# Title of the Project

SMS Spam Classifier

# Objective

The primary objective of the SMS Spam Classifier project is to design and develop an automated system that can accurately classify incoming SMS messages as either 'spam' or 'ham' (not spam). The classification system helps reduce the exposure of users to phishing, promotional messages, and other unwanted content.   
  
This classifier is built using machine learning techniques and natural language processing (NLP) methods, aiming to analyze the text content of each message and make a real-time prediction. The final goal is to implement a lightweight and efficient model that can be integrated into mobile applications, email clients, or telecom infrastructure to protect users from spam messages.

# Dataset Used

The dataset used in this project is the 'SMSSpamCollection', a widely-used dataset available through the UCI Machine Learning Repository. It contains a total of 5,574 English SMS messages labeled as 'spam' or 'ham'.  
  
Key features of the dataset:  
- Two columns: one for the label (spam/ham) and one for the message content.  
- Imbalanced distribution: approximately 13% of the messages are spam, while the remaining are legitimate (ham).  
- Requires significant preprocessing to clean the text and convert it into a suitable format for modeling.

# Model Chosen

The chosen model for this project is the Multinomial Naive Bayes classifier, which is particularly effective for text classification tasks involving word frequency features. The simplicity and efficiency of this model make it a strong choice for real-time SMS filtering.  
  
The model training pipeline includes:  
- Text preprocessing: lowercasing, removing punctuation, stop words, stemming.  
- Feature extraction using TF-IDF (Term Frequency–Inverse Document Frequency) vectorization.  
- Splitting the dataset into training and testing subsets.  
- Fitting the Multinomial Naive Bayes model on the training data.  
- Making predictions on unseen test data.

# Performance Metrics

The following metrics are used to evaluate the performance of the classifier:  
  
- Accuracy: Measures the proportion of total messages correctly classified.  
- Precision: Indicates the percentage of messages classified as spam that are truly spam.  
- Recall: Represents the proportion of actual spam messages that were correctly identified.  
- F1-Score: Harmonic mean of precision and recall, providing a balanced evaluation.  
  
The classifier achieved high precision and recall, making it effective for real-world use cases.

# Challenges & Learnings

Several challenges were encountered during the development of the project:  
  
- Data Preprocessing: SMS data often contains slang, abbreviations, and inconsistent formats, requiring extensive cleaning and normalization.  
- Imbalanced Dataset: The dataset contains far fewer spam messages than ham messages, leading to a class imbalance. Techniques like stratified sampling and careful metric selection helped address this.  
- Model Selection: Exploring different models such as SVM, Logistic Regression, and Random Forest before finalizing Naive Bayes due to its speed and effectiveness with text data.  
-Text Vectorization: Choosing between Count Vectorizer and TF-IDF and fine-tuning parameters for optimal results.  
  
Learnings:  
- Gained hands-on experience with text processing and feature engineering.  
- Understood the trade-offs between various classification algorithms.  
- Learned to evaluate models using appropriate metrics, especially in imbalanced scenarios.  
- Developed a deeper understanding of how spam detection systems work in real-life applications.