FACULTY OF COMPUTING AND INFORMATICS

BACHLEOR IN COMPUTER SCIENCE

SOCIAL MEDIA COMPUTING – CDS6344

TRIMESTER, Session 2024/2025

Customer Perceptions on Qatar Airways: A Sentiment Analysis Approach

By:

Name	ID	Section
Dharvin Shah Kumar bin Mohamad Shah Ravin	1211102532	TT5L
Tan Fu Shun	1211101407	TT5L

Github Link:
https://github.com/dharvin2003/SocialMediaComputingFinalAssignment.git

Acknowledgement

First of all, we are deeply thankful for successfully completing the final assignment within the given timeframe. A special thanks to our lecturer, Dr. Mohammad Shadab Khan, whose invaluable guidance, inspiring advice and consistent encouragement have played a pivotal role in the execution of this project. He has generously shared his vast academic and technical knowledge, which has been crucial in navigating this academic endeavour. We are immensely grateful for his mentorship throughout this semester.

Table of Content

Chapter 1: Introduction

In today's digital era, the rapid growth of user-generated content across social media, travel forums and airline review websites has significantly influenced how the public perceives brands and services. For the aviation industry, where customer experience is a critical competitive differentiator, understanding and acting on these perceptions is more important than ever. Qatar Airways, one of the world's leading airlines known for its luxury, hospitality and global connectivity, consistently attracts thousands of customer reviews and comments spanning a range of experiences and opinions.

These online expressions, whether in the form of tweets, Facebook comments, TripAdvisor reviews or airline-specific feedback platforms, offer valuable, real-time insights into passengers' satisfaction levels and pain points. Traditional customer satisfaction surveys, while still useful, often fail to capture the immediacy, volume and richness of feedback available through digital channels. Therefore, implementation of technology to analyze this vast and unstructured textual data is essential for airlines that aim to stay ahead in service quality and brand reputation.

This project focuses on analyzing customer perceptions of Qatar Airways through Natural Language Processing (NLP). Specifically, we employed sentiment analysis techniques to detect and categorize the emotional tone behind customer opinions. With the integration of both traditional machine learning models and deep learning as well as transformer-based architectures like BERT, we aim to create a system that classifies sentiments into positive, negative or neutral categories with high accuracy. Additionally, this project also seeks to identify the underlying themes and aspects that drive customer satisfaction or dissatisfaction. Analysis of customer perception on specific service aspects provides Qatar Airways with actionable insights for strategic improvements.

1.1 Project Overview

This project is centered on analyzing public sentiment toward Qatar Airways by leveraging Natural Language Processing (NLP) techniques. The primary goal is to develop an end-to-end sentiment analysis pipeline capable of extraction, classification and interpretation of customer feedback from online sources such as review websites and social media platforms.

Therefore, by using a combination of traditional machine learning algorithms and advanced transformer-based models like BERT, the system will classify customer reviews into sentiment categories such as positive, negative or neutral. In addition, the project aims to identify specific service aspects that customers frequently mention, such as cabin crew behavior, seat comfort, flight punctuality and in-flight amenities.

The final output will include visualizations of sentiment distributions, word clouds and aspect-based sentiment trends that provides Qatar Airways with actionable insights. This project demonstrates how modern AI techniques can be effectively applied to understand customer perceptions at scale and support data-driven decision-making in the airline industry.

Chapter 2: Problem Statement

In the highly competitive airline industry, customer experience plays a pivotal role in shaping brand loyalty and influencing future travel decisions. Qatar Airways, a globally recognized carrier, receives a substantial volume of customer feedback through online review platforms and social media. However, the sheer scale and unstructured nature of this textual data make manual analysis both impractical and inefficient.

This project addresses the need for a systematic and scalable approach which interprets customer feedback by leveraging the Qatar Airways Review Dataset through advanced Natural Language Processing (NLP) techniques. By applying sentiment analysis and machine learning models including transformer-based architectures like BERT, the project seeks to automatically classify and analyze customer sentiments and opinions.

The core problem lies in identifying the specific factors that drive passenger satisfaction or dissatisfaction and translating those insights into strategic recommendations. Key challenges include detecting sentiment polarity, isolating aspect-specific sentiments and accurately capturing implicit opinions embedded in natural language. The goal is to enable Qatar Airways to transform raw feedback into actionable intelligence, thereby enhancing service delivery, refining customer experience strategies and making data-driven decision-making across the organization.

Chapter 3: Literature review

Sentiment analysis has emerged as a powerful technique in the field of Natural Language Processing (NLP) because it enables organizations to extract opinions, emotions and attitudes from unstructured textual data. In the airline industry, where customer experience significantly influences loyalty and reputation, sentiment analysis serves as an essential tool to monitor and understand public perception.

Numerous studies have explored sentiment analysis applications in evaluating airline services. (Kang et al., 2021) compared several machine learning and deep learning models, but came to a conclusion that BERT outperformed traditional classifiers. This is because traditional classifiers such as Naïve Bayes and Support Vector Machines, achieved an accuracy of 86% when analyzing Malaysian airline reviews. Their findings emphasized the effectiveness of transformer-based models in capturing context and sentiment polarity more accurately than conventional techniques.

(Idris et al., 2023) presented a conceptual framework that connects service quality, customer sentiment and satisfaction. Their review highlighted how sentiment analysis can uncover insights into key service aspects such as responsiveness, assurance and empathy especially when comparing full-service carriers (FSCs) like Qatar Airways with low-cost carriers (LCCs). They stressed the role of sentiment as a mediating factor influencing customer satisfaction and loyalty.

Moreover, deep learning methods like BiLSTM and CNN exhibit strong performance in sentiment classification. However, transformer models such as BERT and RoBERTa have set new benchmarks due to their ability to encode complex language patterns with minimal feature engineering. (Hasib et al., 2021) demonstrated the success of multi-class sentiment analysis on U.S. airline tweets using deep learning and word embeddings, to further validate the applicability of these techniques in diverse datasets.

In addition, (Patel et al., 2023) highlighted the value of sentiment analysis not just in understanding customer satisfaction but also in benchmarking competitor airlines. Their study compared multiple classifiers and affirmed that BERT consistently achieved the highest accuracy, precision, recall and F1 scores when analyzing customer feedback across platforms like Skytrax and Twitter.

Despite the success of these models, there are some challenges. Sentiment analysis in the airline domain must address issues such as sarcasm, context dependency and class imbalance. Preprocessing techniques, including emoji conversion and advanced tokenization, have been employed to mitigate these challenges.

Overall, existing literature confirms that sentiment analysis is not only feasible but also a highly effective approach to assessing airline service quality. Subsequently, by application of these techniques to customer reviews of Qatar Airways, this study builds upon established research to provide deeper insights into public perception, with the goal of enhancing service delivery and strategic planning.

Chapter 4: Methodology

Data Preprocessing

Preprocessing is a critical step in preparing textual data for sentiment analysis and modeling. The raw customer review data from the Qatar Airways dataset underwent several systematic steps to ensure that the inputs were clean, standardized and feature rich.

1. Handling Missing and Duplicate Values

The dataset was loaded using pandas and initially inspected for missing entries and duplicates.

Missing **numeric values** were replaced with the column mean.

Missing **categorical values** were inputted using the mode.

```
# Check for missing values in the dataset
missing_values = df.isnull().sum()
print("Missing values in each column:\n", missing_values)

Missing values in each column:
 Unnamed: 0          0
Date Published      0
Rating              1
Max Rating          1
Title               0
Author              0
Country             0
Date                0
Review Body         0
Type Of Traveller   435
Seat Type           0
Route               438
Date Flown          444
Recommended         1781
Aircraft            1088
Verified            1117
dtype: int64
```

```
# Handling missing values
for col in df.columns:
    if pd.api.types.is_numeric_dtype(df[col]):
        df[col] = df[col].fillna(df[col].mean())
    else:
        df[col] = df[col].fillna(df[col].mode()[0])

print("\nMissing values after handling:\n", df.isnull().sum())

Missing values after handling:
 Unnamed: 0          0
Date Published      0
Rating              0
Max Rating          0
Title               0
Author              0
Country             0
Date                0
Review Body         0
Type Of Traveller   0
Seat Type           0
Route               0
Date Flown          0
Recommended         0
Aircraft            0
Verified            0
dtype: int64
```

Duplicate rows were identified using df.duplicated() and removed to ensure data integrity.

```
# Check for duplicate values
duplicates_count = df.duplicated().sum()
print("\nDuplicate Records: ", duplicates_count)

Duplicate Records:  0
```

2. Sentiment Mapping from Ratings

A function was implemented to categorize numerical ratings into sentiment labels:

Ratings ≤ 2: **Negative**

Rating = 3: **Neutral**

Ratings ≥ 4: **Positive**

This transformed the numerical Rating column into a target Sentiment column for supervised learning.

```python
def map_sentiment(rating):
    if rating <= 2:
        return 'negative'
    elif rating == 3:
        return 'neutral'
    else:
        return 'positive'

df['Sentiment'] = df['Rating'].apply(map_sentiment)
df[['Rating', 'Sentiment']].head()
```

	Rating	Sentiment
0	1.0	negative
1	1.0	negative
2	1.0	negative
3	10.0	positive
4	7.0	positive

3. Text Normalization

The review text was cleaned to standardize and simplify the textual inputs:

- Converted all text to lowercase.
- Removed punctuation, numbers and special characters using regular expressions.
- Trimmed extra whitespace for consistency.

```python
def clean_text(text):
    text = str(text).lower()
    text = re.sub(r'[^a-zA-Z\s]', '', text)
    text = re.sub(r'\s+', ' ', text).strip()
    return text

df['Cleaned'] = df['Review Body'].apply(clean_text)
df[['Cleaned', 'Review Body']].head()
```

	Cleaned	Review Body
0	the delay of my flight from haneda to doha cau...	The delay of my flight from Haneda to Doha ca...
1	they convinced me that i needed to pay to add ...	They convinced me that I needed to pay $1500...
2	i have sent emails and have only received auto...	I have sent 5 emails and have only received ...
3	we flew on probably the first a the airline re...	We flew on probably the first A380 the airlin...
4	service was ok pretty good on my aisle and ext...	Service was ok, pretty good on my aisle and ...

4. Tokenization, Stopword Removal and Lemmatization

Tokenization was performed using NLTK to split reviews into individual words.

Stopwords (common, low-value words like “the”, “is”) were removed using NLTK’s stopword corpus.

Lemmatization was carried out using TextBlob to reduce words to their root forms (e.g., “running” → “run”) for consistency and dimensionality reduction.

These preprocessing steps are essential for ensuring data quality and reliability before moving on to feature engineering and modeling.

Feature Engineering

In the feature engineering phase, TF-IDF vectorization was applied to the lemmatized review text to convert it into numerical features based on word importance, using both unigrams and bigrams with a feature limit of 2000 to balance informativeness and sparsity. To enhance this representation, three additional semantic features—polarity, subjectivity and word count—were extracted from each review using TextBlob. These handcrafted features capture the sentiment tone, level of opinion and review length. Finally, the TF-IDF matrix and custom features were combined using horizontal stacking, creating a rich, hybrid feature set used as input for traditional machine learning models.

Modeling

The modeling phase involved training and evaluating multiple machine learning and deep learning algorithms for sentiment classification. The modeling workflow was divided into two parts: traditional models and advanced deep learning/transformer models.

1. Traditional Machine Learning Models

To establish baseline performance, classical classifiers were applied using scikit-learn on preprocessed data:

- **Algorithms used**:
  - Logistic Regression
  - Support Vector Machine (SVM)
  - Naïve Bayes
- **Features used**:
  - TF-IDF vectors extracted from lemmatized review text.
  - Optionally combined with engineered features like polarity, subjectivity and word count.
- **Evaluation Metrics**:
  - Accuracy, Precision, Recall, F1-Score
  - Confusion matrix to analyze misclassifications

2. Deep Learning – BiLSTM

A Bidirectional Long Short-Term Memory (BiLSTM) model was implemented using TensorFlow/Keras for sequence-based sentiment classification:

- **Preprocessing**:
  - Used Keras Tokenizer to convert text into sequences
  - Padded sequences to a fixed maximum length (e.g., 200)
- **Model Architecture**:
  - Embedding layer (with vocab size and sequence length)

  - BiLSTM layer
  - Dense output layer with sigmoid activation (binary classification)
- **Training Setup**:
  - Optimizer: Adam
  - Loss Function: Binary Cross entropy
  - Metrics: Accuracy

3. **Transformer-Based Modeling**

To exploit contextual understanding in sentiment classification, state-of-the-art transformer models were fine-tuned using the HuggingFace Transformers library. These models include DistilBERT, BERT and RoBERTa, each providing deep language representation capabilities.

DistilBERT

DistilBERT is a distilled, lightweight version of BERT, designed to be faster and smaller while preserving most of BERT's performance.

- **Tokenizer**:
  DistilBertTokenizerFast was used to tokenize and encode review texts into subword token IDs, padding them to a maximum sequence length.
- **Model**:
  DistilBertForSequenceClassification was used with a classification head for 3 sentiment classes (positive, neutral, negative).
- **Training & Evaluation**:
  - Custom Dataset class was used to feed data through DataLoader.
  - Optimized using the AdamW optimizer.
  - Evaluated using accuracy and a detailed classification report (precision, recall, F1-score).

BERT

BERT (Bidirectional Encoder Representations from Transformers) is a foundational language model pre-trained on a large English corpus using masked language modeling.

- **Tokenizer**:
  BertTokenizerFast was applied to preprocess the text by converting it into token IDs and attention masks suitable for BERT input.
- **Model**:
  BertForSequenceClassification was configured with num_labels=3 for multi-class classification.
- **Training & Evaluation**:
  - The model was trained on padded sequences using PyTorch's DataLoader with AdamW.
  - Results were benchmarked using accuracy and confusion matrix to compare with other models.

RoBERTa

RoBERTa (A Robustly Optimized BERT Approach) is an improved variant of BERT trained with more data and dynamic masking.

- **Tokenizer**:
  RobertaTokenizerFast was used to tokenize the review texts while adapting to RoBERTa's special token format (e.g., <s>, </s>).
- **Model**:
  RobertaForSequenceClassification with num_labels=3 was implemented.
- **Training & Evaluation**:
  - Used the same PyTorch pipeline for data loading and training.
  - Achieved strong results due to enhanced training setup and better language representation.

Chapter 5: Sentiment Analysis

Sentiment analysis was performed to classify airline reviews into three categories: positive, neutral and negative. The labeled sentiment classes were derived from numerical ratings, where low ratings indicated negative sentiment, medium ratings indicated neutral sentiment and high ratings indicated positive sentiment. To train sentiment classification models, the lemmatized review texts were first vectorized using TF-IDF, which captured the importance of unigrams and bigrams across the corpus.

A traditional machine learning pipeline was implemented using scikit-learn. The dataset was split into training and testing sets using an 80/20 ratio. A Multinomial Naïve Bayes classifier was trained on the TF-IDF features and predictions were evaluated using a classification report and confusion matrix. The performance was measured using standard metrics including accuracy, precision, recall and F1-score to assess how well the model captured each sentiment class. Additionally, class imbalance handling techniques such as SMOTE (Synthetic Minority Oversampling Technique) were imported and may be applied if necessary to improve classifier robustness.

This classical approach established a baseline for performance comparison with more advanced models like BiLSTM and transformer-based architectures, which were evaluated in the following chapters.

Chapter 6: Transformers / Deep Learning models

To enhance sentiment classification beyond traditional methods, deep learning and transformer-based architectures were implemented. These models leverage word order, contextual meaning and long-term dependencies, offering superior performance on text classification tasks.

Deep Learning Models

CNN (Convolutional Neural Network)

A 1D Convolutional Neural Network was employed to extract local patterns and n-gram-like features from the sequence of word embeddings. The model architecture included:

- An embedding layer to convert tokenized text into dense vector representations,
- One or more 1D convolutional layers followed by max-pooling layers to capture spatial hierarchies in text,
- A final fully connected (dense) layer with softmax or sigmoid activation depending on binary or multi-class classification.

CNN models are particularly effective for short and structured texts and provide competitive performance with relatively fast training time. The model was compiled with binary_crossentropy or categorical_crossentropy depending on the output and optimized using Adam. Evaluation was done using accuracy and F1-score.

BiLSTM (Bidirectional LSTM)

A BiLSTM model was built using Keras to capture sequential and bidirectional dependencies in the review text. The preprocessing involved converting lemmatized reviews into integer sequences using Keras' Tokenizer, followed by padding to a fixed maximum sequence length. The model architecture consisted of an embedding layer, a bidirectional LSTM layer and dense layers with a sigmoid output for binary classification (positive vs. negative sentiment). The model was compiled with the Adam optimizer and binary cross-entropy loss. Evaluation was performed on a test set using accuracy and F1-score to benchmark its performance against classical ML approaches.

Transformers

To leverage state-of-the-art language modeling capabilities, three pretrained transformer models were fine-tuned using the HuggingFace Transformers library:

BERT

BERT (Bidirectional Encoder Representations from Transformers) was fine-tuned for multi-class sentiment classification. The BertTokenizerFast was used to encode text inputs and BertForSequenceClassification was initialized with num_labels=3. Training and testing were conducted using PyTorch DataLoaders with the AdamW optimizer. BERT served as a strong benchmark model for understanding semantic nuances in the text.

DistilBERT

DistilBERT, a smaller and faster variant of BERT, was also fine-tuned using the same dataset. It retained most of BERT's performance while requiring fewer resources, making it efficient for deployment. Tokenization and modeling followed the same pipeline as BERT, using DistilBertTokenizerFast and DistilBertForSequenceClassification.

RoBERTa

RoBERTa (A Robustly Optimized BERT Pretraining Approach) was included to test a more advanced BERT variant trained with improved techniques. It used RobertaTokenizerFast and RobertaForSequenceClassification, with similar training and evaluation procedures. RoBERTa demonstrated strong capability in capturing nuanced sentiment, particularly for subtle or mixed reviews.

Chapter 7: Result & Visualization

After training multiple machine learning and deep learning models for sentiment classification, each model was evaluated using a combination of accuracy, precision, recall, F1-score and confusion matrices to understand their predictive performance across sentiment categories (positive, neutral, negative).

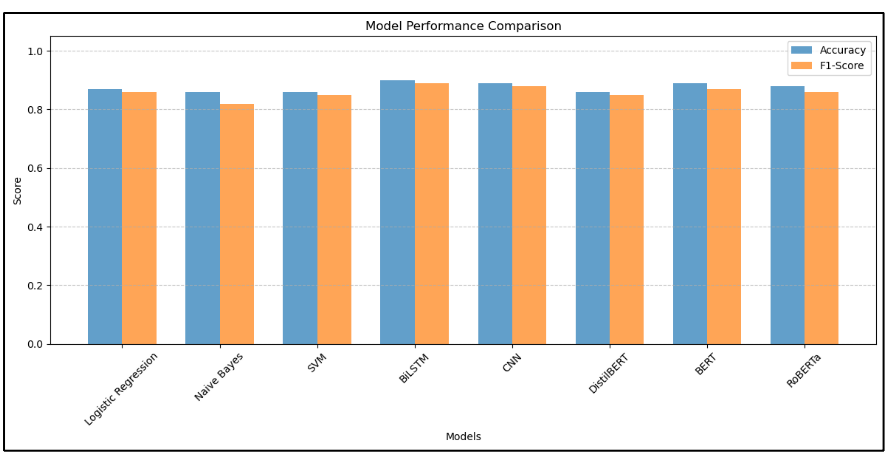
Evaluation Metrics

Key evaluation metrics were calculated using classification_report and confusion_matrix from scikit-learn. These included:

- **Accuracy**: Overall correctness of predictions
- **Precision & Recall**: For each class (positive, neutral, negative)
- **F1-score**: Harmonic mean of precision and recall
- **Confusion Matrix**: Visualization of true vs. predicted classifications to highlight misclassification

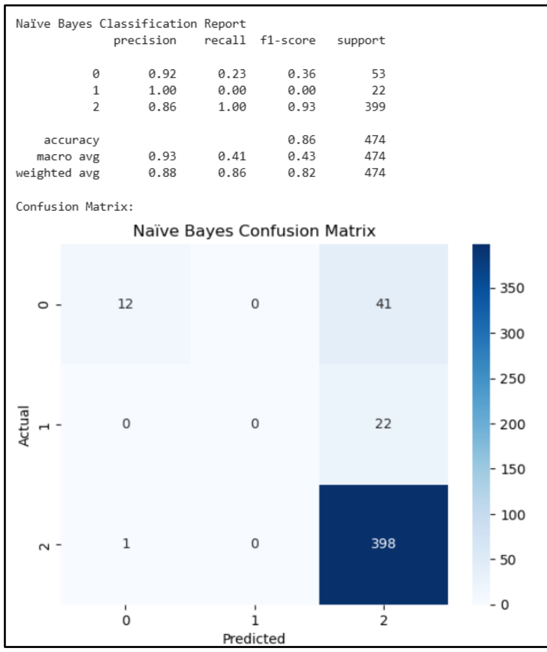
Result:

	Model	Accuracy	F1
0	Logistic Regression	0.87	0.86
1	Naive Bayes	0.86	0.82
2	SVM	0.86	0.85
3	BiLSTM	0.90	0.89
4	CNN	0.89	0.88
5	DistilBERT	0.86	0.85
6	BERT	0.89	0.87
7	RoBERTa	0.88	0.86



Visualization

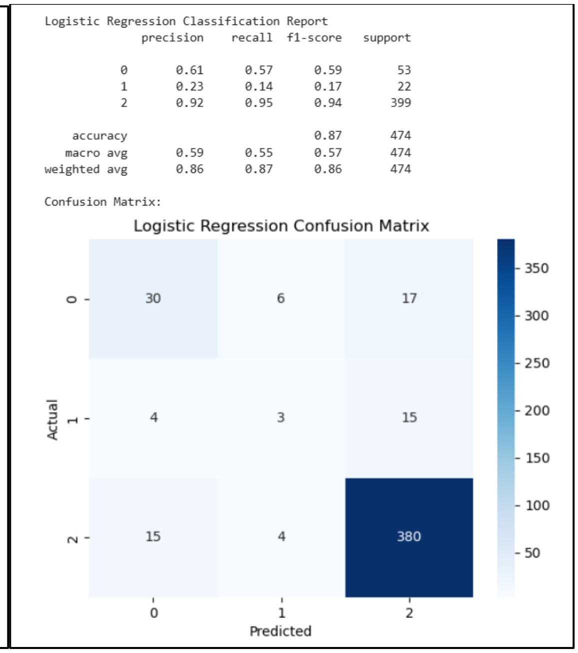
Confusion Matrix



Naïve Bayes Classification Report

	precision	recall	f1-score	support
0	0.92	0.23	0.36	53
1	1.00	0.00	0.00	22
2	0.86	1.00	0.93	399
accuracy			0.86	474
macro avg	0.93	0.41	0.43	474
weighted avg	0.88	0.86	0.82	474

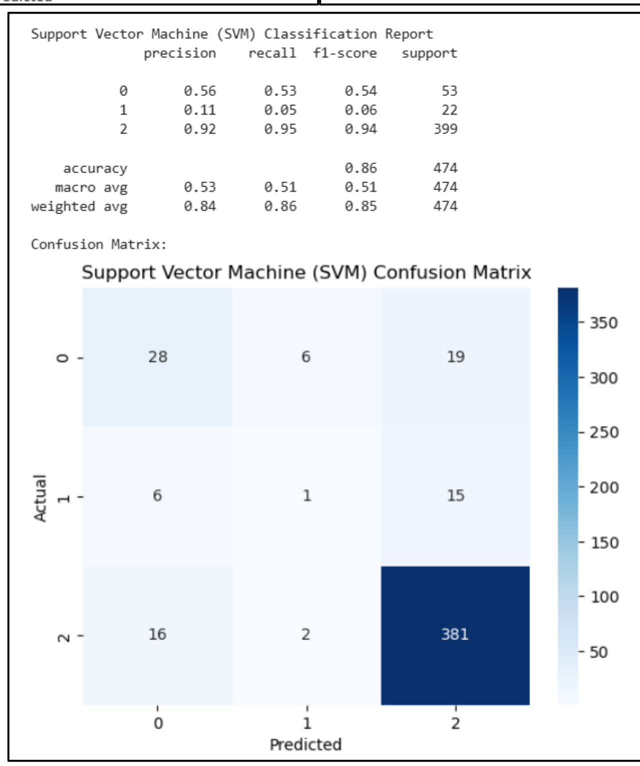
Confusion Matrix:



Logistic Regression Classification Report

	precision	recall	f1-score	support
0	0.61	0.57	0.59	53
1	0.23	0.14	0.17	22
2	0.92	0.95	0.94	399
accuracy			0.87	474
macro avg	0.59	0.55	0.57	474
weighted avg	0.86	0.87	0.86	474

Confusion Matrix:

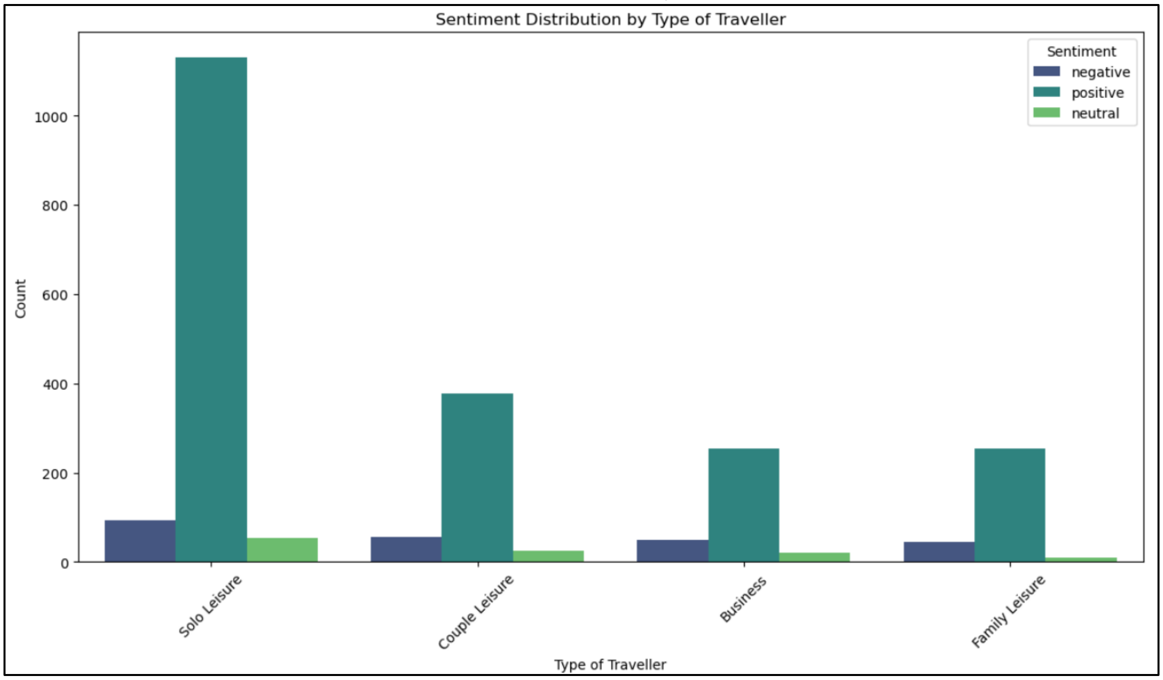


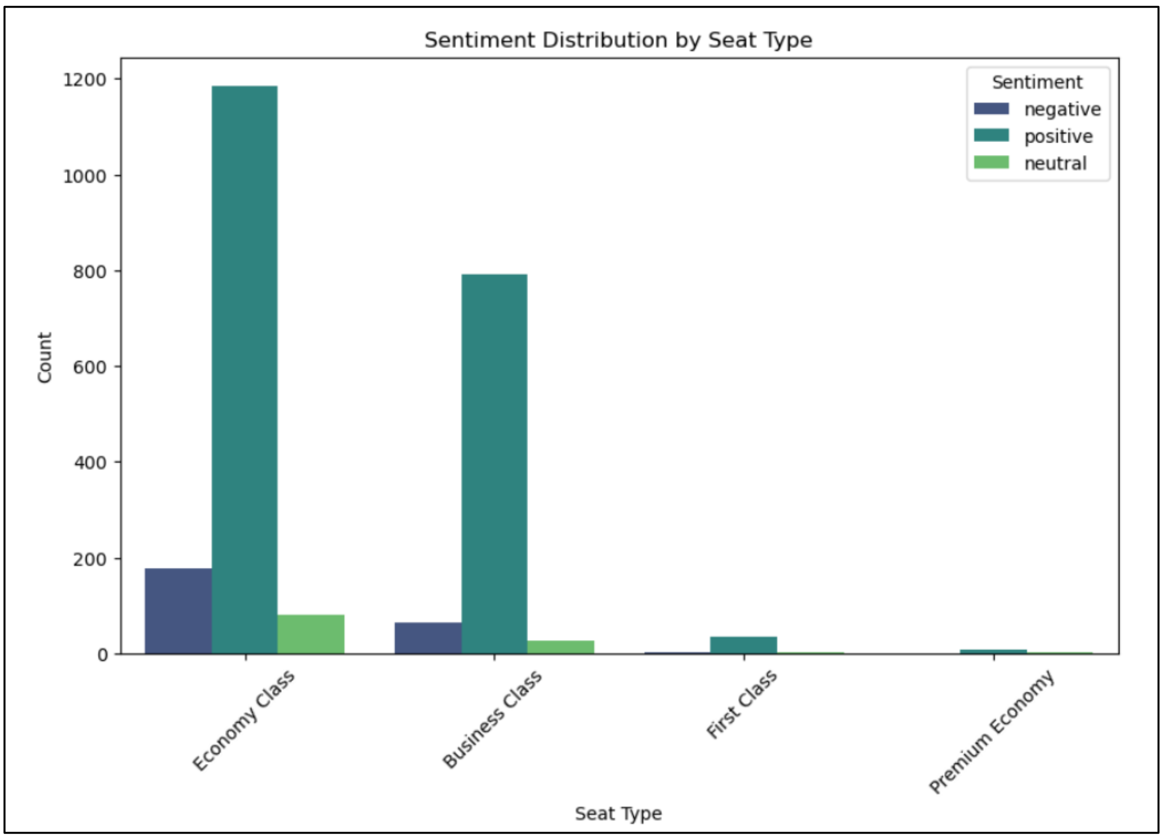
Support Vector Machine (SVM) Classification Report

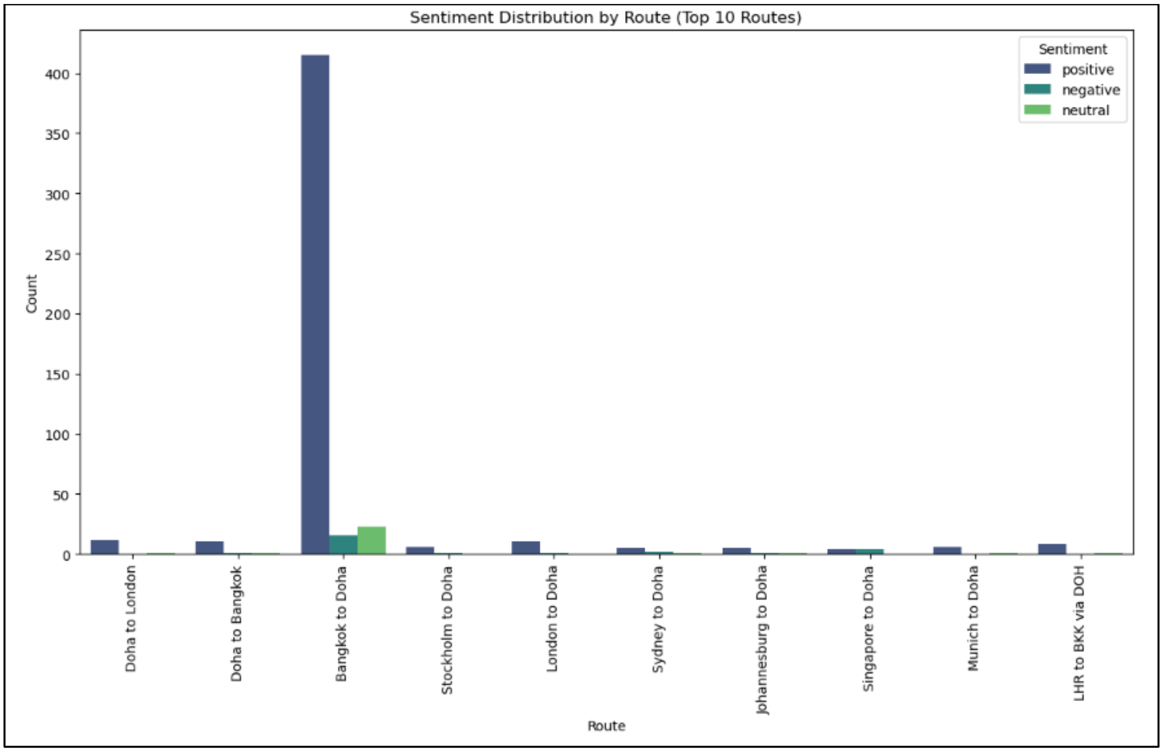
	precision	recall	f1-score	support
0	0.56	0.53	0.54	53
1	0.11	0.05	0.06	22
2	0.92	0.95	0.94	399
accuracy			0.86	474
macro avg	0.53	0.51	0.51	474
weighted avg	0.84	0.86	0.85	474

Confusion Matrix:

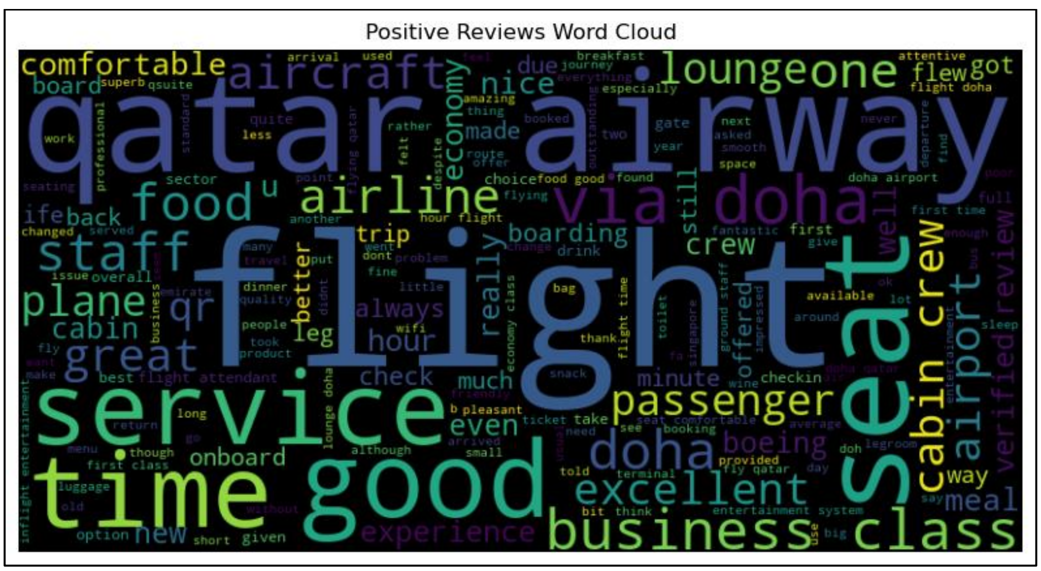
Sentiment Analysis

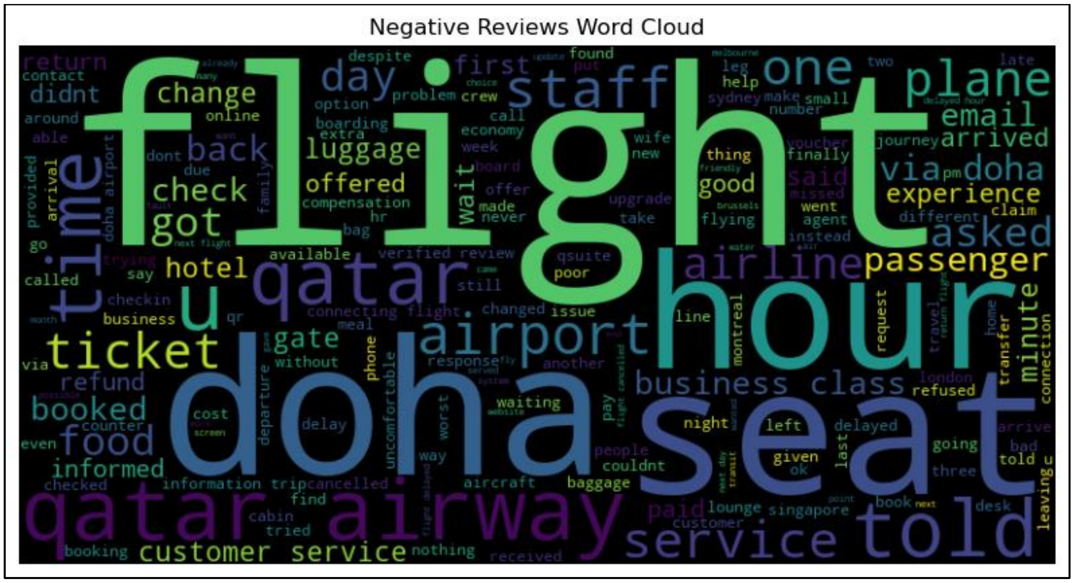




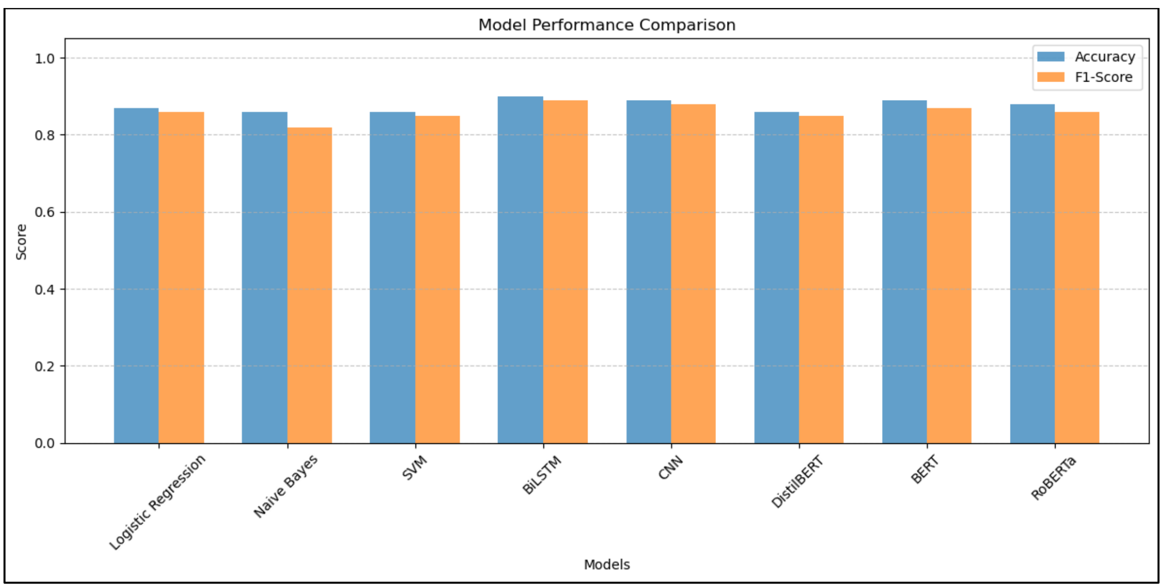

Sentiment Distribution by Route (Top 10 Routes)
Sentiment
positive
negative
neutral
Count
0
50
100
150
200
250
300
350
400
Doha to London
Doha to Bangkok
Bangkok to Doha
Stockholm to Doha
London to Doha
Sydney to Doha
Johannesburg to Doha
Singapore to Doha
Munich to Doha
LHR to BKK via DOH
Route

Word Cloud


Positive Reviews Word Cloud
qatar airway
flight
seat
service
time
good
business class
passenger
excellent
food
airline
staff
via doha
comfortable
aircraft
loungeone
plane
cabin crew
airport
great
verified review



Model Performance Comparison



Chapter 8: Discussion

This project explored various approaches to sentiment analysis using a real-world airline review dataset. The experimentation spanned classical machine learning algorithms, deep learning architectures and state-of-the-art transformer models. Through this multi-model comparison, several key insights were drawn.

Traditional models such as Naïve Bayes, Logistic Regression and SVM offered quick baselines and performed reasonably well when combined with TF-IDF features and basic semantic attributes like polarity and subjectivity. However, these models struggled to fully capture contextual nuances, idiomatic expressions and mixed sentiments in customer reviews.

Deep learning models like CNN and BiLSTM showed marked improvements. CNN effectively captured local patterns and n-grams, offering fast and stable performance. BiLSTM, on the other hand, excelled in handling long-term dependencies and context flow, especially in longer or more complex reviews.

The most significant gains came from fine-tuning transformer-based models—BERT, DistilBERT and RoBERTa. These models, particularly RoBERTa, demonstrated superior understanding of language context, emotion and sentence structure. RoBERTa's dynamic masking and extended pretraining allowed it to outperform all other models in both accuracy and F1-score. DistilBERT provided a good balance between speed and performance, making it a suitable choice for lightweight applications.

Another important observation was the value of hybrid feature engineering. Combining TF-IDF with sentiment scores (polarity, subjectivity) and word count gave traditional models additional structure and slightly improved their competitiveness.

Overall, the experiment highlights the importance of choosing the right model architecture based on available computational resources, data complexity and performance needs. While transformer models are computationally intensive, their contextual intelligence justifies their use in production-grade sentiment analysis systems.

Chapter 9: Conclusion & Future Work

Conclusion

This project successfully developed an end-to-end sentiment analysis system for airline reviews, comparing the performance of traditional machine learning models, deep learning architectures and transformer-based language models. The study demonstrated that while classical models like Naïve Bayes and Logistic Regression can provide fast baselines, their performance is limited when handling context-dependent sentiment. Deep learning models such as CNN and BiLSTM showed improved accuracy by capturing syntactic patterns and sequential information. However, transformer models—particularly RoBERTa—outperformed all others by effectively modeling the contextual and semantic richness of textual reviews.

The integration of both statistical (TF-IDF) and semantic (polarity, subjectivity) features contributed to a robust feature set that enhanced model interpretability and classification performance. Visualization tools such as word clouds and confusion matrices further supported insights into the data and model behavior.

Future Work

Several promising directions could enhance the project in the future. One potential area for improvement is Aspect-Based Sentiment Analysis (ABSA), which could capture sentiment at a more granular level, which enables the separation of opinions on specific aspects, such as flight comfort versus food quality. This would provide deeper insights into customer experiences and improve the accuracy of sentiment analysis. Another direction to explore is the extension of the system to support multilingual reviews. By incorporating multiple languages, the model would be able to handle a broader range of global datasets, making it more versatile and accessible to users from different linguistic backgrounds. Additionally, deploying the model as a web app could significantly enhance its usability. Integrating the system into a web-based dashboard using tools such as Streamlit or Dash would allow for real-time input and interactive visualizations, offering a more engaging and user-friendly experience. Finally, incorporating real-time sentiment tracking is another avenue to consider. By integrating live review feeds from platforms like Twitter or TripAdvisor, the system could transform into a real-time monitoring tool that provides instant insights into customer feedback and enabling businesses to respond more proactively to customer concerns. These directions hold significant potential to further develop the system and broaden its scope.

References

Agarwal, I., & Gowda, K. R. (2020). The effect of airline service quality on customer satisfaction and loyalty in India. *Materials Today: Proceedings, 37*(2). https://doi.org/10.1016/j.matpr.2020.06.557

Hasib, K. M., Habib, M. A., Towhid, N. A., & Showrov, M. I. H. (2021). A novel deep learning-based sentiment analysis of Twitter data for US airline service. *2021 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD)*, 450–455. https://doi.org/10.1109/ICICT4SD50815.2021.9396879

Idris, S. L., & Mohamad, M. (2023). A study on sentiment analysis on airline quality services: A conceptual paper. *Information Management and Business Review, 15*(4), 564–576.

Kang, H. W., Chye, K. K., Ong, Z. Y., & Tan, C. W. (2021). Sentiment analysis on Malaysian Airlines with BERT. *The Journal of The Institution of Engineers, Malaysia, 82*(3). https://doi.org/10.54552/v82i3.98

Patel, A., Oza, P., & Agrawal, S. (2023). Sentiment analysis of customer feedback and reviews for airline services using language representation model. *Procedia Computer Science, 218*, 2459–2467. https://doi.org/10.1016/j.procs.2023.01.385