Proposed Project Name

AI Classification Algorithms Compared to Detect SMS Spam

Team Members

Team Name: Code Crafters

Member Name	Student Number
Gyan Karthik Abhishek Sakibanda	110123230
Karan Anandbhai Dhanwani	110122826
Shashank Kannan	110122650

Problem to be studied

The rise of SMS spam has generated worries about the accessibility and security of mobile messaging systems. This paper explores AI and ML strategies for SMS spam categorization, providing a complete analysis of available methodologies. We investigate classification approaches, with an emphasis on Machine Learning (ML) techniques and explain the crucial significance of feature selection and extraction methods in lowering feature dimensionality. Approaches to SMS spam classification include content-based methods that use character frequency and the "bag of words" model, non-content-based methods that detect anomalies using message size and timestamps, and hybrid methods that combine features from both content and non-content-based approaches. The study intends to contribute to safe mobile message transmission by improving SMS spam identification and categorization using AI approaches.

Keywords: *SMS Spam, SMS Spam Classification, Bag of Words, Text Classification*

Relevant AI History/ AI Techniques

The project on SMS Spam detection encompasses several AI techniques such as:

1. **Logistic Regression**: SMS spam detection with Logistic Regression entails developing a model to categorize text messages as spam or non-spam based on the message content and related characteristics. Each SMS message is represented as a numerical vector in this manner, commonly using techniques such as TF-IDF (Term Frequency-Inverse Document Frequency) or word embeddings, and logistic regression is used as the classification algorithm. The logistic regression model evaluates the likelihood of an incoming SMS being spam, and if the likelihood exceeds a predetermined threshold, the message is classified as spam; otherwise, it is classified as

non-spam. Using a labeled dataset of SMS messages, the logistic regression model is trained by modifying the model's parameters to maximize performance and fine-tuning the decision threshold to meet the desired trade-off between precision and recall. This method uses logistic regression's ability to model binary classification issues to identify SMS spam by learning patterns in text content that distinguish spam from legal texts.

2. **Naïve Bayes Models**: In many situations, SMS spam detection is performed by encoding each SMS message as a numerical vector of features using Naive Bayes models, notably Multinomial Naive Bayes (MultinomialNB). Word frequencies or other text-based characteristics are common examples of these features. The Multinomial Naive Bayes model is trained on a labeled dataset of SMS messages, with each message assigned a spam or non-spam classification. Based on the reported feature frequencies, the model calculates the likelihood of a message belonging to each class (spam or non-spam). When a new SMS is received, the model determines if it is spam or not and assigns the message to the class with the highest likelihood. This methodology provides a simple and effective way for SMS spam identification by utilizing the probabilistic nature of Naive Bayes, since it can capture the specific word patterns and properties associated with spam messages.

3. **Random Forest Classification**: SMS spam detection using Random Forest Classification entails training an ensemble of decision trees known as a Random Forest to categorize text messages as spam or non-spam based on their content and related attributes. Each SMS message is encoded as a numerical vector, frequently using techniques like TF-IDF (Term Frequency-Inverse Document Frequency) or word embeddings, then utilized as input for the Random Forest. The Random Forest approach uses the collective decision-making capability of several decision trees to assess whether an incoming SMS is spam or not. During training, the Random Forest algorithm discovers patterns and correlations in text content that distinguish spam from real communications. By integrating the predictive capabilities of several decision trees, this technique delivers stability and enhanced accuracy in SMS spam detection, making it suitable for real-world applications where distinguishing between spam and non-spam texts may be difficult.

4. **Linear Support Vector Classifier**: The Linear Support Vector Classifier (LinearSVC) is trained using a labeled dataset of SMS texts, with each message tagged as spam or non-spam. The LinearSVC model learns to identify a hyperplane that best separates spam and non-spam feature vectors. When a new SMS comes, its feature vector is utilized to estimate its class, either spam or non-spam, based on where it sits on the hyperplane. LinearSVC is well-suited for high-dimensional data such as text, and it successfully captures complicated decision boundaries, making it a useful tool for SMS spam detection by learning patterns and linguistic qualities

that distinguish spam from real texts.

5. **Stochastic Gradient Descent**: The Stochastic Gradient Descent (SGD) optimization process is used by the classifier to learn the appropriate weights for the characteristics, thereby creating a decision boundary between spam and non-spam messages. When a new SMS comes, the learnt decision boundary is utilized to forecast its class (spam or non-spam). The SGD Classifier is well-known for its training and scaling efficiency, making it an excellent candidate for SMS spam detection by learning and adapting to textual patterns that distinguish spam from authentic texts while rapidly processing big datasets.

6. **Gradient Boosting Classifier**: The Gradient Boosting Classifier is an ensemble model that systematically merges numerous decision trees to increase prediction accuracy. The classifier fits decision trees to the data during training, repeatedly changing the model to remedy mistakes made by earlier trees. This method enables the model to detect intricate correlations and patterns in text content that distinguish spam from authentic communications. When a new SMS arrives, its feature vector is utilized to forecast its class (spam or non-spam) using the ensemble's accumulated knowledge. Gradient Boosting is very useful for SMS spam detection because it adapts to the complexities of the data and gives high prediction accuracy, making it well-suited for real-world applications where the distinction between spam and non-spam texts is subtle.

Expected Workload and Distribution

As a team, we expect to distribute the work as follows:

1. **Data Gathering and Preprocessing (1 Member)**: Responsibilities include collecting and curating the SMS dataset for training and testing, as well as doing data pretreatment activities such as text cleaning, tokenization, and feature extraction. They will also guarantee the accuracy and consistency of the data.

2. **Model Development and Evaluation (2 Members)**: Concentrate on model creation and assessment. Implementing machine learning models such as Logistic Regression, MultinomialNB, Random Forest, LinearSVC, SGD Classifier, and Gradient Boosting Classifier is one of the tasks. The team will focus on optimizing model hyperparameters and fine-tuning model hyperparameters. We will also analyze model performance and generate useful insights from the results.

References

1. Abayomi-Alli, O., Misra, S., Abayomi-Alli, A., & Odusami, M. (2019). A review of soft techniques for SMS spam classification: Methods, approaches and applications. Engineering Applications of Artificial Intelligence, 86, 197–212. https://doi.org/10.1016/j.engappai.2019.08.024

2. WAYS TO STOP SMS SPAM ON iPhone & android - HERE'S HOW… - another website by (CFM) consumer forum Malaysia. (2019, October 16). Another Website by (CFM) Consumer Forum Malaysia. https://consumerinfo.my/ways-stop-sms-spam

3. Gupta, S. D., Saha, S., & Das, S. K. (2021). SMS spam detection using machine learning. Journal of Physics. Conference Series, 1797(1), 012017. https://doi.org/10.1088/1742-6596/1797/1/012017