

# **Project Proposal**

## **Sentiment Analysis of Social Media for Airline Customer Service Improvement**

### **Team Members :**

- Sujith Chandra Dara (A20528384)
- Bhargava Ramani Gorle (A20528476)

### **Project Title: Enhancing Airline Customer Service through Sentiment Analysis of Twitter Posts**

#### **Description of the Problem:**

Customer reviews on social media sites like Twitter have become a vital component of how businesses present themselves and develop their marketing plans in the modern era of social media. A closer look at tweet sentiments can provide airlines with valuable information on customer happiness and areas for development, as customer impression has a direct impact on customer loyalty and business success. Yet, it can be difficult to determine sentiment precisely due to the informal language of social media, which includes slang, typos, and the use of emojis. Further complicating the study is the need to discern between honest and caustic comments.

#### **Brief Survey of Existing Work:**

Natural language processing (NLP) and machine learning (ML) communities have extensively researched sentiment analysis, with applications ranging from political sentiment assessment to market analysis. The development of models that understand social media communication details, tone, and context has been the main focus of prior research. To increase sentiment analysis accuracy, research have particularly used a variety of machine learning (ML) techniques, such as Support Vector Machines (SVM), Naive Bayes classifiers, and deep learning approaches like Long Short-Term Memory (LSTM) networks and Bidirectional Encoder Representations from Transformers (BERT). Nonetheless, there is still a need for customized solutions due to the distinct vocabulary and context of the aerospace sector, making it difficult to apply these sophisticated models to the customer service domain.

#### **Preliminary Plan and Reference:**

##### **Milestones:**

1. **Data Collection and Preprocessing** (March 1-15, 2024): Download the "Twitter US Airline Sentiment" dataset from Kaggle and perform necessary preprocessing steps, including text normalization, handling of missing values, and data augmentation for unbalanced classes.
2. **Literature Review** (March 16-31, 2024): Conduct a thorough review of recent papers on sentiment analysis, focusing on methods applied to social media data.

3. **Model Development** (April 1-15, 2024): Implement various ML algorithms, including classifiers like SVM and Naive Bayes.
4. **Model Evaluation and Optimization** (April 16-22, 2024): Evaluate the models using accuracy, precision, recall, and F1-score metrics. Perform hyperparameter tuning and model optimization based on evaluation results.
5. **Analysis and Reporting** (April 23-27, 2024): Analyze the models' findings, identify trends in airline customer sentiment, and prepare the final project report and presentation.

## **References:**

- Pak, A., & Paroubek, P. (2010). Twitter as a Corpus for Sentiment Analysis and Opinion Mining. *LREC*.
- Wang, X., Wei, F., Liu, X., Zhou, M., & Zhang, M. (2011). Topic Sentiment Analysis in Twitter: A Graph-based Hashtag Sentiment Classification Approach. *Proceedings of the 20th ACM International Conference on Information and Knowledge Management*.
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *arXiv preprint arXiv:1810.04805*.