SMS SPAM AND HAM DETECTION USING NAÏVE BAYES ALGORITM

Maram Sai Charan Reddy
Student, CSED
Kakatiya Institute of Technology and science
(Autonomous under KU)
maramsaicharan2000@gmail.com

ABSTRACT:

ML is termed as Machine Learning; it is the study of computer algorithms which automatically improves through experience. The usage of mobile phones is growing popular in our everyday life. Short Message Service are viewed as most generally applied correspondence administration which is less costly. In any case, this has prompted an increment in cell phones assaults like SMS Spam. Here, Naive bayes algorithm is used in order to differentiate between spam and ham SMS. Spam is the unnecessary fraud messages received whereas ham is legitimate message. The algorithm used here is machine learning classification algorithm, and it is implemented here and can be used in differentiating between spam and ham messages with the help of SMS spam collection data set provided. We train the machine by providing that data set such that it learns from that data and will be able to draw conclusions on its own. Now a days it is very much crucial to identify the spam messages to reduce many frauds happening around the globe. This algorithm can classify with an accuracy of 98.13%.

Keywords: Machine Learning, SMS Spam, Naïve Bayes Classifier, text classification, Python.

1.INTRODUCTION

Short Message Service is measured as most extensively used message facility. It is a technique of sending short text messages from one device to another. The usage of mobiles is growing everyday as they deliver a huge variation of facilities by dropping the rate of amenities. Due to abundant usage of these services, it has led to growth in mobile device outbreaks like SMS junk. Generally, the word

spam refers to the message which is unsolicited. Simply we can state that spam is a junk text message sent from one device to another in the SMS text format. These spam messages can cause threat to personal data stored in the device.

By the enormous growth in population and increase of all these technological aspects have been growing extremely which in turn increases unanimous spreading of such threat due to less effective security control measure and in order to solve such problems many researchers have developed many techniques to solve and protect the devices from such threats in many different ways. The main motive is to provide privacy, convenience and harmony. The classifier used to build these models are Support Vector Machine and Naïve Bayes Classifier, these the two mostly used traditional classifiers. And by convolutional Neural Networks also we can achieve the required model.

Essentially while training the model firstly, we want to consider a data set, here SMS Spam Collection dataset is used and it is divided into train and test data set and the Naïve Bayes Classifier is used in training and evaluation of the model.

2.LITERATURE SURVEY

Ms.D. Karthika, Dr.T. Hamsapriya, Mr. M. Raja, and Ms.P.Lakshmi Surya utilized various AI techniques to direct a relative report over junk identification dependent on supervised Machine learning. The correlation was done utilizing three separate AI characterization calculations, in particular, the Naive Bayes and Multilayer Perceptron (MLP) classifiers. MLP's exactness was acceptable, yet it took a long effort to finish. Albeit the precision of Naive Bayes classifier was less compared to MLP, it is speedy as far as execution and learning. Utilizing FBL include determination and separated Naive Bayes with Bayesian learning, the precision of the Naive Bayes classifier was improved. The refreshed Naive Bayes model had a 91 percent exactness rate.

Harisinghaney, Aman Dixit, Saurabh Gupta (2014) played out a close to assessment on manuscript besides pictures via using KNN, Naïve Bayes Algorithm for electronic mail junk area. The assessment broadside projected a framework aimed at recognizing transcript and junk messages. These people used K-NN, Naïve Bayes and a changed Reverse DBSCAN calculations. Makers used Enron dataset meant for transcript and picture junk plan. They used Google's public library, Tesseract for eliminating disagreements as of pictures. Results show that these three AI estimations provides good results deprived of pre-preparing between which Naïve Bayes computation is significantly exact than different computations.

In the creating time of the Internet, people are including progressively in free online administrations. People will in general share their information on various locales, however that information is bestowed to various associations that spam people to offer their administrations. In today’s world where enormous piece of correspondence happens inside the sort of SMS or messages.

Notwithstanding, on account of publicizing offices and interpersonal interaction sites the vast majority of the SMS coursed contain undesirable information that isn't pertinent to the client. Spam SMS or messages square measure a sort of email correspondence any place the client gets unsought messages by means of email. Spam SMS cause burden and misfortune to the beneficiaries consequently there's a need to channel them and separate them from the real SMS. a few calculations and channels are created to notice the spam messages anyway spammers ceaselessly advance and sophisticate their spamming procedures on account of that the common channels have gotten less viable. the strategy projected during this paper includes making a spam channel exploitation paired and constant probability conveyances. The short electronic informing administration spam is realized the junk or unsolicited messages that we get on mobiles. These SMS spams address an authentic irritation to the versatile supporters. In the paper, they will in general blessing an audit of the by and by possible methodologies, difficulties and future examination bearings on spam identification strategies, sifting and alleviation of versatile SMS spams.

3.ALGORITHM

With this drastically increasing number of mobile phone users the easy and common way of information transfer platform i.e., SMS services are in wide range of usage. SMS services are cheap and best in comparison with other communication services. But with the increase of SMS services, it has become easy for the cyber attackers easy to prey the people by using this service just by sending unsolicited or junk messages.

So, in order to protect people from such threats many solutions are brought up which will classify and let the people know the message is a junk or ham. The proficient classification algorithm which applied here is Naïve Bayes Classifier. Although there are many other algorithms NB classifier always proved its efficiency with its results. The flow chart of NB classifier is as shown in fig.3.

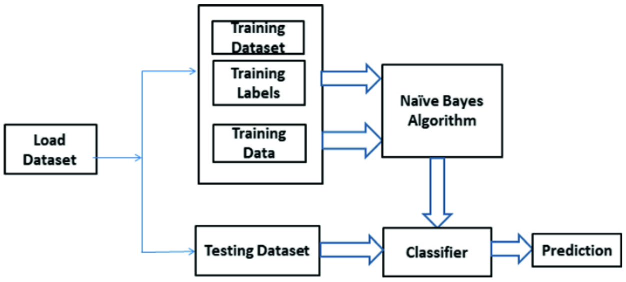


Fig.3. Flowchart of Naïve Bayes

3.1 NAÏVE BAYES CLASSIFIER

It is an algorithm of Machine Learning which comes under classification technique. It's a Supervised technique that uses the Bayes theorem of probability to predict the class of enigmatic datasets. In simple terms, a Naive Bayes algorithm anticipates that the presence of one item in a class is unrelated to the presence of another. Naive Bayes model is not difficult to assemble and especially valuable for enormous informational indexes. It works on the probability principle called and bayes theorem which is shown in fig.3.1.

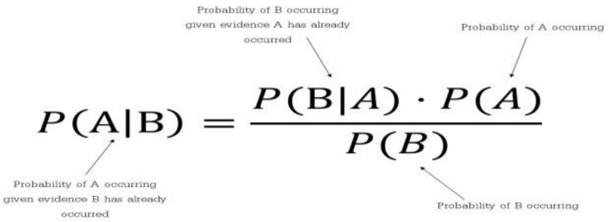
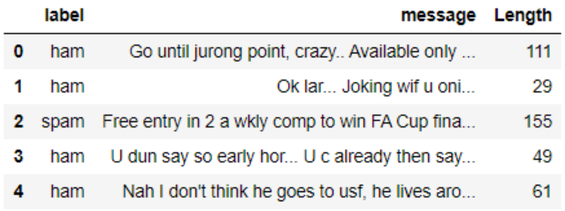

Probability of B occurring given evidence A has already occurred
Probability of A occurring
P(A|B) = P(B|A) · P(A) / P(B)
Probability of A occurring given evidence B has already occurred
Probability of B occurring

Fig.3.1. Bayes Theorem

A high-level form of NB is Multinomial Naive Bayes (MNB). The principal improvement is the autonomy between report length and class. In basic terms, NB classifier includes contingent autonomy of every one of the highlights in the model, though MNB classifier is an extraordinary instance of a NB classifier which utilizes a multinomial conveyance for each element.

3.2 Dataset

In order to train the supervised machine learning model the prior thing to do is collect the dataset. Here the data set considered is smsspamcollection which is taken from UCI machine learning repository. It contains the collected sms with label whether it is spam or ham message. The glimpse of data is as shown in the fig.3.2.



	label	message	Length
0	ham	Go until jurong point, crazy.. Available only ...	111
1	ham	Ok lar... Joking wif u oni...	29
2	spam	Free entry in 2 a wkly comp to win FA Cup fina...	155
3	ham	U dun say so early hor... U c already then say...	49
4	ham	Nah I don't think he goes to usf, he lives aro...	61

Fig.3.2. head of dataset

4. EXPERIMENTS

In the process of building the spam detection model the following operations are conducted to get the evaluation of model done.

4.1 Data pre-processing

Information pre-handling is the technique for getting ready crude information for use in an AI model. It's the first and most significant advance in building an AI model. Genuine information likewise incorporates commotion, missing qualities, and is in an unusable organization that can't be utilized straightforwardly in AI models. Information pre-handling is a fundamental movement for cleaning information and making it appropriate for a machine learning model model, which improves the model's exactness and execution.

It involves cleaning the text by removing unwanted or redundant data by using stop words and noisy data is removed by stemming and lemmatization. Stemming removes the last characters of the words which are leading to the wrong meaning. And lemmatization coverts the words into meaningful base form.

4.2 Vectorization

Here the categorical or text data is converted into integer format. It selects the text from the data set and choose the words which makes the meanings as the features and with the help of these features whole categorical data is converted into integer format. After using CountVectorizer which is imported from sklearn package the data set will become as 0, 1 in matrix format.

4.3 Training and Testing

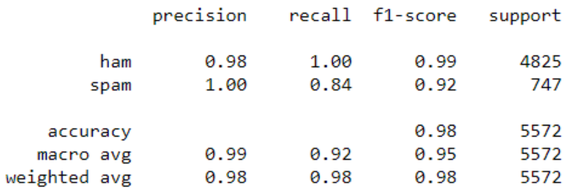
Principally the informational index is separated into two sections i.e., test informational collection and train informational index. Preparing information is utilized to fabricate the AI model and afterward we test it with test informational index to check its exactness and accuracy and numerous different components.

Here Multinomial Naive Bayes used in training the model. Multinomial NB calculation is probabilistic learning technique which is for the most part utilized in Natural Language Processing. The calculation depends on the Bayes hypothesis and predicts the tag of a book like a piece of email or paper article. It figures the likelihood of each tag for a given example and afterward gives the tag with the most noteworthy likelihood as yield. Naive Bayes classifier is an assortment of numerous calculations where every one of the calculations share one basic guideline, and that is each component being arranged isn't identified with some other element. The presence or nonattendance of an element doesn't influence the presence or nonappearance of the other element.

In process of evaluation of result generate confusion matrix as the output which method is imported from sklearn. It is used to calculate accuracy, sensitivity, precision, recall etc.

5. CONCLUSION

In this Project, Naïve Bayes Algorithm analyses based on the factors like precision, recall, f1-score, support. Naive Bayes order calculation is viably valuable for managing clear cut information characterization. The basic hypothesis it utilizes is the Bayes contingent probabilistic model for tracking down a back likelihood given certain conditions. It is classified "Credulous" on the grounds that under the presumption that all highlights (assortments of words) in the dataset are similarly significant and free. Utilizing the Naïve Bayes grouping calculation, the venture got over 98% exactness in foreseeing a spam message dependent on the words it contains which is determined from the confusion matrix which we got as yield.



```
              precision    recall  f1-score   support

         ham       0.98      1.00      0.99      4825
        spam       1.00      0.84      0.92       747

    accuracy                           0.98      5572
   macro avg       0.99      0.92      0.95      5572
weighted avg       0.98      0.98      0.98      5572
```

Fig.5. Result of NB classifier

6. FUTURE ENHANCEMENT

To perform prediction in this project, we only used static data for training and testing the algorithms. To improve the accuracy of the forecasts in the future, the project will need to increase the amount of data in the data collection. Which in turn helps the model to improve its accuracy of prediction.

7. REFERENCES

[1] "Intelligent spam classification for mobile text messages," by K. Mathew and B. Isaac. The 2011 International Conference on Computer Science and Network Technology has published its proceedings.

[2] Ensemble-based classifiers. Rokach L (2010) CrossRefGoogle Scholar

[3] P. Navaney, G. Dubey, and A. Rana (2018). "Supervised Machine Learning Algorithms for SMS Spam Filtering."

[4] “Content-based SMS spam filtering,” in Proceedings of the 2006 ACM Symposium on Document Engineering, October 10-13, 2006, Amsterdam, The Netherlands. “Content based SMS spam filtering,” in Proc. of the 2006 ACM Symposium on Document Engineering, Amsterdam, The Netherlands, October 10-13, 2006. J. M. G. Hidalgo et al., “Content based SMS spam filtering,” in Proc. of the 2006 ACM Symposium on Document Engineering, Amsterdam, The Netherlands, October 10-13, 2006.

[5] K. Nigam and A. McCallum. In 1998, he presented “At the AAAI-98 Workshop on Learning for Text Categorization, he presented “A comparison of event models for nave bayes text classification.”

[6] J. Goodman and W. Tau Yih. Training for discriminative spam filters is available online. The Third International Conference on Email and Anti-Spam is a gathering of experts in the field of email and anti-spam (Mountain View,CA, 2006).

[7] Spam filtering based on SMS content by J. M. G. Hidalgo, G. C. Bringas, E. P. Sanz, and F. C. Garcia. 107–114 in DocEng '06: Proceedings of the 2006 ACMSymposium on Document Engineering, ACM Press, New York, NY, USA, 2006.

[8] (2017): 63-70. "Spam Detection on Social Media Text." G. Jain and B. Manisha Jain.

[9] Shafi'I Muhammad Abdulhamid, “A Review on Mobile SMS Spam Filtering Techniques,” IEEE Access, 2017.

[10] N. K. Nagwani (2017). Using a Bi-Level Text Classification Approach, SMS Spam Filtering and Priority Message Identification The International Arab Journal of Information Technology (IAJIT) has published its 14th issue (4).

[11] T. K. Gannavaram V, R. Bejgam, S. B. Keshipeddi, S. Sunkari and V. K. Aluvala, "Conversion of Sound Energy into Electrical Energy in Highly Populated Areas," 2021 6th International Conference on Communication and Electronics Systems (ICCES), 2021, pp. 32-36, doi: 10.1109/ICCES51350.2021.9489219.

[12] T. K. Gannavaram V and R. Bejgam, "Brief Study and Review on the Next Revolutionary Autonomous Vehicle Technology," 2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), 2021, pp. 34-37, doi: 10.1109/ICACITE51222.2021.9404763.

[13] T. K. Gannavaram V, R. Bejgam, S. B. Keshipeddi, A. Banda and G. Bollu, "Study of Automobile Safety Technology Development using Vehicular Safety Device (VSD)," 2021 6th International Conference on Inventive Computation Technologies (ICICT), 2021, pp. 240-244, doi: 10.1109/ICICT50816.2021.9358670.

[14] T. K. Gannavaram V, U. Maheshwar Kandhikonda, R. Bejgam, S. B. Keshipeddi and S. Sunkari, "A Brief Review on Internet of Things (IoT)," 2021 International Conference on Computer Communication and Informatics (ICCCI), 2021, pp. 1-6, doi: 10.1109/ICCCI50826.2021.9457009.

[15] T. K. Gannavaram V, U. M. Kandhikonda, R. Bejgam, S. B. Keshipeddi and S. Sunkari, "A Brief Review on Internet of Things (IoT)," 2021

International Conference on Computer Communication and Informatics (ICCCI), 2021, pp. 1-6, doi: 10.1109/ICCCI50826.2021.9451163.

[16] Gannavaram V, Tulasi Krishna and Reddy, Ganta Raghotham, IoT Based Electricity Energy Meter (June 28, 2021). Available at SSRN: https://ssrn.com/abstract=3875420 or http://dx.doi.org/10.2139/ssrn.3875420

[17] Tummanapally, Shraddha Shree and Sunkari, Saideep, Smart Vehicle Tracking System using GPS and GSM Technologies (July 12, 2021). Available at SSRN: https://ssrn.com/abstract=3884903 or http://dx.doi.org/10.2139/ssrn.3884903

[18] Tummanapally, Shraddha Shree and Sunkari, Saideep, Traffic Data Collection and Analysis based on Wireless Sensor Network (July 12, 2021). Available at SSRN: https://ssrn.com/abstract=3885102 or http://dx.doi.org/10.2139/ssrn.3885102