A blue text on a white background

AI-generated content may be incorrect.

SmS Spam Detection

Final Project

**AIDI1003 – Machine Learning Frameworks**

**Name:** Paul Udayan Gomez Jayaprakash  
 Ghassan Saba  
 Aishwariya Balaji  
 Udaya Kumar Siva Kumar

**Date:** 11th of August 2025

**Professor Name**: Ethan Davis

# **Project Proposal**

1. **Objective:**

Develop a machine learning spam detection system model capable of classifying SMS messages as either ham or spam. This can reduce risks such as phishing, fraud, and privacy invasion, and can help protect users and improve communication security.

Problem Statement:  
Nowadays, SMS spam messages are a concern worldwide, leading to wasted time, security breaches, and monetary loss. This project aims to use supervised machine learning techniques to analyze a labeled SMS dataset and build a predictive model to automatically filter spam messages.

Goals:

* Clean, preprocess, and analyze the SMS dataset to identify patterns.
* Train and evaluate machine learning models for spam classification.
* Deploy a spam detection system with high accuracy, precision, and recall.

1. **Scope:**

This project:

* Focuses on **binary classification** of SMS messages **(ham vs. spam**) using the **SMS Spam Collection** dataset containing **5,574 labeled messages.**
* Applies text preprocessing techniques including **stop-word removal** and **TF-IDF vectorization** for feature extraction**.**
* Experiments with multiple classification algorithms :**Logistic Regression, Support Vector Machine (SVM)**,and **Multinomial Naive Bayes** and evaluates them using metrics such as **Accuracy, Precision, Recall, F1 Score, and ROC-AUC.**
* Develop a simple deployment interface using Flask.

1. **Intended Approach:**

* **Data Collection & Preparation**  
  Load the dataset, apply lowercasing and stop-word removal, and convert text to numerical features using TF-IDF vectorization (with built-in tokenization).
* **Exploratory Data Analysis (EDA)**  
  Analyze class distribution, visualize message length, and identify frequent words in ham and spam messages.
* **Model Selection & Building**  
  Implement Logistic Regression, SVM, and Naive Bayes classifiers for binary classification.
* **Hyperparameter Tuning**  
  Use GridSearchCV with 5-fold cross-validation to optimize model parameters.
* **Model Evaluation**  
  Evaluate models using accuracy, precision, recall, F1-score, ROC-AUC, and visualize ROC curves.
* **Model Deployment**

Flask API to take SMS input and return prediction.

* **Presentation**

Insights, model performance, and deployment setup with real-world examples.

**4. Final Deliverables**

By the end of the project, deliverables will include:

* A predictive model for classifying SMS messages as spam or ham.
* A documented end-to-end pipeline for data preprocessing, model training, evaluation, and predictions.
* A Flask-based API ready for deployment, enabling real-time spam detection and easy integration into messaging systems.