WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

SMS Classification and Spam Detection by Using RNN

Mr. Ritweek R.[1], Mr. Saurabh K.[2], Mrs. Bhagya. M[3], Mrs. Aruna M G[4], Dr. Malatesh S H[5]

[1,2]Student, [3]Associate Professor, [4]Professor and Head of Department, Department of AI &ML, [5]Professor and Head of Department, Department of CSE
M. S. Engineering College, Bangalore, Karnataka, India

ABSTRACT

SMS, or Short Message Service, is a mobile communication service that enables convenient and cost-effective communication. However, the prevalence of unwanted messages for advertising or harassment purposes has emerged as a significant challenge within this service. Numerous approaches have been developed to identify unsolicited short messages, many of which leverage machine learning techniques. While Neural Networks have been employed for distinguishing between unwanted text messages (spam) and regular short messages (ham) in SMS, to the best of our knowledge, Recurrent Neural Network (RNN) has not been utilized in this context thus far. In this study, we propose a novel method that utilizes RNN to classify ham and spam with sequences of varying lengths, despite employing a fixed sequence length. Our proposed method achieves a noteworthy improvement, with an accuracy of 98.11.

Introduction

The mobile phone market has witnessed significant growth in recent years. In the second quarter of 2013, a total of 432.1 million mobile phones were shipped, reflecting a 6.0% year-over-year increase. With the widespread use of mobile phone devices, the Short Message Service (SMS) has emerged as a thriving commercial industry worth billions of dollars. SMS allows mobile phone users to exchange brief text messages, typically limited to 160 seven-bit characters. It stands as the most widely utilized data application, boasting an estimated 3.5 billion active users, accounting for approximately 80% of all mobile phone subscribers by the end of 2010. However, as the platform's popularity has grown, so has the influx of unsolicited commercial advertisements sent via SMS, commonly referred to as SMS spam. Although SMS spam is not as prevalent as email spam, which constituted around 90% of emails in 2010, it remains a minor issue in North America, accounting for less than 1% of text messages exchanged as of December 2012. Nevertheless, with the increasing popularity among younger demographics and the decreasing costs of text messaging (e.g., in China, the cost of sending a text message is now less than $0.001), SMS spam has been on the rise. In some parts of Asia, up to 30% of text messages in 2012 were classified as spam.

Existing System

Various classification algorithms, including Random Forest (RF), Decision Tree (DT), Neural Network (NN), K-Nearest Neighbor (KNN), and others, have been extensively studied for predicting SMS spam. However, the performance of these algorithms varies significantly, indicating the potential for improvement by exploring different training and testing ratios or combining multiple techniques. Despite ongoing efforts, SMS spam prediction remains a challenging task. Therefore, it is crucial to carefully select appropriate methods for classifying SMS spam within a specific region. In recent years, machine learning algorithms have been proposed to enhance the accuracy of SMS spam prediction.

Table 1:- Literature Summary

S.NO.	TITLE	METHODODLGY	LIMITATIONS
1.	Exploiting Latent Content based Features for Detection of Static SMS Spam (2013)	Random Forest and SVM Algorithms.	It requires large amount of resources.
2.	SMS spam Detection using Machine Learning Approach (2012)	Python using Scikit-learning library.	It requires large amount of resources.
3.	Using the extracted features as input for an average neural network (2020)	Accuracy and F-measure Metrics.	We have studied about the data theft in the recent years.

PROPOSED SYSTEM

The study presents a novel approach for precise prediction of unwanted SMS messages by leveraging the power of Recurrent Neural Networks (RNN). In order to enhance the accuracy of the predictions, the proposed method incorporates preprocessing actions as an integral part of the algorithm. By employing RNNs, the method effectively categorizes short messages as either unwanted or normal. The process involves inputting a sequence of marked text as inputs to the RNN. The final output of the RNN is then utilized for classification, where a value of zero signifies an unwanted message, while a value of one indicates a normal message. This approach offers a robust solution for the accurate identification and classification of unwanted SMS messages using RNNs.

SYSTEM ARCHITECTURE

According to Figure 1.1, SMS spam is similar to email spam in that it involves the dissemination of unsolicited bulk messages with a commercial interest. SMS spam is commonly used for commercial advertising purposes and for spreading phishing links. Commercial spammers often employ malware to send SMS spam, as it helps obscure the origin of the spam and reduces the risk for the spammer, as sending SMS spam is illegal in many countries.

SMS messages are limited in terms of characters, allowing only alphabets, numbers, and a few symbols. Upon examining the content of the messages, a clear pattern emerges. The majority of spam messages prompt users to call a specific number, reply via SMS, or visit a particular URL. This pattern can be observed by analyzing the results obtained from a simple SQL query on the spam corpus, indicating a consistent trend in the nature of SMS spam messages.

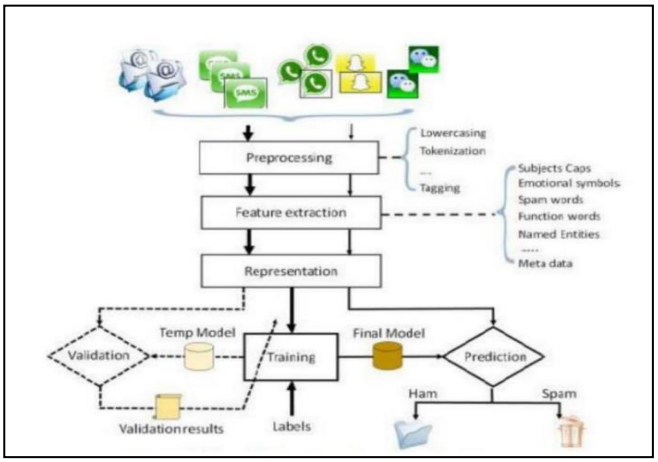

Fig. 1.1 System Architecture

METHODOLOGY

1. Spam detection Exploiting Latent Content base :-

With the rapid proliferation of mobile phones, the prevalence of spams in mobile communication, particularly SMS, has escalated, necessitating further research on SMS spam detection. Most existing techniques for detecting SMS spams have been adapted from methods developed for other contexts, such as emails and the web. However, it is important to recognize that SMS possesses distinct characteristics that warrant specific consideration in spam detection approaches. Therefore, there is a need for research that takes into account the unique attributes of SMS to develop more effective and tailored techniques for identifying and combating SMS spams.

Limitations:- It requires large amount of resources.

2. SMS Spam Detection using Machine Learning Approach :-

In recent years, the widespread adoption of mobile phone devices has led to the exponential growth of the Short Message Service (SMS) industry, which now generates billions of dollars in revenue. However, alongside this growth, there has been an increase in unsolicited commercial advertisements, commonly referred to as SMS spams, being sent to mobile phones.

The reduction in the cost of messaging services has played a role in the proliferation of SMS spams. With lower costs, it has become more feasible for spammers to send large volumes of unsolicited messages, targeting mobile phone users with commercial advertisements.

This issue is particularly prominent in certain regions of Asia, where up to 28% of text messages were classified as spam in 2022. The high prevalence of SMS spams in these areas highlights the significance of addressing this problem and implementing effective measures to combat unwanted messages and protect mobile phone users from intrusive and potentially malicious content.

Limitation:- It requires large amount of resources.

3. Using the extracted features as input for an averaged neural network:-

The remarkable advancements in deep neural networks (DNN) have demonstrated significant improvements in various tasks, motivating our decision to apply DNN specifically for the image classification task. In our study, we utilized a specific type of DNN known as Deep Convolutional Neural Networks (DCNN) for the purposes of feature extraction and image classification.

DCNNs are well-suited for image-related tasks due to their ability to effectively capture and analyze spatial features within images. In our work, we leveraged the power of DCNNs to extract relevant features from images and classify them accurately.

It's worth noting that neural networks can serve multiple purposes, including both classification and feature extraction. In our study, we treated these as two distinct tasks. The first task involved utilizing the DCNN for feature extraction, allowing us to extract meaningful and discriminative features from the images. The second task entailed using the extracted features for image classification, enabling us to accurately classify the images into predefined categories.

By leveraging the capabilities of DCNNs and dividing our work into these two tasks, we aimed to achieve improved performance and accuracy in image classification compared to traditional methods.

Limitation:- We have studied about the data theft in the recent years.

SYSTEM IMPLEMENATION

LIST OF MODULES :-

1. Collecting Dataset
2. Data Pre-processing
3. Training Models

1. Collecting Dataset:-

In our study, we obtained a collection of 425 SMS spam messages from the Grumble text website. This particular website is a UK-based forum where cell phone users publicly share their experiences and complaints about SMS spam messages. It is important to note that most of the users do not report the actual spam message they received, making the identification of the text content of the spam messages a challenging and time-consuming task.

To gather the spam messages, we had to meticulously scan through numerous web pages on the forum, carefully examining the claims made by users to identify instances where they mentioned the text content of the spam messages they received. This manual extraction process required significant effort and attention to detail.

By obtaining this collection of real-world SMS spam messages, we aimed to provide a valuable dataset for further analysis and research in the field of SMS spam detection and mitigation.

2. Data Pre-processing:-

The statement accurately captures the purpose and importance of preprocessing in machine learning. Preprocessing refers to the transformation and preparation of raw data to make it suitable for use in machine learning models. The primary objective of preprocessing is to ensure that the data is structured, clean, and properly formatted, as this facilitates more accurate and effective results from the applied machine learning algorithms.

Data formatting involves converting the data into a consistent and standardized format, which may include handling missing values, resolving inconsistencies, and ensuring uniform units of measurement. Cleaning the data involves removing any noise, outliers, or irrelevant information that could potentially introduce bias or hinder model performance. This step may also involve data normalization or scaling to bring features to a similar scale and prevent certain attributes from dominating the learning process.

Sampling is another preprocessing technique that involves selecting a representative subset of data from a larger dataset. This can be useful when dealing with large or imbalanced datasets, as it helps reduce computational complexity and can improve model performance by providing a more balanced representation of the data.

By performing these preprocessing steps, data scientists can enhance the quality and usability of the data, resulting in more accurate and reliable machine learning models. Proper preprocessing sets the foundation for successful model training, evaluation, and deployment.

3. Training Models:-

The study incorporates a Long Short-Term Memory (LSTM) network to process the extracted features. An LSTM network is an extension of recurrent neural networks (RNNs) that addresses the issue of capturing long-term dependencies in sequential data. In this context, the LSTM units serve as building blocks for the layers of the RNN, resulting in what is commonly referred to as an LSTM network.

The proposed RNN in the study combines both locally recurrent and globally feed-forward structures. This design allows for dynamic properties and clear computational modeling within the proposed RNN. The network includes an input layer, hidden layer(s), and an output layer. It also incorporates feedback connection weights, activation functions, and interconnection weights to facilitate information flow and computations.

By utilizing the internally connected feedbacks, the proposed RNN aims to capture and leverage temporal dependencies and patterns in the data. This enables a more comprehensive understanding and analysis of the sequential nature of the input data.

The mathematical function of each node within the proposed RNN structure is described and defined according to the specific requirements and objectives of the study.

Overall, the combination of LSTM units and the locally recurrent and globally feed-forward structure in the proposed RNN enhances the network's ability to process sequential data and capture long-term dependencies, making it a suitable choice for the SMS spam detection task.

DATA FLOW DIAGRAM

Figure 1.2 illustrates the overall process of the project, which focuses on detecting fake news using machine learning techniques and comparing their performance. The process can be summarized as follows:

1. Input Dataset: The project utilizes a news dataset as the input, which likely contains both real and fake news articles.

2. Preprocessing: The dataset undergoes preprocessing steps to transform and prepare the data for machine learning algorithms. This may involve tasks such as text cleaning, removing stopwords, tokenization, and vectorization to represent the news articles as numerical features.

3. Feature Extraction: The next step involves extracting relevant features from the preprocessed news articles. Various techniques can be employed, such as TF-IDF (Term Frequency-Inverse Document Frequency) or word embeddings like Word2Vec or GloVe. These techniques capture the semantic meaning or statistical properties of the words in the articles.

4. Machine Learning Algorithms: Multiple machine learning algorithms are applied to the extracted features to build predictive models for fake news detection. This could include algorithms such as Naive Bayes, Support Vector Machines (SVM), Random Forest, or Neural Networks. Each algorithm is trained on the labeled dataset, where the labels indicate whether an article is real or fake.

5. Performance Evaluation: The performance of each machine learning algorithm is assessed using appropriate evaluation metrics such as accuracy, precision, recall, and F1-score. These metrics provide insights into how well the algorithms are able to detect fake news in the given dataset.

6. Performance Comparison: The performance of the different machine learning algorithms is compared to identify which algorithm yields the best results in terms of accuracy and other metrics. This allows for the selection of the most effective algorithm for fake news detection.

By following this process, the project aims to efficiently detect fake news using various machine learning algorithms and provide a comprehensive performance comparison to identify the most successful approach. This can contribute to the development of more accurate and reliable systems for automated fake news detection.

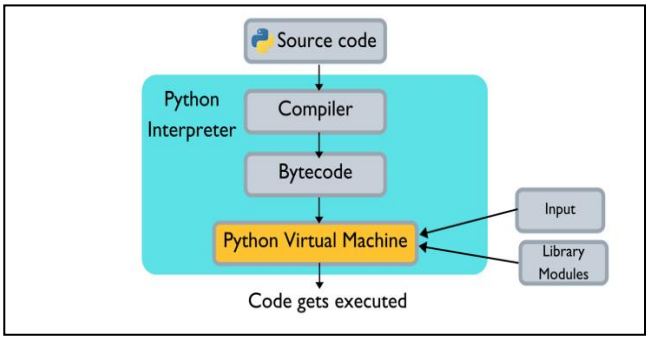

Fig 1.2 :- Level 0 Data Flow Diagram

RESULTS AND DISCUSSION

1. This figure shows the main screen of web application.

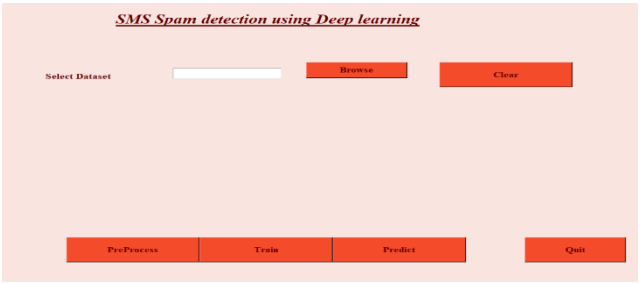

Fig 1.1 :- Main screen of web application

2. This figure shows the sign up page, from here user have to first sign up.

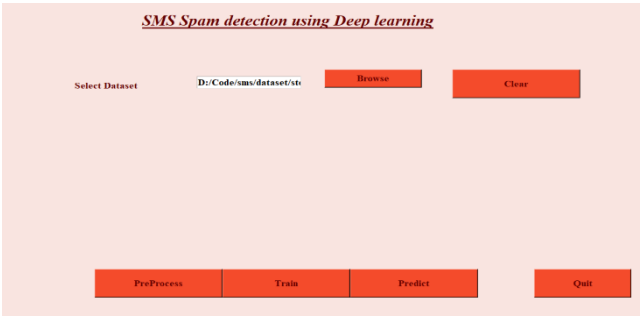

Fig 1.2 :-Sign Up Page

3. This figure shows the preprocessor of our project.

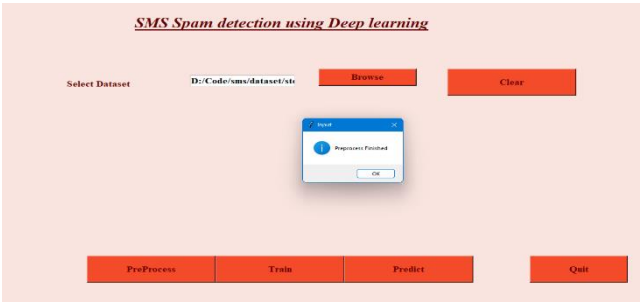

Fig 1.3 :- Preprocessor

4.The below figure shows the Final Prediction of our project.

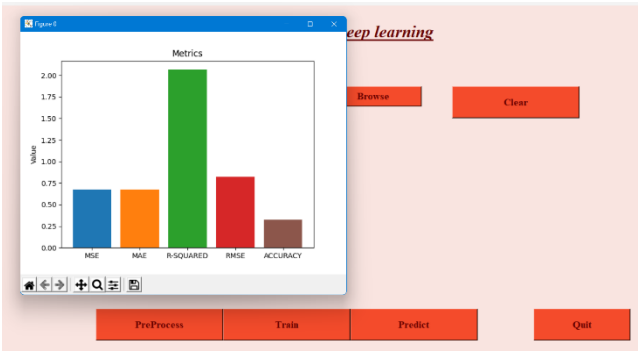

Fig 1.4 :-Prediction Result

CONCLUSION

The project work focuses on developing a classification method for detecting unwanted and normal messages, specifically in the context of SMS spam detection. Previous studies have predominantly utilized SVM or Bayesian methods for this task. However, the proposed method in this project employs Recurrent Neural Networks (RNNs), which is a novel approach in this domain.

To evaluate the effectiveness of the proposed method, test results on standard datasets, such as the UCI SMS spam dataset, are presented. The results demonstrate the efficiency and accuracy of the proposed method. Specifically, when the system reaches a steady state, it achieves a high accuracy rate of around 70%.

One notable aspect of the proposed method is its integration of the pre-processing stage as part of the classification algorithm. This suggests that pre-processing actions are considered within the model itself, rather than being treated as a separate step. This approach leads to higher accuracy compared to recent studies, indicating the effectiveness of the proposed method in addressing the SMS spam detection problem.

Furthermore, the proposed method offers an acceptable runtime, making it a practical and efficient alternative to previous methods. This suggests that the proposed method can be implemented in real-world scenarios, where timely detection of unwanted messages is crucial.

Overall, the project work introduces a novel classification method for SMS spam detection, utilizing RNNs and integrating pre-processing actions within the algorithm. The method achieves high accuracy, outperforming recent studies, and exhibits an acceptable runtime, making it a promising approach for detecting unwanted messages in SMS communication.

REFERENCES

[1] Al Moubayed N., Breckon T., Matthews P., McGough A.S. (2016) SMS Spam Filtering Using Probabilistic Topic Modelling and Stacked Denoising Autoencoder. In: Villa A., Masulli P., Pons Rivero A. (eds) Artificial Neural Networks and Machine Learning ICANN 2016. Lecture Notes in Computer Science, vol 9887.

[2] Karami, Amir and Zhou, Lina. (2014) Improving static SMS spam detection by using new content-based features, Twentieth Americas Conference on Information Systems, Savannah.

[3] Wuying Liu,Ting Wang , Index-based Online Text Classification for SMS Spam Filtering, JOURNAL OF COMPUTERS, VOL. 5, NO. 6, JUNE 2010, doi:10.4304/jcp.5.6.844-851, PP844-851

[4] A comparative study for content-based dynamic spam classification using four machine learning algorithms, B. Yu, Z. Xu, Knowl. Based Syst. (2008), doi: 10.1016/j.knosys.2008.01.001

[5] K. Mathew and B. Issac, "Intelligent spam classification for mobile text message," Proceedings of 2011 International Conference on Computer and Network Technology, Harbin, 2011, pp. 101-105. doi: 10.1109/ICCSNT.2011.618191.

[6] Gordon V. Cormack, José María Gómez Hidalgo, Enrique PuertasSánz, Feature Engineering for Mobile (SMS) Spam Filtering, SIGIR'07, July 23–27, 2007, Amsterdam, The Netherlands. ACM 978-1-59593-597-7/07/0007

[7] Dipak R. Kawade, Dr. Kavita S. Oza , SMS Spam Classification using WEKA, International Journal of Electronics