Journal of Information Technology Management
ISSN: 2008-5893
EISSN: 2423-5059
Summer 2018
Vol. 10, No. 2, PP. 309-330
DOI: 10.22059/jitm.2017.215513.1807

University of Tehran
Faculty of Management

Citation: Samizade, R. & Mahmoudi Saeid Abad, A. (2018). The Application of Machine Learning Algorithms for Text Mining based on Sentiment Analysis Approach. *Journal of Information Technology Management,* 10(2), 309-330.

The Application of Machine Learning Algorithms for Text Mining based on Sentiment Analysis Approach

Reza Samizade [1], Elnaz Mahmoudi Saeid Abad [2]

Abstract: Classification of the cyber texts and comments into two categories of positive and negative sentiment among social media users is of high importance in the research are related to text mining. In this research, we applied supervised classification methods to classify Persian texts based on sentiment in cyber space. The result of this research is in a form of a system that can decide whether a comment which is published in cyber space such as social networks is considered positive or negative. The comments that are published in Persian movie and movie review websites from 1392 to 1395 are considered as the data set for this research. A part of these data are considered as training and others are considered as testing data. Prior to implementing the algorithms, pre-processing activities such as tokenizing, removing stop words, and n-germs process were applied on the texts. Naïve Bayes, Neural Networks and support vector machine were used for text classification in this study. Out of sample tests showed that there is no evidence indicating that the accuracy of SVM approach is statistically higher than Naïve Bayes or that the accuracy of Naïve Bayes is not statistically higher than NN approach. However, the researchers can conclude that the accuracy of the classification using SVM approach is statistically higher than the accuracy of NN approach in 5% confidence level.

Key words: ***Naïve bayes, Neural network, Sentiment analysis, Support vector machine, Text mining.***

1. Assistant Prof. of Industrial Engineering, Alzahra University, Tehran, Iran
2. MSc. Student of Industrial Engineering, Alzahra University, Tehran, Iran

Submitted: 15 / September / 2016
Accepted: 07 / October / 2017
Corresponding Author: Elnaz Mahmoudi Saeid Abad
Email: mahmoudi.fe88@gmail.com

مدیریت فناوری اطلاعات
دانشکدهٔ مدیریت دانشگاه تهران
دورهٔ ۱۰، شمارهٔ ۲
تابستان ۱۳۹۷
صص. ۳۳۰- ۳۰۹

Journal of Information Technology Management
DOI: 10.22059/jitm.2017.215513.1807

کاربرد الگوریتم‌های یادگیری ماشین در متن‌کاوی با رویکرد آنالیز احساس

رضا سمیع‌زاده[۱]، الناز محمودی سعیدآباد[۲]

چکیده: تخصیص نظرها و متن‌های منتشر شدهٔ کاربران در فضای مجازی به طبقاتی با احساسات مثبت یا منفی، در تحقیق‌های مربوط به متن‌کاوی اهمیت بسیار زیادی دارد. هدف این مقاله، استفاده و مقایسهٔ روش‌های یادگیری ماشین در طبقه‌بندی متن‌های فارسی بر اساس احساسات کاربران فعال در فضای مجازی است. داده‌های پژوهش، مجموعه نظرهای منتشرشده دربارهٔ فیلم‌های ایرانی و خارجی در بازهٔ زمانی ۱۳۹۲ تا ۱۳۹۵ در سایت‌های سینمایی و نقد فیلم فارسی زبان است. پیش از به‌کارگیری الگوریتم‌ها، فرایند پیش‌پردازش داده‌ها بر اساس تبدیل آنها به نویسه، حذف ایست‌واژه‌ها و تحلیل چندواژه‌ای انجام گرفت. برای طبقه‌بندی داده‌ها، الگوریتم‌های با نظارت نایوبیز، ماشین بردار پشتیبان و شبکهٔ عصبی استفاده شد. براساس نتایج به‌دست‌آمده، در آزمون خارج از نمونه با وجود دقت عددی بیشتر الگوریتم نایوبیز بر شبکهٔ عصبی و ماشین بردار پشتیبان بر نایوبیز، برتری آماری نایوبیز بر شبکه‌های عصبی و ماشین بردار پشتیبان بر نایوبیز اثبات نشد. با وجود این، نتایج تحقیق گویای برتری معنادار الگوریتم ماشین بردار پشتیبان بر شبکه‌های عصبی در دقت طبقه‌بندی در سطح اطمینان ۵ درصد است.

واژه‌های کلیدی: آنالیز احساس، شبکه‌های عصبی، ماشین بردار پشتیبان، متن‌کاوی، نایو بیز.

۱. استادیار دانشگاه گروه مهندسی صنایع، دانشکدهٔ فنی و مهندسی، دانشگاه الزهرا، تهران، ایران
۲. دانشجوی کارشناسی ارشد مهندسی صنایع، دانشکدهٔ فنی مهندسی، دانشگاه الزهرا، تهران، ایران

تاریخ دریافت مقاله: ۱۳۹۵/ ۰۶/ ۲۵
تاریخ پذیرش نهایی مقاله: ۱۳۹۶/۰۷/۱۵
نویسندهٔ مسئول مقاله: الناز محمودی سعیدآباد
E-mail: mahmoudi.fe88@gmail.com

مقدمه

در جهان امروز انتشار داده‌ها با سرعت چشمگیری در حال افزایش است (نیکنام و نیک‌نفس، ۱۳۹۵). استفادهٔ زیاد مردم از شبکه‌های اجتماعی موجب تولید داده‌های زیادی شده که حجم آنها با شتاب شایان توجهی در حال افزایش است. داده‌هایی که شامل اطلاعات بسیار ارزشمندی دربارهٔ احساس‌ها، عقیده‌ها، علاقه‌ها، دانسته‌ها، پرسش‌ها، شکایت‌ها و سایر خصوصیت‌های فردی کاربران این شبکه‌های اجتماعی هستند و می‌توانند برای صاحبان کسب‌وکارهایی که با این کاربران به‌صورت مستقیم ارتباط دارند، دانسته‌های بسیار ارزشمندی فراهم کنند؛ چرا که کسب‌وکارها قادرند از اهرم رسانه‌های اجتماعی برای رسیدن به موفقیت استفاده کنند (ماسلی، ۲۰۱۲).

یکی از مهم‌ترین رویکردهای نوین داده‌کاوی که به استخراج دانش از حجم وسیعی از داده‌های متنی می‌پردازد، آنالیز احساس[1] نام دارد. آنالیز احساس نوعی زمینهٔ تحقیقاتی است که به تحلیل نظرها، احساس‌ها، ارزیابی‌ها، رفتارها، گرایش‌ها و عاطفه‌های بیان‌شده با یک زبان نوشتاری می‌پردازد. آنالیز احساس به استخراج احساسات و عقاید کاربران از متن‌های منتشر شده در صفحات اینترنتی کمک شایانی می‌کند.

پژوهش‌های بسیاری در حوزهٔ استخراج دانش از داده‌های متنی انجام گرفته است و در تمام آنها تلاش شده که با استفاده از روش‌های طبقه‌بندی احساس، متن‌های منتشر شدهٔ کاربران فضای مجازی بر اساس احساسات ایجادشده در آنها، به گروه‌های مجزا و با ویژگی‌های مشخص طبقه‌بندی شوند. به‌دلیل کمبود پژوهش در زمینهٔ آنالیز احساس در متن با زبان فارسی به‌ویژه در حوزهٔ سینما، هدف این مقاله به‌کار بستن روش‌های آنالیز احساس و بررسی امکان پیاده‌سازی آنها بر متن‌های فارسی منتشر شده در فضای مجازی است. در واقع هدف اصلی این مقاله استفاده از سه الگوریتم نایو بیز، شبکهٔ عصبی و ماشین بردار پشتیبان در فرایند طبقه‌بندی داده‌های متنی منتشر شده به زبان فارسی در حوزهٔ فیلم‌های ایرانی و خارجی در سایت‌های سینمایی و نقد فیلم فارسی زبان و مقایسهٔ عملکرد آنهاست. علاوه بر این، مهم‌ترین اهداف این مقاله عبارت‌اند از: ۱. بررسی فرایند آنالیز احساس روی متن‌های فارسی و ۲. بررسی الگوریتم‌های مختلف آنالیز احساس و فرایند طبقه‌بندی داده‌های متنی منتشر شده به زبان فارسی.

در بسیاری از منابع داده‌ای که اطلاعات و نظر کاربران دربارهٔ محصولات یا خدمات را به کاربران بالقوهٔ دیگر نشان می‌دهند، برای سنجش میزان رضایت از روش‌های نموداری استفاده

---

1. Sentiment analysis

می‌شود (برای مثال ثبت یک تا پنج ستاره در سایت‌های مختلف توسط کاربر به‌منظور نمایش میزان مطلوبیت) و شاید در نگاه اول برای این کار به تحلیل متن‌های منتشر شدهٔ کاربران نیازی نباشد؛ اما استفاده از روش‌های متن‌کاوی در این زمینه راه را برای آنالیز احساس در سطح مشخص هموار می‌کند؛ به‌گونه‌ای که تحلیل متن‌ها به نتایجی منتهی شود که نشان دهد برای مثال کاربران یک محصول از کدام‌یک از ویژگی‌های آن رضایت دارند و کدام ویژگی نتوانسته است مطلوبیت کافی را برای ایشان فراهم آورد. این رضایت یا نارضایتی از مجموعه ویژگی‌ها در نهایت به سطح رضایت کلی از یک محصول منجر می‌شود که این سطح رضایت کلی را می‌توان با روش‌های نموداری نمایش داد. البته در این پژوهش تنها به بررسی آنالیز احساس در سطح کل سند پرداخته شده است. با توجه به اینکه بسیاری از سایت‌های ایرانی (مانند سایت‌های نقد فیلم که موضوع این پژوهش است)، از روش‌های نموداری برای تحلیل مثبت یا منفی بودن کامنت‌ها بهره نمی‌برند، نتایج این پژوهش می‌تواند کمک شایانی به فرایند تحلیل نظر کاربران کند.

توجه به این نکته ضروری است که در مجموعهٔ داده‌های بررسی شده در پژوهش حاضر، روابط گرافیکی وجود ندارد و فقط می‌توان تأیید یا مخالفت (like یا dislike) را با یک کامنت نشان داد؛ به بیان دیگر، روابط گرافیکی نشان نمی‌دهد که چه بخشی از بینندگان یک فیلم را پسندیده‌اند.

در این وب‌سایت‌ها امکان مشخص‌شدن رضایت (نارضایتی) در سطح مشخص وجود ندارد. اما در سایت‌هایی مانند دیجی کالا این امکان برای کاربر فراهم است. در این مورد نیز باید گفت هدف این پژوهش مانند بسیاری از پژوهش‌هایی که در ادبیات موضوع به آنها اشاره شد، آنالیز احساس در سطح سند و کلیت یک نظر دربارهٔ محصول یا فیلم است؛ به گونه‌ای که صاحبان کسب‌وکار نیازی به خواندن کامنت‌ها نداشته باشند و فقط با استفاده از یک سیستم، سهم نظرهای مثبت و منفی را دریابند و حتی به بررسی روابط گرافیکی نیز نیازی نباشد.

پیشینهٔ پژوهش

این بخش از مقاله به بررسی پیشینهٔ نظری و تجربی پژوهش می‌پردازد. در بخش پیشینهٔ نظری به مفاهیم اولیه و پرکاربرد در حوزهٔ این پژوهش اشاره می‌شود و در پیشینهٔ تجربی، پژوهش‌های انجام شدهٔ سایر محققان در حوزهٔ این پژوهش معرفی شده و نتایج آنها بیان خواهد شد.

پیشینهٔ نظری

امروزه کسب‌وکارهای بسیاری برای ارائهٔ خدمات متنوع و تعامل با مشتری از رسانه‌های اجتماعی مانند فیس بوک و توییتر استفاده می‌کنند. برای افزایش مزیت رقابتی و ارزیابی مؤثر محیط

رقابتی کسب‌وکار، شرکت‌ها نیاز دارند که نه تنها ناظر مفاهیم ایجاد شدهٔ مشتریان در سایت‌های مربوط به رسانه‌های اجتماعی خود باشند، بلکه بر اطلاعات متنی ایجاد شدهٔ رقبایشان در سایت‌های اجتماعی نیز نظارت کنند (هی، ژا و لی، ۲۰۱۳).

همان‌طور که اشاره شد، شبکه‌های اجتماعی راه‌های جدیدی برای برقراری ارتباط میان افرادی با فرهنگ‌ها، ارزش‌های اخلاقی و ارزش‌های اجتماعی گوناگون فراهم کرده‌اند. این وب‌سایت‌ها ابزار بسیار قدرتمندی برای ایجاد ارتباط میان افراد و به اشتراک‌گذاری دانش میان آنها هستند. در بسیاری از این شبکه‌های اجتماعی آنچه بیش از هر چیز به چشم می‌آید، استفاده از متن‌ها به‌عنوان ابزار انتقال دانش است. البته توجه به این نکته ضروری است که افراد در زندگی روزمرهٔ خود، در شبکه‌های اجتماعی که جزء جدایی‌ناپذیری از زندگی روزمرهٔ افراد هستند، به نحوهٔ تلفظ و نکات گرامری متن‌ها توجه زیادی نمی‌کنند و همین مسئله استخراج الگوهای منطقی و اطلاعات دقیق از میان این متن‌های منتشر شده و به اصطلاح غیرساخت‌یافته را با مشکل روبه‌رو می‌کند. متن‌کاوی پاسخی به موضوعات ارائه شده در بالاست (عرفان و همکاران، ۲۰۱۲).

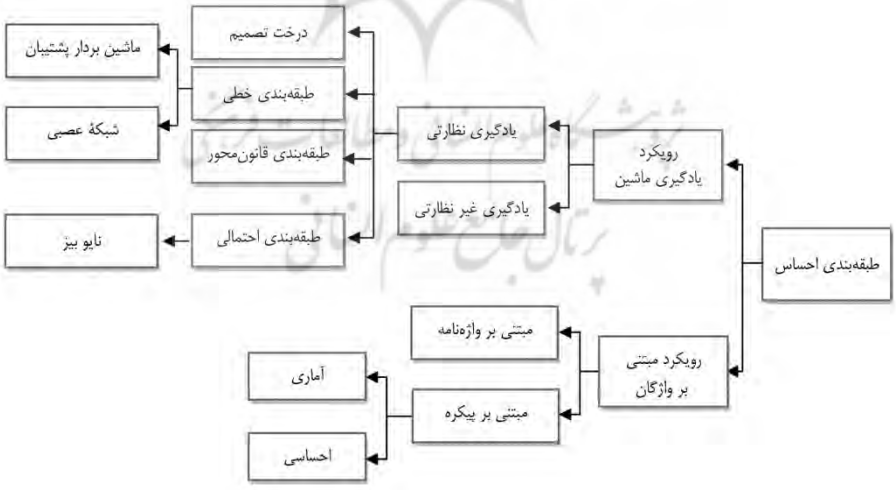
همانند داده‌کاوی که جست‌وجویی برای کشف الگو در داده‌هاست، متن‌کاوی نیز جست‌وجو در متن‌ها برای کشف الگو است. اغلب روش‌های داده‌کاوی بر داده‌های ساخت‌یافته مانند جدول‌ها متمرکزند؛ حال آن که حجم بسیار وسیعی از اطلاعات در دسترسِ دنیای بیرون، در پایگاه داده‌های متنی ذخیره شده‌اند. این پایگاه داده، شامل مجموعهٔ بزرگی از داده‌های متنی مانند کتاب‌ها، مقاله‌ها، کتابخانهٔ دیجیتال و صفحات وب است. امروزه بسیاری از سازمان‌ها اطلاعات خود را در قالب متن بایگانی می‌کنند و این مسئله اهمیت استفاده از داده‌کاوی برای این نوع داده را دوچندان کرده است (اسماعیلی، ۱۳۹۱).

یکی از مهم‌ترین مسائل موجود در متن‌کاوی، استخراج عقیده (نظر) از داده‌های متنی است. بخش شایان توجهی از متن‌های منتشر شده در وب به تولید نظرها و عقیده‌های متفاوت کاربران تعلق دارند. برای مثال می‌توان به نظرهایی که افراد مختلف در خصوص یک محصول جدید از شرکتی خاص منتشر کرده‌اند و می‌تواند مورد استفاده این شرکت خاص قرار گیرد، اشاره کرد. کاوش در متن‌های عقیده‌محور می‌تواند به‌منظور خلاصه‌سازی و آشکار شدن نظر کاربران در موضوع مربوطه استفاده شود. برای مثال می‌تواند انتظاراتی که مشتریان یک شرکت از محصولات آن دارند را مشخص کرده یا احساس رضایت یا نارضایتی کاربران یک محصول مشخص را بررسی کند. این مسئله در پیشینهٔ موضوع با نام آنالیز احساس شناخته می‌شود ژائی و آگاروال، ۲۰۰۲).

آنالیز احساس که با عنوان عقیده‌کاوی (نظرکاوی) نیز شناخته می‌شود، شامل سیستمی برای جمع‌آوری و تحلیل نظر دربارهٔ پست‌ها در بلاگ‌ها، کامنت‌ها، نقدها یا توییت‌هاست. آنالیز

احساس کاربردهای متنوعی دارد؛ برای مثال، در بازاریابی می‌تواند به موفقیت یک کمپین تبلیغاتی یا راه‌اندازی یک محصول جدید کمک کند، محصولات و سرویس‌های خاص و محبوب بین مشتریان را نمایان کرده و حتی مشخص کند کدام جمعیت چه ویژگی‌هایی را پسندیده و چه ویژگی‌هایی را نمی‌پسندند (وینودهینی و چاندراسکاران، ۲۰۱۲). پژوهش‌های بسیاری دربارهٔ تحلیل داده‌هایی که از عقیدهٔ کاربران نشئت گرفته‌اند، وجود دارد. اغلب این پژوهش‌ها به بررسی احساسات و قضاوت دربارهٔ قطبیت نظرهای (مثبت یا منفی بودن نظرها) کاربران می‌پردازند. آنالیز احساس اغلب در سه سطح سند، جمله و منظر (مشخصه) بررسی می‌شود (وینودهینی و چاندراسکاران، ۲۰۱۲). در سطح سند، تمام سند به‌عنوان ورودی در نظر گرفته می‌شود. به بیان دیگر، تمام تحلیل‌های مربوط به آنالیز احساس روی کل سند انجام می‌شود (پارادهان، والا و بلانی، ۲۰۱۶). در سطح جمله، هر جمله از یک سند به‌عنوان ورودی تجزیه و تحلیل می‌شود. در مفهوم دیگر، تحلیل‌های مربوط به آنالیز احساس روی تک‌تک جمله‌های موجود در یک سند انجام شده و در پایان با توجه به خروجی ایجاد شده، دربارهٔ کل سند قضاوت می‌شود (پارادهان و همکاران، ۲۰۱۶). در سطح منظر نیز به بررسی احساسات کاربر از جنبه‌های مختلف یک موضوع مشخص پرداخته می‌شود (مدهات، حسن و کوراشی، ۲۰۱۴).

در پیشینهٔ موضوع، دو رویکرد کلی برای آنالیز احساس وجود دارد؛ رویکرد یادگیری ماشین و رویکرد مبتنی بر واژگان. همچنین روش‌های پردازش زبان طبیعی (NLP) در این زمینه به‌خصوص در تشخیص احساسات سند، استفاده می‌شود. شکل ۱ به معرفی رویکردهای طبقه‌بندی احساس پرداخته است.



شکل ۱. رویکردهای طبقه‌بندی احساس

منبع: مدهات و همکاران (۲۰۱۴)

بر این اساس، تکنیک‌های طبقه‌بندی احساس را می‌توان به سه رویکرد زیر دسته‌بندی کرد:

- رویکرد یادگیری ماشین؛
- رویکرد مبتنی بر واژگان؛
- رویکرد ترکیبی.

رویکرد یادگیری ماشین از الگوریتم‌های معروف یادگیری و مبتنی بر ویژگی‌های زبانی استفاده می‌کند. رویکرد مبتنی بر واژگان بر اساس واژه‌نامهٔ احساسی بنا شده که این واژه‌نامه مجموعه‌ای از لغت‌ها و واژه‌های از پیش تعیین شده است و خود به زیربخش‌های واژه‌نامه‌محور و پیکره‌محور طبقه‌بندی می‌شود و از روش‌های آماری و احساسی برای یافتن قطب احساسی جمله‌ها (مثبت یا منفی بودن جمله) استفاده می‌کند. رویکردهای ترکیبی نیز از هر دو روش فوق بهره می‌برند تا بتوانند بار احساسی جمله‌ها را بررسی کنند (مدهات و همکاران، ۲۰۱۴).

روش‌های طبقه‌بندی احساس که از رویکرد یادگیری ماشین استفاده می‌کنند را می‌توان به دو گروه نظارتی و غیر نظارتی دسته‌بندی کرد. رویکردهای نظارتی زمانی به‌کار می‌روند که حجم بسیار زیادی مستند به‌منظور یادگیری (بخشی از فرایند یادگیری ماشین که با توجه به داده‌های دراختیار به مدل در مورد تشخیص صحیح بردار نتایج با توجه به ورودی‌های منجر شده به نتیجهٔ یاد شده، آموزش داده می‌شود) وجود داشته باشد؛ در غیر این صورت از روش‌های غیرنظارتی استفاده می‌شود.

رویکردهای مبتنی بر واژگان به پیدا کردن واژه‌هایی وابسته‌اند که با خود بار احساسی به همراه دارند. روش مبتنی بر واژه‌نامه، به‌دنبال یافتن برخی واژه‌های خاص در عبارت است و با استفاده از واژه‌نامهٔ مد نظر به یافتن واژه‌های هم‌معنا یا مخالف با واژهٔ یاد شده می‌پردازد. روش پیکره‌محور با فهرستی از تک کلمه‌هایی که بار عقیده‌ای دارند، کار را آغاز کرده و سعی می‌کند واژه‌های دیگری را بیابد که می‌توانند جهت‌گیری مفهومی جمله‌ها را مشخص کنند. (وینودهینی و چاندراسکاران، ۲۰۱۲).

پیشینهٔ تجربی

تحقیقات بسیاری در زمینهٔ آنالیز احساسات و طبقه‌بندی داده‌های متنی بر اساس احساسات ارائه شده، انجام گرفته است. در اغلب پژوهش‌ها منظور از دقت به‌دست آمده، نسبت تعداد متن‌هایی است که احساس موجود در آنها با استفاده از الگوریتم مد نظر به درستی تشخیص داده شده به کل کامنت‌های تجزیه و تحلیل شده‌اند (موراس، والیاتی و نتو، ۲۰۱۲).

در پژوهش کندی (۲۰۰۶) که در خصوص نقدهای منتشر شده روی فیلم‌ها انجام گرفته، دو رویکرد بررسی شده است. در رویکرد اول پس از بررسی تعداد واژه‌های دارای بار مثبت و منفی،

به طبقه‌بندی متن‌ها بر اساس احساسات پرداخته شده و در رویکرد دوم از رویکرد ماشین بردار پشتیبان استفاده شده است.

خو، لیو، لی و سانگ (۲۰۱۱) به بررسی هوش رقابتی (نظارت هدفمند با هدف اخذ تصمیمات راهبردی بر محیط رقابتی‌ای که سازمان‌ها در آن به فعالیت و رقابت می‌پردازند) با استفاده از روش‌های آنالیز احساس پرداختند و آن را یکی از کلیدی‌ترین عوامل در حوزهٔ مدیریت ریسک و سیستم‌های پشتیبان تصمیم‌گیری معرفی کردند.

موراوس و همکارانش (۲۰۱۲) به پیاده‌سازی رویکردهای ماشین بردار پشتیبان و شبکهٔ عصبی مصنوعی روی داده‌های متنی در ارتباط با فیلم‌ها در مجموعه‌های متعادل و نامتعادل[1] پرداختند. نتایج نشان داد شبکهٔ عصبی عملکرد بهتر و حداقل برابر با ماشین بردار پشتیبان در داده‌های متعادل داشته و به‌طور معناداری عملکرد بهتری از ماشین بردار پشتیبان در داده‌های نامتعادل دارد.

اسمئورائون و بوکور (۲۰۱۲) با استفاده از الگوریتم نایو بیز به بررسی فرایند آنالیز احساس روی نظرهای منتشر شدهٔ کاربران دربارهٔ فیلم‌ها در وبلاگ‌های نقد فیلم پرداخت.

جوتیسواران و کوماراسوامی (۲۰۱۳) در پژوهشی ضمن بیان نقش بسیار مهم و اثرگذار آنالیز احساس در مدیریت ارتباط با مشتری و فروشگاه‌های آنلاین، به بررسی موضوع آنالیز احساس با استفاده از الگوریتم نایوبیز پرداخت و به‌منظور استخراج مشخصه از رویکرد درخت تصمیم‌گیری در پیش‌پردازش متن‌ها استفاده کرد.

راویچانداران و کولانتایول (۲۰۱۴) حوزهٔ آنالیز احساس را به محدودهٔ آموزش الکترونیک که یکی از پرطرفدارترین رویکردهای آموزشی مبتنی بر وب است، منتقل کردند. هدف آنها بیان لزوم بررسی میزان رضایت کاربران از آموزش‌های الکترونیک و طبقه‌بندی کردن احساس آنها بود. این پژوهشگران با دقت ۹۵ درصد توانستند به طبقه‌بندی احساسات مربوط به نظر کاربران سیستم‌های آموزش آنلاین بپردازند.

بیهدان، دالال و دوشی (۲۰۱۵) به بررسی مثبت یا منفی بودن نظر کاربران و همچنین روش‌های دومرحله‌ای برای آنالیز احساسات در جنبه‌های مختلف محصول پرداخت و نشان داد که ماشین بردار پشتیبان در طبقه‌بندی متن‌ها بر اساس احساسات، دقت زیادی دارد.

گائو، خو و ونگ (۲۰۱۵) از رویکرد مبتنی بر قانون به‌منظور آنالیز احساس و طبقه‌بندی اطلاعات موجود در میکرو بلاگ‌های چینی استفاده کردند.

---

1. Balanced and unbalanced data

جیاپریا و سلوی (۲۰۱۵) به بررسی جنبه‌های مختلف احساسات گروهی از کاربران در خصوص محصولات پرداخت و به این نکته توجه کرد که کاربران، کدام‌یک از ویژگی‌های محصولات را می‌پسندند و کدام‌یک را نمی پسندند.

روش‌شناسی پژوهش

در این بخش، روش اجرای پژوهش در پنج گام معرفی شده است که عبارت‌اند از: جمع‌آوری و مرتب‌سازی داده‌ها، پیش‌پردازش متن‌ها، آموزش مدل‌های استفاده شده، به‌کارگیری معیارهای ارزیابی عملکرد و آزمون درون نمونه‌ای، آزمون خارج از نمونه به‌عنوان آزمون فرضیه.

جمع‌آوری و مرتب‌سازی داده‌ها

در این پژوهش با هدف اجرای فرایند آنالیز احساس روی متن‌های فارسی، از نظرهای ارائه شدهٔ کاربران دو سایت نقد فارسی[1] و سینما تیکت[2] دربارهٔ فیلم‌های ایرانی و خارجی استفاده شده است. این نظرها در فرمت txt ذخیره شدند. در مجموع ۲۰۱۱ نظر دربارهٔ فیلم‌های سینمایی ایرانی و خارجی جمع‌آوری شد که برای اجرای پژوهش استفاده شدند. پس از دریافت و مطالعهٔ کامنت‌ها از سایت‌های نام برده، در خصوص مثبت یا منفی بودن آنها قضاوت شد و این فرایند به ۱۰۰۱ کامنت مثبت و ۱۰۱۰ کامنت منفی انجامید. به‌منظور آموزش مدل‌های استفاده شده و آزمون آنها، ابتدا ۴۰ کامنت مثبت و ۴۰ کامنت منفی به چهار گروه دسته‌بندی شدند. در هر گروه ۱۰ کامنت مثبت و ۱۰ کامنت منفی قرار گرفت؛ به‌گونه‌ای که مثبت یا منفی بودن آنها برای سیستم از پیش تعیین نشده بود. ۹۷۰ کامنت منفی و ۹۶۱ کامنت مثبت باقی مانده برای آموزش مدل‌ها مد نظر قرار گرفت؛ سپس از مدل‌های آموزش دیده برای طبقه‌بندی چهار گروه آزمایش و سنجش دقت مدل استفاده شد.

پیش‌پردازش متن‌ها

با توجه به بستر مد نظر برای این پژوهش (نرم‌افزار رپیدماینر)، در فرایند پیش‌پردازش داده‌ها از سه فرایند حذف ایست‌واژه‌ها، تجزیه کردن[3] و تحلیل چند واژه‌ای استفاده شده است. در نرم‌افزار براساس فهرستی که کاربر ارائه می‌کند، امکان حذف ایست‌واژه‌ها وجود دارد. در این پژوهش از فهرست ایست‌واژه‌های ارائه‌شده در آزمایشگاه فناوری وب دانشگاه فردوسی مشهد استفاده شده

1. www.naghdefarsi.com
2. www.cinematicket.org
3. Tokenize

است (مشهد، ۱۳۹۵). در این فرایند حتی نیم‌فاصله‌ها نیز می‌توانند به فاصله تبدیل شوند تا واژگانی که مفاهیم شبیه به یکدیگر دارند و تنها ساختار متفاوت را به خود تخصیص داده‌اند، یکسان تلقی شوند.

وظیفهٔ تجزیه، قطعه قطعه کردن جمله به واحدهایی است که بخش[1] نامیده می‌شوند. عمل تجزیه کردن می‌تواند در حالی که جمله‌ها را به بخش‌ها تبدیل می‌کند، برخی نویسه‌های خاص را نیز حذف کند.

یکی از روش‌های پرکاربرد در حوزهٔ پیش‌پردازش متن‌ها، استفاده از تحلیل چندواژه‌ای است. با استفاده از این روش می‌توان به مفاهیم دو یا چند کلمه در کنار یکدیگر نیز اهمیت داد و در تحلیل متن‌ها از مفهوم کلمات پشت سر هم استفاده کرد؛ برای مثال بدون استفاده از فرایند تحلیل چندواژه‌ای، ممکن است عبارت «امروز هوا خوب نیست» به‌دلیل داشتن واژه‌ای با مفهوم مثبت (خوب) مثبت ارزیابی شود، اما وقتی دو واژهٔ «خوب» و «نیست» در کنار هم قرار بگیرند، مفهوم منفی جمله مشخص می‌شود.

آموزش مدل‌های استفاده شده

در این پژوهش از شبکهٔ عصبی، نایو بیز و ماشین بردار پشتیبان برای ایجاد مدل طبقه‌بندی استفاده شده است. این سه الگوریتم از خوشنام‌ترین، موفق‌ترین و پرکاربردترین الگوریتم‌های ماشین یادگیری هستند و نتایج مناسبی ارائه می‌کنند (موراس و همکاران، ۲۰۱۲). در ادامه این الگوریتم‌ها معرفی می‌شوند.

شبکهٔ عصبی

شبکهٔ عصبی، قالب پردازش اطلاعات است که از سیستم بیولوژیک عصبی انسان الهام گرفته است. طریقهٔ کار بدین صورت است که شبکه بر اساس ورود اطلاعات و خروج داده‌های مورد انتظار آموزش می‌بیند؛ پس از آموزش، وزن یال‌های لایهٔ پنهان و لایهٔ خروجی بر اساس تفاوت دادهٔ خروجی مورد انتظار با دادهٔ خروجی واقعی تعیین می‌شود. زمانی که آموزش انجام شد، دادهٔ ورودی جدید بر اساس وزن یال‌ها طبقه‌بندی می‌شود. شبکه‌های عصبی به‌خصوص زمانی که تعداد داده‌های ورودی زیاد است یا داده‌های ورودی با نویز همراه هستند، کاربرد زیادی دارند، اما در کنار این مزیت‌ها، دارای مشکلاتی همچون دشواری در فهم ساختار شبکه و مدت زمان طولانی آموزش نسبت به روش‌های دیگر یادگیری با نظارت هستند.

---

1. Token

نایو بیز

الگوریتم نایو بیز بر اساس احتمال‌های شرطی طراحی شده و مبنای اصلی آن، تئوری بیز است. اگر رخداد B به رخداد A وابسته باشد، یعنی B در شرایطی که رخداد A اتفاق افتاده باشد، رخ دهد؛ براساس تئوری بیز برای محاسبهٔ احتمال وقوع B به شرط A، الگوریتم ابتدا تمام مواردی را که رخدادهای A و B همزمان اتفاق افتاده‌اند را می‌شمارد و بعد به تعداد رخدادهای A که به تنهایی اتفاق افتاده تقسیم می‌کند تا احتمال شرطی مد نظر محاسبه شود.

الگوریتم نایو بیز نیز بر این اساس طراحی شده است. در این الگوریتم، ابتدا درصد رخداد، دو به دو بررسی شده و به درصد رخداد تک به تک تقسیم می‌شود. به‌طور مثال برای بررسی احتمال وقوع کلمهٔ «خوب» در اسناد مثبت، ابتدا تعداد رخدادهای کلمهٔ «خوب» در اسناد مثبت اندازه‌گیری می‌شود؛ سپس بر کل اسنادی که کلمهٔ «خوب» در آنها وجود داشته تقسیم می‌شود تا احتمال وجود کلمهٔ «خوب» در اسناد مثبت به‌دست آید.

ماشین بردار پشتیبان

بردارهای پشتیبان به زبان ساده، مجموعه‌ای از نقاط در فضای n بعدی داده‌ها هستند که مرز دسته‌ها را مشخص می‌کنند و مرزبندی و دسته‌بندی داده‌ها براساس آنها انجام می‌شود و ممکن است با جابه‌جایی یکی از آنها خروجی دسته‌بندی تغییر کند. در فضای دوبعدی بردارهای پشتیبان یک خط را تشکیل می‌دهند، در فضای سه‌بعدی یک صفحه و در فضای n بعدی یک ابر صفحه را شکل خواهند داد.

ماشین بردار پشتیبان داده‌ها را با عبور یک ابر صفحه و با استفاده از به‌کارگیری یک الگوریتم بهینه‌سازی طبقه‌بندی می‌کند؛ بدین‌گونه که ابتدا نمونه‌هایی مرز کلاس‌ها را تشکیل می‌دهند و تعدادی از نقاط آموزشی‌ای که کمترین فاصله را تا مرز تصمیم‌گیری دارند، به‌عنوان بردار پشتیبان در نظر می‌گیرد. در این روش هر چه بعد داده‌ها بیشتر باشد، نتیجهٔ مطلوب‌تری حاصل می‌شود. مرز بین دو کلاس بر اساس موارد زیر انتخاب می‌شود:

۱. تمام نمونه‌های کلاس اول در یک طرف و تمام نمونه‌های کلاس دوم در سمت دیگر مرز واقع شوند.
۲. مرز تصمیم‌گیری به‌گونه‌ای باشد که فاصلهٔ نزدیک‌ترین نمونه‌های آموزشی هر دو کلاس از یکدیگر در راستای عمود بر مرز تصمیم‌گیری در بیشترین مقدار ممکن خود باشد.

به بیان دیگر، در این روش ابتدا فاصلهٔ نزدیک‌ترین نمونه‌های آموزشی در دو کلاس در راستای عمود بر مرزها محاسبه شده و با تحلیل مسئلهٔ بهینه‌سازی مرز بهینه مشخص می‌شود.

دو ابر صفحهٔ موازی در دو طرف مرز تصمیم‌گیری ایجاد می‌شوند؛ به‌گونه‌ای که ابر صفحهٔ مرز، بیشترین فاصله را بین دو ابر صفحهٔ موازی ایجاد کند.

در جدول ۱ پارامترهای استفاده شده برای هر یک از این الگوریتم‌ها مشاهده می‌شود. این پارامترها با توجه به مدت زمان اجرای الگوریتم و دقت نتایج به‌دست آمده، انتخاب شده‌اند.

جدول ۱. پارامترهای الگوریتم‌های استفاده شده

الگوریتم	پارامترهای استفاده شده	
نایو بیز	--	
شبکهٔ عصبی	تعداد لایهٔ پنهان: ۱	دورهٔ آموزش: ۵۰۰
ماشین بردار پشتیبان	تابع کرنل استفاده شده: rbf	مقدار گاما: ۱

استفاده از معیارهای ارزیابی عملکرد و آزمون درون نمونه‌ای

پس از آموزش هر یک از الگوریتم‌های اشاره شده به‌کمک داده‌های آموزشی، آزمون درون نمونه‌ای روی مجموعهٔ آموزش داده شده اجرا شد که خروجی آن در جدول ۲ درج شده است.

جدول ۲. ارزیابی عملکرد الگوریتم‌ها

Accuracy : AC			
	منفی صحیح	مثبت صحیح	Class precision
پیش‌بینی منفی	A	C	CP-N
پیش‌بینی مثبت	B	D	CP-P
Class Recall	CR-N	CR-P	

تعاریف مربوط به پارامترهای جدول ۲ به شرح زیر است:

A: تعداد کامنت‌های منفی که منفی تشخیص داده شده‌اند.

B: تعداد کامنت‌های منفی که مثبت تشخیص داده شده‌اند.

C: تعداد کامنت‌های مثبت که منفی تشخیص داده شده‌اند.

D: تعداد کامنت‌های مثبت که مثبت تشخیص داده شده‌اند.

با توجه به جدول ۲، $a + d$ تعداد تشخیص‌های صحیح مدل را نشان می‌دهد. همچنین در جدول یاد شده برای همهٔ مدل‌ها $a + b =$ ۹۷۰ و $c + d =$ ۹۶۱ است؛ زیرا ۹۷۰ کامنت در مجموعهٔ منفی و ۹۶۱ کامنت در مجموعهٔ مثبت حضور دارند.

روابط زیر، معیارهای ارزیابی عملکرد جدول ۲ را توضیح می‌دهند:

رابطهٔ ۱) $$AC = \frac{a+d}{a+b+c+d}$$

فرمول محاسبهٔ accuracy است که نشان می‌دهد چه نسبتی از مجموعهٔ داده‌های بررسی شده، درست تشخیص داده شده‌اند.

رابطهٔ ۲) $$CR-N = \frac{a}{a+b}$$

رابطهٔ ۲ نشان می‌دهد از کل کامنت‌هایی که منفی هستند، چه تعداد کامنت صحیح تشخیص داده شده است.

رابطهٔ ۳) $$CR-P = \frac{d}{c+d}$$

رابطهٔ ۳ نشان می‌دهد از کل کامنت‌هایی که مثبت هستند، چه تعداد کامنت صحیح تشخیص داده شده است.

رابطهٔ ۴) $$CP-N = \frac{a}{c+a}$$

رابطهٔ ۴ نشان می‌دهد از کل کامنت‌هایی که منفی تشخیص داده شده‌اند، چه تعداد کامنت منفی بوده است.

رابطهٔ ۵) $$CP-N = \frac{d}{b+d}$$

رابطهٔ ۵ نشان می‌دهد از کل کامنت‌هایی که مثبت تشخیص داده شده‌اند، چه تعداد کامنت مثبت بوده است.

آزمون خارج از نمونه‌به‌منزلهٔ آزمون فرضیه

در این مرحله مدل‌های آموزش داده شده روی داده‌هایی که در مجموعه‌های آزمایش قرار داشتند، آزمون شدند و دقت هر یک از مدل‌ها بر اساس رابطهٔ ۱ محاسبه شد. برای هر یک از مدل‌ها که روی چهار مجموعهٔ آزمایش به اجرا درآمد، چهار معیار دقت محاسبه شد که عبارت‌اند از: ۱. دقت مدل روی مجموعهٔ آزمایش اول؛ ۲. دقت مدل روی مجموعهٔ آزمایش دوم؛ ۳. دقت مدل روی مجموعهٔ آزمایش سوم؛ ۴. دقت مدل روی مجموعهٔ آزمایش چهارم.

همچنین به‌منظور یافتن پاسخ این سؤال که کدام یک از مدل‌ها با دقت بیشتری توان طبقه‌بندی متن‌های فارسی را دارند، آزمون مقایسه‌های زوجی انجام شد.

آزمون مقایسه‌های زوجی بررسی می‌کند که آیا میانگین یک متغیر مشخص در دو جامعه، به‌طور معناداری با یکدیگر اختلاف دارند یا خیر. رد شدن فرض صفر در این آزمون به مفهوم تفاوت معنادار آماری در میانگین‌های متغیر مشخص در دو جامعه تعریف شده است. در این پژوهش متغیر مد نظر دقت محاسبه شده برای هر یک از مدل‌ها در آزمون خارج از نمونه است که برای هر مدل روی چهار نمونه آزمایش شده و چهار داده برای دقت هر مدل در اختیار است. آزمون مقایسه‌های زوجی به شکل زیر انجام می‌شود:

$$\begin{cases}\mu_1 = \mu_2 \\ \mu_1 \neq \mu_2\end{cases}$$

مقدار اختلاف دو متغیر در هر نمونه با D_i نمایش داده می‌شود و در این شرایط مقدار آمارهٔ آزمون به‌کمک رابطهٔ ۶ به‌دست می‌آید که از توزیع t-student با $n - ۱$ درجهٔ آزادی تبعیت می‌کند.

$$T_s = \frac{\mu_D}{S_D/\sqrt{n}} \qquad \text{رابطهٔ ۶)}$$

در رابطهٔ ۶، μ_D نشان‌دهندهٔ میانگین اختلاف‌های متغیر تعریف‌شده در هر دو جامعه؛ S_D معرف انحراف استاندارد اختلاف‌های متغیر تعریف‌شده در هر دو جامعه و n تعداد اعضای دو جامعهٔ مقایسه‌شده است.

این آزمون یک آزمون سخت‌گیرانه محسوب می‌شود که در آن رد فرض صفر به معنای این است که میانگین‌های دو جامعه با یکدیگر اختلاف معناداری دارند. در این پژوهش، چنانچه اختلاف معناداری بین میانگین دقت چهار نمونه از یک مدل و مدل دیگر وجود داشته‌باشد، می‌توان گفت که از نظر آماری کدام مدل بر دیگری برتری دارد.

یافته‌های پژوهش

نتایج مرحلهٔ آموزش و آزمون درون نمونه‌ای در جدول‌های ۳، ۴ و ۵ مشاهده می‌شود. جدول ۳ نتایج اجرای الگوریتم نایو بیز را نشان می‌دهد.

جدول ۳. نتایج اجرای نایو بیز

دقت = ٪ ۵۳/٦٨			
	منفی صحیح	مثبت صحیح	Class precision
پیش‌بینی منفی	۶۱۶	۴۴۸	٪ ۵۶/۰۵
پیش‌بینی مثبت	۳۵۴	۴۷۷	٪۵۷/۴۰
Class Recall	٪۵۱/۶۳	٪ ۴۹/۶۹	

همان‌طور که در جدول ۳ مشاهده می‌شود، دقت مدل نایو بیز ۵۶/۶۳ درصد تخمین زده شده و این الگوریتم توانسته است از ۹۷۰ کامنت منفی، ۶۱۶ کامنت را به درستی با عبارت منفی طبقه‌بندی کند و در تشخیص کامنت‌های منفی ۶۳/۵۱ درصد دقیق عمل کرده است. همچنین از ۹۶۱ کامنت مثبت، ۴۷۷ کامنت را به درستی با عبارت مثبت طبقه‌بندی کرده و در تشخیص کامنت‌های مثبت ۴۹/۶۹ درصد دقت داشته است.

از سوی دیگر، از ۱۱۰۲ کامنت قرارگرفته در طبقهٔ منفی، ۶۱۶ کامنت واقعاً منفی بود؛ به بیان دیگر، ارزش اخباری منفی توسط این الگوریتم ۵۶/۰۵ درصد است. از ۸۳۱ کامنتی که در طبقهٔ مثبت قرار گرفته، ۴۷۷ کامنت واقعاً مثبت بود و ارزش اخباری مثبت این الگوریتم ۵۷/۴۰ درصد است.

شایان ذکر است که مدت زمان اجرای الگوریتم نایو بیز با پارامترهای ارائه شده در جدول ۱ روی داده‌های این پژوهش، ۵ دقیقه ثبت شد.

جدول ۴ نتایج اجرای مدل شبکهٔ عصبی را نشان می‌دهد.

جدول ٤. نتایج اجرای مدل شبکهٔ عصبی

دقت = ٤٧/٦٩ ٪			
	منفی صحیح	مثبت صحیح	Class prediction
پیش‌بینی منفی	۵۸۱	۶۰۵	٪ ۵۱/۰۱
پیش‌بینی مثبت	۳۸۹	۳۵۶	٪ ۵۱/۲۱
Class recall	٪ ۶۲/۹۶	٪۴۰/۱۰	

همان‌طور که در جدول ۴ مشاهده می‌شود، دقت مدل شبکهٔ عصبی ۴۷/۶۹ درصد برآورد شده است. این الگوریتم توانست از ۹۷۰ کامنت منفی، ۵۸۱ کامنت را به درستی با عبارت منفی طبقه‌بندی کرده و در تشخیص کامنت‌های منفی ۶۲/۹۶ درصد دقیق عمل کند. همچنین از ۹۶۱ کامنت مثبت، ۳۵۶ کامنت را به درستی با عبارت مثبت طبقه‌بندی کرده و در تشخیص کامنت‌های مثبت ۴۰/۱۰ درصد دقت داشته است.

از سوی دیگر، از ۱۱۸۶ کامنت قرار گرفته در طبقهٔ منفی، ۵۸۱ کامنت واقعاً منفی بود؛ به بیان دیگر، ارزش اخباری منفی این الگوریتم ۵۱/۰۱ درصد است. از ۷۴۵ کامنت قرار گرفته در طبقهٔ مثبت، ۳۵۶ کامنت واقعاً مثبت بود و ارزش اخباری مثبت این الگوریتم ۵۱/۲۱ درصد است.

شایان ذکر است که اجرای این مدل با پارامترهای ارائه‌شده در جدول ۱ روی داده‌های این پژوهش، در مدت ۱۲ ساعت و ۳۵ دقیقه انجام گرفت.

جدول ۵ نتایج اجرای مدل ماشین بردار پشتیبان را نشان می‌دهد.

جدول ۵. نتایج اجرای مدل ماشین بردار پشتیبان

دقت = ٦٦/٩٤ ٪			
	منفی صحیح	مثبت صحیح	Class prediction
پیش‌بینی منفی	۶۶۰	۳۲۹	٪۶۸/۸۰
پیش‌بینی مثبت	۳۱۰	۶۳۲	٪۶۷/۰۹
Class recall	٪۶۸/۰۴	٪۶۵/۸۳	

دقت این مدل ۶۶/۹۴ درصد برآورد شد. این الگوریتم توانست از ۹۷۰ کامنت منفی، ۶۶۰ کامنت را به درستی با عبارت منفی طبقه‌بندی کرده و در تشخیص کامنت‌های منفی ۶۸/۰۴ درصد دقیق عمل کند. همچنین از ۹۶۱ کامنت مثبت، ۶۳۲ کامنت را به درستی با عبارت مثبت طبقه‌بندی کرد و در تشخیص کامنت‌های مثبت ۶۵/۸۳ درصد دقت داشت. از سوی دیگر، از ۹۸۹ کامنت قرار گرفته در طبقهٔ منفی، ۶۶۰ کامنت واقعاً منفی بود؛ به این معنا که ارزش اخباری منفی توسط این الگوریتم ۶۸/۸ درصد است و از ۹۴۲ کامنتی که در طبقهٔ مثبت قرار داشت، ۶۳۲ کامنت واقعاً مثبت بود و ارزش اخباری مثبت این الگوریتم ۶۷/۰۹ درصد است. مدت زمان اجرای این الگوریتم با توجه به پارامترهای ارائه شده در جدول ۱ و داده‌های این پژوهش، ۶ دقیقه محاسبه شد.

نتایج مرحلهٔ آزمون خارج از نمونه در جدول‌های ۷، ۸ و ۹ آورده شده است. در این مرحله، از الگوریتم‌های آموزش داده شده در بخش قبل برای طبقه‌بندی چهار گروه داده‌های از پیش آماده شده استفاده شد.

پیش از بررسی آزمون مقایسه‌های زوجی، خروجی تحلیل با استفاده از الگوریتم ماشین بردار پشتیبان روی چهار نمونهٔ معرفی شده در بالا ارائه می‌شود.

در جدول ۶ مشاهده می‌شود که در هر نمونه، هر نظر چه بُعد احساسی داشته (ستون واقعی) و الگوریتم چه تشخیصی از بعد مثبت و منفی بودن نظر داده است (ستون پیش‌بینی). مشخص است که اگر مقادیر مندرج در ستون‌های واقعی و پیش‌بینی یکسان باشند، در ستون تشخیص مقدار true بوده و در غیر این صورت مقدار False است. این جدول برای دو الگوریتم دیگر نیز تکمیل شد تا محاسبات لازم برای آزمون خارج از نمونه امکان‌پذیر شود.

جدول ٦. نتیجهٔ اجرای الگوریتم شبکهٔ عصبی روی چهار گروه ازمایش

نمونهٔ اول				نمونهٔ دوم				نمونهٔ سوم				نمونهٔ چهارم			
ردیف	واقعی	پیش‌بینی	تشخیص	ردیف	واقعی	پیش‌بینی	تشخیص	ردیف	واقعی	پیش‌بینی	تشخیص	ردیف	واقعی	پیش‌بینی	تشخیص
۱	neg	neg	TRUE	۱	neg	Neg	TRUE	۱	Neg	Neg	TRUE	۱	Neg	pos	FALSE
۲	neg	neg	TRUE	۲	neg	Neg	TRUE	۲	Neg	Neg	TRUE	۲	Neg	neg	TRUE
۳	neg	neg	TRUE	۳	neg	Pos	FALSE	۳	Neg	Neg	TRUE	۳	Neg	neg	TRUE
۴	neg	neg	TRUE	۴	neg	Pos	FALSE	۴	Neg	Neg	TRUE	۴	Neg	neg	TRUE
۵	neg	neg	TRUE	۵	neg	Neg	TRUE	۵	Neg	Neg	TRUE	۵	Neg	neg	TRUE
۶	neg	neg	TRUE	۶	neg	Pos	FALSE	۶	Neg	Neg	TRUE	۶	Neg	pos	FALSE
۷	neg	neg	TRUE	۷	Neg	Neg	TRUE	۷	Neg	Neg	TRUE	۷	Neg	neg	TRUE
۸	neg	neg	TRUE	۸	neg	Neg	TRUE	۸	Neg	Neg	TRUE	۸	Neg	neg	TRUE
۹	neg	neg	TRUE	۹	neg	Neg	TRUE	۹	Pos	Neg	FALSE	۹	Neg	pos	FALSE
۱۰	pos	neg	FALSE	۱۰	neg	Pos	FALSE	۱۰	Pos	Neg	FALSE	۱۰	Neg	neg	TRUE
۱۱	pos	pos	TRUE	۱۱	pos	Pos	TRUE	۱۱	Pos	Pos	TRUE	۱۱	Pos	pos	TRUE
۱۲	pos	pos	TRUE	۱۲	pos	Pos	TRUE	۱۲	Pos	Pos	TRUE	۱۲	Pos	neg	FALSE
۱۳	neg	pos	FALSE	۱۳	pos	Pos	TRUE	۱۳	Pos	Pos	TRUE	۱۳	Pos	Neg	FALSE
۱۴	pos	pos	TRUE	۱۴	pos	Neg	FALSE	۱۴	Pos	Neg	FALSE	۱۴	Pos	Pos	TRUE
۱۵	pos	pos	TRUE	۱۵	pos	Pos	TRUE	۱۵	Pos	Pos	TRUE	۱۵	Pos	Pos	TRUE
۱۶	pos	pos	TRUE	۱۶	pos	Pos	TRUE	۱۶	Pos	Pos	TRUE	۱۶	Pos	Pos	TRUE
۱۷	pos	pos	TRUE	۱۷	pos	Pos	TRUE	۱۷	Pos	Pos	TRUE	۱۷	Pos	Neg	FALSE
۱۸	pos	pos	TRUE	۱۸	pos	Pos	TRUE	۱۸	Pos	Neg	FALSE	۱۸	Pos	Pos	TRUE
۱۹	pos	pos	TRUE	۱۹	pos	Pos	TRUE	۱۹	Pos	Neg	FALSE	۱۹	Pos	Pos	TRUE
۲۰	pos	neg	FALSE	۲۰	pos	Pos	TRUE	۲۰	Pos	Pos	TRUE	۲۰	Pos	Pos	TRUE

جدول ۷ دقت اجرای هر یک از این الگوریتم‌ها را روی هر یک از چهار گروه داده‌های از پیش تعیین شده، نشان می‌دهد. در جدول ۸ نیز، محاسبات مربوط به آزمون مقایسه‌های زوجی ماشین بردار پشتیبان و نایو بیز به نمایش گذاشته شده است.

جدول ۷. نتایج اجرای مدل‌های آموزش دیده روی داده‌های جدید

داده‌ها	دقت مدل		
	مدل ماشین بردار پشتیبان	مدل نایو بیز	مدل شبکهٔ عصبی
گروه اول داده‌ها	٪۸۵	٪۶۵	٪۶۰
گروه دوم داده‌ها	٪۷۵	٪۷۵	٪۶۵
گروه سوم داده‌ها	٪۷۵	٪۶۵	٪۶۵
گروه چهارم داده‌ها	٪۷۰	٪۷۵	٪۶۵

جدول ۸. آزمون مقایسه‌های زوجی روی نتایج ماشین بردار پشتیبان و نایو بیز

	دقت مدل		اختلاف
	مدل ماشین بردار پشتیبان	مدل نایو بیز	
گروه اول داده‌ها	٪۸۵	٪۶۵	٪۲۰
گروه دوم داده‌ها	٪۷۵	٪۷۵	٪۰
گروه سوم داده‌ها	٪۷۵	٪۶۵	٪۱۰
گروه چهارم داده‌ها	٪۷۰	٪۷۵	٪-۵
میانگین اختلاف‌ها			٪۶/۲۵
انحراف استاندارد اختلاف‌ها			٪۱۱/۰۸
مقدار آماره			۱/۱۲۷
ناحیهٔ پذیرش			[-۳/۱۸۲, ۳/۱۸۲]
نتیجهٔ آزمون			دلیلی برای رد فرض صفر وجود ندارد.

نتایج آزمون نشان می‌دهد با وجود اختلاف ۶/۲۵ درصدی بین میانگین دقت ماشین بردار پشتیبان و نایو بیز و برتری ظاهری مدل ماشین بردار پشتیبان، دلیل آماری برای این برتری وجود ندارد. جدول ۹ نتایج محاسبات آزمون مقایسه‌های زوجی ماشین بردار پشتیبان و شبکهٔ عصبی را نشان می‌دهد.

جدول ۹. آزمون مقایسه‌های زوجی روی نتایج ماشین بردار پشتیبان و شبکهٔ عصبی

	دقت مدل		اختلاف
	مدل ماشین بردار پشتیبان	مدل شبکهٔ عصبی	
گروه اول داده‌ها	٪۸۵	٪۶۰	٪۲۵
گروه دوم داده‌ها	٪۷۵	٪۶۵	٪۱۰
گروه سوم داده‌ها	٪۷۵	٪۶۵	٪۱۰
گروه چهارم داده‌ها	٪۷۰	٪۶۵	٪۵
میانگین اختلاف‌ها			٪۱۲.۵
انحراف استاندارد اختلاف‌ها			٪۷.۵
مقدار آماره			۳.۳۳
ناحیهٔ پذیرش			[-۳/۱۸۲, ۳/۱۸۲]
نتیجهٔ آزمون			فرض صفر رد می‌شود

این نتایج نشان می‌دهد مدل ماشین بردار پشتیبان نسبت به مدل شبکهٔ عصبی از لحاظ دقت مدل، عملکرد بهتری دارد. جدول ۱۰ نتایج محاسبات آزمون مقایسه‌های زوجی نایو بیز و شبکهٔ عصبی را نشان می‌دهد.

جدول ۱۰. آزمون مقایسه‌های زوجی روی نتایج نایو بیز و شبکهٔ عصبی

	دقت مدل		اختلاف
	مدل نایو بیز	مدل شبکهٔ عصبی	
گروه اول داده‌ها	٪۶۵	٪۶۰	٪۵
گروه دوم داده‌ها	٪۷۵	٪۶۵	٪۱۰
گروه سوم داده‌ها	٪۶۵	٪۶۵	٪۰
گروه چهارم داده‌ها	٪۷۵	٪۶۵	٪۱۰
میانگین اختلاف ها			٪۶/۲۵
انحراف استاندارد اختلاف ها			٪۴/۱۵
مقدار آماره			۳/۰۱
ناحیهٔ پذیرش			[-۳/۱۸۲, ۳/۱۸۲]
نتیجهٔ آزمون			دلیلی برای رد فرض صفر وجود ندارد.

این آزمون نشان می‌دهد با وجود اختلاف ۶/۲۵ درصدی بین میانگین دقت مدل شبکهٔ عصبی و نایو بیز و برتری ظاهری مدل نایو بیز، از لحاظ آماری دلیلی برای این برتری وجود ندارد.

نتیجه‌گیری و پیشنهادها

در این پژوهش بر اساس داده‌های منتشر شدهٔ کاربران فارسی زبان در فضای مجازی در خصوص فیلم‌های سینمایی ایرانی و خارجی، به بررسی کاربرد سه الگوریتم نایو بیز، شبکهٔ عصبی و ماشین بردار پشتیبان در حوزهٔ طبقه‌بندی احساس پرداخته شد. جدول ۱۱ به‌طور خلاصه نتایج اجرای الگوریتم‌ها را نشان می‌دهد.

جدول ۱۱. نتیجه‌گیری نهایی

	شبکهٔ عصبی	نایو بیز	ماشین بردار پشتیبان
زمان اجرا	۱۲ ساعت و ۳۵ دقیقه	۵ دقیقه	۶ دقیقه
دقت محاسبه شده در بخش آموزش مدل	٪۴۷/۶۹	٪۵۶/۶۳	٪۶۶/۹۴
متوسط دقت در داده‌های آزمایش	٪۶۳/۷۵	٪۷۰	٪۷۶/۲۵

برتری در آزمون سخت‌گیرانه	شبکهٔ عصبی		اثبات نشد	دقت کمتر
	نایو بیز	اثبات نشد		اثبات نشد
	ماشین بردار پشتیبان	دقت بیشتر	اثبات نشد	
برتری در آزمون سهل‌گیرانه	شبکهٔ عصبی		دقت کمتر	دقت کمتر
	نایو بیز	دقت بیشتر		دقت کمتر
	ماشین بردار پشتیبان	دقت بیشتر	دقت بیشتر	

در بخش اول جدول ۱۱، دقت و مدت زمان اجرای الگوریتم‌ها نشان داده شده و در بخش دوم جدول، به بررسی برتری الگوریتم‌ها نسبت به هم بر اساس معیار دقت پرداخته شده است. برای مثال، در آزمون سهل‌گیرانه الگوریتم نایو بیز نسبت به الگوریتم شبه عصبی دقت بیشتری دارد، اما در آزمون سخت‌گیرانه برتری نایو بیز بر شبکهٔ عصبی به اثبات نرسیده است.

در آخر، موضوعات زیر را می‌توان به‌عنوان پیشنهادهایی برای پژوهش‌های بعدی برشمرد:

در این پژوهش سه روش از مجموعه روش‌های با نظارت برای انجام فرایند آنالیز احساس روی کامنت‌های منتشر شده در زبان فارسی بررسی شد؛ می‌توان در پژوهش‌های بعدی از سایر روش‌های طبقه‌بندی احساس، همچون روش‌های یادگیری بدون نظارت یا روش‌های مبتنی بر واژه‌نامه استفاده کرد و بین روش‌های طبقه‌بندی احساس روی داده‌های زبان فارسی مقایسه‌های بیشتری انجام داد.

پیشنهاد دیگر، جمع‌آوری مجموعه داده‌های استاندارد است. افرادی که در حوزهٔ زبان‌شناسی فارسی تخصص دارند، می‌توانند به تولید لغت‌نامه‌ها و فرایندهای لازم برای پیش‌پردازش داده‌ها بپردازند تا از این پایگاه داده برای فرایند آنالیز احساس استفاده شود. همچنین می‌توان مجموعه‌ای از کامنت‌ها را برای محصولات مختلف و برندهای متمایز مانند برندهای تلفن همراه آماده کرد تا برای پژوهش‌های بعدی استفاده شود.

پیشنهاد دیگر برای انجام پژوهش‌های آینده، بررسی روش‌های پیش‌پردازش مانند برگرداندن کلمات به ریشهٔ اصلی در زبان فارسی است که این فرایند نیز به تلاش بیشتر محققان حوزهٔ زبان‌شناسی فارسی نیاز دارد. یکی از مهم‌ترین محدودیت‌های اجرای این پژوهش، فرایند پیش‌پردازش روی داده‌ها بود؛ زیرا نرم‌افزارهایی مانند رپیدماینر، بستری را برای انجام فرایند پیش‌پردازش روی متن‌های انگلیسی فراهم کرده‌اند، اما اجرای آن روی زبان فارسی مشکلاتی را به‌وجود می‌آورد. از این رو بررسی روش‌های پیش‌پردازش داده‌های متنی منتشر شده به زبان فارسی می‌تواند محدودیت‌های پژوهش را از میان بردارد و راه را برای پژوهشگران بعدی هموار کند.

منابع

اسماعیلی، مهدی (۱۳۹۱). *مفاهیم و تکنیک‌های داده‌کاوی*، تهران، نیاز دانش.

نیکنام، فرزاد؛ نیک نفس، علی اکبر (۱۳۹۵). بهبود روش‌های متن‌کاوی در کاربرد پیش‌بینی بازار با استفاده از الگوریتم‌های انتخاب نمونهٔ اولیه. *فصلنامهٔ علمی ـ پژوهشی مدیریت فناوری اطلاعات*، ۸ (۲)، ۴۳۲– ۴۱۵.

References

Aggarwal, C. C., & Zhai, C. (Eds.). (2012). *Mining text data*. Springer Science & Business Media.

Bhadane, C., Dalal, H., & Doshi, H. (2015). Sentiment analysis: measuring opinions. *Procedia Computer Science*, *45*, 808-814.

Esmaili, M. (2012). *Concepts and techniques of data mainig*. Niaz Danesh Perss, Tehran. (*in Persian*)

Gao, K., Xu, H., & Wang, J. (2015). A rule-based approach to emotion cause detection for Chinese micro-blogs. *Expert Systems with Applications*, *42*(9), 4517-4528.

He, W., Zha, S., & Li, L. (2013). Social media competitive analysis and text mining: A case study in the pizza industry. *International Journal of Information Management*, *33*(3), 464-472.

Irfan, R., King, C. K., Grages, D., Ewen, S., Khan, S. U., Madani, S. A., ... & Tziritas, N. (2015). A survey on text mining in social networks. *The Knowledge Engineering Review*, *30*(2), 157-170.

Jeyapriya, A., & Selvi, C. K. (2015, February). Extracting aspects and mining opinions in product reviews using supervised learning algorithm. In *Electronics and Communication Systems (ICECS), 2015 2nd International Conference on* (pp. 548-552). IEEE.

Jotheeswaran, J., & Kumaraswamy, Y. S. (2013). Opinion mining using decision tree based feature selection through manhattan hierarchical cluster measure. *Journal of Theoretical & Applied Information Technology*, *58*(1), 72-80.

Kennedy, A., & Inkpen, D. (2006). Sentiment classification of movie reviews using contextual valence shifters. *Computational intelligence*, *22*(2), 110-125.

Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, *5*(4), 1093-1113.

Moraes, R., Valiati, J. F., & Neto, W. P. G. (2013). Document-level sentiment classification: An empirical comparison between SVM and ANN. *Expert Systems with Applications*, *40*(2), 621-633.

Mosley Jr, R. C. (2012). Social media analytics: Data mining applied to insurance Twitter posts. In *Casualty Actuarial Society E-Forum* (Vol. 2, p. 1).

Niknam, F., Niknafas, A.A. (2016). Improving Text Mining Methods in Market Prediction via Prototype Selection Algorithms. *Jornal of Information Technology Management*, 8(2), 415-434. (*in Persain*)

Pradhan, V. M., Vala, J., & Balani, P. (2016). A survey on Sentiment Analysis Algorithms for opinion mining. *International Journal of Computer Applications*, *133*(9), 7-11.

Ravichandran, M., & Kulanthaivel, G. (2014). Twitter Sentiment Mining (TSM) framework based learners emotional state classification and visualization for

e-learning system. *Journal of Theoretical & Applied Information Technology*, *69*(1), 84-90.

Smeureanu, I., & Bucur, C. (2012). Applying supervised opinion mining techniques on online user reviews. *Informatica economica*, *16*(2), 81-91.

Vinodhini, G., & Chandrasekaran, R. M. (2012). Sentiment analysis and opinion mining: a survey. *International Journal*, *2*(6), 282-292.

Xu, K., Liao, S. S., Li, J., & Song, Y. (2011). Mining comparative opinions from customer reviews for Competitive Intelligence. *Decision support systems*, *50*(4), 743-754.