BUILDING A SMARTER AI-POWERED SPAM CLASSIFIER

Date	**10-10-2023**
Team ID	**1025**
Project Name	**6112-Building a Smarter AI-Powered Spam Classifier**

TABLE OF CONTENTS:

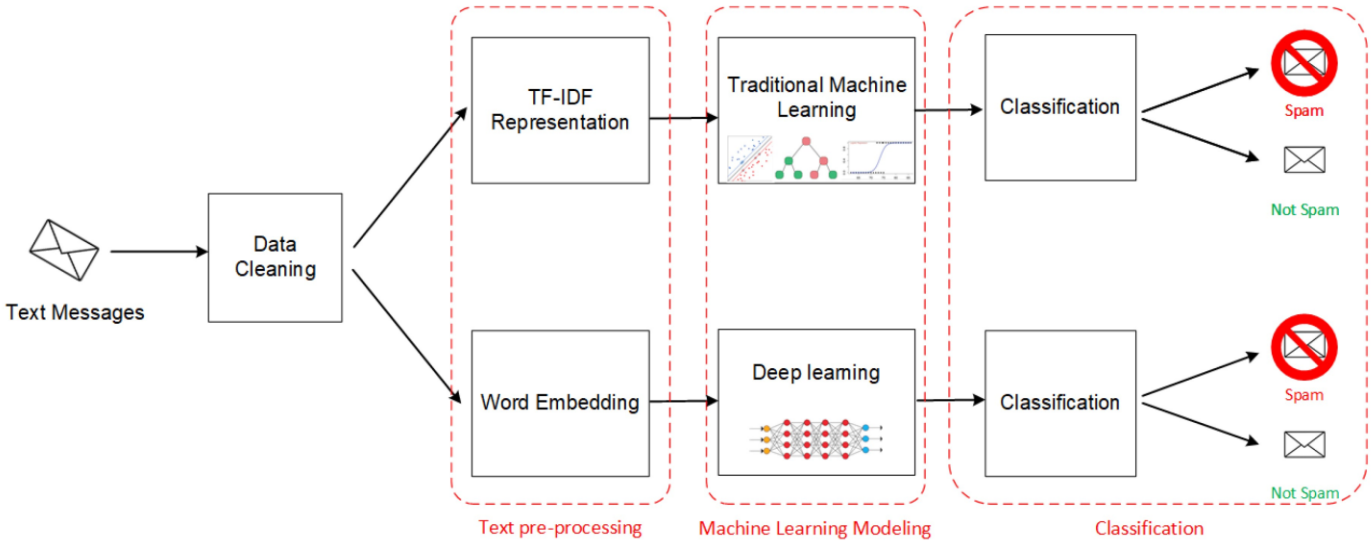
1. INTRODUCTION:

The problem is to build an AI-powered spam classifier that can accurately distinguish between spam and non-spam in emails or text messages. The goals is to reduce the number of false positives (classifying legitimate messages as spam) and false negative (missing actual spam messages) while achieving a high level of accuracy

2. PROBLEM STATEMENT:

Develop an AI-Powered spam classifier using natural language processing (NLP) and Machine learning techniques to accurately distinguish between spam and non-spam messages in email or text messages. The SMS Spam Collection is a set of SMS tagged messages that have been collected for SMS Spam research. It contains one set of SMS messages in English of 5,574 messages, tagged according being ham (legitimate) or spam.

3. DESIGN AND INNOVATION STRATEGIES :



3.1. Data Cleaning :

In filtering of spam, the data cleaning of the textual information is very critical and important. Main objective of data pre-processing in spam detection is to remove data which do not give useful information about the class of the document.

3.2.Text pre-processing :

It is a text pre-processing technique where all words are lowercased so that words like 'cat' and 'CAT' are treated the same way. This technique comes in handy while we are using Bag of words or Tf-Idf for making features out of our natural language data.

3.3 Machine Learning Modeling :

Traditional machine learning approaches have been widely used in spam detection for many years, and they continue to be effective, especially in situations where deep learning models may not be practical due to limited computational resources or when interpretability is a primary concern.

Deep learning techniques have gained popularity in spam classification due to their ability to automatically learn complex patterns and representations from data. They can be highly effective for spam detection when you have access to large datasets and sufficient computational resources.

3.4. Classification (Spam and Non-Spam):

Classification is a fundamental task in machine learning and statistics that involves categorizing data into distinct groups or classes based on their features or attributes. The primary goal of classification is to assign a label or category to a given input data point based on patterns and relationships within the data.

- **Spam messages:**
  - Spam messages are unsolicited and usually irrelevant or inappropriate messages sent to a large number of recipients.
  - They are often sent for commercial purposes, including advertising products, services, or scams.
  - The primary goal of spam messages is to promote products or services, gather personal information, or spread malware.
  - Common types of spam include phishing emails, chain letters, advance-fee fraud, and unsolicited bulk email (UBE).
- **Non-Spam messages:**
  - Non-spam messages are legitimate, desired, and relevant messages sent by individuals, organizations, or businesses to communicate with recipients.
  - These messages include personal emails, work-related emails, newsletters you've subscribed to, social media updates from friends, and more.
  - The content of non-spam messages can vary widely, from personal messages to important work-related information.
  - Recipients generally expect and want non-spam messages, and they form the basis of productive communication.

4. CONCLUSION:

In conclusion, AI-powered spam detection has significantly improved the way we manage unwanted messages in our digital lives. It not only saves time and resources but also enhances the security and overall user experience in our digital communications. As AI technology continues to advance, it holds the promise of even more effective and efficient spam detection solutions in the future.AI-powered spam detection has revolutionized the way we handle unwanted and potentially harmful messages in our digital communication channels.