

# Title: Comprehensive Report on Ticket Management System Analytics

## 1. Introduction:

The Ticket Management System Analytics project aimed to enhance the efficiency and effectiveness of handling support tickets through the implementation of machine learning models. This report provides a detailed overview of our approach, challenges faced, results, potential biases, and suggestions for further improvements.

## 2. Approach and Methodology:

### 2.1 Data Collection:

We started with historical ticket data, including ticket descriptions, response times, and sentiment labels. The dataset was cleaned, and missing values were addressed.

### 2.2 Data Preprocessing:

- Text Cleaning: Implemented a text preprocessing function to clean and tokenize ticket descriptions for better model performance.

- Handling Missing Values: Filled or dropped missing values based on the nature of the data.

### 2.3.1 Ticket Classification:

- Model: Naive Bayes classifier (MultinomialNB)

- Vectorization: TF-IDF Vectorization

- Evaluation Metric: Accuracy

### 2.3.2 Response Time Prediction:

- Model: Random Forest Regression

- Evaluation Metric: Mean Absolute Error (MAE)

### 2.3.3 Sentiment Analysis:

- Model: Assumed to be a classification model

- Evaluation Metrics: Accuracy, Precision, Recall, F1-Score

### 3. Challenges Faced and Solutions:

#### 3.1 UnicodeDecodeError:

Faced issues with character encoding during data loading.

Solution: Specified encoding while loading data using `pd.read_csv(encoding='utf-8')`.

#### 3.2 Categorical Encoding:

Difficulty in encoding categorical variables for machine learning models.

Solution: Utilized Label Encoding and Ordinal Encoding for different models.

### 4. Results and Performance Evaluation:

#### 4.1 Ticket Classification:

Achieved an accuracy of around 73% in categorizing tickets. (accuracy can be increased after performing few other ml techniques like hyperparameter tuning)

Naive Bayes - Accuracy: 0.7205, F1 Score: 0.7024599031802728  
Random Forest - Accuracy: 0.7415, F1 Score: 0.7339697459395756  
SVM - Accuracy: 0.7555, F1 Score: 0.7412601977015878  
K-Nearest Neighbors - Accuracy: 0.7285, F1 Score: 0.7186377085958703  
Logistic Regression - Accuracy: 0.738, F1 Score: 0.7250012651106942  
Decision Tree - Accuracy: 0.7155, F1 Score: 0.7113385587531885

#### 4.2 Response Time Prediction:

Model:Random Forest Regressor Results:

MAE: 146.71508422646926

MSE: 44074.052738980274

R<sup>2</sup>: -0.27466202984292387.

### 5. Potential Biases and Mitigation:

#### 5.1 Data Imbalance:

Biases may be introduced due to imbalances in the dataset.

Mitigation: Employed techniques like oversampling or undersampling to balance class distribution.

#### 5.2 Textual Bias:

Language bias in ticket descriptions and responses.

Mitigation: Regularly updated the text cleaning function to address emerging language patterns.

## 6. Suggestions for Further Improvements and Future Work:

### 6.1 Model Ensemble:

We can Explore the potential of model ensemble techniques for better overall performance.

### 6.2 Advanced Sentiment Analysis:

Implement more sophisticated sentiment analysis models, considering the nuances of customer feedback.

### 6.3 Continuous Monitoring:

Establish a system for continuous model monitoring to adapt to evolving patterns in ticket data.

### 6.4 Feedback Loop:

Incorporate a feedback loop where model predictions are utilized to improve future predictions and model performance.

## 7. Conclusion:

The Ticket Management System Analytics project demonstrates the feasibility and effectiveness of applying machine learning techniques to streamline support ticket processes. Continuous refinement and adaptation will be crucial for sustained success.