Model Development Phase Template

Date	10 April 2025
Team ID	259453
Project Title	SMS Spam Detection using NLP
Maximum Marks	5 Marks

Model Selection Report

In this NLP-based project for SMS spam detection, multiple machine learning models were considered to evaluate their performance in binary text classification tasks. Criteria such as accuracy, simplicity, computational efficiency, and suitability for sparse data were considered during model selection.

Model Selection Report:

Model	Description
Model 1	Multinomial Naive Bayes (MNB) A probabilistic classifier based on Bayes' theorem, particularly effective for text classification with discrete features like TF-IDF. It is fast, interpretable, and performs well on sparse datasets. Ideal for spam detection tasks.
Model 2	Logistic Regression A linear model that estimates probabilities using the logistic function. It can handle large feature spaces and outputs probability scores for classification, making it useful for understanding confidence levels.
Model 3	Support Vector Machine (SVM) SVMs are effective in high-dimensional spaces. With a linear kernel and TF-IDF input, it offers high accuracy but at a slightly higher computational cost than Naive Bayes.

Model 4	LSTM (Optional Deep Learning) An RNN-based model that captures sequential patterns in text. While powerful, it requires more computational resources and is usually used for larger datasets. May be excessive for this task.