Journal of Computer Science and Technology Studies
ISSN: 2709-104X
DOI: 10.32996/jcsts
Journal Homepage: www.al-kindipublisher.com/index.php/jcsts

JCSTS
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

Transforming Customer Experience in the Airline Industry: A Comprehensive Analysis of Twitter Sentiments Using Machine Learning and Association Rule Mining

Maliha Tayaba[1], Eftekhar Hossain Ayon[2] ✉ Md Tuhin Mia[3], Malay Sarkar[4], Rejon Kumar Ray[5], Md Salim Chowdhury[6], Md Al-Imran[7] Nur Nobe[8] Bishnu Padh Ghosh[9], MD Tanvir Islam[10] and Aisharyja Roy Puja[11]

[1]*Department of Computer Science, University of South Dakota, Vermillion, South Dakota, USA*
[2]*Department of Computer & Info Science, Gannon University, Erie, Pennsylvania, USA*
[3,9]*School of Business, International American University, Los Angeles, California, USA*
[4,11]*Department of Management Science and Quantitative Methods, Gannon University, USA*
[5]*Department of Business Analytics Business Analytics, Gannon University, USA*
[6,7]*College of Graduate and Professional Studies Trine University, USA*
[8]*Department of Healthcare Management, Saint Francis College, Brooklyn, New York, USA*
[10]*Department of Computer Science, Monroe College, New Rochelle, New York, USA*

Corresponding Author: Eftekhar Hossain Ayon, **E-mail**: ayon001@gannon.edu

| ABSTRACT

The airline industry places significant emphasis on improving customer experience, and Twitter has emerged as a key platform for passengers to share their opinions. This research introduces a machine learning approach to analyze tweets and enhance customer experience. Features are extracted from tweets using both the Glove dictionary and n-gram methods for word embedding. The study explores various artificial neural network (ANN) architectures and Support Vector Machines (SVM) to create a classification model for categorizing tweets into positive and negative sentiments. Additionally, a Convolutional Neural Network (CNN) is developed for tweet classification, and its performance is compared with the most accurate model identified among SVM and multiple ANN architectures. The results indicate that the CNN model surpasses the SVM and ANN models. To provide further insights, association rule mining is applied to different tweet categories, revealing connections with sentiment categories. These findings offer valuable information to help airline industries refine and enhance their customer experience strategies.

| KEYWORDS

Airline Industry; Twitter Sentiments; Machine Learning; Rule Mining

| ARTICLE INFORMATION

ACCEPTED: 01 December 2023 **PUBLISHED:** 22 December 2023 **DOI:** 10.32996/jcsts.2023.5.4.20

1. Introduction

In the dynamic landscape of the airline industry, enhancing customer experience has become a pivotal focus. With Twitter emerging as a prominent platform for air travelers to voice their opinions and feedback, this study employs a machine learning approach to analyze tweets, aiming to positively impact customer experience. The research involves feature extraction from tweets using both the Glove dictionary approach and the n-gram approach for word embedding. Subsequently, various (khan, 2022,2023,2023,2023) artificial neural network (ANN) architectures, Support Vector Machines (SVM), and Convolutional Neural Networks (CNN) are explored to construct a sentiment analysis classification model, categorizing tweets into positive and negative sentiments.

The study delves into the methodological aspects, detailing the collection of tweets from prominent airlines and the application of techniques such as n-gram models and GloVe for effective feature representation. Sentiment classification, employing SVM, ANN, and CNN, is discussed, with a notable emphasis on the superior performance of the CNN model. Additionally, association rule mining is applied to uncover connections between different tweet categories and sentiment classes, providing intriguing insights for refining customer experience strategies in the airline industry. The results and discussions section unveils the experimental outcomes, including the performance of SVM and ANN configurations on various word embeddings and the commendable performance of CNN on word embedding feature vectors. The study further investigates the impact of different parameters on passengers' emotions during flights through association rule mining.

Air travel stands out as a highly convenient mode of transportation for long-distance journeys, both domestically and internationally. The global presence of numerous airline service providers (ASPs) fosters a competitive environment, compelling airlines to vie for customer attention. Travelers, in turn, carefully consider various factors before choosing an airline, including airfare, travel duration, layovers, baggage allowances, and feedback from existing customers. In response, ASPs are actively enhancing their services across these dimensions, striving to elevate in-flight comfort and attract a broader customer base.

Understanding and addressing customer needs and satisfaction during flights is of paramount importance. Customer feedback serves as a crucial metric for the airline industry, providing insights into the passenger experience. Traditional methods of collecting feedback, such as onboard paper forms, often face challenges as passengers may be disinclined to fill them out, and the questions posed may be limited or biased. Alternative approaches involve online feedback forms on airline websites or mobile applications, post-journey email requests, or SMS messages prompting passengers to rate the service on specific parameters. However, these methods are constrained by predefined parameters. A more contemporary and user-friendly avenue for passengers to share their opinions is through social media platforms. Among these, Twitter emerges as a particularly popular and global platform. Leveraging information from Twitter can offer valuable data for the development of a recommender system. Moreover, passengers find social media, especially Twitter, to be a convenient space for expressing their travel experiences and opinions openly.

The findings underscore the significance of factors such as Cabin Crew Behavior (CCB) and Food Quality (FQL) in influencing sentiments. The study concludes by emphasizing the importance of understanding the underlying factors shaping sentiments and highlights the role of association analysis in offering valuable insights for organizations to enhance customer experience. Despite the valuable contributions, the study acknowledges certain limitations, particularly the exclusive focus on English-language tweets. The potential enrichment of the study by analyzing tweets in other languages is identified as a key area for future research. Overall, this research presents a comprehensive exploration of sentiment analysis in the airline industry, employing advanced machine learning techniques and association rule mining to uncover actionable insights for improving customer satisfaction.

2. Literature Review

Kumar et al. (2019) found that employing a machine learning methodology specifically tailored for tweet analysis can significantly contribute to enhancing the customer experience within the airline industry. The study utilized two distinct approaches, namely word embedding with the Glove dictionary method and the n-gram approach, to extract features from tweets. To develop a classification model for categorizing tweets into positive and negative sentiments, the study employed Support Vector Machines (SVM) and various architectures of Artificial Neural Networks (ANN). Additionally, Convolutional Neural Networks (CNN) were introduced as an alternative classification model, and their performance was compared with the most accurate model identified among SVM and multiple ANN architectures. The results demonstrated that CNN outperformed SVM and ANN models in terms of classification accuracy. Beyond sentiment analysis, the study conducted association rule mining on different tweet categories to establish connections with sentiment categories. The outcomes of this analysis revealed intriguing associations, providing valuable insights for airline industries seeking to enhance their customers' experiences. The study's comprehensive approach, integrating multiple machine learning techniques and association rule mining, offers a nuanced understanding of customer sentiments expressed on social media platforms, particularly Twitter.

Wu et al. (2023) conducted a study to evaluate customer satisfaction by analyzing sentiments expressed in tweets related to airlines. Relevant tweets were collected through Twitter's application programming interface and underwent tokenization and vectorization. These processed vectors were then fed into a pre-trained machine learning classifier to predict sentiments. In addition to sentiment analysis, the study included a lexical analysis of the gathered tweets to model keyword frequencies, providing valuable context for interpreting sentiments. Furthermore, time series methods, such as Bollinger Bands, were employed to detect anomalies in the sentiment data. By examining breakout points on the Bollinger upper and lower bounds, the methodology, based on historical records from January to July 2022, proved effective in capturing sudden and significant shifts in passenger sentiments. The developed methodology has the potential to evolve into an application that assists airlines and other customer-centric businesses in promptly identifying noteworthy changes in customer sentiments, enabling them to take timely and appropriate actions.

Rane et al. (2018) conducted research focusing on a dataset comprising tweets related to six major US Airlines, with a primary objective of performing multi-class sentiment analysis. The methodology involved initial preprocessing techniques to clean the tweets, followed by representing them as vectors using a deep learning concept, specifically Doc2vec, for phrase-level analysis. The analysis incorporated seven distinct classification strategies: Decision Tree, Random Forest, Support Vector Machine (SVM), K-Nearest Neighbors, Logistic Regression, Gaussian Naïve Bayes, and AdaBoost. The classifiers were trained using 80% of the data, with the remaining 20% utilized for testing. The results from the test set provided tweet sentiments (positive/negative/neutral). Accuracy calculations were conducted based on the obtained results to enable comparison among the various classification approaches, and an overall sentiment count was visually represented by aggregating data from all six airlines.

Pak et al. (2010) primarily focus on leveraging Twitter, widely recognized as the most prevalent microblogging platform, specifically for the task of sentiment analysis. The researchers outline the process of automatically compiling a corpus tailored to meet the objectives of sentiment analysis and opinion mining. They conduct a linguistic analysis of the compiled corpus, elucidating the identified phenomena. Using this corpus, the researchers develop a sentiment classifier capable of distinguishing positive, negative, and neutral sentiments within a document. Empirical assessments demonstrate the effectiveness of their proposed techniques, showing superior performance compared to methods introduced in prior studies. Although the investigation concentrated on the English language, it is emphasized that the methodology they propose is adaptable for use with any other language.

Deb et al. (2017) highlight the immense value of sentiment analysis for organizations aiming to assess and enhance their products and services. In the context of the airline industry, leveraging feedback from a discerning data source like Twitter enables sentiment analysis of customer experiences. The advantages of employing Twitter for sentiment analysis remain significant, given that consumers actively seek information on various aspects of daily life. This study specifically concentrates on classifying the sentiment of Twitter messages and presents results obtained through the implementation of a machine learning algorithm using R and Rapid Miner. The tweets undergo extraction and pre-processing before being categorized into neutral, negative, and positive sentiments. The Naive Bayes algorithm is applied to classify the sentiments of recent tweets concerning different airlines, and the results are comprehensively summarized.

3. Methodology

Tweets associated with several prominent airlines, each boasting substantial global followings, were collected from the Twitter server. To achieve this, a Python script utilizing the Twitter API was developed, employing the Tweepy package for retrieving tweets. The selected airlines for analysis encompass Qatar Airways (QTA), Sadia Airlines (SUA), Air New Zealand (ANZ), Turkish Airlines (TKA), JetBlue Airways (JBA), American Airlines (AMA), United Airlines (UTA), British Airways (BTA), Emirates (EMR), and Delta (DLT). Specific hashtags relevant to each airline were carefully chosen to facilitate tweet retrieval, and only English-language tweets were considered for this study. The data collection period extended from March 1, 2019, to March 11, 2019, based on Twitter server availability. Initially formatted in JSON (JavaScript Object Notation), the extracted tweets were converted to CSV format. The total tweet count across all airlines reached 155,000, with AMA having the highest share at 44.13% (Table 1). Following preprocessing, which involved the removal of retweets, the dataset was refined to 120,766 tweets for subsequent analysis.

Airline Name	QTA	SUA	ANZ	TKA	JBA	AMA	UTA	BTA	EMR	DLT
Number Of Tweet	20000	40000	6000	2000	3000	60000	10000	6500	4500	3000

Table 1: Total number of tweets from 1 Mar 2019–1 April 2019

3.1 Feature extraction

In sentiment analysis, a crucial step involves converting raw data into a numerical representation suitable for analysis. Given the Twitter context, where tweets are limited to 140 characters, we utilized n-gram models and the GloVe dictionary approach for word embeddings (WE) to preprocess the dataset. In the analysis of tweets, an effective representation requires converting text into a weighted feature vector. The n-gram model is especially valuable, providing a sequence of words that assigns probabilities to word sequences within a tweet, defined by a fixed window size denoted as 'n.' For example, consider the tweet: "The flight was so pleasant except that food quality can be improved." A 3-gram representation of this tweet would include sequences like "The flight was," "flight was so," "was so pleasant," and so on. It's important to note that a 1-gram model comprises individual words from the tweet. Another method employed in this study is word embeddings (WE), specifically utilizing the GloVe (Global Vectors for Word Representations) algorithm. GloVe is a count-based model that constructs a matrix of word co-occurrences, followed by dimensionality reduction to generate word vectors. Essentially, GloVe transforms the words in tweets into vectors, providing a nuanced understanding of semantic relationships.

3.2 Sentiment Classification and SVM
After extracting preprocessed features, the development of a sentiment analysis classification model requires the application of machine learning methods. Our decision to integrate three prominent approaches—Support Vector Machines (SVM), Artificial Neural Networks (ANNs), and Convolutional Neural Networks (CNN)—was based on a thorough examination of relevant literature. These methodologies were chosen due to their well-established effectiveness in handling sentiment classification tasks. Support Vector Machines (SVM) is a widely acknowledged machine learning technique extensively utilized in sentiment analysis [20, 35]. When dealing with binary class classification problems, SVM distinguishes between two classes using a margin threshold or hyperplane, effectively separating data points belonging to distinct classes. While multiple hyperplanes may exist, SVM aims to identify the optimal one that maximizes the margin between the two classes. In this study, SVM is selected as a sentiment analysis technique due to its widespread use and recognized efficacy.

3.3 Artificial Neural Networks (ANN) and Convolutional Neural Networks (CNN)
Artificial Neural Networks (ANNs) stand out as a widely employed technique for prediction and classification tasks, suitable for both numerical and categorical data. A typical ANN architecture comprises an input layer, one or more hidden layers, and an output layer. The number of neurons in each layer depends on the data type and features within the dataset. Each layer is associated with specific weights, which are multiplied with input values and then transmitted to subsequent layers. BPANN, a variant of ANN, incorporates backpropagation, involving propagating the prediction error (actual output - predicted output) backward to the preceding layers. This propagated error plays a crucial role in adjusting the weights across each layer, aiming to minimize prediction error and enhance accuracy. The effectiveness of this backpropagation approach is influenced by the number of times the network is trained on a specific dataset.

Convolutional Neural Networks (CNNs), originally designed for image data within deep neural networks, have proven their versatility by extending their application to various data types, including speech recognition. The generalized architecture of a CNN tailored for tweet classification is illustrated in Figure 2. In this framework, each tweet is initially represented as a word vector within an n-dimensional space. Assuming d represents the dimension of word vectors, and l denotes the length of the tweet (i.e., the number of concatenated words in the tweet), the tweet matrix's dimension can be defined as proposed by Collobert and Weston, who suggest a perspective where the text matrix is likened to an image matrix, allowing convolution with filters of a size analogous to the region size.

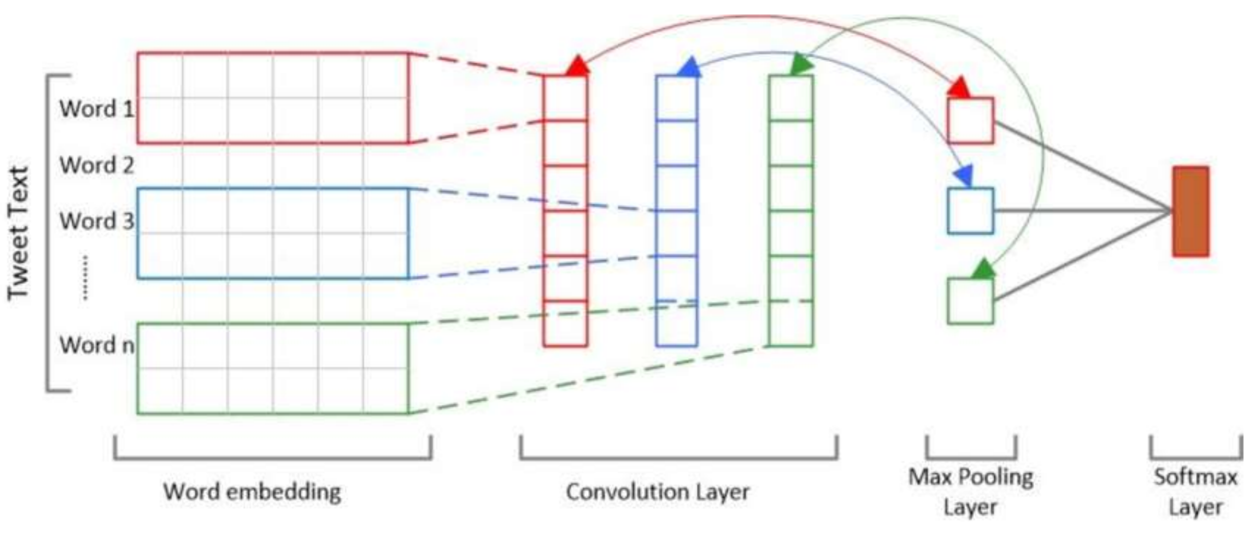

Fig 1: The general architecture of the CNN model

The creation of a comprehensive feature vector involves concatenating tweet vectors. These feature vectors are then combined across all filters, resulting in a feature map matrix. Afterward, a non-linear activation function is applied to the convolution layer parameters before being transmitted to the pooling layer. The primary objective of the max pooling layer is to reduce spatial dimensions, improve computational efficiency and address overfitting. The output from the pooling layer takes the form of a feature map matrix, where 'z' represents the length of the interval over which word vector elements were aggregated. These features are then forwarded to the subsequent layer, acting as a fully connected neural network model. The output from the final hidden layer undergoes the application of the ReLU (rectified linear unit) activation function, ultimately categorizing the respective tweet vector into either the positive or negative class.

This study also explores the examination of various parameters influencing passengers' emotions during flights. Association rule mining emerges as a notable method capable of extracting valuable associations related to different aspects of travel comfort. The Apriori algorithm is utilized to reveal associations between diverse travel-related issues that may impact a passenger's emotional experience during a flight. Association rule mining operates on a market basket-based approach, treating each data instance as a transaction with specific items purchased together. In this context, each tweet serves as a transaction, and the attributes are the categories of words used in the tweets, with each word in the tweet considered as an item. The tweets are categorized based on both positive and negative sentiments, and the Apriori algorithm is applied to the dataset to generate frequent item sets. Subsequently, association rules are formulated for both categories of emotions, namely positive sentiments and negative sentiments.

4. Results and discussion

Experiments were carried out using Python 3.6, utilizing the Scikit-learn library, along with the Keras and TensorFlow modules. A meticulously balanced dataset featuring two emotion categories (positive and negative) was prepared for analysis. Given the binary nature of sentiment classes in the dataset, an analysis was conducted using a linear Support Vector Machine (SVM) classifier, with a penalty parameter set to 1 during training. The performance of the SVM was evaluated on the vectorized tweets employing the Word Embedding (WE) approach. The dataset preparation involved the use of three types of word embeddings: trained, pre-trained, and hybrid. Various configurations of Artificial Neural Networks (ANN) were tested on the data to identify the optimal setup. Different ANN setups in conjunction with SVM were explored, and their performance was assessed using three-word embeddings derived from Glove representation and various dictionaries. Among the 10 ANN configurations used, Table 2 showcases the top-performing six configurations.

Configuration	No. of hidden layers	No. of neurons in layer 1	No. of neurons in layer 2	Batch size
ANN1	1	16	–	256
ANN2	1	32	–	256
ANN3	1	64	–	128
ANN4	2	16	4	256
ANN5	2	32	8	128
ANN6	2	64	16	64

Table 2: Employed ANN configurations.

Figure 3 depicts the testing of all configurations on pre-trained Word Embedding (WE) of tweets. Among these configurations, the highest accuracy was achieved by the ANN4 architecture, featuring 2 hidden layers with 16 and 4 neurons in layers 1 and 2, respectively, and trained with a batch size of 256. Consequently, only ANN4, along with SVM, was tested for the trained and hybrid sets. The results reveal that all ANN configurations with backpropagation surpassed SVM on the pre-trained dataset. Additionally, the performance of SVM and the optimal ANN4 configuration (determined by its performance on word embedding representations) was assessed for various n-gram representations of tweet data, as shown in Figure 4. Six n-gram models were employed to evaluate both ANN4 and SVM. The results demonstrate that the n-gram model with n=4 achieves superior performance for both SVM (88.5%) and ANN4 (89.4%).

Model	**ANN1**	**ANN2**	**ANN3**	**ANN4**	**SVM**	**ANN5**	**ANN6**
Pre-trained	63	67	68	69	60	64	65
Trained				85	86		
Hybrid				88.5	89.4		

Table 3: Performance of SVM and ANN models for different word embeddings

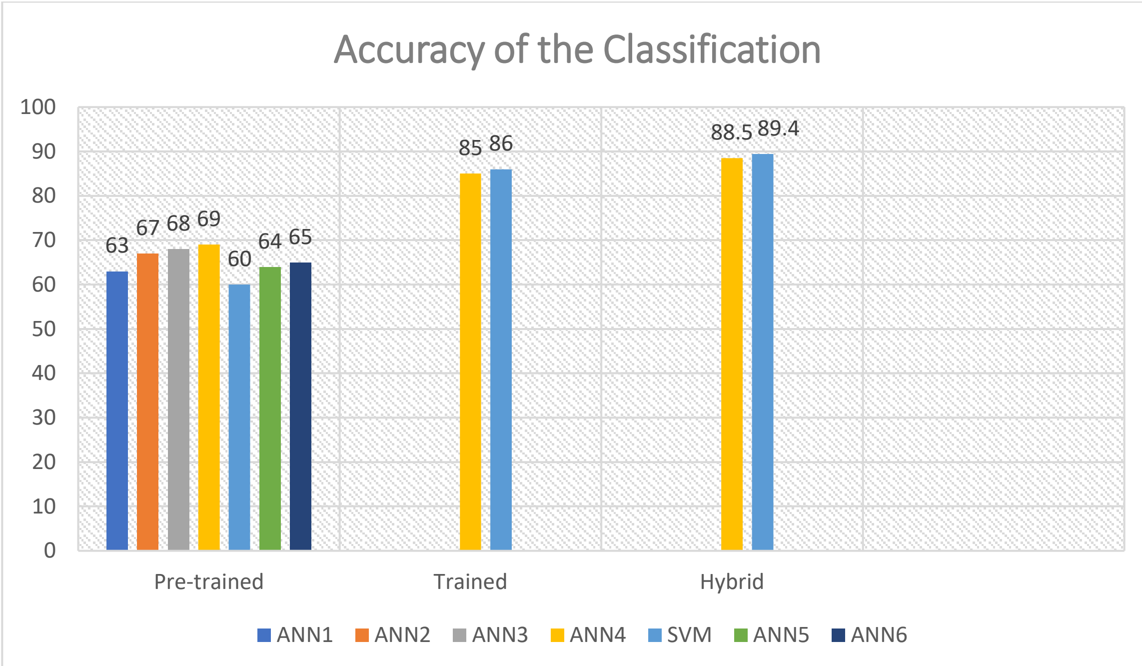

Chart 1: Performance of SVM and ANN models for different word embeddings

Moreover, an examination was undertaken utilizing Convolutional Neural Networks (CNN) on word embedding feature vectors. Various CNN configurations were formulated and assessed for sentiment analysis on tweets. The best performance was identified in a single hidden layer CNN with 128 neurons, 128 filters of sizes 3, 4, and 5, 128 dimensions of Word Embeddings (WE), a dropout rate of 0.5, and a batch size of 128. The CNN architecture's performance on the dataset is illustrated in Figure 2.

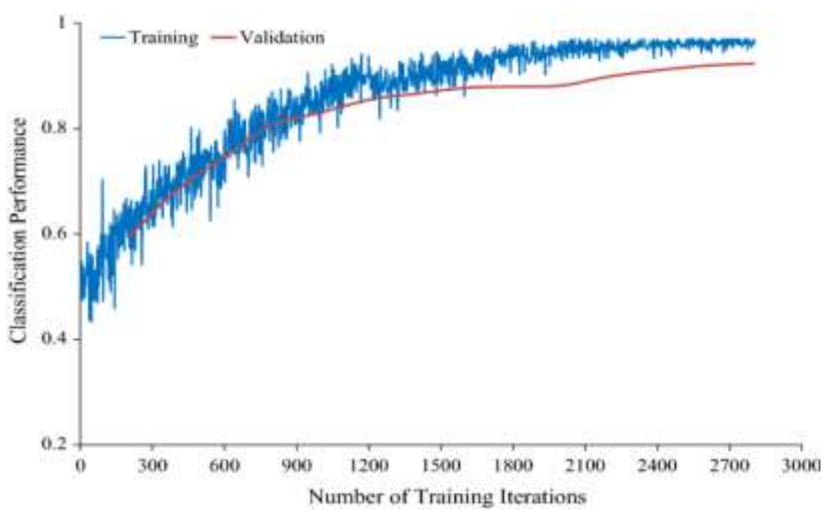

Fig 2: Classification performance of CNN on word embedding feature vectors

Figure 2 illustrates CNN's impressive performance on both the training and test datasets, achieving an accuracy of 92.3% after 2700 iterations on the validation set. This surpasses the results of the ANN4 model, which achieved the highest accuracy of 89.4% among the developed models. This reinforces that CNN demonstrates greater efficacy than traditional ANN, providing more accurate results even for text data. After this, an association analysis was conducted on the preprocessed words extracted from tweets. A dedicated dataset was curated for this analysis, where columns represent different word categories from the tweets, and each row corresponds to an individual tweet. To construct the dataset, words were assigned to various categories of services based on domain expert knowledge, considering multiple words within a tweet to ensure accurate categorization. The categories include Cabin Crew Behavior (CCB), Food Quality (FQL), Cleanliness (CLN), In-Flight Comfort (IFC), Flight Delayed/Cancelled (FDC), and Loss of Baggage (LOB). Table 3 presents these categories along with the respective number of tweets. The prepared dataset for association analysis comprised seven dimensions (six-word categories and one additional sentiment category: 1-positive, 0-

negative), as outlined in Table 3. Each column for every tweet can have a value of either 0 or 1, where 1 indicates the presence of words from a category, and 0 indicates their absence. Subsequently, association rules were generated separately for positive and negative sentiments to comprehend the involvement of each category in sentiment classes.

S. no.	**Category of words**	Number of tweets in the category
1	CCB	72,165
2	FQL	66,589
3	CLN	34,568
4	IFC	55,648
5	FDC	48,457
6	LOB	37847

Table 4: The categories of words and the number of tweets in which they occurred.

Furthermore, we identified strong rules by evaluating support, confidence, and lift values to clarify the influence of all categories on both negative and positive sentiments. Table 3 emphasizes that Cabin Crew Behavior (CCB) and Food Quality (FQL) have the most substantial impact on both negative and positive sentiments. In contrast, Loss of Baggage (LOB) and Cleanliness (CLN) demonstrate the least contribution to positive and negative sentiments, respectively.

Rule no.	**Association rule**	**Support**	**Confidence**	**Lift**
R1	[CCB = 0 AND FQL = 0] → [Sentiment = 0]	0.2	0.88	1.3
R2	[FQL = 0] → [CCB = 0]	0.3	0.95	1.53
R3	[FQL = 1] → [CCB = 1]	0.2	0.96	1.59
R4	[FQL = 1 AND CCB = 1] → [Sentiment = 1]	0.4	0.85	1.45
R5	[FDC = 1] → [LOB = 1]	0.4	0.63	1.05
R6	[FDC = 0 AND CCB = 0] → [IFC = 0]	0.3	0.89	1.54
R7	[IFC = 0] → [Sentiment = 0]	0.5	0.91	2.32
R8	[CCB = 0] → [Sentiment = 0]	0.4	0.85	1.43
R9	[LOB = 1] → [Sentiment = 0]	0.2	0.69	1.06
R10	[FQL = 1] → [Sentiment = 1]	0.2	0.61	0.92

Table 5: Relevant rules illustrating the association between different categories of words.

To investigate the relationship between different word categories, we examined the generated association rules for both positive and negative sentiments. A strong correlation was identified between Cabin Crew Behavior (CCB) and Food Quality (FQL). Table 5 presents the top 10 strong rules, providing insights into the correlation between different categories influencing sentiments.

The impact of Cabin Crew Behavior (CCB) and Food Quality (FQL) on the sentiment category is evident. Positive sentiments are frequently expressed in tweets when both food quality and cabin crew behavior are positive. Table 5, derived from R2 and R3, further emphasizes the strong correlation between food quality and cabin crew behavior. If food quality is subpar, there is a higher probability of negative cabin crew behavior, potentially due to an increase in passenger complaints. Conversely, satisfactory food quality tends to result in positive cabin crew behavior. Both R2 and R3 exhibit substantial confidence and lift values, reinforcing their significance.

Moreover, R4 highlights that favorable food quality and positive cabin crew behavior contribute to positive sentiments. Rule 5 suggests that flight delays or cancellations often lead to traveler complaints about the loss of baggage. Additionally, R6 indicates that in-flight comfort and cabin crew behavior jointly influence passenger emotions. R7 emphasizes that poor in-flight comfort frequently results in negative emotions, with a support value of 0.5. Beyond in-flight scenarios, the loss of baggage (R9) emerges as another significant contributor to negative emotions. R5 and R6 illustrate that flight delays, cancellations, or changes in flight times are associated with negative emotions. Additionally, in cases of flight delays, passengers observe rude behavior from airline staff, prompting them to post negative tweets.

In summary, while developing a sentiment classification model using a robust machine learning approach is essential, understanding the underlying factors influencing sentiments is equally crucial. Association analysis proves invaluable in achieving this goal, offering insights and suggestions for organizations to enhance customer experience.

Despite the valuable findings, certain limitations in the present study deserve discussion and consideration for future work. The study exclusively focuses on English-language tweets, overlooking the diverse array of tweets in various languages directed at different airlines on Twitter. Consequently, the outcomes of the study may not fully represent the opinions of individuals from various nationalities worldwide, as many people express their sentiments in languages other than English. Analyzing tweets in other languages would enrich this study, providing more comprehensive insights into customer satisfaction within the airline industry.

5. Conclusion

In conclusion, this study demonstrates the pivotal role of sentiment analysis in enhancing customer experience within the airline industry, leveraging Twitter as a primary platform for passenger opinions and feedback. The machine learning approach applied here, incorporating word embedding techniques and diverse neural network architectures, provides valuable insights into customer sentiments. The findings reveal the superiority of Convolutional Neural Networks (CNN) over Support Vector Machines (SVM) and Artificial Neural Networks (ANN) in classifying tweet sentiments accurately. Association rule mining adds another layer of depth to the analysis, uncovering meaningful correlations between various aspects of air travel, such as Cabin Crew Behavior (CCB), Food Quality (FQL), and In-Flight Comfort (IFC), and their impact on passenger sentiments. The study highlights the significance of factors like CCB and FQL in influencing both positive and negative sentiments.

The results contribute actionable insights for the airline industry to refine customer experience strategies. Understanding the interplay between different categories of service and passenger sentiments allows airlines to address specific pain points and enhance positive aspects of the travel experience. The study emphasizes the need for a comprehensive approach that combines sentiment analysis, machine learning, and association rule mining to derive nuanced and valuable insights from social media data.

However, the study acknowledges limitations, particularly the focus on English-language tweets. Future research should explore sentiments expressed in other languages to ensure a more inclusive understanding of global customer satisfaction. Overall, this research underscores the dynamic and multifaceted nature of customer sentiments in the airline industry, providing a foundation for further advancements in customer experience management.

Funding: This research received no external funding.
Conflicts of Interest: The authors declare no conflict of interest.

References

[1] Ahmed, A. H., Ahmad, S., Sayed, M. A., Ayon, E. H., Mia, T., & Koli, A. (2023). Predicting the Possibility of Student Admission into Graduate Admission by Regression Model: A Statistical Analysis. *Journal of Mathematics and Statistics Studies*, *4*(4), 97-105.
[2] Deb D D. (2017). IOP Conf. Ser.: Mater. Sci. Eng. 263 042067
[3] Haque, M. S., Amin, M. S., Miah, J., Cao, D. M., & Ahmed, A. H. (2023). Boosting Stock Price Prediction with Anticipated Macro Policy Changes. *Journal of Mathematics and Statistics Studies*, *4*(3), 29-34.
[4] Haque, M. S. (2023). Retail Demand Forecasting Using Neural Networks and Macroeconomic Variables. *Journal of Mathematics and Statistics Studies, 4*(3), 01–06. https://doi.org/10.32996/jmss.2023.4.3.1
[5] Haque M. S., Amin M. S., Ahmad S., Sayed M. A., Raihan A and Hossain M. A., (2023) Predicting Kidney Failure using an Ensemble Machine Learning Model: A Comparative Study, 2023 10th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI), Palembang, Indonesia, 2023, 31-37, doi: 10.1109/EECSI59885.2023.10295641
[6] Haque, M. S., Amin, M. S., Miah, J., Cao, D. M., & Ahmed, A. H. (2023). Boosting Stock Price Prediction with Anticipated Macro Policy Changes. *Journal of Mathematics and Statistics Studies*, *4*(3), 29-34.
[7] Khan, R. H., & Miah, J. (2022, June). Performance Evaluation of a new one-time password (OTP) scheme using stochastic Petri net (SPN). In *2022 IEEE World AI IoT Congress (AIIoT)* (pp. 407-412). IEEE.
[8] Kumar, S., and Zymbler, M. A. (2019) machine learning approach to analyze customer satisfaction from airline tweets. *J Big Data* **6**, 62 (2019). https://doi.org/10.1186/s40537-019-0224-1
[9] Khan, R. H., Miah, J., Rahman, M. M., & Tayaba, M. (2023, March). A Comparative Study of Machine Learning Algorithms for Detecting Breast Cancer. In *2023 IEEE 13th Annual Computing and Communication Workshop and Conference (CCWC)* (pp. 647-652). IEEE.

[10] Khan, R. H., Miah, J., Arafat, S. M. Y., Syeed, M. M. M., & Ca, D. M. (2023). Improving Traffic Density Forecasting in Intelligent Transportation Systems Using Gated Graph Neural Networks. arXiv preprint, arXiv:2310.17729.

[11] Khan, R. H., Miah, J., Arafat, S. M., Syeed, M. M., & Ca, D. M. (2023). Improving Traffic Density Forecasting in Intelligent Transportation Systems Using Gated Graph Neural Networks. *arXiv preprint arXiv:2310.17729*.

[12] Khan R. H., Miah J., Rahman M. M. and Tayaba M. (2023) A Comparative Study of Machine Learning Algorithms for Detecting Breast Cancer, 2023 IEEE 13th Annual Computing and Communication Workshop and Conference (CCWC), Las Vegas, NV, USA, 2023, pp. 647-652, doi: 10.1109/CCWC57344.2023.10099106.

[13] Miah, J., Cao, D. M., Sayed, M. A., & Haque, M. S. (2023). Generative AI Model for Artistic Style Transfer Using Convolutional Neural Networks. *arXiv preprint arXiv:2310.18237*

[14] Miah, J., Ca, D. M., Sayed, M. A., Lipu, E. R., Mahmud, F., & Arafat, S. M. Y. (2023). Improving Cardiovascular Disease Prediction Through Comparative Analysis of Machine Learning Models: A Case Study on Myocardial Infarction. arXiv preprint, 2311.00517.

[15] Md Abu S, Tayaba, M., Islam, M. T., Md Eyasin Ul I P, Md Tuhin M, Eftekhar H A, Nur N, & Bishnu P G. (2023). Parkinson's Disease Detection through Vocal Biomarkers and Advanced Machine Learning Algorithms. *Journal of Computer Science and Technology Studies*, *5*(4), 142–149. https://doi.org/10.32996/jcsts.2023.5.4.14

[16] Miah J., Khan R. H., Ahmed S., and Mahmud M. I., (2023). A comparative study of Detecting Covid 19 by Using Chest X-ray Images– A Deep Learning Approach, 2023 IEEE World AI IoT Congress (AIIoT), Seattle, WA, USA, 2023, pp. 0311-0316, doi: 10.1109/AIIoT58121.2023.10174382.

[17] Miah, J., Khan, R. H., Ahmed, S., & Mahmud, M. I. (2023, June). A comparative study of detecting covid 19 by using chest X-ray images–A deep learning approach. In *2023 IEEE World AI IoT Congress (AIIoT)* (pp. 0311-0316). IEEE.

[18] Pak, A., & Paroubek, P. (2010, May). Twitter as a corpus for sentiment analysis and opinion mining. In *LREc* (10, 2010, pp. 1320-1326).

[19] Rane A. and Kumar A., (2018) Sentiment Classification System of Twitter Data for US Airline Service Analysis, 2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC), Tokyo, Japan, 2018, 769-773, doi: 10.1109/COMPSAC.2018.00114.

[20] Sayed, M. A., Ahamed, S., Cao, D. M., Pavel, M. E. U. I., Sarkar, M., & Mia, M. T. (2023). Parkinson's Disease Detection through Vocal Biomarkers and Advanced Machine Learning Algorithms: A Comprehensive Study. arXiv preprint arXiv:2311.05435.

[21] Syeed, M. M., Khan, R. H., & Miah, J. (2021). Agile Fitness of Software Companies in Bangladesh: An Empirical Investigation. *International Journal of Advanced Computer Science and Applications*, *12*(2).

[22] Wu, S., & Gao, Y. (2023). Machine Learning Approach to Analyze the Sentiment of Airline Passengers' Tweets. *Transportation Research Record*, *0*(0). https://doi.org/10.1177/03611981231172948