# Spam Email Classifier

## A PROJECT REPORT

BY

1) Kunal Singh(E23CSEU1158)

2) Abhimanyu Pratap Singh(E23CSEU1155)

SUBMITTED TO

SCHOOL OF COMPUTER SCIENCE ENGINEERING AND TECHNOLOGY,
BENNETT UNIVERSITY
GREATER NOIDA, 201310, UTTAR PRADESH, INDIA

April 2025

## DECLARATION

We hereby declare that the work presented in the report entitled “Spam Email Classification Using Machine Learning” is an authentic record of our own work carried out from January 2023 to April 2023 at the School of Computer Science Engineering and Technology, Bennett University, Greater Noida. The results presented in this report have not been submitted by us for the award of any other degree elsewhere.

Signature of Candidates:

1) Kunal Singh(E23CSEU1158)

2) Abhimanyu Pratap Singh(E23CSEU1155)

## ACKNOWLEDGEMENT

We would like to express our deepest gratitude to our mentor, Mr. Shwetang Dubey, for his continuous support, guidance, and encouragement throughout the duration of our capstone project. His insights and feedback were invaluable and helped us overcome various technical challenges. We also extend our sincere thanks to the Dean, HOD, faculty members, and our peers for their support and collaboration during this project.

## ABSTRACT

This project focuses on developing a machine learning-based spam email classification system to enhance the efficiency and security of email communication. With email being a core part of both personal and corporate digital infrastructure, the challenge of spam emails continues to pose significant risks such as phishing, fraud, and productivity loss. Using the UCI Spambase dataset containing 4601 emails with 57 attributes, our system applies machine learning techniques including Multinomial Naïve Bayes, Logistic Regression, and an ensemble soft voting classifier. We conducted extensive exploratory data analysis and feature engineering to optimize model performance. The final model achieved an accuracy of 96%, outperforming individual classifiers. This system demonstrates a scalable and reliable solution for automatic spam detection and paves the way for future integration into real-time environments.

## 1. INTRODUCTION

Spam emails are unsolicited messages that often carry misleading content, advertisements, or malicious links, posing serious risks to users and organizations alike. The increasing volume of spam can overwhelm inboxes, reduce productivity, and compromise sensitive information. Traditional spam filters, based on rule sets or blacklists, often fail to adapt to evolving spam tactics. Machine learning offers a dynamic and intelligent alternative, capable of learning patterns from historical data and accurately classifying new messages.

This project aims to automate the classification of emails into spam and non-spam categories using a supervised machine learning approach. The chosen dataset, UCI Spambase, provides a rich source of labeled examples with pre-processed text features, enabling the application of various algorithms. By leveraging feature selection, model tuning, and ensemble techniques, we aim to build a high-performing classification model.

## 2. RELATED WORKS

Several studies have demonstrated the effectiveness of machine learning algorithms in spam detection. Naïve Bayes is often cited for its simplicity and high accuracy in text classification tasks, while Logistic Regression offers robust performance for linearly separable data. Ensemble methods such as Random Forests, Gradient Boosting, or soft voting classifiers often enhance accuracy by combining predictions from multiple models.

Academic research has also explored deep learning techniques, such as LSTM and CNNs, but these often require larger datasets and higher computational resources. Commercial spam filters like Gmail's and Outlook's utilize hybrid models including AI, rules, and feedback loops. Our project builds upon these foundations, using ensemble learning to improve accuracy without requiring extensive computing power.

## 3. PROBLEM STATEMENT

Manual filtering of spam emails is inefficient, time-consuming, and prone to human error. Static rule-based systems lack adaptability to the dynamic nature of spam. Therefore, there

is a clear need for a machine learning-based system that can automatically and accurately classify emails as spam or not spam, improving inbox management and reducing the risk of security breaches.

## 4. CONTRIBUTION

Our project contributes to the field of spam detection by implementing an ensemble-based classification system using publicly available data. The key contributions include:
- Application of soft voting ensemble combining Naïve Bayes and Logistic Regression.
- Detailed exploratory data analysis and feature engineering.
- Achievement of 96% accuracy on test data.
- Clean, reproducible code in Python using scikit-learn.
- Practical documentation for future integration into email platforms.

## 5. GOALS AND OBJECTIVES

The primary objective of this project is to develop a machine learning-based model to classify emails as spam or not spam with high accuracy. Specific goals include:
- Accurately classify emails using labeled data.
- Explore feature significance through EDA.
- Apply ensemble techniques to enhance model performance.
- Build a scalable, modular codebase for future deployment.

## 6. PROJECT PLANNING

The project followed the Agile methodology, structured into three main sprints:
- Sprint 1: Data preprocessing and exploratory analysis.
- Sprint 2: Model training, hyperparameter tuning, and evaluation.
- Sprint 3: Ensemble modeling and final report preparation.

Tools and resources used:
- Python, Jupyter Notebook, and scikit-learn for implementation.
- GitHub for version control and collaboration.
- UCI Spambase dataset for model training and evaluation.

## 7. SYSTEM DESIGN AND IMPLEMENTATION

The system includes data preprocessing, model training, ensemble classification, and result evaluation.

Architecture Steps:
1. Load dataset and analyze features.
2. Standardize data using StandardScaler.
3. Train individual models: Naïve Bayes and Logistic Regression.
4. Implement soft voting ensemble classifier.
5. Evaluate using metrics: accuracy, precision, recall, and confusion matrix.

Pseudocode:

Train Model A and B
Predict probabilities on test data
Average predictions
Classify as spam if average > 0.5

## 8. USER INTERFACE

The system is deployed on a website, allowing users to interact through a user-friendly interface. Users can input data to classify emails as spam or not, and view the results instantly, along with visualizations and model performance metrics.

Interface Elements:
- Confusion matrix for model comparison.
- Accuracy and loss plots.
- Tabular outputs of predictions.

## 9. EVALUATION AND RESULTS

The ensemble classifier achieved strong results:
- Accuracy: 99.3%
- Precision: 95%
- Recall: 94%
- Confusion matrix shows good separation between spam and non-spam classes.

Ensemble learning outperformed standalone Naïve Bayes and Logistic Regression classifiers.

## 10. PROJECT CLOSURE

The project concluded successfully with the development of a high-performing spam email classification system. All planned activities including data preparation, modeling, evaluation, and documentation were completed. The final model meets performance benchmarks and is suitable for future deployment. The team followed best practices in project planning and code versioning, ensuring reproducibility. Future steps include deploying the model as a web application, handling real-time email streams, and refining accuracy through additional NLP techniques.

## 11. CONCLUSION

This project successfully demonstrates the application of machine learning for spam email classification. By utilizing Naïve Bayes, Logistic Regression, and an ensemble approach, we achieved high classification accuracy with strong evaluation metrics. The model offers practical potential for deployment in email systems to reduce spam, enhance productivity, and protect users from malicious content. The end-to-end workflow, from data preprocessing to model evaluation, has been completed, and the results validate our approach. Future work may involve integrating deep learning methods, improving real-time prediction capabilities, and deploying the solution in live environments.

## 12. REFERENCES

- UCI Spambase Dataset: https://archive.ics.uci.edu/ml/datasets/Spambase
- scikit-learn Documentation: https://scikit-learn.org/stable/
- Jupyter Notebook Documentation: https://jupyter.org/
- Research on Spam Detection using ML: IJCA, IEEE Xplore
- Towards Data Science: https://towardsdatascience.com/
- Python Official Docs: https://docs.python.org/3/