iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** V **Month of publication:** May 2024

DOI: https://doi.org/10.22214/ijraset.2024.62908

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue V May 2024- Available at www.ijraset.com

Machine Learning and Streamlit Based Multi Disease Predictor

Dr. N.V Rajasekhar Reddy[1], Sreshta Mannam[2], Suru Satwika[3], Devi Sree M[4]
Department of Information Technology, MLR Institute of Technology Hyderabad, India

Abstract: ***Using machine intelligence, the Automated Multiple Disease Prediction System is chief the hole or door in vessel healthcare with tailor-made situation plans and predicting diagnoses. The answer can call either someone has Diabetes, Heart issues, and Parkinson's disease established the determined manifestations, record of what happened, and results. The system uses progressive algorithms to a degree support vector machines and logistic reversion, detecting early warning signs and diminishing risks. This full of enthusiasm game plan translates patient care, improving results and guaranteeing sustainability in healthcare. In an cycle of promptly progressing healthcare, this science shows a watershed importance in predicting diagnoses and patient comfort. In order to facilitate belongings for completely users, we grown a netting use that admits the consumer to anticipate an disease by just introducing allure attribute principles.***
Keywords: Machine Learning (ML), Support Vector Machine (SVM), Logistic Regression, Classification, Accuracy, Prediction

I. INTRODUCTION

Machine learning has arose as a important tool in healthcare for forecasting diversified diseases, contribution a example shift in full of enthusiasm healthcare management. By controlling enormous datasets including clinical records, behavior determinants, and demographic facts, machine intelligence algorithms authorize healthcare practitioners to forecast the attack and progress of differing diseases. Furthermore, the scalability and changeability of machine intelligence models allow for evident-occasion listening of patient health dossier, happening in early ailment trajectory discovery and fast intervention. Collaborative works middle from two points integrative teams, containing clinicians and dossier chemists, are essential to guaranteeing the righteous development, confirmation, and exercise of machine intelligence-driven predicting models in dispassionate practice. As machine intelligence advances and evolves, it holds huge promise for transforming disease prognosis, redefining healthcare transfer, and promoting a proactive approach to embodied first-contact medical care.

A. Diabetes

One of ultimate common names for diabetes, or diabetes mellitus, is extreme glucose. The hormone insulin arrange dispatching glucose into your cells so they can use it as strength. Diabetes is from either lacking insulin production or vulgar insulin exercise by the physique.

B. Heart Disease

"Heart disease" refers to a type of cardiac environments. Cardiovascular disease (CVD) is the most universal type of essence ailment, affecting ancestry flow to and from the essence. Reduced blood flow can trigger a heart attack. Cardiomyopathy's early stages may not show any symptoms at all. Symptoms that may appear as the illness worsens include light-headedness, dizziness, and fainting, fatigue, having trouble breathing when exercising or sitting still. having trouble breathing while trying to sleep at night or when you get up, irregular heartbeats, enlarged ankles, feet, legs. Some society's health possibly improved by making behavior modifications and communicable cures.

C. Parkinson's Disease

Parkinson's disease is a slowly advance affecting animate nerve organs condition that impairs flexibility. It is mostly from tremors, inflexibility, dullness, and issues with arrangement and balance. It takes place when certain nerve containers concerning the mind, expressly those being the reason for bearing dopamine, embellish injured or decline. Dopamine is a neurotransmitter that plays a important charge in arranging campaign and sensitive answers. As Parkinson's disease progresses, things grant permission also knowledge non-engine manifestations such as intelligent deterioration, color changes, sleep disturbances, and unrestrained politically dysfunction.

While skillful is immediately no cure for Parkinson's disease, situations such as cure, tangible healing, and deep intelligence provocation can help manage syndromes and raise the level of material comfort for those the one are afflicted. Additionally, continuous research works are met on understanding the latent causes of the disease and developing new therapies to slow its progression and alleviate symptom

II. LITERATURE REVIEW

This section briefly presents the works related to disease prediction by K. Arumugam [1] Data mining for healthcare is valuable in assessing the efficacy of medical interventions. Decision tree model outperformed naïve Bayes and supported vector machine models in predicting heart disease in diabetic patients. Hybrid models based on clustering and classification achieved good accuracy for predicting type 2 diabetes. However, this machine learning model predicts only type 2 diabetes and heart disease in diabetic patients.s. Hence, accuracy is low and other diseases can also be added to the model.

Tarigoppula V.S Sriram [2] Machine learning algorithms maybe secondhand for the disease of Parkinson's disease. Various clinical features can provide accurate diagnosis of Parkinson's disease. Neural networks and resolution seedlings have been secondhand for trustworthy disease of Parkinson's disease. Here, only one disease is predicted and diagnosed. Other diseases can also be added to the machine learning model and this model has potential to increase its accuracy.

Valle Harsha Vardhan [3] Application of machine learning for heart disease prediction has significant social impact. Decision tree, Naïve Bayes, Logistic Regression, Random Forest and Ensemble classifier combining strong and weak classifier machine learning methods are used in this work. Extreme Gradient Boosting classifiers had the highest accuracy. It is noteworthy that this system can be extended to incorporate additional diseases in the future, thereby increasing the efficiency and usefulness.

Zong Chen, Dr. Joy Iong [4] A machine learning approach for the early detection of Coronary Artery illness (CAD), a form of heart illness, is presented in this study. Preventing cardiovascular disorders can be aided by early detection of coronary artery disease. Adaptive image-based classification methods for CAD prediction are covered in this work. But it's important to keep in mind that machine learning models aren't always perfect or precise. It is essential to use real data to confirm the model's correctness and have a medical professional review the results in order to ensure patient safety and security.

Nashif Shadman [5] In this work, a machine learning algorithm was chosen based on a substantial dataset to distinguish the presence of heart disease and estimate its probability from among the available algorithms in a Java- based open-access data mining platform (WEKA). Subsequently, an Arduino-based microcontroller system has been suggested for use in a continuous heart monitoring system. The methodology works well, but the processing and memory capacity of Arduino boards is constrained, which could limit the complexity of the real-time monitoring system. It could face security issues with the transfer of private medical data over networks, as well as delays in data transmission that could jeopardize patient treatment schedules.

Rudra A. Godse [6] The paper evaluates the performance of medical disease prediction using machine learning algorithms such as support vector machine, decision tree and naïve bayes classifier. The best algorithm is selected based on model confidence and testing dataset accuracy. The selected algorithm is applied to the testing dataset for accurate results. Even so the model has a very good accuracy score, potential scalability, interpretability to overfitting may pose challenges in understanding and refining the processes. Moreover, while choosing an algorithm, relying solely on model confidence and dataset accuracy testing could leave out other crucial aspects like interpretability, computational efficiency, and robustness to unbalanced and noisy data.

Dr. K Kishore Raj [7] The paper presents a system with three disease prediction models which uses support vector machines, logistic regression, decision tree and KNN machine learning algorithms. The model predicts heart disease, diabetes, and breast cancer. Boosting is used in this model to transform weak learners into strong ones. The results proposed by the system are in binary format which is yes or no. Nevertheless, the system has a scope to be expanded by adding other diseases which can increase the model's utilization and effectiveness.

Rinkal Keniya [8] In this model, the disease prediction is done based on symptoms, age, and gender. An open-source dataset is used to create an excel sheet with symptoms. The RUSBoost algorithm is used to improve the performance of trained datasets. The SubSpace KNN method is used for classification problems with high-dimensional training datasets. Some models have low accuracy due to parameter dependency and this can be increased using effective machine learning models and increasing training datasets.

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

III. PROPOSED METHODOLOGY

By utilizing machine learning techniques like logistic regression and support vector machines, the proposed system presents a paradigm for precisely forecasting a range of diseases. By merging many datasets, we guarantee the highest possible level of prediction precision. Its performance is thoroughly evaluated and compared in order to identify which method works best for prediction. We have created an easy-to-use web tool to streamline communication between physicians and patients. By adding relevant characteristic information, users can quickly predict diseases when using this app. Our team is dedicated to facilitating end users' lives and advancing better health outcomes.

The projected methods for composing various key steps. Firstly, the data group includes gathering inclusive healing datasets containing appropriate visages such as age, grammatical rules applying to nouns that connote sex or animateness, ancestry pressure, cholesterol levels, and disease effects. Subsequently, dossier preprocessing is conducted to wash and preprocess the calm data, manage absent values, climb mathematical features, and encrypt unconditional variables as necessary to guarantee dossier quality and thickness. Following this, feature construction is employed to extract and engineer significant physiognomy from the pre-processed data, seizing appropriate information essential for affliction prophecy while considering rule information and expert insights. Moving forward, model pick includes choosing acceptable machine intelligence algorithms to a degree Support Vector Machine (SVM) and Logistic Regression based on the dataset's traits and the forecast task's nature. These picked models are before trained utilizing the pre-processed and feature-devised data, optimizing their limits to blow up predictive veracity and inference performance. Simultaneously, a common netting interface is grown utilizing Streamlit, enabling consumers to recommend their medical dossier seamlessly and endure real- opportunity prophecies for multiple ailments in a handy manner. Following model unification accompanying the Streamlit interface, bureaucracy bears rigorous experiment and judgment using different experiment datasets to ensure trustworthy forecasting and robustness. Finally, bureaucracy is redistributed to a production surrounding, making it approachable to healthcare professionals and things pursuing affliction prediction aids, while routine maintenance and refurbishes are transported to incorporate new dossier, better prediction algorithms, and address consumer response, ensuring allure persisted effectiveness and pertinence in dispassionate practice and public health pushes.

A. Advantages of the Proposed System

1) The system uses different machine-learning methods to predict multiple diseases.
2) Rapid analysis of large datasets within minimal time frames.
3) Extensive algorithmic accuracy evaluation and comparison to identify the best predictor.

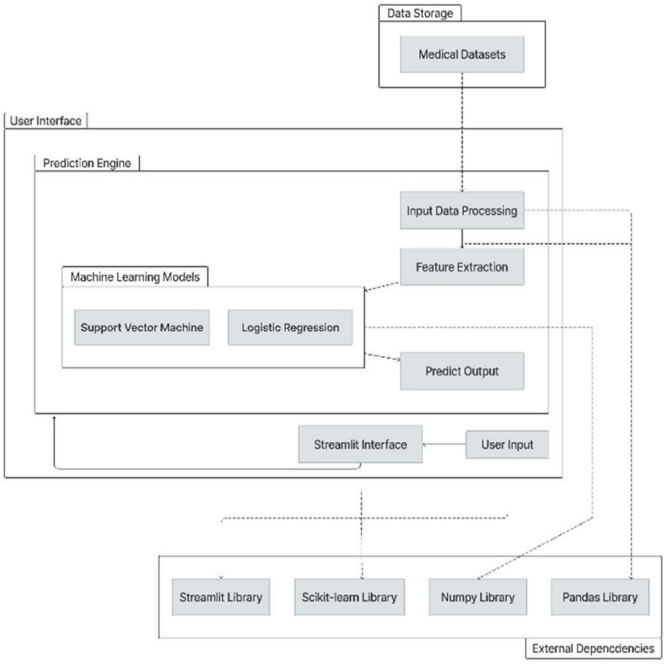


Fig. 1 System Architecture

The architecture shown in Fig. 1, the Streamlit Interface serves as the consumer architectural finish component, allowing consumers to recommend their fitness dossier. This connection seamlessly communicates accompanying the Prediction Engine, which orchestrates the prediction process. The Machine Learning Models piece composes miscellaneous algorithms like Support Vector Machine and Logistic Regression, prepared on Medical Datasets. When a forecast request is taken, the Prediction Engine retrieves the appropriate prepared model from the Trained Models table and fetches the inevitable dossier from the Medical Datasets. The Prediction Engine before promotes the picked Machine Learning Model to act prognoses established the recommendation dossier. Once the prophecy is achieved, the result is restored to the Streamlit Interface for display to the user. This construction guarantees a sleek flow of news, from consumer recommendation to forecast harvest, leveraging machine intelligence methods to support correct forecasting for ailments like Parkinson's disease, diabetes, and ischemic heart disease.

IV. RESULTS AND DISCUSSIONS

The results of the project explain hopeful outcomes mistakenly foreseeing diabetes, heart disease, and Parkinson's disease based on recommendation appearance supported by consumers. Through the implementation of logistic reversion and support heading vehicle (SVM) algorithms, the models achieved extreme levels of veracity, accuracy, recall, and F1 score, indicating their influence in categorization tasks. Evaluation verification to a degree of the area under the recipient operating characteristic curve (AUC- ROC) further justifies the strength of the models in distinguishing between certain and negative instances of each affliction. The user-friendly netting use grown utilizing Streamlit provides an instinctive connection for consumers to input their facts and endure honest-period predictions, embellishing approachability, and utility. Overall, the project's results underscore the potential of machine intelligence methods in promoting early disease discovery and risk appraisal, authorizing individuals to form conversant determinations about their health and welfare.

The site provides a foolproof connection for things to recommend their medical facts and accept original-time indicators for Parkinson's disease, diabetes, and congestive heart failure. Leveraging machine intelligence algorithms such as Support Vector Machine and Logistic Regression, the site offers correct guesses while guaranteeing transparency through descriptive intuitions. Interactive visualizations aid in giving prediction results, while rigid solitude measures safeguard consumers' impressionable health dossier. With allure sensible design and commitment to authorizing conversant administration, the website serves as a valuable source for early ailment discovery and is full of enthusiasm for health administration.

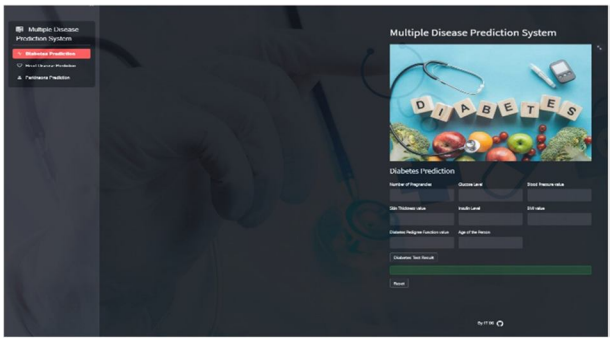


Fig. 2 Diabetes Prediction Page

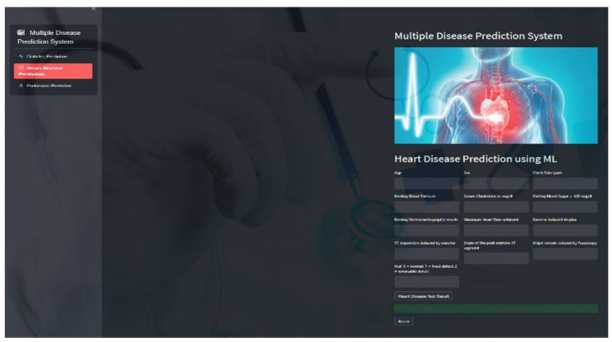


Fig. 3 Heart Disease Prediction Page

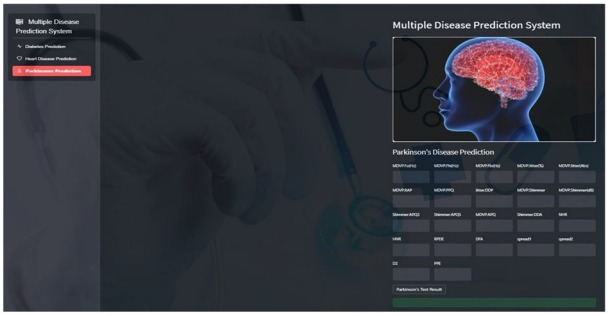


Fig. 4 Parkinson's Disease Prediction Page

With a plain and instinctive design, the website guarantees approachability for all users pursuing embodied health observations.

Fig. 5 Output of Diabetes Prediction

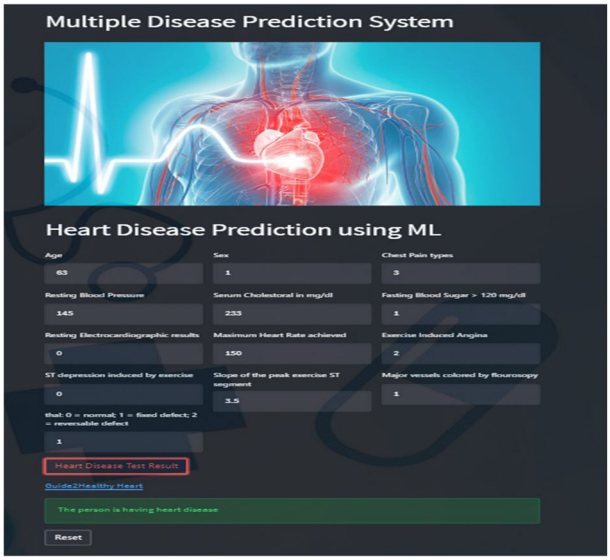


Fig. 6 Output of Heart Disease Prediction

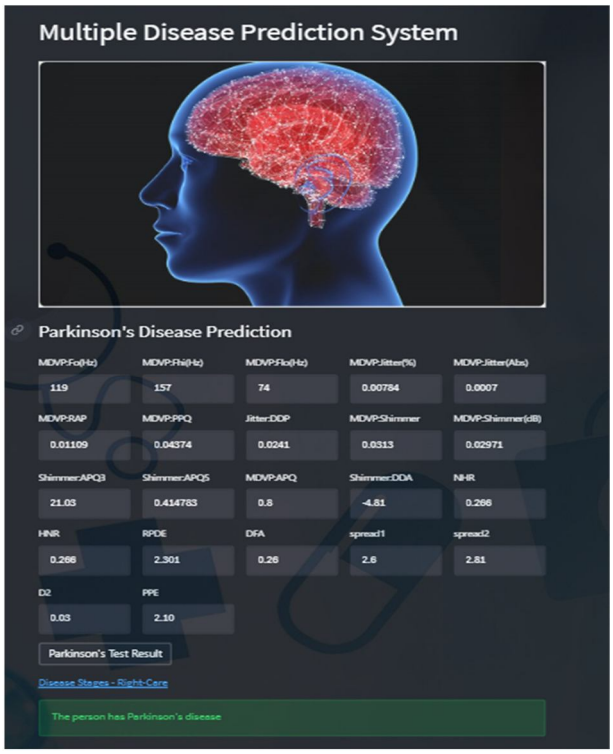


Fig. 7 Output of Parkinson's Disease Prediction

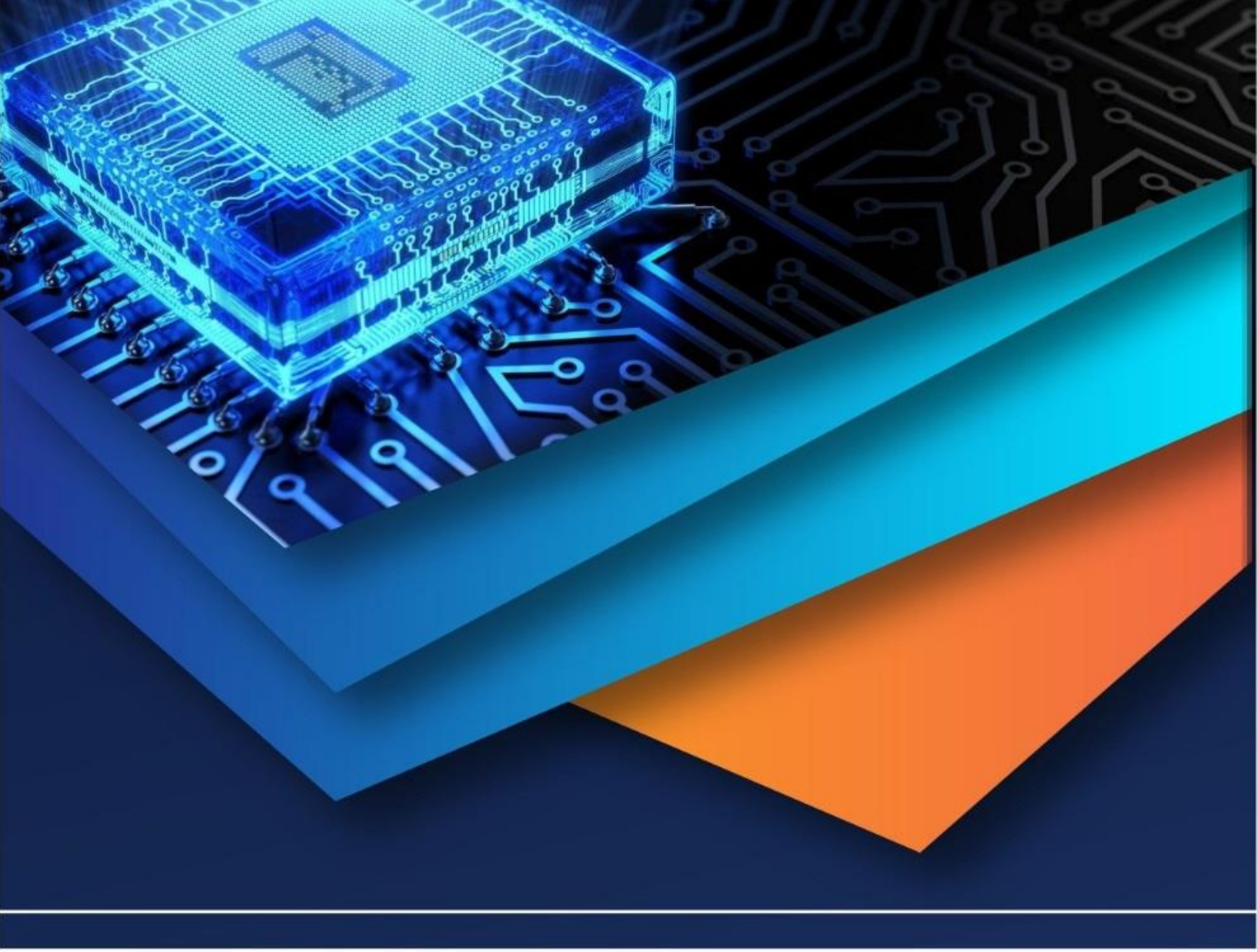
V. CONCLUSION

By mixing advanced algorithms and common imagination capabilities, the terrace offers an adjustable tool for healthcare experts and things alike. Early detection of ailments such as Parkinson's disease, diabetes, and myocardial infarction allows timely mediation and embodied health administration, eventually improving energy consequences and reducing healthcare costs. Through the exercise of logistic reversion and support heading machine (SVM) algorithms, in addition to inclusive dossier preprocessing and model training processes, the project explains the potential of machine intelligence in assisting healthcare experts and things in early affliction detection and risk appraisal. The growth of a netting application utilizing Streamlit supports a shared platform for consumers to recommend their facts and receive certain- opportunity forecasts, enhancing approachability and utility. Overall, the project focal points to the importance of leveraging machine intelligence methods and netting-based sciences to address healthcare challenges and enable things to make conversant resolutions about their fitness. By controlling state-of-the-art electronics like Artificial Intelligence and Deep Learning, bureaucracy guarantees flexibility and scalability in the middle of the active healthcare environment. Its instinctive connection speeds smooth interaction middle from two points victims and healthcare providers, permissive timely invasions, and joint decision-making. With a devoted effort to something approachability and inclusivity, bureaucracy aims to transform affliction administration, embellish patient well- being, and boost healthcare principles in today's active countryside.

REFERENCES

[1] A. Gavhane, G. Kokkula, I. Pandya, and K. Devadkar, "Prediction of heart disease using machine learning," in 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA), 2018, pp. 1275–1278.
[2] Y. Hasija, N. Garg, and S. Sourav, "Automated detection of dermatological disorders through image-processing and machine learning," in 2017 International Conference on Intelligent Sustainable Systems (ICISS),2017, pp. 1047–1051.
[3] S. Uddin, A. Khan, M. E. Hossain, and M. A. Moni, "Comparing different supervised machine learning algorithms for disease prediction,"BMC Medical Informatics and Decision Making, vol. 19, no. 1, pp. 1–16, 2019.
[4] R. Katarya and P. Srinivas, "Predicting heart disease at early stages using machine learning: A survey," in 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), 2020,pp. 302–305.
[5] P. S. Kohli and S. Arora, "Application of machine learning in disease prediction," in 2018 4th International Conference on Computing Communication and Automation (ICCCA), 2018, pp. 14.

[6] M. Patil, V. B. Lobo, P. Puranik, A. Pawaskar, A. Pai, and R. Mishra,"A proposed model for lifestyle disease prediction using support vector machine," in 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT), 2018, pp. 1–6.
[7] F. Q. Yuan, "Critical issues of applying machine learning to condition monitoring for failure diagnosis," in 2016 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM),2016, pp. 1903–1907.
[8] S. Ismaeel, A. Miri, and D. Chourishi, "Using the extreme learning machine (elm) technique for heart disease diagnosis," in 2015 IEEE Canada International Humanitarian Technology Conference (IHTC2015), 2015,pp. 1–3.
[9] D. Dahiwade, G. Patle, and E. Meshram, "Designing disease prediction model using machine learning approach," Proceedings of the 3rd International Conference on Computing Methodologies and Communication,ICCMC 2019, no. Iccmc, pp. 1211–1215, 2019.
[10] S. Jadhav, R. Kasar, N. Lade, M. Patil, and S. Kolte, "Disease Prediction by Machine Learning from Healthcare Communities," International Journal of Scientific Research in Science and Technology, pp. 29–35,2019.
[11] R. Saravanan and P. Sujatha, "A state of art techniques on machine learning algorithms: A perspective of supervised learning approaches in data classification," in 2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS), 2018, pp. 945–949.
[12] M. Marimuthu, M. Abinaya, K. S., K. Madhankumar, and V. Pavithra,"A Review on Heart Disease Prediction using Machine Learning and Data Analytics Approach," International Journal of Computer Applications, vol. 181, no. 18, pp. 20–25, 2018.
[13] A. K. Dwivedi, "Performance evaluation of different machine learning techniques for prediction of heart disease," Neural Computing and Applications, vol. 29, no. 10, pp. 685–693, 2018.
[14] K. Polaraju, D. Durga Prasad, and M. Tech Scholar, "Prediction of Heart Disease using Multiple Linear Regression Model," International Journal of Engineering Development and Research, vol. 5, no. 4, pp.2321–9939, 2017. [Online].
[15] S. Pouriyeh, S. Vahid, G. Sannino, G. De Pietro, H. Arabnia, and J. Gutierrez, "A comprehensive investigation and comparison of machine learning techniques in the domain of heart disease," in 2017 IEEE Symposium on Computers and Communications (ISCC), 2017, pp. 204–207.
[16] H. L. Chen, C. C. Huang, X. G. Yu, X. Xu, X. Sun, G. Wang, and S. J. Wang, "An efficient diagnosis system for detection of Parkinson's disease using fuzzy k-nearest neighbor approach," Expert Systems with Applications, vol. 40, no. 1, pp. 263–271, 2013.
[17] M. Behroozi and A. Sami, "A multiple-classifier framework for Parkinson's disease detection based on various vocal tests," International Journal of Telemedicine and Applications, vol. 2016, 2016.
[18] ¨O. Eskidere, F. Ertas¸, and C. Hanilc¸ i, "A comparison of regression methods for remote tracking of Parkinson's disease progression," Expert Systems with Applications, vol. 39, no. 5, pp. 5523–5528, 2012.
[19] N. Lavesson, Evaluation and Analysis of Supervised Learning Algorithms and Classifiers, 2006.
[20] R. Caruana and A. Niculescu-Mizil, "An Empirical Comparison of Supervised Learning Algorithms Using Different Performance Metrics," Proceedings of the 23rd international conference on Machine Learning, pp. 161–168, 2006.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY