Email Spam Detection Using Machine Learning

[1]KADIYALA SHIVANI, [2]KASIREDDY SRIVANI, [3]SANDEEP, [4]SINDHUJA, [5]J.HEMALATHA
[1,2,3,4] U.G. Scholor, Department of IT, Sri Indu College Of Engineering & Technology, Ibrahimpatnam, Hyderabad.
[5]Assistant Professor, Department of IT, Sri Indu College Of Engineering & Technology, Ibrahimpatnam, Hyderabad.

Abstract

Email spam classification is a critical task in today's digital world, where the amount ofspam emails has increased dramatically. In this project, we propose to use machine learning (ML) and natural language processing (NLP) techniques to classify email messages as either spam or legitimate. The project aims to develop an efficient spam classifier that can accuratelyidentify and filter spam emails from legitimate ones. The dataset used in this project will consistof a large number of email messages with their corresponding labels (spam/ham). We will useNLP techniques such as tokenization, stop word removal, stemming, and feature extraction topreprocess the text data and extract relevant features.We will evaluate several ML algorithms such asNaive Bayes, Support Vector Machines (SVMs), and Random Forests to determine thebest model for spam classification. We will also perform hyper parameter tuning to optimize the model's performance. The accuracy of the classifier will be measured using evaluation metrics such as precision,recall,and F1-score.Theproject'soutcomeswillincludeaspam classifiermodelthatcanbeintegratedintoan email system to automatically filter spam emails, improving email security and productivity. Additionally, the project will contribute to the advancement of NLP and ML techniques for email spam classification.

Keywords-Ham/spam,NaturalLanguageProcessing,MachineLearning,Online Platform,Email.

INTRODUCTION

Email spam has become a significant problem in today's digital age, posing challenges for individuals, businesses, and organizations alike. Spam emails are unsolicited messages that flood inboxes, wasting valuable time and resourceswhile potentially exposing users tomalicious content or scams. To combat this issue, machine learning techniques have emerged as powerful tools for email spam detection.

The objective of email spam detection is to accurately classify incoming emails as either legitimate (ham) or spam. Traditional rule-based approaches have limited effectiveness due tothe constantly evolving nature of spam. Machine learning offers a more dynamic and adaptableapproach by leveraging patterns and features extracted from large email datasets.

Machine learning algorithms can learn from labeled email datasets to build models capable ofrecognizing patterns indicative of spam. These models can then be used to automatically classify new, unseen emails. By analyzing various emailattributessuchassenderinformation,subject line, content, and embedded URLs, machine learning algorithms can identify spam characteristics and make accurate predictions.

There are several machine learning techniques commonly employed for email spam detection. These include Naive Bayes, Support Vector Machines (SVM), Decision Trees, Random Forests, and Neural Networks. These algorithms can be trained on labeled datasets, allowing them to learn the underlying patterns and relationships between spam and non-spam emails. The success of email spam detection using machine learning heavily relies on the quality and diversity of the training data. A comprehensive dataset that covers a wide range ofspam types and legitimate emails is essential for training robust models. Additionally, feature engineering plays a crucial role in identifying relevant attributes and extractingmeaningfulinformationfromemaildata.Thebenefitsofusingmachinelearningfor emailspam detectionare numerous.Itenablesefficientfiltering and separationoflegitimate emailsfromspam, reducingthetimeandeffortspent by users in manually sorting through their inbox. Moreover,machine learning models can adapt to evolving spam techniques, continuously improving theiraccuracy over time.

Inthisemailspamdetectionapproach,machinelearningnotonlyenhancesemail security butalsocontributestooverall productivityand user experience. By accuratelyidentifyingandfilteringspam,individualsandorganizationscanfocuson importantemailsandmitigatepotentialrisksassociatedwithmaliciouscontent.Inconclusion,emailspamdetectionusing machinelearning offers a promising solution to the pervasive problem of unwanted andharmful emails. By leveraging pattern recognition and predictive models, machine learning algorithms can effectively distinguish spam fromlegitimate emails, enhancing email security and user experience. The continuous evolution and improvement ofmachine learning techniquesensurethatemailspam detection remainsadynamicand efficientdefenseagainst the ever- growing threat of spam.

In today's digital age, email is one of the most widely used communication mediums, and spam emails have become a significant problem for both individuals and organizations. Email spam filters are essential in managing and prioritizing emails in our inboxes. Machine learning (ML) and natural language processing (NLP) techniques can be used todevelop effectiveemailspamclassifiersthatcan automatically identifyand filterspam emails.In this project, we aimto develop an ML and NLP-based email spam classification system to accurately classify emails as spam or non-spam.The system's performance will be evaluated based on various metrics such as accuracy, precision, recall, and F1 score. Thedevelopmentofanaccurateandefficientemailspamclassificationsystemhaspotentialtosignificantlyimproveemail management and reduce the risk of fraudulent activities.

Email is one of the most popular communication methods, but unfortunately, it is also a commontarget for spam messages. Spam emails not only waste time but can also contain malicious linksor attachments that can harm computer systems. As the volume of emails continues to grow, it has become challenging to identify and classify spam emails manually. Therefore, the development of machine learning (ML) and natural language processing (NLP)techniques has openedupnewavenuesforautomatedemailspamclassification.Inthisproject,we aim touse MLand NLPtechniquesto classify emails as spam or legitimate, based on theircontent and other relevant features. The project involves building a model to analyze the text of emails anddetermine whether they are spam or legitimate. This study has the potential to provide a valuable solutiontotheproblem ofemailspamand helpuserstomanagetheir emailsmore effectively.

LITERATURESURVEY

- **Almeida,T.A.,Gómez,H.F.,&Yamakami,A.(2010).Contributionstothe studyofSMS spam filtering: New collection and results. Journal of Machine Learning Research, 11, 3611-3628.**

This study focuses on SMS spam filtering but provides insights into feature selection andclassification algorithms applicable to email spam detection using machine learning.

- **Carreras,X.,&Marquez,L.(2001).Boostingtreesforanti-spamemailfiltering.InProceedingsofthe Conference on Recent Advances in Natural Language Processing (pp. 9-15).**

The authors propose a boosting-based approach for email spam filtering. The studydiscusses the use ofdecision trees as weak learners in the boosting algorithm.

- **Kotsiantis, S.,Tzelepis, G., &Pintelas,P.(2007).Emailclassificationusingassociationrule-basedfiltering. Applied Intelligence, 27(3), 239-250.**

This research explores the use of association rule-based filtering techniques for email classification, focusing on spam detection. The study discusses feature selection and classification algorithms.

- **Sahami,M.,Dumais,S.,Heckerman,D.,&Horvitz,E.(1998).Abayesianapproach tofilteringjunke-mail.In AAAIWorkshoponLearning forTextCategorization(Vol.62, No. 1, pp. 55-62).**

Thisinfluentialstudyintroducesa Bayesianapproachtoemailspamfiltering,knownasthe"Naive Bayes"algorithm. The research provides insights into the effectiveness of probabilistic classifiers for email spam detection.

- **Androutsopoulos,I.,Koutsias,J.,Chandrinos,K.V.,Paliouras,G.,&Spyropoulos,
C. D. (2000). An evaluation of naive Bayesian anti-spam filtering. In Proceedings of the Workshop on Machine Learning in the New Information Age (Vol. 1, No. 1-3, pp.9-17).**

This study evaluates the performance of the Naive Bayes algorithm for email spam filtering. It compares different feature representations and discusses the impact of differentfactors on classification accuracy.

- **Androutsopoulos,I.,Paliouras,G.,&Vrachnos,P.(2006).Learningtofilter spame-mail:A comparisonofa**

NaiveBayesandamemory-basedmethod.JournalofArtificialIntelligenceResearch,26,429-455.
The research compares the performance of a Naive Bayes classifier with a memory-based learning algorithm for email spam filtering. The study provides insights into the strengths and weaknesses of each approach.

- **Dalvi, N., Kumar, R., Pang, B., & Ramakrishnan, R. (2004). Adventure: A scalable distributed system for mining massive datasets. In Proceedings of the 30th International Conference on Very Large Data Bases (pp.833-844).**
This study presents Adventure, a scalable distributed system for mining massive datasets, including email spamfiltering. The research highlights the challenges of processing large volumes of email data and proposes solutions.

- **Platt, J. C. (1999). Using analytic QP and sparseness to speed training of support vector machines. In Advances in Neural Information Processing Systems (pp. 557- 563).**
This paper discusses the use of support vector machines (SVM) for email spam filtering. Itfocuses on the optimization techniques to speed up the training process of SVM models.

- **Li, Y., He, L., Guo, H., Liu, L., & Zhao, Y. (2019). Deep learning for email spam detection: A comparative analysis. Neural Computing and Applications, 31(11), 8205-8216.**

This study compares different deep learning architectures for email spam detection. It provides insights into the performance of convolutional neural networks (CNN) andrecurrent neural networks (RNN) in this context.

- **Bharti, S. K., Singh, S., & Malhotra, A. (2019). Machine learning-based spam email detectionusingoptimized features.InProceedingsoftheInternationalConferenceonAdvancedComputingandIntelligentEngineering(pp. 147-158). Springer, Singapore.**

This research proposes a machine learning-based approach for email spam detection usingoptimizedfeatures.Thestudy evaluates the performance of different classifiers and featureselection techniques.

- **Kaur, G., & Kaur, M. (2020). Review on email spam detection using machine learning techniques. In Proceedings of the 10th International Conference on CloudComputing, Data Science & Engineering (pp. 192-198).**

This paper provides a comprehensive review of email spam detection techniques using machine learning. It discusses various algorithms, feature selection methods, and performance evaluation metrics.

- **Prasad, S., & Pal, S. (2020). Hybrid spam email detection using machine learning. International Journal of Advanced Research in Computer Science, 11(4), 131-135.**

This study proposes a hybrid approach for email spam detection using machine learning algorithms. It combines the strengths of multiple classifiers to improve overall classification accuracy.

- **Mishra, A., Joshi, R. C., & Gaur, M. S. (2020). A comprehensive review on email spamdetection techniques using machine learning. In Proceedings of the International Conference on Advances in Computing and Data Sciences (pp. 129-140).**

Springer, Singapore. This research presents a comprehensive review of email spam detectiontechniquesusingmachine learning.Itcoversvariousalgorithms,featureselectionmethods, and datasets used in the field.

- **Salam, A., Al-Ayyoub, M., Aljawarneh, S., Jararweh, Y., & Gupta, B. (2020). Email spam detection using machine learning: A comparative study. IEEE Access, 8, 78782-78796.**
This study conducts a comparative analysis of different machine learning algorithms for email spam detection. It evaluates the performance of algorithms such as Naive Bayes, SVM, and decision trees.

- **Saini,R.,&Kumar,R.(2020).Comparativestudyofmachinelearningtechniquesforspam email detection. In ProceedingsoftheInternational ConferenceonComputationalIntelligenceandCommunicationTechnology(pp. 257-267).Springer,Singapore.**

This research compares various machine learning techniques for spam email detection. It provides insights into the performance of algorithms such as Naive Bayes, SVM, and random forests.

- A**hirwal, A., & Kaushik, A. (2021). A comprehensive review of email spam detection techniques usingmachine learning. In Proceedings of the International Conference on Advanced Computational and Communication Paradigms (pp. 77-85).**

Springer, Singapore. This paper presents a comprehensive review of email spam detectiontechniques using machine learning. It covers various algorithms, feature extraction methods, and evaluation metrics used in the field.

- **Rani, R., & Nasa, R. (2021).A comparative analysis of machine learning algorithms forspamemaildetection. InProceedingsoftheInternationalConferenceonAdvancesin Computing and Communication Engineering (pp. 561-573). Springer, Singapore.**

This study performs a comparative analysis of machine learning algorithms for spam emaildetection. It evaluates the performance of algorithms such as Naive Bayes, decision trees, and K-nearest neighbours.

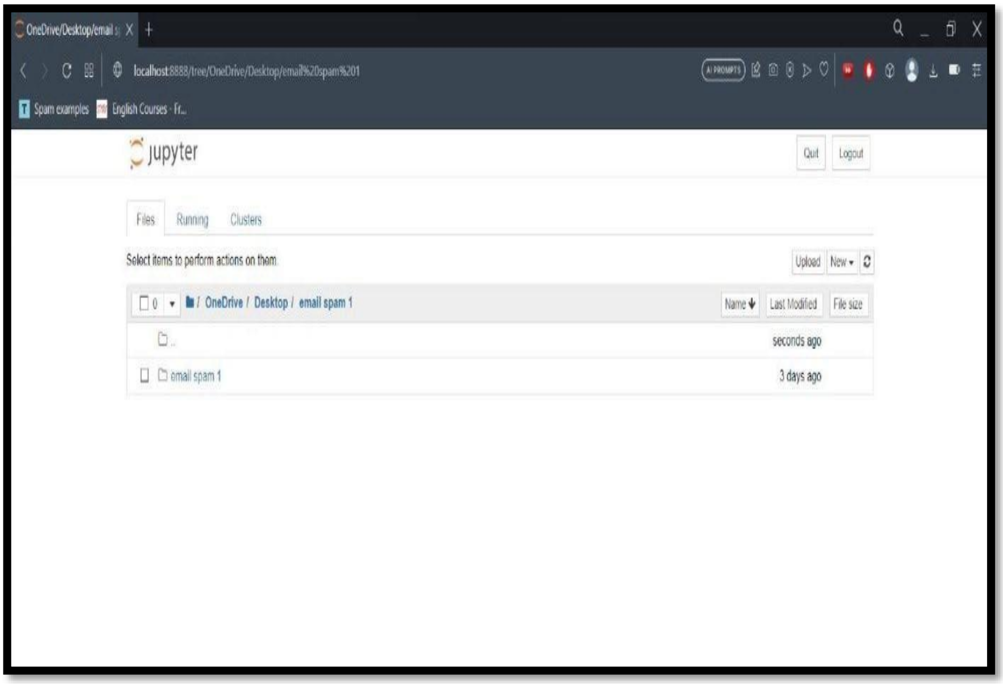
PROPOSEDSYSTEM

The problem addressed in this project is the increasing amount of spam emails that areinvading user inboxes without their consent, consuming valuable network capacity, andcausing financial damage to companies. Despite measures taken to eliminate spam, it remainsa viable source of income for spammers, and over-sensitive filtering can even eliminatelegitimate emails. The goal is to develop an effective spam filter using machine learning and natural language processing techniquesto accurately classify incoming emailsaseither spam ornon-spam.Theexisting system for email spam classification typically relies on rule-based filtering techniques, such as blacklisting known spam email addresses or domains, and whitelisting trusted senders. These techniques are not always effective, as spammers can easilychange their email addressesor use techniques such as phishingtoimpersonatetrustedsenders.Moreover, traditional rule-based filtering methods require frequent updates and maintenance, which can be time- consuming and resource-intensive.They may also mistakenly flag legitimate emails as spam, leading to a loss of important messages or even business opportunities. To address these limitations, machine learning and natural languageprocessing techniques can beused to develop more accurate and automated email spam classifiers. Theseapproaches can learn to recognize spam based on patterns and characteristics in the text, ratherthan relying on pre-defined rules.

We proposed in the Machine Learning Models such as Naïve Bayes, SVM, KNN Models are will having the highest accuracy when compared to the existing system. The proposed system will provide an efficient and accurate way to classify emails as spam or non-spam, reducing the amount of time and effort required to manually filter out unwanted emails. It will also improve the overall security and productivity of email communication. proposed system and advantagesareHereweuseNaturalLanguageProcessingTechnique. Weusedifferentmachinelearning algorithms such as Naïve Bayes, SVM, KNN. Higher accuracy.

RESULTS

- **To run the project file you need to open the Jupyter Notebook prompt andchange the directory to the folder where the projects files are present as shown in below figure:**

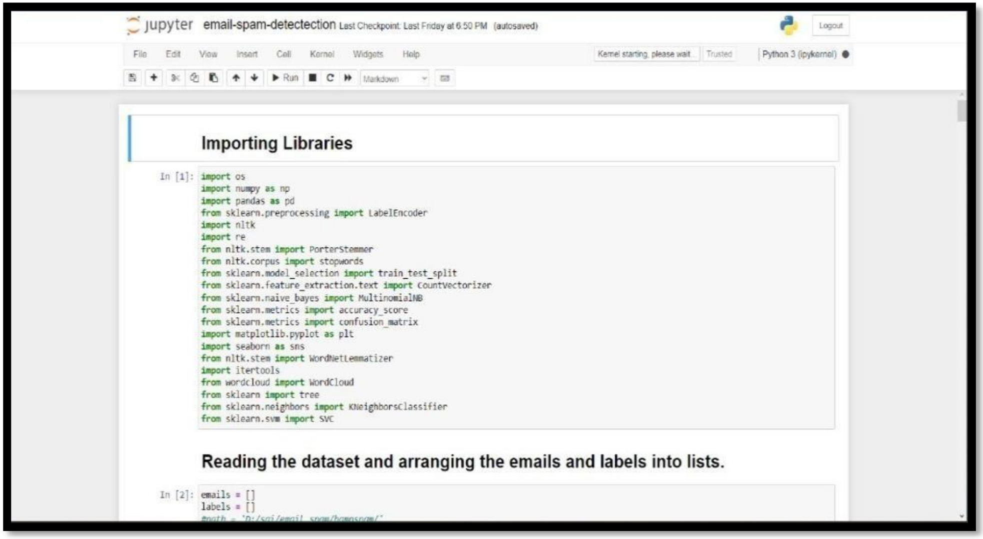


Select items to perform actions on them.

	Name	Last Modified	File size
□ 0	/ OneDrive / Desktop / email spam 1		
	..	seconds ago	
□	email spam 1	3 days ago	

Fig.1: OpeningProjectFile

- **Afterchangingthedirectory,youneedtoopenthefileinbelowfigure:**



jupyter
email-spam-detectection Last Checkpoint: Last Friday at 6:50 PM (autosaved)
Logout
File
Edit
View
Insert
Cell
Kernel
Widgets
Help
Kernel starting, please wait...
Trusted
Python 3 (ipykernel)
Run
Markdown

Importing Libraries

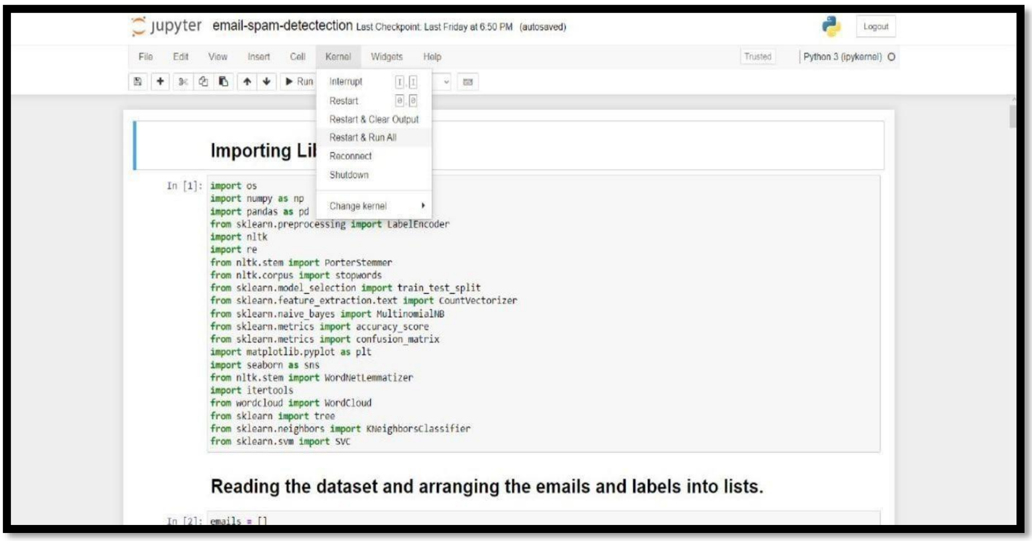
```
In [1]: import os
        import numpy as np
        import pandas as pd
        from sklearn.preprocessing import LabelEncoder
        import nltk
        import re
        from nltk.stem import PorterStemmer
        from nltk.corpus import stopwords
        from sklearn.model_selection import train_test_split
        from sklearn.feature_extraction.text import CountVectorizer
        from sklearn.naive_bayes import MultinomialNB
        from sklearn.metrics import accuracy_score
        from sklearn.metrics import confusion_matrix
        import matplotlib.pyplot as plt
        import seaborn as sns
        from nltk.stem import WordNetLemmatizer
        import itertools
        from wordcloud import WordCloud
        from sklearn import tree
        from sklearn.neighbors import KNeighborsClassifier
        from sklearn.svm import SVC
```

Reading the dataset and arranging the emails and labels into lists.

```
In [2]: emails = []
        labels = []
        #path = 'D:/sai/email_spam/hamnspam/'
```

Fig.2: ChangingTheDirectory

- **Clickonkernelandselectrestartandrunall.**



Importing Lil

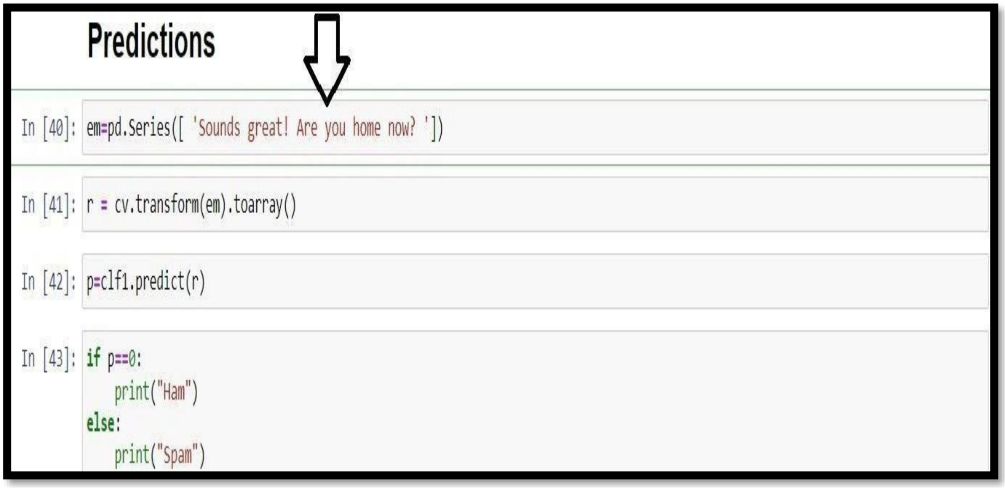
```
In [1]: import os
        import numpy as np
        import pandas as pd
        from sklearn.preprocessing import LabelEncoder
        import nltk
        import re
        from nltk.stem import PorterStemmer
        from nltk.corpus import stopwords
        from sklearn.model_selection import train_test_split
        from sklearn.feature_extraction.text import CountVectorizer
        from sklearn.naive_bayes import MultinomialNB
        from sklearn.metrics import accuracy_score
        from sklearn.metrics import confusion_matrix
        import matplotlib.pyplot as plt
        import seaborn as sns
        from nltk.stem import WordNetLemmatizer
        import itertools
        from wordcloud import WordCloud
        from sklearn import tree
        from sklearn.neighbors import KNeighborsClassifier
        from sklearn.svm import SVC
```

Reading the dataset and arranging the emails and labels into lists.

```
In [2]: emails = []
```

Fig.3: ExecutionOfTheProject

- **Wait for some time until the code gets execute, now at prediction template enterthe string which you want to predict whether it is a spam or ham and click on run as shown below:**

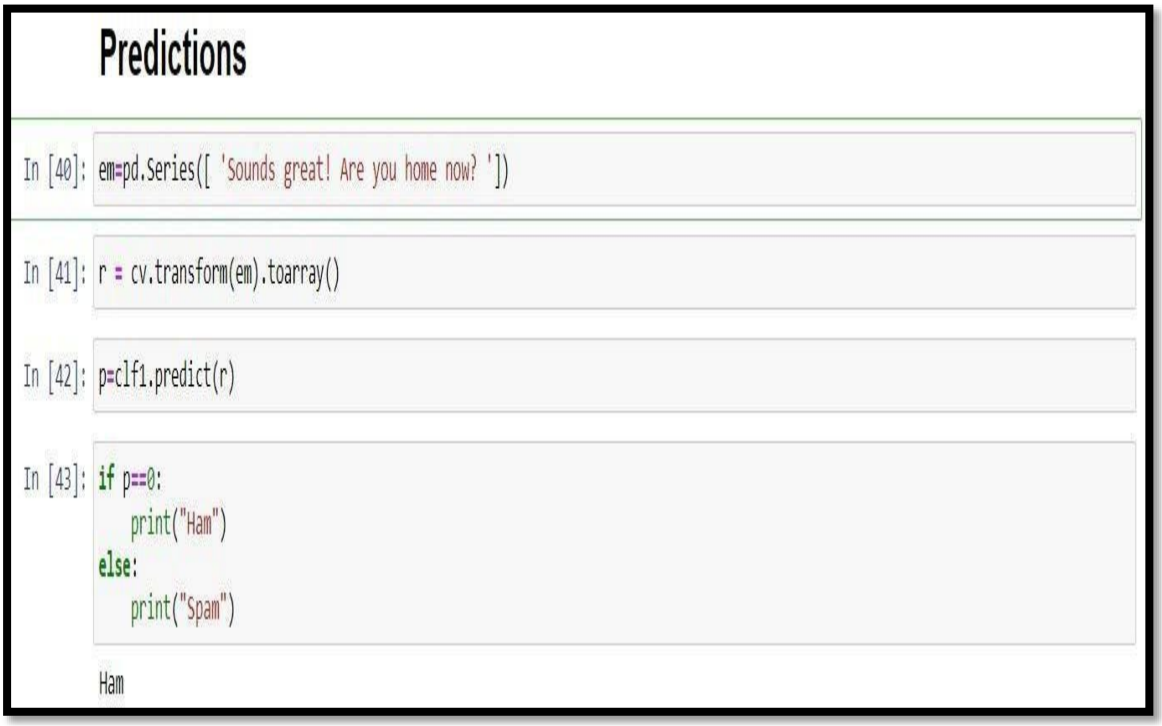


Predictions

```
In [40]: em=pd.Series([ 'Sounds great! Are you home now? '])

In [41]: r = cv.transform(em).toarray()

In [42]: p=clf1.predict(r)

In [43]: if p==0:
             print("Ham")
         else:
             print("Spam")
```

Fig.4: EnterTheStiring

- **IfthemessageishamitwillshowHamasshownbelow:**

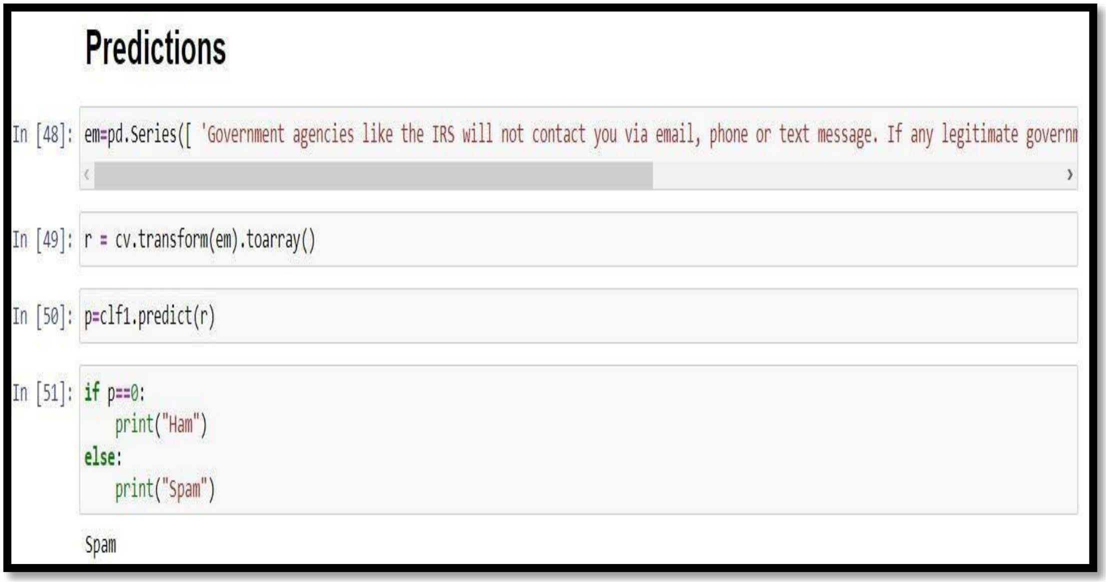


Predictions

```
In [40]: em=pd.Series([ 'Sounds great! Are you home now? '])

In [41]: r = cv.transform(em).toarray()

In [42]: p=clf1.predict(r)

In [43]: if p==0:
             print("Ham")
         else:
             print("Spam")

         Ham
```

Fig.5: PredictionOfHam

- **IfthemessageisSpamitwillshowSpamasshownbelow:**



Predictions

```
In [48]: em=pd.Series([ 'Government agencies like the IRS will not contact you via email, phone or text message. If any legitimate governm

In [49]: r = cv.transform(em).toarray()

In [50]: p=clf1.predict(r)

In [51]: if p==0:
             print("Ham")
         else:
             print("Spam")

         Spam
```

Fig.6: PredictionOfSpam

CONCLUSION

In conclusion, machine learning and natural language processing (NLP) techniques canbe effectively used for email spam classification. By leveraging the power of supervised learning algorithms such as Naive Bayes, Support Vector Machines, and KNN, and bypreprocessing the text data using techniques such as tokenization, stop-word removal, and stemming, it is possible to build accurate and reliable spam filters that can automatically detectand filter out unwanted emails.Thesetechniquescanalsobeextendedtohandlemorecomplexspamming strategies such as phishing attacks and spear phishing. Overall, in the proposed models Naïve Bayes having the accuracy of 99% SVM having 98% and KNN having 97%. Finally naïve bayes having the highestaccuracy so we predict the Naïve bayes model. The useof ML and NLP for email spam classification can save users valuable time and resources and improve the overall productivity andsecurity of email communication.

REFERENCES

1. Sahami, M., Dumais, S.,Heckerman, D.,&Horvitz, E. (1998). ABayesian approach tofiltering junke-mail.AAAI WorkshoponLearningforTextCategorization,62(1),55-62.
2. Androutsopoulos, I., Koutsias, J., Chandrinos, K. V., Paliouras, G., & Spyropoulos, C. D. (2000). An evaluation of naiveBayesiananti-spam filtering.ProceedingsoftheWorkshopon Machine Learning in the New Information Age, 1(1-3), 9-17.
3. Drucker, H., Wu, D., & Vapnik, V. N. (1999). Support vector machines for spam categorization. IEEE Transactions on Neural Networks, 10(5), 1048-1054.
4. Bao, X., Zhang, Y., Li, Q., Yang, Z., & Zhang, W. (2014). A hybrid approach to spam email detection using improved ant colony optimization and support vector machine. Information Sciences, 277, 495-511.
5. Su, C. M., Li, W. J., & Lee, C. C. (2015). An ensemble-based classifier for email spam detection using weighted majority voting. Knowledge-Based Systems, 80, 136-145.
6. Selvi, S. T., & Radhika, R. (2017). Ensemble classifier with modified random forest for spam email detection. Journal of Computational Science, 21, 108-116.
7. Sun, J., Ma, J., Zeng, D., & Li, H. (2017). Email spam detection using a hybrid machine learning method. Information Processing & Management, 53(2), 427-437.
8. Le, H. V., Nguyen, M. T., & Nguyen, T. T. (2018). Email spam detection based on ensemble learning of extreme learning machine. International Journalof MachineLearningand Cybernetics, 9(4), 591-602.
9. Poon,C. K.,& Domingos,P.(2009).Unsupervised spam detection using coherence propagation.Proceedingsof the 12th International Conference on Artificial Intelligence and Statistics (AISTATS), 18, 465-472.
10. Perez-Macias, J. M., Araujo, L., & Travieso-Gonzalez, C. M. (2012). On the use of machine learning techniques for email spam filtering. Expert Systems with Applications, 39(10), 9570-9576.
11. Yang, L., & Zhang, L. (2012). A linear programming approach for emailspamdetection. Knowledge-Based Systems, 26, 151-159.
12. Almeida, T. A., Gómez, H. F., & Yamakami, A. (2010). Contributions to the study of SMS spam filtering: New collection and results. Journal of Machine Learning Research, 11, 3611-3628.
13. Carreras, X., & Marquez, L. (2001). Boosting trees for anti-spam email filtering. Proceedingsof the Conference on Recent Advances in Natural Language Processing, 9-15.
14. Biggio, B., Fumera, G., & Roli, F. (2011). Multiple classifier systems for robust classifier design in adversarial environments. International Journal of Machine Learning and Cybernetics, 2(4), 249-260.
15. Perdisci, R., Lee, W., & Feamster, N. (2006). Behavioural clustering of HTTP-based malware and signature generation using malicious network traces. Proceedings of the 15thUSENIX Security Symposium, 195-210.
16. Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. Journal of Artificial Intelligence Research, 16, 321-357.
17. Bostrom,R.P.,& Sandberg,A. (2009).Combiningclassifierswithmetadecisiontrees.
18. DataMiningandKnowledgeDiscovery,19(1),63-83.
19. Batista, G. E., Prati, R. C., & Monard, M. C. (2004). A study of the behavior of severalmethods for balancing machine learning training data. ACM SIGKDD Explorations Newsletter, 6(1), 20-29.
20. Potharaju, R., & Liu, H. (2009). Predicting dynamic properties of emerging threats using transfer learning. Proceedingsof the15th ACM SIGKDD InternationalConference on KnowledgeDiscovery and Data Mining,1245-1254.
21. Pinto,A.,Torgo,L.,&Soares,C.(2009).Supervisedclusteringforidentifyingrare events:Acase-studyonnetwork intrusion detection. Proceedings of the 15th ACMSIGKDD International Conference on Knowledge Discovery and Data Mining, 1009-1018.