

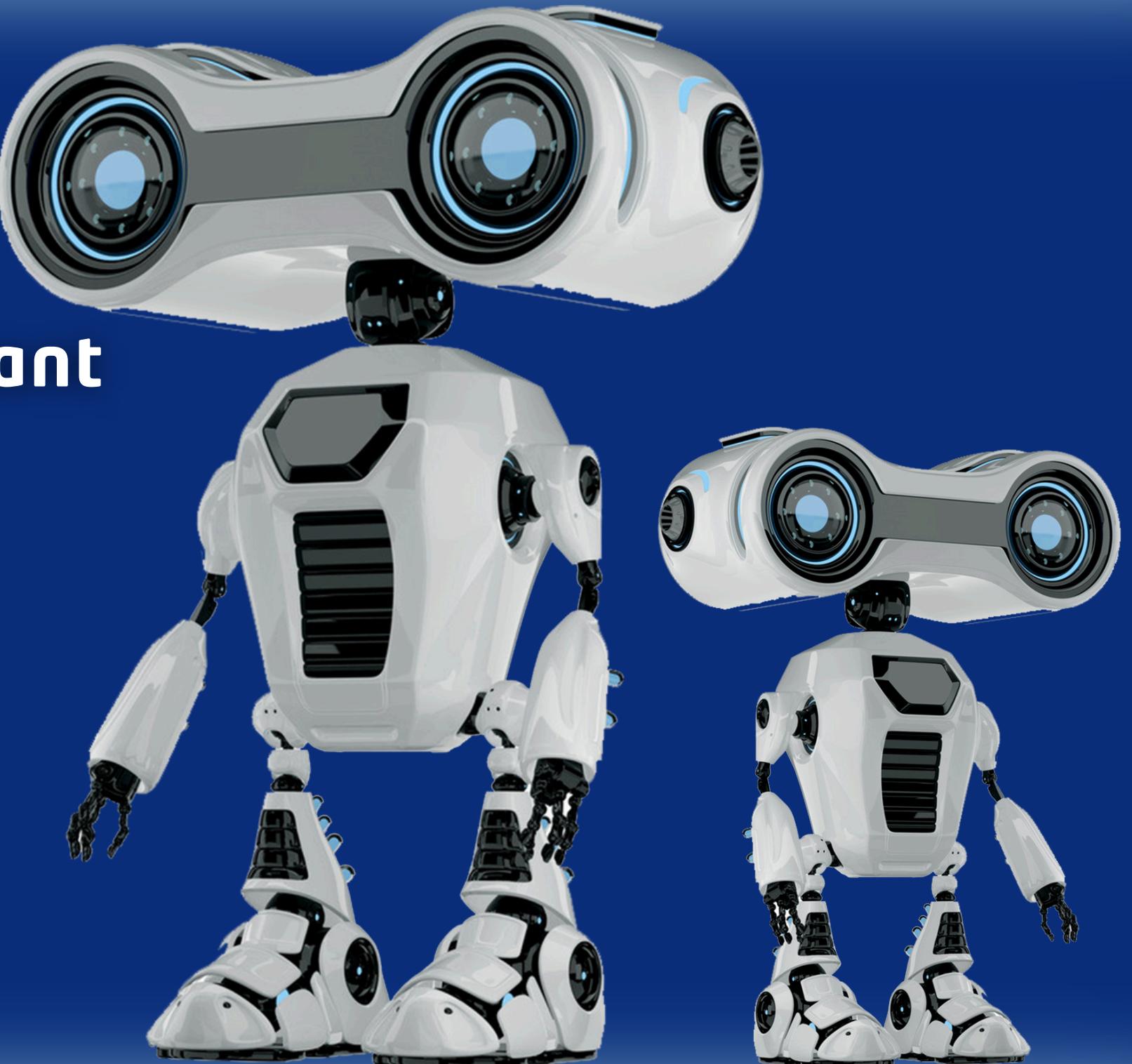
SPAM MAIL PREDICTION

Using machine learning

MEKHAZNI NIHAD
AIDAOUI YOUSRA CHAHINEZ

PLAN:

- ❖ Why “Spam Detection” is important
- ❖ What is Machine Learning
- ❖ Steps to create a ML system
- ❖ Algorithms used
- ❖ Model Evaluation



Why “Spam Detection” is Important



Over 50% of daily
email traffic is spam



Spam often contains
phishing links



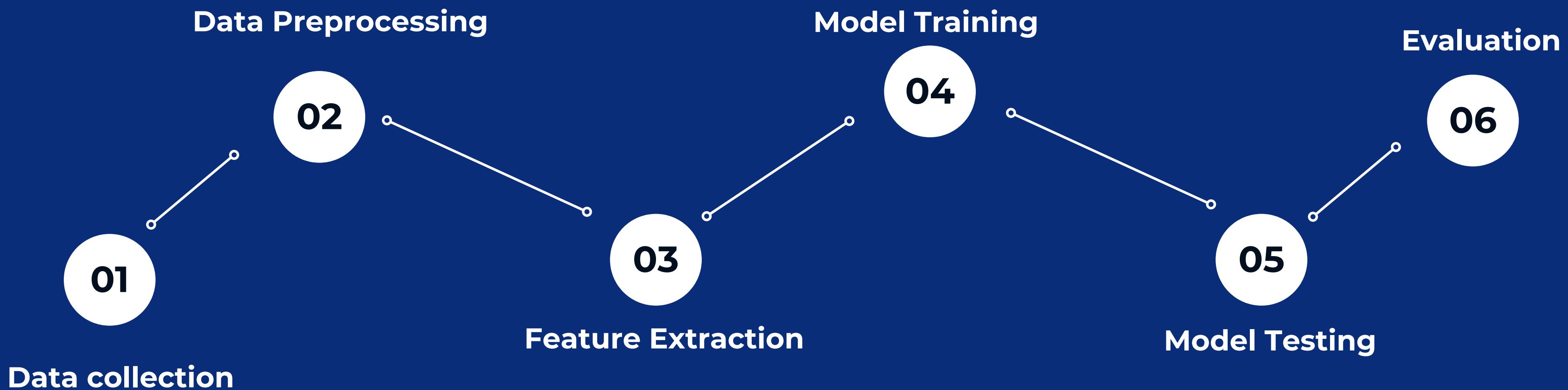
Protects users from
identity theft, fraud,
and malware

what is machine learning

- Machine Learning (ML) is a branch of Artificial Intelligence (AI)
- It allows computers to learn from data without being explicitly programmed
- ML algorithms detect patterns and make predictions or decisions
- In this project, ML is used to detect whether an email is spam or not



Steps :



Algorithms used :

Logistic Regression

Logistic Regression is a statistical method used for classification tasks, where the goal is to predict a binary outcome based on one or more input variables.

Despite its name, it is not a regression technique in the traditional sense, but rather a classification algorithm.

XGboost

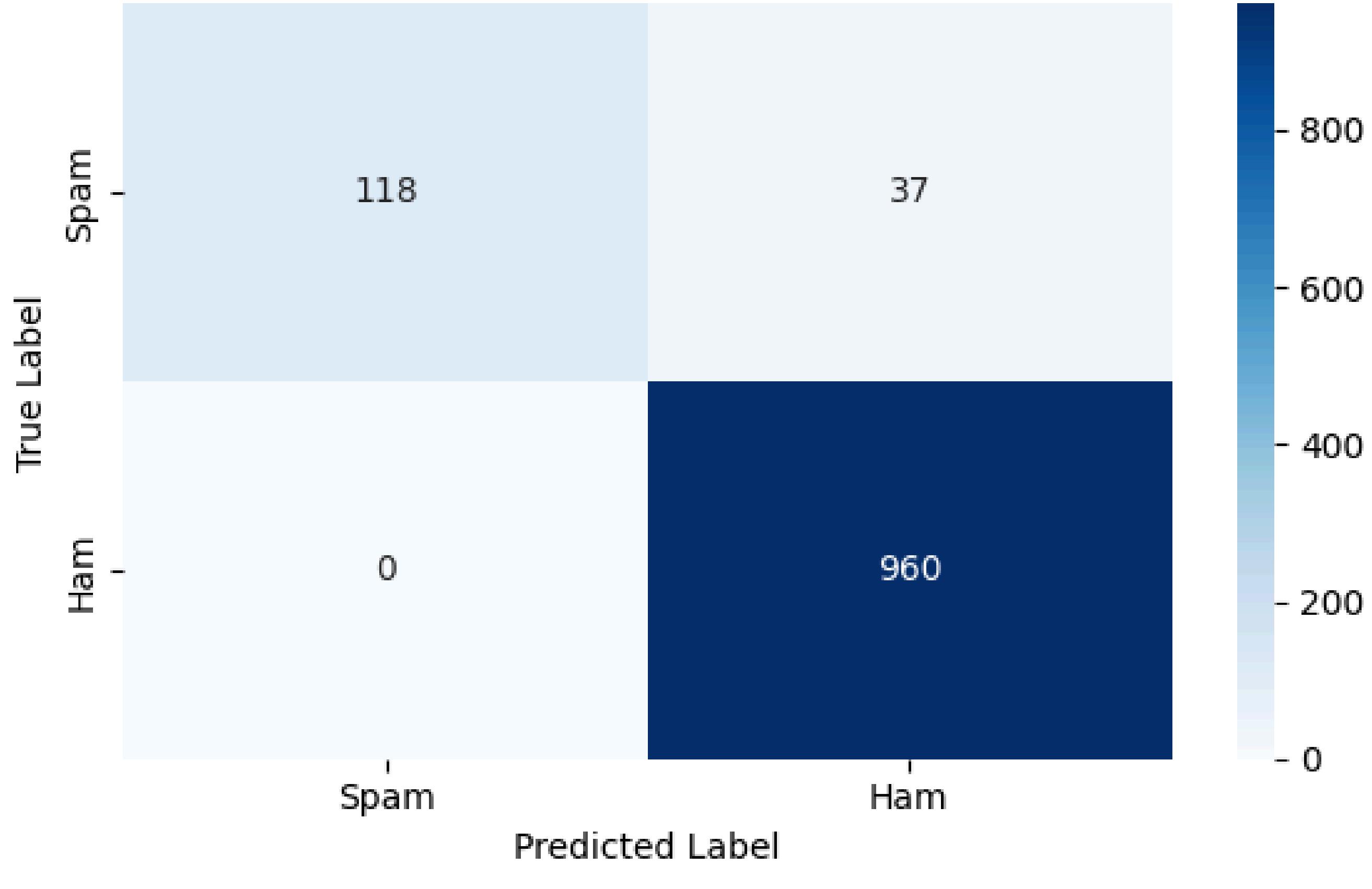
XGBoost is a powerful machine learning algorithm designed for both classification and regression tasks.

It is an implementation of gradient boosting, where multiple decision trees are trained sequentially – each tree correcting the errors of the previous one.

Model Evaluation

Dataset	Accuracy (%)
Training Data	96.77%
Test Data	96.68%

Confusion Matrix - Logistic Regression



Conclusion



- This project demonstrated how machine learning can effectively detect spam emails using real-world data.
- By applying “ Logistic Regression”, we achieved:
 - 96.77% accuracy on training data
 - 96.68% accuracy on test data
- The model is simple, fast, and provides high performance with minimal errors.
- This approach can be integrated into email systems to reduce spam, improve user experience, and increase security.

**THANK YOU FOR
YOUR ATTENTION**