ISSN: 2349-6002

Voice-Based Depression & Anxiety Detection using Quantum NLP

Mrs. Veerendeswari J[1], Sangeetha P[2], Snega R[3], Ilakiya M[4], Deepika S[5]

[1]*Head of the Department, Department of Information Technology, Rajiv Gandhi College of Engineering & Technology Puducherry, India*

[2,3,4,5]*UG, Department of Information Technology, Rajiv Gandhi College of Engineering & Technology Puducherry, India*

doi.org/10.64643/IJIRTV12I11-201065-459

***Abstract*—Mental health disorders such as depression and anxiety have become increasingly prevalent in modern society, affecting individuals across different age groups and backgrounds. These conditions have a significant impact on emotional wellbeing, productivity, and overall quality of life. Despite their seriousness, early detection remains a major challenge due to factors such as social stigma, lack of awareness, and limited access to professional healthcare services. As a result, many individuals remain undiagnosed until the condition becomes severe. Traditional diagnostic approaches primarily rely on self-reporting, psychological assessments, and clinical interviews. While these methods are effective, they are often subjective and depend heavily on the individual's willingness to share personal experiences. In many cases, symptoms may be overlooked or misinterpreted, leading to delayed diagnosis and treatment. This highlights the need for automated, objective, and non-invasive methods for early detection of mental health conditions. In this project, a voice-based depression and anxiety detection system is proposed, which utilizes speech signals as the primary source of input. Human voice carries rich emotional and psychological information, including variations in tone, pitch, speech rate, and energy. By analyzing these acoustic features, it is possible to identify patterns associated with emotional states. Additionally, speech-to-text conversion enables the extraction of linguistic features, providing deeper insight into the semantic content of speech. The proposed system integrates Quantum Natural Language Processing (QNLP) to enhance the analysis of textual data. QNLP leverages principles of quantum mechanics, such as superposition and entanglement, to represent and process language in high dimensional spaces. This approach allows for better modeling of contextual relationships, ambiguity, and complex semantic structures compared to traditional Natural Language Processing techniques. Furthermore, the system employs a hybrid approach that combines classical machine learning models with quantum inspired representations. This integration improves the system's ability to capture both acoustic and linguistic patterns effectively. As a result, the model can achieve higher accuracy and better contextual understanding when identifying signs of depression and anxiety from speech data. Experimental results demonstrate that the proposed system performs effectively in detecting emotional states, showing improved accuracy and reliability compared to conventional methods. The findings suggest that voice-based analysis, combined with Quantum NLP, has strong potential for real-world applications in mental health monitoring. This approach can contribute to the development of scalable, accessible, and efficient tools for early diagnosis and intervention.**

***Index Terms*—Depression Detection, Anxiety Detection, Voice Analysis, Quantum NLP, Machine Learning**

I. INTRODUCTION

Mental health disorders such as depression and anxiety have become significant global health concerns, affecting millions of individuals worldwide. These conditions not only impact emotional well-being but also influence physical health, social interactions, and overall quality of life. Despite their severity, many cases remain undiagnosed due to social stigma, lack of awareness, and limited access to mental healthcare services. Early detection is crucial, as it enables timely intervention and reduces the risk of long-term psychological complications. Conventional methods for diagnosing mental health conditions primarily rely on clinical interviews, self-report questionnaires, and psychological assessments conducted by trained professionals. While these approaches are widely accepted, they are often subjective and depend on the

 ISSN: 2349-6002

individual's willingness to share personal experiences. In many situations, individuals may hesitate to seek help or may not recognize their own symptoms, leading to delayed diagnosis and treatment. With the rapid advancement of artificial intelligence and machine learning technologies, there has been growing interest in developing automated systems for mental health assessment. These systems aim to provide objective, scalable, and efficient solutions that can assist healthcare professionals in early diagnosis. By leveraging computational techniques, it is possible to analyze large amounts of data and identify patterns that may not be easily detectable through traditional methods. Human speech is a rich and informative source of emotional and psychological cues. Variations in tone, pitch, speech rate, rhythm, and intensity can reflect underlying mental states. By analyzing these acoustic features, researchers can detect subtle changes in speech that may indicate depression or anxiety. In addition to acoustic analysis, linguistic features derived from speech content provide valuable insights into an individual's thoughts and emotional expressions.

In this project, a voice-based detection system is proposed that integrates both acoustic and linguistic analysis to identify mental health conditions. The system utilizes speech processing techniques to extract meaningful features from voice signals and applies Natural Language Processing (NLP) to analyze textual data obtained through speech-to-text conversion. This combined approach enhances the system's ability to capture both surface-level and deep contextual information. Furthermore, the proposed system incorporates Quantum Natural Language Processing (QNLP) to improve contextual understanding and semantic representation. Unlike traditional NLP models, QNLP leverages quantum-inspired principles to model complex relationships within language. This enables more accurate interpretation of emotional states and improves the overall performance of the detection system. The integration of voice analysis and QNLP presents a promising approach for developing advanced, reliable, and accessible mental health monitoring solutions.

II. LITERATURE SURVEY

Speech-based emotion recognition has gained significant attention in recent years as a non-invasive method for identifying human emotional states. Early research in this domain primarily focused on classical machine learning techniques such as Support Vector Machines (SVM), k-Nearest Neighbors (k-NN), and Random Forest classifiers. These methods rely heavily on handcrafted acoustic features, including pitch, energy, Mel Frequency Cepstral Coefficients (MFCC), and speech rate. While these approaches have demonstrated reasonable performance, their effectiveness is often limited by feature selection and the complexity of emotional variations in speech.

With the advancement of deep learning technologies, more sophisticated models such as Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks, and Convolutional Neural Networks (CNNs) have been widely adopted for speech and emotion analysis. These models are capable of capturing temporal dependencies and complex patterns in sequential data, leading to improved classification accuracy. Additionally, transformer-based architectures have introduced attention mechanisms that enable better contextual understanding within textual data derived from speech. Despite these advancements, classical and deep learning models still face challenges in handling ambiguity, sarcasm, and deeper semantic meaning in natural language. Emotional expression in speech is often subtle and context-dependent, making it difficult for traditional models to accurately interpret underlying psychological states. This limitation highlights the need for more advanced approaches that can effectively model complex relationships within language.

Quantum Natural Language Processing (QNLP) has emerged as a promising paradigm that combines principles of quantum mechanics with linguistic theory. In QNLP, words and sentences are represented as quantum states in high dimensional Hilbert spaces, allowing for the modeling of superposition and entanglement. This representation enables a more expressive and flexible framework for capturing contextual dependencies, ambiguity, and semantic relationships compared to classical NLP methods. Recent research studies have demonstrated that QNLP-based models can achieve improved performance in various language understanding tasks, including sentiment analysis and text classification. These models show a better ability to handle complex linguistic structures and contextual variations, making them particularly suitable for applications involving emotional and psychological analysis.

The integration of QNLP with speech-based emotion recognition systems presents a novel approach for mental health detection. By combining acoustic feature analysis with advanced quantum-inspired linguistic modeling, researchers aim to develop more accurate and context-aware systems. This approach holds significant potential for improving the detection of depression and anxiety, thereby contributing to the advancement of automated mental health monitoring technologies.

III. PROPOSED METHODOLOGY

The proposed system is designed to detect depression and anxiety using voice signals by integrating speech processing techniques with Quantum Natural Language Processing (QNLP). The methodology is structured into multiple stages, including data collection, preprocessing, feature extraction, model processing, and classification. Each stage plays a critical role in transforming raw voice input into meaningful insights for accurate mental health assessment.

In the initial stage, voice samples are collected from individuals representing different emotional states such as depression, anxiety, and normal conditions. These samples may be obtained from publicly available datasets or recorded in controlled environments. Once collected, the audio data undergoes preprocessing, which includes noise reduction, normalization, and segmentation to improve signal quality and ensure consistency across samples. Following preprocessing, the system performs feature extraction to capture both acoustic and linguistic characteristics of speech. Acoustic features such as pitch, energy, speech rate, and Mel Frequency Cepstral Coefficients (MFCC) are extracted to represent emotional variations in voice. These features provide valuable insights into the speaker's mental state by analyzing vocal patterns and speech dynamics. In addition to acoustic analysis, the speech signals are converted into textual form using automatic speech recognition techniques. The resulting text is processed to extract linguistic features such as word usage, sentence structure, and semantic content. These features are crucial for understanding the contextual meaning and emotional expressions conveyed through speech.

The extracted textual data is then analyzed using a Quantum Natural Language Processing framework. In this stage, sentences are encoded into quantum-inspired representations that capture complex semantic relationships and contextual dependencies. This approach enhances the system's ability to interpret ambiguous and nuanced language, which is often associated with mental health conditions.

Finally, a hybrid classification model is employed to categorize the input data into predefined classes such as depression, anxiety, or normal emotional state. The classifier combines classical machine learning techniques with quantum-inspired models to leverage the strengths of both approaches. This integrated methodology improves classification accuracy and ensures reliable detection of mental health conditions, making the system suitable for real-world applications.

IV. DATA COLLECTION

The data collection phase is a fundamental step in the proposed system, as the quality and diversity of the dataset directly influence the performance of the model. Voice samples are collected from individuals representing different emotional states, including depression, anxiety, and normal conditions. These samples may be obtained from publicly available datasets or recorded in controlled environments to ensure consistency and clarity.

During the data acquisition process, participants are asked to speak on specific topics or respond to guided prompts. This approach helps in capturing natural speech patterns and emotional variations. The recordings are performed using standard audio devices such as microphones or mobile phones, ensuring that the collected data reflects real-world scenarios. Care is taken to minimize background noise and maintain consistent recording conditions.

Each voice sample is carefully labeled based on the emotional state of the speaker. Labeling can be performed using self-assessment methods, clinical evaluations, or standardized psychological questionnaires. Accurate labeling is essential to train the model effectively, as incorrect or inconsistent labels can significantly impact classification performance. To improve the robustness of the system, the dataset is designed to include diversity in terms of

ISSN: 2349-6002

age, gender, language, and speaking style. This diversity ensures that the model can generalize well across different user groups and does not become biased toward a specific demographic. A balanced dataset is also maintained to prevent overfitting and ensure equal representation of all emotional classes.
Ethical considerations play an important role in the data collection process. Participants' consent is obtained before recording, and their privacy and confidentiality are strictly maintained. Sensitive personal information is anonymized to protect user identity. These measures ensure that the data collection process adheres to ethical standards and promotes responsible use of technology.
Overall, the collected dataset serves as the foundation for subsequent stages such as preprocessing, feature extraction, and classification. A well-structured and accurately labeled dataset significantly enhances the reliability and effectiveness of the proposed voice-based mental health detection system.

V. FEATURE EXTRACTION

Feature extraction is a critical stage in the proposed system, as it converts raw voice signals into meaningful representations that can be used for analysis and classification. The process begins after preprocessing, where noise is reduced and the audio signal is normalized. The cleaned signal is then analyzed to extract various acoustic features that reflect the emotional and psychological state of the speaker. One of the primary acoustic features considered is pitch, which represents the frequency of the voice and varies with emotional expression. Energy is another important feature that indicates the intensity or loudness of speech. Additionally, speech rate is analyzed to understand the speed and rhythm of speaking, which can provide insights into mental conditions such as anxiety or depression. These features collectively help in capturing the dynamic characteristics of human speech. Mel Frequency Cepstral Coefficients (MFCC) are widely used in speech processing due to their ability to effectively represent the spectral properties of audio signals. MFCCs mimic the human auditory system by focusing on perceptually relevant frequency components. They provide a compact representation of the speech signal and are particularly useful in identifying subtle variations in voice that may correspond to emotional changes.
In addition to acoustic features, linguistic features are extracted by converting speech into text using automatic speech recognition techniques. This conversion enables the system to analyze the content of speech, including word choice, sentence structure, and semantic meaning. Linguistic analysis provides deeper insights into the speaker's thoughts and emotional expressions, complementing the acoustic analysis.
The extracted textual data is further processed using Quantum Natural Language Processing (QNLP), which enhances the understanding of contextual relationships within language. QNLP allows the system to capture complex semantic patterns and handle ambiguity more effectively than traditional NLP methods. This improves the detection of subtle emotional cues present in spoken language.
By combining both acoustic and linguistic features, the system achieves a comprehensive representation of speech signals. This integrated approach significantly improves the accuracy and reliability of the classification process, making it more effective in detecting depression and anxiety from voice inputs.

VI. QUANTUM NLP MODEL

The Quantum Natural Language Processing (QNLP) model plays a crucial role in enhancing the semantic analysis of textual data derived from speech. In this approach, text obtained from speech-to-text conversion is processed using quantum-inspired techniques that go beyond the capabilities of classical Natural Language Processing methods. QNLP provides a more expressive framework for representing and understanding language, especially in emotionally rich and context-dependent scenarios.
In QNLP, words are represented as vectors in high dimensional Hilbert spaces, which are mathematical structures commonly used in quantum mechanics. These representations allow words to exist in superposition states, enabling the model to capture multiple meanings simultaneously. This is particularly useful in natural language, where words often carry different meanings depending on context. Such representations help in modeling ambiguity more effectively compared to traditional NLP techniques.
Furthermore, relationships between words and phrases are modeled using concepts such as entanglement,

which allows the system to capture dependencies between different parts of a sentence. This enables a more holistic understanding of sentence structure and meaning. Instead of processing text sequentially, the QNLP model considers the entire sentence as a unified entity, preserving compositional relationships between words.

The textual data is encoded into quantum circuits, where linguistic structures are mapped into quantum operations. These circuits process the encoded data to extract meaningful semantic information. Although current implementations are largely quantum-inspired and run on classical hardware, they still provide significant improvements in contextual representation and interpretation. One of the key advantages of using QNLP is its ability to capture subtle emotional cues embedded in language. Expressions related to depression and anxiety are often complex and context-dependent, making them difficult to detect using traditional models. QNLP enhances the system’s ability to identify such patterns by providing a richer and more flexible representation of language.

Overall, the integration of the Quantum NLP model significantly improves the system’s capability to analyze linguistic data. By capturing deeper semantic relationships and contextual dependencies, the model contributes to more accurate detection of mental health conditions, thereby strengthening the overall performance of the proposed system.

VII. CLASSIFICATION

The classification stage is a critical component of the proposed system, responsible for determining the emotional state of the individual based on the extracted features. In this stage, both acoustic and linguistic features obtained from previous steps are combined to form a comprehensive feature set. These features serve as input to the classification model, enabling it to learn patterns associated with different mental health conditions such as depression and anxiety.

A hybrid classification approach is adopted, which integrates classical machine learning algorithms with quantum inspired models. Traditional classifiers such as Support Vector Machines (SVM), Random Forest, and Artificial Neural Networks are used to process acoustic features effectively. At the same time, quantum-inspired representations generated from the Quantum NLP module are utilized to enhance the analysis of linguistic data, allowing the system to capture deeper contextual relationships.

The use of a hybrid model provides several advantages, as it combines the strengths of both classical and quantum approaches. Classical models are efficient in handling structured numerical data, while quantum-inspired models excel in representing complex semantic relationships. This combination results in improved classification accuracy and better handling of subtle emotional variations in speech.

The system classifies each input sample into one of the predefined categories: depression, anxiety, or normal emotional state. The classification process is carried out by training the model on labeled data and then testing it on unseen samples to evaluate its performance. Proper training ensures that the model can generalize well across different speakers and speech conditions. To assess the effectiveness of the classifier, standard evaluation metrics such as accuracy, precision, recall, and F1-score are used. These metrics provide a comprehensive measure of the model’s performance, ensuring that it not only achieves high accuracy but also maintains a balance between false positives and false negatives.

Overall, the hybrid classification approach enhances the reliability and robustness of the system. By leveraging both acoustic and linguistic information along with quantum-inspired modeling, the system achieves more accurate and consistent detection of depression and anxiety, making it suitable for real-world mental health monitoring applications.

VIII. SYSTEM ARCHITECTURE

The system architecture of the proposed model is designed to provide an efficient and structured workflow for detecting depression and anxiety using voice signals. It consists of several interconnected modules, including voice input acquisition, preprocessing, feature extraction, Quantum NLP processing, and classification. Each module performs a specific function, ensuring that the raw audio data is systematically transformed into meaningful predictions. The process begins with the voice input module, where speech signals are captured using devices such as microphones, smartphones, or recording systems. These inputs may be collected in real-time or from pre-recorded datasets. The captured

 ISSN: 2349-6002

audio data serves as the primary input for the system and reflects the emotional and psychological state of the speaker.

Once the audio is collected, it is passed to the preprocessing module. This stage involves noise reduction, signal normalization, and segmentation to improve the quality of the audio signal. Preprocessing ensures that irrelevant disturbances such as background noise or recording inconsistencies do not affect the accuracy of feature extraction and subsequent analysis.

After preprocessing, the feature extraction module processes the cleaned audio signal to extract both acoustic and linguistic features. Acoustic features such as pitch, energy, speech rate, and MFCC are derived directly from the audio signal. Simultaneously, speech-to-text conversion is performed to obtain textual data, which is used for linguistic analysis. This dual feature extraction approach provides a comprehensive representation of the input. The extracted textual data is then processed using the Quantum Natural Language Processing (QNLP) module. In this stage, sentences are encoded into quantum-inspired representations to capture complex semantic relationships and contextual dependencies. This module enhances the system's ability to interpret emotional cues embedded in language.

Finally, the processed features are passed to the classification module, where a hybrid model predicts the emotional state of the individual as depression, anxiety, or normal. The architecture ensures a smooth and efficient flow of data across all stages, resulting in accurate and reliable predictions. This modular design also allows for scalability and easy integration into real-world applications such as mobile health systems and intelligent assistants.

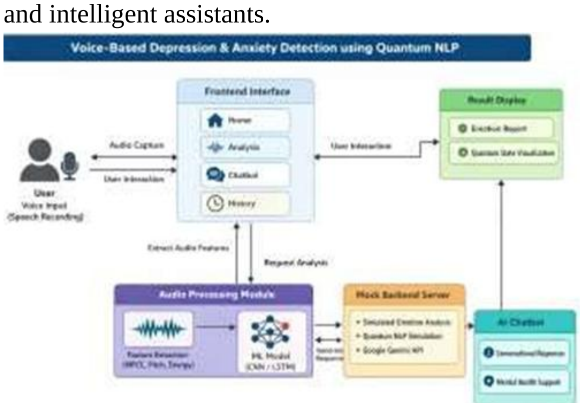

Fig. 1. Proposed System Architecture for Voice-Based Depression and Anxiety Detection

IX. RESULTS AND DISCUSSION

The performance of the proposed voice-based depression and anxiety detection system was evaluated using standard metrics such as accuracy, precision, recall, and F1-score. These metrics provide a comprehensive assessment of the model's ability to correctly classify emotional states while maintaining a balance between false positives and false negatives. The evaluation was carried out on a labeled dataset containing speech samples representing different psychological conditions.

The experimental results indicate that the system achieves a high level of accuracy in identifying depression and anxiety from voice inputs. The use of both acoustic and linguistic features contributed significantly to the overall performance. Acoustic features captured variations in tone, pitch, and speech dynamics, while linguistic features provided insights into the semantic content of speech. This combination enabled the system to analyze emotional states more effectively. A comparative analysis was conducted between the proposed Quantum Natural Language Processing (QNLP)-based model and traditional NLP approaches. The results show that the QNLP model demonstrates superior contextual understanding, particularly in handling complex and ambiguous language patterns. This improvement is attributed to the ability of quantum-inspired representations to model deeper semantic relationships and contextual dependencies.

Furthermore, the hybrid classification approach played a crucial role in enhancing system performance. By integrating classical machine learning techniques with quantum-inspired models, the system was able to leverage the strengths of both approaches. This resulted in improved robustness and consistency across different speech samples and emotional categories. The results also highlight the importance of feature integration in achieving accurate predictions. Systems that relied solely on acoustic or linguistic features showed comparatively lower performance, whereas the combined approach provided better classification results. This confirms that both types of features are essential for capturing the full spectrum of emotional information present in speech.

Overall, the findings suggest that voice-based detection systems are a promising solution for early identification of mental health conditions. The incorporation of Quantum NLP further enhances performance, making the system more effective in real-world scenarios. These results demonstrate the potential of the proposed approach for developing scalable and reliable mental health monitoring applications.

Overall, the proposed system offers a reliable, efficient, and user-friendly solution for mental health monitoring. Its non-invasive nature, scalability, and improved accuracy make it a promising tool for real-world applications in healthcare, research, and personal well-being management.

X. ADVANTAGES

The proposed voice-based depression and anxiety detection system offers several advantages compared to traditional mental health assessment methods. One of the primary benefits is that it is non-invasive, as it relies solely on voice input without requiring physical examinations or intrusive procedures. This makes the system comfortable and easy to use for individuals, encouraging regular monitoring without causing any inconvenience. Another significant advantage is the capability for early detection of mental health conditions. By continuously analyzing speech patterns, the system can identify subtle changes in emotional state at an early stage. This allows timely intervention and reduces the risk of severe psychological complications. Early detection also supports preventive healthcare by enabling individuals to seek help before conditions worsen.

The system is highly scalable and can be integrated into various platforms such as mobile applications, web-based systems, and wearable devices. This scalability ensures that the technology can be deployed in real-world environments, reaching a large number of users across different regions. It also makes the system suitable for remote monitoring, particularly in areas with limited access to mental health professionals. The integration of Quantum Natural Language Processing (QNLP) further enhances the accuracy and effectiveness of the system. QNLP provides improved contextual understanding and semantic analysis, allowing the model to capture complex emotional cues in speech. This results in better classification performance compared to traditional NLP-based systems. In addition, the system combines both acoustic and linguistic features, providing a comprehensive analysis of speech. This dual-feature approach improves reliability and reduces the chances of misclassification. It ensures that both vocal characteristics and spoken content are considered in the detection process.

XI. LIMITATIONS

Despite the promising performance of the proposed system, there are several limitations that need to be considered. One of the primary challenges is the dependency on high-quality voice data. Variations in recording environments, background noise, and device quality can significantly affect the accuracy of feature extraction. Poor audio quality may lead to incorrect interpretation of acoustic features, ultimately impacting the classification results.

Another limitation is the requirement for large and well labeled datasets. Machine learning and quantum-inspired models rely heavily on training data to achieve high accuracy. However, obtaining large-scale, balanced, and accurately labeled datasets for mental health detection is difficult. Limited data availability can lead to overfitting and reduced generalization capability of the model.

The system may also face challenges in handling variations in language, accent, and speaking style. Since speech patterns differ across individuals and regions, the model may not perform consistently across diverse populations without extensive training. This highlights the need for more diverse datasets to improve adaptability and fairness. In addition, Quantum Natural Language Processing (QNLP) is still an emerging field, and most implementations are based on simulations rather than real quantum hardware. This limits the practical deployment and scalability of the system in real-time applications. As quantum computing technology continues to evolve, these limitations may be reduced in the future.

Another concern is the computational complexity involved in processing both acoustic and linguistic features along with quantum-inspired models. This may increase processing time and resource requirements, making it challenging to deploy the system on low-power devices without optimization.

Overall, while the system demonstrates strong potential, these limitations indicate the need for further research and development. Improvements in data quality, model optimization, and advancements in quantum computing are essential to enhance the reliability and scalability of the proposed approach.

XII. APPLI CATIONS

The proposed voice-based depression and anxiety detection system has a wide range of applications, particularly in the healthcare sector. One of its primary uses is in early diagnosis and continuous monitoring of mental health conditions. Healthcare professionals can utilize this system as a supportive tool to identify early signs of depression and anxiety, enabling timely intervention and improving patient outcomes.

The system can also be integrated into telemedicine platforms, allowing remote mental health assessment for individuals who may not have easy access to healthcare services. This is especially beneficial in rural or underserved areas, where access to trained professionals is limited. By providing an automated and accessible solution, the system can help bridge the gap between patients and healthcare providers.

Another important application is in mobile health (mHealth) applications. The system can be embedded into smartphone apps to enable users to monitor their mental health regularly through voice inputs. Such applications can provide feedback, alerts, or recommendations based on detected emotional patterns, encouraging users to seek help when necessary.

In addition, the system can be integrated with wearable devices and smart assistants, enabling continuous and real-time monitoring of emotional states. Devices such as smartwatches and voice-enabled assistants can collect speech data during daily interactions, making the monitoring process seamless and unobtrusive. This integration enhances user convenience and promotes proactive mental health management.

The technology can also be applied in educational institutions and workplaces to monitor stress levels and emotional well-being. Organizations can use this system to identify individuals who may require support, thereby promoting a healthier and more productive environment.

Furthermore, the system can support research in psychology and behavioral science by providing valuable insights into speech patterns and emotional states. Researchers can use the data to study mental health trends and develop more effective intervention strategies. Overall, the proposed system has the potential to make a significant impact across multiple domains by improving mental health awareness and accessibility.

XIII. CONCLUSION

This project presents a voice-based depression and anxiety detection system that integrates speech processing techniques with Quantum Natural Language Processing (QNLP). The system focuses on analyzing both acoustic and linguistic features of speech to identify emotional and psychological patterns. By utilizing voice as the primary input, the approach offers a non-invasive and accessible method for mental health assessment.

The proposed methodology combines feature extraction techniques with advanced quantum-inspired models to improve contextual understanding. Acoustic features such as pitch, energy, and MFCC capture vocal characteristics, while linguistic analysis provides insight into the semantic content of speech. The integration of these features enhances the system's ability to accurately detect signs of depression and anxiety.

The implementation of QNLP plays a significant role in improving the performance of the system. By representing language in high-dimensional quantum states, the model is able to capture complex semantic relationships and contextual dependencies. This results in better interpretation of subtle emotional cues compared to traditional NLP approaches.

Experimental results demonstrate that the system achieves promising accuracy and reliability in classifying emotional states. The hybrid approach, which combines classical machine learning with quantum-inspired techniques, contributes to improved robustness and consistency across different types of speech data.

The proposed system shows strong potential for real-world deployment in applications such as healthcare, telemedicine, and mobile health monitoring. Its non-invasive nature and scalability make it suitable for continuous and remote mental health assessment, helping to bridge the gap between individuals and healthcare services.

Overall, this project highlights the effectiveness of combining voice analysis with Quantum NLP for mental health detection. With further improvements in data quality, model optimization, and quantum computing advancements, the system can evolve into a powerful tool for early diagnosis and mental health support.

XIV. FUTURE SCOPE

The proposed system can be further enhanced by incorporating real-time monitoring capabilities, enabling continuous analysis of voice inputs during daily interactions. This would allow the system to detect changes in emotional states dynamically and provide timely feedback or alerts. Real-time processing can significantly improve the effectiveness of early detection and intervention for mental health conditions. Another important direction for future work is the deployment of the system on mobile platforms. Developing user-friendly mobile applications would make the technology easily accessible to a larger population. Users could regularly monitor their mental health using simple voice inputs, making the system practical for everyday use. Integration with cloud-based services can also support data storage and advanced processing.

The system can be further improved by integrating it with wearable devices such as smartwatches and voice-enabled assistants. These devices can continuously collect speech data during normal conversations, enabling seamless and nonintrusive monitoring. This integration enhances usability and ensures that mental health assessment becomes a part of routine activities without requiring additional effort from users. Advancements in Quantum Natural Language Processing (QNLP) present another promising area for future research. As quantum computing technology continues to evolve, implementing true quantum models instead of quantum-inspired simulations could significantly improve processing speed and accuracy. This would allow the system to handle more complex linguistic structures and larger datasets more efficiently.

In addition, expanding the system to support multiple languages and diverse accents would improve its applicability across different regions and populations. Incorporating larger and more diverse datasets can enhance the model's generalization capability and reduce bias. This is particularly important for developing a globally applicable mental health monitoring system.

Overall, future developments aim to make the system more efficient, scalable, and user-friendly. By combining advancements in artificial intelligence, speech processing, and quantum computing, the proposed system can evolve into a comprehensive solution for real-time mental health monitoring and support.

REFERENCES

[1] B. Schuller et al., “Recognising Realistic Emotions and Affect in Speech: State of the Art and Lessons Learnt,” Speech Communication, vol. 53, no. 9, pp. 1062–1087, 2011.

[2] S. R. Livingstone and F. A. Russo, “The Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS),” PLOS ONE, vol. 13, no. 5, 2018.

[3] Z. Huang, M. Dong, Q. Mao, and Y. Zhan, “Speech Emotion Recognition Using CNN,” in Proc. ACM Multimedia Conf., 2014.

[4] A. Graves, A. Mohamed, and G. Hinton, “Speech Recognition with Deep Recurrent Neural Networks,” in Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP), 2013, pp. 6645–6649.

[5] J. Devlin et al., “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,” in Proc. North American Chapter of the Association for Computational Linguistics (NAACL), 2019, pp. 4171–4186.

[6] B. Coecke, M. Sadrzadeh, and S. Clark, “Mathematical Foundations for a Compositional Distributional Model of Meaning,” Linguistic Analysis, vol. 36, no. 1–4, pp. 345–384, 2010.

[7] B. Coecke et al., “Quantum Natural Language Processing,” arXiv preprint arXiv:2001.03622, 2020.

[8] World Health Organization, Depression and Other Common Mental Disorders: Global Health Estimates, Geneva, Switzerland, 2017.

[9] T. Baltrusaitis, C. Ahuja, and L.-P. Morency, “Multimodal Machine Learning: A Survey,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 41, no. 2, pp. 423–

ISSN: 2349-6002

443, 2019.
[10] S. Cummins et al., “A Review of Depression and Suicide Risk Assessment Using Speech Analysis,” Speech Communication, vol. 71, pp. 10–49, 2015.