# NLP implementation for sentimental analysis

Vipul Mishra

August 8, 2024

## 1 Introduction

This report presents the performance analysis of a machine learning model designed to classify customer reviews of Singapore Airlines based on their ratings. The dataset includes reviews and corresponding ratings from 1 to 5, and the model's performance is evaluated using accuracy and a classification report.

## 2 Dataset Overview

The dataset comprises 2,000 reviews, each associated with a rating from 1 to 5. The following are the columns present in the dataset:

- **published_date**: The date and time when the review was published.
- **published_platform**: The platform on which the review was published (e.g., Desktop).
- **rating**: The rating given by the reviewer.
- **type**: The type of review (e.g., review).
- **text**: The content of the review.
- **title**: The title of the review.
- **helpful_votes**: The number of helpful votes the review received.

## 3 Model Performance

The model achieved an accuracy of **0.59**, indicating that it correctly classified 59% of the reviews in the test set.

### 3.1 Classification Report

Table 1 presents the detailed classification report, which includes precision, recall, and F1-score for each rating class.

## 4 Analysis

The model's performance varies significantly across different rating classes:

- The model performs best on the 5-star rating class, achieving a precision of 0.60 and a recall of 0.99, resulting in an F1-score of 0.74.
- The model struggles with classes 2, 3, and 4, where the F1-scores are very low, indicating poor performance in these classes.
- The low performance in these classes could be due to an imbalanced dataset, where the majority of reviews belong to the 5-star category.

Table 1: Classification Report

| Class | Precision | Recall | F1-Score | Support |
|---|---|---|---|---|
| 1 | 0.64 | 0.52 | 0.58 | 206 |
| 2 | 0.00 | 0.00 | 0.00 | 110 |
| 3 | 0.48 | 0.05 | 0.09 | 211 |
| 4 | 0.24 | 0.02 | 0.04 | 397 |
| 5 | 0.60 | 0.99 | 0.74 | 1076 |
| **Accuracy** | | 0.59 | | |
| **Macro avg** | 0.39 | 0.31 | 0.29 | 2000 |
| **Weighted avg** | 0.48 | 0.59 | 0.48 | 2000 |

# 5 Conclusion

While the model shows reasonable performance in predicting 5-star ratings, it performs poorly in predicting other ratings. To improve the model's performance, future work could involve balancing the dataset, applying different machine learning algorithms, or using more advanced techniques such as ensemble methods.