Scientific Journal of Electronic & Cyber Defense
Vol. 11, No. 4, Winter 2024, Serial No. 44
ISSN: 2322-4347 , E-ISSN: 2980-8979

Presenting a New Sentiment Analysis Method Based on Multi-objective Archimedes Optimization Algorithm and Machine Learning

Elham Kazemi, Mehdi, Sadeghzadeh*, Ali Taei zadeh
*Assistant Professor, Imam Hossein Comprehensive University (AS), Tehran, Iran

(Received: 2023/07/25, Revised: 2023/11/26, Accepted: 2023/12/16, Published: 2024/01/18)

DOR: https://dorl.net/dor/20.1001.1.23224347.1402.11.4.12.9

ABSTRACT

With the expansion and popularity of social networks among people, special attention has been focused on the activities, reactions, and feelings of people in these networks compared to other circles. Analyzing this volume of unstructured textual information of users requires new and optimal methods of text mining and natural language processing. The purpose of sentiment analysis is to examine a large volume of opinions about an entity by the machine and provide a summarized report of the sentiment expressed in it to the user. To achieve this goal, statistical techniques, data mining, and natural language processing are used. In this article, a new method based on Archimedes' optimization algorithm is presented so that users' opinions can be obtained in less time and with higher accuracy. Also, in order to eliminate one of the main challenges in sentiment analysis, in the feature extraction phase, all ironic sentences are collected and entered into the collection as sentences with a specific class, in order to eliminate this big challenge. This method can be applied to different languages and tries to increase the accuracy and speed of previous algorithms. The evaluation results of the data set show that the proposed method has an accuracy of 0.967.

Keywords: Text Mining, Sentiment Analysis, Archimedes Optimization Algorithm

.

This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

Publisher: Imam Hussein University © Authors

*Corresponding Author Email: mehdi.sadeghzadeh@iau.ac.ir

نشریه علمی "پدافند الکترونیکی و سایبری"
سال یازدهم، شماره ۴، زمستان ۱۴۰۲، ص۱۴۵-۳۴
شاپا الکترونیکی: ۸۹۷۹-۲۹۸۰ شاپا چاپی: ۴۳۴۷-۲۳۲۲

علمی – پژوهشی

ارائه یک روش جدید تحلیل احساسات مبتنی بر الگوریتم بهینه‌سازی ارشمیدس چندهدفه و یادگیری ماشین

الهام کاظمی [۱]، مهدی صادق زاده [۲*]، علی طائی زاده [۳]

۱.پژوهشگر،گروه کامپیوتر، دانشگاه آزاد اسلامی، واحد قم، قم، ایران. ۲. استادیار، گروه کامپیوتر، دانشگاه آزاد اسلامی، واحد علوم و تحقیقات، تهران، ایران. ۳. استادیار، گروه کامپیوتر، دانشگاه آزاد اسلامی، واحد قم، قم، ایران.

(دریافت: ۱۴۰۲/۰۵/۰۳، بازنگری: ۱۴۰۲/۰۹/۰۵، پذیرش: ۱۴۰۲/۰۹/۲۵، انتشار: ۱۴۰۲/۱۰/۲۸)

DOR: https://dorl.net/dor/20.1001.1.23224347.1402.11.4.12.9

* این مقاله یک مقاله با دسترسی آزاد است که تحت شرایط و ضوابط مجوز Creative Commons Attribution (CC BY) توزیع شده است.

ناشر: دانشگاه جامع امام حسین (ع) نویسندگان

چکیده

گسترش و محبوبیت شبکه‌های اجتماعی در بین افراد، توجه متخصصان بیش‌ازپیش به فعالیت‌ها، واکنش‌ها و احساسات افراد در این شبکه‌ها نسبت به سایر محافل معطوف گشته است. تحلیل این حجم از اطلاعات متنی غیرساخت‌یافته کاربران نیازمند روش‌های نوین و بهینه متن‌کاوی و پردازش زبان طبیعی است. هدف تحلیل احساسات این است که حجم انبوهی از نظرات پیرامون یک موجودیت توسط ماشین مورد بررسی قرار گیرد و گزارش خلاصه شده‌ای از احساس بیان شده در آن به کاربر ارائه گردد. برای دستیابی به این هدف تکنیک‌های آماری، داده‌کاوی و پردازش زبان طبیعی مورداستفاده قرار می‌گیرند. در این مقاله روش جدیدی مبتنی بر الگوریتم بهینه‌سازی ارشمیدس ارائه شده است تا بتوان در زمان کمتر و دقتی بالاتر نظرات کاربران را دررابطه‌با موضوع موردنظر به دست آورد. همچنین جهت ازبین‌بردن یکی از چالش‌های اصلی در تحلیل احساسات در مرحله استخراج ویژگی تمام جملات کنایه‌آمیز جمع شده و به‌عنوان جملات باکلاس مشخص وارد مجموعه‌داده می‌شود تا این چالش بزرگ را از بین ببرند. این روش قابلیت اعمال بر روی زبان‌های گوناگون را داشته و سعی بر بالابردن دقت و سرعت الگوریتم‌های پیشین دارد. نتایج ارزیابی از مجموعهٔ داده‌ها نشان می‌دهد که روش پیشنهادی دارای دقت ۰,۹۶۷ می‌باشد.

کلیدواژه‌ها :متن‌کاوی، تحلیل احساسات، الگوریتم بهینه‌سازی ارشمیدس.

۱- مقدمه

تحلیل احساسات[1] که به دنبال پیدا کردن نگرش، نظر یا عقیده نویسنده متن درباره‌ی یک موضوع، موجودیت، سازمان و نهاد می‌باشد[۱]. تجزیه و تحلیل احساسات به طور گسترده‌ای در موارد صدای مشتری مانند بررسی و پاسخ نظر سنجی، آنلاین و رسانه‌های اجتماعی و در موارد مراقبت بهداشت برای کاربردهایی از بازاریابی تا خدمات به مشتریان برای پزشکی اعمال می‌شود. تحلیل احساسات فرآیند خودکاری است که قطبیت متن را بر اساس مثبت، منفی و یا خنثی بودن پیش‌بینی می‌کند. مساله تحلیل احساسات را بر حسب میزان جزئیات مورد بررسی مساله می‌توان در سه سطح؛ سند، جمله و جنبه تقسیم‌بندی کرد. در این تقسیم‌بندی هر چه به سمت پایین‌تر می‌رویم میزان بررسی جزئیات تحلیل احساسات بیشتر و با مساله سخت‌تری سروکار داریم[۲, ۳].

پردازش دستی برای تجزیه‌وتحلیل احساسات مصرف‌کنندگان نیاز به یک عنصر انسانی دارد که باتوجه‌به تفسیر پیچیدگی‌های زبان متن، ابهامات، طنز و مانند آن استفاده از این روش از دقت بیشتری برخوردار است؛ اما به دلیل هدررفتن زمان و هزینه مقرون‌به‌صرفه نمی‌باشد. روش دیگر تحلیل کلمات کلیدی با استفاده از یک الگوریتم است که احساسات را به کلمات خاص اختصاص می‌دهد. به‌عنوان‌مثال عالی مثبت است، و افتضاح منفی است. درحالی‌که این روش نتایج به‌سرعت و ارزان عرضه می‌شود، اما برای پیچیدگی‌هایی مانند منفی دوگانه، کلمات با معنای

[1] Sentiment analysis

*Corresponding Author E-mail:mehdi.sadeghzadeh@iau.ac.ir

چندگانه مناسب نمی‌باشد و دارای دقت کمی می‌باشد. تحلیل احساسات در حوزه‌های مختلف می‌تواند استفاده شود در واقع از تحلیل احساسات مصرف‌کننده دیدگاهی به دست می‌آید که می‌تواند جهت بهبود کسب‌وکار به روش‌های مختلفی استفاده شود: ۱- تصمیم‌گیری ۲ -برجسته‌کردن مزیت رقابتی ۳- پیش‌بینی عمر چرخه محصول ۴- بهبود تجربه مشتری.

استفاده از تحلیل احساسات در حوزه‌های جدید، کماکان پتانسیل تحقیق و توسعه بسیاری دارد. پژوهش‌های زیادی دررابطه‌با تحلیل احساسات بر روی نظرات کاربران، مستندات و مقالات انجام شده است. بهترین نتایج به‌دست‌آمده در طبقه‌بندی احساسات از تکنیک‌های یادگیری ماشین مثل بیز ساده و ماشین بردار پشتیبان حاصل شده است. در این پژوهش به ارائه روشی برای تجزیه‌وتحلیل احساسات با الگوریتم بهینه‌سازی ارشمیدس پرداخته می‌شود. در این راستا سعی شده با تمرکز بر مرحله انتخاب ویژگی با الگوریتم ارشمیدس و طبقه‌بندی متن به‌وسیله الگوریتم بیز ساده[1] تا حدودی تحلیل احساسات را بهبود داد. همچنین جهت ازبین‌بردن یکی از چالش‌های اصلی در تحلیل احساسات در مرحله استخراج ویژگی تمام جملات کنایه‌آمیز جمع شده و به‌عنوان جملات باکلاس مشخص وارد مجموعه‌داده می‌شود تا چالش بزرگ استفاده کاربران از جملات کنایه‌آمیز را از بین ببرد و این جملات بادقت بالا تشخیص داده شوند.

نتایج نشان می‌دهد که استفاده از الگوریتم فراابتکاری ارشمیدس و دسته‌بند بیز ساده در مرحله طبقه‌بندی متن بهبود در نتایج به‌دست‌آمده را به دنبال دارد. در ادامه این مقاله به بررسی ادبیات پژوهش و کارهای مرتبط پرداخته می‌شود

۲- روش تحقیق

۲-۱ مروری بر ادبیات پژوهش و کارهای مرتبط

با هوشمند شدن کامپیوترها، مطمئناً آنها می‌توانند طیف گسترده‌ای از احساسات مختلف انسانی را رمزگشایی کرده و تشخیص دهند[۱, ۳]. از یک سو رایانه‌ها ماشین‌های هوشمندی هستند، اما نه می‌توانند هیچ احساسی را ببینند و نه احساس کنند، و تنها ورودی آنها به شکل صفر و یک است. از سوی دیگر، رایانه‌ها در یک چیزی که انسان‌ها با آن دست و پنجه نرم می‌کنند، برتری دارند: پردازش سریع و مؤثر در حجم زیادی از داده‌ها. بنابراین، از نظر تئوری، اگر بتوانیم به ماشین‌ها آموزش دهیم که چگونه احساسات پشت متن ساده را شناسایی کنند، می‌توانیم واکنش احساسی به یک محصول خاص را با تجزیه و تحلیل صدها هزار بررسی یا توییت تجزیه و تحلیل و ارزیابی کنیم. این به نوبه خود، بازخورد ارزشمندی را به شرکت ها ارائه می دهد و به آنها کمک می کند تا محصول بعدی خود را متناسب با نیازهای بازار بهتر طراحی کنند. هدف تجزیه و تحلیل احساسات این است که متن داده شده را به عنوان مثبت، منفی یا خنثی طبقه بندی کند. علاوه بر این، اطلاعات ذهنی مربوط به آن متون را با کمک پردازش زبان طبیعی، تجزیه و تحلیل متن، زبان‌شناسی محاسباتی و یادگیری ماشینی شناسایی می‌کند.

رویکردهای موجود برای تحلیل احساسات را می‌توان به تکنیک‌های مبتنی بر دانش، روش‌های آماری و رویکردهای ترکیبی طبقه‌بندی کرد. هدف تکنیک‌های دانش‌محور شناسایی کلمات یا عبارات نشان‌دهنده احساسات در متن ورودی توسط پایگاه‌های دانش است. این پایگاه‌های دانش از مجموعه‌های زبانی، مانند واژگان و WordNet تشکیل شده‌اند. این تکنیک‌ها علاوه بر پایگاه‌های دانش، قواعد زبان را برای شناسایی بهتر بافت کلمات نیز بررسی می‌کنند؛ بنابراین، این تکنیک‌ها به‌شدت بر منابع زبانی باکیفیت بالا برای دستیابی به نتایج خوب متکی هستند.

روش های آماری از الگوریتم های یادگیری ماشینی استفاده می کنند که از ویژگی های زبانی استخراج شده از متون آموزش می بینند[۴]. آنها قادر به شناسایی کلمات کلیدی مرتبط با احساسات و همچنین فراوانی همزمانی آنها با سایر کلمات کلیدی هستند. با این حال، به طور کلی، آنها از نظر معنایی ضعیف هستند و به داده های زیادی برای آموزش نیاز دارند. رویکردهای ترکیبی هم تکنیک های مبتنی بر دانش و هم روش های آماری را به عنوان راهی برای بهره برداری از بهترین هر رویکرد ترکیب می کنند. یادگیری عمیق یک روش آماری قدرتمند است که از شبکه‌های عصبی مصنوعی چند لایه استفاده می‌کند که برای وظایف یادگیری اعمال می‌شود و برای تحلیل احساسات توسط ژانگ [۵] و وظایف پردازش زبان طبیعی[2] توسط یانگ [۶] محبوب شده است. یادگیری عمیق می تواند به طور خودکار ویژگی های مفید را در وظایف پردازش زبان طبیعی استخراج کند، برخلاف مهندسی ویژگی های دستی که در یادگیری ماشین سنتی استفاده می شود. بسیاری از مدل های یادگیری عمیق در پردازش زبان طبیعی از اسنادی استفاده می کنند که به عنوان جاسازی کلمه به عنوان ویژگی های ورودی نمایش داده می شوند. ایده اصلی این است که کلمات را به‌عنوان بردارهای ویژگی نشان دهیم، که می‌توانند وابستگی‌های معنایی یا نحوی را با استفاده از یک فضای برداری متراکم با ابعاد پایین‌تر ثبت کنند. مدل های اصلی نمایش کلمه توزیع شده شامل Word2Vec ، GloVe، ELMo، FastTextو BERT. هستند.

تانگ و همکاران [۷] یک مدل موضوع احساسات را بر اساس جنبه مشترک پیشنهاد می کند و از حداکثر آنتروپی برای بهبود دقت و عملکرد تجزیه و تحلیل احساسات دسته بندی جنبه

[1] Naïve Bayes

[2] Natural Language Processing

استفاده می کند. با این حال، روش‌های یادگیری غیرعمیق در به تصویر کشیدن همبستگی‌های غیرخطی پیچیده بین ویژگی‌ها و قطبیت احساسات برتری ندارند و نتایج به شدت به کیفیت ویژگی‌ها بستگی دارد که به راحتی می‌تواند به گلوگاه عملکرد برسد. در ادامه برخی از مطالعات مورد بحث قرار گرفته است که از مدل های یادگیری عمیق برای تجزیه و تحلیل احساسات استفاده شده و همچنین برخی مطالعات را برجسته کردند که ثابت کنند مدل های یادگیری عمیق بهتر از مدل های یادگیری ماشین سنتی عمل می کنند.

مدل های یادگیری عمیق مرسوم مانند شبکه عصبی کانولوشن و LSTMهستند. مدل سی‌ان‌ان، یک شبکه عصبی پیش‌خور است. مدل LSTM یکی از مدل‌های شبکه عصبی بازگشتی است که می‌تواند داده‌های متوالی، توالی کلمات را در یک کار پردازش زبان طبیعی یاد بگیرد. به طور معمول، یک مدل LSTM داده های متوالی را در جهت رو به جلو می آموزد، اما در برخی موارد، یک الگو باید در جهت عقب آموخته شود، بنابراین BLSTM برای مدیریت آن توسعه داده شد. اویانگ و همکاران (۲۰۱۵) از یک مدل شبکه عصبی کانولوشن برای انجام طبقه‌بندی احساسات بر روی مجموعه داده‌های مرور فیلم [۸] استفاده کرد و نشان داد که مدل شبکه عصبی کانولوشن دقیق‌تر از الگوریتم‌های طبقه‌بندی سطحی مانند بیز ساده یا ماشین بردار پشتیبان [۹] است. نواک و همکاران LSTMرا با مدل های BLSTMدر طبقه بندی احساسات مجموعه داده های مرور کتاب آمازون مقایسه کرد و دریافت که مدل BLSTM در این کار دقیق تر از مدل LSTM بود[۱۰]. علاوه بر مطالعات تحقیقاتی با استفاده از مدل‌های فردی، مطالعاتی نیز وجود دارد که از ترکیبی از مدل‌ها برای بهبود عملکرد استفاده می‌کنند.

لایما و همکاران یک شبکه تعاملی برای تحلیل متون را پیشنهاد کردند که از دو مدل LSTM برای یادگیری جداگانه کلمات متنی و هدف استفاده می‌کند نتایج روش پیشنهادی نشان از بهبود معیارها نسبت به روش‌های پیشین را دارد[۱۱].

حسن و همکاران [۱۲] یک معماری مبتنی بر شبکه عصبی ConvLstmکه از شبکه عصبی کانولوشن و LSTM برای تحلیل احساسات استفاده میکند را پیشنهاد کردند. سپس معماری پیشنهادی را بر روی مجموعه داده بررسی فیلم IMBD و مجموعه داده(SSTb) Stanford Sentiment Treebank پیاده سازی کرده که نتیجه این مقاله عملکرد بهتر مدل آنها را در مقایسه با سایر مدل‌های یادگیری عمیق و یادگیری ماشین سنتی نشان داد. Çano و Morisio یک معماری یادگیری عمیق جدید به نام ngramCNN برای تجزیه و تحلیل احساسات اسناد متنی طولانی معرفی کردند[۱۳]. این مقاله به ترتیب۹۱/۲، ۷۵/۶ و ۹۵/۹ درصد دقت را در مجموعه داده‌های مرورهایIMDB ، مجموعه داده‌های اشعار آهنگ، و مجموعه داده‌های بررسی تلفنی گزارش کرده است. عملکرد مدل جدید بهتر از سایر مدل های پایه یادگیری ماشینی بود که آنها در نظر گرفتند. همچنین، گولاراس و کامیس گروه‌ها و ترکیب‌های شبکه عصبی کانولوشن و LSTM را مقایسه کردند و به‌طور تجربی نشان دادند که شبکه عصبی کانولوشن و LSTM با هم بهتر از زمانی که به تنهایی استفاده می‌شوند، عمل می‌کنند[۱۴]. علاوه بر این ژانگ و همکاران[۵] به بررسی انواع مختلف مدل‌های شبکه عصبی[1] که برای تحلیل احساسات متن و همچنین تصاویر استفاده می‌شوند، پرداخت. لی و لی برخی از معماری های پیشرفته یادگیری عمیق را برای تجزیه و تحلیل احساسات توضیح دادند. این مقاله همچنین نشان داد که مرورهای کوتاه با خوانایی بالا می‌توانند عملکرد بهتری نسبت به ترکیبی از سطوح تعداد کلمات و خوانایی داشته باشند[۱۵].

در کارهای پیشین که موردمطالعه قرار گرفته اکثر روش‌های پیشنهادی مبتنی بر الگوریتم‌های یادگیر بوده و در برخی کاهش ویژگی‌ها به‌منظور بهبود نتایج صورت‌گرفته است. از تفاوت‌های روش پیشنهادی در این مقاله چندهدفه بودن و همچنین پرداختن به جملات کنایه‌آمیز در مجموعه‌داده‌ها به‌منظور بهبود تحلیل احساسات کاربران است.

۲-۲ -الگوریتم بهینه سازی ارشمیدس

امروزه طراحی الگوریتم‌های بهینه‌سازی برای حل مشکلات در زمینه‌های مختلف علمی بسیار محبوب است. الگوریتم‌های بهینه‌سازی معمولاً با الهام از رفتار طبیعی عامل که می‌تواند انسان، حیوانات، گیاهان یا عوامل فیزیکی یا شیمیایی باشد، الهام گرفته شده است. بیشتر الگوریتم‌های ارائه شده در دهه گذشته از رفتار حیوانات الهام گرفته شده است. الگوریتم بهینه‌سازی ارشمیدس[2] یک الگوریتم مبتنی بر جمعیت است که از قانون طبیعی الهام گرفته شده است[۱۶]. در الگوریتم بهینه سازی ارشمیدس افراد جمعیت اجسام غوطه ور هستند. مانند سایر الگوریتمهای فراابتکاری مبتنی بر جمعیت، الگوریتم بهینه سازی ارشمیدس نیز فرآیند جستجو را با جمعیت اولیه اشیا (راه حلهای نامزد) با حجم، چگالی و شتاب تصادفی آغاز می‌کند. در این مرحله، هر جسم نیز با موقعیت تصادفی در مایع آغاز می‌شود. پس از ارزیابی تناسب جمعیت اولیه، الگوریتم بهینه سازی ارشمیدس تکرار می‌شود تا زمانی که شرط خاتمه برآورده شود. در هر تکرار، الگوریتم بهینه سازی ارشمیدس چگالی و حجم هر

[1] Neural Network
[2] Archimedes optimization algorithm

شی را به روز می‌کند. شتاب جسم براساس شرایط برخورد آن با هر جسم همسایه به روز می‌شود. چگالی، حجم، شتاب به روز شده موقعیت جدید یک جسم را تعیین می‌کند. در زیر بیان دقیق ریاضی مراحل الگوریتم بهینه سازی ارشمیدس آورده شده است.

۲،۲،۱ مراحل الگوریتمی.

در این بخش، تئوری ریاضی الگوریتم بهینه‌سازی ارشمیدس معرفی می‌شود. از نظر تئوری، الگوریتم بهینه‌سازی ارشمیدس می‌تواند به‌عنوان یک الگوریتم بهینه‌سازی جهانی در نظر گرفته شود؛ زیرا فرآیندهای اکتشاف و بهره‌برداری را در بر می‌گیرد. شکل (۱) شبه کد الگوریتم بهینه‌سازی ارشمیدس را بیان می‌کند[۱۶]. از جمله مقداردهی اولیه جمعیت، ارزیابی جمعیت و بروزرسانی پارامترها. از نظر ریاضی، مراحل الگوریتم بهینه سازی ارشمیدس پیشنهادی به شرح زیر است.

مرحله ۱ - مقداردهی اولیه: موقعیت های همه اشیا با استفاده از رابطه (۱) مقداردهی اولیه می شود:

رابطه (۱) $O_i = lb_i + rand \times (ub_i - lb_i);\ i = 1, 2, \ldots, N$

که Oi شی i ام در جمعیت n جسمی است. lb_i و ub_i به ترتیب مرزهای پایینی و بالایی فضای جستجو هستند[۱۶]. حجم (vol) و چگالی (den) را برای هر شی ith با استفاده از رابطه (۲) مقدار اولیه دهید:

رابطه (۲)
$$den_i = rand$$
$$vol_i = rand$$

که در آن rand یک بردار D بعدی است که به طور تصادفی عددی بین [۰،۱] ایجاد می‌کند. و سرانجام، شتاب (acc) شی ithرا با استفاده از رابطه (۳) شروع کنید:

رابطه (۳) $$acc_i = lb_i + rand \times (ub_i - lb_i)$$

در این مرحله، جمعیت اولیه را ارزیابی کرده و جسمی را با بهترین مقدار برازندگی انتخاب کنید. xbest، denbest، volbest و accbest را مقداردهی می‌کند.

مرحله ۲ - بروزرسانی چگالی، حجم: چگالی و حجم شیi برای تکرار $t+1$ با استفاده از رابطه (۴) و رابطه (۵) به روز می‌شود:

رابطه (۴) $$den_i^{t+1} = den_i^t + rand \times (den_{best} - den_i^t)$$

رابطه (۵) $$vol_i^{t+1} = vol_i^t + rand \times (vol_{best} - vol_i^t)$$

که در آن volbest و denbest حجم و چگالی مرتبط با بهترین جسمی است که تاکنون پیدا شده است و rand به طور یکنواخت تعداد تصادفی توزیع می‌شود.

مرحله ۳ - اپراتور انتقال و ضریب چگالی: در ابتدا، برخورد بین اجسام رخ می دهد و پس از مدتی، اجسام سعی می-کنند به حالت تعادل برسند. این در AOA با کمک اپراتور انتقال TF که جستجو را از اکتشاف به بهره برداری تبدیل می کند، با استفاده از رابطه (۶) تعریف شده است.

رابطه(۶) $$TF = exp(\frac{t-t_{max}}{t_{max}})$$

که انتقال TF باگذشت زمان و رسیدن به ۱ تدریجاً افزایش می‌یابد. در اینجا t و t_{max} به ترتیب تعداد تکرار و حداکثر تکرار هستند. به طور مشابه، عامل کاهش چگالی d همچنین به AOA در جستجوی جهانی به محلی کمک می‌کند. با استفاده از زمان رابطه (۷) این کاهش می‌یابد:

رابطه (۷) $$d^{t+1} = exp[\left(\frac{t-t_{max}}{t_{max}}\right) - \left(\frac{t}{t_{max}}\right)]$$

که d^{t+1} باگذشت زمان کاهش می‌یابد که توانایی همگرایی در منطقه امیدوارکننده را دارد. توجه داشته باشید که استفاده صحیح از این متغیر تعادل بین اکتشاف و بهره‌برداری در الگوریتم بهینه‌سازی ارشمیدس را تضمین می‌کند.

مرحله ۴،۱ - مرحله اکتشاف (برخورد بین اشیا): اگر $TF \leq 0.5$، تصادم بین اشیا رخ داد، یک ماده تصادفی (mr) را انتخاب کنید و شتاب شی را برای تکرار $t+1$ با استفاده از رابطه (۸) به روز کنید:

رابطه (۸) $$acc_i^{t+1} = \frac{den_{mr} + vol_{mr} \times acc_{mr}}{den_i^{t+1} \times vol_i^{t+1}}$$

که در آن den_i، vol_i و acc_i چگالی، حجم و شتاب جسم i هستند. درحالی‌که acc_{mr}، den_{mr} و vol_{mr} شتاب، چگالی و حجم مواد تصادفی هستند. ذکر این نکته مهم است که $TF \leq 0.5$ اکتشاف در طول یک‌سوم تکرارها را تضمین می‌کند. اعمال مقدار غیر از ۰،۵ باعث تغییر رفتار اکتشافی-بهره برداری می‌شود.

مرحله ۴،۲ - مرحله بهره برداری (عدم برخورد بین اشیا): اگر $TF > 0.5$ باشد، هیچ تصادمی بین اشیا وجود ندارد، شتاب را برای تکرار $t+1$ با استفاده از رابطه (۹) به روز کنید:

رابطه (۹) $$acc_i^{t+1} = \frac{den_{best} + vol_{best} \times acc_{best}}{den_i^{t+1} \times vol_i^{t+1}}$$

که acc_{best} شتاب بهترین جسم است.

مرحله ۴،۳ - عادی سازی شتاب: عادی سازی شتاب برای محاسبه درصد تغییر با استفاده از رابطه (۱۰) به روز می‌شود:

$$acc_{i-norm}^{t+1} = u \times \frac{acc_i^{t+1} - min(acc)}{max(acc) - min(acc)} + l \quad \text{رابطه (۱۰)}$$

که در آن u و l محدوده نرمال‌سازی هستند و به ترتیب روی ۰،۹ و ۰،۱ تنظیم می‌شوند. acc^{t+1} درصد گام تغییر هر عامل را تعیین می‌کند. اگر جسم i از بهینه جهانی فاصله داشته باشد، مقدار شتاب زیاد خواهد بود - به این معنی که جسم در مرحله اکتشاف قرار خواهد گرفت. در غیر این صورت، در مرحله بهره برداری. این نشان می‌دهد که چگونه جستجو از مرحله جستجو به مرحله بهره برداری تبدیل می‌شود. در حالت عادی، ضریب شتاب با مقدار زیادی شروع می‌شود و با گذشت زمان کاهش می‌یابد. این به نمایندگان جستجو کمک می‌کند تا به سمت بهترین راه حل جهانی حرکت کنند و در عین حال از راه حل های محلی فاصله می‌گیرند. اما قابل ذکر است که ممکن است چند عامل جستجو باقی بمانند که برای اقامت در مرحله اکتشاف به زمان بیشتری نیاز دارند تا موارد عادی از این رو الگوریتم بهینه سازی ارشمیدس به تعادل بین اکتشاف و بهره برداری می‌رسد.

مرحله ۵ - به روز رسانی موقعیت: اگر $TF \leq 0.5$ (مرحله اکتشاف)، موقعیتi_{th} برای تکرار بعدی $t+1$ با استفاده از رابطه (۱۱) به روز می‌شود:

$$x_i^{t+1} = x_i^t + C_1 \times rand \times acc_{i-norm}^{t+1} \times d \times (x_{rand} - x_i^t) \quad \text{رابطه (۱۱)}$$

جایی که $C1$ ثابت است برابر با ۲ است. در غیر این صورت، اگر $TF > 0.5$ (مرحله بهره برداری) باشد، اشیا موقعیت خود را با استفاده از رابطه (۱۱) به روز می‌کنند.

$$x_i^{t+1} = x_{best}^t + F \times C_2 \times rand \times acc_{i-norm}^{t+1} \times d \times (T \times x_{best} - x_i^t) \quad \text{رابطه (۱۲)}$$

که در آن C_2 یک ثابت برابر با ۶ است. T با زمان افزایش می‌یابد و نسبت مستقیم با عملگر انتقال دارد و با استفاده از $T = C3 \times TF$ تعریف می‌شود. T با گذشت زمان در دامنه $[C3 \times 0.3, 1]$ افزایش می‌یابد و درصد مشخصی را از بهترین موقعیت در ابتدا می‌گیرد. این کار با درصد کم شروع می‌شود زیرا این تفاوت زیادی بین بهترین موقعیت و موقعیت فعلی دارد، بنابراین اندازه گام تصادفی زیاد خواهد بود. با ادامه جستجو، این درصد به تدریج افزایش می‌یابد تا اختلاف بین بهترین موقعیت و موقعیت فعلی کاهش یابد. این امر منجر به دستیابی به تعادل مناسب بین اکتشاف و بهره برداری می‌شود. F پرچمی برای تغییر جهت حرکت با استفاده از رابطه (۱۳) است:

$$F = \begin{cases} +1 \; if \; P \leq 0.5 \\ -1 \; if \; P > 0.5 \end{cases} \quad \text{رابطه (۱۳)}$$
$$Where \; P = 2 \times rand - C_4$$

مرحله ۶ - ارزیابی هر یک از اشیا: اشیا را با استفاده از تابع هدف f ارزیابی کنید و بهترین راه حل موجود را به خاطر بسپارید. x_{best}، den_{best}، vol_{best} و acc_{best} را اختصاص دهید.

شبه کد الگوریتم بهینه‌سازی ارشمیدس

Pseudocode of Archimedes Optimization Algorithm

Procedure AOA (Population size N , tmax , Con1, Con2, Con3, Con4)
Start individual's population with random locations, volumes, and densities utilizing the equations (1), (2), and (3), respectively.
Assess the primary population and choose one individual with the best value.
Set iteration counter t = 1
While (t ≤ tmax)
For every individual i
Revise volume and density of every individual utilizing equation (4).
Revise transfer and density declining factors TF and d utilizing equations (5) and (6) respectively.
If (TF ≤ 0.5) I Exploration phase
Revise acceleration utilizing (7) and utilize equation (9) to normalize acceleration.
Revise location utilizing equation (10) else I Exploitation phase
Revise acceleration utilizing (8) and utilize equation (9) to normalize acceleration.
Revise direction flag utilizing (12)
Revise location utilizing equation xt+1
End If
End For
Evaluate every individual and select the one with the finest value
Set t = t + 1
End While
Return individual with finest fitness value.
End Procedure

۳-۲- روش پیشنهادی در تحلیل احساسات

دو روش اصلی برای تحلیل احساسات وجود دارد: یادگیری ماشینی و مبتنی بر واژگان. روش یادگیری ماشینی از داده‌های برچسب‌گذاری شده توسط انسان برای آموزش طبقه‌بندی‌کننده متن استفاده می‌کند و آن را به یک روش یادگیری تحت نظارت تبدیل می‌کند[۱]. رویکرد مبتنی بر واژگان یک جمله را به کلمات تقسیم می کند و جهت گیری معنایی هر کلمه را بر

اساس فرهنگ لغت نمره می دهد. سپس امتیازهای مختلف را جمع می کند تا به نتیجه برسد.

در این مقاله یک روش جدید تحلیل احساسات مبتنی بر الگوریتم بهینه‌سازی ارشمیدس چندوجهی و رویکرد یادگیری ماشینی ارائه می‌شود. مراحل مختلف تحلیل احساسات در این روش به‌صورت زیر خواهد بود.

فاز اول پیش‌پردازش: پیش‌پردازش در جهت بالابردن دقت تحلیل احساسات الزامی است؛ ولی باتوجه‌به اینکه از داده آماده استفاده شده است در پیاده‌سازی عملاً این فاز حذف شده است؛ ولی اگر داده‌ها آماده نباشند باید این فاز اجرا شود.

فاز استخراج ویژگی: این مرحله نیز در مجموعه‌داده‌های موردنظر اجرا شده است؛ ولی ما به جهت بهبود در این مرحله یک سری جملات را که به‌عنوان کنایه استفاده می‌شوند به‌عنوان ویژگی جدید به ویژگی‌های استخراج شده اضافه می‌کنیم.

قبل از پرداختن به توضیحات جزئی این مرحله باید چالش-های موجود در پیاده‌سازی تحلیل هیجانات نویسنده متن را یادآوری کنیم که هیچ‌کدام از این چالش‌ها تا کنون در زمینهٔ پردازش زبان طبیعی به‌کلی حل نشده‌اند و ما باید در روش پیشنهادی سعی کنیم این موارد را تحت شعاع قرار دهیم. برخی از این چالش‌ها عبارت‌اند از[۱]:

۱. کاربر به طور مستقیم احساس خود را بروز نمیدهد و ممکن است به شیوه های متعدد (افعال یا صفات منفی) نظر خود را بروز دهد. جملات پیاپی ممکن است معنی هم را تکمیل کنند. برای مثال: امروز حالم به هیچ وجه خوب نبود یا اگر فکر میکنید حالم خوبه در اشتباهید.

۲. کاربر ممکن است از کنایه-طعنه-شوخی استفاده کند. هر کدام از این موارد در ساختار جمله مفهومی کاملا متفاوت با مفهوم واقعی دارند.

برای مثال به جملات دقت کنید: امروز همان قدر خوشحال بودم که به بره تو سلاخ‌خانه خوشحاله ! این کامپیوتر کاربردش در حد یه پاره‌آجر هم نیست.

۳. کاربران معمولا از ساختار های جمله ای پیچیده ای استفاده می‌کنند که به دلیل رسمی نبودن الزاماً در چهارچوب کلی ساختار زبان قرار نمی‌گیرند.

۴. بنابه دلایل فوق بررسی تنها لغت به لغت میتواند باعث ایجاد نتایج غلط شود.

چالش دوم که استفاده از کنایه - طعنه - شوخی در متون هست. اکثر روش‌های موجود در تشخیص این متون اشتباه می‌کنند. برای حل این مشکل مرحله استخراج ویژگی را مجدد در مجموعه‌داده انجام می‌دهیم. برای انجام این کار چون مجموعه‌داده از اول مرحله استخراج انجام داده ما یک جدول شامل کلیه جملات و اصطلاحات موجود در زبان انگلیسی را جمع کردیم و در جدولی قرار دادیم و کلاس هر کدام را مشخص کردیم. برای بهتر پاسخ‌دادن الگوریتم هنگام خواندن متون اگر به هر کدام از این اصطلاحات در متن مواجه بشویم شماره آن اصطلاح را در مجموعه‌داده که در یک ستون به‌صورت موقت ایجاد شده است درج می‌کنیم. در مرحله انتخاب ویژگی از این ویژگی‌ها حتماً استفاده شود تا دقت روش بیشتر باشد. چند تا از این جملات کنایه‌آمیز در جدول شماره ۱ نشان‌داده‌شده است.

جدول (۱). جملات کنایه آمیز که تشخیص کلاس آنها مشکل است

	اصطلاح	معنی	جایگاه در جمله	کلاس
۱	Haste makes waste	در صورت عجله برای چیزی اشتباه خواهی کرد(عجله کار شیطونه)	به‌تنهایی	منفی
۲	Speak of the devil	فردی که الان در مورد حرف می‌زدیم ظاهر می‌شود(چه حلال‌زاده!)	به‌تنهایی	مثبت
۳	A snowball's chance in hell	هیچ شانسی ندارد(شانس یک گلولهٔ برفی در جهنم)	به‌عنوان بخشی از جمله	منفی
۴	A blessing in disguise	چیز خوبی که در ابتدا بد به نظر می‌رسید(نعمتی نهانی)	به‌عنوان بخشی از جمله	مثبت
۵	Add insult to injury	تشدید یک شرایط بد(عذر بدتر از گناه)	به‌عنوان بخشی از جمله	منفی
۶	On the ball	موفق‌بودن در کاری	به‌عنوان بخشی از جمله	مثبت
۷	Bite off more than you can chew	قبول پروژه‌ای که توانایی به اتمام رساندش را ندارید(لقمه بزرگ‌تر از دهن خود برداشتن)	به‌عنوان بخشی از جمله	منفی
۸	Break the ice	احساس راحتی به دیگران دادن(یخ کسی آب‌شدن)	به‌عنوان بخشی از جمله	مثبت

فاز سوم انتخاب ویژگی: پس از فاز استخراج ویژگی نوبت به فاز انتخاب ویژگی‌های مؤثر و مفید در تحلیل احساسات می‌رسد این فاز توسط الگوریتم فراابتکاری ارشمیدس انجام می‌شود. به دلیل اینکه مسئله انتخاب ویژگی یک مسئله گسسته و چندهدفه می‌باشد در نتیجه لازم است ابتدا این الگوریتم تبدیل به الگوریتم گسسته شود و سپس چندهدفه گردد. جزئیات گسسته و چندهدفه کردن به‌صورت زیر می‌باشد.

در این قسمت به معرفی یک روش جدید باینری برای الگوریتم بهینه‌سازی ارشمیدس بر پایه تابع sigmoid پرداخته می‌شود. الگوریتم بهینه‌سازی ارشمیدس در فضای پیوسته حرکت می‌کند و به همین دلیل تمام راه‌حل‌های موجود در جمعیت این

الگوریتم شامل اعداد پیوسته هستند. باید راه‌حل باینری جدید شامل اعداد ۰ و ۱ باشد که ۱ نشان دهنده انتخاب یک ویژگی جهت دیتاست جدید و صفر نشان دهنده عدم انتخاب یک ویژگی است. برای این کار از تابع sigmoid یا همان[۱۷, ۱۸] S-shaped برای حرکت دادن فرآیندهای الگوریتم بهینه سازی بیوه سیاه در فضای باینری بهره خواهیم برد. بنابراین در این مدل پیشنهادی از تابع sigmoid برای تغییر موقعیت پیوسته راه حل ها در الگوریتم بهینه سازی ارشمیدس به حالت باینری به صورت رابطه (۱۴) استفاده می شود:

$$sg\left(AOA_i^d(t)\right) = \frac{1}{1+e^{-AOA_i^d(t)}} \quad \text{رابطه (۱۴)}$$

AOA_i^d مقدار پیوسته راه حل i ام در جمعیت الگوریتم بهینه سازی ارشمیدس در بعد d ام در تکرار t است. خروجی تابع انتقال sigmoid هنوز در یک حالت پیوسته بین ۰ و ۱ است و در نتیجه برای تبدیل به یک مقدار باینری باید آستان گیری شود، که آستانه تصادفی که در رابطه (۱۴) ارائه شد جهت تبدیل به راه حل مقدار باینری برای انتخاب ویژگی در تابع sigmoid اعمال می شود:

$$AOA_i^d(t+1) = \begin{cases} 0 & if\ rand < sg\left(AOA_i^d(t)\right) \\ 1 & if\ rand \geq sg\left(AOA_i^d(t)\right) \end{cases} \quad \text{رابطه (۱۵)}$$

AOA_i^d نشان دهنده موقعیت راه حل i ام در جمعیت الگوریتم بهینه سازی بیوه سیاه در تکرار t در بعد d ام است، $AOA_i^d(t+1)$ بهترین موقعیت بعدی در بعد d را نشان می دهد. همچنین $rand$ نشان دهنده یک عدد بین صفر و یک از نوع توزیع یکنواخت است. بدین ترتیب راه حل های موجود در جمعیت الگوریتم بهینه سازی ارشمیدس با استفاده از روابط (۱۴) و (۱۵) مجبور می شوند در یک فضای جستجوی باینری حرکت کنند. در ادامه این روابط را به طور دقیق تر در الگوریتم بهینه سازی ارشمیدس جایگذاری شده است.

پس از گسسته کردن الگوریتم ارشمیدس نوبت به چندهدفه کردن الگوریتم ارشمیدس می‌رسد. برای این کار ابتدا تابع هدف انتخاب ویژگی را برای الگوریتم پیشنهادی و سایر الگوریتم‌های فراابتکاری در این مقاله تشریح می‌کنیم. چرا که انتخاب ویژگی می‌تواند به‌عنوان یک مسئله بهینه‌سازی چندهدفه در نظر گرفته شود که در آن دو هدف متضاد به دست می‌آید که شامل حداقل تعداد ویژگی‌های انتخاب شده و دقت دسته‌بندی بالاتر است؛ بنابراین برای تعریف تابع هدف مسئله انتخاب ویژگی نیازمند یک الگوریتم دسته‌بندی هستیم و از ساده‌ترین روش دسته‌بند یعنی دسته‌بندی‌کننده بیز ساده بهره می‌برند، ما نیز برای تعریف تابع هدف مسئله انتخاب ویژگی از این دسته‌بندی استفاده کردیم؛ بنابراین در روش پیشنهادی ما از دسته‌بندی‌کننده بیز ساده برای ارزیابی درست‌تر ویژگی‌های انتخاب شده توسط الگوریتم پیشنهادی و سایر الگوریتم‌ها استفاده کردیم. هر راه‌حل بر اساس تابع چندهدفه پیشنهادی ارزیابی می‌شود که به دسته‌بندی‌کننده بیز ساده بستگی دارد. در تابع چندهدفه پیشنهادی به‌منظور تعادل بین تعداد ویژگی‌های انتخاب شده در هر راه‌حل (حداقل) و دقت دسته‌بندی (حداکثر)، تابع تناسب در رابطه (۱۶) برای ارزیابی یک راه‌حل در هر الگوریتم‌های فراابتکاری استفاده می‌شود.

$$Fitness = \alpha\gamma_R(D) + \beta\frac{|R|}{|N|} \quad (۱۶)$$

که در آن $\alpha\gamma_R(D)$ نرخ خطای دسته بندی یک روش دسته بندی را نشان می دهد و همچنین $|R|$ چند خطی بودن زیر مجموعه انتخاب شده است و $|N|$ تعداد کل ویژگی های موجود در مجموعه داده است، همچنین پارامتر α اهمیت کیفیت دسته بندی و پارامتر β طول زیر مجموعه است. مقادیر این دو پارامتر نیز به صورت $\alpha \in [0, 1]$ و $\beta = (1-\alpha)$ می باشد. در این مطالعه مقدار اولیه α را ۰,۹۹ در نظر گرفته شده است و لذا مقدار B نیز ۰,۰۱ بدست خواهد آمد.

باتوجه‌به اینکه الگوریتم اصلی به نام AOA می‌باشد، گسسته آن را BAOA نام‌گذاری کرده و چندهدفه این روش را MBAOA نامیده‌ایم.

فاز چهارم دسته‌بندی مجموعه‌داده بر اساس احساسات نویسنده متن: هدف تحلیل احساسات از این مجموعه، تعیین احساسات مربوط به نویسنده متن می‌باشد. برای این کار از روش بیز ساده استفاده می‌شود. بیز ساده یک سند d را به کلاس c تخصیص می‌دهد که $p(c|d)$ را با استفاده از قانون Bayes در رابطه (۱۷) حداکثر می‌کند[۱۹].

$$p(c|d) = \frac{p(c)p(d|c)}{p(d)} \quad (۱۷)$$

۳- نتایج و بحث

روش پیشنهادی در Mat lab R2015a با کامپیوتر دارای سی‌پی‌یو $i5$ ۶۴ بیتی و حافظه $4G$ شبیه‌سازی‌شده است. برای این کار دو مجموع داده به نام‌های ISEAR و Sentiment polarity datasets $v2.0$ بکار رفته است که ۲۰٪ داده‌ها را آموزشی و ۸۰٪ بقیه را برای تست استفاده می‌کنیم. همچنین در هنگام استفاده از الگوریتم ارشمیدس در مرحله انتخاب ویژگی هر جسم راه‌حل احتمالی و تعداد اجسام جمعیت اولیه و ویژگی‌ها هم بردار مربوط به هر جسم که ویژگی‌ها تو این بردار قرار می‌گیرد، در نظر گرفته می‌شود. هر جسم یک مقدار شایستگی دارد که توسط تابع شایستگی محاسبه می‌شود. هر چه جسم در

فضای جستجو به هدف نزدیک‌تر باشد، شایستگی بیشتری دارد. در شروع به اجرای الگوریتم جمعیت اولیه به‌صورت تصادفی به وجود می‌آیند و با به‌روز کردن مقدار شایستگی بر اساس الگوریتم ارشمیدس، نسل‌های بعدی به وجود می‌آیند و سعی در یافتن راه‌حل بهینه می‌نمایند. جدول ۲ کارایی طبقه‌بندی‌کننده متن پیشنهادی را با الگوریتم‌های پایه موجود نشان می‌دهد.

جدول(۲). مقایسه کارایی الگوریتم‌های طبقه‌بندی‌کننده متن پایه با الگوریتم پیشنهادی

Dataset	Feature Selection	Accuracy	Sensitivity	Specificity	Precision	f-measure
ISEAR	MBAOA	۰.۹۹۱۲	۰.۹۶۵۹	۰.۹۶۱۳	۰.۹۶۸۵	۰.۹۶۷۹
	BGWO	۰.۹۸۹۲	۰.۹۲۸۵	۰.۹۶۱۲	۰.۹۶۸۵	۰.۹۲۵۹
	BBOA	۰.۹۴۲۱	۰.۹۰۱۴	۰.۹۲۱۷	۰.۹۳۱۲	۰.۹۱۲۸
	BWOA	۰.۹۸۱۴	۰.۹۲۱۷	۰.۹۵۰۲	۰.۹۶۸۷	۰.۹۳۰۴
	BFFA	۰.۹۳۰۱	۰.۸۹۲۵	۰.۹۲۰۱	۰.۹۲۸۷	۰.۹۲۰۹
Sentiment polarity datasets v2.0	MBAOA	۰.۹۸۸۳	۰.۹۷۰۱	۰.۹۷۰۵	۰.۹۷۰۵	۰.۹۷۲۵
	BGWO	۰.۹۸۰۱	۰.۹۶۲۱	۰.۹۶۹۸	۰.۹۶۸۷	۰.۹۶۰۴
	BBOA	۰.۹۱۱۴	۰.۸۹۱۷	۰.۸۹۰۲	۰.۸۹۳۳	۰.۸۹۹۱
	BWOA	۰.۹۸۱۴	۰.۹۶۰۱	۰.۹۶۹۸	۰.۹۶۲۱	۰.۹۶۸۷
	BFFA	۰.۹۲۱۰	۰.۸۷۰۱	۰.۸۹۰۱	۰.۸۹۴۱	۰.۹۱۰۲

باتوجه‌به جدول(۲) می‌توان فهمید که الگوریتم پیشنهادی یعنی الگوریتم AOA چندهدفه عملکردی بهتری در انتخاب ویژگی دارد. در ادامه در جدول (۳) نتایج الگوریتم پیشنهادی و سایر الگوریتم‌ها با توجه میانگین تعداد انتخاب ویژگی را نشان می‌دهد.

جدول (۳): نتایج رویکرد پیشنهادی MBAOA و سایر الگوریتم ها با توجه میانگین تعداد انتخاب ویژگی

Dataset	GA	BBOA	BGWO	BWOA	BFFA	BDA	MBAOA
ISEAR	۵	۶.۲	۵	۵	۵.۲	۷.۱	۴.۵
Sentiment polarity datasets v2.0	۱۶	۱۵.۵	۱۴.۱	۱۵.۸	۱۶.۲	۱۸.۳	۱۳.۱

جدول (۳) نتایج روش پیشنهادی MBAOA و سایر الگوریتم ها از لحاظ میانگین تعداد انتخاب ویژگی را نشان می دهد که روش MBAOA عملکرد بسیار خوبی از خود به نمایش گذاشته است. به طوری که این رویکرد در اکثر مجموعه ها داده برتری خود را ثابت کرده است. البته باید گفت که در تابع هدف هم تعداد ویژگی و هم دقت دسته بندی در نظر گرفته شده است در نتیجه نمی توان تعداد ویژگی کمتر در تمام مجموعه داده را بدست آورد. در ادامه روش MBAOA و سایر الگوریتم های مقایسه ای را از لحاظ دقت دسته بندی مقایسه و در جدول (۴) نشان داده شده است.

جدول (۴). نتایج الگوریتم پیشنهادی و سایر الگوریتم ها از لحاظ تابع هدف

Dataset	GA	BBOA	BFFA	BWOA	BGWO	BDA	MBAOA
ISEAR	۰.۸۴۷۸	۰.۸۸۴۶	۰.۹۱۸۴	۰.۹۲۲۲	۰.۹۲۷۵	۰.۸۲۷۴	۰.۹۴۰۱
Sentiment polarity datasets	۰.۸۱۰۸	۰.۸۷۴۹	۰.۹۰۱۶	۰.۹۱۷۵	۰.۹۲۱۵	۰.۸۲۰۵	۰.۹۳۹۸

جدول (۴) نتایج روش MBAOA و سایر الگوریتم ها را از لحاظ دقت دسته بندی نشان می دهد که روش MBAOA علاوه بر انتخاب ویژگی بهتر که در جدول (۲) نشان داده شد، از لحاظ دقت دسته بندی نیز کارایی بالای از خود به نمایش گذاشته است. این رویکرد در اکثر مجموعه ها داده از لحاظ دقت دسته بندی بهتر عمل کرده است. دلیل این نتایج این است که در تابع هدف هم انتخاب ویژگی و هم دقت دسته بندی در نظر گرفته شده است.

۴- نتیجه گیری

باتوجه‌به اینکه می‌توان AOA به‌عنوان یک الگوریتم مناسب برای حل مشکلات مختلف بهینه‌سازی در نظر گرفت. در این مقاله AOA تبدیل به الگوریتم گسسته شده سپس چندهدفه شده و در تجزیه‌وتحلیل احساسات متن مورداستفاده قرار گرفته است. همچنین جهت ازبین‌بردن چالش اصلی تحلیل احساسات یعنی تشخیص جملات طنز و کنایه‌آمیز در مرحله استخراج ویژگی یک ویژگی به‌صورت موقت جهت ازبین‌بردن این چالش اضافه می‌شود تا در مرحله انتخاب ویژگی و دسته‌بندی احساسات از این ویژگی جهت بالابردن دقت دسته‌بندی استفاده شود. برای ارزیابی نتایج از نظر دقت دسته‌بندی با روش‌های مرتبط مورد ارزیابی واقع شده که نتایج حاکی از بهتر بودن روش پیشنهادی می‌باشد. در

ارزیابی دیگری میانگین تعداد انتخاب ویژگی مورد بررسی قرار گرفته که نشان می‌دهد روش پیشنهادی با میانگین تعداد ویژگی کم به نتایج بهتری می‌رسد. همچنین از لحاظ تابع هدف نیز عملکرد الگوریتم پیشنهادی بهتر از روش‌های دیگر می‌باشد. به طور خلاصه برتری‌های روش جدید شامل گسسته کردن AOA ، چندهدفه کردن روش AOA ، استفاده از انتخاب ویژگی‌های روان‌شناسی تحلیل احساسات و ازبین‌بردن چالش استفاده از طنز و کنایه در متون می‌باشد.

۹. مراجع

1. Hosseinalipour, A., et al., A novel binary farmland fertility algorithm for feature selection in analysis of the text psychology. Applied Intelligence, 2021: p. 1-36.
2. Vazan, M. and J. Razmara, Jointly Modeling Aspect and Polarity for Aspect-based Sentiment Analysis in Persian Reviews. arXiv preprint arXiv:2109.07680, 2021.
3. Hosseinalipour, A., et al., Toward text psychology analysis using social spider optimization algorithm. Concurrency and Computation: Practice and Experience, 2021. **n/a**(n/a): p. e6325.
4. Hosseinalipour, A. and R. Ghanbarzadeh, A novel approach for spam detection using horse herd optimization algorithm. Neural Computing and Applications, 2022: p. 1-15.
5. Zhang, L., S. Wang, and B. Liu, Deep learning for sentiment analysis: A survey. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2018. **8**(4): p. e1253.
6. Young, T., et al., Recent trends in deep learning based natural language processing. ieee Computational intelligenCe magazine, 2018. **13**(3): p. 55-75.
7. Tang, F., et al., Aspect based fine-grained sentiment analysis for online reviews. Information Sciences, 2019. 488: p. 190-204.
8. Pang, B. and L. Lee, Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales. arXiv preprint cs/0506075, 2005.
9. Ouyang, X., et al. Sentiment analysis using convolutional neural network. in 2015 IEEE international conference on computer and information technology; ubiquitous computing and communications; dependable, autonomic and secure computing; pervasive intelligence and computing. 2015. IEEE.
10. Nowak, J., A. Taspinar, and R. Scherer. LSTM recurrent neural networks for short text and sentiment classification. in International Conference on Artificial Intelligence and Soft Computing. 2017. Springer.
11. Ma, D., et al., Interactive attention networks for aspect-level sentiment classification. arXiv preprint arXiv:1709.00893, 2017.
12. Hassan, A. and A. Mahmood. Deep learning approach for sentiment analysis of short texts. in 2017 3rd international conference on control, automation and robotics (ICCAR). 2017. IEEE.
13. Çano, E. and M. Morisio. A deep learning architecture for sentiment analysis. in Proceedings of the International Conference on Geoinformatics and Data Analysis. 2018.
14. Goularas, D. and S. Kamis. Evaluation of deep learning techniques in sentiment analysis from twitter data. in 2019 International Conference on Deep Learning and Machine Learning in Emerging Applications (Deep-ML). 2019. IEEE.
15. Li, R., et al., A local search algorithm with tabu strategy and perturbation mechanism for generalized vertex cover problem. Neural Computing and Applications, 2017. **28**(7): p. 1775-1785.
16. Hashim, F.A., et al., Archimedes optimization algorithm: a new metaheuristic algorithm for solving optimization problems. Applied Intelligence, 2021. **51**(3): p. 1531-1551.
17. Arora, S. and P. Anand, Binary butterfly optimization approaches for feature selection. Expert Systems with Applications, 2019. **116**: p. 147-160.
18. Hussien, A.G., et al., S-shaped binary whale optimization algorithm for feature selection, in Recent trends in signal and image processing. 2019, Springer. p. 79-87.
19. McCallum, A., Graphical Models, Lecture2: Bayesian Network Represention. PDF). Retrieved, 2019. 22.