WWW.IJRPR.COM

International Journal of Research Publication and Reviews

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

Machine Learning and Deep Learning Techniques for SMS Spam Detection, Accuracy Check and Comparative Study

Dr. Sarika Raga[1], Mrs Chaitra B L[2]

[1]Program Coordinator and Associate Professor, Department of Electronics and Communication Engineering*Visvesvaraya Technological University, PG Centre, Bangalore Region, Muddenahalli, Chickballapur-562101. India

[2]II year M.tech student Department of Electronics and Communication Engineering*Visvesvaraya Technological University, PG Centre, Bangalore Region, Muddenahalli, Chickballapur-562101. India

A B S T R A C T

The number of people using mobile devices increasing drastically every year. SMS (short message service) is a text message service available in smartphones as well as keypad phones. So, the traffic of SMS increased drastically. The spam messages also increased. The spammers try to send spam messages for their financial or business benefits like market growth, to obtain personnel information, credit card information,Bank transactions, etc. So, spam classification has special attention and its very important to classify HAM and SPAM for the benefit of every individual in the society. In this paper, we used different machine learning techniques and deep learning technique for SMS spam detection. we used a dataset from UCI repository and build a spam detection model. Our experimental results have shown that our LSTM model outperforms previous models in spam detection with an accuracy of 95% to 98.5%. We used python for all implementations.

Keywords:Machine learning, Deep learning,LSTM, Accuracy, HAM and SPAM.

1.Introduction

In 2022 there are 5.31 billion mobile users in the world. The top three countries using more mobiles are China, India, US. Short Message Service or SMS is a text messaging service available for the last several years. SMS service can be availed without internet also. So, SMS service is available in smartphones and basic mobiles also.

A spammer is a person/company which is responsible for unsolicited messages. For their organization benefits or personal benefits, spammers sending a vast number of messages to the users. These messages are called spam messages. These SPAM messages may cause a huge loss to individuals , as their personnel information and financial information may be accessed using these spam messages. So it is very important to classify Which message is HAM and which message is SPAM. we used various machine learning and deep learning techniques for SMS spam detection.
Machine Learning is a technology, where machines learn from previous data and made a prediction on future data. Nowadays, machine learning and deep learning can be applied to solve most of the real world problems in all sectors like health, security, market analysis, etc.

1.1 Literature survey

Applying ML and DL techniques for spam detection is not a new era. Previously, various researchers applied ML techniques for classification SMS spam. Nilam Nur Amir Sjarif[3] et.al applied the TF-IDF technique in combination with a random forest classifier and achieved an accuracy of 97.5%.TF-IDF is a method used to quantify the words in a document by using two measures Term Frequency and Inverse Document Frequency.A.Lakshmanarao[4] et.al applied four machine learning classifiers Decision Trees, Naive Bayes, Logistic Regression, Random Forest for email spam filtering, and achieved an accuracy of 97% with random forest classifier. PavasNavaney[5] et.al proposed various machine learning algorithms and achieved an accuracy of 97.4%

. Tel.: +91 9964700645,+91 7019584115
E-mail address:chaitrabadum@gmail.com

with support vector machines. Luo GuangJun [6] et.al applied various shallow machine learning algorithms and achieved a good accuracy rate with logistic regression classifier. Tian Xia[7] et.al proposed the Hidden Markov Model for the detection of SMS spam. Their model used the information about the order of words thereby solving issues with low term frequency. They achieved an accuracy of 98% with their proposed HMM model. M. Nivaashini [8] et.al applied a deep neural network for SMS spam detection and achieved an accuracy of 98%. They also compared DNN performance with NB, Random Forest, SVM, and KNN. Mehul Gupta[9] et.al compared various spam detection machine learning models with deep learning models and shown that deep learning models achieved a high accuracy rate in SMS spam detection. Gomatham Sai Sravya[10] et.al compared various machine learning algorithms for SMS spam detection and achieved the best accuracy with the Naive Bayes classification model. M.RubinJulis[11] et.al applied various machine learning classifiers and achieved an accuracy of 97% with a support vector machine. K. Sree Ram Murthy [12] et.al proposed Recurrent Neural Networks for SMS spam detection and achieved a good accuracy rate. S. Sheikh[13] proposed SMS spam detection using feature selection and the Neural Network model and achieved a good accuracy rate. AdemTekerek[14] et.al applied various machine learning classification models for SMS spam detection and achieved an accuracy of 97% with a support vector machine classifier.

1.2 Methodology

First, we collected an SMS spam dataset from the UCI ML repository [15]. Later, we applied different text preprocessing techniques to clean the dataset. Then we applied, we applied various machine learning algorithms and LSTM

1.2.1 Data Collection : In the current research, we have used a dataset consists of 5,574 text messages classified as spam and ham (legitimate), which is publicly available in the UCI machine learning repository . 747 messages are spam while 4,827 of them were labelled as ham. The format of the dataset is text where each line describes a message that has two parts, the message string, and its category.

1.2.2 Feature Extraction and Selection : The feature extraction phase is a very important phase and its a critical task in detection of spam messages as the choice of features can significantly affect the performance of machine learning techniques. So, in most cases, it is a challenging and difficult task to discover the most useful features that can efficiently classify SMS spam messages. Hence, features with the best correlation should be selected to improve the detection rate and produce a shorter process time . In the feature extraction phase, the C# .net framework is used to read lines from the dataset and to extract features from text messages ,emails and save them in a new format. Further, in order to find the most efficient features for SMS spam detection, we have investigated different characteristics of spam messages and selected features which are essential and helpful for identification of these kind of messages in our dataset.

1.2.3 Text preprocessing

Text preprocessing is a method to clean the text data and prerpare it to feed data to the model. Text data contains noise in various forms like emotions, punctuation, text in a different case like upper and lower case.It helps to get rid of unuseful parts of the data, or noise, by converting all characters to lowercase, removing punctuations marks, and removing stop words and typos. Removing noise comes in handy when you want to do text analysis on bundle of data like comments or tweets.

Step 1: Data Preprocessing. Tokenization — convert sentences to words.

Step 2: Feature Extraction.

Step 3: Choosing ML Algorithms.

In natural language processing, text preprocessing means cleaning and preparing text data. NLTK and r are generally used Python libraries used for handling many text preprocessing tasks.

Tokenization is a special technique of tokenizing or splitting a string, text into a list of tokens. One can think of token as small parts like a word is a token in a sentence, and a sentence is a token in a paragraph.Tokenization is one of the very common tasks when it comes to working with text data. ... Tokenization is essentially splitting a phrase, sentence, paragraph, or an entire text document into smaller units, such as individual words or terms. These smaller units are called tokens.

1.2.4 ALGORITHMS USED:

1.2.4.1 Naive Bayes (NB) :

It is one classification technique based on 'Bayes' theorem that assumes independence among predictors. This classifier of Bayesian will tell that there is no relationship between presence of a specific feature in one class and the existence of any other features. If there are dependencies between the existence of the feature with each other, this classifier will treat all properties as independent that is important to determine the probability score. The Naive Bayes classifier is still simple and tough in the case of the dimensionality of desired input is high. Multinomial Naive Bayes (MNB) is a new advanced version of Naive Bayes classifier. The basic improvement is the presence of independency among document class and length. this type classifier includes multinomial distribution that works well for data type that is countable like the words inside a text or document. So, the classifier of NB is a conditional independency between each the feature in the model, while classifier of MNB is a special case of a Naive Bayes algorithm that utilize a multinomial distribution for each and every feature. In our reviewing process, we found that using Naive Bayes algorithm in SMS spam filtering the primary and most prominent technique used by researchers .

It is a classification technique based on theorem of 'Bayes' that assumes there is no dependance among predictors. This classifier of Bayesian illustrates that there is no relationship between presence of a specific feature in one class and the existence of any other features. Even if there are dependencies between the existence of the feature with each other, this classifier will treat all desired properties as independent that contribute in the probability score.

The Naive Bayes classifier is still simple and tough in the case of the dimensionality of desired input is high. Multinomial Naive Bayes (MNB) is a very new advanced version of Naive Bayes classifier. The basic advancement is the presence of independency among document class and length. this classifier includes multinomial distribution that works well for data type that is countable like the words inside a text or document. So, with classifier of NB is a conditional independency between each of the feature in the model, while classifier of MNB is a particular case of a Naive Bayes algorithm that use a multinomial distribution for each and every feature.

1.2.4.2 Decision Tree

DT is an algorithm based on supervised machine learning that is usually used for tasks related to classification. This algorithm works with both continuous and categorical variables. Initially, the algorithm will split the population into many homogeneous groups that is done according to the basis of independent variables or significant attributes. As decision tree is non-parametric, the requirement for examining existence of outlier or separation data linearity is not essentially needed.

1.2.4.3.Random Forest

It refers for a crew of Decision Trees. Meaning that, the RF classifier is a crew learning mode involving set of decision trees. This classifier works as vote for specific class by each tree to classify a new object. The class that have the greatest votes number will deciding the label for classification.

1.2.4.4 Support Vector Machine

It is a characteristic classifier that is widely utilized for the task of classification. The algorithm of SVM plots all items in n-dimensional space as a point supposing each feature value as a specific coordinate value. Then it consists a line which divides the full data into two variously data groups. The adjacent points in these groups will be the furthest from the dividing line.

1.2.4.5 LSTM

LSTM networks are best suited to classifying, processing and making predictions based on time series data, because there can be lags of unknown duration between important events in a time series. LSTMs were developed to solve vanishing gradient problem that can be encountered when training traditional RNNs. Long Short Term Memory Network is an advanced RNN, a sequential network, that allows information to persist. It can handle the vanishing gradient problem faced by RNN. A recurrent neural network is also known as RNN is used for persistent memory.

1.2.5 Experimentation and results

Evaluation parameters:

Different parameters like precision, recall and accuracy are tested for all machine learning and deep learning algorithms.
Recall: Ratio of correct results obtained to desired no of correct results
Precision: Ratio of correct results obtained to total results
Accuracy: Closeness of calculated value to actual value

Table 1:Comparitive results of machine learning and deep learning algorithms

Algorithm	Accuracy	Precision	Recall
NB	67%	50%	73%
DT	94%	82%	84%
RF	95%	83%	85%
SVM	97%	99%	83%
LSTM	98.5%	99%	89%

Result:

In Our paper we mainly concentrated on discussing and evaluating machine learning and deep learning techniques for spam SMS detection. We ran out comparisons among four different ML classifiers and one Deep learning technique. The results obtained from our evaluation of the classifiers show that LSTM Classifier achieves the highest accuracy of around 98% for dataset. Although RNN has been generally used in the classifications of text related data, it has shown significant improvement against the traditional classifiers of machine learning and achieves the highest accuracy among them for textual data as well. This achievement of LSTM has wide opened the research aspect of its application over text related classification problems which involve review classification and sentiment prediction. The machine learning algorithms , Support vector machine , Naive Bias,Random Forest and Decision Tree show good results, but LSTM outperforms all these for the dataset.

In our paper different parameters like precision, recall ,fscore, support, misclassification, false positives, specificity and accuracy are also tested for all machine learning and deep learning algorithms. The best results are obtained in LSTM and the comparative results are tabulated.

Significant results have been obtained from this work, corresponding to which this research can be taken to real world application level for detection of spam SMS.

Acknowledgements

This Acknowledgement is to express my deep gratitude and respect to all the people who filled me with strength and confident that helped me to complete the task effectively

I like to express my gratitude to Dr. Sarika Raga. Head of The Department for helping us providing us with adequate facilities, ways and means by which we were able to complete this paper.

I express my immense pleasure and thanks to all the teachers and staff of the Department of DECS, VTU for their cooperation and support.

Last but not least, I thank to all others, and especially our classmates and our family members who helped me in one way or another in the successful completion of this work.

REFERENCES

K. Yadav, S. K. Saha, P. Kumaraguru, and R. Kumra, "Take control of your smses: Designing an usable spam sms filtering system," in 2012 IEEE 13th International Conference on Mobile Data Management. IEEE, 2012, pp. 352–355.

S. J. Warade, P. A. Tijare, and S. N. Sawalkar, "A novel approach for detection of sms spam."

A. Narayan and P. Saxena, "The curse of 140 characters: evaluating efficacy of sms spam detection on android," in Proceedings of the Third ACM workshop on Security and privacy and protection in smartphones & mobile devices. ACM, 2013, pp. 33–42.

A. S. Onashoga, O. O. Abayomi-Alli, A. S. Sodiya, and D. A. Ojo, "An adaptive collective and collaborative server side sms spam filtering scheme using artificial immune system," Information Security Journal: A Global Perspective, vol. 24, no. 4-6, pp. 133–145, 2015.

J. W. Yoon, H. Kim, and J. H. Huh, "Hybrid spam filtering technique for mobile communication," computers & security, vol. 29, no. 4, pp. 446–459, 2010.

S. J. Delany, M. Buckley, and D. Greene, "Sms spam filtering: methods and text data," Expert Systems with Applications, vol. 39, no. 10, pp. 9899–9908, 2012.

Q. Xu, E. W. Xiang, Q. Yang, J. Du, and J. Zhong, "Sms spam detection using non content features of the message ," IEEE Intelligent Systems, vol. 27, no. 6, pp. 44–51, 2012.

G. V. Cormack, J. M. G. Hidalgo, and E. P. S´anz, "Feature engineering for mobile messages and email spam filtering," in Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval. ACM, 2007, pp. 871–872.

S. Keele, "Important Instructions and Guidelines for performing systematic literature reviews in software engineering," in Technical report, Ver. 2.3 EBSE Technical Report. EBSE, 2007.

T. M. Mahmoud and A. M. Mahfouz, "Sms spam filtering methods and technologies depending on artificial immune system," IJCSI International Journal of Computer Science Issues, vol. 9, no. 1, pp. 589–597, 2012.