

# YOUTUBE SPAM COMMENT DETECTION

**Vardha Pranathi, P. Poornika, k. charishma**

Under the esteemed guidance of

**Mr. Anil Kumar**

Assistant Professor



Bachelor of Technology

Department of Information Technology

**BVRIT HYDERABAD College of Engineering for Women**

November 16, 2023

# Overview

- 1 Introduction
- 2 Literature Survey
- 3 Problem Statement
- 4 Proposed Method
- 5 Dataset
- 6 Modules
- 7 Results
- 8 Conclusion & Future Scope
- 9 References
- 10 Thank you

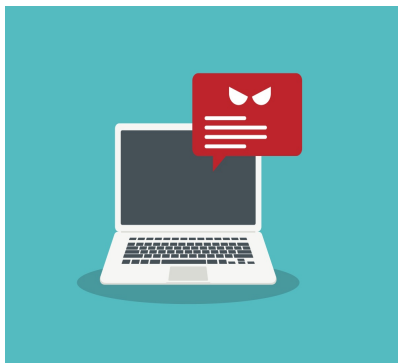
# Introduction

- **What are Spam comments?**

Spam comments are considered as repetitive, fraudulent or malicious scam that can be either in the form of Normal text, Links, Blogs, Emails.

- **Spam comments in Youtube**

YouTube is a popular social networking platform that allows users to upload, watch and comment on videos. However, YouTube also attracts spammers who post unwanted or abusive comments on videos, which can affect the quality and credibility of the content



# Literature Survey

S. No	Title of the paper	Author(s)	Description
1	A Research on YouTube Spam Comments Detection and Deletion	Darshan Bhavsar, Stephen Dcruz, Abhishek Chandekar - 2023	Developed a machine learning algorithm that can automatically identify and flag spam comments on YouTube videos using natural language processing (NLP) algorithms such as Rapid Fuzz and Levenshtein distance.
2	YouTube Spam Comments Detection Scheme Using Cascaded Ensemble Machine Learning Model	Ms.Hayoung Oh - 2022	Used 'Ensemble with Hard voting' and Ensemble with Soft voting' for spam detection and developed a machine learning model with ensemble methods.

# Literature Survey

S. No	Title of the paper	Author(s)	Description
3	Detection of Spam in YouTube Comments Using Different Classifiers	Rama Krushna Das, Sweta Shree Dash - 2021	This paper proposes a technique to detect spam comments on YouTube. YouTube spam comment screening and conducted classification experiments with six different machine learning techniques.
4	Ontological insecurity, nostalgia, and social media: Viewing YouTube videos of old TV commercials re-establishes continuity of the self over time	Mrs.Areni, Mrs.Momeni, Mr.Reynolds - 2020	Established that "Sharing own experiences in YouTube comments of old TV commercials is a way for users to cherish one's own past experiences".

# Problem Statement

- Detection of spam comments on YouTube is a common problem in content moderation and online community management and The primary goal of this project is to develop a machine learning model that can accurately detect comments on YouTube videos as spam or legitimate.

## Spam Comments Detection with Machine Learning



# Proposed Method

- The proposed method involves pre-processing of dataset which is a collection of Youtube comments. Text vectorization is then performed using TF-IDF to convert comments into numerical features. Two machine learning models, Multinomial Naive Bayes and SVM, are trained on the preprocessed data, and their performance is evaluated using various metrics. Additionally, a user-friendly GUI application is created using tkinter, allowing users to interact with the trained model for spam comment classification.

		Predicted Class	
		Spam	Non-Spam
Actual Class	Spam	TP=45	FN=20
	Non-Spam	FP=5	TN=30

# Dataset

- The dataset appears to be a collection of YouTube comments with corresponding metadata. Each row in the dataset represents a comment made on a YouTube video, and the columns provide information about the comment, the commenter, and the video.

Dataset columns	Description
COMMENT ID	A unique identifier for each comment
AUTHOR	The name or username of the person who wrote the comment
DATE	The timestamp indicating when the comment was posted
CONTENT	The actual text content of the comment
CLASS	A binary label indicating whether the comment is classified as spam(1) or non-spam(0).



# Modules

- Loading the Dataset
- Data Preprocessing
- Splitting of datasets into training and testing sets.
- Text Vectorization (TF-IDF)
- Training Machine Learning Models.
- Model Evaluation and sample prediction.
- Creating GUI Application.

# Results

```
Accuracy: 0.9406952965235174  
ROC AUC for Multinomial NB: 0.91  
ROC AUC for SVM: 0.94  
Sample: Your the best!♥  
Predicted class: Not Spam  
PS C:\Users\prana\Desktop\random code> |
```

Figure: Sample Prediction

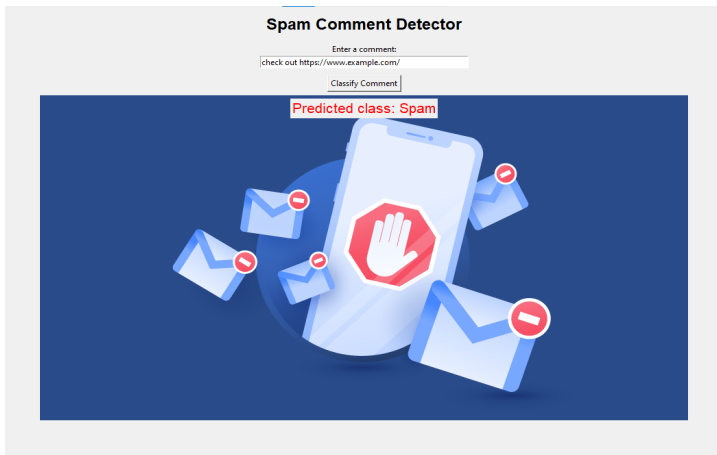


Figure: GUI Application output

# Conclusion & Future Scope

- **Conclusion** : The developed YouTube spam comment detection system employs a systematic approach. By utilizing TF-IDF for feature extraction and training machine learning models, the system demonstrates efficacy in distinguishing between spam and non-spam comments. This system contributes to enhancing user experience and content quality by mitigating the impact of spam comments on online platforms.
- **Future Scope** : There is a potential for enhanced model optimization by exploring advanced machine learning techniques and fine-tuning algorithms to achieve even higher accuracy. Real-time monitoring capabilities could be integrated to identify and manage spam comments as they are posted. Ensuring adaptability to evolving spamming trends will be crucial for maintaining the system's effectiveness in combating emerging forms of online spam

# References

- Moller, Baumgartner, K"uhne Peter, (2021). "Sharing the fun? How social information affects viewers' video enjoyment and video evaluations". Human Communication Research.
- R. K. Das, S. S. Dash, K. Das, and M. Panda, (2020) "Detection of spam in Youtube comments using different classifiers,".

# References

- N. M. Samsudin, C. F. B. Mohd Foozy, N. Alias, P. Shamala, N. E Othman (2019) “YouTube spam detection using Naïve Bayes and logistic regression algorithms”.
- S. Aiyar and N. P. Shetty,(2018) “N-gram assisted Youtube spam comment detection”.

# Thankyou