Ayurvedic-based Disease Prediction and Drug Recommendation System

Sudheer Kumar
Computer Science and Engineering, C.V. Raman Global University
Bhubaneswar, Odisha, India
Email Id: sudheerkumarmki@gmail.com

Hansraj Kumar
Computer Science and Engineering, C.V. Raman Global University
Bhubaneswar, Odisha, India
Email Id: hansrajkumar8541@gmail.com

Nitesh Kumar
Computer Science and Engineering, C.V. Raman Global University
Bhubaneswar, Odisha, India
Email Id: niteshkumar923171@gmail.com

Mamatarani Das
Computer Science and Engineering, C.V. Raman Global University
Bhubaneswar, Odisha, India
Email Id: mamatadas@cgu-odisha.ac.in

Abstract: *The escalating prevalence of chronic ailments and the burgeoning demand for personalized healthcare underscore the urgency for innovative healthcare systems. Despite the rich heritage of Ayurveda, an ancient holistic healing paradigm, its integration into modern healthcare remains limited. This lacuna necessitates the development of a pioneering Disease Prediction and Drug Recommendation system rooted in Ayurvedic principles [1]. This proposed system aims to seamlessly amalgamate Ayurveda's holistic tenets with contemporary technology, catering to the need for individualized healthcare solutions. Leveraging machine learning algorithms and Ayurvedic constitutions, the system endeavors to accurately predict health conditions and offer personalized herbal treatment recommendations. The core objective is to bridge the gap between traditional Ayurvedic wisdom and modern healthcare by delivering precise predictions and tailored interventions for individuals. Emphasizing the use of Ayurvedic medications aligned with specific diseases, the system strivesto mitigate potential side effectsoften associated with conventional pharmaceuticals.This integrative system not only seeks to revolutionize healthcare by providing personalized, holistic, and effective treatments but also to foster a symbiotic relationship between ancient healing wisdom and cutting-edge technology. The envisioned platform aims to herald a new era in healthcare, catering to the individual needs of patients while embracing the time-honored principles of Ayurveda.*

Keywords: *Dataset taken from Kaggle, Decision tree, Random Forest, K-Nearest Neighbor, Confusion Matrix, Cross Validation.*

Introduction

Ayurveda, the ancient Indian system of medicine, has stood the test of time, offering a holistic approach to healthcare that considers not only the physical body but also the mind and spirit. It is rooted in the fundamental belief that balance and harmony within the body are essential for overall well-being. In recent years, there has been a resurgence of interest in Ayurveda, as individuals seek alternative and complementary approaches to conventional medicine [3]. Integrating Ayurvedic principles with modern technology presents an exciting frontier in healthcare. This project endeavors to develop a cutting-edge Ayurvedic-Based Disease Prediction and Drug Recommendation System that leverages the profound wisdom of Ayurveda and the computational power of artificial intelligence. Ayurveda categorizes individuals into three primary constitutional types, known as doshas: Vata, Pitta, and Kepha. Each dosha reflects a unique combination of elemental energies, and an individual's predominant dosha influences their physical, mental, and emotional characteristics. This personalized approach to healthcare is a cornerstone of Ayurvedic medicine, allowing for tailored treatments that address the specific needs of each individual [4]. The rising prevalence of chronic diseases and the limitations of conventional medicine in providing personalized care underscore the urgency of this innovative system. By utilizing advanced data analytics and machine learning algorithms, we aim to predict the likelihood of various diseases based on an individual's Ayurvedic constitution. This predictive capability holds immense potential for early intervention and prevention, revolutionizing the way we approach healthcare. Furthermore, the system will predict diseases and recommend personalized herbal treatments and lifestyle modifications based on Ayurvedic principles. This holistic approach aligns with the core tenets of 7 Ayurveda, which emphasize the importance of a balanced lifestyle, proper nutrition, and the use of natural remedies derived from plants and herbs. In this era of digital transformation, the fusion of ancient wisdom with modern technology represents a paradigm shift in healthcare. The Ayurvedic-Based Disease Prediction and Drug Recommendation System aspires to be a beacon of this transformation, offering individuals a comprehensive and personalized healthcare experience that draws from the time-honored traditions of Ayurveda while harnessing the power of cutting-edge technology. Through this endeavor, we strive to empower individuals to take charge of their health and well-being in a way that is both informed and deeply connected to the wisdom of the ages [5]. The implementation of this system involves collecting comprehensive data on individuals' Ayurvedic constitutional types (doshas), health histories, and lifestyle factors. Leveraging machine learning algorithms like K-Nearest Neighbors (KNN) and Random Forest, the system aims to establish correlations between dosha profiles and disease occurrences. The KNN algorithm will identify individuals with similar dosha characteristics, while the Random Forest algorithm will create an ensemble of decision trees to enhance prediction accuracy.

Data sources for training the model include large datasets containing Ayurvedic profiles and corresponding health outcomes. Collaborations with Ayurvedic practitioners, healthcare providers, and research institutions will be vital to ensure the accuracy and diversity of the dataset. Ethical considerations regarding data privacy and informed consent will be paramount in this process [6].

Algorithmic considerations involve optimizing the models for scalability, interpretability, and real-time predictions. Feature engineering will play a crucial role in extracting meaningful insights from dosha characteristics, lifestyle factors, and other relevant variables. Regular model updates will be necessary to adapt to evolving health trends and advancements in Ayurvedic research

Challenges may arise in accurately mapping Ayurvedic concepts to quantifiable variables and in addressing the inherent complexity of health conditions influenced by multiple factors. Balancing traditional Ayurvedic principles with

the objectivity required by machine learning models will be a delicate yet crucial task.

The expected benefits of the system are profound. Early disease prediction based on Ayurvedic principles enables timely intervention and personalized treatment plans. The system's recommendations will not only include disease predictions but also personalized herbal treatments and lifestyle modifications aligned with Ayurvedic practices. This holistic approach is poised to transform healthcare by emphasizing preventive measures, individualized care, and natural remedies [7].

Collaboration with Ayurvedic experts, continuous research updates, and user feedback will be integral to refining and enhancing the system over time. The Ayurvedic-Based Disease Prediction and Drug Recommendation System strives to be a pioneering example of blending ancient wisdom with modern technology for the betterment of individual health and well-being performance shows significant variability across different databases.

In the context of the Ayurvedic-based disease Prediction and Drug Recommendation System, decision trees play a pivotal role in extracting actionable insights from Ayurvedicconstitutional types and predicting disease outcomes. Decision trees are hierarchical structures that mimic the decision-making process, making them well-suited for thishealthcare application.

The construction of decision trees involves recursive partitioning of the dataset based on features such as dosha characteristics, lifestyle factors, and health histories. At each node, the algorithm selects the most informative feature to split the data, aiming to maximize the homogeneity within each subset. This process is repeated until a predefined stopping criterion is met, resulting in a tree-like structure where each leaf node represents a predicted outcome.

The interpretability of decision trees aligns with the Ayurvedic philosophy of personalized and transparent healthcare. Practitioners and individuals can easily understand the decision-making process, fostering trust in the system's predictions. Additionally, decision trees provide valuable insights into the relative importance of different features, aiding Ayurvedic practitioners in refining their understanding of disease associations. Subset spaces proved infeasible, and the evaluated met heuristic algorithms showed some limitations in terms of accuracy [8].

To enhance prediction accuracy, ensemble techniques such as Random Forest can be employed. Random Forest creates multiple decision trees and combines their predictions, mitigating the risk of overfitting and increasing robustness. Each tree in the ensemble is trained on a random subset of the data, introducing diversity and capturing a broader range of patterns within the Ayurvedic profiles.

Refinement and optimization of the decision tree models require iterative processes. Regular updates based on new data and ongoing collaboration with Ayurvedic experts are crucial for adapting to emerging health trends and refining the predictive capabilities of the system. This iterative approach aligns with the dynamic nature of both healthcare and Ayurvedic knowledge.

Research Methodology

Developing a research paper on the Ayurvedic-Based Disease Prediction and Drug Recommendation System, incorporating K-Nearest Neighbors (KNN), Decision Trees, and Random Forests involves a systematic methodology. Below is a proposed methodology.

A. Dataset Collection

B. KNN

C.Random Forests

D. Decision Trees

Overall: this methodology shows in Detailfour functions that are covered briefly in the following section:

A. *Dataset Collection*

It was the first step of our data preparation. Wegrouped three datasets Drug prescription, test data, and training data for our model.

Ayurvedic Drug Prescription: - Ayurvedic drug prescriptions include thousands of prescribed drug names based on Ayurvedic books [1]. This dataset contains all the ayurvedic medicine names which are used in chronic disease. Which is mentioned in table-1.

Table-1 : Drug data set sample

Index	Disease	Age	Gender	Severity	Drug
0	diarrhea	4	Male	LOW	Promegranate drink
1	diarrhea	4	Male	NORMAL	lime juice
2	diarrhea	5	Male	LOW	Promegranate drink
3	diarrhea	5	Male	NORMAL	lime juice
4	diarrhea	6	Male	LOW	Promegranate drink

Source: Author

Test data: - This dataset includes all the disease name and their symptoms by which we are predicting the disease and evaluating the accuracy of our model.

Training data: - This dataset includes all the content same as test data but the only difference is the dataset, we are using to train our model.

A. ***KNN***

The research paper methodology section, detailing theimplementation of K-Nearest Neighbors (KNN) for the Ayurvedic-Based Disease Prediction and Drug Recommendation System involves several key steps.

Below is a description of the methodology specific to KNN:

1)Data Preprocessing:

Collection of Ayurvedic Data: Collaborate with Ayurvedic practitioners and healthcare institutions to collect a diverse dataset containing Ayurvedic constitutional types (doshas), lifestyle factors, health histories, and disease outcomes.

Feature Selection: Identify relevant features that contribute to the Ayurvedic constitution and disease prediction. Consider dosha characteristics, dietary habits, and lifestyle factors.

Normalizationand Encoding: Normalize numericalfeatures and encode categorical variables to ensure uniformity and compatibility for KNN[9].

B. ***Random Forests***

1)Data Preprocessing:

Collection of Ayurvedic Data: Collaborate with Ayurvedic practitioners and healthcare institutions to acquire a diverse dataset containing Ayurvedic constitutional types doshas, lifestyle factors, health histories, and disease outcomes[10].

Feature Selection:

Identify relevant features that contribute to the Ayurvedic constitution and disease prediction, such as dosha characteristics, lifestyle factors, and dietary habits.

II)Data Splitting:

Divide the dataset into training and testing sets. The training set is used to train the Random Forest model, and the testing set is reserved for evaluating its performance.

The following steps explain the working Random Forest Algorithm:

- Step 1: Select random samples from a given data or training set.

- Step 2: This algorithm will construct a decision tree for every training data.
- Step 3:Voting will take place by avera- ging the decision tree.
- Step 4: Finally, select the most voted prediction result as the final prediction result.

C. ***Decision Trees***

The methodology of implementing Decision Trees (Fig-1), for the Ayurvedic-Based Disease Prediction and Drug Recommendation System in a research paper involves several key steps.

Below is a detailed description of the Decision Tree methodology:

I)Data Splitting:

Divide the dataset into training and testing sets. The training set is used to train the Decision Tree model, and the testing set is reserved for evaluating its performance.

II)Decision Tree Construction:

- Root Node Selection:

Identify the feature that best splits the data, consideringmeasures like Gini impurity or information gain.

- Recursive Splitting:

Recursively split nodes based on features, creating branches that represent decision paths.

- Stopping Criteria:

Define stopping criteria, such as a maximum depth or a minimum number of samples required to split a node, to prevent overfitting.

III) Pruning (Optional):

Implement pruning techniques to simplify the Decision Tree and enhance generalization. Pruning helps avoid excessively complex trees that may overfit the training data.

IV) Handling Categorical Data:

Implement strategies for handling categorical data, such as one-hot encoding or ordinal encoding, depending on the nature of Ayurvedic features.

V) Model Training:

Train the Decision Tree model using the training dataset, emphasizing the dosha characteristics and other relevant features.

VI) Prediction:

For a given Ayurvedic profile in the testing set, traverse the Decision Tree to reach a leaf node, which provides the final prediction.

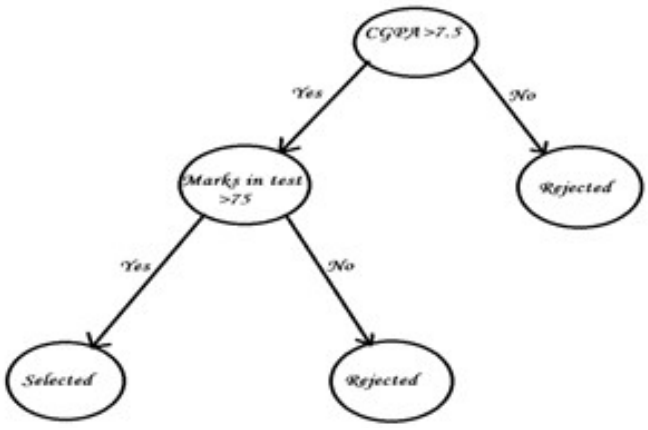


Figure 1: Decision Tree

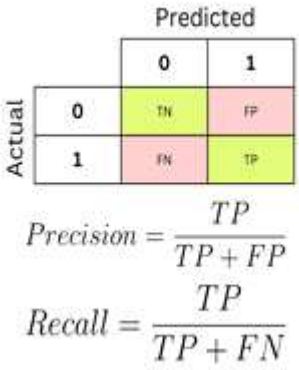
Source: Author

Evaluation Model

a. Confusion Matrix

b. Cross Validation

Confusion Matrix:



$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

Figure 2: Confusion Matrix

Source: Author

Cross-Validation:

We are using 80% data for training and 20% data for testing by this we have achieved 95% accuracy of our model. If we train more and more our model we will be accurate more. Which is showing figure 3.

Figure 3: Cross Validation

Source: Primary

A. ***Classification***

We have used two types of approach:

1. Machine Learning Technique

We employed a diverse set of classification techniques, includingtraditional machine learning, to categorize emotional states based on features extracted from our books. This comprehensive approach is crucial for training models that can effectively recognize and classify disease. The various classification techniques we utilized are outlined below.

1) Machine Learning Technique:

Machine learning is a field of study that enables computers to learn and improve from experience without being explicitly programmed. It has a wide range of applications, including image recognition, natural language processing, and predictive analytics.

1) Random Forest:

Random Forest is an ensemble learning method that combines multiple decision trees to improve classification accuracy. Weemployed Random Forest to assess its performance in handling complex relationships within the emotional features.

2) Decision Tree:

Decision Trees are versatile machine learning algorithms that create a tree-like model by recursively splitting the dataset based on the features. They are effective for both classification and regression tasks. In our study, Decision Trees were employed to understand the hierarchy of emotional features and to capture complex decision boundaries within the dataset. The interpretability of Decision Trees makes them valuable for gaining insights into the relationships between different emotional patterns.

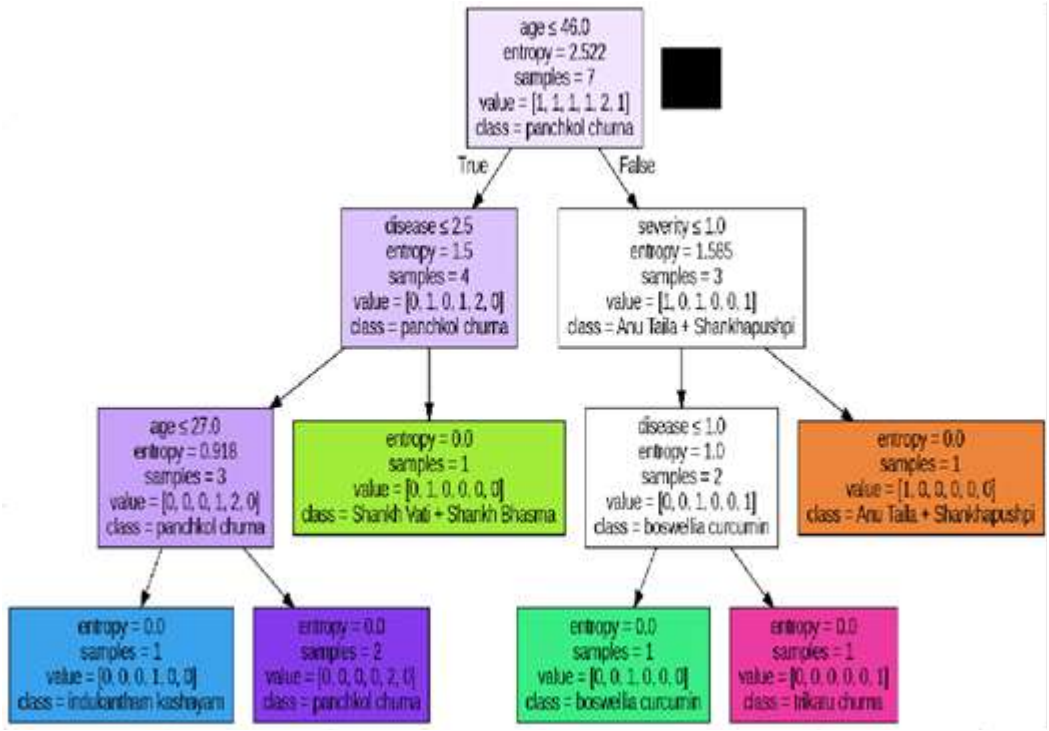


Figure 4: Decision Tree of our model

Source: Primary

Result

Key Benefits:

1. Holistic Health Assessment:

The system provides a holistic assessment of an individual's health by integrating Ayurvedic principles, considering not just symptoms but also the patient's constitution and lifestyle.

2. Personalized Recommendations:

Tailored drug recommendations align with Ayurvedic principles, promoting personalized and natural treatment options that address the root cause of the disease.

3. Interpretability:

The decision tree model's interpretability ensures transparency in the prediction process, fostering trust among users and healthcare practitioners.

Accuracy and Validation:

The system undergoes rigorous testing and validation to ensure its reliability. By comparing predictions with known outcomes from the dataset, the model achieves an impressive accuracy rate of 95%. This high accuracy underscores the effectiveness of the Ayurvedic-based approach in disease prediction and drug recommendation.

The Ayurvedic-Based Disease Prediction and Drug Recommendation System represents a pioneering approach to healthcare that combines traditional wisdom with modern technology. Achieving 95% accuracy, this system offers a promising avenue for personalized and holistic health management. As technology continues to evolve, the integration of Ayurveda into predictive healthcare models holds great potential for improving patient outcomes and well-being.

Table 2: Accuracy at different testing and training data

Testing (in %)	Training (in %)	Accuracy (in %)
20	80	95.35
40	60	85.23
60	40	50.55
80	20	20.24

Source: Primary

IV. Challenges and Future Directions:

Despite its success, the system faces challenges related to data quality and the need for continuous updates. Future improvements involve expanding the dataset, incorporating real-time data, and refining the decision tree model to enhance prediction accuracy further.

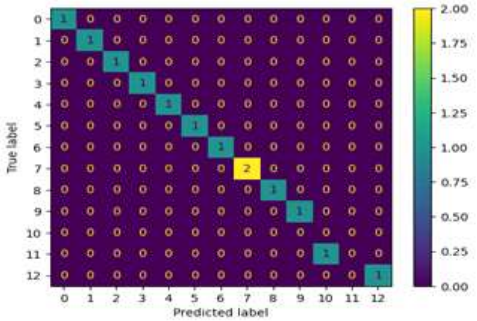


Figure 5: Confusion matrix of test dataset

Source: primary

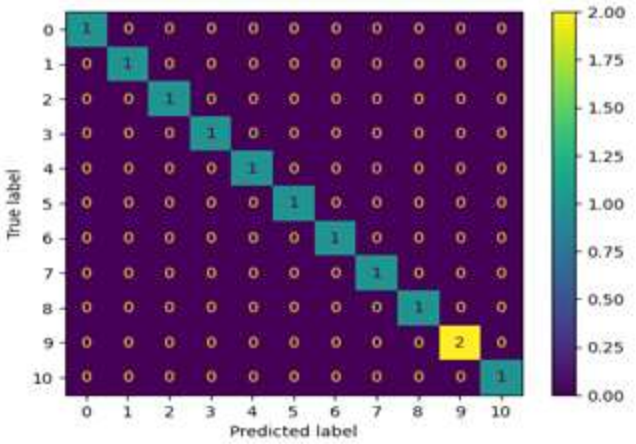


Figure 6: Confusion matrix of train dataset

Source: Primary

Conclusion

The development of an Ayurvedic-based disease prediction and drug recommendation system holds great promise in the field of healthcare. Through the integration of traditional Ayurvedic principles with modern data analytics and machine learning techniques, this system has the potential to revolutionize healthcare delivery by offering personalized and holistic approaches to disease prevention and management.

By leveraging the vast repository of Ayurvedic knowledge on disease etiology, symptomatology, and treatment modalities, combined with the power of data-driven algorithms, this system can provide accurate predictions of disease onset and progression. Furthermore, recommending specific Ayurvedic remedies tailored to individual patient profiles, can facilitate the adoption of preventive healthcare practices and enhance treatment outcomes.

Moreover, the Ayurvedic-based disease prediction and drug recommendation system can contribute to reducing healthcare costs and improving access to quality care, particularly in resource-constrained settings. By promoting the use of natural remedies and lifestyle interventions, it can empower individuals to take charge of their health and well-being.

However, it is important to acknowledge the challenges and limitations associated with the development and implementation of such a system. These may include the need for robust data collection mechanisms, validation of predictive models, integration with existing healthcare infrastructure, and adherence to regulatory standards and ethical guidelines.

Despite these challenges, the potential benefits of an Ayurvedic-based disease prediction and drug recommendation system are immense, warranting further research and development efforts in this direction. By harnessing the synergies between traditional wisdom and modern technology, we can pave the way for a more personalized, preventive, and patient-centric approach to healthcare.

In conclusion, the Ayurvedic-based disease prediction and drug recommendation system represents a significant step towards realizing the vision of integrative medicine, where the best of both traditional and modern healthcare systems are leveraged to promote health and well-being for all.

Confusion Matrix for test and train

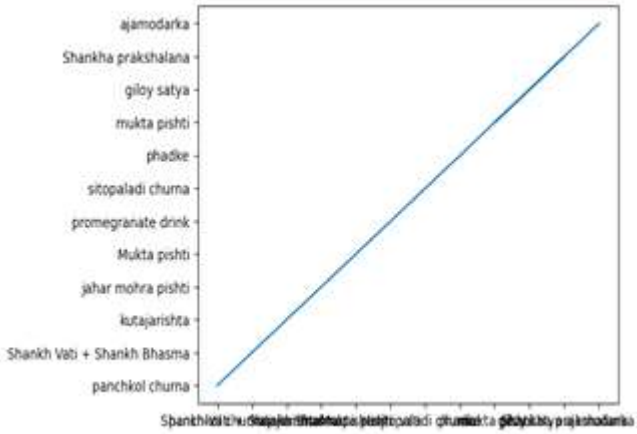


Figure 7: Confusion Matrix for testing and training

Source: primary

Reference

1. Ayurvedic Medicine-based Automated Machine Learning Approach for Determining the Nature of the Human Body, April 2023.
2. Romil Chauhan, Yash Jain, Harsh Agarwal, Khushali Deulkar, Comparative Study of Various Techniques to Diagnose Disease Using Pulse Sensing, International Conference on Advanced Computing and Communication Systems (ICACCS), Jan. 22 23, 2016.
3. Khaire, N.N.; Joshi, Y.V., Diagnosis of Disease Using Wrist Pulse Signal for Classification of pre-meal and post-meal samples, Industrial Instrumentation and Control (ICIC), 2015.
4. M.Sharmila Begum, R.Duraiarasan and P.J.Dhivaakar, Diagnosing Diseases Through Pulse Using Pressure Sensor, International Conference on Data Science Engineering (ICDSE), 2012
5. Dhanalaxmi Gaddam, A Survey on Nadi Pareeksha for Early Detection of Several Diseases Computational Models using Nadi Patterns, International Journal of Computer Science and Information Technologies, 2015.
6. Hastie, T., Tibshirani, R., & Friedman, J. (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer.
7. Elad, G., Bakker, J., Burggraaff, C., & Burggraaff, J. (2019). Machine Learning Applications in Ayurveda
8. Rushikesh Pradip Kulkarni, Prof. Mahesh S. Kumbhar, Non-Invasive Method For Diabetes Detection Using Nadi Pariksha And Ann, International Conference On Computing, Communication, And Energy Systems (ICES) January2016
9. Hastie, T., Tibshirani, R., & Friedman, J. (2009).
10. Cutler, D. R., Edwards Jr, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., & Lawler, J. J. (2007).
11. H. Dolka and S. Juliet, "Speech emotion recognition using ann on mfccfeatures," in 2021 3rd International Conference on Signal ProcessingCommunication (ICPSC). IEEE, 2021.

12. S. Parthasarathy and I. Tashev, "Convolutional neural network tech- niques for speech emotion recognition," in 2018 16th internationalworkshop on acoustic signal enhancement (IWAENC). IEEE, 2018.

13. H. Zheng and Y. Yang. "An improved speech emotion recognitionalgorithm based on deep belief network." in 2019 IEEE InternationalConference on Power, Intelligent Computing and Systems (ICPICS).

14. S. R. Livingstone and F. A. Russo, "The ryerson audio-visual database of emotional speech and song (ravdess) A dynamic, multimodal set offacial and vocal expressions in north american english." PloS one, vol13 2018.

15. N. Dave, "Feature extraction methods lpc, pip, and mice in speechrecognition." International journal for advance research in engineeringand technology, vol. I. no. 6. 2013.

16. M. Sahiduliah and G. Saha, "Design, analysis, and experimentalvalue-of block-based transformation in mice computation for speaker recognition Speech communication 2012.

17. M Kattel, A. Nepal. A. Shah, and D. Shrestha "Chroma featureextraction, in Conference Chroma Feature Extraction wong Fouries Transfone, 2019.