Model Development Phase Template

Date	8-10-2024
Team ID	LTVIP2024TMID25000
Project Title	SMS Spam Detection
Maximum Marks	5 Marks

Model Selection Report

In this project, multiple machine learning models are evaluated to identify the best approach for classifying SMS messages as **Spam** or **Not Spam**. The models are compared based on performance, accuracy, complexity, and computational requirements. The goal is to select the most suitable model for real-time spam detection in a web-based application.

Model Selection Report:

Model	Description
Model 1: Naive Bayes	A probabilistic classifier that uses Bayes' theorem with strong independence assumptions. This model is commonly used for text classification and works well with small datasets. Low computational cost and good performance for spam detection tasks.
Model 2: Decision Tree	A non-parametric supervised learning model used for classification tasks. It works by splitting data into subsets based on feature values. It's easy to interpret but can lead to overfitting.
Model 3: Logistic	A linear model used for binary classification. It predicts probabilities and outputs binary outcomes. Efficient and

Regression	interpretable but may struggle with complex data relationships.
Model 4: Support Vector Machine (SVM)	A powerful classifier that finds a hyperplane separating spam from non-spam messages. SVM is highly accurate but computationally expensive and may require more time to train.

Conclusion

After evaluating these models, the **Naive Bayes classifier** was selected due to its simplicity, low computational cost, and strong performance on text data. The model offers an optimal balance between accuracy and efficiency, making it ideal for real-time SMS spam detection in a web-based application.