



# SMS Spam Detection Model

This presentation outlines the development of an SMS spam detection model. The goal is to create a system that accurately classifies messages. This leads to improved user experience by filtering unwanted spam.



# Project Goals and Objectives



## High Accuracy

Distinguish spam from legitimate SMS messages.



## Minimize False Positives

Avoid classifying legitimate messages as spam.



## Maximize True Positives

Correctly identify and flag spam messages.



## Real-World Application

Deploy the model for spam filtering purposes.

# Dataset Overview

## Source

Combination of publicly available SMS datasets.

## Size

5,574 SMS messages

## Label Distribution

- Non-spam (Ham): 4,827 messages
- Spam: 747 messages

## Example Messages

- Spam: URGENT! Win £1000 prize now!
- Non-spam: Okay, I will be there soon.

# Approach

1

## **Data Preprocessing**

Clean text, Over Sampling, extract features (TF-IDF).

2

## **Training/Testing Split**

80% training, 20% testing data.

3

## **Logistic Regression**

Machine Learning model

4

## **Frameworks**

Python, scikit-learn, and NLTK.



# Performance Metrics

**96.23%**

Overall Model Accuracy

**0.96**

Non-Spam Precision

**1.00**

Spam Precision

**0.72**

Spam Recall



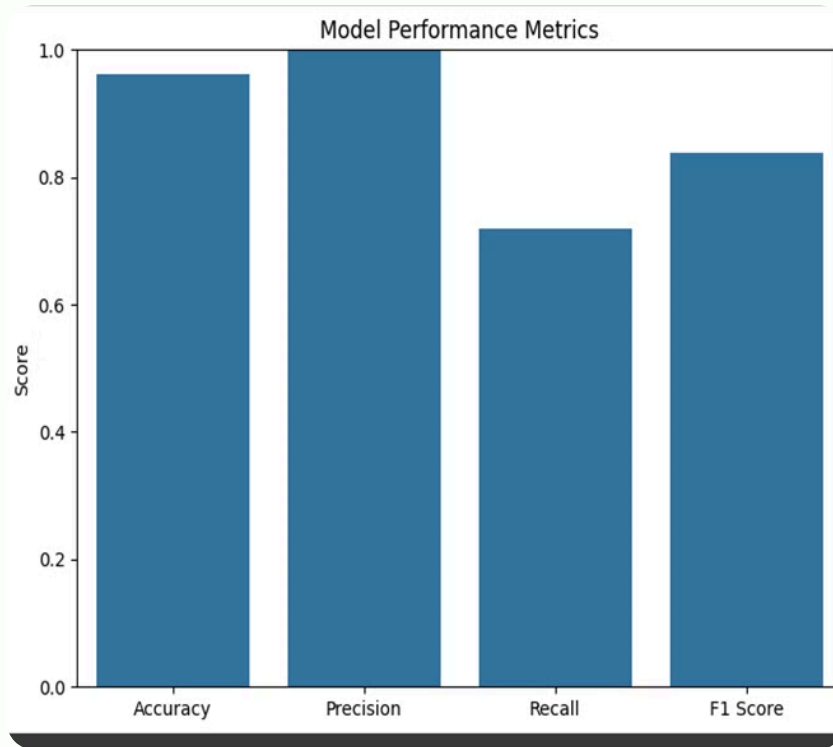


# Classification Report

Metric	Non-spam (o)	Spam (1)
Precision	0.96	1.00
Recall	1.00	0.72
F1-Score	0.98	0.84

This detailed table shows the precision, recall and F1-score for each class.

# Performance Metrics



- **Bar Chart Visualization** - Plots **Accuracy, Precision, Recall, and F1 Score** as bars.
- **Score Range** - Y-axis values range from **0 to 1** for better readability.
- **Title & Labels** - Titled *"Model Performance Metrics"*, with labeled X and Y axes.
- **Color & Aesthetics** - Uses `seaborn` styling for a clean and professional look.

**THANK YOU**