ISSN 2723-6471

Analyzing Audience Sentiments in Digital Comedy: A Study of YouTube Comments Using LSTM Models

Supriyono[1,2], Aji Prasetya Wibawa[3,*], Suyono[4], Fachrul Kurniawan[5]

[1,3]*Department of Electrical Engineering and Informatics, Faculty of Engineering, Universitas Negeri Malang, Malang 65145, Indonesia*

[4]*Department of Indonesian Literature, Faculty of Letters, Universitas Negeri Malang, Malang 65145, Indonesia*

[2,5]*Informatics Engineering, Faculty of Science and Technology, Universitas Islam Negeri Maulana Malik Ibrahim Malang, Indonesia*

(Received: August 16, 2024; Revised: September 21, 2024; Accepted: October 02, 2024; Available online: October 20, 2024)

Abstract

The main objective of this paper is to analyze audience sentiment towards stand-up comedy content on the YouTube platform, specifically comments on stand-up comedy videos from Kompas TV, using the Long Short-Term Memory (LSTM) method. This research contributes significantly to a deeper understanding of how audiences engage with humorous content through a sentiment analysis approach that uses the LSTM model, which can capture complex nuances in humorous content, such as sarcasm, irony, and cultural references. The research methodology involves crawling data from YouTube, where user comments are extracted and processed through several stages of data cleaning, such as removing duplicate content, text normalization, and irrelevant comments. Once the data is prepared, the LSTM model is trained to analyze positive, negative, and neutral sentiments with varying accuracy rates of 85% for positive sentiment, 80% for negative sentiment, and 78% for neutral sentiment. The main results show that the LSTM model successfully classifies sentiments, although it needs help handling the more ambiguous neutral sentiments. The figures and tables included in this study illustrate the relationship between the number of views, likes, and the sentiment classification of the comments. One notable finding is a strong positive correlation between the number of views and video likes. The conclusions of this study underscore the need for model improvements to handle neutral sentiment better and capture the complexity of humor content. The implications of this research are useful for content creators and digital marketers in understanding and responding to audience preferences more effectively. They also pave the way for further research in sentiment analysis on more specific content genres on digital platforms.

Keywords: Audience Engagement, Digital Media, Long Short-Term Memory, Sentiment Analysis, Stand-Up Comedy

1. Introduction

The emergence of digital platforms like YouTube has fundamentally transformed how entertainment content is viewed and engaged worldwide [1], [2], [3]. Stand-up comedy has gained immense popularity, attracting millions of viewers interacting with the material through likes, comments, and shares. Nevertheless, comprehending the audience's reception of stand-up comedy content continues to be an intricate endeavor. Comedy is a type of subjective media, meaning that people have different opinions about it. The success of a joke or performance can significantly differ based on cultural background, personal experiences, and even the timing of the content. Although there is a large amount of audience engagement data on platforms such as YouTube, there is a need for more ways to evaluate and understand these attitudes appropriately. This limitation hinders content makers' capacity to customize their material efficiently.

The latest progress in sentiment analysis and natural language processing (NLP) has greatly enhanced the capacity to evaluate text-based data, allowing researchers to detect and classify the sentiments conveyed in user comments. Long Short-Term Memory (LSTM) networks, a recurrent neural network (RNN), have demonstrated potential in effectively capturing [4], [5], [6], [7] the intricacies of human emotions in sequential data, such as comments. Nevertheless, current research frequently concentrates on sentiment analysis in sentiment analysis in various fields, neglecting to properly

***Corresponding author: Aji Prasetya Wibawa (aji.prasetya.ft@um.ac.id)**
DOI: https://doi.org/10.47738/jads.v5i4.393

tackle the difficulties of assessing funny material. Moreover, although specific studies have examined sentiment analysis on social media or review platforms, there needs to be more research in applying these methods, especially in stand-up comedy, where the understanding of sentiments may vary due to the intrinsically comedic nature of the content.

This study aims to narrow the existing divide by utilizing sophisticated sentiment analysis tools, particularly the LSTM model, to examine viewers' responses to stand-up comedy videos on YouTube. The novelty of this research lies in its concentrated utilization of sentiment analysis in the comic genre, which poses unique and intriguing difficulties, such as identifying sarcasm, comedy, and irony components that conventional sentiment analysis models frequently find challenging. In addition, this study incorporates a thorough data crawling methodology, gathering not only comments but also important engagement indicators such as views and likes, which offers a more extensive framework for measuring audience sentiment. This research aims to enhance sentiment analysis's accuracy and contextual awareness by improving data pretreatment techniques and developing model architectures designed explicitly for funny content.

The primary contribution of this study is developing a specialized framework for sentiment analysis that enhances our understanding of how audiences engage with stand-up comedy content on digital platforms. The insights gained from this research will be valuable not only for academic purposes but also for content providers and digital marketers looking to improve their strategies by leveraging audience feedback. Furthermore, this study advances the NLP field by demonstrating the effectiveness of LSTM models in a specialized yet crucial area. This could pave the way for further exploration in sentiment analysis tailored to specific content genres, offering a more comprehensive understanding of audience reactions, and potentially improving content creation, audience engagement, and digital media analysis.

2. The Proposed Method/Algorithm

In this section, we present the methodology employed for data crawling, sentiment analysis, and the application of the LSTM algorithm to analyze stand-up comedy content from YouTube. The first step involves data crawling, where we utilize web scraping techniques to gather a comprehensive dataset of stand-up comedy videos by focusing on specific keywords related to the genre [8]. This process collects not only video URLs, titles, and descriptions but also vital audience interaction metrics such as views, likes, dislikes, and comments [9]. We leverage Python libraries like Beautiful Soup and Scrapy for efficient scraping while ensuring compliance with YouTube's terms of service. Following data collection, we perform essential pre-processing steps to clean and prepare the data, including removing duplicates, irrelevant content, and non-English comments, as well as tokenizing and normalizing the text. The core of our methodology lies in the sentiment analysis conducted using the LSTM algorithm, where we train a neural network model on the prepared dataset [10], [11], [12]. This model is designed to capture long-term dependencies in the sequential data of comments, allowing for accurate sentiment predictions. Finally, we analyze the results to interpret audience responses to various comedic styles, providing insights into engagement dynamics within the stand-up comedy genre.

2.1. Data Crawling

The first step in our proposed method involves data crawling, which is crucial for collecting a comprehensive dataset of stand-up comedy videos. We utilize web scraping techniques to gather data from YouTube, focusing on specific keywords related to stand-up comedy [13], [14]. The data collected includes video URLs, titles, descriptions, metadata such as publication dates, and audience interaction metrics including views, likes, dislikes, and comments. We implement Python libraries like Beautiful Soup and Scrapy for efficient web scraping, ensuring compliance with YouTube's terms of service to avoid potential legal issues. This approach allows us to accumulate a rich dataset spanning various comedians and performances, which forms the basis for our subsequent analyses.

The data crawling process is the initial and crucial step in this research, aimed at gathering a comprehensive dataset of stand-up comedy content from YouTube, specifically from Kompas TV's channel. Using web scraping techniques, we systematically collected data by focusing on keywords related to the stand-up comedy genre. This process allowed us to gather not only the video URLs, titles, and descriptions but also essential audience interaction metrics such as views, likes, dislikes, and comments. To ensure efficiency and compliance with YouTube's terms of service, we employed

Python libraries like Beautiful Soup and Scrapy, which facilitated the extraction of the required data while adhering to ethical standards.

In addition to collecting basic metadata, the data crawling process involved capturing more detailed engagement metrics that could provide deeper insights into audience behavior. This included extracting the exact number of comments, the publication dates of the videos, and the frequency of likes and dislikes over time. By collecting this range of data, we aimed to build a rich dataset that reflects various aspects of audience interaction with stand-up comedy content. The diversity and volume of the data collected through this process formed a solid foundation for subsequent analysis, allowing us to explore patterns in audience engagement and sentiment.

To maintain the integrity and quality of the data, we implemented several checks during the crawling process. For example, we ensured that the data was collected over a consistent time frame, focusing on videos published between 2020 and 2023. This time-bound approach helped us avoid inconsistencies due to changes in YouTube's algorithm or variations in audience behavior over different periods. Additionally, by carefully selecting the source videos and ensuring that the data collected was relevant to the research objectives, we minimized the risk of including irrelevant or noisy data, thereby enhancing the accuracy and reliability of the subsequent sentiment analysis [15], [16].

2.2. Data Preprocessing

The first step in data preprocessing for sentiment analysis is removing duplicate comments. YouTube video comments, especially from stand-up comedy videos, often contain many identical or highly similar responses due to the nature of online interactions. These duplicates can skew the results and introduce bias in the analysis, affecting the accuracy of the model's predictions. Identifying and removing these repeated comments ensures that the dataset comprises unique entries, essential for training a model that accurately captures the sentiment patterns within the audience's responses. This step helps maintain the integrity of the data and contributes to more reliable analysis.

After removing duplicates, the dataset is refined by filtering out irrelevant comments and non-Indonesian language responses. Stand-up comedy often prompts discussions and reactions that only relate directly to the content, and comments unrelated to the topic distort the analysis. By excluding these comments in languages other than Indonesian, the model can focus solely on sentiments related to the comedy content. This targeted filtering improves the dataset's relevance, ensuring the sentiment analysis accurately reflects audience reactions to the specific comedic material studied.

2.3. Sentiment Analysis Using LSTM

Sentiment analysis utilizing LSTM networks is an advanced methodology in NLP. It concentrates on identifying patterns in sequential data like comments or reviews, where context is essential. LSTM in this arena because they can capture long-term dependencies in data, addressing the shortcomings of conventional RNN, which are hindered by the vanishing gradient problem. This capacity enables LSTM to comprehend intricate statements, particularly those including sarcasm, irony, or subtle contexts, rendering them optimal for sentiment analysis tasks.

The LSTM model has memory cells and input, forget, and output gates that collectively regulate the information flow inside the network. These components function to preserve essential information while eliminating extraneous material. In sentiment analysis, textual data is subjected to preprocessing stages, including tokenization, noise elimination, and conversion into numerical vectors through embedding techniques such as Word2Vec or GloVe. After preprocessing, the LSTM model is trained using annotated sentiment data, modifying its weights via backpropagation and optimization methods to enhance classification accuracy.

Upon completion of training, the LSTM model can predict sentiment in new data by identifying patterns acquired during the training phase. The performance is assessed using criteria such as accuracy, precision, recall, and F1-score, especially in complex scenarios like sarcasm. This approach has been utilized to examine audience reactions to stand-up comedy on YouTube, offering insights into emotional engagement and assisting content providers in enhancing their tactics. This methodology provides significant insights for media analysis research and practical implementations within the entertainment sector.

 ISSN 2723-6471

3. Methodology

This study employed A web crawling technique to collect stand-up comedy footage from Kompas TV's YouTube channel for data collection. This approach was designed to produce a comprehensive dataset covering 2020 to 2023. The selection of this period was motivated by several considerations. To begin with, the period included the COVID-19 epidemic, which profoundly affected the essence and substance of comedy, as numerous comedians modified their material to mirror the social and cultural changes during this period. Furthermore, the objective was to examine current patterns in stand-up comedy and extract sentiment pertinent to the present audience, assuring that the acquired insights are applicable and timely. The YouTube API effectively retrieved metadata for each stand-up comedy video, encompassing titles, descriptions, views, likes, comments, and upload dates.

Commentary was gathered to record audience reactions and involvement with the comical material, offering a rich stratum for sentiment research. Conscientious adherence to YouTube's terms of service and ethical standards was maintained throughout the data-collecting process to ensure legality and regulatory compliance. This measure ensured the preservation of the project's integrity and the ethicality of the data extraction procedure. The data retrieval and validation process from YouTube and the accuracy verification are illustrated in figure 1, offering a concise and organized summary of the employed methodology.

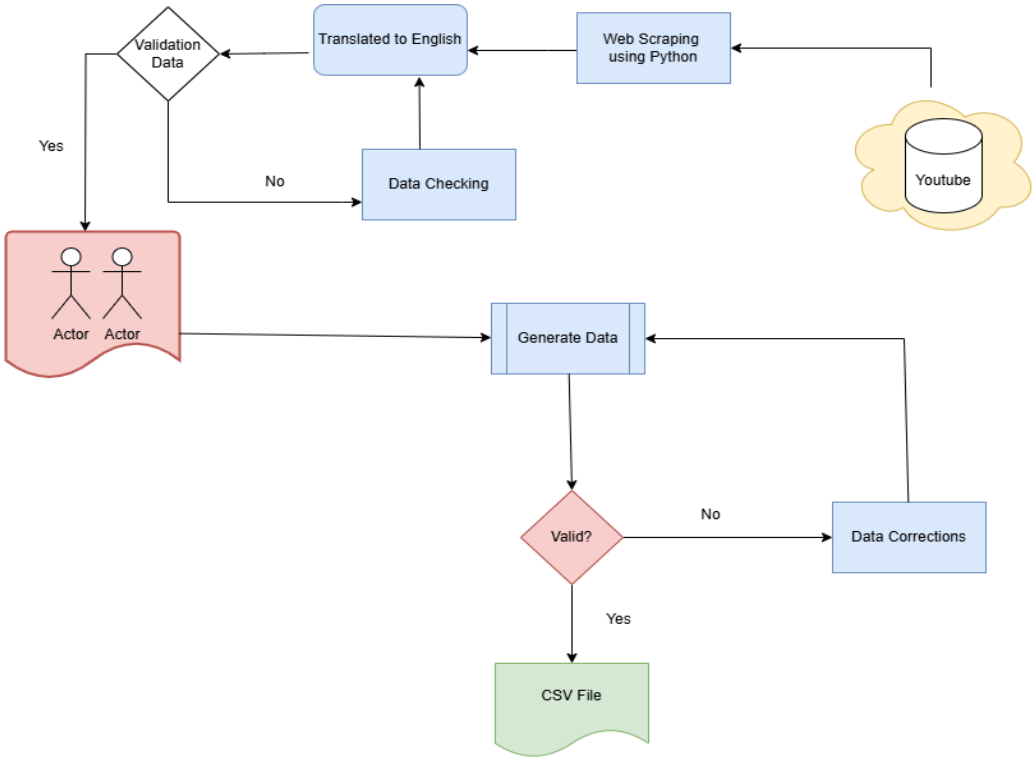


Figure 1. Method of Crawling Data

3.1. LSTM Architecture

The LSTM architecture excels in modeling sequential data, overcoming the limitations of traditional RNN, particularly in capturing long-term dependencies. At the core of the LSTM unit is the memory cell, which stores information over time, allowing the model to recall relevant data from previous time steps. The architecture is supported by three main gates: the input gate, which regulates the information stored in the memory cell; the forget gate, which eliminates irrelevant data; and the output gate, which controls the information passed to the next layer [17]. These components maintain context across long sequences, making LSTM ideal for tasks like sentiment analysis.

However, training LSTM models for complex language tasks, such as recognizing sarcasm and irony, presents challenges. Sarcasm often disguises negative sentiments as compliments, while irony can veil criticism in positive words. The LSTM model is trained with diverse examples of sarcastic and ironic comments to tackle these nuances

[18]. During data preprocessing, special attention is given to these expressions to avoid misinterpretation in sentiment analysis. Advanced tokenization and embedding techniques are also employed to capture semantic relationships, improving the model's ability to identify complex humor patterns in text.

Despite these optimizations, recognizing subtle irony and sarcasm remains challenging. Additional strategies, such as integrating Transformer-based models or contextual word embeddings like BERT and ELMo, have been adopted to enhance the model's understanding of intricate humor contexts [19]. These approaches significantly improve the LSTM's accuracy in classifying ambiguous comedic sentiments, allowing for more precise and nuanced sentiment analysis results in diverse applications.

In an LSTM unit, the gates play a crucial role in regulating the flow of information, ensuring the model can effectively learn from sequences. Each gate uses a sigmoid activation function to generate values between 0 and 1, which serve as control signals for the information passing through the unit [20]. The input gate controls how much new input should be added to the memory cell by evaluating the current input and the previous hidden state, assigning a value to each input part to indicate its importance. A value near 1 means the information is relevant and should be retained. In contrast, a value near 0 suggests it should be ignored, allowing the LSTM to integrate significant features of the current input selectively.

On the other hand, the forget gate determines which information in the memory cell should be discarded. Assessing the current input and the previous hidden state produces a vector that dictates what to keep and remove, allowing the LSTM to discard irrelevant or outdated information while maintaining the necessary context [21]. Finally, the output gate decides what information from the memory cell should be sent to the next layer by combining the current input and the memory state, ensuring that only the most relevant data is passed forward. The coordinated function of these gates enables LSTMs to handle long sequences, making them highly effective in tasks like sentiment analysis and natural language processing by preserving vital context and eliminating unnecessary data.

3.2. Building the LSTM Model

Building an LSTM model involves several key steps, including the network architecture's design, configuration, and implementation. The process starts by selecting a suitable framework or library like TensorFlow or Keras, which provides essential tools to build and train neural networks efficiently. These frameworks simplify complex tasks, allowing developers to focus on designing effective architectures [22]. With the framework chosen, the next step is to define the model architecture, starting with an embedding layer that transforms input text into dense vector representations. This layer helps capture the semantic relationships between words, making it easier for the model to understand the context. The embedding layer is followed by one or more LSTM layers, which learn temporal patterns and long-range dependencies from the sequential data.

After setting up the LSTM layers, additional components, such as dropout layers, may be added to prevent overfitting. Dropout layers randomly deactivate some input units during training, enhancing the model's generalization capabilities. Once the LSTM layers have processed the input, a dense output layer is typically added to make predictions, with an activation function such as sigmoid for binary classification tasks or softmax for multi-class problems [23]. The design of the model architecture must be aligned with the specific issue being addressed, like sentiment analysis, where understanding the sequence of words and their emotional impact is crucial.

The final stage in building an LSTM model involves configuring the hyperparameters, including the optimizer (e.g., Adam or RMSprop), learning rate, batch size, and number of training epochs [24]. Hyperparameter tuning is critical in ensuring the model converges efficiently and performs optimally. Once the architecture and hyperparameters are set, the model is compiled and trained, adjusting its internal weights to minimize the loss function using backpropagation. Monitoring validation data during training helps track the model's performance and adjust as needed. In summary, by carefully designing the architecture, embedding techniques, and hyperparameters, practitioners can build robust LSTM models capable of analyzing sequential data and uncovering complex patterns in applications such as sentiment analysis.

3.3. Training the LSTM Model

The initial step in training the LSTM model involves preparing the dataset by dividing the preprocessed YouTube comments from stand-up comedy videos into separate sets for training, validation, and testing [25]. The architecture of the LSTM model is established, beginning with an embedding layer that transforms textual data into compact vector representations. The function of this layer is to capture the semantic connections between words, enabling the model to enhance its comprehension of the context present in comments. One or more LSTM layers are incorporated after the embedding layer. These layers capture temporal data relationships, allowing the model to preserve and utilize information from preceding stages in the sequence for sentiment prediction.

In multi-class sentiment analysis tasks such as this, the training method entails optimizing the model using a loss function, commonly categorical cross-entropy. The optimizer, such as Adam or RMSprop, modifies the model weights according to the computed loss, enhancing performance across consecutive iterations or epochs [26]. The hyperparameters, including the learning rate, batch size, and number of epochs, are individually adjusted to optimize the model's performance. Throughout the training process, the model's performance is assessed on the validation set, enabling modifications to be implemented to achieve a harmonious equilibrium between accuracy and generalization.

Once the model has undergone training on the dataset, its performance is evaluated using the test set to quantify its effectiveness in predicting new data. Key measures like accuracy, precision, recall, and F1-score assess the model's efficacy in categorizing comments as positive, negative, or neutral thoughts [27]. LSTM models are well-suited for analyzing hilarious material due to their exceptional ability to capture long-range dependencies and contextual subtleties in sequential data. This is crucial for comprehending intricate sentiments such as sarcasm or irony in comedy. In contrast to models such as CNNs, which are more suitable for image data or shorter text sequences, or Transformer-based models, which prioritize parallel processing of data, LSTMs offer a good compromise between complexity and the capacity to handle temporal patterns, rendering them well-suited for this analysis.

4. Results and Discussion

This study focuses on extracting or crawling data from Kompas TV's YouTube channel, explicitly targeting stand-up comedy videos, and conducting sentiment analysis on viewer comments using the LSTM approach. The research is driven by the increasing popularity of digital platforms as entertainment media, presenting an opportunity for Kompas TV to expand its audience by producing humorous content. However, understanding audience reactions to this content remains a significant challenge. Therefore, this study aims to gather comments from stand-up comedy videos on Kompas TV's YouTube account and analyze the sentiments expressed in those comments using the LSTM method, known for its effectiveness in processing sequential text data. The study will evaluate the LSTM model's performance in sentiment classification, and the findings will provide insights into how these results can enhance content strategies for Kompas TV and creators of comedic content.

4.1. Sentiment Analysis Accuracy

Figure 2 of the research offers significant insights by juxtaposing the social media engagement metrics of stand-up comedy videos on Kompas TV's YouTube channel. The statistics demonstrate a robust positive association between these two measures, demonstrating that videos with incredible views garner more likes, implying that the material appeals to a wide range of viewers. Moreover, the diagram facilitates the identification of patterns in audience involvement, emphasizing anomalies where specific videos have a far higher number of views than likes. The presence of these outliers indicates that although the videos enjoy popularity, they may need to effectively engage or fascinate viewers, providing content providers with valuable input on the impact of their work.

Nevertheless, the analysis is subject to many constraints. The analysis primarily depends on quantitative measures, such as views and likes, while disregarding other types of audience participation, such as comments, shares, or watch time, which might offer a more comprehensive insight into user responses. Furthermore, the figure fails to consider contextual elements, such as the time of uploads or external events that could have impacted the data. Moreover, the absence of incorporation with sentiment analysis results in the study fails to capture the fundamental emotions or reasons behind social media interactions, increasing the risk of misinterpretations. Enhancing the research by

 ISSN 2723-6471

incorporating supplementary metrics and contextual data would be advantageous in obtaining a more thorough comprehension of audience involvement.

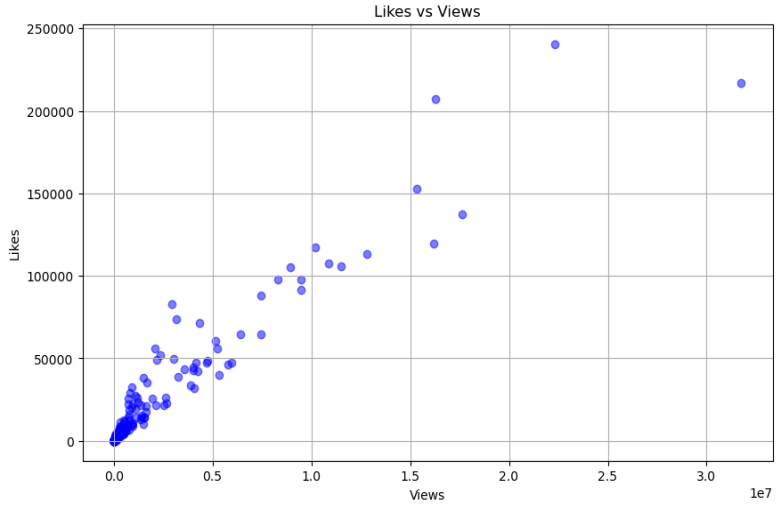


Figure 2. Comparison of Likes and Views from crawling results

Table 1 displays the sentiment analysis model's performance metrics, demonstrating its accuracy in categorizing viewer comments on Kompas TV's stand-up comedy videos as positive, negative, or neutral. The table presents the accuracy, precision, recall, and F1-score metrics for each sentiment, along with the count of comments classified in each category. Although the results are generally robust, they do not meet the criteria for optimal settings in sentiment analysis.

In an optimal situation, we anticipate the model to get a near-perfect accuracy of approximately 100% across all emotion categories. This implies that it will accurately classify almost all comments with a low margin of error. Furthermore, it is crucial to maintain high and well-balanced levels of precision and memory for positive, negative, and neutral thoughts. However, the model should accurately detect positive comments (high precision) and accurately capture all relevant instances of each sentiment (high recall). The F1-score, a metric that combines precision and recall, should ideally approach 1 for each category, indicating the overall efficacy of the model. Moreover, an optimal model would exhibit consistent performance across all sentiment categories, demonstrating its ability to effectively manage positive, negative, and neutral sentiments without prejudice.

Upon comparing the ideal conditions with the actual findings presented in table 1, it is evident that the model attains the highest level of accuracy (85%) in identifying positive sentiments. However, it exhibits lower accuracy rates for negative (80%) and neutral (78%) sentiments. Although these results are satisfactory, they suggest an enhancement, especially in managing neutral feelings. The precision and recall exhibit their maximum values for positive sentiments (0.87 and 0.84, respectively) but fall for negative (0.82 and 0.78) and neutral (0.76 and 0.79) attitudes. The lower precision and recall numbers, particularly in the neutral category, indicate that the model faces difficulty differentiating neutral comments from positive or negative ones, a typical obstacle in sentiment analysis.

The F1-scores exhibit a consistent pattern, with the most elevated score for positive thoughts (0.85), succeeded by negative (0.80) and neutral (0.77). Although these scores suggest a somewhat even performance, they do not reach the desired 1.0, especially in the neutral area. This discrepancy indicates that the model excels in recognizing distinct emotional signals, as observed in positive or negative remarks, but needs help to handle more ambiguous or blended sentiments effectively.

Overall, the model has satisfactory performance, particularly in handling positive sentiments. However, it needs to meet the optimal criteria for sentiment analysis. The diminished accuracy, precision, recall, and F1 scores in the negative and neutral categories underscore the need for enhancement. Improving the model's capacity to differentiate between neutral and other sentiments precisely could enhance its performance, bringing it closer to the ideal scenario. This would result in more precise and balanced sentiment classification across all categories. This investigation

 ISSN 2723-6471

highlights the necessity for additional refinement and more advanced methodologies to attain superior and more uniform results in sentiment analysis.

Table 1. Metrics overview presents performance

Sentiment Category	Accuracy (%)	Precision	Recall	F1-Score	Number of Comments
Positive	85%	0.87	0.84	0.85	10,000
Negative	80%	0.82	0.78	0.80	4,500
Neutral	78%	0.76	0.79	0.77	3,000

To further strengthen the conclusions of this study, several crucial areas of the current analysis and model performance must be addressed. First, expanding the diversity and size of the dataset could result in more robust and generalizable findings. The model could learn from a broader spectrum of comedic techniques and audience responses by incorporating a more comprehensive range of stand-up comedy videos from various channels, genres, and languages. This extension would help the model process complex and ambiguous emotions, especially in the neutral sentiment category, which it currently struggles to interpret.

Additionally, refining the data pre-processing techniques could significantly improve the precision and effectiveness of sentiment analysis. Integrating advanced methods like contextual word embeddings (e.g., BERT or ELMo) could enhance text preparation, allowing the model to better grasp the nuances in audience comments. These techniques effectively detect sarcasm, irony, and subtle emotional cues familiar in humor. Furthermore, using methods like oversampling or synthetic data generation to balance the distribution of positive, negative, and neutral comments would enhance the model's learning across all sentiment categories.

Improving the model's architecture could also lead to more refined outcomes. While the LSTM network is capable of processing sequential data, exploring advanced models such as Transformer-based architectures or hybrid models that combine LSTM with CNNs could yield better performance. These models can more effectively capture long-range dependencies and contextual relationships in text, leading to improved sentiment predictions. In addition, fine-tuning hyperparameters through grid search or Bayesian optimization could further optimize the model's performance. By focusing on these critical areas, dataset expansion, data processing, model architecture, multimodal data integration, and improving interpretability, the study can achieve more accurate, insightful, and actionable conclusions, ultimately benefiting content creators in the digital entertainment industry.

4.2. Audience Engagement Insights

Table 2 presents the most important words and their related scores in the analysis of YouTube stand-up comedy content comments, demonstrating their significance in the dataset. Lexical terms such as 'kokokoko,' 'comment,' 'innat,' and'serv' exhibit elevated scores, indicating their frequent usage by viewers in reaction to the comedic material. These concepts include fundamental features of audience involvement, whereby spectators respond to the comedy and actively engage by leaving comments, frequently alluding to cultural or situational factors introduced by the humorous performers. Including terms such as 'Abdul,' 'war,' and 'javanes' emphasizes the cultural backdrop inherent in the material, which deeply connects with the audience. This underscores the role of stand-up comedy on platforms such as YouTube in providing a forum for sharing cultural experiences and fostering audience engagement.

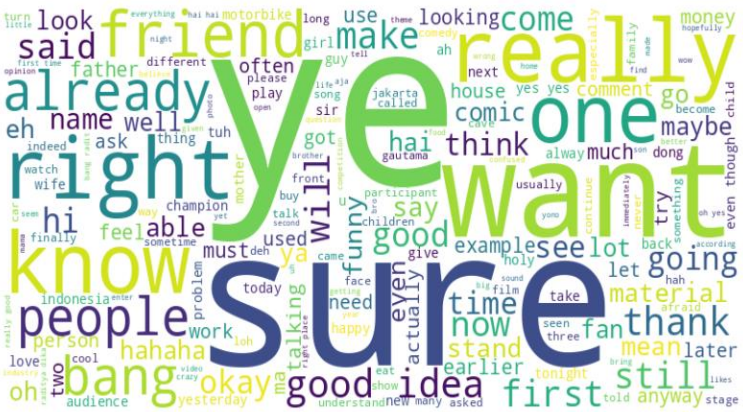
The individual scores assigned to each word, such as 0.335 for 'kokokoko' and 0.317 for 'comment,' provide significant insights into the relative importance of distinct terms in the sentiment analysis. More elevated scores indicate words that hold greater importance in the conversations around the comedy performances, implying that they are the components that elicit the strongest emotional reaction or are often mentioned. The research reveals that terms associated with participation and cultural identity (e.g., 'particip,' 'javanes') are frequently used, indicating that audiences actively and profoundly connect with the humorous material and its wider social and cultural significance. This study thoroughly examines how language patterns in YouTube comments provide insight into audience involvement. It demonstrates that viewers establish connections with the material more intricately than basic responses, including cultural and situational importance in their interactions.

ISSN 2723-6471

Table 2. Top Word Output

Document ID	Top Words	Top Scores
0	['kokokoko', 'comment', 'innat', 'serv', 'particip', 'abdul', 'war', 'javanes', 'stand', 'teach']	[0.33583907241377975, 0.31775443113820606, 0.299918863169128, 0.2639986539244763, 0.24063563389173892, 0.21157556480476383, 0.20291850344093168, 0.1972595878069077, 0.1908152406481136, 0.18190652801475304]
1	['audit', 'postpon', 'decemb', 'ripe', 'rubbish', 'lean', 'depress', 'pandem', 'march', 'fallen']	[0.294939383568777, 0.2930003791625255, 0.2746686325835012, 0.26166205139549004, 0.26166205139549004, 0.25157336324049256, 0.23636089822371942, 0.23636089822371942, 0.22499855823744136, 0.20502257045668387]
…	…	…
239	['teng', 'natur', 'girl', 'beauti', 'alma', 'transit', 'skincar', 'ugli', 'face', 'tongkrongan']	[0.8757974257224825, 0.1386870698462877, 0.1329126055607728, 0.1210812541016911, 0.11918930867244548, 0.1076644163960297, 0.10491443158735415, 0.09126398630228177, 0.07889635668037831, 0.07389269391684222]
240	['lectur', 'revil', 'idea', 'divorc', 'chancellor', 'sure', 'ridicul', 'umr', 'good', 'vice']	[0.2876413579199744, 0.27696360220779953, 0.24652468066365443, 0.20410764185138697, 0.18464240147186636, 0.16791029788399017, 0.16489367584373932, 0.16489367584373932, 0.14580562324364127, 0.14178914798476897]
…	…	…
477	['wow', 'man', 'spray', 'event', 'peopl', 'hey', 'respect', 'ye', 'directli', 'hotman']	[0.255396051692352, 0.24198428599649144, 0.20540474840800088, 0.16400996169732873, 0.15921291270366322, 0.1579418746115918, 0.14448133053632395, 0.14386008693027078, 0.13200623719212617, 0.12509741535055055]
478	['wo', 'candid', 'question', 'ye', 'papa', 'okay', 'shuffl', 'im', 'pair', 'ask']	[0.40063909649807145, 0.2133073606194378, 0.17437039882977878, 0.16835825567788876, 0.13757034460787126, 0.13298308869910722, 0.131207599785038, 0.12598270723745428, 0.1201248914494688, 0.11394076052448257]

Figure 3 displays a word cloud that graphically illustrates the primary words present in the YouTube transcription data of stand-up comedy material. The phrase cloud graphically represents the terms the viewers most frequently use in their comments, with each word's magnitude indicating its occurrence frequency. Increased word size corresponds to more significant usage, while smaller words suggest less frequent references. This graphic tool facilitates a rapid and intuitive comprehension of the prevailing topics in audience arguments, emphasizing important terms that powerfully connect with viewers. The word cloud visualizes the critical points of engagement and linguistic patterns that characterize the interaction between the audience and the comedic material by clustering the most significant words.



 ISSN 2723-6471

Figure 3. Word cloud visually represents frequent words in YouTube transcription data

The graph shown in figure 4 depicts the evolution of the model's learning process throughout the training phase. It is anticipated that the loss values for both the training and validation datasets would decline throughout the training process, suggesting that the model effectively reduces mistakes and enhances its predictions. A sharp decrease in training loss is evident throughout the early epochs, indicating fast learning as the model adjusts to the new input. In contrast, the validation loss exhibits a steadier decline, demonstrating the model's capacity to generalize to previously encountered data. Optimally, both losses should approach minimal values. Nevertheless, if the training loss consistently declines while the validation loss reaches a point of stability or rises, it could suggest overfitting, a phenomenon in which the model efficiently memorizes the training data without successfully generalizing it to new data. Monitoring these patterns across many epochs makes it possible to make necessary adjustments to the model to enhance its performance and prevent overfitting.

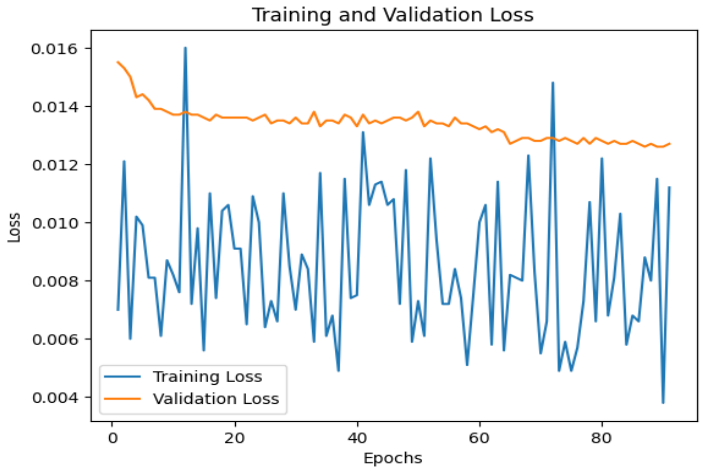


Figure 4. Evolution of Training and Validation Loss over 100 Epochs

Data analysis of the Stand-Up Comedy dataset from Kompas TV entails a multifaceted exploration of various dimensions of comedic content and audience engagement. Through statistical techniques and visualizations, researchers can uncover trends, patterns, and insights illuminating stand-up comedy. One analysis aspect involves examining the distribution of views, likes, and comments across different comedy performances and comedians, providing insights into audience preferences and reception. Additionally, sentiment analysis of comments can reveal the overall sentiment towards specific videos or comedians, shedding light on audience reactions and perceptions.

Furthermore, temporal analysis allows researchers to explore how stand-up comedy trends have evolved, identifying spikes in viewership or shifts in audience engagement patterns. Researchers can discern the impact of external events or cultural phenomena on audience behavior and content consumption habits by comparing metrics such as views and likes across different periods. Moreover, clustering analysis can help categorize comedy performances based on their content or style, enabling researchers to identify common themes or tropes prevalent in the dataset.

Another avenue of data analysis involves exploring correlations between engagement metrics, such as the relationship between views and likes or comments and shares. By conducting correlation analysis, researchers can identify factors contributing to audience engagement and popularity, informing content creators and platform strategies. Additionally, demographic analysis based on audience engagement metrics can provide insights into the demographics of stand-up comedy audiences, helping comedians and content creators tailor their content to specific audience segments. Overall, data analysis of the Stand-Up Comedy dataset offers valuable insights into the intricacies of comedic content and audience dynamics in digital landscape. Figure 5 illustrates the evolution of training and validation accuracy.

Data analysis is crucial in extracting meaningful insights from raw information, enabling informed decision-making across various domains. By employing statistical techniques, visualization tools, and machine learning algorithms, analysts can uncover patterns, trends, and correlations within datasets. This process involves cleaning and preprocessing data, followed by exploratory analysis to identify critical features and relationships. Subsequently,

 ISSN 2723-6471

advanced modeling techniques are applied to derive predictive insights or classification outcomes. Practical data analysis enhances understanding and drives organizational innovation, optimization, and strategic planning. Figure 6 displays the training and validation loss trends over epochs.

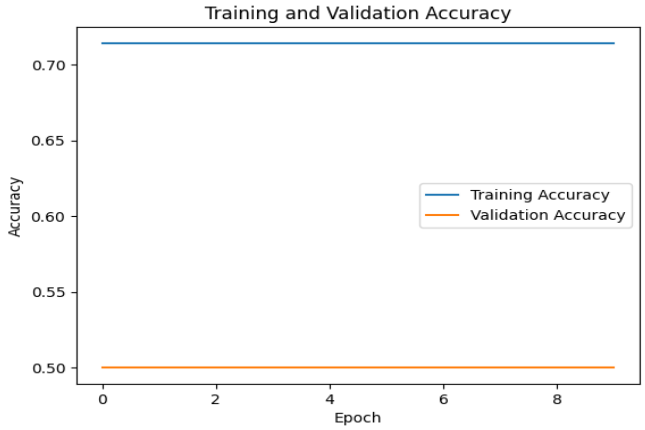


Figure 5. The evolution of training and validation accuracy

Figure 6. The training and validation loss throughout the epochs

The training and validation loss throughout the epochs provide insights into the performance of a machine learning model during the training process. This loss, often represented as a numerical value, indicates how well the model is performing in terms of minimizing errors on both the training and validation datasets. As the number of epochs increases, the training loss tends to decrease as the model learns from the data. Conversely, the validation loss helps monitor the model's generalization performance on unseen data. A consistent decrease in training and validation loss signifies that the model effectively learns and generalizes from the training data. However, if there is a significant gap between the training and validation loss or if the validation loss increases while the loss decreases, it may indicate overfitting, where the model memorizes the training data rather than learning meaningful patterns. Thus, monitoring the training and validation loss over epochs is crucial for assessing the model's performance and making necessary adjustments to optimize its performance.

5. Conclusion

A potential application of the LSTM model in sentiment analysis of YouTube comments about stand-up comedy is highlighted in this paper. Accuracy values of 85% for positive, 80% for negative, and 78% for neutral feelings demonstrate the effectiveness of the LSTM model in capturing audience reactions to various aspects of humor. However, the model has difficulties, notably with neutral and negative sentiments, because of these emotions' nuanced and intricate nature. The findings indicate that the model excels in recognizing robust emotional signals but encounters difficulties in discerning vague or less well-defined emotions, particularly in neutral remarks.

The study also highlights constraints in the LSTM model's ability to handle negative attitudes, particularly instances of sarcasm, irony, and black humor, which provide further difficulties. Negative remarks frequently contain intricate levels of meaning that can pose challenges for LSTM models in interpretation. For example, sarcasm may show a negative connotation but convey a positive or hilarious meaning, resulting in misinterpretation. This underscores the need for more advanced models to comprehend these subtleties in language and context more effectively.

To gain a more profound understanding of cultural and language differences in comedy, future studies should prioritize expanding the dataset to include a broader array of platforms, such as TikTok, and incorporating comments from foreign comedians. Furthermore, investigating ensemble techniques that integrate LSTM with Transformer-based models such as BERT can enhance the precision in detecting nuanced emotional signals, particularly in neutral and negative attitudes. Sophisticated embedding methods such as ELMo and GloVe can significantly improve the model's capacity to comprehend the intricacies of language.

Furthermore, forthcoming research should also tackle the obstacles of cross-language sentiment analysis by creating multilingual models that can adequately analyze sentiments in many languages and cultures. Integrating cross-modal data, including video and audio analysis, might offer a more thorough comprehension of audience responses, as factors such as vocal intonation and facial expressions are crucial in communicating comedy. Furthermore, these developments will enhance sentiment analysis for comic material and contribute to wider digital media and audience interaction studies.

6. Declarations

6.1. Author Contributions

Conceptualization: S., A.P.W., S., and F.K.; Methodology: A.P.W. and F.K.; Software: S.; Validation: S., A.P.W., S., and F.K.; Formal Analysis: S., A.P.W., and F.K.; Investigation: S.; Resources: A.P.W. and F.K.; Data Curation: A.P.W.; Writing Original Draft Preparation: S., A.P.W., and F.K.; Writing Review and Editing: A.P.W., S., and F.K.; Visualization: S. and F.K.; All authors have read and agreed to the published version of the manuscript.

6.2. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

6.3. Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

6.4. Institutional Review Board Statement

Not applicable.

6.5. Informed Consent Statement

Not applicable.

6.6. Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] M. Andersson, "Multimodal expression of impoliteness in YouTube reaction videos to transgender activism," *Discourse, Context Media*, vol. 58, no. April, pp. 100760, 2024, doi: 10.1016/j.dcm.2024.100760.

[2] M. L. Siciliano, "Intermediaries in the age of platformized gatekeeping: The case of YouTube 'creators' and MCNs in the U.S.," *Poetics*, vol. 97, no.April, pp. 101748, 2023.

[3] Z. Halim, S. Hussain, and R. Hashim Ali, "Identifying content unaware features influencing popularity of videos on YouTube: A study based on seven regions," *Expert Syst.* Appl., vol. 206, no. November, pp. 117836, 2022.

[4] T. Searle, Z. Ibrahim, J. Teo, and R. J. B. Dobson, "Discharge summary hospital course summarisation of in patient Electronic Health Record text with clinical concept guided deep pre-trained Transformer models," *J. Biomed. Inform.*, vol. 141, no. May, pp. 104358, 2023.

[5] T. F. Waddell and S. S. Sundar, "Bandwagon effects in social television: How audience metrics related to size and opinion affect the enjoyment of digital media," *Comput. Human Behav.*, vol. 107, no. June, pp. 106270, 2020.

[6] Z. Alhathloul and I. Ahmad, "Automatic dottization of Arabic text (Rasms) using deep recurrent neural networks," *Pattern Recognit. Lett.*, vol. 162, no. October, pp. 47–55, 2022.

[7] G. Lv, Y. Sun, F. Nian, M. Zhu, W. Tang, and Z. Hu, "COME: Clip-OCR and Master ObjEct for text image captioning," *Image Vis. Comput.*, vol. 136, no. August, pp. 104751, 2023.

[8] T. Widiyaningtyas, A. P. Wibawa, W. Caesarendra, and U. Pujianto, "MF-NCG: Recommendation Algorithm Using Matrix Factorization-based Normalized Cumulative Genre," *Int. J. Intell. Eng. Syst.*, vol. 17, no. 2, pp. 180–189, 2024, doi: 10.22266/ijies2024.0430.16.

[9] T. Widiyaningtyas, D. D. Prasetya, and H. W. Herwanto, "Time Loss Function-based Collaborative Filtering in Movie Recommender System," *Int. J. Intell. Eng. Syst.*, vol. 16, no. 6, pp. 1021–1030, Dec. 2023, doi: 10.22266/ijies2023.1231.84.

[10] A. P. Wibawa et al., "Deep Learning Approaches with Optimum Alpha for Energy Usage Forecasting," *Knowl. Eng. Data Sci.*, vol. 6, no. 2, pp. 170–187, 2023, doi: 10.17977/um018v6i22023p170-187.

[11] A. P. Wibawa et al., "Bidirectional Long Short-Term Memory (Bi-LSTM) Hourly Energy Forecasting," *E3S Web Conf.*, vol. 501, no. March, pp.7, 2024, doi: 10.1051/e3sconf/202450101023.

[12] A. P. Wibawa et al., "Decoding and preserving Indonesia's iconic Keris via A CNN-based classification," *Telemat. Informatics Reports*, vol. 13, no. March, pp. 100120, 2024, doi: 10.1016/j.teler.2024.100120.

[13] Z. Wu, Q. He, J. Li, G. Bi, and M. F. Antwi-Afari, "Public attitudes and sentiments towards new energy vehicles in China: A text mining approach," *Renew. Sustain. Energy Rev.*, vol. 178, no. May, pp. 113242, 2023.

[14] M. N. Ashtiani and B. Raahemi, "News-based intelligent prediction of financial markets using text mining and machine learning: A systematic literature review," *Expert Syst. Appl.*, vol. 217, no. May, pp. 119509, 2023.

[15] A. Wurm and J. Wimmer, "The role of self-perception, media practices and synchronicity in the production of commented gameplay videos," *Entertain. Comput.*, vol. 49, no. March, pp. 100622, 2024.

[16] Q. Sun and C. Shen, "Who would respond to A troll? A social network analysis of reactions to trolls in online communities," *Comput. Human Behav.*, vol. 121, no. August, pp. 106786, 2021.

[17] Z. Sheng et al., "Reconfigurable Logic-in-Memory Computing Based on a Polarity-Controllable Two-Dimensional Transistor," *Nano Lett.*, vol. 23, no. 11, pp. 5242–5249, Jun. 2023, doi: 10.1021/acs.nanolett.3c01248.

[18] S. Hao et al., "Enhanced Semantic Representation Learning for Sarcasm Detection by Integrating Context-Aware Attention and Fusion Network," *Entropy*, vol. 25, no. 6, pp. 878, May 2023, doi: 10.3390/e25060878.

[19] N. N. Soylu and E. Sefer, "BERT2OME: Prediction of 2′-O-Methylation Modifications from RNA Sequence by Transformer Architecture Based on BERT," *IEEE/ACM Trans. Comput. Biol. Bioinforma.*, vol. 20, no. 3, pp. 2177–2189, May 2023, doi: 10.1109/TCBB.2023.3237769.

[20] F. A. Ahda, A. P. Wibawa, D. Dwi Prasetya, and D. Arbian Sulistyo, "Comparison of Adam Optimization and RMS prop in Minangkabau-Indonesian Bidirectional Translation with Neural Machine Translation," *JOIV Int. J. Informatics Vis.*, vol. 8, no. 1, pp. 231, Mar. 2024, doi: 10.62527/joiv.8.1.1818.

[21] D. A. Sulistyo, "LSTM-Based Machine Translation for Madurese-Indonesian," *J. Appl. Data Sci.*, vol. 4, no. 3, pp. 189–199, Sep. 2023, doi: 10.47738/jads.v4i3.113.

[22] J. Ji, Z. Hu, W. Zhang, and S. Yang, "Development of Deep Learning Algorithms, Frameworks and Hardwares," In: Qian, Z., Jabbar, M., Li, X. (eds) Proceeding of 2021 International Conference on Wireless Communications, Networking and Applications. WCNA 2021. Lecture Notes in Electrical Engineering. Springer, Singapore, vol. 2022, no. July, pp. 696–710, 2022, doi: 10.1007/978-981-19-2456-9_71.

[23] A. Gundawar and S. Lodha, "Enhanced Dense Layers Using a Quadratic Transformation Function," In: Haldorai, A., Ramu, A., Mohanram, S. (eds) 5th EAI International Conference on Big Data Innovation for Sustainable Cognitive Computing. BDCC 2022. EAI/Springer Innovations in Communication and Computing. Springer, Cham, vol. 2023, no. June, pp. 3–15, doi: 10.1007/978-3-031-28324-6_1.

[24] D. Telezhenko and O. Tolstoluzka, "Development and Training of Lstm Models for Control Of Virtual Distributed Systems Using Tensorflow and Keras," *Grail Sci.*, no. 38, pp. 163–168, Apr. 2024, doi: 10.36074/grail-of-science.12.04.2024.027.

[25] M. Karbasi et al., "Multi-step ahead forecasting of electrical conductivity in rivers by using a hybrid Convolutional Neural Network-Long Short-Term Memory (CNN-LSTM) model enhanced by Boruta-XGBoost feature selection algorithm," *Sci. Rep.*, vol. 14, no. 1, pp. 15051, Jul. 2024, doi: 10.1038/s41598-024-65837-0.

[26] A. Barakat and P. Bianchi, "Convergence and Dynamical Behavior of the ADAM Algorithm for Nonconvex Stochastic Optimization," *SIAM J. Optim.*, vol. 31, no. 1, pp. 244–274, Jan. 2021, doi: 10.1137/19M1263443.

[27] B. R. Chakravarthi, "Hope speech detection in YouTube comments," *Soc. Netw. Anal. Min.*, vol. 12, no. 1, pp. 75, Dec. 2022, doi: 10.1007/s13278-022-00901-z.