The 9th International Conference on Web Research
ICWR2023
May 3th-4th, 2023, Tehran, Iran

A Hybrid Method for Aspect-based Sentiment Analysis in Social Networks

[1]Ali Salehi, [2]Alireza Rezvanian*

[1]Master of Information Technology, Department of Computer Engineering, University of Science and Culture, Tehran, Iran
a.s.dittany@gmail.com

[2] Assistant Professor, Department of Computer Engineering, University of Science and Culture, Tehran, Iran
rezvanian@usc.ac.ir

Abstract

In this paper, due to the importance of analyzing users' sentiments in social networks, a hybrid method is proposed in 5 steps. In this regard, after pre-processing the text, the VADER dictionary is used to determine the polarity, then BoW and TF-IDF are used for feature extraction. Also, an ensemble classifier based on logistic regression, decision tree, and Naïve Bayes is designed, and finally, several clustering methods are used for aspect-based sentiment analysis. For evaluation, the proposed algorithm is conducted on the Amazon, Twitter, and Reddit datasets the results are reported in terms of run-time and precision.

Keywords: Social Networks, Sentiment Analysis, Machine Learning, Clustering, Ensemble Classification.

The 9th International Conference on Web Research
ICWR2023
May 3th-4th, 2023, Tehran, Iran

یک روش ترکیبی برای تحلیل احساسات زمینه‌ای در شبکه‌های اجتماعی

علی صالحی[۱]، علیرضا رضوانیان*[۲]

[۱] کارشناسی ارشد مهندسی فناوری اطلاعات، گروه مهندسی کامپیوتر، دانشگاه علم و فرهنگ، تهران، ایران
a.s.dittany@gmail.com

[۲] استادیار، گروه مهندسی کامپیوتر، دانشگاه علم و فرهنگ، تهران، ایران
rezvanian@usc.ac.ir

چکیده

در این مقاله با توجه به اهمیت تحلیل احساسات کاربران در شبکه‌های اجتماعی، یک روش ترکیبی در ۵ مرحله پیشنهاد شده است، به طوریکه پس از پیش پردازش متن، از فرهنگ لغت VADER برای تعیین قطبیت استفاده شده است، سپس با استفاده از BoW و TF-IDF برای استخراج ویژگی استفاده می‌شود و برای طبقه بندی آنسامبلی از رگرسیون لاجستیک، درخت تصمیم و بیز ساده طراحی شده است و در نهایت برای تحلیل زمینه‌ای احساسات از چندین روش خوشه بندی استفاده شده است. برای ارزیابی روش پیشنهادی از داده‌های آمازون، توییتر و ردیت استفاده شده است و از نظر زمان اجرای الگوریتم و دقت کارایی آن گزارش شده است.

کلمات کلیدی

شبکه‌های اجتماعی، تحلیل احساسات، یادگیری ماشین، خوشه‌بندی، طبقه‌بندی گروهی.

۱- مقدمه

شبکه‌های اجتماعی اغلب به‌عنوان منابع وسیع داده شمرده می‌شوند، زیرا در آن ها به‌صورت روزانه داده ها، در مقیاس‌های مختلف تولید می‌شوند. شبکه‌های اجتماعی همچنین نقش حیاتی در ارتباطات انسانی ایفا می‌کند و کاربردهای تجاری را بهبود می‌بخشد. در این دنیای پویا، متن بزرگ‌ترین راه ارتباطی است و بیش از ۱۸.۲ میلیون پیام متنی در یک دقیقه ارسال می‌شود [1]. بازاریابی عاطفی به عنوان یک مفهوم کلیدی بازاریابی در رویکردهای معاصر بازاریابی با تمرکز بر اهمیت رابطه عاطفی بین شرکت ها و مصرف کنندگان توسعه یافته است. این مفهوم امکان شناسایی و مبادله واقعی احساسات بین شرکت‌ها و مصرف کنندگان آنها را فراهم می‌کند. انجام تحلیل موثر داده‌ها چالش دیگری است که با ابداع روش‌های خودکار برای استخراج احساسات یا نظرات مورد بررسی قرار گرفته است. تحلیل احساسات به عنوان بخشی از بازاریابی عاطفی را می‌توان به صورت تکنیکی برای استخراج خودکار عواطف و احساسات از نظرات و نوشتار کاربران تعریف کرد. مزیت اصلی تحلیل احساسات شناسایی و طبقه‌بندی مؤثر نگرش‌ها و احساسات (شامل ۳ حالت مثبت، منفی یا خنثی) کاربران در نوشتار برای تعیین نگرش آنها نسبت به محصولات، موضوعات یا خدمات است، که می‌تواند نقشی کلیدی در بازاریابی عاطفی داشته باشد.

امروزه برانگیختن مستقیم احساسات مثبت در کاربران می‌تواند باعث افزایش نرخ تعامل بشود و اهمیت این موضوع مورد توجه بازاریابان قرار گرفته است. تحقیقات موجود در این زمینه نشان می‌دهند؛ روش‌های ارائه شده برای اندازه گیری احساسات مشتریان و کاربران ابعاد مختلفی را شامل می شود

یکی از محدودیت اصلی مطالعات ارائه شده این است که فقط برای توییت‌هایی با موضوعی خاص و بدون درنظرگرفتن موضوعات دیگر بررسی شده اند. در برخی موارد، دقت آن پایین است و این موضوع بیشتر به این دلیل است که توزیع بین برچسب‌ها در منابع داده متعادل نیست. رویکردها عموماً می‌تواند به از تحلیل زمینه‌ای توییت‌ها پشتیبانی کند، اما نتایج نشان می دهد که این موضوع دائمی و بهینه نیست.

تحلیل احساسات یک زمینه تحقیقاتی در جریان در زمینه متن کاوی بازاریابی عاطفی است. تحلیل احساسات شناخت محاسباتی نظرات، احساسات و ذهنیت موجود متن است [2]. تحلیل احساسات را می توان بر اساس مجموعه داده های مورد استفاده برای پردازش دسته بندی کرد. تحلیل احساسات، به عنوان یکی از محبوب ترین ابزارها در حوزه پردازش زبان طبیعی، کاوش های زیادی را به خود جلب کرده است [3]. تحلیل احساسات نه تنها برای بررسی محصول اعمال می شود، بلکه می تواند در پیش بینی سهام، بررسی فیلم، مقالات خبری یا بحث های سیاسی نیز اعمال شود [4]. جدای از تحلیل احساسات در سطح جمله، تحلیل احساسات زمینه‌ای نیز، به طور گسترده موردمطالعه قرار گرفته است. این تحلیل وظیفه از دو کار فرعی شامل استخراج اصطلاحات و طبقه‌بندی احساسات در سطح زمینه‌ها تشکیل شده است. هدف استخراج اصطلاحات استخراج نهادهای صاحب‌نظر از بررسی‌ها است. مطالعات اولیه استخراج اصطلاحات عمدتاً بر روی استفاده از تکنیک‌های مهندسی ویژگی برای به‌دست‌آوردن ویژگی‌های مبتنی بر زبان متمرکز است. هدف طبقه‌بندی احساسات در سطح زمینه‌ای، تعیین قطبیت احساسات عبارات زمینه‌ای در بررسی‌ها است.

در این مقاله تلاش شده است تا رویکرد تحلیل زمینه ای در قالب یک روش ترکیبی ارائه شود. از این جهت، در این مقاله با استفاده از فرهنگ لغت VADER، الگوریتم طبقه بندی گروهی، روش کیسه کلمات و الگوریتم های خوشه بندی کاندید به تحلیل زمینه ای داده های کاربران پرداخته شده است و سپس کارایی روش پیشنهادی براساس معیارهای مختلف مورد بررسی قرار گرفته است.

۲- پیشینه پژوهش

برخی از مهمترین تحقیقات و مطالعاتی که در سالهای اخیر برای تحلیل احساسات در شبکه های اجتماعی بیشتر با تمرکز بر بازاریابی انجام شده است، در ادامه مورد بررسی قرار گرفته است.

شاروئی و همکاران در [5] و [6] روش جدیدی با نام الگوریتم تحلیل زمینه‌ای خودکار و خوشه‌بندی گروهی (ACAEC) برای تحلیل احساسات کاربران معرفی نمودند. در این روش با استفاده از ترکیب خوشه بندی متوالی پنجره‌ها در [5] و خوشه‌بندی موازی پنجره‌ها در [6] با روش K-MEANS به تحلیل احساسات کاربران می‌پردازند و برای اثبات کارایی مدل خود وجود احساسات در نظرات مشتریان را بررسی کرده‌اند.

راکلج و همکاران در [7] تاثیرات احساسی تجربه کاربران در محصولات مصرفی را با انجام ۷ مطالعه بررسی کرده اند. در سه مدل از مطالعات داده هایشان را خودشان جمع آوری کردند و در ۴ مدل دیگر از پایگاه داده استفاده کرده اند تا نشان بدهند مشتریان در عطش خود برای رسیدن به تخصص، ممکن است خود را از احساساتی که مصرف به ارمغان می آورد محروم کنند. تحقیق آنها در درک جنبه احساسی‌تر کسب و به کارگیری تخصص و در درک بهتر اینکه چگونه تخصص مصرف کنندگان تجربه مصرف آن ها را شکل می‌دهد نتایج قابل توجهی داشته است. راکلج و همکاران همچنین در [8] بیان می‌کنند که محققان، بازاریاب‌ها و مصرف کنندگان باور دارند با تقویت محتوای احساسی می‌توانند پخش اطلاعات و تصمیم خرید مشتریان تاثیر گذاشت. فهم اینکه این موضوع چرا و چه زمانی برعکس عمل می کند. می‌تواند افراد را قادر سازد تا بتوانند پیش بینی کنند کدام نظرات محبوب خواهند شد. آنها همچنین مقایسه مناسبی از کارهای گذشته ارائه داده اند و در مطالعه‌شان برای بررسی میزان احساسات از روش فرهنگ لغت استفاده کردند.

ژیانگ و همکاران در [9] بیان می کنند که مجموعه‌های دارای حاشیه‌نویسی مناسب برای تحلیل احساسات مبتنی بر یادگیری ماشین ضروری هستند و تقویت داده‌های لغوی یک رویکرد ساده اما کارآمد برای فزونی داده های برچسب‌گذاری‌شده خاص دامنه برای آموزش است. در تحقیقات داده فزونی قبلی، مطالعه روشنی در مورد چگونگی (الف) انتخاب کلمات مناسب برای جایگزینی و (ب) تعیین جایگزین از همسایگان معنایی آنها برای دستیابی به بهترین عملکرد وجود ندارد. آنها برای پرداختن به این دو مشکل، یک رویکرد جایگزینی واژگانی متمرکز بر PLSDA، پیشنهاد می کنند. این روش مجموعه داده آموزشی را با نمونه‌های صحیح نحوی و معنایی غنی می‌کند. آنها همچنین برای بهینه سازی عملکرد، اثرات استراتژی های مختلف داده فزونی را نیز به طور عمیق مورد بررسی قرار می دهند. تیان و همکارانشان در [3] یک مدل جدید برای تحلیل احساسات مبتنی بر تحلیل زمینه ای میان دامنه ای ارائه کرده اند. آنها از Bert به عنوان لایه جاسازی استفاده کرده اند و بازنمایی های متنی برای کلمات ایجاد می کند. آنها از Bi-LSTM و CRF مانند لایه ABSA متقاطع برای تحلیل احساسات خروجی روش خود استفاده کرده اند. همچنین، قابلیت دامنه محور بودن روش شان با تولید یک شبکه پارامتر برای دستیابی به یادگیری انتقالی مبتنی بر حلت، از طریق استراتژی یادگیری چندگانه معرفی می‌شوند. از منظر آزمایش‌ها، آنها اهمیت و اثربخشی همه مؤلفه‌ها را در کار ABSA بین دامنه‌ای نشان دادیم. (از مطالعه موردی و مطالعه توسعه یافته شده، برای تشخیص دانش دامنه توسط مدل پیشنهادیشان استفاده کرده اند).

بهرا و همکارانشان در [4] یک معماری شبکه عصبی متشکل از CNN و LSTM برای پیش بینی احساسات نظرات مشتریان پیشنهاد داده اند. مزیت اصلی مدل آن ها این است که به یک دامنه خاص محدود نمی شود. بنابراین، همان مدل را می توان برای بررسی محصول و همچنین بررسی خدمات بدون تاثیر در عملکرد مدل آموزش داد. در این مدل به مهندسی ویژگی های دستی پیچیده نیاز نیست، بنابراین از دامنه محور بودن اجتناب می شود و این به دلیل استفاده از مدل تعبیه کلمه از پیش آموزش برای تبیین بردار ویژگی ورودی است. مدل آنها در مرحله بعد، با استفاده از لایه CNN قبل از شبکه LSTM به شناسایی ویژگی های مهم تنها از طریق بردار تعبیه شده اقدام می کنند، بنابراین زمان عملکرد آموزش را به شدت بهبود می بخشد و از این رو محاسبات آن را امکان پذیر می کنند. در مرحله آخر، با استفاده از لایه شبکه

LSTM به ساخت مدل با مطالعه ترتیبات متوالی اقدام به بررسی می کنند که این روش را جایگزین روش معمول در نظر گرفتن کلمات یا عبارات تنها بیان می کنند. بنابراین این مدل نه تنها در بررسی خود زمینه موضوعی را در بر می گیرد و بلکه در مورد زمینه هایی مانند نفی و همچنین طعنه نیز عملکرد بهتری دارد. در [10] چالش بیان احساس پیچیده برای جمله های طولانی یا چند جنبه‌ای مطرح شده است و سپس یک مدل شبکه عصبی گرافی پیشنهاد شده است که با در نظر گرفتن ساختار گرافی، ارتباط اسامی را نیز می‌تواند لحاظ کند. در این حالت نتایج شبیه سازی حاکی از موفقیت بیشتر برای پاسخ‌های احساسات خاص است. با توجه به آنکه بررسی بخش زیادی از متون در نواحی شمال آفریقا و خاورمیانه شبکه‌های اجتماعی مربوط به زبان عربی است، تحلیل احساسات زمینه‌ای در شبکه های اجتماعی برای زبان عربی نیز مورد توجه قرار گرفته است و در [11] مروری بر چالش‌ها، پیچیدگی‌ها، گرامر و ساختار زبان عربی انجام شده است و سپس مطالعات انجام شده بر روی این زمینه مورد بررسی قرار گرفته است.

با وجود آنکه تحقیقات مختلفی در زمینه تحلیل احساسات انجام شده است، پشتیبانی از تحلیل زمینه ای به عنوان یک موضوع کمتر کاوش شده باقی مانده است. همچنین می توان مشاهده کرد که روش های ارائه شده برای اندازه گیری احساسات برای دنبال کنندگان و مشتریان ابعاد مخلفی را شامل می شود که شامل محدودیت هایی هستند. یکی از محدودیت اصلی مطالعات ارائه شده این است که تحقیقات عموما در مجموعه داده های تک دامنه ای انجام شده است و این موضوع باعث می شود تا نیاز به پژوهش های زمینه ای در مجموعه داده های میان دامنه ای بیشتر احساس بشود. همچنین در مطالعات میان دامنه ای، نتایج از لحاظ دقت و زمان اجرا نیازمند بهبود هستند و این بیشتر به این دلیل است که توزیع بین برچسب‌ها در مجموعه داده ها متعادل نیست. علاوه بر این برچسب گذاری کاری مشکل و پردردسر بوده، رویکردها نمی‌توانند همیشه به طور مفیدی از تحلیل زمینه‌ای داده‌ها پشتیبانی کنند.

در این مقاله تلاش شده است تا رویکرد تحلیل زمینه ای در قالب یک روش ترکیبی ارائه شود. از این جهت، در این مقاله با استفاده از فرهنگ لغت VADER، الگوریتم طبقه بندی گروهی، روش کیسه کلمات و الگوریتم های خوشه بندی کاندید به تحلیل زمینه ای داده های کاربران پرداخته شده است و سپس کارایی روش پیشنهادی براساس معیارهای مختلف مورد بررسی قرار گرفته است..

۳- روش پیشنهادی

روش پیشنهادی در ۵ مرحله شامل مراحل زیر است که در ادامه همراه با جزئیات شرح داده شده است.

۳-۱- پیش پردازش داده‌ها

در این مرحله داده ها برای استفاده در مراحل بعدی آماده می شوند، به این نحو که ابتدا داده های خالی حذف شده، سپس تنها داده هایی که زبان انگلیسی نوشته نشده اند حذف می‌شوند، سپس علائم و داده های اضافی از جملات حذف می شود و در نهایت نیز فواصل بین کلمات و جملات اصلاح می شود؛ تا به‌این‌ترتیب جملات مرتب و آماده استفاده بشوند

۳-۲- تعیین قطبیت جملات با استفاده از فرهنگ لغت Vader

در این مرحله، از واژگان احساسات تخصصی VADER [12] استفاده می‌شود که نتایج موفقی نیز برای آن گزارش شده است. ازآنجایی‌که از SentiWordNet بر اساس synsetهای WordNet استفاده می‌شد، همان کلمه در SentiWordNet می‌تواند امتیازهای متفاوتی داشته باشد، زیرا ممکن است در چندین synset ظاهر شود. استفاده از مقدار قطبیت یک synset مجزا نیاز به رویکرد تحلیل ابهام متن دارد، در این روش با استفاده از قطبیت سه گانه که به‌صورت خودکار توسط معیار compound فرهنگ لغت VADER تولید می‌شود (که برآیند معیارهای سه‌گانه قطبیت مثبت، منفی و خنثی است.) جملات برچسب زنی می شوند، سپس در اجرای فازهای طبقه بندی گروهی و خوشه‌بندی از این معیار به این نحو که عدد مثبت نماینده احساسات مثبت و عدد منفی نماینده احساسات منفی و عدد صفر نماینده خنثی بودن احساسات است، استفاده می‌شود.

۳-۳- ساخت مدل برداری با استفاده از BoW یا TF-IDF

در این مرحله مدل فضای برداری یک نمایش متداول در پردازش متن است که در آن جملات ویژگی محسوب شده و اسناد مشاهده کننده هستند. در الگوریتم پیشنهادی، اسناد با صفت ها و قیدهای خود نشان داده می شوند که بخش بیان کننده احساسات در گفتار هستند. برای به‌دست‌آوردن دقیق ترین نتایج، از انواع نمایش های ماتریسی استفاده شده است. بنابراین از روش BoW برای استخراج فراوانی کلمات استفاده شده است و از روش امتیازدهی TF-IDF استفاده شده است که در این روش، اهمیت یک کلمه خاص را نه تنها در سند بلکه در بدنه از طریق فراوانی معکوس سند را نیز می‌سنجد.

۳-۴- استفاده از طبقه‌بندی گروهی برای استخراج ویژگی

طبقه‌بندی گروهی رأی‌گیری یکی از روش‌هایی است که در طبقه‌بندی گروهی دسته ای استفاده می‌شود که در آن چندین مدل با هم ترکیب می‌شوند و پیش‌بینی نهایی بر اساس رأی اکثریت مدل‌ها انجام می‌شود. از این روش به‌ویژه برای زمانی که داده های هدف پیچیده هستند و می توانند به اشکال مختلف نمایش داده شوند، استفاده می شود. این روش می تواند با ارائه مدلی با تنوع بالا باعث افزایش دقت و قدرت تعمیم مدل شود. در مدل پیشنهادی از این روش برای استخراج ویژگی–های مستقل از دامنه مدل فضای برداری استفاده می شود که در آن هر مدل مجموعه داده را در یک طرح واره وزن منحصربه فرد نشان می‌دهد. یادگیرنده های استفاده شده در مدل طبقه بندی گروهی پیشنهادی عبارت‌اند از روش های طبقه بندی رگرسیون لاجستیک، درخت تصمیم و بیز ساده. در این روش با استفاده از مدل فضای برداری که با

ماتریس حضـور کار می کند که توسـط روش TF-IDF یا BagofWord تولید می شوند، کار طبقه بندی و استخراج ویژگی را نسبت به برچسب های تولید شده در مرحله قبل انجام دهد. در تهایت در این مرحله طبقه بندی توسط رای گیری انجام می شود و از رأی اکثریت (یا میانگین احتمالات پیش‌بینی‌شده (در رأی گیری نرم) برای طبقه بندی استفاده میشود.

۳-۵- خوشه بندی برای تحلیل احساسات زمینه‌ای

در نهایت در این مرحله، خوشه بندی داده ها برای تحلیل زمینه ای احساسات انجام می‌شـود. باتوجه به اینکه اگر از خروجی مدل طبقه بندی گروهی به عنوان ویژگی در الگوریتم خوشـه بندی اسـتفاده شـود، خوشـه بندی می تواند برای شناسایی الگوها در داده ها و گروه بندی موارد مشابه با هم استفاده شود. در مـدل پیشـــنهادی این تحقیق از خروجی مـدل طبقـه بنـدی برای ایجـاد خوشـه‌هایی از ویژگی های جملات برای تحلیل زمینه ای احسـاسـات اسـتفاده می‌کند. برای این منظور اســتفاده از روشـــهای K-MEANS الگوریتم‌های GMM و Birch و Spectral و DBSCAN پیشنهاد شده است.

۴- شبیه سازی و ارزیابی

۴-۱- مجموعه داده‌ها

مجموعه داده‌های مورد اسـتفاده برای شـبیه سـازی در این مقاله به شـرح زیر است

۴-۱-۱- بررسی‌های آمازون برای تحلیل احساسات

این مجموعه داده [13] شـــامل چند میلیون بررســـی مشـــتری آمازون (متن ورودی) و رتبه‌بندی ستاره (برچسب‌های خروجی) برای یادگیری نحوه آموزش سریع متن برای تحلیل احساسات است. ایده اصلی شکل‌گیری این دیتاست این است که مجموعه‌داده‌ها را به چیزی بیش از یک اسباب‌بازی مبدل کند؛ با فراهم‌کردن داده های تجاری واقعی در مقیاس معقول، اما همچنان بتوان از آن در عرض چند دقیقه با استفاده از یک لپ تاپ ساده برای آموزش‌دادن مدل ها استفاده کرد. این دیتاست تاکنون بیش از ۵۵۸۸۳ بار برای مصارف مختلف استفاده شده است.

۴-۱-۲- مجموعه‌داده تحلیل احساسات توییتر و ردیت

این مجموعه داده [14] به‌عنوان بخشـی از پروژه دانشـگاه اسـتنفورد در مورد تحلیل احسـاسـاتی در بسـترهای شـبکه های اجتماعی چند منبعی با اسـتفاده از PySpark ایجاد شـده اسـت. در این دیتاسـت دو مجموعه‌داده وجود دارد که یکی شـامل توییت‌هایی از توییتر با برچسـب احسـاسـاتی و دیگری از ردیت است که شامل نظرات با برچسب احساسی آن است.

۴-۲- معیارهای ارزیابی

در این پژوهش سـعی شـده اسـت تا با اسـتفاده معیارهای ارزیابی از معیارهای شناخته شده به بررسی دقت الگوریتم پرداخته شود که این معیارها شامل دقت، صـــحت، فراخوانی و هارمونی می شـــوند. در واقع، معیارهای نام‌برده‌شـــده بر حسـب مثبت واقعی (TP) محاسـبه می‌شـوند که تعداد نمونه‌های مثبت واقعی است که به‌درستی به‌عنوان مثبت طبقه‌بندی شده‌اند. منفی‌های واقعی (TN)، یعنی تعـداد نمونـه هـای منفی واقعی کـه بـه‌درســـتی بـه‌عنوان منفی طبقه‌بندی‌شـــده‌اند. موارد مثبت کاذب (FP)، به‌عنوان‌مثال، تعداد نمونه‌های منفی واقعی که به‌اشـتباه به‌عنوان مثبت طبقه‌بندی شـده‌اند. منفی‌های کاذب (FN)، که تعداد نمونه‌های مثبت واقعی اســـت که به‌اشـــتباه به‌عنوان منفی طبقه‌بندی شده‌اند

۴-۲-۱- معیار دقت (Pr)

معیار صحت یکی از معیارهای رایج برای بررسی عملکرد الگوریتم یادگیری ماشین استفاده می شود که در آن بیان می شود چه تعدادی از نمونه های مثبت پیش بینی شده واقعاً مثبت هستند. فرمول ریاضی معیار صحت در فرمول (۱) بیان شده است.

$$Pr = \frac{TP}{TP + FP} \quad (۱)$$

۴-۲-۲- معیار فراخوانی (Re)

معیار فراخوانی که یکی دیگر از معیارهای رایج برای بررسی یادگیری ماشین بوده که مکمل معیار صحت است و در آن بیان می شود که چه تعداد از نمونه های مثبت درست در کل نمونه ها موجود است. فرمول ریاضی معیار فراخوانی در فرمول (۲) بیان شده است.

$$Rec = \frac{TP}{TP + FN} \quad (۲)$$

۴-۲-۳- معیار هارمونیک (F-Score)

معیار هارمونیک میانگین معیارهای فراخوانی و صـحت اسـت. معیار ریاضـی هارمونیک در فرمول (۳) بیان شده است

$$F - \text{Score} = 2 * \frac{Pr * Re}{Pr + Re} \quad (۳)$$

۴-۲-۴- معیار صحت (Acc)

معیار دقت مهم‌ترین معیار بررسی عملکرد الگوریتم یادگیری ماشین است که در آن بررسی می شود چه میزانی از داده ها به‌درستی دسته بندی شده اند. فرمول ریاضی معیار دقت در فرمول (۴) بیان شده است

$$Acc = \frac{TP + TN}{TP + FP + FN + TN} \quad (۴)$$

۳-۴- نتایج آزمایشها

در ادامه نتایج آزمایشها در قالب ۳ آزمایش مختلف گزارش شده است.

۱-۳-۴- آزمایش ۱: بررسی زمان اجرای مراحل

هدف از اجرای این آزمایش بررسی زمان اجرای الگوریتم پیشنهادی نسبت به زمان اجرای الگوریتم معیار است. در این آزمایش مدتزمان اجرای مراحل دو الگوریتم به‌صورت محاسبه شده است. در این تحقیق میانگین ده بار اجرا به‌عنوان زمان اجرا معیار در نظر گرفته می شود. در این تحقیق در روش معیار الگوریتم K-Means با مقادیر n_clusters = 2, max_iter = 500, n_init = 10 در نظر گرفته شده است. همچنین در روش پیشنهادی مقادیر الگوریتم Birch مقادیر threshold=0.01, n_clusters=2 در نظر گرفته شده است. همچنین در روش پیشنهادی برای طبقه بندی گروهی از الگوریتم های LogisticRegression و RandomForestClassifier و GaussianNB استفاده شده است که در آن از رأی‌گیری سخت استفاده می شود. همچنین در این روش مقادیر الگوریتم LogisticRegression multi_class='multinomial', random_state=1 در نظر گرفته شده است و مقادیر الگوریتم RandomForestClassifier n_estimators=50, random_state=1 در نظر گرفته شده است. در جدول (۱) زمان اجرای مراحل روش ACAEC و روش پیشنهادی را می توانید مشاهده کنید. عددها بر حسب ثانیه هستند. همان‌طور که مشاهده می‌شود، زمان اجرای الگوریتم پیشنهادی نسبت به الکوریتم ACAEC بهبود پیدا کرده است.

۲-۳-۴- آزمایش ۲: بررسی زمان اجرای خوشه‌بندی

در این آزمایش به بررسی زمان اجرای الگوریتم خوشه بندی کاندید در مدل پیشنهادی پرداخته شده است. در این آزمایش مدتزمان اجرای مراحل برای هر روش به‌صورت جداگانه محاسبه شده است. در این تحقیق میانگین ده بار اجرا به‌عنوان زمان اجرا معیار در نظر گرفته می شود. در این آزمایش در الگوریتم GaussianMixture میزان n_components=2 در نظر گرفته شده است. همچنین در الگوریتم SpectralClustering میزان n_clusters=2 در نظر گرفته شده است. همچنین در الگوریتم Birch مقادیر threshold=0.01, n_clusters=2 در نظر گرفته شده است. همچنین در الگوریتم DBSCAN مقادیر eps=10, min_samples=2 در نظر گرفته شده است. همچنین برای طبقه بندی گروهی در الگوریتم پیشنهادی از الگوریتم های LogisticRegression و RandomForestClassifier و GaussianNB استفاده شده است که در آن از رأی‌گیری سخت استفاده می شود. همچنین در للگوریتم LogisticRegression multi_class='multinomial', random_state=1 در نظر گرفته شده است. همچنین در الگوریتم RandomForestClassifier n_estimators=50, random_state=1 در نظر گرفته شده است. در جدول (۲) زمان اجرای پروسه های الگوریتم پیشنهادی بعد از تغییرات در الگوریتم خوشه بندی را می توانید مشاهده کنید.

۳-۳-۴- آزمایش ۳: بررسی معیارهای دقت

هدف از اجرای این آزمایش بررسی دقت روش پیشنهادی نسبت به‌دقت روش‌های بوستینگ AdaBoostClassifier و GradientBoostingClassifier در مجموعه داده های میان دامنه ای است. در این بخش برای الگوریتم Birch مقادیر threshold=0. 1, n_clusters=2 در نظر گرفته شده است. همچنین برای طبقه بندی گروهی در للگوریتم پیشنهادی از الگوریتم های LogisticRegression و RandomForestClassifier و GaussianNB استفاده شده است که در آن از رأی‌گیری سخت استفاده می شود. همچنین در للگوریتم LogisticRegression مقادیر multi_class='multinomial', random_state=1 در نظر گرفته شده است. همچنین در الگوریتم AdaBoostClassifier مقادیر n_estimators=100, random_state=0 در نظر گرفته شده است. همچنین برای الگوریتم GradientBoostingClassifier مقادیر n_estimators=100, learning_rate=1.0, max_depth=1, random_state=0 در نظر گرفته شده است.

در جدول (۳) مقایسه‌ای از دقت الگوریتم ها بیان شده است که در آن EC نماینده طبقه بندی گروهی و V نماینده فرهنگ لغت vader و BOW نماینده روش کیسه کلمات و BI نماینده الگوریتم خوشه بندی Birch و Ada نماینده طبقه بندی AdaBoostClassifier و GBCنماینده طبقه بندی GradientBoostingClassifier است.

جدول (۱): مقایسه زمان اجرای اجزای مدل پیشنهادی و معیار

	Steps	Twitter&Reddit	Amazon
روش پیشنهادی	1:Preprocessing	1.06±5.10	0.37±10.76
	2:SentiWordNet	1.93±15.22	2.25±70.66
	3:TF-IDF	2.20±2.64	2.95±4.36
	4:K-Means	6.23±260.54	6.63±1236.49
	5:Interpretation	0.18±3.23	0.36±6.21
ACAEC	1:Preprocessing	1.06±5.10	0.37±10.76
	2:Vader	0.53±1.11	0.25±2.12
	3:TF-IDF	2.72±2.64	2.95±4.36
	4:Ensemble Classification	3.23±60.69	7.03±136.37
	5:Birch	2.27±60.42	5.12±103.81
	6:Interpretation	0.42±2.09	0.29±6.99

جدول (۲): مقایسه زمان اجرای اجزای روش های خوشه بندی

	Methods	Twitter&Reddit	Amazon
Clustering	K-Means	6.23±260.54	6.63±1236.49
	DBSCAN	4.76±71.32	7.25±146.32
	Birch	2.27±60.42	5.12±103.81
	GMM	1.41±76.36	4.32±150.61
	Spectral	7.52±162.71	7.54±123.76

جدول (۳): مقایسه معیارهای دقت الگوریتم پیشنهادی در برابر روش‌های بوستینگ

	ACACE + EC + V + BOW + BI	ACACE + V + BOW + GBC	ACACE + V + BOW + Ada
F-score	92.79	71.30	64.85
Rec	92.12	70.17	64.22
Pr	93.49	72.47	65.50
Ac	94.12	73.49	60.74

۵- نتیجه

در این مقاله، با توجه به اهمیت تحلیل احساسات کاربران در شبکه‌های اجتماعی، یک روش ترکیبی در ۵ مرحله پیشنهاد شد. به طوریکه در مرحله اول؛ پیش پردازش متن انجام شد. در مرحله دوم، از فرهنگ لغت VADER برای تعیین قطبیت استفاده شد. در مرحله سوم، از BoW و TF-IDF برای استخراج ویژگی استفاده شد. در مرحله چهارم برای طبقه بندی آنسامبلی از رگرسیون لاجستیک، درخت تصمیم و بیز ساده طراحی شد. و در نهایت در مرحله پنجم؛ برای تحلیل زمینه‌ای احساسات از چندین روش خوشه بندی استفاده شد.. نتایج ارزیابی روش پیشنهادی براساس داده‌های آمازون، توییتر و ردیت نشان دهنده آن بود که از نظر زمان اجرای الگوریتم و دقت کارایی، روش پیشنهادی مورد قبول است.

مراجع

[1] Y. K. Dwivedi, E. Ismagilova, N. P. Rana, and R. Raman, "Social media adoption, usage and impact in business-to-business (B2B) context: A state-of-the-art literature review," *Inf. Syst. Front.*, pp. 1–23, 2021.

[2] W. Medhat, A. Hassan, and H. Korashy, "Sentiment analysis algorithms and applications: A survey," *Ain Shams Eng. J.*, vol. 5, no. 4, pp. 1093–1113, 2014.

[3] Y. Tian, L. Yang, Y. Sun, and D. Liu, "Cross-domain end-to-end aspect-based sentiment analysis with domain-dependent embeddings," *Complexity*, vol. 2021, pp. 1–11, 2021.

[4] R. K. Behera, M. Jena, S. K. Rath, and S. Misra, "Co-LSTM: Convolutional LSTM model for sentiment analysis in social big data," *Inf. Process. Manag.*, vol. 58, no. 1, p. 102435, 2021.

[5] M. T. AL-Sharuee, F. Liu, and M. Pratama, "Sentiment analysis: an automatic contextual analysis and ensemble clustering approach and comparison," *Data Knowl. Eng.*, vol. 115, pp. 194–213, 2018.

[6] M. T. AL-Sharuee, F. Liu, and M. Pratama, "Sentiment analysis: dynamic and temporal clustering of product reviews," *Appl. Intell.*, vol. 51, pp. 51–70, 2021.

[7] M. D. Rocklage, D. D. Rucker, and L. F. Nordgren, "Emotionally numb: Expertise dulls consumer experience," *J. Consum. Res.*, vol. 48, no. 3, pp. 355–373, 2021.

[8] M. D. Rocklage and R. H. Fazio, "The enhancing versus backfiring effects of positive emotion in consumer reviews," *J. Mark. Res.*, vol. 57, no. 2, pp. 332–352, 2020.

[9] R. Xiang, E. Chersoni, Q. Lu, C.-R. Huang, W. Li, and Y. Long, "Lexical data augmentation for sentiment analysis," *J. Assoc. Inf. Sci. Technol.*, vol. 72, no. 11, pp. 1432–1447, 2021.

[10] B. Liang, H. Su, L. Gui, E. Cambria, and R. Xu, "Aspect-based sentiment analysis via affective knowledge enhanced graph convolutional networks," *Knowl.-Based Syst.*, vol. 235, p. 107643, 2022.

[11] R. Bensoltane and T. Zaki, "Aspect-based sentiment analysis: an overview in the use of Arabic language," *Artif. Intell. Rev.*, vol. 56, no. 3, pp. 2325–2363, 2023.

[12] C. Hutto and E. Gilbert, "Vader: A parsimonious rule-based model for sentiment analysis of social media text," in *Proceedings of the international AAAI conference on web and social media*, 2014, pp. 216–225.

[13] "A. Bittlingmayer. 'amazon review.'" https://www.kaggle.com/datasets/bittlingmayer/amazonreviews

[14] "p. kumar. 'Twitter and Reddit Sentimental analysis Dataset.'" https://www.kaggle.com/ds/429085