Project Initialization and Planning Phase

Date	10 April 2025
Team ID	259453
Project Title	SMS Spam Detection using NLP
Maximum Marks	3 Marks

Project Proposal (Proposed Solution) template

This project proposal outlines a solution to address a specific problem. With a clear objective, defined scope, and a concise problem statement, the proposed solution details the approach, key features, and resource requirements, including hardware, software, and personnel.

Project Overview	
Objective	To develop a machine learning model that detects and classifies SMS messages as spam or ham using natural language processing (NLP) techniques.
Scope	The project focuses on analyzing text-based SMS data, applying preprocessing, training a classification model, and allowing real-time user input to detect spam. It does not cover image-based or voice message filtering.
Problem Statement	
Description	Spam SMS messages are increasing daily and often trick users into clicking malicious links or sharing sensitive information. Manual filtering is inefficient and outdated.
Impact	Automating spam detection improves communication safety, prevents phishing, and enhances user trust by minimizing exposure to harmful content.
Proposed Solution	
Approach	Use Python-based NLP techniques to clean and process SMS text, extract features using TF-IDF, and train a Multinomial Naive Bayes model for binary classification (spam vs ham).
Key Features	- Automated SMS spam classification

Resource Requirements

Resource Type	Description	Specification/Allocation
Hardware		
Computing Resources	CPU/GPU specifications, number of cores	Above i5 8th Gen
Memory	RAM specifications	Above 4GB Ram
Storage	Disk space for data, models, and logs	Above 128 Gb Hdd or SSD(for faster interaction)
Software		
Frameworks	Python frameworks	Flask
Libraries	Additional libraries	tensorflow , pandas, scikit-learn, nltk
Development Environment	IDE, version control	Jupyter Notebook, Google Colab
Data		
Data	Source, size, format	Kaggle dataset given on SmartBridge course platform 5572 sms with 5 columns.