



SMS SPAM DETECTION

GROUP – F

Guided by: Meysam Effati



GROUP MEMBER

Alwin Kannyakonil Scaria [c0894287]

Anisha Susan Mathew [c0907393]

Ashna Viji Alex [c0901082]

Jobin Philip [c0895950]

Mohamed Afthab [c0891945]



Introduction

- The problem of spam detection is significant in email and SMS communication.
- This project aims to build a predictive model to classify SMS messages as spam or not using natural language processing (NLP) and machine learning.



Dataset Overview

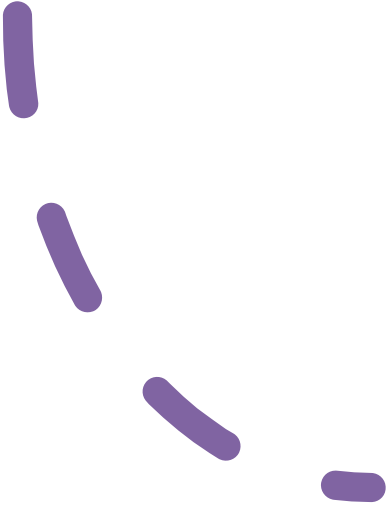
The dataset is sourced from the SMS Spam Collection. It contains SMS messages labeled as either 'ham' (legitimate) or 'spam'.





Data Preprocessing

Data Cleaning

- Dropped unnecessary columns
 - Renamed column:
 - As class and text: where spam = 1 and ham = 0.
- 

Text Preprocessing

Text Cleaning

- Stop words removal
- Stemming
 - PorterStemmer()



TRANSFORMATION

- **Vectorization**

Text to Numerical Feature Vectors.

- **Data Splitting:**

Splitting Data into X and y:

- X represents the input features (i.e., the text of SMS messages).
- y represents the target labels (i.e., spam or ham).



FEATURE ENGINEERING

Balancing Our Dataset with SMOTE:

- Our SMS dataset has many more "ham" messages than "spam."
- This imbalance can cause the model to be biased toward "ham."

Why Balance the Data?

- To improve spam detection, we need a balanced dataset.



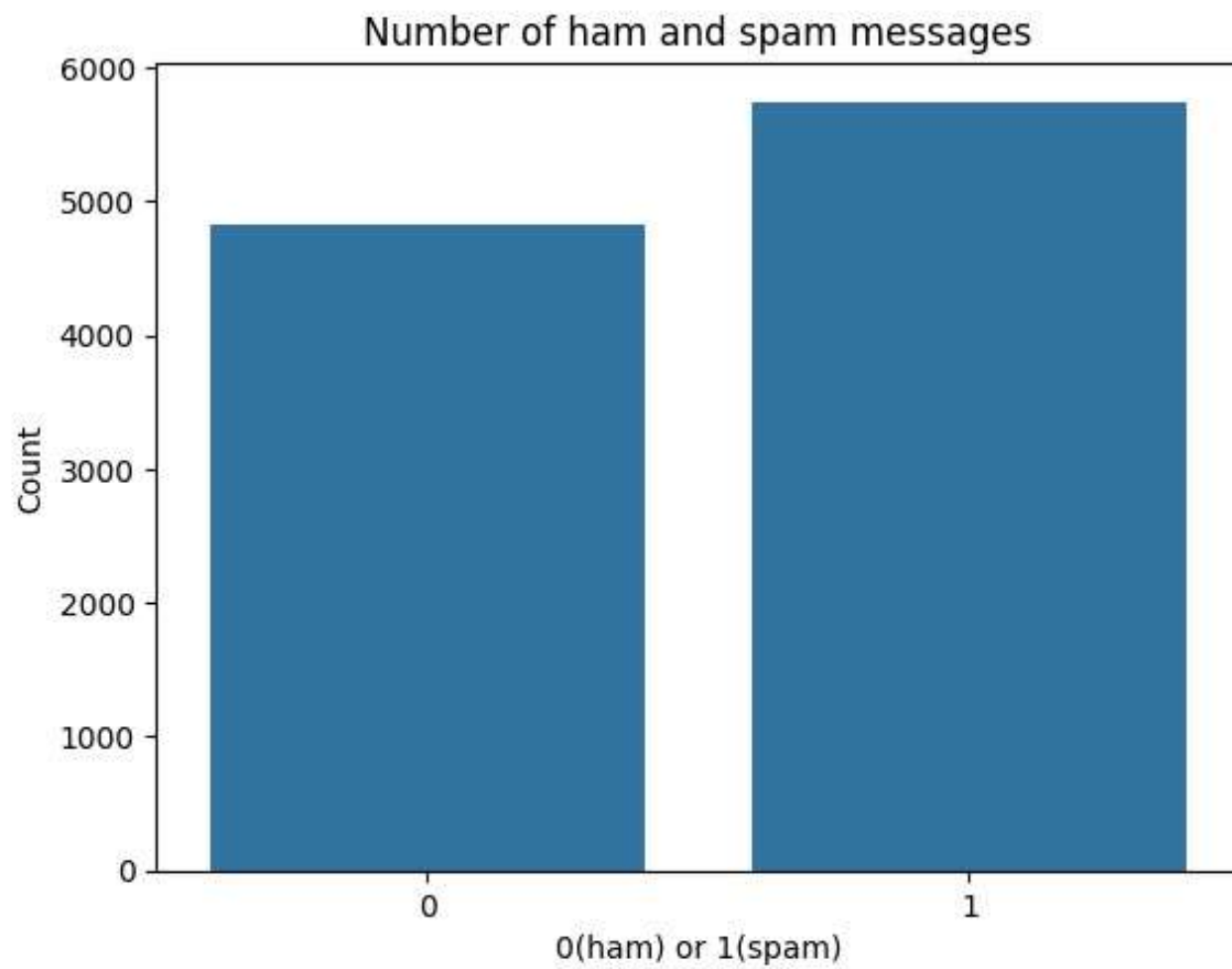
SMOTE Technique:

Purpose: SMOTE (Synthetic Minority Over-sampling Technique)

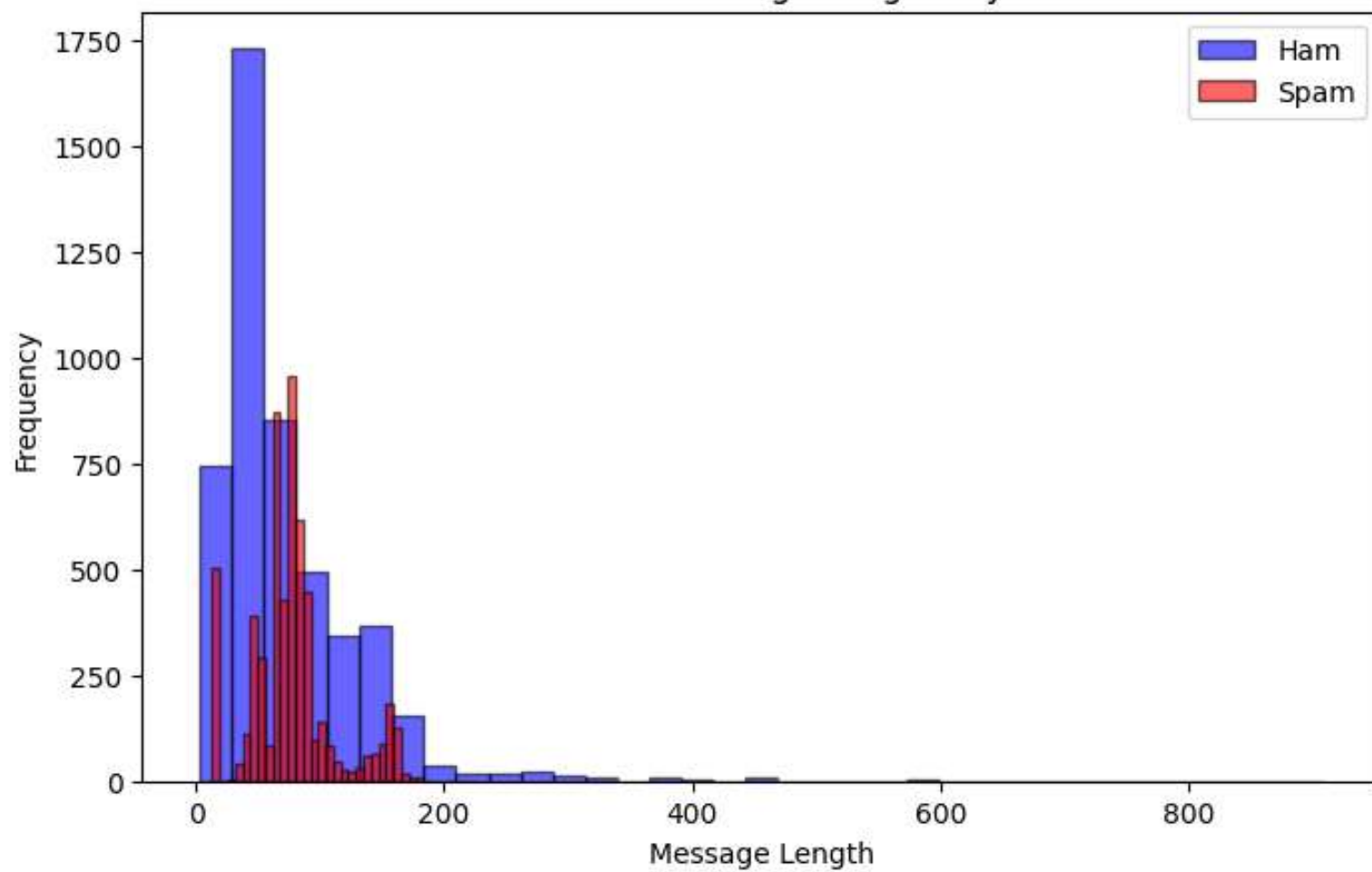


It generates synthetic "spam" messages by combining existing ones and balancing the dataset.





Distribution of Message Lengths by Class



MODEL TRAINING


Classification models:

1. Logistic Regression
2. Naive Bayes Classifier
3. Random Forest Classifier
4. Gradient Boosting
5. SVM (Support Vector Machine)
6. Stochastic Gradient Descent.





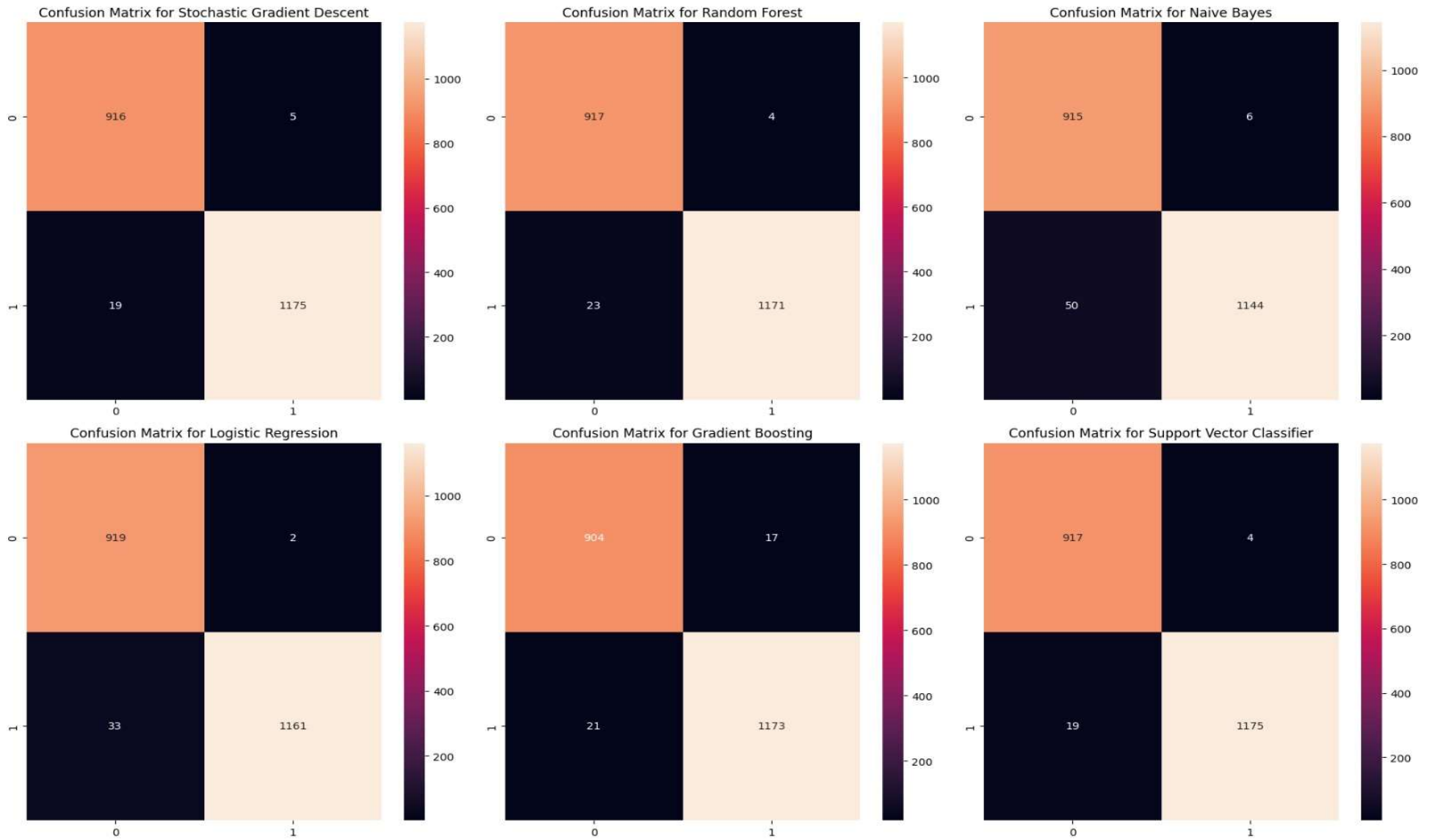
Model Evaluation Metrics

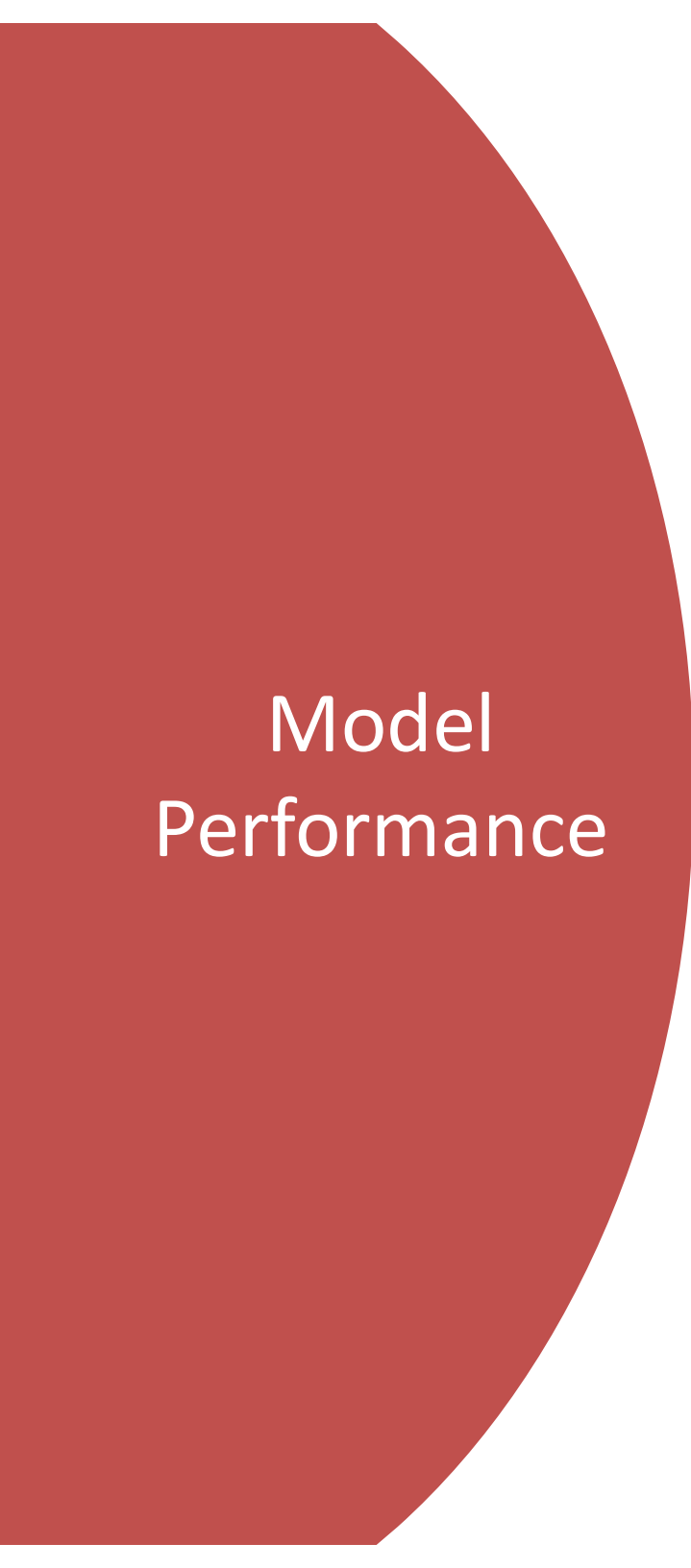


Models were evaluated using accuracy, precision, recall, and F1-score.

Confusion matrices were used to understand the performance in detail.

Comparative Analysis





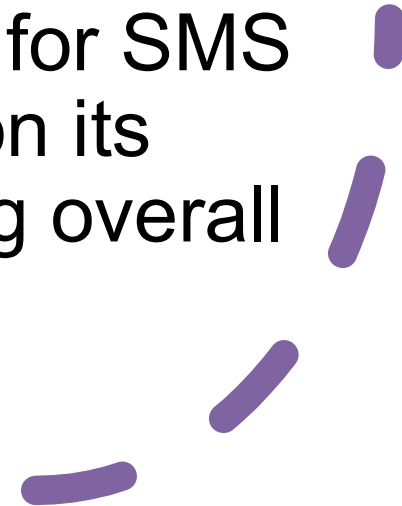
Model Performance

Best Model: Support Vector Classifier

Key Metrics:

- **Accuracy:** 98.91%
- **Precision:** 98.82%
- **Recall: 98.99%** (Best Performance)

The Support Vector Classifier is the most reliable model for SMS spam detection based on its highest recall and strong overall performance.



Limitations

- The code struggles to accurately detect ambiguous messages that could be classified as either spam or ham.
- Example: Messages like "Click here for iPhone" might be challenging for the model to categorize correctly due to the subtlety of the language and context.

Future Work

Advanced models like BERT, deeper NLP techniques, and further optimizing the real-time processing capabilities of the system can be implemented.



Conclusion

A machine-learning SMS spam detection model was developed and deployed using the flask application.

The Support Vector Classifier emerged as the most effective model, achieving the highest recall.

Addressed key challenges like class imbalance using SMOTE and improved model performance through feature engineering.

Future plans include implementing advanced models like BERT.

References

- Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique.
- Pedregosa, F., Varoquaux, G., Gramfort, A., et al. (2011). Scikit-learn: Machine Learning in Python.

Thank
You

