NLP Approach To Sentimental Analysis: A Machine Learning Approach For Analysing Tweets

Dr.P.Sumathi[1], Dr.T.R.Lekhaa[2], R.Swathiramya[3], Karthik A[4], Nivash S[5], Mukilesh Kumar S[6], Nithyasri R[7]
[1]Associate Professor, Dept of Artificial Intelligence and Data Science
[2, 3]Assistant Professor, Dept of Artificial Intelligence and Data Science
[4, 5, 6, 7]Dept of Artificial Intelligence and Data Science
[1, 2, 3, 4, 5, 6, 7] SNS College of Engineering Coimbatore, Tamil Nadu, India

Abstract- *Sentiment analysis a technique or method that is commonly used to analyse the sentiment or emotions of people. This is a method that is applicable to all occasions and this is also applicable to the global pandemic of the new coronavirus (COVID-19) is driving unprecedented digital conversations on social media. Often expressed through tweets, these conversations provide valuable insight into public feelings, concerns, and attitudes toward the epidemic. In this study, we propose a novel approach to explore public sentiment towards COVID19 by using the bidirectional encoder representation from the transformer (BERT) model. Our research involves collecting an extensive dataset of coronavirus-specific tweets from social media platforms covering a key period. We then preprocess and clean the data to remove unnecessary noise and information. Then, we use this prepared dataset to fine-tune the BERT model, enabling us to understand the nuances and context of discussions about COVID-19. We use this refined model to analyse sentiment on a collection of tweets. Sentiment analysis provides a comprehensive view of public emotional responses to different aspects of the epidemic by classifying tweets into positive, negative, and neutral sentiments Furthermore, we use natural language processing techniques to advanced use to extract key topics and trends from tweets, so that public discourse A deeper understanding is possible. Furthermore, we examine temporal changes in sentiment and information during the pandemic, identifying significant changes in public sentiment as events unfold. These longitudinal surveys help monitor changing public perceptions and concerns, which can be valuable for policy makers and health professionals.*

Keywords- Data collection, Annotation, Preprocessing, Model architecture, Collaboration with psychology experts, Language preferences, Continuous monitoring, Tweet Analysis, Sentiment Analysis, BERT Sentiment Analysis.

I. INTRODUCTION

The advent of social media has changed the way we communicate and share information. In recent years, social media platforms have become veritable repositories of data, which captures the attention of billions of people around it, the world of emotions and their collective minds. As countries grappled, individuals took to social media to access information, express their concerns and communicate with others in social distance and isolation and shared the same. Public sentiment as expressed on social media has proven to be a valuable insight into understanding public responses to crises. It provides a real-time window into the collective emotions, fears, and hopes of a diverse global population. Exploring these sentiments can help governments, health professionals and researchers gauge public opinion, identify emerging trends and better respond to changing circumstances. We explore the interesting area of public sentiment analysis on tweets, exploiting the potential of a bidirectional encoder representation from transformer (BERT) model. We will explore the methods, applications, and significance of this research effort. Sentiment analysis, a key component of this research, categorizes tweets into positive, negative, or neutral sentiments, providing a comprehensive view of the public's emotional response towards different aspects of the pandemic. Moreover, sentiment analysis is not limited to gauging public reactions but extends to identifying sentiment drivers. By pinpointing the factors that influence sentiment—such as government policies, vaccine distribution, or public health communication—policymakers can tailor their strategies and messaging to address specific concerns and build trust within communities.

II. EXISTING SYSTEM

Existing approaches to Sentimental Analysis are Knowledge-based techniques (lexical-based approach), Statistical methods, and Hybrid approaches. Knowledge-based techniques make use of prebuilt lexicon sources containing polarity of sentiment words SentiWordNet (SWN) for determining the polarity of a tweet. Lexicon based approach suffers from poor recognition of sentiment. Statistical methods involve machine learning(such as SVM) and deep learning approaches, both approaches require labelled training data for polarity detection. Hybrid approach of sentiment analysis exploits both statistical methods and knowledge-based methods for polarity detection. It inherits high accuracy from machine learning(statistical methods) and stability from the lexicon-based approach.

2.1 Key Issues

1.Data Noise: Twitter data often contains noise in the form of typos, abbreviations, slang, and emojis, which can affect the accuracy of sentiment analysis.

2.Sarcasm and Irony: Identifying and interpreting sarcastic or ironic tweets presents a significant challenge due to the subtlety and context-dependent nature of such expressions.

3.Bias and Fairness: Bias in training data and algorithms can lead to unfair outcomes, requiring careful consideration and mitigation strategies to ensure fairness in sentiment analysis results.

4.Scalability: As the volume of Twitter data continues to grow rapidly, scalable AI and ML solutions are needed to process large datasets efficiently.

5.Ethical Considerations: Privacy concerns, data usage policies, and the potential for unintended consequences highlight the importance of ethical considerations in Twitter sentiment analysis projects.

2.2 Objective

To address this solution, we should collect the data from various sources like different websites, pdfs and word document. After collecting the data we will convert it into csv file then, we will break the data into individual sentences. Then by using Natural Language Processing (NLP) we eliminate stop words. Stop words are those words which are referred as useless words in the sentence or the extra data which are of no use. For example, "the", "a", "an", "in", are some of the examples of stop words in English. After that the algorithm naïve bayes is used to train the model. ANN algorithm works in backend to generate pickle file. Confusion matrix is used for validation technique and Accuracy is used for model defect

2.3 Scope

The scope of the project is to develop an advanced AI and ML-powered system for comprehensive Twitter sentiment analysis. It involves gathering extensive real-time Twitter data and applying robust preprocessing techniques to handle noise, language variations, and cultural nuances. Leveraging Natural Language Processing (NLP) and machine learning algorithms, the project aims to create models adept at accurately categorizing sentiments as positive, negative, or neutral. Overcoming challenges like sarcasm, slang, contextual sentiments, and multilingual content is crucial for nuanced analysis. The system will provide real-time analysis, delivering timely insights into public opinions, emotions, and trends related to specific topics, brands, events, or products on Twitter. Scalability is a key consideration, necessitating models capable of handling large data volumes and adapting to evolving language trends and user behaviors. Evaluation metrics will assess model accuracy and performance, presenting sentiment analysis results in an accessible format for easy interpretation and utilization of the derived insights

III. SOFTWARE REQUIRMENTS

3.1 Programming Language

Python serves as the primary coding language for its versatility, ease of use, and extensive library support. Python's simplicity and readability make it an ideal choice for developing complex applications.Python facilitates the preprocessing of data, feature extraction, model training, and inference. With its simplicity and versatility, Python enables the creation of robust models capable of accurately classifying tweets into positive, negative, or neutral categories, empowering insightful decision-making in various domains.

3.2 System Requirements

RAM: 4GB or 8GB
Windows 10
Intel Core i5/i7 processor
At least 60 GB of Usable Hard Disk Space

3.3 Libraries

1.PANDAS: Pandas is a Python library used for data manipulation and analysis. It offers easy-to-use data structures and tools for handling structured data, facilitating tasks such as cleaning, transforming, slicing, and analysing tabular data, making it efficient for data exploration and manipulation in data science and analysis workflows.Pandas is a powerful Python library used for data manipulation and analysis. It provides data structures and functions to efficiently work with structured data,

2.NUMPY: NumPy is the module of the Python. The numpy word basically shows Numerical Python and it is utilized. This is the module which is basically written in c language and is said as expansion module .Numpy guarantee remarkable execution speed. Numpy is mostly used for performing calculations, tasks using certain functions it provides like multiply, divide, power etc.

3.TENSORFLOW: TensorFlow is an open-source machine learning library developed by Google. It enables building and training machine learning models using neural networks. TensorFlow provides tools for creating complex computational graphs, allowing efficient numerical computation, primarily used for tasks like deep learning, neural network implementation, and model deployment across various platforms.

4.SEABORN:Seaborn is a powerful Python data visualization library built on Matplotlib. It provides a high-level interface for creating aesthetically pleasing statistical graphics. With concise syntax, Seaborn allows the generation of informative and complex visualizations such as scatter plots, histograms, heatmaps, and regression plots. It simplifies data exploration and presentation, offering tools for statistical estimation and enhancing the quality of visual representations in data analysis and research.

5.EMOJI: Emoji package in Python allows easy handling and manipulation of emojis within text data. It provides functions to find, extract, and manipulate emojis, convert emojis to their corresponding Unicode representation, and vice versa. This package enables tasks like counting emojis, replacing emojis with text, and adding emojis to strings, facilitating emoji-related processing and manipulation in text analysis, social media applications, or data processing workflows.

6.NLTK : The NLTK (Natural Language Toolkit) in Python serves as a comprehensive platform for working with human language data. It offers tools, libraries, and resources for tasks such as tokenization, stemming, lemmatization, part-of-speech tagging, parsing, and semantic reasoning. NLTK provides functionalities for text classification, language modelling, and corpora management, making it a valuable resource for natural language processing (NLP), text analysis, and machine learning applications in linguistic research and text processing tasks.

3.4 System Integration and Testing

The Twitter sentiment analysis project using AI and ML involves several crucial phases for system integration and testing. These phases include the development of individual components/modules for data collection, preprocessing, feature extraction, model training, evaluation, and inference, which are then integrated into a cohesive pipeline for sentiment analysis on Twitter data. This ensures seamless communication and data flow between different modules, maintaining the integrity of the system.

Unit Testing :Unit testing is conducted to verify the functionality of each component/module, including input validation, data processing, and output generation. Testing frameworks like pytest or unittest automate and streamline the testing process. Integration testing validates the functionality and interoperability of integrated modules within the sentiment analysis pipeline, ensuring that each module produces expected outputs and handles errors gracefully.

Performance Testing :Performance testing assesses the system's performance in terms of speed, memory usage, and accuracy under various conditions. Benchmarking the system against performance metrics and requirements can identify potential bottlenecks or optimization opportunities. End-to-end testing is conducted to validate the entire sentiment analysis process, using representative datasets covering diverse topics, sentiments, and user demographics. Edge cases, such as tweets with ambiguous sentiment or unusual language constructs, are tested to ensure robustness.

Error handling and logging :Error handling and logging mechanisms are implemented to capture and manage errors effectively during testing and deployment. Monitoring system logs and error reports helps identify and address issues proactively. Finally, validation and verification are conducted to validate the accuracy and reliability of sentiment analysis results against ground truth labels or human annotations, and to solicit feedback from stakeholders and domain experts.

These algorithms serve as the intelligence behind the system, enabling it to interpret received data and determine the tweets.

3.5 Deployment and Maintenances

Deploying a Twitter sentiment analysis system involves a series of intricate steps to ensure its effectiveness and efficiency over time. Initially, the development of such a system begins with the collection of a substantial amount of Twitter data, which can be achieved through Twitter's API, allowing for the gathering of tweets relevant to the chosen topic or keyword. Preprocessing this data is crucial, as tweets often contain slang, typos, and emojis, necessitating a cleaning process to standardize the dataset for analysis. Once the data is cleaned, the next step involves feature extraction, where techniques such as tokenization and vectorization are applied to transform textual data into a format that machine learning models can understand.

The core of the system relies on selecting the right machine learning algorithm; common choices include Naive Bayes, Logistic Regression, or more complex models like Neural Networks, each with their strengths in handling natural language data. Training the chosen model requires dividing the dataset into training and testing sets, a crucial step to evaluate the model's performance accurately. Hyperparameter tuning and cross-validation are then employed to refine the model, ensuring it can generalize well to unseen data. After training, the model's performance is evaluated using metrics such as accuracy, precision, recall, and F1 score to ensure it meets the expectations for sentiment analysis.

Deploying the trained model into a production environment marks a significant phase where the model starts analysing real-time Twitter data. This deployment can be facilitated through cloud services or on-premises servers, depending on the scale and resources available. It's essential to implement an API around the model to enable seamless integration with other applications or services, allowing for live sentiment analysis of tweets. Maintenance becomes a continuous task post-deployment, involving monitoring the system's performance, updating the datasets with new Twitter data, and retraining the model to adapt to the evolving nature of language and sentiments expressed on social media.

Over time, feedback loops can be established to capture inaccuracies or areas where the model may not perform well, feeding this information back into the training process for continuous improvement. Adjusting to the fast-paced changes in social media requires keeping the sentiment analysis model updated with the latest trends, slang, and emojis that emerge on Twitter. Scalability is another critical aspect, ensuring that the system can handle increasing volumes of data without compromising performance. Automating the retraining process can help in maintaining the model's accuracy over time, using updated datasets to reflect the current state of language use on Twitter.

IV. IDEATE

4.1 Proposed System

4.1.1 Module 1: Data Collection and Preprocessing Gather a comprehensive dataset of coronavirus-specific tweets from various sources, including twitter API and other relevant sources we have Pre-processed the data to ensure quality and uniformity, including text cleaning, tokenization, and the removal of noise such as retweets and irrelevant hashtags. Data is gathering is done in a specific time that is only during the covid times. Data gathering is an essential phase in the research process, involving the systematic collection of information relevant to a particular topic or study. This crucial step involves the acquisition of data from various credible sources, such as surveys, interviews, observations, databases, and academic literature. The collected data is then carefully organized and analysed to draw meaningful insights and conclusions. It is imperative to ensure the accuracy, relevance, and reliability of the data through proper validation and verification methods, maintaining the highest ethical standards throughout the process. Effective data gathering lays the foundation for informed decision-making and successful research outcomes.

4.1.2 Module2 : Data Labelling Annotate the dataset with sentiment labels (positive, negative, neutral) to create a ground truth for supervised sentiment analysis as we use human annotators or pre-existing sentiment labels if available. Positive negative and natural are considered as sentiments of humans. Data labelling is a critical task in the realm of data processing and analysis, involving the annotation or tagging of raw data to enhance its usability and applicability. This process entails assigning specific labels, categories, or tags to data points based on predefined criteria or guidelines. These labels provide context and meaning to the data, making it understandable and usable for machine learning algorithms or analytical purposes. Careful and accurate labelling is vital to ensure the quality and reliability of the labelled dataset, improving the performance and effectiveness of subsequent machine learning models and data-driven applications. Adhering to consistent and clear labelling guidelines is essential to mitigate errors and maintain data integrity throughout the labelling process.

4.1.3 Module 3: BERT Model Preparation Objective: Prepare the BERT model for fine-tuning. This includes loading the pretrained BERT model, and configuring it for the sentiment analysis task. Preparing the token embeddings, segment embeddings, and position embeddings. Preparing a BERT model involves several crucial steps to ensure its optimal performance and accuracy in natural language processing tasks. Initially, it requires selecting a suitable BERT variant, considering the task at hand and the available computational resources. Once the BERT variant is chosen, the pre-trained model weights and architecture are obtained from reputable sources like Hugging Face or TensorFlow. Next, tokenization is applied to convert raw text data into BERT-compatible tokenized format, breaking down sentences into sub words or tokens. This step ensures uniformity and compatibility with BERT's input requirements .After tokenization, input data is organized into appropriate formats, such as tensors, to be fed into the BERT model. This includes creating token embeddings, segment embeddings, and positional embeddings, which are crucial for BERT's understanding of context and structure in the input text. xFine-tuning the pre-trained BERT model is the subsequent step, involving additional training on specific tasks using task-specific datasets. This process fine-tunes the model's parameters to adapt it for the desired NLP task, such as sentiment analysis, named entity recognition, or question-answering.Throughout the preparation process, it's important to ensure proper hyperparameter tuning, model evaluation, and validation techniques to optimize the BERT model's performance and achieve the desired accuracy for the targeted NLP application.

4.1.4 Module 4: Fine-Tuning BERT Fine-tune the BERT model on the labelled coronavirus-specific tweet dataset. Train the model to classify tweets into positive, negative, or neutral sentiments. The fine tuning process takes data and train the model into positive negative and neutral as the sentiments concerns. Fine-tuning is a crucial step in leveraging pre-trained models like BERT to adapt them for specific natural language processing (NLP) tasks. The process involves taking a pre-trained model (like a pretrained BERT model) and training it further on a task-specific dataset to customize its parameters and make it proficient in performing a particular NLP task.

4.1.5 Module 5: Model Evaluation and Validation Evaluate the performance of the fine-tuned BERT model using various sentiment analysis metrics, such as accuracy, precision and recall. Validate the model's generalization on a separate test dataset. Model evaluation is a crucial step in machine learning and data science, determining the performance, accuracy, and reliability of a trained model. It helps assess how well the model generalizes to unseen data and whether it's suitable for its intended task. Various metrics and techniques are employed to evaluate models, ensuring they meet the desired performance criteria.

4.1.6 Module 6: Real-Time Sentiment Analysis Gaining data insights from the trained model from the BERT analysis. Implement a realtime sentiment analysis module that continuously monitors and classifies incoming coronavirusrelated tweets using the fine-tuned BERT model. Sentiment analysis, also known as opinion mining, is a natural language processing (NLP) technique used to determine and classify the sentiment expressed in a piece of text. It involves analysing and understanding the attitudes, opinions, emotions, and feelings expressed by individuals towards a particular subject or topic. Sentiment analysis is widely used in various domains, including marketing, customer service, social media monitoring, product reviews, and public opinion analysis.

4.1.7 Module 7: Predictive Modelling Predictive modelling for sentiment analysis using BERT involves utilizing the pretrained BERT model and fine-tuning it on labelled sentiment analysis datasets. To perform sentiment analysis using BERT, start by obtaining a pre-trained BERT model and the associated tokenizer. Next, prepare a labelled dataset for sentiment analysis, where each text sample is labelled with its corresponding sentiment (e.g., positive, negative, neutral). Tokenize the text data using the BERT tokenizer, ensuring proper formatting for BERT input. Then, fine-tune the pre-trained BERT model on the labelled dataset. Modify the model architecture to suit the sentiment analysis task and add a suitable classification layer (e.g., softmax layer) for sentiment prediction. Train the model on the labelled data, adjusting hyperparameters and employing techniques like crossvalidation to optimize performance .Develop a predictive model that uses sentiment data to forecast future sentiment trends, enabling proactive responses to potential public concerns and crises. Training the model for future predictions and also for further considerations.

4.1.8 Module 8: Reporting and Insights Generate regular reports summarizing sentiment trends, key concerns, and actionable insights for policymakers, healthcare authorities, and the public. These reports play a crucial role in making informed decisions and adjusting communication strategies to address evolving sentiments surrounding COVID-19. These proposed work modules provide a structured approach to conducting sentiment analysis of coronavirus-specific tweets using BERT. Each module contributes to the overall goal of understanding public sentiment, monitoring trends, and providing actionable insights during the COVID-19 pandemic

4.2 Advantages

The utilization of pertinent information plays a crucial role in facilitating more judicious decisions regarding resource allocation, enhancing organizational efficiency, refining products and services, and ultimately elevating the quality of life for citizens and fostering more harmonious human interactions. This, in turn, contributes to the overarching objective of cultivating a more advanced society. A prime illustration of such application resides in the profound impact derived from monitoring public sentiment concerning various products, services, and events. This practice equips enterprise managers with invaluable insights and benchmarks, thus significantly informing their decision-making processes. Furthermore, administrators within city councils stand to benefit markedly from this approach. Through a deeper understanding of communal sentiments, these officials are better positioned to augment the services rendered to citizens, and to tackle the pressing issues of development and sustainability with heightened efficacy.

In the contemporary epoch, social media platforms have emerged as the predominant forums for the collection and analysis of public sentiment. Individuals are afforded the liberty to express and disseminate their thoughts, opinions, and reactions to a wide array of subjects, ranging from personal experiences to prevalent societal issues. The wealth of data accessible via social media is not confined to surface-level content; it extends to hidden metadata, which includes, but is not limited to, the language of the operating system, the type of device employed, the time of content capture, and the geographical positioning. This metadata, when analyzed, can unveil additional layers of context, thereby enriching the understanding of public sentiment and enhancing the precision of subsequent actions taken by organizations and governmental entities alike.

V. RESULT AND SCREESNSHOTS

5.1 Input

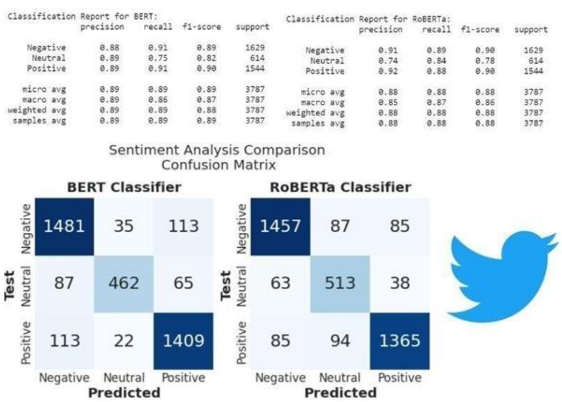

Fig 1: Results of analysis of BERT

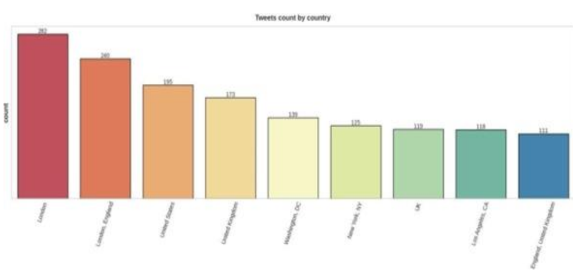

Fig 2: Dataset Analysis – Tweets by Country

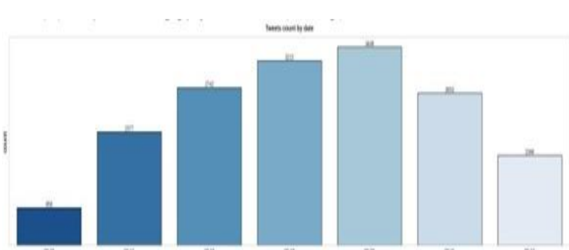

Fig 3: Dataset Analysis – Tweet Counts

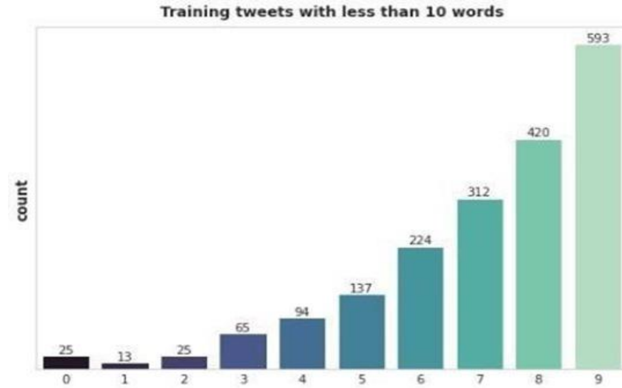

Fig 4 :Data visualisation of Tweets

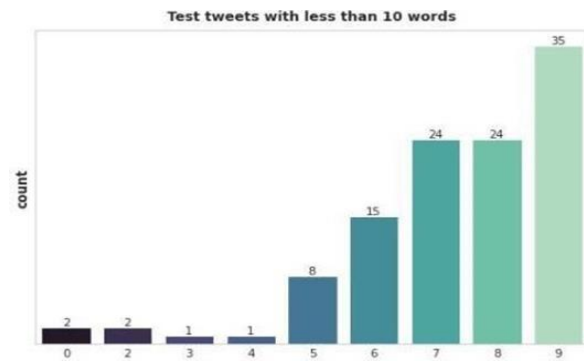

Fig 5 : Data Visualisation of Test tweets

Classification Report for Naive Bayes:

	precision	recall	f1-score	support
Negative	0.70	0.78	0.74	1629
Neutral	0.60	0.47	0.53	614
Positive	0.73	0.72	0.73	1544
accuracy			0.70	3787
macro avg	0.68	0.66	0.66	3787
weighted avg	0.70	0.70	0.70	3787

Fig 6 :Classification Report

5.2 Output

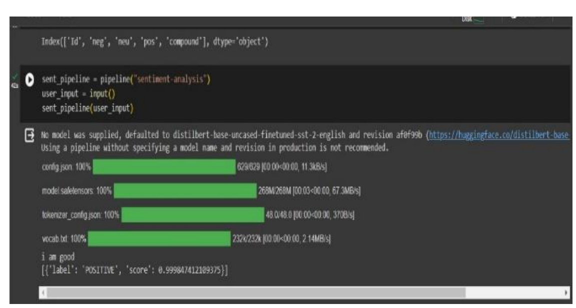

Fig 7: Output Page-1

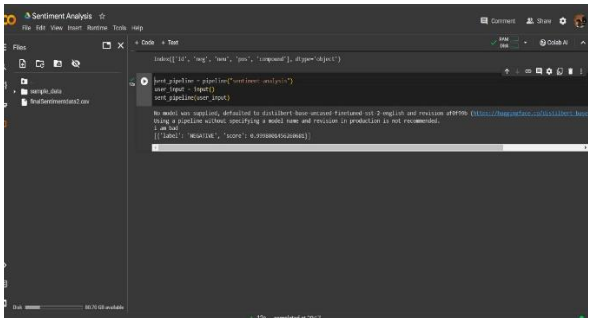

Fig 8: Output Page-2

VI. CONCLUSION

In conclusion, the use of BERT to analyse public sentiment A COVID-specific tweet has provided valuable insight public sentiment is complex and constantly changing during the prevalence pandemic. Systematically, including data collection, preprocessing, optimization of the BERT model, and real-time sensitivity analysis. We got a subtle understanding of emotion, and we got started with positivity and optimism despair and anxiety. This emotional analysis is not only It helps manage public sentiment but also allows them to determine what is important Issues and concerns expressed by the public. The ability to measure the impact of A variety of events and sensory interventions has helped here Refine public health strategies and communication efforts. Moreover, we. The approach extends beyond mere research, as we often predict Modelling, predicting potential changes in perception based on upcoming events. This prompt positioning empowers decisionmakers to respond quickly and effectively It ensures problem-solving with greater flexibility for emerging societal problems draft. Although BERT has become and continues to be a powerful tool for emotional analysis It is not without challenges. Changing models of language, addressed Biases in data, and support for ethical considerations in handling emotions remain an ongoing concern. But our determination to. Continuous improvement and the ethical application of emotional research He remains unmoved. Provide actionable reports and insights on a regular basis. Our goal for stakeholders is to help them make informed and effective decisions through public engagement throughout the COVID-19 pandemic.

VII. FUTURE SCOPE

Sentiment analysis, also known as mind mining, is a method of determining the emotions or feelings behind information, such as social media posts, reviews, customer feedback or artificial intelligence (AI) and in natural language processing With rapid advances in (NLP) technology, the future of cognitive analytics is poised for major changes, making it more accurate, efficient and insightful than ever before Traditionally, sentiment analysis has relied on specialized methods such as keyword matching, where specific words or phrases are associated with positive or negative emotions but this approach has limitations, as it is often not capturing people-based understanding of language nuances, such as sarcasm, humour, or context It also uses advanced machine learning algorithms and NLP techniques to accurately identify underlying emotions. A key advancement in AI-enabled sentiment analysis has been the use of deep learning models, such as recurrent neural networks (RNNs) and transformers These models can process large amounts of data and they recognize complex structures and relationships between words and sentences Consequently, they are better able to capture linguistic context and nuance, making sense analysis more accurate and reliable. Another important advancement in AI-powered sentiment analysis is the ability to analyse text in multiple languages. and important due to the increasing globalization of industry and the internet. Another important advancement in AI-powered sentiment analysis is the ability to analyse text in multiple languages. Globalization and the rise of the Internet have made the need for multilingual sensitivity analysis more important than ever. AI-powered sentiment analytics tools are now able to automatically search and analyse content in different languages, giving companies insights into consumer moods and sentiments across markets and regions Furthermore, AI-powered sentiment analysis tools perform well in many contexts. 33 AI-powered sentiment analysis tools can rapidly process and analyse vast amounts of data, providing real-time insights and enabling businesses to make data-driven decisions faster. With these advancements followed by AI-powered sentiment analysis Even the ability to identify and analyse emotions beyond their original categories of positive, negative and neutral. Using advanced NLP techniques and machine learning algorithms, AI-powered sentiment analysis tools can now identify and categorize a wide range of emotions such as happiness, anger, fear and shock thereby enabling businesses and researchers to navigate and understand the emotional state of their target audience and be able to make informed decisions.

REFERENCES

[1] G. G. Jayasurya, S. Kumar, B. K. Singh, and V. Kumar, "Analysis of Public Sentiment on COVID-19 Vaccination Using Twitter," in IEEE Transactions on Computational Social Systems, vol. 9, no. 4, pp. 1101-1111, Aug. 2022, doi: 10.1109/TCSS.2021.3122439.

[2] N. Rifat, M. Ahsan, R. Gomes and M. Chowdhury, "COVID-19 Sentiment Analysisapplying BERT," 2022 IEEE International Conference on Electro-Information

Technology(eIT), Mankato, MN, USA, 2022, pp. 417-422, doi: 10.1109/eIT53891.2022.9813777.

[3] D. SakethNath, H. Kaur and A. Singh, "Emotion Detection using Natural Language Processing," 2022 5th International Conference on Contemporary Computing and Informatics (IC3I), Uttar Pradesh, India, 2022, pp.

[4] A. Joshy and S. Sundar, "Analyzing the Performance of Sentiment Analysis using BERT, DistilBERT, and RoBERTa," 2022 IEEE International Power and Renewable Energy Conference (IPRECON), Kollam, India, 2022, pp.

[5] R. Olusegun, T. Oladunni, H. Audu, Y. Houkpati and S. Bengesi, "Text Mining and Emotion Classification on Monkeypox Twitter Dataset: A Deep Learning-NaturalLanguage Processing (NLP) Approach," in IEEE Access, vol. 11, pp.

[6] Rustum F. Khalid M, Aslam W, Rupapara V, Mehmood A, Choi GS (2021) A performance comparison of supervised machine learning models for Covid-19 tweets sentiment analysis. PLoS ONE 16(2): e0245909. doi:10.1371/journal.pone.0245909

[7] İ. Aygün, B. Kaya and M. Kaya, "Aspect-Based Twitter Sentiment Analysis on Vaccination and Vaccine Types in COVID-19 Pandemic With Deep Learning," in IEEE Journal of Biomedical and Health Informatics, vol. 26, no. 5, pp. 2360-2369, May 2022, doi:10.1109/JBHI.2021.3133103.

[8] [4] T. Wang, K. Lu, K. P. Chow and Q. Zhu, "COVID-19 Sensing: Negative Sentiment Analysis on Social Media in China via BERT Model," in IEEE Access, vol. 8, pp. 138162138169, 2020, doi: 10.1109/ACCESS.2020.3012595.