

# Sms Spam Detection Using NLP

## ● Overview

SMS Spam Detection is a machine learning model that takes an SMS as input and predicts whether the message is a spam or not spam message. The model is built using Python and deployed on the web using Streamlit.

## ● Technology Used

1. Python
2. Scikit-learn
3. Pandas
4. NumPy
5. Streamlit

## ● Features

1. Data collection
2. Data cleaning and preprocessing
3. Exploratory Data Analysis
4. Model building and selection
5. Web deployment using Streamlit

## ● Data Collection

The SMS Spam Collection dataset was collected from Kaggle, which contains over 5,500 SMS messages labeled as either spam or not spam.

You can access the dataset from

[here](<https://www.kaggle.com/datasets/uciml/sms-spam-collection-dataset>)

## ● Data Cleaning and Processing

The data was cleaned by handling null and duplicate values, and the "type" column was label-encoded. The data was then preprocessed by converting the text into tokens, removing special characters, stop words and punctuation, and stemming the data. The data was also converted to lowercase before preprocessing.

## ● Exploratory Data Analysis

Exploratory Data Analysis was performed to gain insights into the dataset. The count of characters, words, and sentences was calculated for each message. The correlation between variables was also calculated, and visualizations were created using pyplots, bar charts, pie charts, 5 number summaries, and heatmaps. Word clouds were also created for spam and non-spam messages, and the most frequent words in spam texts were visualized.

## ● Model Building and Selection

Multiple classifier models were tried, including NaiveBayes, random forest, KNN, decision tree, logistic regression, ExtraTreesClassifier, and SVC. The best classifier was chosen based on precision, with a precision of 100% achieved.

## ● Web Deployment

The model was deployed on the web using Streamlit. The user interface has a simple input box where the user can input a message, and the model will predict whether it is spam or not spam.

## ● Demo

To try out the SMS Spam Detection model, visit this link  
[Streamlit](#)

## ● Usage

To use the SMS Spam Detection model on your own machine, follow these steps:

- ❖ Clone this repository.
- ❖ Install the required Python packages using

```
pip install -r requirements.txt.
```

- ❖ Run the model using

```
streamlit run app.py.
```

## ● Contributions

Contributions to this project are welcome. If you find any issues or have any suggestions for improvement, please open an issue or a pull request on this repository.

Created By Aryan Wankhade

Contact :

[aryanwankhade0@gmail.com](mailto:aryanwankhade0@gmail.com)

Project Mentor Names : Jay Rathod

Contact : [JayRathod341997 \(Jay Rathod\)](#)