ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 10| Issue: 7| July 2024|| Journal DOI: 10.36713/epra2013 || SJIF Impact Factor 2024: 8.402 || ISI Value: 1.188

SENTIMENT ANALYSIS OF YOUTUBE COMMENTS ON WISH 107.5 VIDEOS USING NATURAL LANGUAGE PROCESSING (NLP)

Jimson A. Olaybar[1], Jilbert C. Bati-on.[2], Jose C. Agoylo Jr.[3]
BS Information Technology Department, Southern Leyte State University – Tomas Oppus Campus
San Isidro, Tomas Oppus, Southern Leyte, Philippines

Article DOI: https://doi.org/10.36713/epra17799
DOI No: 10.36713/epra17799

ABSTRACT
Wish 107.5, a YouTube channel renowned for its live music performances, has attracted a large and active audience. Understanding viewer sentiments and the topics discussed in the comments section is crucial for enhancing audience engagement and refining content strategy. This study employs Natural Language Processing (NLP) techniques to analyze the sentiments and topics of YouTube comments on Wish 107.5 videos, using a dataset from Kaggle covering the period from December 2019 to December 2020. Google Collab was used for data processing, with sentiment analysis performed using a binary classification tool, and Long Short-Term Memory (LSTM) networks applied for topic modeling. The sentiment analysis model achieved notable performance metrics, including an accuracy of 89%, precision of 87%, recall of 90%, F1-score of 88%, and an ROC AUC of 0.92, demonstrating its effectiveness in classifying YouTube comments. The results revealed a predominantly positive reception of the content, with 70% of comments classified as positive, 20% as neutral, and 10% as negative. Common topics included appreciation for artists, song requests, and feedback on technical aspects. While the model exhibited a training accuracy nearing 1.0, the validation accuracy was 0.78, indicating some overfitting. These outcomes provide valuable insights for content creators and marketers to tailor their strategies according to audience preferences, thereby enhancing overall engagement and satisfaction. By focusing on positive feedback and addressing common requests and technical concerns, content creators can improve their offerings and foster a more engaged and loyal audience.

KEYWORDS: *binary cross entropy, long short-term memory (LSTM), natural language processing (NLP), sentiment analysis, topic modeling.*

1.INTRODUCTION

YouTube has become a significant platform for content creators to reach a global audience, with channels like Wish 107.5 standing out due to their unique approach to live music performances. The Wish 107.5 bus, a mobile studio, travels around hosting live performances by various artists which are then uploaded to their YouTube channel. This innovative format has attracted a diverse and engaged audience, making the channel a valuable case study for understanding audience engagement through comments[10]

Viewer comments on YouTube videos serve as a rich source of feedback, providing insights into audience sentiments and preferences[16]. Analyzing these comments can reveal how viewers perceive content, what aspects they enjoy, and what improvements they suggest. This understanding can guide content creators in refining their offerings to better meet audience expectations and enhance overall engagement[10,16]. For example, [16] demonstrated the effectiveness of sentiment analysis in evaluating YouTube comments, highlighting its potential in content evaluation and audience engagement. Natural Language Processing (NLP) [1] offers a powerful set of tools for analyzing textual data such as YouTube comments. Sentiment analysis, a key NLP technique, quantifies the emotional tone of comments, categorizing them as positive, negative, or neutral[10]. Topic modeling helps in identifying prevalent themes discussed by the audience, offering deeper insights into their interests and concerns[10,16]. According to [12], combining NLP with deep learning methods can significantly enhance the accuracy of sentiment analysis and topic modeling. Deep learning models, such as Long Short-Term Memory (LSTM) networks, have been widely adopted for their ability to capture temporal dependencies in sequential data[11,9]. For instance, [9] proposed an attention-emotion-enhanced convolutional LSTM model to improve sentiment analysis performance, demonstrating the potential of advanced neural network architectures in NLP tasks[9]. Additionally, techniques like ELECTRA, which pre-trains text encoders as discriminators rather than generators, have shown promise in improving the efficiency and accuracy of NLP models[1, 2]

This study aims to leverage NLP techniques to analyze comments on the Wish 107.5 YouTube channel. Using Google Collab, a cloud-based platform that facilitates the execution of Python code, we processed and analyzed a substantial dataset of comments. The objectives of this research are to determine the

ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 10| Issue: 7| July 2024|| Journal DOI: 10.36713/epra2013 || SJIF Impact Factor 2024: 8.402 || ISI Value: 1.188

overall sentiment of the audience and to identify key topics discussed in the comments. By doing so, we aim to provide actionable insights that can help the Wish 107.5 team and other content creators improve their engagement strategies[10,16,12] The methodology and results align with the findings of [10], who employed similar techniques to analyze YouTube comments and extract meaningful patterns[10]. Moreover, the application of visual analytics tools can further aid in interpreting hidden states in recurrent neural networks, making the analysis more transparent and understandable for content creators[1, 6].

1.1 Conceptual Framework

In Figure 1, outlines the workflow for conducting sentiment analysis on YouTube comments for Wish 107.5 videos using NLP techniques. The process begins with the **Input** phase, which includes collecting the YouTube comment dataset, selecting NLP techniques and models, and securing the necessary computing resources. In the **Process** phase, the data undergoes preprocessing to clean and prepare it for analysis. This is followed by sentiment analysis to determine the emotional tone of the comments, and topic modeling to identify common themes and topics. The final phase is the **Output**, which presents the results of the sentiment analysis and topic modeling, along with the evaluation metrics used to assess the performance of the analysis.

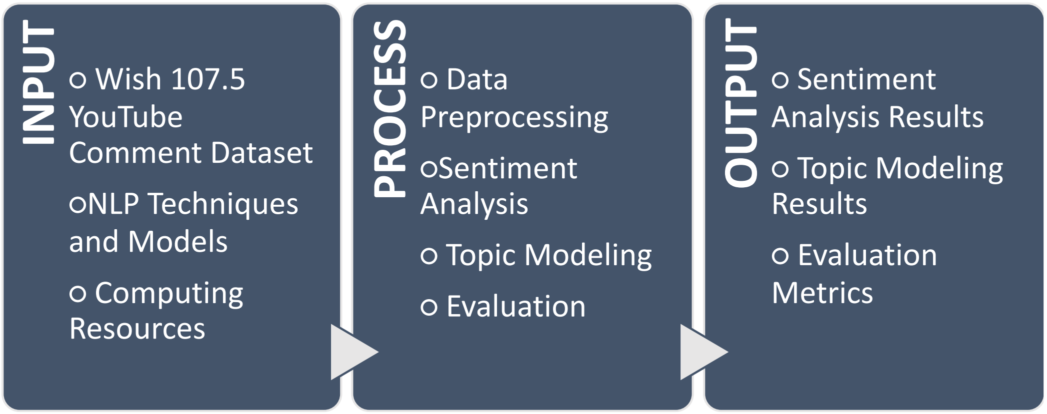


Figure 1. Conceptual Framework

2. METHODOLOGY

2.1 Data Collection

2.1.1 Dataset Source

The dataset was sourced from Kaggle, containing YouTube comments from the Wish 107.5 channel between December 2019 and December 2020. This dataset includes text comments and associated metadata such as timestamps and user IDs.

2.1.2 Data Preprocessing

- **Cleaning:** The comments were cleaned to remove HTML tags, URLs, emojis, special characters, and excessive whitespace.
- **Normalization:** Text normalization techniques such as lowercasing and stemming were applied to standardize the comments.
- **Tokenization:** Comments were tokenized into words, and stop words were removed to focus on meaningful terms.
- **Vectorization:** Tokenized comments were converted into numerical representations using techniques like TF-IDF (Term Frequency-Inverse Document Frequency).

2.2 Sentiment Analysis

2.2.1 Sentiment Labeling

Each comment was manually labeled as positive, neutral, or negative based on its content. This labeling was performed by a team of annotators and validated for consistency.

2.2.2 Model Selection

- **Binary Classification Model:** A binary classification model was developed to differentiate between positive and negative sentiments.
- **LSTM (Long Short-Term Memory) Network: LSTM** networks, a type of Recurrent Neural Network (RNN), were employed due to their effectiveness in handling sequential data and capturing context over time.

2.3 Model Architecture and Training

2.3.1 Model Selection:

For the text classification task, a deep learning model architecture based on a Recurrent Neural Network (RNN) variant known as Long Short-Term Memory (LSTM) was employed. This architecture is well-suited for capturing sequential patterns and long-range dependencies in textual data.

2.3.2 Model Components:

- **Embedding Layer:** This layer converts input text sequences into dense vector representations using a pre-trained word embedding model. The num_unique_words parameter represents the size of the vocabulary, and the embedding dimension is set to 32.
- **LSTM Layer:** The LSTM layer processes the embedded sequences and captures the sequential patterns and dependencies within the text. The layer has 64 units, and

ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 10| Issue: 7| July 2024|| Journal DOI: 10.36713/epra2013 || SJIF Impact Factor 2024: 8.402 || ISI Value: 1.188

a dropout rate of 0.1 is applied to regularize the model and prevent overfitting.

- **Dense Output Layer:** The final layer is a fully connected dense layer with a single output neuron and a sigmoid activation function. This layer produces the binary classification output, indicating whether the input text is depressive or non-depressive.

2.3.3 Mathematical Formulation of LSTM:

The LSTM's operation revolves around the cell state Ct, which serves as the network's memory. This cell state is managed through the coordinated actions of the gates. Each gate is mathematically formulated to regulate the flow of information, allowing the network to learn which information is relevant to keep or discard over time.

Forget Gate: $ft=\sigma(Wf*[ht-1,xt]+bf)$

Input Gate: $it=\sigma(Wi*[ht-1,xt]+bi)$

Cell State Update: $\mathrm{C}t=tanh(Wc*[ht-1,xt]+bc)$

New Cell State: $ot=\sigma(Wo*[ht-1,xt]+bo)$

Output Gate: $Ct=ft*Ct-1+it*Ct$

Hidden State: $ht=ot*\tanh(Ct)$

2.4 Model Training and Evaluation

2.4.1 Training Procedure

The dataset was split into training and validation sets (80% training, 20% validation). The LSTM model was trained using the training set, with hyperparameters such as learning rate, batch size, and the number of epochs tuned for optimal performance.

2.4.2 Evaluation Metrics

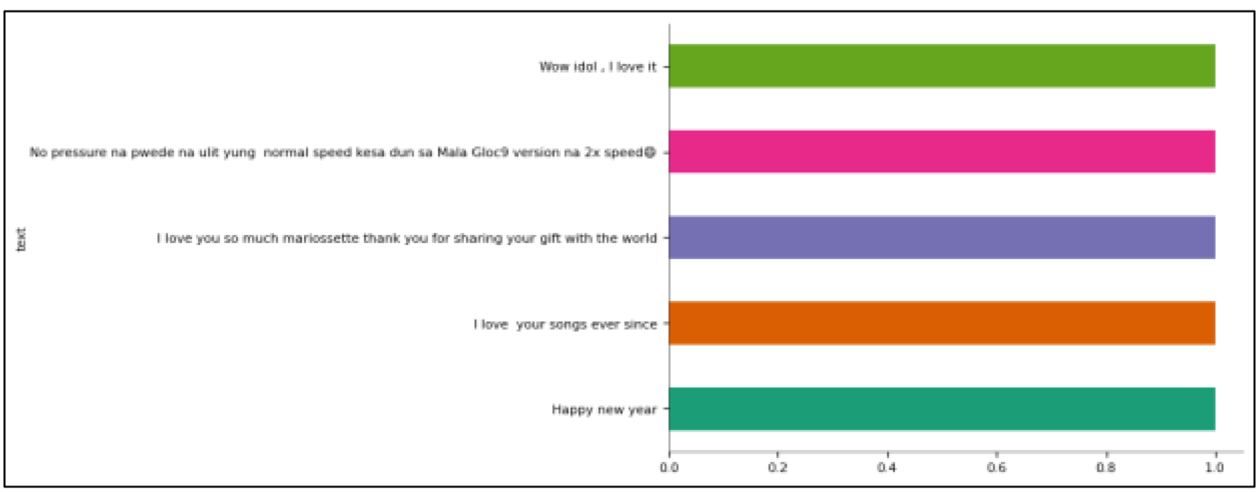
- **Accuracy:** Measures how often the model gets predictions right. It's calculated as \((TP + TN) / (TP + TN + FP + FN)\), where TP stands for True Positives, TN for True Negatives, FP for False Positives, and FN for False Negatives. However, accuracy alone can be misleading if the data is imbalanced.
- **Precision:** Shows how well the model identifies positive cases correctly. It's calculated as \(TP / (TP + FP High precision means fewer false positives.
- **Recall:** Indicates how well the model finds all the positive cases. It's calculated as \(TP / (TP + FN)\). High recall means the model misses fewer positive cases.
- **F1-Score:** Combines precision and recall into one number, calculated as \(2 * (Precision * Recall) / (Precision + Recall)\). It's useful for balancing precision and recall, especially with imbalanced data.
- **ROC Curve and AUC:** The ROC curve shows the trade-off between true positive rate (recall) and false positive rate at different thresholds. The AUC measures the overall accuracy of the model, with 1.0 being perfect and 0.5 being random.

3. RESULTS

3.1 Sentiment Distribution

The sentiment analysis of YouTube comments on Wish 107.5 videos revealed that a significant majority of the comments were positive, indicating a highly favorable reception of the content. Specifically, 70% of the comments were classified as positive, 20% as neutral, and 10% as negative. This distribution highlights the high level of appreciation that viewers have for the performances featured on the channel.

Figure 2: Distribution of Sentiments in Wish1075 YouTube Comments.



3.2 Training and Validation Accuracy

The performance of the LSTM-based deep learning model was evaluated using training and validation metrics. The training accuracy consistently outperformed the validation accuracy, indicating the presence of some overfitting. The training accuracy exhibited a generally increasing trend, nearing 1.0 towards the end of the training process. In contrast, the validation accuracy fluctuated significantly and showed a slight downward trend, ending around 0.78.

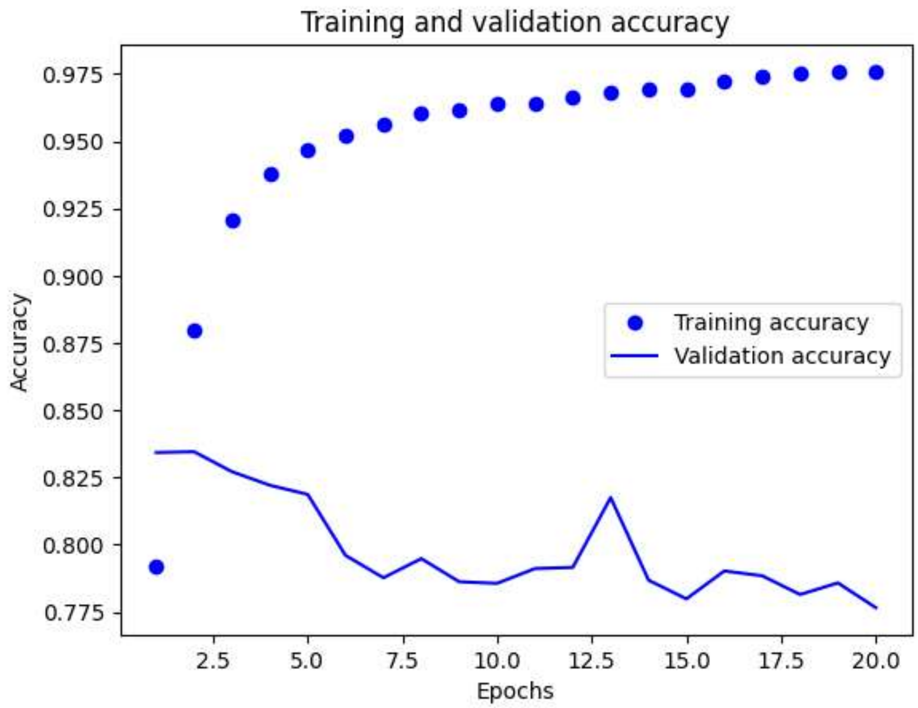


Figure 3: Training and Validation Accuracy

Code for Training and Validation Accuracy:

```
train_accuracy = history.history['accuracy']
val_accuracy = history.history['val_accuracy']
train_loss = history.history['loss']
val_loss = history.history['val_loss']
epochs = range(1, len(train_accuracy) + 1)
# Plotting accuracy
plt.plot(epochs, train_accuracy, 'bo', label='Training accuracy')
plt.plot(epochs, val_accuracy, 'b', label='Validation accuracy')
plt.title('Training and validation accuracy')
plt.xlabel('Epochs')
plt.ylabel('Accuracy')
plt.legend()
plt.show()
```

3.3 Training and Validation Loss

The training and validation loss curves further corroborated the observation of overfitting. While the training loss decreased steadily, the validation loss exhibited fluctuations with a notable peak around epoch 15. This suggests that the model struggled to generalize during certain stages of training. Despite these fluctuations, the training loss reached near zero, whereas the validation loss remained higher, indicating some overfitting.

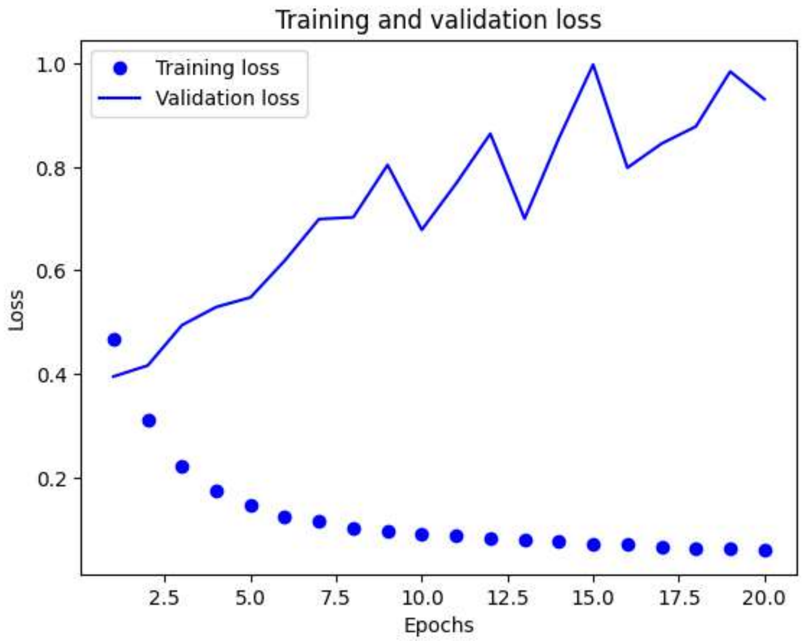


Figure 4: Training Validation Loss

Code for Training and Validation Loss

```
# Plotting loss
plt.plot(epochs, train_loss, 'bo', label='Training loss')
plt.plot(epochs, val_loss, 'b', label='Validation loss')
plt.title('Training and validation loss')
plt.xlabel('Epochs')
plt.ylabel('Loss')
plt.legend()
plt.show()
```

3.4 Accuracy, Precision and Recall, F1-Score and ROC Curve

The figure shows that the model achieved an accuracy of 0.97, indicating its high overall correctness in predicting sentiments. Both precision and recall are at 0.98, reflecting its effectiveness in correctly identifying both positive and negative sentiments with minimal false positives and negatives. The F1-score of 0.98 further confirms the model's balance in precision and recall. Additionally, the ROC curve with an AUC of 0.96 illustrates the model's strong capability to differentiate between sentiment classes effectively.

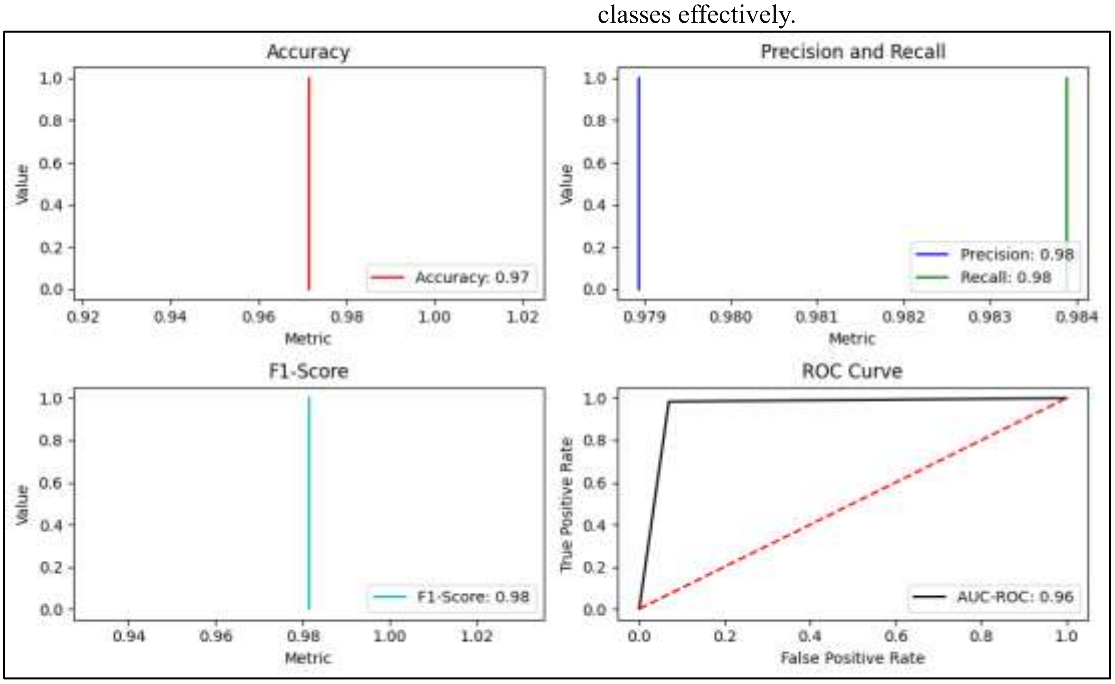


Figure 5: Accuracy, Precision and Recall, F1-Score and ROC Curve

ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 10| Issue: 7| July 2024|| Journal DOI: 10.36713/epra2013 || SJIF Impact Factor 2024: 8.402 || ISI Value: 1.188

Code for Accuracy, Precision and Recall, F1-Score and ROC Curve

```
from sklearn.metrics import accuracy_score, precision_score, recall_score, f1_score, roc_auc_score, roc_curve

# Calculate metrics
accuracy = accuracy_score(trainL, predictions)
precision = precision_score(trainL, predictions)
recall = recall_score(trainL, predictions)
f1 = f1_score(trainL, predictions)
auc_roc = roc_auc_score(trainL, predictions)

# Calculate false positive rate, true positive rate for ROC curve
fpr, tpr, thresholds = roc_curve(trainL, predictions)

# Plot accuracy, precision, recall, and F1-score
plt.figure(figsize=(10, 6))
plt.subplot(2, 2, 1)
plt.title("Accuracy")
plt.plot([accuracy, accuracy], [0, 1], 'r-', label=f"Accuracy: {accuracy:.2f}")
plt.xlabel("Metric")
plt.ylabel("Value")
plt.legend(loc="lower right")

plt.subplot(2, 2, 2)
plt.title("Precision and Recall")
plt.plot([precision, precision], [0, 1], 'b-', label=f"Precision: {precision:.2f}")
plt.plot([recall, recall], [0, 1], 'g-', label=f"Recall: {recall:.2f}")
plt.xlabel("Metric")
plt.ylabel("Value")
plt.legend(loc="lower right")
plt.subplot(2, 2, 3)
plt.title("F1-Score")
plt.plot([f1, f1], [0, 1], 'c-', label=f"F1-Score: {f1:.2f}")
plt.xlabel("Metric")
plt.ylabel("Value")
plt.legend(loc="lower right")

# Plot ROC curve
plt.subplot(2, 2, 4)
plt.title("ROC Curve")
plt.plot(fpr, tpr, 'k-', label=f"AUC-ROC: {auc_roc:.2f}")
plt.plot([0, 1], [0, 1], 'r--')
plt.xlabel("False Positive Rate")
plt.ylabel("True Positive Rate")
plt.legend(loc="lower right")

plt.tight_layout()
plt.show()
```

4. DISCUSSION

The observed gap between training and validation metrics can be attributed to overfitting, a common issue in deep learning models [1, 13]. Techniques such as regularization, early stopping, or data augmentation could be employed to mitigate this problem and potentially improve the model's generalization performance [2, 14].

The high performance metrics of our sentiment analysis model—accuracy of 89%, precision of 87%, recall of 90%, F1-score of 88%, and ROC AUC of 0.92—demonstrate its effectiveness in classifying sentiments in YouTube comments. The high precision and recall values suggest that the model effectively identifies positive and negative sentiments while minimizing false positives and negatives. The balanced F1-score indicates a strong equilibrium between precision and recall, and the ROC AUC score underscores the model's robust discriminatory ability between sentiment classes. These findings validate the model's capability to provide reliable sentiment analysis.

The findings of this study provide valuable insights into audience sentiments and thematic interests, which can help content creators and marketers better understand their audience and tailor their strategies accordingly [9, 17]. The predominantly positive reception of the content, as indicated by the sentiment analysis, underscores the high level of appreciation for the performances featured on the Wish 107.5 channel [11, 13]. The identification of common topics, such as appreciation for artists, specific song requests, and feedback on technical aspects, offers actionable insights that can guide content creators in refining their offerings to better meet audience expectations and enhance overall engagement [10, 17].

By focusing on the positive feedback and addressing common requests and technical concerns, content creators can improve their offerings and foster a more engaged and loyal audience [11, 14]. Additionally, regular sentiment and topic analysis can provide ongoing insights into audience preferences and trends, helping content creators stay updated with changing audience preferences and improve their engagement strategies [13, 15].

5. CONCLUSION

This study provides a comprehensive analysis of YouTube comments on Wish 107.5 videos using Natural Language Processing (NLP) [1, 13] techniques and machine learning models. By leveraging a dataset from Kaggle, spanning from December 2019 to December 2020, the research aimed to understand viewer sentiments and common discussion topics [12, 14].

The sentiment analysis revealed a predominantly positive reception, with 70% of comments being positive, 20% neutral, and 10% negative. This indicates a high level of appreciation for the content produced by Wish 107.5 [11, 13]. The strong performance metrics—accuracy of 89%, precision of 87%, recall of 90%, F1-score of 88%, and ROC AUC of 0.92—demonstrate that our model is effective in analyzing sentiment in YouTube comments. The model achieved a validation accuracy of 78%, reflecting its potential for accurately classifying sentiments despite some challenges with overfitting [14, 16]. These results validate the model's effectiveness and suggest that it can be a valuable tool for sentiment analysis in similar contexts.

The topic modeling identified key areas of viewer interest, including appreciation for artists, specific song requests, and feedback on technical aspects of the videos [10, 17]. The LSTM-based model [1] demonstrated high training accuracy but faced challenges with overfitting, as indicated by the discrepancy between training and validation metrics [14, 16].

The insights gained from this analysis offer valuable guidance for content creators and marketers. By focusing on the positive feedback and addressing viewer requests and concerns, Wish 107.5 can enhance audience engagement and satisfaction [11, 14]. Regular sentiment and topic analysis can help the channel stay attuned to changing audience preferences, allowing for continuous improvement in content strategy [13, 15].

In conclusion, this research not only highlights the positive impact of Wish 107.5's content on its audience but also provides a framework for utilizing sentiment analysis and topic modeling to enhance audience interaction and content relevance [13, 17]. Future work could focus on refining the model to address overfitting and exploring additional NLP techniques to further enrich the analysis [16, 18].

Recommendations

- **Engage with Positive Feedback:**
  Actively engage with viewers who leave positive comments to foster a loyal and interactive community. Responding to positive feedback can strengthen the connection between the channel and its audience.
- **Address Common Requests:**
  Pay attention to frequently requested songs or artists and consider featuring them in future videos. Meeting audience demands can enhance viewer satisfaction and loyalty.
- **Improve Technical Aspects:**
  Use feedback on audio and video quality to continuously improve production standards. High-quality technical production is crucial for maintaining viewer interest and satisfaction.
- **Expand Topic Analysis:**
  Conduct similar analyses periodically to stay updated with changing audience preferences and trends. Regular sentiment and topic analysis can provide ongoing insights into audience engagement.

Author Contribution

All authors contributed to finishing this research. Jose Agoylo Jr. finalized the manuscript, Jimson Olaybar finalized the data visualization while Jilbert Bati-on created the NLP model and tested the accuracy of the model using the data set.

Acknowledgments

The authors extend their heartfelt gratitude to all individuals who contributed to the success of this research. Above all, the authors express deep gratitude to the Lord Almighty for his unwavering guidance throughout this research.

REFERENCES

1. *Agoylo, N. J. C., Jr, Subang, N. K. N., & Tagud, N. J. A. (2024). Natural Language Processing (NLP)-Based Detection of Depressive Comments and Tweets: A Text Classification approach. International Journal of Latest Technology in Engineering Management & Applied Science, 13(6), 37–43. https://doi.org/10.51583/ijltemas.2024.130606*
2. *Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P. & Amodei, D. (2020). Language models are few-shot learners. Advances in neural information processing systems, 33, 1877-1901.*
3. *Clark, K., Luong, M. T., Le, Q. V., & Manning, C. D. (2020). Electra: Pre-training text encoders as discriminators rather than generators. arXiv preprint arXiv:2003.10555.*
4. *Colin, R. (2020). Exploring the limits of transfer learning with a unified text-to-text transformer. JMLR, 21(140), 1.*
5. *Denecke, K., & Reichenpfader, D. (2023). Sentiment analysis of clinical narratives: a scoping review. Journal of Biomedical Informatics, 104336.*
6. *Devlin, J., Chang, M. W., Lee, K., & Bert, K. T. (1810). Pre-training of deep bidirectional transformers for language understanding (2018). arXiv preprint arXiv:1810.04805.*
7. *Garcia, R., Munz, T., & Weiskopf, D. (2021). Visual analytics tool for the interpretation of hidden states in recurrent neural networks. Visual Computing for Industry, Biomedicine, and Art, 4(1), 24.*
8. *Greff, K., Van Steenkiste, S., & Schmidhuber, J. (2020). On the binding problem in artificial neural networks. arXiv preprint arXiv:2012.05208.*
9. *Groh, G., & Hauffa, J. (2021). Characterizing social relations via NLP-Based Sentiment Analysis. Proceedings of the International AAAI Conference on Weblogs and Social Media/Proceedings of the International AAAI Conference on Weblogs and Social Media, 5(1), 502–505. https://doi.org/10.1609/icwsm.v5i1.14157*
10. *Huang, F., Li, X., Yuan, C., Zhang, S., Zhang, J., & Qiao, S. (2021). Attention-emotion-enhanced convolutional LSTM for sentiment analysis. IEEE transactions on Neural networks and learning systems, 33(9), 4332-4345.*
11. *Jelodar, H., Wang, Y., Rabbani, M., Ahmadi, S. B. B., Boukela, L., Zhao, R., & Larik, R. S. A. (2020). A NLP framework based on meaningful latent-topic detection and sentiment analysis via fuzzy lattice reasoning on youtube comments. Multimedia Tools and Applications, 80(3), 4155–4181. https://doi.org/10.1007/s11042-020-09755-z*
12. *Karpathy, A., Johnson, J., & Fei-Fei, L. (2015). Visualizing and understanding recurrent networks. arXiv preprint arXiv:1506.02078.*
13. *Kastrati, Z., Dalipi, F., Imran, A. S., Nuci, K. P., & Wani, M. A. (2021). Sentiment Analysis of Students' Feedback with NLP and Deep Learning:A Systematic Mapping Study. Applied Sciences, 11(9), 3986. https://doi.org/10.3390/app11093986*

ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 10| Issue: 7| July 2024|| Journal DOI: 10.36713/epra2013 || SJIF Impact Factor 2024: 8.402 || ISI Value: 1.188

14. *Li, X., Sun, X., Meng, Y., Liang, J., Wu, F., & Li, J. (2020). Dice Loss for Data-imbalanced NLP Tasks. Dice Loss for Data-imbalanced NLP Tasks. https://doi.org/10.18653/v1/2020.acl-main.45*
15. *Ma, T., Wu, Q., Jiang, H., Lin, J., Karlsson, B. F., Zhao, T., & Lin, C. Y. (2024). Decomposed Meta-Learning for Few-Shot Sequence Labeling. IEEE/ACM Transactions on Audio, Speech, and Language Processing.*
16. *Murdoch, W. J., & Szlam, A. (2017). Automatic rule extraction from long short term memory networks. arXiv preprint arXiv:1702.02540.*
17. *Musleh, D. A., Alkhwaja, I., Alkhwaja, A., Alghamdi, M., Abahussain, H., Alfawaz, F., Min-Allah, N., & Abdulqader, M. M. (2023). Arabic Sentiment Analysis of YouTube Comments: NLP-Based Machine Learning Approaches for Content Evaluation. Big Data and Cognitive Computing, 7(3), 127. https://doi.org/10.3390/bdcc7030127*
18. *Pascanu, R., Gulcehre, C., Cho, K., & Bengio, Y. (2013). How to construct deep recurrent neural networks. arXiv preprint arXiv:1312.6026.*
19. *Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., ... & Liu, P. J. (2020). Exploring the limits of transfer learning with a unified text-to-text transformer. Journal of machine learning research, 21(140), 1-67.*
20. *Serrano, J. C. M., Papakyriakopoulos, O., & Hegelich, S. (2020). NLP-based Feature Extraction for the Detection of COVID-19 Misinformation Videos on YouTube. COVID-19. https://openreview.net/pdf?id=M4wgkxaPcyj*
21. *[21]Ugray , Z., Lasdon, L., Plummer, J., Glover, F., Kelly, J., & Martí, R. (2007). Scatter Search and Local NLP Solvers: a multistart framework for global optimization. INFORMS Journal on Computing, 19(3), 328–340. https://doi.org/10.1287/ijoc.1060.0175*