International Journal of Scientific Research in Science and Technology
Available online at : www.ijsrst.com
OPEN ACCESS
Print ISSN: 2395-6011 | Online ISSN: 2395-602X
doi : https://doi.org/10.32628/IJSRST

Random Tree Classifier: A Machine Learning Spam Comment Detection on YouTube

Mr. N Venkatramana[1], C. Jashnavi[2], J. Jithesh Guptha[2], P. Chandra Sekhar[2], M. Chandra[2]
[1]Assistant Professor, Department of CSE, Faculty, AITS, Tirupati, Andhra Pradesh, India
[2]Student, Department of CSE, Annamacharya Institute of Technology and Sciences, Tirupati, Andhra Pradesh, India

ARTICLEINFO

Article History:
Accepted: 03 March 2024
Published : 28 March 2024

Publication Issue :
Volume 11, Issue 11
March-April-2024

Page Number :
311-317

ABSTRACT

In this paper, we propose a machine learning model for accurately classifying spam comments on youtube. The project aims to develop a machine learning model for detecting spam comments on YouTube videos. The objective of this project is to develop a machine learning based system for accurately detecting and filtering out spam comments on youtube videos, enhancing user experience by promoting genuine interactions and safeguarding the platform's content quality and integrity. One of the commonly employed approaches in ML for spam comment detection is the Extra Tree classification algorithm. By training the algorithm on a labeled dataset of spam and non-spam comments, it can learn to recognize patterns and generalize its understanding to new, unseen comments. Ensemble methods, deep learning, and natural language processing (NLP) techniques are among the advanced ML approaches gaining attention in this domain. One crucial aspect of an effective spam detection system is its adaptability and responsiveness to emerging spam tactics.

Keywords: ML evaluation, ML techniques, Random Forest, Extra Tree, SVC, LSTM classifier.

I. INTRODUCTION

Machine Learning (ML) has revolutionized various domains, and its application to online platforms is notably transforming content moderation. One such prominent arena is the detection of spam comments on YouTube. With the exponential growth of user-generated content, ensuring a safe and engaging environment for users has become a paramount concern. ML-based spam comments detection on YouTube leverages advanced algorithms to swiftly and accurately identify and filter out undesirable and misleading comments. This

innovative approach harnesses the power of ML to automatically learn patterns and characteristics of spam comments from vast datasets.

By analyzing linguistic cues, syntactical irregularities, and user engagement metrics, ML models can distinguish between genuine and spam comments. These models are trained on diverse samples of comments,

enabling them to adapt and evolve alongside evolving spam tactics. The impact of ML-based spam comments detection is multifaceted. It not only safeguards the user experience by curbing the visibility of spam but also enhances the authenticity and credibility of discussions. By reducing the noise generated by spam, the quality of interactions among users is elevated, fostering meaningful conversations and community growth. While ML-based detection is an effective tool, continuous refinement and updates are essential to stay ahead of sophisticated spam techniques. Collaborative efforts between ML engineers and domain experts are crucial to fine-tune models, ensuring minimal false positives and negatives. As YouTube continues to evolve as a dynamic platform, Machine Learning plays a pivotal role in maintaining a vibrant, safe, and trustworthy online ecosystem.

II. RECENT WORKS

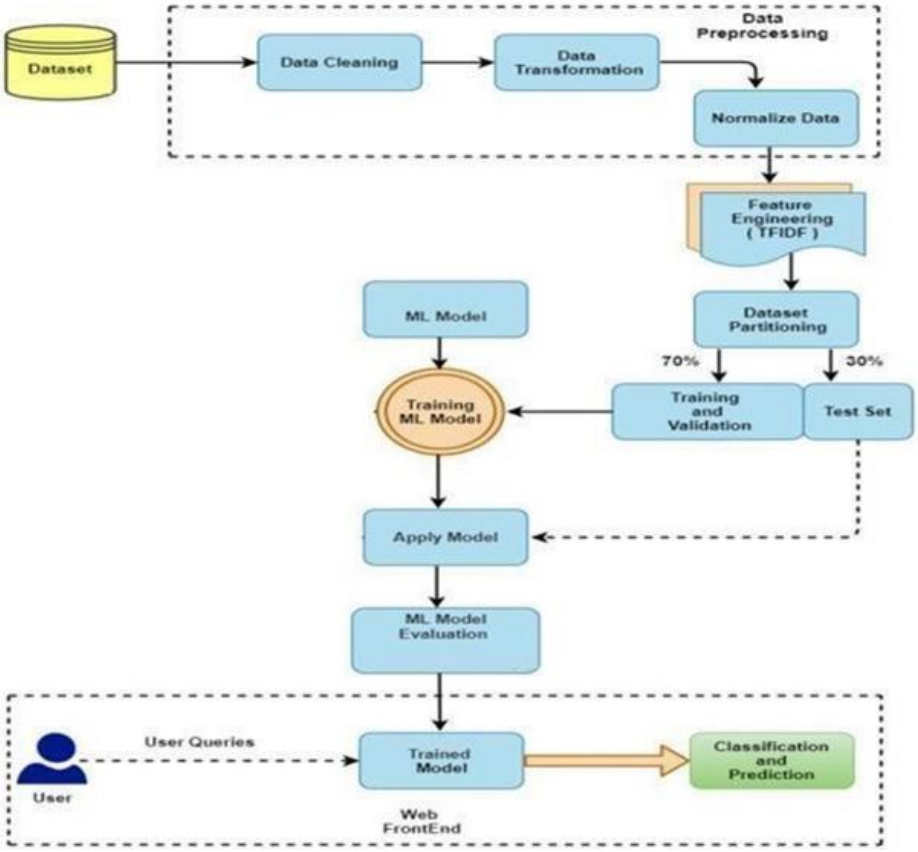
[1] Sah, U. K., & Parmar, N. (2017). An approach for Malicious Spam Detection in Email with comparison of different classifiers. In this paper, today one of the cheapest form of communication in the world is email, and its simplicity makes it vulnerable to many threats. One of the most important threats to email is spam; unsolicited email, especially when advertising agency send a mass mail. Spam email may also include malware as scripts or other executable file. . In this paper, we aim to improve detection of malicious spam through feature selection. We propose a model that employs a novel dataset for the process of feature selection, a step for improving classification in later stage. Feature selection is expected to improve training time and accuracy of malicious spam detection. This paper also shows the comparison of various classifier used during the process.
[3] Alsaleh, M., Alarifi, A., Al-Quayed, F., & Al-Salman, A. (2016). Combating comment spam with machine learning approaches. Proceedings - 2015 IEEE 14th International Conference on Machine Learning and Applications, ICMLA 2015, 295– 300. https://doi.org/10.1109/ICMLA.2015.192. The paper addresses the significant issue of comment spam within online platforms and introduces a novel approach utilizing machine learning techniques for its detection and mitigation. The authors acknowledge the growing concern of comment spam, which undermines the quality of user-generated content and affects user experience. They highlight the need for efficient and automated methods to combat this problem. The study proposes a comprehensive framework that leverages machine learning to effectively identify and filter out comment spam.
[4] Ekta Chhatar, Heeral Chauhan, Shubham Gokhale, Sompurna Mukherjee, Prof. Nikhil Jha, “Survey on Student Attendance Management System”, S.B. Jain Institute of Technology, Management and Research, Nagpur, 2016. The paper titled "Leave a Comment! An In-Depth Analysis of User Comments on YouTube," authored by Scheltus, P., Dorner, V., and Lehner, F., and published in Wirtschaftsinformatik in 2013, delves into a comprehensive exploration of user comments within the context of the popular online platform, YouTube. The authors recognize the pivotal role that user comments play in enhancing user engagement and interaction on YouTube. They undertake a detailed investigation of these comments to gain insights into the nature, patterns, and characteristics of usergenerated content in the form of comments. The study employs a rigorous analytical approach, leveraging a diverse dataset of comments sourced from YouTube. The authors discuss the methodological framework used to collect and analyze these comments, encompassing both

qualitative and quantitative techniques. Through the application of content analysis and sentiment analysis, the paper sheds light on the themes, sentiments, and trends prevalent within the comments.

III.PROPOSED WORK EXPLANATION

The planned system aims to improve the precision of detecting spam comments on YouTube by employing an amalgamation of advanced machine learning algorithms. These techniques encompass Support Vector Machine with Radial Basis Function kernel (SVM-RBF), Random Forest (RF), Extra Trees (ET), and Long Short-Term Memory (LSTM), which is a deep learning methodology. By capitalizing on the unique capabilities of these algorithms, the system intends to address the shortcomings of the Naive Bayes approach and attain superior performance in distinguishing between spam and authentic comments. This collaborative endeavor strives to enhance detection accuracy, fostering a more secure and reliable digital milieu for YouTube users. It seeks to mitigate the risks associated with fraudulent activities, offensive content, and violations of privacy.

3.1. work Flow of Proposed system



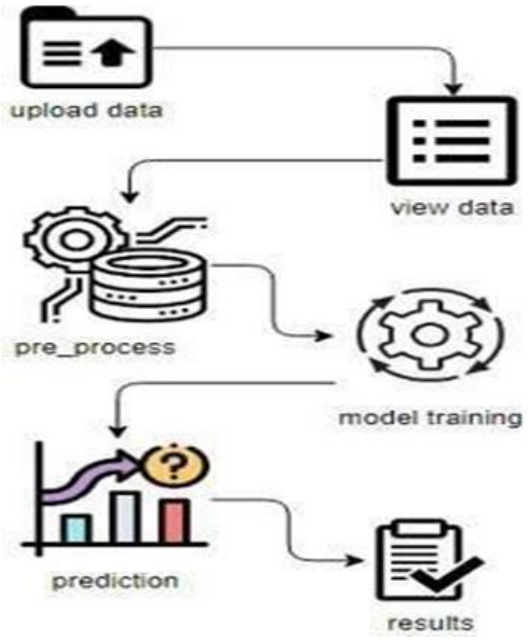
3.2. Block diagram Description:

To detect and classify YouTube comments as spam or not-spam, we have devised a comprehensive pipeline:

1. Data Cleaning: We begin by removing irrelevant and redundant data, ensuring comments are free from noise and anomalies.
2. Data Transforming: The comments are then transformed into a structured format, suitable for analysis.

3. Data Preprocessing: In data preprocessing for YouTube comments spam detection, several essential techniques are applied:
   - 1. Tokenizer: Tokenization involves breaking down comments into individual words or tokens. This process allows the algorithm to work with individual words as features, making it easier to analyze the text.
   - 2. Stemming: Stemming is the process of reducing words to their root or base form. It helps in treating words with the same root as identical, reducing the feature space and simplifying text analysis.
   - 3. Lemmatization: Lemmatization is similar to stemming but considers the context to find a word's base form. It results in more accurate word representations, preserving the meaning of the words.
   - 4. Vectorizer: Vectorization converts text data into numerical form. Techniques like TF-IDF (Term Frequency-Inverse Document Frequency) or word embeddings (e.g., Word2Vec or GloVe) are used to represent comments as numerical vectors, allowing machine learning models to process them.
4. Normalize the Data: We ensure uniformity in the dataset by scaling and normalizing the features. 5. Feature Engineering: Techniques like filling null values, label encoding, and leveraging NLP methods are employed to make the data more conducive for modeling.
5. Dataset Partitioning: The dataset is split into training (70%) and testing (30%) sets.
6. Machine Learning Modeling: Algorithms such as SVM-RBF, Random Forest, Extra Trees, and LSTM are implemented to discern patterns in the data.
7. Model Evaluation: The chosen models undergo rigorous evaluation to gauge their accuracy and efficiency.
8. Classification: The final model classifies comments into 'spam' or 'not-spam' categories.
9. Web Framework: An interactive website interface is built to showcase the project, allowing users to test the system's efficacy in real-time.

3.3. Architecture:



3.4. Algorithm

3.4.1. Support Vector machine

Support Vector Machine with Radial Basis Function kernel (SVM-RBF) is a machine learning algorithm used for classification and regression tasks. It employs a non-linear transformation to map data into a higher-dimensional space, where a hyperplane is established to maximize the margin between different classes. The Radial Basis Function kernel calculates similarity between data points, determining their influence on classification. This kernel's flexibility enables SVM-RBF to effectively handle complex, non- linear relationships in data. It's widely used for its ability to capture intricate patterns and achieve accurate results in various applications, such as image recognition, text categorization, and bioinformatics.

3.4.2. Random Forest

Random Forest is a powerful machine learning algorithm that assembles multiple decision trees to make accurate predictions. Each tree is trained on a subset of data and votes on the final prediction, resulting in improved accuracy and robustness. It mitigates overfitting and handles complex relationships in data by averaging predictions from different trees. Random Forest is versatile, handling classification and regression tasks effectively. It's widely used due to its ability to capture intricate patterns in data, making it suitable for various domains such as finance, healthcare, and image analysis. Its ensemble nature enhances generalization and makes it a popular choice for predictive modeling.

3.4.3. Extra Trees

Extra Trees, short for Extremely Randomized Trees, is an ensemble machine learning algorithm used for classification and regression tasks. It's an extension of the Random Forest method, where multiple decision trees are built using bootstrapped samples and random feature subsets. However, Extra Trees takes randomness a step further by making decisions at each split point based on random thresholds, resulting in a broader exploration of feature space. This increases diversity among trees, reducing overfitting and improving generalization. By aggregating predictions from individual trees, Extra Trees enhances accuracy and robustness, making it suitable for complex datasets and improving overall predictive performance.

3.4.4. Long Short-Term Memory

Long Short-Term Memory (LSTM) is a specialized type of recurrent neural network (RNN) architecture in deep learning. It excels in processing and retaining sequential data by utilizing memory cells with various gates to regulate information flow. LSTMs are adept at capturing longrange dependencies, making them ideal for tasks like text analysis, speech recognition, and time series prediction. The architecture's key components include input, forget, and output gates, along with a cell state that can store and control information over extended sequences. This enables LSTMs to effectively model intricate patterns and relationships within sequential data, leading to enhanced performance in various applications.

3.5. Evaluation parameter:

In machine learning, the evaluation parameters are used to assess the performance of a model on a given dataset, Common evaluation parameters included in this project are:

Accuracy:

The proportion of correctly classified instances out of the total instances. It is evaluated as Accuracy=(TP+TN) / (TP+TN+FP+FN)

Precision:

The proportion of true positive instances (correctly predicted positive) out of all instances predicted as positive. Precision = True Positives / (True Positives+False Positives)

Recall:

Recall also known as sensitivity. The proportion of true positive instances predicted as positive out of all actual positive instances.

Recall = True Positives / (True Positives+False Negatives)

F1 Score:

The harmonic mean of precision and recall, giving a balance between the two metrics.

Table:

Algorithm	Accuracy
Random Forest Classifier	94.54
Support Vector machine	90
Extra Tree Classifier	92
LSTM	64

IV. CONCLUSION

The rise of spam comments on platforms like YouTube necessitates effective solutions, and Machine Learning (ML) provides a potent tool for combating this issue. The Extra Tree classification algorithm, along with other techniques such as Random Forest, Support Vector Machines (SVC), and LSTM classifiers, showcases promising results in accurately detecting spam comments. However, the everevolving nature of spam tactics requires continuous exploration of advanced ML approaches, including ensemble methods and natural language processing (NLP). Achieving adaptability and responsiveness in spam detection systems is crucial for staying ahead of emerging threats. As researchers and developers continue to refine and innovate ML techniques, the goal remains to enhance accuracy and robustness, ensuring a safer and more engaging online environment for users.

V. REFERENCES

[1]. G. Shi, F. Luo, Y. Tang, and Y. Li, "Dimensionality reduction of hyperspectral image based on local constrained manifold structure collaborative preserving embedding," Remote Sens., vol. 13, no. 7, p. 1363, Apr. 2021, doi: 10.3390/rs13071363.

[2]. R. K. Das, S. S. Dash, K. Das, and M. Panda, "Detection of spam in Youtube comments using different classifiers," in Advanced Computing and Intelligent Engineering, 2020, pp. 201–214, doi:10.1007%2F978- 981-15-1081-6_17.

[3]. N. M. Samsudin, C. F. B. Mohd Foozy, N. Alias, P. Shamala, N. F. Othman, and W. I. S. Wan Din, "Youtube spam detection framework using Naïve Bayes and logistic regression," Indonesian J. Electr. Eng. Comput. Sci., vol. 14, no. 3, p. 1508, Jun. 2019, doi: 10.11591/ijeecs.v14.i3.pp1508- 1517.

[4]. N. Hussain, H. Turab Mirza, G. Rasool, I. Hussain, and M. Kaleem, "Predilection decoded: Spam review detection techniques: A systematic literature review," Appl. Sci., vol. 9, no. 5, p. 987, Mar. 2019, doi: 10.3390/app9050987.

[5]. S. Jain and D. M. Patel. (2019). Analyzing User Comments of Learning Videos From Youtube UsingMachineLearning.[Online].Available: http://www.ijirset.com/upload/2019/august/50_Analyzing_DJ.PDF

Author Biography

Mr. N Venkatramana, M.Tech, is Currently working as Assistant Professor in the Department of CSE, Annamacharya Institute of Technology and sciences, Tirupati .

Ms. C Jashnavi is currently pursuing B. Tech in the stream of Computer Science and Engineering from Annamacharya Institute of Technology and Sciences, Tirupati.

Mr. J Jithesh Guptha is currently pursuing B. Tech in the stream of Computer Science and Engineering from Annamacharya Institute of Technology and Sciences, Tirupati.

Mr. P Chandra Sekhar is currently pursuing B. Tech in the stream of Computer Science and Engineering from Annamacharya Instituteof Technology and Sciences, Tirupati.

Mr. M Chandra is currently pursuing B. Tech in the stream of Computer Science and Engineering from Annamacharya Institute of Technology and Sciences, Tirupati.