

Email/SMS Spam Classifier

MINI PROJECT REPORT

Submitted by

K GOKULAKRISHNAN (203002031)

G KARTHICK RAJ (203002045)

UEC1605

MACHINE LEARNING



**Department of Electronics and Communication
Engineering**

Sri Sivasubramaniya Nadar College of Engineering

(An Autonomous Institution, Affiliated to Anna University)

Rajiv Gandhi Salai (OMR), Kalavakkam – 603 11

EVEN SEM 2022-2023

Sri Sivasubramaniya Nadar College of Engineering
(An Autonomous Institution, Affiliated to Anna University)

BONAFIDE CERTIFICATE

Certified that this mini project titled “**Email/SMS Spam Classifier**” is the bonafide work of “**K GOKULAKRISHNAN (203002031), G KARTHICK RAJ (203002045)** of **VI Semester Electronics and Communication Engineering Branch** during **Even Semester 2022 – 2023** for **UEC1605 Machine Learning**

Submitted for examination held on _____09.05.2023_____

INTERNAL EXAMINER

ABSTRACT

This project focuses on developing an Email/SMS spam classifier using the Naive Bayes theorem. The Naive Bayes algorithm is a simple and effective method for building classifiers, and it works well for text classification tasks such as spam detection. The goal of this project is to classify a given message as either spam or ham using the Naive Bayes classifier. The classifier is trained using a dataset of labeled messages, where each message is tagged as either spam or ham. The Naive Bayes classifier is then used to predict the classification of new messages based on their features. The performance of the classifier is evaluated using metrics such as accuracy, precision, recall, and F1-score. The results show that the Naive Bayes classifier is an effective method for email/SMS spam classification, with high accuracy and low false positive rates.

INTRODUCTION

The problem of email and SMS spam has been a persistent issue since the advent of electronic communication. Spam messages not only waste valuable time and resources, but they can also be dangerous if they contain malicious links or attachments. Therefore, developing an effective spam filter is essential to protect users from these unwanted messages. One popular approach to spam classification is the Naive Bayes algorithm, which is a simple yet powerful method for building classifiers. The Naive Bayes algorithm assumes that the features of a message are independent of each other, which simplifies the computation of the probability of a message belonging to a particular class. In this project, we aim to develop an email/SMS spam classifier using the Naive Bayes algorithm and evaluate its performance using metrics such as accuracy, precision, recall, and F1-score. This project will contribute to the development of better spam filters to protect users from unwanted and potentially harmful messages. The motive of this project is to build a model based on Naive Bayes Classifier Algorithm to detect spam messages and classify it.

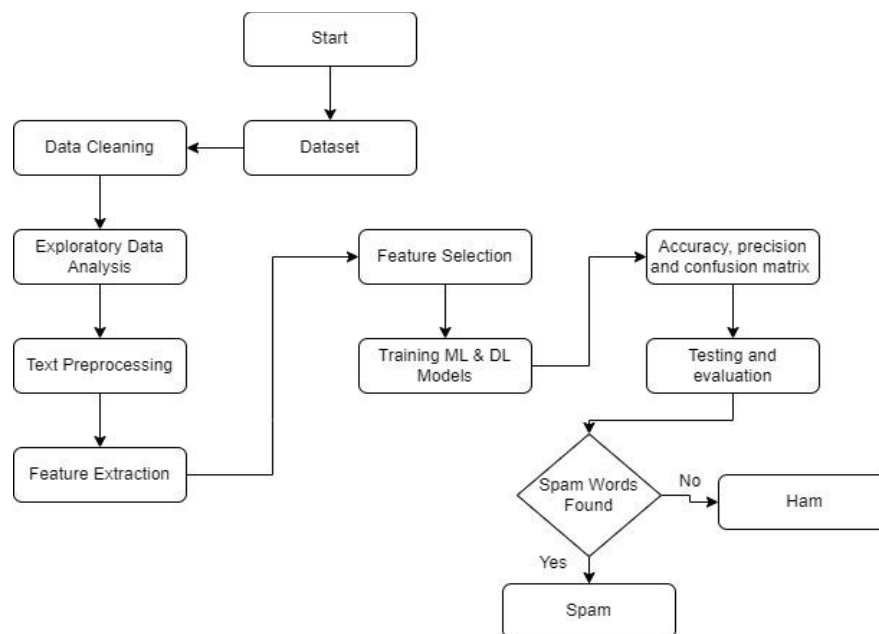
OBJECTIVE

The objectives of identification of spam emails/SMS are :

- To give knowledge to the user about the fake emails and relevant emails.
- To classify the emails/SMS as spam or not using machine learning algorithms as spam fills inbox with unwanted messages.
- To build a model based on Naïve Bayes Classifier Algorithm to detect spam messages and classify the messages as ham or spam.

SYSTEM /MODEL DEVELOPED

- In this project, we are using the Naive Bayes algorithm to implement email/sms spam detection.
- We have chosen this model because it has the best precision score among all the other models that we tested and its accuracy score is around 97%. (Multinomial NB)
- Naive Bayes classifiers work by correlating the use of tokens (typically words, or sometimes other things), with spam and non-spam e-mails and then using Bayes' theorem to calculate the probability that an email is or is not spam.
- Calculation of the probability is based on the Bayes formula and the components of the formula are calculated based on the frequencies of the words in the whole set of messages.



DATASET DETAILS

- In this proposed system, a dataset from “Kaggle” website is used as a training dataset. (SMS Spam Collection Dataset)
- Link: <https://www.kaggle.com/datasets/uciml/sms-spam-collection-dataset>
- The SMS Spam Collection is a set of SMS tagged messages that have been collected for SMS Spam research. It contains one set of SMS messages in English of 5,574 messages, tagged according to being ham (legitimate) or spam.
- The files contain one message per line. Each line is composed by two columns: v1 contains the label (ham or spam) and v2 contains the raw text.
- The data has been collected from free or free for research sources from the Internet and is the most commonly used dataset for SMS spam detection.

RESULTS & DISCUSSION

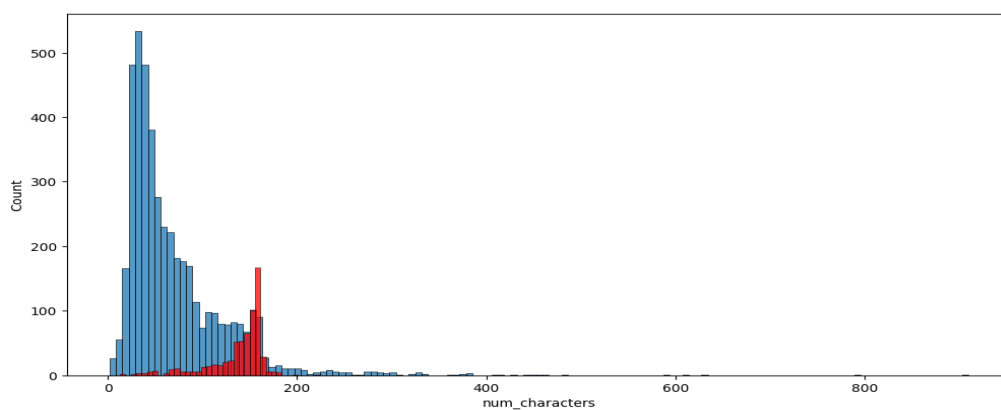


Figure 1: Graph for number of characters with respect to count. (Blue – ham, Red – spam)

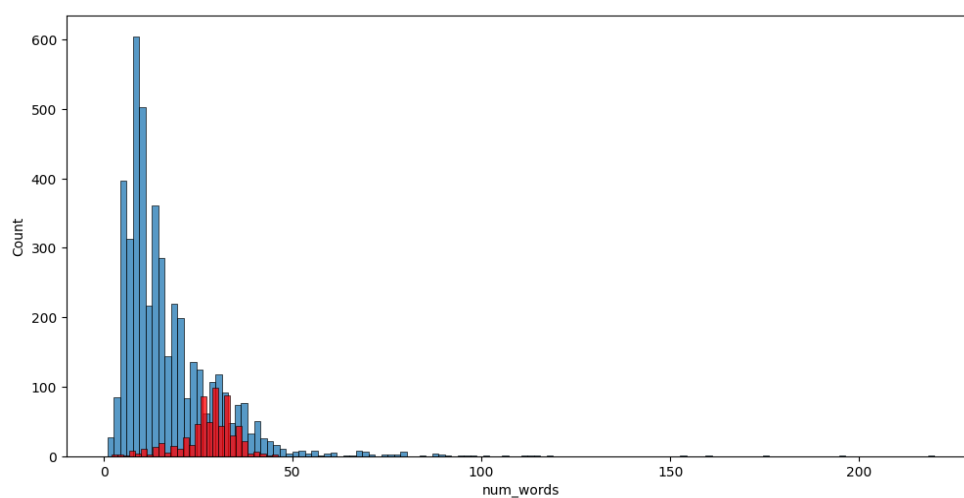


Figure 2: Graph for number of words with respect to count. (Blue – ham, Red – spam)

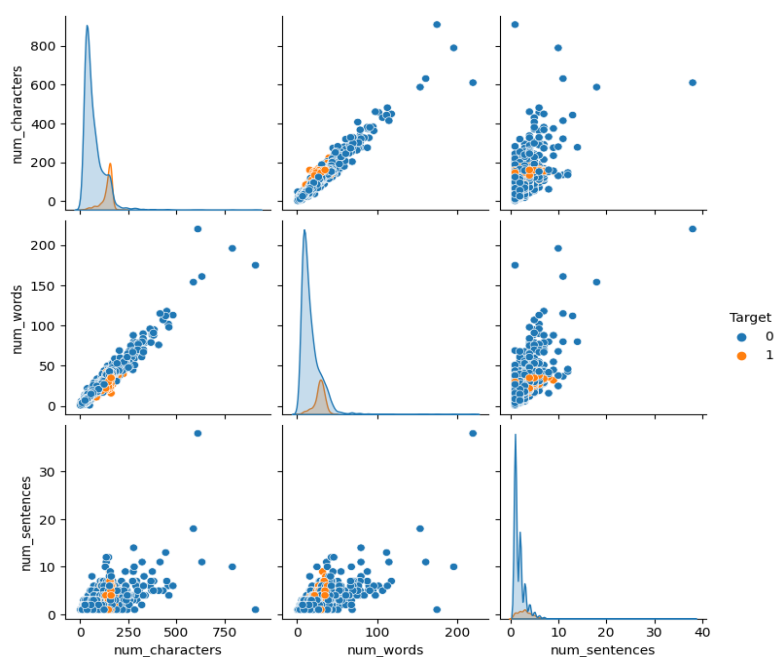


Figure 3: Pair plot for various features in the data set

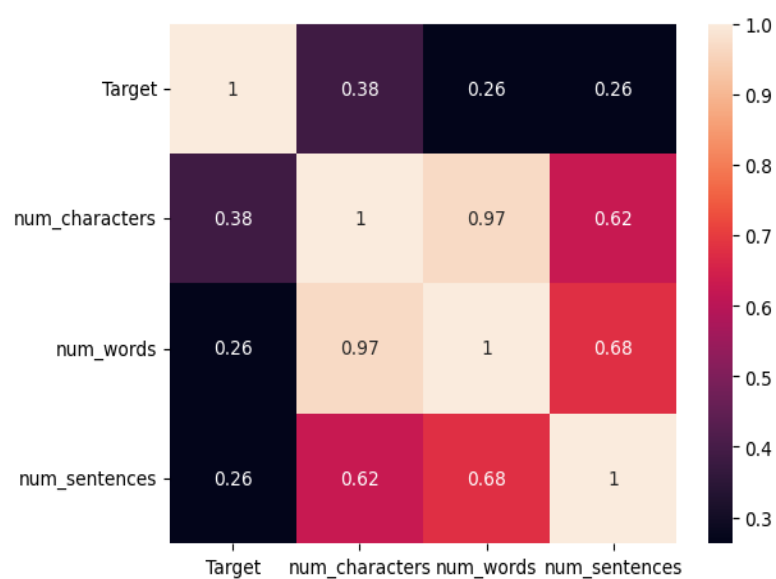


Figure 4: Heat map of the different features used in spam and ham messages.

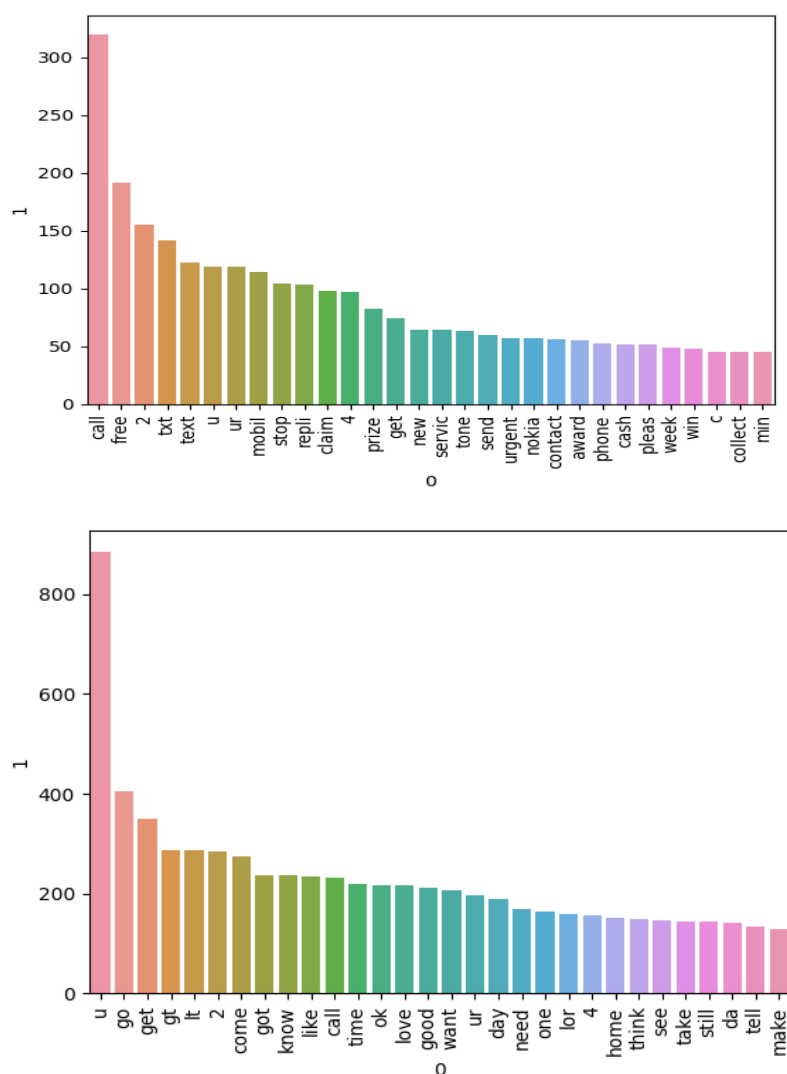


Figure 5 & 6: Bar plot for frequent words in ham and spam messages

Evaluation metrics for different types of Naive Bayes models (Accuracy score, confusion matrix and precision score)

- Gaussian Naïve Bayes

```
0.8694390715667312
[[788 108]
 [ 27 111]]
0.5068493150684932
```

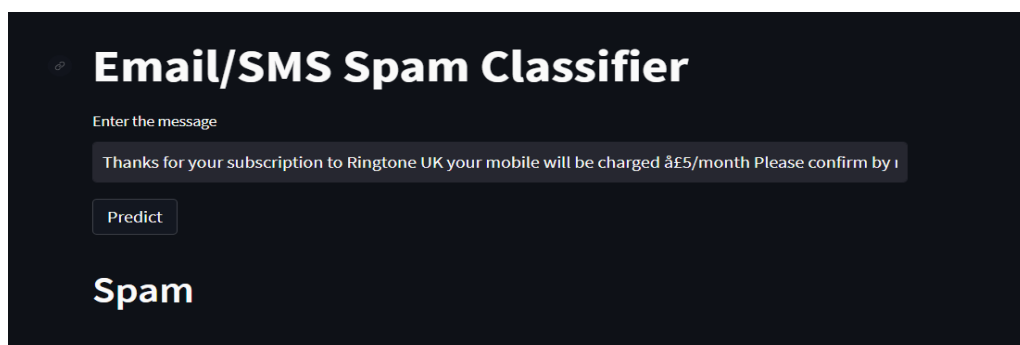

- **Multinomial Naïve Bayes**

```
0.9709864603481625
[[896  0]
 [ 30 108]]
1.0
```

- **Bernoulli Naïve Bayes**

```
0.9835589941972921
[[895  1]
 [ 16 122]]
0.991869918699187
```

Results displayed in App



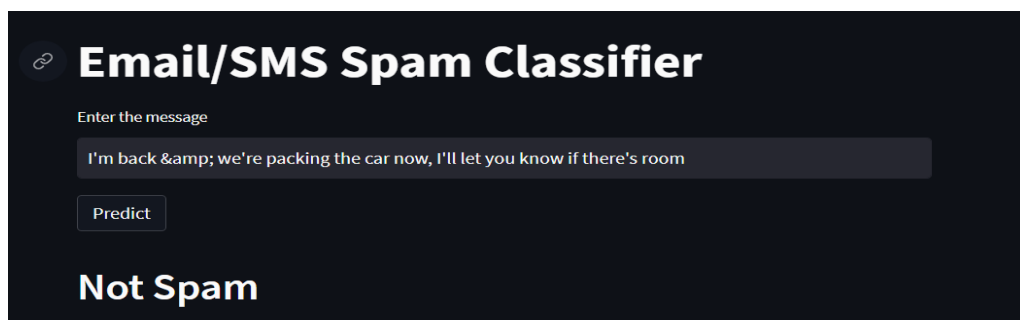
Email/SMS Spam Classifier

Enter the message

Thanks for your subscription to Ringtone UK your mobile will be charged £5/month Please confirm by i

Predict

Spam



Email/SMS Spam Classifier

Enter the message

I'm back & we're packing the car now, I'll let you know if there's room

Predict

Not Spam

Table 1 Different models and their accuracy and precision score

Model	Accuracy (%)	Precision (%)
Gaussian Naïve Bayes	86.94	50.68
Multinomial Naïve Bayes	97.09	100
Bernoulli Naïve Bayes	98.35	99.18

CONCLUSION

- In conclusion, Naive Bayes Algorithm has proven to be an effective method for classifying emails as spam or not spam. The algorithm works by calculating the probability of an email belonging to a particular class based on the occurrence of certain words or features. This approach is relatively simple, but it is surprisingly accurate in practice.
- During the project, we trained a Naive Bayes model on a dataset of emails labeled as spam or not spam. We then tested the model on a separate set of emails and evaluated its performance using metrics such as precision and accuracy score.
- Our results showed that the Naive Bayes model was able to achieve a high level of accuracy in classifying emails. This indicates that the model is effective in identifying spam emails while minimizing false positives and false negatives.
- Overall, this project highlights the potential of machine learning algorithms like Naive Bayes for detecting spam emails. By using these techniques, we can improve email security and reduce the amount of unwanted emails that users receive.

REFERENCES

- [1] Pritesh A. Patil, Prayag P. Bhosale. (2022). Literature Survey on Spam Email Detection. [Online] Available at:
<https://ijrpr.com/uploads/V3ISSUE11/IJRPR8167.pdf>
- [2] Email Spam Detection Using Machine Learning Algorithms. (2021). [Blog] Available at : <https://jpinfotech.org/email-spam-detection-using-machine-learning-algorithms/>
- [3] Sakshi Gupta (2021). Email Spam Filtering Using Naïve Bayes Classifier. [Blog] Available at: <https://www.springboard.com/blog/data-science/bayes-spam-filter/>