University of Essex

School of Mathematics, Statistics
and Actuarial Science

MA981 DISSERTATION

HEART DISEASE PREDICTION USING COMBINED ML AND DL MODELS

KIRTHIK RAMADOSS [2213380]

Supervisor: **Sithole,Terry**

March 20, 2024

Colchester

Contents

1 INTRODUCTION **2**

1.1 Overview of Heart Disease as a Health Concern 2

1.2 Significance of Early Detection in Heart Disease 3

1.3 Issue articulation 3

1.4 Aim 4

1.5 Investigative query 4

1.6 Objectives 4

1.7 Contemporary Issues 5

1.8 Feasibility Evaluation 6

2 LITERATURE REVIEW **7**

2.1 Need for heart disease prediction 7

2.2 Recent Research in heart disease prediction using combined ML and DL models 7

2.3 Linkage to aim 13

3 METHODOLOGICAL STRATEGY **14**

3.1 Selection of algorithms in predicting heart disease 14

3.2 Software Requirements 15

3.3 Investigative Techniques 16

3.4 Dataset description 16

3.5 Insights into Collected Data 17

3.6 Removing Null and Duplicates 21

3.7 Visualizing the Data 22

3.8 Splitting the data for Implementation 25

3.9 Algorithms used in predicting heart disease 25

3.9.1 Decision Tree . . . 25
3.9.2 Support Vector Machine . . . 27
3.9.3 Random Forest . . . 29
3.9.4 Multilayer Perceptron . . . 30
3.9.5 Voting Classifier . . . 31
3.10 Assessment Measurement . . . 33

4 EXPERIMENTAL STUDY AND OUTCOMES ANALYSIS 36
4.1 Implementing Individual ML Heart Disease Prediction Models . . . 36
4.1.1 Evaluating Individual ML Models for Dataset-1 . . . 37
4.1.2 Evaluating Individual ML Models for Dataset-2 . . . 42
4.1.3 Computational Complexity of Individual Models . . . 46
4.2 Implementing Combined ML Heart Disease Prediction Models . . . 47
4.2.1 Evaluating Combined ML Models for Dataset-1 . . . 49
4.2.2 Evaluating Combined ML Models for Dataset-2 . . . 50
4.2.3 Computational Complexity of Combined Models . . . 52
4.3 In-Depth Analysis of Models for Predicting Heart Disease . . . 53

5 CONCLUSIONS 57
5.1 Thorough analysis . . . 57
5.2 Achievement of Research Objectives . . . 58
5.3 Conclusion . . . 58
5.4 Guidance for future research . . . 59

List of Figures

2.1 Results of [19] 9
2.2 Results of [23] 12

3.1 Head and Tail View of Dataset 18
3.2 Columns in Heart disease Datasets 19
3.3 Descriptive statistics of Dataset 2 21
3.4 Count Plot - Heart Disease Prediction 22
3.5 Count Plot - Heart Disease Prediction 22
3.6 Line Plot regarding Columns 23
3.7 Line Plot regarding Columns 24
3.8 Data Splitting 25
3.9 The Decision Tree Structure [31] 26
3.10 Structure of SVM[34] 27
3.11 Representatipn of Random Forest[37] 29
3.12 Structure of Multilayer Perceptorn[40] 30
3.13 Structure of Voting Classifier[42] 32

4.1 Tested parameters 37
4.2 Selected Parameters for dataset 1 37
4.3 Validation Matrix-RF 40
4.4 Testing Matrix-RF 41
4.5 Selected parameters for dataset 2 42
4.6 Validation Matrix-RF 45
4.7 Testing Matrix of RF and DT 46
4.8 Combined ML Modles 47
4.9 Configuration for Dataset 1 48

4.10 Configuration for Dataset 2 48
4.11 Best performing combo modles for dataset 1 49
4.12 Testing Matrix-SVM and RF 50
4.13 Validation Matrix-MLP and DT 51
4.14 Testing Matrix-MLP and DT 52
4.15 Individual Vs Combined Outstanding Modles 56

List of Tables

2.1 Results of [17] 8
2.2 Result of [18] 9
2.3 Results of [20] 10
2.4 Results of [22] 11
2.5 Results of [24] 12

3.1 Removing Null and Duplicates 21
3.2 Confusion Matrix 35

4.1 Validation Metrics for Dataset 1 40
4.2 Testing Metrics for Dataset 1 41
4.3 Validation Matrix-Dataset 2 44
4.4 Testing Matrix-Dataset 2 45
4.5 Time Taken by Individual Models for Dataset 1 46
4.6 Time Taken by Individual Models for Dataset 2 47
4.7 Validation Metrics - Dataset 1 49
4.8 Testing Metrics - Dataset 1 50
4.9 Validation Metrics - Dataset 2 51
4.10 Testing Metrics - Dataset 2 52
4.11 Time Taken by Combined Models for Dataset 1 53
4.12 Time Taken by Combined Models for Dataset 2 53
4.13 Comparison of Individual ML Modles for Dataset 1 54
4.14 Comparison of Individual ML Modles for Dataset 2 54
4.15 Comparison of Combined ML Modles for Dataset 1 54
4.16 Comparison of Combined ML Modles for Dataset 2 55
4.17 Comparison of Performance Duration for Dataset 1 55

4.18 Comparison of Performance Duration for Dataset 2 56

Abstract

Heart disease continues to be one of the major diseases around the world and need to be identified early to save lives. On the other side, this kind of complexity in which the symptoms are varied and the risks are different always poses challenges to fast diagnosis, management, and early-stage recognition which makes predictive models in healthcare even more important. The goal of this project will be to apply individual machine learning (ML) and deep learning (DL) algorithms and then combining together for the prediction of coronary heart disease in order to increase both the model's reliability and effectiveness. Data from both the UCI Heart Disease Dataset and Heart Disease Dataset from Kaggle were thoroughly researched and then collected, preprocessed. Various ML algorithms, such as Decision Trees, SVM, Random Forests, and MLP are implemented on the datasets and the results show in almost all cases that the Random Forest has the highest prediction accuracy, compared to the other techniques considered for the two datasets. Significantly, in combination of ML and DL models, SVM & RF indicated the highest result of 85% accuracy in the dataset 1, while in dataset 2, MLP & DT performed best with the result of 83% accuracy. Taking into account the time required in Dataset 1, the SVM & RF combination takes training and testing time within range 0.14s to 0.34s, but in the test of Dataset 2, the MLP & DT combination has faster computing time with training times in 0.14s to 0.34s and validation/test in 0.01s to 0.04s. These findings underscore the importance of carefully combining ML and DL models tailored to each dataset for accurate heart disease prediction.

CHAPTER 1

INTRODUCTION

The human heart is the most vital organ because it distributes oxygen-rich blood throughout the body via the circulatory system. cardiac disease refers to any condition that causes abnormal cardiac function. There are numerous forms of heart disorders such as congenital heart disease, coronary artery disease, arrhythmia, etc [1]. Several risk factors make it challenging to diagnose heart disease. Human heart disease severity has been determined using a number of data mining and neural network methods[2].

1.1 Overview of Heart Disease as a Health Concern

Heart failure (HF) is a serious global health problem that affects over 26 million individuals in high-income countries (HICs) and low- and middle-income countries (LMICs). Increased mortality, hospitalizations, and healthcare costs are only some of the negative outcomes linked with heart failure. The incidence of HF is nearly nine times greater in those aged 65 and above[3]. It seems from the current evidence that HF is linked with significant morbidity and death, protracted hospitalization, and a heavy social cost. Epidemiological changes and rising rates of risk factors like hypertension, diabetes mellitus, dyslipidemia, and obesity as well as changes in behavior like less exercise, more alcohol consumption, and tobacco use are driving the rising tide of heart failure[4]. Health systems are ill-equipped to handle the communicable and non-communicable diseases (NCDs), and cardiovascular diseases (CVDs) such heart failure (HF) are a prominent challenge[11].

1.2 Significance of Early Detection in Heart Disease

Globally, cardiovascular disease has garnered a reputation as one of the most lethal chronic ailments. When the cardiac system is afflicted with illness, it often exhibits inadequate pumping capacity, resulting in compromised blood circulation to various bodily organs, so impeding their normal physiological functions[5]. Coronary artery disease (CAD) stands as the primary etiology behind the development of heart failure. The cardiovascular system is dependent on the continuous supply of oxygenated blood provided by the coronary arteries. Physical debility, dyspnea, edema of the lower extremities, weariness, and associated symptoms, among others, are the most often seen indications of cardiovascular disease[6]. Heart rhythm problems, congestive heart failure, congenital heart disease, and cardiovascular disease (CVD) include further manifestations of heart illness. The first use of conventional investigative techniques for the detection and identification of heart ailments encountered significant challenges[7]. The diagnosis and treatment of heart disease pose significant challenges due to limited access to diagnostic tools and a shortage of well qualified medical personnel, especially in less developed countries. Nevertheless, obtaining an accurate diagnosis of cardiovascular disease in its early stages is of paramount importance in order to mitigate potential problems[8]. The prevalence of mortality due to cardiovascular disease is rapidly increasing in both developed and developing countries. The annual increase in the number of victims is substantial[9]. Based on a study conducted by the European Society of Cardiology (ESC), it has been determined that a total of 26.5 million persons are affected by heart disease, and an annual increment of 3.8 million new cases is seen [10]. Early identification of predicting heart disease is therefore important to reduce the death rate[13].

1.3 Issue articulation

In order to predict heart disease, combination of machine learning (ML) as well as deep learning (DL) models can be used by using various patient data that can include demographics, personal and medical history, diagnostic results, lifestyle and genetic factors. DL will only allow deep level representations of data. Therefore, there is a

need for blending traditionally ML algorithms with DL architectures to increase the predictive precision as well as interpretability. The approach involves data gathering and integration, deriving insightful features and then model development, training and evaluation. Also, the planning and conducting of the project will require pledging to ethic and regulatory aspects to pursue patients' privacy, data security and the compliance of healthcare regulations. Finally, the main problem to be sorted is to predict heart disease by combining ML and DL algorithms in an efficient way.

1.4 Aim

The aim of the project is to build a model that predicts heart diseases with the help of combining both machine learning and deep learning algorithms in order to obtain higher accuracy with a minimum amount of time.The project compares how the combination works on different datasets of heart disease.

1.5 Investigative query

What is the effect of combined machine learning and deep learning models performing on different heart disease datasets in terms of accuracy and time?

1.6 Objectives

- To conduct a detailed background research for the heart disease prediction, role of ML and DL algorithms, how combinations work in increasing accuracy and previous methods and techniques used.
- To collect and preprocess 2 datasets containing heart disease data, ensuring the data is devoid of errors, accurately labeled, and correctly scaled.
- To implement individual machine learning algorithms, including Decision Tree, Support Vector Machines (SVM), Random Forests, and MLP classifier for heart disease prediction.
- To implement combinations of the above-mentioned algorithms in order to identify how the combinations predict heart diseases with different datasets.

- To evaluate and compare the precision and accuracy of integrated machine learning algorithms with individual algorithms for heart disease prediction.
- To determine the most effective combined machine learning algorithm for the detection of heart diseases.

1.7 Contemporary Issues

The hurdle of applying ML and DL combined models simultaneously in the framework of heart disease prediction is the capacity to utilize diverse data sources and still be able to have a scalable and interpretable model[13]. By blending various data groups like medical records, genetic information, and way of life as well as the use of machine learning systems, feature engineering techniques and data processing of quality are required[11]. These require methods of managing incomplete values, noise, and class imbalance in the right way. Additionally, one of the key issues of merging the ML-DL systems is the interpretability and transparency of the final combined model use in the clinical settings which should be highly valued. And above all that, data privacy, security and ethic concerning the use of sensitive medical data in these systems make the development of these models a bit complicated to be achieved[12].

1.8 Feasibility Evaluation

The technical feasibility of the proposed project to predict heart diseases by using combination of machine learning (ML), and deep learning (DL) algorithms for enhanced accuracy and efficiency, is high. There are enough publications about heart disease prediction and ML/DL algorithms, which makes conduct a comprehensive background research easy. It is very simple to conduct collection and preprocessing of two Kaggle datasets using Python, and there are plenty of resources that can help. Python's established libraries are used to implement individual ML algorithms (Decision Tree, SVM, Random Forests, MLP classifier) and combining them. Measures of evaluation such as precision, accuracy, and execution time can be calculated, and selection of the most appropriate combined algorithm is moderately feasible through systematic comparison. Python is used mostly in machine learning and deep learning tasks, and suitable documentation tools are available, the technical execution of the project is deemed possible, meaning that the development and evaluation of the predicted heart disease model can be successful.

CHAPTER 2

LITERATURE REVIEW

2.1 Need for heart disease prediction

Heart disease, a globally prevalent cause of death, frequently manifests inconspicuous symptoms, making timely detection crucial for saving lives. Predictive algorithms and risk assessment instruments offer a means to pinpoint individuals at elevated risk, enabling proactive interventions and lifestyle adjustments[13]. Furthermore, in the face of increasing cardiovascular disease rates and the strain on healthcare resources, accurate prediction facilitates resource allocation, alleviating the economic load linked to advanced-stage heart disease management. Leveraging the capabilities of data and technology, predictive models usher in a transition toward preventive, individualized healthcare, thereby enhancing patient results and elevating the general standard of healthcare provision[14].

2.2 Recent Research in heart disease prediction using combined ML and DL models

In[17] research paper, an innovative approach that employs machine learning methods to identify crucial features is presented, thereby enhancing the accuracy of cardiovascular disease prediction. The heart disease data used in this work was sourced from the UCI machine learning repository, specifically the Cleveland database. This partic-

ular database was chosen owing to its extensive usage among the machine learning research community, since it is known for its comprehensive and thorough data records. The predictive model is constructed using various feature combinations and multiple established classification techniques. Through the utilization of the hybrid random forest with a linear model (HRFLM), an improved predictive performance achieving an accuracy of 88.4% is attained in heart disease prediction.

Models	**Accuracy**	**Classification**	**Precision**	**F-Measure**	**Sensitivity**	**Sepcificity**
Naive Bayes	75.8	24.2	90.5	84.5	79.8	60.0
Generalized Linear Model	85.1	14.9	88.8	91.6	94.9	20.0
Logistic Regression	82.9	17.1	89.6	90.2	91.1	25.0
Deep Learning	87.4	12.6	90.7	92.6	95	33.3
Decision Tree	85	15.0	86	91.8	98.8	0.0
Random Forest	86.1	13.9	87.1	92.4	98.8	10.0
Gradient Boosted Trees	78.3	21.7	94.1	86.8	80.7	60.0
Support Vector Machine	86.1	13.9	86.1	92.5	100	0.0
VOTE	87.41	12.59	90.2	84.4	-	-
HRFLM(proposed)	**88.4**	**11.6**	**90.1**	**90**	**92.8**	**82.6**

Table 2.1: Results of [17]

[18] study centers on the diagnosis of heart disease by leveraging historical data and information. To accomplish this, the researchers have developed SHDP (Smart Heart Disease Prediction) using the Naive Bayesian approach to forecast heart disease risk factors. The primary focus of this research is to establish SHDP, a Smart Heart Disease Prediction system that combines NB classification and AES encryption for heart disease prediction. The necessary data is collected and standardized. Utilizing the UCI dataset, medical information from patients with heart diseases is gathered. To predict a patient's likelihood of heart disease, several attributes are extracted from their medical profiles, including age, blood pressure, cholesterol levels, gender, and blood sugar, among others. These collected attributes serve as input for Naive Bayesian classification to predict heart disease. The results reveal that, in terms of accuracy, the proposed technique outperforms Naive Bayes, achieving an accuracy rate of 89.77%, despite reducing the number of attributes.

S.NO	No.of Techniques	Accuracy(%)	Time(s)
1	Sequential Minimal Optimization(SMO)	84.07	0.02
2	Bayes Net(BN)	81.11	0.02
3	Multi Layer Perception(MLP)	77.4	0.75
4	Navies Bayesian(NB)	89.77	0.01

Table 2.2: Result of [18]

In [19] research, the aim was to discover machine learning classifiers that exhibit the highest accuracy in achieving accurate heart disease predictions. They applied various supervised machine-learning algorithms, comparing their performance and accuracy in predicting heart disease. Feature importance scores were estimated for each feature in all the applied algorithms, excluding MLP and KNN. The study utilized a heart disease dataset sourced from Kaggle and conducted three-classification using KNN, DT, and RF algorithms. Notably, the RF method achieved a remarkable 100% accuracy, demonstrating both 100% sensitivity and specificity. This discovery highlights that a relatively straightforward supervised machine learning algorithm can be employed to make highly accurate predictions regarding heart disease, showcasing its significant potential utility.

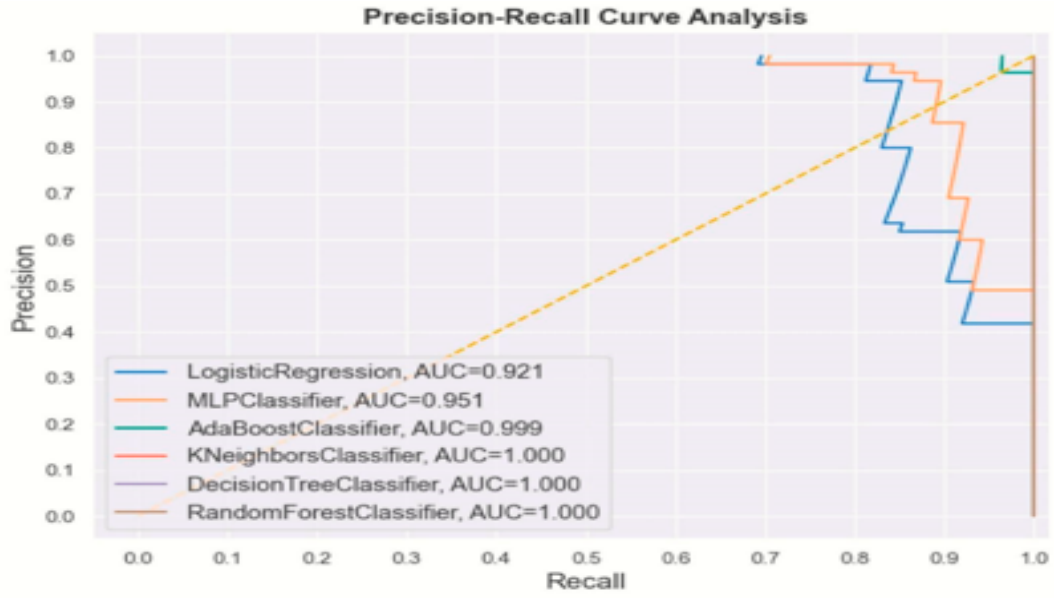

Figure 2.1: Results of [19]

[20] research introduces an innovative machine learning approach for heart disease prediction. The study makes use of the Cleveland heart disease dataset and employs data mining techniques, including regression and classification. Notably, ML techniques, specifically RF and DT, are applied to this dataset. The novel machine learning model is carefully crafted, and the implementation involves the use of three distinct machine learning algorithms: RF, DT and and a Hybrid model that combines RF and DT. The experimental outcomes demonstrate an impressive accuracy level of 88.7% when utilizing the hybrid model for heart disease prediction. Furthermore, an intuitive user interface is designed to collect input parameters from users for heart disease prediction, and this interface harnesses the power of the hybrid model, which combines Decision Tree and Random Forest, to make predictions.

Algorithm	**Accuracy(%)**
Decision Tree	79
Random Forest	81
Hybrid(Decision Tree+Random Forest)	88

Table 2.3: Results of [20]

[21] article conducts an in-depth analysis of various machine learning techniques to evaluate their efficiency in predicting, diagnosing, and treating diverse heart diseases. The study delves into a variety of machine learning approaches used for predicting heart disease, including SVM, DT, NB, KNN and ANN, among others. Subsequently, the research computes the average prediction accuracy for each method to determine their comparative performance. The results indicate that the ANN achieved the highest average prediction accuracy, reaching 86.91%. In contrast, the C4.5 decision tree technique demonstrated the lowest average prediction accuracy, with a value of 74.0%.

[22] paper, to improve the accuracy of heart disease assessment, an Internet of Things (IoT) framework has been introduced. This framework employs a Modified Deep CNN. Patients wear a smartwatch and a heart monitoring device, continuously monitoring

their blood pressure and ECG data. The MDCNN is used to classify the sensor data into normal or abnormal readings. For the assessment of the system's performance, the proposed MDCNN is compared against established DL NN and LR techniques. The results demonstrate that the heart disease prediction system based on the proposed MDCNN surpasses other methods. Furthermore, the findings indicate that, for the majority of records, the MDCNN achieves an accuracy rate of 98.2%, outperforming existing classifiers.

Models	ACC	Error	Precision	F1	Recall	Specificity
DLNN	83.8	16.2	89.6	90.2	91.1	60.2
LR	88.3	11.7	90.1	90	92.8	74.2
MDCNN	98.2	1.8	95.1	95	97.8	92.6

Table 2.4: Results of [22]

[23]research aim to improve the accuracy of heart disease prediction using machine learning techniques. Six distinct algorithms, including RF, KNN, LR, NB, GB, Adaboost are employed. Two datasets, one from Cleveland and the other from IEEE Dataport, serve as sources of data. To optimize model accuracy, GridsearchCV and five-fold cross-validation strategies are implemented. Notably, in the Cleveland dataset, logistic regression outperforms other algorithms with a remarkable accuracy of 90.16%, while AdaBoost shines in the IEEE Dataport dataset, achieving an accuracy of 90%. Further enhancing accuracy, a soft voting ensemble classifier that combines all six algorithms results in an accuracy of 93.44% for the Cleveland dataset and 95% for the IEEE Dataport dataset.

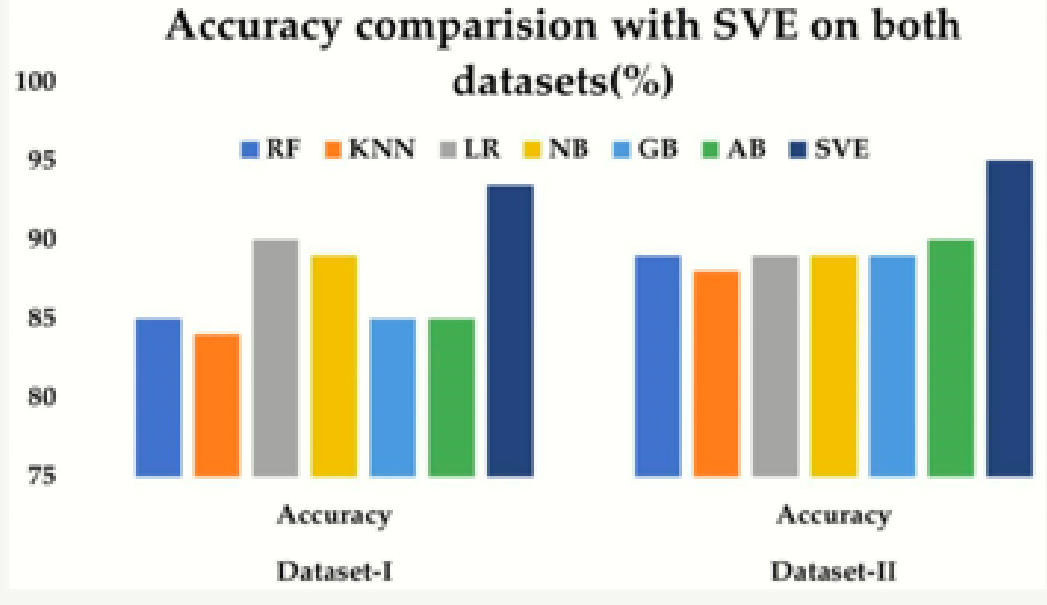

Figure 2.2: Results of [23]

[24] paper introduces an innovative approach involving the utilization of k-modes clustering with Huang initialization to improve classification accuracy. The research employs a variety of ML models, including RF, DT, MLP, and XGB. To fine-tune the models, hyperparameter tuning is conducted through GridSearchCV. The proposed methodology is applied to a real-world dataset comprising 70,000 instances sourced from Kaggle. Notably, the proposed models exhibit impressive AUC values, with DT achieving 0.94, XGBoost at 0.95, RF reaching 0.95, and MLP also scoring 0.95. The research concludes that, in terms of accuracy, the MLP with cross-validation outperforms all other algorithms, achieving the highest accuracy rate of 87.28%.

	Accuracy		Precision		Recall		F1-Score		AUC
Model	Without CV	CV	Without CV	CV	Without CV	CV	Without CV	CV	
MLP	86.94	87.28	89.03	88.70	92.95	84.85	85.88	86.71	0.95
RF	86.92	87.05	88.82	89.42	83.46	83.43	85.91	86.32	0.95
DT	86.53	86.37	90.10	89.58	81.17	81.61	85.40	85.42	0.94
XGB	87.02	86.87	89.62	88.93	82.11	83.57	86.30	86.18	0.95

Table 2.5: Results of [24]

Mohan[17] aimed to enhance the accuracy of cardiovascular disease prediction by utilizing the UCI machine learning repository and implementing a hybrid random forest with a linear model (IRFLM), achieving an 88.7% accuracy rate. Repaka[18] focused on

the diagnosis of heart disease by analyzing historical data from the UCI dataset with a Naive Bayes algorithm, attaining an 88.9% rate of accuracy.

Ali[19] achieved a perfect accuracy rate of 100% by employing KNN, DT, and RF algorithms on a Kaggle dataset for cardiovascular disease predictions. Kashitij[20] introduced an innovative ML approach for heart disease prediction using the Cleveland heart disease dataset and a hybrid model that combines various algorithms, resulting in an 88.7% accuracy rate.

Riyaz[21] set out to diagnose heart disease with a dataset comprising Indian patients, utilizing algorithms like KNN, SVM, DT, and ANN, and achieved an 89.1% accuracy rate. Lastly, Khan[22] aimed to improve the accuracy of heart disease assessment using blood pressure and electrocardiogram (ECG) data from a modified deep convolutional neural network (MDCNN), which led to an accuracy rate of 98.2%.

2.3 Linkage to aim

The above literature review deals with various models for heart disease prediction by integrating both machine learning and deep learning algorithms to enhance accuracy and minimize processing time. It evaluates the performance of these models across various heart disease datasets. However, a notable research gap lies in the limited exploration of the impact of these combined models on diverse heart disease datasets in terms of accuracy and efficiency. This gap implies a need for comprehensive investigations into how the combined machine learning and deep learning algorithms adapt to different data sources, potentially yielding variable results. Further research could shed light on the adaptability and effectiveness of these integrated models in real-world scenarios and the factors that influence their performance with diverse datasets.

CHAPTER 3

METHODOLOGICAL STRATEGY

Heart disease is a critical problem that can be anticipated using both individual and combined machine learning methods. The algorithms were explored singly and in combination as hybrid models for this goal. The data is gathered and cleaned in order to be implemented effectively. The precision, accuracy, recall, and f1 score results from the algorithms during validation and testing, along with how fast they can do calculations, help find the best model for predicting heart disease.

3.1 Selection of algorithms in predicting heart disease

Selecting appropriate algorithms for predicting heart disease involves considering factors such as data characteristics, model interpretability, computational efficiency, and predictive performance. Given the aim to combine both machine learning (ML) and deep learning (DL) algorithms for higher accuracy and efficiency, a diverse set of algorithms can be considered. Here's a selection of algorithms suitable for predicting heart disease:

1. Decision Trees (DF):

- Decision trees are capable of handling both categorical and numerical data.
- They're interpretable, capable of capturing non-linear relationships, and can be easily visualized[25].

2. Support Vector Machine (SVM):

- SVM is a powerful algorithm suitable for both linear and non-linear classification tasks.
- It works well with high-dimensional data and can capture complex decision boundaries.
- SVMs are effective when dealing with smaller datasets but may require tuning for optimal performance[26].

3. Random Forest (RF):

- Random Forests algorithm builds numerous decision trees during training for a more reliable forecast by assembling the outputs[27].

4. Multilayer Perceptron (MLP):

- MLP is capable of learning complex patterns from data and can capture non-linear relationships effectively.
- MLPs can be trained using gradient-based optimization techniques, such as back-propagation, to minimize prediction errors[28].

5. Voting Classifier (VC):

- The Voting Classifier uses the majority voting mechanism to combine the predictive power of Decision Tree, SVM, and Random Forests and therefore improves overall accuracy in heart disease prediction[29].

Experimenting with the above-selected algorithms and combinations therefore will help identify the most effective approach for predicting heart disease in different datasets.

3.2 Software Requirements

Software:

- Programming Language: Python
- Development Environment: Jupyter Notebook
- Data Science Platform: Anaconda

3.3 Investigative Techniques

Dataset Collection and Preprocessing: Gather two heart disease datasets from Kaggle, namely the UCI Heart Disease Dataset and Heart Disease Dataset. Ensure that the collected datasets are devoid of errors, accurately labeled, and correctly scaled. Preprocess the datasets to handle missing values, normalize numerical features, and encode categorical variables as necessary.

Implementation of Individual ML Algorithms: Implement various individual ML algorithms such as Decision Tree, Support Vector Machines (SVM), Random Forests, and Multilayer Perceptron (MLP) classifier for heart disease prediction. Train these algorithms on the collected datasets separately.

Implementation of Combined ML and DL Models Combine the aforementioned ML algorithms with DL models to construct hybrid models for heart disease prediction. Experiment with different combinations of ML and DL algorithms to identify the most effective hybrid model for detecting heart diseases.

Evaluation and Comparison: Evaluate the precision and accuracy of both individual ML algorithms and combined ML-DL models using appropriate evaluation metrics such as accuracy, precision, recall, and F1-score. Compare the performance of these models on different heart disease datasets to assess their effectiveness in terms of accuracy and time efficiency.

Determination of Most Effective Model: Determine the most effective combined ML-DL model for heart disease detection based on the evaluation results. Select the model that achieves the highest accuracy with minimal computational time.

3.4 Dataset description

Dataset 1

The dataset available on Kaggle is a multivariate dataset containing 14 attributes related to various factors associated with heart disease. These attributes include age, sex, chest pain type, resting blood pressure, serum cholesterol, fasting blood sugar, resting electrocardiographic results, maximum heart rate achieved, exercise-induced angina, ST

depression induced by exercise relative to rest (oldpeak), the slope of the peak exercise ST segment, number of major vessels, and Thalassemia.

Dataset 2

The heart disease dataset available on Kaggle is widely utilized in the training of machine learning models for the purpose of predicting occurrences of heart disease. The dataset comprises a total of 76 variables, encompassing several categories such as demographic information, medical history, and clinical measurements. These attributes were collected for a sample size of 303 patients. Indicating whether or not the patient has heart disease, the target variable is a binary variable. The dataset is partitioned into two distinct subsets: a training set and a test set. The training dataset consists of 202 individuals, whereas the test dataset comprises 101 individuals. The training set is utilized for the purpose of training the machine learning model, whereas the test set is employed to assess the performance of the trained model.

3.5 Insights into Collected Data

The acquired dataset 1 comprises 1025 rows and 14 columns, while the acquired dataset 2 comprises 920 rows and 16 columns. The first and last three columns of each dataset are shown in the table below, which serves as confirmation.

Dataset 1

Head and Tail

	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal	target
0	52	1	0	125	212	0	1	168	0	1.0	2	2	3	0
1	53	1	0	140	203	1	0	155	1	3.1	0	0	3	0
2	70	1	0	145	174	0	1	125	1	2.6	0	0	3	0
1022	47	1	0	110	275	0	0	118	1	1.0	1	1	2	0
1023	50	0	0	110	254	0	0	159	0	0.0	2	0	2	1
1024	54	1	0	120	188	0	1	113	0	1.4	1	1	3	0

Dataset 2

Head and Tail

	id	age	sex	dataset	cp	trestbps	chol	fbs	restecg	thalch	exang	oldpeak
0	1	63	Male	Cleveland	typical angina	145.0	233.0	True	lv hypertrophy	150.0	False	2.3
1	2	67	Male	Cleveland	asymptomatic	160.0	286.0	False	lv hypertrophy	108.0	True	1.5
2	3	67	Male	Cleveland	asymptomatic	120.0	229.0	False	lv hypertrophy	129.0	True	2.6
917	918	55	Male	VA Long Beach	asymptomatic	122.0	223.0	True	st-t abnormality	100.0	False	
918	919	58	Male	VA Long Beach	asymptomatic	NaN	385.0	True	lv hypertrophy	NaN	NaN	
919	920	62	Male	VA Long Beach	atypical angina	120.0	254.0	False	lv hypertrophy	93.0	True	

Figure 3.1: Head and Tail View of Dataset

When looking at the information in the datasets, it is seen that Dataset 1 has 1025 entries organized into 14 columns, and a lot of them are integer data types. The data structure consists of columns that pertain to cardiovascular conditions, including age, sex, and cholesterol levels. Dataset 2, comprising 16 columns and 920 entries, comprises a variety of data types, some of which contain absent values.

Dataset -1	`Index(['age', 'sex', 'cp', 'trestbps', 'chol', 'fbs', 'restecg', 'thala 'exang', 'oldpeak', 'slope', 'ca', 'thal', 'target'], dtype='object')`
Dataset -2	`Index(['id', 'age', 'sex', 'dataset', 'cp', 'trestbps', 'chol', 'fbs', 'restecg', 'thalch', 'exang', 'oldpeak', 'slope', 'ca', 'thal', 'nu dtype='object')`

Figure 3.2: Columns in Heart disease Datasets

Dataset -1

- 'age': Age of the patient
- 'sex': Sex of the patient (1 = male, 0 = female)
- 'cp': Chest pain type (1 = typical angina, 2 = atypical angina, 3 = non-anginal pain, 4 = asymptomatic)
- 'trestbps': Resting blood pressure
- 'chol': Serum cholesterol (in mg/dl)(milligrams/deciliter)
- 'fbs': Fasting blood sugar (1 = true, 0 = false)
- 'restecg': Resting electrocardiographic results (0 = normal, 1 = having ST-T wave abnormality, 2 = showing probable or definite left ventricular hypertrophy)
- 'thalach': Maximum heart rate achieved
- 'exang': Exercise-induced angina (1 = yes, 0 = no)
- 'oldpeak': ST depression induced by exercise relative to rest
- 'slope': Slope of the peak exercise ST segment (1 = upsloping, 2 = flat, 3 = downsloping)
- 'ca': Number of major vessels colored by fluoroscopy

- 'thal': Thalassemia (3 = normal, 6 = fixed defect, 7 = reversible defect)
- 'target': Diagnosis of heart disease (0 = no disease, 1 = disease)

  The dtype='object' specifies that the data type of the index is object (string).

Dataset -2

- 'id': Unique identifier for each instance
- 'age', 'sex', 'cp', 'trestbps', 'chol', 'fbs', 'restecg', 'thalach', 'exang', 'oldpeak', 'slope', 'ca', 'thal': Same as explained above
- 'dataset': Indicator of the dataset (1 or 2)
- 'nu': Additional feature (purpose not specified)

  The purpose of creating these indexes is to facilitate data manipulation, analysis, and modeling tasks by providing a convenient way to access and refer to specific columns in the dataset.

As shown in the 3.3. The document comprises data pertaining to cardiovascular ailments, delineated in columns such as age and cholesterol. It is necessary to pay attention to the absent data in Dataset 2. Both datasets contain heart-related information that is of great value, providing possibilities for exploratory data cleansing, modeling, and analysis in order to draw sensible inferences.

	age	ca	chol	oldpeak	thalch	trestbps
count	920.000000	309.000000	890.000000	858.000000	865.000000	861.000000
mean	53.510870	0.676375	199.130337	0.878788	137.545665	132.132404
std	9.424685	0.935653	110.780810	1.091226	25.926276	19.066070
min	28.000000	0.000000	0.000000	-2.600000	60.000000	0.000000
25%	47.000000	0.000000	175.000000	0.000000	120.000000	120.000000
50%	54.000000	0.000000	223.000000	0.500000	140.000000	130.000000
75%	60.000000	1.000000	268.000000	1.500000	157.000000	140.000000
max	77.000000	3.000000	603.000000	6.200000	202.000000	200.000000

Figure 3.3: Descriptive statistics of Dataset 2

In the above illustration of Dataset 2, descriptive statistics such as mean, standard deviation, and quartiles for numerical columns are revealed. Initial values in the 'ca' column for heart conditions are 0 to 3. After reorganization, it is refined to binaries (1: presence, 0: absence), revealing 509 occurrences with cardiac problems and 411 instances without.

3.6 Removing Null and Duplicates

Steps in Preprocessing	Dataset 1	Dataset 2
Initial Structure	(1025, 14)	(920, 16)
Null Count	0	1759
Structure After Removing Null	(1025, 14)	(299, 16)
Duplicates Count	(723, 14)	0
Structure After Removing Duplicates	(302, 14)	(299, 16)

Table 3.1: Removing Null and Duplicates

3.7 Visualizing the Data

Dataset 1

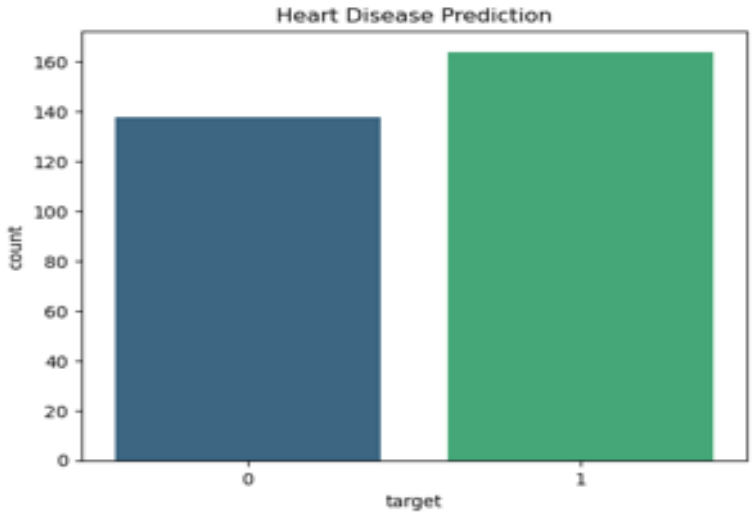

Figure 3.4: Count Plot - Heart Disease Prediction

The resulting countplot depicts the occurrence of heart disease in each patient, which is depicted as either '1' for those with it or '0' for those without it. The graphic representation gives an insight into how the two classes balance each other. This is useful for heart disease prediction and analysis.

Dataset 2

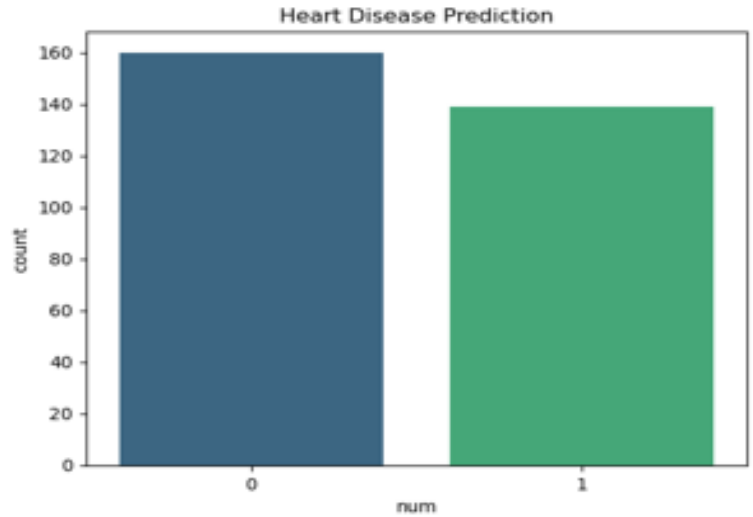

Figure 3.5: Count Plot - Heart Disease Prediction

The count plot presents the distribution of the 'num' variable within the dataset, showing the intensity of heart illnesses (0 indicates no illness, while 1 depicts the presence of illness). It acts as a summary of the dataset for heart disease prediction and helps analyze the problem.

Dataset 1

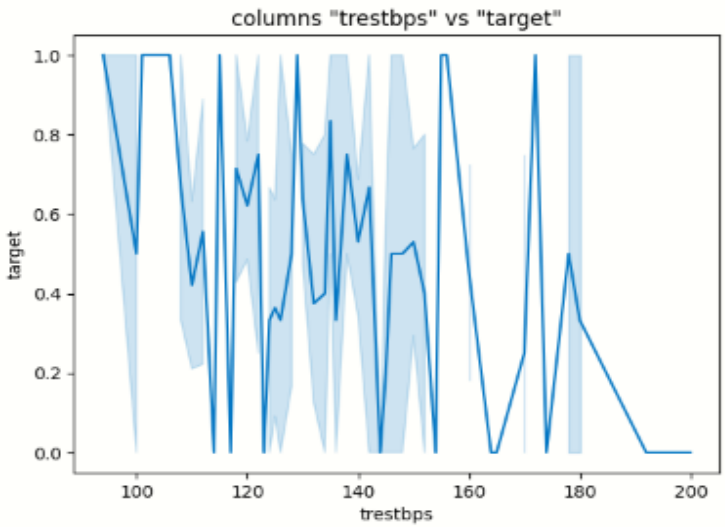

Figure 3.6: Line Plot regarding Columns

The line plot illustrates the correlation visually between the 'trestbps' column and the 'target' variable in the dataset.

- Since the blue line is positioned lower, it suggests that the likelihood of getting heart disease is quite low at very low resting blood pressure values (between 100 and 120 mmHg).
- Based on the peaks in the blue line, it appears that the likelihood of getting heart disease increases when resting blood pressure rises (towards the 130–160 mmHg range).
- As the blue line dips, the likelihood of heart disease looks to be reduced even more for extremely high resting blood pressure readings (over 180 mmHg).

Dataset 2

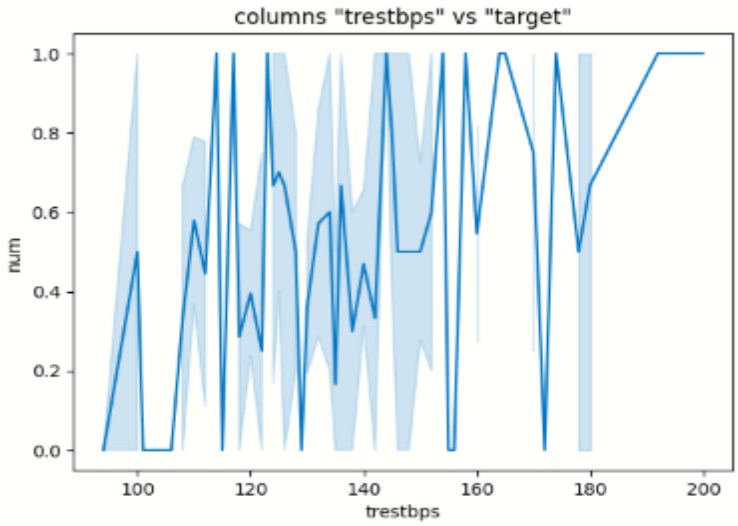

Figure 3.7: Line Plot regarding Columns

The line plot illustrates the correlation visually between the 'trestbps' column and the 'target' variable in the dataset.

- The lower position of the blue line indicates that the likelihood of having heart disease looks to be relatively low at lower blood pressure readings (about 100-120 mmHg).
- The blue line's peaks indicate a discernible increase in the likelihood of heart disease as blood pressure rises to between 130 and 160 mmHg.
- The chance of heart disease appears to rise dramatically with very high blood pressure values (over 180 mmHg), as seen by the blue line's abrupt upward trend towards the conclusion of the figure.

After visualizing both datasets, the datasets undergo "encoding." This is one of the important final steps in preparing the data for implementation. Both datasets, named 1 and 2, were encoded using the label encoder. Finally, after encoding, the resultant dataset 1 with the structure (302, 14) and dataset 2 with the structure (299, 16) were saved as.csv files.

3.8 Splitting the data for Implementation

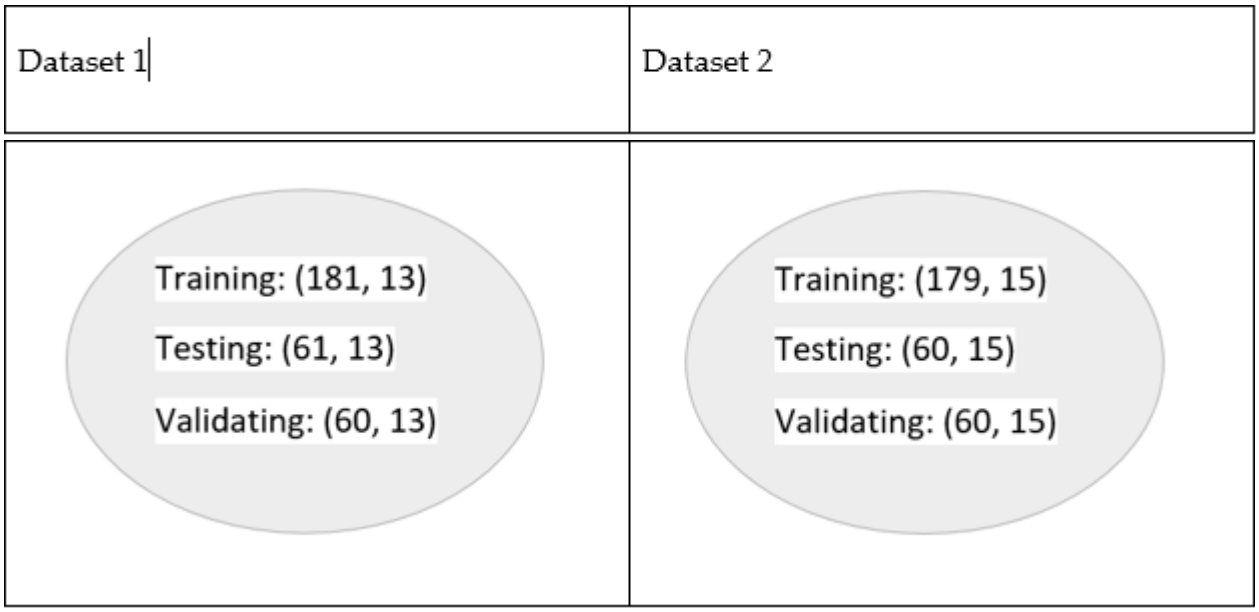

Figure 3.8: Data Splitting

The heart attack data is split into percentages: 60% for training, 20% for validation, and 20% for testing. The table provided here illustrates the exact division of two datasets in depth.

3.9 Algorithms used in predicting heart disease

3.9.1 Decision Tree

Decision Tree is a good example of a machine learning algorithm that describes possible outcomes depending on a set of conditions. Decision Trees are essential in heart disease prediction with combined ML and DL models. The trees divide data based on health indicators such as cholesterol levels, blood pressure among others to help determine patterns relating to heart disease [30].

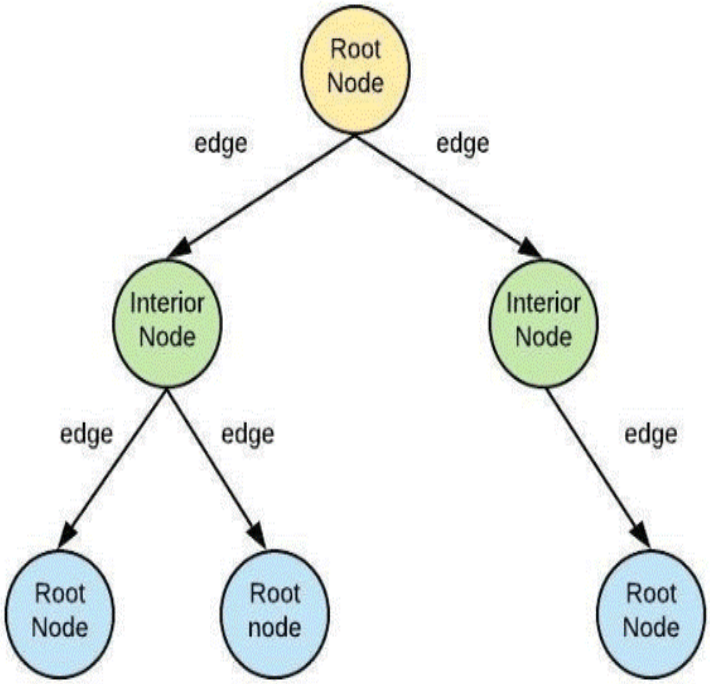

Figure 3.9: The Decision Tree Structure [31]

Node n: if X < threshold, then go to node left ↓ else go to node right ↓
Here,

- X↓ is the value of feature i,
- threshold↓ is the threshold for node n,
- left↓ and right↓, are the child nodes based on the decision rule,

The Gini Impurity for a node t in a classification tree is calculated as follows:

$$\text{Gini}(t) = 1 - \sum_{i=1}^{C} p(i|t)^2 \tag{3.1}$$

where C is the number of classes, and $p(i|t)$ is the proportion of instances of class i at node t.

Decision Trees are notably advantageous concerning interpretability and handling of numerical and categorical data. We incorporate them into collective ML and DL models, in a way that allows us to employ the best of DL towards complex pattern recognition, while rationally using Decision Trees for human-interpretable decision-making. These fusions increase the accuracy of the model thus resulting in more reliable predictions

which in turn helps medical practitioners to understand the main features that lead to heart disease. The ensemble approach with Decision Trees improves not only the predictive performance but also the reliability of the model's recommendation in critical healthcare cases [32].

3.9.2 Support Vector Machine

The Support Vector Machine (SVM) is a supervised ML technique that is used for both classification and regression tasks. This is achieved by locating the best hyperplane that partitions different classes in dataset thereby; increasing the margin. SVMs are very suitable for high dimension feature space and can cope with problems that involve complex distributions relating to data in the feature space Î¦(x). They provide a way of classifying patients on the basis of their diverse features, thereby helping to identify different patterns which may suggest that a patient is suffering from heart related conditions. By providing strong classification boundaries, SVMs complement the abilities ML and DL models [33].

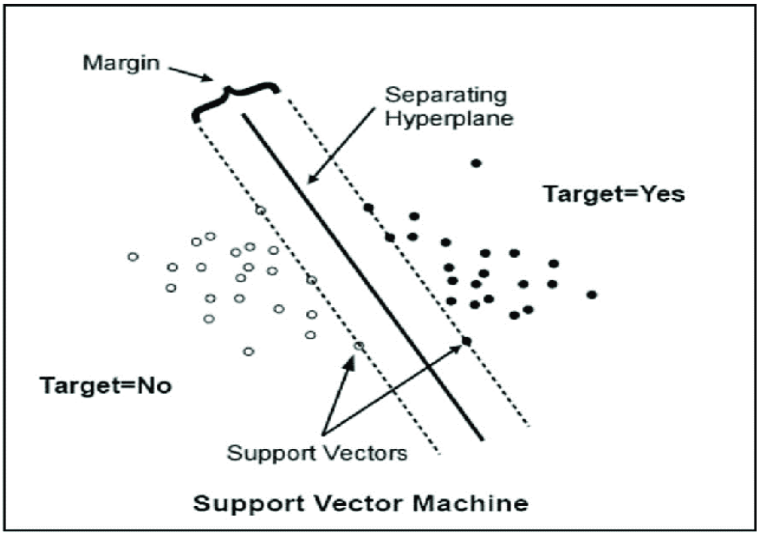

Figure 3.10: Structure of SVM[34]

1. **Hyperplane Equation:** The decision boundary in SVM is represented by a hyperplane equation of the form:

$$f(x) = \text{sign}(wx + b) \tag{3.2}$$

where:

- w is the weight vector.
- x is the input feature vector.
- b is the bias term.

2. **Linear SVM:** For a linear SVM, the decision function for classifying a new data point is given by:

$$f(z) = \text{sign}(wx + b) \tag{3.3}$$

where,

- w is the weight vector.
- x is the input feature vector.
- b is the blas term.

denotes the dot product.

3. **Non-Linear SVM:** For non-linear SVM, the Input features can be mapped into a higher-dimensional space using a kernel function. The decision function becomes:

$$f(x) = \text{sign}(ayK(xx) + b) \tag{3.4}$$

where,

- N is the number of support vectors.
- a is the Lagrange multiplier associated with the i-th support vector.
- Is the class label of the i-th support vector.

K(x, x) is the kernel function that computes the inner product in the higher-dimensional.

In this case, SVMs are preferable due to their suitability for the non-linear data, which is necessary for the recognition of complex patterns connected to heart disease. For example, the use of SVMs together with ML and DL models improves the accuracy of prediction making it more holistic and more effective ML techniques in predicting heart disease [35].

3.9.3 Random Forest

Random Forest is a robust ensemble learning algorithm used often in machine learning for predictive modeling. In predicting heart disease, the fusion of ML and DL models with RF is critical for accuracy and reliability. In the case of classification problems, RF works by generating several DT's during training and gives outputs in the form of modes of the classes. This way, it is especially useful for healthcare as it is adapted to handle big datasets, feature selection, and overfitting [36].

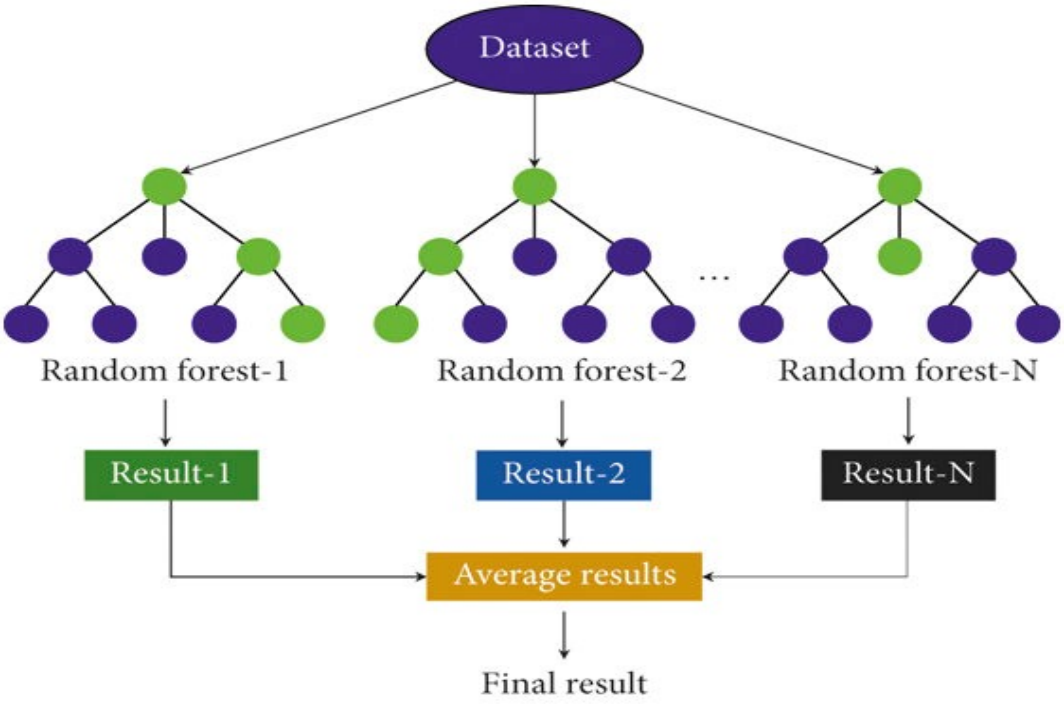

Figure 3.11: Representatipn of Random Forest[37]

Classification: Given N decision trees in the forest and K classes:

$$\text{Final Prediction for Classification} = \text{mode}(C_1, C_2, \ldots, C_N) \tag{3.5}$$

where C_i is the class predicted by the i-th decision tree.

Given a new data point x:

$$\text{Random Forest Prediction}(x) = \text{Majority Vote}(\text{Tree}_1(x), \text{Tree}_2(x), \ldots, \text{Tree}_n(x)) \quad (3.6)$$

where, $\text{Tree}_i(x)$ is the prediction of the i-th tree for the input x.

Combining Random Forest with other ML and DL models results in a more inclusive analysis of the different data type sources in the field of heart disease prediction. While ML models have the ability to capture complicated patterns, DL models work better with complex hierarchical characteristics. Random Forest combines these merits, constituting a strong predictor. Among benefits of this integration include enhanced accuracy, sensitivity, and specificity, which are critical in accurate prognosis and prevention of heart disease [38].

3.9.4 Multilayer Perceptron

MLP is an ANN with multiple layers; namely input layer, one or more hidden layers, and the output layer. Within each layer are nodes or artificial neurons, connected by weights. For example, MLP are trained through supervised learning using backpropagation for weight optimization which allows them to learn complex patterns and relations in data [39].

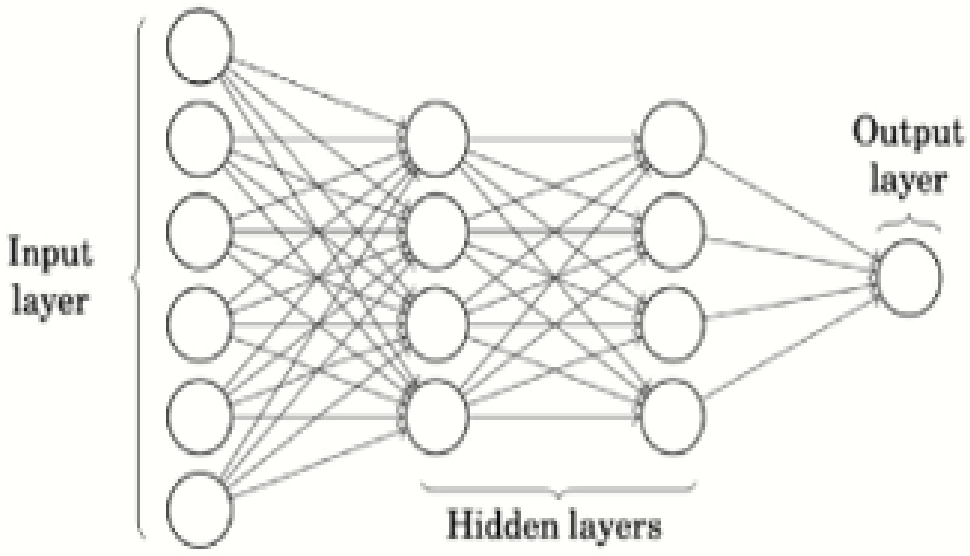

Figure 3.12: Structure of Multilayer Perceptorn[40]

Given an input vector x and a set of parameters (weights and biases), the forward propagation in an MLP can be represented as follows:

For a single layer without activation function:

$$z = Wx + b \tag{3.7}$$

For a single layer with an activation function:

$$a = \sigma(Wx + b) \tag{3.8}$$

For a network with multiple hidden layers:

$$a^{(l)} = \sigma(w^{(l)} \cdot a^{(l-1)} + b^{(l)}) \tag{3.9}$$

where $a^{(l-1)}$ is the output of the previous layer.

For instance, ML and DL models such as MLPs have increased predictive accurateness in predicting heart disease. MLPs have excellent capacity to extract complex features out of a wide array of patient data, contributing to better risk assessments. Their strengths are in determination of non-linear patterns as well as interdependencies, an important aspect in a complex health sector. Therefore, using ML and DL methods, these models provide an all-inclusive strategy for analyzing a multitude of data sources to enhance early identification and customized treatment plan for heart disease [17].

3.9.5 Voting Classifier

A Voting Classifier is an ensemble learning technique that integrates various machine learning (ML) methods to enhance the model's predictive capability. In such studies of heart disease prediction by combining ML with deep learning (DL) models, a Voting Classifier combines the outputs from multiple algorithms, which may include DT, SVM and NN. The goal of this merger is improving the overall precision and generalization. [41].

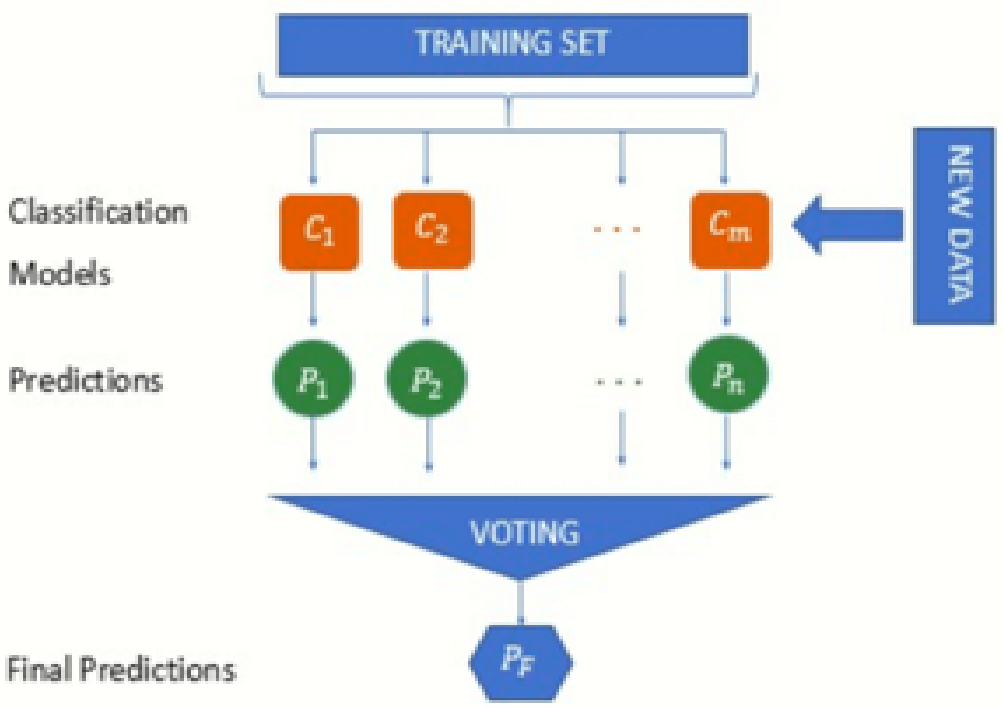

Figure 3.13: Structure of Voting Classifier[42]

The final prediction, $\hat{y}$, is given by:

$$\hat{y} = \text{argmax} \sum_{i=1}^{N} \mathbf{1}(y_i^c = 1) \tag{3.10}$$

where,

- $\hat{y}$ is the final predicted class.
- N is the number of models.
- c represents the class (0 or 1).
- $\mathbf{1}$ is the indicator function.

$y_i^c = 1$ is the prediction of the i-th model for class c.

There are remarkable benefits that come with a Voting Classifier in the prediction of heart diseases. It offsets the weaknesses within any one particular model, a fact that makes it more dependable and hardened. However, this ensemble method is unique in addressing complicated relations within medical datasets and can provide an all-inclusive insight on risk factors. Also, the fact that decision trees are an ensemble

helps to mitigate overfitting and subsequently enhance model generalization. In the health prediction game, where the most accurate prediction wins, the Voting Classifier is a winner, combining a crowd of various models into one highly credible model for predicting heart disease [43].

3.10 Assessment Measurement

- **Precision:** In heart disease prediction, precision, as a measure of accuracy of positive predictions, identifies the proportion of predicted positive cases that are correctly predicted among all predicted positives [44]

  The formula for precision is:

$$\text{Precision} = \frac{\text{True positives}}{\text{True positives} + \text{False positives}} \tag{3.11}$$

  Where:

  True Positives (TP) are the instances correctly predicted as positive (e.g., correctly identifying patients with heart disease).

  False Positives (FP) are the instances incorrectly predicted as positive (e.g., incorrectly identifying patients without heart disease as having heart disease).

- **Recall:** Recall determines whether the model can retrieve all important occurrences of heart disease, given as the number of true-positive cases divided by the overall number of actual positive cases [45].

  The formula for recall is:

$$\text{Recall} = \frac{\text{True positives}}{\text{True positives} + \text{False positives}} \tag{3.12}$$

  Where:

  True Positives (TP) are the number of instances correctly predicted as positive (i.e., correctly identified as having heart disease).

  False Negatives (FN) are the number of instances incorrectly predicted as negative (i.e., incorrectly identified as not having heart disease when they actually do).

- **F1 Score:** F1 Score is a valuable metric of the overall performance of a heart disease prediction model, which combines precision and recall to provide a balance between them [19].

  The formula for calculating the F1 score is:

$$\text{F1 Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{3.13}$$

  The F1 score ranges from 0 to 1, where a higher score indicates better model performance. A perfect classifier achieves an F1 score of 1, while a model that predicts all instances as the majority class would have an F1 score of 0.

- **Accuracy:** In this case, accuracy constitutes the percentage of accurately predicted cases among the positive and negative instances in general relative to the total count of predictions [46].

$$\text{Accuracy} = \left(\frac{\text{Number of correctly predicted instances}}{\text{Total number of instances}} \right) \times 100\% \tag{3.14}$$

  Higher accuracy values indicate better performance of the model in correctly classifying individuals, while lower accuracy values suggest a less reliable model.

- **Confusion matrix:** The confusion matrix is a table that is often used to evaluate the performance of a classification model, such as those used in heart disease prediction. It compares the predicted values from the model with the actual values in the dataset and provides a comprehensive picture on a model's performance in heart disease prediction [47].

  - **True Positive (TP):** The cases where the model correctly predicts the positive class (in this case, presence of heart disease).
  - **False Positive (FP):** The cases where the model incorrectly predicts the positive class (predicts heart disease when it's not present).
  - **True Negative (TN):** The cases where the model correctly predicts the negative class (in this case, absence of heart disease).
  - **False Negative (FN):** The cases where the model incorrectly predicts the negative class (fails to predict heart disease when it's actually present).

	Predicted Negative	**Predicted Positive**
Actual Negative	True Negative (TN)	False Positive (FP)
Actual Positive	False Negative (FN)	True Positive (TP)

Table 3.2: Confusion Matrix

- **Validation time:** The time required to validate is the duration to evaluate and tune the model to avoid the model from being overfitting with different datasets [48].
- **Testing time:** Testing time pertains to the necessary duration for assessing the model's proficiency on fresh data, simulating the presence of circumstances in heart disease prediction [49].
- **Training time:** This refers to the different amounts of time required to train the machine learning and deep learning models to train on the simulated heart disease dataset, which affects the overall effectiveness of the prediction system [50].

CHAPTER 4

EXPERIMENTAL STUDY AND OUTCOMES ANALYSIS

4.1 Implementing Individual ML Heart Disease Prediction Models

Initially, models for ML were implemented separately. The ML algorithm's SVM, RF, MLP and DT are studied in depth. A screenshot illustrating the evaluated parameters for each heart disease prediction model is appended to the table that follows.

SVM Classifier	`Ht_De_Ptn_SVChp = { 'C': [0.1, 0.3, 0.5], 'kernel': ['linear', 'rbf', 'poly', 'sigmoid'] 'gamma': ['scale', 'auto']}`
RF Classifier	`Ht_De_Ptn_RFChp = { 'n_estimators': [100, 200, 300], 'max_depth': [None, 10, 20, 30], 'min_samples_split': [2, 5, 10], 'max_features': ['auto', 'sqrt', 'log2']}`

MLP Classifier	`Ht_De_Ptn_MLPhp = {`
	`'activation': ['relu', 'tanh', 'logistic'],`
	`'solver': ['sgd', 'adam'],`
	`'alpha': [0.0001, 0.001, 0.01],`
	`'learning_rate': ['constant', 'adaptive'],`
	`'max_iter': [100, 200, 300]}`

DT Classifier	`Ht_De_Ptn_DThp = {'criterion': ['gini', 'entropy'],`
	`'max_depth': [10, 20, 30, 40],`
	`'min_samples_split': [3, 5, 7],`
	`'min_samples_leaf': [1, 2, 4]}`

Figure 4.1: Tested parameters

4.1.1 Evaluating Individual ML Models for Dataset-1

The best parameters chosen for the implementation of ML models for dataset 1 are mentioned in the figure below as 4.2

SVM	`Fitting 2 folds for each of 24 candidates, totalling 48 fits`
	`{'C': 0.3, 'gamma': 'scale', 'kernel': 'poly'}`
	`The GSV score of SVM Model: 0.56`

RF	`Fitting 2 folds for each of 108 candidates, totalling 216 fits`
	`{'max_depth': 10, 'max_features': 'auto', 'min_samples_split': 10, 'n_estimators': 20`
	`The GSV score of RF Model: 0.44`

MLP	`Fitting 2 folds for each of 108 candidates, totalling 216 fits`
	`{'activation': 'tanh', 'alpha': 0.001, 'learning_rate': 'adaptive', 'max_iter': 300, 'solver': 'ada`
	`The GSV score of MLP Classifier: 0.53`

DT	`Fitting 2 folds for each of 72 candidates, totalling 144 fits`
	`{'criterion': 'gini', 'max_depth': 10, 'min_samples_leaf': 1, 'min_samples_split':`
	`The GSV score of DT Model: 0.55`

Figure 4.2: Selected Parameters for dataset 1

Support vector machine (SVM):

Parameter (C) for Regularisation 0.3: The regularisation parameter (C) has an optimum value of 0.3. At 0.3, the model favours a softer margin moderately, suggesting that generalisation takes precedence over an excessively close match to the training set.

Gamma is a scale: Scale is the optimal value for the gamma parameter. Selecting 'scale' for gamma indicates that the algorithm will use the inverse of the feature count ($1/n_{\text{features}}$) as the gamma value automatically. By adjusting for different feature scales, this scaling helps prevent overfitting.

Kernel: poly: Poly, which stands for a polynomial kernel, is the best kernel that has been chosen. To learn complex decision boundaries appropriate for the purpose of heart disease prediction, the SVM model makes use of a polynomial kernel.

Interpretation of GSV Score: Evaluate the SVM model's ability to forecast heart disease using the found hyperparameters. Grid Search Cross-Validation (GSCV), which repeatedly divides the dataset into training and validation subsets, is used to assess the performance of the model. The GSV score is the average accuracy over these partitions. With a GSV score of 0.56, the SVM model accurately predicts heart disease in roughly 56% of cases on average. The findings demonstrate the effectiveness of the SVM approach in predicting cardiac disease when the identified hyperparameters are optimized. The subject of heart disease prediction is advanced by these results, which offer useful information regarding the best SVM model configurations.

Random forest (RF):

Parameter Tuning Process: A variety of hyperparameter combinations, such as "max_depth," "max_features," "min_samples_split," and "n_estimators," are investigated by the method using Grid Search Cross-Validation (GSCV). It determines which settings produce the optimal performance by methodically analyzing every combination. The program determines the ideal parameter configuration by analyzing 108 distinct combinations over two folds of cross-validation:

- 'max depth': 10
- 'max features': 'auto'
- 'min samples split': 10
- 'n estimators': 200

GSV Score Interpretation: Grid Search Cross-Validation (GSCV), which divides the dataset into training and validation subsets, is used to evaluate the performance of the model. The GSV score is the average accuracy over these partitions. With a GSV score of 0.44, the Random Forest model predicts heart disease with an average accuracy of 44% over several cross-validation folds.

Multilayer perceptron(MLP):

Hyperparameter Optimization Process: The program determines the optimal parameter configuration after analyzing 108 distinct combinations over two folds of cross-validation:

- "learning_rate": "adaptive"
- "max_iter": 300
- "activation": "tanh"
- "alpha": 0.001

GSV Score Interpretation: Grid Search Cross-Validation (GSCV), which divides the dataset into training and validation subsets, is used to assess the performance of the model ['solver': 'adam']. The GSV score is the mean accuracy for each of these divisions. With various cross-validation folds, the MLP Classifier predicts heart disease with an average accuracy of 53%, as indicated by its GSV score of 0.53.

Dicision tree (DT):

Hyperparameter Tuning Process: The algorithm systematically investigates a variety of hyperparameter combinations, including 'criterion,' 'max_depth,' 'min_samples_leaf,' and 'min_samples_split,' by utilizing Grid Search Cross-Validation (GSCV). To determine which settings produce the greatest results, it assesses every combination. The program determines the optimal parameter configuration after analyzing 72 distinct combinations over two folds of cross-validation:

- Criterion: 'gini'
- Max_depth: 10
- Min_samples_leaf: 1
- Min_samples_split: 7

With a GSV score of 0.55, the Decision Tree model predicts heart disease with an average accuracy of 55% over several cross-validation folds.

As shown in the below table 4.1. The assessment of heart disease prediction models for Dataset 1 revealed that the RF and MLP models achieved notable levels of accuracy (73% and 72%, respectively), while the SVM and DT models demonstrated a comparable accuracy of 65%. The validation set also exhibited the highest scores for additional metrics, including precision recall and f1 score, for the RF heart disease prediction model.

Individual	Precision	Recall	F1 Score	Accuracy
SVM	0.66	0.65	0.63	0.65
RF	0.73	0.73	0.73	0.73
MLP	0.72	0.72	0.71	0.72
DT	0.66	0.65	0.65	0.65

Table 4.1: Validation Metrics for Dataset 1

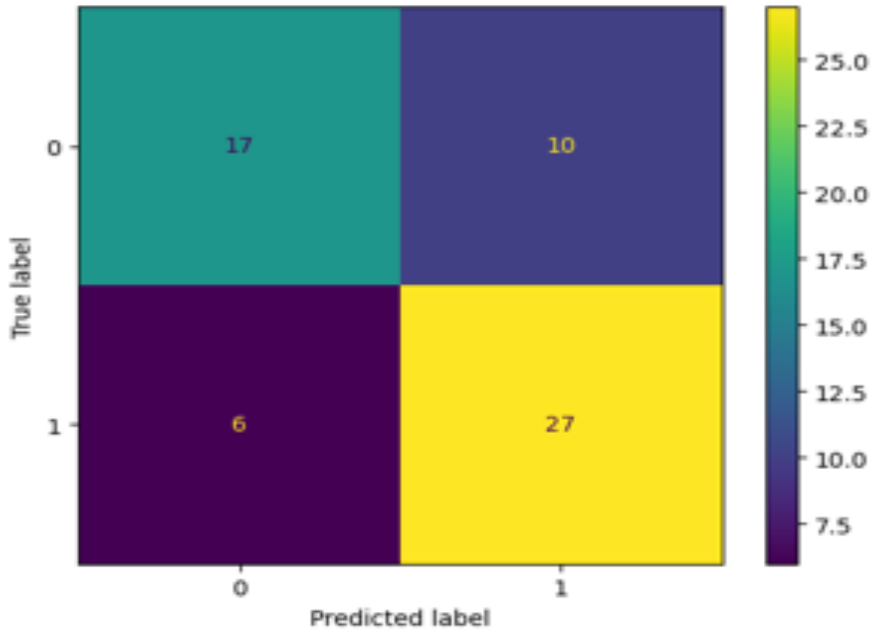

Figure 4.3: Validation Matrix-RF

With 44 instances correctly classified and 16 instances incorrectly classified, the RF heart disease prediction model achieved the highest validation accuracy of 73% in the validation set.

Individual	Precision	Recall	F1 Score	Accuracy
SVM	0.69	0.61	0.58	0.61
RF	0.82	0.80	0.80	0.80
MLP	0.75	0.74	0.74	0.74
DT	0.74	0.74	0.74	0.74

Table 4.2: Testing Metrics for Dataset 1

The evaluation metrics of the heart prediction models for dataset 1 indicate that the RF model, which also performed admirably in the validation set, achieved an outstanding accuracy of 80% in the testing set. MLP, on the other hand, demonstrated an accuracy of 74%. In this case, the DT model achieved an accuracy of 74%, which is significantly higher than the RF model's greatest score in the validation set. The SVM, meanwhile, exhibited subpar performance with an accuracy of 61%.

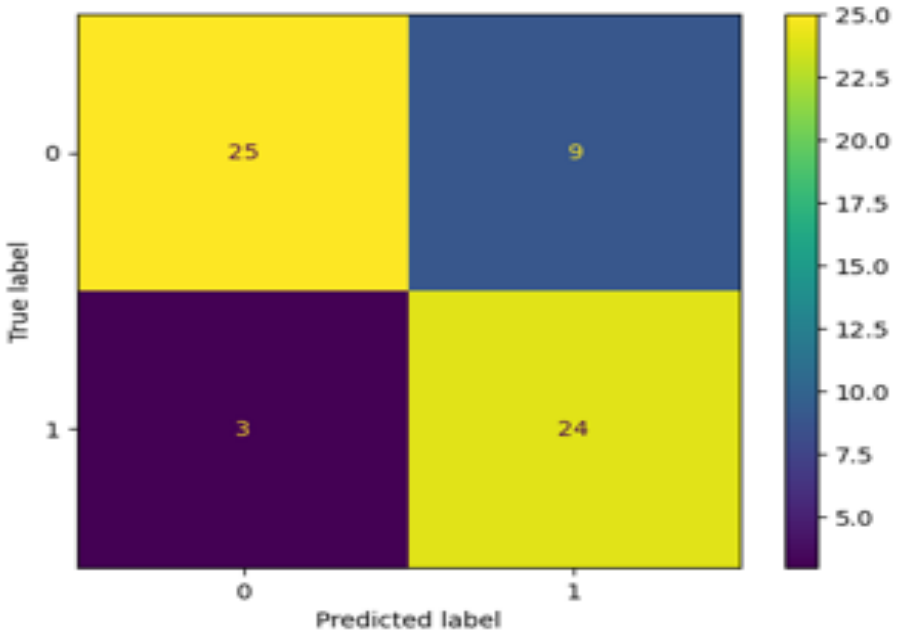

Figure 4.4: Testing Matrix-RF

The excellent-performing RF model has accurately classified 49 instances correctly and 12 instances incorrectly, which leads to a testing accuracy of 80%.

4.1.2 Evaluating Individual ML Models for Dataset-2

The best parameters chosen for the implementation of ML models for dataset 2 are mentioned in the figure below as 4.5.

SVM	Fitting 2 folds for each of 24 candidates, totalling 48 fits {'C': 0.1, 'gamma': 'scale', 'kernel': 'rbf'} The GSV score of SVM model: 0.5700000000000001
RF	Fitting 2 folds for each of 108 candidates, totalling 216 fits {'max_depth': None, 'max_features': 'sqrt', 'min_samples_split': 2, 'n_estimators': 200} The GSV score of RF model: 0.5900000000000001
MLP	Fitting 2 folds for each of 108 candidates, totalling 216 fits {'activation': 'tanh', 'alpha': 0.0001, 'learning_rate': 'adaptive', 'max_iter': 300, 'solver': 'sgd'} The GSV score of MLP classifier: 0.59
DT	Fitting 2 folds for each of 72 candidates, totalling 144 fits {'criterion': 'gini', 'max_depth': 20, 'min_samples_leaf': 2, 'min_samples_split': 3} The GSV score of DT Classifier: 0.5900000000000001

Figure 4.5: Selected parameters for dataset 2

Support vector machine (SVM):

Hyperparameter Tuning Process: The program determines the optimal parameter configuration after analyzing 24 distinct combinations over two folds of cross-validation:

- 'C' = 0.1
- 'gamma' = 'scale'
- 'kernel' = 'rbf'

Interpreting the GSV Score: Grid Search Cross-Validation (GSCV) divides the dataset into subsets for training and validation in order to assess the performance of the model.

The GSV score is the mean accuracy for each of these divisions. With various cross-validation folds, the SVM model's average accuracy in predicting heart disease is 57%, as indicated by its GSV score of 0.57.

Random forest (RF):

Hyperparameter Tuning Process: Following two rounds of cross-validation evaluation of 108 distinct combinations, the method determines the optimal parameter configuration:

- 'max_depth': Null
- 'max features': 'sqrt'
- 'min samples split': 2
- 'n estimators': 200

Interpreting the GSV Score: Grid Search Cross-Validation (GSCV), which divides the dataset into training and validation subsets, is used to assess the performance of the model. The GSV score is the average accuracy over these partitions. With a GSV score of 0.59, the Random Forest model predicts heart disease with an average accuracy of 59% over several cross-validation folds.

Multilayer perceptron (MLP):

The program determines the ideal parameter configuration after analyzing 108 distinct combinations over two folds of cross-validation:

- "Activation": "tanh";
- "Alpha": 0.0001
- "learning_rate": "adaptive";
- "max_iter": 300
- 'solver': sgd

Interpreting the GSV Score: Grid Search Cross-Validation (GSCV) divides the dataset into training and validation subsets to assess the performance of the model. The GSV score is the mean accuracy over these divisions. With an average accuracy of 59% in

predicting heart disease across various cross-validation folds, the MLP classifier obtains a GSV score of 0.59.

Decision tree (DT):

Following the assessment of 72 distinct configurations over two cross-validation folds, the algorithm determines the optimal parameter configuration:

- criteria: 'gini'
- max_depth: 20
- min_samples_leaf: 2
- min_samples_split: 3

Interpreting the GSV Score: Grid Search Cross-Validation (GSCV), which divides the dataset into training and validation subsets, is used to assess the model's performance. The GSV score is the average accuracy over these partitions. With a GSV score of 0.59, the DT classifier predicts heart disease with an average accuracy of 59% across several cross-validation folds.

As shown in the 4.3. The RF heart disease prediction model achieved the highest accuracy of 80% according to the validation metrics presented here for dataset 2. 80% accuracy is the highest accuracy that the same model scored for testing in dataset 1. MLP and DT showed good performance with 72% and 73% accuracy, respectively. SVM showed the worst performance, producing 55% accuracy.

Individual	Precision	Recall	F1 Score	Accuracy
SVM	0.30	0.55	0.39	0.55
RF	0.80	0.80	0.80	0.80
MLP	0.73	0.72	0.71	0.72
DT	0.74	0.73	0.73	0.73

Table 4.3: Validation Matrix-Dataset 2

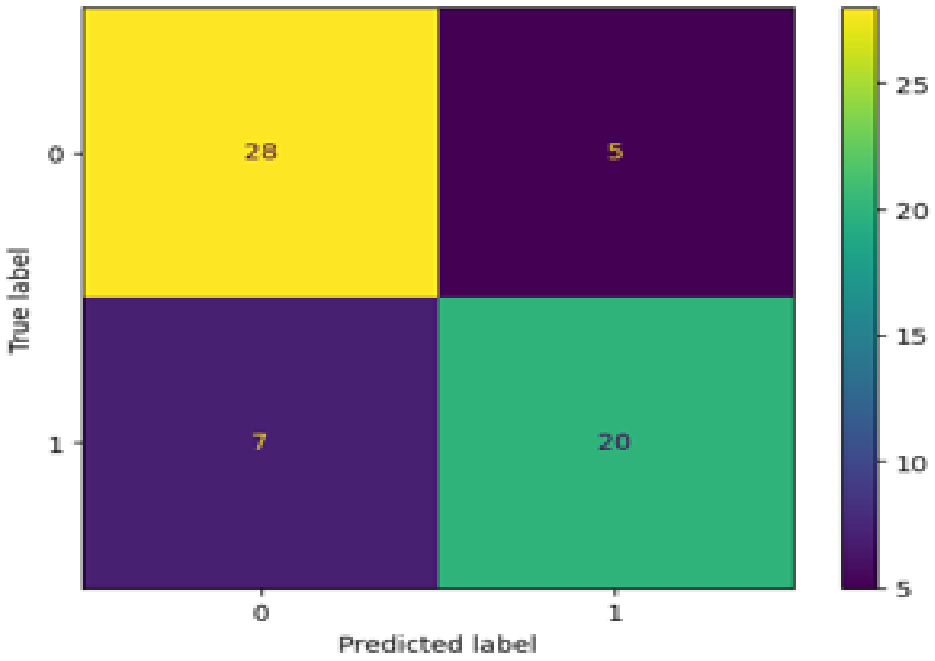

Figure 4.6: Validation Matrix-RF

The validation set yielded the maximum accuracy for the RF heart disease prediction model, at 80%, with 48 instances classified correctly and 12 instances classified incorrectly.

Individual	Precision	Recall	F1 Score	Accuracy
SVM	0.28	0.53	0.37	0.53
RF	0.91	0.90	0.90	0.90
MLP	0.75	0.75	0.75	0.75
DT	0.88	0.88	0.88	0.88

Table 4.4: Testing Matrix-Dataset 2

The testing metrics from dataset 2 are the only metrics that yielded the two highest accuracy values obtained in this experiment when examining individual ML heart disease prediction models: 90% for the RF model and 88% for the DT model. Also effective, MLP demonstrated 75% accuracy. In contrast, SVM demonstrated the least effective performance, scoring 53% accurately.

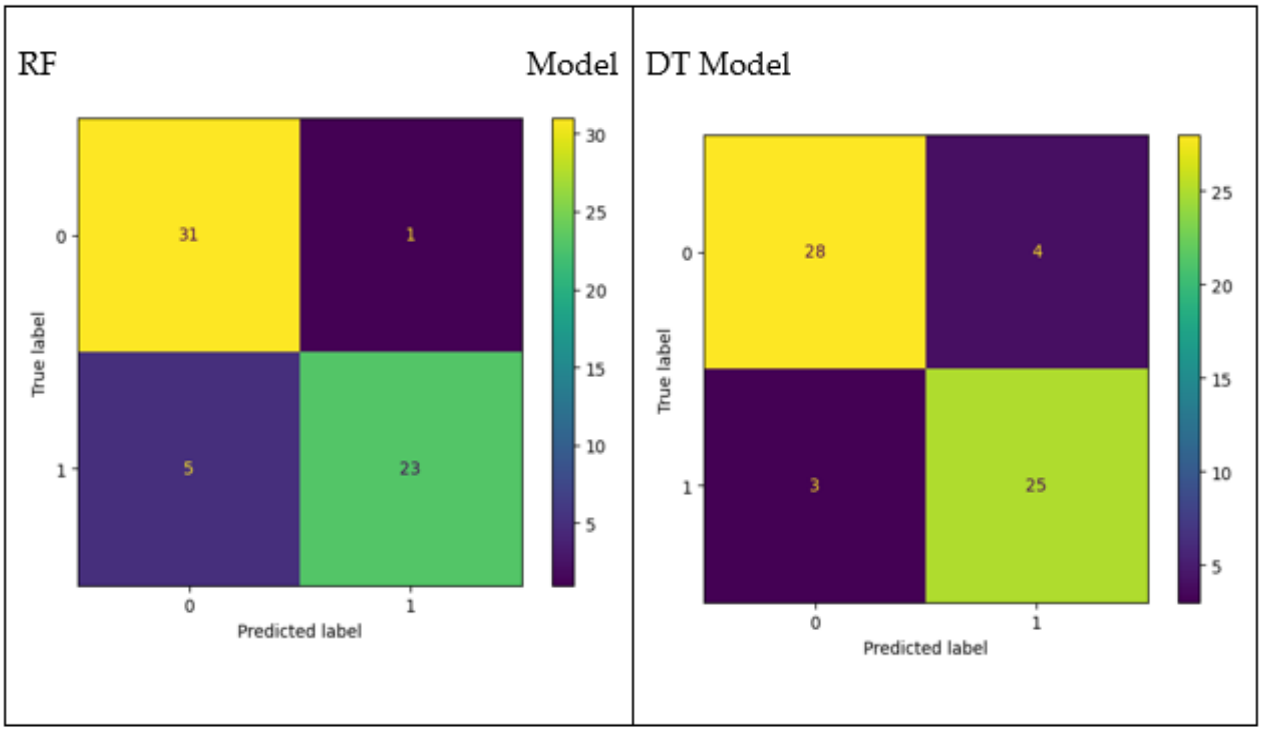

Figure 4.7: Testing Matrix of RF and DT

With 54 instances classified precisely and 6 instances classified incorrectly, the extraordinary RF model has achieved a testing accuracy of 90%. In contrast, the DT model has classified 53 instances accurately and 7 instances incorrectly, for a total accuracy of 88% during testing.

4.1.3 Computational Complexity of Individual Models

Individual ML Models	Training Time	Validation Time	Testing Time
SVM	0.0	0.01	0.02
RF	0.22	0.04	0.03
MLP	0.17	0.003	0.02
DT	0.002	0.01	0.02

Table 4.5: Time Taken by Individual Models for Dataset 1

Individual ML Models	Training Time	Validation Time	Testing Time
SVM	0.01	0.01	0.01
RF	0.23	0.01	0.02
MLP	0.13	0.01	0.01
DT	0.0	0.0	0.0

Table 4.6: Time Taken by Individual Models for Dataset 2

Training, validating, and testing ML models all have different computing requirements. Random Forest (RF) exhibits significantly longer training times, whereas Decision Trees (DT) display modest computational needs throughout all stages in both datasets, demonstrating processing efficiency. In general, all the models exhibit low durations for training, ranging from 0.0 to 0.23 seconds, and for validation and testing, the range is 0.0 to 0.03 seconds.

4.2 Implementing Combined ML Heart Disease Prediction Models

The ML models for predicting heart disease were integrated to produce hybrid models. Tables containing the voting classifiers and additional parameters of each hybrid model developed to assess datasets 1 and 2 are provided below.

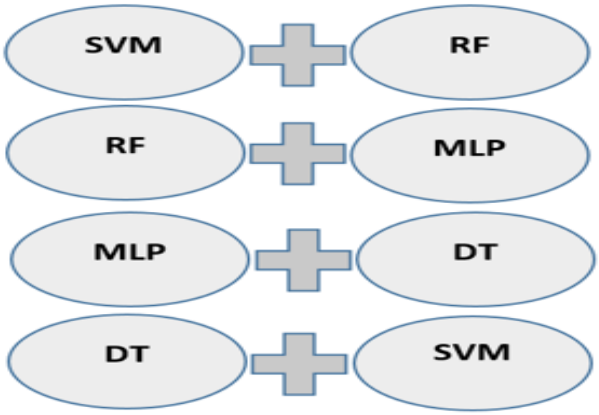

Figure 4.8: Combined ML Modles

Combined	Configurations	
SVM & RF	`SVM_params = {'C': [0.3], 'gamma': ['scale'],'kernel': ['poly']}` `RF_params = {'max_depth': [10], 'max_features': ['auto'], 'min_samples_split': [10], 'n_estimators':`	
	Voting : Hard	GSV Score : 0.88
RF & MLP	`RF_params = {'max_depth': [10], 'max_features': ['auto'], 'min_samples_split': [10], 'n_estimators': [200]}` `MLP_params = {'activation': ['tanh'], 'alpha': [0.001], 'learning_rate': ['adaptive'], 'max_iter': [300], 'solver': ['`	
	Voting : Hard	GSV Score : 0.82
MLP & DT	`DT_params = {'criterion': ['gini'], 'max_depth': [10], 'min_samples_leaf': [1], 'min_samples_split': [7]}` `MLP_params = {'activation': ['tanh'], 'alpha': [0.001], 'learning_rate': ['adaptive'], 'max_iter': [300], 'solver': ['`	
	Voting : Hard	GSV Score : 0.86
DT& SVM	`DT_params = {'criterion': ['gini'], 'max_depth': [10], 'min_samples_leaf': [1], 'min_samples_split':` `SVM_params = {'C': [0.3], 'gamma': ['scale'], 'kernel': ['poly']}`	
	Voting : Hard	GSV Score : 0.88

Figure 4.9: Configuration for Dataset 1

Combined	Configurations	
SVM & RF	`SVM_params = {'C': [0.1], 'gamma': ['scale'],'kernel': ['rbf']}` `RF_params = {'max_depth': [None], 'max_features': ['sqrt'], 'min_samples_split': [2], 'n_estimators':`	
	Voting : Hard	GSV Score : 0.64
RF & MLP	`RF_params = {'max_depth': [None], 'max_features': ['sqrt'], 'min_samples_split': [2], 'n_estimators': [200]}` `MLP_params = {'activation': ['tanh'], 'alpha': [0.0001], 'learning_rate': ['adaptive'], 'max_iter': [300], 'solver': [`	
	Voting : Hard	GSV Score : 0.78
MLP & DT	`DT_params = {'criterion': ['gini'], 'max_depth': [20], 'min_samples_leaf': [2], 'min_samples_split': [3]}` `MLP_params = {'activation': ['tanh'], 'alpha': [0.0001], 'learning_rate': ['adaptive'], 'max_iter': [300], 'solver': [`	
	Voting : Hard	GSV Score : 0.76
DT& SVM	`DT_params = {'criterion': ['gini'], 'max_depth': [20], 'min_samples_leaf': [2], 'min_samples_split':` `SVM_params = {'C': [0.1], 'gamma': ['scale'], 'kernel': ['rbf']}`	
	Voting : Hard	GSV Score : 0.64

Figure 4.10: Configuration for Dataset 2

4.2.1 Evaluating Combined ML Models for Dataset-1

Combined	Precision	Recall	F1 Score	Accuracy
SVM and RF	0.72	0.72	0.72	0.72
RF and MLP	0.72	0.72	0.72	0.72
MLP and DT	0.61	0.58	0.58	0.58
DT and SVM	0.67	0.65	0.65	0.65

Table 4.7: Validation Metrics - Dataset 1

The hybrid models of both SVM and RF and RF and MLP showed 72% highest accuracy in validation. The DT and SVM hybrid models also showed average performance with 65% accuracy. The MLP and DT hybrid models showed poor performance with 58%.

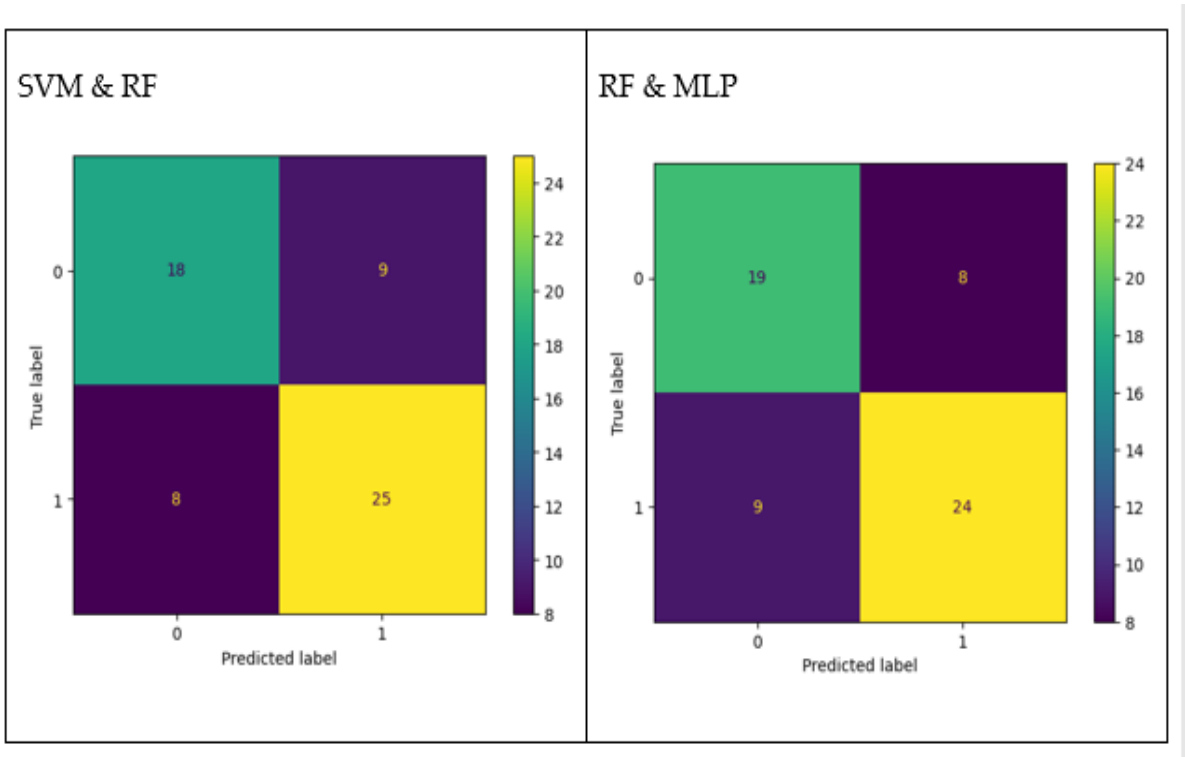

Figure 4.11: Best performing combo modles for dataset 1

With 43 instances classified precisely and 17 instances classified incorrectly, the hybrid SVM and RF model has achieved a validation accuracy of 72%. The same 72% is produced by the RF and MLP models, which also classified 43 instances accurately and 17 instances incorrectly.

Combined	Precision	Recall	F1 Score	Accuracy
SVM and RF	0.86	0.85	0.85	0.85
RF and MLP	0.84	0.84	0.84	0.84
MLP and DT	0.75	0.75	0.75	0.75
DT and SVM	0.75	0.75	0.75	0.75

Table 4.8: Testing Metrics - Dataset 1

The ML combination model SVM and RF again showed the highest accuracy of 85% in testing dataset 1. The combo models (RF and MLP) also showed good performance with 84% accuracy. 75% of the metrics, which is more than the highest metrics in validating dataset 1, are shown by the other two hybrid models in this case.

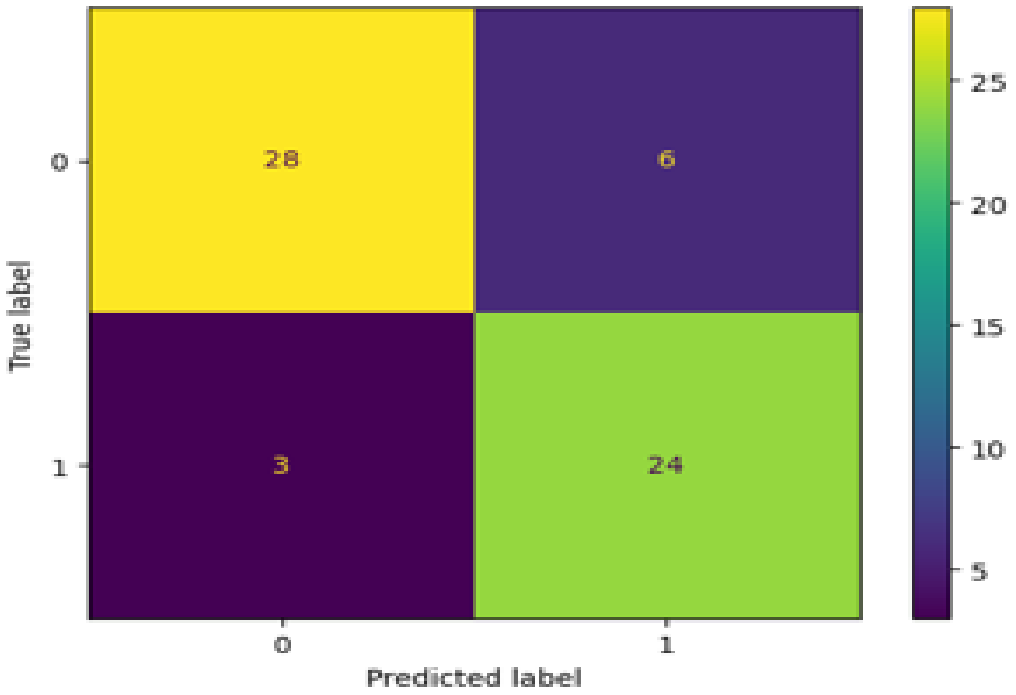

Figure 4.12: Testing Matrix-SVM and RF

The SVM and RF combo model has accurately classified 52 instances correctly and 9 instances incorrectly, which leads to a testing accuracy of 85%.

4.2.2 Evaluating Combined ML Models for Dataset-2

Combined	Precision	Recall	F1 Score	Accuracy
SVM and RF	0.30	0.55	0.39	0.55
RF and MLP	0.73	0.70	0.68	0.70
MLP and DT	0.77	0.77	0.77	0.77
DT and SVM	0.30	0.55	0.39	0.55

Table 4.9: Validation Metrics - Dataset 2

The combo model MLP and DT showed a maximum validation accuracy of 75% in dataset 2, which is a good one when compared to its worst performance in dataset 1. The maximum 75% accuracy is in fact more than the highest validation accuracy of 72%, which two hybrid models produced using dataset 1.

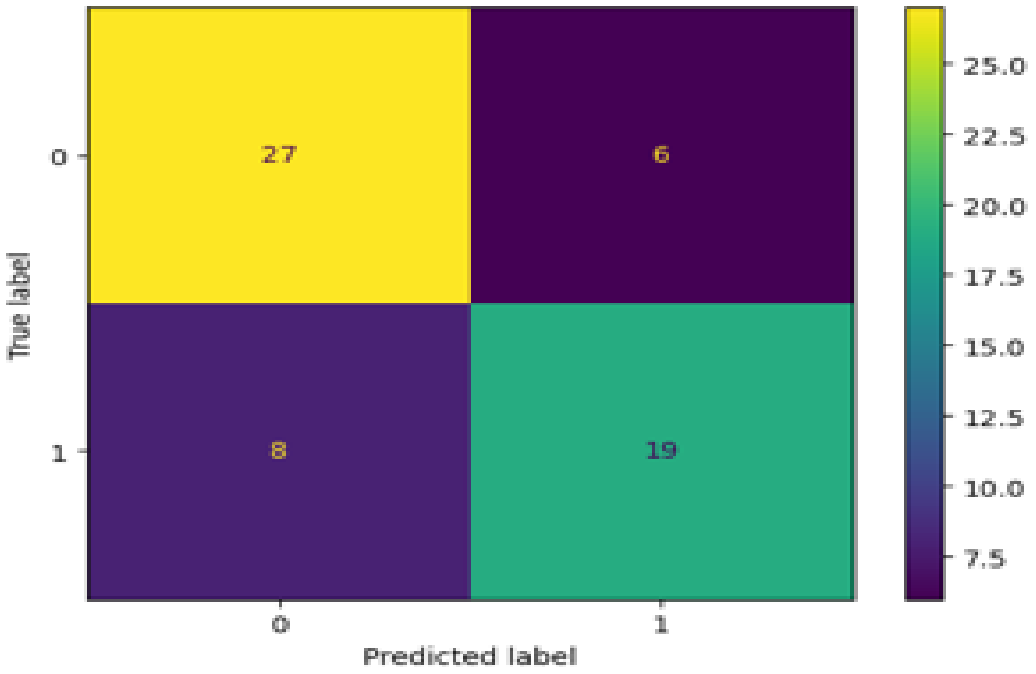

Figure 4.13: Validation Matrix-MLP and DT

With 14 instances classified incorrectly and 46 instances correctly classified, the validation accuracy of the MLP and DT hybrid models is 77%.

Combined	Precision	Recall	F1 Score	Accuracy
SVM and RF	0.28	0.53	0.37	0.53
RF and MLP	0.83	0.80	0.79	0.80
MLP and DT	0.84	0.83	0.83	0.83
DT and SVM	0.28	0.53	0.37	0.53

Table 4.10: Testing Metrics - Dataset 2

According to the testing metrics, the hybrid models, MLP and DT and RF and MLP showed the highest accuracy of 83% and 80% respectively. Meanwhile, the combo models say SVM and RF and DT and SVM showed the worst performance with 53% accuracy.

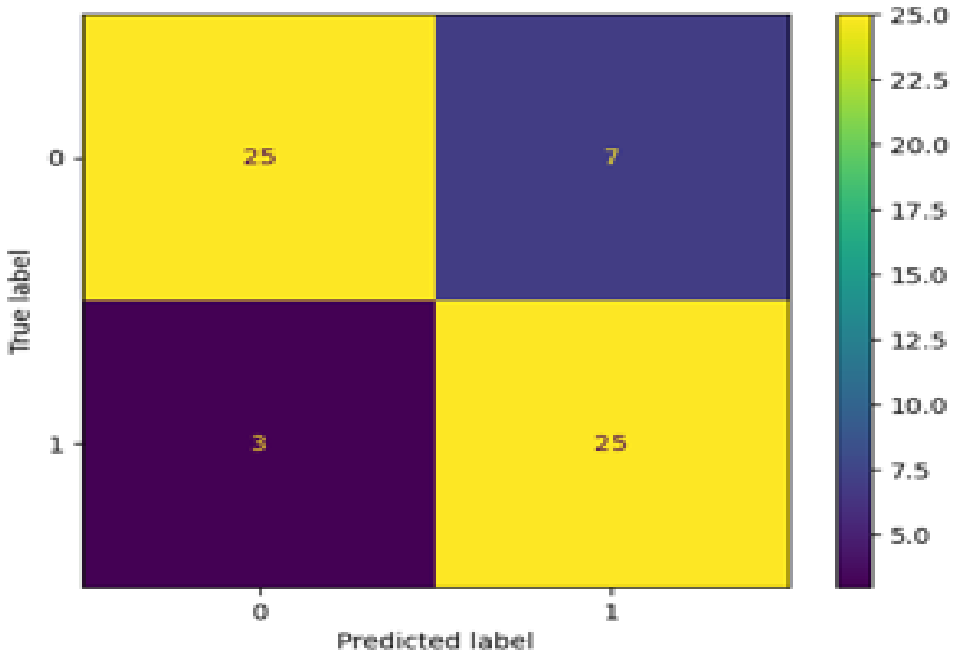

Figure 4.14: Testing Matrix-MLP and DT

The MLP and DT hybrid models achieved a validation accuracy of 83%, with 10 instances classified incorrectly and 50 instances correctly classified.

4.2.3 Computational Complexity of Combined Models

Combined ML Models	Training Time	Validation Time	Testing Time
SVM & RF	0.20	0.03	0.03
RF & MLP	0.34	0.03	0.04
MLP & DT	0.15	0.01	0.01
DT & SVM	0.01	0.01	0.01

Table 4.11: Time Taken by Combined Models for Dataset 1

Combined ML Models	Training Time	Validation Time	Testing Time
SVM & RF	0.20	0.03	0.02
RF & MLP	0.32	0.04	0.04
MLP & DT	0.14	0.01	0.01
DT & SVM	0.01	0.01	0.01

Table 4.12: Time Taken by Combined Models for Dataset 2

Combo ML models have a low computational complexity, with training times between 0.01 and 0.34 seconds. The duration of validation and testing procedures is likewise held constant between 0.01 and 0.04 seconds. In particular, the hybrid models DT and SVM use the fewest computer resources, with times in both datasets as little as 0.01 seconds.

4.3 In-Depth Analysis of Models for Predicting Heart Disease

As part of the thorough examination of heart disease prediction models, individual and combined ML models are assessed according to accuracy and computational demands on two datasets, namely dataset 1 and dataset 2.

Individual ML Modles	Validation Accuracy	Testing Accuracy
SVM	0.65	0.61
RF	0.73	0.80
MLP	0.72	0.74
DT	0.65	0.74

Table 4.13: Comparison of Individual ML Modles for Dataset 1

Individual ML Modles	Validation Accuracy	Testing Accuracy
SVM	0.55	0.53
RF	0.80	0.90
MLP	0.72	0.75
DT	0.73	0.88

Table 4.14: Comparison of Individual ML Modles for Dataset 2

When looking at separate ML models, Random Forest (RF) had the best testing accuracy in Dataset 1, at 80%. It was closely followed by MLP, which got 74%. The reduced accuracy of Decision Tree (DT) and SVM suggests that algorithm selection has a significant bearing. With a testing accuracy of 90% in Dataset 2, again RF exhibited superior performance compared to other models. In contrast, SVM encountered difficulties, achieving the lowest accuracy of 53%. In general, RF exhibited superior performance with consistent accuracy across both datasets, whereas SVM demonstrated the lowest accuracy.

Combined ML Modles	Validation Accuracy	Testing Accuracy
SVM & RF	0.72	0.85
RF & MLP	0.72	0.84
MLP & DT	0.58	0.75
DT & SVM	0.65	0.75

Table 4.15: Comparison of Combined ML Modles for Dataset 1

Combined ML Modles	Validation Accuracy	Testing Accuracy
SVM & RF	0.55	0.53
RF & MLP	0.70	0.80
MLP & DT	0.77	0.83
DT & SVM	0.55	0.53

Table 4.16: Comparison of Combined ML Modles for Dataset 2

In the meantime, a study of mixed machine learning disease prediction models for heart disease shows that different model combinations and datasets give mixed results. The highest validation accuracy of 72% and testing accuracy of 85% were obtained by the SVM and RF combination in Dataset 1, demonstrating the efficacy of combining conventional ML algorithms. On the contrary, the integration of ML models into MLP and DT resulted in a diminished validation performance of 58% accuracy while maintaining a commendable performance of 75% accuracy during testing. In Dataset 2, the MLP and DT which is a heart disease prediction model, outperformed other combinations with validation and testing accuracy of 77% and 83 respectively. The combo RF and MLP, also achieved the second-greatest scores of 80% in testing and 70% in validation. This suggests that some combo ML models might work well together. This shows how important it is to combine models in a way that makes sense for each dataset in order to accurately predict heart disease.

Algorithms	Train	Validate	Test
RF	0.22	0.04	0.03
SVM & RF	0.20	0.03	0.03
RF & MLP	0.34	0.03	0.04

Table 4.17: Comparison of Performance Duration for Dataset 1

Algorithms	Train	Validate	Test
RF	0.23	0.01	0.02
MLP & DT	0.14	0.01	0.01
RF & MLP	0.32	0.04	0.04

Table 4.18: Comparison of Performance Duration for Dataset 2

When analyzing the excellent performers in heart disease prediction, in Dataset 1, the combo model of SVM and RF outperforms others in terms of training and testing times, whereas in Dataset 2, the combo model of MLP and DT exhibits shorter computations compared to others. Overall, the time requirement is low; for training, the range is 0.14 to 0.34 seconds, and for validation and testing, it is 0.01 to 0.04 seconds.

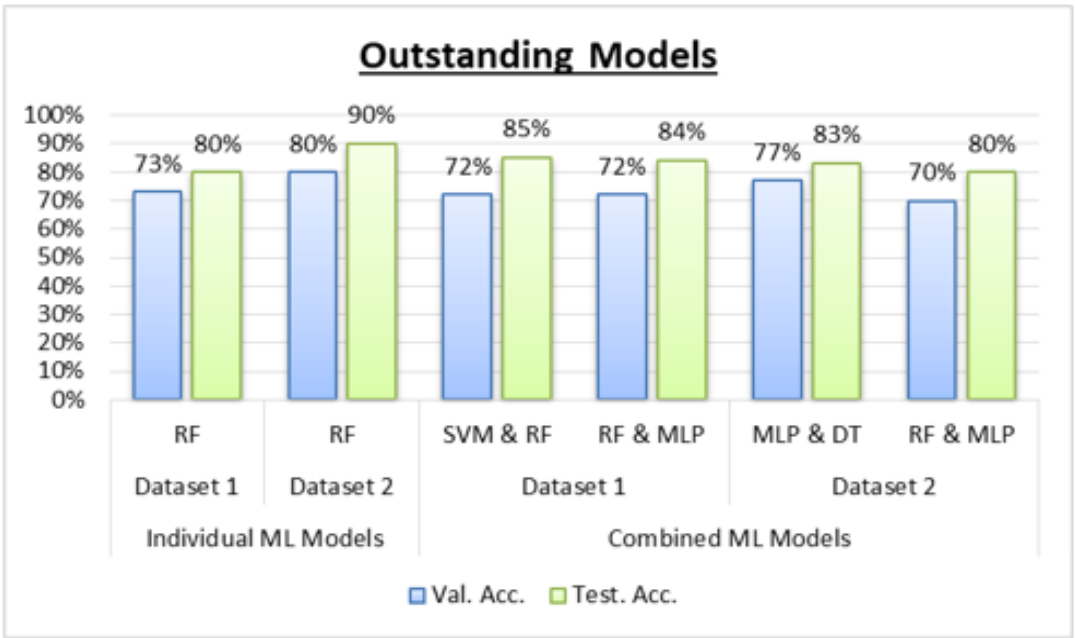

Figure 4.15: Individual Vs Combined Outstanding Modles

In conclusion, the fact that individual models, particularly Random Forest, exhibit consistent accuracy across both datasets highlights the significance of algorithm selection. Some mixes of machine learning models do, however, make predictions more accurate. This shows how important it is to use dataset-specific model integration for accurate heart disease predictions.

CHAPTER 5

CONCLUSIONS

5.1 Thorough analysis

I have learned various crucial lessons from this project on "Heart Disease Prediction Using Combined ML and DL Model." First, the project was an eye-opener in terms of the complex relationship between machine learning (ML) and deep learning (DL) architectures in the context of heart diseases anticipation. Personally, the background research proved to me that it is critical to combine these two approaches for improved accuracy while reducing prediction time.

The hands-on experience resulted from practical implementation of individual ML algorithms such as Decision Tree, Support Vector Machines, Random Forests, and MLP classifier, thereafter, integrated into combined ML and DL models. Therefore, this procedure not only developed my technical skills in implementing, fine-tuning a variety of algorithms, but also enhanced my aptitude to preprocess and manipulate real-world datasets for precise and dependable predictions.

This project has, in summary, developed my technical skills in machine learning and deep learning, as well as given me insight on practical concerns and difficulties of using these models in reality, in the context of medical science. This project has definitely increased my knowledge and understanding of data science which will no doubt go a long way in ensuring that I approach other future projects with a more informed mind.

5.2 Achievement of Research Objectives

In the Identification of Heart Disease Prediction Using Combined ML and DL Model project, the question on the efficacy of a combined machine learning and deep learning model in predicting heart diseases has been comprehensively dealt with. The objectives were systematically followed, which began by conducting a literature review on the prediction of heart disease using ML and DL algorithms. A rigorous review of the previous methods and technique used in the field was undertaken.

The data that was used for analysis was the data on heart disease, which was obtained from the two heart disease datasets. Data quality was ensured by utilizing vigorous preprocessing procedures. Some provisions were made to remove errors, achieve correct labeling, and resolve improper scaling as these required validation of datasets for proper model learning.

Use of diverse separate ML algorithms for prediction of heart disease. The algorithms used included Decision Tree, Support Vector Machines (SVM), Random Forests, and MLP classifier. The algorithms were applied one after another to determine their performance. Integrated Machine Learning models were formed by combining the individual ML algorithms. These combinations are designed to exploit complementarities among alike algorithms to boost prediction accuracy of heart diseases. The emphasis was on establishing the performance in varied cases of the heart disease datasets.

The research was able to determine the best combined machine learning algorithm for the heart disease detection through evaluation and comparison process. Results enunciate the combinations that outperformed others instructing the future applications in the heart disease prediction.

5.3 Conclusion

Accurate prediction of heart disease is important as it enables early intervention, potentially saving lives and reducing healthcare costs. By employing advanced ML models, the individuals at risk are identified, allowing for targeted preventive measures and personalized treatment. This study highlights the importance of combining machine learning and deep learning algorithms for accurate heart disease prediction across dif-

ferent datasets. The datasets used for this research are "UCI Heart Disease Dataset" and "Heart Disease Dataset" which is available on Kaggle. Initially, individual algorithms like Decision Tree, Support Vector Machines (SVM), Random Forests, and MLP classifier are implemented, following which the combination of ML and DL models is carried on by the voting classifier. The research findings showcase that individual model, Random Forest (RF) yielded superior accuracy of 80% and 90% in dataset 1 and 2 respectively. Results obtained by combining ML and DL models show that in Dataset 1, the SVM RF combination achieved the highest validation accuracy of 72% and testing accuracy of 85%, demonstrating the effectiveness of integrating conventional ML algorithms. Conversely, in Dataset 2, the MLP DT combination outperformed others with validation and testing accuracy of 77% and 83%, respectively, highlighting the importance of selecting appropriate model combinations for specific datasets. Also, in Dataset 1, the SVM RF combination demonstrates superior performance in both training and testing times, while in Dataset 2, the MLP DT combo shows shorter computation durations compared to other models, with overall time requirements ranging from 0.14 to 0.34 seconds for training and 0.01 to 0.04 seconds for validation and testing. These results underscore the potential of strategic integration of machine learning and deep learning techniques to improve heart disease prediction accuracy, thereby facilitating more effective diagnosis and treatment strategies in clinical settings.

5.4 Guidance for future research

The heart disease prediction model using combined machine learning (ML) and deep learning (DL) algorithms could be improved in several ways by further research. First, more experiments with ML and DL techniques combined in various ways could give the ideas on synergistic effects on accuracy and efficiency. Investigating other DL models including convolutional neural networks (CNNs) or recurrent neural networks (RNNs) beside classic ML models would reveal new ways of enhancing prediction accuracy.

Additionally, an advanced analysis of hyperparameter tuning and feature selection techniques will improve the model's performance more. Choosing proper features and fine tuning hyperparameters significantly influence the accuracy of the model and

shorten the time to train. Moreover, exploring the advantages of data augmentation approaches, like synthetic minority oversampling technique (SMOTE) or generative adversarial networks (GANs), would fix imbalances of the dataset and enhance model generalization.

In addition, increasing the coverage of the research by using larger and more varied datasets from different sources might improve the robustness and generalization of the model. Including datasets that represent diverse demographic, geographic, and medical condition characteristics may offer a better insight into the model's performance in different populations and healthcare settings. Generally, more research into these directions could help in the progress of heart disease prediction and better patient care.

Bibliography

[1] P. Rani, R. Kumar, N. M. S. Ahmed, and A. Jain, "A decision support system for heart disease prediction based upon machine learning," Journal of Reliable Intelligent Environments, vol. 7, no. 3, pp. 263-275, 2021.

[2] A.Alqahtani, S. Alsubai, M. Sha, L. Vilcekova, and T. Javed, "Cardiovascular disease detection using ensemble learning," Computational Intelligence and Neuroscience, vol. 2022.

[3] Ziaeian, Boback, and Gregg C. Fonarow, "Epidemiology and aetiology of heart failure." Nature Reviews Cardiology 13, no. 6: 368-378, 2016.

[4] J. Ren, N.N. Wu, S. Wang, J.R. Sowers, and Y. Zhang, "Obesity cardiomyopathy: evidence, mechanisms, and therapeutic implications. Physiological reviews", 101(4), pp.1745-1807, 2021.

[5] R. Mohebi, C. Chen, N.E. Ibrahim, C.P. McCarthy, H.K. Gaggin, D.E. Singer, E.P. Hyle, J.H. Wasfy and J.L. Januzzi Jr, "Cardiovascular disease projections in the United States based on the 2020 census estimates". Journal of the American College of Cardiology, 80(6), pp.565-578, 2022.

[6] M. Al-Kofahi, A. Yun, J.W., Minagar and J.S. Alexander, "Anatomy and roles of lymphatics in inflammatory diseases". Clinical and Experimental Neuroimmunology, 8(3), pp.199-214, 2017.

[7] A.G. Goodwill, G.M. Dick, A.M. Kiel, and J.D. Tune, "Regulation of coronary blood flow". Comprehensive Physiology, 7(2), p.321, 2017.

[8] A.G. Atanasov, S.B. Zotchev, V.M. Dirsch and C.T. Supuran, "Natural products in drug discovery: advances and opportunities". Nature reviews Drug discovery, 20(3), pp.200-216, 2021.

[9] K.K Wong, G. Fortino and D. Abbott, "Deep learning-based cardiovascular image diagnosis: a promising challenge". Future Generation Computer Systems, 110, pp.802-811, 2020.

[10] A. Timmis, N. Townsend, C.P. Gale, A. Torbica, M. Lettino, S.E. Petersen, E.A. Mossialos, A.P. Maggioni, D. Kazakiewicz, H.T. May and D. De Smedt, "European Society of Cardiology: cardiovascular disease statistics 2019", European heart journal, 41(1), pp.12-85, 2020.

[11] V. N. Agbor, N. A. Ntusi, and J. J. Noubiap, "An overview of heart failure in low-and middle-income countries," Cardiovascular Diagnosis and Therapy, vol. 10, no. 2, p. 244, 2020.

[12] Y. Muhammad, M. Tahir, M. Hayat, and K. T. Chong, "Early and accurate detection and diagnosis of heart disease using intelligent computational model," Scientific reports, vol. 10, no. 1, p. 19747, 2020.

[13] D. Shah, S. Patel, and S. K. Bharti, "Heart disease prediction using machine learning techniques," SN Computer Science, vol. 1, pp. 1-6, 2020.

[14] V. V. Ramalingam, A. Dandapath, and M. K. Raja, "Heart disease prediction using machine learning techniques: a survey," International Journal of Engineering and Technology, vol. 7, no. 2.8, pp. 684-687, 2018.

[15] M. Marimuthu, M. Abinaya, K. S. Hariesh, 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.

[16] A. Singh and R. Kumar, "Heart disease prediction using machine learning algorithms," in 2020 international conference on electrical and electronics engineering (ICE3), pp. 452-457,IEEE,2020.

[17] S. Mohan, C. Thirumalai, and G. Srivastava, "Effective heart disease prediction using hybrid machine learning techniques," IEEE access, vol. 7, pp. 81542-81554, 2019

[18] A.N.Repaka , S. D. Ravikanti, and R. G. Franklin, "Design and implementing heart disease prediction using naives Bayesian," in 2019 3rd International conference on trends in electronics and informatics (ICOEI), pp. 292-297, IEEE, April 2019.

[19] M. M. Ali, B. K. Paul, K. Ahmed, F. M. Bui, J. M. Quinn, and M. A. Moni, "Heart disease prediction using supervised machine learning algorithms: Performance analysis and comparison," Computers in Biology and Medicine, vol. 136, p. 104672, 2021.

[20] M. Kavitha, G. Gnaneswar, R. Dinesh, Y. R. Sai, and R. S. Suraj, "Heart disease prediction using hybrid machine learning model," in 2021 6th international conference on inventive computation technologies (ICICT), pp. 1329-1333, IEEE, January 2021.

[21] L. Riyaz, M. A. Butt, M. Zaman, and O. Ayob, "Heart disease prediction using machine learning techniques: a quantitative review," in International Conference on Innovative Computing and Communications: Proceedings of ICICC 2021, Volume 3, pp. 81-94, Springer Singapore, 2022.

[22] M. A. Khan, "An IoT framework for heart disease prediction based on MDCNN classifier," IEEE Access, vol. 8, pp. 34717-34727, 2020.

[23] N. Chandrasekhar and S. Peddakrishna, "Enhancing Heart Disease Prediction Accuracy through Machine Learning Techniques and Optimization," Processes, vol. 11, no. 4, p. 1210, 2023.

[24] C. M. Bhatt, P. Patel, T. Ghetia, and P. L. Mazzeo, "Effective heart disease prediction using machine learning techniques," Algorithms, vol. 16, no. 2, p. 88, 2023.

[25] J. M. Luna, E. D. Gennatas, L. H. Ungar, E. Eaton, E. S. Diffenderfer, S. T. Jensen, C. B. Simone, J. H. Friedman, T. D. Solberg, and G. Valdes, "Building more accurate decision trees with the additive tree," Proceedings of the national academy of sciences, vol. 116, no. 40, pp. 19887-19893, 2019.

[26] S. Chidambaram and K. G. Srinivasagan, "Performance evaluation of support vector machine classification approaches in data mining," Cluster Computing, vol. 22, pp. 189-196, 2019.

[27] J. Dou, A. P. Yunus, D. T. Bui, A. Merghadi, M. Sahana, Z. Zhu, C. W. Chen, K. Khosravi, Y. Yang, and B. T. Pham, "Assessment of advanced random forest and decision tree algorithms for modeling rainfall-induced landslide susceptibility in the Izu-Oshima Volcanic Island, Japan," Science of the total environment, vol. 662, pp. 332-346, 2019.

[28] A. Al Bataineh and S. Manacek, "MLP-PSO hybrid algorithm for heart disease prediction," Journal of Personalized Medicine, vol. 12, no. 8, p. 1208, 2022.

[29] S. Kumari, D. Kumar, and M. Mittal, "An ensemble approach for classification and prediction of diabetes mellitus using soft voting classifier," International Journal of Cognitive Computing in Engineering, vol. 2, pp. 40-46, 2021.

[30] R. Ahuja, A. Chug, S. Gupta, P. Ahuja, and S. Kohli, "Classification and clustering algorithms of machine learning with their applications," in Nature-inspired computation in data mining and machine learning, pp. 225-248, 2020.

[31] F. M. J. Shamrat, R. Ranjan, K. M. Hasib, A. Yadav, and A. H. Siddique, "Performance evaluation among id3, c4. 5, and cart decision tree algorithm," in Pervasive Computing and Social Networking: Proceedings of ICPCSN 2021, pp. 127-142, Springer Singapore, 2022.

[32] V. G. Costa and C. E. Pedreira, "Recent advances in decision trees: An updated survey," Artificial Intelligence Review, vol. 56, no. 5, pp. 4765-4800, 2023.

[33] D. A. Pisner and D. M. Schnyer, "Support vector machine," in Machine learning, pp. 101-121, Academic Press, 2020.

[34] E. H. A. Rady and A. S. Anwar, "Prediction of kidney disease stages using data mining algorithms," Informatics in Medicine Unlocked, vol. 15, p. 100178, 2019.

[35] J. P. Li, A. U. Haq, S. U. Din, J. Khan, A. Khan, and A. Saboor, "Heart disease identification method using machine learning classification in e-healthcare," IEEE access, vol. 8, pp. 107562-107582, 2020.

[36] L. Zhao, X. Wu, R. Niu, Y. Wang, and K. Zhang, "Using the rotation and random forest models of ensemble learning to predict landslide susceptibility," Geomatics, Natural Hazards and Risk, vol. 11, no. 1, pp. 1542-1564, 2020.

[37] H. Fu and K. Qi, "Evaluation Model of Teachers' Teaching Ability Based on Improved Random Forest with Grey Relation Projection," Scientific Programming, 2022, pp. 1-12.

[38] M. M. Ahsan and Z. Siddique, "Machine learning-based heart disease diagnosis: A systematic literature review," Artificial Intelligence in Medicine, vol. 128, p. 102289, 2022.

[39] A.J. Mohammed, M. H. Arif, and A. A. Ali, "A multilayer perceptron artificial neural network approach for improving the accuracy of intrusion detection systems," IAES International Journal of Artificial Intelligence, vol. 9, no. 4, p. 609, 2020.

[40] S. Ghoshal and S. Roberts, "Thresholded ConvNet ensembles: neural networks for technical forecasting," Neural Computing and Applications, vol. 32, pp. 15249-15262, 2020.

[41] K. Raza, "Improving the prediction accuracy of heart disease with ensemble learning and majority voting rule," in U-Healthcare Monitoring Systems, pp. 179-196, Academic Press, 2019.

[42] N. Shakya, R. DUBEY, and L. Shrivastava, "Stress Detection using EEG Signal Based on Fast Walsh Hadamard transform and Voting Classifier," 2021.

[43] A.Mehmood, M. Iqbal, Z. Mehmood, A. Irtaza, M. Nawaz, T. Nazir, and M. Masood, "Prediction of heart disease using deep convolutional neural networks," Arabian Journal for Science and Engineering , vol. 46, no. 4, pp. 3409-3422, 2021.

[44] A.Gupta, A. Saha, N. S. Reddy, and U. D. Acharya, "Cardiac Disease Prediction using Supervised Machine Learning Techniques," in Journal of Physics: Conference Series, vol. 2161, no. 1, p. 012013,IOP Publishing, 2022.

[45] R. Aggrawal and S. Pal, "Sequential feature selection and machine learning algorithm-based patient's death events prediction and diagnosis in heart disease," SN Computer Science, vol. 1, no. 6, p. 344, 2020.

[46] D. M. Powers, "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation," arXiv preprint arXiv:2010.16061, 2020.

[47] R. Bharti, A. Khamparia, M. Shabaz, G. Dhiman, S. Pande, and P. Singh, "Prediction of heart disease using a combination of machine learning and deep learning," Computational intelligence and neuroscience, 2021.

[48] T. Xia, Y. Song, Y. Zheng, E. Pan, and L. Xi, "An ensemble framework based on convolutional bi-directional LSTM with multiple time windows for remaining useful life estimation," Computers in Industry, vol. 115, p. 103182, 2020

[49] J. Sekar, P. Aruchamy, H. Sulaima Lebbe Abdul, A.S. Mohammed, and S. Khamuruddeen, "An efficient clinical support system for heart disease prediction using TANFIS classifier," Computational Intelligence, vol. 38, no. 2, pp. 610-640, 2022.

[50] S. Tuli, N. Basumatary, S.S. Gill, M. Kahani, R.C. Arya, G.S. Wander, and R. Buyya, "HealthFog: An Ensemble Deep Learning based Smart Healthcare System for Automatic Diagnosis of Heart Diseases in integrated IoT and fog computing environments," Future Generation Computer Systems, vol. 104, pp. 187-200, 2020.

Dataset-1

Here is my link: Touch Here

Dataset-2

Here is my link: Touch Here