Aspect-Oriented Sentiment Analysis For Airlines

Lakshmi Priya Irava
Department of Computer Science
Georgia State University
Atlanta, GA, USA
lirava1@student.gsu.edu

Mario M. Kubek
Department of Computer Science
Georgia State University
Atlanta, GA, USA
mkubek@gsu.edu

Abstract

This study utilizes five classification algorithms for aspect-oriented sentiment analysis on airline-related tweets, aiming to scrutinize opinions about diverse airline services. It includes preprocessing, aspect extraction, and classifier training, evaluating them on the complete Airline Sentiment dataset taken from Platform X. Performance metrics (precision, recall, accuracy, F1 score) are employed, providing insights into the effectiveness of categorization algorithms. The goal is to assist airlines in service enhancement based on sentiment analysis. Post-sentiment analysis, the project correlates sentiments with specific features, identifying positively and negatively rated components linked to relevant services. This phase offers airlines a profound understanding of customer feedback, facilitating targeted service improvements.

1 Introduction

Analyzing sentiments in text data is vital for comprehending emotions, opinions, and attitudes. The airline industry relies significantly on social media platform X (formerly known as Twitter) for customer service, making it a valuable resource for sentiment research (Pak and Paroubek, 2010). Unlike previous projects focusing solely on sentiment analysis without specific classifiers, our innovative approach introduces advancements. Employing diverse machine learning classifiers—Multinomial Naive Bayes, Multi-Layer Perceptron, XGBoost, Random Forest, and Support Vector Machine—this project pioneers comprehensive aspect-oriented sentiment analysis. The methodology involves robust preprocessing, aspect extraction, and extensive classifier training, evaluated on the complete Platform X Airline Sentiment dataset. The innovative post-sentiment analysis correlates sentiments with specific features, providing profound insights into customer feedback and facilitating targeted service improvements in the airline industry.

2 Methodology

Most sentiment analyses use classification, categorizing the target content as positive, neutral, or negative (Liu, 2012). Several scholars have used a variety of machine learning algorithms to examine sentiment analysis in previous years (Pak and Paroubek, 2010) (Kiritchenko and Mohammad, 2018). The following illustrates the organization of our sentiment analysis process (see Figure 1) (Pak and Paroubek, 2010). The process begins by collecting tweets from platform X, which are designated for both training and testing data. Subsequently, we perform preprocessing and data cleaning to enhance data quality. Following this, aspect extraction is conducted to identify relevant features. A model is then constructed to classify tweets into positive, negative, and neutral categories. The construction of the model involves utilizing various machine-learning classifiers. This comprehensive approach ensures effective sentiment classification analysis.

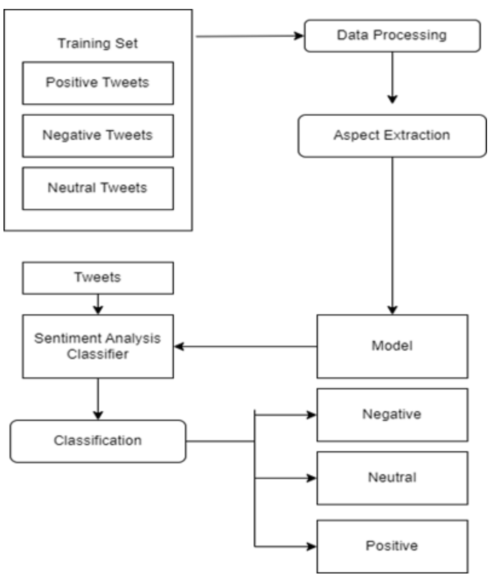

Figure 1: Sentiment Analysis Process Flowchart.

Classifier	Precision	Recall	F1 Score	Accuracy
Multinomial Naive Bayes (MNB)	0.75	0.76	0.76	0.76
Multilayer Perceptron (MPC)	0.74	0.73	0.73	0.73
XGBoost (XGB)	0.75	0.76	0.74	0.76
Random Forest (RFC)	0.76	0.77	0.76	0.77
Support Vector Machine (SVM)	0.77	0.78	0.77	0.78

Table 1: Classification Results of each classifier.

Despite the advancements introduced in our aspect-oriented sentiment analysis methodology, several challenges were encountered during the implementation. The complexity of discerning nuanced sentiments from diverse aspects of airline-related tweets posed a considerable obstacle. We navigated these challenges through rigorous preprocessing techniques, careful selection of features, and continuous refinement of the classifiers. These efforts contribute to the robustness of our approach, addressing the intricacies inherent in sentiment analysis on social media platforms.

3 Improving Sentiment Analysis

3.1 Sentiment Analysis Framework

The sentiment analysis framework begins with robust data preprocessing using regular expressions to eliminate stopwords, URLs, and special characters (Liu, 2012), followed by the transformation of data into a bag of words. Subsequently, the dataset is split into training and test sets (Liu, 2012).

3.2 Classifier Training and Evaluation

The classifiers are then trained on the prepared training set, and their performance is assessed using a comprehensive set of metrics, including accuracy, precision, recall, and F1 score. The results highlight the distinctive performance of each classifier, with SVM emerging as the top performer. SVM achieved an accuracy rate of 0.78, precision of 0,77 recall of 0.78, and an F1 score of 0.77. In contrast, the Multinomial Naive Bayes classifier demonstrated an accuracy of 0.76, precision of 0.75, recall of 0.76, and an F1 score of 0.76. The classification results, as shown in Table 1, elucidate the nuances of each classifier's performance. SVM's superiority, particularly in discerning elements like sarcasm and irony prevalent in social media discourse, is evident. The classifiers' varying performance underscores the importance of selecting an algorithm attuned to the intricacies of sentiment within the context of airline-related tweets.

Among the various aspects analyzed in our project, certain airline features stand out with impressive scores, reflecting their significance in shaping overall sentiment. Noteworthy aspects include 'comfort,' 'staff,' and 'cleanliness,' which consistently achieve commendable performance across classifiers. These aspects play a pivotal role in influencing sentiments, as indicated by their top scores. Future research into cutting-edge natural language processing methods and real-time data integration is encouraged by the methodology's success, which will have a big impact on the aviation sector as we continue to improve our understanding of and ability to respond to changing customer sentiments on social media platforms.

Future work involves integrating Large Language Models (LLMs) like BERT for enhanced aspect extraction and sentiment classification, offering a nuanced understanding of linguistic patterns in user opinions. Considering the rapid developments in the field, exploring pre-trained models on platforms like Hugging Face can be beneficial. These platforms often provide a variety of models that can be fine-tuned for specific tasks, including aspect-oriented sentiment analysis. However, it's essential to underline the inherent difficulty of aligning aspects and their sentiments, which remains a significant challenge in sentiment analysis tasks. This alignment intricacy is a key area that warrants further investigation and innovation to improve the accuracy and applicability of sentiment analysis in real-world scenarios.

Limitations

While our project aims to advance aspect-oriented sentiment analysis in the airline industry, it faces limitations. Reliance on publicly available data from Platform X may introduce biases, and the classifiers' performance depends on training data quality. The study acknowledges potential challenges in capturing evolving language trends and cultural context. Generalizing findings to specific airline services may be influenced by variations in user engagement and communication styles. Despite these limitations, our study lays the groundwork for future research to refine sentiment analysis methodologies for airline-related posts on Platform X.

Ethics Statement

This research aligns with the ACL Code of Ethics[1], emphasizing fairness, transparency, and accountability. Ethical standards were maintained while using Platform X data, ensuring privacy through anonymization. The study focuses on general sentiments about airline experiences while avoiding individual user identification. Acknowledging potential biases in machine learning, the research commits to responsible model training and evaluation.

References

M. Taboada, J. Brooke, M. Tofiloski, K. Voll, and M. Stede. 2011. Lexicon-Based Methods for Sentiment Analysis.

A. Pak and P. Paroubek. 2010. Twitter as a Corpus for Sentiment Analysis and Opinion Mining.

A. Bermingham and A. Smeaton. 2012. On using Twitter to monitor political sentiment and predict election results.

M. J. Bing Liu. 2012. Sentiment Analysis and Opinion Mining.

S. Kiritchenko and S. Mohammad. 2018. Examining the use of machine learning techniques for sentiment analysis.

[1] https://www.aclweb.org/portal/content/acl-code-ethics