IIETA International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Ingénierie des Systèmes d'Information
Vol. 31, No. 1, January, 2026, pp. 109-121
Journal homepage: http://iieta.org/journals/isi

Fake Video Detection on YouTube via Metadata-Driven Sentiment Analysis and GAN-Based Data Augmentation

Shraddha Kalbhor[1*], Dinesh Goyal[2], Kriti Sankhla[1]

[1] Computer Science & Engineering, Poorinma University, Jaipur 302022, India
[2] Computer Science & Engineering, Poorinma Institute of Engineering and Technology, Jaipur 302022, India

Corresponding Author Email: shraddha.kalbhor000@gmail.com

https://doi.org/10.18280/isi.310111

Received: 12 October 2025
Revised: 28 December 2025
Accepted: 15 January 2026
Available online: 31 January 2026

Keywords:
fake video detection, YouTube metadata analysis, sentiment analysis, generative adversarial networks, topic modeling, social media analytics, data augmentation

ABSTRACT

The rapid growth of user-generated content on social media platforms has significantly increased the spread of misleading and manipulated videos. On platforms such as YouTube, fake trailers, edited clips, and deceptive multimedia content can easily mislead viewers and distort public perception. Therefore, developing effective techniques for detecting fake videos has become an important research challenge. This study proposes a machine-learning framework for detecting fake YouTube videos by combining metadata analysis, sentiment analysis of user comments, and generative adversarial network (GAN)-based data augmentation. First, a custom dataset of YouTube video metadata and user comments is constructed using the YouTube Data API. After preprocessing and text normalization, sentiment features are extracted from user comments using natural language processing techniques. Topic modelling using latent Dirichlet allocation (LDA) is further employed to capture latent semantic patterns in user feedback. To address class imbalance and enhance model robustness, a conditional generative adversarial network (CTGAN) is used to generate synthetic training samples. Several machine-learning classifiers, including Random Forest (RF), Decision Tree (DT), K-Nearest Neighbors (KNN), and Naïve Bayes (NB), are evaluated on both the publicly available US Trending YouTube Video Dataset (USVideo) and a self-constructed YouTube review dataset. Experimental results demonstrate that GAN-based data augmentation improves classification performance by approximately 5–6% across most models. Among the evaluated classifiers, RF achieves the best performance, reaching 98% accuracy on the USVideo dataset and 88% accuracy on the self-constructed dataset. These results indicate that integrating metadata features, sentiment analysis, and GAN-based data augmentation provides an effective approach for detecting misleading video content on social media platforms.

1. INTRODUCTION

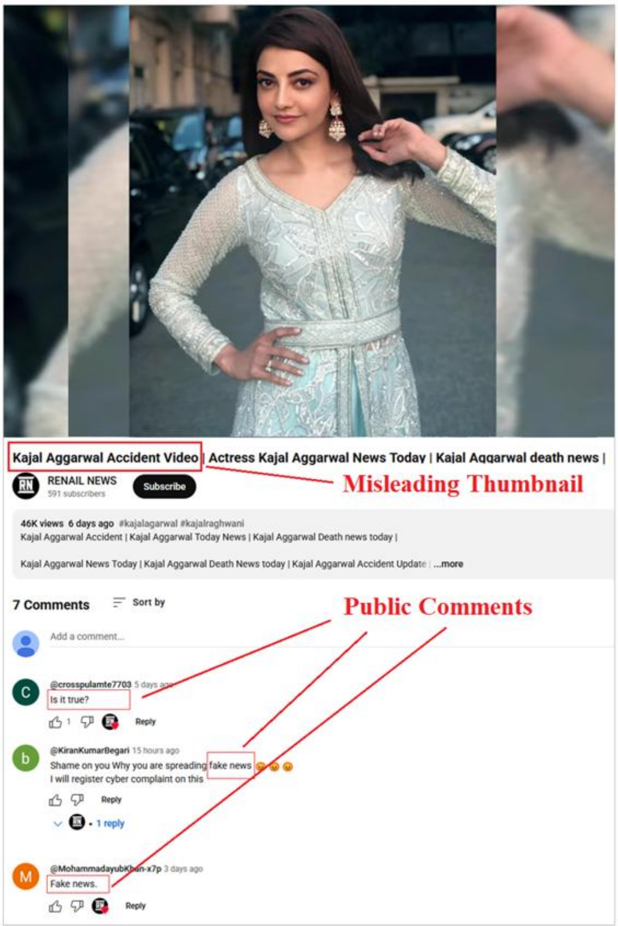
With the rapid growth of digital media, platforms like YouTube have evolved into major repositories of user-generated videos, covering a wide range of categories such as gaming, lifestyle, educational topics, and entertainment [1]. As more and more the number of resources available is growing, analyzing and generating value from YouTube video content becomes highly important. In order to protect individuals and communities, the fake or misleading information on social media must be controlled or stopped [2]. These agents compromise the safety, and financial security of innocent people by disseminating fraud. Moreover, the dissemination of misinformation also undermines a democracy's very bedrock and creates distrust in sources of information [3]. In addition, polarization creates hostility and conflict between groups and is based on false information on social groups. Incorrect information can affect the consumer's behavior, market stability and investor appraisal among other aspects of the economy. Therefore, it is not only a personal responsibility, but also a way to protect democratic values, social harmony and integrity of the economy.

In this study, the term "fake video" specifically refers to edited or manipulated movie clips, trailers, and other online videos that have been intentionally altered to mislead viewers. These videos often combine unrelated scenes, modify original content, or use deceptive editing techniques to create false impressions about events, characters, or storylines.

The rapid dissemination of false information is one of the main issues in the digital world. On social media platforms and WhatsApp or social apps, where algorithms often interact with precision, false information spreads quickly [4]. For example, some content creator adds (fake) videos of movie trailer for upcoming, unpublished movies. The goal of these fake trailers is to manipulate fan interest and expectations to drive higher channel traffic. Similarly, other fake video can spread false details about many subjects in an attempt to speak or cause misunderstanding. The material that can be produced online and can be spread online has made it challenging to consider between authentic and scams - fraud videos, which also create problems for public and material platforms equally to maintain the accuracy and dependence of the information provided.

Consequently, shocking or scam data spreads broadly and generates "eco-importunateness", where people see information that supports their pre-existing beliefs. It had a significant impact by reducing social division, political stress and institutional trust. By fostering skewed beliefs and inhibiting critical thought, eco-chambers and filter bubbles obliterate the issue. An illustration of a deceptive thumbnail and user comments can be found in Figure 1. The misleading thumbnail and user comments make it obvious that the information being circulated is untrue.

Numerous techniques have been investigated for the identification of fake news and reviews, such as the study by Sudhakar and Kaliyamurthie [5], which uses classifiers including K-Nearest Neighbor, Decision Tree (DT), SVM, Linear SVM, and SGD in addition to the N-gram and TF-IDF approaches for feature extraction.



Kajal Aggarwal Accident Video | Actress Kajal Aggarwal News Today | Kajal Aqgarwal death news |

RENAIL NEWS
591 subscribers

46K views 6 days ago #kajalagarwal #kajalraghwani
Kajal Aggarwal Accident | Kajal Aggarwal Today News | Kajal Aggarwal Death news today |

Kajal Aggarwal News Today | Kajal Aggarwal Death News today | Kajal Aggarwal Accident Update | ...more

7 Comments

@crosspulamte7703 5 days ago
Is it true?

@KiranKumarBegari 15 hours ago
Shame on you Why you are spreading fake news 😡😡😡
I will register cyber complaint on this

@MohammadayubKhan-x7p 3 days ago
Fake news.

Figure 1. Fake video detection using comments [https://www.youtube.com/watch?v=zwKFOC2HlXA]

Using the publicly accessible Fake News dataset, Sudhakar and Kaliyamurthie [5] created a Linear SVM machine learning (ML) classifier that produced a 98% accuracy rate. A slightly different approach was taken in a related study by Ozbay and Alatas [6], who used TF-IDF as their feature extraction technique. They also assessed the efficacy of 23 different ML classifiers for identifying fake news and reviews.

Given the success of DL across a range of fields, DL-based approaches have lately been put forth and generated a lot of interest. First, while modelling the features of input news, deep learning (DL) can effectively exploit its expressive power without feature engineering. For instance, Chen et al. [7] used TSNN (Topic and Structure Aware Neural Network) to illustrate the sequential relationship between news stories. Madani et al. [8] used a two-phase model using natural language processing; they extracted two new structural features along with other key features and used a hybrid algorithm to detect fake news. Meanwhile, Chang et al. [9] employed a GNN model that integrated a top-down and bottom-up method to analyze the transmission and propagation of rumors. Similarly, Geto et al. [10] proposed CNN with attention-based BiLSTM model for multimodal Amharic fake news detection.

Most studies categorize the data based on feature opinions, even though many works already focus on detecting fake videos. Recently, unsupervised and weakly supervised methods have gained attention as ways to address issues caused by a deficiency of labeled data [11]. This is important for real-world applications because of how fast information spreads on social media, With the rapid growth of digital content creation and sharing, authenticating web videos is imperative. Comments are powerful and can both provide valuable context as well social proof, but they don't do a great job of telling you if a video is real. By conducting a comprehensive review, including checking sources, metadata identification and visual analysis combined with expert opinion, researchers are able to establish whether or not footage is genuine even if there are only a minimum of comments [11].

Automated systems that can reliably analyze the bogus videos and opinions stated in user evaluations on social media are becoming more and more necessary. This work presents an enhanced model for collecting and categorizing YouTube video metadata across multiple content categories. Our approach initially extracts comments then perform data processing and sentiment analysis of YouTube videos using the NLP and then classify videos as fake or valid using ML methods. We also use domain-specific features and relevant data to enhance sentiment analysis's effectiveness and goal across various video genres. To validate the proposed method, the publicly available US Trending YouTube Video Dataset (USVideo) is used. The data imbalance issue in the YouTube video and USVideos dataset is addressed using a generative adversarial network (GAN), specifically the Conditional Tabular Generative Adversarial Network (CTGAN) [12], which improves the overall performance of the ML model. Even though there are a lot of comments, the proposed algorithms are able to identify fake videos in addition to analyzing sentiment. This two-in-one tool gives users and content producers helpful information by making it simple help discern between authentic and fraudulent YouTube video evaluations. The following is a description of this research effort's benefit.

1.1 Creating YouTube review dataset

Scrape user comment along with metadata using YouTube API key to create a large corpus of YouTube reviews. By allowing strong identification of fake videos by elaborate contextual analysis combined with a smaller number of user reviews, the goal is to improve over the classification accuracy of video authenticity.

1.2 Data balancing with generative adversarial network

Create and use enhanced sentiment categorization methods utilizing GANs models and customized metadata.

1.3 Establish a multi-stage analysis methodology

Establish a systematic, multi-phase approach that includes the generation of unique data, comprehensive feature extraction, sentiment classification using ML.

1.4 Verify results using publicly available datasets

The suggested methodology is validated on publicly accessible datasets, including the YouTube US video dataset, with performance assessed through standard evaluation metrics such as F1-score, precision, recall, and accuracy.

The article is organized as follows: Section 2 reviews related work, Section 3 describes the self-generated dataset and sources, Section 4 presents the proposed methodology, Section 5 reports the tools, performance metrics, and experimental results, and Section 6 concludes the study.

2. LITERATURE REVIEW

The rise of fake and misleading videos on social media has led to extensive research on automated detection methods. Early studies focused mainly on feature-based classification using traditional ML algorithms. Recently, unsupervised, weakly supervised, and GAN-based approaches have gained importance due to the lack of labeled data.

Recent works have shown the use of GANs in correcting data imbalance and improving classification performance. This has led to advancements in research in the areas of false video detection and sentiment analysis [12]. The features that were taken from the video need to be carefully examined. Since they only concentrate on content-based solutions, the current research has not adequately addressed this issue. Bronakowski et al. [13] have introduced an automatic clickbait detector and used gradient-boosted DT to predict clickbait from news headlines. To train the classifier, they also use a variety of manually created features, like n-grams and bag-of-words. YouTube's impact on scientific research is examined by Welbourne et al. [14], who also provided metrics to measure it. Metrics like views, likes, and comments are covered, along with how they connect to more conventional measures of research impact like citations. In an effort to improve model performance and efficiency, Gul and Bashir [15] investigated a number of feature selection strategies in sentiment analysis. The problem of choosing the most instructive features from big datasets is discussed by the authors. In contrast, Sharma et al. [16] proposed an effective randomized feature selection approach aimed at improving computational efficiency and accuracy in feature subset selection. Their method employs a Boolean operator–based Particle Swarm Optimization (PSO) model to identify optimal feature subsets.

A context-specific heterogeneous graph convolutional network was suggested by Zuo et al. [17] with the intention of addressing the issues associated with implicit sentiment analysis. Additionally, the network was designed to successfully capture contextual relationships and sentiment expressions across a variety of domains. Li and Zou [18] focused on sentiment analysis for short texts using DL techniques. They brought attention to the challenges that are related with the length of such data and the very limited context in which it is presented. Balli et al. [19] applied NLP techniques in order to conduct an analysis of the sentiment included within content on Twitter that was written in Turkish. This article makes a contribution to our understanding of the specific emotional patterns that are exhibited by users of social media platforms who speak Turkish. The research conducted by Kumar et al. [20] proposes a multi-domain approach to the classification of sentiments in online social networks. The method utilizes several data variables. The study by Jayaraman et al. [21] tries to enhance the sentiment polarity identification by concentrating on locating some targets in texts. The authors apply the ML algorithms, including a Random Forest (RF) and Gradient Boosting to classify fake news in order to increase the detection rate and improve the reliability of information [22].

According to Ramezani et al. [23], a CNN-based model can be used to detect misleading videos; the authors suggest that the presence of misinformation can be detected by using DL. To achieve a finer representation of the nuanced expressions of emotions, Yang et al. [24] integrated sentiment lexicons with the DL methods in the name of carrying out sentiment analysis on Chinese product reviews of online retailers. In the meantime, Li et al. [25] presented a hybrid approach that combines CNN, BiGRU, and Attention Mechanisms (AT), TIG-CNN-BiGRU to better lyrics recognition ability, which may be used in the lyrics sentiment analysis.

In order to obtain realistic synthetic data to enhance the research on cyber security, Hermawan et al. [26] presented a comparative investigation of GAN-based tabular synthetic data generation methods, aiming to clarify how different GAN variants perform when generating realistic structured (tabular) data. The study is motivated by the growing need for synthetic data in scenarios where real datasets are limited by privacy, access restrictions, imbalance, or scarcity, while still requiring data that preserves key statistical properties and supports downstream analytics.

In this line of work, Deressa et al. [27] proposed GenConViT, a generative convolutional transformer vision and is a system that uses ConvNeXt and Swin Transformer as the surrounding backbones and utilizes autoencoder / variational autoencoders as image generators to enhance the ability to resist a variety of Deepfake generation models.

Petmezas et al. [28] improved video deepfake detection through the incorporation of 3D Morphable Models (3DMMs) into a hybrid CNN–LSTM–Transformer framework, which captures detailed facial spatial cues as well as temporal instabilities across both short- and long-term sequences.

In addition to these visual methods, Wang et al. [29] introduced ERF-BA-TFD+, a multimodal audio-visual Deepfake detector, which combines the extended receptive-field visual features with audio features on datasets like DDL-AV and LAV-DF, and emphasizes the value of cross-modal inconsistencies used.

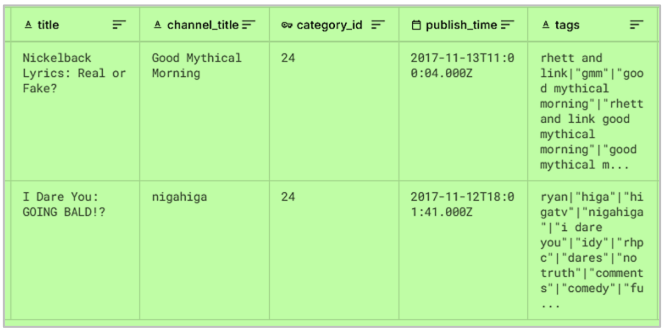
A recent survey by Alrashoud [30] systematically surveys such methods of Deepfake video detection and categorizes them into CNN-based, frequency-domain, and multimodal techniques and highlights the fast rate of development of such models and the accompanying robustness issues.

In contrast to these Deepfake-oriented studies, which primarily focus on AI-generated or heavily manipulated facial content, our work targets a different but practically important class of edited fake movie clips, trailers, and videos, typically produced using conventional video editing rather than generative face-swapping. Moreover, while most Deepfake detectors rely on frame-level visual artifacts or audio–visual inconsistencies, our approach exploits YouTube metadata and user comments (including emotional cues) as key signals, and employs GAN-based augmentation to enhance traditional ML models for fake video detection. This positions our contribution as complementary to Deepfake detection research, addressing a less explored but highly relevant form of misleading video content on social platforms.

3. DATA SOURCES

3.1 US Trending YouTube Video Dataset

YouTube ranks its trending videos using various user engagement metrics, including comments, views, shares and likes, instead of depending solely on view counts, according to Variety magazine. The USVideo dataset [31] captures daily records of these trending videos across various countries, including the US, UK, India, etc. It includes up to 200 videos per day, covering several months of data. Figure 2 depicts a sample taken from the dataset that is available on USVideo.



title	channel_title	category_id	publish_time	tags
Nickelback Lyrics: Real or Fake?	Good Mythical Morning	24	2017-11-13T11:00:04.000Z	rhett and link|"gmm"|"good mythical morning"|"rhett and link good mythical morning"|"good mythical m...
I Dare You: GOING BALD!?	nigahiga	24	2017-11-12T18:01:41.000Z	ryan|"higa"|"higatv"|"nigahiga"|"i dare you"|"idy"|"rhpc"|"dares"|"no truth"|"comments"|"comedy"|"fu ...

Figure 2. US Trending YouTube Video Dataset sample

Some of the most important metadata, such as the video title, the channel name, the publish time, tags, likes, dislikes, views, comment count, description, and `category_id`, are maintained separately for each region. Users are able to reference the regional JSON files that are connected with the `category_id` in order to map it to actual categories.

3.2 Self-created YouTube video metadata dataset

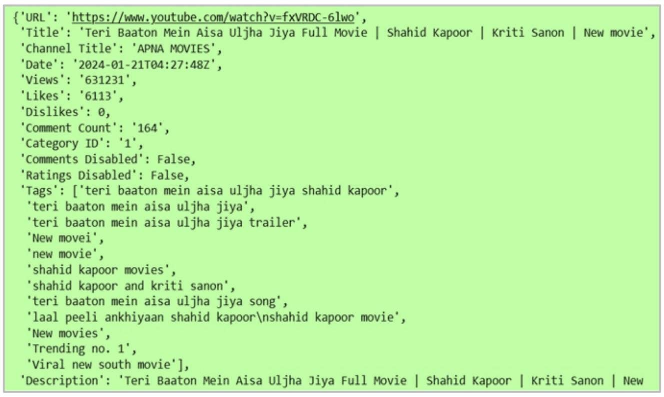
We used the YouTube Data API to generate comprehensive metadata for our dataset as part of our process. A collection of more than 2000 different YouTube video URLs that were scattered across several categories, including politics, entertainment, sports, and more. The detail dataset creation methodology is explained below;

YouTube MetaData Creation Process

(1) *Generate API Key:* YouTube data should be sent to API through Google Developers Console to get the API key required to generate data sets related to comments on YouTube videos. When you arrive, start building a new project, establishing an API key with the required rights and turning on the YouTube Data API, such as reading video comments. This key allows important applications to validate and access the API of YouTube, so that video commentary can be organized and examined to generate a dataset.

(2) *Enter YouTube Video URL:* In the data aggregation tool, enter URL of YouTube video relevant to making a dataset related to the comments on YouTube videos. Then you can get comments related to each video URL using YouTube Data API. You can create your own dataset for further examination or analysis of YouTube videos by compiling and reviewing these comments.

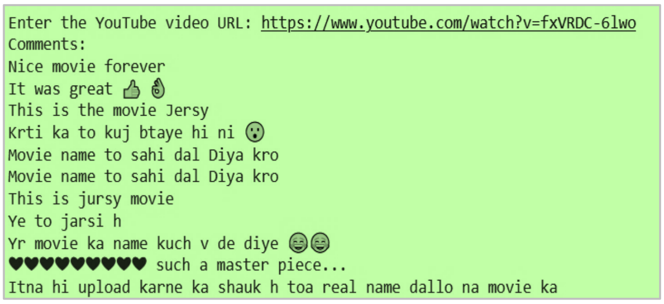
(3) *Metadata extraction from YouTube Videos:* YouTube video includes Metadata removal, tag, video URL, title, Upload date and number of ideas. It also includes landscapes that may not include metadata, for example without title or tag. To ensure the perfection and purity of the dataset, it is completed using these examples and perhaps other data sources. This is especially useful for checking comments on YouTube videos. Figure 3 shows the extracted YouTube Video Metadata.



```
{'URL': 'https://www.youtube.com/watch?v=fxVRDC-6lwo',
 'Title': 'Teri Baaton Mein Aisa Uljha Jiya Full Movie | Shahid Kapoor | Kriti Sanon | New movie',
 'Channel Title': 'APNA MOVIES',
 'Date': '2024-01-21T04:27:48Z',
 'Views': '631231',
 'Likes': '6113',
 'Dislikes': 0,
 'Comment Count': '164',
 'Category ID': '1',
 'Comments Disabled': False,
 'Ratings Disabled': False,
 'Tags': ['teri baaton mein aisa uljha jiya shahid kapoor',
  'teri baaton mein aisa uljha jiya',
  'teri baaton mein aisa uljha jiya trailer',
  'New movei',
  'new movie',
  'shahid kapoor movies',
  'shahid kapoor and kriti sanon',
  'teri baaton mein aisa uljha jiya song',
  'laal peeli ankhiyaan shahid kapoor\nshahid kapoor movie',
  'New movies',
  'Trending no. 1',
  'Viral new south movie'],
 'Description': 'Teri Baaton Mein Aisa Uljha Jiya Full Movie | Shahid Kapoor | Kriti Sanon | New
```

Figure 3. YouTube video metadata extraction

(4) *Download YouTube Video Comments:* A dataset is created using the YouTube Data API, specifically the comments feature. To download the comments for the selected video, enter the video ID. Using this API request, you may get all of the comments associated with the video, including the metadata, the time stamp of the reactions, and the user name of the commenter. This lets you obtain a lot of information for your dataset about YouTube video comments. Figure 4 shows the extracted comments from YouTube Video.



```
Enter the YouTube video URL: https://www.youtube.com/watch?v=fxVRDC-6lwo
Comments:
Nice movie forever
It was great 👍 👌
This is the movie Jersy
Krti ka to kuj btaye hi ni 😮
Movie name to sahi dal Diya kro
Movie name to sahi dal Diya kro
This is jursy movie
Ye to jarsi h
Yr movie ka name kuch v de diye 😄😄
🖤🖤🖤🖤🖤🖤🖤🖤🖤 such a master piece...
Itna hi upload karne ka shauk h toa real name dallo na movie ka
```

Figure 4. YouTube video extracted reviews

(5) *Store Comments and Metadata in CSV:* Once you've downloaded the comments from a YouTube video, make sure to save them in a CSV file along with a unique identifier (URL_ID). This ID is YouTube Videos Unique ID. By assigning each comment a unique ID, we created a well-organized dataset that makes it much easier to analyze and link comments to the right video for research or insights.

4. PROPOSED METHODOLOGY

The proposed system architecture diagram of fake video detection and sentiment analysis using YouTube Video Metadata is shows in Figure 5. A dataset based on YouTube video comments is created at the beginning of the procedure (details are given in section 3). Data pre-processing, which includes standardizing textual data and eliminating duplicates and superfluous whitespace, is done first after the dataset has been gathered. Further dataset cleaning involves eliminating emoji's, punctuation, stop-words and stemming to enhance data quality. NLP processing and Feature extraction is then applied to get the key attributes related to content, including category sentiment. To strengthen the dataset and remove data imbalance problem GAN (CTGAN) model is utilized, ensuring robustness and scalability. After downloading YouTube video comments, be sure to store them in a CSV file with a unique identity (URL_ID). This ID can be the URL or YouTube ID for the video. It will be much simpler to examine and connect comments to the appropriate films for your study or insights if you give each comment a unique ID.

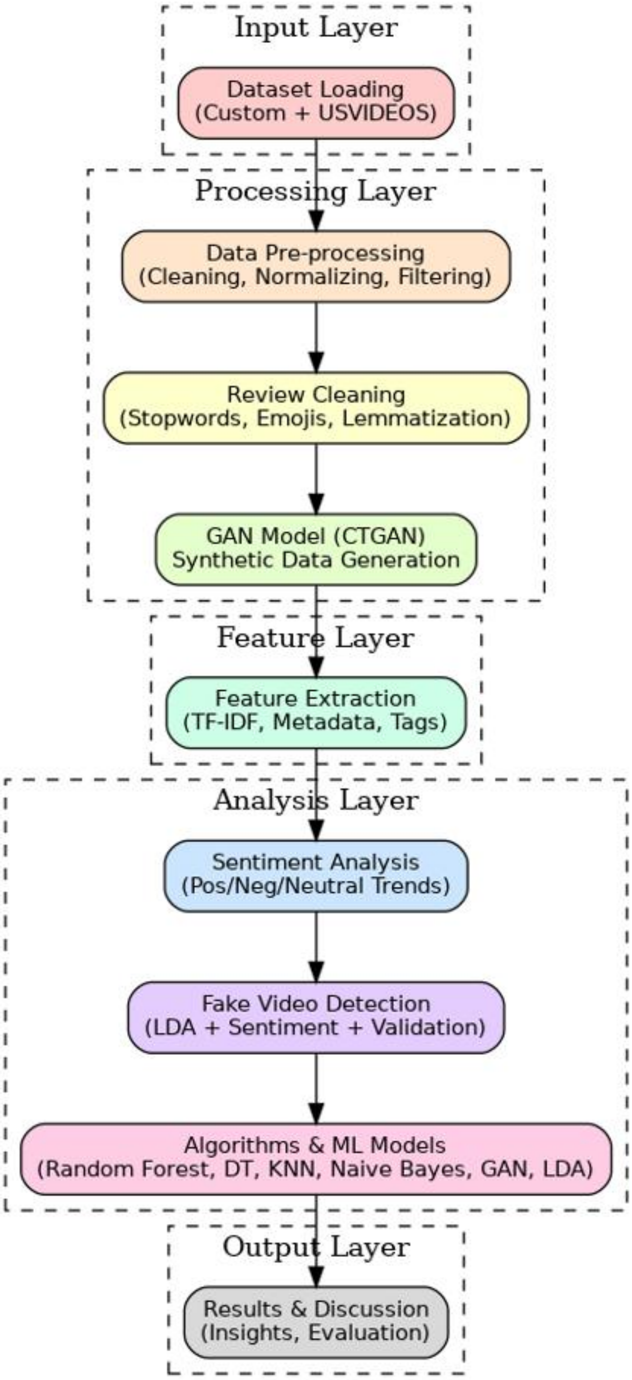


Figure 5. Proposed system architecture diagram

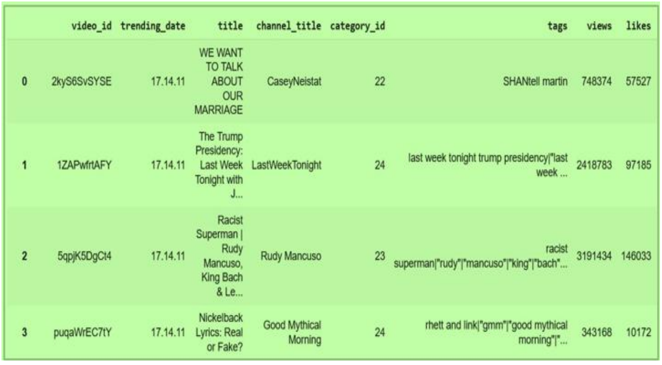
4.1 Methodology

4.1.1 Dataset loading

This work utilizes two datasets: (1) a self-created custom YouTube video review, and (2) publicly available USVideos review dataset. Prior to analysis, both datasets require thorough pre-processing to ensure proper formatting and organization for related modelling and classification tasks. The self-created YouTube dataset consist of comprehensive metadata for each video, including titles, reviews, URLs, view counts, like/dislike ratios, and user ratings, etc. The initial data preparation phase involves loading the CSV files, performing necessary data cleaning operations, and structuring the information into an analyzable format suitable for ML models.

4.1.2 Data pre-processing

Data Pre-processing involves several steps, including eliminating unnecessary columns to focus on relevant features, removing duplicate comments, normalizing text, removing whitespace and converting text to lowercase, implementing a minimum view count threshold to maintain content relevance. These pre-processing steps transform the raw data into a more structured and uniform format, significantly enhancing its suitability for relevant modelling and analytical tasks. Figure 6 presents the original metadata.

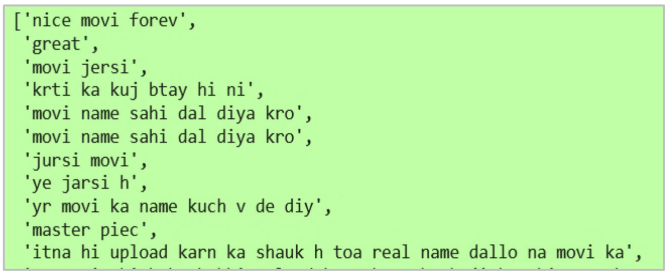


	video_id	trending_date	title	channel_title	category_id	tags	views	likes
0	2kyS6SvSYSE	17.14.11	WE WANT TO TALK ABOUT OUR MARRIAGE	CaseyNeistat	22	SHANtell martin	748374	57527
1	1ZAPwfrtAFY	17.14.11	The Trump Presidency: Last Week Tonight with J...	LastWeekTonight	24	last week tonight trump presidency|"last week ...	2418783	97185
2	5qpjK5DgCt4	17.14.11	Racist Superman | Rudy Mancuso, King Bach & Le...	Rudy Mancuso	23	racist superman|"rudy"|"mancuso"|"king"|"bach"...	3191434	146033
3	puqaWrEC7tY	17.14.11	Nickelback Lyrics: Real or Fake?	Good Mythical Morning	24	rhett and link|"gmm"|"good mythical morning"|"...	343168	10172

Figure 6. Metadata of self-created YouTube video reviews

4.1.3 Review cleaning

During the cleaning process, special characters, punctuation, and emojis are eliminated to increase readability. Furthermore, the stop-words are removed. Additionally, pre-processed text is lemmatized. Figure 7 displays some reviews of a self-created dataset.



```
['nice movi forev',
 'great',
 'movi jersi',
 'krti ka kuj btay hi ni',
 'movi name sahi dal diya kro',
 'movi name sahi dal diya kro',
 'jursi movi',
 'ye jarsi h',
 'yr movi ka name kuch v de diy',
 'master piec',
 'itna hi upload karn ka shauk h toa real name dallo na movi ka',
```

Figure 7. Clean text of self-created YouTube video reviews data

4.1.4 Generative adversarial network model

The implementation of GAN involves a process comprising four key steps: data-preparation, model selection, training, and evaluation. After cleaning and pre-processing of dataset, a

CTGAN architecture specifically designed to generate high-quality synthetic data is applied. The technical details of CTGAN implementation, including its architecture and training dynamics, are outlined in the Algorithm section. Figure 8, compares sample distributions before and after GAN processing on the USVideo dataset.

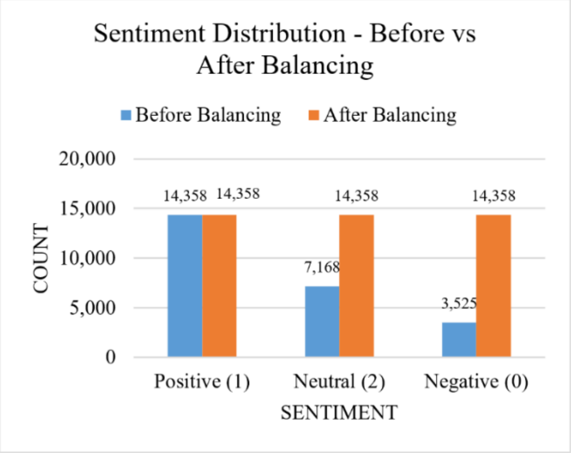


Figure 8. Without data balancing and with GAN based data balancing

Note: GAN = generative adversarial network.

4.1.5 Feature extraction

When it comes to YouTube video reviews, there are several features to consider, such as categories, the percentage of likes and dislikes, capitalization, the number of words in the title, and tag phrases. These are important things that assist us in exploring and learning about what your users are doing/asking within the body of data. These cues would arguably offer useful navigational aids to the textual and interactive content of YouTube videos, too, in terms of making an informed determination as to its content's authenticity or lack thereof. To improve the numerical representation of text, we also use topic modelling approaches such as TF-IDF vectorization [32]. TF-IDF calculates weighted scores for words and phrases to identify the most important words in each review comment. The method allows the identification of important subjects and related topics from the reviews, through which insights are derived toward video content and audience involvement. Through the identification of salient phrases and patterns, we characterize discussion and sentiment by viewers, who might later augment the dataset to uncover new insights.

4.1.6 Sentiment analysis

Sentiment analysis analyses how emotionally positive or negative the reviews of YouTube videos in the dataset are. In sentiment analysis, NLP techniques are applied in order to categorize the sentiment as either positive, neutral, or negative.

Figure 9 is an illustration of the sentiment Trend analysis that was performed on the USVideo dataset. This categorizes comments and gives us a helpful indication of how the audience is feeling or what they are thinking about the material.

4.1.7 Fake video detection

One useful tool for dealing with YouTube video reviews is latent Dirichlet allocation (LDA) [33], which integrates sentiment analysis and topic modelling. For a topic-modelling an LDA approach is applied to uncover the hidden topics in the review corpus. These subjects consist of words that tend to co-occur. In addition to providing content creators and researchers with a systematic understanding of the debates that take place within the review, these themes also indicate shared elements, such as the quality of the video, the involvement of the audience, the relevancy of the contents, and the production value.

In our proposed framework, LDA is used as a feature-extraction and content-verification component that strengthens the fake video classification process. To determine the sentiment related to each topic, LDA is combined with sentiment analysis for topic identification. After identifying the subjects, we apply sentiment lexicons to obtain polarity ratings, allowing us to determine whether the reviews express positive, negative or neutral attitudes on multiple aspects of the video such as delivery style, content and overall viewing experience. Validation of the videos' titles against review content and detection of frequently used keywords are also used to uncover fake reviews. This approach enables us to understand detailed topics-distribution and sentiment-distribution of user feedback. The sentiment score of the YouTube video data set Figure 10 illustrated that the video is false and has poor response from different viewers as well.

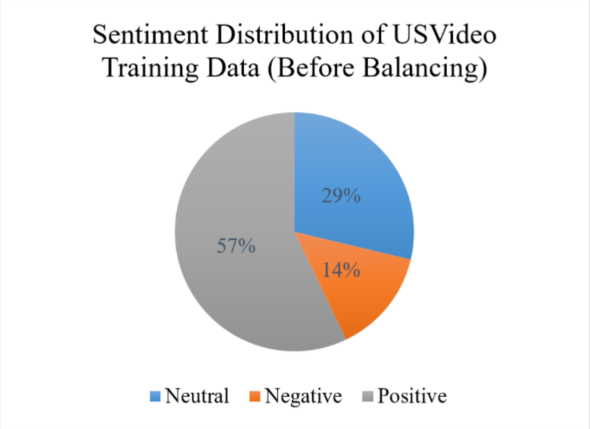


Figure 9. Sentiment analysis (USVideo dataset)

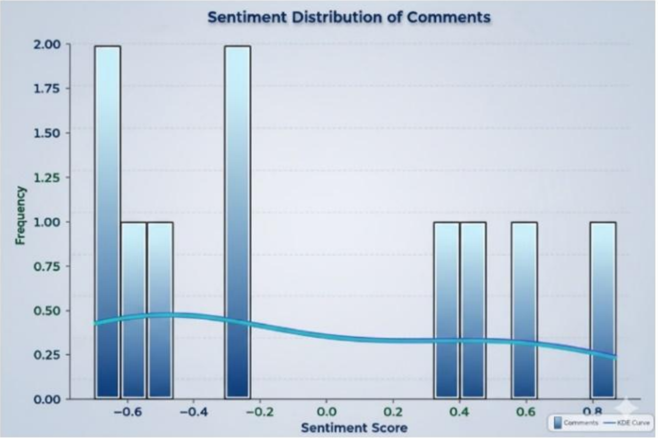


Figure 10. Sentiment score (YouTube video dataset)

The output of LDA (topic distribution) combined with sentiment scores forms a set of high-level semantic features representing how viewers perceive different aspects of the video. These features are transformed into vectors and fed into the machine-learning classifiers.

4.2 Algorithms

The purpose of this part is to provide a more in-depth analysis of a variety of ML strategies, such as RF, DT, Naïve Bayes (NB), K-Nearest Neighbors (KNN) and LDA among others. These techniques are specifically designed for assessing the feelings expressed in reviews and identifying films that are fraudulent. You'll find a comprehensive explanation of each algorithm in the following paragraphs.

Algorithm 1: CTGAN

CTGAN is a customized GAN architecture, with a generator and discriminator being its two main parts. The discriminator plays the role of a classifier to determine whether the input samples are generated or real ones, and the generator is trained to produce synthetic data that mimics the distributions of the actual data. During training, these two components play an adversarial game: the discriminator gives feedback by punishing outputs that don't look realistic, in that way driving the generator towards producing more realistic examples. CTGAN synthesizes data efficiently and it preserves the statistical properties of original dataset by utilizing both conditional transformation methods and classic GAN principles. You can visualize how the GAN model works in Figure 11.

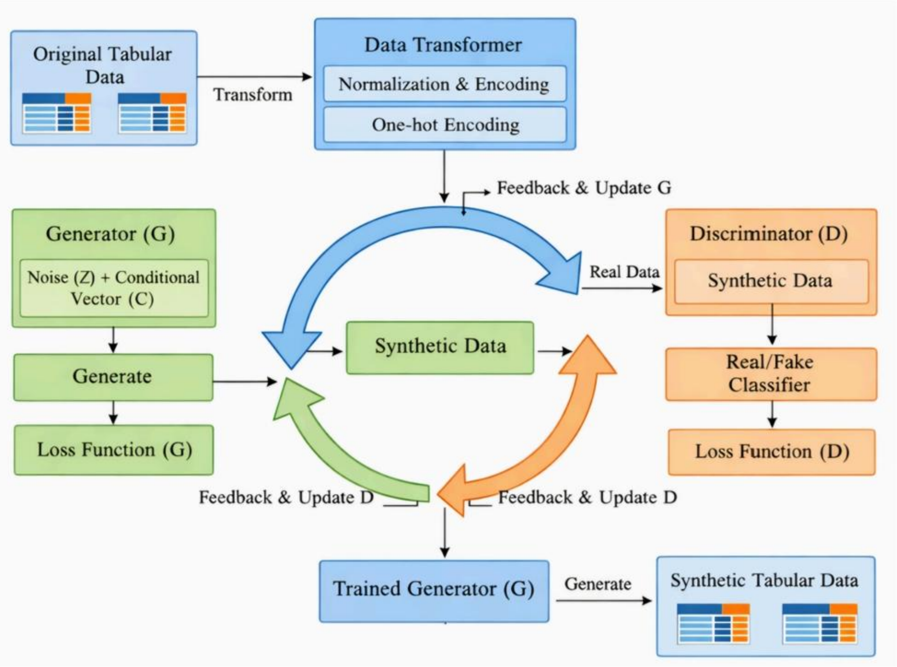


Figure 11. Working of GAN model
Note: GAN = generative adversarial network.

The Conditional GAN (CT model) gains power to transact with the users by attaching them with conditional information such as some features, properties of samples thereby allowing more user control over the resulting features of generated instances. This technique makes it possible to create synthetic data that is remarkably lifelike and closely resembles real datasets. This has applications in ML and the development of artificial test cases for training and privacy testing. How CTGAN works The Conditional GAN Builds upon the traditional GANs and gives users a more direct control of the features of the generated samples by leveraging conditional information that is related to certain data traits or attribute. The approach enables to generate extremely robust synthetic data closely resembling the characteristics of original data population, with various applications such as privacy training, text generation, and ML. This is a thorough description of how CTGAN functions:

- The model receives the original dataset after it has been cleaned, normalized, and feature-selected.
- A generator creates samples, which the discriminative network compares to the actual data distribution.
- By using artificial candidates that closely resemble actual data, training advances to improve network reliability.
- The discriminator uses the original training dataset as a reference.
- Training keeps going until dataset samples reach the required accuracy.
- Given random latent vectors as input, the generator progressively learns to synthesize plausible outputs that aim to fool the discriminator.
- Both the discriminator and the generator go through backpropagation; the discriminator gets better at telling real images from fakes, while the generator enhances image quality.
- The discriminator is a CNN, and the generating network is a de-convolutional neural network.
- The hyper-parameters of CTGAN model are shown in Table 1.

Algorithm 2: Latent Dirichlet Allocation

LDA operates on the fundamental idea that every document is essentially a mix of hidden topics, each represented as a probability distribution over words. This method tackles a significant challenge in NLP: the automatic detection and organization of thematic patterns within large sets of text data. By analyzing how words co-occur, LDA works to reverse-engineer the process that likely generated the documents we see. With the explosion of written content in our digital world—think news articles and social media posts—there's a growing need to automate how we classify and summarize this information. LDA is an essential component in enhancing information retrieval, recommending content, and analyzing topic structures across extensive datasets. Its independent nature makes it particularly useful for exploratory data analysis, as it doesn't require specific labels when the data structure is unclear. LDA is a useful framework for investigating and comprehending the underlying subject structure of lengthy texts in the field of natural language processing.

$$P(z_i = j|z_{-i}, w_i, d_i, \cdot) \propto \frac{C^{WT}_{w_i j} + \eta}{\sum_{w=1}^{W} C^{WT}_{wj} + W_\eta} \frac{C^{DT}_{d_i j} + \alpha}{\sum_{t=1}^{T} C^{DT}_{d_i t} + T\alpha}$$

where, Vocabulary length, or the total number of unique words, is represented by W.

- One hyper-parameter is α, and T is the number of patients. Any document that has only a few major topics will be given greater weight if its alpha number is low; a high alpha value will yield many more dominant subjects.
- One hyper parameter is η. A low η value (eta) indicates that each topic should consist of only a few prominent words.

By pre-processing, analyzing, and classifying the sentiment of YouTube videos, each of these algorithms enhances the dataset and offers insightful information about audience involvement and content perception.

Table 1. CTGAN hyperparameters

Parameters	Value	Description
Epochs	50	Number of training iterations
Batch size	256	Mini-batch size used during training
generator_dim	(128, 128)	Hidden layer sizes of the Generator network
discriminator_dim	(128, 128)	Hidden layer sizes of the Discriminator network
embedding_dim	56	Embedding dimension for categorical variables
generator_lr	2×10^{-4}	Learning rate for the Generator
discriminator_lr	2×10^{-4}	Learning rate for the Discriminator
generator_decay	1×10^{-6}	L2 regularization for the Generator
discriminator_decay	1×10^{-6}	L2 regularization for the Discriminator
verbose	True	Enables training progress output

Note: CTGAN = Conditional Tabular Generative Adversarial Network.

5. RESULTS AND DISCUSSION

Performance metrics, experimental designs, analysis, and findings are presented in this section when utilizing YouTube video reviews and metadata to detect fake videos.

5.1 Experimental setup

Python 3.10 was used in an experimental setup in Google Colab, several libraries were used to apply ML methods. Scikit-Learn was one of the main libraries used to create classifies as RF, DT, KNN, NB for classification of fake news. Tokenization and elimination hours for stop words are two examples of run -up tasks completed using the Natural Language Toolkit (NLTK). Data uses measures such as pre-pro-rose, model training, data dating, classification of ML and evaluation accuracy, accuracy, recall and F1 score, which was part of the experimental pipeline to measure model performance.

5.2 Performance parameters

Accuracy, Precision, Recall and F1-score are used to calculate the performance of models. Formulas are listed below, Where True positives (TP) are the values, we correctly predicted to be positive. On the flip side, true negatives (TN) are those we expected to be negative and were right about. Now, a false positive (FP) happens when we think something is a yes, but it turns out to be a no. And then there are false negatives (FN), which occur when the real class is a yes, but our prediction says it's not there at all.

$$Precision(Pre) = \frac{TP}{TP + FP}$$

$$Accuracy(Acc) = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Recall(Rec) = \frac{TP}{TP + FN}$$

$$F1 - Score = 2 * \frac{Pre * Rec}{Pre + Rec}$$

5.3 Result

To assess the performance, the performance analysis phase compares the results of several ML methods, including RF, NB, KNN and DT. Self-created YouTube video set and publicly available USVideo dataset was employed to train and assess the model in order to determine how well they performed in review classifications. In addition, data set flexibility was reinforced using a generative Accuracy, accuracy, and memory are all significantly improved by the undesirable network model GAN. This study showed how leveraging generic side effects to create false data can improve the performance of ML algorithms. The performance analysis comprises two key components: evaluation on a custom YouTube video dataset and validation on the USVideo benchmark dataset.

5.3.1 Confusion matrix

Class 0 – Fake: Videos that are intentionally manipulated, fabricated, or artificially generated, including deepfakes or synthetically altered content.

Class 1 – Valid: Authentic and legitimate videos with no evidence of manipulation or misleading alterations.

Class 2 – Unidentified: Videos whose authenticity cannot be confidently determined—either due to incomplete metadata, ambiguous patterns, or unclear content characteristics.

Table 2 illustrates the confusion matrix results for models incorporating and excluding the GAN model, tested on the USVideo dataset and the self-developed YouTube Video Review dataset. The confusion matrix results for the USVideos dataset clearly demonstrate that the RF classifier performs significantly better than the other ML models, both with and without GAN augmentation. Without GAN-based data generation, RF already shows strong classification performance, correctly identifying most Valid (Class 1) and Unidentified (Class 2) samples, though some confusion remains in distinguishing Fake (Class 0) from the other categories. After applying GAN augmentation, the RF confusion matrix shows a substantial improvement, with Fake, Valid, and Unidentified classes being predicted with far higher precision and fewer misclassifications. In particular, Valid videos achieve exceptionally high recognition (2737 correct), and misclassification rates across all classes drop sharply, demonstrating that GAN-generated samples help the RF model learn more discriminative boundaries.

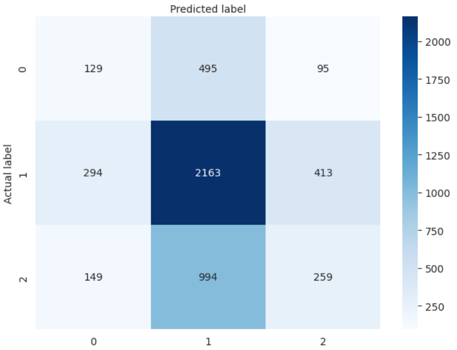
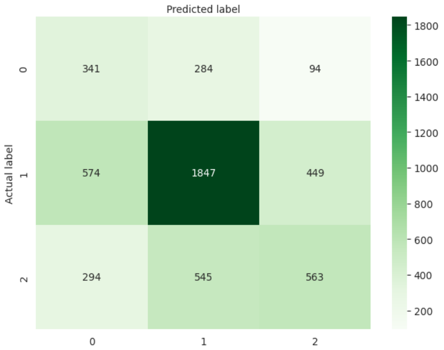
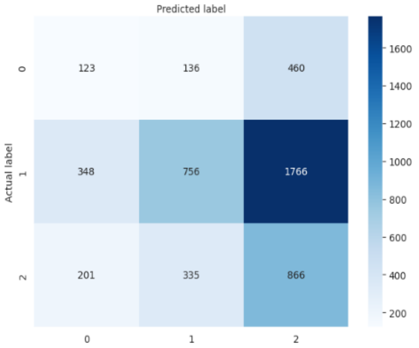
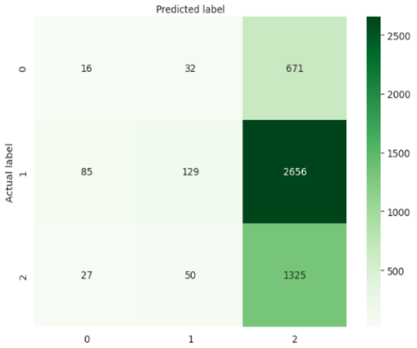
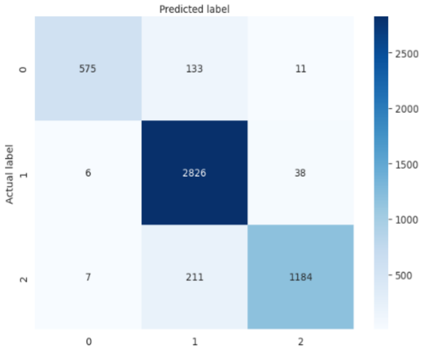
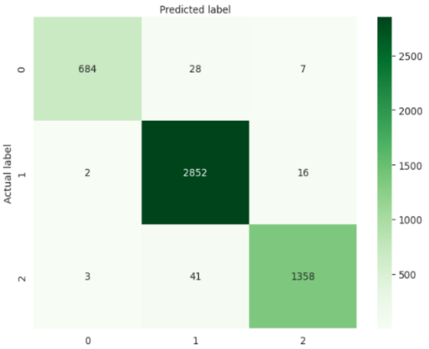
By comparison, the other ML models show weaker performance.

Naïve Bayes (NB) performs the worst due to its strong independence assumptions, which do not hold for the highly correlated metadata features in the video dataset, resulting in high confusion between Fake and Valid classes.

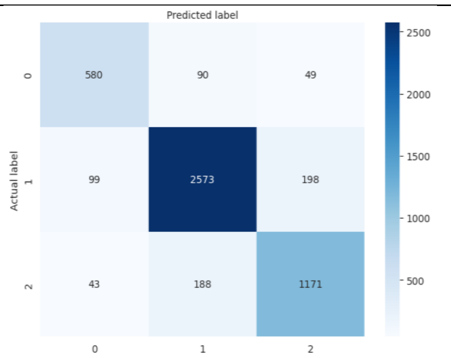
DT and K-Nearest Neighbors (KNN) perform moderately well but still fall short of RF. DT tends to overfit and struggles with ambiguous borderline cases, while KNN is sensitive to feature scaling and distance distortions, causing misclassification when class distributions overlap. Overall, RF's ensemble learning, robustness to noise, and ability to capture nonlinear feature interactions make it the most effective and reliable model for distinguishing Fake, Valid, and Unidentified videos in the USVideos dataset, especially when enhanced with GAN-based augmentation.

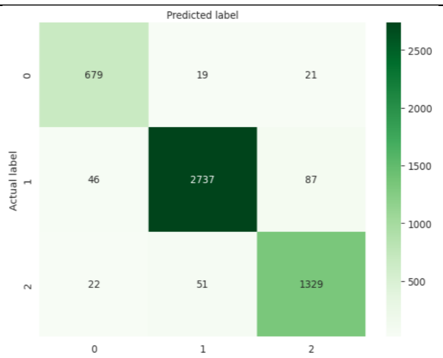
The RF confusion matrix with GAN augmentation on the Self-Created YouTube dataset shows strong performance across all classes. The model correctly identifies most Fake videos (828), achieves high accuracy for Valid videos (972), and reliably classifies Unidentified videos (835). Although some confusion remains between Valid and Unidentified classes, overall, the GAN-enhanced RF model provides clear improvements in distinguishing Fake, Valid, and Unidentified content.

Table 2. Confusion Matrix

Model	Confusion Matrix (Without GAN)	Confusion Matrix (With GAN)
	USVideo Dataset	
KNN		
NB		
DT		

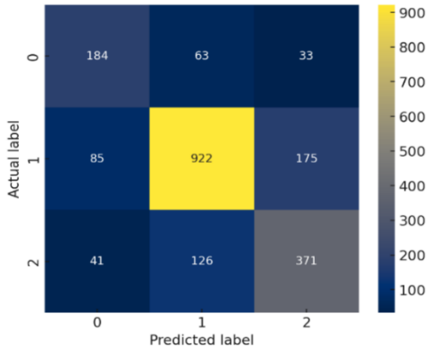
RF

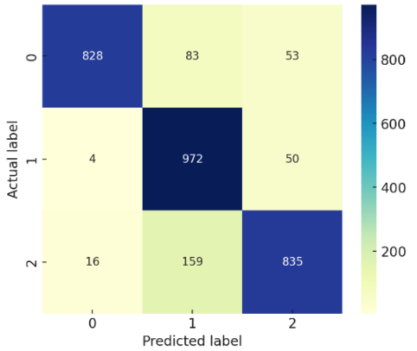




Self-created YouTube Dataset

RF





Note: USVideo = US Trending YouTube Video Dataset; GAN = generative adversarial network; KNN = K-Nearest Neighbors; NB = Naïve Bayes; DT = Decision Tree; RF = Random Forest.

5.4 Comparative analysis

Our study carried out a thorough assessment of several ML algorithms, methodically contrasting their performance across important criteria such as F1-score, accuracy recall, and precision. The empirical findings from this extensive comparison provide insightful information for choosing the best strategies for applications using video analysis. Table 3 shows the ML models comparative Analysis.

Table 3. ML models comparative analysis

Dataset	Metric	DT	KNN	NB	RF
USVideo Dataset (With GAN Model)	Accuracy	96	56	30	98
	F1-Score	97	58	16	98
	Precision	95	56	45	99
	Recall	97	56	31	98
USVideo Dataset (Without GAN Model)	Accuracy	89	50	35	93
	F1-Score	86	50	37	94
	Precision	87	48	44	92
	Recall	87	49	34	90
YouTube_Self_Created Dataset (With GAN Model)	Accuracy	76	45	34	89
	F1-Score	78	47	19	87
	Precision	77	47	32	90
	Recall	76	46	32	87
YouTube_Self_Created Dataset (Without GAN Model)	Accuracy	76	51	36	81
	F1-Score	69	35	33	79
	Precision	70	36	37	89
	Recall	70	36	41	72

Note: ML = machine learning; DT = Decision Tree; KNN = K-Nearest Neighbors; NB = Naïve Bayes; RF = Random Forest. USVideo = US Trending YouTube Video Dataset; GAN = generative adversarial network.

5.4.1 US Trending YouTube Video Dataset result analysis

In Figure 12, the accuracy comparison graph is displayed for ML models that include the GAN model and those that do not include it. As we can see in Figure 12, nearly all ML algorithms experience about a 5% boost in accuracy when they incorporate the GAN model, with the exception of the KNN model. The RF model stands out with an impressive accuracy of 98%, surpassing all the other models, which suggests that the GAN model really enhances the overall performance of ML models.

5.4.2 Self-created YouTube (Sentiment) dataset result analysis

Furthermore, the findings of our experiments show that the incorporation of a GAN-based model resulted in an improvement in model performance across the majority of ML techniques. With the exception of KNN model, all classifiers showed ~5% accuracy gains when trained on GAN-augmented data from the YouTube Video dataset. This consistent improvement suggests that GAN-generated samples effectively enhance the discriminative capability of ML models by addressing potential data limitations. Among all evaluated approaches, RF achieved superior performance with

88% accuracy, outperforming other algorithms in both standard and GAN-augmented configurations. Figure 13 provides a comprehensive visual comparison of these accuracy measurements, contrasting baseline performance against GAN-enhanced results across all models.

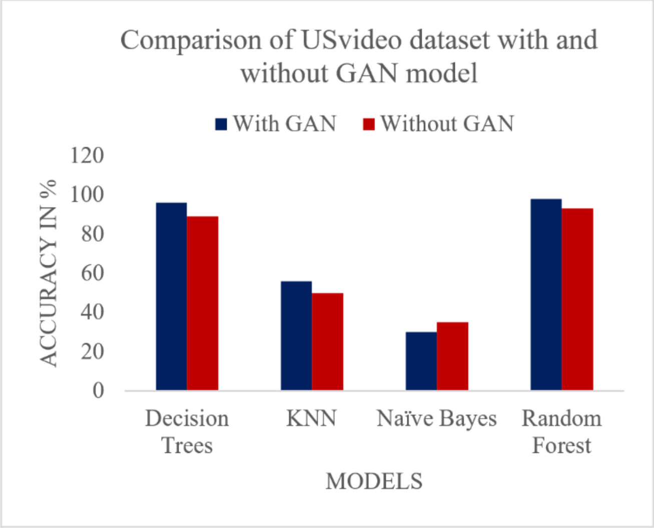


Figure 12. Performance comparison of USVideo dataset with and without GAN model
Note: USVideo = US Trending YouTube Video Dataset; GAN = generative adversarial network.

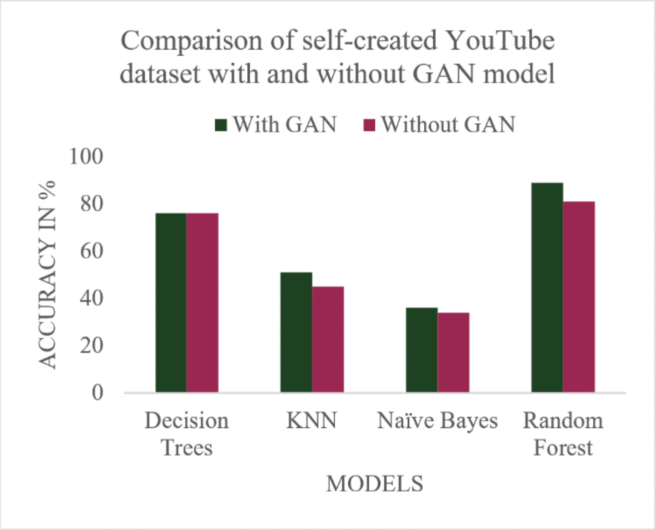


Figure 13. Performance comparison of self-created YouTube video review dataset with and without GAN model
Note: GAN = generative adversarial network.

These findings highlight both the value of GAN-based data augmentation for video classification tasks and the particular effectiveness of ensemble methods like RF for this application domain.

5.5 Discussion

Our experimental results demonstrated significant performance disparities amongst a number of ML algorithms when used to video sentiment analysis. Using common metrics like precision, accuracy, recall, and F1-score, we thoroughly assessed four algorithms: KNN, NB, DT, and RF. The confusion matrix investigation yielded particularly useful information regarding the classification performance of each system for YouTube video evaluations across both datasets. Notably, the incorporation of GANs demonstrated measurable improvements, enhancing all evaluation metrics compared to traditional ML approaches alone. This performance boost underscores GANs' effectiveness in both dataset augmentation and sentiment analysis tasks. Our findings highlight two critical considerations for video sentiment analysis: (1) the importance of algorithm selection, with RF emerging as particularly effective, and (2) the value of advanced data augmentation techniques like GANs. The comparative analysis of different algorithmic approaches and dataset configurations has yielded actionable insights that can inform future sentiment analysis projects, particularly in understanding the strengths and limitations of various methodologies for video content analysis.

6. CONCLUSIONS

Key developments in sentiment analysis techniques created especially for extracting and annotating metadata from a variety of YouTube video genres are highlighted in this study. Our suggested methodology has demonstrated notable gains in sentiment analysis performance and fake video detection through thorough investigation on a variety of datasets, especially when GAN is integrated. The approach's resilience and flexibility have been improved by using both a self-made dataset and the publicly accessible USVideo dataset. The accuracy and F1-scores of all ML models significantly improved by approximately 5 to 6% once GANs were added. With 98% accuracy on the USVideo dataset and 88% accuracy on the YouTube review dataset generated by the algorithm itself, the RF algorithm performed better than the others. These outcomes highlight how well the suggested system does sentiment analysis from a variety of data sources. Furthermore, faster and more effective detection of bogus videos is made possible by the methodology's capacity to collect information and analyze comments, especially in the case of limited reviews. It is a useful tool for controlling and enhancing the user experience on websites like YouTube since it facilitates more effective content management and recommendation systems.

REFERENCES

[1] Appel, G., Grewal, L., Hadi, R., Stephen, A.T. (2020). The future of social media in marketing. Journal of the Academy of Marketing Science, 48(1): 79-95. https://doi.org/10.1007/s11747-019-00695-1

[2] Berger, L.M., Kerkhof, A., Mindl, F., Münster, J. (2025). Debunking “fake news” on social media: Immediate and short-term effects of fact-checking and media literacy interventions. Journal of Public Economics, 245: 105345. https://doi.org/10.1016/j.jpubeco.2025.105345

[3] Oprea, S.V., Bara, A. (2025). Fake news in elections: Leveraging large language models using semantic analyses to extract insights from academic research. Connection Science, 37(1): 2587447. https://doi.org/10.1080/09540091.2025.2587447

[4] Dekker, C.A., Baumgartner, S.E., Sumter, S.R. (2025). For you vs. for everyone: The effectiveness of algorithmic personalization in driving social media

engagement. Telematics and Informatics, 101: 102300. https://doi.org/10.1016/j.tele.2025.102300

[5] Sudhakar, M., Kaliyamurthie, K.P. (2024). Detection of fake news from social media using support vector machine learning algorithms. Measurement: Sensors, 32: 101028. https://doi.org/10.1016/j.measen.2024.101028

[6] Ozbay, F.A., Alatas, B. (2020). Fake news detection within online social media using supervised artificial intelligence algorithms. Physica A: Statistical Mechanics and its Applications, 540: 123174. https://doi.org/10.1016/j.physa.2019.123174

[7] Chen, Z., Wang, L., Zhu, X., Dietze, S. (2023). TSNN: A topic and structure aware neural network for rumor detection. Neurocomputing, 531: 114-124. https://doi.org/10.1016/j.neucom.2023.02.016

[8] Madani, M., Motameni, H., Roshani, R. (2024). Fake news detection using feature extraction, natural language processing, curriculum learning, and deep learning. International Journal of Information Technology & Decision Making, 23(3): 1063-1098. https://doi.org/10.1142/S0219622023500347

[9] Chang, Q., Li, X., Duan, Z. (2024). A novel approach for rumor detection in social platforms: Memory-augmented transformer with graph convolutional networks. Knowledge-Based Systems, 292: 111625. https://doi.org/10.1016/j.knosys.2024.111625

[10] Geto, A.D., Emiru, E.D., Seid, N.E., Tessema, A.B., Ahmed, B.Y. (2025). Multimodal based Amharic fake news detection using CNN and attention-based BiLSTM. Scientific Reports, 15(1): 34447. https://doi.org/10.1038/s41598-025-17579-w

[11] Saifullah, S., Dreżewski, R., Dwiyanto, F.A., Aribowo, A.S., Fauziah, Y., Cahyana, N.H. (2024). Automated text annotation using a semi-supervised approach with meta vectorizer and machine learning algorithms for hate speech detection. Applied Sciences, 14(3): 1078. https://doi.org/10.3390/app14031078

[12] Mahalakshmi, V., Shenbagavalli, P., Raguvaran, S., Rajakumareswaran, V., Sivaraman, E. (2024). Twitter sentiment analysis using conditional generative adversarial network. International Journal of Cognitive Computing in Engineering, 5: 161-169. https://doi.org/10.1016/j.ijcce.2024.03.002

[13] Bronakowski, M., Al-Khassaweneh, M., Al Bataineh, A. (2023). Automatic detection of clickbait headlines using semantic analysis and machine learning techniques. Applied Sciences, 13(4): 2456. https://doi.org/10.3390/app13042456

[14] Welbourne, D.J., Grant, W.J. (2016). Science communication on YouTube: Factors that affect channel and video popularity. Public Understanding of Science, 25(6): 706-718. https://doi.org/10.1177/0963662515572068

[15] Gul, R., Bashir, M. (2024). Feature selection for sentiment analysis using hybrid multiobjective evolutionary algorithm. Journal of Intelligent & Fuzzy Systems, 46(4): 8917-8932. https://doi.org/10.3233/JIFS-234615

[16] Sharma, H.D., Budaraju, R.R., Kumar, N., Kumar, V., Rathore, N.C., Babu, G.R., Dhaka, A. (2025). Sentiment classification via improved feature selection using Boolean operator-based particle swarm optimization. Scientific Reports, 15(1): 38923. https://doi.org/10.1038/s41598-025-22894-3

[17] Zuo, E., Zhao, H., Chen, B., Chen, Q. (2020). Context-specific heterogeneous graph convolutional network for implicit sentiment analysis. IEEE Access, 8: 37967-37975. https://doi.org/10.1109/ACCESS.2020.2975244

[18] Li, Z., Zou, Z. (2024). Punctuation and lexicon aid representation: A hybrid model for short text sentiment analysis on social media platform. Journal of King Saud University-Computer and Information Sciences, 36(3): 102010. https://doi.org/10.1016/j.jksuci.2024.102010

[19] Balli, C., Guzel, M.S., Bostanci, E., Mishra, A. (2022). Sentimental analysis of Twitter users from Turkish content with natural language processing. Computational Intelligence and Neuroscience, 2022(1): 2455160. https://doi.org/10.1155/2022/2455160

[20] Kumar, S., Mundra, K., Verma, R. (2024). Sentiment classification of multidomain reviews using machine learning models. In International Conference on Emerging Trends in Expert Applications & Security, Springer, Singapore, pp. 93-104. https://doi.org/10.1007/978-981-97-3991-2_8

[21] Jayaraman, A.K., Trueman, T.E., Ananthakrishnan, G., Murugappan, A., Cambria, E., Mitra, S. (2025). Aspect category and sentiment polarity detection using a permutation language-based transformer model. Procedia Computer Science, 258: 4179-4189. https://doi.org/10.1016/j.procs.2025.04.668

[22] TS, S.M., Sreeja, P.S., Ram, R.P. (2022). Fake news article classification using random forest, passive aggressive, and gradient boosting. In 2022 International Conference on Connected Systems & Intelligence (CSI), Trivandrum, India, pp. 1-6. https://doi.org/10.1109/CSI54720.2022.9924131

[23] Ramezani, E.B. (2025). Sentiment analysis applications using deep learning advancements in social networks: A systematic review. Neurocomputing, 634: 129862. https://doi.org/10.1016/j.neucom.2025.129862

[24] Yang, L., Li, Y., Wang, J., Sherratt, R.S. (2020). Sentiment analysis for E-commerce product reviews in Chinese based on sentiment lexicon and deep learning. IEEE Access, 8: 23522-23530. https://doi.org/10.1109/ACCESS.2020.2969854

[25] Li, P., Liang, T., An, N., Zhang, L., Wu, X., Wang, X., Lu, J. (2023). A Sentiment classification model based on attention TIG-CNN-BiGRU. In 2023 International Conference on Culture-Oriented Science and Technology (CoST), Xi'an, China, pp. 78-83. https://doi.org/10.1109/CoST60524.2023.00025.

[26] Hermawan, S., Fatichah, C., Kamal, I.M. (2025). A comparative study of GAN-based methods for tabular synthetic data generation. In 2025 15th International Conference on Information & Communication Technology and System (ICTS), Denpasar, Indonesia, pp. 1-6. https://doi.org/10.1109/ICTS67612.2025.11369547

[27] Deressa, D.W., Mareen, H., Lambert, P., Atnafu, S., Akhtar, Z., Van Wallendael, G. (2025). GenConViT: Deepfake video detection using generative convolutional vision transformer. Applied Sciences, 15(12): 6622. https://doi.org/10.3390/app15126622

[28] Petmezas, G., Vanian, V., Konstantoudakis, K., Almaloglou, E.E., Zarpalas, D. (2025). Video deepfake detection using a hybrid CNN-LSTM-Transformer model for identity verification. Multimedia Tools and Applications, 84(33): 40617-40636.

https://doi.org/10.1007/s11042-024-20548-6

[29] Wang, L., Zhao, J., Zhang, X., Guo, X., et al. (2025). ERF-BA-TFD+: A multimodal model for audio-visual deepfake detection. Vicinagearth, 2(1): 1-12. https://doi.org/10.1007/s44336-025-00021-0

[30] Alrashoud, M. (2025). Deepfake video detection methods, approaches, and challenges. Alexandria Engineering Journal, 125: 265-277. https://doi.org/10.1016/j.aej.2025.04.007

[31] Hasan, M.S., Sarker, B., Shrestha, D., Shrestha, R., Shrestha, S.N. (2023). Trending YouTube video analysis. https://doi.org/10.21203/rs.3.rs-2548456/v1

[32] Onan, A., Korukoğlu, S., Bulut, H. (2016). Ensemble of keyword extraction methods and classifiers in text classification. Expert Systems with Applications, 57: 232-247. https://doi.org/10.1016/j.eswa.2016.03.045

[33] Albalawi, R., Yeap, T.H., Benyoucef, M. (2020). Using topic modeling methods for short-text data: A comparative analysis. Frontiers in Artificial Intelligence, 3: 42. https://doi.org/10.3389/frai.2020.00042