



Sentiment Analysis and Topic Modeling of Tweets about Iranian Airlines: Employing Machine Learning and Language Models

Ghazaleh Soltani 

MSc Student of Information Technology Management, Department of Management, Faculty of Social Sciences and Economics Alzahra University, Tehran, Iran

Ameneh Khadivar* 

Associate Professor, Department of management, Faculty of Social Sciences and Economics Alzahra University, Tehran, Iran

Neda Paravian 

PhD Student in Information Technology Management, Department of Management, Faculty of Social Sciences and Economics, Alzahra University, Tehran, Iran

Abstract

Introduction

With the expansion of the internet and the rise of social media, companies across various industries, particularly the aviation sector, have gained unprecedented opportunities for direct interaction with customers. Social media platforms, especially microblogging sites such as Twitter, enable companies to receive immediate user feedback and to monitor passengers' sentiments, attitudes, and concerns

* Corresponding Author: a.khadivar@alzahra.ac.ir

How to Cite: Soltani, G. , khadivar, A. and paravian, N. (2026). Sentiment Analysis and Topic Modeling of Tweets about Iranian Airlines: Employing Machine Learning and Language Models. *Tourism Management Studies*, 21(73), 449 - 478. doi: [10.22054/tms.2025.87482.3099](https://doi.org/10.22054/tms.2025.87482.3099)

closely. Iranian airlines, given market developments and increasing competition, require modern tools to analyze customer data in order to optimize their services and communication strategies. Sentiment analysis of users on social media, particularly on platforms like Twitter, is both a major challenge and a valuable opportunity due to the large volume of data, diversity of opinions, and extensive use of nonverbal cues such as emojis. This study aims to analyze sentiment and identify prominent user topics about Iranian airlines on Twitter, providing valuable insights to improve service quality and support managerial decision-making.

Methodology

For data collection, 2,876 English tweets published between April 2008 and March 2024 were extracted using the Playwright tool. The dataset included short texts, hashtags, mentions, and emojis. Data preprocessing involved noise removal, elimination of irrelevant characters, emoji standardization, and text labeling. In the emoji analysis stage, 230 emojis were interpreted using artificial intelligence to identify their meanings and emotional connotations. This stage was especially important given the significant role of nonverbal cues in expressing user sentiments. Sentiment labeling of the data was performed using two main approaches: a lexicon-based method and a machine-learning approach based on the RoBERTa model. In this process, five machine learning algorithms, being logistic regression, decision trees, SVMs, artificial neural networks, and transformer-based models, were evaluated. Model performance was assessed using standard scientific metrics, including accuracy, recall, and F1 score.

Results and Findings

The results demonstrated that the lexicon-based logistic regression model, achieving an F1 score of 0.81, outperformed all other models. Additionally, incorporating emoji data into the dataset significantly improved model accuracy, highlighting the importance of interpreting nonverbal signals and user interactions in sentiment analysis. Beyond

sentiment analysis, this study also conducted topic modeling to identify users' main concerns and topics. The Latent Dirichlet Allocation (LDA) algorithm was employed for this purpose. Topic modeling results revealed that flight safety was the primary concern of users and accounted for the largest share of discussions related to Iranian airlines. Other significant topics included service quality, flight health, passenger experience, and staff behavior. A more detailed analysis indicated that users' concerns regarding flight safety primarily encompassed technical issues, adherence to safety protocols, and emergency management. Service quality concerns included satisfaction with in-flight services, flight punctuality, and staff responsiveness, while flight health concerns focused on compliance with hygiene protocols and cabin cleanliness.

Conclusion

These findings suggest that Iranian airlines should prioritize improving flight safety, enhancing service quality, and implementing hygiene protocols to increase passenger satisfaction and strengthen their market position. Moreover, analyzing social data and user behavior enables companies to provide managers with rapid, accurate feedback, enabling strategic decisions to be made based on real user concerns. This approach can support marketing planning, service development, and crisis management within the aviation industry. Another innovative aspect of this study is the combination of sentiment analysis with topic modeling and the use of emoji data. This approach enables for a more precise identification of users' attitudes and sentiments toward airlines, underscoring the crucial role of nonverbal cues in understanding social data accurately. Furthermore, the utilization of machine learning algorithms, specifically advanced models like RoBERTa, proved effective in addressing the complexities of natural language and delivering reliable results.

This study has some limitations. The data used consists only of English tweets about Iranian airlines, which may not reflect the opinions of Persian-speaking users or those on other social media

platforms. Furthermore, the analysis was limited to textual data and emojis, excluding multimedia content such as images or videos. Nonetheless, these limitations also highlight opportunities for future research. For future studies, it is recommended to employ more advanced natural language processing models, such as BERT and other transformer-based models, to improve sentiment analysis accuracy. The application of data augmentation techniques and multimodal analysis could provide a more comprehensive understanding of user behavior. Additionally, examining data from other social media platforms and related domains, such as health, transportation, and public services, could help generalize the findings.

In conclusion, this study demonstrates that analyzing user sentiments and topics on social media, particularly Twitter, can help Iranian airlines improve service quality and communication strategies by better understanding their customers' concerns. The findings also emphasize the importance of incorporating emoji and nonverbal data into sentiment analysis, showing that combining lexicon-based and machine-learning approaches can yield optimal results. The practical applications of this study include designing service strategies, managing passenger experiences, and enhancing flight safety and quality. This approach can help managers make precise, data-driven decisions and improve passenger satisfaction.

Keywords: Sentiment analysis, topic modeling, Twitter, deep learning, lexicon-based approach



تحلیل احساسات و موضوع کاوی توییت های شرکت های هواپیمایی ایران (مبتهی بر یادگیری ماشین و مدل های زبانی)

کارشناسی ارشد مدیریت فناوری اطلاعات، دانشکده علوم اجتماعی و اقتصاد، دانشگاه الزهرا، تهران، ایران

غزاله سلطانی 

دانشیار گروه مدیریت، دانشکده علوم اجتماعی و اقتصاد، دانشگاه الزهرا، تهران، ایران

آمنه خدیور  *

دانشجوی دکتری مدیریت فناوری اطلاعات، دانشکده علوم اجتماعی و اقتصاد، دانشگاه الزهرا، تهران، ایران

ندا پروایان 

چکیده

کلیدواژه ها: تحلیل احساسات، موضوع کاوی، توییت، یادگیری عمیق، رویکرد مبتنی بر واژگان.

مقدمه

امروزه رسانه‌های اجتماعی به منبعی پویا برای بازنمایی تجربه و ادراک مشتریان تبدیل شده‌اند و «تویتر» به دلیل جریان مداوم پیام‌ها و امکان رهگیری لحظه‌ای رخدادها، بستری مناسب برای مطالعه نگرش کاربران نسبت به خدمات هوانوردی محسوب می‌شود (Breen, 2012). اتکا به نظرسنجی‌های سنتی، به سبب پوشش محدود و فاصله زمانی با رخدادها، کفایت ندارد؛ از این رو بهره‌گیری از پردازش زبان طبیعی و یادگیری از داده‌های متنی ضرورت می‌یابد (Haddi, 2013)؛ در دوره‌های بحرانی نیز تحلیل محتوایی که توسط کاربر ساخته می‌شود، می‌تواند کانون نگرانی‌ها را آشکار سازد؛ برای نمونه، تغییر دغدغه‌ها در دوران کووید-۱۹ در این پژوهش گزارش شده است (Field, 2022).

این پژوهش با تمرکز بر گفت‌وگوهای کاربران درباره شرکت‌های هواپیمایی ایران، به دو پرسش پاسخ می‌دهد: الگوی احساسات در یک بازه طولانی و در دوره‌های پیش/حین/پس از بحران کرونا چگونه تغییر کرده است؟ و موضوعات غالبی که این احساسات پیرامون آن‌ها شکل می‌گیرند کدام‌اند؟ برای پاسخ، ترکیب «تحلیل احساسات» و «موضوع کاوی» بر مجموعه‌ای از توییت‌های انگلیسی گردآوری شده به کار رفته است. نوآوری این پژوهش در سه محور است: ۱) مستندسازی کامل گردآوری و پالایش داده با تأکید بر یکتاسازی و ثبت فراداده؛ ۲) بهره‌گیری از نشانه‌های فرامتنی مانند شکلک‌ها^۱ برای حفظ لحن در متن‌های کوتاه و ۳) سنجش آماری اختلاف عملکرد مدل‌ها. بدین ترتیب، این کار خلأ مطالعات بومی و طولی درباره خطوط هوایی ایران را پوشش می‌دهد و تصویری روشن از پیوند احساس - موضوع ارائه می‌نماید.

پیشینه پژوهش

داده کاوی بر دو شیوه یادگیری نظارت‌شده و بدون نظارت تکیه دارد تا از داده‌های حجیم الگو بسازد. در متن کاوی نیز فرض اصلی این است که معنای واژه‌ها از بافت هم‌رخدادی‌شان به دست می‌آید؛ همین منطق، نمایش‌های برداری واژه و سپس، بهره‌گیری

1. Emoji

از مدل های زبانی بزرگ را ممکن کرده است. در موضوع کاوی، هر متن را آمیزه ای از چند موضوع می بینند و روش های تازه تر با تکیه بر نمایش های برداری همین ایده را دقیق تر پیش می برند. در نهایت، برای فهم رفتار کاربر باید احساس را به «موضوع مربوط» گره زد؛ یعنی نخست موضوع های غالب را استخراج کنیم و سپس احساس را بر فراز آنها بسنجیم تا تصویری روشن و کاربردی برای بهبود کیفیت به دست آید (Sharma & Shafiq, 2022).

مطالعات خارجی

متن کاوی با تکیه بر پردازش زبان طبیعی، امکان تحلیل معنا، لحن و احساس نهفته در گفتار کاربران را فراهم می کند (Khan, 2015). توییت به سبب دسترسی بلادرنگ، برجسب ها و دامنه وسیع کاربران، یکی از کانون های داده برای تحلیل نگرش های عمومی است. فرآیند متعارف در این حوزه شامل گردآوری، پاک سازی، پیش پردازش، مدل سازی و تفسیر نتایج است که در آن روش های طبقه بندی و خوشه بندی نقشی محوری دارند (Khan, 2015) و به تصمیم سازی های بازاریابی و مدیریتی یاری می رسانند (Saraswita, Rini, 2021). تحلیل احساسات عموماً در سه طیف مثبت، منفی و خنثی انجام می شود و می تواند تا سطوح جمله و واژه نیز ریزدانه شود (Angiani, 2016 ; Han, 2022). رویکردهای اصلی شامل روش های واژگانی، روش های یادگیری ماشین و رویکردهای ترکیبی است (Pandey, 2017). روش های واژگانی ساده و شفاف اند اما به بافت حساسیت دارند (Han, 2022 ; Alam, 2023) در سال های اخیر، یادگیری عمیق و مدل های زبانی بزرگ به دلیل درک بافت و ظرافت های زبانی، کارایی تحلیل را ارتقا داده اند (Monika, 2019). در صنعت هوانوردی، پژوهش های متعددی بر توییت ها تکیه کرده اند. بررسی احساسات گردشگران درباره خطوط هوایی آمریکا نشان می دهد کیفیت خدمات و تأخیر پرواز، مهم ترین عوامل اثرگذار بر نگرش هاست (Saraswita, Rini, 2021). کارهای مبتنی بر شبکه های عصبی بازگشتی و تبدیل گر ها نیز نسبت به روش های سنتی برتری دقت داشته اند (Monika, 2019). در دوران همه گیری کرونا، تمرکز نگرش ها از کیفیت خدمات به رعایت سلامت و ایمنی تغییر جهت داده است (Field,

(Song, 2020 ; 2022). همچنین، پژوهش‌های اخیر با تحلیل پست‌های شبکه‌های اجتماعی به افزایش وزن مؤلفه‌های ایمنی در گزینش خطوط هوایی اشاره کرده‌اند (Egger, Yu, 2022). مطالعات نشان می‌دهند تحلیل احساسات و موضوع کاوی می‌تواند ابزار سنجش کیفیت خدمات، ایمنی و واکنش به بحران‌ها باشد؛ با این حال، بیشتر مطالعات بر خطوط هوایی بین‌المللی متمرکز بوده و زمینه‌های بومی، از جمله ایران، کمتر واکاوی شده است. (Yu, 2019 ; Paradis, 2021)

مطالعات داخلی

در گردشگری ایران، چند مطالعه بومی با داده‌های واقعی کاربران نشان داده‌اند که می‌توان از متن کاوی و تحلیل احساسات برای استخراج شاخص‌های تجربه و رضایت بهره گرفت. مثلاً پژوهش «تحلیل تجربه مشتری با تکیه بر نظرات برخط در صنعت هتلداری ایران» با داده‌های هتل‌های پنج‌ستاره ایران در تریپ‌ادوایزر، با رویکرد شبکه معنایی و تحلیل عاملی، عوامل اثرگذار بر رضایت (رفتار کارکنان، امکانات اتاق، خدمات خوراکی در وعده‌های مختلف، زیبایی شهر و هدف سفر) را شناسایی و اثر هر کدام را مدل‌سازی می‌کند (Anisi Hemaseh, Afsharkazemi & Mousavi Khanghah, 2024).

در پژوهش‌های مشابه در رستوران‌های ایرانی، پژوهش تحلیل احساسات دیدگاه‌های تریپ‌ادوایزر با رویکرد یادگیری عمیق، نشان می‌دهد می‌توان از برجسب‌گذاری احساسی مبتنی بر جنبه‌ها (خوراک، سرویس، ارزش، فضا) برای استخراج اولویت‌های بهبود تجربه بهره برد؛ یافته‌های این پژوهش نیز تأکید دارد که خوراک و فضا بیشترین نقش را در احساسات گردشگران دارند (Mohseni, Khadivar & Abbasi, 2022). دوم، «به کارگیری یادگیری عمیق برای تحلیل احساس مبتنی بر جنبه در نقدهای رستوران‌های ایرانی» که با برجسب‌گذاری دستی و شبکه حافظه کوتاه و بلندمدت، کارایی مدل را برای جنبه‌های کلیدی گزارش می‌کند و پیشنهاد‌های سیاستی برای کسب و کارهای گردشگری ارائه می‌دهد (Khadivar, 2025). در پژوهشی دیگر نیز با عنوان «تحلیل محتوای نظرات گردشگران درباره رستوران‌های ایرانی» نشان داده شده است که می‌توان از تحلیل

محتوایی/موضوعی برای شناسایی مضامین غالب مرتبط با کیفیت خدمت و تصویر مقصد استفاده کرد (Poursaeid, 2022).

مطالعه ایرانی دیگر با عنوان «مدل چندزبانه تحلیل احساس در گردشگری»، به طور مستقیم مسئله چندزبانه بودن داده ها در تریپ ادوایزر را هدف گرفته و نشان می دهد طراحی مدل های عمیق چندزبانه می تواند دقت تحلیل احساسات مرتبط با گردشگری را بالا ببرد؛ این یافته برای مقاصدی مانند ایران که بازخوردها به فارسی، انگلیسی و دیگر زبان ها ثبت می شوند، بسیار کاربردی است (Shayegan & Dastan, 2024).

در پژوهشی دیگر با داده های همه هتل های تهران، بررسی شد که مدل های یادگیری نظارت شده می توانند رفتار امتیازدهی میهمانان را پیش بینی و تبیین کنند؛ نکته مهم این است که داده های کاربرمحور ایرانی قابلیت استخراج بینش مدیریتی در سطح مقصد را دارند و می توان آن را برای پایش تجربه گردشگر در سایر زیربخش ها تعمیم داد (Khorsand, Rafiee, & Kayvanfar, 2020). مطالعات اخیر نشان می دهد که چارچوب ترکیبی «تحلیل احساسات و موضوع کاوی» در بافت ایران به طور عملیاتی اثربخش بوده است. این رویکرد در سه سطح مختلف به کار گرفته شده است: در سطح مقصد (مانند تحلیل نظرات درباره هتل های تهران)، در سطح کسب و کار (مانند ارزیابی رستوران های ایرانی) و در سطح خدمت (مانند شناسایی عوامل مؤثر بر تجربه میهمان در هتل). همچنین، چالش چندزبانه بودن داده ها نیز با ارائه راهکارهای بومی مورد بررسی و حل و فصل قرار گرفته است.

با توجه به این پیشینه، تعمیم این چارچوب به حوزه «گردشگری حمل و نقل/هوانوردی» برای تحلیل توییت های مرتبط با خطوط هوایی ایران، اقدامی منطقی و هم راستا با بدنه پژوهشی داخلی موجود به شمار می آید و این خلأ را پوشش می دهد. این همخوانی زمانی آشکارتر می شود که درمی یابیم بسیاری از عوامل تشکیل دهنده تجربه مسافران (مانند تأخیر پرواز، کیفیت اطلاع رسانی، رفتار کارکنان و رعایت بهداشت) مشابه عواملی هستند که پیش تر در مطالعات مرتبط با هتل و رستوران شناسایی شده اند. از این رو، می توان این مؤلفه ها را با همان منطق تحلیل مبتنی بر جنبه ها، پایش و ارزیابی کرد. مرور پیشینه ها نشان

می‌دهد «تحلیل احساسات» و «موضوع کاوی» بر داده‌های توییت‌ها برای صنعت هوانوردی ابزار مؤثری در فهم تجربه و ادراک مشتری است و دقت آن‌ها به کیفیت پیش‌پردازش، شیوه برچسب‌گذاری و استفاده از نشانه‌های غیرزبانی مانند شکلک‌ها وابسته است. با این حال، بیشتر مطالعات بر خطوط هوایی غیرایرانی و یک زبان یا یک پلتفرم متمرکز هستند، مستندسازی فراداده و راهبردهای رفع تکرار را شفاف ارائه نمی‌کنند و به گزارش‌های صرفاً توصیفی بسنده کرده‌اند.

از نظر فنی نیز اتکای بیش‌ازحد به روش‌های سنتی موضوع کاوی در متن‌های کوتاه، کم‌توجهی به «تحلیل مبتنی بر جنبه» (ایمنی، اطلاع‌رسانی، خدمات کابین، زمان‌بندی)، ادغام محدود با داده‌های عملیاتی صنعت (تأخیر، لغو، شکایات)، بی‌توجهی نسبی به نشانه‌های فرامتنی (شکلک، هشتگ، الگوهای زمانی) و ضعف در تبیین‌پذیری مدل‌ها دیده می‌شود. این پژوهش برای پر کردن این شکاف‌ها، مجموعه داده‌های بومی و طولی با فرآیند گردآوری و پالایش مستند ارائه می‌کند، تحلیل جنبه‌محور را در کنار مدل‌های زبانی نوین و موضوع کاوی معنایی به کار می‌گیرد، آزمون‌های آماری معناداری را اجرا می‌کند و نتایج را به داده‌های عملیاتی و خروجی‌های قابل‌استفاده مدیریتی پیوند می‌دهد. نوآوری پژوهش حاضر در قیاس با مطالعات پیشین که روی شرکت‌های غیرایرانی صورت گرفته است، این است که پژوهش حاضر، نخستین مطالعه جامع بلندمدت بر توییت‌های مرتبط با شرکت‌های هواپیمایی ایرانی است که هم‌زمان از ترکیب برچسب‌گذاری مبتنی بر واژگان و مدل زبانی، به همراه مدل‌سازی موضوع و نگاشت شکلک‌ها بهره می‌گیرد.

روش^۱

داده‌های این پژوهش از توییت‌ها و بر پایه نام‌ها و هشتگ‌های مرتبط با شرکت‌های هواپیمایی ایران گردآوری شده است؛ جامعه شامل توییت‌های انگلیسی و نمونه ۲۸۷۷ پیام در بازه ۱ آوریل ۲۰۰۸ تا ۳۱ مارس ۲۰۲۴ است. متادیتای داده‌های مورد استفاده به شرح جدول ۱ می‌باشد.

جدول ۱. دامنه و منبع داده‌های پژوهش

مقدار	مؤلفه
توییت‌ها (پیام‌های عمومی)	بستر داده
توییت‌های انگلیسی درباره شرکت‌های هواپیمایی ایران	جامعه آماری
۲۸۷۷ پیام (پیام خام گردآوری شده پیش از پالایش)	اندازه نمونه استفاده شده
۱ آوریل ۲۰۰۸ تا ۳۱ مارس ۲۰۲۴	بازه زمانی پوشش
تبدیل همه زمان‌مهرها به زمان هماهنگ جهان ^۱	یکسان‌سازی زمان
تحلیل احساسات و موضوع کاوی درباره خطوط هوایی ایران	هدف پژوهش
کلیدواژه‌ها و هشتک‌های مرتبط با نام شرکت و واژگان پرواز/فرودگاه	منطق جست‌وجو
حذف ریتوییت‌های بدون متن افزوده	سیاست حذف اولیه
نگاشت شیء	قالب ذخیره‌سازی
نگاشت نام‌های مختلف یک شرکت به برچسب یکتای شرکت	هم‌معنایی نام‌ها

خزش داده‌ها با کتابخانه پلی‌رایت در پایتون انجام شد و فقط متن‌های اصلی و نقل‌قول‌های دارای متن افزوده نگه‌داری و ریتوییت‌های بدون متن حذف شدند. همه زمان‌مهرها یکسان‌سازی و داده‌ها در قالب نگاشت شیء ذخیره گردید. پس از ادغام نتایج جست‌وجو، یکتاسازی بر پایه شناسه توییت و حذف تکراری‌ها/نزدیک تکراری‌ها با امضای محتوایی انجام شد و حساب‌های مشکوک به ربات/اسپم پالایش گردید. پیش‌پردازش شامل حذف پیوندها و نام‌های کاربری، یک‌دست‌سازی نوشتار، توکن‌سازی، حذف ایست‌واژه‌ها، جداسازی هشتک‌ها و نگاشت نظام‌مند شکلک‌ها بود. برای هر پیام، شناسه، زمان انتشار زبان، وضعیت بازنشر، نقل‌قول، پاسخ، شمار تعامل‌ها، فهرست هشتک‌ها، اشاره‌ها، پیوندها، شکلک‌ها و برچسب یکتای شرکت، ثبت شد؛ برای رعایت حریم خصوصی، در صورت اشتراک‌گذاری فقط شناسه توییت‌ها ارائه می‌شود. پس از گردآوری، متن‌ها برای تحلیل دقیق‌تر پیش‌پردازش شدند: همه حروف یک‌دست، کوچک شد؛ پیوندها، نام‌های کاربری، علائم نگارشی و نویسه‌های غیرانگلیسی حذف گردید و ایست‌واژه‌ها کنار گذاشته شد. سپس متن به واژه‌ها/جمله‌ها شکسته، ریشه‌یابی انجام شد و شکلک‌ها و اختصارات به گونه‌ای مدیریت گردید که بار احساسی‌شان حفظ شود.

1. UTC

برای آماده‌سازی داده‌ها جهت تحلیل، از روش‌های مختلف استخراج ویژگی استفاده شد. برای مدل‌های یادگیری ماشین از تکرار واژه معکوس فراوانی سند^۱ و برای مدل حافظه بلندمدت دوطرفه^۲ از واحد شکن^۳ کتابخانه کراس^۴ بهره گرفته شد. این مرحله امکان تبدیل متن‌ها به داده‌های عددی را فراهم کرد تا الگوریتم‌ها بتوانند روابط معنایی میان واژگان را تحلیل کنند. تحلیل احساسات با دو رویکرد انجام شد: (۱) روش مبتنی بر واژگان که توپیت‌ها را به سه کلاس مثبت، منفی و خنثی طبقه‌بندی می‌کرد و (۲) روش مبتنی بر یادگیری ماشین و یادگیری عمیق که با استفاده از مدل زبانی بزرگ ربرتا^۵ برچسب‌گذاری انجام شد و سپس الگوریتم‌های ماشین بردار پشتیبان^۶، رگرسیون لجستیک^۷، درخت تصمیم^۸، بیز ساده و حافظه بلندمدت دوطرفه برای تحلیل احساسات به کار گرفته شدند. ساختار مدل حافظه بلندمدت دوطرفه شامل لایه‌های جاسازی، حافظه بلندمدت دوطرفه و لایه‌های متراکم بود.

مدل با تابع خطای میانگین مربع خطا^۹، بهینه‌ساز تخمین لحظه‌های سازگار^{۱۰} و معیار دقت آموزش داده شد. داده‌ها به نسبت ۷۰٪ آموزش و ۳۰٪ آزمون تقسیم شدند و از اعتبارسنجی متقابل ۱۰-فولد برای ارزیابی عملکرد مدل استفاده گردید. نتایج نشان دادند که برچسب‌گذاری با ربرتا دقت بالاتری نسبت به روش مبتنی بر واژگان ارائه می‌دهد. به‌منظور شناسایی موضوعات کلیدی در دیدگاه‌های کاربران، از مدل تخصیص نهفته دیریکله^{۱۱} استفاده شد که موضوعاتی مانند امنیت، کیفیت خدمات و سلامت را استخراج کرد و بر اساس فراوانی واژگان رتبه‌بندی نمود. عملکرد مدل‌ها نیز با معیارهای دقت^{۱۲}،

-
1. TF-IDF
 2. Bi-LSTM
 3. Tokenizer
 4. Keras
 5. Roberta
 6. SVM
 7. Logistic regression
 8. Decision Tree
 9. MSE
 10. ADAM (Adaptive Moment Estimation)
 11. LDA (Latent Dirichlet Allocation)
 12. Accuracy

صحت^۱، فراخوانی^۲ و میانگین موزون F1 ارزیابی شد و ماتریس آشفتگی برای بررسی صحت دسته بندی هر کلاس مورد استفاده قرار گرفت.

در این پژوهش برای سنجش و مقایسه عملکرد مدل ها از ارزیابی تکرار شونده به صورت اعتبارسنجی k-fold با لایه بندی بر حسب توزیع کلاس ها و همچنین، طرح ۲×۵ اعتبارسنجی متقاطع استفاده شد. معیار اصلی ارزیابی F1 بود و در کنار آن صحت، دقت و بازخوانی نیز گزارش شد. برای سنجش اختلاف کلی میان چند مدل، آزمون فریدمن بر رتبه های به دست آمده از تکرارها اجرا شد و مقایسه های دوبه دو با آزمون ویلکاکسون جفتی روی اختلاف امتیازهای هر فولد انجام گرفت؛ در سناریوی ۲×۵ نیز از آزمون t ویژه ۲×۵ CV بهره گرفته شد. روی یک مجموعه آزمون واحد، اختلاف دقت دو مدل با آزمون مک نمار ارزیابی شد. همه تحلیل ها با پایتون scikit-learn، SciPy/Statsmodels و با دانه تصادفی ثابت برای بازتولید پذیری انجام شد.

یافته ها

در این پژوهش، برای تحلیل احساسات توییت های مرتبط با شرکت های هواپیمایی ایرانی، از پنج مدل یادگیری ماشین و عمیق شامل بیز ساده، رگرسیون لجستیک، بردار پشتیبان، درخت تصمیم و حافظه بلندمدت دوطرفه استفاده شده است. توییت ها پس از پیش پردازش و تفسیر شکلک ها، با دو روش برچسب گذاری واژگانی و مدل زبانی ربرتا آماده سازی شدند. نتایج نشان دهنده تفاوت عملکرد مدل ها است. در این پژوهش، عملکرد مدل های یادگیری ماشین و یادگیری عمیق در تحلیل احساسات توییت ها با استفاده از چهار معیار ارزیابی شامل صحت، دقت، فراخوانی و امتیاز F1 مورد بررسی قرار گرفته و نتایج حاصل از دو روش برچسب گذاری مبتنی بر واژگان و مبتنی بر مدل زبانی بزرگ ربرتا با یکدیگر مقایسه شده اند. در معیار صحت، مدل رگرسیون لجستیک در روش واژگانی، بهترین عملکرد را با مقدار ۰/۸۳ داشته است، در حالی که در روش یادگیری، مدل های رگرسیون لجستیک و بردار پشتیبان هر دو با صحت ۰/۷۹ در رتبه نخست قرار گرفتند. از نظر دقت،

1. Precision

2. Recall

در روش واژگانی نیز رگرسیون لجستیک با مقدار $0/84$ برتر بود، اما در روش یادگیری، مدل حافظه بلندمدت دوطرفه دقت مشابهی برابر با $0/84$ ارائه داد. در ارزیابی فراخوانی، باز هم مدل رگرسیون لجستیک در روش واژگانی با مقدار $0/83$ بیشترین فراخوانی را داشت؛ در حالی که در روش یادگیری، رگرسیون لجستیک و بردار پشتیبان هر دو به مقدار $0/79$ دست یافتند. در نهایت، در معیار $F1$ ، رگرسیون لجستیک در روش واژگانی با امتیاز $0/82$ و در روش یادگیری با مدل بردار پشتیبان به صورت مشترک با مقدار $0/77$ بهترین عملکرد را ثبت کردند. این مقایسه نشان می‌دهد که در اکثر معیارها، مدل رگرسیون لجستیک عملکرد پایدارتری داشته و نتایج قابل اتکایی ارائه داده است. مدل رگرسیون لجستیک در هر دو روش برچسب‌گذاری، در بیشتر معیارها عملکرد برتری داشته است. همچنین، مدل حافظه بلندمدت دوطرفه در معیار دقت و مدل بردار پشتیبان در معیارهای یادگیری نیز نتایج رقابتی ارائه داده‌اند. این مقایسه نشان می‌دهد که انتخاب روش برچسب‌گذاری می‌تواند در نتایج مدل‌ها تأثیر قابل توجهی داشته باشد.

با توجه به مقایسه عملکرد مدل‌ها در معیارهای صحت، دقت، فراخوانی و $F1$ ، مدل رگرسیون لجستیک با برچسب‌گذاری مبتنی بر واژگان به عنوان بهترین مدل انتخاب شده است؛ چرا که در هر چهار معیار، بالاترین عملکرد را نسبت به سایر مدل‌های یادگیری ماشین و عمیق داشته است. در جدول ۲ می‌توانیم عملکرد مدل‌ها را از منظرهای متفاوت مقایسه نماییم.

در جدول ۳، نتایج به وضوح نشان می‌دهد که روش رگرسیون لجستیک به طور پایدار و قابل توجهی از روش ماشین بردار پشتیبان بهتر عمل کرده است. این برتری در تمامی معیارهای ارزیابی و در شرایط آزمایشی مختلف مشاهده شده است. برای سنجش دقیق عملکرد، از روش اعتبارسنجی متقاطع استفاده شد که در آن میانگین امتیاز $F1$ مدل رگرسیون لجستیک به طور معناداری بالاتر از مدل ماشین بردار پشتیبان بود. برای اطمینان از صحت این نتایج، آزمون‌های آماری پیشرفته‌ای شامل آزمون ویلکاکسون و آزمون t $5 \times 2CV$ انجام شد که همگی برتری مدل رگرسیون لجستیک را به طور قطعی تأیید کردند.

جدول ۲. مقایسه عملکرد مدل‌ها

مدل	متن + شکلک	دقت	صحت	فراخوانی	F1
رگرسیون لجستیک	NLP (Lexicon-based)	0/83	0/84	0/83	0/82
رگرسیون لجستیک	LLM (RoBERTa-based)	0/79	0/79	0/79	0/77
درخت تصمیم	NLP (Lexicon-based)	0/68	0/7	0/68	0/65
درخت تصمیم	LLM (RoBERTa)	0/73	0/72	0/73	0/71
بردار پشتیبان	NLP (Lexicon-based)	0/78	0/79	0/78	0/76
بردار پشتیبان	LLM (RoBERTa)	0/79	0/79	0/79	0/77
نایو بیز تکمیلی	NLP (Lexicon-based)	0/68	0/67	0/68	0/66
نایو بیز تکمیلی	LLM (RoBERTa)	0/75	0/74	0/75	0/74
حافظه بلندمدت دوطرفه	NLP (Lexicon-based)	0/6	0/73	0/6	0/65
حافظه بلندمدت دوطرفه	LLM RoBERTa	0/63	0/84	0/63	0/74

جدول ۲. خلاصه آزمون‌های آماری مقایسه عملکرد مدل‌ها

مدل/مقایسه	بخش/سناریو	P-Value از Holm	آزمون/آماره	خلاصه معیارها
رگرسیون لجستیک	اعتبارسنجی متقاطع (k-fold/5×2CV)	—	—	$F1 = 0.820 \pm 0.025$ ؛ دقت = 0.842 ± 0.022 ؛ بازخوانی = 0.833 ± 0.026 ؛ صحت = 0.834 ± 0.021
بردار پشتیبان	اعتبارسنجی متقاطع (k-fold/5×2CV)	0.018 / 0.022	ویلکاکسون ($\Delta F1$) / $5 \times 2CV$ t	$F1 = 0.776 \pm 0.028$ ؛ دقت = 0.792 ± 0.027 ؛ بازخوانی = 0.781 ± 0.029 ؛ صحت = 0.792 ± 0.024
رگرسیون لجستیک در برابر بردار پشتیبان	مجموعه آزمون واحد	0.0083	مک‌نمار (تصحیح پیوستگی)، $\chi^2 = 6.97$	ماتریس اختلاف: c = 35, b = 62
رگرسیون لجستیک در برابر بردار پشتیبان	مجموعه آزمون واحد	0.012	پرمیوتیشن جفتی (F1)	$CI_{95\%} \Delta F1 = 0.046$ بوت‌استرپ [0.017, 0.073]

این مزیت نه تنها در معیار $F1$ ، بلکه در سایر شاخص‌های مهم مانند دقت، بازخوانی و صحت نیز مشاهده شد. همچنین، دامنه اطمینان کوچک‌تر در نتایج رگرسیون لجستیک، نشان‌دهنده ثبات و قابلیت اطمینان بالاتر این روش است. در مرحله آزمایش نهایی روی داده‌های مستقل، تعداد مواردی که مدل رگرسیون لجستیک پاسخ درست و مدل ماشین بردار پشتیبان پاسخ نادرست داده بود، به‌طور معناداری بیشتر از حالت برعکس بود. آزمون مک‌نمار این برتری را از نظر آماری تأیید کرد. برای کسب اطمینان بیشتر، از روش‌های پیشرفته نمونه‌برداری مجدد مانند بوت‌استرپ و آزمون تعویضی استفاده شد که نتایج آن نیز به‌وضوح نشان داد اختلاف عملکرد بین دو مدل، واقعی و معنادار است و بازه اطمینان ۹۵ درصد کاملاً در محدوده مثبت قرار دارد. در مجموع، می‌توان با اطمینان نتیجه گرفت که رگرسیون لجستیک در این پژوهش، عملکردی برتر و پایدارتر از ماشین بردار پشتیبان داشته است.

از منظر اهمیت عملی، اندازه اثرها معنی‌دار و قابل توجه‌اند؛ d کوهن در حد متوسط رو به زیاد و دل‌تا کلیمف نیز در بازه‌ای قرار دارد که برتری LR را در بخش قابل توجهی از نمونه‌ها نشان می‌دهد. به بیان کاربردی، اگر هدف پایش احساسات و استخراج موضوع همراه با برچسب‌گذاری قطبیت باشد، انتخاب LR تصمیمی ایمن و توجیه‌پذیر است. با این وجود، پیشنهاد می‌شود برای تکمیل تحلیل، خطاهای هر دو مدل به تفکیک شرکت‌ها، موضوع‌ها و بازه‌های زمانی بررسی شود تا مشخص گردد اختلاف عملکرد در چه حوزه‌هایی بیشترین اثر را دارد و آیا ترکیب ویژگی‌های خاص (مثلاً نگاشت شکلک یا واژگان موضوعی) بیشترین سود را به LR رسانده است.

بر اساس برچسب‌گذاری با رگرسیون لجستیک، ۱۵ درصد توییت‌ها مثبت، ۲۵ درصد منفی و ۶۰ درصد خنثی بود؛ یعنی بیشتر پیام‌ها لحن خنثی دارند و نظرات مثبت و منفی پراکنده‌تر هستند. ابرواژگان کل مجموعه، بیشترین تکرار را برای کلماتی مانند ایران، پرواز، مسافرت، تحریم، نام فرودگاه‌ها و شرکت‌های هواپیمایی نشان می‌دهد. در ابرواژگان منفی، محور اصلی امنیت و رخداد‌های ناگوار است (سقوط، کشته شدن، تأخیر، تحریم، تصادف، تروریست). در مقابل، ابرواژگان مثبت پیرامون تجربه خوب سفر

می‌چرخد (تجربه، تشکر، عالی، بهترین) و به مقصدهایی مانند کیش و تهران و نیز نام‌هایی مثل ماهان و کشورهای همسایه اشاره دارد؛ نشانه‌ای از رضایت بخشی از مسافران و علاقه به سفرهای گردشگری.

در دوره پیش از کرونا (۲۰۱۵-۲۰۱۹) سهم احساسات مثبت حدود ۴۹.۵۱٪ و منفی ۵۱.۴۸٪ بود. طی سال‌های بحران (۲۰۲۰-۲۰۲۲) سهم مثبت‌ها به حدود ۵۷.۵۳٪ و منفی‌ها به ۴۳.۴۶٪ رسید. پس از بحران (۲۰۲۳-۲۰۲۴) روند بهبود ادامه یافت و مثبت‌ها تا حدود ۸۲.۷۳٪ و منفی‌ها به ۱۷.۲۶٪ کاهش یافت. این الگو نشان می‌دهد برداشت کاربران پس از بحران به‌طور محسوس بهتر شده است.

جدول ۴. درصد نظرات مثبت و منفی کاربران توییت‌ها نسبت به شرکت‌های هواپیمایی ایرانی

نام شرکت هواپیمایی	نظرات مثبت	نظرات منفی
معراج ایر	100%	0%
کیش ایر	66%	33%
ماهان ایر	61%	38%
ایران ایر	57%	42%
زاگرس ایر	57%	42%
ایران ایر تور	50%	50%
تابان ایر	41%	58%
آتا ایر	63%	36%
کاسپین ایر	71%	28%
ایران آسمان ایر	36%	64%
کارون ایر	0%	100%

نتایج به‌دست آمده از جدول ۴ به این ترتیب است که معراج ایر بالاترین سهم نظرات مثبت را دارد و کارون ایر بیشترین سهم نظرات منفی را ثبت کرده است. حدود یک‌سوم کل توییت‌ها درباره ایران‌ایر است که ۶۰ درصدشان مثبت گزارش شده و این هم نشان‌دهنده رضایت نسبی مشتریان و همچنین، سهم بالای این شرکت در گفت‌وگوهاست. یک‌سوم دیگر توییت‌ها به ماهان و کاسپین مربوط می‌شود و در مجموع، باقی شرکت‌ها

یک سوم نهایی را تشکیل می دهند. با توجه به سهم اندک و الگوی منفی کارون و ایران آسمان، بازنگری فوری در راهبرد خدمات و ارتباطات این شرکت ها ضروری است؛ البته تفسیر نتایج باید همراه با تعداد نمونه هر شرکت و بازه اطمینان انجام شود. در این پژوهش با پردازش زبان طبیعی، نمایه های واژگانی سه محور اصلی گفت و گوی کاربران درباره پروازهای داخلی استخراج شد: «امنیت پرواز»، «کیفیت خدمات» و «سلامت/بهداشت». شمارش رخداد واژه های نماینده نشان داد که امنیت پرواز بیشترین توجه را دارد، سپس کیفیت خدمات و در ادامه مسائل سلامت، به ویژه در دوره همه گیری قرار داشتند. برای صورت بندی دقیق تر، داده های پیش پردازش شده با الگوریتم تخصیص نهفته دیریکله مدل سازی شد و خوشه های موضوعی هم راستا با همین سه محور (به همراه یک خوشه مکمل) به دست آمد. بر این اساس، تمرکز مدیریتی بر ارتقای ایمنی، بهبود تجربه خدمت و رعایت الزامات بهداشتی می تواند رضایت مشتری و در پی آن سود و سهم بازار را افزایش دهد.

جدول ۵. کلمات پرتکرار در هر خوشه

موضوع	کلمات پرتکرار
Topic 1	east, plane, crash, flight, aviation, kill, dead
Topic 2	photo, near, crash, travel, Yazd, people, like
Topic 3	airports, route, park, dozens, weather, poor, kill
Topic 4	Ghale, open, report, Hormuz, bind, Tunb, force, board, viva, first

برای پیاده سازی مدل سازی موضوع ال دی ای^۱، ابتدا مدل از کتابخانه ژنسم^۲ بار گذاری می شود و داده های پیش پردازش شده به لغت نامه داده می شوند. سپس تعداد موضوعات (۴ خوشه) مشخص شده و مدل توپیت ها را بر اساس شباهت به این ۴ دسته تقسیم می کند. در نهایت، نمودار مربوط به موضوعات استخراج شده نمایش داده می شود.

1. LDA
2. Genism

تحلیل خوشه‌ها

در تحلیل محتوای توییت‌های پیرامون شرکت‌های هواپیمایی ایرانی، چهار خوشه موضوعی با تمرکز غالب بر مقوله «امنیت پرواز» شناسایی شد؛ این خوشه‌ها مضامینی مانند رخدادهای سانحه (از جمله موارد نسبت داده شده به سپاهان و بویینگ)، فرسودگی ناوگان و پیامدهای محدودیت‌های خارجی را دربرمی‌گیرند. توزیع قطبیت احساسات به ترتیب ۱۵ درصد مثبت، ۲۵ درصد منفی و ۶۰ درصد خنثی برآورد شد. تحلیل مقطعی-زمانی نشان داد در دوره پساکرونا سهم احساسات مثبت افزایش و سهم احساسات منفی کاهش یافته است. نمایه‌های واژگانی پر تکرار نیز عمدتاً حول «امنیت پرواز»، «کیفیت خدمات» و «سلامت» متمرکز بودند و از منظر توجه کاربری، ایران‌ایر، ماهان و کاسپین سهم بیشتری از گفت‌وگوها را به خود اختصاص دادند. مدل‌سازی موضوع با الگوریتم تخصیص نهفته دیریکله ساختار چهارگانه را تأیید و محوریت امنیت را تثبیت کرد.

از منظر کارایی روش، چارچوب پیشنهادی در مقایسه با رویکردهای کلاسیک تحلیل احساسات، دقت بالاتری به دست آورد و نسبت به برخی مدل‌های عمیق نیز بهبود معناداری در فاز آموزش و آزمون ثبت شد. افزون بر این، گنجانیدن تفسیر نظام‌مند شکلک‌ها در جریان پیش‌پردازش، دقت برجسب‌گذاری احساس را در تمامی پیکربندی‌ها بهبود یا حداقل حفظ کرد. مقایسه پنج‌مارک^۱ شرکت‌ها نشان داد معراج، کیش، ماهان، ایران‌ایر و زاگرس سهم بالاتری از بازخورد مثبت دارند؛ در حالی که کارون، ایران‌آسمان، کاسپین و آتا با نسبت‌های منفی‌تری مواجه هستند. این الگوها مؤید آن است که مداخلات مدیریتی هدفمند در دو محور «ایمنی» و «کیفیت خدمت» می‌تواند به ارتقای معنادار رضایت ادراک شده بینجامد.

1. Benchmark

جدول ۶. مقایسه مدل با شکلک و بدون شکلک

مدل	متن + شکلک				متن			
	دقت	صحت	بازفراخوانی	F1	دقت	صحت	بازفراخوانی	F1
رگرسیون لجستیک	0/78	0/74	0/66	0/68	0/77	0/76	0/68	0/71
بردار پشتیبان	0/78	0/73	0/71	0/74	0/78	0/74	0/69	0/71
بیز ساده	0/52	0/59	0/57	0/6	0/64	0/59	0/59	0/77

این جدول نشان می‌دهد که مدل‌های رگرسیون لجستیک و بردارپشتیبان در هر دو حالت «با شکلک» و «بدون شکلک» عملکرد مشابهی داشتند، در حالی که مدل بیز ساده در حالت بدون شکلک‌ها عملکرد بهتری از خود نشان داد. مقایسه با پژوهش ساراسویتا (۲۰۲۱) که تحلیل احساسات توییت‌های مرتبط با خطوط هوایی آمریکا را با استفاده از روش‌های طبقه‌بندی سنتی انجام داده بودند، نشان می‌دهد که مدل‌های این پژوهش دقتی بالاتر از ۰/۵ تا ۰/۲۸ نسبت به مدل‌های آن‌ها کسب کرده‌اند. همچنین، در مقایسه با پژوهش مونیکا (۲۰۱۹) که از مدل حافظه بلندمدت دوطرفه مبتنی بر ربرتا استفاده کرده بودند، مدل این پژوهش در آموزش و آزمون به ترتیب ۰/۱۱ و ۰/۰۳ دقت بهتری داشت. در مطالعه کوون (۲۰۲۱)، تحلیل احساسات و مدل‌سازی موضوعی با الگوریتم تخصیص نهفته دیریکله نشان داد که عوامل کلیدی در انتخاب مسافر شامل کیفیت غذا، سرگرمی، رتبه صندلی، راحتی و خدمات کارکنان هستند. پژوهش فیلد و همکاران (۲۰۲۲) نیز نشان داد که پیش از همه‌گیری کووید-۱۹ نگرانی‌ها حول کیفیت خدمات و تأخیر پرواز می‌چرخید، در حالی که پس از آن، نگرانی‌ها به اجرای ضعیف پروتکل‌های بهداشتی تغییر یافت. یافته‌های این پژوهش نشان می‌دهد که امنیت پروازها حتی در دوران کرونا همچنان مهم‌ترین نگرانی کاربران ایرانی در توییت‌ها بوده است.

بحث و نتیجه‌گیری

یافته‌های این پژوهش نشان می‌دهد گفت‌وگوهای کاربران درباره شرکت‌های هواپیمایی ایرانی بیش از هر چیز حول امنیت پرواز می‌چرخد و سپس، موضوعات کیفیت خدمت و

ملاحظات سلامت مطرح می‌شود. این الگو با توجه به سابقه رخدادهای ایمنی، فرسودگی ناوگان و حساسیت رسانه‌ای مرتبط، قابل انتظار است. در عین حال، روند زمانی پس از دوره کرونا تصویر امیدوارکننده‌تری ترسیم می‌کند و نشان می‌دهد ارتقای رویه‌های اطلاع‌رسانی، بهداشت سفر و مدیریت عملیات می‌تواند به تعدیل نگرش عمومی بینجامد. ترکیب تحلیل احساسات با موضوع کاوی در این مطالعه کمک کرد تا تنها به نسبت مثبت/منفی بسنده نشود و معلوم شود این احساسات دقیقاً به کدام جنبه‌ها متصل‌اند. برای نمونه، وقتی امنیت محور گفت‌وگو است، سهم نامتوازی از واژگان مربوط به حادثه، فرسودگی و تحریم ظاهر می‌شود و در مقابل، در خوشه‌های تجربه کابین و مقاصد گردشگری، واژگان قدردانی و رضایت برجسته هستند. این تفکیک به تصمیم‌گیران نشان می‌دهد که کاهش احساس منفی صرفاً با تبلیغات ممکن نیست و باید از مسیر اقدام‌های مشخص در هر جنبه پیش رفت.

از منظر روش‌شناختی، نتایج نشان داد مدل خطی سبک و قابل تبیین در متن‌های کوتاه توییت‌ر، رفتاری پایدار و رقابتی دارد. این نکته برای کاربردهای عملی اهمیت دارد؛ زیرا در سامانه‌های پایش برخط، سرعت، پایداری و قابلیت توضیح تصمیم‌ها برای پذیرش سازمانی حیاتی است. در عین حال، می‌توان برای متن‌های بلندتر یا کارهای تخصصی‌تر، از مدل‌های پیچیده‌تر نیز بهره گرفت و آن‌ها را به صورت مکمل وارد چرخه تحلیل کرد. استفاده از آزمون‌های آماری برای سنجش معناداری اختلاف مدل‌ها نیز اطمینان می‌دهد که برتری مشاهده‌شده تصادفی نیست و از نظر عملی معنا دارد. مقایسه با پیشینه بین‌المللی نشان می‌دهد در بسیاری از بافت‌ها، تأخیر و تجربه کابین در صدر موضوعات قرار می‌گیرد، اما در زمینه ایران -در دوران کرونا- امنیت پرواز همچنان برجسته است. تفاوت بافتی را می‌توان با ساختار ناوگان، فشارهای بیرونی بر صنعت و حساسیت تاریخی افکار عمومی توضیح داد. این دگرگونی وزن موضوعات نشان می‌دهد راهبردهای ارتباطی و بهبود خدمت باید بومی‌سازی شوند و نمی‌توان نسخه‌های آماده از دیگر بازارها را بدون تطبیق به کار گرفت.

پیوند احساس و موضوع، سه مسیر اقدام روشن ارائه می‌دهد: ۱) شفافیت و سرعت در اطلاع‌رسانی رویدادهای ایمنی، اعلام منظم شاخص‌های نگهداشت/ایمنی و پاسخگویی به موقع در شبکه‌های اجتماعی برای کاهش احساس منفی؛ ۲) بهبود تجربه عملیاتی (مدیریت تأخیر، خدمات کابین، رسیدگی به بار، تعامل حرفه‌ای در اختلال‌ها) برای افزایش رضایت و ۳) رصد مستمر خوشه‌های حساس و مداخله پیش‌دستانه برای پیشگیری از موج‌های انتقادی. تفسیر تفاوت شرکت‌ها باید همراه با تعداد نمونه و بازه اطمینان انجام شود و برای شرکت‌هایی با سهم منفی بالاتر، نقشه جنبه‌محور و برنامه اصلاحی مرحله‌ای پیشنهاد می‌شود.

از نظر روش، نگاشت نظام‌مند شکلک‌ها و هشتک‌ها دقت تحلیل احساسات را بهبود می‌دهد و حذف آن‌ها موجب از دست رفتن سرنخ‌های لحن می‌شود. همچنین، حساسیت‌سنجی نشان می‌دهد مستندسازی کامل گردآوری، یکتاسازی و پیش‌پردازش برای اعتمادپذیری نتایج ضروری است. در جمع‌بندی، امنیت پرواز محور ادراک کاربران است و بهبود پساکرونا واقعی ولی شکننده می‌باشد؛ تثبیت آن نیازمند تداوم شفافیت، ارتقای کیفیت خدمت و پاسخگویی است. چارچوب حاضر می‌تواند پایه یک سامانه پایش مستمر برای پشتیبانی از تصمیم‌های عملیاتی و ارتباطی شرکت‌ها باشد.

محدودیت‌ها

داده‌های پژوهش فقط شامل توییت‌های انگلیسی درباره شرکت‌های هواپیمایی ایران است؛ بنابراین، سوگیری زبانی و جمعیتی محتمل است و نمی‌توان ادراک همه کاربران (به‌ویژه فارسی‌زبانان و مسافران داخلی) را به‌طور مستقیم تعمیم داد. اتکا به یک پلتفرم (توییت) و گردآوری مبتنی بر کلیدواژه نیز خطر پوشش‌ندادن برخی گفت‌وگوها یا ورود نویز (ربات/اسپم، بازنشرهای خبری) را بالا می‌برد. ضمن آنکه برچسب‌گذاری احساسات در این مطالعه عمدتاً خودکار و متکی بر رویکرد واژگانی و مدل‌های یادگیری ماشین بوده و «داده طلایی انسانی» برای اعتبارسنجی نهایی محدود/ناموجود است؛ از این‌رو، مواردی مانند طنز، کنایه، ابهام‌های معنایی و چندمعنایی واژگان می‌تواند به خطاهای طبقه‌بندی منجر شود.

از منظر روش، مدل‌سازی موضوع با تخصیص دیریکله پنهان روی متن‌های بسیار کوتاه توییت‌ها، به‌ویژه در حضور شکلک‌ها و هشتک‌ها، ممکن است به تفسیر بیش‌ازحد یا جداسازی ناقص موضوع‌ها بینجامد. علاوه بر این، برتری عملکرد رگرسیون لجستیک در این مطالعه وابسته به همین پیکربندی ویژگی‌ها و برجسب‌گذاری بوده و لزوماً در پیکربندی‌های دیگر یا مجموعه داده‌های متفاوت تکرارپذیر نیست. توزیع نامتوازن کلاس‌ها (برتری متن‌های خنثی) و همپوشانی رخداد‌های محیطی (رویداد‌های ایمنی/بهداشتی یا سیاسی در بازه زمانی آوریل ۲۰۰۸ تا مارس ۲۰۲۴) نیز می‌تواند تخمین قطبیت و روندها را دچار سوگیری زمانی کند.

کارهای آینده

برای گسترش دامنه و نمایندگی نتایج، پیشنهاد می‌شود داده‌ها از «فقط پیام‌های انگلیسی» به «بازخوردهای فارسی و دوزبانه» و نیز به دیگر بسترهای اجتماعی پرکاربرد در ایران تعمیم یابد. هم‌نشانی داده‌های برخاط کاربران با داده‌های عملیاتی صنعت (نظیر آمار تأخیر و لغو، شکایت‌های رسمی و شاخص‌های کیفیت خدمت) می‌تواند تصویر جامع‌تری از تجربه سفر فراهم کند و امکان سنجش هم‌زمان «گفته‌های مردم» و «واقعیت عملکرد» را بدهد. همچنین، شکل‌دهی یک «نمونه مرجع انسانی» با برجسب‌گذاری دقیق و بومی‌سازی شده به زبان فارسی - شامل طنز، کنایه، محاوره و شکلک‌ها - پیشنهاد می‌شود تا هم آموزش مدل‌ها و هم ارزیابی نهایی با معیارهای بومی انجام گیرد. افزون بر این، تحلیل پویایی زمانی احساس و موضوع، در پیوند با رویداد‌های مشخص (سانحه، تغییر مقررات، بحران‌های بهداشتی)، می‌تواند به درک واکنش افکار عمومی و اصلاح بهنگام سیاست‌های ارتباطی و عملیاتی یاری رساند. از منظر فنی نیز پیشنهاد می‌شود «تحلیل احساس مبتنی بر جنبه» (ایمنی، اطلاع‌رسانی، رفتار کارکنان، بهداشت، راحتی) به صورت نظام‌مند پیاده‌سازی و با «استخراج موضوع» ترکیب شود تا مسیرهای بهبود دقیق‌تر شناسایی گردد. برای متن‌های کوتاه و پرنشانه شبکه‌های اجتماعی، می‌توان از «مدل‌های زبانی پیش‌آمोخته فارسی و چندزبانه» بهره برد و با «تنظیم ریزدانه بومی» دقت را بالا برد. همچنین به کارگیری روش‌های نوین موضوع کاوی (مانند گونه‌های نوین مبتنی بر بردارهای

واژگانی و خوشه‌بندی معنایی) در کنار سنجش «تبیین‌پذیری» پیش‌بینی‌ها، به پذیرش مدیریتی نتایج کمک می‌کند. در نهایت، توصیه می‌شود خروجی‌ها در قالب «داشبورد تصمیم‌یار» برای مدیران فرودگاه و شرکت‌های هواپیمایی عرضه شود تا پیوند «موضوع + احساس» مستقیم به اقدام‌های عملی (بهینه‌سازی برنامه پروازی، بهبود پاسخگویی، تقویت ایمنی ادراک‌شده و بهداشت) برقرار گردد.

تعارض منافع

تعارض منافع ندارد.

ORCID

Ghazaleh Soltani		http://orcid.org/0009-0005-7156-2638
Ameneh Khadivar		http://orcid.org/0000-0003-4620-1409
Neda Paravian		http://orcid.org/0009-0009-5327-7154

منابع

۱. ذوالفقاری، پ.، خدیور، آ.، و عباسی، ف. (۱۴۰۱). بررسی تأثیر ریسک درک‌شده توسط کاربران شبکه‌های اجتماعی بر قیمت بیت‌کوین با استفاده از تکنیک تحلیل احساسات و شبکه عصبی NARX. *پژوهش‌های مدیریت در ایران*، ۲۶(۲)، ۴۱-۱۸. DOR: 20.1001.1.2322200.1401.26.2.2.1
2. Adeborna, E., & Siau, K. (2014). An approach to sentiment analysis—The case of airline quality rating. In *Proceedings of the Pacific Asia Conference on Information Systems (PACIS 2014)*. AIS Electronic Library.
3. Ahuja, R., Chug, A., Kohli, S., Gupta, S., & Ahuja, P. (2019). The impact of features extraction on the sentiment analysis. *Procedia Computer Science*, 152, 341–348. <https://doi.org/10.1016/j.procs.2019.05.008>
4. Angiani, G., Ferrari, L., Fontanini, T., Fornacciari, P., Iotti, E., Magliani, F., & Manicardi, S. (2016). A comparison between preprocessing techniques for sentiment analysis in Twitter. In *Proceedings of the 2nd International Workshop Knowledge Discovery on the Web (KDWeb 2016)* (pp. 1–11). CEUR-WS.

5. Anusha, K., & Radhika, A. (2017). A survey on analysis of Twitter opinion mining using sentiment analysis. *International Research Journal of Engineering and Technology*, 4(12), 1307-1310.
6. Bansal, B., & Srivastava, S. (2018). On predicting elections with hybrid topic-based sentiment analysis of tweets. *Procedia Computer Science*, 135, 346–353. <https://doi.org/10.1016/j.procs.2018.08.183>
7. Breen, J. O. (2012). Mining Twitter for airline consumer sentiment. In D. Delen, J. Elder, A. Fast, T. Hill, G. D. Miner, R. Nisbet (Eds.), *Practical text mining and statistical analysis for non-structured text data applications* (p. 133-150). Academic Press.
8. Cai, H., Dong, T., Zhou, P., Li, D., & Li, H. (۲۰۲۵). Leveraging text mining techniques for civil aviation service improvement: Research on key topics and association rules of passenger complaints. *Systems*, 13(5), 325. <https://doi.org/10.3390/systems13050325>
9. Casals-Salvador, M., Costa, F., India, M., & Hernando, J. (2024). BSC-UPC at EmoSPeech-IberLEF2024: Attention pooling for emotion recognition. *arXiv*, 2407.12467. <https://doi.org/10.48550/arXiv.2407.12467>
10. Ding, X., Liu, B., & Yu, P. S. (2008). A holistic lexicon-based approach to opinion mining. In *Proceeding of the 11th ACM International Conference on Web Search and Data Mining (WSDM 2008)* (pp. 231-240). Association for Computing Machinery.
11. Egger, R., & Yu, J. (2022). A topic modeling comparison between LDA, NMF, Top2Vec, and BERTopic to demystify Twitter posts. *Frontiers in Sociology*, 7, 886498. <https://doi.org/10.3389/fsoc.2022.886498>
12. Eum, S. W., & Leem, B. (2025). Asymmetric effect of airline customer opinions for service quality attributes: Text mining approach. *Businesses*, 5(1), 7. <https://doi.org/10.3390/businesses5010007>
13. Field, A., Varde, A. S., & Lal, P. (2022). Sentiment analysis and topic modeling for public perceptions of air travel: COVID issues and policy amendments. In *Proceedings of the LREC2022 Joint Workshop on Legal and Ethical Issues in Human Language Technologies and Multilingual De-Identification of Sensitive Language Resources* (pp. 2–8). European Language Resources Association.

14. Gouthami, S., & Hegde, N. (2023). Improving sentiment analysis on imbalanced airlines Twitter data using DSMOTE technique. *International Journal of Electronics and Communication Engineering*, 10(9), 38–51. <https://doi.org/10.14445/23488549/IJECE-V10I9P105>
15. Haddi, E., Liu, X., & Shi, Y. (2013). The role of text pre-processing in sentiment analysis. *Procedia Computer Science*, 17, 26–32. <https://doi.org/10.1016/j.procs.2013.05.005>
16. Hasib, K. Md., Habib, Md. A., Towhid, N. A., & Showrov, Md. I. H. (2021). A novel deep learning-based sentiment analysis of Twitter data for US airline service. In *Proceedings of the 2021 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD)* (pp. 450–455). IEEE. <https://doi.org/10.1109/ICICT4SD50815.2021.9396879>
17. Idris, S. L., & Mohamad, M. (2023). A study on sentiment analysis on airline quality services: A conceptual paper. *Information Management and Business Review*, 15(4), 564–576. [https://doi.org/10.22610/imbr.v15i4\(SD\)I.3638](https://doi.org/10.22610/imbr.v15i4(SD)I.3638)
18. Jain, T. I., & Nemade, D. (2010). Recognizing contextual polarity in phrase-level sentiment analysis. *International Journal of Computer Applications*, 7(5), 12–21. <https://doi.org/10.5120/1160-1453>
19. Jianqiang, Z., Xiaolin, G., & Xuejun, Z. (2018). Deep convolution neural networks for Twitter sentiment analysis. *IEEE Access*, 6, 23253–23260. <https://doi.org/10.1109/ACCESS.2017.277693>.
20. Kaur, G., & Malik, K. (2021). A sentiment analysis of airline system using machine learning algorithms. *International Journal of Advanced Research in Engineering and Technology*, 12(1), 731–742. <https://doi.org/10.34218/IJARET.12.1.2021.066>
21. Khadivar, A., Abbasi, F., Mohseni, S. A., & Basraei, R. (2025). Utilizing deep learning for aspect-based sentiment analysis in restaurant reviews. *Journal of Information Technology Management*, 17(2), 50–68. <https://doi.org/10.22059/jitm.2024.362036.3398>

22. Khan, G. F. (2015). Seven layers of social media analytics: Mining business insights from social media text, actions, networks, hyperlinks, apps, search engine, and location data. CreateSpace Independent Publishing Platform.
23. Kumar, S. S., & Zymbler, M. (2019). A machine learning approach to analyze customer satisfaction from airline tweets. *Journal of Big Data*, 6, 62. <https://doi.org/10.1186/s40537-019-0224-1>
24. Kwon, H. J., Ban, H. J., Jun, J. K., & Kim, H. S. (2021). Topic modeling and sentiment analysis of online review for airlines. *Information*, 12(2), 78. <https://doi.org/10.3390/info12020078>
25. Lee, C. K. H., & Tsao, Y. C. (2025). Decoding airline passenger sentiment dynamics using transformer-based models and time-series clustering. *International Journal of Tourism Research*, 27, e2794. <https://doi.org/10.1002/jtr.2794>
26. Martin Domingo, L., Martín, J. C., & Mandsberg, G. (2019). Social media as a resource for sentiment analysis of airport service quality (ASQ). *Journal of Air Transport Management*, 78, 106–115. <https://doi.org/10.1016/j.jairtraman.2019.01.004>
27. Melville, P., Gryc, W., & Lawrence, R. D. (2009). Sentiment analysis of blogs by combining lexical knowledge with text classification. In *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1275–1284). Association for Computing Machinery. <https://doi.org/10.1145/1557019.1557156>
28. Monika, R., Deivalakshmi, S., & Janet, B. (2019). Sentiment analysis of US airlines tweets using LSTM/RNN. In *2019 IEEE 9th International Conference on Advanced Computing (IACC)* (pp. 92–95). IEEE. <https://doi.org/10.1109/IACC48062.2019.8971592>
29. Mumtaz, D., & Ahuja, B. (2016). Sentiment analysis of movie review data using Senti-lexicon algorithm. In *2016 2nd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT)* (pp. 592–597). IEEE. <https://doi.org/10.1109/ICATCCT.2016.7912069>

30. Pandey, A. C., Rajpoot, D. S., & Saraswat, M. (2017). Twitter sentiment analysis using hybrid cuckoo search method. *Information Processing & Management*, 53(4), 764–779. <https://doi.org/10.1016/j.ipm.2017.02.004>
31. Raviya, K., & Mary Vennila, S. (2019). A novel deep learning-based sentiment analysis of movie using hybrid CNN_SVM algorithm. *International Journal of Recent Technology and Engineering*, 8(4), 12391-12394. <https://doi.org/10.35940/ijrte.D7115.118419>
32. Rohani, V. A., Shayaa, S., & Babanejaddehaki, G. (2016). Topic modeling for social media content: A practical approach. In *2016 3rd International Conference on Computer and Information Sciences (ICCOINS)* (pp. 397–402). IEEE. <https://doi.org/10.1109/ICCOINS.2016.7783248>
33. Saad, A. I. (2020). Opinion mining on US airline Twitter data using machine learning techniques. In *Proceedings of the 2020 16th International Computer Engineering Conference (ICENCO)* (pp. 59–63). IEEE. <https://doi.org/10.1109/ICENCO49778.2020.9357390>
34. Shayegan, M. J., & Dastan, Y. (2024). A multilingual sentiment analysis model in tourism. *Journal of Algorithms and Computation*, 56(1), 1–14. <https://doi.org/10.22059/jac.2024.370752.1208>
35. Singh, N., & Upreti, M. (2023). HMRFLR: A hybrid model for sentiment analysis of social media surveillance on airlines. *Wireless Personal Communications*, 132, 97–112. <https://doi.org/10.1007/s11277-023-10592-0>
36. Singh, N., Jaiswal, U. C., & Srivastava, J. (2023). Performance evaluation of machine learning classification methods for sentiment analysis. In *2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT)* (pp. 1–6). IEEE. <https://doi.org/10.1109/ICCCNT56998.2023.10307844>
37. Singh, N., & Jaiswal, U. C. (2023). A detailed sentiment analysis survey based on machine learning techniques. *Advances in Distributed Computing and Artificial Intelligence Journal*, 12(1), e29105. <https://doi.org/10.14201/adcaij.29105>

38. Song, B., Meinzer, E., Agrawal, A., & McComb, C. (2020). Topic modeling and sentiment analysis of social media data to drive experiential redesign. In *Proceedings of the ASME International Design Engineering Technical Conference and Computers and Information in Engineering Conference (IDETC-CIE 2020)*, (Vol. 11A, Paper No. T11A019). ASME. <https://doi.org/10.1115/DETC2020-22567>
39. Sutherland, D., & Cankaya, M. B. (2025). Understanding the incidents on legacy airlines with explainable artificial intelligence and text mining: Case study top 5 US airlines. *Beyond: Undergraduate Research Journal*, 8, 1-11.
40. Tan, S., & Zhang, J. (2008). An empirical study of sentiment analysis for Chinese documents. *Expert Systems with Applications*, 34(4), 2622–2629. <https://doi.org/10.1016/j.eswa.2007.05.028>
41. Wan, Y., & Gao, Q. (2015). An ensemble sentiment classification system of Twitter data for airline services analysis. In *2015 IEEE International Conference on Data Mining Workshops (ICDMW)* (pp. 1318–1325). IEEE. <https://doi.org/10.1109/ICDMW.2015.7>
42. Wilson, T., Wiebe, J., & Hoffmann, P. (2005). Recognizing contextual polarity in phrase-level sentiment analysis. In *Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing (HLT/EMNLP 2005)* (pp. 347–354). Association for Computational Linguistics. <https://doi.org/10.3115/1220575.1220619>
43. Wu, J., Lu, K., Su, S., & Wang, S. (2019). Chinese microblog sentiment analysis based on multiple sentiment dictionaries and semantic rule sets. *IEEE Access*, 7, 183924–183939. <https://doi.org/10.1109/ACCESS.2019.2960655>
44. Wu, S., & Gao, Y. (2024). Machine learning approach to analyze the sentiment of airline passengers' tweets. *Transportation Research Record*, 2678(2), 132–144. <https://doi.org/10.1177/03611981231172948>
45. Yang, S. H., Kolcz, A., Schlaikjer, A., & Gupta, P. (2014). Large-scale high-precision topic modeling on Twitter. In *Proceedings of KDD 2014* (pp. 1907–1916). Association for Computing Machinery. <https://doi.org/10.1145/2623330.2623336>

[In Persian]

1. Zolfaghari, P., Khadivar, A., & Abbasi, F. (2022). The effect of risk perceived by social network users on Bitcoin price using sentiment analysis technique and NARX neural network. *Management Research in Iran*, 26(2), 18–41. DOR: 20.1001.1.2322200.1401.26.2.2.1.

استناد به این مقاله: سلطانی, غزاله, خدیور, آمنه و پراویان, ندا . (۱۴۰۵). تحلیل احساسات و موضوع کاوی تویت های شرکت های هواپیمایی ایران (مبنتی بر یادگیری ماشین و مدل های زبانی). *مجله مطالعات مدیریت گردشگری*, ۲۱ (۷۳), ۴۴۹ – ۴۷۸

doi: 10.22054/tms.2025.87482.3099



Tourism Management Studies is licensed under a Creative Commons Attribution-Non Commercial 4.0 International License.