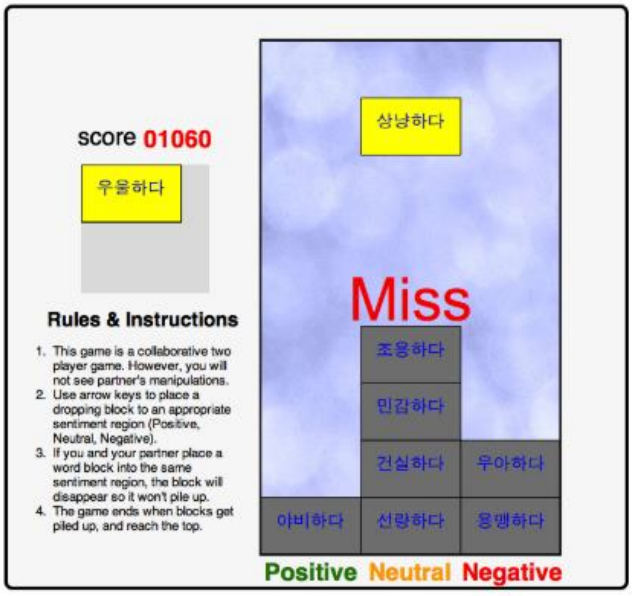
۳-۴ روش‌های تولید واژه‌نامه

ازآنجایی‌که ساختن یک واژه‌نامه‌ی احساسی کار بسیار فشرده و زمان‌بری است[10]، بسیاری از تحقیقات صورت پذیرفته بر روی واژه‌نامه‌های احساسی بر روی نحوه‌ی ساختن واژه‌نامه و تطبیق آن با حوزه‌های مختلف صورت پذیرفته است. در[88] چهار روش برای ساخت واژه‌نامه‌ی احساسی معرفی شده است که عبارت‌اند از:

- ساخت دستی شامل جمع سپاری[1] [10,107] و بازی‌سازی[2][108,109].
- Bootstrapping از مجموعه‌ای از کلمات بذر[3] و همچنین با استفاده از واژه‌نامه، استفاده از روابط نحوی یا روابط وابسته[110,111].
- تطبیق یک واژه‌نامه از حوزه‌ی دیگر با استفاده از نوعی از انتقال یادگیری[112,113].
- یادگیری ماشین یا یادگیری آماری [114,115]، بر مبنای کدگذاری انسانی احساسات یا امتیازدهی ستاره‌ای قطعات بزرگ متن مانند جملات، دیدگاه‌ها یا پُست‌های رسانه‌های اجتماعی.

۳-۴-۱ روش‌های دستی

یک واژه‌نامه‌ی احساسی فهرستی از ویژگی‌های لغوی (به‌عنوان‌مثال کلمات) است که به‌طورکلی بر اساس جهت‌گیری معنایی آن‌ها به‌صورت مثبت یا منفی برچسب‌گذاری می‌شود[116]. هرچند که ایجاد دستی چنین لیست‌هایی از ویژگی‌های حاوی عقیده یکی از قدرتمندترین روش‌های برای تولید واژه‌نامه‌های احساسی قابل‌اطمینان محسوب می‌شود، یکی از زمان‌برترین و پرهزینه‌ترین روش‌ها نیز است[10].



شکل ۴. نمایی از بازی Tower of Babel

[1] Crowdsourcing
[2] Gamification
[3] Seed Word

برخی از محققان [10,107] برای افزایش سرعت ساخت دستی واژه‌نامه از روش جمع‌سپاری استفاده نموده‌اند. در این روش با استفاده از پلتفرم‌هایی از قبیل Amazon's Mechanical Turk[1] یا CrowdFlower تعدادی از افراد از طریق اینترنت برای حل مسئله استخدام می‌شوند. یکی دیگر از روش‌های تولید واژه‌نامه که سرگرم کننده‌تر و دقیق‌تر نیز می‌باشد بازی‌سازی[2] است این روش غالباً با استفاده از بازی‌های چند نفره صورت می‌پذیرد. برخی از محققان[108,109] با طراحی بازی‌هایی که غالباً چند نفره می‌باشند سعی در امتیازدهی به کلمات احساسی داشتند. هنگ و همکاران در[109] یک بازی دونفره با عنوان "Tower of Babel" طراحی نمودند که در آن بازیکنان با تغییر موقعیت بلوک‌های درحال سقوط، قطبیت احساسی نسبت می‌دهند آن‌ها دریافتند که استفاده ازاین‌روش به‌جای روش کدگذاری دستی سریع‌تر است. واژه‌نامه‌هایی از قبیل GI، WordNet، ANEW، LIWC و VADER به روش دستی[3] و با مشارکت انسانی تولید شده‌اند.

۳-۴-۲ Bootstrapping یک واژه‌نامه‌ی احساسی با استفاده از یک مجموعه از کلمات بذر

در این روش با استفاده از یک مجموعه از کلمات بذر و همچنین باتوجه‌به روابط معنایی مابین کلمات بذر و کلمات دیگر موجود در یک منبع واژگانی (به‌عنوان‌مثال WordNet یا فرهنگ جامع)، یا روابط نحوی یا وابستگی با کلمات دیگر درون یک پیکره، یک واژه‌نامه‌ی احساسی تولید می‌گردد. محققان از روابط معنایی مابین کلمات بذر و دیگر کلمات درون WordNet برای انتشار مقادیر احساسات استفاده نموده‌اند. استفاده از وابستگی‌های لغت[4] بر مبنای مجاورت (یا همپوشانی در یک اندازه پنجره خاص) با یک کلمه بذر یا بر مبنای روابط نحوی با کلمات بذر در داخل یک پیکره، یکی دیگر ازاین‌گونه روش‌ها است[88].

[1] https://www.mturk.com/
[2] Gamification
[3] Manual
[4] Word Associations

۳-۴-۳ تطبیق یک واژه‌نامه از حوزه‌ی دیگر با استفاده از نوعی از انتقال یادگیری

تجزیه‌وتحلیل احساسات یک مسئله وابسته به حوزه و زبان است. قطبیت یک کلمه باتوجه‌به حوزه مورداستفاده در آن می‌تواند تغییر نماید. دلیل این امر وابستگی بسیاری از حوزه‌ها به کلمات فنی و اصطلاحات خاص مورداستفاده در آن حوزه است[115]. به همین دلیل عملکرد الگوریتم‌های تجزیه‌وتحلیل احساسات در هنگام استفاده از واژگان احساسی عمومی یا استفاده از واژگان احساسی نامربوط به حوزه مورداستفاده کاهش می‌یابد. تطبیق دامنه فرایندی ضروری برای تبدیل یک واژه‌نامه احساسی عمومی به یک واژه‌نامه احساسی خاص دامنه است. به‌طورکلی دوروش مختلف برای تطبیق دامنه وجود دارد. روش نخست مبتنی بر انتقال یادگیری و روش دوم مبتنی بر دستکاری واژه‌نامه احساسی است. روش‌های مبتنی بر انتقال یادگیری سعی بر یادگیری یک ویژگی نهفته مشترک در حوزه‌های مختلف دارند. این مدلها بر این فرض استوار هستند که ویژگی‌های احساسی دارای یک توزیع جهانی بر روی حوزه‌های مختلف هستند. ازاین‌رو در برخی از مواقع که شکاف وسیعی بین دامنه‌ها وجود دارد نتایج حاصل از این مدل تطبیق، عملکرد بدتری نیز از خود نشان می‌دهند که به انتقال منفی معروف هستند[117].

۴-۴-۳ روش‌های مبتنی بر یادگیری ماشین

در روش‌های مبتنی بر یادگیری ماشین معمولاً محتوای تولید شده توسط کاربر با استفاده از طبقه‌بندی‌کننده‌های عمومی نظیر نایوبیز، ماکزیمم بی‌نظمی و ماشین بردار پشتیبان به دو کلاس مثبت و منفی طبقه‌بندی می‌گردند. طبقه‌بندی‌کننده‌ها برای فاز آموزش و تست نیازمند پایگاه‌های داده‌ی برچسب-گذاری شده‌ی بزرگی هستند. از سوی دیگر عملکرد این روش‌ها وابسته به مجموعه آموزشی مورداستفاده است. همچنین این روش‌ها از لحاظ پردازش، حافظه و زمان آموزش/ طبقه‌بندی پرهزینه هستند. از سوی دیگر قابلیت تعمیم و تفسیر این روش‌ها و همچنین قابلیت اصلاح و گسترش آنها دشوارتر است[10].

۳-۵ معرفی برخی از واژه‌نامه‌های موجود

واژه‌نامه‌های احساسی را ازلحاظ نحوه‌ی تعریف ویژگی‌های عبارات موجود به سه دسته‌ی زیر می‌توان تقسیم‌بندی نمود:

- واژه‌نامه‌های با جهت‌گیری معنایی مبتنی بر قطبیت لغات
- واژه‌نامه‌های با ویژگی‌های مبتنی بر شدت احساسات
- واژه‌نامه‌های آگاه از زمینه

در این بخش برخی از واژه‌نامه‌های معروف و پرکاربرد موجود در حوزه‌ی تجزیه‌وتحلیل احساسات را در غالب دسته‌بندی فوق به‌صورت مختصر معرفی می‌نماییم.

۳-۵-۱ واژه‌نامه‌هایی با جهت‌گیری معنایی «مبتنی بر قطبیت لغات»

۳-۵-۱-۱ واژه‌نامه‌ی LIWC[1]

این واژه‌نامه[84] برای مطالعه اجزای هیجانی، شناختی، ساختاری در نمونه‌های مختلف متون طراحی شده است. ایده‌ی اولیه‌ی طراحی این واژه‌نامه، شناسایی و گردآوری گروهی از کلمات دارای پایه احساسی و ابعاد شناختی است که غالباً در روان‌شناسی اجتماعی، بهداشتی و شخصیتی مورد بررسی قرار می‌گرفت. این واژه‌نامه شامل تقریباً ۴۵۰۰ کلمه و ریشه‌ی کلمه است که در ۱۰۰ دسته طبقه‌بندی شده‌اند به نحویکه یک کلمه در یک یا بیشتر از یک دسته قرارداده شده است[۲]. این دسته‌بندی به‌صورت آبشاری[۳] بوده و یک دسته دسته‌های دیگری را می‌تواند شامل شود[118].

در این واژه‌نامه از ریشه‌ی کلمات استفاده شده است و با تمامی کلمات مشتق شده از یک ریشه به یک نحو برخورد می‌گردد. انتخاب واژه‌های این واژه‌نامه در طی چند گام مختلف و در طی چند سال صورت پذیرفته

[1] Linguistic Inquiry And Word Count

[۲] به عنوان مثال کلمه Cried بخشی از پنج دسته : غم، احساسات منفی، عواطف سراسری، فعل و فعل زمان گذشته است.

[۳] به عنوان مثال تمام کلمات خشم جزء دسته احساسات منفی و کلمات احساسی سراسری نیز محسوب می‌شوند

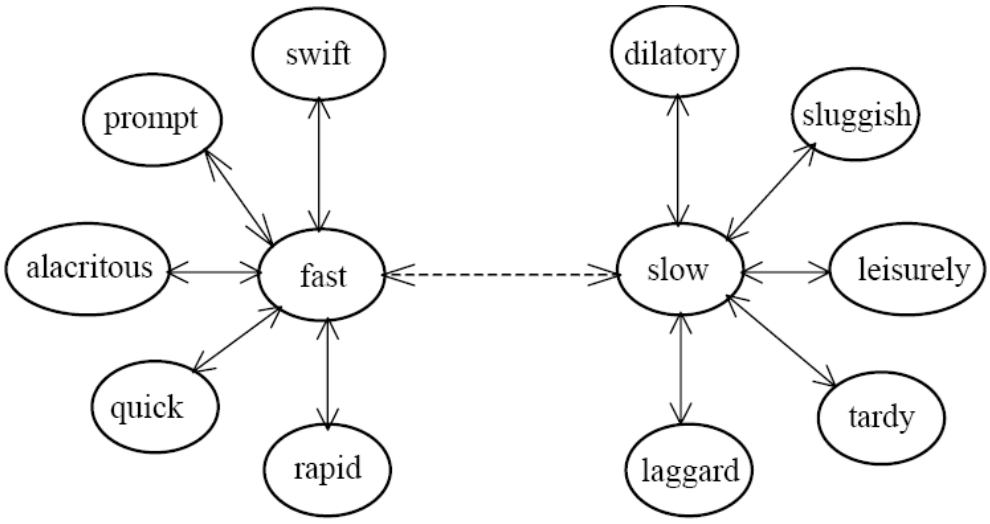
است. کلمات حداقل توسط سه شخص خبره بررسی و دسته‌بندی شده‌اند. ازآنجایی‌که این واژه‌نامه به طور وسیعی مورد تائید واقع شده است لذا برای جامعه‌شناسان، روان‌شناسان، زبان‌شناسان و دانشمندان کامپیوتر بسیار جذاب است[10]. LIWC از راهبرد شمارش واژه استفاده می‌نماید، بدین ترتیب که با تکیه‌بر واژه‌نامه‌ی داخلی خود که شامل بیش از ۱۰۰ گروه واژگانی بوده و قبلاً توسط داوران درجه‌بندی و دسته‌بندی شده است، در هر فایل متنی به جستجوی واژه به واژه و قراردادن هر واژه در زیرگروه خاص می‌نماید.

اگرچه LIWC2015 حاوی تعدادی از کلمات پرتکرار مورداستفاده در اغلب رسانه‌های اجتماعی و پیام‌های متنی و همچنین تعداد زیادی از شکلک‌های مبتنی بر متون نقطه‌گذاری همچون ":)" و ",(;" است اما از صورتک‌ها[1] و شکلک‌هایی‌که[2] به طور وسیعی در رسانه‌های اجتماعی مورداستفاده قرار می‌گیرند حمایت نمی‌کند[89]. همچنین این واژه‌نامه قادر به تشخیص شدت احساسات کلمات نیست[10] . از این واژه‌نامه در[119] برای پیش‌بینی انتخابات با استفاده از توییتر،[120] پیش‌بینی شروع افسردگی با استفاده از متن‌های موجود در رسانه‌های اجتماعی و[121] توصیف تغییرات هیجانی مادران باردار از طریق پست‌های توییتر استفاده شده است. این واژه‌نامه یک واژه‌نامه تجاری[3] بوده و گزینه‌های بهینه‌سازی از قبیل امکان استفاده از واژه‌نامه-های سفارشی‌سازی شده به‌جای واژه‌نامه‌ی استاندارد را برای کاربران فراهم می‌نماید.

۳-۵-۱-۲ واژه‌نامه‌ی Hu & Liu

این واژه‌نامه[122] حاوی ۶۷۹۰ کلمه بوده که در دودسته‌ی کلمات احساسی مثبت با ۲۰۰۶ کلمه و کلمات احساسی منفی با ۴۷۸۳ کلمه و بدون برچسب‌گذاری اجزای واژگانی کلامی دسته‌بندی شده است. این واژه‌نامه به‌صورت خودکار و با استفاده از فرایند bootstrapping با استفاده از واژه‌نامه‌ی wordNet و بهره‌گیری از تکنیک‌های یادگیری ماشین مبتنی بر نقدهای مشتری از حوزه‌های مختلف در طول چند سال گردآوری شده است.

[1] Emoji
[2] Emoticon
[3] Www.Liwc.Net



شکل ۵. ساختار دوبعدی صفات (مترادف متضاد) در WordNet

تولیدکنندگان این واژه‌نامه از یک مجموعه‌ی بسیار کوچک «تقریباً ۳۰ صفت» از صفات شناخته شده و مستقل از دامنه به‌عنوان مجموعه بذر استفاده نمودند. این مجموعه به‌صورت دستی به دو قطب مثبت و منفی برچسب‌گذاری گردیده، تولیدکنندگان این واژه‌نامه با استفاده از مجموعه Synsetهای موجود در wordNet و بهره‌گیری از ویژگی مترادفی/متضادی اقدام به توسعه‌ی این مجموعه‌ی کوچک و گسترش این مجموعه لغات نمودند. این واژه‌نامه قادر به اقتباس احساسات از شکلک‌ها و مخفف‌های اختصاری نیست[10]. باتوجه‌به نتایج به‌دست‌آمده، برخی از نویسندگان[88] استفاده از این واژه‌نامه را در حوزه‌ی نقد محصولات پیشنهاد کرده‌اند.

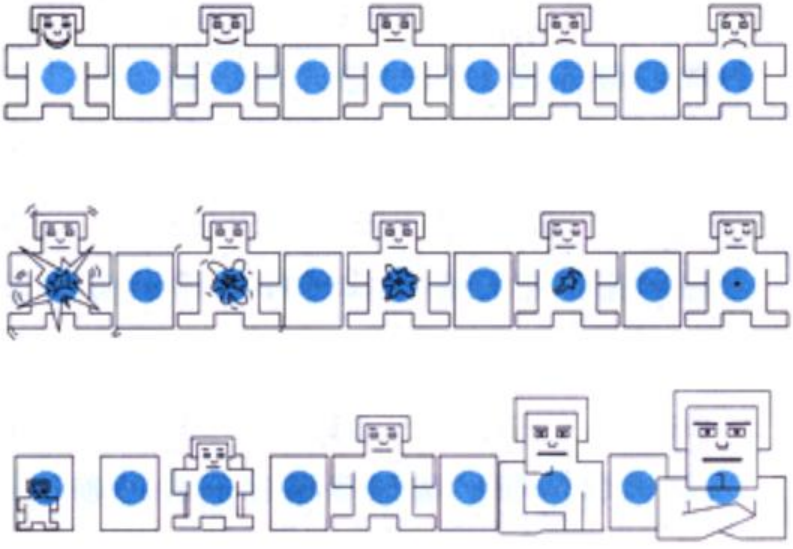
۳-۵-۲ واژه‌نامه‌های با شدت احساسات «valance based»

۳-۵-۲-۱ واژه‌نامه‌ی ANEW[1]

این واژه‌نامه[123] از ۱۰۴۰ کلمه انگلیسی تشکیل شده است. هدف از طراحی این واژه‌نامه فراهم نمودن مجموعه‌ای از منابع کلامی رتبه‌بندی شده بر مبنای معیارهای خوش‌آیندی، انگیختگی و تسلط برای تکمیل

[1] Affective Norms For English Words

سیستم بین‌المللی تصاویر عاطفی IAPS[1] و سیستم بین‌المللی صداهای عاطفی IADS[2] که به ترتیب مجموعه‌ای از محرک‌های تصویری و صوتی هستند، بود. طراحان این واژه‌نامه برای ارزیابی سه بعد خوشایندی، انگیختگی و تسلط از یک سیستم رتبه‌بندی عاطفی تحت عنوان آزمون خودارزیابی تصاویر آدمک SAM[3] استفاده نمودند.SAM سه نوع احساسات مختلف را نشان می‌داد: خوشحال در مقابل ناراضی، هیجان در مقابل آرام و کنترل شده در مقابل بدون کنترل و از کاربران خواسته شد در مدت بسیار کوتاهی بدون آنکه زمان زیادی را صرف فکر کردن در مورد هر کلمه نمایند، بر اساس اولین واکنش اقدام به علامت‌گذاری پاسخ‌ها نمایند.



شکل ۶. اشکال گرافیکی آزمون SAM برای سه بعد خوشایندی (اشکال بالا) برانگیختگی (اشکال وسط) و تسلط (اشکال پایین)

در این واژه‌نامه به‌جای تقسیم‌بندی ساده و دوسطحی با قطبیت مثبت و منفی، کلمات به‌صورت دستی بر مبنای مدل PDA[4] به لحاظ خوشایندی، برانگیختگی و تسلط و از لحاظ شدت احساسات بین ۱ تا ۹ (نقطه مرزی خنثی ۵) درجه‌بندی شده‌اند. ازآنجایی‌که این دیکشنری پیش از ظهور میکرو بلاگ‌ها تولید شده است عدم وجود کلمات عامیانه یکی از نقاط ضعف این دیکشنری محسوب می‌شود.

[1] International Affective Picture System
[2] International Affective Digitized Sounds
[3] Self-Assessment Manikin
[4] مدل حالت پاد(Pad Emotional State Model) ، که در توصیف و مشخص کردن حالت‌های هیجانی کاربرد دارد، از سه بعد عددی برای نمایش تمامی هیجان‌ها استفاده می‌کند .سه بعد پاد عبارتند از لذت، برانگیختگی و سلطه .(Dominance) درجات لذت، برانگیختگی، و سلطه، به ترتیب، میزان خوشایندی، شدت و ذات کنترلی هیجان‌ها را مشخص می‌کنند.

Description	Word No.	Valence Mean(SD)	Arousal Mean(SD)	Dominance Mean (SD)	Word Frequency
excellence	151	8.30 (0.82)	4.70 (2.26)	7.40 (2.17)	15
excitement	152	6.87 (2.36)	6.93 (2.43)	5.60 (1.84)	32
excuse	153	3.82 (1.88)	4.88 (2.64)	3.71 (2.31)	27
execution	154	3.27 (2.46)	5.53 (2.77)	4.27 (2.22)	15
exercise	155	6.93 (1.91)	7.13 (2.20)	6.07 (2.74)	58
fabric	742	5.21 (1.44)	3.95 (1.82)	4.79 (1.44)	15
face	556	6.12 (1.76)	4.76 (2.37)	5.28 (1.51)	371
failure	156	2.27 (1.22)	4.00 (2.54)	2.67 (1.95)	89
fall	743	3.86 (2.22)	4.81 (2.54)	4.29 (2.15)	147
FALSE	744	3.24 (1.58)	3.27 (2.23)	4.38 (1.66)	29

شکل ۷. نمونه ای از امتیازدهی کلمات در دیکشنری ANEW

۳-۵-۲-۲ واژه‌نامه‌ی Senti-WordNet

این واژه‌نامه که توسعه‌یافته‌ی واژه‌نامه‌ی WordNet محسوب می‌شود از ۱۴۷۳۰۶ عدد Synset تشکیل شده است. این دیکشنری با این فرض که "یک اصطلاح می‌تواند ویژگی‌های مختلف احساسی باتوجه‌به مفاهیم مختلف داشته باشد " طراحی شده است. این واژه‌نامه صفات، افعال، اسامی و سایر گروه‌های گرامری را به مجموعه‌هایی از مترادف‌ها با نام Synset دسته‌بندی می‌کند و به هر Synset سه امتیاز عددی شامل میزان مثبت بودن، میزان منفی بودن و میزان خنثی بودن اختصاص‌داده‌شده است[124]. محدوده‌ی این امتیازات مابین ۰ تا ۱ بوده و مجموع این سه امتیاز برای هر Synset برابر با یک است. این امتیازات با بهره‌گیری از ترکیب پیچیده‌ای از الگوریتم‌های نیمه نظارتی محاسبه شده است[10].

نسخه‌ی ۱ این واژه‌نامه با استفاده از یک الگوریتم یادگیر شبه ناظر در چهار مرحله تشکیل شده است که به طور مختصر در ادامه توصیف شده است:

- **گام اول توسعه مجموعه بذر:** در این مرحله پس از برچسب‌گذاری دستی تعداد محدودی از اصطلاحات دو مجموعه‌ی کوچک از اصطلاحات بذر تولید شده و باتوجه‌به روابط معنایی (هم‌معنی و متضاد) موجود در شبکه‌ی واژگان بر روی کلمات و گروه‌های هم‌معنی مرتبط بسط داده می‌شود.
- **گام دوم آموزش طبقه‌بند**: در این مرحله از دو مجموعه از synsetهای تولید شده از مرحله قبل و مجموعه‌ای از گروه‌های هم‌معنی فاقد بارحسی که در این مرحله به‌صورت دستی برچسب‌گذاری گردیده است به‌عنوان داده‌های آموزشی، برای استفاده در یادگیری با استفاده از یک روش دسته‌بندی نیمه نظارتی، در یک دسته‌بند سه‌تایی[1] استفاده می‌شود.
- **گام سوم طبقه‌بندی synsetها:** در این گام با استفاده از الگوریتم دسته‌بندی بقیه گروه‌های هم‌معنی برچسب‌گذاری حسی می‌شوند. در این مرحله برای کاهش خطای الگوریتم‌های دسته‌بندی، گروه‌های هم‌معنی که در مرحله دو آماده شده بودند برای آموزش تعدادی از الگوریتم‌های دسته‌بندی مورداستفاده قرار می‌گیرند. سپس نتایج آن‌ها با یکدیگر ترکیب می‌شوند. آموزش با استفاده از تعاریف و مثال‌های موجود در هر گروه synsetها صورت می‌پذیرد؛ و هرکدام از آن‌ها با مجموعه‌ای متفاوت از داده‌ها و ابزارها تحت آموزش قرار می‌گیرند
- **گام چهارم ترکیب طبقه‌بندها:** در این گام میانگین بار حسی به‌دست‌آمده از الگوریتم‌های مختلف دسته‌بندی محاسبه و به عددی در بازه‌ی صفر تا یک نگاشت می‌شود.

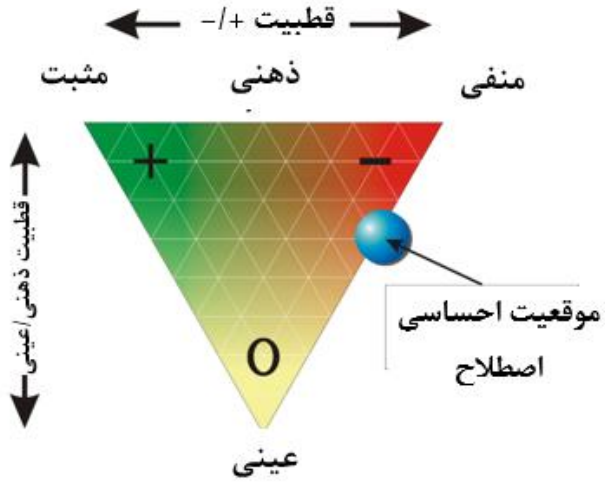
در فرایند امتیازدهی به هر Synset از یک مجموعه از هشت طبقه‌بند سه‌تایی استفاده شده است که می‌توانند مثبت بودن یا منفی بودن و یا خنثی بودن هر Synset را مشخص نماید. هرکدام از این هشت طبقه‌بند، از الگوریتم‌های یادگیری مختلف و داده‌های آموزشی متفاوتی استفاده نموده‌اند. همچنین آموزش آن‌ها نیز با استفاده از یک مجموعه آموزشی متفاوت صورت پذیرفته است . چنانچه تمامی طبقه‌بندها برچسب

[1] یک طبقه‌بند n-ary طبقه‌بندی است که به هر شئی دقیقا یکی از N برچسب از پیش تعریف شده را نسبت می‌دهد.

یکسانی را به یک Synset نسبت دهند، این برچسب حداکثر امتیاز را برای آن Synset خواهد داشت، در غیر این صورت هر برچسب دارای نمره‌ی متناسب با تعداد طبقه‌بندهایی است که به آن اختصاص‌یافته است. در این واژه‌نامه یک مدخل به‌صورت زیر نمایش داده می‌شود:

"<n, ۷۴۷۸۳۱۸, ۰, ۰,۷۵, ۰,۲۵, Syncope\#۱ swoon\#۱, faint\#۱, deliquim\#۱, a spontaneous loss of consciousness caused by insufficient blood to the brain>".

این مدخل اسم را با امتیاز ۰، ۰,۷۵ و ۰,۲۵ به ترتیب برای قطبیت‌های مثبت، منفی و خنثی نشان می‌دهد و شامل لغات "سنکوپ"، "غش‌وضعف" می‌باشد همچنین مثال ارائه شده برای این مجموعه مترادف "ازدست‌دادن خودبه‌خود آگاهی ناشی از ناکافی بودن خون به مغز" بیان شده است. شکل ۸ مدل گرافیکی این واژه‌نامه را برای نمایش نمرات حاصل از یک مجموعه Synset نشان می‌دهد.



شکل ۸. مدل گرافیکی sentiWordNet برای یک Synset

در نسخه‌ی سوم این واژه‌نامه نتایج نسخه‌ی قبل با استفاده از الگوریتم تکراری گام تصادفی بر روی گراف شبکه واژگان PWN اصلاح گردیده است و قابل‌استفاده برای عموم می‌باشد[1]. در این واژه‌نامه تعداد کثیری از Synsetها فاقد قطبیت مثبت یا منفی هستند و همچنین این واژه‌نامه فاقد ویژگی‌های مرتبط با متن مخصوص میکرو بلاگ‌ها است[125].

[1] http://sentiwordnet.isti.cnr.it/

۳-۲-۵-۳ واژه‌نامه‌ی VADER[1]

این واژه‌نامه[10] حاوی ۷۵۱۶ کلمه‌ی احساسی است که بین ۴+ و ۴- امتیازبندی شده‌اند. این واژه‌نامه با استفاده از لغات موجود در واژه‌نامه‌های LIWC، ANEW و GI و همچنین لیست کامل شکلک‌ها[2]، مخففها و کلمات ابتکاری[3] و اصطلاحات عامیانه احساسی ساخته شده است. در مرحله‌ی نخست از پیاده‌سازی این واژه‌نامه ۹۰۰۰ کلمه استخراج شده و با استفاده از روش جمع‌سپاری و از طریق Amazon Mechanical Turk (AMT) و بهره‌گیری از ده شخص مستقل در مقیاس ۴+ تا ۴- امتیازدهی گردید. سپس کلماتی که دارای امتیاز غیر صفر بوده و میزان انحراف استاندارد آن‌ها کمتر از عدد ۲,۵ محاسبه شده بود نگه داشته شده و بقیه کلمات حذف گردیدند به‌این‌ترتیب تنها حدود ۷۵۰۰ کلمه که هم قطبیت و هم شدت آن در مقیاس ۴+ تا ۴- محاسبه گردیده بود استخراج گردید.

جدول ۱. اثر مساوی تشدیدکننده‌ها بر روی کلمات احساسی در دیکشنری VADER

Sentiment phrase	Score
Good	0.4404
deeply good	0.4927
very good	0.4927
deeply deeply good	0.5379
very deeply good	0.5379

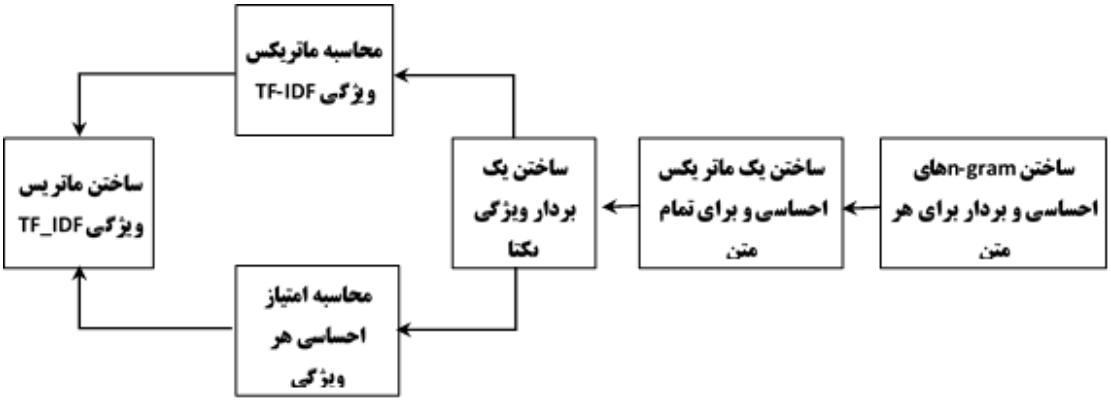
این واژه‌نامه یک واژه‌نامه‌ی عمومی بوده و از وب‌سایت[4] نویسندگان مقاله قابل‌دسترس است. این واژه‌نامه علاوه بر ۷۵۰۰ لغت احساسی فوق شامل برخی از مهم‌ترین تشدیدکننده‌های مورداستفاده در میکرو بلاگ‌ها که در محاسبه‌ی شدت احساسات یک عبارت احساسی مؤثر هستند نیز می‌باشد. همچنین در الگوریتم محاسبه شدت احساساتی که توسط تهیه‌کنندگان این واژه‌نامه ارائه گردیده پنج قانون نحوی و دستوری برای محاسبه‌ی

[1] Valence Aware Dictionary For Sentiment Reasoning
[2] http://en.wikipedia.org/wiki/list_of_emoticons#western
[3] http://en.wikipedia.org/wiki/list_of_acronyms
[4] http://comp.social.gatech.edu/papers/

دقیق‌تر قطبیت جملات معرفی و مورداستفاده قرار گرفته است. این پنج قانون به‌طورکلی به اثر نقطه‌گذاری‌ها[1]، بزرگ‌نویسی‌ها[2]، تأثیر تشدیدکننده‌ها در جمله، تأثیر کلمه But در جمله و همچنین اثر معکوس‌سازی منفی‌کننده‌ی Not در Trigramها می‌پردازد. این دیکشنری تنها از ۶۵ تشدیدکننده و تضعیف‌کننده پشتیبانی می‌نماید و به‌تمامی آنها امتیاز ثابت ±0.293 را تخصیص می‌دهد. جدول ۱ چند نمونه از امتیازات تخصیص‌داده‌شده توسط الگوریتم VADER را نشان می‌دهد.

۴-۲-۵-۳ واژه‌نامه‌ی Lexical TF-IDF

این واژه‌نامه[79] از یک رویکرد جدید به نام Lexical TFIDF برای ایجاد بردار ویژگی‌ها استفاده نموده است.



شکل ۹ .فلوچارت مراحل ایجاد واژه‌نامه TF_IDF

برخلاف روش‌های مرسوم که تنها از نرخ TF-IDF برای نمایش بردارهای ویژگی Unigram یا Ngram استفاده می‌نمایند و هیچ محدودیتی در ساخت کلمات Ngram وجود ندارد، در این روش در مرحله نخست و پس از حذف ایست‌واژه‌ها، نقطه‌گذاری‌ها، هایپرلینک‌ها، سمبل‌ها و اعداد از متن، Ngramهای احساسی با این شرط که یک یا دو تشدیدکننده یا منفی‌کننده پیش از یک کلمه احساسی Unigram ظاهر گردد، با جمع‌آوری از نقدها و استفاده از واژه‌نامه‌ی احساسی و همچنین بهره‌گیری از تقویت‌کننده‌ها و منفی‌کننده‌های موجود در [106] ساخته شده‌اند.

[1] Punctuation
[2] Capitalization

به‌این‌ترتیب برادر متناظر برای هر متن تولید شده و با استفاده از این بردارها ماتریس احساسی یکپارچه‌ای برای تمامی متن‌های مورداستفاده تولید می‌گردد. در گام بعد یک بردار یکتا از ماتریس مرحله‌ی قبل استخراج گردیده و پس از تولید و محاسبه‌ی ماتریس TF-IDF با استفاده از بردار فوق، امتیازات کلمات احساسی، تقویت‌کننده‌ها و منفی‌کننده‌ها از منابع مختلفی[10] استخراج شده و آنگاه با استخراج امتیاز کلمات Unigram از واژه‌نامه‌ی VADER و Ngramها از واژه‌نامه‌ی SO-CAL، امتیاز حاصل در رتبه‌ی TF-IDF آن ضرب می‌شود تا وزن مشخصی را تعیین نماید. سپس از نتایج حاصل در دو طبقه‌بندی کننده‌ی ماشین‌های بردار پشتیبان و روش حداکثر بی‌نظمی برای اعتبارسنجی میان‌دامنه‌ای استفاده گردیده که طبق ادعای نویسندگان آن بهبود قابل‌توجهی در دقت و سایر معیارهای عملکرد از خود نشان داده است.

۳-۵-۲-۵ واژه‌نامه‌ی SCL-OPP[1]

این واژه‌نامه[73] با بهره‌گیری از ۱۱ میلیون توئیت جمع‌آوری شده از توییتر که حاوی ":)" یا ":(" بودند ساخته شده است. برای تولید این واژه‌نامه تمامی Bigramها و Trigramهای موجود در پیکره مرجع که حداقل دارای یک کلمه مثبت و یک کلمه منفی بوده‌اند (بعنوان مثال عبارت Happy accident یا best winter break) استخراج گردیده و سپس به‌صورت دستی و با استفاده از روش بهترین –بدترین مقیاس BWS[2] قطبیت کلمات Ngram با استفاده از واژه‌نامه‌های همچون[122] Hu and Liu، [68] MPQA، [126] NRC و واژه‌نامه‌ی خودکار مبتنی بر توییتر[6] NRC[3] بررسی گردیده[4].

پس از استخراج عبارات Bigram و Trigram، امتیاز عبارات استخراج شده با استفاده از روش خودکار مبتنی بر PMI ارائه شده در[6] تخمین زده شد. در مرحله بعد محدوده‌ی کامل مقادیر احساسی به پنج بخش تقسیم شده و از هر بخش- به‌جز بخش میانی که فقط در حدود ۵۰ عدد Bigram و Trigram از آن استخراج گردید-

[1] Sentiment Composition Lexicon For Opposing Polarity Phrases
[2] Best–Worst Scaling
[3] Nrc's Automatically Generated Twitter-Specific Lexicon
[4] کلماتیکه در دو واژه‌نامه دارای قطبیت مختلف بودند حذف شده‌اند

در حدود ۱۰۰ عبارت احساسی Bigram و ۱۰۰ عدد عبارت احساسی Trigram انتخاب گردید. علاوه بر Bigram و Trigramهای فوق تمامی Unigramهای ظاهرشده در مجموعه‌ی انتخابی بالا نیز استخراج و انتخاب گردید. در این مرحله مجموعاً ۱۶۶۱ کلمه و اصطلاح استخراج گردید که ۸۵۱ اصطلاح آن Bigram و Trigram و ۸۱۰ عدد آن Unigram بودند. سپس با استفاده از عبارات استخراج شده‌ی فوق تعدادی سؤال چهارگزینه‌ای تولید و با استفاده از پلتفرم جمع‌سپاری CrowdFlower[1] از کاربران خواسته شد تا با استفاده از روش BSW، مثبت‌ترین و یا منفی‌ترین عبارت را از بین چهار گزینه‌ی ارائه شده انتخاب نمایند.

هرکدام از این چهار گزینه‌ای‌ها حداقل توسط هشت نفر که زبان مادری آن‌ها انگلیسی بوده و در ایالات متحده زندگی می‌کردند و همچنین در آزمون اولیه مورد پذیرش قرار گرفته بودند ارائه شده. آنگاه با استفاده از یک روش شمارش ساده [127,128] [2] این پاسخ‌ها به امتیاز عددی تبدیل گردید. درنهایت پس از پالایش مجدد کلمات و حذف عبارات Bigram و Trigramایکه با کلمات Unigram خود هم امتیاز بودند، ۱۱۷۸ عبارت احساسی به‌صورت دستی در دو سطح عبارات با قطبیت مثبت و عبارات با قطبیت منفی استخراج گردیده‌اند. در این واژه‌نامه یک مدخل به‌صورت زیر نمایش داده می‌شود:

happy accident ۰.۵۹۴ A+N ۱۴

که ستون اول عبارت احساسی و ستون دوم امتیاز تخصیص‌داده‌شده به عبارت که عدی بین ۱ و ۱- را در بر می‌گیرد را شامل می‌شود. در ستون سوم برچسب واژگان کلامی کلمات تشکیل‌دهنده عبارت و در نهایت ستون آخر تعداد تکرار عبارت موردنظر در پیکره حاوی ۱۱ میلیون توییت نشان‌داده‌شده است. همان‌طور که در جدول ۲ نشان‌داده‌شده است، این واژه‌نامه شامل ۶۰۲ عدد Unigram، ۳۱۱ عدد Bigram و ۲۵۶ عدد Trigram است که به دودسته مثبت با ۵۳۱ عبارت و منفی با ۶۴۷ عبارت دسته‌بندی شده است[3].

[1] https://www.figure-eight.com/

[2] درصد زمانهاییکه عبارت به عنوان مثبت‌ترین اصطلاح انتخاب می‌شد منهای درصد زمانهاییکه به عنوان منفی‌ترین اصطلاح انتخاب شده بود.

[3] http://www.saifmohammad.com/webpages/scl.html#opp

جدول ۲. تعداد یک گرمی‌ها، دو گرمی‌ها و سه گرمی‌ها در واژه‌نامه‌ی SCL-OPP

	مثبت	منفی	مجموع
یک گرمی	۲۹۲	۳۱۰	۶۰۲
دو گرمی	۱۳۵	۱۷۶	۳۱۱
سه گرمی	۱۰۴	۱۶۱	۲۶۵
مجموع	۵۳۱	۶۴۷	۱۱۷۸

۳-۵-۲-۶ واژه‌نامه‌ی SentiStrength

هسته‌ی این واژه‌نامه[1][13] از ۲۳۱۰ کلمه و ریشه‌ی کلمه تشکیل شده است و این کلمات از واژه‌نامه‌های LIWC، GI و برخی کلمات اضافه شده در طی آزمایش واژه‌نامه خصوصاً برای کلمات جدید مرتبط با کامپیوتر[2] استخراج گردیده است. ازآنجائی‌که یک متن می‌تواند هم شامل کلمات مثبت و هم کلمات منفی باشد نتیجه خروجی این واژه‌نامه برای هر متن شامل یک امتیاز مثبت مابین ۱ تا ۵ و یک امتیاز منفی مابین ۱- تا ۵- است که مقدار خروجی ۱،۱- نشان‌دهنده‌ی متن خنثی است. در این واژه‌نامه از ریشه‌ی کلمات استفاده شده است. به‌عنوان‌مثال تمامی کلماتی که با پیشوند Amaze شروع می‌گردند یکجا در نظر گرفته شده‌اند. این واژه‌نامه قادر به تشخیص شدت احساسات مثبت و منفی در متن‌های کوتاه و حتی، برای زبان غیررسمی است. به استثنای متون سیاسی دقت این واژه‌نامه برای متن‌های وب اجتماعی در زبان انگلیسی در حد دقت انسانی است[26]. علاوه بر این با تغییر ورودی‌ها امکان تنظیم این واژه‌نامه برای زبان‌های دیگر نیز فراهم شده است. ازجمله نقاط ضعف این واژه‌نامه عدم توجه به برچسب‌گذاری واژگان کلامی محسوب می‌شود[13].

این واژه‌نامه برای محاسبه‌ی بهتر نتایج، علاوه بر لیست کلمات احساسی از لیست کلمات منفی‌کننده، لیست کلمات تقویت‌کننده، لیست اصطلاحات و لیست شکلک‌ها نیز بهره‌برداری می‌نماید. در این واژه‌نامه نیز همانند واژه‌نامه‌ی SO-CAL تکرار علامت‌های نقطه‌گذاری همچون "!" و "؟" و تکرار برخی از حروف کلمات

[1] http://sentistrength.wlv.ac.uk
[2] Computer Mediated Communication

«حداقل دو بار» در امتیازدهی نهایی متن مؤثر است. به‌طورکلی دو نسخه از این واژه‌نامه وجود دارد. نسخه‌ی اولیه این واژه‌نامه تنها برای پیام‌های غیر رسمی دوستانه در MySpace مورد آزمایش قرارگرفته بود در حالی کهه نسخه جدید برای بررسی انواع متنوع‌تری از متون (BBC، Dig، Twitter، Runners World، MySpace، YouTube) طراحی شده است. نسخه‌ی اولیه از ۲۹۸ کلمه‌ی مثبت و ۴۶۵ کلمه منفی تشکیل شده بود و نسخه‌ی جدید از ۲۳۱۰ واژه پشتیبانی می‌نماید. ازآنجائی‌که مهم‌ترین نقطه‌ضعف نسخه‌ی اولیه در تشخیص قدرت احساسات منفی[76] بود لذا مهم‌ترین تغییر جدید، توسعه‌ی لیست لغات منفی واژه‌نامه با استفاده از کلمات منفی موجود در واژه‌نامه‌ی GI [85] است.

در این واژه‌نامه تمامی کلمات ابتدا به‌صورت دستی امتیازدهی شده‌اند و سپس در فاز آموزش، امتیازدهی آن‌ها به‌صورت خودکار تغییر پیدا نموده است. برخی از کلمات استفاده شده در این واژه‌نامه کلمات استاندارد انگلیسی بوده و برخی دیگر از آن‌ها کلمات بدون استانداردی است که به‌صورت متداول در MySpace (به‌عنوان‌مثال luv, xox, lol, haha, muah) استفاده می‌گردید. الگوریتم آموزش اقدام به بهینه‌سازی شدت کلمات احساسی و همچنین قطبیت آن‌ها می‌نماید. در این روش تأثیر تغییر امتیاز کلمات احساسی بر روی افزایش دقت طبقه‌بندی در پیکره‌ی متنی کدگذاری شده توسط انسان مورد بررسی قرارگرفته و چنانچه باعث افزایش دقت طبقه‌بندی می‌گردید، این تغییر شدت امتیاز، بر روی واژه‌نامه اعمال می‌شد. این الگوریتم تازمانیکه تمامی کلمات موجود در واژه‌نامه مورد بررسی قرار نگرفته باشد و یا اعمال مجدد الگوریتم، باعث تغییری در امتیازات تولید شده نگردد تکرار می‌گردید. همچنین در این نسخه لیست اصطلاحات نیز با بهره-گیری از اصطلاحات جدید توسعه پیدا نموده و لیست منفی‌کننده‌ها نیز افزوده شده است. این نسخه جدید تحت عنوان SentiStrength 2 شناخته می‌شود. این واژه نامه تنها از ۲۸ کلمه تشدیدکننده که در سه گروه با امتیازات ۲ ، ۱ و ۱- امتیازدهی شده‌اند پشتیبانی می‌نماید.

۳-۵-۲-۷ واژه‌نامه‌ی DAL[1]

این واژه‌نامه[129,130] یک واژه‌نامه با برچسب‌گذاری دستی است که توسط دانشگاه لورنیته باهدف فراهم کردن منابع برای ارزیابی احساسات در زبان ساخته شده است. این واژه‌نامه بر مبنای اصول روان‌شناسی تهیه شده است و حاوی فهرستی از رایج‌ترین واژه‌های موجود در دهه ۱۹۶۰ است. کلمات استخراج شده از پیکره تقریباً توسط ۲۰۰ داوطلب جوان که بین ۲۰ تا ۲۵ سال داشته‌اند امتیازدهی و به طور دستی برچسب‌گذاری شده است. این واژه‌نامه از ۸۷۴۲ آیتم احساسی تشکیل شده است که به هر آیتم بر مبنای ابعاد احساسی آزگود دو مقدار دلپذیری[2] و فعالیت[3] در مقیاس بین ۱ تا ۳ نسبت‌داده‌شده است.

جدول ۳. چند نمونه از کلمات موجود در دیکشنری DAL و امتیازات آنها

کلمه	میزان تصویرسازی ذهنی	فعالیت	دلپذیری
Good	۱.۰	۱.۹۱۶۷	۲.۷۵۰۰
Pretty	۲.۶	۱.۶۱۵۴	۲.۸۴۶۲
Beautiful	۲.۸	۱.۶۴۲۹	۳.۰
Rough	۲.۲	۲.۰	۱.۶۲۵
Ugly	۲.۲	۱.۸۳۳۳	۱
Bad	۲.۲	۲.۱۱۱۱	۱.۲۸۵۷
Very	۱.۰	۱.۸۰۰۰	۱.۸۳۳۳
Not	۱.۰	۱.۸۷۵۰	۱.۱۲۵۰

همچنین امتیازی تحت عنوان تصویرسازی به یک کلمه تخصیص می‌دهد که میزاد سادگی کلمه برای تشکیل یک تصویر ذهنی را مشخص می‌نماید. در این واژه‌نامه اشکال مختلف گرامری از یک کلمه از لحاظ زمانی، تعداد و... متفاوت در نظر گرفته شده‌اند. جدول زیر چند نمونه از کلمات احساسی موجود به همراه امتیازات تخصیص‌داده‌شده به آنها نشان می‌دهد.

[1] Dictionary Of Affect In Language
[2] Pleasantness
[3] Activation

۳-۵-۲-۸ واژه‌نامه‌ی senticNet[1]

این واژه‌نامه[131] از واژه‌نامه SentiWordNet الهام گرفته شده، یک منبع معنایی و احساسی قابل دسترس عمومی برای تجزیه‌وتحلیل احساسات و نظرکاوی در سطح معنایی است[132]. هدف از SenticNet استنتاج قطبیت مفاهیم حسی از متن‌های نوشته شده به زبان طبیعی در یک سطح معنایی به‌جای سطح نحوی است. این واژه‌نامه از تکنیک‌های پردازش زبان طبیعی، هوش مصنوعی، وب‌معنایی، دانش اجتماعی و محاسبات عاطفی برای تعریف قطبیت مفاهیم حسی استفاده می‌نماید[133]. حاوی ۱۴۲۴۴ مفهوم از قبیل خشم، غم و تعجب بوده که یک مقدار عددی مابین ۱+ تا ۱- به آن نسبت‌داده‌شده است و مفاهیم خنثی و یا مفاهیمی که بسیار نزدیک به خنثی می‌باشند از آن حذف شده‌اند. این واژه‌نامه یک واژه‌نامه‌ی نیمه‌خودکار[2] بوده و با استفاده از محاسبات سنتیکت که از تکنیک‌های هوش مصنوعی و وب‌معنایی برای پردازش نظرات زبان طبیعی از طریق تکنیک‌های کاهش ابعاد و تحلیل گراف[2] بهره می‌برد ساخته شده است. این واژه‌نامه برای عموم قابل‌استفاده است.

۳-۵-۲-۹ واژه‌نامه‌ی Senti-N-gram

هسته‌ی اصلی این واژه‌نامه[2] از واژه‌نامه‌ی[10] VADER تشکیل شده است. این واژه‌نامه حاوی ۷۵۱۶ عدد Unigram، ۲۴۱۵ عدد Bigram و ۳۰۵عدد Trigramای است. این واژه‌نامه یک واژه‌نامه‌ی نیمه‌خودکار بوده که از ترکیب کلمات احساسی با تشدیدکننده‌ها و منفی‌کننده‌ها بر مبنای یک پیکره از نظرات امتیازدهی شده توسط کاربران از پایگاه‌داده آمازون تولید شده است. برای ساخت این واژه‌نامه نخست تمامی Unigramها، Bigramها و Trigramهای موجود در پیکره نظرات پایگاه‌داده آمازون که از ترکیبی از کلمات احساسی موجود در واژه‌نامه‌ی VADER و تشدیدکننده‌ها و منفی‌کننده‌ها ساخته شده‌اند، به همراه امتیاز داده شده به آن‌ها استخراج گردیده و پس از محاسبه درصد کاهش یا افزایش Bigram نسبت به Unigramهای استخراج شده و

[1] http://sentic.net/
[2] Graph-Mining

همین‌طور Trigram نسبت به Bigramهای استخراج شده مقدار این افزایش یا کاهش در امتیاز اصلی ضرب و امتیاز Bigram و Trigramهای جدید محاسبه گردیده است.

۳-۵-۳ واژه‌نامه‌های آگاه از زمینه «context- awareness»

چه از واژه‌نامه‌های دودویی مبتنی بر قطبیت استفاده کنیم و چه از واژه‌نامه‌های مبتنی بر شدت احساسات، امکان بهبود کارایی تجزیه‌وتحلیل احساسات به‌وسیله درک عمیق‌تر از ویژگی‌های لغوی (به‌عنوان‌مثال pos) و یا حوزه مورداستفاده برای آگاهی بیشتر از متن وجود دارد. به‌عنوان‌مثال، یک واژه‌نامه می‌تواند بر مبنای فرایند word-sense disambiguation[1] تنظیم گردد[134].

۱-۳-۵-۳ واژه‌نامه‌ی GI[2]

واژه‌نامه‌ی[85] GI در اواخر دهه ۱۹۶۰ توسط فلیپ استون و همکاران در هاروارد و به‌صورت دستی ساخته شده و باهدف کشف و ردیابی موضوعات روان‌شناسی تولید گردیده است. این دیکشنری یکی از قدیمی‌ترین دیکشنری‌هایی تولید شده به‌صورت دستی است که هنوز هم مورداستفاده قرار می‌گیرد. این واژه‌نامه حاوی ۸۶۴۱ کلمه احساسی است که برخی از کلمات جنبه‌های احساسی چندگانه دارند. مجموع کل کلمات احساسی این واژه‌نامه ۱۱۷۸۹ کلمه بوده که در ۱۸۲ دسته‌ی مرتبط با حوزه مورداستفاده، قطبیت، طبقات اجتماعی و غیره گروه‌بندی شده‌اند. به‌عنوان‌مثال کلمات علاوه بر دسته‌بندی دوقطبی مثبت و منفی در دسته‌بندی‌های احساسی از قبیل لذت، تحریک، احساس و درد طبقه‌بندی شده‌اند.۱۹۱۵ کلمه در این مجموعه دارای برچسب مثبت و ۲۲۹۱ کلمه برچسب منفی خورده‌اند. دویت و احمد در[135] شرح داده‌اند که این واژه‌نامه از لیست‌های کلمات پرتکرار از واژه‌نامه‌ی[136] Harvard IV و واژه‌نامه‌ی[137] Lasswell در دوره‌های زمانی مختلف ساخته شده است. عملیات اعتبارسنجی و برچسب‌گذاری این واژه‌نامه به‌صورت دستی و برطبق مجموعه وسیعی از مقوله‌های معنایی توسط نظریه‌های روان‌شناسی و تحلیل محتوایی انجام پذیرفته

[1] فرآیند شناسایی احساس کلمه درون جمله زمانیکه کلمه دارای چند معنی مختلف است

[2] General Inquirer

است. این واژه‌نامه شدت جهت‌گیری احساسی مابین کلمات احساسی را نادیده می‌گیرد[10]. همچنین این واژه‌نامه فاقد ویژگی‌های لغوی مرتبط با احساس متداول مورداستفاده در متون اجتماعی بوده و همچنین قادر به تشخیص میزان تفاوت شدت کلمات احساسی مختلف نیست[138].

۲-۳-۵-۳ واژه‌نامه‌ی SO-CAL[1]

این واژه‌نامه[139] یک واژه‌نامه‌ی ساخته شده به‌صورت دستی است که توسط افراد بومی[2] کدگذاری و توسط کمیته‌ای از سه محقق بازبینی شده است. منابع اصلی اصطلاحات این واژه‌نامه عبارت‌اند از:

- صفات از یک پیکره حاوی ۴۰۰ نقد از هشت طبقه شامل: کتاب‌ها، اتومبیل‌ها، کامپیوترها، وسایل آشپزی، هتل‌ها، فیلم‌ها، موسیقی و تلفن استخراج شده‌اند[140,141].
- یک زیرمجموعه حاوی ۱۰۰۰ نقد فیلم از پایگاه‌داده‌ی قطبی[71].
- لغات مثبت و منفی از[85] GI.

به هرکدام از اصطلاحات موجود در واژه‌نامه امتیازی در بازه‌ی ۵+ تا ۵- داده شده است و کلمات خنثی و دارای امتیاز صفر از این مجموعه حذف شده‌اند. نسخه‌ی قبلی این واژه‌نامه[140,141] برای تشخیص قطبیت یک متن از یک واژه‌نامه که تنها حاوی صفات بود استفاده می‌نمود. در نسخه‌ی ابتدایی از روش ساده جمع و میانگین‌گیری برای محاسبه‌ی متوسط امتیاز یک سند استفاده شده بود. به‌این‌ترتیب که امتیازات یکایک صفات موجود در یک سند با یکدیگر جمع شده و سپس بر تعداد کل صفات آن سند تقسیم می‌گردید در نسخه‌ی جدید، روش امتیازدهی بر مبنای روش‌های پیشرفته محاسباتی و با درنظرگرفتن منفی‌کننده‌ها، تشدیدکننده‌ها و irrealis moods و همچنین بن‌واژه‌سازی[3] افعال و اسامی صورت می‌پذیرد و زمان افعال و یا مفرد و جمع بودن اسامی و افعال در امتیازدهی بدون تأثیر است.

[1] Companion Semantic-Orientation Calculator
[2] Native
[3] Lemmatized

به‌استثنای تعداد معدودی از قیود که به‌صورت دستی اضافه یا تغییر داده شده‌اند، بقیه قیدهای تعریف شده در این واژه‌نامه به‌صورت خودکار و با استفاده از واژه‌نامه صفات تولید و ساخته شده‌اند. در هنگام مواجه با قیدهایی که در لیست قیدهای واژه‌نامه وجود نداشته باشد ابتدا عملیات ریشه‌یابی[1] صورت پذیرفته و سپس ریشه‌ی حاصل در لیست صفات واژه‌نامه مورد جستجو قرار می‌گیرد. در این واژه‌نامه به‌جای تغییر قطبیت، کلمات احساسی ترکیب شده با منفی‌کننده‌هایی از قبیل not، none، nobody، never و... از روش شیفت دادن استفاده شده است. به‌این‌ترتیب که امتیاز احساسی کلمات مثبت با مقدار ۴- و امتیاز احساسی کلمات منفی با ۴+ جمع می‌گردد. این واژه‌نامه موجب تقویت عبارات منفی[2] موجود در متون می‌گردد و شدت اصطلاحاتی را که به طور مکرر اتفاق می‌افتند را کاهش می‌دهد. این واژه‌نامه حاوی ۲۸۲۸ صفت، ۹۱۳ قید، ۱۵۴۹ اسم و ۱۱۴۲ فعل است همچنین این مجموعه ۵۶۳ اصطلاح چندکلمه‌ای و ۱۶۹ الگوی عبارات منظم را در برمی‌گیرد. همچنین این واژه‌نامه از ۲۱۶ تشدیدکننده نیز استفاده می‌نماید.

۳-۵-۳-۳ واژه‌نامه‌ی احساسی WKWSCI

این واژه‌نامه[89] به‌صورت دستی و توسط گروهی از دانشجویان مقطع کارشناسی دانشکده ارتباطات و اطلاعات Wee Kim Wee، دانشگاه فنّاوری Nanyang تولید شده است. این واژه‌نامه از ۲۹۷۱۸ کلمه، شامل ۳۱۲۱ کلمه‌ی مثبت، ۷۱۰۰ کلمه‌ی منفی و ۱۹۵۰۰ کلمه‌ی خنثی تشکیل شده است. فرایند تولید این واژه‌نامه از سال ۲۰۱۲ شروع شده و تا ۲۰۱۵ ادامه داشته است در هرسال از خدمات ۶ دانشجو برای امتیازدهی به کلمات استفاده گردیده است که بیشترین امتیاز را در آزمون زبان کسب نموده بودند. به کلمات احساسی موجود امتیازاتی مابین ۳ تا ۳- نسبت‌داده‌شده است و کلمات در سه گروه مثبت، منفی و خنثی گروه‌بندی شده‌اند. همچنین کلمات احساسی در چهار گروه صفات، اسامی، قیدها و افعال دسته‌بندی شده‌اند و بجز کلمات موجود در بخش اسامی تعدادی از کلمات در گروه‌های دیگر با هم مشترک هستند. کلماتی که در بخش اسامی با دیگر گروه‌ها مشترک بودند از مجموعه حذف شده‌اند.

[1] Stemming

[۲] امتیاز کلمات منفی را در عدد 1.5 ضرب می‌نماید.

- ۴۷۳ کلمه در هر دو گروه صفات و فعل مشترک هستند.
- ۳۷۴ کلمه در هر دو گروه صفات و قید مشترک هستند.
- ۸۳ کلمه در هر دو گروه قید و فعل قید مشترک هستند.
- ۶۵ کلمه در هر سه گروه صفات و قیدها و فعل مشترک هستند.

جدول ۴ تعداد کلمات موجود در این واژه‌نامه را به تفکیک امتیاز و بر مبنای برچسب‌گذاری واژگان کلامی نشان می‌دهد.

جدول ۴. تعداد کلمات در واژه‌نامه‌ی WKWSCI با انواع برچسب‌گذاری واژگان کلامی و مقدار شدت احساسات

قطبیت	مثبت			خنثی	منفی			جمع کل
شدت احساسات	۳	۲	۱	۰	-۱	-۲	-۳	
صفت	۶۰	۶۹۲	۷۵۰	۳۷۳۶	۱۴۰۰	۱۰۴۰	۳۴	۷۷۱۲
قید	۲۴	۳۲۶	۲۵۰	۱۱۵۴	۲۹۹	۴۴۶	۱۳	۲۵۱۴
فعل	۴	۶۵	۲۳۶	۴۵۳۱	۹۷۰	۴۱۱	۱۲	۶۲۲۹
اسم	۲۳	۴۵۵	۲۳۴	۱۰۰۷۶	۱۳۳۸	۱۱۲۷	۱۰	۱۳۲۶۳
جمع کل	۳۱۲۱			۱۹۴۷۹	۷۱۰۰			۲۹۷۱۸

۴-۳-۵-۳ واژه‌نامه‌ی MPQA[1] Subjectivity

این واژه‌نامه[2][68] از ۸۲۲۲ کلمه‌ی Unigram تشکیل شده است که شامل ۲۷۱۹ کلمه‌ی مثبت، ۴۹۱۴ کلمه‌ی منفی و ۵۱۹ کلمه‌ی خنثی می‌باشد. این کلمات شامل صفات، قیدها، افعال، اسامی و تمامی برچسب‌های کلامی موجود می‌باشند. این واژه‌نامه از منابع متنوعی که برخی به‌صورت دستی و برخی دیگر به‌صورت خودکار تولید شده‌اند، ساخته شده است. بیشتر موجودیت‌های این واژه‌نامه در یک پروژه به‌وسیله[142] Riloffو Wiebe ایجاد گردیده است. همچنین در این واژه‌نامه از کلمات ذهنی دارای قطبیت مثبت و منفی موجود در واژه‌نامه‌ی[85] GI نیز استفاده شده است. این کلمات به‌صورت دستی برچسب‌گذاری شده‌اند. کلماتی که در اکثر زمینه‌ها و در اغلب موارد ذهنی بودند به‌صورت *strongsubj* و کلماتی که تنها در

[1] Multi-Perspective Question Answering
[2] http://mpqa.cs.pitt.edu/lexicons/subj_lexicon/

برخی از زمینه‌های مورداستفاده معنای ذهنی داشتند با *weaksubj* برچسب‌گذاری گردیدند. قطبیت کلمات استخراج شده از واژه‌نامه‌ی *GI* حفظ گردید و بقیه لغات به‌صورت دستی به سه دسته‌ی مثبت، منفی یا خنثی دسته‌بندی گردید. در زیر نمونه‌ای از نحوه نمایش اطلاعات یک کلمه‌ی احساسی توسط این واژه‌نامه نشان‌داده‌شده است.

جدول ۵. بخشی از واژه‌نامه‌ی ذهنی MPQA

Strength	Length	Word	POS	Stemmed	Polarity
type=weaksubj	len=1	word1=abandoned	pos1=adj	stemmed1=n	priorpolarity=negative
type=weaksubj	len=1	word1=abandonment	pos1=noun	stemmed1=n	priorpolarity=negative
type=weaksubj	len=1	word1=abandon	pos1=verb	stemmed1=y	priorpolarity=negative
type=strongsubj	len=1	word1=abase	pos1=verb	stemmed1=y	priorpolarity=negative
type=strongsubj	len=1	word1=abasement	pos1=anypos	stemmed1=y	priorpolarity=negative
type=strongsubj	len=1	word1=abash	pos1=verb	stemmed1=y	priorpolarity=negative

۵-۳-۵-۳ واژه‌نامه‌ی NRC Word-Sentiment Association

موجودیت‌های این واژه‌نامه[72] از سه منبع مختلف استخراج شده است که عبارت‌اند از:

- ۲۰۰ Unigram و Bigram اول پرتکرار از مجموعه‌های هر برچسب کلامی شامل (صفات، قیدها، اسامی و افعال) از فرهنگ جامع[143] Macquarie[1] که این اصطلاحات پرتکرار با استفاده از پیکره Ngram گوگل[144] شناسایی شده بود.
- ۶۴۰ اصطلاح که از واژه‌نامه‌ی WordNet Affect استخراج شده است.
- ۸۰۰۰ اصطلاح گرفته شده از GI.

[1] http://www.macquarieonline.com.au/

هر اصطلاح در این واژه‌نامه در هشت گروه احساسی شامل خشم، ترس، انتظار، اعتماد، تعجب، غم، شادی و انزجار و در دو گروه کلی مثبت و منفی طبقه‌بندی شده‌اند[126]. کلماتی که بیش از یکبار در *Macquarie Thesaurus* تکرار شده بود از لیست کلمات پرتکرار حذف شده است[145].

جدول ۶. تعداد عناصر هرکدام از دسته‌بندی‌های موجود در واژه‌نامه‌ی NRC

	مثبت	منفی	خشم	ترس	انتظار	اعتماد	تعجب	غم	شادی	انزجار
تعداد کلمات	۲۳۱۲	۳۳۲۴	۱۲۴۷	۸۳۹	۱۰۵۸	۱۴۷۶	۶۸۹	۱۱۹۱	۵۳۴	۱۲۳۱

این دسته‌بندی با استفاده از پلتفرم Mechanical Turk آمازون صورت پذیرفته است و هر اصطلاح حداقل توسط سه نفر اعتبارسنجی شده است. تنها کسانی مجاز به امتیازدهی به کلمات احساسی بودند که نسبت به معنای کلمه آشنایی کامل داشته باشند و این مسئله با استفاده از یک سیستم سؤال و پاسخ خودکار مورد بررسی قرار گرفت. این واژه‌نامه به طور عمومی قابل‌دسترس می‌باشد[1]. نسخه‌ی نهایی این واژه‌نامه شامل ۱۴۱۸۳ اصطلاح احساسی است. جدول زیر تعداد هرکدام از گروه‌های احساسی موجود را به تفکیک نشان داده است. در این واژه‌نامه لغات احساسی به ۴۰ زبان مختلف ترجمه شده‌اند و مورداستفاده قرار می‌گیرند.

۶-۳-۵-۳ واژه‌نامه‌ی WAL[2]

این واژه‌نامه[146] از ۲۸۷۴ عدد Synset و ۴۷۸۷ عدد کلمه بر مبنای واژه‌نامه[3] WordNet[147] DOMAINSکه یک افزونه چندزبانه از WordNet بوده و توسط ITC-irst تولید شده است، ساخته شده است. در مرحله‌ی نخست با استفاده از واژه‌نامه یک مجموعه‌ی اولیه از کلمات احساسی که Affected نام‌گذاری شده بود به‌صورت دستی ایجاد گردید. این مجموعه اولیه شامل ۱۹۰۳ اصطلاح که به‌صورت مستقیم یا غیرمستقیم

[1] http://www.purl.org/net/emolex
[2] https://wndomains.fbk.eu/wnaffect.html
[3]. در Wordnet Domains هر کدام از Synsetها با حداقل یک برچسب دامنه (به عنوان مثال ورزش، تصاویر،پزشکی و... که از مجموعه‌ای از حدود ۲۰۰ برچسب مرتب شده به‌صورت سلسله مراتبی حاشیه نویسی شده است،

به حالت‌های ذهنی (مثلاً عاطفی) اشاره داشت است. جدول ۷ لیست هرکدام از این دامنه‌ها را به همراه تعداد عناصر هر دامنه نشان می‌دهد.

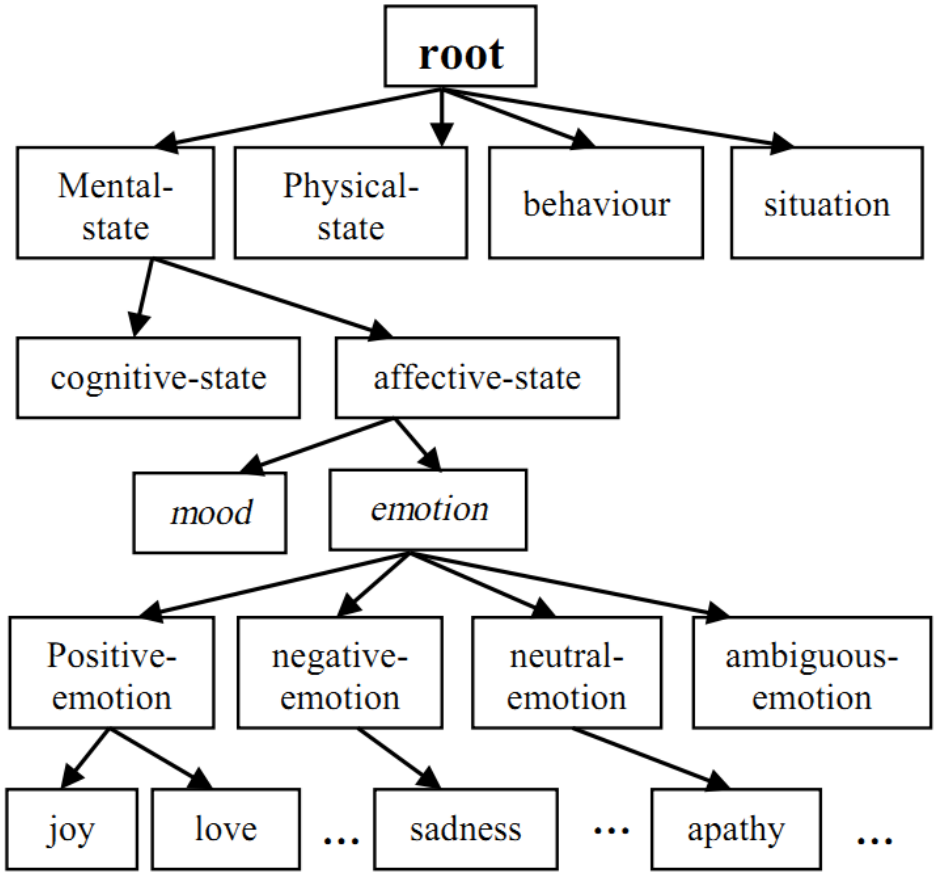
جدول ۷ . حوزه‌های احساسی واژه‌نامه‌یWAL

Domain	**Number**	**Examples**
Attitude (att)	۸۷۰	Intolerance, belligerent
Behaviour (beh)	۴۸۴	Approval, inhibited
Cognitive state (cog)	۶۸۵	Confusion, wistful
Edonic signal (eds)	۱۰۵	Carsick, gracious
Emotion (emo)	۲۰۴۵	Anger, fear
Mood (moo)	۷۱	Animosity, amiable
Physical state (phy)	۲۲۰	Depression, alive
Emotional response (res)	۵۵	Palpitation, livid
Sensation (sen)	۱۲۶	Pleasure, thirsty
Emotion-eliciting situation (sit)	۲۸۲	Quietude, vivacious
Trait (tra)	۱۵۹۸	Superiority, itchy

این مجموعه از ۵۳۹ عدد اسم و ۵۱۷ عدد صفت تشکیل شده است و تعداد کمی از افعال و تعداد بسیار اندکی از قیدها (۱۵ قید) را شامل می‌شود. برای هرکدام از آیتم‌ها قالبی جهت افزودن اطلاعات لغوی همچون میزان همبستگی بین اصطلاحات انگلیسی و ایتالیایی، برچسب‌گذاری نحوی، تعاریف، مترادف‌ها و متضادها ایجاد گردید.

با نگاشت احساسی اصطلاحات درون مجموعه AFFECT به Synsetهای موجود در واژه‌نامه‌ی WordNet، هسته‌ی عاطفی این واژه‌نامه ایجاد گردید. Synsetها به‌صورت شبکه‌ای از کلمات هم‌معنی به‌جای کلمه سازمان‌دهی شده‌اند و واحد پایه در واژه‌نامه‌ی WordNet محسوب می‌شوند. طراحان واژه‌نامه‌ی WAL

زیرمجموعه‌ای از تمامی Synsetهایی که حداقل یکی از کلمات موجود در لیست AFFECT اولیه در آن وجود داشت را از واژه‌نامه‌ی WordNet انتخاب نموده و بقیه Synsetها دور انداخته شد.



شکل ۱۰. گراف جزئی از حوزه‌ها و سلسله‌مراتب WAL

سپس این مجموعه‌ی جدید با اطلاعات احساسی به‌صورت دستی برچسب‌گذاری گردیده و برای توسعه‌ی هسته اولیه آماده گردید. در مرحله‌ی بعد ازآنجایی که تعداد ثابتی از روابط لغوی و معنایی در واژه‌نامه‌ی WordNet وجود دارد با استفاده ازاین‌روابط هسته‌ی عاطفی تشکیل شده در مرحله‌ی قبل گسترش داده شد. در این توسعه چنانچه رابطه‌ی موردنظر (به‌عنوان‌مثال روابطی چون مترادف/متضاد، شباهت، مشتق از و...) حفظ‌کننده‌ی معنای عاطفی[1] بود و Synset در واژه‌نامه وجود نداشت آن Synset به واژه‌نامه اضافه می‌گردید.

[1] یعنی اگر رابطه ای که به مجموعه ای از Wordnet-Affect اعمال میشود، Synsetهایی تولید نماید ه هنوز مفاهیم عاطفی را تولید می‌کنند

در واقع نگاشت بر مبنای یک برچسب‌گذاری سلسله‌مراتبی عواطف، به‌صورت خودکار و مستقل از دامنه با تأکید بر روابط موجود در WordNet انجام می‌پذیرد که بخشی از آن درشکل ۱۰ نشان‌داده‌شده است.

برخلاف GI و DAL این منبع پیکره محور نیست بلکه از طریق هر دو فرایند بصری و خودکار حاصل شده است[135]. نسخه‌ی اولیه این واژه‌نامه از ۲۸۷۴ عدد Synset تشکیل شده است و شامل ۴۷۸۷ کلمه است که ٪۵۱ از این کلمات صفات، ٪۲۷ اسامی، ٪۱۱ قید و ٪۱۱ بقیه فعل هستند. این واژه‌نامه اساساً متفاوت از دیگر واژه‌نامه‌ها است و احساسات را بر مبنای ابعاد اَزگود بوسیله قطبیت احساسی آن‌ها مقداردهی نمی‌نماید.

جدول ۸. کلمات و Synsetهای احساسی دسته‌بندی شده بر اساس برچسب‌گذاری نحوی

	مجموع	قیود	افعال	صفات	اسامی
Synsets	۲۸۷۴	۳۲۷	۳۲۲	۱۴۶۲	۷۶۳
Words	۴۷۸۷	۵۵۲	۶۵۷	۲۲۹۳	۱۲۸۵

۳-۵-۳-۷ واژه‌نامه‌ی SentiFul

نسخه‌ی نهایی این واژه‌نامه[11] که دارای ۱۲۹۰۰ اصطلاح احساسی است حاوی ۱۳۷ تغییردهنده، ۲۴۰ کلمه‌ی تقویت‌کننده/معکوس‌کننده و ۷۴ عملگر مودال است. این واژه‌نامه به‌صورت خودکار و با استفاده از روابط هم‌معنایی و تضاد، شُمول معنایی[1]، اشتقاق[2] و ترکیب با لغات شناخته شده، ساخته شده است. در این واژه‌نامه به هر کلمه علاوه بر قطبیت مثبت و منفی از لحاظ وزنی نیز در بازه‌ی ۰ تا ۵ امتیازدهی شده است. هسته‌ی اولیه‌ی این واژه‌نامه شامل ۲۴۳۸ کلمه‌ی احساسی برگرفته از کلمات حسی موجود در پایگاه‌داده‌ی افکت[148] است که ۹۱۷ صفت، ۲۴۳ قید، ۹۰۰ اسم و ۳۷۷ عدد فعل را شامل می‌شود.

[1] Hyponymy
[2] Derivation

جدول ۹. نمونه ای از کلمات بهمراه مقادیر آنها در واژه‌نامه Sentiful

مقادیر غیر صفر از بردار احساسی پایگاه‌داده Affected		امتیازات قطبی		امتیازات وزنی		برچسب	کلمه احساسی
		امتیاز مثبت	امتیاز منفی	وزن مثبت	وزن منفی		
تعجب: ۱.۰، شادی:۰.۵، ترس:۰.۱		۰.۷۵	۰.۱	۰.۶۷	صفت	صفت	**Tremendous**
غم:۰.۲، علاقه:۰.۱		۰.۱	۰.۲	۰.۵	قید	قید	**Pensively**
شادی:۰.۹،علاقه:۰.۶،تعجب:۰.۵		۰.۶۷	۰.۰	۱.۰	اسم	اسم	**Success**
گناه:۰.۲،غم:۰.۱		۰.۰	۰.۱۵	۰.۰	فعل	فعل	**Regret**

با بهره‌گیری از ویژگی‌های احساسی هرکدام از کلمات موجود در پایگاه‌داده‌ی افکت که با استفاده از نُه حس، خشم، نفرت، ترس، گناه، علاقه، شادی، غم، شرم و تعجب کدگذاری و مقداردهی گردیده بود، بردار عاطفی هر کلمه تولید و سپس باتوجه‌به این موضوع که اساساً سه حس (علاقه، شادی و تعجب) دارای جهت‌گیری مثبت و شش حس (خشم، نفرت، ترس، گناه، غم، شرم) دارای جهت‌گیری احساسی منفی می‌باشند و باتوجه‌به مقادیر بردار عاطفی هرکدام از کلمات احساسی، قطبیت کلمات و وزن قطبی هر کلمه مشخص گردید. سپس با استفاده از شبکه‌ی واژگان حسی انگلیسی SentiWordNet اقدام به ابهام‌زدایی و آزمون درستی کلمات حسی گردید. آنگاه با بهره‌گیری از روابط معنایی مختلف موجود در شبکه‌ی واژگان، مجموعه‌ی لغات واژه‌نامه گسترش داده شده است. این فرایند در سه‌گام به‌صورت زیر انجام پذیرفت:

گام اول: برای هر کلمه‌ی موجود در واژه‌نامه‌ی SentiFul تمامی Synsetهای حاوی کلمه‌ی مربوطه استخراج گردید.

گام دوم: برای هرکدام از Synsetهای استخراج شده در مرحله‌ی قبل، کلماتی که قبلاً درون واژه‌نامه‌ی SentiFul وجود داشت استخراج شده و امتیازات و وزن‌های درون Synset برای آن‌ها محاسبه گردیده و این امتیاز و وزن جدید به مابقی کلماتی که در Synset وجود داشت منتسب گردید.

جدول ۱۰. دسته‌بندی انواع پیوست‌های استفاده شده برای تولید کلمات جدید در SentiFul

Type of affix	**Prefix (+*class of base lexeme*); (*class of base lexeme*+) suffix**
	Adjective formation
Propagating	pro- (+*a*); (*a*+) -ish; (*v*+) {-able, -ant, -ent, -ible,-ing}; (*n*+) {-al, -en, -ful, -ic, -like, -type, -y}; (*v*/*n*+) {-ate, -ed, -ive, -ous}
Reversing	{a-, ab-, an-, anti-, contra-, counter-, de-, dis-, dys-, il-, im-, in-, ir-, mal-, mis-, non-, pseudo-, un-, under-} (+*a*); (*n*+) -less
Intensifying	{extra-, hyper-, mega-, super-, ultra-} (+*a*)
Weakening	semi- (+*a*)
	Adverb formation
Propagating	pro- (+*adv*); (*a*+) -ly; (*n*+) {-wise, -wards}
Reversing	{a-, ab-, an-, anti-, contra-, counter-, de-, dis-,dys-, il-, im-, in-, ir-, mal-, mis-, non-, pseudo-, un-, under-} (+*adv*);
Intensifying	{extra-, hyper-, mega-, super-, ultra-} (+*adv*)
Weakening	semi- (+*adv*)
	Noun formation
Propagating	{neo-, re-} (+*n*); (*v*+) {-age, -al, -ant, -ation, -ent, -ication, -ification, -ion, -ment, -sion, -tion, -ure}; (*a*+) {-ity, -ness}; (*n*+) {-ful, ist, -ship}; (*v*/*a*+) {-ance, -ence, -ee}; (*v*/*n*+) {-er, -ing, -or}; (*a*/*n*+) {-cy, -dom, -hood}; (*v*/*n*/*a*) {-ery, -ry}
Reversing	{anti-, counter-, dis-, dys-, in-, mal-, mis-, non-, pseudo-, under-} (+*n*)
Intensifying	{arch-, hyper-, mega-, super-, ultra-} (+*n*)
Weakening	{mini-, semi-} (+*n*); (n+) {-ette, -let}
	Verb formation
Propagating	{be-, co-, fore-, inter-, pre-, pro-, re-, trans-} (+*v*); {em-, en-} (+*n*/*a*); (*n*/*a*+) {-ate, -en, -fy, -ify, -ise, -ize}
Reversing	{de-, dis-, dys-, mis-, un-, under-} (+*v*)
Intensifying	{out-, over-} (+*v*)

گام سوم: پس از تکمیل گام اول و گام دوم برای تمامی موجودیت‌های اصلی Sentiful تمامی کلمات جدید تکراری که از Sensetهای مختلف و با استفاده از کلمات متفاوت موجود در SentiFul به‌دست‌آمده‌اند حذف شده و امتیاز جدیدی که برابر با متوسط امتیازات نسبت‌داده‌شده به آن‌ها است به آن‌ها نسبت داده شد.

در این مرحله ۴۱۹۰ کلمه‌ی جدید با درنظرگرفتن رابطه‌ی مترادفی مستقیم به‌صورت خودکار از واژه‌نامه‌ی WordNet استخراج گردید که شامل ۱۱۲۲ صفت، ۱۰۷ قید، ۱۷۳۱ اسم و ۱۲۳۰ فعل بود. در مرحله‌ی بعد با استفاده از روش‌های ریخت‌شناسی و دست‌کاری ریشه کلمات، با پیشوندها و پسوندهای مختلف

کلمات جدید تولید و امتیازدهی گردید. آن‌ها چهار نوع پیوست را باتوجه‌به نقشی که در ارتباط با ویژگی‌های احساسی بازی می‌کردند تعریف نمودند که عبارت بود از پیوست‌های منتشر شونده[1]، پیوست‌های معکوس کننده، پیوست‌های تقویت‌کننده و پیوست‌های تضعیف‌کننده.

جدول ۱۱. لیست کلمات موجود در واژه‌نامه‌ی SentiFul به تفکیک منبع

منبع	صفت	قید	اسم	فعل	مجموع
کلمات اصلی استخراج شده از پایگاه‌داده افکت	۹۱۷	۲۴۳	۹۰۰	۳۷۷	۲۴۳۸
کلمات استخراج شده با استفاده از روابط معنایی از WordNet	۱۱۲۲	۱۰۷	۱۷۳۱	۱۲۳۰	۴۱۹۰
کلمات تولید شده با استفاده از ریشه و پیوست‌های موجود	۱۴۰۵	۴۸۴	۱۸۰۰	۳۴۰	۴۰۲۹

در این مرحله ریشه‌ی کلمات موجود در SentiFul استخراج گردیده و سپس با استفاده از این پیوست‌ها کلمات جدید تولید گردید و پس از تأیید صحت کلمه‌ی تولید شده امتیازات کلمات جدید به‌صورت زیر محاسبه گردید.

چنانچه از پیوست منتشر شونده استفاده شده بود امتیاز ریشه به کلمه‌ی جدید منتقل می‌گردید و چنانچه از پیوست معکوس کننده استفاده شده بود امتیاز ریشه معکوس شده و به کلمه‌ی جدید انتساب داده می‌شد.

[1] Propagating Affixes

برای تقویت‌کننده و تضعیف‌کننده امتیاز ریشه به ترتیب در ۲ و ۰٫۵ ضرب و امتیاز حاصل به کلمه‌ی جدید نسبت داده می‌شد. جدول ۱۰ انواع پیوست‌های استفاده شده در واژه‌نامه‌ی sentiFul را نشان می‌دهد.

با استفاده از رویکرد ریخت‌شناسی ۴۰۲۹ کلمه‌ی جدید شامل ۱۴۰۵ صفت، ۴۸۴ قید، ۱۸۰۰ اسم و ۳۴۰ فعل به‌صورت خودکار تولید گردید. جدول ۱۱ لیست کلمات موجود در واژه‌نامه اولیه را به تفکیک منبع نشان می‌دهد.

۶-۳ روش‌های پایه آنالیز احساسات با استفاده از Unigram و Ngram

در آنالیز احساسات مبتنی بر Unigram محاسبه‌ی امتیاز یک جمله به روشی بسیار سرراست و با استخراج امتیازات احساسی کلمات از واژه‌نامه و جمع این امتیازات برای محاسبه‌ی امتیاز جمله به دست می‌آید[10]. برخی از نویسندگان با افزودن قوانین وابسته به انواع خاصی از کلمات، نقطه‌گذاری‌ها و کلمات عامیانه[1] این روش‌ها را بهبود داده‌اند[10,12].

چند نویسنده نیز برای درک بهتر احساس متن استفاده از ساختار مبتنی بر معانی بیانی[2] را پیشنهاد کرده‌اند[149,150]. در تعداد زیادی از روش‌های جدید آنالیز احساسات، برای یافتن قطبیت اسناد متنی از امتیاز Unigramهای احساسی، استخراج شده از واژه‌نامه، همراه با تقویت‌کننده‌ها و منفی‌کننده‌ها استفاده می‌شود[151]. به‌عنوان‌مثال امتیاز یک کلمه Unigram مانند good می‌تواند با یک کلمه تقویت‌کننده همچون very افزایش یافته و یا با یک کلمه تضعیف‌کننده همچون slightly تضعیف گردد. همچنین می‌تواند با یک کلمه منفی‌کننده همچون not منفی شود. در مقایسه با روش‌های مبتنی بر Unigram استفاده از این گونه روش‌های مبتنی بر Ngram بهتر عمل می‌نمایند[12,96].

در بین این روش‌ها، SCL-OPP دستی بوده و از تفسیرهای انسانی برای امتیازدهی۳ به Ngramها استفاده می‌نماید. بقیه روش‌ها: Contextual Valence Shifter، SO-CAL، SO-CAL in Inheritance-based

[1] Slang
[2] Rhetorical
[3] Human Annotators

approach، Senti-N-gram و TF-IDF Lexicon از روش‌های مبتنی بر قوانین نیمه‌خودکار برای تولید امتیازات یک Ngram احساسی استفاده نموده‌اند که در جدول ۱۲ نشان‌داده‌شده است.

جدول ۱۲. روش‌های نیمه‌خودکار مبتنی بر واژگان برای اصلاح نمره Ngramها با افزودن تقویت‌کننده‌ها و منفی‌کننده‌ها

روش	منفی‌کننده	تضعیف‌کننده	تقویت‌کننده
Contextual Valence Shifter	قطبیت Unigram معکوس می‌شود	یک امتیاز از امتیاز Unigram کم می‌شود	یک امتیاز به امتیاز Unigram اضافه می‌شود
SO-CAL	امتیاز Unigram به‌اندازه‌ی ۴ واحد در جهت عکس قطبیت جابه‌جا می‌شود	امتیاز Unigram به‌وسیله درصد از پیش تعریف شده‌ای از تضعیف‌کننده کاهش می‌یابد	امتیاز Unigram به‌وسیله درصد از پیش تعریف شده‌ای از تقویت‌کننده افزوده می‌شود
SO-CAL in Inheritance-based	امتیاز Unigram به‌اندازه‌ی ۴ واحد در جهت عکس قطبیت جابه‌جا می‌شود	امتیاز Unigram به‌وسیله درصد از پیش تعریف شده‌ای از تضعیف‌کننده ضرب می‌شود	امتیاز Unigram به‌وسیله درصد از پیش تعریف شده‌ای از تقویت‌کننده افزوده می‌شود
Senti-N-gram	امتیاز Ngram از جدول از پیش تعریف شده استخراج می‌گردد	امتیاز Ngram از جدول از پیش تعریف شده استخراج می‌گردد	امتیاز Ngram از جدول از پیش تعریف شده استخراج می‌گردد
TF-IDF Lexicon	با استفاده از Unigram و ضریب به‌دست‌آمده از TF-IDF امتیاز محاسبه می‌گردد	با استفاده از Unigram و ضریب به‌دست‌آمده از TF-IDF امتیاز محاسبه می‌گردد	با استفاده از Unigram و ضریب به‌دست‌آمده از TF-IDF امتیاز محاسبه می‌گردد
VADER	قطبیت Unigram معکوس می‌شود	امتیاز Unigram با عدد ثابت 0.293 تفریق می‌شود	امتیاز Unigram با عدد ثابت 0.293 جمع می‌شود

در جدول زیر نیز برخی از ویژگی‌های آماری واژه‌نامه‌های مطرح شده فوق را ارائه نموده است.

جدول ۱۳. اطلاعات آماری برخی از واژه‌نامه‌های احساسی

نام واژه‌نامه	منفی				خنثی				مثبت			
	اسم	فعل	قید	صفت	اسم	فعل	قید	صفت	اسم	فعل	قید	صفت
WKWSCI	۲۴۷۵	۱۳۹۳	۷۵۸	۲۴۷۴	۱۰۰۷۶	۴۵۳۱	۱۱۵۴	۳۷۳۶	۷۱۲	۳۰۵	۶۰۲	۱۵۰۲
GI	۷۶۱	۷۰۲	-	۸۰۰	-	-	-	-	۶۸۱	۴۰۶	-	۷۷۱
MPQA	۱۳۴۶	۸۶۹	۱۸۳	۱۸۳۹	۱۴۴	۷۶	۱۹	۲۳۵	۶۷۷	۳۸۰	۱۲۸	۱۱۷۱
SO-CAL	۱۰۰۵	۷۹۱	۴۲۹	۱۵۷۶	۰	۰	۰	۱	۵۴۴	۳۵۱	۴۴۸	۱۲۵۰
NRC	۳۳۲۴								۲۳۱۲			
Hu &Liu				۴۸۷۳								۲۰۰۶
SentiWordNet	۱۰۴۲۰	۱۶۶۴	۴۳۹	۶۲۷۳	۹۹۷۰۶	۸۵۰۸	۱۹۶۴	۱۰۴۰۱	۷۶۷۲	۱۳۵۷	۲۰۷۸	۴۸۰۵

۳-۷ جمع‌بندی

در این بخش سعی نمودیم در ابتدا روش‌های طبقه‌بندی احساسات را به طور خلاصه توصیف کرده و نگاهی گذرا بر انواع روش‌های تولید واژه‌نامه‌های احساسی بیندازیم و سپس با جزئیات بیشتری روش‌های تولید واژه‌نامه را موردبحث قرار دهیم و به‌مرور مهم‌ترین و شاخص‌ترین کارهای انجام شده در این حوزه در زبان انگلیسی پرداخته و برخی از مهم‌ترین واژه‌نامه‌های احساسی موجود را مورد بررسی دقیق‌تر قرار دهیم و به طور مختصر مراحل تولید آن‌ها را بررسی نماییم. همان‌طور که در جدول ۱۴ داده شده است سه ستون اول روش‌های بیان شده را بر مبنای این موضوع که آیا یک واژه‌نامه خاص ایجاد شده است یا خیر، و همچنین نوع دیکشنری بر مبنای Unigram، Bigram و Trigram از هم متمایز می‌نماید. سه ستون آخر نیز روش تولید این واژه‌نامه‌ها را بر مبنای سه روش دستی، نیمه دستی، و اتوماتیک مشخص نموده است .

با دقت در جدول فوق مشاهده می‌نمایید که از بین تمامی واژه‌نامه‌های موجود تنها واژه‌نامه‌های Senti-N-gram، Lexical TF-IDF و SCL-OPP از عبارات احساسی Ngram پشتیبانی می‌نمایند. در این میان واژه‌نامه Lexical TF-IDF یک واژه‌نامه حوزه محور و اتوماتیک بوده که با گذشت زمان و ورود جملات جدید

نیازمند به‌روزرسانی مجدد دارد و دو دیکشنری senti-N-gram و SCL-OPP نیز از تعداد بسیار معدودی از کلمات Bigram و Trigram حمایت می‌نمایند که این مسئله یکی از بزرگترین مشکلات واژه‌نامه‌های ارائه شده فعلی بوده.

جدول ۱۴. مقایسه واژه‌نامه‌های موجود بر مبنای توسعه واژگان و روش امتیازدهی

نام واژه‌نامه	توسعه واژگان			محاسبه امتیاز بر مبنای Ngram		
	Unigram	Bigram	Trigram	دستی	نیمه دستی	اتوماتیک
SO-CAL	بله	خیر	خیر	خیر	بله	خیر
ANEW	بله	خیر	خیر	بله	خیر	خیر
VADER	بله	خیر	خیر	بله	خیر	خیر
Lexical TF-IDF	خیر	بله	بله	خیر	خیر	بله
DAL	بله	خیر	خیر	بله	خیر	خیر
GI	بله	خیر	خیر	بله	خیر	خیر
SCL-OPP	خیر	بله	بله	بله	خیر	خیر
DAL	بله	خیر	خیر	بله	خیر	خیر
Senti-N-gram	خیر	بله	بله	خیر	خیر	بله
Hu & Liu	بله	خیر	خیر	خیر	خیر	بله
WKWSCI	بله	خیر	خیر	بله	خیر	خیر
MPQA	بله	خیر	خیر	خیر	بله	خیر

فصل چهارم

بررسی اثر موقعیت مکانی

تشدیدکننده‌ها نسبت به کلمه احساسی

فصل ۴– بررسی اثر موقعیت مکانی تشدیدکننده‌ها نسبت به کلمه احساسی

۴-۱ مقدمه

باتوجه‌به مرور انجام شده و توضیحات فصول قبل ازآنجائی‌که امکان تولید خودکار واژه‌نامه‌های احساسی بزرگ‌تر با بهره‌گیری از واژه‌نامه‌های احساسی تولید شده به‌صورت دستی و با استفاده از تکنیک‌های نیمه نظارتیی وجود دارد[126,152] و باتوجه‌به نقاط ضعف روش‌هایی که قبلاً در این حوزه کارشده بود، در این پایان‌نامه سعی شده است تا یک واژه‌نامه‌ی مبتنی بر Ngram برای بالابردن درصد درستی تشخیص احساس در متون غیر ساختار یافته پیشنهاد شود. باتوجه‌به اینکه در بیشتر روش‌های مبتنی بر واژه‌نامه تنها از صفات به‌عنوان شاخصی برای جهت‌گیری معنایی متن استفاده شده است و تنها تعداد بسیار معدودی [2,11,96] به نقش تقویت‌کننده‌ها در جهت‌گیری معنایی متن پرداخته‌اند هدف از این فصل تولید یک واژه‌نامه‌ی احساسی با توانایی کاربست کلمات تقویت‌کننده و منفی‌کننده متوالی و تخصیص امتیاز به عبارات احساسی چندکلمه‌ای به جهت افزایش میزان دقت و صحت تشخیص احساسات در میان متون غیر ساختاریافته و جملات احساسی است. یکی از دلایل انتخاب این موضوع به‌عنوان پایه پروژه، ضعف بسیار بالای واژه‌نامه‌های Ngram موجود در پوشش تمامی حالات مختلف قرارگیری تشدیدکننده‌ها و منفی‌کننده‌های متوالی و همچنین عدم توجه به تأثیر قطبیت کلمه احساسی بر تشدیدکننده پیش از آن بوده است.

در یک محاسبه ساده چنانچه تعداد کلمات احساسی را ۷۵۰۰ عدد (تقریباً معادل کلمات احساسی موجود در واژه‌نامه VADER) و حدود ۲۰۰ تشدیدکننده در نظر بگیریم ۱۵۰۰۰۰۰ عدد Bigram مختلف تولید خواهد شد و ۳۰۰۰۰۰۰۰۰ عدد Trigram درحالی‌که در روش senti-N-gram تنها ۲۴۱۵ عدد Bigram معتبر و ۳۰۵ عدد Trigram تولید شده است این مقدار برای الگوریتم SCL-OPP نیز ۳۱۱ عدد Bigram و ۲۶۵ عدد Trigram است. از سوی دیگر امکان امتیازدهی دستی این حجم عظیم از ترکیبات مختلف و متنوع علاوه بر صرف زمان بسیار زیاد و هزینه فوق‌العاده نیازمند استفاده از تعداد زیادی از افراد کاملاً متخصص در ارزیابی

عبارات احساسی خواهد بود و عملاً این کار غیرممکن خواهد بود. علاوه‌برآن پردازش موردنیاز برای بررسی Ngramها در پایگاه‌داده‌های مختلف ازجمله مواردی است که بر پیچیدگی‌های کار می‌افزاید. در این بخش ابتدا مراحل آماده‌سازی داده‌ها را به‌اختصار بیان کرده و سپس پایگاه‌داده‌های موجود به‌اختصار معرفی شده و در ادامه رویکرد مدنظر، چالش‌ها و فرضیات این حوزه را مورد بررسی قرار خواهیم داد. در نهایت نتایج به‌دست‌آمده از شبیه‌سازی‌های ارائه خواهد شد.

۲-۴ آماده‌سازی داده‌ها[1]

آماده‌سازی داده‌ها یکی از مهم‌ترین و زمان‌برترین بخش‌ها در فرایند تجزیه‌وتحلیل احساسات محسوب می‌شود. ازآنجایی‌که واژه‌های مورداستفاده در یک متن غالباً دارای تفاوت‌های ساختاری هستند ازاین‌رو آماده‌سازی متن برای انجام عملیات پردازشی اصلی یکی از مهم‌ترین بخش‌ها محسوب می‌شود. به‌منظور استخراج دانش‌پایه، در گام نخست هر متن می‌بایست مورد پیش‌پردازش قرار گیرد. به‌طورکلی آماده‌سازی متن به چهاربخش اصلی جداسازی اجزای متن یا قطعه‌بندی، پاک‌سازی داده، نرمال‌سازی و حاشیه‌نویسی تقسیم می‌شود. در طی فرایند پاک‌سازی قسمت‌های کم کاربرد متن از طریق عملیات‌هایی همچون حذف نویز، Case Folding و حذف ایست‌واژه‌ها از متن پالایش می‌شوند. فرایند نرمال‌سازی فرایند ترجمه یا نگاشت اصطلاحات یا کاهش حالات مختلف کلمات از طریق ریشه‌یابی[2] و بن‌واژه‌سازی[3] و یا سایر اشکال استانداردسازی است. جداسازی اجزای متن یا قطعه‌بندی فرایندی است که رشته‌های طولانی متن را به قطعات کوچک‌تر یا نشانه‌ها تقسیم می‌نماید در این مرحله قطعات متن را می‌توان به‌صورت جملات و جملات را به‌صورت کلمات یا عبارات Ngram تبدیل نمود. در نهایت حاشیه‌نویسی بر اساس نوع کاربرد از برچسب‌گذاری[4]، اضافه نمودن علایم یا برچسب‌گذاری واژگان کلام[5] استفاده می‌نماید.

[1] Data Preprocessing
[2] Stemming
[3] Lemmatization
[4] Labeling
[5] Pos Taging

- **حذف نویز:** نخستین مرحله از پیش‌پردازش متن پاک‌سازی متن است. در طی فرایند پاک‌سازی داده کلیه اعداد، آدرس‌ها و لینک وب‌سایت‌ها (**www،url ،http**) و کاراکترهای غیرالفبایی همچون (؟،،،،. ،! و.. .) و همچنین کاراکترهای اضافی -کاراکترهایی که مفهومی به متن اضافه نمی‌کنند ("،}،)،[و..)-و همچنین کاراکترهای تکراری (goooooood) حذف گردید. این عملیات باعث افزایش سرعت پردازش گردیده و بار پردازشی موردنیاز برای عملیات اصلی را کاهش می‌دهد.

- **Case Folding :** فرایند تبدیل تمامی کاراکترهای درون یک متن به یک حالت را case-folding می‌نامند. در این مرحله کلیه‌ی کلمات از نظر کوچک بودن و یا بزرگ بودن حروف به یک صورت در می‌آیند. این مرحله به دلیل ممانعت از ایجاد مدل‌های یکسان از کلمات صورت می‌پذیرد و بار پردازشی و ذخیره‌سازی سیستم را کاهش می‌دهد. در این پژوهش تمامی حروف کلمات پایگاه‌داده به حروف کوچک تبدیل گردید.

- **حذف ایست‌واژه‌ها:** ایست‌واژه‌ها کلماتی هستند که با وجود تکرار فراوان در متن از جهت معنایی اهمیت کمی دارند. در تمامی حوزه‌های مرتبط با پردازش متن حذف این‌گونه از کلمات باعث کاهش بار محاسباتی می‌گردد. این کلمات که شامل ضمایر، حروف تعریف، حروف اضافه و حروف ربط می‌شوند، حاوی اطلاعات نیستند. همچنین حروف تعریفی همچون "a" و "the" و ضمایر مانند "he"، "I" و "they" معمولاً حاوی هیچ‌گونه اطلاعات احساسی نیستند.

- **جداسازی متن یا قطعه‌بندی:** نشانه‌گذاری یا جداسازی متن، تقسیم‌بندی متن به واحدهای کوچک‌تری چون جملات، اصطلاحات[1]، کلمات و نمادها است که در اصطلاح نشانه[2] نامیده می‌شوند در این پژوهش کلیه نظرات به واحدهای کوچک‌تری تقسیم شدند و درنهایت کلمات و اصطلاحات به‌دست‌آمده است.

[1] Terms
[2] Tokens

- **استخراج Ngramها** :در این پیاده‌سازی به‌جای درنظرگرفتن هر کلمه به‌تنهایی، کلمات متوالی به‌منظور درنظرگرفتن عبارات احساسی مورداستفاده قرار گرفته است. مدل Ngram به‌صورت گسترده‌ای در متن-کاوی و پردازش زبان طبیعی استفاده می‌شود Ngramها یک توالی به‌هم‌پیوسته از n عنصر یک متن یا سخنرانی هستند. در Ngram اندازه‌ی ۱ به Unigram، اندازه‌ی ۲ به Bigram و اندازه‌ی ۳ به Trigram اشاره دارد. پس از تولید تمامی Unigramها و Bigramها و Trigramهای موجود در متن تنها Ngramهای احساسی موجود، استخراج و مجموع امتیازات ثبت شده توسط کاربران و تعداد تکرار هرکدام در کل پیکره به تفکیک ثبت گردید.

- **ریشه‌یابی کلمات:** ریشه‌یابی کلمات عبارت است از استخراج ریشه‌ی کلمات از طریق حذف پسوندها و پیشوندها به نحویکه کلمات با ریشه‌ی یکسان دارای شکل یکسانی شوند. در یک زبان هر واژه باتوجه‌به نقش معنایی و نحوی خود به شکل متفاوتی در یک جمله ظاهر می‌شود. این اشکال متفاوت با آنکه تماماً از یک ریشه مشتق شده‌اند اما دارای بار معنایی هستند، بااین‌حال از نظر معنایی نزدیکی بسیار زیادی با یکدیگر دارند. هدف از ریشه‌یابی، استاندارد نمودن کلمات است. در این مرحله کلیه‌ی کلمات به فرمت ریشه‌ی اصلی خود درمی‌آیند. الگوریتم‌های ریشه‌یابی پیشوندها و پسوندهای کلمه را بر اساس قواعد گرامری حذف کرده و ریشه‌ی کلمه را تولید می‌کنند.

- **برچسب‌گذاری اجزاء واژگانی کلام:** در زبان‌شناسی پیکره، برچسب‌گذاری بخشی از گفتار (POStagging) که به آن برچسب‌گذاری گرامری نیز گفته می‌شود. علامت‌گذاری کلمه‌ای در یک متن (پیکره) تشخیص نقش گرامری آن کلمه بر مبنای رابطه با کلمات مجاور و مرتبط و همچنین تعریف آن کلمه است. ازآنجایی‌که برخی از کلمات دارای نقش‌های دستوری مختلفی هستند و بار معنایی متفاوتی را تولید می‌نمایند برچسب‌گذاری اجزاء واژگان کلام نقش بسیار مهمی در تشخیص بار احساسی کلمات ایفا می‌نماید. در این روش به هر کلمه‌ای بر اساس نقش دستوری آن کلمه در متن، برچسبی چون فعل صفت یا قید، تخصیص می‌دهند.

۳-۴ پایگاه‌داده

یکی از مهم‌ترین موارد در پیاده‌سازی روش پیشنهادی، استفاده از پایگاه‌داده‌های اصولی برچسب خورده و متداول در مقالات می‌باشد که معیار مقایسه و استخراج واژه‌نامه‌ی نیمه‌خودکار خواهد بود. بر اساس مرور مطالب انجام شده و مقالات پایه‌ی این حوزه، تعدادی پایگاه‌داده استاندارد وجود دارد که بسته به حجم لغات و پردازش موردنیاز برای تحلیل، از آن‌ها می‌توان در موارد مختلف استفاده کرد. اولین و بزرگ‌ترین پایگاه‌داده در این حوزه پایگاه‌داده سایت آمازون[1] است که دادگان بسیار عظیمی از نظرات در موضوعات مختلف را در خود گردآوری کرده و برچسب‌گذاری نموده است. این پایگاه‌داده، بیست و چهار حوزه مختلف را در دسته‌بندی خود گردآوری کرده و از نظر تعداد نظرات نسبت به همه‌ی پایگاه‌داده‌های مشابه بزرگ‌تر است. در مجموع در حدود هجده میلیون نظر و هشتاد میلیون محتوا در این پایگاه‌داده وجود دارد که از این نظر آن را به یکی از بزرگ‌ترین و جامع‌ترین پایگاه‌داده‌های موجود تبدیل کرده است. نظرات موجود در این پایگاه‌داده حاوی یک سیستم رتبه‌بندی در مقیاس ۵ بوده و در قالب فایل‌های با فرمت JSON ذخیره شده‌اند. محتوای موضوع و تعداد نظر هر قسمت در جدول ۱۵ آورده شده تا خواننده با فرم کلی این پایگاه‌داده آشنا گردد. این پایگاه‌داده محوریت اصلی استخراج واژه‌نامه‌ی نیمه‌خودکار پیشنهادی در این پایان‌نامه خواهد بود.

به‌عنوان دومین پایگاه‌داده از [106] برای استخراج لیست تشدیدکننده‌ها استفاده شده است. این مرجع یک پایان‌نامه‌ی دکترا در حوزه‌ی پردازش زبان است که در انتهای آن به‌صورت پیوست لیست تشدیدکننده‌های زبان انگلیسی آورده شده است. بر اساس مرور انجام شده در مقالات لیست داده شده در این پایان‌نامه نسبت به سایر انواع موجود کامل‌تر بوده و از این نظر معیار استخراج تشدیدکننده‌ها در این پایان‌نامه قرار خواهد گرفت. لازم به توضیح است امتیاز اختصاص داده شده توسط [106] به دلیل رویکرد متفاوت آن مورداستفاده نبوده و مجدداً امتیازدهی و قواعد مناسب برای اعمال اثر تشدیدکننده‌ها در متن در طرح روش پیشنهادی مدنظر قرار خواهد گرفت.

[1] Http://Jmcauley.Ucsd.Edu/Data/Amazon/

جدول ۱۵ . محتوا و نظر هر بخش موضوعی از پایگاه‌داده آمازون

حوزه	تعداد محتوا	تعداد نظر
Books	۲۲,۵۰۷,۱۵۵	۸,۸۹۸,۰۴۱
Electronics	۷,۸۲۴,۴۸۲	۱,۶۸۹,۱۸۸
Movies and TV	۴,۶۰۷,۰۴۷	۱,۶۹۷,۵۳۳
CDs and Vinyl	۳,۷۴۹,۰۰۴	۱,۰۹۷,۵۹۲
Shoes and Jewelry,Clothing	۵,۷۴۸,۹۲۰	۲۷۸,۶۷۷
Home and Kitchen	۴,۲۵۳,۹۲۶	۵۵۱,۶۸۲
Kindle Store	۳,۲۰۵,۴۶۷	۹۸۲,۶۱۹
Sports and Outdoors	۳,۲۶۸,۶۹۵	۲۹۶,۳۳۷
Cell Phones and Accessories	۳,۴۴۷,۲۴۹	۱۹۴,۴۳۹
Health and Personal Care	۲,۹۸۲,۳۲۶	۳۴۶,۳۵۵
Toys and Games	۲,۲۵۲,۷۷۱	۱۶۷,۵۹۷
Video Games	۱,۳۲۴,۷۵۳	۲۳۱,۷۸۰
Tools and Home Improvement	۱,۹۲۶,۰۴۷	۱۳۴,۴۷۶
Beauty	۲,۰۲۳,۰۷۰	۱۹۸,۵۰۲
Apps for Android	۲,۶۳۸,۱۷۲	۷۵۲,۹۳۷
Office Products	۱,۲۴۳,۱۸۶	۵۳,۲۵۸
Pet Supplies	۱,۲۳۵,۳۱۶	۱۵۷,۸۳۶
Automotive	۱,۳۷۳,۷۶۸	۲۰,۴۷۳
Grocery and Gourmet Food	۱,۲۹۷,۱۵۶	۱۵۱,۲۵۴
Lawn and Garden,Patio	۹۹۳,۴۹۰	۱۳,۲۷۲
Baby	۹۱۵,۴۴۶	۱۶۰,۷۹۲
Digital Music	۸۳۶,۰۰۶	۶۴,۷۰۶
Musical Instruments	۵۰۰,۱۷۶	۱۰,۲۶۱
Amazon Instant Video	۵۸۳,۹۳۳	۳۷,۱۲۶
مجموع	۸۰,۷۳۷,۵۶۱	۱۸,۱۸۶,۷۳۳

لیست کلمات احساسی از واژه‌نامه‌ی VADER استخراج شده که در بخش۳-۵-۲-۳ راجع به آن توضیح داده شد. مجموعاً حدود ۷۵۰۰ لغت احساسی در واژه‌نامه VADER وجود دارد که امتیاز آن‌ها در این واژه‌نامه به‌صورت دستی محاسبه و در واژه‌نامه پیشنهادی از آن استفاده شده است. منفی‌سازها بر اساس استاندارد دانشگاه کمبریج[1] در نظر گرفته شده است.

"the tips of my fingers often get pretty dry and cracked, especially when i've been playing a-lot. the problem, at least for me, is that pieces of the dead skins sometimes catch on guitar strings making slides, bends, and pull-offs less fluid. a-few strokes of ghs makes the fretboard give me a much smoother playing experience. an alternative would be using vaseline on my fingers, but that would probably be terrible for my fretboard. also, i sometimes use an emory board to smooth my fingertips which helps, but i'd rather do that as little as possible to keep callouses strong. ghs allows me to minimize that.", "overall": 5.0.

شکل ۱۱. نمونه ای از نظرات ثبت شده در پایگاه داده آمازون

در کنار تشدیدکننده‌ها، منفی‌سازها و جملات احساسی، برای تحلیل درستی روش‌های پیشنهادی نیاز به پایگاه‌داده‌های مشخصی است که به‌صورت استاندارد در مقالات مورد استناد قرار می‌گیرد. بر اساس مرور مقالات صورت گرفت دو پایگاه‌داده برای مقایسه‌ی واژه‌نامه‌ی پیشنهادی و صحت آن با روش‌های موجود مدنظر قرار گرفت که هر دو آن‌ها در مقالات مختلف با معیارهای متفاوت مورد تحلیل قرار گرفته‌اند.

یکی از دو پایگاه‌داده، پایگاه‌داده Taboada است که مجموعه‌ای از هشت حوزه‌ی مختلف و از هر حوزه پنجاه نظر مثبت و منفی را در خود جمع کرده است (مجموعاً چهارصد متن برچسب خورده) که به‌عنوان مقایسه‌ی کارایی واژه‌نامه‌ها معمولاً از این نظرات استفاده می‌شود. این پایگاه‌داده با کمک دانشگاه استنفورد در سال ۲۰۰۴ گردآوری و برچسب‌گذاری شده است[141]. پایگاه‌داده دوم[2] پایگاه‌داده‌ای مشهور Pang & Lee است که مربوط به نقد فیلم‌های سینمایی است[67]. در این پایگاه‌داده ۱۰۰۰ نقد مثبت و ۱۰۰۰ اظهارنظر

[1] https://dictionary.cambridge.org/grammar/british-grammar/questions-and-negative-sentences/negation
[2] http://www.cs.cornell.edu/people/pabo/movie-review-data/

منفی گردآوری شده که همگی راجع به فیلم‌های سینمایی هستند. این پایگاه‌داده نیز به خاطر تنوع لغات و گستره‌ی عبارات احساسی تحت پوشش به‌عنوان معیار مقایسه بین روش‌های مختلف معمولاً مورد استناد و استفاده قرار می‌گیرد.

۴-۴ معیارهای مقایسه

در مرحله پایانی، روش پیشنهادی با دو دیکشنری نسبتاً جدید در این حوزه شامل VADER, Senti-N-gram مقایسه شد. مقایسه در دو پایگاه‌داده Taboada و Pang & Lee که در بخش ۴-۳ راجع به آنها توضیح داده شده صورت گرفت تا کارایی واژه‌نامه پیشنهادی نسبت به دو واژه‌نامه معرفی شده در بخش قبل نشان داده شود. برای مقایسه از پارامترهای متداول True positives (TP) (اسناد با قطبیت مثبت درست برچسب خورده)، True negatives (TN) (اسناد با قطبیت منفی درست برچسب خورده)، False positives (FP) (اسناد با قطبیت منفی که به‌اشتباه مثبت برچسب‌خورده‌اند) و False negatives (FN) (اسناد با قطبیت مثبتی که به‌اشتباه منفی برچسب‌خورده‌اند) استفاده شده است. با استفاده از این معیارها ۴ پارامتر Precision, Recall, F-measure, Accuracy به دست آمد. روابط مربوط به این چهار معیار در فرمول‌های ۴-۱ تا ۴-۴ آورده شده است.

معیار Precision: صحت یک کلاس، نسبت اسناد به‌درستی طبقه‌بندی شده به کل اسناد متعلق به آن کلاس است. به‌طورکلی این پارامتر دقت ارزیابی روش پیشنهادی در حوزه پاسخ نهایی مثبت، را نشان می‌دهد که در اکثر الگوریتم‌ها این معیار عدد کمتری نسبت به Recall است چرا که در تشخیص اسناد منفی به سبب پیچیدگی بالا، تعداد تشخیص اشتباه نسبت به اسناد مثبت زیادتر است، لذا مقدار FP در مخرج رابطه ۴-۱ بالا رفته و مقدار Precision کاهش می‌یابد.

$$Precision\ (p) = \frac{TP}{TP + FP} \qquad (4-1)$$

معیار:Recall نسبت اسناد به‌درستی طبقه‌بندی شده مثبت نسبت به تمام اسناد متعلق به کلاس مثبت است. این معیار پاسخ مناسب‌تری نسبت به Precision از خود نشان می‌دهد چرا که ارزیابی روش در اسناد مثبت ورودی است که اکثر روش‌ها در تشخیص این‌گونه اسناد نسبت به اسناد منفی عملکرد بهتری دارند.

$$Recall\ (R)\ = \frac{TP}{TP + FN} \qquad (4-2)$$

معیار F-measure : میانگین هارمونیک Precision و Recall است.

$$F - measure\ (F)\ = \frac{2 * Precision\ * Recall}{Precision\ + Recall} \qquad (4-3)$$

معیار Accuracy: دقت، قابلیت درستی مجموعه سیستم را در کل اسناد مثبت و منفی نشان می‌دهد. از نظر ایده‌آل هر ۴ معیار درصورتی‌که به سمت یک میل کنند مناسب و هرچه فاصله آن‌ها تا عدد یک بیشتر باشد بیانگر ضعف عملکرد محسوب می‌شود.

$$Accuracy\ (A) = \frac{TP + TN}{TP + TN + FP + FN} \qquad (4-4)$$

برای آنکه ارزیابی در شرایط یکسان صورت گیرد دو دیکشنری نسبتاً جدید VADER, Senti-N-gram با هر دو الگوریتم تشخیص قطبیت سند که در بخش ۴-۵-۳ توضیح داده شده است پیاده‌سازی شده تا شرایط مقایسه با این دیکشنری‌ها کاملاً یکسان باشد و از نسخه‌های اصلاح شده، توابع دلخواه جریمه یا برخی تغییرات هنگام ارزیابی نهایی قطبیت سند استفاده نشده است.

۴-۵ روش پیشنهادی

روش پیشنهادی در این تحقیق برای ایجاد یک دیکشنری جدید در دو قسمت نسبتاً متفاوت طراحی شده است. در مرحله اول پس از پالایش داده‌ها، تمامی ترکیبات احساسی شامل Unigram, Bigram, Trigram از پایگاه‌داده آمازون استخراج شد و با استفاده تحلیل‌های آماری و برخی محاسبات ریاضی یک الگوی فازی به آنها تخصیص داده می‌شود. سپس امتیاز به‌دست‌آمده مورد تصحیح و پالایش قرار می‌گیرد. در مرحله دوم بر

اساس الگوی فازی ترکیبات احساسی، به هر تشدیدکننده یک ضریب تأثیر نهایی تخصیص داده می‌شود که در محاسبه این ضریب چند مرحله جبران‌سازی خطا صورت‌گرفته تا ضریب به‌دست‌آمده به واقعیت و نظر کاربر انسانی نزدیک‌تر بوده و به‌درستی مجموعه سیستم کمک کند. لازم به توضیح است در این پروسه از امتیازات واژه‌نامه VADER که به‌صورت دستی محاسبه شده است کمک گرفته شده که این موضوع نیز به دلیل تأثیر تصمیمات انسانی در تخصیص امتیاز نهایی، سطح کارایی ضرایب محاسبه شده را افزایش می‌دهد.

۴-۵-۱ استخراج الگوی فازی عبارات احساسی

در اولین گام طراحی دیکشنری، با استفاده از لغات احساسی واژه‌نامه VADER و لیست تشدیدکننده‌های [106] کلیه عبارات احساسی Unigram , Bigram, Trigram که از یک کلمه احساسی یا ترکیب یک کلمه احساسی موجود در واژه‌نامه VADER و تشدیدکننده‌ها تشکیل می‌شد، از پایگاه‌داده آمازون استخراج گردید. این ترکیبات، بهمراه امتیازات جمله یا متنی که این ترکیب در آن وجود داشت به‌صورت یک لیست ذخیره شد.

در غالب مقالات موجود برای محاسبه ضریب تأثیر از ساده‌ترین عملگر ممکن یعنی تقسیم میانگین دو جامعه بر هم بدون توجه به اختلاف تعداد نمونه‌ها استفاده کرده‌اند و از ابهامات و خطای احتمالی در این محاسبه صرف‌نظر می‌کنند. در رویکرد پیشنهادی در این پایان‌نامه برای آنکه دقت در محاسبه ضریب تأثیر تشدیدکننده بیشتر شود از یک رویکرد فازی کمک گرفته شده است.

در رویکرد پیشنهادی ما ابتدا تعداد نمونه‌های جامعه‌ای که قصد مقایسه آن را داریم، یکسان‌سازی شده و سپس از عملگر تقسیم میانگین نمونه‌ها برای محاسبه ضریب تأثیر تشدیدکننده استفاده شده است. فرایند تولید نمونه‌های جدید در سه مرحله صورت می‌پذیرد. ابتدا الگوی احتمالاتی امتیازات استخراج شده از نظرات کاربران برای هر Bigram و Trigram به دست می‌آید. این الگوی به‌دست‌آمده توسط آزمون‌هایی که در زیر به آنها اشاره شده ارزیابی و اعتبارسنجی می‌شوند و سپس بر اساس این الگو نمونه‌های جدید تولید می‌شود. طبیعی است اضافه‌کردن داده‌ها بر اساس الگوی فازی احتمالاتی داده‌های موجود، بهترین و معتبرترین روش

افزودن نمونه محسوب می‌گردد و اعتبار تقسیم و محاسبه ضریب تشدیدکننده را بالا برده و خطای احتمالی رخ‌داده را کاهش خواهد داد. علاوه‌برآن در محاسبه الگو و به‌تبع آن تولید نمونه‌های جدید عملاً پراکندگی نظرات نیز لحاظ می‌گردد که این امر در بالابردن دقت بسیار حائز اهمیت است. لازم به ذکر است در روش محاسبه مرسوم بر اساس میانگین ساده و تقسیم، پراکندگی امتیازات مربوط به نظرات داده شده به‌هیچ‌وجه در محاسبه ضریب تأثیر تشدیدکننده تأثیر نخواهد داشت.

باتوجه‌به توضیح داده شده در مرحله اول باتوجه‌به آنکه امتیازات داده شده به عبارات توسط کاربر انسانی فازی محسوب می‌شود (به آن دلیل که نظرات انسانی از الگوی فازی تبعیت می‌کنند) سعی می‌شود تابع چگالی احتمال نظرات مربوط به Bigramهای شامل هر Unigram به دست آید. این تابع چگالی احتمال علاوه بر میانگین شامل پراکندگی نظرات و امتیازات داده شده به جملات شامل آن عبارت احساسی نیز می‌گردد چرا که تابع چگالی احتمال علاوه بر میانگین که گشتاور اول آماری محسوب می‌شود شامل یک بیان کامل از کلیه خواص متغیر تصادفی نیز هست. درصورتی‌که تعداد نمونه‌های ذخیره شده در لیست مربوط به Trigram یا Bigram از تعداد خاصی کمتر باشد استخراج الگوی آماری دارای خطای زیادی است و عملاً نمی‌توان رابطه معنی‌داری استخراج کرد، لذا در مواردی که تعداد نمونه موجود از یک سطح آستانه مشخص کمتر بوده است (سی نمونه) از میانگین‌گیری ساده استفاده شده و به سراغ استخراج الگوی فازی نرفتیم.

در سیستم‌های دارای تعداد نمونه بالا که در آن نمونه‌ها نقطه‌نظرات انسانی هستند (مشابه کار ما) معمولاً تابع چگالی احتمال گوسی یا نرمال نتایج مناسبی تولید می‌کند. به دلیل کاربرد بالای این تابع چگالی احتمال در جوامع آماری شامل نظرات انسانی، آزمون‌های زیادی برای بررسی صحت فرض الگوی نرمال برای متغیرهای تصادفی موجود پیشنهاد شده‌اند. لذا در این تحقیق به‌صورت پیش‌فرض با تابع چگالی احتمال نرمال، الگوی نمونه‌ها ارزیابی شد و پس از آن آزمون معنی‌داری برازش باهدف بررسی اینکه آیا این الگوی آماری مناسب نمونه‌ها بوده است مورداستفاده قرار گرفت.

این نکته ضروری است که هیچ توزیع آماری نمی‌تواند دقیقاً بر داده‌های مشاهده، برازش خوبی داشته باشد و به این منظور عموماً از سه روش آزمون نیکویی برازش شامل آزمون کای اسکوئر، آزمون کولموگروف–اسمیرنف و آزمون اندرسون–دارلینگ به‌عنوان روش تشخیص مناسب بودن تابع انتخاب شده استفاده می‌شود.

۴-۵-۱-۱ آزمون کای اسکوئر

این آزمون برای بررسی معنی‌داری الگو شدن مجموعه‌ای از داده‌ها با یک تابع توزیع احتمال مورداستفاده قرار می‌گیرد و رابطه آن به‌صورت زیر است.

$$\chi^2 = \sum_{i=1}^{NC} \frac{(f_i - \hat{f}_i)^2}{\hat{f}_i} \qquad (4-5)$$

دررابطه‌بالا f مقدار تابع توزیع احتمال گسسته داده‌های واقعی، $\hat{f}$ مقدار محاسبه شده بر اساس توزیع مورد آزمون، i شماره دسته‌بندی‌ها و NC تعداد دسته‌بندی‌ها است. برای استفاده از این آزمون باید تعداد نمونه‌های موجود در هر بازه حداقل ۵ باشد. χ^2 توزیع خی دو با (NC-NP-1) درجه آزادی است. NP تعداد پارامترهایی است که باید بر اساس توزیع برآورد شود که در تابع چگالی نرمال برابر عدد دو است. پاسخ رابطه بالا با مقدار توزیع کای اسکوئر در سطح معنی داری در نظر گرفته شده (۹۵٪) مقایسه می‌شود و چنانچه مقدار χ^2 به‌دست‌آمده کمتر از مقدار توزیع خی دو بر اساس جدول این توزیع باشد فرض درست بودن برازش در سطح معنی داری فرض شده پذیرفته می‌شود

۴-۵-۱-۲ آزمون کولموگروف – اسمیرنف

آزمون کولموگروف -اسمیرنوف مشابه آزمون خی دو برای بررسی معنی‌داری الگو شدن مجموعه‌ای از داده‌ها با یک تابع توزیع احتمال مورداستفاده قرار می‌گیرد. تفاوت این آزمون با آزمون خی دو در عدم نیاز به حداقل ۵ نمونه در هر بازه است. این آماره به‌صورت زیر محاسبه می‌شود.

$$D = max\left|F(x) - \hat{F}(x)\right| \qquad (4-6)$$

که در آن F تابع توزیع تجمعی داده واقعی و $\hat{F}$ تابع توزیع تجمعی الگوی پیشنهادی است. مراحل انجام این آزمون به شرح زیر است

۱. تابع توزیع تجمعی داده‌ها برای بازه‌های مختلف محاسبه می‌شود.
۲. تابع توزیع تجمعی برای بازه‌های مختلف با استفاده از توزیع آماری تئوری محاسبه می‌شود.
۳. قدر مطلق تفاضل دو تابع در هر بازه محاسبه می‌شود
۴. بیشترین مقدار تفاضل به‌دست‌آمده از مرحله ۳ به‌عنوان D انتخاب می‌شود

با انتخاب مقدار α به‌عنوان میزان خطا و برای تعداد نمونه n مقدار جدول کولموگروف-اسمیرنوف با D مقایسه می‌شود. اگر D از مقدار جدول کمتر باشد آنگاه فرض نیکویی برازش پذیرفته شده و در غیر این صورت رد می‌شود.

۳-۱-۵-۴ آزمون اندرسون – دارلینگ

آزمون اندرسون-دارلینگ نیز مشابه دو آزمون قبل برای بررسی معنی‌داری الگو شدن مجموعه‌ای از داده‌ها با یک تابع توزیع احتمال مورداستفاده قرار می‌گیرد. از نظر ریاضی می‌توان اثبات کرد که آماره اندرسون-دارلینگ بهترین آماره‌ای است که می‌تواند مطابقت توزیع داده‌ها را با توزیع نرمال انجام دهد. به‌منظور بررسی مطابقت توزیع داده‌ها با توزیع نرمال در این آزمون چهار حالت مختلف وجود دارد:

حالت اول: میانگین و واریانس توزیع مشخص است.

حالت دوم: میانگین نامشخص ولی واریانس مشخص است.

حالت سوم: میانگین مشخص ولی واریانس نامشخص است.

حالت چهارم: میانگین و واریانس هر دو نامشخص هستند.

که ما در حالت چهارم قرار داریم.

در این حالت ابتدا میانگین $\bar{X}$ و واریانس $\hat{\sigma}^2$ نمونه‌های موجود محاسبه می‌شود. سپس داده‌ها با استفاده از رابطه $Y_i = \frac{X_i - \bar{X}}{\hat{\sigma}}$ نرمال می‌شوند.

آماره آزمون اندرسون- دارلینگ در نهایت به‌صورت زیر است:

$$A^2 = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)(\ln \Phi(Y_i) + \ln(1 - \Phi(Y_{n+1-i}))) \qquad (4-7)$$

دررابطه‌بالا منظور از Φ تابع توزیع تجمعی نرمال استاندارد است. پس از محاسبه رابطه باید مقدار A2 به‌صورت زیر اصلاح شود.

$$A^{*2} = A^2\left(1 + \frac{4}{n} - \frac{25}{n^2}\right) \qquad (4-8)$$

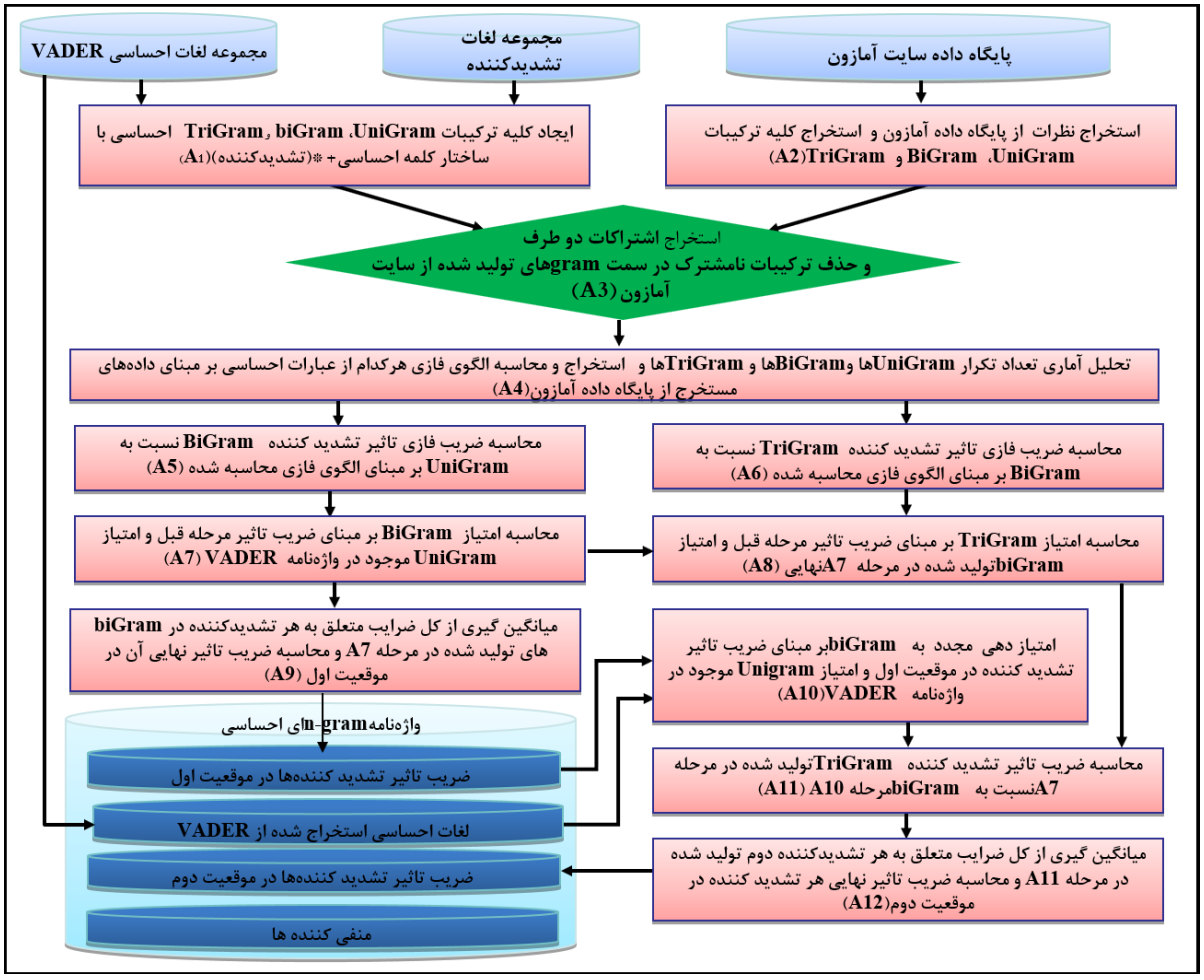
اگر مقدار آماره آزمون در این روش از مقدار بحرانی جدول توزیع بر اساس سطح معنی‌داری در نظر گرفته شده کوچک‌تر باشد، فرض نیکویی برازش تأیید می‌شود.

بر اساس توضیحات فوق، با استفاده از امتیازات ذخیره شده در لیست برای هر ترکیب احساسی ,Bigram Trigram یک توزیع آماری نرمال به دست آمد. نمونه توزیع‌های محاسبه شده بر اساس آزمون نیکویی برازش در سه حالت مورد بررسی قرار گرفتند و در تمامی موارد، آزمون، نیکویی برازش را تأیید کرد.

در مرحله آخر با استفاده از تابع چگالی احتمال به‌دست‌آمده برای متغیرهای فازی که همان امتیازات ذخیره شده در لیست Bigramها و Trigramها بودند می‌توان به‌صورت نامحدود نمونه جدید با همان الگو تولید کرد. در این مرحله این تعداد برابر با تعداد Unigram برای Bigram ها و تعداد Bigram برای Trigramها انتخاب شد. پس از آن این مجموعه نمونه‌های جدید، برای تولید ضرایب اولیه تشدیدکننده‌ها استفاده شد.

با ذکر این نکته که مقالات این حوزه برای تخصیص ضریب یا امتیاز مؤثر برای ترکیب احساسی از عملگرهای آماری ایستا استفاده کرده‌اند و هیچ یک از رویکردهای طراحی دیکشنری بر اساس آخرین جستجوی انجام شده به دنبال تخصیص امتیاز فازی به عبارات احساسی نبوده است و باتوجه‌به آنکه ذهن انسان هنگام محاسبه

قطبیت و احساس متون رویکرد فازی دارد، رویکرد پیشنهادی، ضریب تأثیر دقیق‌تر و نزدیک‌تر به نظرات انسانی را محاسبه خواهد کرد. لازم به ذکر است این مرحله ابتدایی محاسبه ضرایب است و در سیستم نهایی تصمیم‌گیری با یک سطح آستانه ثابت و همانند سیستم‌های کلاسیک انجام خواهد شد.



شکل ۱۲ . فلوچارت روش پیشنهادی اول

در مرحله محاسبه ضرایب در اولین گام با استفاده از الگوی فازی به‌دست‌آمده در تحلیل آماری Unigram, Bigram, Trigramها تأثیر تشدیدکننده موقعیت اول و دوم با تولید نمونه‌های تصادفی با تعداد یکسان برای دو جامعه محاسبه می‌گردد. برای نمونه اگر تعداد Unigram در پایگاه‌داده آمازون برای یک کلمه احساسی ۱۰۰۰ و Bigram شامل همان کلمه احساسی ۷۵۰ باشد با استفاده از تابع تعلق فازی استخراج شده برای Bigram در مرحله قبل، تعدادی نمونه تصادفی تولید می‌شود تا جامعه Bigram از نظر تعداد نمونه با جامعه Unigram مساوی گردد (یعنی ۱۵۰ نمونه جدید Bigram تولید می‌کنیم) (رابطه ۴-۹ و ۴-۱۰).

$\mu_{U_k}, \sigma_{U_k}, N_{U_k}$

$$\mu_{B_{jk}}, \sigma_{B_{jk}}, N_{B_{jk}} \implies PDF\left(\mu_{B_{jk}}, \sigma_{B_{jk}}, N_{B_{jk}}\right) \quad (4-9)$$

$$\mu_{T_{ijk}}, \sigma_{T_{ijk}}, N_{T_{ijk}} \implies PDF\left(\mu_{T_{jk}}, \sigma_{T_{ijk}}, N_{T_{ijk}}\right) \quad (4-10)$$

توضیحات پارامترها در جدول ۱۶ آورده شده است.

جدول ۱۶. نمادهای مربوط به واژه‌نامه احساسی پیشنهادی اول

نماد	توصیف
μ_{U_k}	میانگین آماری کلمه احساسی k ام واژه‌نامه VADER
$\mu_{B_{jk}}$	میانگین آماری Bigram شامل کلمه احساسی k و تشدیدکننده j
$\mu_{T_{ijk}}$	میانگین آماری Trigram شامل کلمه احساسی k و تشدیدکننده موقعیت اول jو تشدیدکننده موقعیت دوم i
$\tilde{\mu}_{B_{jk}}$	میانگین فازی Bigram شامل کلمه احساسی k و تشدیدکننده j
$\tilde{\mu}_{T_{ijk}}$	میانگین فازی Trigram شامل کلمه احساسی k و تشدیدکننده موقعیت اول jو تشدیدکننده موقعیت دوم i
σ_{U_k}	واریانس آماری Unigram شامل کلمه احساسی k
$\sigma_{B_{jk}}$	واریانس آماری Bigram شامل کلمه احساسی k و تشدیدکننده j
$\tilde{\sigma}_{B_{jk}}$	واریانس آماری Bigram جامعه جدید تولید شده بر اساس الگوی فازی کلمه احساسی k و تشدیدکننده j
$\sigma_{T_{ijk}}$	واریانس آماری Trigram شامل کلمه احساسی k و تشدیدکننده موقعیت اول jو تشدیدکننده موقعیت دوم i
$\tilde{\sigma}_{T_{jk}}$	واریانس آماری Trigram جامعه جدید تولید شده بر اساس الگوی فازی کلمه احساسی k و تشدیدکننده موقعیت اول jو تشدیدکننده موقعیت دوم i
N_{U_k}	تعداد Unigram احساسی k ام از واژه‌نامه VADER
$N_{B_{jk}}$	تعداد Bigram احساسی شامل کلمه احساسی k و تشدیدکننده موقعیت اول j
$N_{T_{ijk}}$	تعداد Trigram شامل کلمه احساسی k و تشدیدکننده موقعیت اول jو تشدیدکننده موقعیت دوم i
$CO_{I_{jk}}$	ضریب تشدیدکننده jدر هنگام قرارگیری پیش از کلمه احساسی K در Bigram
$CO_{I_{ijk}}$	ضریب تشدیدکننده در هنگام قرارگیری پیش از کلمه احساسی K و تشدیدکننده jدر Trigram
So_{VADER_k}	امتیاز کلمه احساسی k ام از واژه‌نامه VADER
$So_{B_{jk}}$	امتیاز اولیه Bigram شامل کلمه احساسی k و تشدیدکننده j
$So_{T_{ijk}}$	امتیاز Trigram شامل کلمه احساسی k و تشدیدکننده موقعیت دوم jو تشدیدکننده موقعیت اول i
$So_{BN_{jk}}$	امتیاز نهایی Bigram شامل کلمه احساسی k و تشدیدکننده j
$CO_{I_{jk}}$	ضریب تشدیدکننده موقعیت اول jدر هنگام قرارگیری پیش از کلمه احساسی k
$CO_{I_{ijk}}$	ضریب تشدیدکننده موقعیت دوم i در هنگام قرارگیری پیش از کلمه احساسی k و تشدیدکننده موقعیت اول j

۴-۵-۲ محاسبه ضریب تأثیر تشدیدکننده‌ها

بعد از یکسان نمودن تعداد نمونه‌ها با استفاده از الگوی فازی بیان شده در بخش قبل، در این مرحله با استفاده از میانگین‌گیری ساده برای این نمونه‌های جدید، امتیاز Bigram محاسبه و با تقسیم عدد حاصل از میانگین‌گیری به میانگین حاصل از جامعه Unigram ضریب تأثیر اولیه تشدیدکننده موقعیت اول موجود در عبارت احساسی Bigram محاسبه می‌گردد. (روابط ۴-۱۱ و ۴-۱۲).

این امر به طور مشابه با استفاده از میانگین حاصل از Trigram وBigram برای محاسبه ضریب تأثیر تشدیدکننده موقعیت دوم هم انجام می‌شود.

$$\begin{matrix}(\mu_{U_k}, & \sigma_{U_k}, & N_{U_k})\\ (\tilde{\mu}_{B_{jk}}, & \tilde{\sigma}_{B_{jk}}, & N_{U_k})\\ (\tilde{\mu}_{T_{ijk}}, & \tilde{\sigma}_{T_{ijk}}, & N_{U_k})\end{matrix} \Longrightarrow \begin{cases} CO_{I_{jk}} = \dfrac{\tilde{\mu}_{B_{jk}}}{\mu_{U_k}} & (4-11)\\ CO_{I_{ijk}} = \dfrac{\tilde{\mu}_{T_{ijk}}}{\tilde{\mu}_{B_{jk}}} & (4-12)\end{cases}$$

سپس برای آنکه پاسخ به داده‌های انسانی نزدیک‌تر و از هرگونه خطای احتمالی در آن اجتناب شود ضریب فازی تأثیر تشدیدکننده محاسبه شده به‌ازای هر کلمه احساسی k در رابطه ۴-۱۱ که بر مبنای پایگاه‌داده آمازون محاسبه شده است، در امتیاز کلمه احساسی k که از دیکشنری VADER استخراج شده ضرب گردیده است. به‌عبارت‌دیگر امتیاز Bigram با استفاده از امتیاز موجود در واژه‌نامه VADER بروز رسانی شد و به‌این‌ترتیب خطای احتمالی موجود در محاسبه امتیازات مستخرج از پایگاه‌داده یک مرحله کاهش داده شده است. با استفاده از این روش می‌توان اطمینان داشت، امتیاز به‌دست‌آمده برای با Bigram به امتیاز واقعی عبارت احساسی نسبت به حالتی که از پایگاه‌داده آمازون مستقیماً استخراج می‌شد نزدیک‌تر است (رابطه ۴-۱۳).

$$So_{B_{jk}} = CO_{I_{jk}} * P_{VADER_k} \qquad (4-13)$$

در مرحله بعد با استفاده از قواعد منطقی تعریف شده در جدول ۱۷ عملیات پالایش امتیازات صورت می‌پذیرد. در طی عملیات پالایش برخی از امتیازات استخراج شده از پایگاه‌داده آمازون حذف شده تا به‌این‌ترتیب عملیات تصحیح ضرایب به‌دست‌آمده صورت پذیرد. در مرحله نخست برای تصحیح Bigramها

چهار شرط و دو حالت در نظر گرفته شده است. و در صورت عدم موفقیت این شرایط تابع refiner برای تصحیح عملیات فراخوانی می‌شود. به‌عنوان‌مثال در یک Bigram که از یک کلمه احساسی مثبت و تضعیف‌کننده تشکیل شده است انتظار داریم امتیاز حاصل کوچک‌تر از امتیاز کلمه احساسی باشد چنانچه نتایج به‌دست‌آمده این انتظار را برآورد ننماید تابع پالایش فراخوانی می‌گردد. در طی فرایند پالایش فرض می‌شود که کلیه مقادیر بالا ناشی از وجود تعصب مشتری بوده و یا به علت طولانی بودن جمله در نتیجه ترکیب با کلمات احساسی دیگر حاصل شده است -شکل ۸ نمونه ای از یک نظر استخراج شده در پایگاه‌داده آمازون را نشان می‌دهد- لذا این امتیازات بصورت گام به گام از فرایند محاسبه حذف می‌گردد. این فرایند برای حالت دوم نیز صورت پذیرفته و در طی فرایند پالایش مشاهدات مغرضانه (خیلی زیاد یا خیلی کم) حذف می‌گردد. فرایند بررسی و تصحیح تا برآورده شدن شرایط تکرار می‌گردد و در نهایت ضریب اصلاح شده حاصل برای مرحله بعد تولید شده و یا در صورت عدم برآورده شدن شرایط موردنیاز عبارت احساسی از مجموعه حذف می‌گردد. جدول ۱۷ حالات مختلف مورداستفاده جهت بررسی و تأیید امتیازات استخراج شده برای اصطلاحات Bigram حاوی یک کلمه احساسی و تشدیدکننده را نشان می‌دهد.

جدول ۱۷. جدول حالات مختلف بررسی و تأیید امتیازات استخراج شده برای اصطلاحات Bigram

حالات اول			
تشدیدکننده	**قطبیت کلمه احساسی**	**تشدیدکننده**	**قطبیت کلمه احساسی**
تقویت‌کننده	منفی	تضعیف کننده	مثبت
شرط ۱		شرط ۲	
حالت دوم			
تشدیدکننده	**قطبیت کلمه احساسی**	**تشدیدکننده**	**قطبیت کلمه احساسی**
تقویت‌کننده	مثبت	تضعیف کننده	منفی
شرط ۱		شرط ۲	

همچنین شکل ۱۳ بخشی از شبه کد مورداستفاده جهت عملیات اعتبارسنجی و تصحیح را نشان می‌دهد. پس از پایان مرحله اعتبار سنجی و تصحیح با استفاده از امتیاز محاسبه شده در رابطه ۴-۱۳ و ضریب به‌دست‌آمده در رابطه ۴-۱۲ امتیاز اولیه Trigramها مجدداً محاسبه می‌شود. (رابطه ۴-۱۴).

$$So_{T_{ijk}} = CO_{I_{ijk}} * So_{B_{jk}} \qquad (4-14)$$

در این مرحله نیز مجدداً عملیات اعتبارسنجی و تصحیح همانند مرحله قبل صورت می‌پذیرد تا به‌این‌ترتیب کلیه خطاهای احتمالی حذف و اصلاح گردد. امتیازات به‌دست‌آمده در رابطه ۴-۱۳، برای کلمات احساسی

حالت اول

if (($So_{VADER_k} <$ ۰ *AND Intensifire IS* Reinforcers) *OR* ($So_{VADER_k} >$ ۰ *AND Intensifire IS* downtoner)) {

case=1;

while ($So_{VADER_k} < So_{B_{jk}}$) {

$So_{B_{jk}} = Refiner(Bigram, So_{VADER_k}, So_{B_{jk}}$, HighestRating, case);

HighestRating=HighestRating-1;

}

}

حالت دوم

if (($So_{VADER_k} >$ ۰ *AND Intensifire IS* Reinforcers) *OR* ($So_{VADER_k} <$ ۰ *AND Intensifire IS* downtoner)) {

case=۲;

while ($So_{VADER_k} < So_{B_{jk}}$) {

$So_{B_{jk}} = Refiner(Bigram, So_{VADER_k}, So_{B_{jk}}$, LowestRating, case);

LowestRating= LowestRating+ ۱;

}

}

شکل ۱۳. شبه کد مربوط به فرآیند بررسی اعتبار امتیازات Bigram

مختلف به‌ازای هر تشدیدکننده میانگین‌گیری می‌شود تا ضریب نهایی تشدیدکننده موقعیت اول طبق رابطه ۴-۱۵ به دست آید.

$$\forall j \longrightarrow Co_{IN1_j} = avg\left(Co_{I_{jk}}\right) \qquad (4-15)$$

با مشخص‌شدن ضریب نهایی تأثیر تشدیدکننده موقعیت اول مجدداً امتیاز Bigramها به‌روزرسانی شده و امتیاز جدید برای استخراج ضریب تأثیر تشدیدکننده Trigram مدنظر قرار گرفته و مجدداً ضریب تأثیر تشدیدکننده‌ها در موقعیت دوم طبق رابطه ۴-۱۷ و ۴-۱۸ محاسبه می‌گردد.

$$So_{BN_{jk}} = CO_{IN1_j} * P_{VADER_k} \qquad (4-16)$$

$$CO_{I_{ijk}} = \frac{So_{T_{ijk}}}{So_{BN_{jk}}} \qquad (4-17)$$

$$\forall i \longrightarrow Co_{IN2_i} = avg\left(Co_{I_{ijk}}\right) \qquad (4-18)$$

در نهایت ضرایب نهایی تشدیدکننده‌ها در موقعیت اول و دوم ترکیبی از منطق فازی، روش‌های آماری، پایگاه‌داده آمازون و واژه‌نامه دستی VADER بوده است. باتوجه‌به نتایج به‌دست‌آمده، ترکیب مراحل فوق در عمل می‌تواند خطای احتمالی در محاسبه ضریب تأثیر تشدیدکننده‌ها را به کمترین حد ممکن برساند. در نهایت ضرایب تأثیر به‌دست‌آمده برای تشدیدکننده موقعیت اول و دوم در دیکشنری نهایی ذخیره و دیکشنری نهایی شامل کلمات احساسی، ضریب تأثیر تشدیدکننده در موقعیت اول و موقعیت دوم تولید می‌گردد.

$$So_{U_k}=So_{VADER_k} \qquad (4-19)$$

$$So_{BN_{jk}} = Co_{IN1_j} * So_{VADER_k} \qquad (4-20)$$

$$So_{TN_{ijk}} = Co_{IN2_i} * Co_{IN1_j} * So_{VADER_k} \qquad (4-21)$$

باتوجه‌به روابط ۴-۱۹ تا ۴-۲۱ و امتیازات "enjoy" و "bad" در واژه‌نامه احساسی VADER که به ترتیب 2.3 و 2.5- هستند و با استفاده از جدول ۱۸ می‌توان امتیازات عبارات احساسی "very big bad" و " very deeply enjoyed" را به‌صورت زیر محاسبه نمود.

جدول ۱۸. ضرایب محاسبه شده برای برخی از تشدیدکننده‌ها باتوجه‌به موقعیت آنها در عبارت احساسی در روش پیشنهادی اول

intensifier	Second position	First position
deeply	1.805689	2.490525
very	1.452547	2.152212
big	1.200473	2.086358

very big bad=1.45*2.08*(-2.5)=-7.54

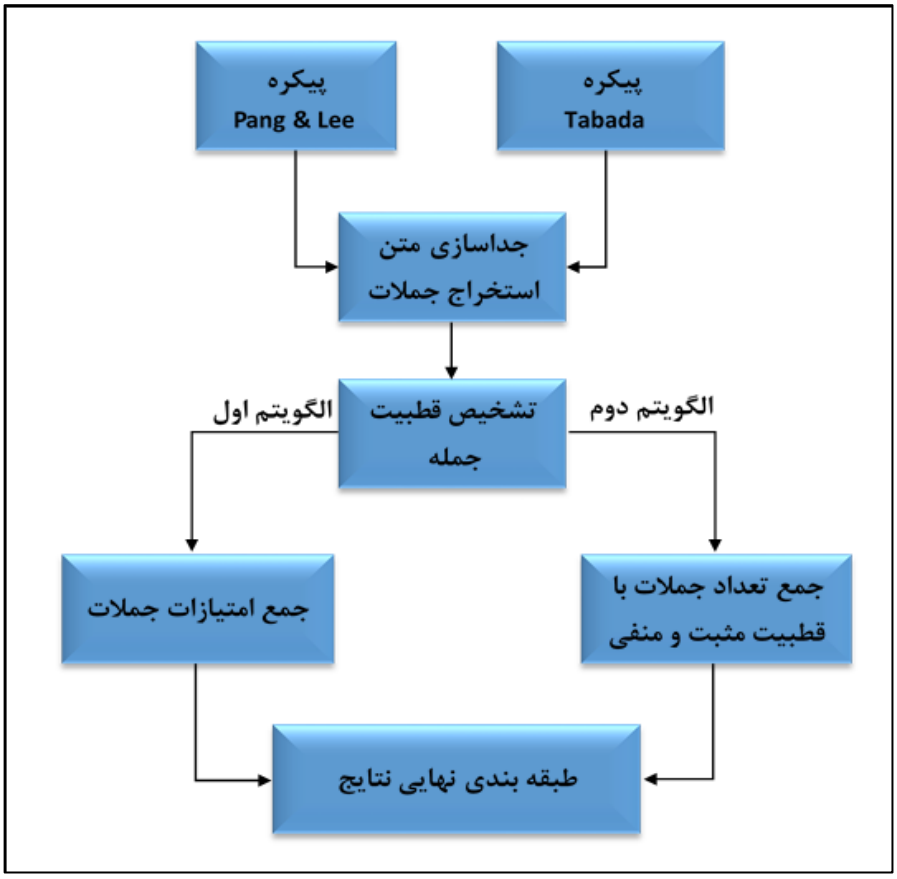
very deeply enjoyed=1.45*2.5*2.3=8.34

۳-۵-۴ الگوریتم‌های مورداستفاده برای ارزیابی کارایی روش‌های پیشنهادی

توجه به این نکته ضروری است که طراحی یک دیکشنری احساسی فرایندی کاملاً مستقل از طراحی الگوریتم‌های تشخیص احساس است. در طی فرایند طراحی دیکشنری احساسی هدف، استخراج کلمات احساسی، تعیین قطبیت و تخصیص شدت احساسی به این کلمات و در مراحل پیشرفته‌تر تخصیص ضرایب مناسب به تشدیدکننده‌ها و منفی‌کننده‌های موجود است. درحالی‌که الگوریتم‌های ارزیابی ضمن استفاده از داده‌های موجود در این قبیل از دیکشنری‌ها با استفاده از روابط و قوانین نحوی، معنایی و یا وابستگی بین کلمات و برخی از قوانین اکتشافی اقدام به ارزیابی متون احساسی در سطوح مختلفی همچون سطح جمله یا سند می‌نمایند

در این رساله به منظور بررسی کارایی دیکشنری طراحی شده نسبت به آخرین دیکشنری‌های موجود از دو الگوریتم طبقه‌بندی سند مبتنی بر سطح-جمله استفاده شده است. الگوریتم‌ها به نحوی انتخاب شده‌اند تا تأثیر توجه به اثر تشدیدکننده‌ها را به‌صورت واضح آشکار نمایند و از بررسی موارد دیگری همچون تأثیر بزرگ‌نویسی کلمات احساسی، بررسی اثر کلمه But در جملات و یا بررسی بلوک‌های irrealise و موارد موثر در تعیین قطبیت یک سند احساسی صرف‌نظر شده است. در الگوریتم نخست[153] کلمات احساسی و عبارات

احساسی هر جمله به تفکیک استخراج و میانگین امتیاز عبارات احساسی موجود در هر جمله محاسبه می‌شود. در این محاسبه ضریب تأثیر ذخیره شده در دیکشنری پیشنهادی برای تشدیدکننده‌ها و امتیازات موجود در

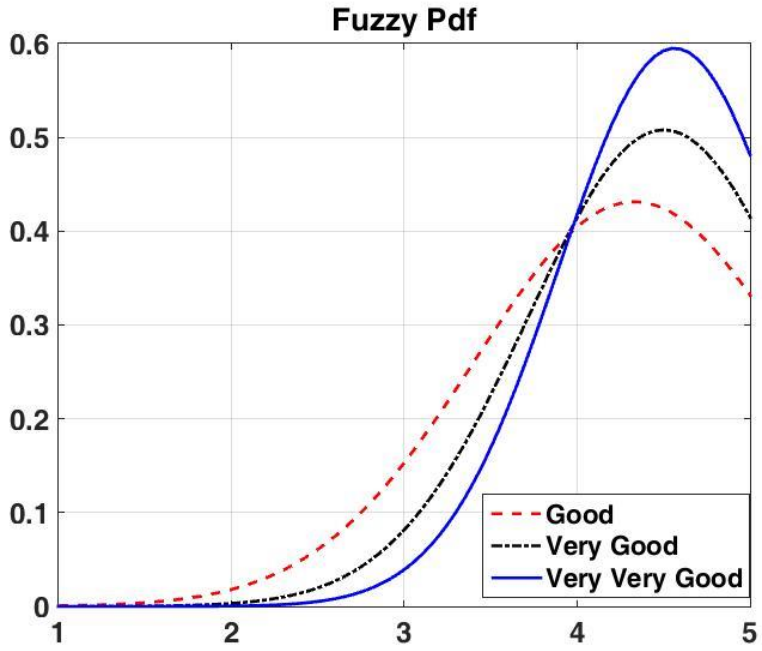


شکل ۱۴. فرآیند تعیین قطبیت اسناد

واژه‌نامه VADER مدنظر بوده است. پس از این مرحله میانگین امتیازات عبارات احساسی موجود در هر جمله به‌عنوان امتیاز هر جمله محاسبه شده و در نهایت مجموع امتیاز این جملات قطبیت سند مورد ارزیابی را مشخص می نماید. در الگوریتم پیشنهادی در [2] نیز مشابه الگوریتم نخست کلمات و عبارات احساسی هر جمله به تفکیک استخراج و قطبیت هر جمله محاسبه می‌شود. پس از این مرحله به‌جای جمع کردن میانگین‌ها با هم، تعداد جملات مثبت و منفی شمارش و تعداد بالاتر، معیار قطبیت سند قرار می‌گیرد به عبارت ساده‌تر در این مرحله رای‌گیری ملاک عمل خواهد بود. در حالت برابر بودن تعداد جملات مثبت و منفی، قطبیت سند، منفی ارزیابی می‌شود.

۶-۴ نتایج

بخش نتایج به دو قسمت مختلف تقسیم شده است. در بخش اول نتایج حاصل از تابع چگالی احتمال فازی و آزمون‌های آماری فازی در محاسبه امتیازات اولیه، برای چند Ngram نمونه آورده شده و کارایی الگوی فازی به‌دست‌آمده و معنی‌دار بودن توابع چگالی احتمال فازی نشان‌داده‌شده است.



شکل ۱۵ . نمودار فازی برای سه عبارت good، Very good و Very very good

در بخش دوم بر اساس معیارهای مختلف درستی مانند Recall, precision-measure, accuracy دیکشنری پیشنهادی در برابر آخرین روش‌های موجود بررسی شده و نشان داده شد روش پیشنهادی در غالب معیارها نسبت به آخرین دیکشنری‌های این حوزه عملکرد بسیار مطلوب‌تری از خود نشان داده است.

۷-۴ الگوی فازی

در مرحله اول بر اساس الگوی آماری به‌دست‌آمده و تعداد نمونه‌های موجود در Unigram، Bigram و Trigram باید الگوی فازی Bigram, Trigram به‌دست‌آمده و سپس با تولید تعداد نمونه کافی و یکسان‌سازی

نمونه‌ها مجدداً پارامترهای ارزیابی Bigram, Trigram محاسبه شود (روابط ۱-۴ و ۲-۴). علاوه بر این انحراف معیار آنها نیز محاسبه شد، تا بر اساس تحلیل صورت‌گرفته درستی و صحت نتایج مشخص شود.

در شکل ۱۵ الگوی فازی محاسبه شده برای ترکیبات good, very good, very very good نشان‌داده‌شده است. خروجی الگو باتوجه‌به امتیازدهی واقعی جملات در پایگاه‌داده آمازون بین بازه ۱ تا ۵ محدود شده است. همان‌طور که در شکل ۱۵ می‌توان دید در Bigram به دلیل نزدیکی امتیازات موجود در جملات، انحراف معیار نسبت به Unigram کاهش پیدا کرده است.

علاوه‌برآن به دلیل نزدیکی امتیاز جملاتی که در آنها Trigram مدنظر بوده است، انحراف معیار نسبت به Bigram کاهش پیدا کرده است. برای نشان‌دادن تفاوت هنگام یکسان‌سازی تعداد نمونه در میانگین به‌دست‌آمده آماری و فازی، در جدول ۱۹ میانگین آماری اولیه و میانگین فازی متناظر توزیع پس از یکسان-سازی نمونه‌ها آورده شده تا تغییر صورت‌گرفته نسبت به وضعیت مرسوم نشان داده شود.

جدول ۱۹. نتایج آماری و فازی یک نمونه Unigram, Bigram, Trigram

Sentiment phrase	mean		std
	statistical	fuzzy	
Good	4.3306	4.3306	0.925
Very good	4.5013	4.3992	0.785
Very very good	4.5634	4.4776	0.670

بر اساس نتایج جدول ۱۹ می‌توان دید که در حدود ۲ تا ۵ درصد، ضریب یا میانگین به‌دست‌آمده تغییر داشته که حکایت از بهتر شدن پاسخ و نزدیک شدن جواب الگوی فازی به جواب واقعی است. پس از اینکه میانگین بر اساس تولید نمونه کافی با الگوی فازی متناظر در Bigram, Trigram اصلاح شد، ضریب تأثیر اولیه محاسبه می‌شود.

با این کار عملاً نسبت به الگوی ساده آماری ضرایب تأثیر اولیه اصلاح شده خطای کمتری داشته و از نظر تعداد نمونه نیز یکسان‌سازی شده‌اند. این مرحله معادل روابط ۹-۴ و ۱۰-۴ محسوب می‌شود و پاسخ

به‌دست‌آمده در ستون دوم و سوم جدول ۱۹ معادل میانگین و انحراف معیار نمونه‌ها یعنی $\tilde{\mu}_{T_{jk}}$ ، $\tilde{\mu}_{B_{jk}}$ ، $\tilde{\sigma}_{B_{jk}}$ ، $\tilde{\sigma}_{T_{jk}}$ در روابط ۹-۴ و ۱۰-۴ خواهد بود.

۸-۴ ارزیابی دیکشنری پیشنهادی

در این مرحله دیکشنری پیشنهادی با ۲ دیکشنری نسبتاً جدید و متداول در این حوزه شامل VADER و Senti-N-gram مقایسه شد. مقایسه در دو پایگاه‌داده Taboada و Pang & Lee که در بخش ۴-۳ راجع به آنها توضیح داده شده صورت گرفت تا کارایی واژه‌نامه پیشنهادی نسبت به دو واژه‌نامه معرفی شده در ۳-۲-۵-۳ و ۳-۵-۲-۹ نشان داده شود.

برای آنکه شرایط مقایسه کاملاً یکسان باشد، پایگاه‌داده‌های استاندارد و معیار مقایسه مشابه برای هر سه واژه‌نامه انتخاب شده است. در زمینه روش‌های تخصیص امتیاز، امتیازات واژه‌نامه VADER به دلیل دستی بودن، ثابت بوده اما در واژه‌نامه Senti-N-gram نمونه‌ها مشابه با روش پیشنهادی در نظر گرفته شده تا کارایی الگوریتم پیشنهادی در شرایط کاملاً مشابه با روش استفاده شده در واژه‌نامه Senti-N-gram مقایسه شود. دو دیکشنری VADER, Senti-N-gram و دیکشنری پیشنهادی با هر دو الگوریتم تشخیص قطبیت سند که در فصل ۴-۵-۳ توضیح داده شد ارزیابی گردید تا شرایط مقایسه با این دیکشنری‌ها کاملاً یکسان باشد و از نسخه‌های اصلاح شده، توابع دلخواه جریمه یا برخی تغییرات هنگام ارزیابی نهایی قطبیت سند استفاده نشده است.

جدول ۲۰ درصد درستی حاصل از اعمال دو الگوریتم طبقه‌بندی اسناد در سطح جمله با استفاده از دو دیکشنری احساسی VADER و Senti-N-gram و دیکشنری ساخته شده در روش پیشنهاد اول را بر روی دو پایگاه‌داده معرفی شده در بخش ۴-۳ نشان می‌دهد.

جدول ۲۰. مقایسه پارامتر دقت بر مبنای دوروش طبقه‌بندی اسناد در سطح جمله با استفاده از واژه‌نامه‌های VADER، Senti-N-gram و واژه‌نامه پیشنهادی اول

Datasets	طبقه‌بندی اسناد در سطح جمله «الگوریتم اول»			طبقه‌بندی اسناد در سطح جمله «الگوریتم دوم»		
	VADER	Senti-N-gram	روش پیشنهادی اول	VADER	Senti-N-gram	روش پیشنهادی اول
Epinions	62	61.75	**67.75**	63.5	63.75	**69**
Movie review	63.5	63.75	**67.05**	63.95	63.8	**66.35**

روش پیشنهادی ما نسبت به‌تمامی دیکشنری‌های مورد ارزیابی به طور میانگین ۴.۱۸ ٪ از نظر درستی، پاسخ بهتری در هر دو پایگاه‌داده از خود نشان داده است. درصورتی‌که افزایش درصد درستی به‌صورت نسبی بیان شود باتوجه‌به بهترین درصدهای موجود روش‌های گذشته، تقریباً ۹.۳ درصد بهبود نسبی داشته‌ایم که بسیار مناسب محسوب می‌گردد.

در جدول ۲۱ درصد درستی اسناد منفی و مثبت به تفکیک حوزه گزارش شده تا کارایی دیکشنری‌ها به تفکیک حوزه‌ها نشان داده شود. همان‌طور که می‌توان دید روش پیشنهادی در تمامی موارد نسبت به روش‌های

جدول ۲۱. مقایسه پارامتر دقت بر مبنای روش طبقه‌بندی اسناد در سطح جمله[153] **«الگوریتم اول» با استفاده از واژه‌نامه‌های VADER،Senti-N-gram و واژه‌نامه پیشنهادی نخست در حوزه‌های مختلف**

Datasets	VADER			Senti-N-gram			proposed		
	Positive	Negative	overall	Positive	Negative	overall	Positive	Negative	overall
BOOKS	**100**	36	68	92	32	62	88	**56**	**72**
CARS	**100**	28	64	**100**	**32**	**66**	**100**	**32**	**66**
COMPUTERS	**100**	32	66	**100**	36	68	**100**	**44**	**72**
COOKWARE	**100**	**16**	**58**	**100**	12	56	**100**	**16**	**58**
HOTELS	**100**	8	54	**100**	20	60	**100**	**36**	**68**
MOVIES	**96**	**32**	**64**	92	28	60	**96**	**32**	**64**
MUSIC	92	24	58	92	20	56	**96**	**32**	**64**
PHONES	92	36	64	92	40	66	**100**	**56**	**78**

جدول ۲۲. مقایسه پارامترهای Precision، Recall و F-measure برمبنای روش طبقه‌بندی اسناد در سطح جمله[153] «الگوریتم اول» با استفاده از واژه‌نامه‌های VADER،Senti-N-gram و واژه‌نامه پیشنهادی نخست در حوزه‌های مختلف

Datasets	VADER			Senti-N-gram			الگوریتم پیشنهادی ۱		
	P	R	F	P	R	F	P	R	F
BOOKS	0.61	**1**	**0.76**	0.58	0.92	0.71	**0.67**	0.88	**0.76**
CARS	0.58	**1**	0.74	**0.60**	**1**	**0.75**	**0.60**	**1**	**0.75**
COMPUTERS	0.60	**1**	0.75	0.61	**1**	0.76	**0.64**	**1**	**0.78**
COOKWARE	0.54	**1**	**0.70**	0.53	**1**	0.69	**0.54**	**1**	**0.70**
HOTELS	0.52	**1**	0.69	0.56	**1**	0.71	**0.61**	**1**	**0.76**
MOVIES	0.59	**0.96**	**0.73**	0.56	0.92	0.70	**0.59**	**0.96**	**0.73**
MUSIC	0.55	0.92	0.69	0.54	0.92	0.68	**0.59**	**0.96**	**0.73**
PHONES	0.59	0.92	0.72	0.61	0.92	0.73	**0.69**	**1**	**0.82**

پیشین خصوصاً در حوزه تشخیص قطبیت اسناد منفی که پیچیدگیهای بیشتری را نیز باخود دارد کارایی مناسب‌تری از خود نشان می‌دهد. در بهترین حالت شاهد بهبود ۳۵۰ ٪ نتایج نسبت به دیکشنری VADER در تشخیص قطبیت اسناد منفی مرتبط با حوزه HOTELS هستیم . نتایج حاصل از این آزمایش تقریباً ۹٪ بهبود نسبی در کل نتایج و در حوزه اسناد منفی میانگین بهبود نسبی معادل ۳۸.۲٪ را در ارزیابی‌های انجام شده بر مبنای الگوریتم نخست نشان می‌دهد. بیشترین افزایش در حوزه PHONES با ۱۸٪ بهبود نسبی و ۴۰٪ بهبود در حوزه اسناد منفی بوده است.

در جدول ۲۲ برای تکمیل مقایسه مقادیر Precision, Recall, F-measure در سه دیکشنری مقایسه شده است. همان‌طور که قابل مشاهده است مقدار پارامتر Precision در تمامی حوزه‌ها -به جز سه حوزه CARS و COOKWARE و MOVIES که ثابت و بدون تغییر مانده در بقیه حوزه‌ها افزایش از خود نشان داده است.

همان‌طور که در جدول ۲۳ نشان‌داده‌شده است، واژه‌نامه پیشنهادی تعداد تشخیص اسناد منفی را در شش حوزه افزایش می‌دهد. همچنین خطای موجود در اسناد مثبت را در حوزه‌های MOVIES و MUSIC به صفر

جدول ۲۳. مقایسه پارامتر دقت بر مبنای روش طبقه‌بندی اسناد در سطح جمله[2] «الگوریتم دوم» با استفاده از واژه‌نامه‌های VADER،Senti-N-gram و واژه‌نامه پیشنهادی نخست در حوزه‌های مختلف

Datasets	VADER			Senti-N-gram			proposed		
	Positive	Negative	overall	Positive	Negative	overall	Positive	Negative	overall
BOOKS	**100**	36	68	96	40	68	96	**48**	**72**
CARS	**100**	28	64	**100**	32	**68**	**100**	**36**	**68**
COMPUTERS	**96**	**40**	**68**	96	36	66	92	**40**	66
COOKWARE	**100**	12	56	**100**	**20**	**60**	**100**	**20**	**60**
HOTELS	**100**	20	60	**100**	16	58	**100**	**28**	**64**
MOVIES	92	36	64	96	24	60	**100**	**40**	**70**
MUSIC	96	20	58	92	20	56	**100**	**36**	**68**
PHONES	**96**	44	70	92	60	76	92	**76**	**84**

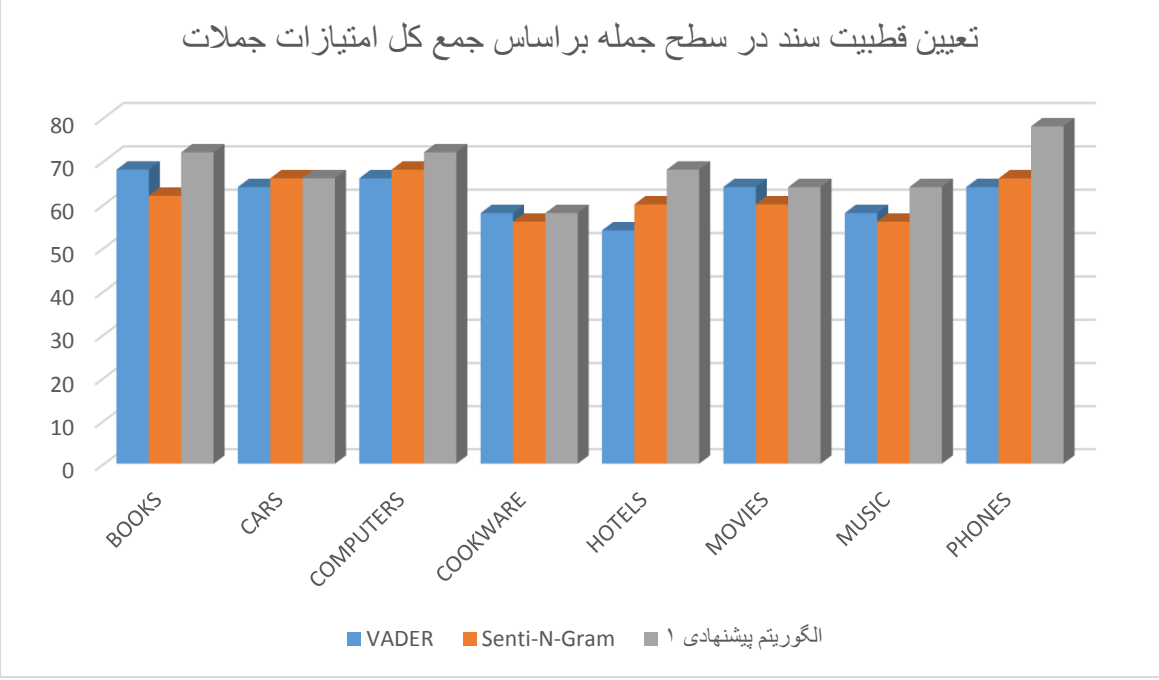
می‌رساند. هرچند که برخی از نتایج متفاوت است بااین‌حال مشابه روش قبلی بیشترین پیشرفت در حوزه PHONES مشاهده می‌شود. به‌عنوان‌مثال، در حوزه PHONES از واژه‌نامه VADER این روش در واژه‌نامه پیشنهادی ۳۲ افزایش در طبقه‌بندی اسناد منفی را نشان می‌دهد درحالی‌که روش طبقه بندی اول فقط ۲۰ افزایش را نشان می‌داد. در این جدول ، روش پیشنهادی نخست به ترتیب در حوزه‌های HOTELS ، MUSIC و BOOKS در اسناد منفی ۸۰ ، ۴۰ و ۲۰ درصد باعث افزایش دقت شده است.

طبق نتایج جدول ۲۴، پارامترهای F-measure و precision روش پیشنهادی در همه حوزه‌ها به جز حوزه COMPUTERS بهبود عملکرد را از خود نشان می‌دهد. در نتیجه کلی ، به طور متوسط روش پیشنهادی در معیار precision به ترتیب ۴٪ و ۴.۶٪ بهبود نتایج را نسبت به دو واژهنامه Senti-n-gram و VADER و در معیار F-measure به طور متوسط ۳.۲درصد بهبود نتایج را نسبت به دو واژه‌نامه نشان می‌دهد. در طبقه بندی اسناد مثبت (معیار) میانگین نتایج واژه‌نامه پیشنهادی مشابه دو واژه‌نامه قبلی است.

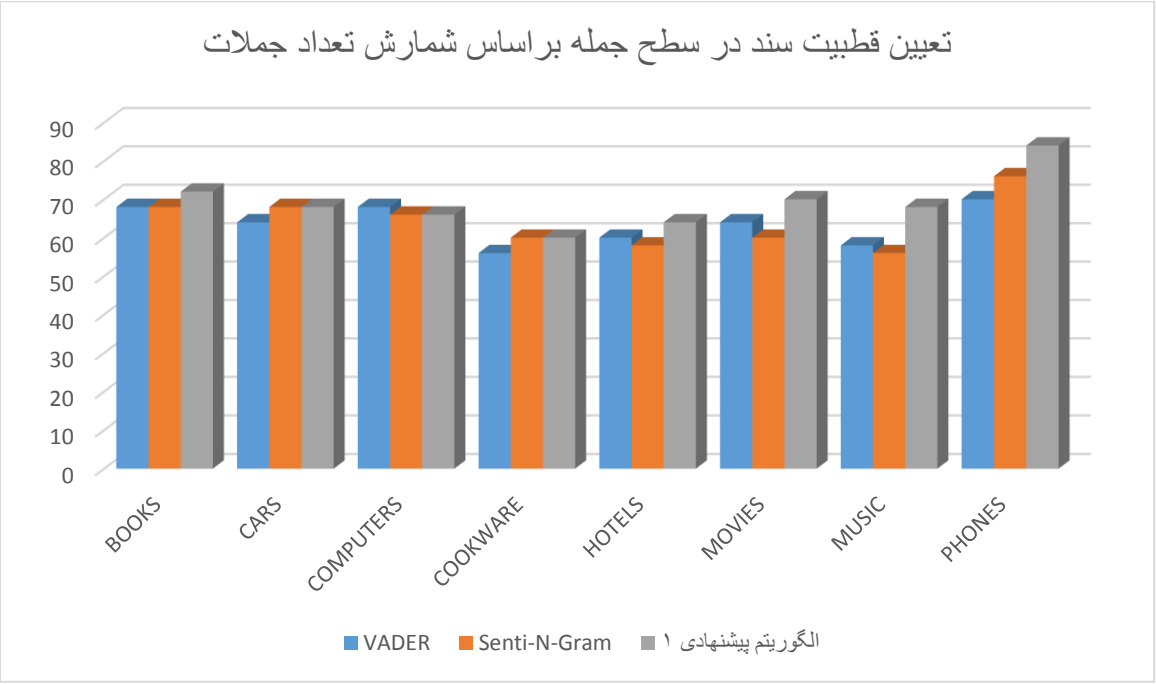
جدول ۲۴. مقایسه پارامترهای Precision، Recall و F-measure برمبنای روش طبقه‌بندی اسناد در سطح جمله[2] «الگوریتم دوم» با استفاده از واژه‌نامه‌های VADER،Senti-N-gram و واژه‌نامه پیشنهادی نخست در حوزه‌های مختلف

	VADER			Senti-N-gram			proposed		
	P	**R**	**F**	**P**	**R**	**F**	**P**	**R**	**F**
BOOKS	0.61	**1.00**	0.76	0.62	0.96	0.75	**0.65**	0.96	**0.77**
CARS	0.58	1.00	0.74	0.60	1.00	0.75	**0.61**	**1.00**	**0.76**
COMPUTERS	**0.62**	0.96	**0.75**	0.60	**0.96**	0.74	0.61	0.92	0.73
COOKWARE	0.53	1.00	0.69	0.56	1.00	0.71	0.56	**1.00**	**0.71**
HOTELS	0.56	1.00	0.71	0.54	1.00	0.70	**0.58**	**1.00**	**0.74**
MOVIES	0.59	0.92	0.72	0.56	0.96	0.71	**0.63**	**1.00**	**0.77**
MUSIC	0.55	0.96	0.70	0.54	0.92	0.68	**0.61**	**1.00**	**0.76**
PHONES	0.63	**0.96**	0.76	0.70	0.92	0.79	**0.79**	0.92	**0.85**

دو شکل ۱۶ و ۱۷ نمودار ستونی پارامتر دقت بر مبنای دو الگوریتم جمع کل امتیازات جملات و همچنین الگوریتم شمارش تعداد جملات (رای اکثریت) را نشان می‌دهد



شکل ۱۶.مقایسه پارامتر دقت بر مبنای روش طبقه‌بندی اسناد در سطح جمله **[153]** «الگوریتم اول» با استفاده از واژه‌نامه‌های VADER،Senti-N-gram و واژه‌نامه پیشنهادی نخست در حوزه‌های مختلف



شکل ۱۷.مقایسه پارامتر دقت بر مبنای روش طبقه‌بندی اسناد در سطح جمله[2] «الگوریتم دوم» با استفاده از واژه‌نامه‌های Senti-N-gram،VADER و واژه‌نامه پیشنهادی نخست در حوزه‌های مختلف

۹-۴ جمع‌بندی و خاتمه

در این فصل یک واژه‌نامه مبتنی بر امتیازدهی ترکیبی برای بالابردن کارایی واژه‌نامه‌های موجود در حوزه تشخیص قطبیت اسناد پیشنهاد شد. واژه‌نامه‌های دستی پیشین به دلیل پیچیدگی و زمان‌بر بودن و محدودیت ارزیابی تعداد واژگان همواره مشکل محدودیت منابع داشتند ولی به دلیل استفاده از امتیازات انسانی رویکردی نزدیک‌تر به نظر افراد دارند. در برابر رویکرد امتیازدهی دستی، روش‌های مبتنی بر یادگیری ماشین باتوجه‌به حوزه محور بودن، نشان داده‌اند که در بسیاری از حالات دارای کارایی بهتری هستند، از نظر حجم برداشت منابع و تشکیل سریع نسبت به واژه‌نامه‌های دستی نیز اولویت بالایی دارند. اما ایجاد الگوهای آموزشی ماشینی با نظرات انسانی فاصله داشته و این موضوع ایجاد خطا می‌کند. انتخاب ویژگی‌های صحیح و مقداردهی عددی به آن‌ها مشکل بوده و ضمن وابستگی کامل به حوزه آموزش نیازمند به بروزرسانی‌های مداوم هستند. در برابر این دو دسته واژه‌نامه‌های نیمه اتوماتیک که به‌صورت ترکیبی از روش‌های یادگیری ماشین و واژه‌نامه‌های دستی موجود استفاده می‌کنند به دلیل درنظرگرفتن هم‌زمان نظرات انسان و الگوریتم‌های یادگیری ماشین می‌توانند کارایی مجموعه را ارتقاء دهند. نقطه‌ضعف اکثر واژه‌نامه‌های موجود، عدم توجه به تشدیدکننده‌ها و

موقعیت قرارگیری آنها بوده است. باتوجه‌به این توضیح در این فصل یک روش تشکیل دیکشنری نیمه اتوماتیک با ترکیب الگوریتم‌های یادگیری ماشین و یکی از معتبرترین واژه‌نامه‌های موجود با درنظرگرفتن موقعیت مکانی تشدیدکننده پیشنهاد شد که علاوه بر لحاظ‌کردن موقعیت مکانی تشدیدکننده‌ها در امتیازدهی و اصلاح ضریب تأثیر تشدیدکننده با ترکیب روش‌های آماری و منطق فازی توانسته الگویی نزدیک‌تر به نظرات انسانی ایجاد کند که باعث افزایش کارایی مجموعه نسبت به آخرین روش‌های موجود در این حوزه گردیده است. علاوه بر این روش پیشنهادی در این فصل می‌تواند بر روی واژه‌نامه‌های دستی مختلف اعمال شده و علاوه بر افزایش دایره عبارات تحت پوشش نسبت به اصلاح کارایی آنها تأثیر مناسبی از خود نشان می‌دهد.

فصل پنجم

بررسی تأثیر موقعیت مکانی تشدیدکننده و اثر متقابل قطبیت کلمه احساسی بر تشدیدکننده‌ها

فصل ۵–بررسی تأثیر موقعیت مکانی تشدیدکننده و اثر متقابل قطبیت کلمه احساسی بر تشدیدکننده‌ها

۵-۱ مقدمه

پدیده‌های زبانی یکی از مؤثرترین عوامل تغییر قطبیت اولیه کلمات و عبارات احساسی در حوزه تجزیه‌وتحلیل احساسات محسوب می‌شوند. از جمله این عوامل می‌توان به نقش منفی‌کننده‌ها و تشدیدکننده‌ها در یک عبارت احساسی اشاره نمود. عدم توجه دقیق به تأثیر منفی‌کننده‌ها و یا تشدیدکننده‌ها و همچنین موقعیت این پدیده‌ها در یک عبارت احساسی از یک سو و عدم توجه به تأثیر قطبیت کلمات احساسی بر این پدیده‌های زبانی یکی از عواملی است که باعث کاهش کارایی واژه‌نامه‌های احساسی می‌شود. برطبق تحقیقات ما تقریباً هیچ‌کدام از این واژه‌نامه‌ها به موقعیت قرارگیری پدیده‌های زبانی پیش از کلمه احساسی و همچنین اثر متقابل قطبیت کلمه احساسی بر روی این پدیده‌های زبانی توجهی نداشته‌اند. باهدف ارتقاء کارایی واژه‌نامه‌های احساسی در این تحقیق، ابتدا عبارات احساسی Bigram و Trigram حاوی کلمات احساسی و تشدیدکننده‌ها مشخص و سپس با تحلیل آماری و روش‌های ریاضی، الگوی فازی هر یک از ترکیبات ایجاد و با استفاده از چند مرحله اصلاح بر اساس واژه‌نامه‌های دستی امتیاز هر تشدیدکننده به فرمی به دست آمد که هم در آن یادگیری ماشین و الگوهای آماری و هم ضرایب و امتیازات واژه‌نامه‌های دستی لحاظ شده است. برای ایجاد انعطاف‌پذیری و بالابردن دقت، برای هر تشدیدکننده بر اساس قطبیت کلمه احساسی و همچنین موقعیت آن در ترکیب احساسی نیز ضرایب متفاوت تعیین شد تا با این کار ضرایب مشخص شده بتواند بهتر بیانگر قطبیت نظرات باشد. نتایج شبیه‌سازی واژه‌نامه حاصل با آخرین دیکشنری‌های موجود در این حوزه بهبودی حداقل ۹٪ در پایگاه‌داده‌های استاندارد از خود نشان می‌دهد.

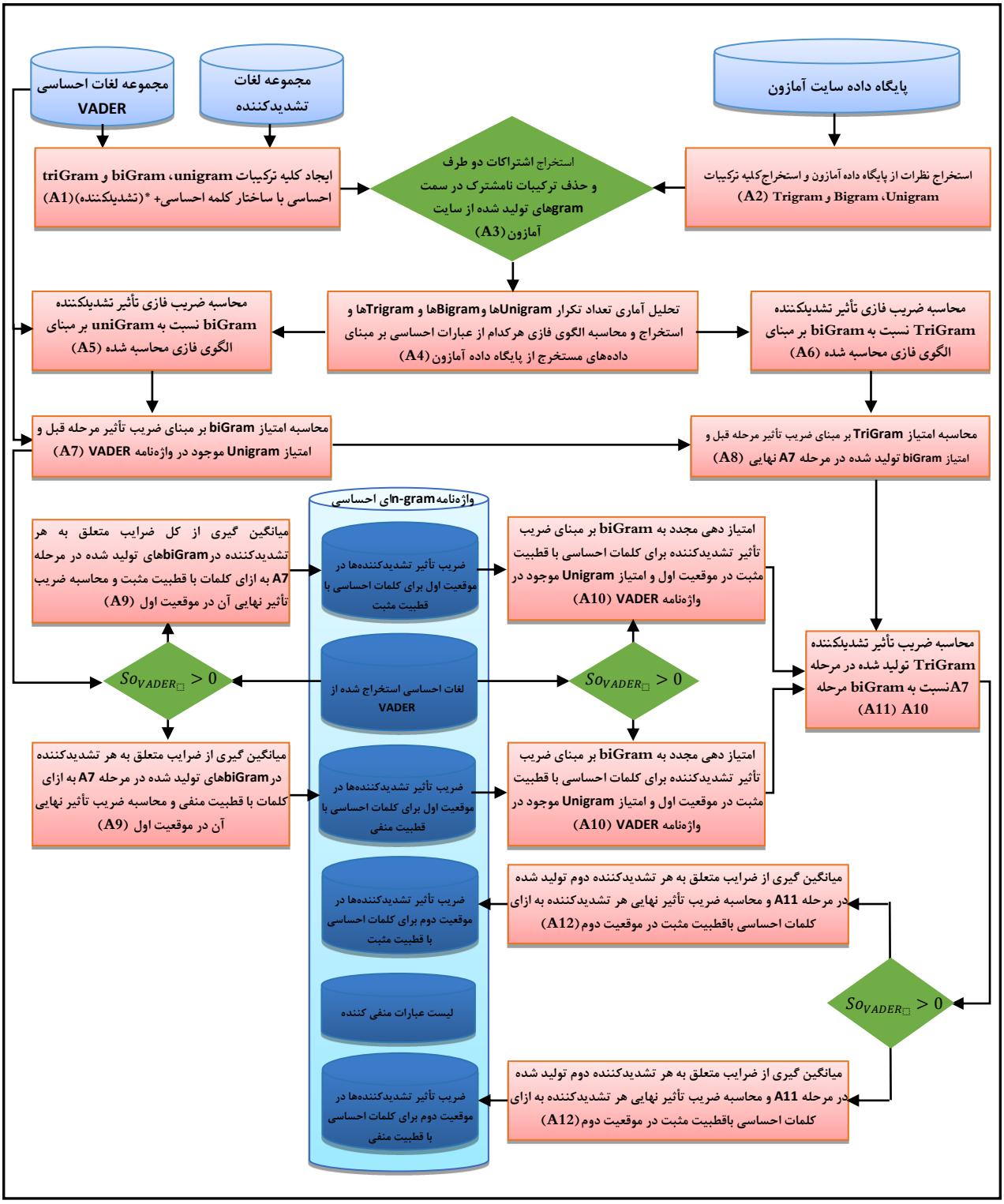
۵-۲ روش پیشنهادی

روش پیشنهادی در این فصل با اصلاح رویکرد پیشنهادی در فصل قبل، علاوه بر جایگاه، اثر قطبیت کلمات احساسی را نیز در امتیازدهی به تشدیدکننده‌ها مدنظر قرار داده است. به‌عبارت‌دیگر امتیاز هر تشدیدکننده علاوه بر جایگاه آن هنگام اتصال به کلمات احساسی مثبت و منفی تغییر می‌کند درحالی‌که قبلاً فقط به جایگاه ارتباط داشت. باتوجه‌به مشابه بودن فرایند آماده‌سازی داده‌ها و همچنین یکسان بودن پایگاه مورداستفاده و پارامترهای مقایسه در این فصل از ذکر مجدد موارد فوق پرهیز نموده و به‌صورت مستقیم به روش پیشنهادی می‌پردازیم. فلوچارت روش پیشنهادی دوم در شکل ۱۸ آورده شده است.

۵-۲-۱ نمونه‌برداری تصادفی طبقه‌ای انتساب متناسب

در اولین گام طراحی واژه‌نامه، با استفاده از لغات احساسی واژه‌نامه VADER و لیست تشدیدکننده‌ها[106] کلیه ترکیبات احساسی Bigram وTrigram که از ترکیب یک کلمه احساسی و تشدیدکننده‌ها قابل تولید بود، ساخته شده و در یک لیست ذخیره شد. پس از این مرحله با استفاده از لیست ایجاد شده و تعدادی نمونه که با استفاده از روش نمونه‌گیری تصادفی طبقه‌ای انتساب متناسب، از تمامی حوزه‌های پایگاه‌داده آمازون استخراج شده بود، کلیه ترکیبات Unigram, Bigram, Trigram مشابه با لیست ذخیره شده مرحله قبل بهمراه امتیاز متن مربوطه جداسازی شد. نمونه‌گیری تصادفی طبقه‌ای با این هدف صورت گرفت که از پیکره بسیار عظیم پایگاه‌داده آمازون تعداد محدودی نظر انتخاب شود که بتوان بر اساس این مجموعه منتخب، راجع به کل حجم پیکره اظهارنظر کرد. انتخاب تعداد محدودی نمونه از پیکره آمازون علاوه بر کاهش مشهود حجم محاسبات، منجر به سهولت استنتاجات آماری و جلوگیری از محاسبات زائد خواهد شد و از طرف دیگر باتوجه‌به اصول تئوری نمونه‌برداری پاسخ‌ها با قطعیت نسبتاً مناسب به پاسخ کل پیکره مشابهت دارد. این کار سخت‌افزار موردنیاز و زمان محاسبات لازم برای ایجاد واژه‌نامه را نیز به‌صورت مشهود کاهش داده و قادر به اعمال به هر

پیکره بسیار بزرگی از نظرات خواهد بود. در نهایت ترکیبات استخراج شده از نمونه نظرات به همراه امتیازات جمله متناظر با آن به‌صورت یک لیست جدید ذخیره گردید.



شکل ۱۸. فلوچارت پیاده سازی دیکشنری پیشنهادی دوم

جدول ۲۵. نمادهای مربوط به الگوریتم پیشنهادی دوم

نماد	توصیف
μ_{U_k}	میانگین آماری کلمه احساسی k ام واژه‌نامه VADER
$\mu_{B_{jk}}$	میانگین آماری Bigram شامل کلمه احساسی k و تشدیدکننده j
$\mu_{T_{ijk}}$	میانگین آماری Trigram شامل کلمه احساسی k و تشدیدکننده موقعیت اول jو تشدیدکننده موقعیت دوم i
$\tilde{\mu}_{B_{jk}}$	میانگین فازی Bigram شامل کلمه احساسی k و تشدیدکننده j
$\tilde{\mu}_{T_{ijk}}$	میانگین فازی Trigram شامل کلمه احساسی k و تشدیدکننده موقعیت اول jو تشدیدکننده موقعیت دوم i
σ_{U_k}	واریانس آماری Unigram شامل کلمه احساسی k
$\sigma_{B_{jk}}$	واریانس آماری Bigram شامل کلمه احساسی k و تشدیدکننده j
$\tilde{\sigma}_{B_{jk}}$	واریانس آماری Bigram جامعه جدید تولید شده بر اساس الگوی فازی کلمه احساسی k و تشدیدکننده j
$\sigma_{T_{ijk}}$	واریانس آماری Trigram شامل کلمه احساسی k و تشدیدکننده موقعیت اول jو تشدیدکننده موقعیت دوم i
$\tilde{\sigma}_{T_{jk}}$	واریانس آماری Trigram جامعه جدید تولید شده بر اساس الگوی فازی کلمه احساسی k و تشدیدکننده موقعیت اول jو تشدیدکننده موقعیت دوم i
N_{U_k}	تعداد Unigram احساسی k ام از واژه‌نامه VADER
$N_{B_{jk}}$	تعداد Bigram احساسی شامل کلمه احساسی k و تشدیدکننده موقعیت اول j
$N_{T_{ijk}}$	تعداد Trigram شامل کلمه احساسی k و تشدیدکننده موقعیت اول jو تشدیدکننده موقعیت دوم i
$CO_{I_{jk}}$	ضریب تشدیدکننده‌jدر هنگام قرارگیری پیش از کلمه احساسی K در Bigram
$CO_{I_{ijk}}$	ضریب تشدیدکننده در هنگام قرارگیری پیش از کلمه احساسی K و تشدیدکننده jدر Trigram
So_{VADER_k}	امتیاز کلمه احساسی k ام از واژه‌نامه VADER
$So_{B_{jk}}$	امتیاز اولیه Bigram شامل کلمه احساسی k و تشدیدکننده j
$So_{T_{ijk}}$	امتیاز Trigram شامل کلمه احساسی k و تشدیدکننده موقعیت اول jو تشدیدکننده موقعیت دوم i
$So_{BF_{jk}}$	امتیاز نهایی Bigram شامل کلمه احساسی k و تشدیدکننده j
$So_{TF_{ijk}}$	امتیاز نهایی Trigram شامل کلمه احساسی k و تشدیدکننده موقعیت اول jو تشدیدکننده موقعیت دوم i
$CO_{I_{jk}}$	ضریب تشدیدکننده موقعیت اول jدر هنگام قرارگیری پیش از کلمه احساسی k
$CO_{I_{ijk}}$	ضریب تشدیدکننده موقعیت دوم i در هنگام قرارگیری پیش از کلمه احساسی k و تشدیدکننده موقعیت اول j
Co_{PIF1_j}	ضریب تشدیدکننده موقعیت اول jبرای کلمات احساسی با قطبیت مثبت
Co_{NIF1_j}	ضریب تشدیدکننده موقعیت اول jبرای کلمات احساسی با قطبیت منفی
Co_{PIF2_i}	ضریب تشدیدکننده موقعیت دوم i برای کلمات احساسی با قطبیت مثبت
Co_{NIF2_i}	ضریب تشدیدکننده موقعیت دوم i برای کلمات احساسی با قطبیت منفی

۵-۲-۲ استخراج الگوی فازی عبارات احساسی

باتوجه‌به لیست ایجاد شده و یک تحلیل آماری از هر ترکیب احساسی Unigram, Bigram, Trigram یک توزیع آماری با چگالی احتمال نرمال برای هر ترکیب احساسی با تعداد نمونه بیشتر از ۳۰ محاسبه شد. درصورتی‌که تعداد نمونه موجود از یک ترکیب کمتر از ۳۰ عدد (که تقریباً تعداد مطمئن برآورد یک تابع چگالی احتمال نرمال باشد)، بود از روش‌های ساده آماری مانند میانگین کمک گرفته‌ایم و از برازش یک الگوی خاص بر داده‌ها یا حذف ترکیب به دلیل احتمال خطای بالا صرف‌نظر کردیم. در بسیاری از تحقیقات مرسوم از عملگر میانگین برای تخصیص ضریب استفاده شده است که بدیهی است استفاده از منطق فازی به‌جای میانگین ساده می‌تواند هنگام محاسبه ضریب نهایی دقت کار را بالا برده و به رویکرد تشخیص قطبیت که رویکردی فازی محسوب می‌شود نزدیک‌تر باشد.

در اولین مرحله محاسبه ضرایب، در صورت وجود الگوی فازی، تعداد نمونه Unigram, Bigram, Trigram متناظر یکسان‌سازی می‌شود. به‌عبارت‌دیگر درصورتی‌که بخواهیم تأثیر تشدیدکننده موقعیت اول و دوم را با یکدیگر بسنجیم این کار با تعداد یکسان نمونه برای دو جامعه صورت می‌گیرد.

برای نمونه برای more easy، با تعداد تقریباً ۱۰۰۰ نمونه به دلیل آنکه تعداد نمونه easy برابر ۸۹ هزار نمونه است، ۸۸۰۰۰ نمونه جدید تصادفی تولید می‌شود و از این نظر، هنگام محاسبه ضریب تأثیر تشدیدکننده، جوامع آماری مدنظر از نظر تعداد نمونه با یکدیگر کاملاً یکسان خواهد بود. بیان ریاضی این کار در روابط ۵-۱ و ۵-۲ و توضیح پارامترها در جدول ۲۵ آورده شده است.

$$\mu_{U_k}, \sigma_{U_k}, N_{U_k}$$

$$\mu_{B_{jk}}, \sigma_{B_{jk}}, N_{B_{jk}} \quad \Longrightarrow PDF\left(\mu_{B_{jk}}, \sigma_{B_{jk}}, N_{B_{jk}}\right) \Longrightarrow \left(\tilde{\mu}_{B_{jk}}, \tilde{\sigma}_{B_{jk}}, N_{U_k}\right) \qquad (5-1)$$

$$\mu_{T_{ijk}}, \sigma_{T_{ijk}}, N_{T_{ijk}} \Longrightarrow PDF\left(\mu_{T_{jk}}, \sigma_{T_{ijk}}, N_{T_{ijk}}\right) \Longrightarrow \left(\tilde{\mu}_{T_{ijk}}, \tilde{\sigma}_{T_{ijk}}, N_{U_k}\right) \qquad (5-2)$$

۵-۳ محاسبه ضریب تأثیر تشدیدکننده‌ها

پس از یکسان‌سازی تعداد نمونه‌ها با الگوی فازی، برای محاسبه ضریب تأثیر تشدیدکننده‌ها، بر اساس روابط ۵-۳ و ۵-۴ استفاده از میانگین‌گیری ساده در نمونه‌های جدید امتیاز ضریب تأثیر اولیه برای موقعیت اول و دوم محاسبه می‌شود.

$$\begin{matrix}(\mu_{U_k}, & \sigma_{U_k}, & N_{U_k})\\ (\tilde{\mu}_{B_{jk}}, & \tilde{\sigma}_{B_{jk}}, & N_{U_k})\\ (\tilde{\mu}_{T_{jk}}, & \tilde{\sigma}_{T_{jk}}, & N_{U_k})\end{matrix} \Longrightarrow \begin{cases} CO_{I_{jk}} = \dfrac{\tilde{\mu}_{B_{jk}}}{\mu_{U_k}} & (5-3)\\ CO_{I_{ijk}} = \dfrac{\tilde{\mu}_{T_{ijk}}}{\tilde{\mu}_{B_{jk}}} & (5-4)\end{cases}$$

پس از این مرحله برای آنکه پاسخ را به واقعیت نزدیک و خطا را در یک واژه‌نامه عمومی در مجموعه‌های مختلف پایگاه‌داده با نظرات گوناگون کمتر کنیم با استفاده از ضریب به‌دست‌آمده تشدیدکننده در موقعیت اول از رابطه ۵-۳ و امتیاز کلمات احساسی موجود در واژه‌نامه VADER، امتیازات ترکیبات Bigram ذخیره شده در لیست استخراج شده به‌روزرسانی می‌شود.

به‌این‌ترتیب در امتیاز استخراج شده جدید هم تأثیر واژه‌نامه دستی VADER و هم محاسبات فازی مرحله قبل دخیل خواهد بود و به این صورت روش یادگیری ماشین پیشنهادی با روش دستی واژه‌نامه VADER ترکیب خواهد شد و خطای احتمالی موجود کاهش داده می‌شود. بیان ریاضی این کار در رابطه ۵-۵ آورده شده است.

$$So_{B_{jk}} = CO_{I_{jk}} * So_{VADER_k} \qquad (5-5)$$

در مرحله بعد با استفاده از قواعد منطقی تعریف شده در جدول ۱۷ عملیات پالایش امتیازات صورت می‌پذیرد. در طی عملیات پالایش برخی از امتیازات استخراج شده از پایگاه‌داده آمازون حذف شده تا به‌این‌ترتیب عملیات تصحیح ضرایب به‌دست‌آمده صورت پذیرد.

با استفاده از ضرایب تأثیر محاسبه شده برای تشدیدکننده موقعیت دوم از رابطه ۵-۴ و امتیاز به‌روزرسانی شده Bigram رابطه ۵-۵ ، مشابه مرحله قبل امتیاز Trigramهای لیست ذخیره شده به‌روزرسانی می‌گردد (رابطه ۵-۶).

$$So_{T_{ijk}} = CO_{I_{ijk}} * So_{B_{jk}} \qquad (5-6)$$

در این مرحله نیز برای تصحیح Trigramها چهار شرط و دو حالت در نظر گرفته شده است و در صورت عدم موفقیت این شرایط تابع refiner برای تصحیح عملیات فراخوانی می‌شود. جدول ۲۶ حالات مختلف مورداستفاده جهت بررسی و تأیید امتیازات استخراج شده برای اصطلاحات Trigram حاوی یک کلمه احساسی و تشدیدکننده اول و دوم را نشان می‌دهد.

جدول ۲۶. جدول حالات مختلف بررسی و تأیید امتیازات استخراج شده برای اصطلاحات Trigram

حالات اول			
تشدیدکننده دوم	**قطبیت Bigram**	**تشدیدکننده دوم**	**قطبیت Bigram**
تقویت‌کننده	منفی	تضعیف کننده	مثبت
شرط ۱		شرط ۲	
حالت دوم			
تشدیدکننده دوم	**قطبیت Bigram**	**تشدیدکننده دوم**	**قطبیت Bigram**
تقویت‌کننده	مثبت	تضعیف کننده	منفی
شرط ۱		شرط ۲	

چنانچه یک Trigram یکی از شرایط حالت نخست را برآورده نماید و لیکن مقدار نهایی محاسبه شده برای آن بزرگتر از مقدار امتیاز حاصل از Bigram محاسبه شده در مرحله قبل باشد در این حالت تابع refiner فراخوانی شده و عملیات پالایش نتایج صورت می‌پذیرد. در طی این فرایند بصورت گام به گام مقادیر بالا با این فرض که کلیه مقادیر بالا ناشی از وجود تعصب مشتری بوده و یا به علت طولانی بودن جمله در نتیجه ترکیب با کلمات احساسی دیگر حاصل شده است حذف می‌گردد. به‌این‌ترتیب تا زمانی که تابع تصحیح با موفقیت شرایط مورد نظر را پاس نکرده باشد این عملیات تکرار می‌گردد و در نهایت در صورت عدم موفقیت عبارت احساسی از مجموعه حذف خواهد شد. این فرایند برای حالت دوم نیز صورت پذیرفته و در طی فرایند

پالایش مشاهدات مغرضانه (خیلی زیاد یا خیلی کم) حذف می‌گردد. فرایند بررسی و تصحیح تا برآورده شدن شرایط تکرار می‌گردد و در نهایت ضریب اصلاح شده حاصل برای مرحله بعد تولید می‌گردد.

در دومین مرحله محاسبه ضریب تأثیر، با استفاده از امتیازات محاسبه شده برای مجموعه ترکیبات Bigram لیست ذخیره شده، باتوجه‌به قطبیت کلمه احساسی، Bigramها به دو دسته ترکیبات شامل کلمه احساسی با وزن مثبت و کلمات با وزن احساسی منفی تقسیم و برای هر تشدیدکننده هنگام اتصال به کلمات احساسی با قطبیت مثبت یک ضریب تأثیر و هنگام اتصال به کلمات احساسی منفی یک ضریب تأثیر مجزا محاسبه می‌شود. به‌این‌ترتیب می‌توان اطمینان داشت ضریب تأثیر تشدیدکننده درBigram الزاماً تابع قطبیت کلمه احساسی نیز بوده که این موضوع باتوجه‌به نتایج حاصل از آزمایش‌ها ما قطعاً در درصد درستی خصوصاً در جملات منفی تأثیر مستقیم و مناسبی از خود نشان خواهد داد (رابطه ۵-۷ و ۵-۸).

$$Co_{PIF_{1j}} = \frac{\sum_{k \in So_{VADER_K} > 0} Co_{Ijk}}{N_{\left(Co_{Ijk}\, j \in Co_{Ijk}, k \in So_{VADER_k} > 0\right)}} \qquad (5-7)$$

$$Co_{NIF_{1j}} = \frac{\sum_{k \in So_{VADER_K} < 0} Co_{I_{jk}}}{N_{\left(Co_{I_{jk}}\, j \in Co_{I_{j_k}}, k \in So_{VADER_k} < 0\right)}} \qquad (5-8)$$

با استفاده از نتیجه مراحل ۵-۷ و ۵-۸ و امتیاز کلمات احساسی موجود در واژه‌نامه VADER،امتیاز Bigramهای لیست ذخیره شده مجدداً به‌روزرسانی می‌شود.

$$So_{BF_{jk}} = \begin{cases} CO_{PIF_{1j}} * So_{VADER_k} & So_{VADER_k} > 0 \\ CO_{NIF_{1j}} * So_{VADER_k} & So_{VADER_k} < 0 \end{cases} \qquad (5-9)$$

حال به‌ازای Bigramهای به‌روزرسانی شده و امتیازات Trigramهای لیست ذخیره شده که بر اساس رابطه ۵-۶ بروز شده بود، ضرایب تشدیدکننده‌ها در موقعیت دوم برای ترکیبات دارای کلمه احساسی با قطبیت مثبت و منفی مجدداً محاسبه می‌شود. سپس با میانگین‌گیری از این ضرایب تأثیر، ضریب تأثیر نهایی هر تشدیدکننده در موقعیت دوم بر اساس قطبیت کلمه احساسی محاسبه خواهد شد (رابطه ۵-۱۰ و ۵-۱۱). این

کار منجر به حداقل کردن خطای احتمالی هم در نمونه مثبت و هم در نمونه منفی خواهد شد و تأثیر تشدیدکننده‌ها در موقعیت دوم را بسیار دقیق‌تر در نظر خواهد گرفت.

$$Co_{PIF_{2i}} = \frac{\sum_{j \in Co_{I_{jk}}} \sum_{k \in So_{VADER_K} > 0} Co_{I_{ijk}}}{N_{(Co_{I_{ijk}}\, i, j \in Co_{I_{ijk}}, k \in So_{VADER_k} > 0)}} \qquad (5-10)$$

$$Co_{NIF_{2i}} = \frac{\sum_{j \in Co_{I_{jk}}} \sum_{k \in So_{VADER_K} < 0} Co_{I_{ijk}}}{N_{(Co_{I_{ijk}}\, i, j \in Co_{I_{ijk}}, k \in So_{VADER_k} < 0)}} \qquad (5-11)$$

برطبق نتایج به‌دست‌آمده توسط ما ترکیب مجموعه مراحل فوق و به‌صورت مشخص الگوی فازی استفاده شده، درنظرگرفتن تأثیر موقعیت تشدیدکننده پیش از کلمه احساسی و اثر قطبیت کلمه احساسی بر تشدیدکننده می‌تواند با ایجاد یک تحلیل آماری نسبتاً کارآمد در ارتقاء کارایی تشخیص قطبیت نظرات نسبت به روش‌های موجود نتایج بهتری از خود نشان دهد.

$$So_{U_k} = So_{VADER_k} \qquad (5-12)$$

$$So_{TF_{ijk}} = \begin{cases} Co_{PIF2_i} * CO_{PIF1_j} * So_{VADER_k} & So_{VADER_k} > 0 \\ Co_{NIF2_i} * CO_{NIF1_j} * So_{VADER_k} & So_{VADER_k} < 0 \end{cases} \qquad (5-13)$$

باتوجه‌به روابط ۵-۹ ، ۵-۱۲ ، ۵-۱۳ و امتیازات "enjoy" و "bad" در واژه‌نامه احساسی VADER که به ترتیب 2.3 و 2.5- هستند و با استفاده از جدول ۲۷ می‌توان امتیازات عبارات احساسی "very big bad" و "very deeply enjoyed" را به‌صورت زیر محاسبه نمود.

جدول ۲۷. ضرایب محاسبه شده برای چند تشدیدکننده نمونه باتوجه‌به موقعیت آنها در عبارت احساسی و قطبیت کلمه احساسی در روش پیشنهادی دوم

تشدیدکننده	قطبیت مثبت کلمه احساسی		قطبیت منفی کلمه احساسی	
	موقعیت اول	موقعیت دوم	موقعیت اول	موقعیت دوم
deeply	1.552856	1.207102	3.0426	3.035896
very	1.356644	1.2476723	2.826969	1.8423642
big	1.432992	1.0424849	2.6126184	1.4095658

very big bad=1.84*2.61*(-2.5)=-12.006

very deeply enjoyed=1.25*1.55*2.3=4.456

۵-۴ شبیه‌سازی و نتایج

۵-۴-۱ ارزیابی روش پیشنهادی

دراین مرحله از شبیه‌سازی، نتایج روش پیشنهادی دوم با دو واژه‌نامه احساسی عمومی VADER و Senti-N-gram که کارایی مناسب‌تری نسبت به واژه‌نامه‌های احساسی موجود از خود نشان داده‌اند مقایسه گردیده است. برای آنکه شرایط مقایسه کاملاً یکسان باشد، پایگاه‌داده‌های استاندارد و معیار مقایسه مشابه برای هر سه واژه‌نامه انتخاب شده است. در زمینه روش‌های تخصیص امتیاز، امتیازات واژه‌نامه VADER به دلیل دستی بودن، ثابت بوده اما در واژه‌نامه Senti-N-gram نمونه‌ها مشابه با روش پیشنهادی در نظر گرفته شده تا کارایی الگوریتم پیشنهادی در شرایط کاملاً مشابه با روش استفاده شده در واژه‌نامه Senti-N-gram مقایسه شود.

پایگاه‌داده‌های انتخاب شده در این مرحله شامل دو پایگاه‌داده Taboada و Pang & Lee است که در بخش‌های قبل راجع به آنها توضیح داده شده و از پایگاه‌داده‌های بسیار متداول برای بررسی کارایی روش‌های تشخیص قطبیت محسوب می‌شوند.

برای ارزیابی کارایی واژه‌نامه‌ها از پارامترهای متداول True positives (TP) (اسناد با قطبیت مثبت درست تفکیک شده)، True negatives (TN) (اسناد با قطبیت منفی درست تفکیک شده)، False positives (FP) (اسناد با قطبیت منفی که به‌اشتباه مثبت تشخیص داده شده‌اند) و False negatives (FN) (اسناد با قطبیت مثبتی که به‌اشتباه منفی تشخیص داده شده‌اند)، استفاده شده است. برای ارزیابی کارایی واژه‌نامه احساسی، با استفاده از مقادیر فوق، معیارهای متداول Precision، Recall،Accuracy و F-measure محاسبه و استفاده شده است.

دقت کنیم که در واژه‌نامه VADER امتیاز عبارات احساسی دستی و ثابت بود، اما برای واژه‌نامه Senti-N-gram محاسبه امتیاز بر اساس نمونه‌برداری انجام شده صورت گرفت تا شرایط مقایسه کاملاً مشابه باشد. در

جداول ۲۷ تا ۳۲ نتایج روش پیشنهادی دوم با دو روش مختلف تعیین قطبیت سند در سطح جمله بر اساس جمع کل امتیازات جملات [154] و بر اساس شمارش تعداد جملات مثبت و منفی و انتخاب اکثریت [2] و دو واژه‌نامه Senti-N-gram و VADER آورده شده است.

جدول ۲۸. مقایسه پارامتر دقت بر مبنای دوروش طبقه‌بندی اسناد در سطح جمله با استفاده از واژه‌نامه‌های VADER، Senti-N-gram و واژه‌نامه پیشنهادی دوم

Datasets	طبقه‌بندی اسناد در سطح جمله «الگوریتم اول»			طبقه‌بندی اسناد در سطح جمله «الگوریتم دوم»		
	VADER	Senti-N-gram	روش پیشنهادی دوم	VADER	Senti-N-gram	روش پیشنهادی دوم
Epinions	62	63.25	**72.00**	63.5	64.50	**70.25**
Movie review	63.5	63.25	**68.70**	63.95	62.80	**68.00**

همان‌طور که می‌توان دید واژه‌نامه پیشنهادی نسبت به هر دو واژه‌نامه VADER و Senti-N-gram در هر دو روش تشخیص قطبیت، به طور متوسط ۹ درصد در کارایی بهبود داشته که درصورتی‌که نسبت به مقدار اولیه درصد کارایی را بسنجیم، در بهترین حالت چیزی در حدود ۱۴ درصد نسبت به روش‌های قبل بهبود داشته‌ایم. این بهبود نشان می‌دهد، روش پیشنهادی در تخصیص امتیاز باتوجه به استفاده از چهار ضریب تأثیر مختلف مبتنی بر قطبیت کلمه احساسی و موقعیت تشدیدکننده، الگوهای فازی، روش‌های مختلف اصلاح امتیاز کارایی بسیار مناسب‌تری نسبت به بهترین روش‌های گذشته از خود نشان داده است.

در جدول ۲۹ و جدول ۳۰ به تفکیک حوزه، کارایی روش پیشنهادی دوم با واژه‌نامه‌های موجود بر اساس جمع کل امتیازات جمله مقایسه شده تا بتوان ارزیابی روشنتری از کارایی روش دوم نسبت به واژه‌نامه‌های این حوزه داشت. همان‌طور که می‌توان دید اگرچه در بعضی موارد در حوزه مثبت، روش پیشنهادی افت کارایی داشته، اما در مجموع اسناد مثبت و منفی در همه موارد روش پیشنهادی کارایی مناسبتری از خود نشان داده است. برای نمونه در حوزه HOTELS که تنها ۸ سند با قطبیت منفی توسط الگوریتم VADER تشخیص داده شده بود، روش پیشنهادی توانسته ۴۰ سند با قطبیت منفی را مشخص کند (بهبود ۴۰۰ درصدی)، که این موضوع اهمیت تحلیل مجزای ضرایب تشدیدکننده‌ها بر اساس قطبیت و موقعیت را نشان می‌دهد.

جدول ۲۹. مقایسه پارامتر دقت بر مبنای روش طبقه‌بندی اسناد در سطح جمله [153] **«الگوریتم اول» با استفاده از واژه‌نامه‌های VADER،Senti-N-gram و واژه‌نامه پیشنهادی دوم در حوزه‌های مختلف**

Datasets	VADER			Senti-N-gram			proposed		
	Positive	Negative	overall	Positive	Negative	overall	Positive	Negative	overall
BOOKS	**100**	36	**68**	92	44	68	80	**64**	**72**
CARS	**100**	28	64	**100**	28	64	96	**48**	**72**
COMPUTERS	**100**	32	66	**100**	36	68	**100**	**52**	**76**
COOKWARE	**100**	16	58	**100**	12	56	**100**	**32**	**66**
HOTELS	**100**	8	54	**100**	24	62	**100**	**40**	**70**
MOVIES	**96**	32	64	96	28	62	88	**44**	**66**
MUSIC	**92**	24	58	92	24	58	92	**52**	**72**
PHONES	92	36	64	88	48	68	**96**	**68**	**82**

جدول ۳۰. مقایسه پارامترهای Precision، Recall و F-measure بر مبنای روش طبقه‌بندی اسناد در سطح جمله [153] **«الگوریتم اول» با استفاده از واژه‌نامه‌های VADER،Senti-N-gram و واژه‌نامه پیشنهادی دوم در حوزه‌های مختلف**

Datasets	VADER			Senti-N-gram			proposed		
	P	R	F	P	R	F	P	R	F
BOOKS	0.61	**1.00**	**0.76**	0.62	0.92	0.74	**0.69**	0.80	0.74
CARS	0.58	**1.00**	0.74	0.58	**1.00**	0.74	**0.65**	0.96	**0.77**
COMPUTERS	0.60	**1.00**	0.75	0.61	**1.00**	0.76	**0.68**	**1.00**	**0.81**
COOKWARE	0.54	**1.00**	0.70	0.53	**1.00**	0.69	**0.60**	**1.00**	**0.75**
HOTELS	0.52	**1.00**	0.69	0.57	**1.00**	0.73	**0.63**	**1.00**	**0.77**
MOVIES	0.59	**0.96**	**0.73**	0.57	**0.96**	0.72	**0.61**	0.88	0.72
MUSIC	0.55	**0.92**	0.69	0.55	**0.92**	0.69	**0.66**	**0.92**	**0.77**
PHONES	0.59	0.92	0.72	0.63	0.88	0.73	**0.75**	**0.96**	**0.84**

به‌عبارت‌دیگر تفکیک امتیازات تخصیص‌داده‌شده به تشدیدکننده‌ها بر اساس قطبیت-موقعیت و پروسه اصلاح ضرایب در چند مرحله و ترکیب آن با الگوی فازی و یکسان‌سازی نمونه توانسته کارایی تخصیص امتیاز را به طرز بسیار مشهودی به واقعیت، خصوصاً در جملات منفی نزدیک‌تر کند و به‌صورت هم‌زمان استفاده از نمونه‌گیری توانسته بار محاسباتی سیستم را کاهش دهد که ترکیب این‌ها با یکدیگر در مجموع علاوه بر کاهش بار محاسباتی، افزایش کارایی امتیازات تخصیص‌داده‌شده را به دنبال داشته است.

در جدول ۳۰ در سه معیار F-measure، Recall, Precision مقایسه صورت‌گرفته که همان طور که از مجموع نتایج روی پایگاه‌داده می‌توان دید روش پیشنهادی در معیار Recall که فقط با اسناد با قطبیت مثبت سر و کار داشته فقط در سه حوزه آنهم به مقدار اندکی نسبت به روش‌های قبل کاهش داشته که این کاهش، در مقایسه با افزایش کارایی روش پیشنهادی در تشخیص اسناد با قطبیت منفی بسیار کمتر است. باتوجه‌به معیار Precision می‌توان دید که روش روش پیشنهادی به طور میانگین در حدود ۱۲ درصد بهبود نسبت

جدول ۳۱. مقایسه پارامتر دقت بر مبنای روش طبقه‌بندی اسناد در سطح جمله [2] **«الگوریتم دوم» با استفاده از واژه‌نامه‌های VADER،Senti-N-gram و واژه‌نامه پیشنهادی دوم در حوزه‌های مختلف**

Datasets	VADER			Senti-N-gram			proposed		
	Positive	Negative	overall	Positive	Negative	overall	Positive	Negative	overall
BOOKS	**100**	36	68	88	40	64	92	**56**	**74**
CARS	**100**	28	64	**100**	32	66	**100**	**48**	**74**
COMPUTERS	96	**40**	**68**	**100**	36	**68**	88	**40**	64
COOKWARE	**100**	12	56	**100**	20	60	**100**	**32**	**66**
HOTELS	**100**	20	60	**100**	20	60	**100**	**36**	**68**
MOVIES	92	36	64	**96**	28	62	92	**40**	**66**
MUSIC	**96**	20	58	**96**	20	58	92	**40**	**66**
PHONES	**96**	44	70	92	64	78	92	**76**	**84**

جدول ۳۲. مقایسه پارامترهایPrecision ، Recallو F-measure بر مبنای روش طبقه‌بندی اسناد در سطح جمله [2] «الگوریتم دوم» با استفاده از واژه‌نامه‌هایVADER ، Senti-N-gramو واژه‌نامه پیشنهادی دوم در حوزه‌های مختلف

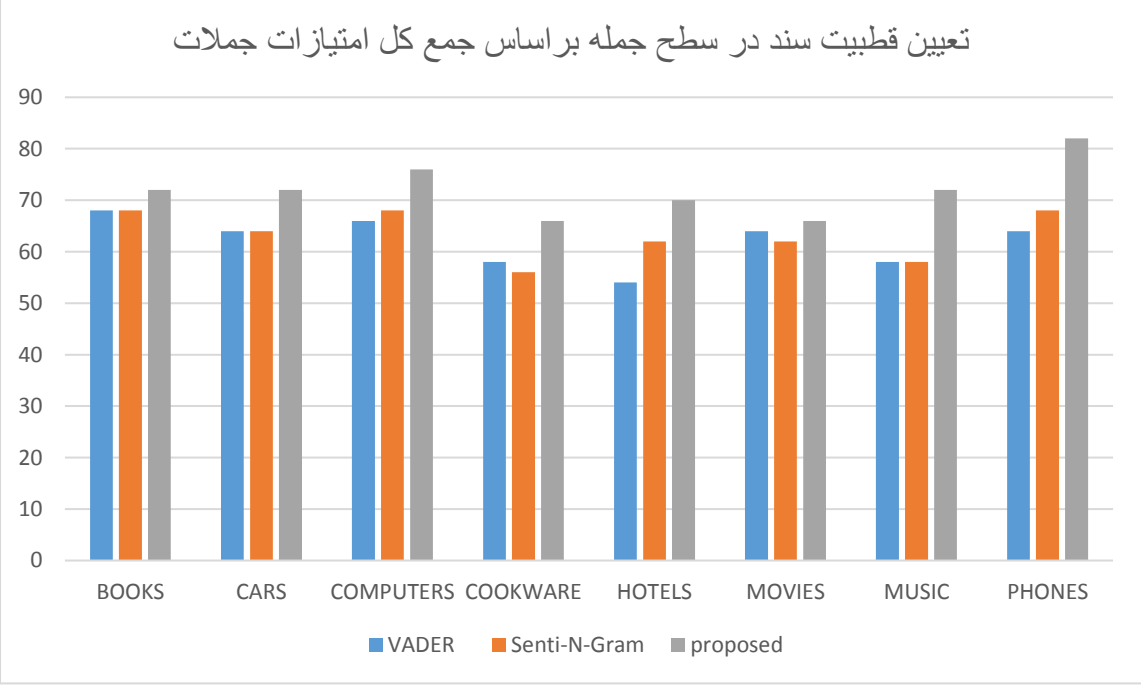
Datasets	VADER			Senti-N-gram			proposed		
	P	R	F	P	R	F	P	R	F
BOOKS	0.61	**1.00**	0.76	0.60	0.88	0.71	**0.68**	0.92	**0.78**
CARS	0.58	**1.00**	0.74	0.60	**1.00**	0.75	**0.66**	**1.00**	**0.79**
COMPUTERS	**0.62**	0.96	0.75	0.61	**1.00**	**0.76**	0.60	0.88	0.71
COOKWARE	0.53	**1.00**	0.69	0.56	**1.00**	0.71	**0.60**	**1.00**	**0.75**
HOTELS	0.56	**1.00**	0.71	0.56	**1.00**	0.71	**0.61**	**1.00**	**0.76**
MOVIES	0.59	0.92	0.72	0.57	**0.96**	0.72	**0.61**	0.92	**0.73**
MUSIC	0.55	**0.96**	0.70	0.55	**0.96**	0.70	**0.61**	0.92	**0.73**
PHONES	0.63	**0.96**	0.76	0.72	0.92	0.81	**0.79**	0.92	**0.85**

به بهترین نتایج حاصل از دیکشنری‌های مورد ارزیابی از خود نشان داده است. در معیار F-measure این بهبود نسبت به بهترین نتایج واژه‌نامه‌ها افزایش تقریبا ۶٪ای را به دنبال داشته که بیانی مشخص از کارایی بهتر روش پیشنهادی دوم در مجموع اسناد با قطبیت مثبت و منفی محسوب می‌گردد.

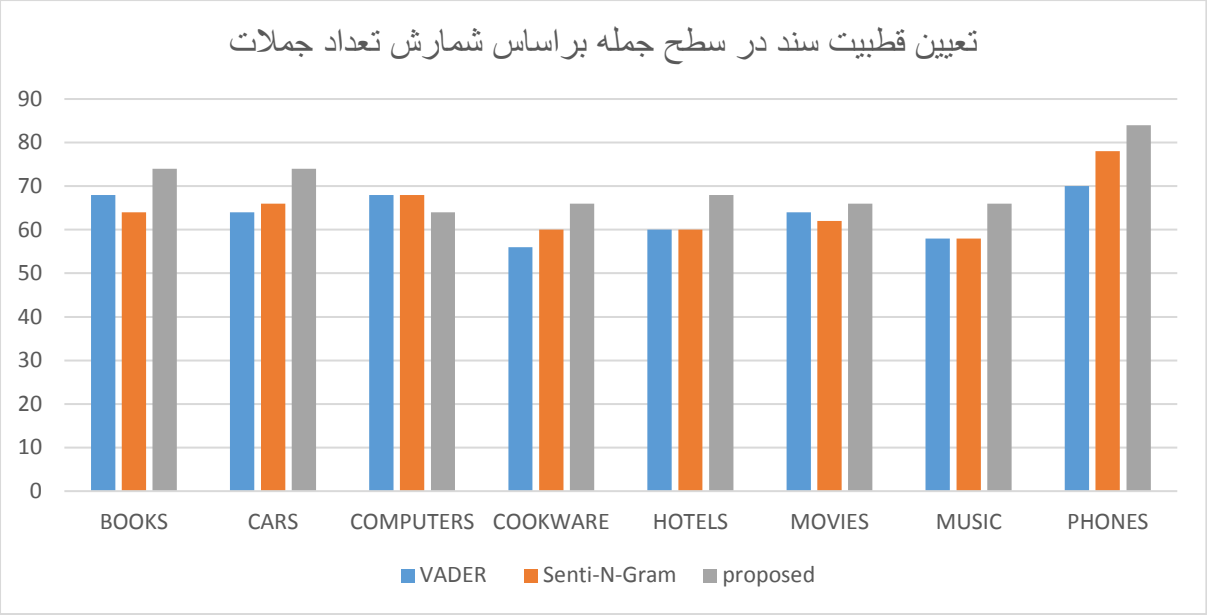
در جداول ۳۱ و ۳۲ نتایج تخصیص امتیاز پیشنهادی با الگوریتم معیار قالب یا رای اکثریت در تشخیص قطبیت اسناد بر روی پایگاه‌داده Taboada گزارش شده است. نتایج تا حد زیادی تأیید کننده نتایج مندرج در جداول ۲۸ و ۲۹ است. نتایج حاصل از جدول ۳۱ در حوزه تشخیص اسناد منفی نشان از عملکرد قابل توجه این الگوریتم خصوصا در حوزه تشخیص اسناد منفی بوده و در تمامی حوزه‌ها افزایش تشخیص اسناد منفی را شاهد هستیم. بیشترین میزان تشخیص نسبی اسناد با قطبیت منفی متعلق به حوزه‌های Music، Hotels و Cookware به ترتیب با مقادیر ۱۰۰٪، ۸۰٪ و ۶۰٪ در این جدول قابل مشاهده است. همچنین در این جداول نیز در قسمت دقت کلی غیر از حوزه کامپیوتر در تمامی حوزه‌ها بهبود داشته‌ایم. بیشترین بهبود در حوزه MUSIC با ۱۳.۸ درصد قابل مشاهده است. علاوه‌برآن به‌صورت میانگین در کل سیستم، درصد درستی تقریباً ۹ درصد نسبت روش‌های مشابه در الگوریتم تعداد غالب از خود بهبود نشان می‌دهد. روش پیشنهادی صرف‌نظر

از نوع تشخیص قطبیت دقت مناسب‌تری نسبت به دیکشنری‌های قبل خواهد داشت. در جدول ۳۲ معیارهای Precision, Recall, F-measure برای دیکشنری پیشنهادی نسبت به واژه‌نامه‌های مورد ارزیابی و با معیار الگوریتم تعداد غالب آورده شده که باز هم مشابه جدول ۲۸ نتایج معیار کارایی روش پیشنهادی خصوصاً در قسمت‌هایی است که اسناد منفی نقش بیشتری در آن داشته‌اند بسیار مشهودتر و بهتر است.

شکل ۱۹ پارامتر دقت کلی روش را بر مبنای روش طبقه بندی اسناد با مجموع امتیازات نشان می‌دهد. باتوجه‌به ارتفاع قسمت خاکستری یعنی روش پیشنهادی می‌توان دید در تمامی حوزه‌ها در این حالت الگوی پیشنهادی نسبت به دیکشنری‌های قبل کارایی مناسب تر از خود نشان داده که بیشترین تغییر در قسمت تلفن در انتهای نمودار قابل مشاهده است. در شکل ۲۰ همین نمودار بر اساس الگوریتم تعداد غالب رسم شده که با وجود کمتر شدن اختلاف‌ها باز هم روش پیشنهادی در اکثر موارد نسبت به دیکشنری‌های قبلی نتیجه مناسب‌تری از خود نشان می‌دهد.



شکل ۱۹. مقایسه پارامتر دقت بر مبنای روش طبقه‌بندی اسناد در سطح جمله [153] «الگوریتم اول» با استفاده از واژه‌نامه‌های VADER،Senti-N-gram و واژه‌نامه پیشنهادی دوم در حوزه‌های مختلف



شکل ۲۰. مقایسه پارامتر دقت بر مبنای روش طبقه‌بندی اسناد در سطح جمله[2] «الگوریتم دوم» با استفاده از واژه‌نامه‌های VADER،Senti-N-gram و واژه‌نامه پیشنهادی دوم در حوزه‌های مختلف

۵-۵ جمع‌بندی

در این فصل رویکردی برای محاسبه امتیاز در واژه‌نامه احساسی عمومی، مبتنی بر امتیازدهی فازی با درنظرگرفتن جایگاه تشدیدکننده نسبت به کلمه احساسی و تأثیر قطبیت کلمه احساسی بر تشدیدکننده پیشنهاد گردید که با ایجاد تنوع در امتیازدهی، کارایی تشخیص قطبیت اسناد و جملات در سیستم‌های پردازش زبان طبیعی را بهبود بخشید.

برای بهتر کردن کارایی رویکرد پیشنهادی نسبت به روش‌های موجود و رویکرد فصل قبل، برای آنکه امتیازدهی به‌صورت درست انجام گیرد علاوه بر چند مرحله اصلاح ضرایب و استفاده از الگوی فازی، جایگاه تشدیدکننده از نظر موقعیت اول یا دوم نسبت به کلمه احساسی و به‌صورت هم‌زمان تأثیر قطبیت کلمه احساسی بر تشدیدکننده در نظر گرفته شد که به‌این‌ترتیب برای هر تشدیدکننده باتوجه‌به موقعیت اول یا دوم و قطبیت کلمه احساسی چهار ضریب مختلف به دست آمد که منجر به ایجاد تنوع امتیازی بسیار کارآمدتر، گستره ترکیبات تحت پوشش بیشتر و در نهایت کارایی مناسب‌تر خصوصاً در حوزه اسناد و جملات با قطبیت منفی نسبت به واژه‌نامه‌های مرسوم شد.

اصلاح چندمرحله‌ای ضریب تأثیر که هم در آن الگوریتم‌های یادگیری ماشین و هم ضرایب انسانی دخالت داشتند نیز در افزایش کارایی مجموعه بی‌تأثیر نبوده است. در نهایت روش پیشنهادی در شرایط یکسان بر روی تعدادی از پایگاه‌داده‌های استاندارد پیاده‌سازی و نتایج آن با دو واژه‌نامه معتبر و بروز در این حوزه مقایسه گردید.

مقایسه‌های انجام شده در همه پارامترها نشان داد روش پیشنهادی باتوجه‌به اختصاص امتیاز به تشدیدکننده‌ها بر اساس قطبیت کلمه احساسی و جایگاه تشدیدکننده، خصوصاً در حوزه اسناد منفی توانسته الگوی نزدیک‌تری به نظرات انسان ایجاد کند که این کار در نهایت به کارایی مجموعه سیستم بسیار کمک کرده و نسبت به آخرین روش‌های این حوزه در حدود ۹.۳ درصد به‌صورت نسبی افزایش کارایی داشتیم.

۶-۵ نتیجه‌گیری و کارهای آینده

در این رساله روشی مبتنی بر قانون برای تولید یک واژه‌نامه احساسی که قابلیت پشتیبانی از تشدیدکننده‌های متوالی را داشته باشد ارائه شده است. هرچند که اهمیت پدیده‌های زبانی بر تغییر شدت قطبیت و حتی تغییر قطبیت عبارات احساسی بر هیچ‌کدام از محققان این حوزه ناشناخته نیست بااین‌حال غالب تحقیقات صورت‌گرفته در حوزه تجزیه‌وتحلیل احساسات به استخراج کلمات احساسی و امتیازدهی به این کلمات تخصیص‌یافته است. در حوزه پدیده‌های زبانی نیز غالب تحقیقات به نقش منفی‌کننده‌ها در هنگام قرارگیری آنها پیش از کلمات احساسی توجه نموده است. باتوجه‌به نقش‌های متفاوتی همچون نقش تاکید یا تضعیف و یا منفی‌کننده‌گی که یک منفی‌کننده در جمله می‌تواند داشته باشد در ساده‌ترین حالت این پدیده‌ها منجر به معکوس نمودن قطبیت عبارت احساسی می‌شوند و در حالات پیچیده اثر متفاوتی را بر مبنای کلمه احساسی بعد از خود می‌توانند داشته باشد. روش پیشنهاد شده در این رساله بر مبنای این پرسش اساسی بنا شده است که آیا تأثیر چند تشدیدکننده متوالی پیش از یک کلمه احساسی با تاثیر همان تشدیدکننده به‌صورت تنها یکسان است و همچنین آیا قطبیت یک کلمه احساسی می‌تواند همانند تأثیری که بر روی منفی‌کننده‌ها اعمال نماید بر تشدیدکننده‌های پیش از خود نیز مؤثر باشد.

۷-۵ جمع‌بندی روش پیشنهادی در رساله

تجزیه‌وتحلیل احساسات فرایند تشخیص قطبیت اسناد با استفاده کلمات احساسی موجود در یک متن است. تشخیص میزان قطبیت و شدت یک متن نه‌تنها وابسته به کلمات احساسی موجود در آن متن است که می‌تواند به کلمات دیگری که فاقد بار عاطفی و احساسی هستند نیز وابسته باشد. این کلمات که تحت عنوان پدیده‌های زبانی شناخته می‌شوند شامل منفی‌کننده‌ها، تشدیدکننده‌ها، عبارات حاوی کلمه"But" و عبارات غیرواقعی[1] یا جملات فرضی هستند که تأثیر قابل توجهی بر ارزیابی و تحلیل احساسات یک متن از خود به‌جای می‌گذارند. عدم توجه به اثر این پدیده‌های زبانی می‌تواند باعث کاهش کارایی و حتی تشخیص اشتباه قطبیت یک سند گردد. هرچند که اثر این پدیده‌های زبانی بر هیچ محققی در حوزه تجزیه‌وتحلیل احساسات پوشیده نیست اما در طی سالیان گذشته به سبب پیچیدگی‌های فراوان موجود تنها به بخش بسیار کوچکی از این پدیده‌ها توجه گردیده است.

باتوجه‌به اهمیت روزافزون تجزیه‌وتحلیل احساسات در جملات و گستره عملکرد این مسئله در حوزه‌های مختلف سیاسی، اجتماعی، اقتصادی، پزشکی و آموزشی در این رساله یک روش ترکیبی جهت تولید ضرایب یک تشدیدکننده پیش از یک کلمه احساسی ارائه شده است. روش فوق قابلیت اعمال بر روی هر نوع دیکشنری احساسی حاوی کلمات احساسی امتیازدهی شده را داشته و باعث بهبود کارایی و عملکرد آن می‌گردد. بدیهی است که باتوجه‌به مدت‌زمان طولانی موردنیاز برای تولید دیکشنری‌های احساسی دستی که تنها حاوی کلمات احساسی باشند و همچنین هزینه فراوان پیاده‌سازی این قبیل از دیکشنری‌ها، امتیازدهی دستی به ترکیبات مختلف عبارات احساسی حاوی حتی یک تشدیدکننده -به‌عنوان‌مثال در یک دیکشنری احساسی حاوی ۷۵۰۰ کلمه احساسی با درنظرگرفتن ۲۰۰ تشدیدکننده تعداد ترکیبات Bigram و Trigram به ترتیب برابر با ۱۵۰۰۰۰۰ و ۳۰۰۰۰۰۰۰۰ حالت مختلف خواهد بود – را نیز ناممکن می‌سازد، لذا لزوم استفاده از روش‌های مبتنی بر یادگیری ماشینی در این حوزه آشکار می‌گردد.

[1] irrealis

برطبق تحقیقات ما و باتوجه‌به کارهای پیشین انجام شده به جرأت می‌توان این ادعا را نمود که تاکنون هیچ روشی که توانایی امتیازدهی تمامی حالات مختلف یک عبارت احساسی شامل چندین تشدیدکننده متوالی را با درنظرگرفتن اثر متقابل قطبیت کلمه احساسی بر تشدیدکننده و موقعیت تشدیدکننده پیش از کلمه احساسی داشته باشد ارائه نشده است. در معدود دیکشنری‌های احساسی بسیار جدید که قابلیت پشتیبانی از تشدیدکننده‌ها را دارا می‌باشند نه‌تنها به تأثیر متقابل یک کلمه احساسی و تشدیدکننده توجهی نشده است که در تقریباً تمامی روش‌های پیشین اثر موقعیت مکانی تشدیدکننده نیز موردتوجه قرار نگرفته است. در این دیکشنری‌ها غالباً اثر تشدیدکننده دوم چشم‌پوشی شده است و یا در صورت توجه به تشدیدکننده دوم اثری برابر با قرارگرفتن آن تشدیدکننده در موقعیت نخست از خود بروز می‌دهند. تنها روش و جدیدترین روش موجود که به‌عنوان مبنای این مقاله در مقایسه نتایج مورداستفاده قرار گرفته است روش senti-N-Gram می‌باشد که باتوجه‌به اشکالات متعدد، و ضعف در نحوه پیاده‌سازی قابلیت استفاده گسترده در عملیات تجزیه‌وتحلیل احساسات را دارا نیست. از جمله این اشکالات به پوشش بسیار ضعیف این دیکشنری برای عبارات Bigram و Trigram می‌توان اشاره کرد که به ترتیب تنها ۲۴۱۵ و ۳۰۵ اصطلاح را پشتیبانی می‌نماید. از سوی دیگر باتوجه‌به ترکیب منفی‌کننده‌ها در اصطلاحات احساسی تولید شده و امتیازدهی یکجای این اصطلاحات، غالب عبارات احساسی حاوی منفی‌کننده‌ای که در این واژه‌نامه امتیازدهی نشده باشند قابل امتیازدهی نیستند.

در طرح پیشنهادی، ضرایب تشدیدکننده طی چندین مرحله با استفاده از الگوی فازی اصلاح شد. همچنین، مکان تشدیدکننده نسبت به کلمه احساسی و اثر متقابل بین قطبیت کلمه احساسی و تشدیدکننده در نظر گرفته شد. بنابراین، برای هر تشدیدکننده چهار ضریب مختلف باتوجه‌به موقعیت اول یا دوم نسبت به کلمه احساسی و منفی یا مثبت بودن کلمه احساسی به دست آمد. این طرح منجر به تخصیص امتیاز کارآمدتر بر اساس مراحل تصحیح خطای فازی و یادگیری ماشین و بازبینی مجدد آن و پوشش طیف وسیع‌تری از ترکیبات و در نهایت افزایش دقت به ویژه در تشخیص اسناد منفی می‌شود. در نهایت، با استفاده از واژه‌نامه پیشنهادی و اعمال الگوریتم تعیین قطبیت اسناد بر روی چندین پایگاه‌داده استاندارد، نتایج حاصل از اعمال این الگوریتم‌ها

آن با دو واژه‌نامه به روز عمومی مقایسه شد. نتایج در همه معیارها از خود بهبود نشان داد. برای نمونه باتوجه‌به معیار Precision بصورت نسبی میتوان دید که روش پیشنهادی به‌صورت میانگین حدود ۱۳ درصد افزایش داشته است و افزایش ۵.۷۳٪ای F-measure را به دنبال داشته که بیانی مشخص از کارایی بهتر روش پیشنهادی دوم در مجموع اسناد با قطبیت مثبت و منفی محسوب می‌گردد.

۵-۸ کارهای آینده

- توسعه دیکشنری با استفاده از پسوندها و پیشوندهای تشدیدکننده و تضعیف‌کننده و امتیازدهی به اصطلاحات ایجاد شده جدید بر مبنای ترکیب روش‌های آماری و یادگیری تقویتی و یادگیری عمیق
- بررسی اثر افعال کمکی بر قطبیت و شدت کلمات احساسی و همچنین تعیین اثر مجزای هر یک از منفی‌کننده‌های مختلف بر یک کلمه احساسی
- بررسی اثر هم‌زمان قطبیت یک کلمه احساسی و منفی‌کننده بر تشدیدکننده در یک عبارت احساسی
- بررسی تأثیر بزرگ‌نویسی کلمات احساسی و همچنین تأثیر کلمات but و although در یک عبارت احساسی
- بررسی تأثیر تشدیدکننده بر یک عبارت احساسی باتوجه‌به جایگاه لغوی کلمه احساسی از لحاظ (فعل، قید، صفت یا اسم بودن) و اثر متقابل کلمه بر تشدیدکننده.
- بررسی تأثیر تکرار یک کلمه احساسی در یک جمله و همچنین بررسی اثر جایگاه نحوی کلمه احساسی بر تشدیدکننده و بررسی اثر بلوک‌های irrealis و ارائه یک الگوریتم امتیازدهی مبتنی بر دیکشنری تولید شده جدید.

فهرست مراجع

[1]. Hung C, Chen SJ. Word sense disambiguation based sentiment lexicons for sentiment classification. *Knowledge-Based Systems* 2016;110:pp. 224–32.

[2]. Dey A, Jenamani M, Thakkar JJ. Senti-N-Gram: An n-gram lexicon for sentiment analysis. *Expert Systems with Applications* 2018;103:pp. 92–105.

[3]. Abdi A, Shamsuddin SM, Hasan S, Piran J. Deep learning-based sentiment classification of evaluative text based on Multi-feature fusion. *Information Processing and Management* 2019;56:pp. 1245–59.

[4]. Kirchner AN. An exploration of a financial lexicon-based approach to sentiment analysis and its application to financial news and reports. *Orthern Illinois University* 2019:pp. 1–22.

[5]. Sun S, Luo C, Chen J. A review of natural language processing techniques for opinion mining systems. *Information Fusion* 2017;36:pp. 10–25.

[6]. Kiritchenko S, Zhu X, Mohammad SM. Sentiment analysis of short informal texts. *Journal of Artificial Intelligence Research* 2014;50:pp. 723–62.

[7]. Jurek A, Mulvenna MD, Bi Y. Improved lexicon-based sentiment analysis for social media analytics. *Security Informatics* 2015;4:pp. 1–13.

[8]. Polanyi L, Zaenen A. Contextual valence shifters. *AAAI Spring Symposium - Technical Report*, vol. SS-04-07, *Berlin/Heidelberg*: Springer-Verlag; 2005, p. 106–11.

[9]. Araque O, Corcuera-Platas I, Sánchez-Rada JF, Iglesias CA. Enhancing deep learning sentiment analysis with ensemble techniques in social applications. *Expert Systems with Applications* 2017;77:pp. 236–46.

[10]. Hutto CJJ, Gilbert EE. VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. *Eighth International AAAI Conference on Weblogs and Social Media*, 2014, p. 216–25.

[11]. Neviarouskaya A, Prendinger H, Ishizuka M. SentiFul: A lexicon for sentiment analysis. *IEEE Transactions on Affective Computing* 2011;2:pp. 22–36.

[12]. Taboada M, Brooke J, Voll K, Tofiloski M, Voll K, Stede M, et al. Lexicon-Based Methods for Sentiment Analysis. *Computational Linguistics* 2011;37:pp. 267–307.

[13]. Thelwall M, Buckley K, Paltoglou G, Cai D, Kappas A. Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology* 2010;61:pp. 2544–58.

[14]. Sarsam SM, Al-Samarraie H, Alzahrani AI, Alnumay W, Smith AP. A lexicon-based approach to detecting suicide-related messages on Twitter. *Biomedical Signal Processing and Control* 2021;65:pp. 1–8.

[15]. Jadon P, Bhatia D, Mishra DK. A New Methodology on Sentiment Analysis. *Advances in Intelligent Systems and Computing*, vol. 1187, Springer Science and Business Media Deutschland GmbH; 2021, p. 617–25.

[16]. Lin X-M, Ho C-H, Xia L-T, Zhao R-Y. Sentiment analysis of low-carbon travel APP user comments based on deep learning. *Sustainable Energy Technologies and Assessments* 2021;44:pp. 101014.

[17]. Alharbi JR, Alhalabi WS. Hybrid approach for sentiment analysis of twitter posts using a dictionary-based approach and fuzzy logic methods: Study case on cloud service providers. *International Journal on Semantic Web and Information Systems* 2020;16:pp. 116–45.

[18]. Mendon S, Dutta P, Behl A, Lessmann S. A Hybrid Approach of Machine Learning and Lexicons to Sentiment Analysis: Enhanced Insights from Twitter Data of Natural Disasters. *Information Systems Frontiers* 2021:pp. 1–24.

[19]. Sailunaz K, Alhajj R. Emotion and sentiment analysis from Twitter text. *Journal of Computational Science* 2019;36:pp. 101003.

[20]. Xu F, Pan Z, Xia R. E-commerce product review sentiment classification based on a naïve Bayes continuous learning framework. *Information Processing and Management* 2020:pp. 102221.

[21]. Tripathy A, Agrawal A, Rath SK. Classification of sentiment reviews using n-gram machine learning approach. *Expert Systems with Applications* 2016;57:pp. 117–26.

[22]. Giang NTP, Dien TT, Khoa TTM. Sentiment Analysis for University Students' Feedback. *Advances in Intelligent Systems and Computing* 2020;1130 AISC:pp. 55–66.

[23]. Vo T, Sharma R, Kumar R, Son LH, Pham BT, Bui DT, et al. Crime rate detection using social media of different crime locations and Twitter part-of-speech tagger with Brown clustering. *Journal of Intelligent and Fuzzy Systems* 2020;38:pp. 4287–99.

[24]. Vasudha Rani V, Sandhya Rani K. Identification of ontologies of prediabetes using SVM sentiment analysis. *Advances in Intelligent Systems and Computing* 2020;1054:pp. 535–50.

[25]. Mu R, Zheng Y, Zhang K, Zhang Y. Research on Customer Satisfaction Based on Multidimensional Analysis. *International Journal of Computational Intelligence Systems* 2021;14:pp. 605–16.

[26]. Yue L, Chen W, Li X, Zuo W, Yin M. A survey of sentiment analysis in social media. *Knowledge and Information Systems* 2018:pp. 1–47.

[27]. Bose R, Dey RK, Roy S, Sarddar D. Analyzing Political Sentiment Using Twitter Data. *Information and Communication Technology for Intelligent Systems Springer, Singapore* 2019;51:pp. 427–36.

[28]. Badaro G, Baly R, Hajj H, El-Hajj W, Shaban KB, Habash N, et al. A survey of opinion mining in Arabic: A comprehensive system perspective covering challenges and advances in tools, resources, models, applications, and visualizations. *ACM Transactions on Asian and Low-Resource Language Information Processing* 2019;18:pp. 1–52.

[29]. Kechaou Z, Ben Ammar M, Alimi A. Improving e-learning with sentiment analysis of users' opinions. *2011 IEEE Global Engineering Education Conference (EDUCON)*, IEEE; 2011, p. 1032–8.

[30]. Cambria E, Grassi M, Hussain A, Havasi C. Sentic Computing for social media marketing 2012:pp. 557–77.

[31]. Liu SM, Chen JH. A multi-label classification based approach for sentiment classification. *Expert Systems with Applications* 2015;42:pp. 1083–93.

[32]. Abdi A, Shamsuddin SM, Aliguliyev RM. QMOS: Query-based multi-documents opinion-oriented summarization. *Information Processing and Management* 2018;54:pp. 318–38.

[33]. de Melo T, da Silva AS, de Moura ES, Calado P. OpinionLink: Leveraging user opinions for product catalog enrichment. *Information Processing and Management* 2019;56:pp. 823–43.

[34]. Georgiadou E, Angelopoulos S, Drake H. Big data analytics and international negotiations: Sentiment analysis of Brexit negotiating outcomes. *International Journal of Information Management* 2020;51:pp. 102048.

[35]. Zhang W, Wang M, Zhu Y chun. Does government information release really matter in regulating contagion-evolution of negative emotion during public emergencies? From the perspective of cognitive big data analytics. *International Journal of Information Management*

2020;50:pp. 498–514.

[36]. Subrahmanian VSS, Reforgiato D. AVA: Adjective-Verb-Adverb Combinations for Sentiment Analysis. *IEEE Intelligent Systems*, vol. 23, 2008, p. 43–50.

[37]. Hiremath BN, Patil MM. Enhancing optimized personalized therapy in clinical decision support system using Natural Language Processing. *Journal of King Saud University - Computer and Information Sciences* 2020:pp. 1–18.

[38]. Greco F, Polli A. Emotional Text Mining: Customer profiling in brand management. *International Journal of Information Management* 2020;51:pp. 1–8.

[39]. Rathore AK, Ilavarasan PV. Pre- and post-launch emotions in new product development: Insights from twitter analytics of three products. *International Journal of Information Management* 2020;50:pp. 111–27.

[40]. Sindhu C, Sasmal B, Gupta R, Prathipa J. Subjectivity detection for sentiment analysis on Twitter data. *Lecture Notes in Networks and Systems* 2021;130:pp. 467–76.

[41]. Wang Y, Subhan F, Shamshirband S, Asghar MZ, Ullah I, Habib A, et al. Fuzzy-based sentiment analysis system for analyzing student feedback and satisfaction. *Computers, Materials and Continua* 2020;62:pp. 631–55.

[42]. You L, Peng Q, Xiong Z, He D, Qiu M, Zhang X. Integrating aspect analysis and local outlier factor for intelligent review spam detection. *Future Generation Computer Systems* 2020;102:pp. 163–72.

[43]. Lazib L, Qin B, Zhao Y, Zhang W, Liu T. A syntactic path-based hybrid neural network for negation scope detection. *Frontiers of Computer Science* 2020;14:pp. 84–94.

[44]. Carrillo-de-albornoz J, Plaza L. An emotion-based model of negation, intensifiers, and modality for polarity and intensity classification. *Journal of the American Society for Information Science and Technology* 2013;64:pp. 1618–33.

[45]. Upadhya BA, Udupa S, Kamath SS. Deep Neural Network Models for Question Classification in Community Question-Answering Forums. *2019 10th International Conference on Computing, Communication and Networking Technologies, ICCCNT 2019* 2019:pp. 1–6.

[46]. Shi Y, Tang Y ran, Long W. Sentiment contagion analysis of interacting investors: Evidence from China's stock forum. *Physica A: Statistical Mechanics and Its Applications* 2019;523:pp. 246–59.

[47]. Noekhah S, Salim N binti, Zakaria NH. Opinion spam detection: Using multi-iterative graph-based model. *Information Processing and Management* 2020;57:pp. 102140.

[48]. Li J, Lv P, Xiao W, Yang L, Zhang P. Exploring groups of opinion spam using sentiment analysis guided by nominated topics. *Expert Systems with Applications* 2021;171:pp. 114585.

[49]. Ligthart A, Catal C, Tekinerdogan B. Analyzing the effectiveness of semi-supervised learning approaches for opinion spam classification. *Applied Soft Computing* 2021;101:pp. 107023.

[50]. Prathap BR, Ramesha K. Twitter Sentiment for Analysing Different Types of Crimes. *2018 International Conference on Communication, Computing and Internet of Things (IC3IoT)*, IEEE; 2018, p. 483–8.

[51]. Abdi A, Hasan S, Shamsuddin SM, Idris N, Piran J. A hybrid deep learning architecture for opinion-oriented multi-document summarization based on multi-feature fusion. *Knowledge-Based Systems* 2021;213:pp. 106658.

[52]. Pang B, Lee L. Opinion mining and sentiment analysis. vol. 2. 2008.

[53]. Langacker RW. Observations and speculations on subjectivity. *BooksGoogleCom* 1985:pp. 109.

[54]. Lyons J. Language, meaning, and context. *Fontana Paperbacks*, 1981.

[55]. Conrad, Susan and Biber D. Adverbial marking of stance in speech and writing. *Evaluation in Text: Authorial Stance and the Construction of Discourse* 2000:pp. 56--73.

[56]. Biber D, Finegan E. Adverbial stance types in English. *Discourse Processes* 1988;11:pp. 1–34.

[57]. Martin J, White P. Appraisal in English: The language of evaluation 2005.

[58]. Scheibman J. Point of view and grammar: Structural patterns of subjectivity in American English conversation. John Benjamins Publishing; 2002.

[59]. Linguistics JW. Tracking Point of View in Narrative. *Computational linguistics*, vol. 20, 1994, p. 233.

[60]. Munezero M, Montero CS, Sutinen E, Pajunen J. Are they different? affect, feeling, emotion, sentiment, and opinion detection in text. *IEEE Transactions on Affective Computing*, vol. 5, 2014, p. 101–11.

[61]. Liu B. Sentiment Analysis and Opinion Mining. vol. 5. Morgan & Claypool Publishers; 2012.

[62]. Katiyar A, Cardie C. Investigating LSTMs for Joint Extraction of Opinion Entities and Relations 2016:pp. 919–29.

[63]. Liao J, Wang S, Li D, Li X. FREERL: Fusion relation embedded representation learning framework for aspect extraction. *Knowledge-Based Systems* 2017;135:pp. 9–17.

[64]. Wiegand M, Bocionek C, Ruppenhofer J. Opinion holder and target extraction on opinion compounds - A linguistic approach. *2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL HLT 2016 - Proceedings of the Conference* 2016:pp. 800–10.

[65]. Poria S, Cambria E, Gelbukh A. Aspect extraction for opinion mining with a deep convolutional neural network. *Knowledge-Based Systems* 2016;108:pp. 42–9.

[66]. Lingjia D, Wiebe J. Recognizing opinion sources based on a new categorization of opinion types. *IJCAI International Joint Conference on Artificial Intelligence* 2016;2016-Janua:pp. 2775–81.

[67]. Pang B, Lee L. A Sentimental Education : Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts. *In Proceedings of the 42nd annual meeting on Association for Computational Linguistics*, 2004, p. 1–8.

[68]. Wilson T, Wiebe J, Hoffmann P. Recognizing contextual polarity in phrase-level sentiment analysis. *Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing - HLT '05* 2005:pp. 347–54.

[69]. Liu MZ, Zhou FY, Chen K, Zhao Y. Co-attention networks based on aspect and context for aspect-level sentiment analysis. *Knowledge-Based Systems* 2021;217:pp. 106810.

[70]. Behdenna S, Barigou F, Belalem G. Document Level Sentiment Analysis : A survey 2017;4:pp. 1–8.

[71]. Pang B, Lee L, Vaithyanathan S, Rd H, Jose S. Thumbs up?: sentiment classification using machine learning techniques. *Proceedings of the ACL-02 conference on Empirical methods in natural language processing - EMNLP '02*, vol. 10, *Morristown, NJ, USA*: Association for Computational Linguistics; 2002, p. 79–86.

[72]. Mohammad SM, Turney PD. Emotions evoked by common words and phrases: using mechanical turk to create an emotion lexicon. *CAAGET '10 Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*, vol. 14, 2010, p. 26–34.

[73]. Kiritchenko S, Mohammad SM. Happy Accident: A sentiment composition lexicon for opposing

polarity phrases. *Proceedings of the 10th International Conference on Language Resources and Evaluation, LREC 2016*, 2016, p. 1157–64.

[74]. Parraga-Alava J, Caicedo RA, Gomez JM, Inostroza-Ponta M. An Unsupervised Learning Approach for Automatically to Categorize Potential Suicide Messages in Social Media. *2019 38th International Conference of the Chilean Computer Science Society (SCCC)*, vol. 38, 2019, p. 1–8.

[75]. Turney PD, D. P. Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews. *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics - ACL '02* 2001:pp. 417.

[76]. Thelwall M, Buckley K, Paltoglou G. Sentiment strength detection for the social web. *Journal of the American Society for Information Science and Technology* 2012;63:pp. 163–73.

[77]. Chae B. Insights from hashtag #supplychain and Twitter analytics: Considering Twitter and Twitter data for supply chain practice and research. *International Journal of Production Economics* 2015;165:pp. 247–59.

[78]. Zhang D, Levene M, Mudinas A, Zhang D, Levene M. Combining lexicon and learning based approaches for concept-level sentiment analysis. *Proceedings of the First International Workshop on Issues of Sentiment Discovery and Opinion Mining* 2012:pp. 1–8.

[79]. Dey A, Jenamani M, Thakkar JJ. Lexical TF-IDF: An n-gram Feature Space for Cross-Domain Classification of Sentiment Reviews. *International Conference on Pattern Recognition and Machine Intelligence*, vol. 10597 LNCS, 2017, p. 380–6.

[80]. Cambria E, Schuller B, Xia Y, Havasi C. New avenues in opinion mining and sentiment analysis. *IEEE Intelligent Systems* 2013;28:pp. 15–21.

[81]. Canhoto AI, Padmanabhan Y. ' We (don' t) know how you feel ' – a comparative study of automated vs . manual analysis of social media conversations. *Journal of Marketing Management* 2015;1376:pp. 1141–57.

[82]. Tirunillai S, Tellis GJ. Mining Marketing Meaning from Online Chatter : Strategic Brand Analysis of Big Data Using Latent Dirichlet Allocation. *Journal of Marketing Research* 2014;51:pp. 463–79.

[83]. Liu B. Opinion Mining and Sentiment Analysis. *Web Data Mining, Berlin, Heidelberg*: Springer Berlin Heidelberg; 2011, p. 459–526.

[84]. Pennebaker J, Boyd R, Jordan K, Blackburn K. The development and psychometric properties of LIWC2015. *LIWC2015* 2015:pp. 1–25.

[85]. Hartman JJ, Stone PJ, Dunphy DC, Smith MS, Ogilvia DM. The General Inquirer: A Computer Approach to Content Analysis. *American Sociological Review* 1967;32:pp. 484–98.

[86]. O'Leary DE. Blog mining-review and extensions: 'From each according to his opinion'. *Decision Support Systems* 2011;51:pp. 821–30.

[87]. Deng S, Sinha AP, Zhao H. Adapting sentiment lexicons to domain-specific social media texts. *Decision Support Systems* 2017;94:pp. 65–76.

[88]. Khoo CSG, Johnkhan SB. Lexicon-based sentiment analysis: Comparative evaluation of six sentiment lexicons. *Journal of Information Science* 2018;44:pp. 491–511.

[89]. Dhaoui C, Webster CM, Tan LP. Social media sentiment analysis: lexicon versus machine learning. *Journal of Consumer Marketing* 2017;34:pp. 480–8.

[90]. Zafra SMJ, Teresa Martin Valdivia M, Camara EM, Alfonso Urena Lopez L. Studying the Scope of Negation for Spanish Sentiment Analysis on Twitter. *IEEE Transactions on Affective Computing* 2015;10:pp. 129–41.

[91]. Benamara F, Chardon B, Mathieu Y, Popescu V, Asher N. How do Negation and Modality Impact on Opinions? *Proceedings of the Workshop on Extra-Propositional Aspects of Meaning in Computational Linguistics. Association for Computational Linguistics*, 2012, p. 10–8.

[92]. Barnes J, Velldal E, Øvrelid L. Improving Sentiment Analysis with Multi-task Learning of Negation. *Natural Language Engineering* 2021;27:pp. 249–69.

[93]. Wu C, Wu F, Liu J, Yuan Z, Wu S, Huang Y. THU_NGN at SemEval-2018 Task 1: Fine-grained Tweet Sentiment Intensity Analysis with Attention CNN-LSTM. *Proceedings of The 12th International Workshop on Semantic Evaluation, Stroudsburg, PA, USA*: Association for Computational Linguistics; 2018, p. 186–92.

[94]. Mudgal P, Khunteta A. Handling Double Intensifiers in Feature-Level Sentiment Analysis Based on Movie Reviews. *5th International Conference on Artificial Intelligence and Applications*, 2020, p. 383–92.

[95]. Lu Q, Zhu Z, Xu F, Zhang D, Wu W, Guo Q. Bi-gru sentiment classification for chinese based on grammar rules and bert. *International Journal of Computational Intelligence Systems* 2020;13:pp. 538–48.

[96]. Satthar FS. Modelling SO-CAL in an inheritance-based sentiment analysis framework. *2015 Imperial College Computing Student Workshop (ICCSW 2015)* 2015:pp. 46–53.

[97]. Liu J, Seneff S. Review sentiment scoring via a parse-and-paraphrase paradigm. *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: Volume 1*, 2009, p. 161–9.

[98]. Cruz NP, Taboada M, Mitkov R. A Machine-Learning Approach to Negation and Speculation Detection for Sentiment Analysis. *Journal of the Association for Information Science and Technology* 2015;67:pp. 2118–36.

[99]. Blanco E, Moldovan D. Some Issues on Detecting Negation from Text. *Proceedings of the 24th Florida Artificial Intelligence Research Society Conference (FLAIRS-24)* 2011:pp. 228–33.

[100]. Mukhtar N, Khan MA, Chiragh N, Nazir S. Identification and handling of intensifiers for enhancing accuracy of Urdu sentiment analysis. *Expert Systems* 2018;35:pp. 1–12.

[101]. Dragut EC, Fellbaum C. The Role of Adverbs in Sentiment Analysis. *Proceedings of Frame Semantics in NLP: A Workshop in Honor of Chuck Fillmore*, 2014, p. 38–41.

[102]. Taboada M. Sentiment Analysis: An Overview from Linguistics. *Annual Review of Linguistics* 2016;2:pp. 325–47.

[103]. Brooke J, Tofiloski M, Taboada M. Cross-linguistic sentiment analysis: From English to Spanish. *International Conference Recent Advances in Natural Language Processing, RANLP*, 2009, p. 50–4.

[104]. Mukhtar N, Khan MA. Urdu sentiment analysis using supervised machine learning approach. *International Journal of Pattern Recognition and Artificial Intelligence* 2018;32:pp. 1–15.

[105]. Asghar MZ, Khan A, Ahmad S, Qasim M, Khan IA. Lexicon-enhanced sentiment analysis framework using rule-based classification scheme. *PLoS ONE* 2017;12:pp. 1–22.

[106]. Brooke J. A Semantic Approach To Automated Text Sentiment Analysis. 2009.

[107]. Chung CK, Pennebaker JW. Linguistic Inquiry and Word Count (LIWC). *Applied Natural Language Processing* 2013:pp. 206–29.

[108]. Thisone CM. Sentiment Analysis Using a Novel Human Computation Game. *Proceedings of the 3rd Workshop on the People's Web Meets NLP: Collaboratively Constructed Semantic Resources and Their Applications to NLP* 2012:pp. 1–9.

[109]. Hong Y, Kwak H, Baek Y, Moon S. Tower of babel: A crowdsourcing game building sentiment

lexicons for resource-scarce languages. *Proceedings of the 22nd International Conference on World Wide Web* 2013:pp. 549–56.

[110]. Giachanou A, Gonzalo J, Crestani F. Propagating sentiment signals for estimating reputation polarity. *Information Processing & Management* 2019;56:pp. 102079.

[111]. Lee J-H, Ko S-K, Han Y-S. SALNet: Semi-Supervised Few-Shot Text Classification with Attention-based Lexicon Construction 2021:pp. 13189–97.

[112]. Dong X, De Melo G. A helping hand: Transfer learning for deep sentiment analysis. *ACL 2018 - 56th Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference (Long Papers)*, vol. 1, 2018, p. 2524–34.

[113]. Mohamad Beigi O, Moattar MH. Automatic construction of domain-specific sentiment lexicon for unsupervised domain adaptation and sentiment classification. *Knowledge-Based Systems* 2021;213:pp. 106423.

[114]. Araque O, Zhu G, Iglesias CA. A semantic similarity-based perspective of affect lexicons for sentiment analysis. *Knowledge-Based Systems* 2019;165:pp. 346–59.

[115]. Xing FZ, Pallucchini F, Cambria E. Cognitive-inspired domain adaptation of sentiment lexicons. *Information Processing and Management* 2019;56:pp. 554–64.

[116]. Liu B. Sentiment analysis and subjectivity. *Handbook of Natural Language Processing* 2010:pp. 627–66.

[117]. Wu F, Huang Y, Yan J. Active sentiment domain adaptation. *ACL 2017 - 55th Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference (Long Papers)* 2017;1:pp. 1701–11.

[118]. Pennebaker J, Chung CK, Ireland M, Gonzales A, Booth R. The Development and Psychometric Properties of LIWC2007. *He Development and Psychometric Properties of Liwc2007* 2007:pp. 1–25.

[119]. Tumasjan A, Sprenger TO, Sandner PG, Welpe IM. Predicting Elections with Twitter : What 140 Characters Reveal about Political Sentiment. *Proceedings of the International AAAI Conference on Web and Social Media*, 2010, p. 178–85.

[120]. Gamon M, Counts S, Choudhury M De, Gamon M, Counts S, Horvitz E. Predicting Depression via Social Media. *Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media*, vol. 2, 2013, p. 128–37.

[121]. De Choudhury M, Counts S, Horvitz E. Predicting postpartum changes in emotion and behavior via social media. *Conference on Human Factors in Computing Systems - Proceedings, New York, New York, USA*: ACM Press; 2013, p. 3267–76.

[122]. Hu M, Liu B, Liu MH and B. Mining and summarizing customer reviews. *Proceedings of the 2004 ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '04* 2004;4425:pp. 168.

[123]. Bradley M, Lang P. Affective norms for English words (ANEW): Instruction manual and affective ratings. *Affective norms for English words (ANEW): Instruction manual and affective ratings*, vol. 30, 1999, p. 25–36.

[124]. Baccianella S, Esuli A, Sebastiani F. SENTIWORDNET 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. *Proceedings of the 7th International Conference on Language Resources and Evaluation, LREC 2010* 2010:pp. 2200–4.

[125]. Garapati A, Sai M. SentiPhraseNet: An extended SentiWordNet approach for Telugu sentiment analysis. *International Journal of Advance Research* 2019;5:pp. 433–6.

[126]. Mohammad SM, Kiritchenko S, Zhu X. NRC-Canada: Building the state-of-the-art in sentiment

analysis of tweets. **SEM 2013 - 2nd Joint Conference on Lexical and Computational Semantics*, vol. 2, 2013, p. 321–7.

[127]. Orme B. Maxdiff analysis: Simple counting, individual-level logit, and hb. *Sawtooth Software*, 2009, p. 1–7.

[128]. Flynn TN. Best Worst Scaling : Theory and Methods. 2014.

[129]. Sweeney K, Whissell C. A Dictionary Of Affect IN Language : I. Establishment AND Preliminary Validation. *Perceptual and Motor Skills* 1984;59:pp. 695–8.

[130]. WHISSELL CM. THE DICTIONARY OF AFFECT IN LANGUAGE. *The Measurement of Emotions*, Elsevier; 1989, p. 113–31.

[131]. Cambria, Erik, Robert Speer, Catherine Havasi AH. SenticNet : A Publicly Available Semantic Resource for Opinion Mining 2010:pp. 14–8.

[132]. Cambria E, Havasi C, Hussain A. SenticNet 2: A semantic and affective resource for opinion mining and sentiment analysis. *Proceedings of the 25th International Florida Artificial Intelligence Research Society Conference, FLAIRS-25* 2012:pp. 202–7.

[133]. Cambria E, Hussain A. Sentic Computing. vol. 2. *Dordrecht*: Springer Netherlands; 2012.

[134]. Akkaya C, Wiebe J. Subjectivity Word Sense Disambiguation. *Proceedings of the 2009 conference on empirical methods in natural language processing*, 2009, p. 190–9.

[135]. Devitt A, Ahmad K. Is there a language of sentiment? An analysis of lexical resources for sentiment analysis. *Language Resources and Evaluation* 2013;47:pp. 475–511.

[136]. Lorge I, Thorndike E. The teacher's word book of 30 000 words. 1963.

[137]. Lasswell H, Namenwirth J. The Lasswell Value Dictionary (3 vols.). New Haven: Yale University. *The Lasswell Value Dictionary (3 Vols) New Haven: Yale University* 1968.

[138]. Hutto CJ. Computing for social science: Characterizing, quantifying, and analyzing social phenomena in technology mediated communications. Georgia Institute of Technology, 2018.

[139]. Taboada M, Brooke J, Tofiloski M, Voll K, Stede M. Lexicon-basedmethods for sentiment analysis. *Computational Linguistics* 2011;37:pp. 267–307.

[140]. Taboada M, Anthony C, Voll K. Methods for Creating Semantic Orientation Databases. *Proceeding of LREC-06, the 5th International Conference on Language Resources and Evaluation* 2006:pp. 427–32.

[141]. Taboada M, Grieve J. Analyzing Appraisal Automatically. *AAAI Spring Symposium - Technical Report* 2004;2004:pp. 158–61.

[142]. Riloff E, Wiebe J. Learning Extraction Patterns for Subjective Expressions. *In Proceedings ofthe 2003 Conference on Empirical Methods in Natural Language Processing (EMNLP-03)*, 2003.

[143]. Bernard JR. Macquarie thesaurus: The. Macquarie Library; 1984.

[144]. Brants, Thornsten and AF. Web 1T 5-gram version 1. Linguistic Data Consortium. *Philadelphia LDC2006T13* 2006.

[145]. Mohammad SM, Turney PD. Crowdsourcing a word-emotion association lexicon. *Computational Intelligence* 2013;29:pp. 436–65.

[146]. Strapparava C, Valitutti A. WordNet-Affect: An affective extension of WordNet. *Proceedings of the 4th International Conference on Language Resources and Evaluation, LREC 2004* 2004;4:pp. 1083–6.

[147]. Magnini B, Cavagli G. Integrating Subject Field Codes into WordNet Subject Field Codes organization. *Lrec* 2000:pp. 1413–8.

[148]. Neviarouskaya A, Prendinger H, Ishizuka M. Textual Affect Sensing for Sociable and Expressive. *International Conference on Affective Computing and Intelligent Interaction Springer, Berlin, Heidelberg* 2007:pp. 218–9.

[149]. Heerschop B, Goossen F, Frasincar F, Kaymak U, Hogenboom A, Jong F De, et al. Polarity Analysis of Texts using Discourse Structure Bas. *Proceedings of the 20th ACM International Conference on Information and Knowledge Management - CIKM '11* 2011:pp. 1061–70.

[150]. Hogenboom A. Using Rhetorical Structure in Sentiment Analysis Categories and Subject Descriptors. *Communications of the ACM* 2015;58:pp. 69–77.

[151]. Heerschop B, Hogenboom A, Frasincar F. Sentiment lexicon creation from lexical resources. *Lecture Notes in Business Information Processing* 2011;87 LNBIP:pp. 185–96.

[152]. Esuli A, Sebastiani F. SentiWordNet: a high-coverage lexical resource for opinion mining. *Evaluation* 2007;17:pp. 1–26.

[153]. Khan A, Baharudin B, Khan K. Sentiment Classification Using Sentence-level lexical Based semantic orientation of Online reviews. *Trends in Applied Sciences Research* 2011;6:pp. 1141–57.

[154]. Khan A, Baharudin B, Khan K. Sentiment classification using sentence-level semantic orientation of opinion terms from blogs. *2011 National Postgraduate Conference - Energy and Sustainability: Exploring the Innovative Minds, NPC 2011* 2011;6:pp. 1141–57.

Abstract

One of the main factors in all text polarity detection methods is the sentimental phrases in that text. Most dictionary-based methods have focused on extracting sentiment words, determining their polarity, and finally assigning the sentimental score to these words. On the other hand, almost all researchers in sentiment analysis emphasize that a text can contain linguistic phenomena that affect text polarity, although they do not contain any sentimental polarity. Negators and intensifiers are two important linguistic phenomena that can increase the accuracy and efficiency of the polarity detection process. Lacking accurate attention to the impact of negators and intensifiers and the position of these phenomena in a sentiment phrase and neglecting the effect of the polarity of sentiment words on these linguistic phenomena are some disadvantages that reduce the efficiency of existing sentiment dictionaries. According to the above details, this study has proposed a new high-coverage semi-supervised sentiment dictionary that provides the ability to calculate senti-n-gram expressions made from a combination of sentimental words and several intensifiers and negators. Sentimental word scores were extracted from the manual VADER dictionary. The intensifiers coefficients were extracted automatically using a hybrid fuzzy-statistical method based on the intensifier location and the sentiment word polarity. In the proposed method, four different coefficients are assigned to each intensifier. The simulation results show that compared to the latest methods in this area, the proposed dictionary has at least 9% improvement in standard databases. The improvement in the field of document with negative polarity is 49.6%.

Keywords

Sentiment analysis, Sentiment dictionary, Intensifier, Intensifier location, consecutive intensifiers, Effect of sentimental word polarity on the intensifier

Shahrood University of Technology

Faculty of Kharazmi Campus

Ph.D. Thesis in Artificial Intelligence

Designing an English sentimental dictionary with the ability to use consecutive intensifiers and negative words

By: Hamed Zargari

Supervisor:
Dr. Morteza Zahedi

Advisor:
Dr. Marziea Rahimi

September 2021