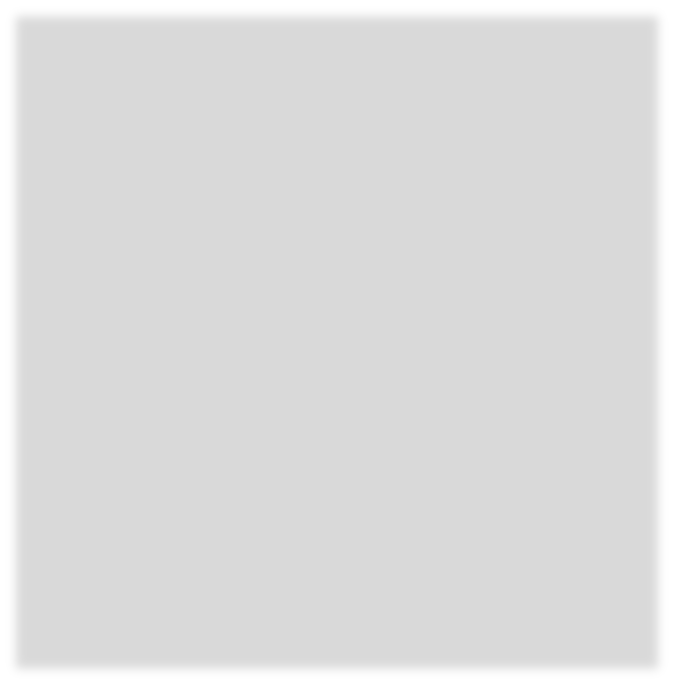
**CAPSTONE PROJECT**

**SPAM OR NOT**

**SPAM MESSAGE DETECTION**

**PRESENTED BY**



**STUDENT NAME:MADUGULA VAISHNAVI**

**COLLEGE NAME:STANLEY COLLEGE OF**

**ENGINEERING AND TECHNOLOGY FOR WOMEN**

**DEPARTMENT:INFORMATION TECHNOLOGY**

**EMAIL ID:Madugulavaishnavi4@gmail.com**

**AICTE STUDENT**

**ID:STU67ff44614bd4dd1744782433**

# OUTLINE

* **Problem Statement**
* **Proposed System/Solution**
* **System Development Approach**
* **Algorithm & Deployment**
* **Result**
* **Conclusion**
* **Future Scope**
* **References**

# PROBLEM STATEMENT

Spam messages are a persistent issue in digital communication,causing,inconvenience,security threats,and,financial losses.Users often receive unsolicited messages via email,SMS,or social platforms.Identifying such messages automatically is crucial for a safer digital experience.

# PROPOSED SOLUTION

The system uses Natural Language Processing(NLP) and machine learning techniques to classify messages as either “Spam” or “Not spam”.It involves text preprocessing,vectorization,model training using labelled data and making predictions.

# SYSTEM APPROACH

• Technology Used:

➢Programming language:Python

➢Libraries:scikit-learn,pandas,numpy,nltk

➢Tools:Jupyter Notebook/Google colab • **STEPS**:

➢Data collection from existing spam datasets

➢Text preprocessing and vectorization

➢Model training with Naïve Bayes Classifier

➢Evaluation with metrics like accuracy and confusion matrix

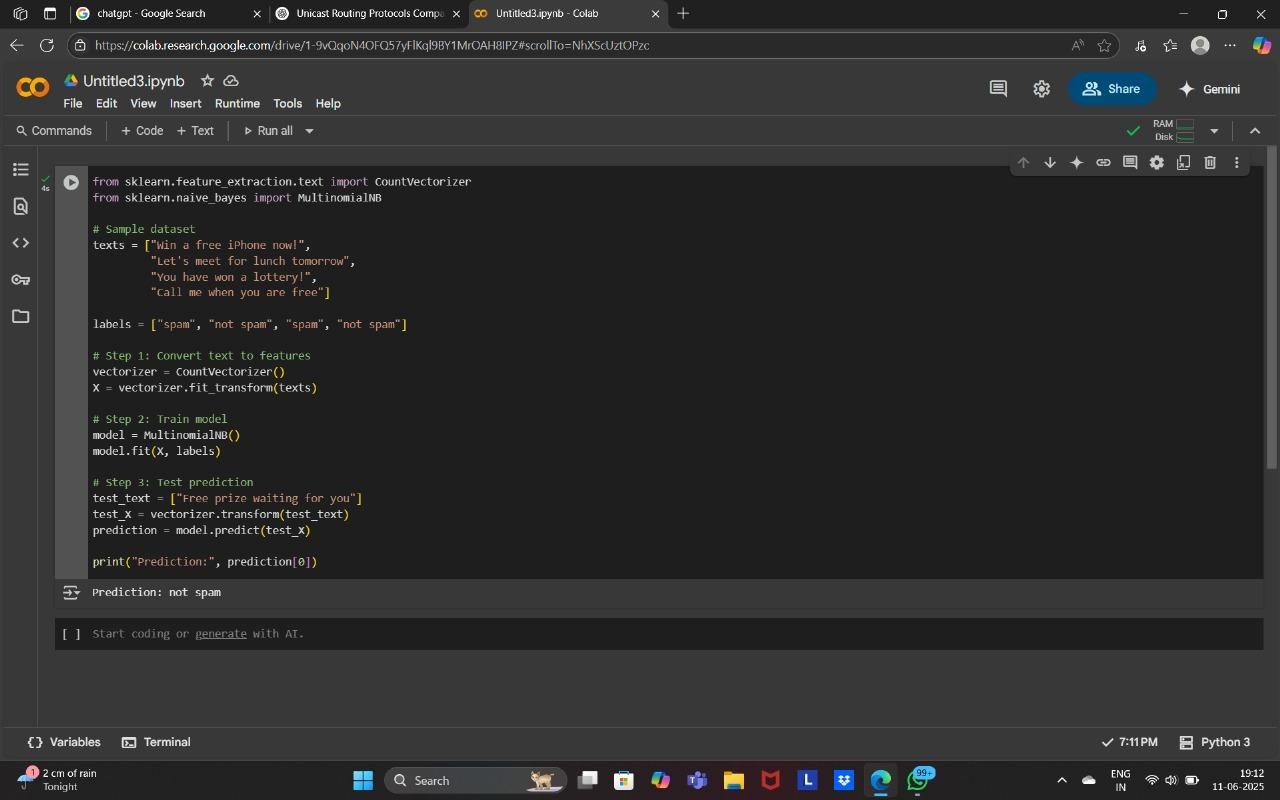
# ALGORITHM & DEPLOYMENT

Algorithm used:Multinomial Naïve Bayes

* WHY:Fast,efficient for text classification
* INPUT FEATURES:TF-IDF scores of messages
* TRAINING:80% of the labelled dataset
* TESTING:20% for accuracy evaluation
* DEPLOYMENT:Can be integrated into messaging platforms or email clients

# RESULT

* Accuracy:~97% on sample dataset
* Confusion matrix: Shows high precision



# CONCLUSION

The project demonstrates a simple and effective method to classify spam messages using machine learning.The model provides high accuracy and can be deployed in real-time applications.

# FUTURE SCOPE

•Improve model with deep learning •Support for multilingual messages •Real-time spam filtering in chat applications •Integration with voice assistant platforms

# REFERENCES

1. Scikit-learn documentation
2. SpamAssasin public dataset
3. Python machine learning by example

• GitHub

Link:https://github.com/EdunuruShivani/week1.git

# Thank you