Journal of Information and Communication Technology

Volume 17, Issue 65, Special Issue on "Artificial Intelligence and data analytics", October 2025, pp 73-87

Analysis of User Reviews on Digikala Store with the Aim of Detecting Deceptive Reviews

Hosein Sarlak[1*], Alireza Sheikh[2]

[1] Master's graduate, Department of Information Technology Engineering, Amirkabir University of Technology, Tehran, Iran

[2] Faculty of Management, Science, and Technology, Amirkabir University of Technology, Tehran, Iran

Received: 22 August 2024 , Revised: 29 April 2025, Accepted: 02 May 2025

Paper type: Research

Abstract

The present study examines and analyzes user reviews on the Digikala store with the goal of detecting deceptive reviews. Initially, user review data were collected and pre-processed, followed by the identification of deceptive reviews using various machine learning models and large language models. The results showed that deceptive reviews were mostly written by users with lower credibility, and reviews that received more dislikes had lower credibility. Additionally, positive reviews were the most frequent, and users who left positive reviews often received more likes. This study demonstrated that using large language models and machine learning helps improve the detection of deceptive reviews, enhances the accuracy of user review monitoring systems, and better identifies valuable and influential users.

Keywords: Deceptive Reviews, Machine Learning, Large Language Models, User Review.

[*] Corresponding Author's email: hosein.sarlak@aut.ac.ir

دوفصلنامه
فناوری اطلاعات و ارتباطات ایران

سال هفدهم، شماره ۶۵، مهر ۱۴۰۴، شمارهٔ ویژهٔ «هوش مصنوعی و تحلیل داده»، صفحه ۷۳ الی ۸۷

تحلیل نظرات کاربران در فروشگاه دیجی‌کالا با هدف تشخیص نظرات فریبنده

حسین سرلک[1*]، علیرضا شیخ[2]

[1] دانش آموخته کارشناسی ارشد، گروه مهندسی فناوری اطلاعات، دانشگاه صنعتی امیرکبیر(پلی تکنیک ایران)، تهران، ایران

[2] استادیار دانشکده مدیریت، علم و فناوری، دانشگاه صنعتی امیرکبیر(پلی تکنیک ایران)، تهران، ایران

تاریخ دریافت: ۱۴۰۳/۰۶/۰۱ تاریخ بازبینی: ۱۴۰۴/۰۲/۰۹ تاریخ پذیرش: ۱۴۰۴/۰۲/۱۲

نوع مقاله: پژوهشی

چکیده

کاربران معتبر، دارایی‌های ارزشمند یک پلتفرم به شمار می‌روند. پژوهش حاضر به بررسی و تحلیل نظرات کاربران در فروشگاه اینترنتی دیجی‌کالا با هدف تشخیص نظرات فریبنده پرداخته است. ابتدا، داده‌های نظرات کاربران جمع‌آوری و پیش‌پردازش شد و سپس با استفاده از مدل‌های مختلف یادگیری ماشین و مدل‌های زبانی بزرگ، شناسایی نظرات فریبنده صورت گرفت. نتایج نشان داد که نظرات فریبنده در بیشتر موارد، توسط کاربرانی با اعتبار پایین‌تر نوشته شده و نظراتی که تعداد دیس‌لایک[1] بیشتری دریافت کردند، دارای اعتبار کمتری هستند. همچنین، نظرات مثبت بیشترین تعداد را دارند و کاربران با نظرات مثبت، در بیشتر مواقع لایک[2]های بیشتری دریافت می‌کنند. این پژوهش نشان داد که استفاده از مدل‌های زبانی بزرگ و یادگیری ماشین به بهبود تشخیص نظرات فریبنده و افزایش دقت سیستم‌های نظارت بر نظرات کاربران و همچنین شناسایی بهتر کاربران باارزش و تأثیرگذار کمک می‌کند.

کلیدواژگان: نظرات فریبنده، یادگیری ماشین، مدل‌های زبانی بزرگ، تحلیل نظرات کاربران، شبکه‌های عصبی عمیق.

* رایانامهٔ نویسنده مسؤول: hosein.sarlak@aut.ac.ir

[1] Dislike
[2] Like

۱- مقدمه

امروزه، با گسترش فناوری اطلاعات در جهان و ورود سریع آن به زندگی روزمره، کسب‌وکار الکترونیکی جایگزین روش‌های سنتی شده است. مطالعات بسیاری نشان می‌دهند که در دنیای رقابتی کنونی، موفقیت شرکت‌ها وابسته به حفظ، نگهداری و ارتباط با مشتری می‌باشد (هسیائو[1]، ۲۰۰۹). به دنبال آن، دنیای بازاریابی شرایط جدیدی را تجربه کرده و در آینده نیز دستخوش تحولات بیشتری خواهد بود. علت اینکه بسیاری از شرکت‌ها در سطح جهان برای ترویج محصولات و خدمات خود به بازاریابی دهان به دهان گرایش نشان دادند، همین واقعیت است. یکی از مقرون به صرفه‌ترین، موثرترین و معتبرترین روش‌های بازاریابی مناسب برای این فضا، بازاریابی الکترونیکی به خصوص تبلیغات در شبکه‌های اجتماعی است (جعفرپیشه[2]، ۲۰۱۲). بنابراین، این روزها مشتریان به‌طور چشمگیری رفتارهایشان را همسو با فناوری و محیط اقتصادی دنیا تغییر می‌دهند. آنها حجم زیادی از اطلاعات را به دست آورده، از محصولات با خبر و با آنها آشنا شده و اعتمادشان را نسبت به تبلیغات از دست می‌دهند. محصولات و خدمات سفارشی را ترجیح داده و کانال‌های خرید خود را تغییر می‌دهند. بنابراین، کسب‌وکارها برای بقا مجبور به اصلاح یا حتی تغییر راهبردهای تبلیغاتی خود هستند تا از عهده تغییرات، حقایق و رفتارهای مشتریان برآیند (سیو[3] و همکاران، ۲۰۱۸).

کسب‌وکار الکترونیک تعریفی گسترده‌تر و عمیق‌تری از تجارت الکترونیک دارد که نه‌تنها شامل خرید و فروش محصولات و خدمات می‌شود بلکه ارائه خدمات به مشتریان، مشارکت شرکای کسب‌وکار، آموزش الکترونیکی و مبادلات الکترونیکی درون‌سازمانی را نیز در برمی‌گیرد. نخستین بار شرکت آی.بی.ام، اصطلاح کسب‌وکار الکترونیک را به صورت زیر تعریف کرد: «رویکردی امن، انعطاف‌پذیر و یکپارچه برای دستیابی به ارزش در کسب‌وکارهای متفاوت به وسیله ایجاد ترکیبی از سیستم‌ها و فرآیندهایی که بر فعالیت اصلی کسب‌وکار با حفظ سادگی و استفاده از فناوری اینترنت استوار هستند». در شرایط کنونی یکی از مهمترین ساختارهای مربوط به کسب اطلاعات، نظرسنجی کاربران از یکدیگر است که در قدرت و اعتباربخشی به محصول و برند و یا تضعیفش نقشی کلیدی و اساسی دارد و به همین دلیل نظرات فریبنده و آسیب‌زا در این زمینه با آسیب به محصول و توسعه آن و شرایط برند همراه است که همین مسأله اهمیت پرداختن به این موضوع را بسیار زیاد می‌کند.

با توجه به رشد روزافزون اخیر در تجارت الکترونیکی و سیستم ارائه خدمات به صورت برنامه‌های موبایلی و ارائه خدمات به صورت ارتباط کاربران با یکدیگر شناخت شرایط و مدل نظرات فریبنده بسیار مهم است. بنابراین، مدل‌سازی تشخیص نظرات در این زمینه و با هدف مقابله با این چالش به قصد رفتاری، اعتماد مشتریان و رضایت الکترونیکی مشتریان به‌ویژه در سیستم برنامه‌های موبایلی لازم و ضروری است و به توسعه و بهبود این فناوری در کشورمان کمک شایانی می‌کند. این ساختار، زمینه بهبود شرایط و محافظت از شرایط سازمان را فراهم می‌کند و منافع بسیاری برای سازمان دارد که می‌توان از آن در امر سیاست‌گذاری در این حوزه استفاده کرد. در کنار این مسأله، خلأ نظری و پژوهشی از دلایل دیگر اهمیت و ضرورت پژوهش حاضر در این حیطه است.

۲- مبانی نظری

۲-۱- نظرات برخط فریبنده

مصرف‌کننده آنلاین در عصر اقتصاد پلتفرم (یعنی شکل جدیدی از اقتصاد که از فناوری‌های دیجیتال و پلتفرم‌های اینترنتی برای هماهنگی و سازمان‌دهی منابع استفاده می‌کند) بسیار مهم است (وانا و لامبرشت، ۲۰۲۱؛ یو و همکاران، ۲۰۲۲). بررسی‌های برخط، اطلاعات ارزشمندی را برای خرید در اختیار مصرف‌کنندگان قرار می‌دهند (لی و همکاران ۲۰۲۱؛ لی و همکاران ۲۰۱۹). برای کسب‌وکارهای با الگوی تجارت الکترونیک، بررسی‌های برخط می‌تواند به درک بهتر نگرانی‌های مصرف‌کنندگان برای بهبود محصولات یا خدمات کمک کند (وانگ و همکاران، ۲۰۲۲). در نتیجه، بررسی‌های برخط به منبع اطلاعاتی مهمی برای مصرف‌کنندگان و کسب‌وکارهای تجارت الکترونیک تبدیل شدند (لمب و همکاران، ۲۰۲۰). بر اساس یک نظرسنجی عمومی، ۹۲٪ مصرف‌کنندگان در واقع از نظرات برخط به‌عنوان مرجع در خرید برخطشان استفاده می‌کنند و ۹۰٪ این مصرف‌کنندگان می‌گویند که نظرات مثبت، بیشتر به استفاده از یک محصول در تصمیم‌گیری‌های خرید به آنها کمک می‌کنند (ژانگ و همکاران ۲۰۲۲). با وجود این، به دلیل عدم بررسی دقیق بررسی‌های برخط، تعداد زیادی از بررسی‌های فریبنده تولید و برای سودجویی استفاده می‌شوند و به مشکل بزرگی در عصر اقتصاد پلتفرم تبدیل شدند (پانگ و همکاران2022 .؛ ژوانگ و همکاران ۲۰۱۸).

[1] Hsiao
[2] Jafarpishe
[3] Cui

نظرات فریبنده شامل دیدگاه‌های گمراه‌کننده‌ای هستند که با هدف تبلیغی یا ضدتبلیغ نگاشته و باعث فریفتگی در نظرکاوی شده و نیز در کسب تحلیل معنایی اختلال ایجاد می‌کنند. در بیشتر موارد، بررسی‌های فریبنده به تجربه‌های بد در خرید برخط منجر می‌شوند (جزیری، ۲۰۱۹؛ لمب و همکاران، ۲۰۲۰)، ضرر اقتصادی برای مصرف‌کنندگان برخط ایجاد می‌کنند و در نتیجه به اعتبار کسب‌وکارهای درگیر و بیشتر اقتصاد پلتفرم در کل آسیب می‌رسانند (ژانگ و همکاران، ۲۰۱۸). برای مثال، تحقیقات Daily Mail نشان داد که شرکت‌های آسیب‌زا در سایت آمازون[1] نظرات فریبنده‌ای را به خرده‌فروشان برخط می‌فروشند که میلیون‌ها مشتری آمازون را با تصمیم‌های خرید بد احتمالی در معرض خطر قرار می‌دهند. از آنجایی که بررسی‌های برخط فریبنده به موضوعی فراگیر در اقتصاد پلتفرم امروزی تبدیل شدند، مطالعه بررسی‌های فریبنده بیش از پیش توجه صنعت و دانشگاه را به خود جلب کرده است (ژانگ و همکاران، ۲۰۲۳).

۲-۲- انگیزه اثرگذار بر نظرات برخط فریبنده

انگیزه بررسی‌های آنلاین فریبنده، ناشی از ارزش تجاری آنهاست که با گسترش سریع این نوع بررسی‌ها در عصر اقتصاد پلتفرم همراه شده است (چاترجی و همکاران ۲۰۲۱؛ پلوتکینا و همکاران ۲۰۲). به‌طور کلی، کمبود چارچوب برای انگیزه‌های بررسی برخط فریبنده در زمینه تجارت الکترونیک وجود دارد.

به‌طور کلی، انگیزه بررسی‌های برخط فریبنده را می‌توان به انگیزه بیرونی و انگیزه داخلی تقسیم کرد (حسین و همکاران، ۲۰۱۸). بررسی‌های فریبنده ناشی از انگیزه خارجی به دستکاری بررسی‌های برخط بوسیله فروشندگان برای انگیزه مالی اشاره دارد و در بیشتر موارد این نظرات فریبنده بوسیله پوسترهای پولی (کومار و همکاران ۲۰۱۹) برای منافع مالی یا مزیت رقابتی ارسال می‌شوند (لی و همکاران ۲۰۱۸). در اینجا، پوسترهای پولی به بازبینان فریبنده‌ای اشاره دارند که بوسیله فروشندگان برای ارسال نظرات برخط جعلی بر اساس اولویت‌های فروشندگان پول می‌گیرند. در فضای رقابتی فزاینده بازار، تعداد بیشتری از مشاغل تجارت الکترونیک نگران هستند که سایر مشاغل بررسی‌های برخط را دستکاری کنند و از این طریق بر عملکرد و سودآوری کسب‌وکار اثر منفی بگذارند (گاسلینگ و همکاران، ۲۰۱۸).

همچنین، مصرف‌کنندگان می‌توانند نظرات فریبنده‌ای را برای دریافت پاداش از فروشندگان و ارضای نیازهای روان‌شناختی خود - انگیزه درونی انجام دهند. سازوکار پاداش، انگیزه مهمی برای مصرف‌کنندگان برای ارسال نظرات منفی در این زمینه است (یو و همکاران، ۲۰۲۲). یک نظرسنجی از مصرف‌کننده نشان می‌دهد که پاداش‌های مالی مصرف‌کنندگان را به سمت ارسال نظرات برخط فریبنده‌تر سوق می‌دهند (وو و همکاران، ۲۰۲۰) تا جایی که به نیازهای روانی مربوط می‌شود. اندرسون و سیمستر (۲۰۱۴)، معتقدند که گاهی بررسی‌های برخط فریبنده برای ارضای نیازهای غیرپولی مصرف‌کنندگان مانند احترام از طرف گروه یا رهایی از خشم ناشی از تجربیات بد ارسال می‌شوند. کاربران برخط، گاهی هویت‌های مجازی متعددی را برای ارسال بررسی‌های برخط فریبنده ایجاد کرده یا از هویت‌های جعلی برای ارسال اطلاعات گمراه‌کننده در فضاهای مجازی برای جلب توجه دیگران استفاده می‌کنند (وانگ و همکاران ۲۰۱۹). مطالعات قبلی پیشنهاد می‌کردند که نیاز به تعامل اجتماعی، میل به انگیزه‌های اقتصادی و مشاوره‌جویی انگیزه‌های مهمی برای مصرف‌کنندگان برای نوشتن نظرات برخط هستند (کاپور و همکاران ۲۰۲۱؛ وانا و لامبرشت، ۲۰۲۱).

۲-۳- پیشینه پژوهش

لی[2] و همکارانش [۱۴]، سعی کردند با استفاده از شبکه‌های عصبی کانوولوشنی (CNN)، عمل یادگیری را انجام دهند و برای تشخیص نظرات فریبنده از این شبکه‌ها استفاده کردند. نتایجی که از این پژوهش به‌دست آمد نشان از کارآمدی شبکه‌های CNN در مسئله تشخیص نظرات فریبنده داشت. رن[3] و ژانگ[4] (۲۰۱۶)، در مطالعه‌ای برای تشخیص نظرات فریبنده از یادگیری در سطح سند استفاده کردند. در این مقاله، ابتدا یک سند به مدل داده می‌شود و با استفاده از CNN که با یک شبکه عصبی بازگشتی گیت‌گذاری‌شده[5] ترکیب شده، جملات، ساختار آنها و نحوه نمایش آنها یادگیری می‌شود و بردارهای سند با این روش استخراج می‌شوند. سپس، این بردارها به‌صورت مستقیم برای یادگیری برای تشخیص نظرات فریبنده استفاده می‌شوند. وانگ[6] و همکارانش در سال ۲۰۱۸، مقاله‌ای را منتشر کردند که در آن با استفاده از شبکه‌های حافظه کوتاه‌مدت طولانی (LSTM) سعی در تشخیص نظرات داشتند. نکته قابل توجهی که در این مقاله وجود دارد این است که آنها با مقایسه روش‌های مختلف یادگیری ماشین و مقایسه آنها با روش‌های جدیدتر و مدل‌های یادگیری عمیق به این نتیجه رسیدند که از بین

[1] www.Amazon.com
[2] Lie
[3] Ren
[4] Zhang
[5] RNN-Gated
[6] Wang

روش‌های مختلف، (وزنی) یادگیری عمیق برای این مسئله کارایی بهتری دارند و شبکه‌های LSTM نسبت به ماشین بردار پشتیبان (SVM) مناسب‌تر هستند. در پژوهش دیگری که در سال ۲۰۲۰ بوسیله ماهالاکشمی[1] و همکارانش انجام شد نیز از شبکه‌های حافظه کوتاه‌مدت طولانی (LSTM) استفاده شد. آنها در این پژوهش از مجموعه داده OpSpam استفاده کرده و با دستیابی به دقتی در حدود ۸۰ تا ۸۵ درصد و مقایسه آن با مدل‌های کلاسیک یادگیری ماشین مانند ماشین بردار پشتیبان، k-نزدیک‌ترین همسایه و مدل‌های دیگر به دقت بهتری دست یافتند.

۲-۴- روش‌شناسی

پژوهش حاضر، با هدف تشخیص نظرات فریبنده فارسی ضمن ایجاد مجموعه داده مناسب پس از بررسی روش‌های منتخب یادگیری ماشین و یادگیری عمیق به اعمال و مقایسه این روش‌ها بر اساس سنجه‌های عملکردی دقت، صحت، امتیاز F1 و ... اقدام می‌کند. در این مطالعه، از مدل‌های یادگیری مختلفی برای انجام پروژه استفاده شد.

۲-۵- مدل اجرایی پژوهش حاضر

این پژوهش، به‌صورت عملیاتی در شرکت دیجی‌کالا انجام شد. مراحل انجام پژوهش به شرح زیر است:

۲-۵-۱- مرحله اول: جمع‌آوری، پیش‌پردازش و برچسب‌گذاری

- جمع‌آوری:

با توجه به وب‌سرویس‌های عمومی‌دیجی‌کالا به دریافت نظر از ۶۰۰ هزار کاربر اقدام و با حذف داده‌های غیرضروری مجموعه داده‌ای مطابق جدول زیر ایجاد شد:

جدول ۱. ایجاد یک مجموعه داده جدید مبتنی بر مجموعه داده موجود

عنوان فی‌لد	توضی‌حات
Comment	متن کامنت افراد
User_ID	شناسه کاربر
Like	تعداد لای‌ک‌های کامنت
Dislike	تعداد دی‌س‌لای‌ک‌های کامنت
User_score	اعتبار کاربر

- پیش‌پردازش:

برای محاسبه اعتبار از فرمول زیر استفاده شد:

$$score = \big((a * L) - (b * D)\big) + (c * C) \quad (۱)$$

برای محاسبه امتیاز اعتبار کاربران از فرمولی استفاده می‌شود که در آن تعداد لایک‌ها (L)، تعداد دیس‌لایک‌ها (D) و محتوای کامنت (C) به‌عنوان ورودی در نظر گرفته می‌شوند. ضرایب a، b و c وزن نسبی هر ویژگی را مشخص می‌کنند. ضریب a اهمیت لایک‌ها، ضریب b اهمیت دیس‌لایک‌ها و ضریب c اهمیت محتوای کامنت را نشان می‌دهد. برای مثال، اگر لایک‌ها نسبت به محتوای کامنت اهمیت بیشتری داشته باشند ضریب a بیشتر از ضریب c در نظر گرفته می‌شود. این ضرایب این امکان را به ما می‌دهند تا تاثیر هر یک از ویژگی‌ها را روی امتیاز نهایی کاربر تنظیم کرده و امتیاز نهایی را بتوان به‌طور دقیق‌تری محاسبه کرد.

پس از محاسبه اعتبار خام کاربران، برای نرمال‌سازی این امتیازات در بازه ۰ تا ۱۰۰ از فرمول زیر استفاده می‌شود. همچنین، این نرمال‌سازی برای مقایسه بهتر و عادلانه‌تر امتیازات کاربران مختلف انجام می‌شود.

$$\frac{(\text{minScore} - Score)}{(\text{min}Score_{\max} - Score)} \times 100 = Normalized\ Score \quad (۲)$$

- برچسب‌گذاری:

یک رابط کاربری برای برچسب‌گذاری نظرات کاربران ایجاد کردیم که به کارشناسان خبره امکان می‌دهد تا با دقت و سرعت نظرات را ارزیابی و برچسب‌گذاری کنند. هر زمان که دکمه «بارگیری مجدد داده‌ها» کلیک شود، مجموعه‌ای از کامنت‌های یک محصول به کارشناس نمایش داده می‌شود. این کامنت‌ها شامل اطلاعاتی چون متن نظر، تعداد لایک‌ها و دیس‌لایک‌ها و امتیاز اعتبار کاربر هستند. در این رابط کاربری، کارشناس می‌تواند فراوانی احساسات نظرات را در سه دسته منفی، خنثی و مثبت مشاهده کند. این اطلاعات به کارشناسان کمک می‌کنند تا الگوهای کلی احساسات کاربران را درک کنند. همچنین، توزیع میانگین اعتبار کاربران در این سه دسته نیز نمایش داده می‌شود تا به کارشناسان نشان دهد که کاربران با چه سطح اعتباری نظرات خود را ارائه کردند. کارشناسان با توجه به این اطلاعات و همچنین با بررسی میزان لایک و دیس‌لایک هر کامنت قادر خواهند بود تا تصمیم‌گیری کنند که آیا یک نظر فریبنده است یا غیرفریبنده. این تصمیم‌گیری بر اساس تحلیل جامع از محتوای نظر، بازخورد کاربران و اعتبار نویسنده نظر انجام می‌شود.

۲-۵-۲- مرحله دوم: پردازش

در این پروژه، از مدل‌های زبان بزرگ (LLM) و به‌ویژه مدل

[1] Mahalakshmi

ParsBert استفاده شد که مدل ترانسفورمری پیشرفته برای زبان فارسی است. این مدل‌ها قابلیت استخراج ویژگی‌های معنایی متن را دارند و به ما کمک می‌کنند تا داده‌های متنی را به بردارهای عددی قابل استفاده در مدل‌های یادگیری ماشین تبدیل کنیم.

نحوه تبدیل متن به بردار به اینصورت است که هر کلمه از متن به یک توکن (Token) تبدیل می‌شود. این فرآیند با استفاده از یک واژه‌نامه خاص انجام می‌شود که BERT از آن استفاده می‌کند. واژه‌نامه BERT شامل کلمات رایج و زیرکلمات است که به مدل کمک می‌کند تا حتی کلمات ناشناخته را به توکن‌های قابل فهم تبدیل کند. برای هر جمله، دو توکن خاص اضافه می‌شود: `[CLS]` در ابتدای جمله و `[SEP]` در انتهای جمله. `[CLS]` مخفف «classification» است و به‌عنوان نماینده کل جمله عمل می‌کند در حالی که `[SEP]` برای جداسازی جملات استفاده می‌شود.

توکن‌های ورودی به بردارهایی با ابعاد ثابت تبدیل می‌شوند که به این فرآیند Token Embedding گفته می‌شود. این بردارها شامل اطلاعاتی درباره خود توکن‌ها و همچنین موقعیت آنها در جمله هستند. بعلاوه، بردارهای تعبیه‌شده شامل Segment Embedding و Position Embedding نیز هستند. Segment Embedding مشخص می‌کند که هر توکن به کدام جمله (اگر بیش از یک جمله وجود داشته باشد) تعلق دارد و Position Embedding موقعیت هر توکن را در جمله نشان می‌دهد.

بردارهای تعبیه‌شده از توکن‌ها از طریق چندین لایه ترنسفورمر عبور می‌کنند. هر لایه شامل سازوکارهای Self-Attention و Feed-Forward Neural Networks است که به مدل کمک می‌کند تا وابستگی‌های پیچیده بین کلمات را درک کند. Self-Attention به مدل اجازه می‌دهد که به تمام کلمات در جمله توجه کند و اهمیت هر کلمه را نسبت به کلمات دیگر درک کند، در حالی که Feed-Forward Neural Networks اطلاعات پردازش شده را تقویت و اصلاح می‌کند.

پس از عبور از لایه‌های ترنسفورمر، بردار نهایی هر توکن استخراج می‌شود. بردار مربوط به توکن `[CLS]` به‌عنوان نماینده کل جمله در نظر گرفته می‌شود و می‌تواند به‌عنوان ورودی برای وظایف مختلف مانند طبقه‌بندی، ترجمه یا تولید متن استفاده شود. این بردار نهایی ترکیبی از تمام اطلاعات پردازش‌شده از جمله است و به مدل‌های یادگیری ماشین کمک می‌کند تا با دقت بیشتری به تحلیل و پیش‌بینی‌های خود بپردازند.

همچنین برای تشخیص بهتر و دقیق‌تر، باید متادیتای بیشتری در اختیار مدل قرار داد. این متادیتا شامل اطلاعات اضافی و مرتبط با هر نظر است که در فهم بهتر و دقیق‌تر نظرات به مدل کمک می‌کند. در پژوهش حاضر، متادیتای اضافه‌شده شامل تعداد لایک‌ها، تعداد دیس‌لایک‌ها، اعتبار کاربران و برچسب نظرات است. این اطلاعات به مدل کمک می‌کنند تا علاوه بر متن نظرات، ویژگی‌های دیگری که ممکن است بر تحلیل نظرات تأثیر بگذارند را نیز مد نظر قرار دهد.

۲-۵-۳- مرحله سوم: یادگیری و آزمایش و ذخیره

- تقسیم داده‌ها به مجموعه‌های آموزشی و آزمایشی

ابتدا داده‌ها به دو مجموعه آموزشی[1] و آزمایشی[2] تقسیم می‌شوند. این کار به منظور ارزیابی دقیق‌تر مدل انجام می‌شود. در این پروژه، ۷۵ درصد از داده‌ها برای آموزش و ۲۵ درصد برای آزمایش مدل استفاده شدند. همچنین، برای حفظ توزیع یکنواخت برچسب‌ها در هر دو مجموعه از روش طبقه‌بندی‌شده (stratified sampling) استفاده می‌شود.

- راه‌اندازی محیط PyCaret

برای ساده‌سازی فرآیند آموزش و ارزیابی مدل‌های یادگیری ماشین، از کتابخانه PyCaret استفاده می‌شود. PyCaret محیط یکپارچه‌ای برای پیش‌پردازش داده‌ها، آموزش مدل‌ها و ارزیابی عملکرد آنها فراهم می‌کند. در این مرحله، محیط PyCaret با مشخص‌کردن متغیر هدف و تنظیمات اولیه راه‌اندازی می‌شود تا فرآیند آموزش مدل‌ها به صورت خودکار انجام شود.

- مقایسه مدل‌ها

PyCaret قابلیت مقایسه خودکار مدل‌های مختلف یادگیری ماشین را فراهم می‌کند. در این مرحله، مدل‌های مختلف با استفاده از اعتبارسنجی متقابل[3] مقایسه می‌شوند تا بهترین مدل بر اساس معیارهای مختلف ارزیابی انتخاب شود. این مقایسه شامل بررسی معیارهایی مانند دقت[4]، صحت[5]، یادآوری[6] و غیره است.

- پیش‌بینی با بهترین مدل

پس از انتخاب بهترین مدل، این مدل برای پیش‌بینی روی مجموعه آزمایشی استفاده می‌شود. در این مرحله، عملکرد مدل در پیش‌بینی داده‌های جدید و ناشناخته ارزیابی می‌شود. مدل با استفاده از

[1] training set
[2] test set
[3] cross-validation
[4] accuracy
[5] precision
[6] recall

داده‌های آزمایشی امتحان می‌شود تا معیارهای مختلفی مانند دقت، دقت مثبت و یادآوری محاسبه و گزارش شوند.

- ترسیم ماتریس سردرگمی

برای بررسی عملکرد مدل در دسته‌بندی صحیح نمونه‌ها، ماتریس سردرگمی رسم می‌شود. این ماتریس نشان می‌دهد که مدل تا چه حد در دسته‌بندی صحیح نمونه‌ها موفق بوده است و تعداد نمونه‌های درست و نادرست دسته‌بندی شده را نمایش می‌دهد. این ارزیابی به درک بهتر نقاط قوت و ضعف مدل کمک می‌کند.

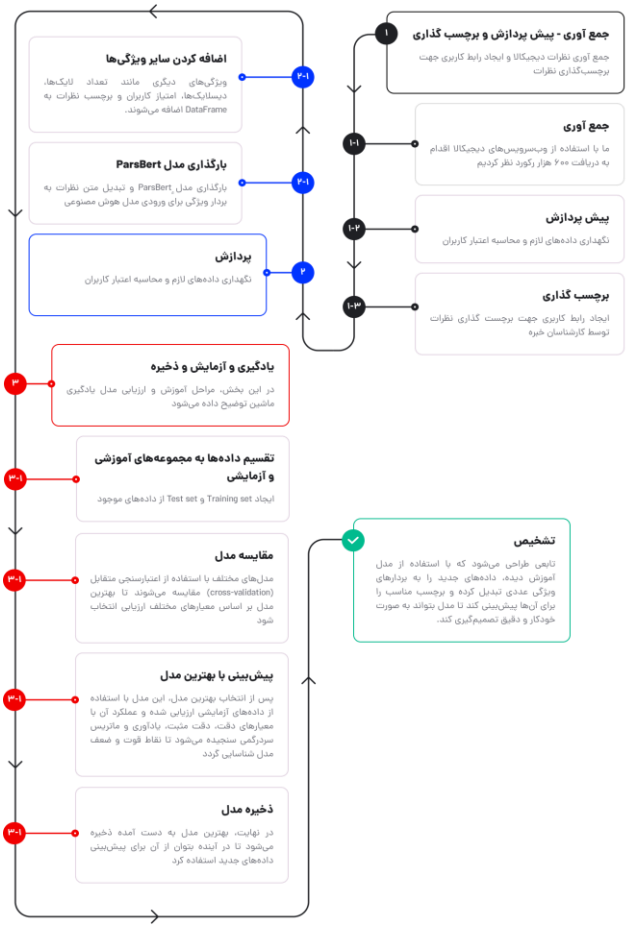
- ذخیره مدل

در نهایت، بهترین مدل به‌دست‌آمده ذخیره می‌شود تا در آینده بتوان از آن برای پیش‌بینی داده‌های جدید استفاده کرد. ذخیره مدل این امکان را به ما می‌دهد که بدون نیاز به آموزش مجدد از مدل آموزش‌دیده در محیط‌های مختلف و برای داده‌های جدید استفاده کنیم.

۲-۵-۴- مرحله چهارم: تشخیص

برای پیش‌بینی داده‌های جدید، تابعی طراحی می‌شود که بتواند با استفاده از مدل آموزش‌دیده، برچسب داده‌های جدید را پیش‌بینی کند. این تابع ابتدا داده‌های ورودی جدید را دریافت کرده و آنها را به بردارهای ویژگی عددی تبدیل می‌کند، همان‌طور که در مرحله پیش‌پردازش برای داده‌های آموزشی انجام شد. سپس، این بردارهای ویژگی به مدل ذخیره‌شده وارد می‌شوند و مدل با استفاده از الگوها و روابطی که در مرحله آموزش یاد گرفته است، برچسب مناسب را برای هر ورودی جدید پیش‌بینی می‌کند.

فلوچارت روش پیشنهادی در شکل ۱ نشان داده شد. مشخصات داده‌ها نیز در جدول ۲ قابل مشاهده است.



شکل ۱. فلوچارت روش پیشنهادی

جدول ۲. مشخصات داده‌های استفاده‌شده

نام فیلد	نوع داده	توضیح
Comment	متن[1]	متن نظر کاربر
Like	عدد صحیح[2]	تعداد لایک‌هایی که نظر دریافت کرده
Dislike	عدد صحیح	تعداد دیس‌لایک‌هایی که نظر دریافت کرده
user_score	عدد اعشاری[3]	امتیاز اعتبار کاربر (مقداری بین ۰ تا ۱۰۰)
sentiment	متن	احساس نظر (مثبت یا منفی)
User_ID	متن	شناسه یکتای کاربر
Date	عدد صحیح	تاریخ ارسال نظر به صورت یونیکس تایم‌استمپ[4]
label	متن[5]	برچسب نظر (گمراه‌کننده یا غیر گمراه‌کننده)
Brand_Id	عدد صحیح	شناسه برند محصول
Product_Id	متن	شناسه یکتای محصول

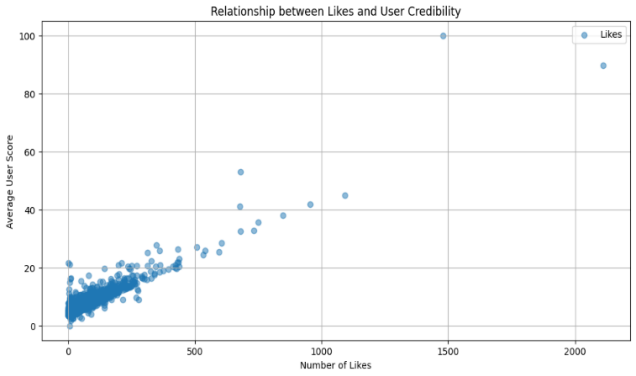
[1] String
[2] Integer
[3] Float
[4] Unix Timestamp
[5] String

۳- یافته‌ها

نظرات در سه سطح مثبت، خنثی و منفی بررسی می‌شوند. بیشترین تعداد نظرات در دسته نظرات مثبت قرار دارند که حدود ۷۰،۰۰۰ نظر را شامل می‌شود. این میزان بالای نظرات مثبت نشان‌دهنده رضایت کلی کاربران از محصولات و خدمات ارائه‌شده است. در مقابل، نظرات خنثی با حدود ۲۰،۰۰۰ نظر در مرتبه دوم قرار دارند که بیانگر عدم تمایل کاربران به ابراز نظر خاصی درباره محصولات یا خدمات است. در نهایت، نظرات منفی با حدود ۱۵،۰۰۰ نظر کمترین تعداد را به خود اختصاص داده‌اند که نمایانگر نارضایتی کاربران از برخی جنبه‌های محصولات یا خدمات است. این توزیع فراوانی نظرات به تحلیل‌گران کمک می‌کند تا درک بهتری از رضایت و نارضایتی کاربران نسبت به برند و محصولات کسب کنند و بتوانند به بهبود کیفیت خدمات و محصولات بپردازند.

همچنین، بررسی تغییرات نظرات نشان داد که نظرات مثبت طی سال‌ها به‌طور کلی روند افزایشی داشته و بیشترین تعداد نظرات مثبت در سال‌های ۲۰۱۹ و ۲۰۲۰ مشاهده شد که نشان‌دهنده افزایش رضایت کاربران در این سال‌ها است. نظرات خنثی نیز طی سال‌ها تا حدودی ثابت بوده و تغییرات کمتری نسبت به نظرات مثبت داشتند. این نظرات بیانگر کاربرانی هستند که نظر خاصی نسبت به محصولات یا خدمات نداشتند و روند آنها در طول زمان پایدار بود. نظرات منفی نیز طی سال‌ها تغییرات زیادی نداشته و تا حدودی ثابت باقی ماندند. با این حال، در برخی از سال‌ها مانند سال ۲۰۲۰، افزایش جزئی در تعداد نظرات منفی مشاهده می‌شود که نشان‌دهنده مشکلات خاصی در آن سال‌ها است. این توزیع فراوانی نظرات بر اساس سال به تحلیل‌گران کمک می‌کند تا تغییرات احساسات کاربران را در بازه زمان بررسی کرده و بتوانند به شناسایی روندها و نقاط قوت و ضعف محصولات و خدمات بپردازند.

نمودار ارتباط بین لایک‌های نظرات و اعتبار کاربر نشان می‌دهد که با افزایش تعداد لایک‌های نظرات، امتیاز اعتبار کاربران نیز به‌طور قابل توجهی افزایش می‌یابد. بیشتر کاربران با تعداد لایک‌های کمتر از ۵۰۰ دارای امتیاز اعتبار پایین‌تری هستند اما برخی کاربران با تعداد لایک‌های بالاتر، امتیاز اعتبار بسیار بالاتری دارند. این رابطه مثبت بین تعداد لایک‌ها و اعتبار کاربران نشان می‌دهد که نظرات مورد تایید بیشتر کاربران (با تعداد لایک بالا) اغلب بوسیله کاربران با اعتبار بالاتر نوشته شده است. این اطلاعات به شناسایی کاربران با ارزش و تاثیرگذار و همچنین به بهبود راهبردهای تعامل با کاربران کمک می‌کنند.



شکل ۲. ارتباط بین لایک‌های نظرات و اعتبار کاربر

۳-۱- مشخصات آزمایش‌های انجام‌شده

جدول عملکرد مدل‌های یادگیری ماشین در پیش‌بینی دسته‌ها (جدول ۳)، شامل معیارهای مختلفی برای ارزیابی مدل است که به تحلیل دقیق‌تر عملکرد مدل‌ها کمک می‌کند. این معیارها شامل دقت[۱]، صحت[۲]، بازخوانی[۳]، امتیازF1[۴]، ضریب کاپا[۵]، ضریب همبستگی متیوز[۶] و زمان اجرا[۷] هستند. دقت نشان می‌دهد که چه نسبتی از کل نمونه‌ها به درستی طبقه‌بندی شدند. دقت نشان می‌دهد که از بین تمام نمونه‌هایی که به‌عنوان مثبت پیش‌بینی شدند، چه نسبتی به درستی مثبت هستند که به‌ویژه در مسائلی با هزینه مثبت کاذب بالا اهمیت دارد. بازخوانی نشان می‌دهد که از بین تمام نمونه‌های مثبت، چه نسبتی به درستی به‌عنوان مثبت پیش‌بینی شدند که زمانی مهم است که هزینه منفی کاذب بالا باشد. امتیاز F1 میانگینی از دقت و بازخوانی است و توازن بین این دو معیار را نشان می‌دهد که به‌ویژه زمانی مفید است که توزیع داده‌ها نامتوازن است. ضریب کاپا معیاری برای سنجش میزان توافق مدل با واقعیت است که اثرات تصادفی را در نظر می‌گیرد. مقدار بالاتر کاپا نشان‌دهنده توافق بیشتر مدل با داده‌های واقعی است. ضریب همبستگی متیوز، یک معیار کلی برای سنجش عملکرد مدل‌های یادگیری ماشین است که هم داده‌های مثبت و هم منفی را در نظر می‌گیرد. این معیار به‌ویژه در مجموعه داده‌های نامتوازن کارایی بالاتری نسبت به دقت دارد و مقدار نزدیک به ۱ نشان‌دهنده عملکرد بهینه مدل است. زمان اجرا مدت‌زمان موردنیاز برای اجرای مدل را نشان می‌دهد و نقش مهمی در انتخاب مدل برای کاربردهای عملی

[1] Accuracy
[2] Precision
[3] Recall
[4] F1-score
[5] Kappa
[6] MCC
[7] TT

دارد. مدل‌هایی با زمان اجرای کوتاه‌تر، در سیستم‌هایی که نیاز به پردازش سریع دارند، ارجحیت بیشتری خواهند داشت. این معیارها به ارزیابی عملکرد مدل‌ها به‌طور جامع و تصمیم‌گیری‌های بهتر برای بهبود مدل‌های یادگیری ماشین کمک می‌کنند.

جدول ۳. عملکرد مدل‌های یادگیری ماشین در پیش‌بینی دسته‌ها

	Model	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC	TT (Sec)
knn	K Neighbors Classifier	0.9555	0.9749	0.9555	0.9569	0.9554	0.9106	0.9122	0.2720
xgboost	Extreme Gradient Boosting	0.9515	0.9753	0.9515	0.9538	0.9513	0.9025	0.9050	2.3200
lightgbm	Light Gradient Boosting Machine	0.9494	0.9708	0.9494	0.9516	0.9493	0.8985	0.9008	6.3420
ada	Ada Boost Classifier	0.9434	0.9730	0.9434	0.9445	0.9433	0.8865	0.8876	2.9060
gbc	Gradient Boosting Classifier	0.9393	0.9739	0.9393	0.9407	0.9392	0.8783	0.8798	9.8160
dt	Decision Tree Classifier	0.9130	0.9131	0.9130	0.9138	0.9130	0.8258	0.8265	0.3140
svm	SVM - Linear Kernel	0.8606	0.9485	0.8606	0.8739	0.8566	0.7185	0.7321	0.1680
rf	Random Forest Classifier	0.8178	0.8934	0.8178	0.8229	0.8174	0.6359	0.6406	0.7680
lr	Logistic Regression	0.7794	0.8816	0.7794	0.7848	0.7782	0.5583	0.5636	1.4760
et	Extra Trees Classifier	0.6336	0.6860	0.6336	0.6396	0.6320	0.2700	0.2740	0.4000
ridge	Ridge Classifier	0.6175	0.6407	0.6175	0.6190	0.6161	0.2339	0.2354	0.1600
nb	Naive Bayes	0.5789	0.6312	0.5789	0.5795	0.5780	0.1561	0.1570	0.2860
lda	Linear Discriminant Analysis	0.5385	0.5343	0.5385	0.5411	0.5346	0.0783	0.0796	0.2580
qda	Quadratic Discriminant Analysis	0.5203	0.5810	0.5203	0.5239	0.4041	0.0091	0.0196	0.1980
dummy	Dummy Classifier	0.5182	0.5000	0.5182	0.2686	0.3538	0.0000	0.0000	0.1380

۳-۱-۱- شرح خروجی

جدول ۳، تحلیل عملکرد مدل‌های مختلف یادگیری ماشین بر اساس معیارهای دقت، AUC، بازخوانی، درستی، امتیاز-F1 ، ضریب کاپا، MCC و زمان آموزش[1] (TT) را ارائه می‌کند. مدل KNN با دقت ۰,۹۵۵۵ و AUC ۰,۹۷۴۹ عملکرد بسیار خوبی دارد و در زمان به نسبت کوتاهی آموزش دید (۰,۲۷۲۰ ثانیه). مدل تقویت گرادیان مضاعف[2] نیز عملکرد بسیار خوبی با دقت ۰,۹۵۱۵ و AUC ۰,۹۷۵۳ دارد اما زمان آموزش آن بیشتر است (۲,۳۲۰۰ ثانیه).

مدل تقویت گرادیان ضعیف[3] نیز با دقت ۰,۹۴۹۴ و AUC ۰,۹۷۰۸ عملکرد بالایی دارد اما زمان آموزشش زیاد است (۶,۳۴۲۰ ثانیه). در مقابل، مدل دسته‌بندی دامی[4] عملکرد بسیار ضعیفی با دقت ۰,۵۱۸۲ و AUC ۰,۵۰۰۰ دارد که نشان می‌دهد این مدل به هیچ وجه قابل اعتماد نیست.

برای ارزیابی مدل‌های یادگیری ماشین در این پژوهش، از تابع ارزیابی مدل[5] در کتابخانه PyCaret استفاده شد. این تابع با ارائه معیارهای جامع عملکرد مدل‌ها مانند دقت، سطح زیر منحنی، بازخوانی، درستی، F1-امتیاز، ضریب کاپا و MCC امکان بررسی و مقایسه عملکرد مدل‌های مختلف را با دقت بالا فراهم می‌کند. همچنین، این تابع نمودارهای ارزیابی متنوعی از جمله ماتریس سردرگمی، منحنی ROC، نمودار صحت - بازخوانی[6] و نمودارهای ارتقاء و بهره[7] را تولید می‌کند که به تحلیل بصری عملکرد مدل‌ها کمک شایانی می‌کند. با استفاده از این ابزار قدرتمند، توانستیم مدل‌های مختلف را به‌طور کامل ارزیابی کرده و بهترین مدل را برای تشخیص نظرات فریبنده انتخاب کنیم. این ارزیابی دقیق، امکان ارتقای مدل‌های یادگیری ماشین را فراهم می‌کند.

۳-۱-۲- هایپرپارامتر[8]های مدل

جدول ۴، مقادیر هایپرپارامترهای مدل را نمایش می‌دهد. هایپرپارامترها تنظیماتی هستند که قبل از شروع فرآیند آموزش مدل تعیین می‌شوند.

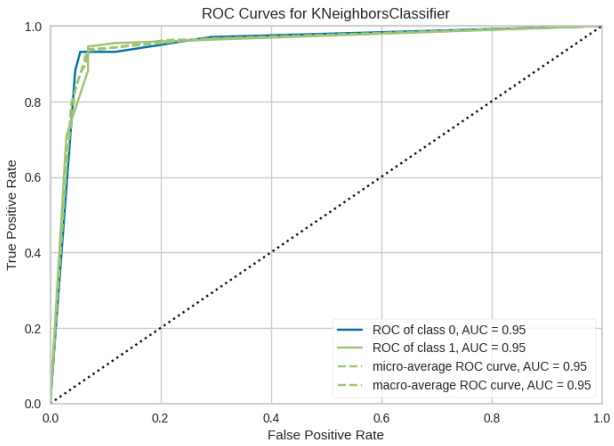
[1] Training time
[2] EXTRA boosting classifier
[3] Light Gradient Boosting Machine
[4] Dummy Classifier
[5] Evaluate model
[6] Precision-Recall
[7] Gain
[8] Hyperparameter

جدول ۴. هایپرپارامترهای مدل

پارامترها	مقدار/حالت
الگوریتم	خودکار[1]
اندازه لیف	۳۰
متریک	مینکوسکی[2]
پارامترهای متریک[3]	ندارد
تعداد شغل[4]	-۱
تعداد همسایگان	۵
p	۲
وزن	یکنواخت

۳-۱-۳- مساحت زیر منحنی[5] (AUC)

نمودار ROC [6] در شکل ۳، منحنی مشخصه عملکرد سیستم یا منحنی عملیاتی گیرنده و مساحت زیر منحنی AUC را نشان می‌دهد که برای سنجش عملکرد کلی مدل در تشخیص درست دسته‌ها است.



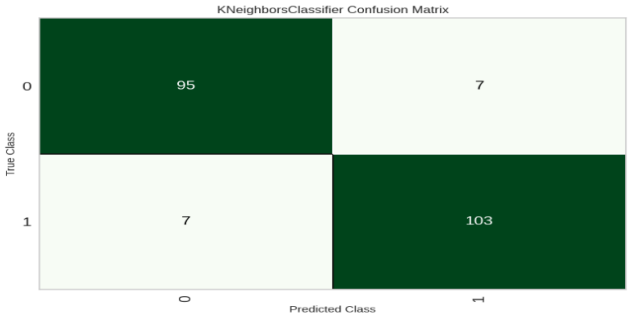
شکل ۳. نمودار ROC برای مدل KNN

نمودار ROC برای مدل KNN- دسته‌بندی نزدیک‌ترین K-همسایگان) را نشان می‌دهد. نمودار ROC ابزار مهمی برای ارزیابی عملکرد مدل‌های دسته‌بندی است. در این نمودار، مقدار AUC برای هر دو کلاس ۰ و ۱ برابر ۰,۹۵ است که نشان‌دهنده عملکرد بسیار خوب مدل در تفکیک کلاس‌ها است. این نمودار نشان می‌دهد که مدل KNN عملکرد بسیار خوبی در طبقه‌بندی داده‌ها دارد. مقدار AUC برابر ۰,۹۵ برای هر دو کلاس نشان می‌دهد که مدل قادر است با دقت بالایی نمونه‌های مثبت و منفی را تفکیک کند. این به معنای آن است که مدل نرخ پایینی از مثبت کاذب و منفی کاذب دارد و بنابراین می‌تواند به خوبی در کاربردهای واقعی استفاده شود.

۳-۱-۴- ماتریس درهم‌ریختگی

ماتریس درهم‌ریختگی مدلKNN ، عملکرد آن را در تفکیک نمونه‌های دو کلاس ۰ و ۱ نمایش می‌دهد. این ماتریس شامل چهار مقدار کلیدی است: مثبت صحیح (TP) که نشان می‌دهد مدل ۹۵ نمونه از کلاس ۰ را به‌درستی شناسایی کرده، منفی صحیح (TN) که نشان می‌دهد ۱۰۳ نمونه از کلاس ۱ به‌درستی پیش‌بینی شده‌اند، مثبت کاذب (FP) که ۷ نمونه از کلاس ۰ را به‌اشتباه کلاس ۱ در نظر گرفته، و منفی کاذب (FN) که ۷ نمونه از کلاس ۱ را نادرست به‌عنوان کلاس ۰ طبقه‌بندی کرده است.

ماتریس درهم‌ریختگی[7] در شکل ۴، تعداد پیش‌بینی‌های درست و نادرست مدل برای هر دسته را نشان می‌دهد.



شکل ۴. ماتریس درهم‌ریختگی مدل KNN

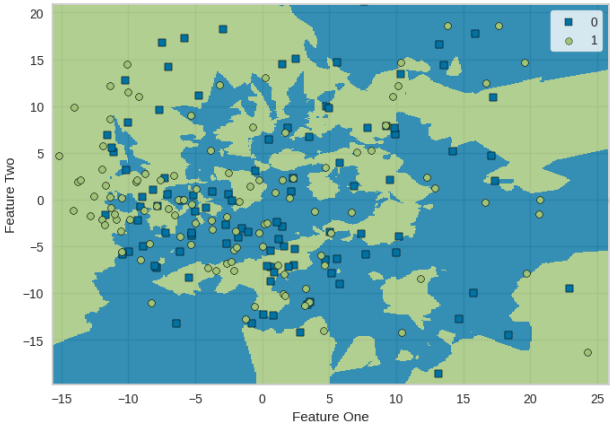
این نتایج، دقت بالای مدل در تشخیص صحیح کلاس‌ها را نشان می‌دهند. نرخ پایین نمونه‌های اشتباه طبقه‌بندی‌شده (FP و FN) نشان می‌دهد که مدل توانایی خوبی در تفکیک کلاس‌ها دارد. این عملکرد مناسب مدل در کاربردهای واقعی استفاده شده و به تصمیم‌گیری‌های دقیق‌تری منجر می‌شود.

۳-۱-۵- نمودار مرز تصمیم‌گیری

نمودار مرز تصمیم‌گیری در شکل ۵، مدل را در فضای ویژگی‌ها نمایش می‌دهد که نشان می‌دهد مدل چگونه داده‌ها را به دسته‌های

[1] Auto
[2] Minkowski
[3] Metric parameters
[4] n_jobs
[5] Area Under the Curve
[6] Receiver Operating Characteristic
[7] Confusion Matrix

مختلف تقسیم می‌کند.



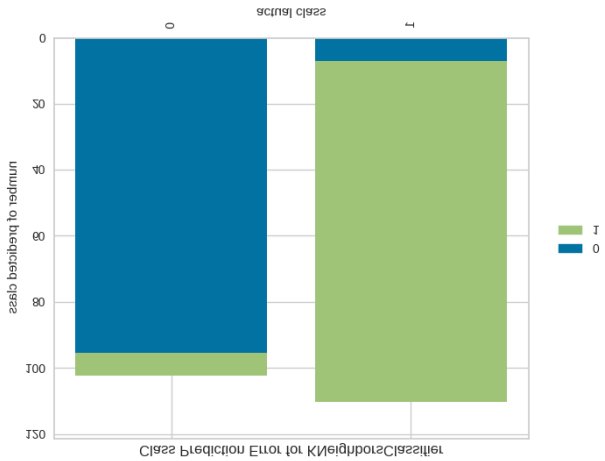
شکل ۵. نمودار مرز تصمیم‌گیری مدل KNN

نمودار مرز تصمیم‌گیری مدل KNN در شکل ۵، نشان می‌دهد که مدل چگونه فضای ویژگی‌ها را برای طبقه‌بندی داده‌ها به دو کلاس ۰ و ۱ تقسیم می‌کند. این مرزها نشان‌دهنده تصمیم‌گیری مدل بر اساس ویژگی‌های داده‌ها هستند. هرچند مدل عملکرد خوبی در طبقه‌بندی داده‌ها دارد اما برخی اشتباهات طبقه‌بندی نیز وجود دارد که در نقاط داده‌های اشتباه طبقه‌بندی شده مشاهده می‌شوند.

در تحلیل دیگر نمودارها نمودار Lift نشان می‌دهد که مدل KNN در پیش‌بینی دسته‌ها به‌طور کلی بهتر از یک مدل تصادفی عمل می‌کند. در ابتدا، مدل توانسته است با دقت بالایی دسته‌بندی‌ها را انجام دهد اما با افزایش تعداد نمونه‌های بررسی‌شده، عملکرد مدل به‌تدریج به مدل تصادفی نزدیک‌تر می‌شود. این نمودار به شناسایی قسمت‌های داده با عملکرد بهتر و قسمت‌های داده نیازمند بهبود کمک می‌کند. نمودار تجمعی بهره نشان می‌دهد که مدل KNN در تشخیص دسته‌ها بسیار بهتر از یک مدل تصادفی عمل می‌کند. منحنی‌های بهره نشان می‌دهند که این مدل تعداد زیادی از نمونه‌ها را به درستی و با سرعت دسته‌بندی می‌کند و با افزایش درصد نمونه‌ها، عملکرد مدل بهبود می‌یابد. نمودار آستانه نشان می‌دهد که چگونه عملکرد مدل KNN با تغییر آستانه تصمیم‌گیری تغییر می‌کند. در اینجا، خط نقطه‌چین عمودی ممکن است نشان‌دهنده آستانه پیشنهادی یا بهینه باشد که بر اساس معیارهای مختلف ارزیابی انتخاب شده است. این نمودار به شما کمک می‌کند تا عملکرد مدل را در سطوح مختلف آستانه درک کنید و تصمیم‌گیری‌های بهتری برای بهبود مدل خود بگیرید. در بررسی نمودار آستانه یافته‌ها بیانگر آستانه‌ای است که بیشترین مقدار آماره K-S ۰,۹۲۴ در آن نقطه مشاهده می‌شود. این نقطه به‌عنوان بهترین آستانه برای تفکیک بین دو کلاس در نظر گرفته می‌شود.

۳-۱-۶- نمودار پیش‌بینی خطا[1]

شکل ۶، پیش‌بینی خطای تفاوت بین مقادیر پیش‌بینی‌شده و مقادیر واقعی را نشان می‌دهد. همچنین، این نمودار نشان می‌دهد که مدل در کدام کلاس‌ها عملکرد بهتری داشته و در کجاها دچار خطا شده است.



شکل ۶. نمودار خطای پیش‌بینی مدل KNN

نمودار خطای پیش‌بینی نشان می‌دهد که مدل KNN عملکرد بسیار خوبی در طبقه‌بندی داده‌ها دارد اما همچنان تعدادی نمونه را به اشتباه طبقه‌بندی کرده است. به‌طور خاص:

این مدل توانسته است بیشتر نمونه‌های کلاس ۰ را به درستی طبقه‌بندی کند اما ۷ نمونه را به اشتباه به‌عنوان کلاس ۱ پیش‌بینی کرد. همچنین، این مدل بیشتر نمونه‌های کلاس ۱ را به درستی طبقه‌بندی کرده است اما ۷ نمونه را به اشتباه به‌عنوان کلاس ۰ پیش‌بینی کرد.

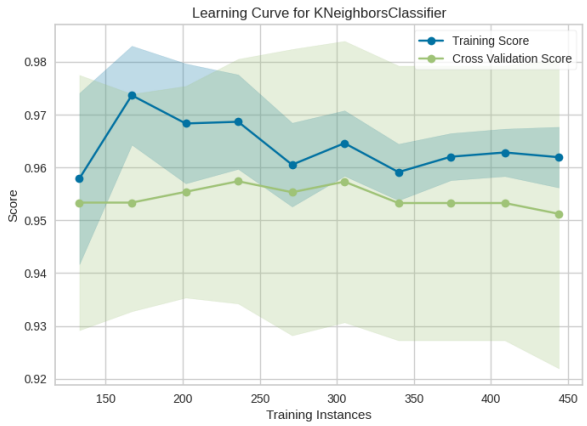
این نمودار نشان‌دهنده نقاط قوت و ضعف مدل در پیش‌بینی دسته‌های مختلف است و می‌تواند به بهبود مدل و کاهش خطاهای پیش‌بینی کمک کند. به‌طور کلی، مدل KNN عملکرد مناسبی داشته و توانسته است بیشتر نمونه‌ها را به درستی طبقه‌بندی کند اما همچنان جای بهبود دارد.

۳-۱-۷- منحنی یادگیری[2]

منحنی یادگیری نحوه تغییر عملکرد مدل با افزایش اندازه داده‌های آموزشی را نشان می‌دهد.

[1] Prediction Error Cruve

[2] Learning Curve

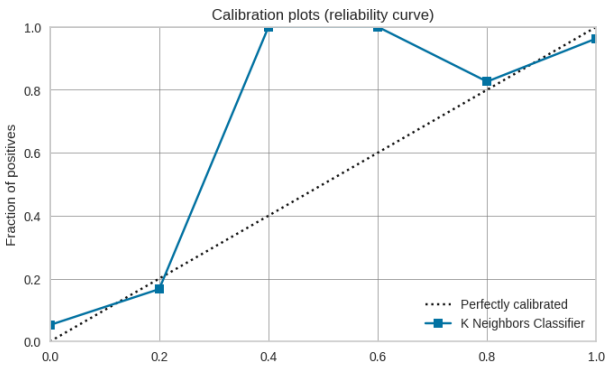


شکل ۷. منحنی یادگیری مدل KNN

نمودار منحنی یادگیری برای مدل KNN نشان می‌دهد که مدل با افزایش تعداد نمونه‌های آموزشی عملکرد پایداری دارد و دقت بالایی در هر دو مجموعه داده‌های آموزشی و اعتبارسنجی نشان می‌دهد. هرچند مقداری بیش‌برازش مشاهده می‌شود اما این بیش‌برازش به حدی نیست که عملکرد مدل را به‌طور جدی تحت تأثیر قرار دهد.

۳-۱-۸- منحنی کالیبرسیون[1]

این نمودار نشان می‌دهد که چگونه احتمالات پیش‌بینی‌شده بوسیله مدل با توزیع واقعی داده‌ها سازگار هستند.



شکل ۸. نمودار کالیبراسیون

خط به‌طور کامل کالیبره‌شده[2]: این خط که با نقطه‌چین سیاه نشان داده شد مدلی را نشان می‌دهد که به‌طور کامل کالیبره شد که در آن احتمالات پیش‌بینی‌شده به‌طور دقیق با فراوانی وقوع واقعی مطابقت دارند.

منحنی آبی (KNN): این منحنی نشان‌دهنده عملکرد مدل KNN است.

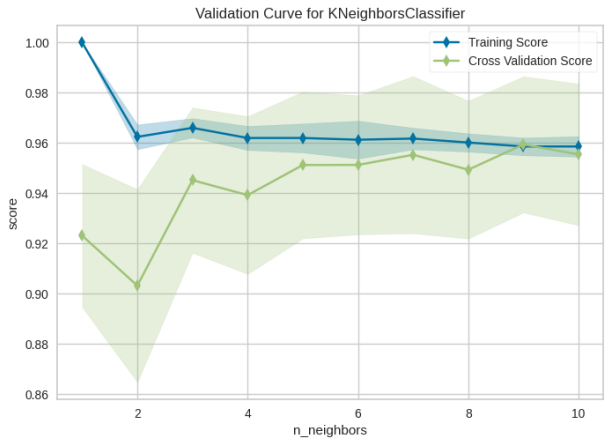
نمودار کالیبراسیون برای مدل KNN نشان می‌دهد که این مدل در برخی از بازه‌های احتمال پیش‌بینی‌شده به خوبی کالیبره شده و در برخی دیگر دچار اشکال است. به‌طور خاص:

در بازه‌های احتمال پایین (نزدیک به ۰,۲)، مدل به درستی کالیبره شده است.

در بازه‌های احتمال بالاتر (نزدیک به ۰,۴ و ۰,۶)، مدل به خوبی کالیبره نشده و تفاوت زیادی بین احتمالات پیش‌بینی‌شده و فراوانی وقوع واقعی وجود دارد.

۳-۱-۹- منحنی اعتبارسنجی[3]

منحنی اعتبارسنجی، عملکرد مدل را در مقابل مقادیر مختلف یک هایپرپارامتر خاص نشان می‌دهد.



شکل ۹. نمودار اعتبارسنجی مدل KNN

نمودار اعتبارسنجی برای مدل KNN نشان می‌دهد که این مدل در تعداد همسایگان دچار بیش‌برازش کمتری می‌شود زیرا دقت آموزشی بسیار بالا و دقت اعتبارسنجی پایین‌تر است. با افزایش تعداد همسایگان، دقت اعتبارسنجی بهبود می‌یابد و مدل تعمیم بهتری پیدا می‌کند. در این نمودار، مقدار بهینه هایپرپارامتر تعداد همسایگان حدود ۵ تا ۷ به نظر می‌رسد زیرا در این محدوده هر دو دقت آموزشی و اعتبارسنجی به سطح به نسبت ثابتی می‌رسند و فاصله بین آنها کمترین است.

۴- نتیجه‌گیری

در این پایان‌نامه، نتایج حاصل از تحلیل داده‌ها و ارزیابی مدل‌های یادگیری ماشین برای مقایسه عملکرد الگوریتم‌های KNN و XGBoost و دیگر مدل‌های رایج یادگیری ماشین استفاده از مدل‌های زبانی بزرگ (LLM) و به ویژه مدل ParsBERT برای

[1] Calibration Curve
[2] Perfectly calibrated
[3] Validation Curve

تحلیل احساسات نظرات کاربران و شناسایی بهترین مدل در تشخیص نظرات فریبنده در فروشگاه‌های برخط منابع باز و وب‌سرویس‌های دیجی‌کالا بررسی شد. داده‌های مورد استفاده در این تحقیق شامل نظرات کاربران با ویژگی‌های مختلفی از جمله متن نظر، نام کاربر، تعداد لایک و دیس‌لایک، امتیاز کاربر، تاریخ ارسال نظر و غیره در فروشگاه‌های برخط است. بر اساس توزیع فراوانی نظرات مبتنی بر احساسات پیام، نظرات مثبت بیشترین تعداد نظرات (حدود ۷۰،۰۰۰ نظر) بود که نشان‌دهنده رضایت کلی کاربران از محصولات و خدمات ارائه شده است. در مقابل، نظرات خنثی با حدود ۲۰،۰۰۰ نظر بیانگر عدم تمایل کاربران به ابراز نظر خاصی درباره محصولات یا خدمات است. نظرات منفی با حدود ۱۵،۰۰۰ نظر کمترین تعداد را به خود اختصاص دادند که نارضایتی کاربران از برخی جنبه‌های محصولات یا خدمات را نشان می‌دهد. وتوزیع فراوانی سالانه نظرات نشان داد نظرات مثبت به‌طور کلی روند افزایشی داشته و بیشترین تعداد نظرات مثبت در سال‌های ۲۰۱۹ و ۲۰۲۰ بود. نظرات خنثی تا حدودی ثابت بوده و تغییرات کمتری نسبت به نظرات مثبت داشتند. نظرات منفی با وجود افزایش جزئی در برخی سال‌ها مانند ۲۰۲۰، تغییرات زیادی نداشته و تا حدودی ثابت باقی ماندند. توزیع فراوانی نظرات بر اساس اعتبار کاربر نشان داد که بیشتر نظرات از سوی کاربران با اعتبار پایین ارسال شدند که می‌تواند به دلیل تعدد کاربران جدید یا غیرفعال با اعتبار پایین باشد. بر اساس توزیع فراوانی لایک‌ها و دیس‌لایک‌ها تعداد لایک‌ها به‌طور کلی بیشتر از تعداد دیس‌لایک‌ها بود و به‌طور کلی روند افزایش نوسانی را نشان می‌داد. نمودار ارتباط بین لایک‌های نظرات و اعتبار کاربر نشان داد که تعداد لایک‌های نظرات تاثیر مستقیمی بر امتیاز اعتبار کاربران داشت. رابطه مثبت بین تعداد لایک‌ها و اعتبار کاربران نشان‌دهنده نظرات مورد تایید بیشتر کاربران (با تعداد لایک بالا) بوسیله کاربران با اعتبار بالاتر بود.

داده‌های عملکرد مدل‌های یادگیری ماشین در پیش‌بینی دسته‌ها نشان داد که مدل KNN با دقت ۰,۹۵۵۵ و AUC ۰,۹۷۴۹ یکی از بهترین مدل‌ها است. مدل Extreme Gradient Boosting نیز با دقت ۰,۹۵۱۵ و AUC ۰,۹۷۵۳ عملکرد بسیار خوبی از خود نشان داد. مدل‌های LightGBM و Ada Boost نیز به ترتیب با دقت‌های ۰,۹۴۹۴ و ۰,۹۴۳۴ و AUC های ۰,۹۷۰۸ و ۰,۹۷۳۰ در رتبه‌های بعدی قرار گرفتند. مدل‌های دیگر نیز بسته به معیارهای مختلف ارزیابی، عملکرد متفاوتی داشتند. به‌طور کلی، مدل‌های Boosting و KNN در این پژوهش عملکرد بهتری نسبت به سایر مدل‌ها داشتند.

طبق نمودار macro-average ROC مدل KNN عملکرد بسیار خوبی در طبقه‌بندی داده‌ها نشان داد. مقدار AUC برابر ۰,۹۵ برای هر دو کلاس نشان می‌دهد که مدل قادر است با دقت بالایی نمونه‌های مثبت و منفی را تفکیک کند و نرخ پایینی از مثبت کاذب و منفی کاذب دارد و بنابراین می‌تواند به خوبی در کاربردهای واقعی استفاده شود. نمودار ماتریس درهم‌ریختگی نیز دقت بالای مدل در تشخیص صحیح کلاس‌ها ۰,۹۳۱٪ نمونه‌های کلاس ۰ و ۰,۹۳۶٪ نمونه‌های کلاس ۱ را نشان داد که تایید بر عملکرد به نسبت خوب مدل KNN در طبقه‌بندی نمونه‌ها بود. نمودار مرز تصمیم‌گیری مدل KNN نشان‌دهنده عملکرد خوب مدل در طبقه‌بندی داده‌ها بود. منحنی‌های بهره مدل KNN به‌طور قابلتوجهی از خط Baseline قرار داشتند که نشان‌دهنده عملکرد بهتر مدل نسبت به مدل تصادفی بود.

نمودار تجمعی بهره نشان داد که مدل KNN در تشخیص دسته‌ها بسیار بهتر از یک مدل تصادفی عمل می‌کند. و با سرعت تعداد زیادی از نمونه‌ها را به درستی دسته‌بندی می‌کند و با افزایش درصد نمونه‌ها، عملکرد مدل بهبود می‌یابد. در سطوح آستانه پایین، دقت مدل بالا بود. با افزایش آستانه، دقت مدل کاهش یافت. در سطوح آستانه پایین، بازخوانی پایین بود زیرا مدل تعداد زیادی از نمونه‌های مثبت واقعی را به‌عنوان منفی طبقه‌بندی می‌کند. با افزایش آستانه، مدل نمونه‌های مثبت بیشتری را به درستی شناسایی می‌کند و بازخوانی افزایش یافت.

نمودار Precision-Recall برای مدل KNN با میانگین دقت برابر ۰,۹۴ حاکی از عملکرد بالای مدل بود.

نمودار خطای پیش‌بینی نیز نشان داد که با وجود طبقه‌بندی اشتباه تعدادی نمونه مدل KNN عملکرد بسیار خوبی در طبقه‌بندی داده‌ها دارد. دقت و بازخوانی بالای مدل (برای نمونه‌های کلاس ۱ برابر ۰,۹۳۶ و برابر ۰,۹۳۱ برای نمونه‌های کلاس ۰) نشان داد مدل KNN در طبقه‌بندی داده‌ها بسیار موثر بوده و توانسته است به خوبی دسته‌بندی‌های صحیح را انجام دهد. دقت مدل در داده‌های آموزشی با افزایش تعداد نمونه‌های آموزشی کمی نوسان دارد اما به‌طور کلی در سطح بالایی (حدود ۰,۹۶ تا ۰,۹۸) باقی می‌ماند که حاکی از عملکرد بسیار خوب مدل در داده‌های آموزشی است و می‌تواند بیشتر نمونه‌های آموزشی را به درستی طبقه‌بندی کند. دقت مدل در داده‌های اعتبارسنجی نیز با افزایش تعداد نمونه‌های آموزشی کمی نوسان دارد اما در سطح پایین‌تری نسبت به دقت آموزشی (حدود ۰,۹۴ تا ۰,۹۶) باقی می‌ماند. در این نمودار، دو منحنی (آموزشی و اعتبارسنجی) به هم نزدیک هستند که بیش‌برازش اندک

مدل را نشان می‌دهد.

نمودار منحنی یادگیری نشان داد با افزایش تعداد نمونه‌های آموزشی مدل KNN عملکرد پایدار و دقت بالا در هر دو مجموعه داده‌های آموزشی و اعتبارسنجی دارد. خوشه‌های مجزا برای کلاس‌های ۰ و ۱ نشان داد که مدل KNN داده‌ها را به خوبی بر اساس ویژگی‌هایشان تفکیک کرد.

زمان اجرای t-SNE نشان‌دهنده سرعت بالای پردازش الگوریتم بود.

نمودار کالیبراسیون برای مدل KNN نشان می‌دهد که این مدل در برخی بازه‌های احتمال پیش‌بینی‌شده به خوبی کالیبره شده و در برخی دیگر دچار اشکال است. در بازه‌های احتمال پایین (نزدیک به ۰,۲)، مدل به درستی کالیبره‌شده و در بازه‌های احتمال بالاتر (نزدیک به ۰,۴ و ۰,۶)، مدل به خوبی کالیبره‌نشده نشان از تفاوت زیاد بین احتمالات پیش‌بینی‌شده و فراوانی وقوع واقعی بود. دقت بسیار بالای مدل در داده‌های آموزشی با کاهش تعداد همسایگان (حدود ۱) و کاهش آن با افزایش تعداد همسایگان (حدود ۰,۹۶) نشان داد مدل تمایل به بیش‌برازش دارد زیرا تمامی نقاط آموزشی را به درستی دسته‌بندی می‌کند. با افزایش تعداد همسایگان، دقت اعتبارسنجی مدل از حدود ۰,۹ به حدود ۰,۹۶ بهبود می‌یابد.

نمودار اعتبارسنجی برای مدل KNN نشان می‌دهد که این مدل در تعداد همسایگان[1] کمتر به دلیل دقت آموزشی بسیار بالا و دقت اعتبارسنجی پایین‌تر دچار بیش‌برازش می‌شود. بر اساس نتایج ارزیابی‌ها، مدل KNN با دقت ۰,۹۵۵۵ و AUC ۰,۹۷۴۹ به‌عنوان یکی از بهترین مدل‌ها شناخته شد. مدل Extreme Gradient Boosting نیز با دقت ۰,۹۵۱۵ و AUC ۰,۹۷۵۳ عملکرد بسیار خوبی از خود نشان داد. مدل‌های LightGBM و Ada Boost نیز به ترتیب با دقت‌های ۰,۹۴۹۴ و ۰,۹۴۳۴ و AUCهای ۰,۹۷۰۸ و ۰,۹۷۳۰ در رتبه‌های بعدی قرار گرفتند. مدل‌های دیگر نیز بسته به معیارهای مختلف ارزیابی، عملکرد متفاوتی داشتند. به‌طور کلی، مدل‌های تقویتی[2] و همسایگی (KNN) در این پژوهش عملکرد بهتری نسبت به مدل‌های دیگر داشتند. در پایان بر اساس یافته‌های بدست‌آمده برای شناسایی نظرات جعلی و فریبنده به کمک هوش مصنوعی پیشنهاد می‌شود سیستم توصیه‌گر محصول بوسیله دستیار هوشمند خرید مبتنی بر نیاز کاربر بوسیله مدل زبان بزرگ ایجاد شود. همچنین استفاده از سیستم هوشمند پرسش و پاسخ در زمینه محصولات مبتنی بر هوش مصنوعی و سیستم خلاصه‌ساز متن نظرات یک محصول بوسیله هوش مصنوعی و سیستم استخراج محورهای اصلی سوالات کاربران حول محصول بوسیله هوش مصنوعی از دیگر موارد پیشنهادشده در این بخش است.

مراجع

[1] Anderson ET, Simester DI (2014) Reviews without a purchase: Low ratings, loyal customers, and deception. J. Mark. Res. 51:249–269. https://doi.org/10.1509/jmr.13.0209

[2] Cao H (2020) Online review manipulation by asymmetrical firms: Is a firm's manipulation of online reviews always detrimental to its competitor? Inf. Manag. 57:103244. https://doi.org/10.1016/j.im.2019.103244

[3] Chatterjee S, Goyal D, Prakash A, Sharma J (2021) Exploring healthcare/health-product ecommerce satisfaction: A text mining and machine learning application. J. Bus. Res. 131:815–825. https://doi.org/10.1016/j.jbusres.2020.10.043.

[4] Chen J, Yang Y, Liu H (2021) Mining bilateral reviews for online transaction prediction: A relational topic modeling approach. Inf. Syst. Res. 32:541–560. https://doi.org/10.1287/ISRE.2020.0981.

[5] Dong X, Lian Y (2021) A review of social media-based public opinion analyses: Challenges and recommendations. Technol. Soc. 67:101724. https://doi.org/10.1016/j.techsoc.2021.101724.

[6] Ho SM, Hancock JT (2019) Context in a bottle: Language-action cues in spontaneous computer-mediated deception. Comput. Human Behav. 91:33–41. https://doi.org/10.1016/j.chb.2018.09.008.

[7] Huang G, Liang H (2021) Uncovering the effects of textual features on trustworthiness of online consumer reviews: A computational-experimental approach. J. Bus. Res. 126:1–11. https://doi.org/10.1016/j.jbusres.2020.12.052.

[8] Hussain S, Guangju W, Jafar RMS, Ilyas Z, Mustafa G, Jianzhou Y (2018) Consumers' online information adoption behavior: Motives and antecedents of electronic word of mouth communications. Comput. Human Behav. 80:22–32. https://doi.org/10.1016/j.chb.2017.09.019

[9] Ke Z, Liu D, Brass DJ (2020) Do online friends bring out the best in us? The effect of friend contributions on online review provision. Inf. Syst. Res. 31:1322–1336. https://doi.org/10.1287/isre.2020.0947

[10] Kumar N, Venugopal D, Qiu L, Kumar S (2019) Detecting anomalous online reviewers: An unsupervised approach using mixture models. J. Manag. Inf. Syst. 36:1313–1346. https://doi.org/10.1080/07421222.2019.1661089

[11] Kumar N, Venugopal D, Qiu L, Kumar S (2018) Detecting review manipulation on online platforms with hierarchical supervised learning. J. Manag. Inf. Syst. 35:350–380. https://doi.org/10.1080/07421222.2018.1440758.

[12] Lamb Y, Cai W, McKenna B (2020) Exploring the complexity of the individualistic culture through social exchange in online reviews. Int. J. Inf. Manage. 54:102198. https://doi.org/10.1016/j.ijinfomgt.2020.102198.

[13] Lee SY, Qiu L, Whinston A (2018) Sentiment manipulation in online platforms: An analysis of movie tweets. Prod. Oper. Manag. 27:393–416. https://doi.org/10.1111/poms.12805.

[1] n_neighbors

[2] Boosting

[14] [14] Lei Z, Yin D, Zhang H (2021) Focus within or on others: The impact of reviewers' attentional focus on review helpfulness. Inf. Syst. Res. 32:801–809. https://doi.org/10.1287/ISRE.2021.1007.

[15] Li L, Lee KY, Lee M, Yang SB (2020) Unveiling the cloak of deviance: Linguistic cues for psychological processes in fake online reviews. Int. J. Hosp. Manag. 87:102468. https://doi.org/10.1016/j.ijhm.2020.102468

[16] Li L, Yang L, Zhao M, Liao M, Cao Y (2022) Exploring the success determinants of crowdfunding for cultural and creative projects: An empirical study based on signal theory. Technol. Soc. 70:102036. https://doi.org/10.1016/j.techsoc.2022.102036

[17] Pang H, Liu J, Lu J (2022) Tackling fake news in socially mediated public spheres: A comparison of Weibo and WeChat. Technol. Soc. 70:102004. https://doi.org/10.1016/j.techsoc.2022.102004

[18] Qiao D, Lee SY, Whinston AB, Wei Q (2020) Financial incentives dampen altruism in online prosocial contributions: A study of online reviews. Inf. Syst. Res. 31:1361–1375. https://doi.org/10.1287/isre.2020.0949

[19] Sporer SL (1997) The less travelled road to truth: verbal cues in deception detection in accounts of fabricated and self-experienced events. Appl. Cogn. Psychol. 11:373–397. https://doi.org/10.1002/(SICI)1099-0720(199710)11:53.0.CO;2-0

[20] Wang H, Du R, Shen W, Qiu L, Fan W (2022a) Product reviews: a benefit, a burden, or a trifle? How seller reputation affects the role of product reviews. MIS Q. Manag. Inf. Syst 46:1243–1272. https://doi.org/10.25300/misq/2022/15660

[21] Wang Q, Zhang W, Li J, Mai F, Ma Z (2022b) Effect of online review sentiment on product sales: The moderating role of review credibility perception. Comput. Human Behav. 133:107272. https://doi.org/10.1016/j.chb.2022.107272

[22] Wang Y, Goes P, Wei Z, Zeng D (2019) Production of online word-of-mouth: Peer effects and the moderation of user characteristics. Prod. Oper. Manag. 28:1621–1640. https://doi.org/10.1111/poms.13007

[23] Wu C, Mai F, Li X (2021) The effect of content depth and deviation on online review helpfulness: Evidence from double-hurdle model. Inf. Manag. 58:103408. https://doi.org/10.1016/j.im.2020.103408

[24] Wu PF (2019) Motivation crowding in online product reviewing: A qualitative study of amazon reviewers. Inf. Manag. 56:103163. https://doi.org/10.1016/j.im.2019.04.006

[25] Xu W, Zhang C(2018) Sentiment, richness, authority, and relevance model of information sharing during social Crises—the case of #MH370 tweets Comput. Human Behav. 89:199–206. https://doi.org/10.1016/j.chb.2018.07.041

[26] Yang, Y, Zhang, K, 2022. sDTM: A Supervised Bayesian Deep Topic Model for Text Analytics. Inf. Syst. Res. https://doi.org/10.1287/isre.2022.1124

[27] Yu Y, Khern-am-nuai W, Pinsonneault A (2022) When paying for reviews pays off: The case of performance-contingent monetary incentives. MIS Q. Manag. Inf. Syst 46:609–626. https://doi.org/10.25300/MISQ/2022/15488

[28] Zhang D, Zhou L, Kehoe JL, Kilic IY (2016) What online reviewer behaviors really matter? Effects of verbal and nonverbal behaviors on detection of fake online reviews. J. Manag. Inf. Syst. 33:456–481. https://doi.org/10.1080/07421222.2016.1205907

[29] Zhang W, Du Y, Yoshida T, Wang Q (2018) DRI-RCNN: An approach to deceptive review identification using recurrent convolutional neural network. Inf. Process. Manag. 54:576–592. https://doi.org/10.1016/j.ipm.2018.03.007

[30] Zhang W, Xie R, Wang Q, Yang Y, Li J (2022) A novel approach for fraudulent reviewer detection based on weighted topic modelling and nearest neighbors with asymmetric Kullback–Leibler divergence. Decis. Support Syst. 157:113765. https://doi.org/10.1016/j.dss.2022.113765

[31] Zhou Y, Lv X, Wang L, Li J, Gao X (2023) What increases the risk of gamers being addicted? An integrated network model of personality–emotion–motivation of gaming disorder. Comput. Human Behav. 141:107647. https://doi.org/10.1016/j.chb.2022.107647

[32] Zhuang M, Cui G, Peng L (2018) Manufactured opinions: The effect of manipulating online product reviews. J. Bus. Res. 87:24–35. https://doi.org/10.1016/j.jbusres.2018.02.016

[33] Lakshami, S., Muthumani, S.(2017). Building Brand Power. IOP Conference Series: Materials Science and Engineering. 197,https://iopscience.iop.org/article/10.1088/1757-899X/197/1/01