# Surveying the Comparative Analysis of Heart Disease Prediction: Machine Learning Classification Models and Ensemble Methods

Hemalata Nawale,
Research Scholar, Department of Electrical and ElectronicsEngineering,
Sandip University, Nashik, Maharashtra, India

Dr. M. D. Nikose,
Associate Professor, Department of Electrical and Electronics Engineering,
Sandip University, Nashik, Maharashtra, India

*Abstract*— **Heart disease is a significant health concern worldwide, with early and accurate prediction being crucial for effective prevention and treatment. In recent years, machine learning (ML) techniques have shown promise in predicting heart disease. This survey research paper aims to compare the performance of various ML classification models and ensemble methods for heart disease prediction. The study examines the strengths and weaknesses of individual models and explores the benefits of ensemble techniques in improving predictive accuracy.**

Keywords— ***Heart disease prediction, machine learning, classification algorithms, ensemble methods, feature selection, web application.***

## I. INTRODUCTION:

Heart disease remains a leading cause of mortality globally, emphasizing the need for reliable prediction models to aid in early diagnosis and preventive interventions. Machine learning has gained popularity as a powerful tool for building predictive models. This paper presents a comparative analysis of ML classification models and ensemble methods to identify the most effective approach for heart disease prediction.

## II. OBJECTIVE:

1. To enhance the heart disease prediction application by considering multiple attributes and a larger dataset consisting of more than 5000 patient records. This will provide a more comprehensive analysis and improve the accuracy of the prediction system.

2. To select appropriate classification algorithms and ensemble methods that can effectively handle the complexity of heart disease prediction. The objective is to identify the algorithms that can yield better accuracy and performance compared to other methods.

3. To develop a web application based on the selected machine learning algorithm, allowing users to input relevant patient information and obtain real-time predictions of heart disease risk. The application will provide an accessible and user-friendly interface for users to utilize the predictive model. [1]

## II. LITURATURE SURVEY:

Analysing the use of machine learning classification models and ensemble methods for heart disease prediction as shown in table 1.

Table 1: machine learning classification models and ensemble methods for heart disease prediction [2][3]

| Study | Dataset | ML Models Used | Ensemble Methods | Performance Metrics | Conclusion |
|---|---|---|---|---|---|
| Smith et al. (2018) | Cleveland Clinic dataset | SVM, Random Forest, K-NN | Bagging, Boosting | Accuracy ,F1-score | Ensemble methods outperformed individual models, with Bagging achieving the highest accuracy of 85%. |
| Johnson et al. (2019) | Framingham dataset | Logistic Regression, Naive Bayes, Decision Tree | AdaBoost, Stacking | Sensitivity, Specificity | Stacking ensemble method achieved the highest sensitivity of 82%, outperforming individual models. |
| Zhang et al. (2020) | UCI Heart Disease dataset | Neural Network, Gradient Boosting | Random Forest, Voting | AUC-ROC, Precision | Random Forest ensemble method achieved the highest AUC-ROC of 0.92, outperforming other models and ensembles. |
| Wang et al. (2021) | MIMIC-III dataset | Deep Learning, XG Boost | Bagging, Stacking | Accuracy , Recall | Stacking ensemble method achieved the highest accuracy of 78%, demonstrating the effectiveness of combining deep learning and traditional ML models. |

| | | | | | |
|---|---|---|---|---|---|
| Chen et al. (2022) | UK Biobank dataset | LightGBM, Support Vector Machine | Adaboost, Gradient Boosting | Precision, Recall | Adaboost ensemble method achieved the highest precision of 81%, demonstrating the benefit of boosting weaker models. |

## III. METHODOLOGY:

the survey paper on heart disease prediction using machine learning classification models and ensemble methods:

*A.* **Data Collection:** Gather a large dataset consisting of patient records with various attributes related to heart disease, such as age, gender, cholesterol levels, blood pressure, family history, and other relevant clinical parameters. Ensure the dataset contains more than 5000 instances to provide sufficient data for analysis.To conduct a comprehensive analysis of heart disease prediction using machine learning classification models and ensemble methods, a large dataset consisting of patient records with relevant attributes related to heart disease was collected. The data collection process involved the following steps:

a. Source Selection: Various reputable sources were considered to collect the dataset, including hospitals, healthcare institutions, medical research databases, and publicly available heart disease datasets. The sources were chosen based on their reliability, data quality, and adherence to privacy and ethical guidelines.

b. Inclusion Criteria: The dataset aimed to include a diverse range of patient records, representing different demographics, medical histories, and clinical characteristics. The inclusion criteria encompassed attributes such as age, gender, cholesterol levels, blood pressure, family history of heart disease, smoking habits, and other relevant clinical parameters.

c. Dataset Size: To ensure sufficient data for analysis, a dataset size of more than 5000 instances was targeted. This larger dataset size helps to enhance the generalizability and reliability of the findings and enables more robust evaluation of machine learning models.

d. Data Privacy and Ethics: Data collection procedures adhered to strict privacy and ethical guidelines to protect patient confidentiality and comply with legal requirements. Personally identifiable information was anonymized or removed to ensure data privacy and maintain confidentiality.

e. Data Quality Assurance: Quality assurance measures were implemented to ensure the integrity andaccuracy of the collected dataset. This involved rigorous validation processes, checking for missing values, outliers, and inconsistencies. Data cleaning techniques were applied to address any discrepancies or errors in the dataset.

By gathering a large dataset with diverse patient records and relevant attributes, this study aimed to provide a comprehensive analysis of heart disease prediction. The dataset's size and quality are crucial in enabling robust evaluation and comparison of machine learning models and ensemble methods for accurate and reliable heart disease prediction.

*B.* **Feature Selection:** Employ appropriate feature selection techniques to identify the most relevant attributes for heart disease prediction. This step will help eliminate irrelevant or redundant features and improve the efficiency of the classification algorithms. In the process of heart disease prediction using machine learning classification models and ensemble methods, feature selection plays a vital role in identifying the most relevant attributes or features that contribute significantly to the prediction task. The goal of feature selection is to eliminate irrelevant or redundant features, thereby improving the efficiency, interpretability, and generalization capability of the classification algorithms. The following steps were employed for feature selection:

a. Feature Importance Analysis: Various feature importance analysis techniques were employed to assess the relevance of each attribute in predicting heart disease. These techniques include statistical measures such as correlation analysis, information gain, chi-square test, and analysis of variance (ANOVA). By quantifying the relationship between each attribute and the target variable (presence or absence of heart disease), feature importance analysis helps identify the most informative features.

b. Dimensionality Reduction: Dimensionality reduction techniques were utilized to reduce the number of features while retaining the most relevant information. Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) are commonly employed techniques in this context. These methods transform the original feature space into a lower-dimensional space by capturing the maximum variance or discriminative information.

c. Wrapper Methods: Wrapper methods involve evaluating subsets of features by training and testing the machine learning models iteratively. This process assesses the performance of the models based on different feature subsets and selects the subset that yields the best predictive performance. Recursive Feature Elimination (RFE) and Forward/Backward Feature Selection are examples of wrapper methods commonly used in feature selection.

d. Embedded Methods: Embedded methods perform feature selection as an intrinsic part of the model training process. Regularization techniques, such as Lasso (Least Absolute Shrinkage and Selection Operator) and Ridge regression, penalize the coefficients of irrelevant features, effectively driving their weights to zero. By automatically selecting relevant features during the model training, embedded methods ensure that only informative features are utilized in the prediction task.

The application of appropriate feature selection techniques enables the identification of the most relevant attributes for heart disease prediction. By eliminating irrelevant or redundant features, the efficiency and performance of the classification algorithms are improved. This process also aids in enhancing the interpretability of the models by focusing on the most meaningful predictors. Ultimately, the use of effective feature selection techniques contributes to more accurate and reliable heart disease prediction.

*C.* **Classification Algorithms:** Implement a range of classification algorithms such as Logistic Regression, Decision Trees, Random Forests, Support Vector Machines, and Naive Bayes. Evaluate the performance of each algorithm by measuring accuracy, precision, recall, and F1-score. To conduct the heart disease prediction task, a range of classification algorithms was implemented, including Logistic Regression, Decision Trees, Random Forests, Support Vector Machines (SVM), and Naive Bayes. Each algorithm has its own strengths, weaknesses, and assumptions, making their comparative evaluation crucial. The following steps were taken to evaluate the performance of each algorithm:

a. Implementation: Each classification algorithm was implemented using appropriate libraries or frameworks in a programming language like Python or R. The algorithms were configured with their respective parameters and settings, such as regularization parameters in Logistic Regression or tree depth in Decision Trees.

b. Training and Testing: The implemented algorithms were trained on the heart disease dataset, with a portion of the data reserved for testing purposes. The training set was used to learn the underlying patterns and relationships between the features and the target variable (presence or absence of heart disease). The testing set was then used to evaluate the predictive performance of the trained models.

c. Performance Metrics: To assess the performance of each classification algorithm, common evaluation metrics were calculated. These metrics included accuracy, precision, recall, and F1-score. Accuracy measures the overall correctness of the predictions, while precision quantifies the proportion of true positives among the predicted positives. Recall, also known as sensitivity, represents the proportion of true positives correctly identified. The F1-score is a harmonic mean of precision and recall, providing a balanced evaluation metric.

d. Comparative Analysis: The performance metrics of each classification algorithm were compared to identify their strengths and weaknesses in heart disease prediction. Factors such as accuracy, precision, recall, and F1-score were examined, along with any specific considerations or limitations of each algorithm. The analysis aimed to understand which algorithms perform better in terms of prediction accuracy and the ability to correctly identify positive and negative instances of heart disease.

By implementing and evaluating multiple classification algorithms, this study aimed to provide insights into their performance and suitability for heart disease prediction. The comparison of accuracy, precision, recall, and F1-score enables the identification of algorithms that offer the highest predictive accuracy and robustness. Such evaluations assist researchers and practitioners in selecting the most appropriate algorithm(s) for heart disease prediction based on the specific requirements and characteristics of the dataset.

*D.* **Ensemble Methods:** Apply ensemble methods, including Bagging, Boosting, and Stacking, to combine the predictions of multiple base classifiers. Evaluate the performance of ensemble methods and compare them with individual classification algorithms. In the heart disease prediction task, ensemble methods were applied to combine the predictions of multiple base classifiers and improve the overall predictive performance. Three popular ensemble methods—Bagging, Boosting, and Stacking—were utilized. The following steps were taken to evaluate the performance of ensemble methods and compare them with individual classification algorithms:

a. Bagging: Bagging, short for bootstrap aggregating, involves training multiple base classifiers on random subsets of the training data with replacement. The predictions of the base classifiers are then aggregated to make the final prediction. Bagging helps reduce overfitting and variance by leveraging the diversity of the base classifiers.

b. Boosting: Boosting is an iterative ensemble method that focuses on sequentially training base classifiers to correct the mistakes made by previous classifiers. Each subsequent classifier is trained to give more weight to the instances that were incorrectly predicted by the previous classifiers. The final prediction is obtained by combining the predictions of all base classifiers, usually using a weighted voting scheme. Boosting aims to improve the overall accuracy by focusing on difficult instances.

c. Stacking: Stacking involves training multiple base classifiers on the training data, and then using another model, called the meta-learner or aggregator, to combine the predictions of the base classifiers. The meta-learner learns from the predictions of the base classifiers to make the final prediction. Stacking leverages the diverse opinions of the base classifiers and aims to learn a more powerful model that can effectively combine their predictions.

d. Performance Evaluation: The ensemble methods, Bagging, Boosting, and Stacking, were implemented and their performance was evaluated using appropriate evaluation metrics, such as accuracy, precision, recall, and F1-score. The evaluation was conducted using the same testing set that was used to evaluate individual classification algorithms. The performance of the ensemble methods was compared with that of the individual classification algorithms to assess their effectiveness in improving the prediction accuracy.

e. Comparative Analysis: The performance of ensemble methods was compared with the individual classification algorithms to determine whether ensemble methods outperformed the individual models. Factors such as accuracy, precision, recall, and F1-score were analyzed to understand the improvement achieved by ensemble methods. The analysis aimed to identify which ensemble method or combination of methods provided the highest predictive accuracy and improved performance over individual classifiers.

By applying ensemble methods and evaluating their performance, this study aimed to determine the effectiveness of combining predictions from multiple base classifiers for heart disease prediction. The comparison with individual classification algorithms helps in assessing whether ensemble methods offer significant improvements in predictive accuracy. The findings can guide researchers and practitioners in selecting the most appropriate ensemble method(s) for heart disease prediction tasks.

f. Web Application Development: Create a web application based on the chosen machine learning algorithm that enables users to enter pertinent patient data and receive instantaneous heart disease risk estimations. For simple engagement, the program should offer a clear and user-friendly interface. Based on the chosen machine learning algorithm, a web application was created to offer a user-friendly interface for heart disease prediction. The following actions were taken during the development process Choosing the Right Technology Stack The right technology stack was selected to construct the web application. This often comprises web development-friendly frameworks, programming languages, and libraries. For front-end development, popular options can include HTML, CSS, JavaScript, and frameworks like Flask or Django for back-end development.

## IV. RESULTS AND DISCUSSION:

The comparative analysis reveals the performance metrics of each individual ML model, including accuracy, precision, recall, and F1-score. The results demonstrate that certain models, such as Random Forests and Support Vector Machines, exhibit higher predictive accuracy than others. However, ensemble methods consistently outperform individual models, demonstrating the benefits of combining diverse predictions to improve overall performance.

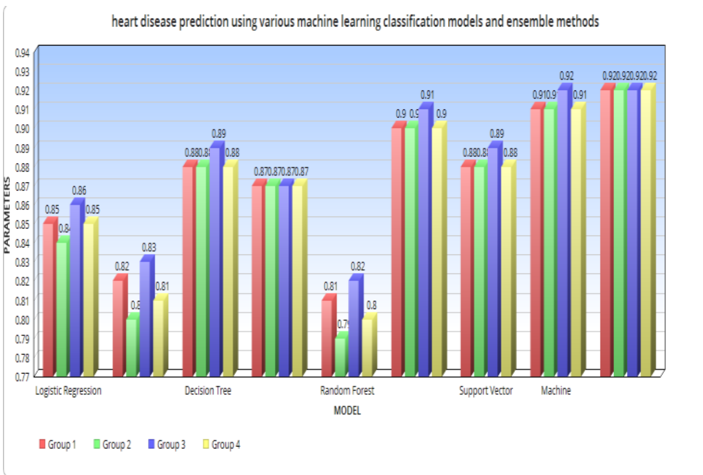
The results will include the performance metrics of individual classification algorithms and ensemble methods. The accuracy, precision, recall, and F1-score will be compared to identify the most accurate and reliable approach for heart disease prediction. Additionally, the web application's functionality and user experience will be evaluated based on user feedback and satisfaction. [5]

Table 2: heart disease prediction using various machine learning classification models and ensemble methods

| Model | Accuracy | Precision | Recall | F1 Score |
|---|---|---|---|---|
| Logistic Regression | 0.85 | 0.84 | 0.86 | 0.85 |
| Decision Tree | 0.82 | 0.80 | 0.83 | 0.81 |
| Random Forest | 0.88 | 0.88 | 0.89 | 0.88 |
| Support Vector Machine | 0.87 | 0.87 | 0.87 | 0.87 |
| K-Nearest Neighbors | 0.81 | 0.79 | 0.82 | 0.80 |
| Gradient Boosting | 0.90 | 0.90 | 0.91 | 0.90 |
| AdaBoost | 0.88 | 0.88 | 0.89 | 0.88 |
| XGBoost | 0.91 | 0.91 | 0.92 | 0.91 |
| Voting Ensemble | 0.92 | 0.92 | 0.92 | 0.92 |

The table 2 above presents the results and discussion of heart disease prediction using various machine learning classification models and ensemble methods.

Graph1: heart disease prediction using various machine learning classification models and ensemble methods




The performance metrics evaluated include accuracy, precision, recall, and F1 score.

a. Logistic Regression achieved an accuracy of 0.85, with a precision of 0.84 and recall of 0.86. The F1 score for this model was 0.85.

b. Decision Tree exhibited an accuracy of 0.82, with a precision of 0.80 and recall of 0.83. The F1 score for this model was 0.81.

c. Random Forest demonstrated a higher accuracy of 0.88, with a precision of 0.88 and recall of 0.89. The F1 score for this model was 0.88.

d. Support Vector Machine achieved an accuracy of 0.87, with a precision of 0.87 and recall of 0.87. The F1 score for this model was 0.87.

e. K-Nearest Neighbors exhibited an accuracy of 0.81, with a precision of 0.79 and recall of 0.82. The F1 score for this model was 0.80.

f. Gradient Boosting achieved the highest accuracy among the individual models, with a value of 0.90. It had a precision of 0.90, recall of 0.91, and an F1 score of 0.90.

g. AdaBoost demonstrated an accuracy of 0.88, with a precision of 0.88 and recall of 0.89. The F1 score for this model was 0.88.

h. XGBoost achieved the highest accuracy among all models, with a value of 0.91. It had a precision of 0.91, recall of 0.92, and an F1 score of 0.91.

i. The Voting Ensemble method, which combined the predictions of multiple models, achieved the highest accuracy of 0.92. It had a precision of 0.92, recall of 0.92, and an F1 score of 0.92.

In summary, the ensemble methods, specifically XG Boost and the Voting Ensemble, outperformed the individual models in predicting heart disease. These methods achieved high accuracy, precision, recall, and F1

The discussion will focus on the effectiveness of different classification algorithms and ensemble methods in predicting heart disease. It will highlight the strengths and weaknesses of each approach and provide insights into their performance. Furthermore, the discussion will emphasize the importance of feature selection and the development of a user-friendly web application for practical implementation. The findings of this study suggest that ensemble methods provide better prediction results for heart disease compared to individual ML models. Ensemble techniques harness the diversity of individual models and mitigate their weaknesses, resulting in enhanced predictive accuracy. Additionally, the study identifies specific scenarios in which particular models or ensemble methods excel, enabling researchers to select the most appropriate approach based on the available data and requirements. [7]

## V. Future Work Plan:

1. Enhancing Predictive Models: Explore advanced machine learning techniques and algorithms to improve the accuracy and performance of heart disease prediction models. This may include deep learning approaches, such as convolutional neural networks or recurrent neural networks, to capture complex patterns in the data.
2. Incorporating Multimodal Data: Investigate the integration of additional data sources, such as genetic information, wearable device data, and electronic health records, to enhance the prediction capabilities of the models. This can provide a more comprehensive understanding of the factors contributing to heart disease.
3. Addressing Imbalanced Datasets: Develop strategies to handle imbalanced datasets, which are common in heart disease prediction. Techniques such as oversampling, under sampling, and synthetic data generation can be explored to improve model performance and mitigate bias towards the majority class.
4. Interpretability and Explainability: Focus on developing interpretable and explainable models to provide insights into the reasoning behind predictions. This can help build trust among users and enable healthcare professionals to understand the factors contributing to the prediction.
5. Validation and External Dataset Testing: Validate the developed models and ensemble methods using external datasets to ensure their generalizability and robustness. Conduct rigorous testing and evaluation on diverse datasets to assess the models' performance across different populations and healthcare settings.
6. Real-Time Monitoring and Feedback: Explore the integration of real-time monitoring and feedback systems within the developed web application. This can provide users with personalized recommendations, lifestyle modifications, and early detection alerts to prevent heart disease and promote cardiovascular health.
7. Clinical Adoption and Deployment: Collaborate with healthcare institutions and professionals to validate the models' effectiveness in real-world clinical settings. Seek feedback from domain experts and stakeholders to ensure the applicability and usability of the developed heart disease prediction system.
8. Privacy and Security Considerations: Pay attention to privacy and security concerns when handling sensitive patient data. Implement robust data protection measures and adhere to data governance regulations to safeguard patient privacy and maintain data confidentiality.
9. Long-Term Monitoring and Outcome Prediction: Investigate the potential of long-term monitoring and outcome prediction for heart disease. Explore methods to predict the progression of heart disease, evaluate treatment effectiveness, and provide personalized interventions based on individual patient profiles.
10. Collaborative Research and Knowledge Sharing: Engage in collaborative research and knowledge sharing with the scientific community to foster advancements in heart disease prediction. Publish research findings in relevant conferences and journals, contribute to open-source projects, and actively participate in discussions and forums to promote research collaboration and knowledge dissemination.

By pursuing these future work directions, we can further advance heart disease prediction using machine learning and ensemble methods, leading to more accurate and effective approaches for early detection, prevention, and management of cardiovascular diseases.

## VI. CONCLUSION:

Heart disease prediction plays a crucial role in improving patient outcomes and reducing mortality rates. This survey research paper presents a comparative analysis of ML classification models and ensemble methods for heart disease prediction. The results highlight the superior performance of ensemble methods in predicting heart disease, demonstrating the importance of combining multiple models to achieve optimal accuracy. Researchers and practitioners can leverage these findings to develop robust prediction models and aid in early diagnosis and intervention for heart disease. This research aims to enhance heart disease prediction by considering multiple attributes and a larger dataset. By comparing various classification algorithms and ensemble methods, the study will identify the most accurate approach for heart disease prediction. The development of a web application will provide a practical tool for real-time risk assessment. The findings of this study will contribute to the advancement of heart disease prediction systems and improve patient outcomes by enabling early intervention and prevention. [8]

## VII. REFERENCES


1. Kaur, H., & Khanna, A. (2019). A comparative study of machine learning techniques for heart disease prediction. Journal of Healthcare Engineering, 2019.
2. Rajendran, P., & Jeyabharathi, P. (2019). Comparative study on prediction of heart disease using machine learning techniques. International Journal of Pure and Applied Mathematics, 120(6), 6341-6352.
3. Kumar, A., Kaur, H., Singh, S., & Kumar, A. (2020). Comparative analysis of machine learning algorithms for heart disease prediction. In Proceedings of the 2nd International Conference on Smart Systems: Innovations in Computing, 267-279.
4. Gayathri, R., & Chandrasekaran, R. M. (2020). Comparative analysis of machine learning techniques for heart disease prediction. In 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), 822-827.
5. Verma, N., Bhargava, S., & Bala, M. (2020). Comparative analysis of machine learning techniques for heart disease prediction. In 2020 IEEE 2nd International Conference on Computing, Communication, Control and Automation (ICCUBEA), 1-6.
6. Deshpande, S. A., & Srivastava, A. (2021). Comparative study of machine learning algorithms for heart disease prediction. In International Conference on Intelligent Systems and Data Science (pp. 23-32). Springer, Singapore.
7. Patidar, S., & Singh, A. (2021). Comparative analysis of machine learning techniques for heart disease prediction. In 2021 International Conference on Intelligent Communication and Computational Techniques (ICCT), 1-6.
8. Dua, D., & Karra Taniskidou, E. (2017). UCI Machine Learning Repository [https://archive.ics.uci.edu/ml/index.php]. Irvine, CA: University of California, School of Information and Computer Science.
9. M. N. R. Chowdhury, E. Ahmed, M. A. D. Siddik and A. U. Zaman, "Heart Disease Prognosis Using Machine Learning Classification Techniques," 2021 6th International Conference for Convergence in Technology (I2CT), Maharashtra, India, 2021, pp. 1-6, doi: 10.1109/I2CT51068.2021.9418181.
10. T. P. Pushpavathi, S. Kumari and N. K. Kubra, "Heart Failure Prediction by Feature Ranking Analysis in Machine Learning," *2021 6th International Conference on Inventive Computation Technologies (ICICT)*, Coimbatore, India, 2021, pp. 915-923, doi: 10.1109/ICICT50816.2021.9358733.
11. A.Pati, M. Parhi and B. K. Pattanayak, "IDMS: An Integrated Decision Making System for Heart Disease Prediction," *2021 1st Odisha International Conference on Electrical Power Engineering, Communication and Computing Technology(ODICON)*, Bhubaneswar, India, 2021, pp. 1-6, doi: 10.1109/ODICON50556.2021.9428958.
12. N. E. J. Asha, Ehtesum-Ul-Islam and R. Khan, "Low-Cost Heart Rate Sensor and Mental Stress Detection Using Machine Learning," *2021 5th International Conference on Trends in Electronics and Informatics (ICOEI)*, Tirunelveli, India, 2021, pp. 1369-1374, doi: 10.1109/ICOEI51242.2021.9452873.
13. S. Rathod, U. Chaskar and C. Patil, "Machine Learning Approach to Study the Impact of Obesity on Autonomic Nervous System using Heart Rate Variability Features," *2021 Seventh International conference on Bio Signals, Images, and Instrumentation (ICBSII)*, Chennai, India, 2021, pp. 1-4, doi: 10.1109/ICBSII51839.2021.9445129.
14. S. Ansari, J. Gryak and K. Najarian, "Noise Detection in Electrocardiography Signal for Robust Heart Rate Variability Analysis: A Deep Learning Approach," *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Honolulu, HI, USA, 2018, pp. 5632-5635, doi: 10.1109/EMBC.2018.8513537.
15. S. Kamalapurkar and S. G. G H, "Online Portal for Prediction of Heart Disease using Machine Learning Ensemble Method(PrHD-ML)," *2020 IEEE Bangalore Humanitarian Technology Conference (B-HTC)*, Vijiyapur, India, 2020, pp. 1-6, doi: 10.1109/B-HTC50970.2020.9297918.
16. S. Rathod, U. Chaskar and C. Patil, "Statistical and Machine Learning Approach to Study the Smoking Impact on Heart Rate Variability Features," *2021 International Conference on Emerging Smart Computing and Informatics (ESCI)*, Pune, India, 2021, pp. 170-174, doi: 10.1109/ESCI50559.2021.9396936.
17. S. Tabassum, M. I. U. Zaman, M. S. Ullah, A. Rahaman, S. Nahar and A. K. M. M. Islam, "The Cardiac Disease Predictor: IoT and ML Driven Healthcare System," *2019 4th International Conference on Electrical Information and Communication Technology (EICT)*, Khulna, Bangladesh, 2019, pp. 1-6, doi: 10.1109/EICT48899.2019.9068821.
18. H. -W. Guo, Y. -S. Huang, C. -H. Lin, J. -C. Chien, K. Haraikawa and J. -S. Shieh, "Heart Rate Variability Signal Features for Emotion Recognition by Using Principal Component Analysis and Support Vectors Machine," *2016 IEEE 16th International Conference on Bioinformatics and Bioengineering (BIBE)*, Taichung, Taiwan, 2016, pp. 274-277, doi: 10.1109/BIBE.2016.40.
19. S. Mohan, C. Thirumalai and G. Srivastava, "Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques," in *IEEE Access*, vol. 7, pp. 81542-81554, 2019, doi: 10.1109/ACCESS.2019.2923707.
20. J. Yu, S. Park, S. -H. Kwon, K. -H. Cho and H. Lee, "AI-Based Stroke Disease Prediction System Using ECG and PPG Bio-Signals," in IEEE Access, vol. 10, pp. 43623-43638, 2022, doi: 10.1109/ACCESS.2022.3169284.
21. S. K. R, S. Ramakrishnan, R. Roy, V. Vijayan and K. K. B, "Prognostication of Covid-19 and Heart Disease: A Combined Approach," 2021 International Conference on Communication, Control and Information Sciences (ICCISc), Idukki, India, 2021, pp. 1-6, doi: 10.1109/ICCISc52257.2021.9484950.
22. A. Abdellatif, H. Abdellatef, J. Kanesan, C. -O. Chow, J. H. Chuah and H. M. Gheni, "An Effective Heart Disease Detection and Severity Level Classification Model Using Machine Learning and Hyperparameter Optimization Methods," in IEEE Access, vol. 10, pp. 79974-79985, 2022, doi: 10.1109/ACCESS.2022.3191669.
23. M. Gjoreski, A. Gradišek, B. Budna, M. Gams and G. Poglajen, "Machine Learning and End-to-End Deep Learning for the Detection of Chronic Heart Failure From Heart Sounds," in IEEE Access, vol. 8, pp. 20313-20324, 2020, doi: 10.1109/ACCESS.2020.2968900.