



International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 14, Issue 7, July 2025

Cardiovascular Analysis Based on Human Physical and Mental Health Conditions

Dr.V.Vasudha Rani, Ms.V. Anupama

Associate Professor, Department of IT, GMR Institute of Technology, Rajam, India

Associate Professor, Department of CSE , Lendi Institute of Engineering and Technology Autonomous, India

ABSTRACT: Among diseases currently people are suffering more, heart disease causes death more in the world. This integration of machine learning algorithms in assessing cardiovascular fitness and its broader implications for physiology and mental health. Machine learning methods unveil the intricate relationship between cardiovascular fitness levels and health across various sources. Health professionals utilize algorithms to extract extensive datasets comprising physical dimensions, lifestyle, and health outcomes. Regular physical activity and maintaining optimal fitness contribute significantly to reducing the risk of developing cardiovascular disease and improving outcomes for individuals already affected by the condition. The dataset provided information on over 3,08,854 patients and included 19 attributes, each representing a potential risk factor for cardiovascular disease. These attributes encompassed behavioral, and medical risk factors, providing a comprehensive basis for predictive modeling and risk assessment. Some algorithms used, include Decision trees, K Nearest Neighbors, Logistic Regression, Random Forest, which facilitate to recognises patterns and relationships in data. Evaluation criteria are accuracy ,Precision ,Recall, F1 score and interpretability metrics employed to gauge algorithm performance. Through the integration of machine learning with cardiovascular and fitness analytics, healthcare providers can tailor preventive care strategies, promote healthy lifestyles, and enhance overall quality of life.

KEYWORDS: Cardiovascular Disease, Physical fitness, Decision trees, Logistic Regression, K Nearest Neighbors, Random Forest.

I. INTRODUCTION

Cardiovascular disease (CVD) is among the global health problems, as evidenced by the mortality rates reported by the World Health Organization (WHO). Around 18 million people lost their lives each year from heart disease, that includes many conditions such as ischemic heart disease, heart attack and stroke. This major loss highlights the critical need for innovative strategies in prevention, early diagnosis, and treatment.

At the heart of this public health crisis is the need to understand the full spectrum of heart disease. Coronary heart disease, one of the most prevalent and deadly conditions, occurs when blood flow to the heart muscle is reduced. Warning signs like chest pain (angina), shortness of breath, fatigue, and dizziness are key indicators of heart damage, impacting individuals globally.

The landscape of cardiovascular diseases is shaped by an array of risk factors, spanning both modifiable and non-modifiable domains. Lifestyle factors such as diet, exercise, smoking, and alcohol use play a major role in increasing the risk of heart disease to CVDs. Simultaneously, genetic predispositions, advancing age, familial history, and comorbidities such as hypertension and diabetes paint a complex portrait of susceptibility to cardiovascular complications.

Yet, the narrative of cardiovascular health extends beyond the confines of physical parameters alone. The intricate interplay between physical and mental health conditions emerges as a pivotal determinant of cardiovascular outcomes. Obesity, hypertension, diabetes, and dyslipidemia cast long shadows over cardiovascular function, escalating the risk of CVDs. Equally impactful are mental health disorders, such as stress, anxiety, and depression, which exert their toll through diverse physiological mechanisms, including inflammation, autonomic nervous system dysregulation, and maladaptive health behaviors.

Acknowledging the symbiotic relationship between physical and mental well-being is crucial in charting a course towards improved cardiovascular care. Embracing comprehensive risk assessment strategies that integrate both realms of health can pave the way for early detection, personalized intervention, and tailored management approaches tailored to individual patient needs.

In this landscape of challenges and opportunities, advancements in machine learning (ML) technologies offer a glimmer of hope. ML algorithms, fueled by diverse datasets encompassing clinical, genetic, lifestyle, and behavioral factors, hold the promise of crafting accurate predictive models. Armed with these insights, healthcare professionals can steer towards proactive and personalized approaches to cardiovascular care, thereby potentially alleviating the burden on healthcare systems and improving patient outcomes.

In embarking on this journey through the realms of cardiovascular analysis, we endeavor to illuminate the intricate interplay between human physical and mental health conditions. By shedding light on this holistic perspective, we aspire to unlock pathways towards enhanced cardiovascular health outcomes and a brighter future for individuals affected by CVDs.

II. LITERATURE REVIEW

Using various machine learning algorithms such as K-Nearest Neighbor (KNN), Random Forest (RF), Logistic Regression (LR), Support Vector Machine (SVM), Gaussian Naive Bayes (GNB) and Decision Trees (DT) to predict ischemic heart disease. These methods have demonstrated their effectiveness in disease prediction by demonstrating high sensitivity, recall, F1 score, sensitivity and area under the curve (AUC) values. [1].

Using a hybrid deep neural network (HDNN) that combines artificial neural networks (ANN), convolutional neural networks (CNN), long-term memory (LSTM), and additional Dense layers, which enhance prediction accuracy for ischemic cardiovascular disease by effectively extracting and learning relevant features from input data [2].

The feature selection techniques such as Fast Correlation-Based Filter (FCBF), Relief, Particle Swarm Optimization (PSO) to identify optimal features for cardiovascular disease detection, and trained machine learning classifiers like Extra Tree and Random Forest on these features [3].

Pre-Generalized Laboratory Progress (GLP) standard training using self-Supervised Learning(SSL) for cardiovascular event detection, allowing the model to learn from the underlying structure of the data without requiring explicit labels. Interpolation methods like linear interpolation and piecewise cubic Hermite interpolation polynomial (PCHIP) can be applied and barycentric interpolation to extract values from known observations. [4].

Using machine learning and hyperparameter optimization to improve heart disease diagnosis and severity classification techniques to improve accuracy and clinical outcomes in healthcare. Advantages include development of efficient and effective decision-making methods to predict the incidence and severity of cardiovascular disease (CVD). [5].

Cardiovascular diagnosis based on machine learning using algorithms such as LDA, RF, GBC, DT, SVM and KNN, with Sequential Feature Selection (SFS) for essential feature selection, achieving high accuracy with Random Forest and Decision Tree classifiers [6].

HDPM model, which includes DBSCAN to anomaly detection, SMOTE-ENN to handling inconsistent data, and by XGBoost prediction achieved 95.90% and 96.40% accuracy on Statlog and Cleveland datasets, respectively, and outperforming other models and previous studies. A prototype of a Heart Disease Clinical Decision Support System (CDSS) was developed for early diagnosis [7].

Machine learning (ML) technology improves results of heart disease prediction and introduces deep learning-enhanced neural networks to improve diagnosis. Additionally, the study employs a Modified Huffman Algorithm for better lossless file compression and secures data transmission using the Advanced Encryption Standard and the Modified Huffman Algorithm. The research lacks a comparison with other existing algorithms or methods for the heart disease prediction [8].

It employs regression analysis, the XGBoost algorithm, and deep learning models to predict and diagnose cardiovascular diseases (CVDs), utilizing K-fold cross-validation techniques for robust evaluation. Regression analysis accurately predicts overlapping symptoms of CVDs and identifies essential features in CVD data using Principal Component analysis (PCA). The study used data collected from UAE hospitals, which included age, symptoms and heart disease data, divided into targets, diagnoses and characteristics for analysis.[9]

It uses logistic regression, deep neural networks, and Light GBM models, along with feature engineering incorporating variables such as age, sex, blood sugar, BMI, and more, to predict cardiovascular disease (CVD) risk. Feature importance analysis using the Shapley additive explanations method enhances CVD risk estimation. ML models demonstrated superior performance compared to ACC/AHA guidelines, with medication information further improving prediction accuracy. However, potential biases may arise from the use of retrospective medical data reliant on physicians' decisions regarding medication usage, and the feature engineering process is not adequately described. Two datasets, representing 2-year and 10-year follow-up analyses from the same subjects, were utilized, with features including age, sex, blood sugar, BMI, blood pressure, among others.[10]

III. PROPOSED METHOD

The methodology for conducting a cardiovascular analysis based on human physical and mental health conditions involves a comprehensive and structured approach. Initially, data collection is performed, utilizing a dataset from Kaggle, which provides a rich source of information for the analysis. Data collection went through a rigorous preliminary process that included data cleaning to remove missing values and eliminate inconsistencies, data visualization to understand relationships and distributions, data encoding to convert categorical variables into numerical formats, and data scaling to standardize numerical features. To facilitate model evaluation, the data is first split into training and testing. Use a variety of algorithms such as logistic regression, decision trees, SVMs, random forests, and K-nearest neighbors to train data model. To ensure performance, models are assessed using metrics such as accuracy, precision, recall and F1 score.

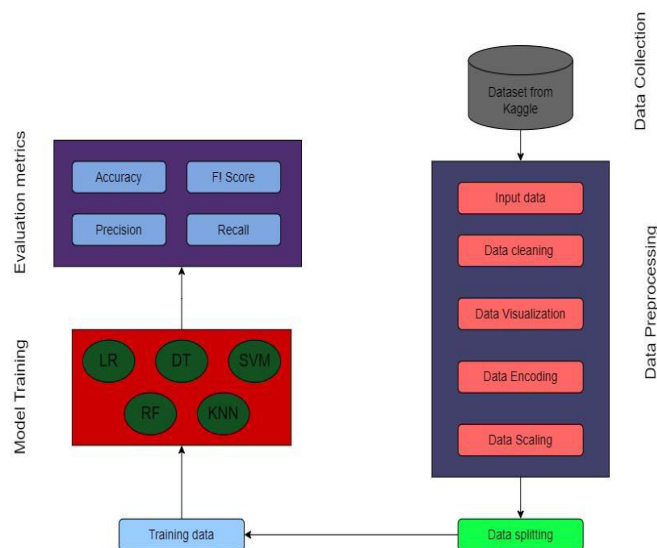


Fig. 1. Methodology

Technologies used in working:

1.Jupyter Notebook :

Jupyter Notebook is an indispensable tool in the development of the Cardiovascular Analysis Based on Human Physical and Mental Health Conditions project. This interactive computing environment allows for seamless integration of code execution, data visualization, and narrative text, making it ideal for the iterative and exploratory nature of data

science projects. In this project, Jupyter Notebooks facilitate the preprocessing of datasets obtained from platforms like Kaggle, enabling the cleaning, visualization, and transformation of data to prepare it for analysis.

2 Numpy

NumPy library plays an vital role in research and development of cardiovascular analysis. Based on Human Physical and Mental Health Conditions project. NumPy, short for Numerical Python, serves as a fundamental tool for scientific computing, offering powerful capabilities for handling multidimensional arrays, conducting mathematical operations, and facilitating high-performance computing tasks. In the context of this project, NumPy enables researchers and data scientists to preprocess vast datasets related to physical and mental health conditions. This includes operations such as data cleaning, normalization, and transformation, which are essential for preparing the data for machine learning models. NumPy's efficient handling of large arrays and matrices is crucial for these preprocessing tasks, ensuring that data is processed quickly and accurately.

3. Scikit-learn Library:

Scikit-learn, commonly referred to as Sklearn, is a fundamental library for machine learning in Python that provide robust tools to develop and evaluate predictive models. In the Cardiovascular Analysis Based on Human Physical and Mental Health Conditions project, Sklearn plays major role in different stages of machine learning pipeline, from initial data to model evaluation. By leveraging Sklearn's comprehensive set of tools for data preprocessing, model training, evaluation, and validation, the Cardiovascular Analysis project can develop highly accurate and reliable predictive models, ultimately aiding in better healthcare decisions based on integrated physical and mental health data.

4. Matplotlib:

Matplotlib is a widely used Python plotting library and framework for data visualization in the Cardiovascular Analysis Based on Human Physical and Mental Health Conditions project. Provides a variety of techniques for creating static, animated and interactive visuals, which are crucial for understanding data patterns and model performance. One of the primary uses of Matplotlib in this project is to perform exploratory data analysis (EDA). Visualization helps in gaining insights into the data, revealing trends, patterns, and outliers. Key features of Matplotlib for data exploration include histograms, Scatter plots, Box plots and line plots. Matplotlib supports the creation of interactive plots through tools like `mpl_toolkits` and integration with interactive environments such as Jupyter Notebooks. This allows for dynamic exploration of the data, where users can zoom, pan, and update plots in real-time, making the analysis process more engaging and insightful. By utilizing Matplotlib for data visualization, the Cardiovascular Analysis project benefits from enhanced data exploration, model performance evaluation, and effective communication of findings. This, in turn, supports more informed and accurate healthcare decisions based on comprehensive physical and mental health data visualizations.

IV. LIST OF SOME COMMON MENTAL HEALTH FACTORS

1. Stress

Impact: Chronic stress causes stress on the body, which results in high levels of hormones such as cortisol. This increases heart rate and blood pressure, leading to development of high blood pressure, other heart diseases.

Mechanism: Chronic stress can lead to unhealthy behaviors such as poor nutrition, physical activity, and smoking, which are risk factors for heart disease.

2. Depression

Impact: Depression is linked to an increased risk of heart conditions, including coronary artery disease and heart failure. People with depression are also more likely to have poor outcomes after a heart attack.

Mechanism: Depression can cause biological changes that can lead to heart disease, such as increased inflammation and abnormal heartbeats. It also affects treatment compliance and healthy lifestyle.

3. Anxiety

Impact: Depression can cause biological changes that can lead to heart disease, such as increased inflammation and abnormal heartbeats. It also affects treatment compliance and healthy lifestyle.

Mechanism: Stress activates the sympathetic nervous system, leading to an elevated heart rate and higher blood pressure. It can also encourage heart-harming habits, such as smoking and shortness of breath.

4. Social Isolation and Loneliness

Impact: Loneliness and isolation have been found to be important risk factors for heart disease. Lack of social support is associated with higher rates of heart disease and death.

Mechanism: Loneliness can increase stress and deteriorate health. It can also affect the immune system, causing illness and leading to heart problems.

5. Chronic Anger and Hostility

Impact: Chronic anger and hostility are linked to a higher risk of heart disease and negative cardiovascular events.

Mechanism: These emotions lead to the release of stress hormones, causing blood pressure and heart rate to increase. This can also damage the cardiovascular system and promote atherosclerosis.

6. Post-Traumatic Stress Disorder (PTSD)

Impact: Post-traumatic stress disorder (PTSD) is linked to a higher risk of cardiovascular disease, which includes heart disease and stroke.

Mechanism: Post-traumatic stress disorder (PTSD) causes the body to respond to stress, resulting in persistent stress hormones, increased inflammation, and poor lifestyle, all of which can lead to heart disease.

7. Sleep Disorders

Impact: Conditions such as insomnia and sleep apnea are associated with a higher risk of high blood pressure, heart disease, and stroke.

Mechanism: Inadequate sleep, lack of sleep can cause stress on the brain, high blood pressure, poor glucose metabolism, all of which contribute to risk of heart disease.

8. Substance Abuse

Impact: Abuse of substances such as alcohol, tobacco and drugs are important risk factors for heart disease.

Mechanism: Drug abuse can cause damage to the cardiovascular system, including high blood pressure, rapid heartbeat, irregular heartbeat. It also promotes unhealthy behaviors and lifestyle choices that increase the risk of heart disease.

List of some common physical health factors**1. High Blood Pressure (Hypertension)**

Impact: High blood pressure exerts pressure on the heart and arteries, raising the likelihood of heart disease and stroke.

Mechanism: It damages the arteries over time, making them more susceptible to atherosclerosis (build-up of plaque), which can result in heart attacks and strokes.

2. High Cholesterol Levels

Impact: High levels of LDL cholesterol, often referred to as "bad" cholesterol, can contribute to plaque accumulation in the arteries, raising risk of coronary artery disease.

Mechanism: LDL cholesterol can build up on artery walls, forming hard, thick deposits called plaques. This narrows the arteries and limits blood flow to the heart, which can result in chest pain (angina) or a heart attack.

3. Obesity

Impact: Obesity is a contributing factor to heart disease, as it increases the likelihood of developing conditions like hypertension and diabetes.

Mechanism: Excess body fat can lead to metabolic changes that contribute to cardiovascular risk, including insulin resistance, inflammation, and dyslipidemia (abnormal blood lipid levels).

4. Diabetes Mellitus

Impact: Diabetes greatly heightens risk of cardiovascular disease.

Mechanism: Prolonged high blood sugar levels can harm blood vessels and nerves that regulate the heart, while also elevating the chances of other risk factors like hypertension, dyslipidemia.

5. Physical Inactivity

Impact: Inadequate regular physical activity is a significant risk factor for heart disease.

Mechanism: Regular exercise aids in maintaining a healthy weight, reducing blood pressure, improving cholesterol levels, and boosting overall cardiovascular health. In contrast, sedentary lifestyles heighten the risk of obesity and other factors associated with heart disease.

6. Poor Diet

Impact: Diets rich in saturated fats, trans fats, cholesterol, and sodium can play a role in the development of heart disease.

Mechanism: An unhealthy diet can result in high blood pressure, elevated cholesterol levels, obesity, and diabetes, all of which raise the risk of cardiovascular disease.

7. Smoking

Impact: Smoking is one of major critical risk factors for heart disease.

Mechanism: The chemicals in secondhand smoke can harm the heart and blood vessels, contributing to atherosclerosis, heart disease, and other cardiovascular issues.

8. Family History of Cardiovascular Disease

Impact: A family history of heart disease raises the likelihood of developing cardiovascular issues..

Mechanism: Genetic factors can make individuals more susceptible to conditions like high blood pressure, elevated cholesterol levels, and coronary artery disease, all of which contribute to cardiovascular risk.

V. RESULTS AND DISCUSSION

Through advanced algorithmic analysis of a comprehensive dataset encompassing both physical and mental health factors, we have gained valuable insights into cardiovascular disease risk assessment and management.

Confusion Matrix :-

It is used for the performance metrics. It is generated True-positive, True Negative, False Positive, False Negative. Predicted values are described as Positive and Negative. Actual values are described as True and False.

True-Positive:- Both model Prediction and actual values are positive for the positive class.

True-Negative:- Both model Prediction and actual values are negative for the Negative class.

False-Positive:- The model is incorrectly predicting the negative class. This is also known as **TYPE-I** error.

False-Negative:- The model is making incorrect predictions for positive class.. This is also called as **TYPE-II** error.

Output for Physical Features:

		True Class	
		Positive	Negative
Predicated Class	Positive	TP	FP
	Negative	FN	TN

Fig 2. Confusion Matrix

		Confusion Matrix - Logistic Regression	
		0	1
True Labels	0	41640	15175
	1	12755	43964
		0	1
		Predicted labels	

Fig 3. Logistic Regression

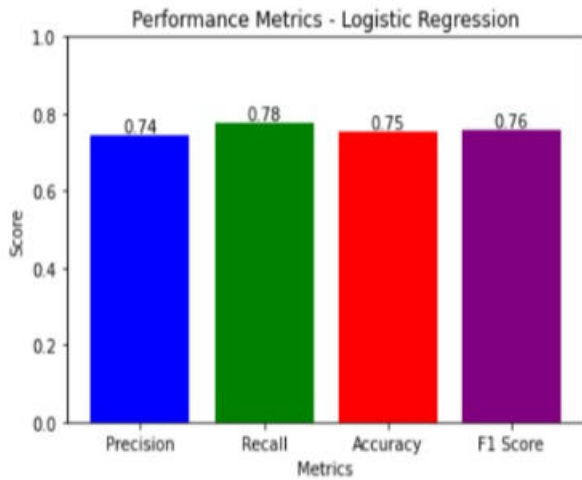


Fig 4. Logistic Regression

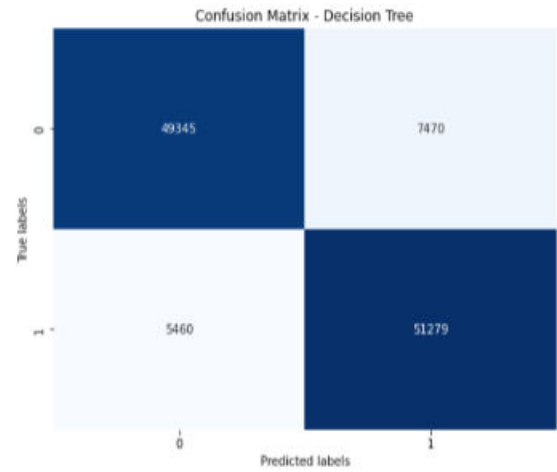


Fig 5. Decision Tree

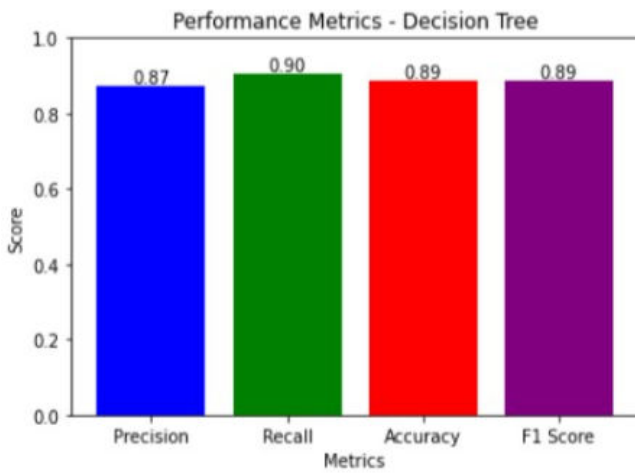


Fig 6. Decision Tree

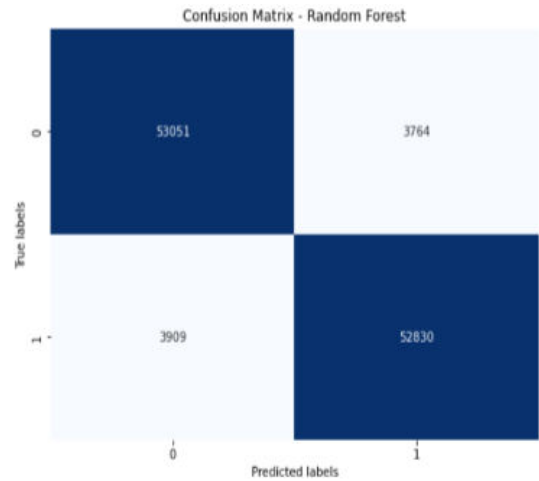


Fig 7. Random Forest

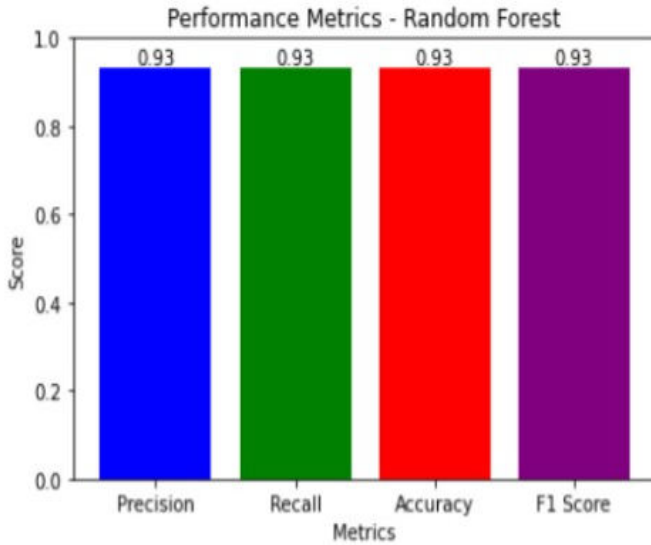


Fig 8. Random Forest

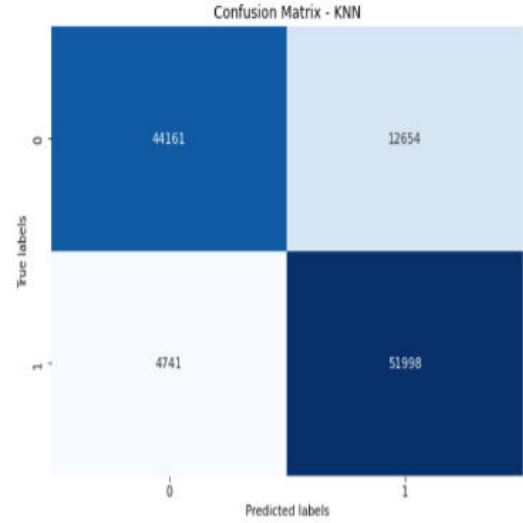


Fig 9. KNN

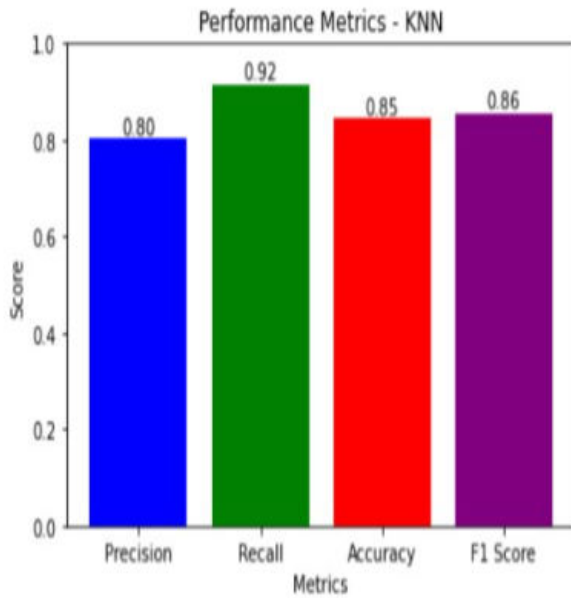
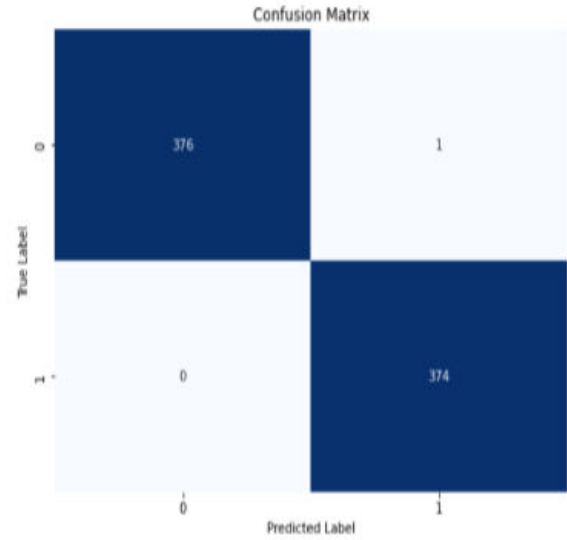


Fig 10. KNN



Accuracy: 0.9986684420772304
Precision: 0.9973333333333333
Recall: 1.0
F1 Score: 0.9986648865153538

Fig 11. KNN

Output for Mental Features

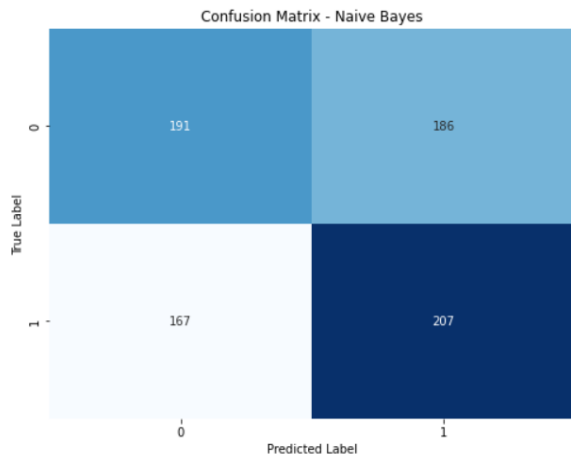


Fig 12. Naïve Bayes

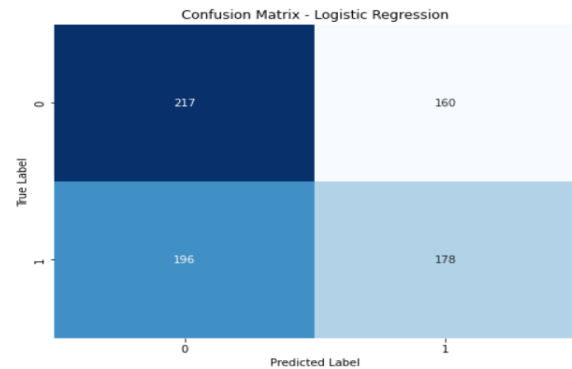


Fig 13. Regression

Algorithm	Accuracy
Logistic Regression	75.4363%
K-Nearest Neighbors	84.7517%
Decision Tree	88.7234%
Random Forest	93.2094%

Fig 14. Metrics for Physical Factors

Algorithm	Accuracy
Logistic Regression	52.5965%
Naive Bayes	52.9960%
K-Nearest Neighbors	99.8668%

Fig 15. Metrics for Mental Factors

VI. CONCLUSION

In conclusion, our project on CVD, incorporating both physical and mental health features, has yielded significant insights into the holistic nature of cardiovascular risk assessment. By integrating ML algorithms, particularly Random Forest, we achieved an impressive accuracy of 93% for the physical features data and for the mental features data the KNN get accuracy of 99.8% . This underscores the effectiveness of leveraging comprehensive datasets that encompass various physical and mental health attributes in predicting cardiovascular outcomes. The inclusion of mental health parameters alongside traditional physical dimensions has enriched our understanding of cardiovascular risk factors. It highlights the intricate interplay between physiological and psychological factors in determining cardiovascular health outcomes. By recognizing and incorporating mental health indicators into predictive modeling and risk assessment, we have advanced towards a more holistic approach to healthcare. Our findings emphasize the importance of tailoring preventive care strategies that address both physical and mental well-being to promote overall cardiovascular health. ML algorithms such as LR, DT, SVM, and RF have proven that invaluable in uncovering patterns , relationships within data. This enhances accuracy and effectiveness of predictive models, empowering healthcare professionals to deliver personalized care and interventions. Ultimately, our project underscores the potential of ML in cardiovascular fitness assessment to drive advancements in preventive medicine and healthcare delivery. By acknowledging the multifaceted nature of CVD and adopting a comprehensive approach that considers both physical and mental health features, we can strive towards better cardiovascular outcomes and enhanced overall well-being for individuals worldwide.

REFERENCES

1. Muhammad, G., Naveed, S., Nadeem, L., Mahmood, T., Khan, A. R., Amin, Y., & Bahaj, S. A. O. (2023). Enhancing Prognosis Accuracy for Ischemic Cardiovascular Disease Using K Nearest Neighbor Algorithm: A Robust Approach. IEEE Access.
2. Ghorashi, S., Rehman, K., Riaz, A., Alkahtani, H. K., Samak, A. H., Cherrez-Ojeda, I., & Parveen, A. (2023). Leveraging Regression Analysis to Predict Overlapping Symptoms of Cardiovascular Diseases. IEEE Access.
3. A. Abdellatif, H. Abdellatef, J. Kanesan, C. -O. Chow, J. H. Chuah and H. M. Ghenni, "An Effective Heart Disease Detection and Severity Level Classification Model Using Machine Learning and Hyperparameter Optimization Methods," in IEEE Access.
4. G. N. Ahmad, S. Ullah, A. Algethami, H. Fatima and S. M. H. Akhter, "Comparative Study of Optimum Medical Diagnosis of Human Heart Disease Using Machine Learning Technique With and Without Sequential Feature Selection," in IEEE Access.
5. T. Ullah et al., "Machine Learning-Based Cardiovascular Disease Detection Using Optimal Feature Selection," in IEEE Access. T. Ullah et al., "Machine Learning-Based Cardiovascular Disease Detection Using Optimal Feature Selection," in IEEE Access.
6. Ahmad, G. N., Ullah, S., Algethami, A., Fatima, H., & Akhter, S. M. H. (2022). Comparative study of optimum medical diagnosis of human heart disease using machine learning technique with and without sequential feature selection. iee access, 10, 23808-23828.
7. Fitriyani, N. L., Syafrudin, M., Alfian, G., & Rhee, J. (2020). HDPM: an effective heart disease prediction model for a clinical decision support system. IEEE Access, 8, 133034-133050.
8. Sarmah, S. S. (2020). An efficient IoT-based patient monitoring and heart disease prediction system using deep learning modified neural network. Ieee access, 8, 135784-135797.
9. Ghorashi, S., Rehman, K., Riaz, A., Alkahtani, H. K., Samak, A. H., Cherrez-Ojeda, I., & Parveen, A. (2023). Leveraging regression analysis to predict overlapping symptoms of cardiovascular diseases. IEEE Access.
10. Joo, G., Song, Y., Im, H., & Park, J. (2020). Clinical implication of machine learning in predicting the occurrence of cardiovascular disease using big data (Nationwide Cohort Data in Korea). IEEE Access, 8, 157643-157653.
11. Ashri, S. E., El-Gayar, M. M., & El-Daydamony, E. M. (2021). HDPF: heart disease prediction framework based on hybrid classifiers and genetic algorithm. iee access, 9, 146797-146809.
12. Baghdadi, N. A., Farghaly Abdelaliem, S. M., Malki, A., Gad, I., Ewis, A., & Atlam, E. (2023). Advanced machine learning techniques for cardiovascular disease early detection and diagnosis. Journal of Big Data, 10(1), 144.
13. Li, J. P., Haq, A. U., Din, S. U., Khan, J., Khan, A., & Saboor, A. (2020). Heart disease identification method using machine learning classification in e-healthcare. IEEE access, 8, 107562-107582.
14. Rossom, R. C., Hooker, S. A., O'Connor, P. J., Crain, A. L., & Sperl-Hillen, J. M. (2022). Cardiovascular risk for patients with and without schizophrenia, schizoaffective disorder, or bipolar disorder. Journal of the American Heart Association, 11(6), e021444.
15. Dalal, S., Goel, P., Onyema, E. M., Alharbi, A., Mahmoud, A., Algarni, M. A., & Awal, H. (2023). Machine learning for cardiovascular disease risk prediction. Computational Intelligence and Neuroscience, 2023, 1-12.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details