UNIVERSITY OF ASIA PACIFIC

Department of Computer Science & Engineering

Course Title: Artificial Intelligence Lab

Course Code: CSE 404

Submitted By:

Name: Moumita Afreen

Reg no: 21201130

Sec: C

Group: C-2

SMS Spam Detection Using Naive Bayes

1. Introduction

In today's digital communication era, SMS is a common medium for personal and professional messages. However, it is often plagued by spam messages, which can be intrusive and potentially harmful. This project aims to build a machine learning model that can automatically detect whether an SMS is spam or not.

2. Dataset Description

The dataset used is the SMS Spam Collection Dataset, which contains 5,574 SMS messages labeled as either 'spam' or 'ham' (not spam). It is a publicly available dataset from the UCI Machine Learning Repository.

3. Methodology

The steps followed in this project include:
- Loading and exploring the dataset
- Text preprocessing and vectorization using CountVectorizer
- Splitting data into training and testing sets
- Training a Naive Bayes classifier
- Evaluating model performance using various metrics

4. Machine Learning Model

The model used in this project is the Multinomial Naive Bayes classifier. It is particularly effective for text classification problems and is based on applying Bayes' theorem with the assumption of feature independence.

5. Performance Metrics

The performance of the model was evaluated using the following metrics:

- Accuracy: 0.99
- Precision: Ham: 0.99, Spam: 0.97
- Recall: Ham: 0.99, Spam: 0.95
- F1 Score: Ham: 0.99, Spam: 0.96

These metrics provide a detailed view of how well the model can identify spam messages. The confusion matrix shows:
- 961 ham messages correctly classified
- 5 ham messages incorrectly labeled as spam
- 141 spam messages correctly classified
- 8 spam messages missed (classified as ham)

Overall, the Naive Bayes model achieved an accuracy of 99%, demonstrating its effectiveness for SMS spam detection.

6. Conclusion
The Naive Bayes classifier was able to effectively classify SMS messages as spam or ham. This model can serve as a foundation for building more sophisticated spam detection systems and can be integrated into messaging apps to improve user experience and safety.