Available online @ https://jjem.jnnce.ac.in
https: www.doi.org/10.37314/JJEM.SP0258
Indexed in International Scientific Indiexing (ISI)
Impact factor: 1.395 for 2021-22
Published on: 08 December 2024

DETECTION OF SPAM SMS

Sneha D [1], Adarsh MJ [2]
[1]Student, [2]Assistant Professor, Department of Computer Applications
JNN College of Engineering, Shivamogga

snehadeepu6@gmail.com,adarshmj@jnnce.ac.in

Abstract

The speedy development of technology and the well-known use of mobile phones have introduced various risks, such as spam and phishing attacks. Machine learning is one of the top extensively used and renowned technologies for detecting spam in communications. This work integration machine learning techniques like logistic regression and other classifiers to build a spam detection model. Various data analysis techniques will be employed to predict and classify spam information from user data, ensuring a clear separation of spam from legitimate information the ulti mate aim is to develop a robust model that enhances information categorization thereby ensuring secure data storage on devices. Feature extraction methods help classify messages as spam or legitimate with high accuracy. This system enhances communication safety and reduces spam impact. It significantly improves the reliability of cellular interactions for consumers.

Keywords: SMS, spam detection, machine learning, algorithms, CNN algorithm

1. Introduction

SMS Spam has been increasingly prevalent in last few years sms spam is explained as any fictitious text message that is distributed via a mobile network without the recipients knowledge they have concerns about users 68% of users have been impacted by sms spam according to a recent survey sms spam can include malicious actions such as smishing a form of smishing is a cybersecurity attack targeting mobile users where spam sms messages are sent containing a link malicious software or both aiming to deceive the recipient it is made up of two words sms and phishing that are joined sms spam can include malicious activities such as smishing a form of smishing is a cybersecurity attack targeting mobile users where spam sms messages aresent containing a link malicious software or both aiming to deceive the recipient sms spam has become increasingly common in last few years it is determined as any fraudulent text message sent over a mobile network without the recipients consent these messages are a significant concern for users. Because of the ubiquity of mobile phones, SMS has emerged as a widely used communication channel. However, this convenience has unfortunately spurred a rise in SMS spam messages, which pose significant annoyances and risks. These messages can be employed for various malicious purposes such as phishing, identity theft, and other harmful activities. Therefore, it is essential to develop effective methods for detecting and filtering out these unwanted SMS spam messages.

2.Related Work

Abhishek Patel et al.,[1] utilized both the svm and nb techniques on the data in this preliminary study the svm-based model proved to be the most effective achieving a precision of 97.64% the nb model was a close second with a preci- sion of 97.50%.

Dhananjay Bhagat et al.,[2] proposed the evalu- ation of the sms spam dataset with various clas- sification models underscored the notable per- formance of multinomial NB with LaPlace smoothing and svm using a linear kernel ac- cording to the original study svm emerged as the top classifier achieving an impressive precision score of 94.89% similarly nb demonstrated strong performance with a precision of 97.07% and exhibited high accuracy in its analysis.

Houshmand Shirani et al.[3] proposed on sms spam filter demonstrated prominent

improvements in the precision of spam detection models by leveraging a database from the uci machine learning repository the project involved comprehensive pre-processing the final results evaluated using 10-fold cross-validation showed a notable reduction in the overall error rate more than halving the error rate of the best model from the original study citing this dataset this implies a substantial enhancement in accuracy suggesting that the best classifier in Shirani Mehrs work achieved an accuracy likely approaching or exceeding 97.5% a significant improvement over traditional models which typically achieve around 90% to 95% accuracy.

Sneha et al.,[4] aim to improves sms spam detection through an efficient implementation of the nb algorithm enhanced by thorough preprocessing steps that boost classifier precision research supports the effectiveness of NB often matching more complex models future studies could investigate hybrid models or incorporate advanced nlp techniques for enhanced spam detection the paper reports a precision of 93.47% meaning 93.47% of sms classified as spam were indeed spam with a recall of 94.69% it accurately identified 94.69% of all spam messages in the test set the f1 score which balances precision and recall is demonstrating the robust presentation of the naive bayes algorithm strating the robust performance of the naive bayes algorithm.

Anikait Kapoor et al.,[5] focuses on utilizing supervised learning and nlp techniques for classifying sms messages they employ 10-fold cross-validation to evaluate model performance and achieve a notable decrease in error rates compared to existing benchmarks these findings underscore the effectiveness of their approach their primary classifier significantly reduces overall error rates showcasing more than a 50 improvement over previous state of the art results highlighting its efficiency in accurate.

Suparna Das Gupta et al., [6] collected a dataset from the Kaggle Repository and thoroughly cleaned it by removing white spaces, standardizing text, eliminating punctuation, tokenizing messages, and stemming words. They then gen- erated testing and training datasets from the cleansed data. Using TF-IDF vectorization, they created word vectors for feature generation. These vectors were used to classify messages as spam or ham. Finally, they tested their model by predicting whether input messages were SPAM or HAM.

Ghourabi et al.,[7] introduced the nlp lstm model which incorporates a long short term memory layer with the convolutional neural network cnn layer the CNN model was weighed by the researchers through comparison with nine traditional machine learning techniques in- cluding the standalone cnn and lstm models their experimental findings showed that the CNN lstm model accurate advanced than any other strategy with a f1 score of 0.91% and a precision of 98.3%.

Shaik Mohammed Imran et al.,[8] proposed the serious problem that the growth of spam com- munications in short message services poses for spam identification classic machine learning methods like logistic regression have been commonly utilized these models have per- formed exceptionally well overall on the dataset the twitter dataset and generated extraordinary f1-scores when applied to the texts spam col- lecting dataset these results show how well the modified transformer model handles sms spam detection

Sheikhi et al.,[9] analysis of the algorithmic strategies for characterizing spam sms gru with lstm models outperform more established neu- ral network models like svm and nb the lstm method in particular earned the greatest preci- sion of 9818 and the highest spam catch rate of
90%to 96% these models were created with cer- tain hyperparameters and trained across ten epochs using the adam optimizer they made use of nlp methods for preliminary processing.

3.Proposed methodology

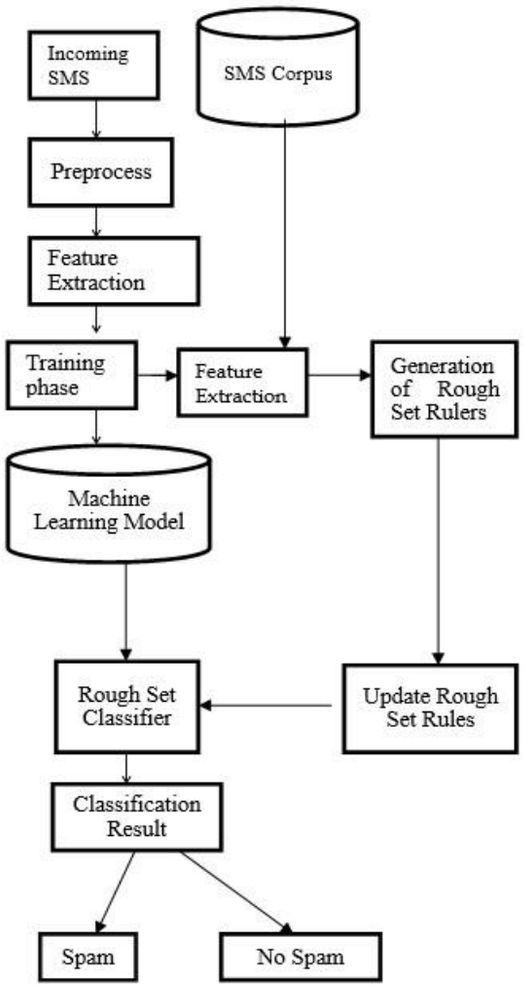

Fig 3.1 Flowchart of proposed methodology

When an SMS arrives the system first cleans and normalizes the text key attributes like key-words and metadata are then extracted the at-tributes are used to train a machine learning model with a big collection of sms to learn the difference between spam and ham the trained model uses these attributes to make predictions to improve accuracy a rough set classifier is used which handles uncertainty and enhances decision-making during training rough set rules are created to update the classifier these rules are continuously refined to adapt to new types of spam ultimately the classifier uses the up-dated rules to label the email as spam or non- spam.

Preprocess:

Here the data demonstrates the preprocessing steps like converting to lowercase, removing punctuation, tokenizing, removing stop words, lemmatizing, and then vectorizing the text into a TF-IDF matrix. The resulting matrix can then be used as input for the Naive Bayes algorithm to train a spam detection model.

Feature Extraction:

In this research the feature extracted is the most significant data from the SMS messages is ex- tracted by this module, including the frequency of certain words and their probability distribu- tions. This module is essential for giving the method known as Naive Bayes the pertinent data it needs to properly analyze the incoming messages.

Training Phase:
Using the features extract from the sms corpus, a machine learning model is trained. This model learns to recognize patterns and features that distinguish spam from non-spam emails. Classification Result: The classifier outputs the result, indicating whether the email is spam or non-spam.

Spam/Non-Spam: Finally, the sms is classified into either the spam or no spam category focused on the classification result.

Word Embeddings: Use pre-trained embed- dings like Word2Vec, Fast Text to capture semantic meaning.

```
from sklearn. feature_extraction.text import
 Count Vectorizer
 cv=
 CountVectorizer(max_features=3500)
 X=cv.fit_transform(corpus).toarray()
 y= pd.get_dummies(message[‘LABEL’])
 y=y.iloc[:,1].values
```

4. Result and Discussion

The CNN model in python the project demonstrate the strength of the CNN lstm model achieving an accuracy of 98.37% accuracy of 95.39% recall of 87 f1-score of 91.48% and an impressive area under the curve of 93.7% these findings show that CNN lstm exemplar surpasses alternative machine learning algorithm in reliably categorizing sms spam this technological breakthrough greatly enhances mobile security by effectively identifying and filtering out unwanted and potentially harmful messages thus reducing the risks associated with smishing attacks across various mobile platforms

The Accuracy formula referred to calculation which is used to determine the accuracy of classification model. Accuracy helps in model evaluation Comparison Results and decision making

The formulae used for Accuracy is

$$\text{Accuracy} = \frac{(TP+TN}{FP+FNTP+TN)} *100$$

Where TP=True Positive, TN=True Negative FP= False Positive, FN=False Negative

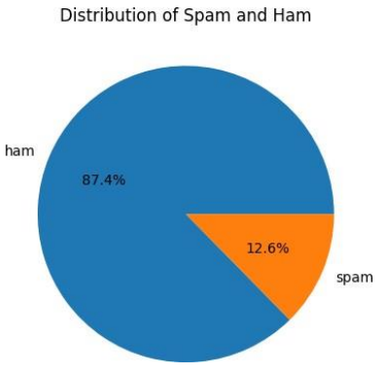

Fig 4.1 Graph is based on distribution of spam and ham

5. Conclusion

The generating a mixture model for sorting spam sms is described which relies on cnn and lstm to address various sms contexts such as mobile network messages to create realistic assessment dataset a collection of communications in Arabic and English was gathered and analyzed by leveraging the strengths of various models and continuously refining detection techniques it is possible to significantly reduce the prevalence of spam and improve the overall user experience in digital communication it used to identify sms spam the experimentalevaluation of the proposed method revealed that the CNN lstm model outperformed other techniques in sms spam classification based on the results our CNN lstm model achieved an accuracy of 98.37% a precision of 95.39% a recall of 87% an f1-score of 91.48% and an overall area under the curve of 93.7% this technology can significantly enhance mobile phone security filtering spam messages and reducing the risks asso- ciated with smishing attacks in mobile environments.

Reference

1 Abhishek Patel, Priya Jhariya, Sudalagunta Bharath, Ankita, “SMS Spam Detection using Machine Learning Approach”, IJCRT, 9, 4, April 2021.

2 Dhananjay Bhagat, Pranali Dhawas, Saicharan Kotichintala, Rohit Patra, Raj Sonarghare, “SMS Spam Detection Web Application Using Naive Bayes Algorithm & Stream lit”, International Journal of Current Science, 13, 2023.

3 Houshmand Shirani Mehr, Mehran Sahami, CS229 Machine Learning Course Reports, Stanford University, 8, 2013.

4 S.N. Sneha, N. Ganesan, Dr .C. Thiyagarajan, Vaideghy, “A Study About Message spam Detection Using Machine Learning Algorithm", International Research Journal of Modernization in Engineering Technology and Science, 05, April2023.

5 Anikait Kapoor, Debavushan Saikia, Ishaan Dhawan, “SMS Spam Detection using Machine learning Approach”, Journal of Research in Science and technology, 14, pp 10-17, Febrauary2024.

6 Suparna Das Gupta, "SMS Spam Detection Using Machine Learning", Journal of Physics”, 10, pp 1-8, 2021.

7 Abdallah Ghourabi, Mahmood Qusay M. Alzubi, “A Hybrid CNN LSTM Model for SMS Spam Detection in Arabic and English Messages” ,12, 2020.

8 Shaik Mohammad, Nythani Harshitha, Mohammad Yasmeen, Mittapelly Ruthika International Journal of Gender, Science and Technology, Vo1 8, pp 222-228, 2022.

9 S. Sheikhi, M. T. Kheirabadi, A. Bazzazi, “An Effective Model for SMS Spam Detection Using Content based Features and Averaged Neural Network”, International Journal of Engineering, vol. 33, 2020