iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 Issue: IV Month of publication: April 2024

DOI: https://doi.org/10.22214/ijraset.2024.59966

Multiple Disease Detection System

Pranay Parbate[1], Dhairyashil Thombare[2], Tushar Yerkal[3], Adarsh Varpe[4], Prof. Suresh Reddy[5]
[1, 2, 3, 4, 5]*Computer Engineering, Savitribai Phule Pune University*

Abstract: In the realm of healthcare, the early detection of multiple diseases presents a formidable challenge but holds immense potential for improving patient outcomes. This paper proposes an integrative approach for the simultaneous detection of four prevalent diseases: heart disease, Parkinson's disease, diabetes, and skin cancer. Leveraging advanced machine learning techniques, our framework encompasses data preprocessing methods for cleaning and normalization, feature selection strategies to extract discriminative features from heterogeneous medical datasets, and ensemble classification models for accurate disease prediction. Real-world datasets encompassing diverse medical conditions are utilized to evaluate the efficacy of the proposed framework. Experimental results demonstrate its superior performance in terms of accuracy, sensitivity, and specificity compared to conventional methods. This framework stands as a testament to the transformative potential of technology in healthcare, facilitating early detection and intervention across multiple diseases, thereby enhancing patient care and quality of life.

Keywords: Machine Learning, Deep Learning, Disease Prediction, App Development, Java, Python, etc.

I. INTRODUCTION

In an era characterized by the omnipresence of technology and an increasing emphasis on personal well-being, the convergence of advanced computational techniques and healthcare has emerged as a pivotal frontier. The contemporary landscape is marked by a surge in demand for intelligent and accurate health solutions amidst a backdrop of pervasive internet usage and digital interconnectedness. However, despite the accessibility of information, individuals often grapple with health concerns, navigating a complex array of symptoms and potential diagnoses.

Against this backdrop, our paper presents an innovative multi-disease prediction system designed to address the diverse healthcare needs of modern society. Rooted in the premise of leveraging technology for proactive health management, our system offers a comprehensive approach to disease prediction and diagnosis across a spectrum of prevalent conditions. Specifically, our framework encompasses the detection and assessment of heart disease, Parkinson's disease, diabetes, and skin cancer – ailments that collectively pose significant challenges to global healthcare systems and individual well-being.

The development of our system is guided by a fundamental principle: to empower individuals with the tools and insights necessary to take charge of their health proactively. Central to our approach is the integration of deep learning methodologies, notably a Convolutional Neural Network (CNN) model, which serves as the cornerstone of our skin cancer diagnosis module. By harnessing the power of artificial intelligence and machine learning, we endeavor to enhance the accuracy, efficiency, and accessibility of healthcare services, transcending the limitations of traditional diagnostic paradigms.

Through a synthesis of interdisciplinary expertise and technological innovation, our paper elucidates the conceptualization, development, and validation of our multi-disease prediction system. By addressing the complexities inherent in disease diagnosis and management, we aspire to contribute to the advancement of healthcare technology and the realization of a future where personalized, data-driven healthcare solutions are accessible to all.

II. LITERATURE SURVEY

1) " Deep Convolution Neural Network based Automatic Multi-Class Classification of Skin Cancer from Dermoscopic Images" (2021) [1] The authors concluded that their proposed DCNN-based approach demonstrated high accuracy in classifying malignant skin cancer types from dermoscopic images. The results were promising and indicated the potential for practical applications in the medical field. The authors suggested that further implementation and validation through clinical trials could establish the approach as a reliable tool for assisting dermatologists in diagnosing skin cancer. Melanoma has been considered the most fatal category of skin cancer. In the past few decades, image processing has shown to be a boon in the biomedical field where such cancerous diseases can be diagnosed well in time. Skin Cancer might not be the deadliest but it's necessary to detect it at its early stage in order to save the lives of many. So finding a technique that gives high accuracy and early detection is very crucial.

Enthusiastic results of image processing in the medical field have convinced trained practitioners to rely on the outputs obtained from computer-vision systems (CVS).

2) " Analysis and Classification of Skin Cancer Images using Convolutional Neural Network Approaches" (2020) [2] The paper concludes that the application of deep learning algorithms, particularly CNNs, can significantly improve the accuracy of skin cancer classification, which is crucial for medical diagnosis. The study acknowledges the computational time required for training the CNN models and suggests that further research could focus on optimizing training efficiency. In this modern era, skin cancer is a serious problem around the globe and so it is the age of technology and it is important to solve this problem through intelligent machines that use different algorithms rather than conventional methods. Intelligent machines use different algorithms to classify skin cancer images in a reliable way to save effort, and time and ease human life. For this purpose, a deep learning (CNN) algorithm is used by intelligent machines to classify skin cancer images according to their types
3) " A Clinical Support System for Prediction of Heart Disease using Machine Learning Techniques" (2020) [3] The authors concluded that the proposed clinical decision support system using machine learning algorithms could enhance the prediction of heart disease risk and help clinicians make more accurate diagnoses. While the accuracy decreased when using cross-validation, they believe it's a robust technique for handling overfitting. The study encourages further research to improve the accuracy of heart disease prediction models and highlights the importance of the choice of dataset, number of attributes, and algorithms used. Heart disease is a leading cause of death worldwide. However, it remains difficult for clinicians to predict heart disease as it is a complex and costly task. Hence, we proposed a clinical support system for predicting heart disease to help clinicians with diagnosis and make better decisions.
4) " Predictive Analytics on Diabetes Data using Machine Learning Techniques" (2021) [4] Precision medicine has gained attention for improving disease treatment and prevention. The research work contributes to the development of diabetes prediction models. Further integration of genetic biomarkers could enhance prediction accuracy and treatment decisions. Diabetes mellitus is an ongoing illness related to anomalous undeniable levels of the sugar glucose in the blood and it is a major public health problem. Diabetes has become the 4th driving reason for death in developed countries. Though several methodologies have been developed to predict this chronic disease, there is a need for innovative approaches that may aid in the early prediction of diabetes and its complications.
5) " A Supervised Machine Learning Approach using Different Feature Selection Techniques on Voice Datasets for Prediction of Parkinson's Disease" (2019) [5] The study demonstrated the effectiveness of supervised machine learning and different feature selection techniques in distinguishing Parkinson's disease using voice data. SVM with GA-based features achieved high accuracy. However, further investigation with larger datasets and exploring other advanced classification algorithms could enhance accuracy and robustness. Future research can explore the integration of multimodal data (voice, gait, etc.) for improved accuracy. Among the neurological diseases, Parkinson's disease is the second most common disease, which affects the old age people over the age of 65 years. It is also mentioned that the number of people affected with Parkinson's disease will increase at a higher rate until 2050, and it will be a rising concern to many developed countries because the cost to the healthcare service of this disease is really high

III. EXISTING SYSTEM

In the current landscape, existing disease prediction systems primarily cater to large-scale diseases such as Heart Disease and Cancer, utilizing relatively small datasets with predefined conditions. However, these systems may fall short of accommodating changes in diseases and their influencing factors over time, potentially leading to inaccuracies in results. Moreover, users often encounter lengthy questionnaires, causing delays in accessing necessary information. As diseases evolve continually, there's a pressing need for more dynamic and efficient systems to keep pace with diagnostic demands.

IV. PROPOSED SYSTEM

Our proposed system aims to address these limitations by offering a streamlined User Interface (UI) and ensuring time efficiency. To achieve this, we plan to implement a specific questionnaire tailored to each user, minimizing the time required for input. Our objective is to serve as a vital link between doctors and patients, leveraging machine learning techniques to deliver accurate predictions. Additionally, our system will feature a Doctor's Consultation functionality, suggesting users seek professional medical advice based on the generated reports. By incorporating this feature, we aim to not only cater to doctors but also gain their trust in the system's reliability and efficacy.

V. SYSTEM ARCHITECTURE

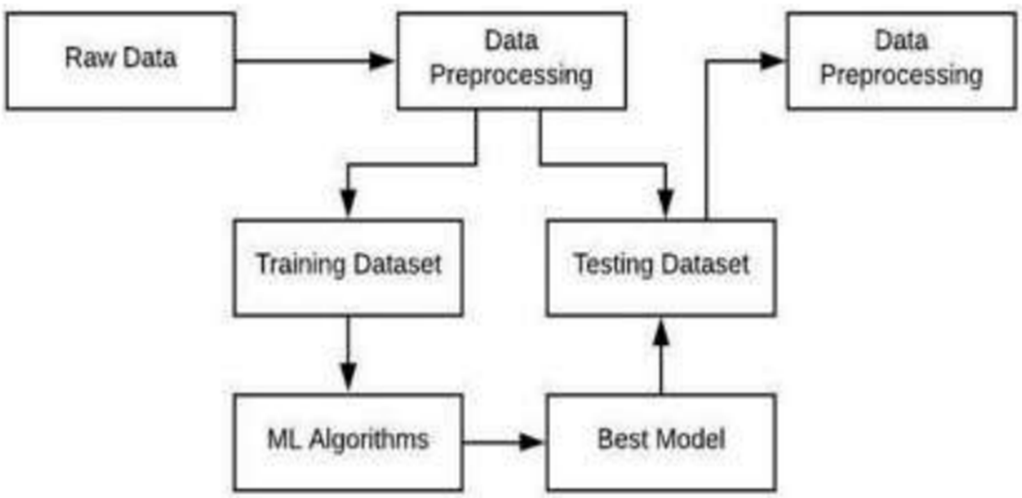


Fig. 1 ML Model Architecture

As depicted in the figure above, the raw data from the original dataset undergoes Data Preprocessing as the initial phase. This process involves cleaning the data of redundancies, missing values, etc., rendering it suitable for training various algorithmic models. The training of models is a fundamental step in machine learning projects, typically involving Supervised Learning or Unsupervised Learning approaches. Our system primarily employs Supervised Learning initially. In Supervised Learning, the system learns from examples in a training set and then predicts new values based on a test set. Effective partitioning of datasets, often using an 80/20 split for training and testing purposes, is crucial for achieving high model accuracy.

In our system, different algorithms are applied to the training dataset, and based on the model's confidence and testing dataset accuracy, the best-performing algorithm is selected for application to the testing dataset, resulting in accurate predictions.

The above diagram illustrates the use cases of our system, involving two actors: the user and the server containing the ML model for disease prediction. Users first register with the system, choose a disease category, and provide input in the form of an image or answers to questionnaires. This input is then forwarded to the ML model for prediction. Subsequently, the system displays the predicted result and suggests appropriate medications based on the predicted disease.

By implementing these components, our proposed system aims to revolutionize disease prediction by prioritizing user experience, accuracy, and efficient healthcare delivery.

VI. METHODOLOGY

A. *Data Collection and Preprocessing*

In our endeavor to achieve precise disease detection, our methodology adopts a multifaceted approach. We commenced with exhaustive data collection, meticulously aggregating diverse datasets pertinent to skin cancer, heart disease, diabetes, and Parkinson's disease. The primary dataset, HAM10000, served as the cornerstone for skin cancer classification. To ensure data reliability, we subjected the datasets to rigorous preprocessing techniques. This entailed cleansing them of noise, standardizing formats, and balancing class distributions. For skin cancer detection, images from HAM10000 underwent preprocessing techniques such as resizing, normalization, and augmentation, optimizing them for effective training of Convolutional Neural Networks (CNNs). This pivotal step significantly enhanced the model's discernment of subtle disease-related patterns.

B. *Deep Learning Model Development*

Central to our methodology was the development of sophisticated deep learning models, specifically Convolutional Neural Networks (CNNs), tailored for disease detection. Leveraging the prowess of CNNs in image-based pattern recognition, our team meticulously designed and trained intricate neural architectures. These models underwent iterative refinement, employing techniques such as transfer learning and ensemble methods. Transfer learning enabled us to harness pre-trained models, customizing them to our healthcare context, thus reducing training time while maintaining accuracy. Ensembles of multiple CNN architectures were devised, amalgamating their individual strengths to create a robust disease detection framework.

Rigorous validation and testing procedures were implemented to evaluate the models' performance, ensuring their efficacy in accurately classifying diseases based on user-input symptoms.

1) Model 1 Accuracy:

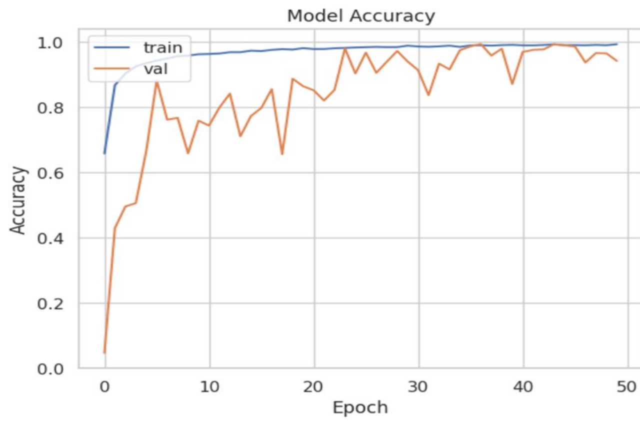


fig. 2 Model 1 Accuracy graph

2) Model 2 Accuracy

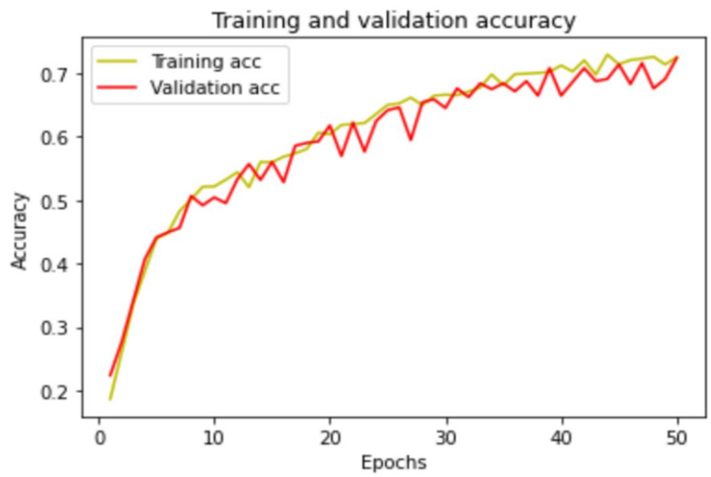


fig. 3 Model 2 Accuracy graph

3) Final Model Accuracy and Loss

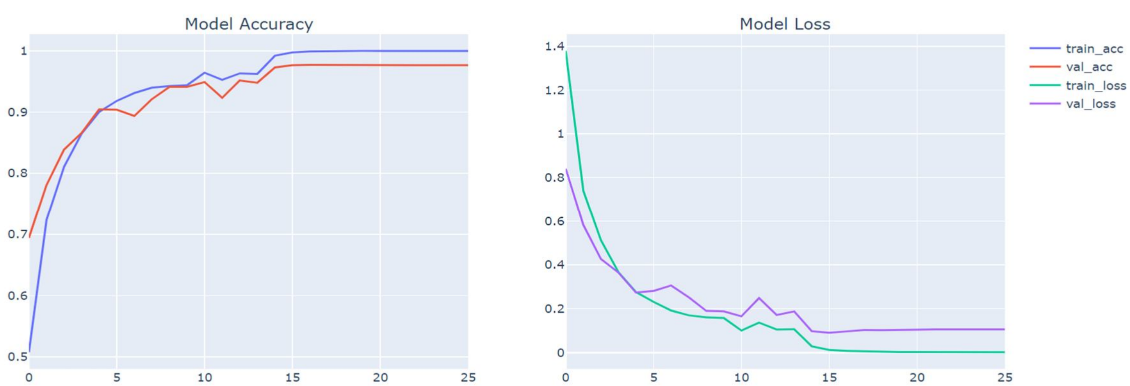


fig.4 Final Model Accuracy and Loss Graph

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue IV Apr 2024- Available at www.ijraset.com

C. User-Driven Interface and Holistic Healthcare Integration:

Beyond technical aspects, our methodology prioritized user experience and holistic healthcare integration. We meticulously designed a Java-based mobile application with an intuitive interface, ensuring ease of use for users from diverse backgrounds. The user-driven interface facilitated seamless symptom input, enabling precise disease predictions. Additionally, the application was augmented with a comprehensive medication database and intelligent algorithms. These algorithms analyzed user data and provided tailored medication suggestions, transforming the app into a holistic healthcare guide. This integration of disease prediction and medication recommendations offered users a comprehensive solution, empowering them to proactively manage their health. Thus, our methodology not only focused on technological innovations but also seamlessly integrated user-centric design and healthcare expertise, culminating in a transformative healthcare tool at the nexus of cutting-edge technology and real-world medical needs.

fig. 5 Home Fragment UI

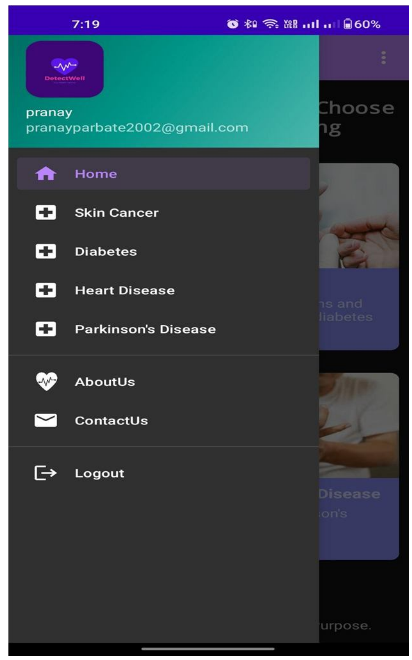


fig. 6 Navigation UI

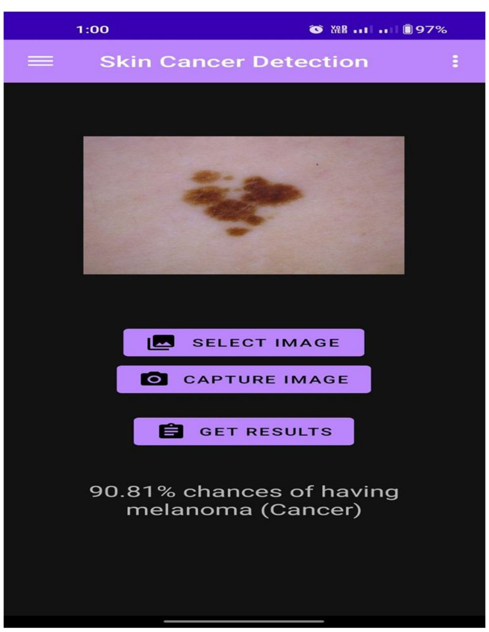


fig. 7 Skin Cancer Detection Fragment UI

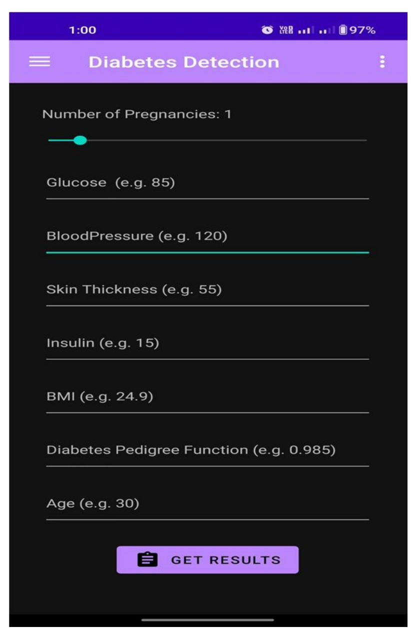


fig. 8 Diabetes Detection Fragment UI

D. Continuous Learning and Ethical Considerations:

An integral aspect of our methodology was establishing a framework for continuous learning and ethical considerations. Our models were designed to adapt and learn from new data, ensuring relevance and accuracy over time. Regular updates and model retraining protocols were implemented to enable the application to evolve alongside emerging medical research and diagnostic advancements. Additionally, ethical considerations remained paramount throughout the project. Stringent privacy measures were integrated into the app, safeguarding user data and ensuring compliance with relevant healthcare regulations. Transparent and interpretable AI techniques were employed, enabling users to comprehend how the app arrived at specific predictions, and fostering trust and confidence. Ethical guidelines were rigorously followed, ensuring fairness, accountability, and transparency in both the development and deployment of the innovative Java-based mobile application. This unwavering commitment to learning, ethical practices, and user trust underscores the comprehensive and responsible approach adopted in redefining the landscape of disease detection and management.

VII. CONCLUSION

In conclusion, the proposed multi-disease prediction system represents a significant advancement in healthcare technology, offering a streamlined and efficient approach to disease detection and management. By addressing the limitations of existing systems, such as limited dataset size and cumbersome user interfaces, our system aims to provide users with accurate and timely predictions while minimizing the time required for input.

Through the integration of machine learning techniques, specifically supervised learning algorithms, our system demonstrates the potential to adapt and evolve alongside emerging medical research and diagnostic advancements. By leveraging sophisticated deep learning models, such as Convolutional Neural Networks (CNNs), we can enhance the accuracy and reliability of disease predictions, thus bridging the gap between patients and healthcare providers.

Furthermore, the inclusion of a Doctor's Consultation feature underscores our commitment to holistic healthcare integration, empowering users to seek professional medical advice based on generated reports. This not only enhances the user experience but also fosters trust and confidence in the system among healthcare professionals.

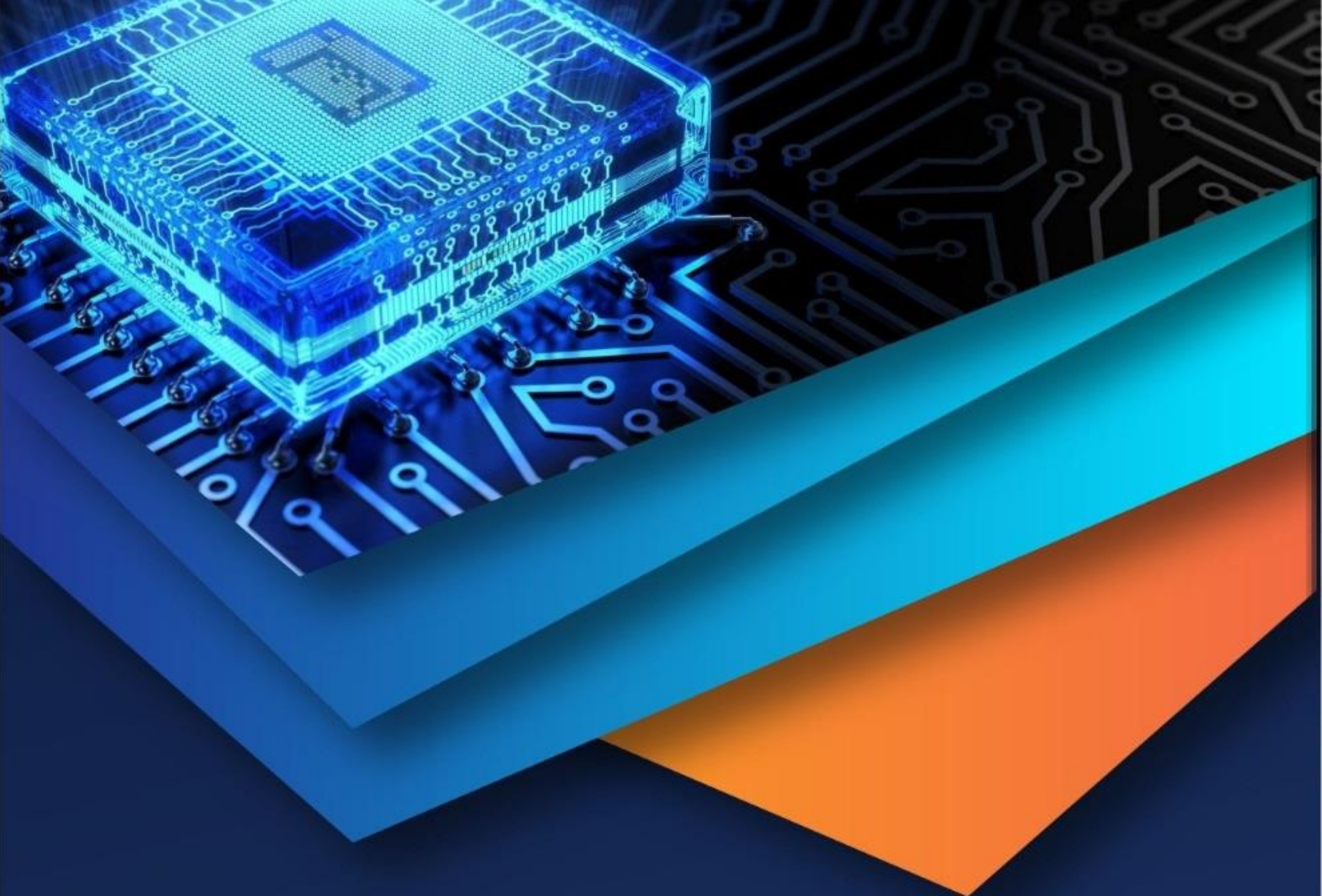
In essence, our proposed system serves as a connecting bridge between patients and doctors, facilitating proactive health management and personalized care. By prioritizing user-centric design, continuous learning, and ethical considerations, we aspire to redefine the landscape of disease detection and management, ultimately improving patient outcomes and transforming healthcare delivery in the digital age.

VIII. ACKNOWLEDGMENT

We extend our heartfelt gratitude to all those who contributed to the successful completion of this research project. First and foremost, we would like to express our sincere appreciation to our supervisor, Prof. Suresh Reddy, for their invaluable guidance, unwavering support, and insightful feedback throughout the duration of this project. Their expertise and mentorship played a pivotal role in shaping the direction and execution of our research endeavors. We are immensely thankful to the authors and researchers whose pioneering work laid the foundation for our study. Their contributions to the fields of machine learning, healthcare technology, and disease prediction have been instrumental in shaping our understanding and approach. We would like to acknowledge the contributions of Pranay Parbate, Dhairyashil Thombare, Tushar Yerkal, and Adarsh Varpe, whose dedication and collaboration greatly enriched the development and implementation of our proposed system. Their collective efforts and expertise were instrumental in overcoming challenges and achieving milestones. We are grateful to the participants and volunteers who generously contributed their time and data, enabling us to conduct thorough experimentation and validation of our system. Their involvement was invaluable in ensuring the accuracy and reliability of our findings. Our sincere thanks are also due to the academic institutions, funding agencies, and research organizations that provided financial support and resources for this project. Their contributions were essential in facilitating our research efforts and enabling the realization of our goals. Lastly, we would like to express our gratitude to our families, friends, and loved ones for their unwavering encouragement, understanding, and patience throughout this journey. Their support provided us with the motivation and inspiration to persevere through challenges and strive for excellence. In conclusion, we acknowledge and appreciate the collective efforts of all individuals and entities who contributed to the successful completion of this research project. Your support and collaboration have been indispensable, and we are deeply grateful for your contributions.

REFERENCES

[1] "Deep Convolution Neural Network based Automatic Multi-Class Classification of Skin Cancer from Dermoscopic Images" Vijaya Mishra, Akash Kumar V, Monika Arora (2021)
[2] "Analysis and Classification of Skin Cancer Images using Convolutional Neural Network Approach" Khalid Zaman, Javed Iqbal Bangash, Sozan Sulaiman Maghdid (2020)
[3] "A Clinical support system for Prediction of Heart Disease using Machine Learning Techniques" Halima EL HAMDAOUI, Saïd BOUJRAF, Nour El Houda CHAOUI, Mustapha MAAROUFI (2020)
[4] "Predictive Analytics on Diabetes Data using Machine Learning Techniques" V Krishnapraseeda M S Geetha Devasena V Venkatesh A Kousalya (2021)
[5] "A Supervised Machine Learning Approach using Different Feature Selection Techniques on Voice Datasets for Prediction of Parkinson's Disease" Satyabrata Aich, Hee-Cheol Kim, Kim Young, Kueh Lee Hui, Ahmed Abdulhakim Al-Absi, Mangal Sain (2019)



10.22214/IJRASET

INDEX COPERNICUS
45.98

IMPACT FACTOR:
7.129

TOGETHER WE REACH THE GOAL
IMPACT FACTOR:
7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY