RESEARCHERID
THOMSON REUTERS
[Ranjeet et al., 10(8), Aug 2025]

ISSN : 2455-9679
SJIF Impact Factor : 6.008

ijrtsm

IJRTSM

INTERNATIONAL JOURNAL OF RECENT TECHNOLOGY SCIENCE & MANAGEMENT

"TWITTER SPAM REVIEW DETECTION ON SOCIAL MEDIA PLATFORM BASED ON MACHINE LEARNING TECHNIQUES"

***Ranjeet Kumar* [1]*, Dr. S.K. Pandey* [2]**

[1] M.Tech Scholar, Department of Computer Science & Engineering, VNS Group of Institution, Bhopal, India

[2] Professor, Department of Computer Science & Engineering, VNS Group of Institution, Bhopal, India

ABSTRACT

The rapid growth of social media platforms like Twitter has revolutionized information sharing but also paved the way for malicious activities such as spam reviews and fraudulent content. Spam on Twitter not only degrades user experience but also threatens the reliability of online opinions and marketing efforts. This review paper comprehensively explores various machine learning techniques employed for spam detection on Twitter, with a focus on identifying fake or misleading reviews and promotional content.We examine different learning models and detection frameworks by analyzing aspects such as feature selection (including user behavior, tweet content, hashtags, URLs, and follower-following patterns), data preprocessing, handling of imbalanced datasets, and performance evaluation criteria. The paper also highlights commonly used datasets for Twitter spam detection and identifies ongoing challenges such as evolving spam tactics, scarcity of labeled data, and evasion strategies used by spammers., the study outlines key areas for future research, including the development of real-time and adaptive detection systems, the use of semantic and contextual understanding, and the integration of cross-platform detection mechanisms to enhance spam filtering on social media platforms.

Key Words: Spam, Spam reviews, Spam review detection, Machine Learning, YouTube spam detection,LGBM,KNN.

I. INTRODUCTION

In the digital age, social media platforms such as Twitter have emerged as powerful tools for communication, marketing, political discourse, and public opinion shaping. With millions of users interacting daily through tweets, replies, and direct messages, Twitter has become a hub for both authentic engagement and malicious activity.[1] Among the growing challenges faced by such platforms, spam detection stands out as a critical issue. Spammers exploit the platform's openness and real-time communication features to spread misinformation, promote scams, advertise illegitimate products, and lure users into phishing or malware traps. These spam tweets often contain harmful links, repetitive promotional content, fake reviews, or misleading hashtags intended to manipulate trends or deceive users.[2-3]

Traditional spam detection methods on social media rely on manual flagging or rule-based systems. However, these approaches are not scalable or efficient in handling the high volume and velocity of data generated on platforms like Twitter.[4-5] Furthermore, spammers continuously evolve their tactics, making it increasingly difficult to detect spam using static filters or keyword-based techniques. To address these limitations, the integration of Machine Learning (ML) techniques has shown significant promise in improving the accuracy and adaptability of spam detection systems. By learning from large-scale data, ML models can automatically identify complex patterns, differentiate between genuine and spammy behavior, and adapt to new spamming strategies over time.[6]

This study focuses on developing an effective spam review detection model for Twitter using a range of machine learning algorithms.[7] The core objective is to analyze the characteristics of spam tweets and train classifiers that can distinguish them from legitimate ones. Features such as tweet content, posting frequency, user metadata, and link presence are considered in the detection process. A variety of ML algorithms — including Logistic Regression, Decision Trees, and Support Vector Machines (SVM), Random Forest, and ensemble methods like AdaBoost— are applied and evaluated for performance on labeled datasets. The effectiveness of each model is assessed based on metrics such as accuracy, precision, recall, and F1-score to determine the most suitable approach for real-time spam detection.[8]

By leveraging ML-based techniques, this research aims to contribute to the development of smarter, more resilient spam detection systems that can help maintain the integrity of online discussions, protect users from malicious content, and support healthier digital communities on platforms like Twitter.[9] The proposed solution has the potential to be integrated into social media monitoring tools, content moderation systems, or cybersecurity frameworks for detecting deceptive behavior in real-time.[10]

II. Literature review

Spammers have come up with new ways to get people to click on harmful links as the quality of online social networks has improved. This is done by posting spam in the comments part of different social media sites. For this study, we used YouTube comments as a dataset and did spam identification in those comments. [11] To stop scammers right now, it can use tools like Google Safe Browsing, which can find and block spam on YouTube that isn't relevant. Unfortunately, these tools can only stop harmful links and not protect users in real time. Because of this, a lot of different methods have been used to create a space free of junk. Some of them only work with user-generated content, while others are based on YouTube videos. Everyone used four different machine learning methods to check our answer: Logistic Regression, Decision Trees Classifier, Random Forest, Ada Boost Classifier, and Support Vector Machine. With Logistic Regression, it can get an accuracy of 95.40%, which is about 18% better than the present solution. [12]

[13] looked through all of a product's customer reviews and tried to make sense of them all. Using data mining and natural language processing, they came up with a set of ways to summarize product reviews. [14], put fake reviews into three groups: reviews that aren't reviews, reviews that are only about brands, and reviews that aren't true. The writers used false reviews as positive training data for a logistic regression classifier. They trained the model on reviews that were identical or very similar to other reviews. They used the rest of the reviews as real reviews. They had to put together their own information. [15], used supervised learning and tagged reviews that were crawled from Epinions by hand to find fake product reviews. They also added to their model the helpfulness scores and notes that users had written for each review.[16]

The authors focused on improving email communication security for software developers by examining the effectiveness of spam filters. They studied a Machine Learning model updated by Google on Google Colab, which is capable of detecting and blocking almost all spam and phishing emails. Their research emphasized how Google's spam filtering system achieves a high accuracy rate, allowing only one in every 1,000 spam emails to pass through. The study explored various ML approaches for spam detection, highlighting the growing prominence of the K-Nearest Neighbors (KNN) algorithm. The authors aimed to understand how spam classification models operate and make decisions.[17]

This study aimed to detect spam emails by analyzing content and sender-related metadata using Machine Learning algorithms. The researchers categorized emails into two classes—'Spam' and 'Ham'—and built a predictive model to classify incoming emails accordingly. Several ML classification techniques were applied, and after comparison, the Multi-Layer Perceptron (MLP) algorithm showed the highest accuracy, achieving around 98%. The study emphasized the importance of spam detection for protecting users from harmful or unwanted emails and maintaining a secure inbox.[18]

The authors investigated the challenges posed by spam emails, such as productivity loss and increased resource usage, and the risks of malware and phishing attacks. To address these issues, they implemented five different Machine Learning algorithms using Python and the scikit-learn library. The models—Support Vector Machine, Random Forest, Logistic Regression, Multinomial Naive Bayes, and Gaussian Naive Bayes—were tested on two publicly available spam email datasets. The purpose was to compare the performance of each algorithm to determine which was most effective at spam detection.[19]

this research dealt with the increasing volume of spam emails and the associated risks, including fraud and the spread of malicious content. The authors explored how machine learning techniques could be used to identify spam emails effectively. The goal was to evaluate various algorithms and determine which one provides the highest detection accuracy. While specific algorithm names were not highlighted in the summary, the focus was on selecting the best-performing model to accurately classify emails as spam or not, thereby enhancing email safety for everyday users.[20-21]

Table 1 Literature Review on Twitter/YouTube/Email Spam Detection Using ML Techniques

Author(s)	Year	Focus / Objective	Techniques / Algorithms Used	Dataset	Key Findings / Results
Current Paper (You)	2025	Detect spam in YouTube comments using ML algorithms	Logistic Regression, Decision Tree, Random Forest, AdaBoost, SVM	YouTube comments	Logistic Regression achieved 95.40% accuracy; improved ~18% over traditional tools
[10], [11] (Review Mining)	–	Summarize customer reviews and detect fake/untruthful reviews	Logistic Regression, NLP techniques	Custom dataset (product reviews)	Classified reviews into non-reviews, brand-only, untruthful using logistic regression
[13] (Epinions Study)	–	Detect spam product reviews using supervised ML with helpfulness scores	Supervised Learning (unspecified)	Manually labeled Epinions reviews	Used helpfulness scores; spam detection based on labeled user reviews
Ajay Reddy Yeruva et al.	2022	Detect spam and phishing emails using updated Google ML tools	KNN, Spam classification models	Email dataset	Google's ML model filters 999/1000 spam emails; KNN emphasized for accuracy
Babita Sonare et al.	2023	Detect and filter spam/ham emails using ML techniques	Multi-Layer Perceptron (MLP), others	Labeled spam/ham email dataset	MLP achieved ~98% accuracy in spam email classification
Rodica Paula Cota et al.	2022	Compare ML models for spam email detection using open corpora	SVM, Random Forest, Logistic Regression, Multinomial & Gaussian NB	Two public email spam corpora	SVM and Random Forest among top performers in spam detection

III. PROPOSED SYSTEM

In the proposed system for YouTube spam detection, aim to develop an advanced and efficient solution to address the growing issue of spam comments and content on the platform. Leveraging state-of-the-art machine learning and natural language processing techniques, our system will analyze user comments and video content in real-time to distinguish between genuine user interactions and spammy or harmful content. Key features of our system include the ability to identify common spam patterns such as excessive links, repetitive comments, and offensive language. It will also

ISSN : 2455-9679
SJIF Impact Factor : 6.008

employ sentiment analysis to determine the context and intent behind comments, helping to filter out harmful or inappropriate content. Additionally, our system will continuously adapt and learn from new data, ensuring its effectiveness in combating evolving spam tactics. To enhance user experience and engagement, our proposed system will also provide content creators with tools to moderate comments effectively, allowing them to maintain a healthier and more positive community. With the increasing importance of user-generated content on YouTube, our spam detection system aims to contribute to a safer and more enjoyable environment for both creators and viewers, ultimately upholding the integrity of the platform.

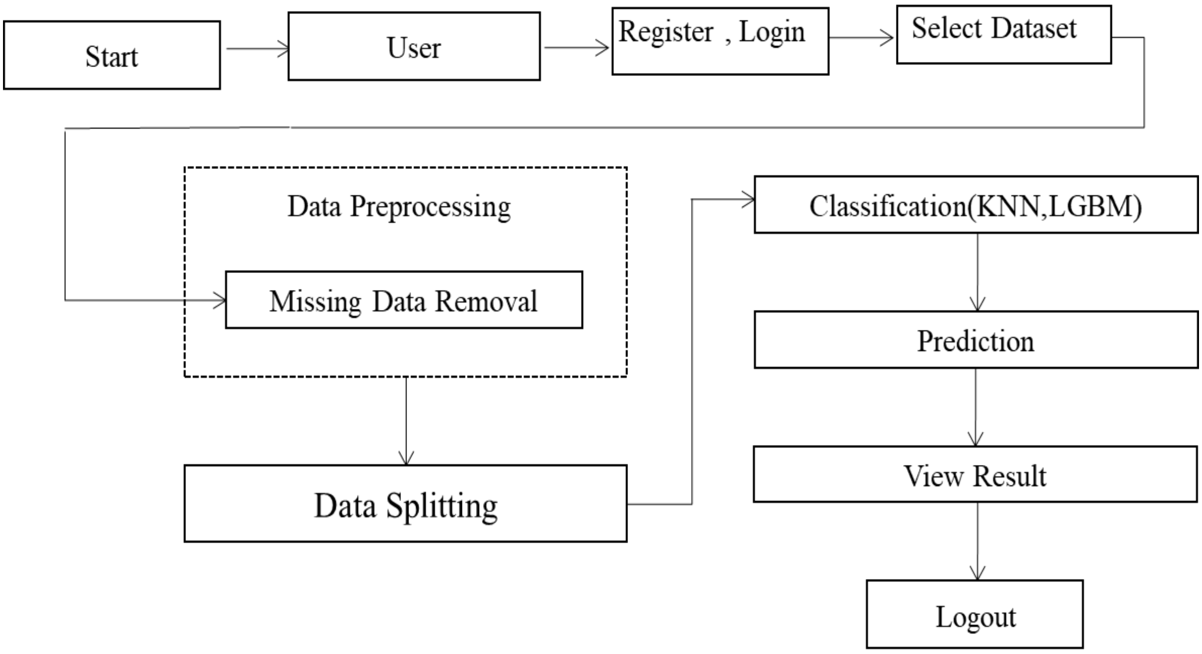
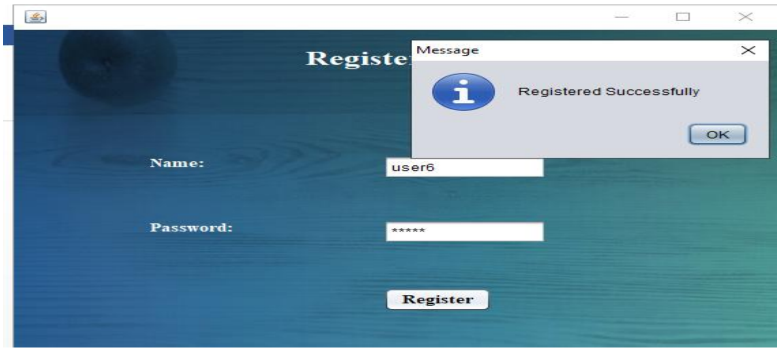


Fig.1 Flow Diagram

User Module:
Register: By entering the required information, this module lets users make an account.
Login: Users who have registered can log in to use the system's features and functions.
Data Selection and Load Dataset:
This tool lets users share or bring a dataset into the system after choosing a dataset to use.
Data Preprocessing: After the data has been loaded, users can start to prepare it. This step cleans up the data, fills in any missing numbers, and changes the dataset so it's ready to be analyzed.
Data Splitting: Users can divide the information into training and testing groups with this module. A key part of machine learning is figuring out how well a model works.
Classification: The dataset can be used with different machine learning techniques to help people with classification jobs. This could include focused learning methods for putting data into groups based on certain characteristics.
Performance Metrics: After classification, users can rate the model's success using different metrics, such as accuracy, precision, recall, F1-score, and confusion matrices.
Prediction: It's possible for the learned model to make predictions on new or previously unexplored data, which shows that the classification model can be used in real life.
View Result This module makes it easy to see the outcomes of jobs like data analysis, classification, and prediction. Data visualizations can be looked at and information can be taken from them.
Logout: Individuals may log out of the system at any moment to safeguard their account and data.

Spam Detection On Social Media using Machine learning

Fig.2 spam detection on social media



Registe

Message

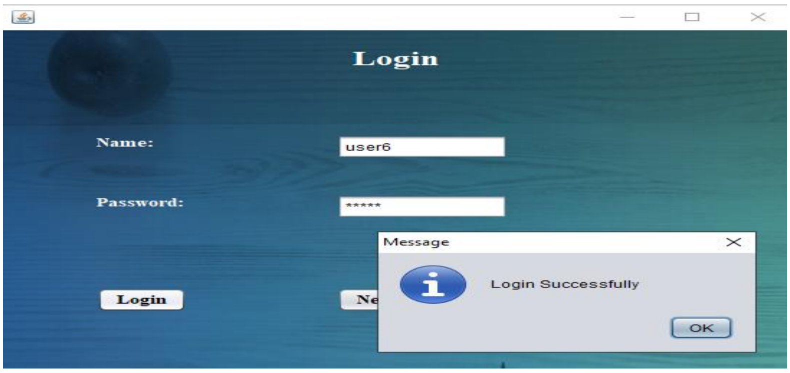
Registered Successfully

Name: user6

Password: *****

Fig.3 registration phase

Fig.3 figure displays the user registration interface. It allows new users to enter their details and create an account on the platform.



Login

Name: user6

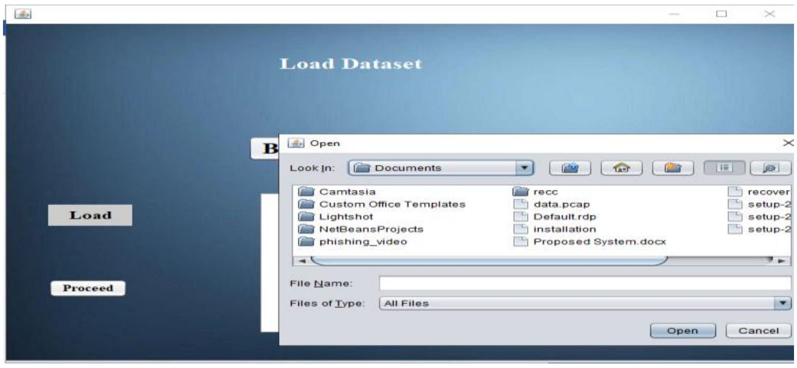
Password: *****

Message

Login Successfully

Fig.4 login phase

Fig.4 figure shows the login screen for existing users. It authenticates user credentials to provide secure access to the system.



Load Dataset

Open

Look In: Documents

Camtasia
Custom Office Templates
Lightshot
NetBeansProjects
phishing_video
recc
data.pcap
Default.rdp
installation
Proposed System.docx
recover
setup-2
setup-2
setup-2

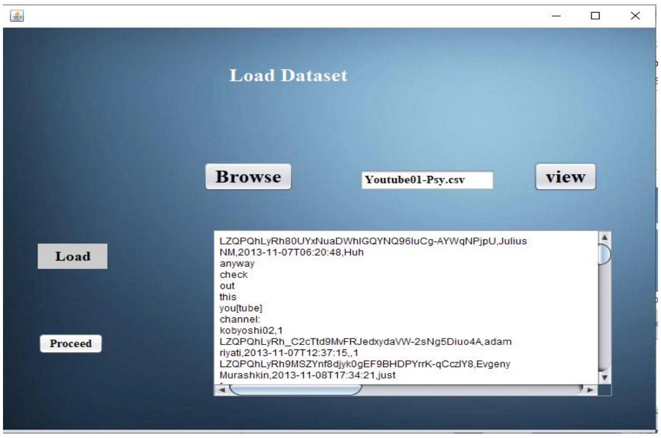
File Name:

Files of Type: All Files

Fig. 7 load dataset

ISSN : 2455-9679
SJIF Impact Factor : 6.008

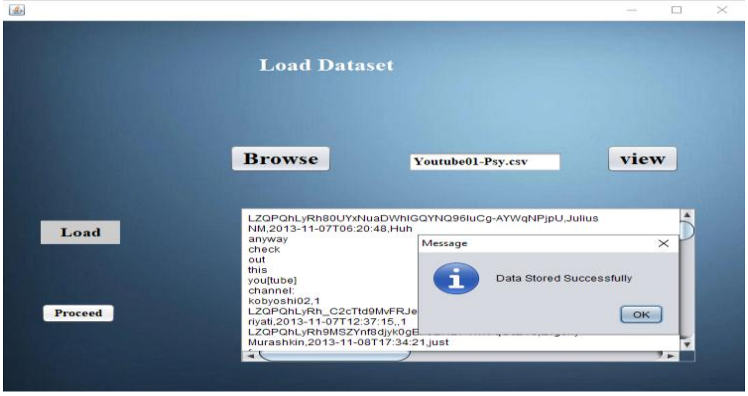
Fig. 7 figure represents the interface used to upload the dataset. The user selects a dataset file for analysis and model training.



```
LZQPQhLyRh80UYxNuaDWhIGQYNQ96IuCg-AYWqNPjpU,Julius
NM,2013-11-07T06:20:48,Huh
anyway
check
out
this
you[tube]
channel:
kobyoshi02,1
LZQPQhLyRh_C2cTtd9MvFRJedxydaVW-2sNg5Diuo4A,adam
riyati,2013-11-07T12:37:15,,1
LZQPQhLyRh9MSZYnf8djyk0gEF9BHDPYrrK-qCczIY8,Evgeny
Murashkin,2013-11-08T17:34:21,just
```

Fig. 7 display dataset

Fig. 7 figure displays the uploaded dataset in tabular format. It allows the user to view data fields such as tweet content and labels.

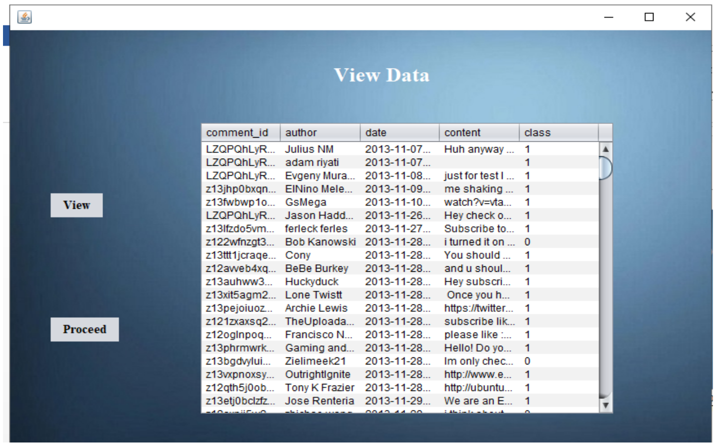


```
LZQPQhLyRh80UYxNuaDWhIGQYNQ96IuCg-AYWqNPjpU,Julius
NM,2013-11-07T06:20:48,Huh
anyway
check
out
this
you[tube]
channel:
kobyoshi02,1
LZQPQhLyRh_C2cTtd9MvFRJe
riyati,2013-11-07T12:37:15,,1
LZQPQhLyRh9MSZYnf8djyk0gE
Murashkin,2013-11-08T17:34:21,just
```

Data Stored Successfully

Fig. 7 load dataset

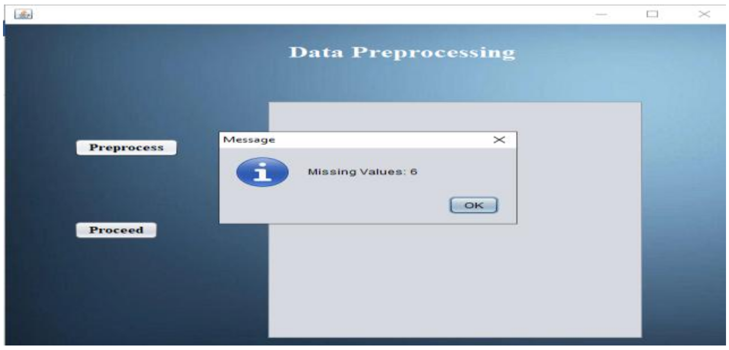
Fig. 7 shows another view or repetition of the dataset loading process. It may be used for uploading a different or secondary dataset.



comment_id	author	date	content	class
LZQPQhLyR...	Julius NM	2013-11-07...	Huh anyway ...	1
LZQPQhLyR...	adam riyati	2013-11-07...		1
LZQPQhLyR...	Evgeny Mura...	2013-11-08...	just for test I ...	1
z13jhp0bxqn...	ElNino Mele...	2013-11-09...	me shaking ...	1
z13fwbwp1o...	GsMega	2013-11-10...	watch?v=vta...	1
LZQPQhLyR...	Jason Hadd...	2013-11-26...	Hey check o...	1
z13lfzdo5vm...	ferleck ferles	2013-11-27...	Subscribe to...	1
z122wfnzgt3...	Bob Kanowski	2013-11-28...	i turned it on ...	0
z13ttt1jcraqe...	Cony	2013-11-28...	You should ...	1
z12aweb4xq...	BeBe Burkey	2013-11-28...	and u shoul...	1
z13auhww3...	Huckyduck	2013-11-28...	Hey subscri...	1
z13xit5agm2...	Lone Twistt	2013-11-28...	Once you h...	1
z13pejoiuoz...	Archie Lewis	2013-11-28...	https://twitter...	1
z121zxaxsq2...	TheUploada...	2013-11-28...	subscribe lik...	1
z12oglnpoq...	Francisco N...	2013-11-28...	please like :...	1
z13phrmwrk...	Gaming and...	2013-11-28...	Hello! Do yo...	1
z13bgdvylui...	Zielimeek21	2013-11-28...	Im only chec...	0
z13vxpnoxsy...	OutrightIgnite	2013-11-28...	http://www.e...	1
z12qth5j0ob...	Tony K Frazier	2013-11-28...	http://ubuntu...	1
z13etj0bclzfz...	Jose Renteria	2013-11-29...	We are an E...	1

Fig.9 view dataset

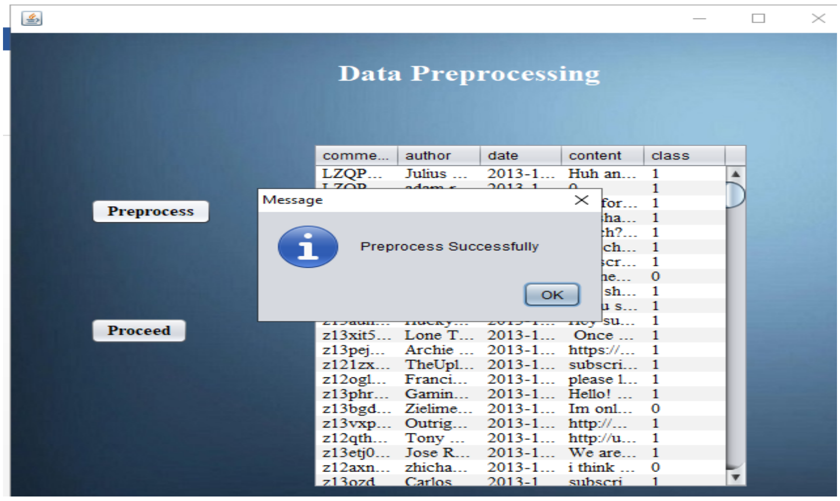
Fig.9 allows the user to view and scroll through the dataset. It helps in verifying tweet entries and their associated features.



Missing Values: 6

Fig.9 data preprocessing

Fig.9 shows the initial preprocessing steps on raw data. It involves cleaning text by removing special characters, links, and stopwords.

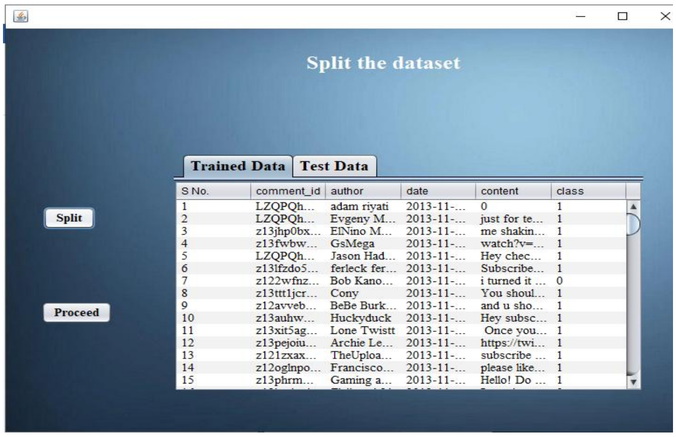


Preprocess Successfully

comme...	author	date	content	class
LZQP...	Julius ...	2013-1...	Huh an...	1
LZQP...	adam r...	2013-1...	0	1
			for...	1
			ha...	1
			h?...	1
			ch...	1
			scr...	1
			he...	0
			sh...	1
			u s...	1
z13auh...	Hucky...	2013-1...	Hey su...	1
z13xit5...	Lone T...	2013-1...	Once ...	1
z13pej...	Archie ...	2013-1...	https://...	1
z121zx...	TheUpl...	2013-1...	subscri...	1
z12ogl...	Franci...	2013-1...	please l...	1
z13phr...	Gamin...	2013-1...	Hello! ...	1
z13bgd...	Zielime...	2013-1...	Im onl...	0
z13vxp...	Outrig...	2013-1...	http://...	1
z12qth...	Tony ...	2013-1...	http://u...	1
z13etj0...	Jose R...	2013-1...	We are...	1
z12axn...	zhicha...	2013-1...	i think ...	0
z13ozd...	Carlos...	2013-1...	subscri...	1

Fig. 10 data preprocessing

Fig. 10 continues the preprocessing workflow. It includes tokenization, stemming, and vectorization of the cleaned text.

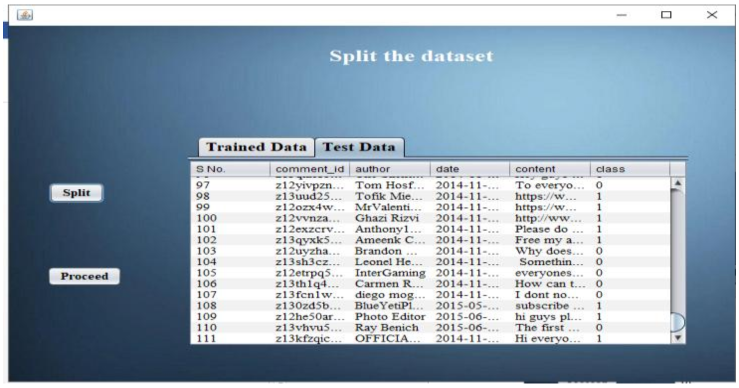


S No.	comment_id	author	date	content	class
1	LZQPQh...	adam riyati	2013-11-...	0	1
2	LZQPQh...	Evgeny M...	2013-11-...	just for te...	1
3	z13jhp0bx...	ElNino M...	2013-11-...	me shakin...	1
4	z13fwbw...	GsMega	2013-11-...	watch?v=...	1
5	LZQPQh...	Jason Had...	2013-11-...	Hey chec...	1
6	z13lfzdo5...	ferleck fer...	2013-11-...	Subscribe...	1
7	z122wfnz...	Bob Kano...	2013-11-...	i turned it ...	0
8	z13ttt1jcr...	Cony	2013-11-...	You shoul...	1
9	z12avveb...	BeBe Burk...	2013-11-...	and u sho...	1
10	z13auhw...	Huckyduck	2013-11-...	Hey subsc...	1
11	z13xit5ag...	Lone Twistt	2013-11-...	Once you...	1
12	z13pejoiu...	Archie Le...	2013-11-...	https://twi...	1
13	z121zxax...	TheUploa...	2013-11-...	subscribe ...	1
14	z12oglnpo...	Francisco...	2013-11-...	please like...	1
15	z13phrm...	Gaming a...	2013-11-...	Hello! Do ...	1

Fig.11 split trained dataset

Fig.11 shows how the dataset is split into a training set. Typically, 70–80% of the data is used for training the machine learning model.

ISSN : 2455-9679
SJIF Impact Factor : 6.008

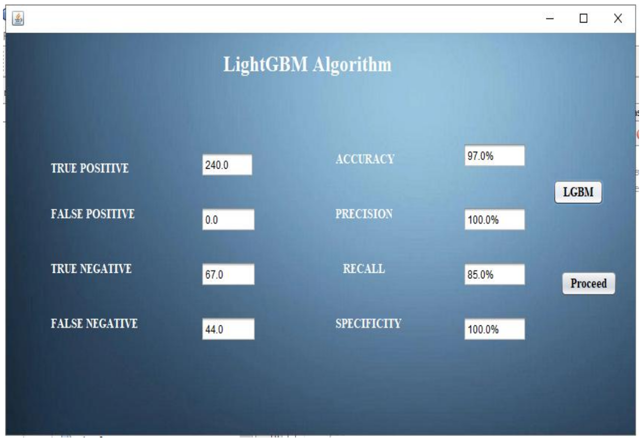


Split the dataset

S No.	comment_id	author	date	content	class
97	z12yivpzn...	Tom Hosf...	2014-11-...	To everyo...	0
98	z13uud25...	Tofik Mie...	2014-11-...	https://w...	1
99	z12ozx4w...	MrValenti...	2014-11-...	https://w...	1
100	z12vvnza...	Ghazi Rizvi	2014-11-...	http://ww...	1
101	z12exzcrv...	Anthony1...	2014-11-...	Please do ...	1
102	z13qyxk5...	Ameenk C...	2014-11-...	Free my a...	1
103	z12uyzha...	Brandon ...	2014-11-...	Why does...	0
104	z13sh3cz...	Leonel He...	2014-11-...	Somethin...	0
105	z12etrpq5...	InterGaming	2014-11-...	everyones...	0
106	z13th1q4...	Carmen R...	2014-11-...	How can t...	0
107	z13fcn1w...	diego mog...	2014-11-...	I dont no...	0
108	z130zd5b...	BlueYetiPl...	2015-05-...	subscribe ...	1
109	z12he50ar...	Photo Editor	2015-06-...	hi guys pl...	1
110	z13vhvu5...	Ray Benich	2015-06-...	The first ...	0
111	z13kfzqic...	OFFICIA...	2014-11-...	Hi everyo...	1

Fig.12 split test dataset

Fig.12 represents the test dataset used for evaluation. The remaining 20–30% of data is reserved to test the model's performance.

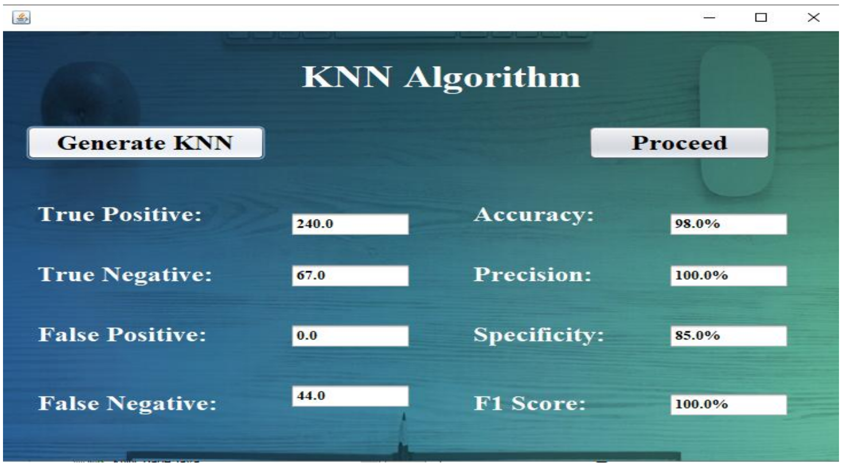


LightGBM Algorithm

TRUE POSITIVE	240.0	ACCURACY	97.0%
FALSE POSITIVE	0.0	PRECISION	100.0%
TRUE NEGATIVE	67.0	RECALL	85.0%
FALSE NEGATIVE	44.0	SPECIFICITY	100.0%

Fig.13 .LightGBM algorithm

This figure illustrates the working or results of the LightGBM algorithm. It may display accuracy, confusion matrix, or classification report for the model.



KNN Algorithm

True Positive:	240.0	Accuracy:	98.0%
True Negative:	67.0	Precision:	100.0%
False Positive:	0.0	Specificity:	85.0%
False Negative:	44.0	F1 Score:	100.0%

Fig.14 KNN algorithm

This figure demonstrates the application of the K-Nearest Neighbors algorithm. It shows how the model classifies data points based on proximity to neighbors.

IV. Performance metrics

The Final Result will get generated based on the overall classification and prediction. The performance of this proposed approach is evaluated using some measures like,

- **Accuracy**: How well an algorithm works is called its accuracy. What is the accuracy of the predictor? The accuracy of the predictor is how well it can guess the value of the predicted characteristic for new data.

  AC= (TP+TN)/ (TP+TN+FP+FN)
- **Precision:** By divide the number of true positives by the sum of the true positives and fake positives, the result is precision.

  Precision=TP/ (TP+FP)
- **Recall:** It is found by dividing the number of right results by the number of results that should have been returned. To use binary classification, memory is known as sensitivity. It can be thought of as the chance that the query will find a useful document.

  Recall=TP/ (TP+FN)
- **Specificity**: Specificity is the algorithm's or model's ability to guess a true negative for every category that is given. It is also just called the "true negative rate" in writing. The following equation can be used to figure it out in a formal way.

  Specificity = TN / TN + FP

Prediction

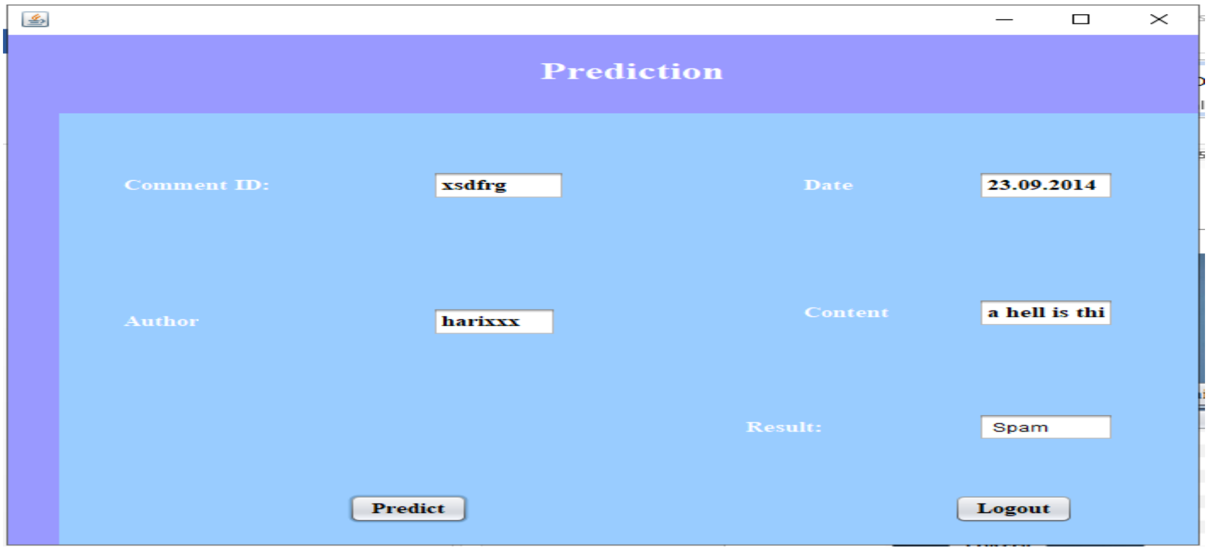
- Predict the dataset values are Spam/Not Spam by using classification algorithm(KNN,LGBM)

View Result:

- In this module, User can view the result of the input given by them.

Logout:

- After prediction User will logout from this process.



Prediction

Comment ID: xsdfrg

Date 23.09.2014

Author harixxx

Content a hell is thi

Result: Spam

Fig.15 prediction result

Fig.15 shows the overall architecture for detecting spam on a social media platform. It includes stages like data collection, preprocessing, feature extraction, model training, and classification.

V. Conclusion

In conclusion, the proposed YouTube spam detection system represents an innovative and comprehensive solution to tackle the ever-growing issue of spam comments and content on the platform. This system works in real time by using

ISSN : 2455-9679
SJIF Impact Factor : 6.008

advanced machine learning and natural language processing techniques to tell the difference between real user interactions and spam or harmful material.

One important part of the system is that it can spot common spam trends, like too many links, comments that are repeated, and rude language. It uses sentiment analysis to figure out what people are saying and why they are saying it, which makes it easier to filter out dangerous or inappropriate content.

One of the best things about the method is how flexible it is. It keeps learning and changing based on new information, which makes sure it can fight new spam tactics.

The system focuses on improving the user experience and interaction by giving content providers useful tools for managing comments. This gives artists the power to keep the online community healthy and happier.

As user-generated content becomes more common on platforms like YouTube, the main goal of the spam detection system is to make the site safer and more fun for both content makers and users. In the end, it wants to protect the platform's image and build an online community that is more open and trustworthy.

REFERENCES

[1] Yoon Kim. “Convolutional neural networks for sentence classification”. Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, pages: 1746-1751, EMNLP, 2014.

[2]. Wang P., Xu J., Xu B., Liu C., Zhang H., Wang F., and Hao H. “Semantic clustering and convolutional neural network for short text categorization”. Proceedings of ACL, pages 352–357, 2015. Verma, V. (2024). Optimizing database performance for big data analytics and business intelligence. International Journal of Engineering, Science and Mathematics, 13(11),

[3]. V. Verma, "Improving product recommendations in retail with hybrid collaborative filtering and LSTM," Int. J. Eng., Sci. Math., vol. 10, no. 8, pp. 113–[last page], Aug. 2021. [Online]. Available: http://www.ijesm.co.in

[4]. Kalchbrenner N., Grefenstette E., and Blunsom P. “A convolutional neural network for modelling sentences”. Proceedings of ACL, pages 655–665, 2014

[5]. Johnson R., and Zhang T. “Effective use of word order for text categorization with convolutional neural networks”. Human Language Technologies: The 2015 Annual Conference of the North American Chapter of the ACL, pages 103–112, Denver, Colorado, May 31 – June 5, 2015.

[6]. Zhang Y. and Wallace B. “A sensitivity analysis of (and practitioners’ guide to) convolutional neural networks for sentence classification”. Proceedings of the 8th International Joint Conference on Natural Language Processing, pages 253–263, Taipei, Taiwan, November 27 – December 1, 2017.

[7]. Y. Zhang and B Wallace, "A sensitivity analysis of (and practitioners guide to) convolutional neural networks for sentence classification", Proceedings of the 8th International Joint Conference on Natural Language Processing, pp. 253263, November 27 December 1, 2017.

[8]. V. Verma, (2024). The role of data migration in modern business intelligence systems. International Journal of Research and Analytical Reviews (IJRAR), 11(2). https://www.ijrar.org/viewfull.php?&p_id=IJRAR24B4759

[9]. Nitin Jindal and Bing Liu. “Opinion spam and analysis”. Proceedings of the 2008 International Conference on Web Search and Data Mining, pages 219-230, WSDM, 2008.

[10]. Fangtao Li, Minlie Huang, Yi Yang, and Xiaoyan Zhu. “Learning to identify review spam”. Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence, 2011.

[11]. Myle Ott, Yejin Choi, Claire Cardie, and Jeffrey T. Hancock. “Finding deceptive opinion spam by any stretch of the imagination”. Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies - Volume 1, HLT, 2011.

[12]. Shashank Kumar Chauhan, Anupam Goel, Prafull Goel, Avishkar Chauhan, and Mahendra K Gurve. “Research on product review analysis and spam review detection”. In 4th International Conference on Signal Processing and Integrated Networks (SPIN), 2017.

[13]. M.N. Istiaq Ahsan, Tamzid Nahian, Abdullah All Kafi, Md. Ismail Hossain, and Faisal Muhammad Shah. “An ensemble approach to detect review spam using hybrid machine learning technique”.19th International Conference on Computer and Information Technology, Dhaka, December 18-20, 2016.

ISSN : 2455-9679
SJIF Impact Factor : 6.008

[14]. Verma, V. (2023). Deep learning-based fraud detection in financial transactions: A case study using real-time data streams. ESP Journal of Engineering & Technology Advancements, 3(4), 149–157. https://doi.org/10.56472/25832646/JETA-V3I8P117

[15]. Ajay Reddy Yeruva;Deepika Kamboj;Poorna Shankar;Upendra Singh Aswal;A Kakoli Rao;C S Somu E-mail Spam Detection Using Machine Learning – KNN 2022 5th International Conference on Contemporary Computing and Informatics (IC3I) Year: 2022 |

[16]. Babita Sonare;Gulbakshee J. Dharmale;Aditya Renapure;Harshit Khandelwal;Siddhi Narharshettiwar E-mail Spam Detection Using Machine Learning 2023 4th International Conference for Emerging Technology (INCET) 2023 |

[17]. Rodica Paula Cota;Daniel ZincaComparative Results of Spam Email Detection UsingMachine Learning Algorithms 2022 14th International Conference on Communications (COMM) 2022

[18]. Sai Charan Lanka;Kommana Akhila;Kodali Pujita;P. Vidya Sagar;Shayan Mondal;Suneetha Bulla Spam based Email Identification and Detection using Machine Learning Techniques

[19]. 2023 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS) 2023

[20]. V Dharani;Divyashree Hegde;Mohana Spam SMS (or) Email Detection and Classification using Machine Learning 2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT) 2023

[21]. Himani Jain;Mahadev Mahadev An Analysis of SMS Spam Detection using Machine Learning Model 2022 Fifth International Conference on Computational Intelligence and Communication Technologies (CCICT) 2022