A Machine Learning Approach to Detect Spam Messages

RASHMI KUMARI
CSE(Cyber Security)
22152146049

What is spam?
Why is spam detection important?
How spam affects users?

- Unwanted or irrelevant messages sent in bulk over the internet. Includes junk emails, fake ads, phishing messages, and harmful links

- Why is spam detection important?
- Protects users from **fraud and phishing attacks**.
- Saves **time and storage space**.
- Prevents **malware or virus infections**.
- Keeps your **inbox clean and organized**.

- How spam affects users
- Leads to **financial losses** through scams.
- Steals **personal or sensitive data**.
- Can **infect computers** with viruses or ransomware.

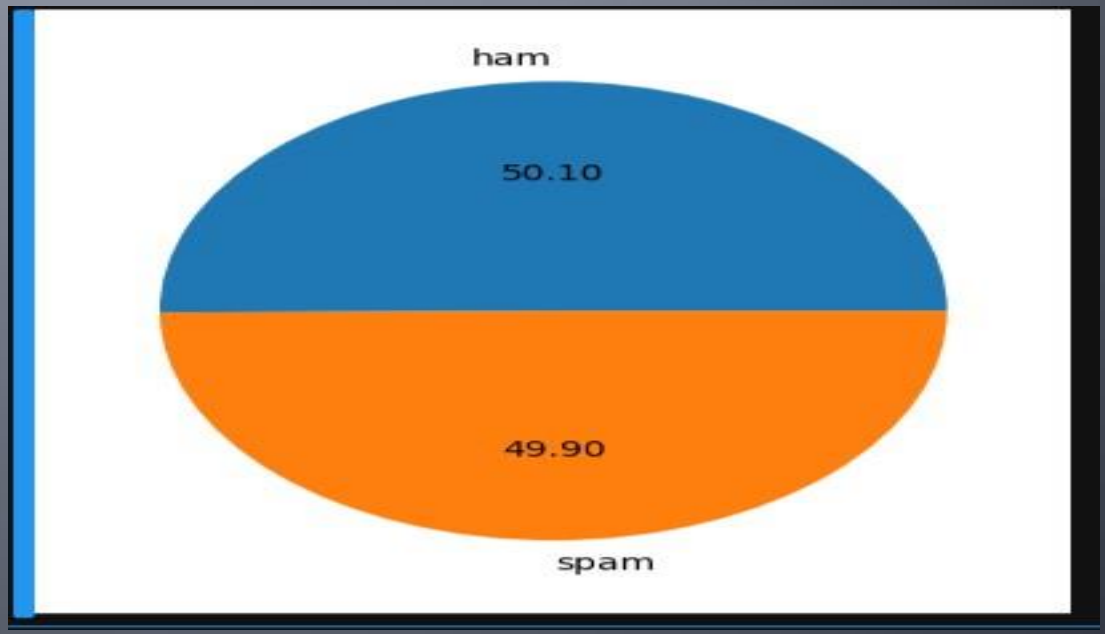
Objective of the Project:

develop an efficient "Spam Detection System "that can automatically identify and classify emails or messages as spam or not ham using Machine Learning techniques

- To classify messages as Spam or Not Spam using Machine Learning.
- To automate the process of spam detection with high accuracy.
- To deploy the trained model using a simple Gradio web app.

Name of dataset : "small_spam_dataset.csv"
Total messages: 500
Two columns: Label (Spam/Ham) and Message (Text).
Around 49% of messages are spam.

pie chart showing % of Spam vs Ham:



Data Preprocessing

- Lowercasing the text.
- Removing punctuation and special characters.
- Tokenization and stopword removal.
- Applying stemming using PorterStemmer.
- Converting text to numerical form using TF-IDF.

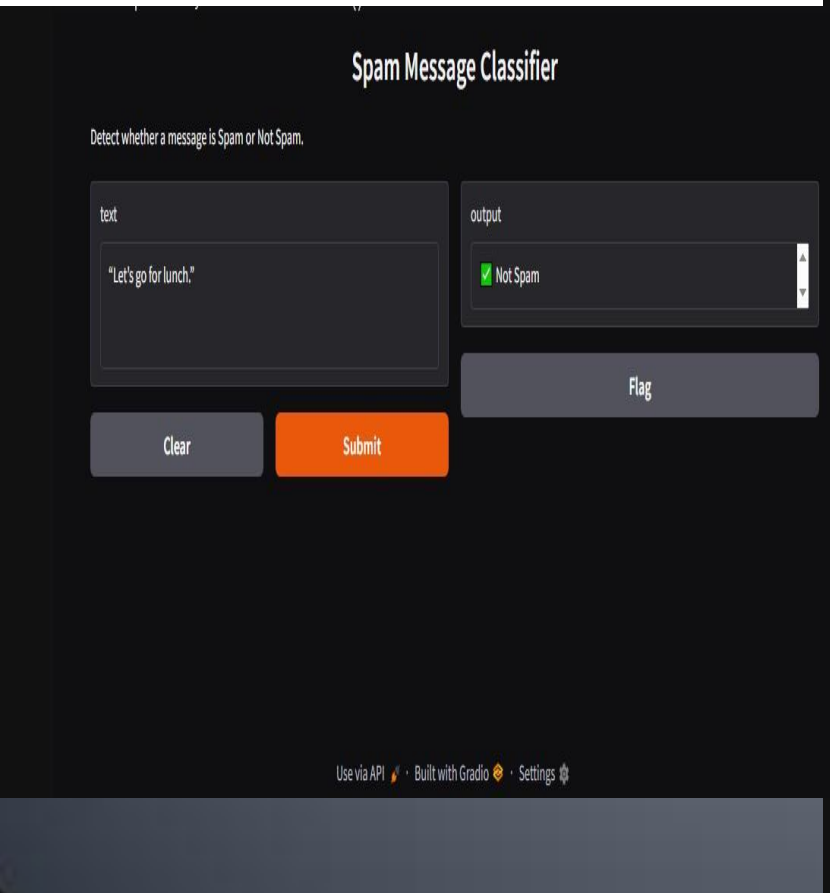
Model Building

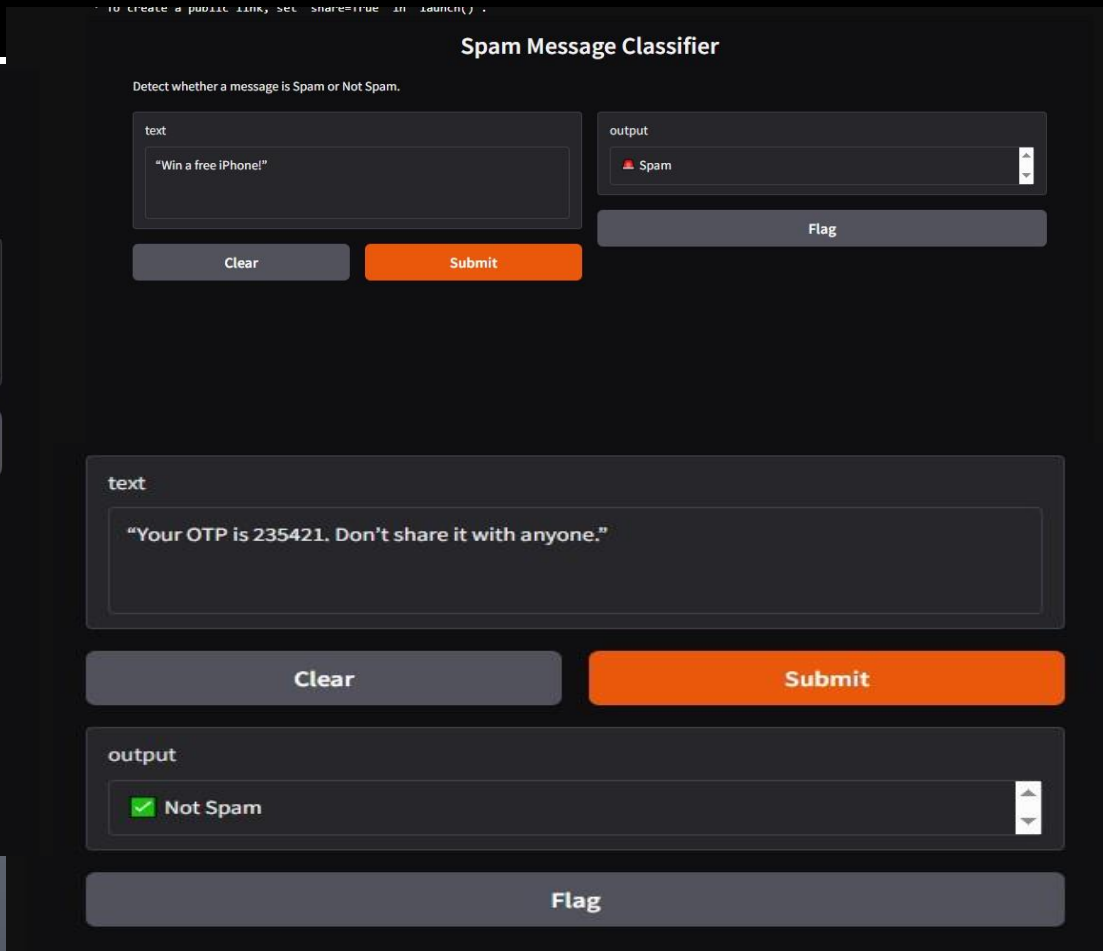
Algorithm used: Naive Bayes.

- TF-IDF Vectorizer used to convert text into feature vectors.
- Dataset split into training and testing sets (50:49).
- SMOTE used to balance spam and ham samples.
- Model trained to achieve 92% accuracy.

Implementation:

by Gradio App :A Gradio app is created for real-time spam detection.The user inputs a message, and the model predicts Spam or Not Spam

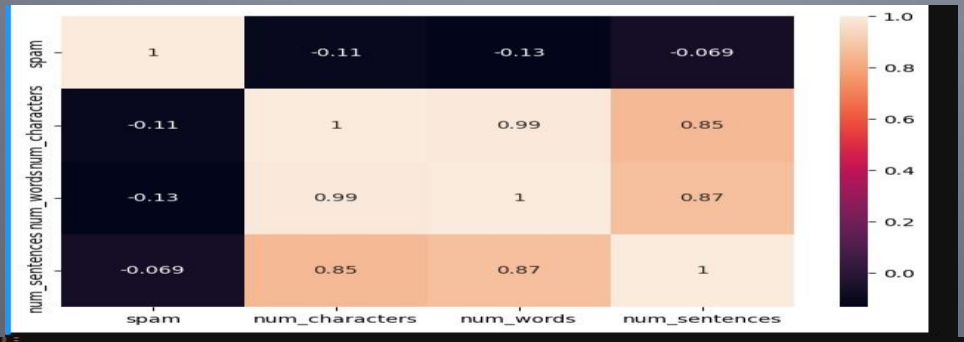


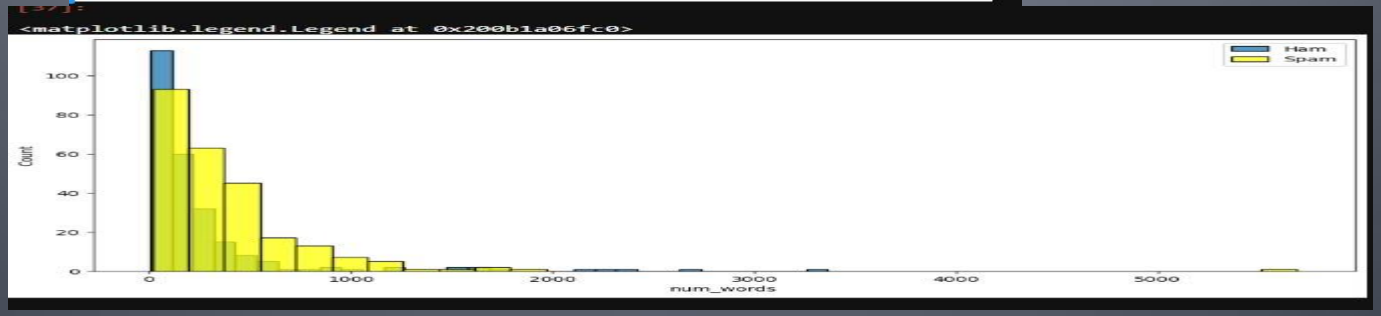


Results & Accuracy

-Accuracy achieved 92%

- Confusion matrix or classification report:





References

- Dataset source: kaggle
- Tools used : Python, Scikit-learn, NLTK, Gradio,etc
- Open source
-IDE: Jupyter Notebook.