ISSN No:-2456-2165

Message Spam Identification by Naive Bayes Classifier Algorithm using Machine Learning

Lokam. Devi Naga Srinu[1]
Meesala. Dhanush Kumar[2]
Mulaparthi. Mani Gopal[3]
Swarnandhra College of Engineering and Technology

Abstract:- With the spread of modern life, messaging has become one of the most important forms of communication. SMS (Short Message Service) is a text messaging service available on all smart phones and mobiles. Facebook, WhatsApp etc. Unlike other chat-based communication applications, SMS does not require any internet connection. SMS traffic has increased significantly and spam has also been increased rapidly. Hackers and spammers are trying to scam over devices through SMSs. As a result, SMS support for mobile devices becomes difficult. Spammers may ask for business expansion, lottery information, credit card information, etc. They also try to send spam emails to obtain financial or commercial benefits such as: attackers attempt to disrupt the system by sending spam links that, when clicked, allow them to control mobile devices. To analyze this communication, the authors developed a system that can analyze malicious messages and determine whether they are RAW or SPAM. Here, we use text classification methods such as Naive Bayes classifier algorithm to classify the texts and determine the message whether it is spam or not.

Keywords:- *Machine Learning, Language Processing, Spam, Ham, SMS, Naive Bayes, Logistic Regression.*

I. INTRODUCTION

Short text messages (SMS) are more than just a conversation. SMS, which was first defined as part of the GSM family of standards in 1985, is a method of sending messages of up to 160 characters to GSM mobile phones. SMS technology is based on international mobile communications standards and is recognized worldwide. Spam is the misuse of email to send illegal, unsolicited messages. While the most common type of spams are SMS spam, the word is also used for similar abuse in other media and news. Spam messages are unsolicited messages that resemble spam, often for commercial benefit. Spam emails are used to market and send phishing links. Commercial spammers use malware to send spam because spam is illegal in many countries. Sending spam from an infected computer reduces the spammer's risk by masking the source of the spam. Text can contain limited characters, including letters, numbers, and some symbols.

Check out the news to see the full model. Nearly all spam emails ask users to call a number, respond to a text message, or visit a URL. This pattern can be seen in the results of a simple SQL query for spam emails. The low cost and high bandwidth of SMS networks attract a large number of spam messages. Every time a spam message arrives in the user's inbox, a notification is sent to the user's phone. Users get angry when they see spam, malicious emails taking up space on their phone storage. The aim of this project is to apply different learning machines to the spam message classification problem, to get an idea and learn more about the problem by comparing their performances and to create an application as one of the algorithms that can be effectively filtered. The truth is that it is spam. Complete the extraction and initial analysis results in MATLAB and then use the scikitlearn library function to implement various machine learning algorithms in python.

Spam and not spam: Spam refers to the content of email and the use of electronic communications to send unsolicited messages especially advertisements and bad links are called spam. Therefore, if you do not know the sender, the message may be spam. Many users do not realize that when downloading free services, software or updates, they are only signing up for certain emails. "Raw" means the email is not spam.

In this project, we apply the Naive Bayes technique to learn a model that may be used to categorize text as spam or non-spam. Frequently used as "free", "win", "win", "money", "gift", etc. It is used in its meanings. These languages exist because they are designed to grab your attention and force you to give them your full attention. Additionally, spam emails contain words written in all capital letters and also use lots of exclamation marks. Spam emails are often easily detected by recipients, and our goal is to train the model to do this for us. The definition of spam is a binary classification problem since messages are either categorized as "spam" or "not spam." This is also a learning problem because we will feed data to the model and it can learn to make predictions about the future.

SMS is a manual communication system that allows mobile phone users to send text messages. This is the most widely used data application and is expected to have 3.5 billion users by the end of 2010; This number accounts for approximately 80% of all mobile phone users [3]. As the importance of detail increases, we also see an increase in the number of jobs sent via mobile notifications. Spam doesn't even look like spam. In 2010, about 90% of emails were spam, and this is not yet a big problem in North America. The change in December 2012 was less than 1% [4]. But due

ISSN No:-2456-2165

to the significant growth of the youth market and the long-term decline in the cost of SMS notifications (the cost of instant messaging in China is now under $0.001), the potential for spam SMS is different. In 2012, this rate increased to 30% in Asia. Instant messages are spam. Some operators are responsible for sending instant messages in the Middle East. Additionally, spam is more problematic than spam because in some countries the cost of spam to recipients can increase. These factors, combined with limited access to mobile spam filters, make instant messaging spam an interesting problem worth studying. There are many differences between instant messages and text messages. The actual number of spam messages is very small compared to the amount of information available. Moreover, due to the current small size of the language, the number of domains available for its preparation is less than the corresponding number in the language.

II. LITERATURE SURVEY

[1] Due to the great popularity of short message services (SMS), spammers have managed to achieve many targets. Spam messages can trick mobile phone users into revealing confidential information, which can lead to serious consequences. The magnitude of this problem increases the need to develop spam filtering solutions. Machine learning algorithms have become the best tools to classify data into text. This explanation is perfect for our case because it separates SMS into two labels: spam or normal. By merging two machine learning techniques—supervised and unsupervised learning algorithms—this paper will show an SMS spam filtering solution. The new hybrid system is designed to improve spam filtering accuracy and F-measure.

[2] In today's world where digitalization is everywhere, messaging has grow to be among the most significant channels of communication. In contrast to alternative social media networks like Facebook and WhatsApp, messaging does not require an internet connection at all. It is known that hackers and spammers trying to hack devices and message supported mobiles had become more vulnerable as attackers try to filter through the system by sending uninvited links and the attacker gains remote access of mobile phones by clicking on such links. Therefore, to identify these messages, the authors have developed a system that can detect these types of messages and detect whether the message is SPAM or not. Authors use the TF - IDF Vectorizer algorithm to create a dictionary containing the entire content of spam messages.

[3] Although today's mobile phones continue to evolve with many different communication mediums, SMS still remains people's option of communication tools. However, today, as the cost of SMS has decreased, SMS spam has also increased, and some people are using SMS as another way to advertising and fraud. From there after, it became a major problem as it affects and harms users, and one of the solutions is to automatic SMS spam filtering. One of the most major problems in spam filtering is accuracy. In this work, we aim to improve SMS spam filtering which is performed by the combination of both information by correlation and classification. FP correlation enhancement is used to examine active SMS samples and here we have Naive Bayes classifier algorithm which is used to classify whether SMS as spam or the regular mail. Training the data using SMS spam collected in the earlier studies. For the SMS Spam Collection v.1 dataset, results in using Naive Bayes and FP-Growth collaboration improve the average maximum accuracy by 98%, 506% and 0.025% respectively and increases the correct score without using FP-Growth. Thus, making the department more efficient.

[4] Short Message Service Spam refers to uninvited links or unwanted messages which will be received on mobile phones. These spam messages are real problem for phone users. This business practice is also worrisome for service providers because it can upset their customers and even cause them to lose customers. To reduce this, researchers have suggested quite many solutions to control and filter out spam messages. In this article, we review the methods existing already, pros and cons, and further research for spam detection, filtering, and how to reduce risk of mobile SMS spam. The body of research literature was examined and evaluated. The most commonly occurring SMS spam detection, filtering, and reduction techniques are compared, including the data used their outcomes, limitations and future research directions are discussed. In this review the main goal is to help researchers to identify areas that require further development.

[5] Spam message analysis is an important task in identifying and filtering spam messages. We can see that; SMS messages are sent each day increases and it becomes more difficult for users to remember the new SMS messages they receive in their inbox and associate them with the messages they have received before. In this study, detection of messages spam and device identification problems are discussed. Plan is made in two phases. In the first phase, binary classification algorithm is used to divide the text into two groups i.e spam text and non-spam text. Then, in the second phase, non-negative matrix factorization and K-means group algorithm are used. The definition of SMS messages according to similar messages, that is, the duration of consecutive communication, is explained and the effect is tried to be analyzed at the beginning of the analysis title. Performance parameters such as accuracy, precision value, regression and F-measure are also studied and evaluated. SMS messages defined in this application can be used by other apps such as SMS message content, distribution in SMS inbox and other related SMS management.

III. METHODOLOGY

After cleansing the dataset, we divide it into training and testing sets. Using the training set of data, the Naive Bayes classifier is trained. Using test results, the efficacy of teacher preparation was evaluated. A Description of the Dataset The dataset SMS Spam Collection v.1 was utilized by us [9]. These data were downloaded from [10]. 5572 text messages, categorized as spam or regular, are included in the dataset. It is divided into two columns, v1 and v2. To indicate if the text in the second v2 column is legitimate email or spam, the first v1 column has two values: ham and spam only.

ISSN No:-2456-2165

The CSV (comma separated values) files for these files are provided. The Grumbletext website, the NUS SMS Corpus (NSC), Caroline Tag's PhD thesis, and the SMS Spam Corpus v.0.1 Large are the sources of the words in this article. In this instance, 747 messages were classified as spam and 4825 messages as regular messages. B. The original data lines v1 and v2 are renamed as class and text, respectively.

After renaming the columns, we shuffle the dataset to reduce over fitting. After shuffling, the data set is cleaned. To clean the file, all text is converted to lowercase and punctuation, and numbers, stopped words, and URLs are removed. Naive Bayes Classifier After data preprocessing, the dataset is divided into training data and testing data. There are a total of 5572 messages in the file, 747 of which are marked as spam and 4825 of which are marked as normal text. The data is divided into two data sets. The tutorial contains 4,000 words; Of these, 3,461 were marked as regular emails and 539 were marked as spam.

The evaluation data included the remaining 1,572 messages, of which 1,364 were normal messages and 208 were marked as spam. For classification, a model or distribution is created which is further used to predict class names [11]. First, we convert the text of the training data into time matrix data and remove words with frequency less than 5. The 0 entry of the time matrix data is replaced with "no ", and the other non-zero entries are replaced with "Yes". So, this information time matrix has only two values: "yes" and "no". Use this matrix of data elements and word lists from the training data to train the Naive Bayes classifier. Similarly, an object database was created for the test data and Naive Bayes classifier was used to predict the text name.

One of the most straightforward and effective algorithms for classifiers is Naive Bayes. The relationship between the possibility of the previous assumption from proof P(A) and the possibility of the ultimate assumption confirmed by proof P(AB) is defined by the Bayes theorem given assumption A and evidence B:

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}$$

Where:
A, B = event
P(AB) = probability of A given that B is true
P(BA) = probability of B given that A is true
P(A), P (B) = independence of A and B

The Naive Bayes classifier in statistics is a straightforward, uncertain classifier that makes use of Bayes' theorem, which is based on a hypothetical decision given the data and some prior information. Despite this basic assumption, which is regularly not the case in practice, the Naive Bayes classifier is widely used in many applications due to its effectiveness and efficiency.

Naive Bayes classifier algorithm is one of the simplest forms of model in Bayesian network. We can achieve high accuracy when combined with fast prediction. This method includes the use of kernel function to calculate the probability of all possible inputs and allowing the scheduler to improve its efficiency under certain conditions. Hence, Naive Bayes is a powerful tool in Machine Learning. It especially works effectively in text classification, filtering out spam messages.

	Category	Message
0	1	go jurong point crazy available bugis n great ...
1	1	ok lar joking wif u oni
2	0	free entry 2 wkly comp win fa cup final tkts 2...
3	1	u dun say early hor u c already say
4	1	nah dont think goes usf lives around though
...	...	...
5567	0	2nd time tried 2 contact u u 750 pound prize 2...
5568	1	b going esplanade fr home
5569	1	pity mood soany suggestions
5570	1	guy bitching acted like id interested buying s...
5571	1	rofl true name

5572 rows × 2 columns

Fig (1): Data Set

IV. RESULT

```
[1]: df['label'] = df.label.map({'ham':0, 'spam':1})
     X_train, X_test, y_train, y_test = train_test_split(df['sms_message'],df['label'],random_state=1)
     count_vector = CountVectorizer()
     training_data = count_vector.fit_transform(X_train)
     testing_data = count_vector.transform(X_test)
     naive_bayes = MultinomialNB()
     naive_bayes.fit(training_data, y_train)
     predictions = naive_bayes.predict(testing_data)
     print('Accuracy score: ', format(accuracy_score(y_test, predictions)))
```

Fig (2): Code

Output:
Accuracy score: 0.9885139985642(OR) 98%

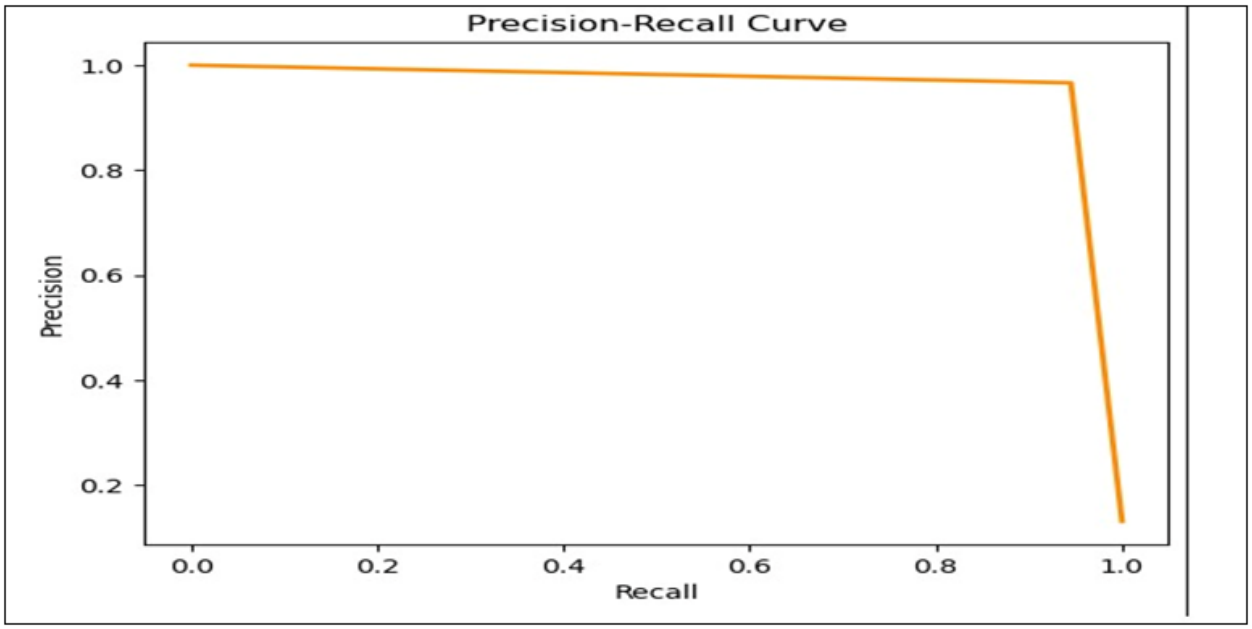

Fig (3): Precision-Recall Curve

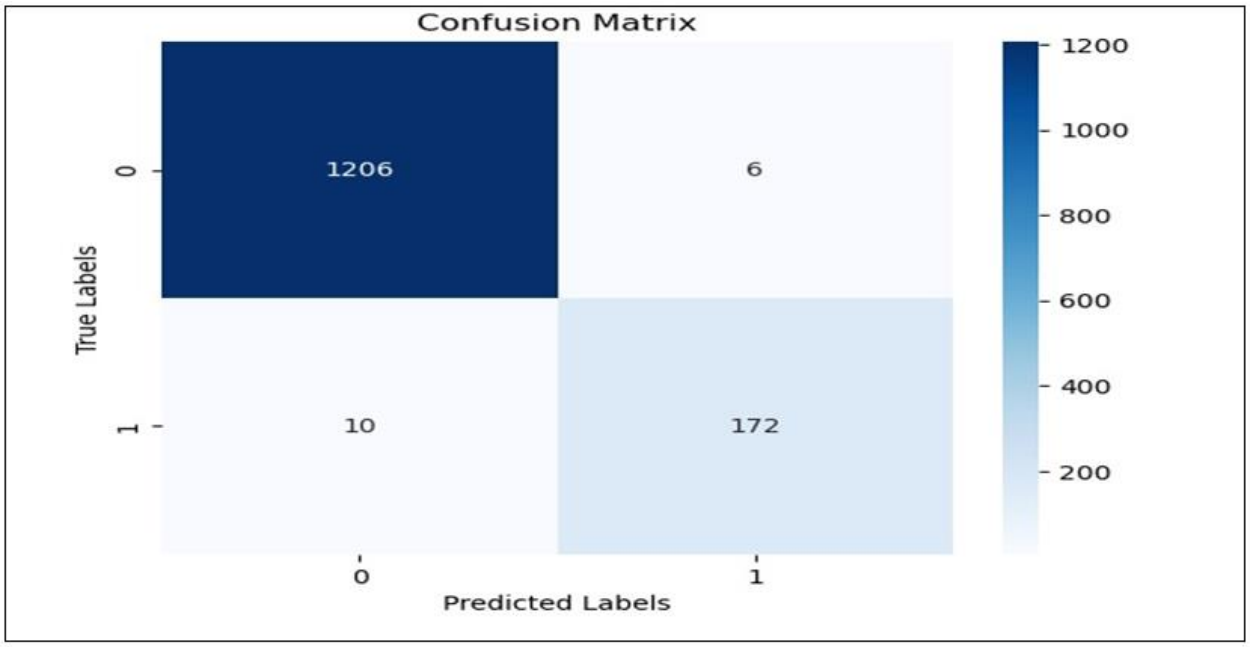

Fig (4): Confusion Matrix

ISSN No:-2456-2165

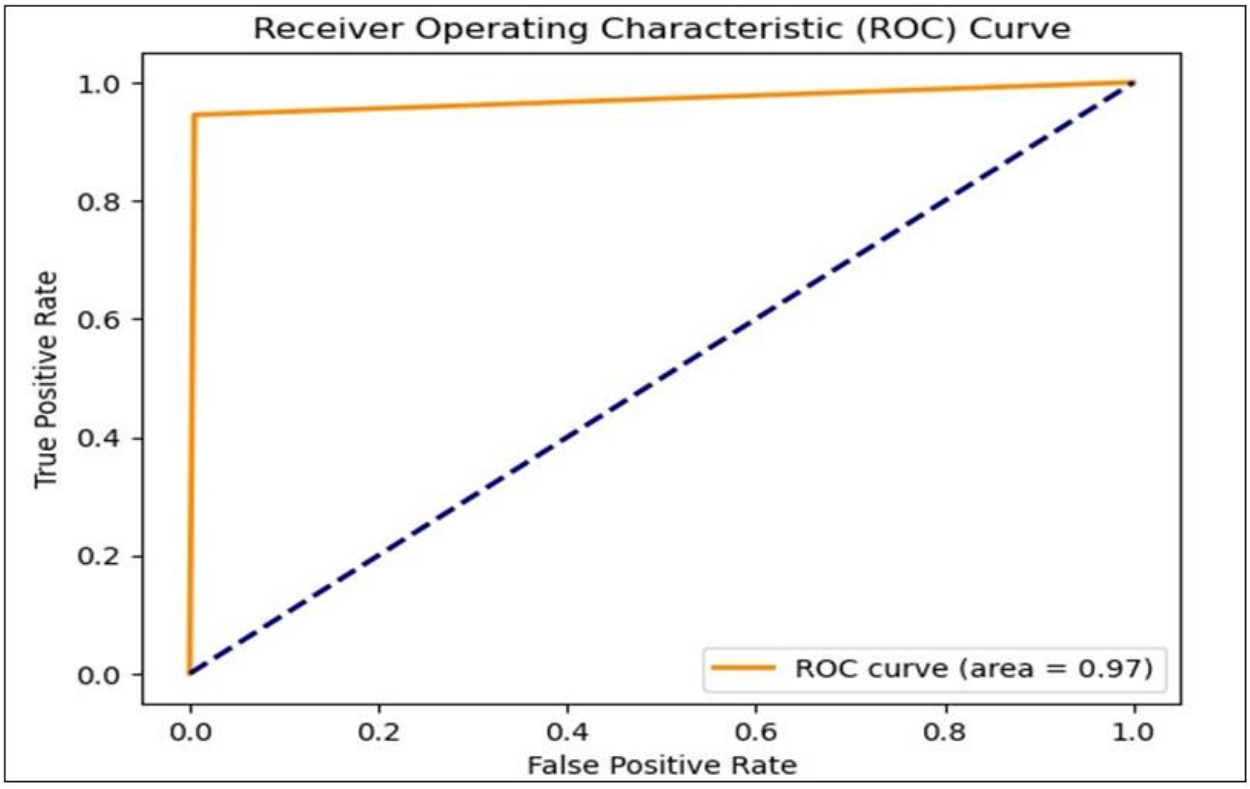

Fig (5): ROC Curve

V. CONCLUSION

We conclude that Naive Bayes algorithm is the best for classification in SMS spam detection and it is worth studying Polynomial Naive Bayes algorithm as it has many applications in many industries and the predictions made in this way of the algorithm are real and fast. Media classification is one of the most popular users of the Naive Bayes algorithm. News political, regional, international etc.

REFERENCES

[1]. Baaqeel, Hind, and Rachid Zagrouba. "Hybrid SMS Spam Filtering System Using Machine Learning Techniques." 2020 21st International Arab Conference on Information Technology (ACIT). IEEE, 2020.
[2]. Gupta, Suparna Das, Soumyabrata Saha, and Suman Kumar Das. "SMS Spam Detection Using Machine Learning." Journal of Physics: Conference Series. Vol. 1797. No. 1. IOP Publishing, 2021.
[3]. Dea Delvia Arifin, Shaufiah Moch and Arif Bijaksana, "Enhancing Spam Detection on mobile phone short message service(SMS) performance using FP-Growth and naive bayes classifier", Wireless and Mobile (APWiMob) 2016 IEEE Asia Pacific Conference, 2016.
[4]. Shafil Muhammad Abdulhamid, "A Review on Mobile SMS Spam Filtering Techniques", IEEE Access, 2017.
[5]. Nagwani Naresh Kumar and Aakanksha Sharaff, "SMS Spam Filtering and thread identification using bi-levcl text classification and clustering techniques", Journal of Information Science, 2017.