

# Project Summary

---

## Project: Spam Message Detection using LinearSVC

### Description

This is a machine learning project that classifies SMS messages as Spam or Not Spam using Natural Language Processing (NLP) and Scikit-learn models. It includes a web interface built with Flask, styled using HTML and CSS, and can be deployed locally.

### Tools & Technologies:

- Python
- Scikit-learn
- NLTK
- Flask (for web interface)
- Google Colab (for training)
- Pandas, NumPy
- HTML, CSS (for frontend)

### Files Included:

- app.py (Flask backend)
- index.html (Frontend UI)
- model\_new.pkl (Trained model)
- vectorizer\_new.pkl (TF-IDF vectorizer)
- spam.csv (Dataset)
- train\_model.py (Training script)
- test\_nltk.py (NLTK check script)

### Dataset Summary:

- Original dataset from Kaggle - SMS Spam Collection
- Preprocessing includes:
  - Lowercasing
  - Tokenization using nltk.word\_tokenize
  - Stopword removal
  - Stemming using PorterStemmer

<https://www.kaggle.com/datasets/uciml/sms-spam-collection-dataset>

## Model: LinearSVC (Support Vector Machine)

- Accuracy: 99.83%
- Precision: 99.89%
- Recall: 99.78%
- F1 Score: 99.83%

```
→  Evaluation Metrics for LinearSVC  
Accuracy: 0.9983397897066962  
Precision: 0.9988888888888889  
Recall: 0.9977802441731409  
F1 Score: 0.9983342587451416
```

### Functionality:

- User enters an SMS message on the web page
- Message is transformed using the same logic as training
- Flask app loads the model & vectorizer, predicts whether message is Spam ✗ or Not Spam ✓

### Output Example:

SPAM

The image shows two side-by-side screenshots of a web application titled "Spam Message Detector". Both screenshots feature a large input field at the top with placeholder text "Enter your message here...". Below the input field is a blue "Check" button. In the first screenshot, the input field contains the text "Congratulations! You've won a \$500 Amazon gift card. Claim it here". A small green checkmark icon is displayed next to the word "Check". In the second screenshot, the input field is empty, and the word "Spam" is followed by a red "✗" symbol.

NOT SPAM

The image shows two side-by-side screenshots of the same web application for "NOT SPAM" messages. Both screenshots have an empty input field with the placeholder "Enter your message here...". Below the input field is a blue "Check" button. In the first screenshot, the input field contains the text "Hi I am Suman". A small green checkmark icon is displayed next to the word "Check". In the second screenshot, the input field is empty, and the text "Not Spam" is followed by a green checkmark "✓" symbol.

## **Folder Structure(VS Code):**

```
spam-detector/
    app.py
    train_model.py
    test_nltk.py
    model_new.pkl
    vectorizer_new.pkl
    spam.csv
    templates/
        index.html
    static/
        style.css
```