SMS SPAM DETECTION USING DEEP LEARNING

Prof. Manjunatha P V[1] ,Sri Narahari C N[2], Sriram Lakshmi Narasimha[3],

Tarun Muthyala[4], Rakshith R[5]

Department of Computer Science and Engineering, SJCIT, Chikkaballapur, Karnataka, India[1-5]

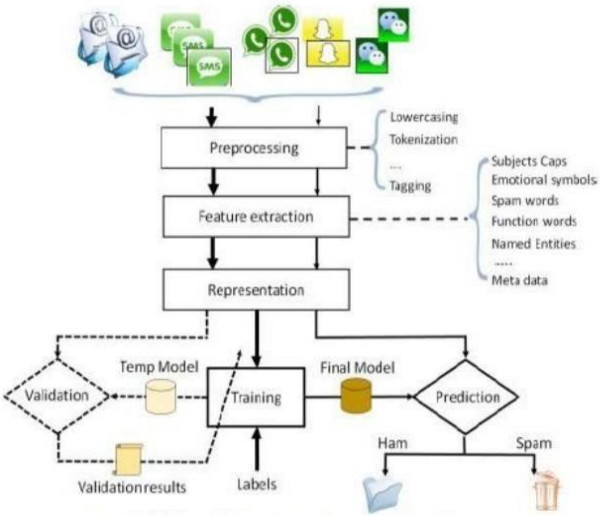
Abstract— Short Message Service (SMS) is one of the mobile messaging apps that enables quick and simple and reasonably priced communication. The main problem in this is creating undesired messages for the purpose of advertising or harassment and sending these messages via SMS. Service. Unwanted short message spam has been detected using a variety of techniques, many of which are machine learning-based. In order to distinguish between SMS's legitimate brief messages (known as ham) and unwelcome text messages (known as spam), neural networks have been used. Recurrent Neural Network (RNN) hasn't yet been applied in this problem, as far as we know. Although we used predetermined sequence lengths in this study, we offered a novel approach that makes use of RNN to distinguish between ham and spam. The accuracy of the suggested method, which was 98.11, shows a significant improvement.

I. INTRODUCTION

In this perspective, unwanted bulk SMS messaging 16 with some corporate interest is very comparable to email spam. SMS spam is used to disseminate phishing URLs and commercial advertising. Because sending SMS spam is typically unlawful, commercial spammers utilise malware to send SMS spam. Spammers run less of a risk when they send spam from a compromised computer since it makes it difficult to determine where the spam came from. SMS messages are only allowed a maximum of seven characters, which are made up of letters, numbers, and a few symbols. A quick scan of the messages reveals a distinct trend. Almost all spam messages invite recipients to contact a number, send a text message in response, or go to a certain URL. The outcomes of a straightforward SQL query on the spam corpus reveal this pattern.

II. METHODOLOGY

This section explains the methods and equipment used to use the 42 different meteorological features for rainfall forecast. The management of missing data, handling of outliers, normalisation, training and testing, the application of the prediction model, and the outcomes of the 38 performances of the models are all part of the classification framework utilised in this article, as illustrated in Figure 1.



III. DESIGN AND IMPLENTATION

The process of applying multiple requirements and enabling their actual realisation is referred to as the "design" of the system. The system is developed using a variety of design principles; the design specification outlines the features of the system, its competitors or constituent parts, and how they will appear to end users. Data flow across a process or system is depicted using a data-flow diagram, which is typically an information system. The DFD additionally gives details about each entity's inputs and outputs as well as the process itself. A data-flow diagram lacks control flow, loops, and decision-making processes. Using a flowchart, certain operations based on the data can be depicted. The project's overall process is described by Level: 0. We are trained as input and are images. The deep learning method will effectively improve the underwater image's quality. Data preparation. Take the stop words out. Stop words are ones that appear a great deal in any text. For instance, the terms "the," "a," "an," "is," "to," etc. We learn nothing about the text's content from these words. Therefore, if we leave these words out of the text, it shouldn't matter. During the pre-process, the unique characters and symbols are eliminated. White spaces and punctuation signs are employed to separate the text; these tools are referred to as tokenizers, and our dataset makes use of one. Tokenizers are used to extract the features, which are then input to the training model.

IV. RESULTS AND CONCLUSION

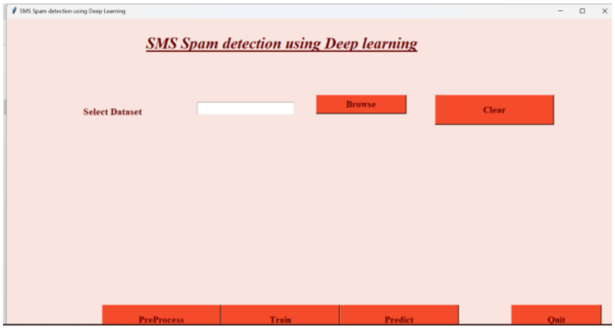
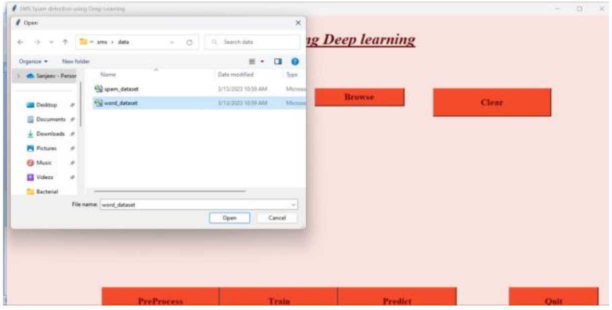

SMS Spam detection using Deep Learning
SMS Spam detection using Deep learning
Select Dataset
Browse
Clear
PreProcess
Train
Predict
Quit

Figure – 1: Snapshots Home page

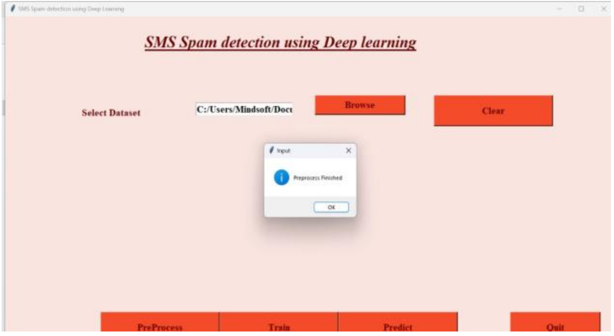


spam_dataset 3/13/2023 10:59 AM Microso

word_dataset 3/13/2023 10:59 AM Microso

File name: word_dataset

Figure – 2: Loading Dataset


SMS Spam detection using Deep Learning
SMS Spam detection using Deep learning
Browse
Clear
Input
OK
PreProcess
Train
Predict
Quit

Select Dataset C:/Users/Mindsoft/Docı

Preprocess Finished

Figure – 3: Pre-process

DOI: 10.17148/IJARCCE.2023.125243

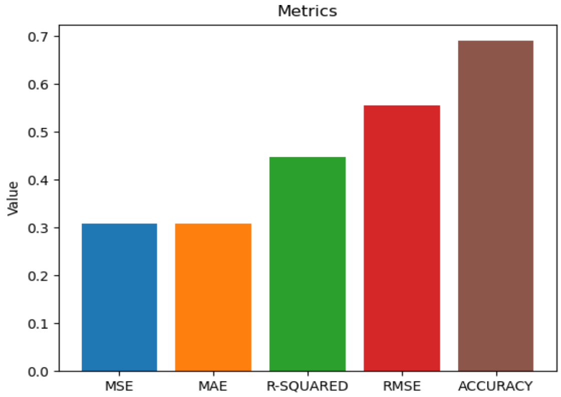


Figure – 4: Training accuracy

Identifying SMS spam is a crucial undertaking that can shield users from unwanted and potentially dangerous communications. There are numerous methods and strategies for identifying SMS spam, including statistical filtering, machine learning, sender reputation, keyword-based filtering, and user input. An efficient spam detection system that can correctly categorise SMS messages as spam or non-spam can be created using a combination of these strategies. Data gathering, data pre-processing, feature extraction, model training, and evaluation are typical processes in the methodology for constructing such a system. The methodology must be kept current because spam strategies and texts change over time. The user experience can be enhanced and users can be shielded from security threats by using an efficient SMS spam detection system.

REFERENCES

[1] "SMS Spam Detection using Machine Learning Techniques," International Journal of Advanced Research in Computer Science and Software Engineering, vol. 9, no. 4, pp. 1-6, April 2019. S. S. Shinde, S. A. Shinde, and S. R. Ganorkar.

[2] "SMS Spam Filtering Based on a Multi-level Feature Extraction Method," Journal of Networks, vol. 13, no. 5, May 2018, pp. 548-555.

[3] M. S. Ahmadi, S. A. Talebi, and S. A. Razavi, "SMS Spam Detection Using Machine Learning Techniques," Proceedings of the 2017 Second Conference on Knowledge-Based Engineering and Innovation (KBEI), Tehran, Iran, December 2017, pp. 315–320.

[4] "SMS Spam Detection Using Machine Learning Techniques: A Comparative Study," in M. M. Islam, M. R. Islam, and M. R. Hasan International Conference on Computer, Communication, Chemical, Materials and Electronic Engineering (IC4ME2) Proceedings, Rajshahi, Bangladesh, April 2018, pp. 1-

[5] "SMS Spam Detection Using Nave Bayes Algorithm," International Journal of Computer Applications, vol. 182, no. 10, pp. 7–12, Feb. 2019, by A. W. Al-Fahadawi, N. A. Hamza, and H. A. Mohammed.

[6] S. Zhang, X. Zhang, & F. Liu (2016). a look at SMS spam filtering methods. 63, 42–60, Journal of Network and Computer Applications.

[7] Alazab, M., Watters, P., Venkatraman, S., & Alazab (2019). Identifying SMS spam: A review of recent studies. IEEE Access 7, pages 46544–46655.

[8] (2016). Fattah, S. A., Muda, Z., and Omar. An analysis of SMS spam filtering methods. 8(10), 145–152, Journal of Telecommunication, Electronic and Computer Engineering.

[9] Zhang, W., Li, W., Li, Y., and Guo, X. (2018). employing machine learning techniques to detect SMS spam. 9(1), 207–218 Journal of Ambient Intelligence and Humanised Computing.

[10] Guzman, J. L., Peralta, & Garcia, F. D. SMS spam filtering with machine learning. 368–375 in Procedia Computer Science, volume 63.