

A Case Study with 8ML Algorithms for Heart Disease Diagnosis

Sujata S Kokil¹ and Sachin Bhoite²

¹Research Scholar, Dr. Vishwanath Karad, MIT World Peace University, Pune, India
Email: sujata.kokil@mitwpu.edu.in

²Dr. Vishwanath Karad, MIT World Peace University, Pune, India
Email: sachintenjuly@gmail.com

Abstract—Now a days, Heart Disease (HD) has become a significant worldwide challenge, creating a need for early detection and intervention. Machine learning (ML) techniques are used effectively for assisting, decision making, and prediction in the healthcare industry. The classification algorithms, have demonstrated promising capabilities in predicting the risk of HD based on patient data. This paper provides a detailed overview of various ML algorithms for early detection of the risk of HD and focus on the study of eight algorithms namely "Logistic Regression", "Naive Bayes", "Support Vector Machine", "K-Nearest Neighbors", "Decision Tree", "Random Forest", "XGBoost", and "Neural Network". It also discusses a comparative study of their performance measures. Additionally, it explores the integration of feature selection and ensemble methods to increase predictive accuracy and generalization. The paper actually worked on these eight algorithms and observed that the accuracy of Random Forest Model – 95.08% is found to be the best accuracy amongst all. This paper aims to assist the researchers for selecting appropriate approaches and methodologies for the studies related to HD. The paper concludes with suggestions for future research directions aimed at advancing the field of HD prediction using ML algorithms.

Index Terms— HD Prediction, Classification Algorithms, Machine Learning, Feature Selection, Evaluation Metrics, Ensemble Methods.

I. INTRODUCTION

HD is a widespread concern worldwide, affecting millions across different ages, genders, and regions. Considering its prevalence and seriousness is very important for finding effective ways to tackle it. [17]As mentioned everywhere "Prevention is better than cure", the detection and prevention at early stage are key weapons to fight against HD. [1][7]Often, heart issues don't show symptoms until they become severe. So, finding a way to catch them early is important. Equally, it is essential to recognize the risk factors and take preventive measure to stop HD before it starts. [2][7][8] In today's digital era of technological advancement, the convergence of healthcare and machine learning leads to the best results to fight against heart diseases. [1][16] Machine learning is like a super-smart computer program that learns from data, previous experiences and make predictions. It is incredibly helpful in identifying who might be at risk of heart problems and providing the help in advance. [2][4] Unlike traditional methods, which focus on just a few factors, machine learning considers many variables at once. This gives doctors a much clearer picture of who might be in danger and what they can do to help. [14] As a result, predictive models fueled by machine learning offer personalized risk assessments,

and help clinicians to come up with the interventions based on the unique needs of each patient. But machine learning doesn't just predict who might get heart disease; it's also great at spotting early signs of trouble. [3][9] By looking at various types of information, like medical records and wearable devices, it can pick up on small changes that could signal a risk. This allows doctors to take action early, preventing serious issues down the line. This research work emphasizes on the applications of classification algorithms for understanding the performance measures for heart disease diagnosis. [1][2][7] Machine learning predictive models include a large variety of algorithms and many researchers used the algorithms such as NB, LR, ANN, SVM, DT, K-NN, and RF etc. to identify the risk of HD. Furthermore, the role of machine learning extends beyond risk prediction to encompass early detection strategies. [10] As per the reports of Institute of Health Technology Transformation (iHT2), the big data coming from healthcare industry has reached around the size of Zettabytes and growing towards the size in Yottabyte scale. The technological development in health industry led to the generation of heterogeneous dataset, offering opportunities to enhance healthcare services. [11] By analyzing disparate data sources that includes patients health records, commonly known as EHR, imaging data, wearable technology, and genetic information — machine learning algorithms can detect subtle deviations indicative of preclinical heart disease. These early warning signals enable doctors to make better decisions, catch the problems early, and keep people healthy.

II. ARTICLE SELECTION

For this review paper, we thoroughly investigated several databases and research papers like as IEEE Explore Digital Library, Google Scholar, Science Direct, PubMed and Springer. The search included the keywords like Heart disease (HD), Machine Learning (ML), Prediction, Classification algorithms from the year 2015 to 2024. A total of 74 research papers were initially identified. A refined selection process was carried out to use review papers (literature review), conference papers, and Journal Papers and the search was limited to heart disease prediction and classification algorithms and its implementation only and excluded the articles based on implementation of AI based techniques, Image processing, Big Data. Finally, 20 articles were chosen and studied meticulously for a detailed study to uncover the gaps and motivation for conducting the research.

Along with the study of research articles and journals, different performance measure metrics are also used for classifiers performance evaluation. This study used eight classification algorithms namely, "Logistic Regression(LR)", "Naive Bayes(NB)", "Support Vector Machine(SVM)", "K-Nearest Neighbors(KNN)", "Decision Tree(DT)", "Random Forest(RF)", "XGBoost(XGB)", and "Neural Network(NN)" and studied the performance measures like Accuracy, Precision, F-measure and Recall using heart dataset available on online resource, the Kaggle repository. At last the study concluded with performances of these algorithms and will be useful for novice researchers as ready reckoner, compares the performance of the algorithms and motivates the researchers to work on heart disease detection. The research study has the following contributions-

- Provides a comprehensive review of algorithms employed for prediction of heart diseases by studying eight different ML algorithms on a dataset in a single research study and concludes with comparative study of its performance metrics.
- The study suggests further research areas in the field of Medical Diagnosis and assist the experts to improve healthcare services.

The remaining part of paper includes –Section 2 discusses about the literature review. The dataset, the concepts of feature selection, machine learning algorithms and performance measures are discussed in section 3. The results and detailed analysis of the experiment are summarized in Section 4, while Section 5 explores the conclusions and future aspects of the research study.

III. LITERATURE REVIEW

As mentioned in many of the research studies, healthcare sector is generating a voluminous data with a variety that includes structured, semi-structured and unstructured data, aiming at improved healthcare quality. [12][17][19] Using specialized technologies like Big Data Analytics, machine learning (ML), deep learning and natural language processing (NLP) helps significantly to enhance the processing and analysis of healthcare data. These advanced tools empower medical professionals to gain valuable insights into various aspects of patient treatment and results and disease trends. [11] This research study presents existing ML algorithms to explain the prediction of the HDs. The primary objective of this study is to get familiar with the terminologies, to apply the classification techniques on existing datasets and understand the performance measuring metrics. [1] Senthilkumar and Thirumalai discussed about DT, SVM, NB alongwith a new technique Hybrid Random Forest

with Linear Model (HRFLM) and compared the performance of the newly introduced HRFLM with earlier classification algorithms. [6] M. S. Al Reshan et al. proposed a prediction system to predict heart disease, Hybrid Deep Neural Networks (HDNNs) that combined diverse neural network structures to extract important features from the heart data and was utilized to leverage feature learning capabilities and non-linear methodologies, aiming to extract intricate patterns and their correlations in Heart datasets, thereby achieving enhanced accuracy. [2] N. L. Fitriyani and colleagues introduced a robust model, the heart disease prediction model (HDPM), that incorporates Density-Based Spatial Clustering of Applications with Noise (DBSCAN) for outlier detection and elimination. [9] Ekta Maini et al studied five algorithms for developing a HD prediction system and suggests that the ML based prediction could serve as screening tool for heart disease diagnosis in primary healthcare centers in India to achieve the sustainable development goals given by WHO. [5] Vaishali Baviskar et al. developed a model H2EDCO to help medical practitioners to diagnose heart diseases better. This model combines different methods like stackingCV classifier, bagging, feature boosting, and augment multiple sets to analyze health data and give more accurate results about heart health. [19] Therefore, it is very essential for the new researchers to first understand the traditional algorithms for the detection and prediction of the diseases.

IV. METHODS AND MATERIALS

A. Machine Learning Algorithms

Machine Learning (ML) algorithms plays an important role in disease diagnosis. These algorithms contribute to the organization of patients into distinct categories based on factors such as symptoms, medical backgrounds, and diagnostic data. Figure 1 provides an overview of the model's workflow. Initially, the data is gathered from the data sources, transformed into a dataset and applied the pre-processing techniques. Then, the model is generated, and the results are subsequently analyzed.

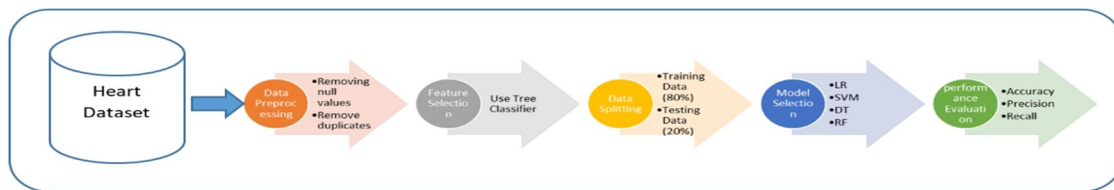


Figure 1: Basic Flow Of Operations Performed on Heart dataset

B. About The Dataset

Heart Disease dataset was sourced from the Kaggle Database, which is widely used to explore, analyze, and share quality data for studying different machine learning algorithms. The dataset comprises of 13 attributes and one output (target variable) class. It consists of 303 records in all with all number and nominal values. The consideration was to study binary classification, and hence we excluded image data, multiclass data and considered only simple dataset. The target variable or output class has 165 entries labelled as the presence (positive class) and remaining 138 entries, labelled as absence (negative class) of heart disease. It has been observed that the heart attack occurs due to severely reduced or blocked flow of blood towards the heart. When there is a deposition of cholesterol, fat and other substances in the heart or more specifically Coronary Arteries, it turns into heart attack. The exploratory data analysis performed concludes with following results-

- 45.54 % of patient have no heart problems and 54.46% patient suffers from heart problem.
- Number of Male facing heart problems are more than that of female.
- Individuals experiencing typical angina chest pain are at a higher risk of heart problems compared to those with asymptomatic chest pain.
- People with resting blood pressure levels categorized as '1' and '0' exhibit a significantly higher likelihood of having heart disease compared to those classified as '2'.
- The people experiencing exercise-induced angina are significantly less prone to develop heart problems.
- The presence of a slope value '2' is associated with a considerably higher likelihood of heart pain compared to slope values '0' and '1'.

V. RESULTS AND DISCUSSION

The presented study was evaluated on the Kaggle dataset (UCI Repository) for heart disease and was used for binary classification. Various algorithms were used to test the performance metrics of each one. For binary

classification, the UCI dataset was composed of 303 records, segmented into an 80:20 ratio. Among these, 242 records were used to train the model, and 61 records were reserved for testing and validation.

Specifically, the binary classification performance, including accuracy, is detailed in Table II, showcasing the accuracy comparison among various models employed for prediction. Comparative analysis of all these algorithms reveals their strengths and weaknesses for predicting heart disease risk. Integration of feature selection techniques and ensemble methods leads to improved model performance and robustness.

Table I represents existing work done by the researchers and Table II represents the study of the 8ML algorithms with the accuracy after running the studied models.

In this study, the actual implementation is performed in Jupyter notebook using Python libraries – scikitlearn, pandas, matplotlib, seaborn, xgboost etc. The paper studied eight prediction models using 13 distinct features and assesses their accuracy across different modeling techniques. Table III represents the studied algorithms with its use and accuracy. Table III outlines a comparison across metrics such as Precision (P), F-measure (F-score), recall(R) and support(S). The Random Forest (RF) classification method emerges with the highest accuracy 95.08% among the existing methods scrutinized in the analysis. The results obtained from these algorithms are data driven and subject to the data size. The performance evaluation metrics observed during this experiment are characterized in Table III.

TABLE I ASSESSMENT OF THE PROPOSED APPROACH IN RELATION TO THE EXISTING RESEARCH BY THE RESEARCHERS

Study	Model	Dataset	Language/Libraries used for classification	Performance metrics
2023[12]	LR, DT, SVM, RF, KNN, GNB	Cleveland Heart Disease dataset (303 instances), Long Beach VA(200), Switzerland (123), Hungarian (294)	Python Libraries	Accuracy (91.8), Recall (91.4), F1 score (91.9), & Precision(92.5) (KNN)
2023[17]	SVM, RF, XGBoost, KNN, SHAP and LIME technique (Proposed)	Cleveland dataset (303 instances)	Python Libraries	F1 Score (88) (SVM +XGBoost)
2023[18]	SVM , LR	Heart and CHD Dataset	Python Libraries	Accuracy (82%)(LR)
2023[19]	MLP, RF, DT, XGBoost	a Kaggle dataset for cardiovascular disease (70,000 instances)	Google Colab Python Libraries	Accuracy(87.02), Precision (89.62), Recall(82.11), F1-score(86.16) XGB
2023 [5]	DNN, RF, XGBoost, SVC, AdaBoost, RNN+GA, RNN+ PSO, LSTM+GA, H2EMDCO (Proposed)	UCI dataset (binary class) Cleveland Heart Disease dataset (303 instances) (multiclass)	Python Libraries	Accuracy(93.44) for binary class Accuracy(53.33) for Multiclass
2022[20]	RSS-KNN (Proposed)	Cleveland, Stalog, and UCI ML Repository	MATAB 2018	Precision(98.95), Recall(99.81), F1-Score(97.24)
2021[9]	KNN, NB, LR, AdaBoost, RF	Actual data collected from South Indian hospitals	Python Libraries	Accuracy (93.8) (RF)
2020[2]	LR, NB, DT, MLP, SVM, RF, HDPM (Proposed)	Statlog and Cleveland dataset;	Python Libraries	Accuracy (95.90), Precision(97.14), F-Measure (95.35) for HDPM
2020[15]	LR, KNN, ANN, DT, SVM(Linear), NB, FCMIM-SVM (proposed)	Cleveland Heart Disease dataset (303 instances)	Python Libraries	Accuracy(92.37) FCMIM-SVM(proposed)
2019 [1]	NB, Generalized Linear Model(LM), LR, DT, RF, GB Trees, SVM, Deep Learning, HRFLM (Proposed)	UCI heart dataset (303 records)	R Studio Rattle	Accuracy (88.4), Precision(90.1), F-Measure (90) for HRFLM
Present Study	LR, NB, SVM, KNN, DT, RF, XGBoost and NN	UCI heart dataset (303 records)	Python Libraries	Accuracy (95.08%) (RF)

As per the results of [12], the study explains the experimentation on six ML algorithms (classifiers) and KNN classifier showed the highest performance with accuracy 91.8%, recall 91.4%, f1-score 91.9% and a precision 92.5%. [17] Putri Sari Asih et al, concluded that their study developed and evaluated multiple models to predict the HDs, they included RF, SVM, XGBoost, and k-NN. As per their study, SVM and XGBoost gives a better predictive accuracy as compare to rest of the models experimented based on f1-score – 88%. The paper [18] experimented on the datasets, heart dataset and CHD dataset, compared the accuracy by using the SVM and LR (82% Accuracy-best one) and Identified the heart disease status of patients through classification. [19] Focused on HD prediction on large datasets comprising of 70000 records, studied the algorithms RF, DT, Multilayer

TABLE II. 8ML ALGORITHMS AND ACCURACY MEASURES ACHIEVED

Sr No	Classification Algorithm	Use of the algorithm	Accuracy score achieved
1	Logistic Regression	Useful for binary classification Provides comprehensive data and explains the relationship between a dependent binary variable and one or more nominal, ordinal, interval, or ratio-level independent variables.	85.25%
2	Naïve Bayes	useful for classification problem is used when there is a boolean type of dependent or target variable present in the dataset	85.25%
3	Support Vector Machine	used for classification, regression and outlier detection	81.97%
4	K-Nearest Neighbour	Helpful for classifications or predictions concerning the grouping of an individual data item.	67.21%
5	Decision Tree	Employed for addressing classification problems and sorting objects based on their learned features.	81.97%
6	Random Forest	used to handle classification and regression problems ensures the most accurate predictions	95.08%
7	XGBoost	used for classification, regression, ranking, and even user-defined prediction challenges	78.69%
8	Neural Network	Is a robust tool in computer science and artificial intelligence, enabling the rapid classification and clustering of data.	80.33%

Perception (MLP), and XGBoost, along with hyper parameter tuning to identify the presence of HD. All these classifiers achieved an accuracy above 86.5%, the accuracy enhanced after hyperparameter tuning with GridSearchCV to 86.90%. The studied algorithms indicated the MLP Model achieving the highest accuracy as 87.23%. The finding from this research demonstrated the potential of using k-modes clustering for predicting the HD accurately and suggested that MLP can be used as a valuable tool to develop the diagnostic system. Vaishali Baviskar et al [5], used the UCI and Cleveland dataset, worked to increase the accuracy of the classifiers by proposing a “Hyper Ensemble Model for Evaluation of Heart Diseases from Clinical Observations (H2EMDCO)” model to predict the HD. The model was evaluated for binary classification and multiple classification, and observed that their model performed well around 3.4% better than the existing models, achieved an accuracy of 93.44% and 60.5% for single and multiple disease class. The model used a combination of ensemble classification and feature augmentation procedure. Kannan S [20], proposed an ensemble model with Random subspace and K-nearest neighbor (RSS-KNN) for early prediction of HD. The model achieved an accuracy of 97.65% (without features), accuracy 98.56 (with features) and precision of 98.95%, recall of 99.81%, F1-score of 97.24%, and AUROC of 97.62%. The values of Precision and recall of the UCI ML repository dataset are 2% higher as compare to that of Cleveland and Statlog datasets, using MATLAB environment.

[9] Ekta Maini used actually collected data from South India comprising of 1670 records to study the performance of 5ML algorithms and suggested that the RF is best performing algorithms attaining an accuracy of 93.08%. The study suggested the potential use of ML algorithms in HD prediction systems. [2] Heart Disease Prediction Model (HDPm) proposed by N. L. Fitriyani and team, integrated DBSCAN, SMOTE-ENN and XGBoost to improve the accuracy in predicting the HD. The model used two datasets, Statlog and Cleveland and achieved an accuracy 95.10% for dataset I, 98.40% for dataset II, precision 97.14% and 98.57%, f-score 95.35 and 98.32%. [15]J. P. Li et al proposed ML based diagnostic system that used the six ML classifiers and four feature selection algorithms and proposed a new FCMIM model to solve the problem of feature selection resulting in an accuracy of 92.37% for SVM with FCMIM. S. Mohan et al [1], introduced a new technique “Hybrid Random Forest with Linear Model (HRFLM)”, that used three important association rules such as apriori, predictive and Tertius used to find the risk factors of HDs using UCI dataset. The performance metrics produced accuracy is 88%, Sensitivity- 92.8%, Specificity- 82.6, precision-82.5%, F-measure-90%.

The present study with 8ML algorithms resulted into RF as the best performing algorithm with accuracy 95.08%, precision 96%, recall-95%, F1-score -95% based on a dataset with 13 attributes and 303 records. One of the important observation about this study is that the results obtained are data driven and are depending on the algorithm that is selected, the size and diversity of the dataset and the platform where we run the models etc., hence the generalization is needed. The study suggests that these traditional algorithms further can be combined with advanced techniques such as feature selection, hyperparameter tuning and so on to improve the predictive accuracy and contributing in the development of diagnosis system. The presented study founds to be a foundational study to the researchers.

TABLE III: EVALUATION METRICS OF 8ML ALGORITHMS OBSERVED DURING MODEL EXECUTION

Logistic Regression					Naïve Bayes				
	precision	recall	f1-score	support		precision	recall	f1-score	support
0	0.85	0.81	0.83	27	0	0.88	0.78	0.82	27
1	0.86	0.88	0.87	34	1	0.84	0.91	0.87	34
avg / total	0.85	0.85	0.85	61	avg / total	0.85	0.85	0.85	61
Support Vector Machine					K Nearest Neighbors				
	precision	recall	f1-score	support		precision	recall	f1-score	support
0	0.83	0.74	0.78	27	0	0.62	0.67	0.64	27
1	0.81	0.88	0.85	34	1	0.72	0.68	0.70	34
avg / total	0.82	0.82	0.82	61	avg / total	0.68	0.67	0.67	61
Decision Tree					Random Forest				
	precision	recall	f1-score	support		precision	recall	f1-score	support
0	0.79	0.81	0.80	27	0	0.90	1.00	0.95	27
1	0.85	0.82	0.84	34	1	1.00	0.91	0.95	34
avg / total	0.82	0.82	0.82	61	avg / total	0.96	0.95	0.95	61
XGBoost					Neural Network				
	precision	recall	f1-score	support		precision	recall	f1-score	support
0	0.75	0.78	0.76	27	0	0.78	0.78	0.78	27
1	0.82	0.79	0.81	34	1	0.82	0.82	0.82	34
avg / total	0.79	0.79	0.79	61	avg / total	0.80	0.80	0.80	61

VI. CHALLENGES AND LIMITATIONS

The challenges in HD prediction include data imbalance issues, model interpretability, and generalization across diverse patient populations. Addressing these challenges requires interdisciplinary collaboration and innovative research approaches.

VII. FUTURE DIRECTIONS

Present study is a basic overview of the classification algorithms. Future research directions include integrating advanced ML techniques including deep learning, reinforcement learning, exploring personalized medicine approaches, and developing real-time monitoring systems to improve the efficiency of prediction system in heart disease diagnosis. Emphasis should be placed on improving model interpretability and generalizability to enhance clinical applicability.

VIII. CONCLUSION

This paper provides an overview and summary of different methods employed for predicting and detecting heart disease. It highlights the significance of machine learning algorithms in this context, as they serve as essential tools for predicting heart disease and facilitating early detection and intervention. The paper focuses on the importance of ongoing research endeavors to improve predictive models and optimize their effectiveness within the healthcare industry, aiming to elevate the quality of healthcare facilities. Moreover, it suggests directing

research efforts towards optimization techniques and performance tuning methods to advance the efficiency of these algorithms towards the correct diagnosis of heart disease.

REFERENCES

- [1] Senthilkumar Mohan, Chandrasegar Thirumalai, and Gautam Srivastava, “ Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques”, IEEE Access Digital Object Identifier 10.1109/Access.2019.2923707
- [2] Norma Latif Fitriyani, Muhammad Syafrudin. Ganjar Alfian, And Jongtae Rhee, “HDPM: An Effective Heart Disease Prediction Model For A Clinical Decision Support System”, IEEE Access Digital Object Identifier 10.1109/Access.2020.3010511
- [3] Yuanyuan Pan, Minghuan Fu, Biao Cheng, Xuefei Tao, and Jing Guo, “Enhanced Deep Learning Assisted Convolutional Neural Network For Heart Disease Prediction On The Internet Of Medical Things Platform”, IEEE Access, Digital Object Identifier 10.1109/Access.2020.3026214
- [4] Shah Nazir, Sulaiman Khan, Habib Ullah Khan, Shaukat Ali, Iván García-Magariño, Rodziah Binti Atan, And Muhammad Nawaz, “A Comprehensive Analysis Of Healthcare Big Data Management, Analytics And Scientific Programming”, IEEE Access, Digital Object Identifier 10.1109/Access.2020.2995572
- [5] Vaishali Baviskara, Aditya Ojhab, Ishaan Vermab, Sahajpal Mahlab, Madhushi Vermab, Pradeep Chatterjeec, Sunil Kumard, “H2EMDCO: Design of A Hyper Ensemble Model For Evaluation Of Heart Diseases From Clinical Observations”, Sciencedirect, Procedia Computer Science 218 (2023) 1288–1294
- [6] Mana Saleh Al Reshan, Samina Amin , Muhammad Ali Zeb, Adel Sulaiman , Hani Alshahrani, and Asadullah Shaikh, “A Robust Heart Disease Prediction System Using Hybrid Deep Neural Networks”, IEEE Access Digital Object Identifier 10.1109/ACCESS.2023.3328909, Volume 11, 2023
- [7] Manabu Torii, Jung-wei Fan, Wei-li Yang, Theodore Lee, Matthew T. Wiley, Daniel S. Zisook, Yang Huang, “Risk Factor Detection For Heart Disease By Applying Text Analytics In Electronic Medical Records”, Journal of Biomedical Informatics 58 (2015) S164–S170
- [8] Chunyan Guo, Jiabing Zhang, Yang Liu, Yaying Xie, Zhiqiang Han and Jianshe Yu, “Recursion Enhanced Random Forest With an Improved Linear Model (RERF-ILM) for Heart Disease Detection on the Internet of Medical Things Platform”, IEEE Access Digital Object Identifier 10.1109/ACCESS.2020.2981159
- [9] Ekta Maini, Bondu Venkateswarlu, Baljeet Maini, Dheeraj Marwaha, “Machine learning based heart disease prediction system for Indian population: An exploratory study done in South India”, Medical Journal Armed Forces India 77 (2021) 302–311, Elsevier
- [10] Fatima Hussain, Muhammad Nauman, Abdullah Alghuried, Adi Alhudhaif, and Nadeem Akhtar, “Leveraging Big Data Analytics for Enhanced Clinical Decision-Making in Healthcare”, IEEE Access Digital Library.
- [11] Awais Ahmed, Rui Xi, Mengshu Hou, Syed Attique Shah, and Sufian Hameed, “Harnessing Big Data Analytics for Healthcare : A comprehensive Review of Frameworks, Implications, Applications, and Impacts”, IEEE Access Digital Library, 2023, Digital Object Identifier 10.1109/ACCESS.2023.3323574
- [12] Ghulam Muhammad, Saad Naveed, Lubna Nadeem, Tariq Mahmood, Amjad R. Khan, Yasar Amin, And Saeed Ali Omer Bahaj, “Enhancing Prognosis Accuracy for Ischemic Cardiovascular Disease Using K Nearest Neighbor Algorithm: A Robust Approach”, IEEE Access Digital Library, 2023, Digital Object Identifier 10.1109/ACCESS.2023.3312046
- [13] Subashish Mohapatra, Sushree Maneesha, Prashanta Kumar Patra, Subhadarshin Mohanty, “Heart Diseases Prediction based on Stacking Classifiers Model”, ScienceDirect, Procedia Computer Science 218 (2023) 1621–1630
- [14] Ali Rizwan, Ahmed Zoha, Rui Zhang, Wasim Ahmad, Kamran Arshad, Najah Abu Ali , Akram Alomainy, Muhammad Ali Imran And Qammer H. Abbasi, “A Review On The Role Of Nano-Communication In Future Healthcare Systems: A Big Data Analytics Perspective”, IEEE Access Digital Library, Digital Object Identifier 10.1109/ACCESS.2018.2859340
- [15] Jian Ping Li, Amin Ul Haq, Salah Ud Din, Jalaluddin Khan, Asif Khan, And Abdus Saboor, “Heart Disease Identification Method Using Machine Learning Classification In E-Healthcare”, IEEE Access Digital Library, Digital Object Identifier 10.1109/Access.2020.3001149
- [16] Tobias Gentner, Timon Neitzela, Jacob Schulze , Felix Gerschner, Andreas Thessler, “Data Lakes in Healthcare: Applications and Benefits from the Perspective of Data Sources and Players”, ScienceDirect, Procedia Computer Science 225 (2023) 1302–1311
- [17] Putri Sari Asih, Yufis Azhara, Galih Wasis Wicaksona, Denar Regata Akbi, “Interpretable Machine Learning Model For Heart Disease Prediction”, ScienceDirect, Procedia Computer Science 227 (2023) 439–445
- [18] Ankit Kumar, Kamred Udham Singh , And Manish Kumar “A Clinical Data Analysis Based Diagnostic Systems For Heart Disease Prediction Using Ensemble Method”, BIG DATA MINING AND ANALYTICS ISSN 2096-0654 10/10 Pp 513– 525 Volume 6, Number 4, December 2023 DOI: 10.26599/BDMA.2022.9020052
- [19] Chintan M. Bhatt, Parth Patel, Tarang Ghetia and Pier Luigi Mazzeo, “Effective Heart Disease Prediction Using Machine Learning Techniques”, MDPI 2023, 16, 88. <https://doi.org/10.3390/a16020088>
- [20] Kannan S., “An automated clinical decision support system for predicting cardiovascular disease using ensemble learning approach”, Concurrency Computat Pract Exper. 2022; 34(18):e7007. doi: 10.1002/cpe.7007