Industrial Management Journal
Print ISSN: 2008-5885
Online ISSN: 2423-5369

A Data-Driven Model for Ranking Airlines Using Multicriteria Decision-making and Sentiment Analysis

Fateme Fazeli
MSc. Student, Department of Industrial Engineering, Faculty of Industrial and Systems Engineering, Tarbiat Modares University, Tehran, Iran. E-mail: fateme_fazeli@modares.ac.ir

Ali Husseinzadeh Kashan*
*Corresponding Author, Associate Prof., Department of Socio-economic Systems, Faculty of Industrial and Systems Engineering, Tarbiat Modares University, Tehran, Iran. E-mail: a.kashan@modares.ac.ir

Jalil Heidary Dahooie
Associate Prof., Department of Industrial Management, Faculty of Industrial Management and Technology, College of Management, University of Tehran, Tehran, Iran. E-mail: heidaryd@ut.ac.ir

Abstract

Objective

Analyzing customer feedback from airline websites is more effective than traditional questionnaire-based methods, as these websites provide highly accurate and comprehensive information. Using big data technology, these websites collect and analyze millions of passenger reviews, offering more accurate information about customer experiences. Online reviews, as an open platform, provide the opportunity for employers to receive criticisms and suggestions, and due to their high volume and widespread dissemination, they can serve as a valuable source for analyzing customer sentiments and needs. Therefore, this study aims to propose a data-driven framework for ranking airlines, combining Multi-Criteria Decision Making (MCDM) methods with Sentiment Analysis (SA) at the aspect level. The main objective of this research is to evaluate the quality of airline services and rank them based on the reviews recorded on the SKYTRAX website from the users' sentiments hidden in their reviews.

Methods

The proposed framework consists of three stages: Stage (1): After collecting the data and preprocessing the text, airline features were extracted using the High Attribute Clustering (HAC) algorithm. Stage (2): Sentiment orientation in each airline was identified to calculate the performance scores for each airline. Stage (3): Airlines were ranked using the TOPSIS

method based on intuitive fuzzy numbers, considering the scores obtained in the second stage. Intuitive Fuzzy Sets (IFS) were used to represent effective customer opinions, including hesitant phrases in the decision matrix. Also, the criteria weights were determined through the entropy method.

Results

The performance of 10 airlines was analyzed, and ranked accordingly. The results show that for economy class airlines, with a weight of 0.17, features such as customer service, legroom, flight delays, and security inspection, each with a weight of 0.11, are equally important to passengers as other features. According to the results, Middle East Airlines demonstrates the highest performance among the ten airlines (Saudi Arabian Airlines, Kuwait Airways, Oman Air, Iran Air, Egyptair, Royal Jordanian Airlines, Middle East Airlines, Pegasus Airlines, flydubai, and Air Arabia) i.e. it has the closest distance to the positive ideal solution of the fuzzy intuitive set and the farthest distance from the negative ideal solution of the fuzzy intuitive set. While Pegasus Airlines has the closest distance to the negative ideal solution and the farthest distance from the positive ideal solution, its performance is the lowest among the four airlines.

Conclusion

This research greatly assists Middle Eastern airlines in seeking areas for improvement and in comparing their performance with their competitors to achieve a better competitive advantage in the market. The data from this research can be used to create a recommendation system for travelers helping them choose airlines that best align with their expectations, preferences, and travel goals. This can take into account factors such as budget, destination, and flight class, which can help airline managers better understand and meet customer needs.

Keywords: Intuitive fuzzy sets, Multi-criteria decision making, Online reviews, Sentiment analysis.

پژوهشگاه علوم انسانی و مطالعات فرهنگی

Citation: Fazeli, Fateme; Husseinzadeh Kashan, Ali & Heidary Dahooie, Jalil (2024). A Data-Driven Model for Ranking Airlines Using Multicriteria Decision-Making and Sentiment Analysis. *Industrial Management Journal*, 16(4), 560- 596. (*in Persian*)

Industrial Management Journal, 2024, Vol. 16, No 4, pp. 560-596
Published by University of Tehran, Faculty of Management
https://doi.org/10.22059/IMJ.2024.373514.1008133
Article Type: Research Paper
© Authors
Received: March 03, 2024
Received in revised form: October 01, 2024
Accepted: November 03, 2024
Published online: December 02, 2024

مدیریت صنعتی
شاپا چاپی: ۲۰۰۸-۵۸۸۵
شاپا الکترونیکی: ۲۴۲۳-۵۳۶۹

یک مدل داده‌محور برای رتبه‌بندی خطوط هوایی با استفاده از تصمیم‌گیری چندمعیاره و تحلیل احساسات

فاطمه فاضلی

دانشجوی کارشناسی ارشد، گروه مهندسی صنایع، دانشکده مهندسی صنایع و سیستم‌ها، دانشگاه تربیت مدرس، تهران، ایران. رایانامه: fateme_fazeli@modares.ac.ir

علی حسین‌زاده کاشان *

* نویسنده مسئول، دانشیار، گروه سیستم‌های اقتصادی و اجتماعی، دانشکده مهندسی صنایع و سیستم‌ها، دانشگاه تربیت مدرس، تهران، ایران. رایانامه: a.kashan@modares.ac.ir

جلیل حیدری دهوئی

دانشیار، گروه مدیریت صنعتی، دانشکده مدیریت صنعتی و فناوری، دانشکدگان مدیریت، دانشگاه تهران، تهران، ایران. رایانامه: heidaryd@ut.ac.ir

چکیده

هدف: تحلیل بازخورد مشتری از طریق وبسایت‌های خطوط هوایی، به‌دلیل دقت زیاد و جامعیت آن در ارائهٔ اطلاعات، بهتر از روش‌های مرسوم جمع‌آوری این بازخوردها از طریق پرسش‌نامه‌هاست. با بهره‌گیری از فناوری کلان‌داده، این وبسایت‌ها میلیون‌ها نظر مسافران را جمع‌آوری و تجزیه‌وتحلیل می‌کنند و اطلاعات دقیق‌تری از تجربه‌های مشتریان را ارائه می‌دهند. نظرهای آنلاین، به‌عنوان یک پلتفرم باز، امکان ارائهٔ انتقادها و پیشنهادها به کارفرماها را فراهم می‌کنند و به‌دلیل حجم زیاد و انتشار گسترده، می‌توانند به‌عنوان منبع ارزشمندی برای تحلیل احساسات و نیازهای مشتریان استفاده شوند. بر این اساس، پژوهش حاضر یک چارچوب داده‌محور برای رتبه‌بندی خطوط هوایی پیشنهاد می‌کند که ترکیبی از روش‌های تصمیم‌گیری چند معیاره (MCDM) و تحلیل احساسات (SA) در سطح جنبه است. هدف اصلی این پژوهش، ارزیابی کیفیت خدمات خطوط هوایی و رتبه‌بندی آن‌ها با استفاده از نظرهای ثبت‌شده در وبسایت SKYTRAX از تجربهٔ مسافران بر مبنای احساسات کاربران است که در نظرهای آن‌ها مستتر است.

روش: چارچوب پیشنهادی شامل سه مرحله است. در مرحلهٔ ۱، پس از جمع‌آوری داده‌ها و پیش‌پردازش متن، ویژگی‌های خطوط هوایی با استفاده از الگوریتم تعداد صفت بالا (HAC) استخراج می‌شوند. در مرحلهٔ ۲، جهت‌گیری احساسات در هر خط هوایی شناسایی می‌شود تا نمره‌های عملکرد مربوط به هر خط هوایی محاسبه شود. در مرحلهٔ ۳، رتبه‌بندی خطوط هوایی با استفاده از روش تاپسیس، بر مبنای اعداد فازی شهودی و با توجه به نمره‌های حاصل از مرحلهٔ دوم انجام می‌شود. مجموعه‌های فازی شهودی (IFS) برای نمایش مؤثر نظرهای مشتری، از جمله عبارات مردد، در ماتریس تصمیم‌گیری استفاده می‌شوند. همچنین وزن معیارها با استفاده از روش آنتروپی تعیین می‌شود. در ادامه، برای نشان دادن جنبهٔ کاربردیِ روش پیشنهادی، مدل ترکیبی در یک نمونهٔ واقعی برای رتبه‌بندی خطوط هوایی منتخب خاورمیانه با استفاده از بررسی نظرهای آنلاین مشتریان (OCR) در وبسایت SKYTRAX آزمایش شده است. ۸۰۱۰ نظر که از ژانویه ۲۰۱۵ تا دسامبر ۲۰۲۱ در سایت SKYTRAX ثبت‌شده، جمع‌آوری شده است.

یافته‌ها: نتایج حاصل از عملکرد ۱۰ خط هوایی تجزیه‌وتحلیل شد و رتبه‌بندی بین آن‌ها صورت گرفت. نتایج نشان داد که خطوط هوایی کلاس اقتصادی با وزن ۰/۱۷ بیشترین اهمیت را دارد و از نظر مسافران، ویژگی‌هایی مانند خدمات مشتری، فضای پا، تأخیر در پرواز و بازرسی امنیتی با وزن ۰/۱۱ در مقایسه با سایر ویژگی‌ها، اهمیت یکسانی دارند. با توجه به نتایج، خط هوایی Middle East Airlines در بین ده خط هوایی (Saudi Arabian Airlines, Kuwait Airways, Oman Air, Iran Air, Egyptair, Royal Jordanian Airlines, Middle East Airlines, Pegasus Airlines, flydubai, Air Arabia) بهترین عملکرد را دارد؛ به این معنا که B_{11} نزدیک‌ترین فاصله از راه‌حل ایدئال مثبت مجموعه فازی شهودی و دورترین فاصله از راه‌حل ایدئال منفی مجموعه فازی شهودی را دارد. در حالی که Pegasus Airlines نزدیک‌ترین فاصله به راه‌حل ایدئال منفی و دورترین فاصله از راه‌حل ایدئال مثبت را دارد و عملکرد آن در بین چهار خط هوایی کمترین است.

نتیجه‌گیری: این پژوهش به مدیران خطوط هوایی کمک می‌کند تا به‌دنبال زمینه‌های بهبود باشند و بتوانند برای دستیابی به مزیت رقابتی بهتر در بازار، عملکردشان را با رقبای خود مقایسه کنند. از داده‌های این پژوهش می‌توان به‌منظور ایجاد یک سیستم توصیه برای مسافران استفاده کرد تا به آن‌ها کمک کند خطوط هوایی‌ای را انتخاب کنند که به بهترین وجه انتظارات، خواسته‌ها و هدف سفرشان را برآورده کند. این می‌تواند عواملی مانند بودجه، مقصد و نوع کلاس پرواز را در نظر بگیرد و به درک بهتر نیازهای مشتریان برای مدیران خطوط هوایی کمک کند.

کلیدواژه‌ها: تحلیل احساسات، تصمیم‌گیری چند معیاره، مجموعه‌های فازی شهودی، نظرهای آنلاین.

استناد: فاضلی، فاطمه؛ حسین‌زاده کاشان، علی و حیدری دهوئی، جلیل (۱۴۰۳). یک مدل داده‌محور برای رتبه‌بندی خطوط هوایی با استفاده از تصمیم‌گیری چندمعیاره و تحلیل احساسات. *مدیریت صنعتی*، ۱۶(۴)، ۵۶۰- ۵۹۶.

مدیریت صنعتی، ۱۴۰۳، دوره ۱۶، شماره ۴، صص. ۵۶۰- ۵۹۶
ناشر: دانشکدهٔ مدیریت دانشگاه تهران
نوع مقاله: علمی پژوهشی
© نویسندگان
تاریخ دریافت: ۱۴۰۲/۱۲/۱۳
تاریخ ویرایش: ۱۴۰۳/۰۷/۱۰
تاریخ پذیرش: ۱۴۰۳/۰۸/۱۳
تاریخ انتشار: ۱۴۰۳/۰۹/۱۲
doi: https://doi.org/10.22059/IMJ.2024.373514.1008133

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

مقدمه

افزایش ثروت باعث رشد تقاضای سفرهای هوایی در سطح جهانی و منطقه‌ای می‌شود. با این‌حال، این رشد چالش‌های خاص خود را به‌ویژه برای تجربه مسافران ایجاد کرده است. رشد سفرهای هوایی، همچنین تقاضا برای خدمات فرودگاهی را افزایش داده و فرایند کارآمدتری از ارائهٔ خدمات به مشتریان را الزامی کرده است (اوم، فو و یو[1]، ۲۰۰۳). شرکت‌های هواپیمایی به‌شدت برای جذب مشتری با یکدیگر رقابت می‌کنند؛ زیرا جذب مشتریان برای تجارت و رشد آن‌ها در این رقابت مهم است (حسین، الناصر و حسین[2]، ۲۰۱۵). برای پایداری در این محیط رقابتی و حفظ مشتریان خود، ارائهٔ خدمات برتر به‌عنوان اولویت اصلی خطوط هوایی در نظر گرفته می‌شود (چو[3]، ۲۰۱۴). مشتریان همچنین در مورد نیازهای خود بسیار دقیق شده‌اند و تمایل دارند درصورتی‌که یک خط هوایی نتواند نیازهای آن‌ها را برآورده کند، به خط هوایی دیگر روی‌آورند (گوپتا[4]، ۲۰۱۸). بسیاری از مطالعات از روش‌های پیمایشی برای اندازه‌گیری کیفیت خدمات در صنعت هواپیمایی استفاده کرده‌اند. با این‌حال، چند مطالعه اخیر، مزایای تحلیل داده‌های نظرهای آنلاین برای مطالعه رضایت مشتریان یا تجربهٔ آن‌ها از خطوط هوایی را برجسته کرده‌اند. نظرهای آنلاین حیاتی هستند؛ زیرا منبع مهمی برای رشد تجاری، عملکرد و بهبود تجربهٔ مشتری محسوب می‌شوند و به شرکت‌های هواپیمایی این امکان را می‌دهند که با مسافران خطوط هوایی ارتباط دوطرفه داشته باشند (بان و کیم[5]، ۲۰۱۹). امروزه، بسیاری از تولیدکنندگان، سازمان‌ها و ارائه‌دهندگان خدمات به نظرها و بازخورد مشتریان برای ارزیابی محصولات و کیفیت خدمات خود نیاز دارند (گریر و لی[6]، ۲۰۱۲). تحلیل بازخورد مشتری از طریق وبسایت‌های خطوط هوایی، به‌دلیل دقت زیاد و جامعیت در ارائهٔ اطلاعات، از روش‌های مرسوم جمع‌آوری بازخوردها از طریق پرسش‌نامه بهتر است. وبسایت‌های مختص عملکرد خطوط هوایی با بهره‌گیری از فناوری کلان‌داده، میلیون‌ها نظر مسافران را جمع‌آوری و تجزیه‌وتحلیل می‌کنند و از تجربه‌های مشتریان را اطلاعات دقیق‌تری ارائه می‌دهند (وان و گائو[7]، ۲۰۱۶). نظرهای آنلاین، به‌عنوان یک پلتفرم باز، امکان ارائهٔ انتقادها و پیشنهادها را به کارفرما فراهم می‌کنند و به‌دلیل حجم زیاد و انتشار گسترده، می‌توانند به‌عنوان منبع ارزشمندی برای تجزیه‌وتحلیل احساسات (SA)[8] و نیازهای مشتریان استفاده شوند (رایت[9]، ۲۰۰۵).

امروزه استفاده از شبکه‌های اجتماعی، پیام‌رسان‌ها و وبسایت‌ها برای دریافت اطلاعات از نظرهای مشتریان رایج است (محتشمی، حسین‌زاده کاشان و حیدری دهوئی، ۱۴۰۱). اما تجزیه و تحلیل داده‌های دریافتی از این منابع به‌دلیل حجم زیاد آن‌ها نیازمند ابزار کارآمدی است. SA در حال تبدیل‌شدن به ابزار ضروری برای تجزیه‌وتحلیل محتوای نظرهای ثبت شدهٔ مشتریان در سایت‌هاست. بر اساس این اطلاعات، شرکت‌های خدماتی مانند خطوط هوایی، می‌توانند یک استراتژی ارتباطی برای بهبود تصویر مشتریان خود از شرکت و خدمات دریافت شده طراحی و اجرا کنند. متن‌کاوی

1. Oum, Yu & Fu
2. Hussain, Al Nasser & Hussain
3. Chow
4. Gupta
5. Ban & Kim
6. Greer & Lei
7. Wan & Gao
8. Sentiment analysis
9. Wright

و تجزیه‌وتحلیل‌های مرتبط با آن، مانند SA می‌تواند به محققانی که روی نظرکاوی کار می‌کنند، کمک کند تا نظرهای مصرف‌کنندگان دربارهٔ محصولات و خدمات را نسبت به قبل، مؤثرتر و به‌موقع‌تر درک کنند (راوی و راوی[1]، ۲۰۱۵). علاوه‌براین، پردازش زبان طبیعی یکی دیگر از رشته‌های هوش مصنوعی است که بر تعامل بین انسان و رایانه از طریق زبان طبیعی تمرکز دارد. اکثر روش‌های پردازش زبان طبیعی که سعی در بازیابی و درک معنای زبان‌های انسانی دارند، مبتنی بر روش‌های یادگیری مصنوعی هستند (رئیسی وانانی و مجیدیان[2]، ۲۰۲۰). داده‌های زیادی در اینترنت وجود دارد؛ اما این داده‌ها زمانی اهمیت دارند که حاوی اطلاعات مفیدی هستند. سازوکار تصمیم‌گیری افراد که از آن برای رسیدن به اهداف خود استفاده می‌کنند، تحت تأثیر مجموعه داده قرار می‌گیرد؛ اما باید اطلاعات مفیدی در داده‌ها برای تصمیم‌هایشان وجود داشته باشد (کایا و اوزتورک[3]، ۲۰۲۰).

پژوهشگران به‌طور فزاینده‌ای از روش‌های تصمیم‌گیری چندمعیاره MCDM برای مدیریت پیچیدگی‌های رتبه‌بندی کیفیت خدمات استفاده می‌کنند. MCDM به‌عنوان یکی از شاخه‌های تحقیق در عملیات، بر حل مسائل تصمیم‌گیری تمرکز دارد که معیارهای متعدد و اغلب متضادی را شامل می‌شود (طاهردوست و معدنچیان[4]، ۲۰۲۳).

مدیران و ذی‌نفعان خطوط هوایی موظف‌اند تا از منظر مشتریان، اساسی‌ترین معیارها را شناسایی کنند. روش MCDM این امکان را فراهم می‌کند که گزینه‌های مختلف بر اساس عملکرد آن‌ها در معیارهای گوناگون ارزیابی و مقایسه شوند و در نهایت، انتخاب بهینه یا مناسب‌ترین گزینه براساس یک مدل ترجیحی مشخص صورت گیرد (ساهو و گوسوامی[5]، ۲۰۲۳).

روش‌های سنتی MCDM با چالش‌هایی مانند محدودیت در دسترسی به داده‌ها، ناکارآمدی در سازوکارهای رتبه‌بندی، عدم ثبات رتبه‌بندی و تکیه بیش از حد بر داده‌های کمی تاریخی مواجهند (سوتوده انواری[6]، ۲۰۲۲). با توجه به چالش‌های موجود، احساس نیاز فوری به بهبود روش‌های سنتی تصمیم‌گیری چندمعیاره با استفاده از منابع اطلاعاتی مکمل احساس می‌شود. تکنیک‌های متن‌کاوی در سال‌های اخیر به‌عنوان ابزاری مؤثر برای استخراج اطلاعات از مجموعه‌های وسیع داده‌های متنی شناخته شده‌اند (آنتونز، گرونوالد، سیچی و سالج[7]، ۲۰۲۰؛ حسنی، بنکی، اونگر، مزینانی و یگانگی[8]، ۲۰۲۰). برای مقابله با چالش‌های تصمیم‌گیری در صنعت هواپیمایی، ناظران و مدیران عملیاتی به‌طور فزاینده‌ای از روش‌های MCDM استفاده می‌کنند تا بتوانند عوامل مؤثر بر کیفیت خدمات را از دیدگاه مشتریان شناسایی نمایند. ادغام جنبه‌های کیفی با MCDM، به تقویت چارچوب تصمیم‌گیری کمک کرده و امکان درک جزئیات مهمی را فراهم می‌آورد که ممکن است در تحلیل‌های صرفاً کمی نادیده گرفته شوند (هاشم خانی، زلفانی و درخشانی[9]، ۲۰۲۰).

تحلیل بازخورد مشتری برای بهبود خدمات خطوط هوایی ضروری است؛ اما روش‌های مرسوم مانند پرسش‌نامه

1. Ravi & Ravi
2. Raeesi Vanani & Majidian
3. Kaya & Öztürk
4. Taherdoost and Madanchian
5. Sahoo and Goswami
6. Sotoudeh-Anvari
7. Antons, Grünwald, Cichy & Salge
8. Hassani, Beneki, Unger, Mazinani & Yeganegi
9. Hashemkhani Zolfani and Derakhti

وقت‌گیر و اغلب نادرست هستند. مشکلاتی مانند عدم جدیت مشتریان در پُر کردن پرسش‌نامه‌ها و ورود داده‌های نویز به تحلیل احساسات، کارایی این روش‌ها را کاهش می‌دهد (وان و گائو[1]، ۲۰۱۶).

بر خلاف پرسش‌نامه‌ها، وبسایت‌های خطوط هوایی منبع بهتری برای جمع‌آوری احساسات مشتریان هستند. فناوری‌های کلان‌داده به جمع‌آوری و تجزیه‌وتحلیل میلیون‌ها نظر مسافران کمک کرده و باعث صرفه‌جویی در هزینه‌ها می‌شود. نظرهای آنلاین اطلاعات دقیق‌تر و جامع‌تری درباره تجربیات مشتریان ارائه می‌دهند و می‌توانند نیازها و ترجیحات آن‌ها را بهتر نمایان کنند (ژانگ و همکاران[2]، ۲۰۲۰). همچنین، این نظرها به‌طور خودجوش ثبت می‌شوند و به کارفرماها امکان ارائهٔ انتقادها و پیشنهادهای بهبود خدمات را می‌دهند و به‌عنوان منبع ارزشمندی برای تحلیل احساسات مشتریان محسوب می‌شوند (رایت[3]، ۲۰۰۵). به این دلیل در این پژوهش به ارائهٔ یک مدل ترکیبی مبتنی بر SA و MCDM برای ارزیابی کیفیت خدمات خطوط هوایی که کاملاً داده محور است، پرداخته می‌شود.

در این پژوهش، به این نکته توجه شده است که نظرهای آنلاین ممکن است نمایندهٔ کامل تجربه‌های تمامی مشتریان نباشند. به‌منظور کاهش این اثر اریبی، تلاش شده است تا نظرها، از منابع مختلف و در بازه‌های زمانی متفاوت جمع‌آوری شود. همچنین، از تکنیک‌های پیشرفته‌ای برای تحلیل داده‌ها استفاده شده است که می‌تواند به شناسایی و تعدیل اریبی‌های موجود کمک کند.

چارچوب پیشنهادی، یک مرحلهٔ پیش پردازش متن و سه مرحلهٔ اصلی را دربرمی‌گیرد. در مرحلهٔ اول، پس از پیش پردازش متن، ویژگی‌های خطوط هوایی با استفاده از الگوریتم تعداد صفت بالا (HAC)[4] استخراج می‌شوند. در مرحلهٔ دوم، جهت‌گیری احساسات در هر خط هوایی شناسایی می‌شود تا نمره‌های عملکرد مربوط به هر خط هوایی محاسبه شود. این شناسایی احساسات از طریق روش SA و با استفاده از فرهنگ لغت احساسات صورت می‌پذیرد. در مرحلهٔ سوم، رتبه‌بندی خطوط هوایی با استفاده از روش تاپسیس فازی شهودی و با توجه به نمره‌های حاصل از مرحلهٔ دوم انجام می‌شود. در این روش، مجموعه‌های فازی شهودی (IFS)[5] برای نمایش مؤثر نظرهای مشتری، از جمله عبارات مردد در ماتریس تصمیم‌گیری استفاده می‌شوند. همچنین وزن معیارها با استفاده از روش آنتروپی تعیین می‌شود. این پژوهش عمدتاً بر تجزیه‌وتحلیل دیدگاه‌های مشتریان در مورد خدمات خطوط هوایی تمرکز دارد. همچنین از داده‌های ثبت‌شده در وبسایت SkyTRAX[6] برای تجزیه‌وتحلیل عملکرد و رتبه‌بندی خطوط هوایی استفاده می‌کند. به‌منظور ارزیابی اثربخشی چارچوب پیشنهادی، از آن برای رتبه‌بندی و انتخاب بهترین خط هوایی خاورمیانه از بین گزینه‌های منتخب استفاده می‌شود. هدف اصلی از این پژوهش، ارزیابی کیفیت خدمات خطوط هوایی و رتبه‌بندی آن‌ها با استفاده از نظرهای ثبت‌شده در وبسایت SKYTRAX از تجربهٔ مسافران، بر مبنای احساسات کاربران است که در نظرهای آن‌ها مستتر است.

1. Wan and Gao
2. Zhang et al.
3. Wright
4. High Adjective Count
5. Intuitionistic fuzzy set

۶. این سایت بستری را برای مسافران خطوط هوایی فراهم می‌کند تا نظرشان را دربارهٔ تجربهٔ پرواز خود در چهار دسته به اشتراک بگذارند: فرودگاه‌ها، خطوط هوایی، سالن‌ها و صندلی‌ها.

با توجه به اهمیت بخش خطوط هوایی در صنعت گردشگری، این مطالعه به‌صورت تجربی به بررسی تجربهٔ مسافران خطوط هوایی از سفرهایشان و رتبه‌بندی خطوط هوایی توسط تجزیه‌وتحلیل داده‌های موجود در وبسایت SKYTRAX می‌پردازد.

در طول این مسیر، خطوط هوایی این فرصت را دارد تا به درک عوامل موجود در وبسایت مورد بررسی دست یابد تا بدین‌وسیله به این بازار نفوذ کرده و استراتژی‌های بازاریابی متناسب با آن را برای مزیت‌های قوی خود ایجاد کند. درک نظرهای آنلاین به‌عنوان تجلی تجربه‌های مسافران می‌تواند به خطوط هوایی در شناسایی ویژگی‌های اصلی مورد نیاز برای دستیابی به احساسات مثبت پس از خرید و به حداقل رساندن احساسات منفی کمک کند. بنابراین، نظرهای آنلاین نه‌تنها راهی مناسب برای شرکت‌های هواپیمایی به‌منظور جمع‌آوری بازخورد از مسافران خطوط هوایی است، بلکه فرصتی برای کشف چگونگی ایجاد احساس مثبت پس از تجربه فراهم می‌کند (بان و کیم[1]، ۲۰۱۹).

پیشینهٔ پژوهش

در این بخش فقط تحقیقاتی بررسی شده است که در زمینهٔ خدماتی مانند هتل‌داری، خدمات بیمارستانی و حمل‌ونقل از دو روش SA و MCDM استفاده کرده‌اند.

شی و همکاران[2] (۲۰۲۴) یک مدل جامع برای ارزیابی کیفیت خدمات خطوط هوایی ایجاد کرده‌اند که از روش‌شناسی LSA-TOPSIS-VIKOR-AISM[3] استفاده می‌کند. این مدل با استخراج داده‌های نظرهای آنلاین آغاز می‌شود و نمره‌های احساسی به آن‌ها اختصاص می‌یابد. سپس با تحلیل کارایی و معیارهای پشیمانی، راه‌حل‌های رتبه‌بندی شده‌ای استخراج می‌شود. همچنین از روش AISM برای بررسی دقیق شاخص‌ها و تولید گراف‌های هیرارشی استفاده کرده‌اند. شواهد تجربی نشان می‌دهد که این روش از یکپارچگی داده‌ها و محاسبات ساده‌ای برخوردار است و به‌عنوان ابزاری مؤثر برای ارزیابی کیفیت خدمات محسوب می‌شود.

پارک[4] (۲۰۲۳) یک روش جدید ارزیابی سطح رضایت مشتری ارائه داده که تحلیل پوششی داده‌ها (DEA) را با متن‌کاوی ترکیب می‌کند تا نظرهای آنلاین را تحلیل کند. این روش از الگوریتم فراوانی واژه ـ معکوس فراوانی سند (TF-IDF) برای شناسایی چندین معیار رضایت استفاده می‌کند و با تحلیل احساسات، این معیارها را کمی‌سازی می‌کند. سپس، با استفاده از مدل DEA، سطح رضایت مشتریان را با توجه به این معیارها ارزیابی می‌کند.

لی، ژانگ و شو[5] (۲۰۲۰) یک مدل MCDM برای ارزیابی محصولات و خدمات با استفاده از بررسی نظرهای آنلاین مشتریان (OCR)[6] و SA معرفی کردند. آن‌ها از PLTS[7] برای تجزیه‌وتحلیل جهت‌گیری احساسات نظرها

1. Ban & Kim
2. Xie et al.
3. Latent Semantic Analysis
4. Park
5. Li, Zhang & Xu
6. Online customer reviews
7. Probabilistic Linguistic Term Sets

استفاده کردند و برای مسائل MCDM با وزن معیارهای ناشناخته، روش PP[1] و MULTIMOORA[2] را ترکیب کردند تا روش توسعه‌یافته‌ای به نام PP-MULTIMOORA ایجاد کنند. سپس، این مدل در ارزیابی کیفیت خدمات پزشکان به‌کار گرفته شد و تجزیه‌وتحلیل مقایسه‌ای برای نمایش اثربخشی آن انجام شد.

ژانگ، لی و وو[3] (۲۰۲۰) روش TODIM[4] را برای رتبه‌بندی محصولات با نظرهای آنلاین در محیط فازی شهودی ارائه دادند. آن‌ها از چارچوب کلمه‌ای احساسات مبتنی بر IF[5] و اندازه‌گیری فازی استفاده کردند و نتایج مطالعه موردی در مورد انتخاب اتومبیل نشان داد که روش پیشنهادی مؤثر است.

ژانگ، شیونگ، کائو، لی و هوانگ[6] (۲۰۲۰) با استفاده از روش ترکیبی با VIKOR، یک روش جدید برای رتبه‌بندی محصولات از طریق بررسی آنلاین طراحی کردند. این روش شامل سه مرحله اصلی بود: ایجاد فهرست محصولات جایگزین، جمع‌آوری نظرهای آنلاین، و پردازش و اندازه‌گیری رضایت مشتری. این مطالعه سه پیشرفت مهم را نشان داد: اول، اعمال تئوری چشم‌انداز برای اندازه‌گیری دقیق‌تر رضایت مشتری؛ دوم، بهبود وزن‌دهی با استفاده از روش وزن آنتروپی و وزن‌های مشخصه فردی و سوم، ارائهٔ یک الگوریتم تحلیل احساسات جدید برای پردازش نظرهای آنلاین. برای اعتبارسنجی روش، مطالعه موردی رتبه‌بندی خودرو و برخی مقایسه‌ها انجام شده است. این روش نه‌تنها زمان و تلاش را صرفه‌جویی می‌کند، بلکه به مشتریان کمک می‌کند تا محصولات مورد نیاز خود را انتخاب کنند.

وانگ، ونگ، پنگ و ونگ[7] (۲۰۲۰) تحلیل مقایسه‌ای را با استفاده از یک مدل پشتیبانی تصمیم‌گیری رفتاری منطقی محدود انجام دادند تا تفاوت‌های انتخاب هتل بین انواع مختلف مسافران را بررسی کنند. آن‌ها از بررسی آنلاین هتل‌ها و استفاده از الگوریتم‌های فرکانس سند فرکانس معکوس (TF-IDF)[8] و تبدیل کلمه به بردار (Word2Vec)[9] برای شناسایی عوامل کلیدی هتل و اهمیت معیار استفاده کردند. نتایج نشان داد که انواع مختلف مسافران در انتخاب هتل متفاوت‌اند؛ اما خانواده‌ها و دوستان نتایج مشابهی دارند. این مطالعه می‌تواند به مدیران هتل کمک کند تا ترجیحات مسافران را درک کرده و وبسایت گردشگری را بهینه‌سازی کنند.

هو، ژنگ، یانگ، لیو و چن[10] (۲۰۲۰) سیستم جدیدی را برای رتبه‌بندی پزشکان با استفاده از روش ویکور (VIKOR)[11] بررسی کردند. این سیستم جدید DRS[12] نام دارد که بر پایه روش تصمیم‌گیری گروهی چند معیاره (MCGDM)[13] طراحی شده است و برای حل مشکلات حفاظت از حریم خصوصی استفاده می‌شود. DRS از اطلاعات متنی برای شناسایی ترجیحات کاربر و معیارهای مکمل استفاده می‌کند و وزن معیارها با استفاده از TF-IDF تعیین

1. Projection Pursuit
2. Multiplicative Multi-Objective Optimization by Ratio Analysis
3. Zhang, Li & Wu
۴. روشی مؤثر بر اساس یک تابع آینده برای گرفتن رفتار روانی در معرض خطر است.
5. Intuitionistic fuzzy
6. Zhang, Xiong, Cao, Li & Huang
7. Wang, Wang, Peng & Wang
8. Term frequencyinverse document frequency
9. Word to vector
10. Hu, Zhang, Yang, Liu & Chen
11. Vise Kriterijumska Optimizacija
12. Doctor-Ranking System
13. Multi-criteria group decision-making

می‌شود. همچنین، IFS برای جایگزینی احساسات برای بیان معیارهای ذهنی کاربر استفاده می‌شود و روش VIKOR برای حل مشکل رتبه‌بندی پزشک به کار گرفته می‌شود. این مطالعه نشان می‌دهد که سیستم DRS دقیق و قابل اطمینان برای رتبه‌بندی پزشکان است.

چالی و بالامان[1] (۲۰۱۹) سیستمی برای رتبه‌بندی محصولات جایگزین با استفاده از معیارها و نظرهای مشتریان مرتبط در وبسایت‌ها برای پیشنهاد مناسب‌ترین جایگزین به مشتریان مورد بررسی قرار دادند. این سیستم دارای دو مرحله است: در مرحلهٔ اول، نظرهای آنلاین مشتریان با استفاده از SA به امتیازهای رضایت مشتری تبدیل می‌شود و در مرحلهٔ دوم، محصولات جایگزین با استفاده از روش IF-ELECTRE یکپارچه با VIKOR رتبه‌بندی می‌شوند. از IFS برای نمایش نظرهای مشتری استفاده می‌شود و وزن معیارها با استفاده از روش آنتروپی تعیین می‌شود. این رویکرد با یک مطالعه موردی که در آن نظرهای مشتریان دربارهٔ تجارب هتل ارزیابی شده و هتل‌های جایگزین رتبه‌بندی می‌شوند، ارزیابی شده است.

شارما، تاندون، کاپور و آگاروال[2] (۲۰۱۹) یک مدل رتبه‌بندی هتل‌ها بر اساس جنبه‌هایی مانند خدمات، تمیزی، ارزش، کیفیت خواب، اتاق و مکان، با استفاده از طبقه‌بندی احساسات مبتنی بر رتبه‌بندی و TOPSIS[3] نوتروسوفیک با ارزش بازه‌ای مورد بررسی قرار دادند. این جنبه‌ها را به احساسات مثبت، خنثی و منفی تبدیل کرده و به اعداد Neutrosophic تبدیل می‌کنند، سپس ماتریس تصمیم نوتروسوفیک با ارزش بازه‌ای را شکل می‌دهند. وزن‌های جنبه با استفاده از روش حداکثر کردن انحراف تعیین می‌شوند. در نهایت، این وزن‌ها و ماتریس تصمیم برای انجام روش مورد نیاز برای رتبه‌بندی پنج هتل جایگزین ترکیب می‌شوند.

لیانگ، لیو و وانگ[4] (۲۰۱۹) مدل حمایتی جدیدی برای تصمیم‌گیری در انتخاب هتل با استفاده از نظرهای آنلاین را بررسی کردند. این مدل شامل دو مرحله است: ۱. تبدیل ترجیحات احساسات گردشگران به قالب توزیع زبانی با تجزیه‌وتحلیل کلمه‌های احساسی در نظرهای آنلاین؛ ۲. استفاده از روش DL-VIKOR[5] برای رتبه‌بندی و انتخاب هتل‌ها بر اساس وزن‌دهی به ویژگی‌های ارزیابی. در نهایت، از یک مورد واقع‌بینانه از TripAdvisor.com برای انتخاب هتل برای نشان دادن عملی و امکان‌سنجی مدل پیشنهادی بهره بردند.

لیانگ و وانگ[6] (۲۰۱۹) مدل پشتیبانی تصمیم‌گیری ابری شهودی زبانی با تحلیل احساسات برای انتخاب محصول در تجارت الکترونیک را بررسی کردند. این مدل شامل سه ماژول است که شامل اکتساب اطلاعات، تبدیل اطلاعات و مدل یکپارچه‌سازی است. ماژول اکتساب اطلاعات برای جمع‌آوری اطلاعات فازی شهودی زبانی (LINC)[7] از طریق تحلیل احساسات استفاده می‌شود، سپس ماژول تبدیل اطلاعات برای تبدیل آن‌ها به ابرهای عادی شهودی زبانی استفاده می‌شود. در نهایت، ماژول یکپارچه‌سازی برای به دست آوردن لیست رتبه‌بندی شده از محصولات جایگزین استفاده

1. Çalı & Balaman
2. Sharma, Tandon, Kapur & Aggarwal
3. Technique for Order of Preference by Similarity to Ideal Solution
4. Liang, Liu & Wang
5. Distribution linguistic VIKOR
6. Liang & Wang
7. linguistic intuitionistic normal clouds

می‌شود. این مدل از یک مطالعه موردی در Taobao.com برای نشان دادن اثربخشی و امکان‌سنجی پیشنهاد، همراه با تحلیل‌های حساسیت و مقایسه، برای تأیید ثبات و برتری آن استفاده کرده است.

لیو و تنگ[1] (۲۰۱۹) روش TODIM زبانی احتمالی توسعه‌یافته (PL-TODIM)[2] را برای ارزیابی محصولات جایگزین از طریق نظرهای مصرف‌کنندگان در خصوص عملکرد محصول بررسی کردند. این مطالعه یک روش جدید MADM[3] را برای رتبه‌بندی محصولات بر اساس OPR[4] ها معرفی می‌کند. برای این منظور، نظریه‌های اساسی PLTS بررسی شده و یک فرمول امکان برای مقایسه PLTS پیشنهاد شده است. همچنین، یک روش وزن‌دهی ترکیبی برای تعیین وزن‌های هدف بر اساس معیارهای آنتروپی متقاطع و آنتروپی توسعه داده شده است. مطالعه موردی خودروهای شاسی‌بلند را شامل می‌شود و در نهایت، مقایسه با سایر روش‌های موجود برای نشان دادن مزایای آن انجام شده است.

جدول ۱ مطالعات اصلی را نشان می‌دهد که خدمات یا محصولات را برای اهداف مختلف با استفاده از OCR و از طریق ترکیبی از روش‌های SA و MCDM مورد بررسی قرار داده‌اند.

جدول ۱. دسته‌بندی مطالعات پیشین با استفاده از SA-MCDM از طریق OCR

نویسنده	روش استخراج ویژگی	روش وزن‌دهی ویژگی‌ها	روش رتبه‌بندی	تحلیل احساسات			مطالعه موردی
				مثبت	خنثی	منفی	
شیه و همکاران، ۲۰۲۴	LSA	TF-IDF	TOPSIS-VIKOR-AIS	✓	✓	✓	خطوط هوایی
پارک، ۲۰۲۳	TF-IDF	TF-IDF	DEA	✓	✓	✓	خطوط هوایی
حیدری دهوئی، رفعت، قربانی و دایم[5]، ۲۰۲۱	الگوریتم HAC	IF-IDOCRIW	IF-MULTIMOORA	✓	✓	✓	گوشی‌های همراه
لی و همکاران، ۲۰۲۰	فرکانس مدت	Projection pursuit	MULTIMOORA	✓	✓	✓	کیفیت خدمات پزشکان
ژانگ و همکاران، ۲۰۲۰	اسم جمعی، صفت، افعال از POS tagging	فرکانس و درجه توجه ویژگی	TODIM	✓	✓	✓	اتومبیل
کومار و پاریمالا[6]، ۲۰۲۰	فرکانس مدت	AHP	TOPSIS	✓	-	✓	گوشی‌های هوشمند
ژانگ و همکاران، ۲۰۲۰	ویژگی‌های مشخص‌شده در پلتفرم	ترکیبی از روش آنتروپی و وزن هدف ارائه‌شده توسط مشتریان	VIKOR	✓	-	✓	اتومبیل

1. Liu and Teng
2. Probabilistic linguistic TODIM
3. Multiple Attribute Decision Making
4. Online Product Reviews
5. Heidary Dahooie, Raafat, Qorbani & Daim
6. Kumar & Parimala

نویسنده	روش استخراج ویژگی	روش وزن‌دهی ویژگی‌ها	روش رتبه‌بندی	تحلیل احساسات			مطالعه موردی
				مثبت	خنثی	منفی	
وانگ و همکاران، ۲۰۲۰	TF-IDF	Word2Vec	TODIM	✓	✓	✓	هتل
یانگ، اویانگ، فو و پنگ[1]، ۲۰۲۰	مدل یادگیری عمیق	عملگرهای q-ROFIWHM	عملگرهای q- ROFIWHM و عملکرد امتیاز	✓	✓	✓	تلفن‌های همراه
هو و همکاران، ۲۰۲۰	فرکانس مدت	TF-IDF	VIKOR	✓	✓	✓	پزشکان
لیو و تنگ، ۲۰۱۹	ترکیبی از استخراج‌کننده‌های پویا و استاتیک	آنتروپی متقابل و معیارهای آنتروپی	TODIM	✓	✓	✓	وسایل نقلیه کاربردی اسپرت
چالی و بالامان، ۲۰۱۹	فرکانس مدت	Entropy	ترکیبی از ELECTRE and VIKOR	✓	✓	✓	هتل
شارما و همکاران، ۲۰۱۹	ویژگی‌های مشخص‌شده در پلتفرم	Maximizing deviation	TOPSIS	✓	✓	✓	هتل
لیانگ و همکاران، ۲۰۱۹	ویژگی‌های مشخص‌شده در پلتفرم	فراوانی کلمه‌ها	VIKOR	✓	✓	✓	هتل
لیانگ و وانگ، ۲۰۱۹	مرور ادبیات	Entropy	LINCNWBM عملگر	✓	✓	✓	محصولات مراقبت از پوست
وو و ژانگ، ۲۰۱۹	با استفاده از نظر کارشناسان و اسامی به دست آمده از POS Tagging	استفاده از فرکانس و درجه توجه	TODIM	✓	✓	✓	تلفن‌های همراه
یانگ و همکاران، ۲۰۱۹	فرکانس مدت	زوال تصاعدی زمان	امتیاز فازی بصری و فاصله طرح‌ریزی عمودی	✓	✓	✓	تلفن‌های همراه
تحقیق حاضر	HAC TF-IDF	Entropy	TOPSIS	✓	✓	✓	خطوط هوایی خاورمیانه

تحقیقات قبلی در مورد SA بر مبنای نظرهای آنلاین در زمینه خطوط هوایی در درجه اول بر استفاده از روش‌های طبقه‌بندی و مدل‌سازی موضوع متمرکز بوده است. با این حال، هنوز مطالعه‌ای در مورد استفاده از SA در ترکیب با روش‌های MCDM برای رتبه‌بندی خطوط هوایی بر اساس نظرهای آنلاین انجام نشده است. انجام این پژوهش به‌طور بالقوه می‌تواند سهم چشمگیری در این حوزه از تحقیقات داشته باشد.

1. Yang, Ouyang, Fu & Peng

همان‌طور که در جدول ۱ مشخص شده است، این مطالعات تأثیر چشمگیری بر رتبه‌بندی خدمات با استفاده از SA داشته‌اند؛ اما خلأهایی نیز در این مطالعات به چشم می‌خورد که در زیر به آن‌ها اشاره شده است:

- در مرحلهٔ استخراج ویژگی‌های محصول یا خدمات، بیشتر مطالعات قبلی روی ویژگی‌های مشخص‌شده در وبسایت‌ها متمرکز شده‌اند (برای مثال: ژانگ و همکاران، ۲۰۲۰؛ شارما و همکاران، ۲۰۱۹؛ لیانگ و همکاران، ۲۰۱۹) یا از رویکردهای مبتنی بر فرکانس استفاده کرده‌اند که ویژگی‌های مهم با فرکانس پایین را نادیده می‌گیرند (به‌طور مثال: لی و همکاران، ۲۰۲۰؛ کومار و پاریمالا، ۲۰۲۰؛ هو و همکاران، ۲۰۲۰؛ چالی و بالامان، ۲۰۱۹؛ یانگ و همکاران، ۲۰۱۹).
- با تغییر تکنولوژی، نیازهای مشتریان تغییر کرده و در نتیجه ویژگی‌های جدید برای آن‌ها مهم می‌شود. از این رو، یکی از راه‌های مهم برای همگام شدن با روندهای بازار، مشارکت مشتری در فرایند توسعه محصول یا خدمات است.
- موضوع مهم دیگر در مرحله شناسایی جهت‌گیری‌های احساسی نظرهای مشتریان است. چند مطالعه سعی کرده‌اند با در نظر گرفتن جهت‌گیری‌های مثبت و منفی در نظرها، موضوع عدم اطمینان مشتری را مدل‌سازی کنند (به‌طور مثال، ژانگ و همکاران، ۲۰۲۰؛ راوی و راوی، ۲۰۲۰؛ کومار و پاریمالا، ۲۰۲۰)، درحالی‌که موضوع جهت‌گیری احساسات خنثی در نظرهای مشتریانی که ابراز تردید می‌کنند، کمتر موردتوجه محققان در این زمینه قرار گرفته است (به‌طور مثال، وانگ و همکاران، ۲۰۲۰؛ شارما و همکاران، ۲۰۱۹؛ چالی و بالامان، ۲۰۱۹).

روش‌شناسی پژوهش

شکل ۱ روند حل چارچوب پیشنهادی بر اساس متدولوژی CRISP-DM[1] جهت رتبه‌بندی خطوط هوایی از طریق نظرهای آنلاین مسافران را نشان می‌دهد که در ادامه تمام مراحل آن شرح داده می‌شود.

استخراج ویژگی‌های خطوط هوایی از OCR

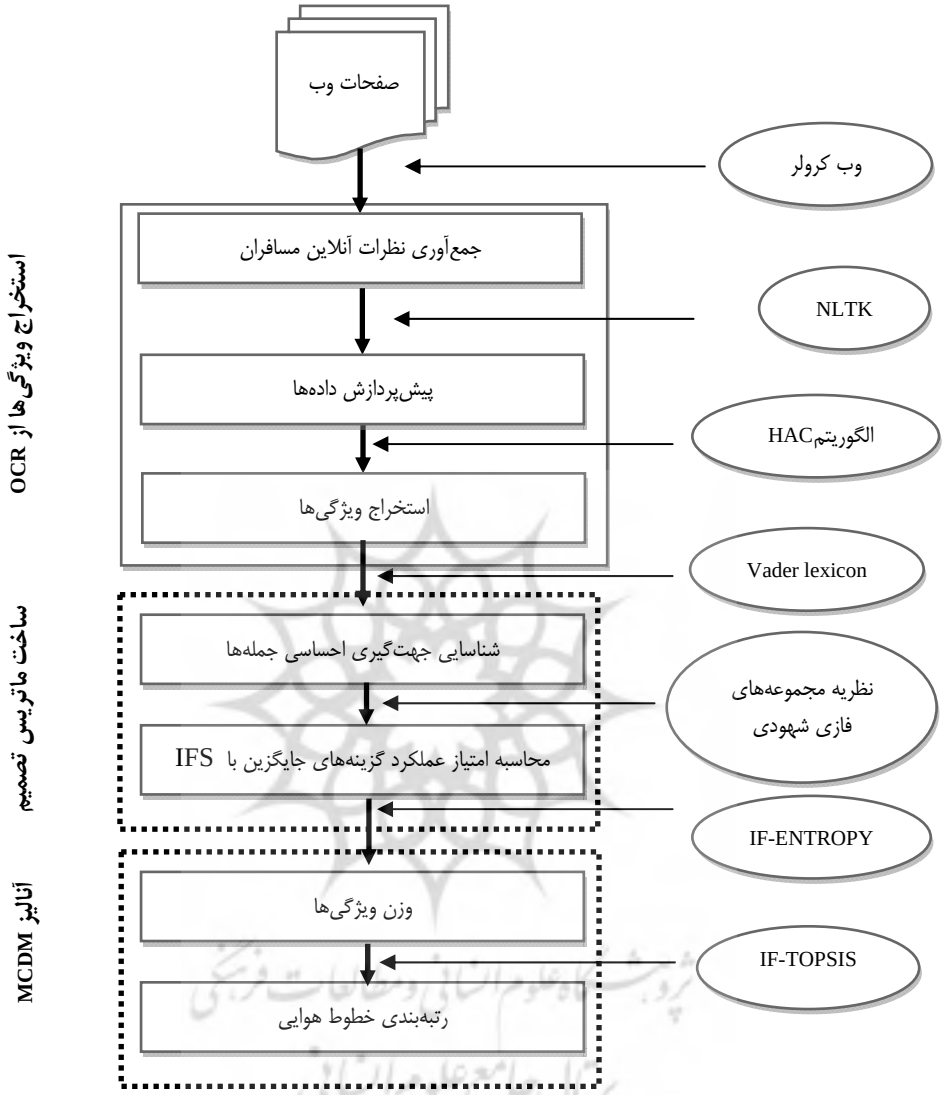
در مرحلهٔ اول، از ابزار خزندهٔ وب برای جمع‌آوری OCRها استفاده می‌شود. جعبه ابزار زبان طبیعی برای پیش‌پردازش OCRها مورد استفاده قرار می‌گیرد و از الگوریتم HAC ویژگی‌های خطوط هوایی استخراج می‌شوند.

جمع‌آوری داده

داده‌های مورد استفاده در این مطالعه از وبسایت SKYTRAX[2] استخراج شده است. SKYTRAX یک شرکت مشاوره مستقر در بریتانیاست که در صنعت هواپیمایی تخصص دارد. در این مطالعه ۸۰۱۰ نظر در وبسایت SKYTRAX روی خطوط هوایی منتخب خاورمیانه، شامل ۱۷ خط هوایی بین ۱ ژانویه ۲۰۱۵ و ۳۱ دسامبر ۲۰۲۱ جمع‌آوری شده است. این مطالعه از کتابخانه درخواست‌های پایتون برای خزیدن در سایت مد نظر و همچنین، از کتابخانه scrapy با توجه به

1. CRoss-Industry Standard Process for Data Mining
2. www.airlinequality.com

ساختار HTML وبسایت، برای تجزیه پاسخ وبسایت به‌دست‌آمده برای دریافت نظرهای کاربران و جزئیات سفر مسافران استفاده می‌کند.



شکل ۱. فرایند اجرای پژوهش

خراشیدن وب با خزندهٔ وب

در این مرحله از یک خزندهٔ وب برای استخراج داده‌های OCR استفاده می‌شود. خزندهٔ وب برنامه‌ای است که معمولاً صفحات وب را برای موتور جست‌وجوی وب بارگیری می‌کند. از آنجایی که تعداد زیادی خزنده برای OCRها توسعه داده شده است، این تحقیق از خراشنده وب[1] استفاده می‌کند که به‌عنوان یک برنامه افزودنی به کروم اضافه شده است و می‌تواند به‌راحتی داده‌های سایت مدنظر را استخراج کند(حیدری دهوئی و همکاران، ۲۰۲۱).

در این مطالعه با استفاده از کتابخانهٔ Scrapy در پایتون فرایند خراشیدن وب پیاده‌سازی شده است.

1. Web Scraper

پیش‌پردازش داده‌ها با استفاده از جعبهٔ ابزار زبان طبیعی

این مرحله در مجموع از ۹ فرایند برای آماده‌سازی متن برای عملیات استخراج ویژگی و تحلیل احساسات تشکیل شده است. این ۹ مرحله به شرح زیر است:

۱. تبدیل حروف بزرگ به کوچک[1]

یکی از رایج‌ترین مراحل پیش‌پردازش است که در آن متن به حروف مشابه ترجیحاً کوچک تبدیل می‌شود. کلمه‌هایی مانند Flight و flight معنی یکسانی دارند؛ اما وقتی به حروف کوچک تبدیل نمی‌شوند، این دو به‌عنوان دو کلمه متفاوت در مدل فضای برداری نمایش داده می‌شوند (در نتیجه ابعاد بیشتری دارند). این کار برای کاهش حجم واژگان داده‌های متنی انجام می‌شود.

۲. حذف فضاهای اضافی[2]

تمامی فضاهای اضافی و خالی بین کلمه‌ها و جمله‌ها در این مرحله حذف می‌شود. از تابع join و split برای حذف تمام فضاهای سفید در یک رشته استفاده می‌شود.

۳. توکن‌سازی[3]

توکن‌سازی فرایندی است که طی آن یک دادهٔ متنی معین به عناصر زبانی کوچک‌تر به نام نشانه‌ها تقسیم می‌شود. کلمه‌ها، اعداد، علائم نگارشی و موارد دیگر از جمله عناصر زبانی هستند که به‌عنوان نشانه شناخته می‌شوند (فرزادنیا و رئیسی وانانی[4]، ۲۰۲۲). در این مطالعه از ماژول word tokenize کتابخانه nltk.tokenize در پایتون استفاده شده است.

٤. حذف اعداد

گاهی به اعداد ذکرشده (مانند تاریخ پرواز، ساعت پرواز) در متن نیازی نیست؛ بنابراین چنین عناصری نیز از متن حذف می‌شود.

۵. حذف کردن کلمه‌های توقف[5]

کلمه‌های توقف کلمه‌هایی هستند که به‌طور مکرر در زبان وجود دارند؛ اما هیچ اطلاعات مهمی در ارتباط با معنای کلی متن ندارند. این کلمه‌ها معمولاً حروف ربط، معرف‌های اسم، افعال کمک‌کننده و حروف اضافه هستند (آگراوال، جاگاناتان و دهلی[6]، ۲۰۲۱). در این مطالعه از فهرست کلمه‌های توقف زبان انگلیسی ارائه‌شده در جعبه ابزار زبان طبیعی پایتون استفاده شده است.

1. Lower casing
2. Remove Extra Whitespaces
3. Tokenization
4. Farzadnia & Raeesi Vanani
5. Removing Stopwords
6. Agrawal, Jagannathan and Delhi

۶. حذف کلمه‌های پراکنده

گاهی لازم است کلمه‌های پراکنده از داده‌های متنی حذف شود. این کلمه‌ها شامل نام افراد، کشورها، شهرها، نام فرودگاه‌ها، نام خطوط هوایی و... می‌شود.

۷. حذف علائم نگارشی[1]

در این مرحله تمام علائم نگارشی از متن حذف می‌شود. پایتون یک کتابخانهٔ رشته‌ای دارد که یک ویژگی string.punctuation را شامل می‌شود و حاوی فهرست از پیش تعریف شده‌ای از علائم نگارشی است.

۸. بن‌واژه‌سازی[2]

این فرایند شبیه به بن‌سازی[3] است که هدف آن کاهش یک کلمه به شکل پایهٔ آن و گروه‌بندی اشکال مختلف یک کلمه یکسان است. بنابراین کلمه‌ها را با معنای مشابه با ریشهٔ آن‌ها استاندارد می‌کند. در این مطالعه بن‌واژه‌سازی به‌جای بن‌سازی انتخاب می‌شود؛ زیرا فرایند بن‌سازی همیشه درست نیست و به‌سادگی پیشوند یا پسوند یک کلمه را حذف می‌کند. درحالی‌که در واژه‌سازی، ریشه‌ها با لم‌های موجود در فرهنگ لغت تطبیق داده می‌شوند و ازاین‌رو نتایج دقیق‌تری به دست می‌آید.

۹. برچسب‌گذاری دستوری[4]

برچسب‌گذاری POS برای ارائهٔ اطلاعات مفید واژگانی با طبقه‌بندی کلمه‌ها در متن بر اساس موقعیت یک کلمه (اسم، فعل، قید و صفت) استفاده می‌شود. در مرحله استخراج ویژگی با استفاده از الگوریتم HAC مورد نیاز است.

استخراج ویژگی‌ها از OCR با استفاده از الگوریتم HAC

هدف از این مرحله شناسایی ویژگی‌های خطوط هوایی بر اساس OCR است (هر نظر مشتری منعکس‌کننده ترجیحات مصرف‌کننده است). در این پژوهش از الگوریتم HAC استفاده شده است که به‌جای استفاده از اصطلاح فرکانس، اسم و صفت را شناسایی می‌کند. الگوریتم HAC به‌جای استفاده از اصطلاح فرکانس، اسم و صفت را شناسایی می‌کند. این الگوریتم توسط ایریناکی، پیسال و سینگ[5] (۲۰۱۲) توسعه داده شده است. برای شناسایی ویژگی‌های بالقوه و به‌عنوان یک الگوریتم یادگیری ماشین با استفاده از رویکرد بدون نظارت در نظر گرفته می‌شود. ایدهٔ اصلی این الگوریتم آن است که اسامی‌ای که مشتریان دربارهٔ آن‌ها نظرهای متفاوتی دارند، احتمالاً مهم‌تر از ویژگی‌های متمایز هستند که کاربران چندان دربارهٔ آن‌ها اظهارنظر نمی‌کنند. در این الگوریتم به‌جای استفاده از اسم‌های تکراری، در ابتدا صفت‌ها و اسم‌ها را در سند شناسایی می‌کند و ابتدا به اسم‌ها نقطه صفر اختصاص می‌دهد. سپس با توجه به اینکه هر صفت به اسمی که نزدیک‌ترین اسم به آن است مربوط می‌شود و آن اسم را توصیف می‌کند، این الگوریتم صفت‌های نزدیک به اسم‌ها را

1. Removing Punctuations
2. Lemmatization
3. Stemming
4. Part of Speech
5. Eirinaki, Pisal & Singh

شناسایی می‌کند و در نتیجه امتیاز اسم‌ها را افزایش می‌دهد. پس از پردازش کل سند، الگوریتم به هر اسم امتیاز می‌دهد. اسامی با امتیاز بالاتر نشان می‌دهد که صفت‌های بیشتری آن‌ها را توصیف کرده‌اند و در نهایت برای فیلتر کردن اسم‌ها و شناسایی ویژگی‌های بالقوه می‌توان اسم‌هایی را انتخاب کرد که امتیازی بالاتر از یک آستانه مشخص دارند این آستانه پارامتری از الگوریتم است و می‌تواند بر اساس آزمایش‌ها و ارزیابی انسانی بر روی مجموعه‌های مختلف داده‌های بررسی انتخاب شود (ایریناکی و همکاران،۲۰۱۲).

مزایای الگوریتم HAC

از صفت‌ها اغلب برای بیان عقاید یا احساسات در مورد یک اسم استفاده می‌شود، بنابراین تعداد صفت بالا می‌تواند نشان‌دهنده یک متن با نظر یا احساسی‌تر باشد. تعداد زیاد صفت می‌تواند نشان‌دهندهٔ این باشد که نویسنده بر توصیف اسم‌ها تأکید دارد و متن را توصیفی‌تر و مفصل‌تر می‌کند. تعداد زیاد صفت می‌تواند نشان دهد که متن از نظر واژگان و کاربرد زبان پیچیده‌تر است. تعداد زیاد صفت می‌تواند نشان دهد که اسم‌هایی که توصیف می‌شوند، در متن مهم یا قابل توجه هستند؛ به‌طور خاص برای شناسایی و گروه‌بندی ویژگی‌ها طراحی شده است؛ قادر است روابط معنایی بین ویژگی‌ها را بهتر شناسایی کند؛ بدون نظارت عمل می‌کند و به داده‌های برچسب‌گذاری شده نیازی ندارد. نتایج HAC معمولاً قابل تفسیرتر هستند و کمک می‌کنند تا الگوهای مشترک شناسایی شوند. این مزایا HAC را برای استخراج ویژگی‌ها از نظرهای آنلاین مناسب می‌سازد (ایریناکی و همکاران، ۲۰۱۲).

ساخت ماتریس تصمیم بر اساس ویژگی‌ها

در این مرحله، برای تشکیل ماتریس تصمیم با استفاده از ابزارVader Lexicon، ابتدا جهت‌گیری هر جمله بر اساس هر ویژگی خطوط هوایی تعیین می‌شود و تعداد جمله‌های مثبت، منفی و خنثی به ازای هر ویژگی محاسبه می‌شود و سپس این مقادیر به اعداد IF تبدیل می‌شود. شایان ذکر است که ابزارهای دیگری جهت تحلیل احساسات مبتنی بر واژگان مانند NLTK و TextBlob نیز وجود دارد که به داده‌های آموزشی نیاز دارند و عملکردشان در شرایط خاص (مانند متون غیررسمی یا غیرساختاریافته) ممکن است ناپایدار باشد. در حالی که VADER بدون نیاز به آموزش قبلی و با استفاده از واژگان معتبر به‌طور خاص برای تحلیل احساسات در رسانه‌های اجتماعی طراحی شده است. با این توضیحات، می‌توان نتیجه‌گیری کرد که VADER به‌دلیل کارایی و دقت در تحلیل احساسات در متون رسانه‌های اجتماعی، بهترین گزینه برای این تحقیق بوده است.

شناسایی جهت‌گیری احساسی جمله‌ها دربارهٔ هر ویژگی با استفاده از واژگان Vader

با توجه به مزایای روش VADER و نکات ذکرشده در این مطالعه برای تحلیل احساسات مبتنی بر واژگان از این ابزار استفاده می‌شود.

محاسبه امتیاز عملکرد گزینه‌های جایگزین با IFS

IFS به‌دلیل قابلیت آن در مدل‌سازی هم‌زمان درجات عضویت، عدم عضویت و میزان تردید، برای مدیریت ارزیابی‌های

مبتنی بر نظرهای خبرگان ابزاری بسیار قوی است (نصرتی ملک جهانی، حسین‌زاده کاشان و سجادی[1]، ۲۰۲۴). IFS امکان می‌دهد که علاوه‌بر ارزیابی کیفیت خدمات، عدم قطعیت‌های مربوط به هر ارزیابی نیز به‌طور مستقیم مدل‌سازی شود. این ویژگی به‌ویژه در مسائل رتبه‌بندی و تصمیم‌گیری چندمعیاره که با داده‌های غیرقطعی سر و کار دارند، اهمیت دارد (حسین، احمد و علم[2]، ۲۰۱۲).

فرض کنید $X = \{x_1, x_2, \dots, x_n\}$ است.

$X \to [0,1]$ و بردار وزن $w = (w_1, w_2, \dots, w_n)^T$ به ازای هر عنصر $x_j (j = 1,2, \dots, n)$ که $w_j \geq 0$ و $\sum_{j=1}^{n} w_j = 1$ است.

A یک مجموعه فازی است که با تابع عضویت μ_A مشخص می‌شود (زاده[3]، ۱۹۶۵).

رابطهٔ ۱) $$A = \{\langle x_j, \mu_A(x_j)\rangle \mid x_j \in X\}$$

که در آن $\mu_A(x_j)$ درجه عضویت عنصر x_j را در مجموعه A نشان می‌دهد. آتاناسوف[4] (۱۹۸۶) یک مجموعه فازی تعمیم‌یافته به نام IFS را معرفی کرد که به‌صورت زیر نشان داده شده است. A یک IFS در X به شکل زیر است:

رابطهٔ ۲) $$A = \{\langle x_j, \mu_A(x_j), v_A(x_j)\rangle \mid x_j \in X\}$$

نظریهٔ IFS توسعه‌یافته توسط آتاناسوف (۱۹۸۶) ابزار مناسبی را برای مقابله با ابهام و تردیدهایی که معمولاً در تصمیم‌گیری‌های زندگی واقعی با آن مواجه می‌شوند، ارائه می‌دهد که با یک تابع عضویت μ_A و یک تابع غیرعضو v_A مشخص می‌شود، که در آن:

رابطهٔ ۳) $$\mu_A: X \to [0,1], \qquad \&\&x_j \in X \to \mu_A(x_j) \in [0,1]$$

رابطهٔ ۴) $$v_A: X \to [0,1], \qquad \&\&x_j \in X \to v_A(x_j) \in [0,1]$$

با شرط:

رابطهٔ ۵) $$\mu_A(x_j) + v_A(x_j) \leq 1, \text{ for all } x_j \in X$$

برای هر A مجموعه فازی شهودی در X، اگر

رابطهٔ ۶) $$\pi_A(x_j) = 1 - \mu_A(x_j) + v_A(x_j)$$

سپس $\pi_A(x_j)$ درجه نامعین بودن x_j به A نامیده می‌شود. به خصوص اگر:

رابطهٔ ۷) $$\pi_A(x_j) = 1 - \mu_A(x_j) + v_A(x_j) = 0 \quad \text{for each } x_j \in X$$

IFNها[5] با درجه عضویت و درجه عدم عضویت یک عنصر، که اجزای اصلی یک IFS هستند، تعریف می‌شوند.

1. Nosrati Malekjahan, Husseinzadeh Kashan & Sajadi
2. Husain, Ahmad and Alam
3. Zadeh
4. Atanassov
5. Intuitive fuzzy numbers

امتیاز عملکرد هر جایگزین با توجه به هر معیار را می‌توان با توجه به IFNها نشان داد (چالی و بالامان، ۲۰۱۹):

- درصد جمله‌های مثبت مربوط به هر معیار از هر جایگزین به‌عنوان بازنمایی رضایت مشتری
- درصد جمله‌های منفی مربوط به هر معیار از هر جایگزین به‌عنوان نمایش نارضایتی مشتری
- درصد جمله‌های خنثی مربوط به هر معیار از هر جایگزین به‌عنوان نماینده درجه تردید
- μ_{ij} :نشان‌دهنده میزان رضایت از جایگزین i با توجه به معیار j ام است.
- ϑ_{ij} :نشان‌دهنده درجه نارضایتی از جایگزین i با توجه به معیار j ام است.
- π_{ij} :درجه تردید جایگزین i را با توجه به معیار j ام نشان می‌دهد.

فرمول‌بندی برای محاسبه این شاخص‌ها به شرح زیر است:

رابطه ۸)
$$\mu_{ij} = \frac{T_{ij}^{pos}}{T_{ij}^{pos} + T_{ij}^{neg} + T_{ij}^{neu}}$$

رابطه ۹)
$$\nu_{\mathrm{ij}} = \frac{T_{ij}^{neg}}{T_{ij}^{pos} + T_{ij}^{neg} + T_{ij}^{neu}}$$

رابطه ۱۰)
$$\pi_{\mathrm{ij}} = \frac{T_{ij}^{\mathrm{neu}}}{T^{pos} + T^{neg} + T^{neu}}$$

رابطه ۱۱)
$$\mu_{ij} + \nu_{\mathrm{ij}} + \pi_{\mathrm{ij}} = 1\,, i = 1,2, \dots, m\,, j = 1,2, \dots, n$$

T_{ij}^{pos} تعداد جمله‌های مثبت در مورد معیار j در بررسی i ام؛ T_{ij}^{neg} تعداد جمله‌های منفی در مورد معیار j در بررسی iام؛ T_{ij}^{neu} تعداد جمله‌های خنثی در مورد معیار j در بررسی i ام.

ENTROPY و TOPSIS فازی شهودی

۱. ایجاد ماتریس تصمیم مجموعه فازی شهودی

فرض کنید $C = \{C_1, C, \dots, C_n\}$ مجموعه معیارها و $A = \{A_1, A_2, \dots, A_m\}$ مجموعه‌ای از گزینه‌ها باشد. نمره‌های عملکرد مربوط به هر گزینه A_i با توجه به هر معیار C_j را می‌توان با یک $IFN(\mu_{ij}, v_{ij})$ مشخص کرد، که μ_{ij} نشان‌دهندهٔ درجه گزینه A_i که معیار C_j را برآورده می‌کند، v_{ij} نشان‌دهندهٔ درجهٔ گزینهٔ A_i که معیار C_j را برآورده نمی‌کند و $0 \le \mu_{ij} + v_{ij} \le 1, \mathrm{i} = 1,2,\dots,\mathrm{m}, \mathrm{j} = 1,2,\dots,\mathrm{n}, \pi_{ij} = 1 - \mu_{ij} + v_{ij}$ نشانگر درجه تردید است (چالی و بالامان، ۲۰۱۹).

ماتریس تصمیم فازی شهود "IFN "Z هایی را گرد هم می‌آورد که نمره‌های عملکرد هر گزینه را بر اساس هر معیار بیان می‌کند (محاسبهٔ آن قبلاً توضیح داده شده است). ردیف‌های ماتریس تصمیم هر گزینه را نشان می‌دهد، در حالی‌که ستون‌های ماتریس تصمیم هر معیار را نشان می‌دهد. به‌طور مثال، $z_{11} = (\mu_{11}, v_{11}, \pi_{11})$ نشان‌دهندهٔ نمرهٔ عملکرد A_1 مطابق با C_1 است (چالی و بالامان، ۲۰۱۹).

$$Z=\begin{bmatrix} z_{11} & \cdots & z_{1n} \\ \vdots & \ddots & \vdots \\ z_{m1} & \cdots & z_{mn} \end{bmatrix}=\begin{bmatrix} (\mu_{11},v_{11},\pi_{11}) & \cdots & (\mu_{1n},v_{1n},\pi_{1n}) \\ \vdots & \ddots & \vdots \\ (\mu_{m1},v_{m1},\pi_{m1}) & \cdots & (\mu_{mn},v_{mn},\pi_{mn}) \end{bmatrix}$$

۲. محاسبهٔ وزن معیارها با استفاده از روش آنتروپی فازی شهودی

مفهوم آنتروپی در نظریه اطلاعات توسط شانون[1] (۱۹۴۸) به‌عنوان معیار عدم قطعیت در اطلاعات بر اساس نظریه احتمال پیشنهاد شده است. زلنی[2] (۱۹۷۶) اندازه‌گیری آنتروپی شانون را به‌عنوان ابزاری مناسب برای تعیین اهمیت نسبی یک معیار معرفی کرد؛ زیرا امکان اندازه‌گیری شدت کنتراست نسبی معیارها را برای نمایش میانگین اطلاعات ذاتی ارسال‌شده به تصمیم‌گیرنده فراهم می‌کند. اندازه‌گیری آنتروپی بالاتر به این معنی است که تصمیم‌گیرنده مقادیر ارزیابی مشابهی را برای هر گزینه بر اساس معیار تعیین می‌کند. اندازه‌گیری آنتروپی معیار سطح اطلاعات منتقل‌شده توسط معیار را نشان می‌دهد. هرچه معیار آنتروپی کمتر باشد، اهمیت نسبی بیشتری دارد. از معیار IF-Entropy توسعه‌یافته توسط ولاخوس و سرگیادیس[3] (۲۰۰۷) برای محاسبه وزن معیارها استفاده می‌شود که در زیر توضیح داده شده است (هانگ و چن[4]، ۲۰۰۹).

برای ساخت ماتریس تصمیم فازی شهودی $z=[z_{ij}]_{m*n}$ اندازهٔ آنتروپی معیار C_j توسط رابطهٔ ۱۲ قابل محاسبه است. هر معیار C_j به‌عنوان یک IFS در نظر گرفته می‌شود.

$$E_{LT}^{IFS}(C_j)=-\frac{1}{m\ln 2}\sum_{i=1}^{m}\left[\mu_{ij}\ln\mu_{ij}+v_{ij}\ln v_{ij}-(1-\pi_{ij})\ln(1-\pi_{ij})-\pi_{ij}\ln 2\right] \quad \text{رابطه ۱۲)}$$

به‌طوری‌که $j = 1,2,\ldots,n$

اگر $\mu_{ij}=0, v_{ij}=0, \pi_{ij}=1$ پس:

$$\mu_{ij}\ln\mu_{ij}=0, v_{ij}\ln v_{ij}=0, (1-\pi_{ij})\ln(1-v_{ij})=0 \quad \text{رابطه ۱۳)}$$

به‌این‌ترتیب، $E_{LT}^{IFS}(C_j)$ شامل درجه دودلی و درجه فازی C_j IFS است. به عبارت دیگر، هر معیار C_j به‌عنوان یک IFS در نظر گرفته می‌شود.

مقدار آنتروپی محاسبه‌شده توسط رابطه ۱۲ مبهم بودن و شهودگرایی اطلاعات منتقل‌شده توسط هر معیار C_j را اندازه‌گیری می‌کند.

ولاخوس و سرگیادیس (۲۰۰۷) توضیح دادند که چگونه معادله آنتروپی IF توسعه می‌یابد. سپس، درجه واگرایی میانگین اطلاعات ذاتی ارائه‌شده توسط نمرات عملکرد مربوطه بر اساس معیار C_j را می‌توان با استفاده از رابطه ۱۴ محاسبه کرد.

1. Shannon
2. Zeleny
3.Vlachos and Sergiadis
4. Hung and Chen

رابطه ۱۴) $$d_j = 1 - E_{LT}^{IFS}(C_j), \qquad j = 1,2,\dots,n$$

در نهایت، وزن آنتروپی معیارها بر اساس مقدار آن محاسبه می‌شود که نشان‌دهندهٔ شدت کنتراست ذاتی معیار C_j است.

رابطه ۱۵) $$w_j = \frac{d_j}{\sum_{j=1}^{n} d_j}$$

که در آن w_jوزن آنتروپی معیار C_j و را نشان می‌دهد و $\sum_{j=1}^{n} w_j = 1$.

۳. تعیین راه‌حل ایدئال مثبت (بهترین راه‌حل) و راه‌حل ایده آل منفی (بدترین راه‌حل)

A^+ : راه‌حل ایدئال مثبت مجموعه فازی شهودی

A^- : راه‌حل ایدئال منفی مجموعه فازی شهودی

رابطه ۱۶) $$A^+ = (\langle \mu_1^+, v_1^+, \pi_1^+ \rangle, \langle \mu_2^+, v_2^+, \pi_2^+ \rangle, \cdots, \langle \mu_n^+, v_n^+, \pi_n^+ \rangle)$$

رابطه ۱۷) $$A^- = (\langle \mu_1^-, v_1^-, \pi_1^- \rangle, \langle \mu_2^-, v_2^-, \pi_2^- \rangle, \cdots, \langle \mu_n^-, v_n^-, \pi_n^- \rangle)$$

که در آن:

رابطه ۱۸) $$\mu_j^+ = \max_{1 \le i \le m} \{\mu_{ij}\}$$

رابطه ۱۹) $$\mu_j^- = \min_{1 \le i \le m} \{\mu_{ij}\}$$

رابطه ۲۰) $$\gamma_j^+ = \min_{1 \le i \le m} \{v_{ij}\}$$

رابطه ۲۱) $$\gamma_j^- = \max_{1 \le i \le m} \{v_{ij}\}$$

رابطه ۲۲) $$\pi_j^+ = 1 - \mu_j^+ - v_j^+$$

رابطه ۲۳) $$\pi_j^- = 1 - \mu_j^- - v_j^-$$

٤. محاسبه فاصلهٔ نرمال‌شدهٔ اقلیدسی وزنی هر طرح $\mathbf{A_i}$ تا راه‌حل ایدئال مثبت و منفی مجموعه فازی شهودی

رابطه ۲۴) $$d(A_i, A^+) = \sqrt{\frac{1}{2}\sum_{j=1}^{n} \omega_j \left[\left(\mu_{ij} - \mu_j^+\right)^2 + \left(v_{ij} - v_{ij}^+\right)^2 + \left(\pi_{ij} - \pi_j^+\right)^2 \right]}$$

رابطه ۲۵) $$d(A_i, A^-) = \sqrt{\frac{1}{2}\sum_{j=1}^{n} \omega_j \left[\left(\mu_{ij} - \mu_j^-\right)^2 + \left(v_{ij} - v_{ij}^-\right)^2 + \left(\pi_{ij} - \pi_j^-\right)^2 \right]}$$

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

۵. محاسبه درجه نزدیکی نسبی طرح Ai به راه‌حل ایدئال مثبت مجموعه فازی شهودی

$$C_i = \frac{d(A_i, A^-)}{d(A_i, A^+) + d(A_i, A^-)}, i = 1,2, \dots, m \qquad 0 \le C_i \le 1$$ رابطه (۲۶)

بر اساس درجه نزدیکی نسبی طرح A_i به راه‌حل ایدئال مثبت مجموعه فازی شهودی مرتب می‌شود. C_i هر چه بزرگ‌تر باشد، طرح A_i نزدیک‌تر به راه‌حل ایدئال مثبت است؛ بنابراین طرح بهتر است و برعکس.

یافته‌های پژوهش

ویژگی‌های استخراج شده از OCR

همان‌طور که گفته شد، در این مرحله OCRها از وبسایت SKYTRAX برای رتبه‌بندی خطوط هوایی منتخب خاورمیانه جمع‌آوری می‌شود. در جدول ۲ نام هر خط هوایی به همراه کد و کشور مربوط به آن نشان داده شده است.

جدول ۲. نام خطوط هوایی به همراه کد و کشور مربوطه

کشور	نام خط هوایی	کد
عربستان سعودی	Saudi Arabian Airlines	$B_۱$
کویت	Kuwait Airways	$B_۲$
عمان	Oman Air	$B_۳$
ایران	Iran Air	$B_۴$
مصر	Egyptair	$B_۵$
اردن	Royal Jordanian Airlines	$B_۶$
لبنان	Middle East Airlines	$B_۷$
ترکیه	Pegasus Airlines	$B_۸$
امارات متحده عربی	flydubai	$B_۹$
امارات متحده عربی	Air Arabia	$B_{۱۰}$

خراشیدن وب با scrapy

در این مرحله، OCRهای مربوط به هر خط هوایی با استفاده از کتابخانه scrapy و بر اساس ساختار HTML وبسایت SKYTRAX در پایتون در قالب یک فایل json ذخیره شده است و سپس تمامی فایل‌های json به فرمت csv تبدیل شده است. شایان ذکر است تمام نظرهایی که به زبان‌های دیگری جز انگلیسی نوشته شده بودند، در این مرحله حذف شدند.

پیش‌پردازش داده‌ها با استفاده از جعبهٔ ابزار زبان طبیعی

در این پژوهش از پایتون و کتابخانه NLTK برای پیش‌پردازش متن OCRها استفاده شده است. مراحل پیش‌پردازش شامل تبدیل متن به حروف کوچک، حذف فضاهای اضافی، توکن‌سازی، حذف اعداد، حذف کلمه‌های توقف، حذف

کلمه‌های پراکنده، یافتن ریشهٔ کلمه‌ها، شکستن متن به جمله‌ها، برچسب‌گذاری نقش دستوری و برچسب‌گـذاری POS[1] برای هر جمله با استفاده از ماژول‌های فرعی sent_tokenize و pos_tag در کتابخانهٔ NLTK است. این مراحل بـرای آماده‌سازی متن برای مرحله بعدی فرایند پژوهش انجام می‌شود.

استخراج ویژگی‌ها از OCRها با استفاده از الگوریتم HAC

در این مرحله ویژگی‌های خطوط هوایی موجود با استفاده از الگوریتم HAC استخراج شده است. جـدول ۳ ویژگـی‌هـای استخراج‌شده توسط الگوریتم HAC با در نظر گرفتن threshold = ۴۰۰ را نشان می‌دهد (تنهـا اسـم‌هـایی کـه تعـداد صفات آن‌ها بیشتر از ۴۰۰ عدد است در نظر گرفته شده‌اند) که به‌صورت نزولی با توجه به امتیـاز هـر یـک از ویژگـی‌هـا مرتب شده‌اند.

جدول ۳. ویژگی‌های استخراج‌شده توسط الگوریتم HAC

ویژگی	امتیاز	ویژگی	امتیاز
flight	۷۲۸۴	leg	۷۲۶
seat	۳۹۲۵	lounge	۷۰۷
time	۳۱۴۴	customer	۶۸۴
service	۲۶۱۲	return	۶۴۹
staff	۲۴۳۵	check	۶۲۷
food	۱۹۴۵	thing	۵۷۲
experience	۱۷۹۲	selection	۵۲۲
airline	۱۵۷۸	trip	۴۹۹
hour	۱۵۶۹	choice	۴۹۱
meal	۱۳۵۰	people	۴۸۹
airport	۱۳۰۴	year	۴۸۲
cabin	۱۰۵۹	class	۴۶۵
passenger	۹۹۷	quality	۴۶۵
business	۹۶۷	hotel	۴۶۴
entertainment	۹۳۰	drink	۴۵۸
economy	۸۵۶	luggage	۴۵۰
crew	۸۱۲	price	۴۱۶
planes	۸۰۵	ticket	۴۰۳

همان‌طور که از جدول ۳ مشخص است کلمه‌هایی ماننـد flight، seat، time، sevice وstaff بیشـترین امتیـاز را دارند. همچنین تعدادی از این ویژگی‌ها با بررسی نظرهای مسافران به‌منظور آن‌ها پی برده شده است (برای مثـال کلمـه

1. POS TAGING

leg که با بررسی نظرهای مسافران متوجه شدیم به Iegroom[1] اشاره می‌کند، کلمه check به دو موضوع Check-in[2] و security check[3] برمی‌گردد، کلمه time به دو موضوع Flight delay و waiting time می‌پردازد و کلمه Class به دو موضوع Business Class و economy Class اشاره می‌کند). در نهایت با توجه به ویژگی‌های به‌دست آمده توسط الگوریتم HAC، بررسی نظرهای خطوط هوایی و توجه به‌منظور کلمه‌ها و با کمک نظرهای خبرگان ویژگی‌های نهایی در جدول ۴ همراه با کد در نظر گرفته‌شده، نشان داده شده است.

پس از استخراج ویژگی‌ها با HAC، گروهی از خبرگان صنعت هوایی و بازاریابی آن‌ها را از نظر کاربردی و اهمیت تأیید و نهایی کردند تا اطمینان حاصل شود که ویژگی‌ها معیارهای مهم مشتریان را به‌درستی منعکس می‌کنند. از دو نفر خبرگان صنعت هوایی با حداقل ۵ سال تجربه استفاده شد. انتخاب این خبرگان بر اساس تجربه مرتبط آن‌ها در ارزیابی خدمات مشتری و تحلیل داده‌های کیفی بود.

جدول٤. ویژگی‌های نهایی

کد	ویژگی	معادل فارسی
C_1	ticket price	قیمت بلیت
C_2	food/meal	غذا / وعده غذایی
C_3	customer service	خدمات مشتری
C_4	seat	صندلی
C_5	staff	کارکنان
C_6	crew	خدمه
C_7	entertainment	سرگرمی
C_8	check-in	فرایند پذیرش
C_9	luggage/baggage	بار مسافر
C_{10}	Airport lounge	سالن استراحت فرودگاه
C_{11}	drink	نوشیدنی
C_{12}	legroom	فضای پا
C_{13}	Flight delay	تأخیر پرواز
C_{14}	waiting time	زمان انتظار
C_{15}	Business Class	کلاس تجاری
C_{16}	economy Class	کلاس اقتصادی
C_{17}	security check	بازرسی امنیتی

ویژگی‌های نشان داده شده در جدول ۵ همه عواملی هستند که می‌توانند هنگام ارزیابی یک خط هوایی در نظر گرفته شوند.

۱. فضای پا
۲. فرایند پذیرش
۳. بازرسی امنیتی

«check-in» به فرایند پذیرش اشاره دارد. «بار مسافر» به فرایند چک کردن یا حمل چمدان، مقدار مجاز چمدان، هزینهٔ چمدان‌های تحویل داده‌شده و دستی و جابه‌جایی چمدان‌های گم‌شده یا آسیب‌دیده اطلاق می‌شود. «سالن فرودگاه» به یک منطقه VIP در دسترس مسافران خاص اشاره دارد. «فضای پا» به فضای موجود برای پاهای مسافر اشاره دارد. «کلاس تجاری» و «کلاس اقتصادی» به سطوح مختلف خدمات و امکانات موجود در یک پرواز اشاره دارد. منظور از «بازرسی امنیتی» اقدامات ایمنی و امنیتی است که در طول پرواز و در فرودگاه وجود دارد.

ساخت ماتریس تصمیم بر اساس ویژگی‌ها

حال با توجه به ویژگی‌های نهایی ذکر شده در مرحله قبل، ماتریس تصمیم ساخته می‌شود که مراحل ایجاد ماتریس تصمیم در ادامه به‌طور کامل شرح داده می‌شود.

شناسایی جهت‌گیری احساسی جمله‌ها دربارهٔ هر ویژگی با استفاده از Vader Lexicon

Vader-Lexicon یک ابزار تحلیل احساسات مبتنی بر قانون است که برای تعیین جهت احساسات ویژگی‌های محصول در جمله‌هایی از OCR که شامل ویژگی‌های تعریف‌شده هستند، استفاده می‌شود. برای بهبود نتایج، قوانین خاصی برای رسیدگی به موارد خاص تعریف شد. برای مثال، اگر یک جمله شامل یک حرف ربط مانند «but»، «although» و «though» باشد، احساس در هر طرف از حرف ربط جدا تلقی می‌شود. علاوه‌براین، هنگامی که کلمه‌های منفی بدون صفت ظاهر می‌شوند، احساس به‌درستی شناسایی نمی‌شود، بنابراین قانونی تعریف شد که تعداد این جمله‌ها را از تعداد جمله‌های خنثی کم کرده و آن‌ها را به تعداد جمله‌های منفی اضافه می‌کند. قاعدهٔ دیگری برای تغییر جهت احساسات تعریف شد، بدین‌صورت که اگر حروف ربط خاصی مانند «exclude»، «except» و «wish» در جمله وجود داشت، می‌بایست برای احساس جمله تغییر کند (حیدری دهوئی و همکاران، ۲۰۲۱).

مثال:

The food was bad but the staff was great. غذا بد بود؛ اما کارکنان عالی بودند.

در این جمله جهت احساسات در دو طرف کلمه «اما» متفاوت است، پس باید به دو جملهٔ جداگانه، یعنی «غذا خوب بود» و «کارکنان عالی بودند» تبدیل شود و جهت احساس برای هر جمله جداگانه تعیین می‌شود.

The customer service number is never reachable, when it rings nobody picks up.

شمارهٔ خدمات مشتری هرگز قابل دسترسی نیست، وقتی زنگ می‌خورد کسی آن را برنمی‌دارد.

این جمله به‌صورت خنثی شناخته می‌شود؛ زیرا جمله بدون صفت است و Vader-Lexicon قادر به تشخیص درست احساسات نیست. پس تعداد این‌گونه جمله‌ها محاسبه می‌شود و از تعداد جمله‌های خنثی کم و به تعداد جمله‌های منفی اضافه می‌شود.

Except the ground staffs of Istanbul airport everything was good.

به‌جز کارکنان زمینی فرودگاه استانبول همه‌چیز خوب بود.

در این جمله ازآنجایی‌که صفت «خوب» در جمله وجود دارد، به اشتباه با وجود کلمه «به‌جز» آن را نسبت به

کارکنان در نظر می‌گیرد، در حالی‌که احساس جمله نسبت به کارکنان منفی است؛ پس قاعده به این صورت تعریف شد که برای جمله‌های حاوی ویژگی اگر این‌گونه کلمه‌ها در جمله‌ها دیده شد، احساس جمله را برعکس کند.

در نهایت، تعداد جمله‌هایی که احساسات مثبت، منفی و خنثی در مورد هر ویژگی خطوط هوایی بر اساس این قوانین دارند، در جدول ۵ نشان داده شده است.

جدول ۵. تعداد جمله‌های حاوی احساسات مثبت، منفی و خنثی در مورد هر ویژگی خطوط هوایی

C_5			C_4						C_2			C_1			
T_{ij}^{neu}	T_{ij}^{neg}	T_{ij}^{pos}	T_{ij}^{neu}	T_{ij}^{neg}	T_{ij}^{pos}	T_{ij}^{neu}	T_{ij}^{neg}	T_{ij}^{pos}	T_{ij}^{neu}	T_{ij}^{neg}	T_{ij}^{pos}	T_{ij}^{neu}	T_{ij}^{neg}	T_{ij}^{pos}	
۱۲	۴۷	۳۲	۲۵	۴۴	۵۷	۱	۱۴	۳	۴۲	۳۷	۹۴	۸	۳	۵	B_1
۷	۶۲	۳۰	۲۲	۴۴	۴۸	۲	۱۵	۶	۲۰	۴۲	۶۹	۳	۱۰	۱۱	B_2
۱۸	۵۱	۶۷	۴۵	۴۶	۹۹	۳	۲۸	۱۱	۲۴	۶۲	۱۳۹	۷	۱۰	۲۱	B_3
۴	۳	۱۶	۹	۵	۱۲	۰	۲	۰	۹	۹	۳۱	۰	۰	۱	B_4
۷	۲۳	۳۵	۳۱	۵۶	۶۲	۹	۱۲	۶	۱۶	۴۹	۷۷	۲	۵	۱۴	B_5
۱۲	۵۵	۲۹	۲۴	۴۰	۵۱	۰	۱۸	۴	۲۹	۳۹	۶۶	۲	۳	۹	B_6
۳	۷	۱۷	۷	۱۹	۳۲	۰	۱	۰	۱۹	۵	۴۶	۲	۱	۳	B_7
۱۳	۷۴	۵۱	۲۲	۴۹	۴۷	۰	۳۲	۷	۲۸	۳۴	۳۱	۸	۱۴	۲۲	B_8
۱۵	۵۲	۲۵	۴۳	۶۲	۴۲	۰	۲۴	۵	۲۸	۳۶	۴۹	۱۰	۱۵	۱۸	B_9
۶	۱۹	۱۱	۴	۷	۱۳	۱	۱۷	۱	۳	۱۳	۱۰	۱	۲	۲	B_{10}
c_{10}			c_9			c_8			c_7			c_6			
T_{ij}^{neu}	T_{ij}^{neg}	T_{ij}^{pos}	T_{ij}^{neu}	T_{ij}^{neg}	T_{ij}^{pos}	T_{ij}^{neu}	T_{ij}^{neg}	T_{ij}^{pos}	T_{ij}^{neu}	T_{ij}^{neg}	T_{ij}^{pos}	T_{ij}^{neu}	T_{ij}^{neg}	T_{ij}^{pos}	
۷	۹	۱۷	۱۹	۴۰	۱۷	۱۰	۱۷	۲۲	۰	۲۰	۵۲	۱۰	۲۶	۵۰	B_1
۳	۴	۴	۱۵	۳	۱۷	۱۵	۲۱	۲۲	۰	۱۹	۴۸	۱۰	۳۵	۳۷	B_2
۱۱	۵	۳۴	۱۱	۱۴	۷	۲۴	۴۱	۵۲	۰	۹	۶۰	۱۷	۳۳	۷۶	B_3
۵	۱	۱	۱	۰	۶	۲	۲	۵	۰	۴	۱۶	۷	۹	۱۸	B_4
۶	۲۰	۱۹	۱۷	۲۰	۲۰	۱۵	۲۱	۳۵	۰	۱۰	۳۳	۱۳	۳۴	۳۹	B_5
۱۵	۹	۱۷	۱۴	۲۸	۱۱	۸	۳۳	۲۵	۰	۷	۲۸	۱۶	۱۷	۴۱	B_6
۱	۱	۳	۱۱	۸	۶	۷	۱۱	۱۴	۰	۳	۲۱	۳	۸	۳۶	B_7
۰	۱	۱	۳۰	۶۶	۲۹	۴۰	۶۱	۳۳	۳	۳	۸	۱۱	۳۹	۲۹	B_8
۵	۱	۵	۲۲	۴۵	۲۰	۳۲	۴۲	۳۶	۰	۱۷	۳۰	۱۲	۳۹	۴۰	B_9
۰	۱	۱	۱۰	۱۹	۲۰	۹	۱۰	۷	۰	۱	۴	۰	۷	۹	B_{10}
c_{15}			c_{14}			c_{13}			c_{12}			c_{11}			
T_{ij}^{neu}	T_{ij}^{neg}	T_{ij}^{pos}	T_{ij}^{neu}	T_{ij}^{neg}	T_{ij}^{pos}	T_{ij}^{neu}	T_{ij}^{neg}	T_{ij}^{pos}	T_{ij}^{neu}	T_{ij}^{neg}	T_{ij}^{pos}	T_{ij}^{neu}	T_{ij}^{neg}	T_{ij}^{pos}	
۸	۱۴	۱۳	۱۶	۲۰	۷	۳	۵۱	۷	۱	۱	۴	۸	۲	۹	B_1
۸	۹	۸	۹۶	۱۸۶	۱۲۲	۰	۹۹	۱۱	۶	۰	۹	۶	۳	۶	B_2
۲۰	۱۲	۴۳	۲۴	۱۶	۱۳	۲	۵۷	۴	۵	۳	۱۱	۱۲	۱۸	۲۲	B_3
۲	۱	۴	۰	۱	۲	۱	۱۱	۴	۲	۰	۱	۰	۰	۷	B_4
۲۱	۲۰	۲۴	۸	۱۶	۸	۰	۳۸	۶	۲	۱	۸	۷	۶	۹	B_5
۹	۸	۱۳	۱۷	۱۵	۱۱	۱	۶۹	۹	۱	۰	۳	۱۴	۱۲	۱۵	B_6
۵	۳	۵	۴	۱	۱	۰	۹	۴	۱	۰	۴	۵	۰	۶	B_7
۰	۱	۰	۲۳	۴۱	۱۳	۲	۱۰۶	۱۴	۲	۲	۳	۱۰	۱۸	۱۸	B_8
۱۱	۱۱	۱۱	۲۵	۳۰	۷	۱	۷۲	۷	۵	۷	۸	۴	۹	۱۰	B_9
۰	۰	۱	۵	۹	۸	۱	۲۷	۴	۰	۰	۴	۲	۴	۲	B_{10}

	c_{16}			c_{17}		
	T_{ij}^{pos}	T_{ij}^{neg}	T_{ij}^{neu}	T_{ij}^{pos}	T_{ij}^{neg}	T_{ij}^{neu}
B_1	۶	۱	۲	۱۱	۲	۰
B_2	۲	۰	۱	۲۲	۵	۰
B_3	۳	۱	۲	۱۱	۵	۰
B_4	۲	۰	۰	۰	۱	۰
B_5	۲	۰	۲	۱۵	۱	۱
B_6	۴	۰	۱	۱۳	۴	۰
B_7	۲	۰	۰	۱	۰	۰
B_8	۰	۱	۰	۳	۳	۱
B_9	۵	۳	۲	۵	۱	۰
B_{10}	۰	۱	۰	۴	۱	۰

محاسبه امتیاز عملکرد گزینه‌های جایگزین با IFS

پس از تعیین جهت‌گیری احساسات مسافر در مورد هر ویژگی خطوط هوایی، اعداد IF بر اساس تعداد جهت‌گیری‌های احساسی محاسبه می‌شوند. سپس ماتریس تصمیم‌گیری بر اساس اعداد IF تشکیل می‌شود که امتیاز عملکرد هر خط هوایی را بر اساس هر ویژگی نشان می‌دهد. بنابراین، با تبدیل جهت‌گیری‌های احساسات به اعداد IF، که اجزای اصلی IFS هستند، احساسات خنثی در نظرها نیز در نظر گرفته شده است. سپس از این ماتریس در مرحله بعدی برای به دست آوردن وزن ویژگی‌ها و رتبه‌بندی خطوط هوایی استفاده می‌شود. ماتریس تصمیم به دست آمده برای خطوط هوایی در جدول ۶ نشان داده شده است.

جدول ٦. ماتریس تصمیم خطوط هوایی دسته ۲

	C_1	C_2	C_3	C_4	C_5	C_6
B_1	(۰/۳۱,۰/۱۹,۰/۵)	(۰/۵۴,۰/۲۱,۰/۲۴)	(۰/۱۷,۰/۷۸,۰/۰۶)	(۰/۴۵,۰/۳۵,۰/۲)	(۰/۳۵,۰/۵۲,۰/۱۳)	(۰/۵۸,۰/۳,۰/۱۲)
B_2	(۰/۴۶,۰/۴۲,۰/۱۲)	(۰/۵۳,۰/۳۲,۰/۱۵)	(۰/۲۶,۰/۶۵,۰/۰۹)	(۰/۴۲,۰/۳۹,۰/۱۹)	(۰/۳,۰/۶۳,۰/۰۷)	(۰/۴۵,۰/۴۳,۰/۱۲)
B_3	(۰/۵۵,۰/۲۶,۰/۱۸)	(۰/۶۲,۰/۲۸,۰/۱۱)	(۰/۲۶,۰/۶۷,۰/۰۷)	(۰/۵۲,۰/۲۴,۰/۲۴)	(۰/۴۹,۰/۳۸,۰/۱۳)	(۰/۶,۰/۲۶,۰/۱۳)
B_4	(۱/۰,۰/۰,۰/۰)	(۰/۶۳,۰/۱۸,۰/۱۸)	(۰/۰,۱/۰,۰/۰)	(۰/۴۶,۰/۱۹,۰/۳۵)	(۰/۷,۰/۱۳,۰/۱۷)	(۰/۵۳,۰/۲۶,۰/۲۱)
B_5	(۰/۶۷,۰/۲۴,۰/۱)	(۰/۵۴,۰/۳۵,۰/۱۱)	(۰/۲۲,۰/۴۴,۰/۳۳)	(۰/۴۲,۰/۳۸,۰/۲۱)	(۰/۵۴,۰/۳۵,۰/۱۱)	(۰/۴۵,۰/۴,۰/۱۵)
B_6	(۰/۶۴,۰/۲۱,۰/۱۴)	(۰/۴۹,۰/۲۹,۰/۲۲)	(۰/۱۸,۰/۸۲,۰/۰)	(۰/۴۴,۰/۳۵,۰/۲۱)	(۰/۳,۰/۵۷,۰/۱۲)	(۰/۵۵,۰/۲۳,۰/۲۲)
B_7	(۰/۵,۰/۱۷,۰/۳۳)	(۰/۶۶,۰/۰۷,۰/۲۷)	(۰/۰,۱/۰,۰/۰)	(۰/۵۵,۰/۳۳,۰/۱۲)	(۰/۶۳,۰/۲۶,۰/۱۱)	(۰/۷۷,۰/۱۷,۰/۰۶)
B_8	(۰/۵,۰/۳۲,۰/۱۸)	(۰/۳۳,۰/۳۷,۰/۳)	(۰/۱۸,۰/۸۲,۰/۰)	(۰/۴,۰/۴۲,۰/۱۹)	(۰/۳۷,۰/۵۴,۰/۰۹)	(۰/۳۷,۰/۴۹,۰/۱۴)
B_9	(۰/۴۲,۰/۳۵,۰/۲۳)	(۰/۴۳,۰/۳۲,۰/۲۵)	(۰/۱۷,۰/۸۳,۰/۰)	(۰/۲۹,۰/۴۲,۰/۲۹)	(۰/۲۷,۰/۵۷,۰/۱۶)	(۰/۴۴,۰/۴۳,۰/۱۳)
B_{10}	(۰/۴,۰/۴,۰/۲)	(۰/۳۸,۰/۵,۰/۱۲)	(۰/۰۵,۰/۸۹,۰/۰۵)	(۰/۵۴,۰/۲۹,۰/۱۷)	(۰/۳۱,۰/۵۳,۰/۱۷)	(۰/۵۶,۰/۴۴,۰/۰)

	$C_۷$	$C_۸$	$C_۹$	$C_{۱۰}$	$C_{۱۱}$	$C_{۱۲}$
$B_۱$	(۰/۷۲,۰/۲۸,۰/۰)	(۰/۴۵,۰/۳۵,۰/۲)	(۰/۲۲,۰/۵۳,۰/۲۵)	(۰/۵۲,۰/۲۷,۰/۲۱)	(۰/۴۷,۰/۱۱,۰/۴۲)	(۰/۶۷,۰/۱۷,۰/۱۷)
$B_۲$	(۰/۷۲,۰/۲۸,۰/۰)	(۰/۳۸,۰/۳۶,۰/۲۶)	(۰/۴۹,۰/۰۹,۰/۴۳)	(۰/۳۶,۰/۳۶,۰/۲۷)	(۰/۴,۰/۲,۰/۴)	(۰/۶,۰/۰,۰/۴)
$B_۳$	(۰/۴,۰/۶,۰/۰)	(۰/۴۴,۰/۳۵,۰/۲۱)	(۰/۲۲,۰/۴۴,۰/۳۴)	(۰/۶۸,۰/۱,۰/۲۲)	(۰/۴۲,۰/۳۵,۰/۲۳)	(۰/۵۸,۰/۱۶,۰/۲۶)
$B_۴$	(۰/۸,۰/۲,۰/۰)	(۰/۵۶,۰/۲۲,۰/۲۲)	(۰/۸۶,۰/۰,۰/۱۴)	(۰/۱۴,۰/۱۴,۰/۷۱)	(۱/۰,۰/۰,۰/۰)	(۰/۳۳,۰/۰,۰/۶۷)
$B_۵$	(۰/۷۷,۰/۲۳,۰/۰)	(۰/۴۹,۰/۳,۰/۲۱)	(۰/۳۵,۰/۳۵,۰/۳)	(۰/۴۲,۰/۴۴,۰/۱۳)	(۰/۴۱,۰/۲۷,۰/۳۲)	(۰/۷۳,۰/۰۹,۰/۱۸)
$B_۶$	(۰/۸,۰/۲,۰/۰)	(۰/۳۸,۰/۵,۰/۱۲)	(۰/۲۱,۰/۵۳,۰/۲۶)	(۰/۴۱,۰/۲۲,۰/۳۷)	(۰/۳۷,۰/۲۹,۰/۳۴)	(۰/۷۵,۰/۰,۰/۲۵)
$B_۷$	(۰/۸۸,۰/۱۲,۰/۰)	(۰/۴۴,۰/۳۴,۰/۲۲)	(۰/۲۴,۰/۳۲,۰/۴۴)	(۰/۶,۰/۲,۰/۲)	(۰/۵۵,۰/۰,۰/۴۵)	(۰/۸,۰/۰,۰/۲)
$B_۸$	(۰/۵۷,۰/۲۱,۰/۲۱)	(۰/۲۵,۰/۴۶,۰/۳)	(۰/۲۳,۰/۵۳,۰/۲۴)	(۰/۵,۰/۵,۰/۰)	(۰/۳۹,۰/۳۹,۰/۲۲)	(۰/۴۳,۰/۲۹,۰/۲۹)
$B_۹$	(۰/۶۴,۰/۳۶,۰/۰)	(۰/۳۳,۰/۳۸,۰/۲۹)	(۰/۲۳,۰/۵۲,۰/۲۵)	(۰/۴۵,۰/۰۹,۰/۴۵)	(۰/۴۳,۰/۳۹,۰/۱۷)	(۰/۴,۰/۳۵,۰/۲۵)
$B_{۱۰}$	(۰/۸,۰/۲,۰/۰)	(۰/۲۷,۰/۳۸,۰/۳۵)	(۰/۴۱,۰/۳۹,۰/۲)	(۰/۵,۰/۵,۰/۰)	(۰/۲۵,۰/۵,۰/۲۵)	(۱/۰,۰/۰,۰/۰)

	$C_{۱۳}$	$C_{۱٤}$	$C_{۱۵}$	$C_{۱۶}$	$C_{۱۷}$
$B_۱$	(۰/۱۱,۰/۸۴,۰/۰۵)	(۰/۱۶,۰/۴۷,۰/۳۷)	(۰/۳۷,۰/۴,۰/۲۳)	(۰/۶۷,۰/۱۱,۰/۲۲)	(۰/۸۵,۰/۱۵,۰/۰)
$B_۲$	(۰/۱,۰/۹,۰/۰)	(۰/۳,۰/۴۶,۰/۲۴)	(۰/۳۲,۰/۳۶,۰/۳۲)	(۰/۶۷,۰/۰,۰/۳۳)	(۰/۸۱,۰/۱۹,۰/۰)
$B_۳$	(۰/۰۶,۰/۹,۰/۰۳)	(۰/۲۵,۰/۳,۰/۴۵)	(۰/۵۷,۰/۱۶,۰/۲۷)	(۰/۵,۰/۱۷,۰/۳۳)	(۰/۶۹,۰/۳۱,۰/۰)
$B_۴$	(۰/۲۵,۰/۶۹,۰/۰۶)	(۰/۶۷,۰/۳۳,۰/۰)	(۰/۵۷,۰/۱۴,۰/۲۹)	(۱/۰,۰/۰,۰/۰)	(۰/۰,۱/۰,۰/۰)
$B_۵$	(۰/۱۴,۰/۸۶,۰/۰)	(۰/۲۵,۰/۵,۰/۲۵)	(۰/۳۷,۰/۳۱,۰/۳۲)	(۰/۵,۰/۰,۰/۵)	(۰/۸۸,۰/۰۶,۰/۰۶)
$B_۶$	(۰/۱۱,۰/۸۷,۰/۰۱)	(۰/۲۶,۰/۳۵,۰/۴)	(۰/۴۳,۰/۲۷,۰/۳)	(۰/۸,۰/۰,۰/۲)	(۰/۷۶,۰/۲۴,۰/۰)
$B_۷$	(۰/۳۱,۰/۶۹,۰/۰)	(۰/۱۷,۰/۱۷,۰/۶۷)	(۰/۳۸,۰/۲۳,۰/۳۸)	(۱/۰,۰/۰,۰/۰)	(۱/۰,۰/۰,۰/۰)
$B_۸$	(۰/۱۱,۰/۸۷,۰/۰۲)	(۰/۱۷,۰/۵۳,۰/۳)	(۰/۰,۱/۰,۰/۰)	(۰/۰,۱/۰,۰/۰)	(۰/۴۳,۰/۴۳,۰/۱۴)
$B_۹$	(۰/۰۹,۰/۹,۰/۰۱)	(۰/۱۱,۰/۴۸,۰/۴)	(۰/۳۳,۰/۳۳,۰/۳۳)	(۰/۵,۰/۳,۰/۲)	(۰/۸۳,۰/۱۷,۰/۰)
$B_{۱۰}$	(۰/۱۲,۰/۸۴,۰/۰۳)	(۰/۳۶,۰/۴۱,۰/۲۳)	(۰/۰,۱/۰,۰/۰)	(۱/۰,۰/۰,۰/۰)	(۰/۸,۰/۲,۰/۰)

تحلیل MCDM

در این مرحله با توجه به ماتریس تصمیم تشکیل شده در مرحله قبل وزن ویژگی‌ها با استفاده از روش آنتروپی بر مبنـای اعداد فازی شهودی محاسبه می‌شود و در نهایت خطوط هوایی با استفاده از روش تاپسیس بر مبنای اعداد فازی شهودی رتبه‌بندی می‌شوند (حیدری دهوئی و همکاران، ۲۰۲۲).

محاسبه وزن هر ویژگی با استفاده از IF-ENTROPY

در این مرحله، وزن هر ویژگی خطوط هوایی با استفاده از ماتریس تصمیم ساخته‌شده محاسبه می‌شود. برای این منظور از روش IF-ENTROPY برای تعیین وزن‌ها استفاده می‌شود. وزن‌های به دست آمده برای هر ویژگی برای خطوط هوایی در جدول ۷ آورده شده است.

جدول ۷. وزن ویژگی‌های خطوط هوایی ۲

ویژگی	وزن (Wj)	ویژگی	وزن (Wj)	ویژگی	وزن (Wj)
$C_{۱}$	۰/۰۴	$C_{۷}$	۰/۰۵	$C_{۱۳}$	۰/۱۱
$C_{۲}$	۰/۰۲	$C_{۸}$	۰/۰۱	$C_{۱۴}$	۰/۰۲
$C_{۳}$	۰/۱۱	$C_{۹}$	۰/۰۴	$C_{۱۵}$	۰/۰۷
$C_{۴}$	۰/۰۱	$C_{۱۰}$	۰/۰۲	$C_{۱۶}$	۰/۱۷
$C_{۵}$	۰/۰۲	$C_{۱۱}$	۰/۰۵	$C_{۱۷}$	۰/۱۱
$C_{۶}$	۰/۰۲	$C_{۱۲}$	۰/۱۱		

نتایج نشان می‌دهد که از برای خطوط هوایی، کلاس اقتصادی با وزن ۰/۱۷ بیشترین اهمیت و ویژگی‌هایی مانند خدمات مشتری، فضای پا، تأخیر در پرواز و بازرسی امنیتی با وزن ۰/۱۱ از نظر مسافران نسبت به سایر ویژگی‌ها اهمیت یکسانی دارند.

تعیین رتبهٔ نهایی با استفاده از IF-TOPSIS

ماتریس‌های تصمیم تشکیل‌شده و وزن‌های به‌دست‌آمده با روش IF-ENTROPY (جدول ۷) در این مرحله به‌عنوان ورودی برای روش IF-TOPSIS استفاده می‌شود. به‌کمک نتایج به‌دست آمده و با توجه به درجهٔ نزدیکی نسبی هر طرح به راه‌حل ایدئال مثبت مجموعه فازی شهودی، می‌توان رتبه‌بندی عملکرد خطوط هوایی را به دست آورد. رتبه‌بندی گزینه‌ها در نهایت بر اساس شاخص شباهت (C_i) به دست می‌آید. هرچه عدد C_i بزرگ‌تر باشد، رتبه گزینه بالاتر است. جدول ۱۲ رتبه بندی نهایی خطوط هوایی دسته ۱ را با استفاده از روش IF-TOPSIS نشان می‌دهد.

جدول۸. رتبه‌بندی نهایی خطوط هوایی

Alternatives	C_i	Rank	Ranking
$B_{۱}$	۰/۵۹	۵	$C_{۷} > C_{۵} = C_{۶} > C_{۲} > C_{۱۰} > C_{۱} > C_{۳} > C_{۴} > C_{۹} > C_{۸}$
$B_{۲}$	۰/۶۳	۳	
$B_{۳}$	۰/۵۷	۶	
$B_{۴}$	۰/۵۶	۷	
$B_{۵}$	۰/۶۵	۲	
$B_{۶}$	۰/۶۵	۲	
$B_{۷}$	۰/۷	۱	
$B_{۸}$	۰/۲۷	۹	
$B_{۹}$	۰/۵۳	۸	
$B_{۱۰}$	۰/۶۱	۴	

با توجه به نتایج جدول ۸، خط هوایی Middle East Airlines در بین ده خط هوایی (Saudi Arabian Airlines, Kuwait Airways, Oman Air, Iran Air, Egyptair, Royal Jordanian Airlines, Middle East

Airlines, Pegasus Airlines, flydubai, Air Arabia) بالاترین عملکرد را دارد؛ به این معنا که B_{11} نزدیک‌ترین فاصله از راه‌حل ایدئال مثبت مجموعه فازی شهودی و دورترین فاصله از راه‌حل ایدئال منفی مجموعه فازی شهودی را دارد. در حالی که Pegasus Airlines نزدیک‌ترین فاصله به راه‌حل ایدئال منفی و دورترین فاصله از راه‌حل ایدئال مثبت دارد، عملکرد آن در بین چهار خط هوایی کمترین است.

نتیجه‌گیری و پیشنهادها

این دستاورد ما را بر آن داشت تا با استفاده از این اطلاعات، توصیه‌هایی را برای مدیریت هر شرکت در رابطه با خدمات ارائه‌شده بیان کنیم که در جدول ۹ به‌طور خلاصه آمده است.

جدول ۹. توصیه‌هایی به خطوط هوایی مورد مطالعه

خط هوایی	توصیه‌ها
Middle East Airlines	رضایت مسافران از خدمات این شرکت هواپیمایی در سطح خوبی است، بااین‌حال بیشتر نظرهای منفی مربوط به تأخیر زمانی و انتقاد از کیفیت خدمات در موضوعاتی مانند «خدمات مشتری و «غذا» است. این خدمات باید توسط شرکت بهبود یابد تا رضایت مشتری افزایش یابد.
Egyptair	ازجمله نقاط ضعف این شرکت هواپیمایی را می‌توان به موضوعاتی مانند «خدمات مشتری»، «صندلی»، «سالن استراحت فرودگاه»، «کلاس تجاری»، «زمان انتظار» و «خدمه» نسبت داد که با اتخاذ تصمیمات درست می‌توان آن‌ها را بهبود بخشید.
Royal Jordanian Airlines	عمده‌ترین نقاط ضعف این شرکت هواپیمایی بیشتر در زمینه خدمات مشتری، تأخیرهای زمانی، کارکنان، بار مسافر و مسائل تغذیه است که نیاز به بهبود دارد. توصیه می‌شود مدیران این شرکت هواپیمایی با سنجیدن کمی و کیفی رضایت مشتریان با تمرکز بر نقاط ضعف شناسایی‌شده تصمیمات مناسب اتخاذ کنند.
Kuwait Airways	در مبحث وعده غذایی، قیمت بلیت، خدمات مشتری، کارکنان و تأخیرهای زمانی ضعیف است که محل بازنگری و بهبود به‌ویژه در حوزه تغذیه دارد. توجه به ذائقه مسافران در پروازها در برخی مناطق خاص می‌تواند به‌عنوان راهی برای افزایش رضایت مشتریان مؤثر باشد.
Air Arabia	این شرکت هواپیمایی در موضوعاتی مانند «فرایند پذیرش»، «نوشیدنی»، «غذا»، «خدمات مشتری»، «کارکنان» امتیازهای پایینی دارد که نیاز به بهبود دارند.
Saudia	مدیران باید توجه بیشتری به طراحی صندلی سرنشینان به‌ویژه در کلاس اقتصادی داشته باشند. ارتقای کیفیت غذای ارائه‌شده به مسافران و کاهش تأخیر در ارائه خدمات از توصیه‌های دیگر به مدیران است.
Oman Air	بیشتر نظرهای منفی مربوط به تأخیر زمانی و انتقاد از کیفیت خدمات در موضوعاتی مانند «خدمات مشتری»، «بار مسافر»، «فرایند پذیرش» و «غذا» است. این خدمات باید توسط شرکت بهبود یابد تا رضایت مشتری افزایش یابد.
Iran Air	اکثر نظرهای منفی در موضوعاتی مانند «خدمات مشتری»، «سالن استراحت»، «زمان تأخیر»، «نوشیدنی» است که با اتخاذ تصمیمات مناسب توسط مدیران می‌توان آن‌ها را بهبود بخشید و باعث جلب رضایت مشتریان شد.
flydubai	خدمات ارائه‌شده در عناوین «خدمات مشتری»، «فرایند پذیرش»، «صندلی»، «کارکنان»، «خدمه»، «تأخیر زمانی»، «نوشیدنی» و «بار مسافر» باعث نارضایتی مسافران شده است که نیاز به بررسی دارد.
Pegasus Airlines	ازجمله موضوعاتی که اکثر مسافران نسبت به آن‌ها اعلام نارضایتی داشتند می‌توان به مواردی مانند «غذا»، «خدمات مشتری»، «صندلی»، «کارکنان»، «خدمه»، «بار مسافر»، «نوشیدنی»، «زمان تأخیر و انتظار» و «فرایند پذیرش» اشاره کرد که می‌بایست مورد توجه مدیران جهت اخذ تصمیمات مناسب برای بهبود قرار گیرد.

همچنین، با جمع‌آوری نظرهای مسافران و به‌کارگیری پیشنهادهای مورد استفاده توسط خطوط هوایی برتر جهان، می‌توان راه‌کارهای مناسبی به‌منظور بهبود ویژگی‌های مختلفی که در این پژوهش برای خطوط هوایی مورد بررسی قرار گرفته‌اند، پیشنهاد داد که در جدول ۱۰ مشخص شده است.

جدول ۱۰. راه‌کارهای بهبود ویژگی‌ها برای خطوط هوایی مورد مطالعه

ویژگی	راه‌کار بهبود ویژگی‌ها
قیمت بلیت	استفاده از قیمت‌گذاری پویا برای تنظیم قیمت بلیت بر اساس تقاضا و سایر عوامل مانند زمان رزرو، روز هفته و رویدادهای خاص تخفیف برای رزروهای اولیه، خرید عمده و رفت‌وبرگشت برنامه وفاداری برای پاداش دادن به مسافران مکرر با ارائهٔ تخفیف و هدیه اجازه خریداری بلیت‌هایی به مشتریان برای خدمات اضافی مانند چمدان، انتخاب صندلی، و وعده‌های غذایی در پرواز با هزینه اضافی ارائهٔ نرخ‌های تخفیفی برای دانشجویان، کارکنان نظامی و سایر گروه‌های دارای نیازهای ویژه
غذا	ارائهٔ گزینه‌های گیاه‌خواری، وگان و بدون گلوتن برای رفع محدودیت‌های غذایی مختلف استفاده از بسته‌بندی‌ها و ظروف سازگار با محیط‌زیست برای کاهش ضایعات پیش سفارش کردن گزینه‌های وعده غذایی برای مسافرانی با نیازها یا ترجیحات غذایی خاص توجه به تنوع فرهنگی مسافران ازنظر ترجیحات غذایی ارائهٔ اطلاعات واضح آلرژی‌زا و مواد تشکیل‌دهنده برای همه مواد غذایی
خدمات مشتری	ارائهٔ خدمات مشتریان چندزبانه برای پاسخ‌گویی به پایگاه مشتریان استفاده از رسانه‌های اجتماعی برای تعامل با مشتریان، پاسخ به سؤالات و حل مشکلات ارائهٔ اطلاعات واضح و قابل درک درباره خطمشی‌ها، رویه‌ها و خدمات شرکت هواپیمایی
صندلی	ارائهٔ یک صندلی راحت با فضای پا زیاد و تکیه‌گاه سر قابل تنظیم ارائهٔ گزینه‌هایی به مسافران برای انتخاب صندلی‌های راحت‌تر با پرداخت هزینه اضافی ارائهٔ پریزهای برق و پورت‌های USB کنار صندلی برای شارژ دستگاه‌های الکترونیکی
کارکنان	تشویق آموزش و توسعه مداوم برای اطمینان از به‌روز بودن کارکنان در مورد آخرین روندهای صنعت و بهترین شیوه‌ها شناسایی و پاداش به کارکنانی با ارائهٔ خدمات بهتر به مشتریان پیشنهاد کارمندان چند زبانه برای پاسخ‌گویی به پایگاه مشتریان مختلف
خدمه	شناسایی و پاداش به اعضای خدمه‌ای با ارائهٔ خدمات بهتر به مشتریان پیشنهاد خدمه چند زبانه برای پاسخ‌گویی به پایگاه مشتریان مختلف
سرگرمی	ارائهٔ انواع بازی‌ها برای سرگرم نگه‌داشتن مسافران فراهم کردن هدفون حذف نویز برای مسافران در طول پرواز ارائهٔ طیف گسترده‌ای از گزینه‌های صوتی و تصویری برای مسافرانی با اختلال بینایی یا شنوایی ارائهٔ یک سرویس پخش زنده تلویزیونی برای به‌روز نگه‌داشتن مسافران از اخبار و رویدادهای ورزشی
فرایند پذیرش	ارائهٔ خطوط اختصاصی برای پذیرش به مسافرانی با نیازهای ویژه، مانند خانواده‌هایی با کودکان خردسال یا مسافران دارای معلولیت

ویژگی	راه‌کار بهبود ویژگی‌ها
	در اولویت قرار دادن خدمات پذیرش برای مسافران کلاس ممتاز یا اعضای برنامه پروازهای مکرر استفاده از فناوری بیومتریک، مانند تشخیص چهره، برای ساده‌سازی فرایند پذیرش و بهبود امنیت
بار مسافر (چمدان)	ارائهٔ یک سیستم ردیابی چمدان برای ردیابی مکان چمدان مسافران ارائهٔ خدمات بسته‌بندی چمدان برای محافظت از وسایل شکننده در حین حمل‌ونقل ارائهٔ گزینه بیمه چمدان با هزینه اضافی
سالن فرودگاه	ارائهٔ خدمات متنوع مانند خدمات حمام، آبگرم و ماساژ ارائهٔ یک برنامه وفاداری برای پاداش دادن به کاربران مکرر سالن با ارائهٔ هدیه و تخفیف
نوشیدنی	ارائهٔ مجموعه‌ای از نوشیدنی‌های محلی و منطقه‌ای برای پاسخ‌گویی به مشتریان مختلف ارائهٔ اطلاعات شفاف آلرژی‌زا و مواد تشکیل‌دهنده برای همه نوشیدنی‌ها
فضای پا	ارتقا صندلی با هزینه‌های اضافی برای تقاضای مسافرانی با فضای پای بیشتر ارائهٔ اطلاعات واضحی در مورد فضای پا و ابعاد صندلی در طول فرایند رزرو فراهم کردن فضای پای کافی برای مسافران دارای معلولیت
تأخیر پرواز	ارائهٔ نوشیدنی‌ها و میان وعده‌های رایگان برای مسافرانی که پروازشان با تأخیر انجام می‌شود. ارائهٔ برنامه جبران خسارت برای مسافرانی که پروازشان با تأخیر است یا لغو شده است. ارائهٔ بازپرداخت، رزرو مجدد یا کوپن برای لغو یا تأخیر پرواز
زمان انتظار	ارائهٔ اولویت فرایند پذیرش و غربالگری امنیتی برای مسافران کلاس ممتاز یا اعضای برنامه پروازهای مکرر برای به حداقل رساندن زمان انتظار برای آن‌ها ارائهٔ اطلاعات واضح و دقیقی در مورد زمان انتظار برای پذیرش، امنیت و سوارشدن به هواپیما
کلاس تجاری و اقتصادی	یک خدمه اختصاصی برای ارائهٔ خدمات شخصی و توجه به مسافران کلاس تجاری و کلاس اقتصادی ارائهٔ یک برنامه وفاداری برای پاداش دادن به مسافران مکرر کلاس تجاری و اقتصادی با هدیه و تخفیف
بررسی امنیتی	در نظر گرفتن خطوط جداگانه‌ای برای خانواده‌هایی با فرزندان خردسال یا مسافران دارای معلولیت فراهم کردن خطوط امنیتی اختصاصی برای مسافران کلاس ممتاز یا اعضای برنامه پروازهای مکرر استفاده از فناوری‌هایی مانند غربالگری بیومتریک و اسکن اشعه ایکس برای ساده‌سازی فرایند امنیتی و بهبود امنیت

این پژوهش به خطوط هوایی خاورمیانه کمک زیادی می‌کند تا به دنبال زمینه‌های بهبود باشند و به‌راحتی می‌توانند عملکرد خود را با رقبای خود برای دستیابی به مزیت رقابتی بهتر در بازار مقایسه کنند. از داده‌های این پژوهش می‌توان برای ایجاد یک سیستم توصیه برای مسافران استفاده کرد تا به آن‌ها کمک کند که خطوط هوایی‌ای را انتخاب کنند که به بهترین وجه انتظارات، خواسته‌ها و هدف سفرشان را برآورده کند. این می‌تواند عواملی مانند بودجه، مقصد و نوع کلاس پرواز را در نظر بگیرد که می‌تواند به درک بهتر نیازهای مشتریان برای مدیران خطوط هوایی کمک کند.

در این پژوهش، رویکرد ابجکتیو برای وزن‌دهی به‌دلیل ماهیت داده‌های کلان (نظرهای آنلاین) و نیاز به کاهش سوگیری انسانی انتخاب شده است. رویکردهای ذهنی برای سناریوهایی که نیاز به قضاوت‌های دقیق خبرگان باشد مناسب هستند؛ اما در اینجا حجم داده‌های جمع‌آوری‌شده و نیاز به تحلیل دقیق و بدون سوگیری، انتخاب رویکرد ابجکتیو را توجیه می‌کند و درنهایت از بین ویژگی های بدست آمده از الگوریتم HAC از نظرهای خبرگان نیز استفاده

شده است. برای بهبود نتایج در پژوهش‌های آینده می‌توان از رویکردهای ترکیبی استفاده کرد که از هر دو رویکرد عینی و ذهنی بهره‌برداری کنند. برای مثال، می‌توان ارزیابی‌های ذهنی خبرگان را نیز با استفاده از داده‌های کیفی جمع‌آوری کرد و در کنار رویکردهای ابجکتیو، از روش‌های ذهنی استفاده کرد تا ترکیب بهتری از دقت داده‌ها و قضاوت‌های انسانی به دست آید.

در زیر، زمینه‌هایی برای تحقیقات آینده پیشنهاد شده است:

- این مطالعه فقط ویژگی‌های صریح بیان‌شده توسط کلمه‌های صریح در بررسی آنلاین مشتریان (OCR) را در نظر گرفته و ویژگی‌های ضمنی را پوشش نداده است. بدین منظور در تحقیقات آینده می‌توان ضمن بررسی ویژگی‌های ضمنی به بررسی تأثیر آن‌ها بر رضایت مشتری پرداخت.
- در این پژوهش جهت‌گیری احساسات برای هر ویژگی فقط از نظر تعداد بررسی شده است، به این معنا که تعداد جمله‌های مثبت، منفی و خنثی برای هر ویژگی محاسبه می‌شود. از سوی دیگر، شدت احساسات به درجه یا قدرت احساسات بیان‌شده در یک متن اشاره دارد. این جنبه می‌تواند خاص‌تر از جهت‌گیری احساسات باشد؛ زیرا می‌تواند تفاوت‌های ظریفی مانند درجه مثبت یا منفی را به تصویر بکشد.
- استفاده از رویکردهای داده محور مانند نقشه شناختی فازی و قواعد انجمنی که رابطه بین ویژگی‌ها و شناسایی الگوها و هم‌بستگی‌ها در داده‌ها را در نظر می‌گیرد و محیط تصمیم‌گیری مؤثرتری را فراهم می‌کند، زمینهٔ دیگری برای تحقیقات آینده است.
- استفاده از روش مدل‌سازی موضوع[1] بدین‌صورت که ابتدا موضوعات را با استفاده از مدل‌سازی موضوع با روش‌های مختلفی مانند NMF[2]، LDA[3] و LSI[4] شناسایی کرد و سپس احساسات مسافر را در مورد موضوعات شناسایی‌شده تجزیه‌وتحلیل کرد.
- انجام تحلیل مقایسه‌ای بین روش‌های مختلف تحلیل احساسات و استخراج ویژگی‌ها (مثل مقایسه HAC با الگوریتم‌های دیگر مثل TF-IDF یا Word2Vec)
- استفاده از تکنیک‌های پیشرفته‌تر پیش‌پردازش متن، مانند اصلاح اشتباه‌های املایی و بررسی نحوهٔ برخورد با کلمه‌های هم‌معنا[5] و متضاد[6] در داده‌ها.

برای بهبود اعتبار تحلیل‌ها، توصیه می‌شود بررسی‌های بیشتری روی بایاس‌های ممکن انجام شود. به‌ویژه، تحلیل‌هایی برای شناسایی و تعدیل نظرهای جعلی و تبلیغاتی و بررسی دقت و صحت نظرها می‌تواند به بهبود کیفیت نتایج کمک کند.

1. Topic modeling
2. Nonnegative Matrix Factorization
3. Latent Dirichlet Allocation
4. Latent Semantic Indexing
5. Synonyms
6. Antonyms

منابع

محتشمی، ملیکا؛ حسین‌زاده کاشان، علی و حیدری دهوئی، جلیل (۱۴۰۱). تبیین الگوی تصمیم‌گیری داده‌محور برای شناسایی عوامل مؤثر بر رضایت معامله‌گران صرافی‌های برخط با استفاده ازنقشه شناختی فازی (مورد مطالعه: گوگل پلی و اپ استور). *فصلنامه علوم مدیریت ایران*، ۸۰، ۱۵۷-۱۸۰.

References

Agrawal, A.K., Jagannathan, M. & Delhi, V.S.K. (2021). Control Focus in Standard Forms: An Assessment through Text Mining and NLP. *Journal of Legal Affairs and Dispute Resolution in Engineering and Construction*, 13(1). https://doi.org/10.1061/(asce)la.1943-4170.0000441.

Antons, D., Grünwald, E., Cichy, P. & Salge, T. O. (2020). The application of text mining methods in innovation research: current state, evolution patterns, and development priorities. *R&D Management*, 50(3), 329-351.

Atanassov, K.T. (1986). Intuitionistic fuzzy sets., *Fuzzy Sets and Systems*, 20(1), 87–96. https://doi.org/10.1016/S0165-0114(86)80034-3.

Bakır, M., Akan, Ş. & Durmaz, E. (2019). Exploring Service Quality of Low-Cost Airlines in Europe: an Integrated MCDM Approach., *Economics and Business Review*, 5(2), 109–130. https://doi.org/10.18559/ebr.2019.2.6.

Ban, H. J., Kim, H. S. (2019). Understanding customer experience and satisfaction through airline passengers. *online review. Sustainability*, 11(15), 4066.

Çalı, S. & Balaman, Ş.Y. (2019). Improved decisions for marketing, supply and purchasing: Mining big data through an integration of sentiment analysis and intuitionistic fuzzy multi criteria assessment., *Computers and Industrial Engineering*, 129, 315–332. https://doi.org/10.1016/j.cie.2019.01.051.

Chow, C.K.W. (2014). Customer satisfaction and service quality in the Chinese airline industry. *Journal of Air Transport Management*, 35, 102–107. https://doi.org/10.1016/j.jairtraman.2013.11.013.

Eirinaki, M., Pisal, S. & Singh, J. (2012). Feature-based opinion mining and ranking., *Journal of Computer and System Sciences*, 78(4), 1175–1184. https://doi.org/10.1016/j.jcss.2011.10.007.

Farzadnia, S. & Vanani, I.R. (2022). Identification of opinion trends using sentiment analysis of airlines passengers' reviews. *Journal of Air Transport Management*, 103, 102232.

Greer, C.R. & Lei, D. (2012). Collaborative Innovation with Customers: A Review of the Literature and Suggestions for Future Research. *International Journal of Management Reviews*, 14(1), 63–84. https://doi.org/10.1111/j.1468-2370.2011.00310.x.

Gupta, H. (2018). Evaluating service quality of airline industry using hybrid best worst method and VIKOR. *Journal of Air Transport Management*, 68, 35–47. https://doi.org/10.1016/j.jairtraman.2017.06.001.

Hashemkhani Zolfani, S. & Derakhti, A. (2020). Synergies of text mining and multiple attribute decision making: a criteria selection and weighting system in a prospective MADM outline. *Symmetry*, 12(5), 868.

Hassani, H., Beneki, C., Unger, S., Mazinani, M. T. & Yeganegi, M. R. (2020). Text mining in big data analytics. *Big Data and Cognitive Computing*, 4(1), 1.

Heidary Dahooie, J. Raafat, R., Qorbani, A. R. & Daim, T. (2021). An intuitionistic fuzzy data-driven product ranking model using sentiment analysis and multi-criteria decision-making. *Technological Forecasting and Social Change*, 173, 121158. https://doi.org/10.1016/j.techfore.2021.121158

Heidary Dahooie, J., Husseinzadeh Kashan, A., Shoaei Naeini, Z., Vanaki, A. S., Zavadskas, E. K. & Turskis, Z. (2022). A hybrid multi-criteria-decision-making aggregation method and geographic information system for selecting optimal solar power plants in Iran. *Energies*, *15*(8), 2801. https://doi.org/10.3390/en15082801

Hu, J., Zhang, X., Yang, Y., Liu, Y. & Chen, X. (2020). New doctors ranking system based on VIKOR method. *International Transactions in Operational Research*, *27*(2), 1236-1261. https://doi.org/10.1111/itor.12569

Hung, C. C. & Chen, L. H. (2010). A multiple criteria group decision making model with entropy weight in an intuitionistic fuzzy environment. *Intelligent automation and computer engineering*, 17-26. https://doi.org/10.1007/978-90-481-3517-2_2

Husain, S., Ahmad, Y. & Alam, M. A. (2012). A study on the role of intuitionistic fuzzy set in decision making problems. *International journal of computer applications*, 48(0975-888), 35-41.

Hussain, R., Al Nasser, A. & Hussain, Y.K. (2015) Service quality and customer satisfaction of a UAE-based airline: An empirical investigation. *Journal of Air Transport Management*, 42, 167–175. https://doi.org/10.1016/j.jairtraman.2014.10.001.

Kaya, T. & Öztürk, Z. K. (2020). Feature analysis for multi-criteria rating values of airline companies. *Mühendislik Bilimleri ve Tasarım Dergisi*, 8(2), 333-344.

Kumar, G. & Parimala, N. (2020). An Integration of Sentiment Analysis and MCDM Approach for Smartphone Recommendation, *International Journal of Information Technology and Decision Making*, 19(4), 1037–1063. https://doi.org/10.1142/S021962202050025X.

Li, Y., Zhang, Y. & Xu, Z. (2020). A Decision-Making Model Under Probabilistic Linguistic Circumstances with Unknown Criteria Weights for Online Customer Reviews, *International Journal of Fuzzy Systems*, 22, 777–789. https://doi.org/10.1007/s40815-020-00812-1.

Liang, R. & Wang, J. Q. (2019). A Linguistic Intuitionistic Cloud Decision Support Model with Sentiment Analysis for Product Selection in E-commerce, *International Journal of Fuzzy Systems*, 21, 963–977. https://doi.org/10.1007/s40815-019-00606-0.

Liang, X., Liu, P. & Wang, Z. (2019). Hotel selection utilizing online reviews: A novel decision support model based on sentiment analysis and DL-VIKOR method, *Technological and Economic Development of Economy*, 25(6), 1139–1161. https://doi.org/10.3846/tede.2019.10766.

Liu, P. & Teng, F. (2019). Probabilistic linguistic TODIM method for selecting products through online product reviews, *Information Sciences*, 485, 441–455. https://doi.org/10.1016/j.ins.2019.02.022.

Mohtashamimaali, M., Hoseinzade Kashan, A. & Heidary, J. (2023). The data-driven decision-making model to identify the factors affecting the satisfaction of traders of online exchanges: A study of Google Play and Apple Store. *Iranian journal of management sciences, 17*(68), 157-180. (*in Persian*)

Nosrati Malekjahan, A., Husseinzadeh Kashan, A. & Sajadi, SM. (2024). A novel sequential risk assessment model for analyzing commercial aviation accidents: Soft computing perspective, *Risk Analysis*. Available at: https://doi.org/10.1111/risa.14486.

Oum, T.H., Yu, C. &Fu, X. (2003). A comparative analysis of productivity performance of the world's major airports: Summary report of the ATRS global airport benchmarking research report – 2002. *Journal of Air Transport Management*, 9(5), 285–297. https://doi.org/10.1016/S0969-6997(03)00037-1.

Ravi, K. & Ravi, V. (2015). A survey on opinion mining and sentiment analysis: Tasks, approaches and applications, *Knowledge-Based Systems*, 89, 14–46. https://doi.org/10.1016/j.knosys.2015.06.015.

Sahoo, S. K. & Goswami, S. S. (2023). A comprehensive review of multiple criteria decision-making (MCDM) Methods: advancements, applications, and future directions. *Decision Making Advances*, 1(1), 25-48.

Shannon, C.E. (1948). A Mathematical Theory of Communication, *Bell System Technical Journal*, 27(3), 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x.

Sharma, H., Tandon, A., Kapur, P. K. & Aggarwal, A. G. (2019). Ranking hotels using aspect ratings based sentiment classification and interval-valued neutrosophic TOPSIS. *International Journal of System Assurance Engineering and Management*, 10(5), 973-983. https://doi.org/10.1007/s13198-019-00827-4

Sotoudeh-Anvari, A. (2022). The applications of MCDM methods in COVID-19 pandemic: A state of the art review. *Applied Soft Computing*, 126, 109238.

Taherdoost, H. & Madanchian, M. (2023). Multi-criteria decision making (MCDM) methods and concepts. *Encyclopedia*, *3*(1), 77-87.

Vanani, I. R. & Majidian, S. (2020). Meta-Heuristic Algorithms: A Concentration on the Applications in Text Mining. *In Big Data, IoT, and Machine Learning* (pp. 113-132). CRC Press. https://doi.org/10.1201/9780429322990-6

Vlachos, I.K. & Sergiadis, G.D. (2007). Intuitionistic fuzzy information - Applications to pattern recognition, *Pattern Recognition Letters*, 28(2), 197–206. https://doi.org/10.1016/j.patrec.2006.07.004.

Wan, Y. & Gao, Q. (2016). An Ensemble Sentiment Classification System of Twitter Data for Airline Services Analysis, in *Proceedings - 15th IEEE International Conference on Data Mining Workshop, ICDMW 2015*, 1318–1325. https://doi.org/10.1109/ICDMW.2015.7.

Wang, L., Wang, X. K., Peng, J. J. & Wang, J. Q. (2020). The differences in hotel selection among various types of travellers: A comparative analysis with a useful bounded

rationality behavioural decision support model. *Tourism management*, 76, 103961.

Wright, K.B. (2005) Researching internet-based populations: Advantages and disadvantages of online survey research, online questionnaire authoring software packages, and web survey services, *Journal of Computer-Mediated Communication*, 10(3). https://doi.org/10.1111/j.1083-6101.2005.tb00259.x.

Wu, C. & Zhang, D. (2019). Ranking products with IF-based sentiment word framework and TODIM method, *Kybernetes*, 48(5), 990–1010. https://doi.org/10.1108/K-01-2018-0029.

Yang, Z., Xiong, G., Cao, Z., Li, Y. & Huang, L. (2019). A Decision Method for Online Purchases Considering Dynamic Information Preference Based on Sentiment Orientation Classification and Discrete DIFWA Operators, *IEEE Access*, 7, 77008–77026. https://doi.org/10.1109/ACCESS.2019.2921403.

Yang, Z., Ouyang, T., Fu, X. & Peng, X. (2020). A decision-making algorithm for online shopping using deep-learning–based opinion pairs mining and q-rung orthopair fuzzy interaction Heronian mean operators, *International Journal of Intelligent Systems*, 35(5), 783–825. https://doi.org/10.1002/int.22225.

Zadeh, L.A. (1965). Fuzzy sets. *Information and Control*, 8(3), 338–353. https://doi.org/10.1016/S0019-9958(65)90241-X.

Zeleny, M. (1976) Attribute-Dynamic Attitude Model (Adam). *Management Science*, 23(1), 12–26. https://doi.org/10.1287/mnsc.23.1.12.

Zhang, C., Tian, Y. X., Fan, L. W. & Li, Y. H. (2020). Customized ranking for products through online reviews: a method incorporating prospect theory with an improved VIKOR, *Applied Intelligence*, 50(6), 1725–1744. https://doi.org/10.1007/s10489-019-01577-3.

Zhang, D., Li, Y. & Wu, C. (2020). An extended TODIM method to rank products with online reviews under intuitionistic fuzzy environment. *Journal of the Operational Research Society*, 71(2), 322–334. https://doi.org/10.1080/01605682.2018.1545519.

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی