Sentiment Analysis of YouTube Comments Using Bidirectional Encoder Representations from Transformers Neural Network Model

Pramila R. Gadyanavar[1][a], Mahantesh Laddi[2][b], Prashant S. Jadhav[3][c], Vaishali S. Katti[1][d], Jambukeshwar Pujari[1][e] and Azeem Javed Jamadar[1][f]

[1]*Department of CSBS, Kolhapur Institute of Technology's College of Engineering (Autonomous), Kolhapur, India*
[2]*Department of CSE, Bharatesh Institute of Technology, Belagavi, India*
[3]*Department of ME, Rajarambapu Institute of Technology, Rajaramnagar, Shivaji University, Kolhapur, India*

Keywords: BERT (Bidirectional Encoder Representations), AMB (Adapted Multimodal BERT), TCN (Temporal Convolutional Network), CNN (Convolutional Neural Network), NLP (Natural Language Processing).

Abstract: Everything in the today's world based on Sentiment. Sentiments are the Feelings that is based on Socially, Mentally, Economically, Psychologically Factors of the audience. Suppose you are multinational brand, and you want to know more about your Consumers Sentiments by figuring out by Looking at the people comment in your video at YouTube. It's very hard to analyse comments line by line, word by word. Practically it's not possible at that stage, because dealing with n numbers of comments are not possible. To Overcome these technical Situation, we are Introducing Our Sentiment Model that can Filtered the Audience Comments or sentiments. Sentiment Analysis is a natural Languages Processing Technique that is use know about the Sentiments in the text Mainly Positive Negative and Our model will Classify the YouTube comments Outcomes into five different Labels 1: "Very Negative", 2: "Negative" ,3: "Neutral" ,4: "Positive", 5: "Very Positive". It will also give some Insights from the data using some Visualization Technique like Bar Graph, Pie chart and Word Cloud.

1 INTRODUCTION

Everything in the today's world based on Sentiment. Sentiments are the Feelings that is based on Socially, Mentally, Economically, Psychologically Factors of the audience. Suppose you are multinational brand, and you want to know more about your Consumers Sentiments by figuring out by Looking at the people comment in your video at YouTube. It's very hard to analyse comments line by line, word by word. Practically it's not possible at that stage, because dealing with n numbers of comments are not possible. Not only have you had to deal you also wanted to know more about consumer's Sentiments.

This Problem is not only limited to Brands, when it comes to YouTube Content creator they have the huge impact of audience Sentiments. Many YouTube Creators Makes videos on the basis based on audience polarity or Sentiments, if they themselves did not have the knowledge of Audience How they will boost their channels and audience. Similarly, it applicable on constraints.

Comments are included in local language with multiple emoji or gifs, so classifying such comments using normal technique of sentiment analysis is efficient one. Because is sentiment analysis we have work with "context-based model". And all traditional methods are aspect-based models so here in this paper

[a] https://orcid.org/0009-0009-0586-918X
[b] https://orcid.org/0009-0002-9119-0657
[c] https://orcid.org /0000-0002-3102-6460
[d] https://orcid.org/0009-0004-4198-8171
[e] https://orcid.org/0000-0003-3422-4515
[f] https://orcid.org/0009-0001-1834-4667

Gadyanavar, P. R., Laddi, M., Jadhav, P., Katti, V. S., Pujari, J. and Jamadar, A. J.
Sentiment Analysis of YouTube Comments Using Bidirectional Encoder Representations from Transformers Neural Network Model.
DOI: 10.5220/0013640000004664

In *Proceedings of the 3rd International Conference on Futuristic Technology (INCOFT 2025) - Volume 3*, pages 683-689
ISBN: 978-989-758-763-4

we used sentence BERT technique for efficient classification of YouTube comments.

2 LITERATURE SURVEY

Text classification, named entity recognition (NER), and sentiment analysis benefit from the ability to assess contextual meanings in language using bidirectional context. In contrast, the autoregressive structure of the GPT model enables it to generate more coherent content and achieve effective results in tasks such as chatbots and creative writing. However, applying these models to agglutinative and morphologically complex languages, like Turkish, presents significant challenges. (M. Salıcı, 2024)

An Aspect-based sentiment Analysis method based on transformer-based deep learning models. It uses the Hugging Face Transformers library, which gives access to cutting-edge pre-trained models. For sentiment categorization, the approach used the BERT model, a potent transformer architecture. This study adds to the expanding field of sentiment analysis by offering a scalable and reliable transformer-based learning method for ABSA. This approach is a useful tool for applications in product reviews, and customer feedback analysis. (G. P. M S, 2024)

RoBERT sentiment analysis generates negative, positive and neutral sentiment in the analysed text. These scores are described independently, it is very difficult to complete understanding about the opinions in the comments. To propose a more reliable analysis of mixed emotions found in unstructured data, a composite sentiment summarizer combines positive, negative, and neutral scores into one. This composite score gives a more accuracy and reliable representation of the sentiment elaborated in comments. (Z. -Y. Lai, 2023)

For multimodal tasks, the Adapted Multimodal BERT (AMB) is a BERT-based model that integrates adapter modules with intermediate fusion layers. These fusion layers perform task-specific, layer-by-layer integration of audio-visual data and textual BERT representations. Meanwhile, the adapter adjusts the pre trained language model to suit the specific task at hand. This approach enables fast and parameter-efficient training by keeping the parameters of the pre trained language model frozen during the adaptation process. Research shows that this method yields effective models that are resilient to input noise and can outperform their fine-tuned counterparts. (O. S. Chlapanis, 2023)

The architecture comprising a temporal convolutional network (TCN), a convolutional layer, a bidirectional long short-term memory (BiLSTM), and robustly optimized bidirectional encoder representations from transformers pre-training approach (RoBERTa) is proposed to address issues. Dual branch feature coding network based on RoBERTa (DBN-Ro) is the name of the suggested architecture. The stitched vectors undergo dimensionality compression via the convolutional layer. (F. Wang, 2021)

To assess the performance of lexicon-based and sentence-BERT sentiment analysis models used as code-mixed, low-resource texts as input, it summarizes the results of experiments. Some code-mixed texts in Javanese and Bahasa Indonesia are utilized as a sample of low-resource code-mixed languages in this study. Google Machine Translation is used first to translate the raw dataset into English. The input text is translated into English and then classified using a pre-trained Sentence-BERT model. The dataset used in this study is divided into positive and negative categories. The experimentation found that the combined Google machine translator and Sentence-BERT model achieved 83 % average accuracy, 90 % average precision, 76 % average recall, and 83 % average F1 Score. (C. Tho, 2021)

Classifying sentiment is a crucial step in figuring out how individuals feel about a good, service, or subject. Sentiment classification and numerous models for natural language processing have been put forth. But most of them have focused on categorizing sentiment into two or three groups. The model tackles the fine-grained sentiment categorization task using a promising deep-learning model named BERT. Without a complex design, experiments demonstrate that the model performs better than other well-known models for this task. (M. Munikar, 2019)

Sentiment classification used as an Indonesian dataset, was explored with the problem utilizing a two-step procedure: sentiment classification and aspect detection. The bag-of-words vector, which is handled by a fully connected layer, and the word embedding vector, which is handled by a gated recurrent unit (GRU), are two deep neural network models with different input vectors and topologies for aspect detection that are compared. They also contrast two deep neural network methods for sentiment classification. Word embedding, sentiment lexicon, and POS tags are input vectors in the first method, which has a bi-GRU-based architecture. In the second, the word embedding vector is rescaled using an aspect matrix as the input vector, and the topology is based on a convolutional neural network (CNN). (AIlmania, 2018)

Four sentiment categories of data will be of the form positive, negative, neutral, and mixed in text messages using sentiment analysis and a soci al adaptive fuzzy similarity-
based categorization technique. It can also determine which emotion categories are most prevalent in the messages, such as happiness, excitement, rage, sadness, anxiety, and satisfaction. Additionally, it is integrated into a comprehensive social media analysis system that can gather, filter, categorize, and analyse text data from social media platforms and present a dashboard of descriptive and predictive metrics for a particular idea. The suggested approach has been created and is prepared for user licensing. (Z. Wang, 2018)

3 METHDOLOGY

Using Visualization and sentiments reports any Technical or non-technical person can have the Audience's Point of View what they are conveying in the comments and what they believe in your Videos. If these Sentiment Reports are used properly and analysed efficiently then it can be helping Brand and Content creator to boost their YouTube channel. Cause every Comment in the video have the specific value to the data.

Types of data used, and purpose means why you choose. For The Sentiment Report the Model is just getting ' Video ID ' as the input then Model will Automatically generate the Original Comments CSV file and then it will Generate another CSV file that contains Sentiments reports and Shows Visualization Using Graph as a output. For the sentiments we require the NLP model so that our Comments can be analysed on the Dataset, so we used “Bert-base-multilingual-uncased-sentiment " which is the variant of BERT fine-tuned specifically for multilingual sentiment analysis. It was developed by NLP Town, this model can be used for six different languages English, Dutch, German, French, Spanish, and Italian. And it predicts the sentiment of the review as a number of stars (between 1 and 5) which I have described earlier.

3.1 Sentiment Analysis

Main Methods for sentiments start with the Initializing Sentiment Pipelines, Basically Pipeline are nothing but the single function or module contains different processes of analysis, it is generally the subpart of scikit-learn which is a free and open-source machine learning library In our use case we require only two passing Parameters first is method name that is "Sentiment-Analysis" in our case and second we require to pass model which we used " BERT-base-multilingual-uncased-sentiment " that developed by NLP town Community.

After initializing and passing the parameters in pipelines function. The processor further

Will be working on the Sentiment Mapping. In our case we have described 1 star to 5 star

Based on Sentiments Score.

"1 ": "1 (Very Negative)",
"2 ": "2 (Negative)",
"3 ": "3 (Neutral)",
"4 ": "4 (Positive)",
"5 ": "5 (Very Positive)".

Now the mapping process has been also finished. Now we need to handle 'Truncate of tokens' Suppose if the length of the comment is to large. Then to avoid tokenisation limit we have to shorten the input texts to the first 512 characters to prevent issues with the model's limit on text length. This is the number of product reviews used for fine-tuning this model:

Language	Number of reviews
English	150k
Dutch	80k
German	137k
French	140k
Italian	72k
Spanish	50k

Using this multilingual sentiment, we can easily conclude sentiments of any comments. This model claims Accuracy of 67% the exact match for the number of stars. And When Accuracy Off by 1 it predicates 95% of a accuracy numbers were given by the human reviews.

3.1.1 Sentence-BERT (Bidirectional Encoder Representation from transformers) technique:

A transformer-based machine learning model called BERT (Bidirectional Encoder Representations from Transformers) was created for problems involving natural language processing (NLP). Google researchers created it, and its capacity to pre-train on large volumes of text and then fine-tune for tasks has transformed natural language processing (NLP) and resulted in notable advancements in tasks such as sentiment analysis, linguistic inference, and question answering. Following are the variants of BERT technique with varying numbers of layers (Encoders), attention heads, and hidden units.

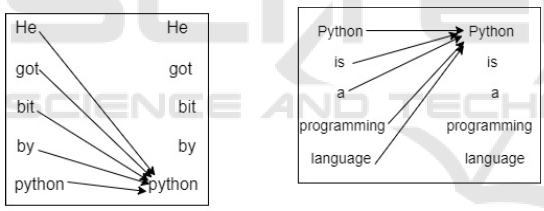
Table 1: Variants of BERT technique with varying numbers of layers

Variants	BERT Base	BERT Large	BERT Tiny	BERT Mini	BERT Small
Encoders (L)	12	24	2	4	4
Attention Heads (A)	12	16	2	4	4
Hidden (H)Units	768	1024	128	256	512

Working of BERT will encode the sentences with all possible keys in same sentence, so it is having unique encoding code for all words. And it also generates different code for same words in two different sentences. In following sentences, the word “python” means different. First sentence python is snake, and second sentence python is programming language.

Sentence 1: He got bit by python.

Sentence 2: Python is a programming language.



By Sentence- BERT technique can generate two different codes for two pythons, so it is context based model. And for sentiment analysis it is efficient model. In BERT it encodes each word to a vector it can be called as word2vec. Each of the word from the sentence carries unique encoded vector.

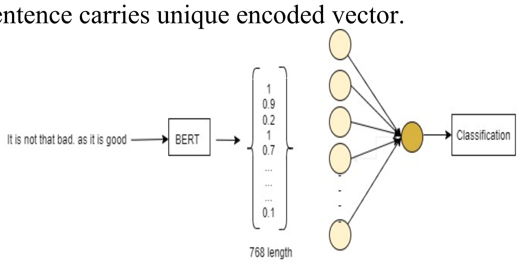


Figure. 1: Architecture of Sentence-BERT

Truncation can also refer to the process of converting data into a new record with smaller field lengths than the original. Truncates tokens that exceed a specified character limit. This limit defaults to 10 but can be customized using the length parameter. In fact, Many NLP couldn't Support Large Tokens in our Use Case BERT have a maximum token (Word or characters) limit that they can access in a single input. If the input text exceeds this limit, it can cause errors or lead to incomplete processing. By truncating, we ensure that The Model Can Process the Text Smoothly Staying within the model's toset next pageken limit avoids errors. Efficiency shortening long texts improves processing speed without significantly affecting sentiment accuracy.

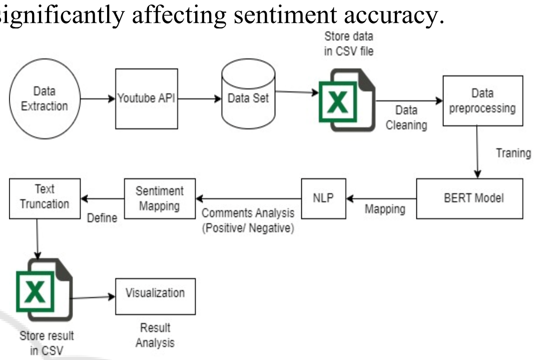


Figure. 2: Architecture Diagram

3.2 Data Cleaning

Most Important Part of any Sentiment Analysis is Data Cleaning Without cleaning the data, Model is imperfect. There are multiple data cleaning technique, but we have chosen the most appropriated and simple Technique that fits in our model that is "Regex". Regular Expressions, regex or regexp in short is extremely and amazingly powerful in searching and manipulating text strings, particularly in pre-processing text. Using this python library, we can easily check if a string contains the specified search pattern. Python has a built-in package called re, which can be used to work with Regular Expressions. In our Use Case, YouTube comments can contain URL’s, numbers, special characters, punctuations and some white spaces which do not play any role in our Analysis, so we need to remove all unessential part of comments, So We used "re" to manipulate pre-process Comments. Cleaning is essential because it increase model's Accuracy. And re is simplest version to get fitted.

3.3 Data Preparation

Data Preparation technique start with our first step in Research where we are trying to fetch All Possible YouTube Comments Using YouTube API. Using API, we are fetching Top Level Comments that can play an important our analysis. We are not taking comments reply,

Comment deleted, Spam Comment because it doesn't play an important role in sentiment analysis. In top levels comments that are organized by the YouTube we are fetching 'author' that is username of the account, 'published at' date when it was originally published, 'updated at' date when the comment was edited, 'like count' no of likes counts, 'text' is the Main comment. There is a limit of fetching a comment in one request so in our case it was 100 so to overcome this problem we used "Pagination" technique that can fetch all comments in one request. First Just we have to set " maxResults=100 " next it will check and get the result in one output. YouTube API, It stand for Application Programming Interference that allows developer to embed videos and offer other YouTube functionalities on your code and YouTube functionalities on your application. It functions as a client-server model, where the API acts as the intermediary between the client (your application) and the YouTube servers. This contract defines how the two communicate with each other using requests and Google Itself provides an environment to develop Such Applications. We just need to generate "Developer key" and do specific tasks.

3.4 Statistical Method

When the process of sentiment analysis has finished, we required some statistical tools to demonstrate our progress and visualize the reports and conclude some results using this visualization. For the visualization of Sentiments Currently we have used "Matplotlib" library of python which is used to create static, animated and interactive visualization in any applications. Using matplotlib we have drawn Bar Graph and Pie chart. Bar Graph typically Shows Sentiments Distribution of comments that shows how many comments fall in each category positive, negative or neutral.

Similarly, we have also draw pie chart that use to Show the proportion of each sentiment as a percentage. It exactly defines percentage of each Sentiments. Here each sentiment is represented as a slice of the pie. After Visualizing with Bar Graph and Pie chart we have also used Word cloud which represents the Most Frequent Word clouds.

4 RESULT AND DISCUSSION

YouTube API is fetched the comments, these comments will be encoded and classified using BERT.

Figure. 3: Original comment Dataset and Sentiment Reports

Following visualizations will describe very positive comments are above 55, positive comments are 15 in number, neutral comments near about 30, very negative comments are above 45, and negative comments are near about 5 in dataset.

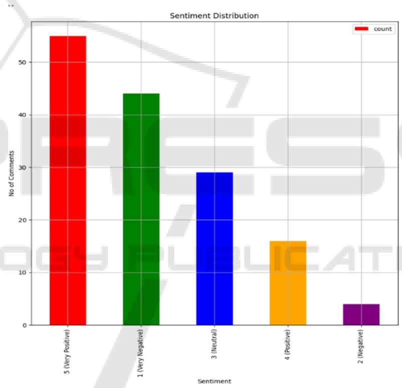


Figure. 4: Bar Graph

Each word's magnitude in a word cloud, which is a visual representation of text data, reflects how frequently or how important it is in the dataset. Larger font sizes are used for words that occur more frequently in the text, whereas smaller font sizes are used for words that occur less frequently. To make it easier to see which terms are most common, word clouds are frequently used to summarize, analyse, or illustrate textual data.

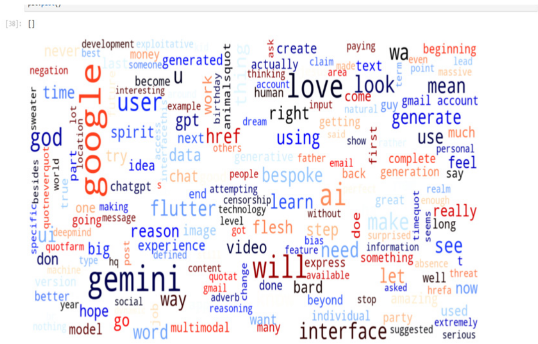


Figure 5: Word Cloud

Heatmap representing the confusion matrix, offering a detailed analysis of the model's classification performance and accuracy. For very negative statements model has predicted 4 comments accurately, for negative 48 comments are predicted accurately, foe neutral 48 comments are predicted accurately, for positive 12 comments are predicted accurately, very positive 67 comments are predicted accurately.

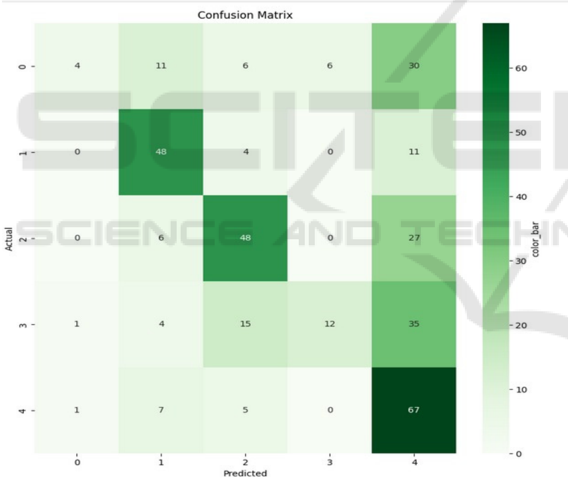


Figure 6: Confusion Matrix

5 CONCLUSIONS

The YouTube Comment Sentiment Analysis system successfully demonstrated the application of machine learning and natural language processing (NLP) techniques to analyse audience feedback. By automating the retrieval and sentiment classification of comments, the system provided actionable insights into how content is perceived. BERT algorithm is provided the accuracy for classification of average comments. We used fine tuning to increase the accuracy of model. If we go with the optimization, then again accuracy of the model will get increased.

REFERENCES

M. Salıcı and Ü. E. Ölçer, (2024) "Impact of Transformer-Based Models in NLP: An In-Depth Study on BERT and GPT," *2024 8th International Artificial Intelligence and Data Processing Symposium (IDAP)*, Malatya, Turkiye, 2024, pp. 1-6, doi: 10.1109/IDAP64064.2024.10710796.

G. P. M S, P. G. J, R. Kumar, P. K. H R, V. S. K and A. Dahiya, (2024) "Sentiment Analysis Using Transfer Learning for E-Commerce Websites," *2024 International Conference on Advances in Modern Age Technologies for Health and Engineering Science (AMATHE)*, Shivamogga, India, 2024, pp. 1-5, doi: 10.1109/AMATHE61652.2024.10582143.

Z. -Y. Lai, L. -Y. Ong and M. -C. Leow, (2023) "A Composite Sentiment Summarizer Score for Patient Reviews: Extending RoBERTa," *2023 11th International Conference on Information and Communication Technology (ICoICT)*, Melaka, Malaysia, 2023, pp. 405-410, doi: 10.1109/ICoICT58202.2023.10262598

O. S. Chlapanis, G. Paraskevopoulos and A. Potamianos, (2023) "Adapted Multimodal Bert with Layer-Wise Fusion for Sentiment Analysis," *ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Rhodes Island, Greece, 2023, pp. 1-5, doi: 10.1109/ICASSP49357.2023.10094923.

F. Wang, G. Liu, Z. Wang and X. Wu, (2021) "Sentiment Analysis of Movie Reviews based on Pre training and Dual Branch Coding," *2021 11th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)*, Cracow, Poland, 2021, pp. 1051-1055,doi: .1109/IDAACS53288.2021.9661049.

C. Tho, Y. Heryadi, I. H. Kartowisastro and W. Budiharto,(2021) "A Comparison of Lexicon-based and Transformer-based Sentiment Analysis on Code-mixed of Low-Resource Languages," *2021 1st International Conference on Computer Science and Artificial Intelligence (ICCSAI)*, Jakarta, Indonesia, 2021,pp. 81-85, doi: 10.1109/ICCSAI53272.2021.9609781

M. Munikar, S. Shakya and A. Shrestha,(2019) "Fine-grained Sentiment Classification using BERT," *2019 Artificial Intelligence for Transforming Business and Society (AITB)*, Kathmandu, Nepal, 2019, pp. 1-5, doi: 10.1109/AITB48515.2019.8947435.

AIlmania, Abdurrahman, S. Cahyawijaya and A. Purwarianti, (2018) "Aspect Detection and Sentiment Classification Using Deep Neural Network for Indonesian Aspect-Based Sentiment Analysis," *2018 International Conference on Asian Language Processing (IALP)*, Bandung, Indonesia, 2018, pp. 62-67, doi: 10.1109/IALP.2018.8629181.

Z. Wang, C. S. Chong, L. Lan, Y. Yang, S. Beng Ho and J. C. Tong,(2016) "Fine-grained sentiment analysis of social media with emotion sensing," *2016 Future Technologies Conference (FTC)*, San Francisco, CA, USA, 2016, pp. 1361-1364, doi: 10.1109/FTC.2016.7821783.

Talaat, A.S, (2023) Sentiment analysis classification system using hybrid BERT models. *J Big Data* **10**, 110. https://doi.org/10.1186/s40537-023-00781-w

Manuel-Ilie, Dorca & Gabriel, Pitic & George, Crețulescu. (2023). Sentiment Analysis Using Bert Model. International Journal of Advanced Statistics and IT&C for Economics and Life Sciences. 13. 59-66. 10.2478/ijasitels-2023-0007.

Alaparthi, S., Mishra, M. (2021) BERT: a sentiment analysis odyssey. *J Market Anal* **9**, 118–126 (2021). https://doi.org/10.1057/s41270-021-00109-8

R. Man and K. Lin, (2021) "Sentiment Analysis Algorithm Based on BERT and Convolutional Neural Network," *2021 IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC)*, Dalian, China, 2021, pp. 769-772, doi: 10.1109/IPEC51340.2021.9421110.

P. R. Gadyanavar and D. A. Kulkarni,(2018)"Temperature measurement using soft computing techniques based on computer vision," *2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS)*, Chennai, India, 2017, pp. 896-900, doi: 10.1109/ICECDS.2017.8389566.

Sudharsan Ravichandiran, (2021) *Getting Started with Google BERT: Build and train state-of-the-art natural language processing models using BERT*, Packt Publishing, 2021.

J. Yadav, D. Kumar and D. Chauhan,(2020)"Cyberbullying Detection using Pre-Trained BERT Model," *2020 International Conference on Electronics and Sustainable Communication Systems (ICESC)*, Coimbatore, India, 2020, pp. 1096-1100, doi: 10.1109/ICESC48915.2020.9155700.