AI BASED MENTAL HEALTH CHATBOT USING SENTIMENT ANALYSIS AND MULTICLASS TEXT CLASSIFICATION

Chinmai Tarphe[1], Varsha Mharse[2], Sandhya Kaprawan[3]

[123]Department of Information Technology, University of Mumbai, India.

Abstract - *Mental health problems like stress, anxiety and depression are growing fast and emotional help is still not readily available because social stigma and inadequate care systems hinder timely assistance for mental health issues. This research work presents an Artificial Intelligence based Mental Health Chatbot which can offer immediate, non-judgemental and continuous emotional support to individuals through conversational interaction. It uses Natural Language Processing (NLP) and Machine Learning algorithms to process the text input by the user, detect emotional patterns, and provide relevant replies. Data preprocessing was done and then loaded to the data set building algorithms such as tokenization, stop words removal, and vectorization. The system's emotional state awareness will be improved by applying machine learning algorithms, such as XG Boost and NLP classification. The model works well in emotion recognition and giving proper supportive responses, and based on our observation and testing, it is able to improve the accuracy of the conversation, as well as creating dynamic responses and overall effect of conversation better than rule-based methods using the experiments. The chatbot is not designed to diagnose, treat or replace a mental health professional, rather it is a tool that is easy to access as a first help aid that helps individuals learn to voice themselves freely and seek help*.

Key Words: Mental Health, Sentiment Analysis, NLP, Depression and Stress Detection, AI Chatbots, Machine Learning .

1. INTRODUCTION

Stress, anxiety, depression and emotional distress are mental health disorders that have reached the large scale, and are now a major health concern for millions of people around the world. AI, NLP and ML techniques are growing in sophistication to develop intelligent conversational systems providing scalable, accessible mental health support [21]. Mental health competencies-based AI chatbots developed into a realistic and practical digital environment where someone can converse with a human-like chatbot, receive an analysis of their emotions, and receive a response in seconds. They use both sentiment detection and contextual understanding, along with deep learning architectures, to identify feelings based on text input and transform it into a personalized response. According to the research studies, sentiment analysis combined with NLP techniques has significantly improved the chatbot's capacity to identify the sentiment of the user and provide context-relevant assistance.

2. RESEARCH OBJECTIVES

- To create a mental health chatbot powered by Artificial Intelligence (AI) and leveraging Natural Language Processing (NLP) and Machine Learning (ML) methods.
- To use multi-class classification methods to classify multiple emotional states of stress, anxiety, depression, loneliness and neutral states.
- To enhance the accuracy of chatbots' responses and their understanding of the context thanks to deep learning models like Bi-LSTM and NLP-based classifiers.
- To offer accessible, real-time and non-judgemental mental health assistance with conversational AI.
- Assess the framework's performance on the basis of the accuracy of classification, the relevance of the responses and the ability to detect the emotions.

3. RESEARCH QUESTIONS

1. What are the different methods for multiclass classification and how well do they differentiate between different psychological disorders, when analysed from user dialogue?

2. Can NLP and deep learning based models generate responses more personal and in context than the rule based chatbots?
3. How do AI-driven conversational systems gives an improved gateway for the initial mental health support?
4. How effective is it in the proposed structure to detect such emotions as anxiety stress depression and loneliness?
5. What are the main challenges and ethical concerns with the use of AI-powered chatbots in mental health?

4. LITERATURE REVIEW

AI and NLP are revolutionizing digital mental health platforms by allowing for the creation of chatbots that comprehend and engage with human emotions. Numerous research papers have investigated different strategies for the creation of mental health talkbots that offer emotional support, emotional assistance, and sentiment-based communication.

Abd-Alrazaq et al. [1] did a thorough review of chatbot systems in mental health care applications. Their research underscored the benefits of AI-driven conversational applications for improving access to mental health care. The authors determined that a chatbot system can be an immediate helper and lessen the stigma of traditional therapy. Due to privacy issues, emotional issues, and ethical concerns, the study also uncovered some limitations.

In A. R. et al. [2] proposed an AI-powered mental health care chatbot using sentiment analysis techniques. This system employed two Bidirectional Long Short-Term Memory (Bi-LSTM) models that could be used to categorize emotional states based on the text. The model was able to accurately identify emotional patterns, but it was found to be highly sensitive to the quality of the data sets and the methods used for preprocessing the data.

Mmbayi et al. [3] developed a sentiment analysis-based monitoring model for mental health chatbots to assess user engagement with the chatbots. The study highlighted the importance of the use of NLP techniques in enhancing emotional interaction and conversational quality. Results showed that sentiment based systems could be used to effectively track user behavior and engagement over time.

Balcombe [5] investigated the use of AI chatbots in the field of digital mental health and examined Human–Artificial Intelligence (HAI) approaches that emphasize empathy, ethics, and usercentered design. The study emphasized trust, privacy security, and ethical utilization of AI in mental health apps.

Casu et al. [6] conducted a scoping review of AI chatbot systems in mental healthcare and investigated their effectiveness, feasibility and realistic application. The study's findings indicated that AI chatbots can deliver scalable and cost-effective support mechanisms. But, there were some limitations like emotional misunderstanding, contextually misunderstanding and lack of long term clinical validation.

Recent research by Siddals et al. [7] and Manole et al. [8] also highlighted the increasing capabilities of generative AI chatbots for managing anxiety and providing emotional support. These research papers highlighted the significance of sophisticated NLP models in comprehending the emotional context and creating tailored responses.

The emotion prediction proposed by Mehta et al. [17] is given by the following confusion matrix. Overall the model achieves the multiclass classification performance with an accuracy of 90% as reported in [17].

Table-1: Comparative Analysis of AI-Based Mental Health Chatbot Models

Reference number	Model used	Accuracy(%)	Data used
[1]	Rule-Based & AI Based Mental Health Chatbots (CBT-based Chatbots)	Not Mentioned /It is Scoping review.	Mental health chatbot studies, user interaction datasets, depression/anxiety/autism/PTSD related data
[2]	Bidirectional LSTM	80.88 %	Tweet dataset
[3]	BERT,Random Forest, SVM, Logistic Regression, SGD, Naïve Bayes	BERT: 99.18%, Random Forest: 98.18%	82,102 mental health chatbot reviews collected from Google Play Store and Apple App Store.
[5]	AI Chatbots, NLP, ML, CBT-based Conversational Agents	Not mentioned	Literature review studies, digital mental health datasets, chatbot interaction studies, Woebot & Wysa case studies
[9]	CBT-based AI Chatbot / NLP Conversational Agent	Not Mentioned	Mental health chatbot interaction data and app studies
[12]	Generative AI Companion Chatbots (Cleverbot, Simsimi, GPT-3)	Recognition Accuracy: 61.9%	Conversation datasets from Cleverbot, Simsimi, and GPT-3 chatbot interactions)

[13]	Dialogflow NLPbased Emotional Disclosure Chatbot (Woebot-inspired CBT Chatbot)	**behavioral/user study paper:** User Satisfaction Score:5.41/7, Reuse Intention: 5.34/7	348 participant responses collected through Amazon Mechanical Turk mental health counseling interactions.
[20]	AI-Based Mental Health Chatbots / NLP Conversational Agents	Not Mentioned it is framework and literature review paper.	34 research papers on mental health chatbots, public service chatbot studies, COVID-19 mental health datasets, user interaction studies
[22]	LSTM-RNN (Deep Learning Model)	99.9%	Twitter Dataset (Sentiment140, TWINT scraped tweets, Google Word2Vec)
[23]	NLP and Machine Learning-based Sentiment Analysis Model	88%	10,000 social media posts from Twitter, Instagram, Facebook, and Reddit

Table-2: Comparative Analysis of Existing Mental Health Chatbots.

Chatbot	Therapeutic Approach	Prompt Strategies	Target Audience	Limitations
Woebot	CBT-focused, anxiety and stress management	Prescripted conversation flows	Adults	Limited adaptive personalization; minimal role assignment customization
Wysa	CBT-focused, mindfulness, anxiety, and stress management	Empathetic but generic prompts	Teens, adults	Restricted context integration; lacks dynamic real-time updates
Replika	CBT-focused, mindfulness, general supportive talk	Passive, user-led queries	General	Inconsistent therapy alignment; uncertain crisis management features

Tess	CBT and motivational interviewing	Thematically structured modules	General	Basic prompt cues; lacks deeper user-specific context analysis
Talkspace Chatbot	Guided self-help and therapy access facilitation	Guided conversation flows with service integration prompts	General	Limited direct therapeutic intervention; primarily a gateway to paid services; not a replacement for therapy
Rasa	Modular framework (open source)	Customizable NLU and dialogue flows; partial script expansions	Developers	Requires specialized design; does not inherently provide integrated therapeutic cues
Youper	AI-driven CBT and mindfulness	Adaptive daily check-ins	General	May lack advanced role assignment; personalization limited to user mood surveys
Sanvello	CBT, mindfulness, relaxation	Prebuilt guided journeys and journaling prompts	General	Primarily static prompts; expansions often generic, lacking advanced user context
MindShift	CBT for anxiety	Topic-focused modules, psychoeducation	General	Narrow domain focus (anxiety); conversation logic not strongly customized
Ada AI	Symptom checker + chat support	Basic triage-style prompts; limited mental health focus	General	Lacks robust therapy alignment; prompt structure is primarily biomedical

The above table is referred from [21].

As per the literature reviewed, it can be seen that the existing chatbot systems are primarily concerned about simple sentiment analysis and fixed routines. Some systems still struggle with accurately identifying intricate emotions, preserving the context, and guaranteeing ethical AI conduct. Thus, the aim of the proposed system is to enhance the emotion detection accuracy and conversational quality by using a hybrid sentiment analysis and multi-class classification technique.

4.1 Research Gap

Previous mental health chatbot systems have demonstrated significant advancements in delivering automated conversational assistance to emotional support. But there are a number of things that are missing from current approaches. Most of the current systems are rule-based and use a pre-defined set of responses, restricting their capacity to comprehend intricate emotional situations and give individualized interactions.

Many studies are limited to binary sentiment classification, which distinguishes between emotions of positive and negative. These methods are not adequate for detecting a range of psychological conditions like stress, depression, anxiety and loneliness. This constraint compromises emotional assessment accuracy and usefulness.

There is also no efficient multiclass text classification mechanism that can discern between a number of emotional states with greater accuracy. Often, the responses given by the chatbots are generic and do not adjust to the user's emotional state, which decreases interaction and support quality.

Another issue that needs special attention is the lack of attention to ethical issues such as privacy protection, secure user data handling and responsible response generation.

The constraints underscore the importance of having an AI-powered mental health chatbot with the ability to incorporate sentiment analysis and multiclass text classification to properly identify a variety of emotional states, and to provide context-responsive, personalized, and ethical answers.

5. METHODOLOGY/PROPOSED SYSTEM

This virtual companionship has been recently brought to the fore because they are available at all times and non-judgemental. This version has built in appreciation of technology that gives neutral responses, but has incorporated mood detection with cognitive pattern analysis to determine personal experiences. Their arrival is unobtrusive, their presence is un asking, their presence is informing.

The emotional input to the user is via a simple text-only window. The system processes the message, eliminates noise and reorganizes structure after it has been inputted. Here, additional words fall away like snow. Clean up step often makes fuzzy emotional language come into focus. Better input results in better machine analysis of mood states [7],[8].

Emotion patterns are created after the cleaning procedure using sentiment analysis. A model developed from a variety of labeled examples identifies signals linked to anxiety, depression, or stress. Previous work with such sentiment-based tools and diverse data demonstrated strong performance with regards to psychological well-being monitoring [7],[9].

If the system detects emotional clues, responses arrive with empathy, woven into each response. It does not seek to diagnose, but to recognise the feelings in a gentle manner, following the guidance of previous research that has addressed supportive chat tools [3] and [5]. Access to chatbot is free and there are no fees of any kind, in particular no premium tiers hiding under the hood. The bot is quiet, but it works whenever it is required without any need for a response.

5.1 System Architecture and Workflow

Workflow of Chatbot Processing

- Once the user enters text into the input, the chat bot starts processing the input.
- The first step is to detect emotion, in which the system is able to detect the emotional tone of the message.
- The second part of the process is context analysis—analyzing context words and meaning to understand.
- Response planning – post-interpretation, choosing the most appropriate reply according to the detected emotion.
- The language formulation stage is used to convert a planned response into natural language text.
- Finally, output delivery returns the created response to the user.
- Each stage is followed sequentially to ensure continuity of the conversation.
- A feedback loop enables the chatbot to refine its reactions in future interactions based on the outcomes of its interactions.

User Interaction

- The chatbot is used in typed text messages.
- Communication is all verbal, not using verbal gestures.
- The chatbot offers a user-friendly and easily accessible platform for mental health services.

Text Pre-processing

- The pre-processing step is to make everything lowercase.
- This guarantees consistency when processing text.
- The second step is to eliminate the extraneous words in the input (stop words).
- Identification and extraction of delimiters and special symbols.
- These cleaning processes are used to make the text ready for further analysis.
- In this, the accuracy of sentiment analysis and classification tasks is enhanced.

Sentiment Analysis

- The chatbot processes text to classify the sentiment as positive, negative, or neutral.
- This stage identifies the user's emotional tone in their message.
- It additionally looks for emotional connections in the text.
- Sentiment analysis is used to determine the user's state of mind.

Mental Health Classification

- The cleaned text is classified into categories such as:
  - Anxiety
  - Depression
  - Stress
- Classification involves matching the input patterns to trained data labels.
- The system predicts the most relevant mental health category.
- Accuracies of classification will depend upon the quality of the training data.

Response Generation

- The chatbot gives a suitable reply according to the mental health level classification.
- Answers will be empathetic, supportive and context sensitive.
- The system promotes meaningful interaction by the user.

Feedback Loop

- User feedback is taken following response delivery.
- Feedback aids in improving future responses from the chatbot.
- Continuous feedback ensures that the system learns and the conversations are more effective.

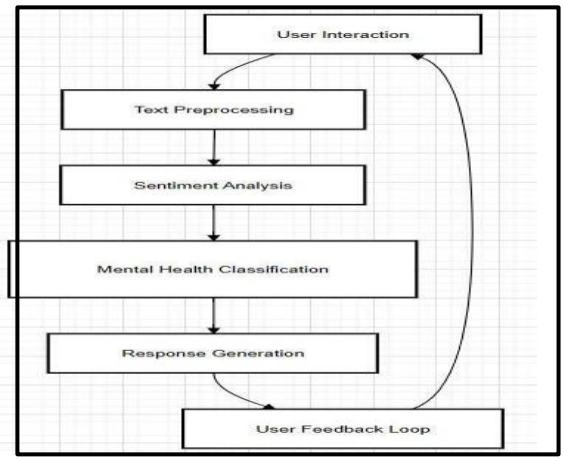

Fig.1. Proposed Framework for AI-Based Mental Health Chatbot

The dataset used in this research was collected from the public data repository available on Kaggle. The dataset file named "Combined Data.csv" contains text-based mental health statements and their corresponding psychological condition labels.

1. Total Records: 53,043
2. Dataset Format: CSV
3. Main Features:

- Statement → textual user statement
- Status → mental health category/class label

The dataset combines multiple mental health-related textual samples for multi-class classification tasks in AI-based mental health analysis systems.

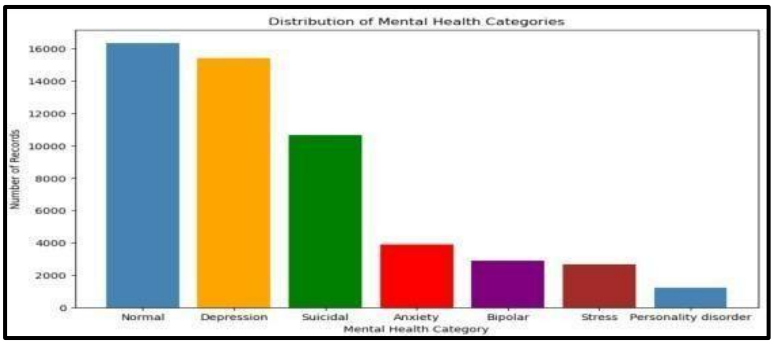

Fig.2. Distribution of Mental Health Categories in the Dataset.

This diagram shows the number of samples of various mental health categories available in the data set on which this experiment is performed. The data set has 7 number of classes namely Normal Depression Suicidal Anxiety Bipolar, Stress and Personality Disorder. From the diagram we can see that the sample numbers in the classes Normal and Depression are higher than other classes while the samples in the class Personality Disorder are low. From this we can see that the data is reasonably imbalanced.

6. DISCUSSION

Looking at the issue from another perspective, it's clear that this chatbot fills the gap in today's research not by doing separate sentiment analyses, but by combining the recognition of the mood with the monitoring of the emotional patterns at a deeper level. The most remarkable thing is that it meets the already identified user requirements: availability at any time, confidentiality, and answers that truly resonate with the user - all inspired by previous works. Unlike conventional methods, where each interaction is considered as an independent event, the new model, through the use of both the affective indicators and the contextual changes, is able to maintain the emotional state along different exchanges.

Previously, systems lacked the ability to understand subtlety; this method compensates for it through inherent features.Joining the datasets leads to fewer misses in emotion recognition, allowing the creation of more accurate models for changing conditions, which is very important for the continuous monitoring of mood by automated means [7][9]. Besides confidentiality, the correct and responsible implementation of AI is very important, with an emphasis on the system being a tool to support the human care rather than replacing it [3][6].

7. CONCLUSION AND FUTURE SCOPE

This project aims to develop an AI-powered mental health chatbot that uses sentiment analysis combined with multi-class classification to discern a user's emotional state and offer a comforting response in a conversation. The paper shows that AI and NLP techniques can be very helpful in recognizing signs of stress, anxiety, and depression through text-based communication.

The chatbot envisioned here acts as an easily reachable, instantly available, and non-judgmental medium of emotional support with an added feature of motivating users to get professional help if needed. Test results reveal that conversation systems, prompted by AI, can not only enhance emotional surveillance but also increase user's willingness to interact.

Introducing highly sophisticated deep learning algorithms, offering multilingual interface, implementing speech communication, incorporating recognition of facial emotions, and allowing users to monitor their mood in real time can be considered as the next steps for development. Seeking the advice of mental health practitioners will help make the artificially generated responses more dependable and suitable for real healthcare settings.

Mental health chatbots based on AI did well in multi-class classification but some limitations remain. For one thing the dataset was got from publicly available online sources that may not be representative of actual clinical scenarios or of the population of people who might use this system. Besides, the dataset was a little imbalanced and so the prediction of uncommon classes like Personality Disorder and Stress may not be very accurate.

The AI based mental health chatbots mainly use text input and so may miss the other ways of perception like tone, facial expressions, or behavior context overall. Because of this, the prediction generated by the model can only be taken as a guide and not a diagnosis.

Another critical ethical concern when creating and implementing AI-powered mental health chatbots is the problem of safety and bias. AI programs might suggest incorrect or harmful advice, leading to a safety risk for the user In particular when the model does not offer different alternatives. To be exact, a mental health chatbot should not become a part of a therapy or medication plan since it cannot replace a professional's judgment. Users must understand that a chatbot supplies only emotional support and cannot work as a professional helping medium. Bias present in the training data can skew the model's predictions leading to unfair treatment toward some groups of people. It is useful to keep the deployed model under constant review. Human intervention and careful watch are essential.

8. REFERENCES

[1] A. A. Abd-alrazaq, M. Alajlani, A. A. Alalwan, B. M. Bewick, P. Gardner, and M. Househ, "An overview of the features of chatbots in mental health: A scoping review," International Journal of Medical Informatics, vol. 132, p. 103978, Dec. 2019, doi: 10.1016/j.ijmedinf.2019.103978.

[2] A. R. et al., "Artificial Intelligence Powered Chatbot for Mental Healthcare Based on Sentiment Analysis," in Proceedings of the 5th International Conference on Advances in Science and Technology (ICAST), Dec. 2022, doi: 10.1109/ICAST58049.2022.10039548.

[3] I. I. Mmbayi, C. Gakii, and F. M. Musyoka, "Sentiment analysis-based model for monitoring user engagement with mental health chatbots," Engineering Reports, vol. 7, no. 6, Jun. 2025, doi: 10.1002/eng2.70247.

[4] A. Basith, "AI-powered mental health chatbot: A survey," PhilPapers, 2025. [Online]. Available: PhilPapers.

[5] L. Balcombe, "AI chatbots in digital mental health," Informatics, vol. 10, no. 4, p. 82, Dec. 2023, doi: 10.3390/informatics10040082.

[6] M. Casu, S. Triscari, S. Battiato, L. Guarnera, and P. Caponnetto, "AI chatbots for mental health: A scoping review of effectiveness, feasibility, and applications," Applied Sciences, vol. 14, no. 13, p. 5889, Jul. 2024, doi: 10.3390/app14135889.

[7] S. Siddals, J. Torous, and A. Coxon, "'It happened to be the perfect thing': Experiences of generative AI chatbots for mental health," npj Mental Health Research, vol. 3, no. 1, pp. 1–9, Oct. 2024, doi: 10.1038/s44184-024-00097-4.

[8] A. Manole, R. Cârciumaru, R. Brînzaş, and F. Manole, "Harnessing AI in anxiety management: A chatbot-based intervention for personalized mental health support," Information, vol. 15, no. 12, p. 768, Dec. 2024, doi: 10.3390/info15120768.

[9] S. Gupta et al., "AI-based and digital mental health apps: Balancing need and risk," IEEE Journal of Biomedical and Health Informatics, 2023, doi: 10.1109/JBHI.2023.10063146.

[10] A. Lal, P. Dhadwal, D. Gade, and J. Joy, "TherapAI: AI-powered therapy chatbot with sentiment analysis, mood tracking, and automated reporting for enhanced mental health support," in Proceedings of the IEEE International Conference, Dec. 2024, pp. 1–8, doi: 10.1109/ICAST.2024.10932101.

[11] M. V. Heinz et al., "Randomized trial of a generative AI chatbot for mental health treatment," NEJM AI, vol. 2, no. 4, Mar. 2025, doi: 10.1056/aioa2400802.

[12] J. De Freitas, A. K. Uğuralp, Z. Oğuz Uğuralp, and S. Puntoni, "Chatbots and mental health: Insights into the safety of generative AI," Journal of Consumer Psychology, vol. 34, no. 3, Oct. 2023, doi: 10.1002/jcpy.1393.

[13] G. Park, J. Chung, and S. Lee, "Effect of AI chatbot emotional disclosure on user satisfaction and reuse intention for mental health counseling: A serial mediation model," Current Psychology, vol. 42, Nov. 2022, doi: 10.1007/s12144-022-03932-z.

[14] W. Pichowicz, M. Kotas, and P. Piotrowski, "Performance of mental health chatbot agents in detecting and managing suicidal ideation," Scientific Reports, vol. 15, no. 1, Aug. 2025, doi: 10.1038/s41598-025-17242-4.

[15] T. G. Heckman, J. C. Markowitz, and B. D. Heckman, "A generative AI chatbot for mental health treatment: A step in the right direction?" NEJM AI, vol. 2, no. 9, Aug. 2025, doi: 10.1056/aip2500453.

[16] I. Gratch and T. Essig, "A letter about 'Randomized trial of a generative AI chatbot for mental health treatment,'" NEJM AI, vol. 2, no. 9, Aug. 2025, doi: 10.1056/aip2500390.

[17] V. Mythily, D. Vinoparkavi, P. Sukumar, C. V., A. M., and B. M., "AI-Driven Chatbot for Mental Health Support Using Sentiment Analysis," in Proceedings of the 1st International Conference on Research and Development in Information, Communication, and Computing Technologies, 2025, pp. 292–297, doi: 10.5220/0013881700004919.

[18] B. M. Chaudhry and H. R. Debi, "User perceptions and experiences of an AI-driven conversational agent for mental health support," mHealth, Jan. 2023, doi: 10.21037/mhealth2355.

[19] Y. Cheng and H. Jiang, "AI-powered mental health chatbots: Examining users' motivations, active communicative action and engagement after mass-shooting disasters," Journal of Contingencies and Crisis Management, vol. 28, no. 3, pp. 339–354, Sep. 2020, doi: 10.1111/1468-5973.12319.

[20] N. Damij and S. Bhattacharya, "The role of AI chatbots in mental health related public services in a (post)pandemic world: A review and future research agenda," IEEE Xplore, Apr. 2022. [Online]. Available: IEEE Xplore.

[21] L. Wang, T. Bhanushali, Z. Huang, J. Yang, S. Badami, and L. Hightow-Weidman, "Evaluating Generative AI in Mental Health: Systematic Review of Capabilities and Limitations," JMIR Mental Health, vol. 12, p. e70014, May 2025, doi: 10.2196/70014.

[22] M. L. Joshi and N. Kanoongo, "Emotional Artificial Intelligence and Machine Learning Techniques for Depression Detection and Analysis," Materials Today: Proceedings, vol. 58, pp. 217–226, 2022.

[23] S. Swapna, "AI-Powered Sentiment Analysis in Real-Time Brand Monitoring," International Journal of Integrated Research and Practice (IJIRP), vol. 1, no. 7, pp. 67–74, Oct. 2025.