IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 3, March 2024

DOI: 10.17148/IJARCCE.2024.13307

SMS SPAM DETECTION USING MACHINE LEARNING

Shreya Menthe[1], Kanish Rawal[2], Mrudula Hirave[3], A. J. Patil[4]

Student Department of Computer Technology, Bharati Vidyapeeth's Jawaharlal Nehru Institute of Technology (Poly), Pune, Maharashtra, India[1-3]

Guide, Department of Computer Technology, Bharati Vidyapeeth's Jawaharlal Nehru Institute of Technology (Poly), Pune, Maharashtra, India[4]

Abstract: The proliferation of mobile users has led to a significant increase in mobile messaging, resulting in a rise in SMS (Short Message Service) spam. Unlike other messaging platforms such as Facebook and Whats-app, SMS does not necessitate an active internet connection. Spam SMS messages, which are unwanted and potentially harmful to users, pose a substantial challenge in mobile communication. These messages are primarily aimed at distributing electronic messages for commercial or financial gain. Consequently, combating SMS spam is crucial for preserving the integrity of mobile communication channels. However, existing email filtering algorithms may underperform due to factors such as the lack of real databases for SMS spam, limited features, and informal. This study proposes an approach utilizing Machine Learning techniques to address SMS spam. The approach encompasses various components, including data-set combinations, data cleaning, exploratory data analysis, and feature engineering. Additionally, several machine learning algorithms, such as Naive Bayes and Support Vector Machine, are assessed for model building. The ultimate aim of SMS spam detection is to protect users from spam-related issues.

Keywords: Spam SMS, Facebook, Whats-app, Internet Connection, Financial gain, Data-sets, Data cleaning, Feature engineering, Naive Bayes, Model building.

I. INTRODUCTION

The widespread availability and simplicity of SMS have made it an appealing target for malicious users, resulting in unwarranted expenses for mobile users and compromising Secure Mobile Message Communication. Many individuals and organizations exploit this feature to disseminate unsolicited bulk messages, commonly known as Spam SMS. This project aims to create a robust SMS spam detection system utilizing Machine Learning algorithms. We will investigate various ML algorithms, such as Naive Bayes, Support Vector Machines (SVM), and Random Forests, to assess and categorize SMS messages based on their content, linguistic features, and other pertinent attributes. Through thorough training and evaluation procedures, our goal is to construct a highly precise and efficient spam detection model capable of identifying subtle patterns and characteristics intrinsic to spam messages. Machine Learning presents a promising approach by facilitating the automated detection of spam messages through the recognition of patterns and characteristics learned from labeled data. Multiple ML models, including Naive Bayes, Support Vector Machines, and neural networks, can be trained on features extracted from the text, such as word frequencies, n-grams, and semantic features. Furthermore, feature engineering techniques and preprocessing steps, such as tokenization, and TF-IDF normalization, are pivotal in enhancing the performance of spam detection systems. By continuously updating and refining these models with new data, SMS spam detection systems can adapt to evolving spamming techniques, furnishing users with a dependable defense against unwanted messages while preserving efficient communication and user experience.

II. RELATED WORK

SMS spam detection using Machine Learning has attracted considerable attention recently due to the proliferation of unwanted text messages. One prevalent approach involves employing Natural Language Processing (NLP) techniques to preprocess and analyze message content. Methods like tokenization, which divides text into individual units, and the creation of feature vectors using techniques like the bag-of-words model or TF-IDF, are utilized. These techniques facilitate the transformation of text data into a suitable format for machine learning algorithms to identify patterns indicative of spam.

Various machine learning algorithms have been utilized in SMS spam detection, including Support Vector Machines (SVM), Naive Bayes classifiers, and decision trees. SVMs excel at distinguishing between spam and legitimate messages

DOI: 10.17148/IJARCCE.2024.13307

by separating classes in high-dimensional spaces using extracted features. Naive Bayes classifiers, leveraging their assumption of feature independence and computational efficiency, are well-suited for handling large data-sets. Decision trees offer interpret-ability and can capture complex decision boundaries, thereby enhancing the robustness of spam detection models.

In addition to algorithmic approaches, feature engineering significantly enhances the effectiveness of SMS spam detection systems. Techniques such as word embedding, representing words as continuous vectors, and syntactic features, capturing structural nuances in text, have been explored to extract informative features from SMS messages.

Furthermore, model evaluation metrics such as precision, recall, and F1-score play a crucial role in assessing the performance of spam detection models and guiding the selection of optimal approaches for real-world deployment.

III. METHODOLOGY

A.Feature Extraction: Examine text features such as word frequency, length, and the presence of specific keywords. Utilize methods like TF-IDF (Term Frequency-Inverse Document Frequency) to quantify word importance.

B.Text Preprocessing: Perform tokenization and legitimization to normalize text. Remove stop words and punctuation. Machine Learning Models: Utilize traditional algorithms like Naive Bayes, Support Vector Machines (SVM), or logistic regression.

C.Evaluation Metrics: Assess performance using metrics like precision, recall, F1 score, and accuracy. Employ techniques such as cross-validation for robust evaluation.

D.Handling Imbalanced Data: Tackle class imbalance in spam detection data sets through methods like oversampling, under sampling, or synthetic data generation.

E.Ensemble Methods: Combine multiple models to enhance performance. Consider ensemble techniques such as bagging or boosting.

F.Cross-Validation: Use cross-validation to ensure the model's generalization to unseen data.
G. Real-time Deployment: Investigate options for deploying the model for real-time SMS spam detection, such as employing lightweight frameworks like Flask or Fast API.Background:

IV. BACKGROUND

A. Data Collection: Assemble a data-set comprising SMS messages categorized as spam or non-spam (ham). Explore publicly accessible data-sets, including:
- SMS Spam Collection Data-set: A data-set featuring SMS messages labeled as spam or ham, widely utilized in SMS spam detection studies.
- NUS SMS Corpus: Another data-set containing spam and ham-labeled SMS messages, frequently employed for text classification research.
- Kaggle Data-sets: Platforms like Kaggle offer various data-sets for machine learning tasks, including SMS spam detection. Search for relevant data-sets while ensuring data integrity and accurate labeling.

B. Data Preprocessing:
- Tokenization: Divide each SMS message into individual words or tokens.
- Text Cleaning: Eliminate stop words, punctuation, and irrelevant characters to reduce noise.
- Address Missing Values: Although missing values may not be prevalent in SMS data, ensure appropriate handling if present.
- Vectorization: Convert text data into numerical form using techniques like TF-IDF or word embedding.

C. Feature Engineering: Extract pertinent features from text data to aid in distinguishing between spam and non-spam messages. Features may encompass:
- Message Length
- Presence of Specific Words or Patterns
- Frequency of Certain Words or Phrases

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 3, March 2024

DOI: 10.17148/IJARCCE.2024.13307

D. Model Selection: Opt for a machine learning algorithm suitable for text classification tasks, such as:
- Naive Bayes
- Support Vector Machines (SVM)
- Logistic Regression
- Decision Trees
- Random Forests
- Gradient Boosting Machines (GBM)

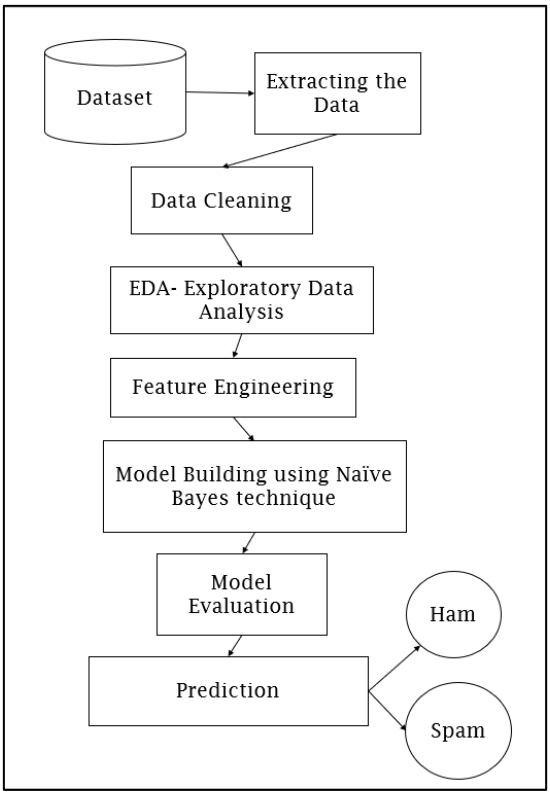
E. Model Training: Partition the data-set into training and testing sets. Train the chosen machine learning model using the training data.

F. Model Evaluation: Assess the performance of the trained model using metrics like accuracy, precision, recall, F1-score, and ROC-AUC score on the testing data-set.

G. Hyperparameter Tuning: Refine the parameters of the machine learning algorithm to optimize performance. Employ techniques like cross-validation to evaluate different hyperparameter configurations, mitigating overfitting and enhancing generalizability.

H. Model Deployment: Deploy the trained model in a production environment for classifying incoming SMS messages as spam or non-spam. Enable users to access classification results and offer feedback on message accuracy. Implement reporting capabilities for generating insights and analytics on message classification performance.

V. SYSTEM MODEL



A. Data-sets:
In SMS spam detection, the data-set forms the backbone for training the machine learning algorithm. It comprises labeled SMS messages, categorized as either spam or non-spam (ham). During training, the algorithm assimilates patterns and features from the data-set to differentiate between spam and non-spam messages. These features encompass word frequencies, character n-grams, and more. Subsequently, the model employs these learned patterns to predict the spam status of new, unseen messages. Continuously updating and enhancing the data-set aids in refining the model's accuracy and adaptability over time.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 3, March 2024

DOI: 10.17148/IJARCCE.2024.13307

B. Data Extraction:
In SMS spam detection utilizing machine learning, data extraction involves procuring a data-set containing SMS messages tagged as spam or non-spam. This data undergoes preprocessing to cleanse and prepare it for analysis. Features are extracted from the text, typically employing techniques like Bag-of-Words or TF-IDF. The data-set is partitioned into training and testing sets, and a suitable machine learning model is selected and trained on the training data. The model's efficacy is evaluated using the testing data, with potential fine-tuning conducted to optimize performance. Finally, the trained model is deployed to classify new SMS messages as spam or non-spam.

C. Exploratory Data Analysis:
Exploratory Data Analysis (EDA) involves visually and statistically examining data-sets to comprehend their fundamental characteristics, often preceding more formal modeling approaches. It encompasses summarizing key data-set attributes, frequently through visual aids such as histograms, scatter plots, or box plots. EDA aids in identifying patterns, anomalies, relationships, and trends within the data, thereby guiding subsequent modeling decisions. Essentially, it serves as an initial exploration to garner insights and steer further analysis.

D. Feature Engineering:
Feature engineering entails selecting, creating, or transforming features from raw data to enhance the performance of machine learning models. It encompasses tasks such as selecting pertinent features, addressing missing data, encoding categorical variables, scaling numerical features, and generating new features through methods like polynomial features or feature crosses. The objective of feature engineering is to furnish the model with the most relevant and informative input variables, thereby ameliorating its predictive accuracy and robustness.

E. Model Building:
Model building employing Naive Bayes entails training a classification model grounded on Bayes' theorem, under the "naive" assumption of feature independence. It computes the probability of a given sample belonging to each class, predicated on the probability of features given the class. The model undergoes training using a labeled data-set, where each instance comprises features and a corresponding class label. During prediction, the model computes the probability of each class for a new instance and selects the class with the highest probability as the predicted class. Naive Bayes finds particular utility in text classification tasks, such as spam detection, sentiment analysis, and document categorization, owing to its simplicity and efficacy with high-dimensional data.

F. Model Evaluation:
Model evaluation in SMS spam detection involves gauging the performance of a machine learning model in accurately categorizing SMS messages as spam or non-spam. This typically encompasses metrics like accuracy, precision, recall, and F1-score, gauging the model's capacity to correctly classify spam and non-spam messages while minimizing false positives and false negatives. Evaluation aids in determining the model's effectiveness in identifying spam messages, crucial for deploying a dependable spam detection system.

G. Predictions:
In the prediction phase of SMS spam detection using machine learning, the trained model processes new, unseen SMS messages to predict their spam status. This prediction relies on features extracted from the text and patterns learned from the training data. The model assigns a probability or class label to each message, indicating whether it is classified as spam or non-spam. This predictive process is indispensable for real-time spam detection, where swift classification of incoming messages is imperative for user safeguarding.

VI. CONCLUSION

Spam detection is critical for safeguarding message and email communication. Achieving accurate spam detection poses a significant challenge, prompting numerous researchers to propose various detection methods.

Employing machine learning for SMS spam detection offers a versatile and resilient solution that can be customized to specific requirements. Through ongoing enhancements and model refinement, we can effectively address the continually evolving landscape of SMS spam, ensuring users benefit from enhanced security and a more enjoyable messaging experience.

These models can attain remarkable levels of accuracy and precision in distinguishing spam messages, thereby serving as invaluable assets for SMS filtering.

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 3, March 2024

DOI: 10.17148/IJARCCE.2024.13307

REFERENCES

[1]. *Nagre, S.*, "Mobile SMS Spam Detection using Machine Learning Techniques," JETIR, December 2018, Volume 5, Issue 12.

[2]. *Julis, M. R., Alagesan, S.*, "Mobile SMS Spam Detection using Machine Learning Techniques," INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH, Volume 9, Issue 02, February 2020.

[3]. *Kavya, P., Dr. Rengarajan, A.*, "A Comparative Study for SMS Spam Detection," International Journal of Trend in Scientific Research and Development (IJTSRD), Volume 5, Issue 1, November-December 2020.

[4]. *Lota, L. N., Hossain, B. M. M.*, "A Systematic Literature Review on SMS Spam Detection Techniques," I.J. Information Technology and Computer Science, 2017, 7, 42-50, Published Online July 2017 in MECS (http://www.mecs-press.org/), DOI: 10.5815/ijitcs.2017.07.05.

[5]. *Patel, A., Jhariya, P., Bharath, S., Wadhawan, A.*, "SMS Spam Detection using Machine Learning Approach," IJCRT, Volume 9, Issue 4, April 2021, ISSN: 2320-2882.

[6]. *Warade, S. J., Tijare, P. A., Sawalkar, S. N.*, "An approach for SMS spam detection.