Technical Journal, University of Engineering and Technology (UET) Taxila, Pakistan Vol. 30 No. 3-2025
ISSN:1813-1786 (Print) 2313-7770 (Online)

Unveiling Ambivalence in Reviews: Using Sentence-BERT and K-NN for Airline Recommendations

M. U. Javeed[1], S. M. Aslam[2], S. Aslam[3], M. Iqbal[4], M. Farhan[5], M. Javed[6], A. Raza[7]

[1,3,5,6,7] Department of Computer Science, COMSATS University of Islamabad, Sahiwal Campus, Pakistan
[2] School of Computer Science, Shaanxi Normal University, Xi'an, Shaanxi, China
[3] Department of Computer Science, University of Engineering and Technology, Lahore, Pakistan

[1]usmanjavveed@gmail.com

Abstract- There are a number of benefits to travel, when it comes to crossing foreign borders. People are free to share opinions by posting reviews on websites or other online platforms. Customer interactions are directly impacted by reviews. These viewpoints may be presented in a single review positive, negative or conflicted. The studies proved that conflicting internet reviews, has gained enormous attention in recent years. In along with assisting consumers to choose an appropriate airline, these ratings also help airline companies find and fix problems with their services. In order to fill this gap, we propose a study that conceptualizes the features of contradictory airline evaluations, identifying the perceptions of travelers that lead to their attitudinal ambiguity or doubt while formulating their actions. We studied to see how travelers' attitudinal ambivalence, which results in indecision, is triggered by contradicting aspects of airline evaluations. In this paper, we proposed many strategies to eliminate this ambivalence and offer suggestions to the customers as well as airlines. We firstly pre-processed traveler evaluations in the recommendation system using NLP (Natural Language Processing) approaches. After pre-processing Sentence-BERT is used to convert the textual reviews into vectors and then K-NN is used to recommend the airline based on those vectors. This machine learning technique suggests a suitable airline to customers. The machine learning model's recommendation accuracy is increased by this novel approach to using online social networks to advertise low-cost flights to travelers. Our proposed model achieved an accuracy of 94.7%, precision of 95.24%, F1 score of 95.4% and recall of 95.5%.

Keywords- Airline, Reviews, Scraping, Ambivalence, S-BERT, K-NN

I. INTRODUCTION

Effective online and offline client relationship management is essential to the air transport industry's intense competition in order to sustain passenger satisfaction and develop future revenue [1]. Recently, the airlines have played a significant role in the development and growth of the world economy. Airlines are the primary sponsor of the tourist and travel business, which has led to significant growth in this sector. Buses, trains, and other forms of transportation compete with air travel, which is the fastest mode of transportation worldwide [2]. Flying enables planes to avoid natural and man-made barriers such and mountains, oceans and rivers. In the late 20th century, flying was considered an expensive luxury. Larger passenger planes and technological advancements have made it a reasonably priced form of transportation, and in certain places, costs are constantly dropping. Additionally, travelers are offered economic incentives, Wi-Fi, food and drink, safety, and a nice environment throughout their flight. Because of the speed and lower accident rate compared to other forms of transportation, travelers prefer flying [3]. The fastest way to go around is by air, which may save you hours or even days of travel time. Air travel's speed also makes it possible to transfer important supplies to individuals in need, such as medications and organs. Pilots, mechanics, stewards, customer service representatives, and designers are all employed in the aviation sector.
Travelers share their thoughts and insights via blogs, tweets, reviews, and sentimental analysis. Among other things, it helps the business, organization, and industry as well as people who wish to go to their company, purchase under their brand, and use their application. They can help customers compare shopping outcomes and increase the credibility of retailers. Benefits for e-commerce (profit gain, useful market research); benefits for end users (improved trust in the product and the business, which provides customers a voice). Qatar Airways, Turkish Airlines, PIA, Emirates, and so forth are examples of different airline types. Every airline is operating well in comparison to one another. The traveler finds it difficult to choose an airline. Social media is becoming a significant source of

evaluations and opinions. According to earlier studies on an average day, hardly 15% of flights (out of a maximum of 34%) produced contrail [4]. Between June and September, 63 percent of the total Contrail Along Track Distance was produced, with the majority of contrails being created in the southeast and along the Pacific coast.

A review serves as an opportunity for us to share our ideas, opinions, and motivations with the business or customer. These reviews are not simple. There are other kinds of reviews, but we'll concentrate on two here. Consumer reviews, comments, and ideas about a product, event, or entity are among the vast amounts of fresh data generated by the quickly growing internet resources, blogs, discussion groups, and forums, which is why people are leaving comments. Customers and decision-makers profit from reviews of any business, airline, bank, hotel, and online shopping items, such as books, smartphones, and laptops [5]. Internal reviews are the first, and they contain both positive and negative comments. Across reviews are the second, where varying viewpoints are expressed. In essence, these reviews are a collection of reviews. As a instance, "This chocolate is amazing, but the feelings that I get after eating it are not good."

Due of social distance, social media has been more popular and rated higher since COVID. Everyone loves to purchase, travel, and enjoy entertainment online. The world is not simultaneously interconnected. Views, suggestions, ratings, video testimonials, tweets, photos, and blog postings are all examples of electronic word-of-mouth (eWOM) [6-7]. According to Filieri, people today rely heavily on electronic word of mouth (eWOM) to distribute views while making travel plans because of the widespread usage of the internet [8-9]. As a result, one of the significant and prominent forms of eWOM is online reviews [11]. According to study, customers are more likely to process information incorrectly when reviews are bivalent (i.e., good and negative), inconsistent, or conflicting practices by airlines and consumers and this may cause them to experience ambivalence and emotions of doubt [12], [13]. Therefore, we argue that paradoxes that coexist in an item (like online reviews) cause customers to develop attitudinal ambivalence by giving them a confusing and ambiguous attitude toward the object, which prevents them from making a decision.

Despite the extensive use of sentiment analysis in online reviews, most existing studies focus on classifying reviews into simple positive or negative categories. However, airline reviews often contain ambivalence, where travelers express both satisfaction and dissatisfaction within the same review (e.g., "The staff was friendly, but the flight was delayed for hours"). Such mixed opinions present a challenge for conventional models, which may misclassify or oversimplify the sentiment, leading to unreliable recommendations. Addressing this gap is essential, as inaccurate interpretation of reviews can misguide passengers in selecting airlines and limit the ability of companies to respond effectively to customer needs. To overcome this limitation, our study leverages Sentence-BERT (S-BERT), which captures semantic and contextual nuances in text, combined with K-Nearest Neighbors (K-NN), which is effective in classifying reviews based on semantic similarity. This approach enables us to better identify and interpret ambivalent reviews, thereby enhancing the accuracy of airline recommendation systems.

This study investigates and estimates recommendations made by reviewers in online evaluations of travel-related services, particularly airline services, from a number of perspectives. We gathered information and reviews from the websites of the top ten airlines. Using the dataset pre-processing approach, textual and non-textual data are discovered. After that, the cleaned reviews are then transformed into numerical vectors that represent their semantic content using Sentence-BERT. Then, we developed a model by using machine learning algorithm K-NN. By identifying reviews with similar meanings, the K-NN (K-Nearest Neighbors) algorithm utilizes these vectors to promote airlines. This enables the system to provide airline recommendations based on users who had similar experiences. In this study we examine in this work the effectiveness of our novel proposed model for suggestion-based airline systems using a K-NN.

II. Literature Review

A. Online Reviews

There are several methods to characterize online platforms. Examples of online communities that have been reviewed include social networking platforms, blogs, forums, online commerce websites, and online review sites. Online customer ratings and reviews (opinions in text format) are usually submitted by customers who have first-hand, usually recent experience with a product or service [16]. It has also been discovered that customer-focused product development is significantly impacted by online user reviews [17]. According to Zhang [18], Online reviews and advancements in smartphone functionality have a strong connection. Researchers have acknowledged the significance of online user reviews in product development, and more research is focusing on product related issues. A complex LDA model was created by Qiao [14] in order to identify and gather important data on product issues from online reviews. In the US, the majority of people have read (82%) and published (61%) online reviews.

Timoshenko and Hauser [15] extracted customer requirements from Amazon reviews. The demands

of consumers are a statement outlining the benefits that they want, including both expressing dissatisfaction with current features and promoting new ones. It is probably true that online reviews with a bigger number of review votes correlate more strongly with sales. Numerous components, including language characteristics, opinions, semantic elements, and other information sources, make up online evaluations. Saurabh Bahulikar's suggested method can be seen by the pseudo-code [16]. Since no currently available datasets could provide all the necessary tuples within the dataset, To complete the research, we had to create our own training dataset with the necessary filtering and, subsequently, suggestion-generating criteria. The dataset consists of 480 flights, each with a unique set of content-based indicators [10]. Coordinate validation is offered by Google as part of the "google API." The results are highly accurate and may be found using precision, recall, or the F measure.
Dr. Swagato's [17] work focuses on pricing and marketing strategies that companies may utilize in a multi-period game. In the first period, they create positive online reviews, and in the second period, they take advantage of the positive impact those reviews have on their pricing, sales, and earnings. Overall sentiment was calculated using (Pos neg) / (pos + neg), where pos is the frequency of positive terms and neg is the frequency of negative keywords in the text. Good and negative emotions have also been demonstrated to be proportionately relevant, and customer satisfaction is an excellent indicator of customer outcomes and services.

B. eWOM

In order to help travelers and advertisers make wise decisions, Lobel trong Thuy Tran [18] has worked to develop advanced automated data-driven prediction algorithms that spot important trends and insights from extensive online customer evaluations. In a research, A. Navitha Sulthana [10] compared the importance of eWOM (electronic word-of-mouth) marketing to traditional marketing. More people sign up for online social networking sites as a result of eWOM. Michell, P. Yu. The author examined customer reviews, reviewer personality, review website characteristics, product review characteristics, environmental influence, and interpersonal to determine which factors had the most impact on buy intent. sA survey of 337 college students was conducted to gather data. The author found that purchase intent was positively impacted by six eWOM features. The most significant factor influencing a consumer's intention to purchase is customer reviews. Gilda Antonelli and Francesca Di Virgilio [19] claim that trust and eWOM act as moderating factors for the intention to make an online purchase. Web 2.0 technologies allow users to create and share user-generated content. internet word of mouth, or eWOM, is the dissemination of information about goods and services across many internet platforms [10].

C. Polar Reviews on Airlines

Negative sentiment expression can also be predicted by positive sentiment exposure. Both negative and positive customer evaluations of a service or product are a powerful marketing tool since they are made available to a large number of people and organizations online. The amount, quality, and timeliness of negative online evaluations have a favorable impact on consumers' opinions of usefulness and usability. We tried to gather information from social media i.e. Twitter, but we found that these platforms only provide data from the last several weeks. To quantify information about the words, the Text Parsing node examines a set of texts. We categorized the data as either positive or negative during collection depending on the customer rating value.

D. Conflicting Reviews

Wang claim that information about a product or service that only contains positive information or that simultaneously contains both positive and negative information will bore people's attention. In the past, it has been shown that mixed information such as conflicting assessments can lead to positive, negative, or ambiguous feelings. In Lee's opinion as they are subjective, online reviews might provide conflicting facts. According to earlier research, conflicting evaluations or other mixed information might produce unclear, negative, or positive feelings. These reviews cause a client to feel attitude ambivalence, which is a conflicted and bivalent emotional state.

E. Attitude Ambivalence

The current body of research ignores the influence of both online and offline factors on conflicting online reviews and unclear attitudes that influence consumer behavior. Univalent (i.e., positive or negative) evaluations document clashing or conflicted thoughts regarding a stimulus, in contrast to attitude, which records attitude ambivalence. Boukamcha [19] claimed in another important study that attitude ambivalence leads to a conflict between actions and attitude. More motivation and cognitive effort are needed to close the research gap. Ambivalence is the existence of both favorable and unfavorable assessments of an attitude object in a decision-making situation. It may result in an emotional state that is unpleasant.

III. MATERIALS AND METHODS

To improve the quality of service and to meet the expectations of customer, the industry must firstly identify its problems. In order to achieve this, businesses must first determine which tourist destinations result in poor publicity for their businesses. The study expands on the findings of the earlier research. to determine if the information presented in internet reviews will predict the services that reviewers would suggest. There is a study gap in the area of getting consumer suggestions from internet reviews. We get data from the website Skytrax, located at www.airlinequality.com, offers airline ratings and reviews. During the pre-processing stage, some components that have less of an impact are eliminated in order to increase the study's efficiency. The text is separated, and identified by a vector of conventional word weightings that show how much each word in the review is worth. In this study, we are building a recommendation system for travelers' convenience based on their online reviews. For this purpose, we developed a methodology, which is explained below as Fig.1

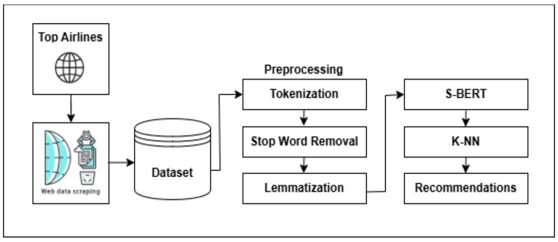
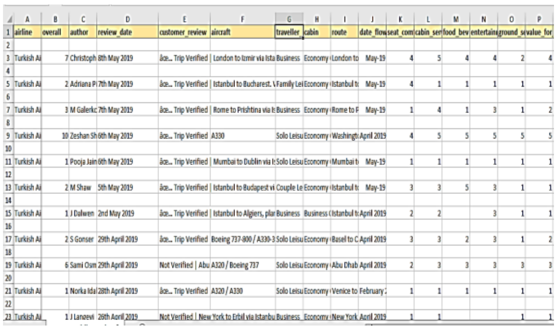


Figure 1. Methodology

A. Data Collection

Data collection is the process of gathering reviews, measuring it, and evaluating accurate research insights using developed, recognized methodologies. The top 10 websites are scraped for data collection including TripAdvisor Flights, Cheapair, Priceline, Skyscanner, Kayak, Expedia, Orbitz, Travelocity, Hotwire, and Momondo are some of these websites. This recommendation system is dealing with a huge volume of data. Due to the massive volume of data for this recommendation system, we removed unnecessary information and simplify the dataset to satisfy our requirements. By following to the state of the art 14 features are selected. Figure. 2 illustrate the sample of data.



airline	overall	author	review_date	customer_review	aircraft	traveller	cabin	route	date_flow	seat_com	cabin_ser	food_bev	entertain	ground_s	value_for
Turkish Ai	7	Christoph	8th May 2019	âœ… Trip Verified | London to Izmir via Ista		Business	Economy	London to	May-19	4	5	4	4	2	4
Turkish Ai	2	Adriana P	7th May 2019	âœ… Trip Verified | Istanbul to Bucharest. V		Family Lei	Economy	Istanbul t	May-19	4	1	1	1	1	1
Turkish Ai	3	M Galerkc	7th May 2019	âœ… Trip Verified | Rome to Prishtina via Is		Business	Economy	Rome to P	May-19	1	4	1	3	1	2
Turkish Ai	10	Zeshan Sh	6th May 2019	âœ… Trip Verified	A330	Solo Leisu	Economy	Washingt	April 2019	4	5	5	5	5	5
Turkish Ai	1	Pooja Jain	6th May 2019	âœ… Trip Verified | Mumbai to Dublin via I		Solo Leisu	Economy	Mumbai t	May-19	1	1	1	1	1	1
Turkish Ai	2	M Shaw	5th May 2019	âœ… Trip Verified | Istanbul to Budapest vi		Couple Le	Economy	Istanbul t	May-19	3	3	5	3	1	1
Turkish Ai	1	J Dalwen	2nd May 2019	âœ… Trip Verified | Istanbul to Algiers, plar		Business	Business	Istanbul t	April 2019	2	2		3	1	1
Turkish Ai	2	S Gonser	29th April 2019	âœ… Trip Verified	Boeing 737-800 / A330-3	Solo Leisu	Economy	Basel to C	April 2019	3	3	2	3	1	2
Turkish Ai	6	Sami Osm	29th April 2019	Not Verified | Abu	A320 / Boeing 737	Solo Leisu	Economy	Abu Dhab	April 2019	2	3	3	3	3	3
Turkish Ai	1	Norka Ida	28th April 2019	âœ… Trip Verified	A320 / A330	Solo Leisu	Economy	Venice to	February	1	1	1	1	1	1
Turkish Ai	1	J Langevi	26th April 2019	Not Verified | New York to Erbil via Istanbu		Business	Economy	New York	April 2019	1	1			1	1

Figure 2. Sample Dataset

The dataset used in this study was collected from publicly available airline review websites. Only textual reviews that were openly accessible without login restrictions were scraped. No Personally Identifiable Information (PII), such as usernames, email addresses, or booking details, was collected. The scraping process complied with the terms of use of the respective platforms, and all data were anonymized before analysis. This study strictly uses the dataset for academic and non-commercial research purposes, ensuring ethical handling and privacy preservation.

B. Data Description

The information from several airline websites makes up the dataset used in this study. Dataset is not publicly available it is scrapped from different websites and then converted into a csv file. Many factors, including airline, compartment, value for money, food, and beverage, are included in datasets. It consists of 14 characteristics and 131,895 records. Table 1 explains the features and the detail of dataset.

Table 1. Attributes and Description of Dataset

No	Attributes	Description
1	Review Data	Reviews might be conducted across, inside, or both.
2	Customer reaction	Select the airline or ignore it.
3	Aircraft Type	The Jets. Light jets. Medium-sized jets
4	Compartment	Good or bad
5	Route of Plan	Airline routes, destination
6	Date of Flight	Date of booking flight
7	Seat Comfort	Safety or comfort
8	Cabin Service	Good or bad
9	Food Beverage	Water, soft drinking, fast food
10	Entertainment	Listen to a music, play a game, write a journal, watch a movie
11	Traveler kind	Education platform, socially conscious, kids enjoy
12	Ground Service	Lounges cabin, baggage, presentation, executive
13	Value for Money	Economy or business class
14	Recommendat ion	Yes or no

C. Data Preprocessing

The next stage is Pre-Processing. One of the most important step of the procedure is this. A crucial step in creating a machine learning model is data pre-processing, and the effectiveness of this process determines the quality of the data. Machine learning models demonstrate improved classification accuracy when data is pre-processed. The dataset is subjected to lemmatization, tokenization, and stop word removal.

D. Data Analysis

Data analysis is the first and most important step in any prediction or recommendation process. Data

Analysis is the process of utilizing summary statistics and graphical representations to find patterns, anomalies, test theories, and validate presumptions. We are examining the airlines' frequency or values, which are displayed below in Fig 3.

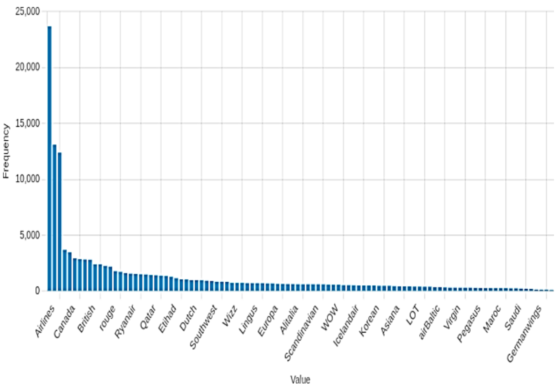


Figure 3. Frequency of Airlines

Despite EDA (Exploratory Data Analysis) is most frequently utilized while working with data to explore what data can tell us without the need for formal modeling or hypothesis testing; it may be utilized with or without a statistical technique. Statistical charts and other data visual analytics are frequently used in exploratory data analysis to highlight key elements of a data collection. All of this can be aided by exploratory data analysis. It helps you understand and make greater sense of data by eliminating anomalies and unnecessary numbers. Fig 4. shows the insights of data analysis.

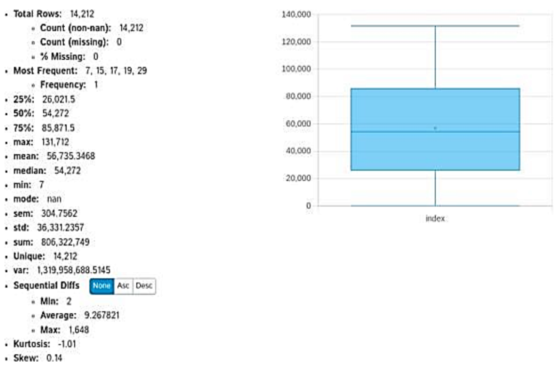


Figure 4. Data Insights

E. Data Cleaning

Removing unnecessary values and variables or data with abnormalities from your dataset, is referred to as data cleansing. These abnormalities have the potential to significantly bias the data, which could affect the results. The following actions can be taken to clean data:

i. Null Values Removal

A specific SQL marker known as a NULL value shows that the value is not present. In another way, it works mainly as a placeholder for unknown or missing information. Our dataset has a large number of null values, which we are eliminating using the cleaning values approach. There are 65,947 null values in our dataset. Fig 5. shows the sample of data after removal of null values.

Figure 5. Dataset after Removal of Null Values

ii. Descriptive Statistical Analysis

In some instances, descriptive analytics and inferential statistics were different. Simply said, descriptive statistics explain what the data shows or is. Descriptive statistics include factors like a variable's mean, standard deviation, and frequency. The characteristics of a sample or data collection are defined or summarized using descriptive statistics.

iii. Heat Map

The term "heat maps" refers to the use of color to represent magnitude in two-dimensional data visualizations. It would demonstrate to the reader how the event is set up or how it unfolds over time. The color shift might be in hue or intensity. Each row and column in a matrix represent a distinct phenomenon or category, and each cell's size is fixed. Fig 6. illustrate the heatmap of our dataset.

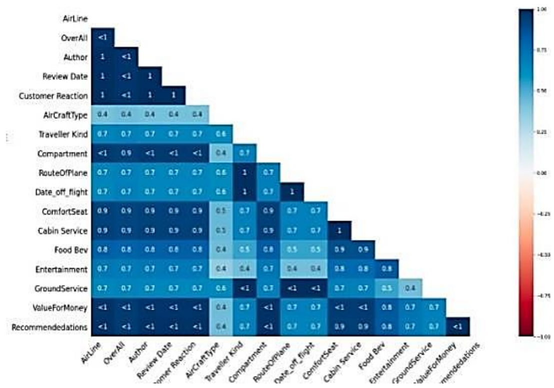


Figure 6. Heatmap

iv. Bar Chart

The bar chart is a type of graph or chart that shows data using rectangular bars whose lengths or heights correspond to the values they belong. It's normal to use bar graphs and bar charts interchangeably. The bars can be arranged in either a vertical or horizontal arrangement. Another name for a vertical bar chart is a column chart. The relationship between each attribute is depicted in a bar chart, such as the one shown below as Fig 7.

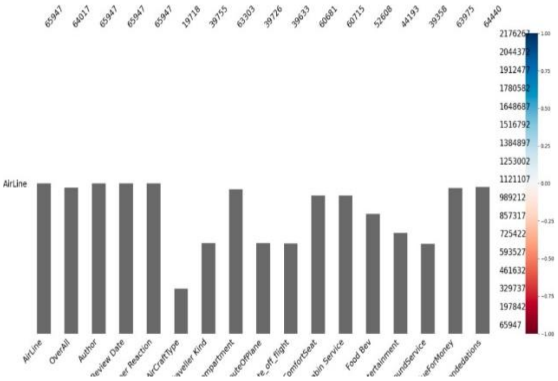


Figure 7. Bar Chart

v. Histogram

Graphical representations of a set of data points divided into user-specified ranges are called histograms. The histogram is similar to a bar graph and may be used to reduce data series into a visually understandable format. Fig 8. illustrate the histogram of dataset.

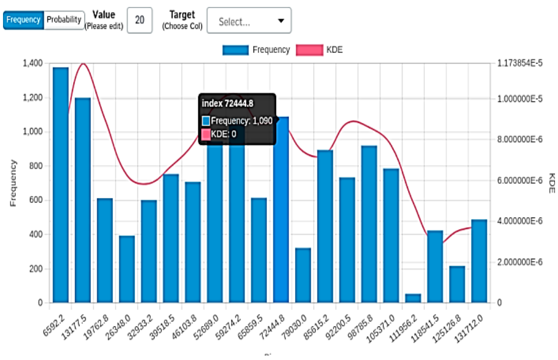


Figure 8. Histogram

IV. Results & Discussions

In this study to unveil ambivalence in airline reviews, we used two step approach by using S-BERT and K-NN classifier. Initially after pre-processing of dataset S-BERT is used to produce vectors of pre-processed tokens of reviews. It is done by generating high-dimensional embeddings of every review and capturing balanced conceptual meaning every review is converted into fixed-length vector. This made it possible to analyze reviews semantically using distance-based techniques. After generation of vectors, to find reviews that expressed different points of view K-NN method is used by examining the sentiment of each review's closest neighbors in the embedding space. If a review showed a mixture of feelings both positive and negative, or different sentiment polarity then it is considered as ambivalent. This approach offers more insight into complicated client experiences by effectively highlighting evaluations that included both compliments and criticism.

4.1 Training and Testing of Data

To avoid overfitting, in which the model learns the training data too well and performs poorly on new data, it is essential to keep training and testing data separate. Dataset is split into two subsets one is named as training set and second as testing set. Training set is used to train the model so that it can learn patterns of data and then testing set is used for the evaluating the performance of the model over unseen data. By testing the model's performance on unseen data, we get a more realistic measure. To ensure that our model is performing well on unseen data we adopted the K-Fold technique. It is a resampling technique in which dataset is divided into equal sized of folds. In this study we divided dataset into five equal- sized subgroups at random by utilizing 5-K fold cross validation. The model is constructed using the first four folds and then applied to the fifth fold. Each fold of the dataset is used as a validation set once during the five iterations of this process. Fig 9. illustrate the iterations of 5-K Fold cross validation.

Test-Train Split					
Split 1	Train 80%				Test 20%
Split 2	Train 80%			Test 20%	
Split 3	Train 80%		Test 20%		
Split 4		Test 20%	Train 80%		
Split 5	Test 20%	Train 80%			

Figure 9. 5-K Fold Cross Validation Method

In addition to applying 5-fold cross-validation for robust evaluation, we set the hyperparameter *k* in the K-Nearest Neighbors (K-NN) classifier to **5**. This means that the classification of each review is determined by considering the five most semantically similar reviews in the S-BERT embedding space. The majority class among these neighbors is then assigned as the predicted label. Choosing k = 5 provides a balanced trade-off, reducing the risk of overfitting that may occur with very small *k* values, while avoiding the excessive smoothing effect of larger *k* values.

4.2 Evaluation Matrix

Recent applications of machine learning and deep learning across various domains emphasize the importance of robust evaluation using accuracy, recall, F1-score, and AUC-ROC [23-34]. We used Accuracy, Recall, Precision, and F1 Score to evaluate the performance of our model. Accuracy is the ratio of accurately predicted instances to total predictions, which gives an estimate of how frequently the model generates accurate predictions. The ability of the model to determine every relevant instance of a given class is known as recall. For a fair evaluation we used F1 score that is the harmonic mean of precision and recall. This is very beneficial when working with imbalanced data sets. When used, all above metrics those offer an in-depth overview of the performance. All these metrics can be determined by using statistical formulas shown as Eq(i-iv).

$Accuracy = TP + TN / TP + TN + FP + FN$ Eq i

$Recall = TP / TP + FN$ Eq ii

$Precion = TP / TP + FP$ Eq iii

$F1 = 2(Precison.\ Recall) / \ Precision + Recall$ Eq iv

4.3 Results

As discussed earlier we used S-BERT for generation of vectors of pre-processed reviews. After that K-NN classifier is utilized for prediction. Fig 10. illustrate the confusion matrix of proposed model. Confusion matrix shows 960 True Positive, 720 True Negative, 48 False Positive and 46 False Negative predicted instances.

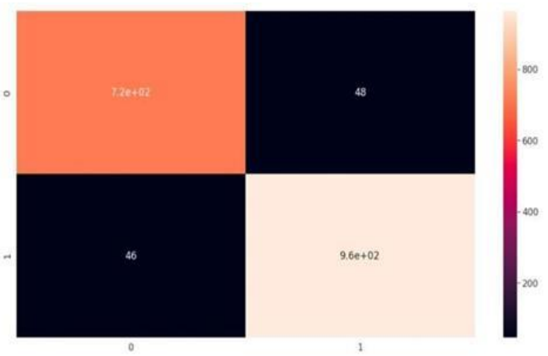


Figure 10. Confusion Matrix of K-NN

Table 2. Results

Matrix	Score
Accuracy	94.7 %
Recall	95.5%
Precision	95.24%
F1 Score	95.4%

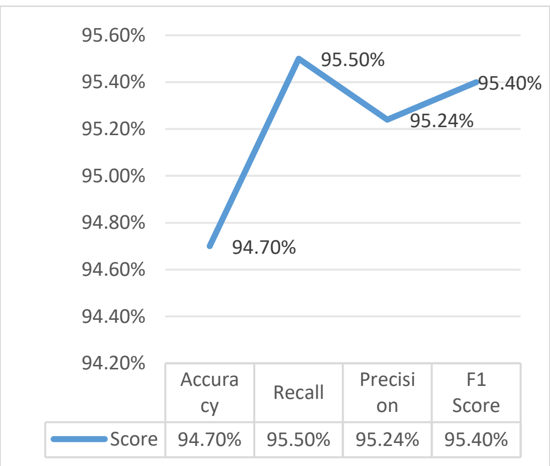


Figure 11. Comparison of Evaluation Metrics

Table.2 and Fig.11 illustrate the comparison of performance metrics. Our model demonstrate a balanced result on all the evaluation metrics. Proposed K-NN model achieved an accuracy of 94.7%, precision of 95.24%, F1 score of 95.4% and recall of 95.5% as highest score among all the indicators. These results shows that the proposed K-NN model is reliable and performed very well. However, there are some limitations of this model. K-NN can perform worst on a large scale dataset and is also sensitive to irrelevant features, class imbalance and noisy data.

Table 3. Comparison with Literature Review

Study	Methodology	Accuracy	Precision	Recall	F1-Score
Khan [5]	Sentiment classification (lexical + sentence structure)	82%	80%	78%	79%
Heidari & Rafatirad [4]	CNN + Transfer Learning (airline ticket recommendation)	85%	83%	82%	82%
Proposed Study	S-BERT + K-NN (airline review ambivalence & recommendation)	94.70 %	95.24 %	95.50 %	95.40 %

Table 3 presents a comparative summary of previous studies alongside the proposed approach. Khan [5] applied sentiment classification using lexical and syntactic features, achieving moderate results with an accuracy of 82% and an F1-score of 79%. Heidari & Rafatirad [4] employed CNN with transfer learning for airline ticket recommendation, which slightly improved performance, reaching an accuracy of 85% and F1-score of 82%. In contrast, our proposed S-BERT + K-NN framework significantly outperformed these methods, attaining 94.70% accuracy, 95.24% precision, 95.50% recall, and 95.40% F1-score. This demonstrates the robustness of semantic embeddings with S-BERT combined with K-NN in effectively detecting and analyzing ambivalence in airline reviews.

V. LIMITATIONS

Although the proposed S-BERT and K-NN framework achieves strong performance, several limitations remain. First, K-NN suffers from scalability challenges when applied to very large datasets, as distance computation becomes increasingly expensive. Second, the model's performance may be influenced by class imbalance, which can bias predictions toward the majority class. Finally, the approach may struggle to capture nuanced cases such as sarcasm, irony, or contradictory statements, which are common in online reviews. Addressing these challenges presents opportunities for future improvements through advanced sampling techniques, scalable architectures, and context-aware language models.

VI. Conclusion

The main goal of this study is to examine the reviews of different airlines used by tourists. We examined the process of the feature extraction process, the recommendation system, and the impact of feature extraction on performance. In this study S-BERT is used to generate vectors from preprocessed reviews and then K-NN model utilized for recommendation. Evaluation matrices such as F-measure, Precision, and Recall are used to improve outcomes. We a machine learning technique to suggests a suitable airline to customers. In this study, the machine learning model's recommendation accuracy is increased by this novel approach to using online social networks to advertise low-cost flights to travelers. Our proposed model achieved an accuracy of 94.7%, precision of 95.24%, F1 score of 95.4% and recall of 95.5%. In future more large and authentic datasets can be deployed over more modern and robust models for improved results.

References

[1] H. Arasli, M. B. Saydam, and H. Kilic, "Cruise travelers' service perceptions: a critical content analysis," Sustainability, Vol. 12, No. 17, p. 6702, 2020.

[2] J. Kuljanin and M. Kalić, "Exploring characteristics of passengers using traditional and low-cost airlines: A case study of Belgrade Airport," J. Air Transp. Manag., Vol. 46, pp. 12–18, 2015.

[3] A. Brochado, P. Rita, C. Oliveira, and F. Oliveira, "Airline passengers' perceptions of service quality: Themes in online reviews," Int. J. Contemp. Hosp. Manag., 2019.

[4] M. Heidari and S. Rafatirad, "Using transfer learning approach to implement convolutional neural network model to recommend airline tickets by using online reviews," in 2020 15th Int. Workshop Semantic and Social Media Adaptation and Personalization (SMA), IEEE, 2020, pp. 1–6.

[5] A. Khan, B. Baharudin, and K. Khan, "Sentiment classification from online customer reviews using lexical contextual sentence structure," in Int. Conf. Softw. Eng. Comput. Syst., Springer, 2011, pp. 317–331.

[6] R. P. de Oliveira, A. V. Oliveira, and G. Lohmann, "A Network-Design Analysis of Airline Business Model Adaptation in the Face of Competition and Consolidation," Transp. Sci., Vol. 55, No. 2, pp. 532–548, 2021.

[7] N. Donthu, S. Kumar, N. Pandey, and A. Mishra, "Mapping the electronic word-of-mouth (eWOM) research: A systematic review and bibliometric analysis," J. Bus. Res., vol. 135, pp. 758–773, 2021.

[8] R. Filieri, S. Alguezaui, and F. McLeay, "Why do travelers trust TripAdvisor? Antecedents of trust towards consumer-generated media and its influence on recommendation adoption and word of mouth," Tourism Manag., vol. 51, pp. 174–185, 2015.

[9] Y. Yang, "Characteristics and formation mechanism of continuous hazes in China: a case study during the autumn of 2014 in the North China Plain," Atmospheric Chemistry and Physics, vol. 15, No. 14, pp. 8165–8178, 2015.

[10] A. N. Sulthana and S. Vasantha, "Influence of electronic word of mouth eWOM on purchase intention," Int. J. Sci. Technol. Res., vol. 8, No. 10, pp. 1–5, 2019.

[11] A. Audrezet and B. Parguel, "Using the Evaluative Space Grid to better capture manifest ambivalence in customer satisfaction surveys," J. Retail. Consum. Serv., vol. 43, pp. 285–295, 2018.

[12] R. Filieri and F. McLeay, "E-WOM and accommodation: An analysis of the factors that influence travelers' adoption of information from online reviews," J. Travel Res., vol. 53, No. 1, pp. 44–57, 2014.

[13] D. Wang, F. L. Weisstein, S. Duan, and P. Choi, "Impact of ambivalent attitudes on green purchase intentions: The role of negative moods," Int. J. Consum. Stud., vol. 46, No. 1, pp. 182–199, 2022.

[14] G. D. Moody, P. B. Lowry, and D. F. Galletta, "It's complicated: explaining the relationship between trust, distrust, and ambivalence in online transaction relationships using polynomial regression analysis and response surface analysis," Eur. J. Inf. Syst., vol. 26, No. 4, pp. 379–413, 2017.

[15] L. Yang and H. R. Unnava, "Ambivalence, selective exposure, and negativity effect," Psychol. Mark., vol. 33, No. 5, pp. 331–343, 2016.

[16] M. Zhang, B. Fan, N. Zhang, W. Wang, and W. Fan, "Mining product innovation ideas from online reviews," Inf. Process. Manag., vol. 58, No. 1, p. 102389, 2021.

[17] Y. Jin, A. Compaan, T. Bhattacharjee, and Y. Huang, "Granular gel support- enabled extrusion of three-dimensional alginate and cellular structures," Biofabrication, vol. 8, No. 2, p. 025016, 2016.

[18] H. Zhang, H. Rao, and J. Feng, "Product innovation based on online review data mining: a case study of Huawei phones,"

Electron. Commer. Res., vol. 18, No. 1, pp. 3–22, 2018.

[19] Z. Qiao, X. Zhang, M. Zhou, G. A. Wang, and W. Fan, "A domain oriented LDA model for mining product defects from online customer reviews," 2017.

[20] M. U. Javeed, M. S. Ali, A. Iqbal, M. Azhar, S. M. Aslam, and I. Shabbir, "Transforming Heart Disease Detection with BERT: Novel Architectures and Fine-Tuning Techniques," in 2024 Int. Conf. Frontiers of Info. Tech. (FIT), Islamabad, Pakistan, 2024, pp. 1–6. doi: 10.1109/FIT63703.2024.10838424.

[21] M. Javeed, S. Aslam, F. M., and M. Iqbal, "An Enhanced Predictive Model for Heart Disease Diagnoses Using Machine Learning Algorithms," Tech. J., vol. 28, No. 4, pp. 64–73, 2023. [Online]. Available: https://tj.uettaxila.edu.pk/index.php/technical-journal/article/view/1828

[22] S. Aslam, M. Javeed, M. M. Iqbal, H. Ahmad, and A. Tariq, "Personality Prediction of the Users Based on Tweets through Machine Learning Techniques," J. Comput. Biomed. Inform., vol. 8, No. 2, 2025. [Online]. Available: https://www.jcbi.org/index.php/Main/article/view/796

[23] M. U. Javeed, S. Aslam, H. A. Sadiqa, A. Raza, M. Iqbal, and M. Akram, "Phishing Website URL Detection Using a Hybrid Machine Learning Approach," J. Comput. Biomed. Inform., vol. 9, No. 1, 2025. [Online]. Available: https://jcbi.org/index.php/Main/article/view/989

[24] M. U. Javeed, H. A. Sadiqa, M. Jaffar, S. Aslam, M. Khadim Hussain, Z. Raza, and M. Azhar, "A Deep Learning Approach For Securing Iot Systems With Cnn-Based Prediction of Worst-Case Response Time," Spectrum Eng. Sci., vol. 3, No. 7, pp. 376–385, 2025. [Online]. Available: https://www.sesjournal.com/index.php/1/article/view/599

[25] H. Shakeel, "LncRNAs Disease: A text mining Approach to Find the role of lncRNA in Aging," JCBI, vol. 9, No. 1, Jun. 2025.

[26] M. Jaffar, "ontology - based sentiment analysis for real-time product reputation modeling," SES, vol. 3, No. 7, pp. 648–667, Jul. 2025.

[27] "Deep Transfer Learning for COVID-19 Screening: Benchmarking ResNet50, VGG16, and GoogleNet on Chest X-Ray Images," IJACET, vol. 1, No. 2, pp. 69–83, Aug. 2025. [Online]. Available: https://ijacet.com/index.php/IJACET/article/view/18

[28] "Optimized Deep Convolutional Neural Network for Robust Occluded Facial Expression Recognition," ABBDM, vol. 5, No. 3, pp. 62–80, Jul. 2025. doi: 10.62019/dpfhnf43

[29] "Predicting Customer Loyalty from E-Commerce Reviews Using Aspect-Based Sentiment Analysis and ANN," ABBDM, vol. 5, No. 3, pp. 49–61, Jul. 2025. doi: 10.62019/3akt6733

[30] "Deep Learning in Hematology: Automated Counting of Blood Cells Using YOLOv5 Object Detection," IJACET, vol. 1, No. 3, Sep. 2025. [Online]. Available: https://ijacet.com/index.php/IJACET/article/view/21

[31] "Intelligent Image Gallery System Using Deep Learning for Automated Fruit and Vegetable Classification," IJACET, vol. 1, No. 3, Sep. 2025. [Online]. Available: https://ijacet.com/index.php/IJACET/article/view/20

[32] "A Hybrid Machine Learning Framework for Personalized News Recommendation," IJACET, vol. 1, No. 3, pp. 49–62, Sep. 2025. doi: 10.64796/qwjyd222

[33] A. Raza, S. Zongxin, G. Qiao, M. Javed, M. Bilal, H. H. Zuberi, and M. Mohsin, "Automated classification of humpback whale calls in four regions using convolutional neural networks and multi scale deep feature aggregation (MSDFA)," Measurement, vol. 118, p. 1038, 2025.

[34] A. Raza, M. Javed, A. Fayad, and A. Y. Khan, "Advanced deep learning-based predictive modelling for analyzing trends and performance metrics in stock market," J. Account. Fin. Emerg. Econom., vol. 9, No. 3, pp. 277–294, 2023.