মেট্রোপলিটন ইউনিভার্সিটি
Metropolitan University

Project Report

SMS Spam Detection System

Submitted By

Ishrak Saleh Chowdhury

ID: 232-134-034

Course Name

Artificial Intelligence Lab

Submitted To

Al Akram Chowdhury

Senior Lecturer

Date of Submission: 17/12/25

Project Overview

Title: SMS Spam Detector – An NLP-Based Text Classification System

The SMS Spam Detection System is a Natural Language Processing (NLP) and Machine Learning–based application designed to automatically classify SMS messages as **Spam** or **Not Spam (Ham)**. The system analyzes the textual content of incoming messages and predicts whether they are malicious, promotional, or unwanted spam messages.

This project integrates:

- Text preprocessing using NLP techniques
- Feature extraction using TF–IDF vectorization
- A supervised machine learning classifier (Logistic Regression)
- A user-friendly Streamlit-based web interface

The system provides both a final classification and a probability score, enabling users to understand how confident the model is about its prediction.

Purpose of the System

The primary objective of the SMS Spam Detector is to automate the identification of spam messages. Manual filtering of SMS messages is inefficient and unreliable, especially with the increasing volume of unsolicited and fraudulent communications. This system aims to:

- Reduce exposure to spam and scam messages
- Automate message filtering using AI
- Provide fast and consistent classification results
- Assist users in identifying potentially harmful messages

Intended Users

- General mobile phone users
- Students and researchers learning NLP
- Developers building message-filtering systems
- Cybersecurity and fraud prevention enthusiasts

System Architecture Overview

The SMS Spam Detection System operates in three main stages:

1. **Data Preprocessing and Feature Engineering**
2. **Model Training and Evaluation**
3. **Prediction and Streamlit Web Application**

Each stage transforms raw SMS text into meaningful insights through a structured machine learning pipeline.

Stage 1: Dataset Loading and Text Preprocessing

1. Dataset Loading

The dataset used for training the SMS Spam Detector is a labeled SMS dataset stored in a CSV file (spam.csv). Each record contains:

- **Message text** – the raw SMS content
- **Label** – indicating whether the message is *Spam* or *Ham*

This dataset serves as the foundation for training the machine learning model.

2. Text Cleaning and Normalization

Before training, each SMS message undergoes a structured preprocessing pipeline to remove noise and irrelevant information. The steps include:

- Conversion of text to lowercase
- Tokenization using NLTK
- Removal of punctuation and non-alphanumeric characters
- Removal of English stopwords (e.g., *the*, *is*, *and*)
- Stemming using the Porter Stemmer

This process ensures that the model focuses on meaningful words rather than unnecessary symbols or common filler terms.

Output: Cleaned and normalized SMS text

3. Text Vectorization (TF–IDF)

After preprocessing, the text data is converted into numerical format using **TF–IDF (Term Frequency–Inverse Document Frequency)** vectorization. This technique assigns importance to words based on how frequently they appear in a message relative to the entire dataset.

The TF–IDF vectors allow the machine learning model to mathematically analyze text patterns and distinguish spam-related terms from normal messages.

Stage 2: Model Training and Evaluation

1. Train–Test Split

The dataset is divided into two parts:

- **Training set:** Used to train the model
- **Testing set:** Used to evaluate performance

This separation ensures that the model is tested on unseen data, providing a realistic evaluation of its accuracy.

2. Model Selection

The classifier used in this project is **Logistic Regression**, a supervised learning algorithm well-suited for binary classification tasks such as spam detection.

Reasons for choosing Logistic Regression include:

- Efficient performance on text-based features
- Ability to output probability scores
- Simplicity and interpretability

3. Model Evaluation

The trained model is evaluated using standard classification metrics, including:

- Accuracy
- Precision
- Recall
- F1-score

These metrics measure how effectively the model identifies spam messages while minimizing false positives and false negatives.

4. Model Persistence

After training, the following components are saved using Python's pickle module:

- Trained Logistic Regression model (model.pkl)
- TF–IDF vectorizer (vectorizer.pkl)

These saved files allow the system to make predictions without retraining the model each time the application runs.

Stage 3: Prediction and Streamlit Application

1. Model Loading

The Streamlit application loads the pre-trained model and vectorizer from disk. This ensures fast startup time and efficient predictions.

2. User Input Processing

Users can enter any SMS message into the application interface. The input message undergoes the same preprocessing steps as the training data to maintain consistency.

3. Spam Prediction

The processed text is:

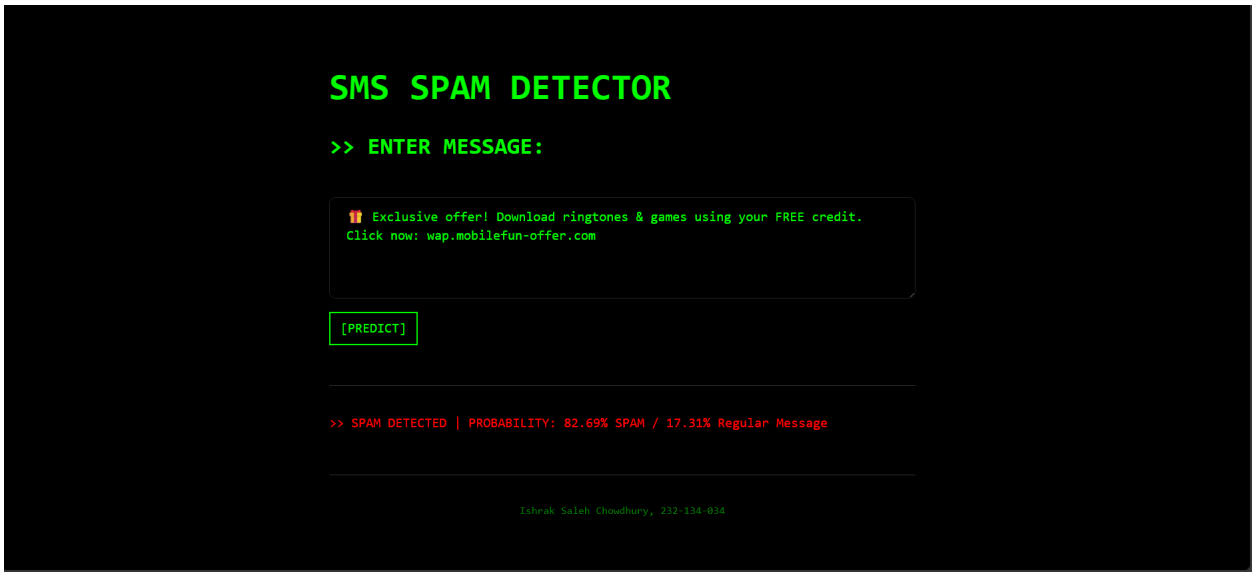
1. Transformed into TF–IDF features
2. Passed to the Logistic Regression model
3. Classified as **Spam** or **Not Spam**

The system also calculates the probability of the message being spam.

Output Presentation

The prediction results are displayed using a modern Streamlit interface featuring:

- Clear spam or not-spam status
- Probability percentages for spam and regular messages
- Cyber-themed dark UI for enhanced readability



Why This System Is Necessary

Manual SMS filtering suffers from several limitations:

- Time-consuming message review
- High risk of human error
- Increasing sophistication of spam content

The SMS Spam Detection System addresses these issues by:

- Automating message classification
- Providing fast, reliable predictions
- Reducing exposure to scams and phishing messages

Summary

The SMS Spam Detection System is a complete end-to-end NLP and Machine Learning application that:

- Preprocesses raw SMS text using NLP techniques
- Converts text into TF–IDF feature vectors
- Uses Logistic Regression for binary classification
- Predicts spam messages with probability scores
- Provides a user-friendly Streamlit-based interface

This project demonstrates the practical application of NLP, supervised learning, and model deployment, making it suitable for academic evaluation and real-world spam detection use cases.

Technology Stack

- Python
- NLTK

- Scikit-learn
- Pandas
- Streamlit
- Pickle

Model: Logistic Regression + TF–IDF