International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2022): 7.942

Comparative Analysis for Predicting Cardiovascular Diseases Using Machine Learning and Deep Learning Approaches

Chidozie Louis Uzoegwu[1], Farah Ahmed[2], Honglei Li[3]

[1]MSc Computer Science and Digital Technologies Project, Department of Computer and Information Sciences at Northumbria University, Newcastle, United Kingdom

[2]Supervisor, Department of Computer and Information Sciences at Northumbria University, Newcastle, United Kingdom

Abstract: ***This research investigates the potential of using physiological signs, including respiratory rate, blood pressure, body temperature, heart rate, and oxygen saturation, to predict cardiovascular disease (CVD) in humans. Machine learning (ML) and deep learning (DL) models were employed to determine the most effective prediction model by comparing their performance metrics to a previous study conducted by Ashfaq et al. in 2022. Ashfaq's research utilized three parameters (body temperature, heart rate, and oxygen saturation) and achieved a top performance of 96% using K-Nearest Neighbour (KNN). The analysis utilized a dataset obtained from the MIMIC-III clinical database. Four models were evaluated: Random Forest (RF), K-Nearest Neighbour (KNN) as part of the ML approach, and Multi-Layer Perception (MLP) and Convolutional Neural Network (CNN) as part of the DL approach. Performance evaluation was conducted using five measurement metrics, namely accuracy, precision, recall, F1-score, and ROC AUC. The findings demonstrate significant performance by all models, with MLP exhibiting the highest overall performance measures, including an accuracy of 99%, precision of 99%, recall of 99%, F1-score of 98%, and ROC AUC of 98%. The RF model closely followed MLP in terms of performance. This research provides valuable insights for medical researchers, individuals, academies, analysts, and artificial intelligence enthusiasts, informing them about research ideas and areas for improvement, particularly in the health sector, specifically in the management of CVD in humans. Furthermore, the integration of these models into monitoring systems using body sensors could facilitate prompt emergency intervention for CVD patients. In comparison to the previous study by Ashfaq et al., this research expands the parameter set to include five body parameters, enhancing the accuracy and effectiveness of CVD prediction. The utilization of advanced ML and DL models highlights the potential for significant improvements in the field of cardiovascular disease prediction and management.***

Keywords: cardiovascular diseases, machine learning, deep learning, ML, DL, comparative analysis, CVD

1. Introduction

This chapter presents the background of cardiovascular disease (CVD) research that focuses on this topic and highlights the issues and motivations of this research. Several research questions surrounding CVD are explored and key questions examining research gaps are addressed. Based on the problems identified in this topic, the purpose and objectives of the study are determined. Research methods for solving problems identified in research are indicated, and tools and methods for solving the problems are highlighted. Finally, the structure of the study will be organized according to the main elements of the study and each element focuses on the topic.

1.1 Background

Cardiovascular disease (CVD) encompasses various heart and circulatory disorders, such as irregular heartbeat, coronary heart disease, myocardial infarction, hypertension, and stroke (British Heart Foundation, 2019; National Heart, Lung, and Blood Institute, 2022; NHS, 2022). CVD has emerged as a significant global health concern, with the number of deaths from CVD surpassing 31 million by 2020, mainly affecting low- and middle-income countries (WHO, 2020; Tsao et al., 2022b). It is predicted that CVD will become the leading cause of death worldwide, accounting for over 40% of all deaths by 2030 (WHO, 2021). Unhealthy lifestyle choices, genetics, age, and underlying health conditions like high cholesterol and diabetes contribute to CVD development (WHO, 2019; CDC, 2019; Cleveland Clinic, 2021). Despite research on lifestyle factors, cultural and environmental factors in low-income regions like Sub-Saharan Africa and Asia hinder effective CVD management (Hamid, Groot, and Pavlova, 2019; Mishra et al., 2019).

Addressing CVD requires better interventions, especially in low- and middle-income countries with limited healthcare access. Artificial Intelligence (AI) has emerged as a transformative tool to aid in the prevention, detection, treatment, and monitoring of CVD through machine learning (ML) and deep learning (DL) models. These models, such as Decision Trees, Random Forest, Support Vector Machines, and Convolutional Neural Networks, can analyze large medical datasets to predict CVD likelihood and customize treatment strategies for patients (Hassani et al., 2020; Aurelien Géron, 2019; Karatzia, Aung, and Aksentijevic, 2022). ML and DL have the potential to revolutionize medicine by monitoring human physiological vital signs, aiding early diagnosis, and providing insights into individual fitness and health levels (Muthukumar, Dhanagopal, & Ramesh, 2019; Kebe et al., 2020).

The application of ML and DL to predict CVD using essential body signs has gained significant interest in recent years. The focus of this paper is on enhancing the precision, accuracy, and reliability of utilizing these symptoms and vital human physiological body parameters in the prediction of CVD (Karatzia, Aung, and Aksentijevic, 2022). By leveraging AI technologies, there is a potential to improve CVD management and reduce its impact on global health, particularly in underserved regions (Witayanukorn et al., 2013; Jacob et al., 2022; Malo et al., 2017). However, non adherence to medication and the lack of adequate healthcare facilities remain as challenges that need to be addressed (Goldstein, Gathright, and Garcia, 2017; Abegaz et al., 2017; Hsu, Warren, & Riddle, 2022).

In conclusion, CVD is a significant cause of global mortality, with most deaths occurring in low- and middle-income countries. Unhealthy lifestyle choices and underlying health conditions contribute to CVD, and there is a need for improved interventions and healthcare access. AI technologies, particularly ML and DL, offer promising solutions for predicting, preventing, and treating CVD by utilizing vital human physiological body parameters. Despite challenges like non adherence to medication and limited healthcare facilities, the application of AI in CVD management holds the potential to enhance disease control and improve patients' well-being globally.

1.2 Research Questions

1) What are the many viewpoints, findings, and methods taken by other academics in the use AI for bettering the treatment of CVD, as well as what are the key aspects which need to be considered when utilising ML and DL methods to forecast CVD?
2) What ML and DL model is more effective in predicting CVD, and how do their performances compare?

1.3 Research Aims

CVD is the leading cause of death worldwide. Clinical professionals are always searching for novel CVD treatments. Additionally, AI has enhanced CVD diagnosis, management, and observation. Despite several publications, CVD prediction can be enhanced. Comparative research on ML and DL for predicting CVDs is required due to the presence of additional target parameters and large datasets. This study compares the performance and efficacy of ML and DL in predicting CVD through these indicators and symptoms by examining more accurate and efficient structures, algorithms, and methodologies using a large clinical dataset containing vital human physiological signs associated with CVD. Using ML and DL algorithms, this study aims to expand knowledge of CVD prediction.

1.4 Research Objectives

A set of objectives has been defined and is being implemented to achieve the expected study findings:

1) Research and critically evaluate important CVD research publications, their origins, obstacles, and achievements, as well as the intervention of AI via ML and DL predictions, relevant algorithms, and prior implementation of these models in existing research and datasets.
    - Critically reviewing relevant research articles to learn more about CVD causes, risk factors, and treatment choices.
    - Examining and critically analysing relevant research publications to identify CVD research gaps and future research possibilities.
2) Review the literature to identify significant body parameters for CVD prediction and medically accepted good and bad symptoms and measurement.
    - Review the literature for CVD prediction indicators and symptoms.
    - Read medical papers to learn about normal vital indicator ranges.
    - Create CVD prediction models using appropriate body signals.

3) Comparing datasets for project selection.
- Consider data access's ethical, legal, societal, professional, and security implications.
- Evaluate several datasets and sources.
- Data management should meet industry standards.
- The dataset should contain all relevant variables.

4) Process and analyse data.
- Data collection and data pre-processing.
- Perform exploratory data analysis.
- Perform data transformation.
- Visualize the data characterization.

5) ML/DL model comparison.
- Select ML and DL models for comparison.
- Use ML and DL algorithms to predict, categorise, or assess outputs based on input data.
- Explain the analysis's findings via visuals, reports, or other methods.
- Use performance measurements to evaluate deployed models.
- Compare the results to verify the analysis.

6) Critical evaluation for research work
- Visualise project results.
- Critically evaluate the research aim and compare project results to existing research.
- Applaud the successes and critique the methods.

7) Viva analysis, general report, and publication
- Prepare the publication for KF7029 and supervisor approval.
- Prepare the MSc Computer Science and Digital Technologies (KF7029) final project VIVA.

1.5 Research Approach

The methodology for the study will include descriptive, exploratory, predictive, and comparative techniques.

A descriptive technique comprises obtaining and analysing data using methods such as observations and case studies to give a comprehensive and in-depth picture of the issue under investigation (Siedlecki, 2020). Exploratory technique is a research strategy that assists researchers in investigating and gaining a preliminary understanding of a certain occurrence or subject (Elman, Gerring, and Mahoney, 2020). As a result, the study will employ these tools to better understand

CVD, their causes, therapies, and the role of AI in combating them.

The present study will use a predictive strategy of the two models between ML and DL. According to (Dookie et al., 2018), a prediction strategy is a quantitative research method that employs both past and current data to foresee likely happenings or results.

In addition, both models will be evaluated in a comparative analysis to find similarities and differences between the two targeted models, ML and DL. The comparable approach, as defined by (Ivanov and Webster, 2018), is a research technique that includes comparing two or more events, groups, or situations to uncover similarities and contrasts to understand their underlying causes.

2. Literature Review

2.1 Cardiovascular Diseases and their Prevalence

Cardiovascular disease (CVD) encompasses various heart and blood vessel conditions, often caused by the accumulation of fatty deposits in arteries, leading to blood clots and potential damage to organs like the brain, heart, kidneys, and eyes (NHS, 2022; WHO, 2022). CVD-related illnesses, including coronary heart disease, cerebrovascular disease, and rheumatic heart disease, collectively affect the heart and blood arteries (WHO, 2022). CVD is the leading cause of global deaths, claiming approximately 17.6 million lives annually (WHO, 2022; NHS, 2022; Balakumar et al., 2016; Jagannathan et al., 2019; Mensah et al., 2019; CDC, 2022). Low- and middle-income countries (LMICs) face higher CVD death rates due to lifestyle changes like reduced exercise, smoking, and obesity, with LMICs experiencing three times the death rate of high-income countries (HICs) (Jagannathan et al., 2019; Mensah et al., 2019). CVD affects both developed and developing nations (Balakumar et al., 2016; Townsend et al., 2021).

Europe accounts for around 60 million CVD-related deaths yearly, while Asia has 58% of the global CVD deaths, presenting challenges in prevention and management due to the continent's size, cultural diversity, and healthcare disparities (Movsisyan et al., 2020; Zhao, 2021; OECD and Organisation, 2022). Sub-Saharan Africa (SSA) experiences a significant CVD outbreak, with hypertension, obesity, diabetes mellitus, and dyslipidemia contributing to a notable number of deaths and disabilities (Amegah, 2018). In 2013, SSA accounted for one million CVD-related deaths, making up 5.5% of global CVD deaths and 11.3% of African mortality. CVD is the primary cause of death in several nations, affecting specific racial and ethnic groups more, such as African Americans, American Indians, Alaska Natives, Hispanics, and white males, who have higher heart disease mortality rates than other diseases (CDC, 2022). Among specific populations like Pacific Islanders, Asian Americans, American Indians, Alaska Natives, and Hispanic women, cancer is the only disease more prevalent than heart disease (CDC, 2022). Accurate data on CVD prevalence and epidemiology in Asia is crucial for effective pandemic policy design (Zhao, 2021; OECD and Organisation, 2022).

Table 2.1: Proportion of deaths attributed to heart disease in 2020, categorized by ethnicity, race, and gender.

Race of Ethnic Group	% of Deaths	Male, %	Female, %
American Indian or Alaska Native	14.2	15.5	12.7
Asian	18.9	20.0	17.8
Black (Non-Hispanic)	20.7	21.0	20.3
Native Hawaiin or Other Pacific Islander	20.8	21.9	19.4
White (Non-Hispanic)	21.3	22.7	19.8
Hispanic	15.8	15.8	15.8
All	20.6	21.6	19.5
More Than One Race	18.2	19.2	16.9

Source: CDC (2022)

2.2 Risk Factors of CVDs

In general, risk factors are situations or activities that increase the probability of acquiring a disease (Wein, 2017; Health and Safety Executive (HSE), 2017; National Cancer Institute 2019, British Heart Foundation (BHF), 2019). Several cardiac and circulation illnesses are caused by risk factors that may be managed, corrected, or altered (BHF, 2019). CVD risk factors, according to (Ada's Medical Knowledge Team, 2022), are specific trends of actions, habits, or illnesses that might increase an individual's probability of developing CVD. Inadequate physical activity, poor eating habits, cigarette use, diabetes, old age, and a family history of CVD are all risk factors (BHF, 2019; CDC, 2020; National Institute of Health and Care Excellence (NICE), 2023).

To further explain on the risk factors impacting CVD, (NICE, 2023) classified them as non-modifiable, modifiable, and comorbidities. Age, race, ethnic background, and heritage are non-modifiable factors; cigarette use, elevated cholesterol levels, inadequate exercise, improper nutrition, drinking too much alcohol, and becoming overweight are modifiable factors; and comorbidities include diabetes, hypertension, and schizophrenia.

Human biophysical body characteristics, sometimes known as symptoms or indicators, may detect, identify, and reduce risk factors (Reber et al. 2019).

2.3 Human Bodily Symptoms and Parameters Measurements

Human physiological or biological body symptoms are essential indicators of a live body's overall status and play a crucial role in recognizing and monitoring medical disorders. Healthcare practitioners rely on vital signs, including heart rate, respiration rate, temperature, blood pressure, and oxygen saturation, to assess bodily functions (MedlinePlus, 2015; Johns Hopkins Medicine, 2019). These vital indicators are vital in determining a patient's clinical status and are routinely assessed in medical settings, households, and emergency situations. Several studies have highlighted the significance of physiological characteristics, such as blood pressure, ECG, arterial stiffness, ankle-brachial blood pressure index, and blood glucose measurements, in assessing cardiovascular disease (CVD) risk (Lin, Zhang, and Zhang, 2013). For CVD monitoring and intervention, ML algorithms have been employed to predict vital signs, such as body temperature, heart rate, and oxygen saturation (Ashfaq et al., 2022).

Normal vital sign ranges can vary depending on age, gender, weight, exercise capacity, and overall health. However, for adults in good health at rest, typical ranges include blood pressure between 90/60 mmHg and 120/80 mmHg, respiratory rate of 12 to 18 breaths per minute, heart rate of 60 to 100 beats per minute, and body temperature of 97.8°F to 99.1°F (36.5°C to 37.3°C) (MedlinePlus, 2015; Lindsey, 2022; Johns Hopkins Medicine, 2019; Sapra, Malik, and Bhandari, 2021). Oxygen saturation levels below 95% at rest are considered abnormal, with 95% being the normal threshold (Europe PMC, 2019; Khan et al., 2021; Windisch et al., 2023). Monitoring vital signs remains a crucial aspect of healthcare, providing valuable information for assessing and maintaining a person's well-being and preventing potential complications.

Table 2.2: Shows normal vital sign measurement in adults. Source: (Lindsey, 2022) of Healthline

NORMAL VITAL SIGNS IN ADULTS

CORE TEMPERATURE	98.6°F (37°C)
HEART RATE	60–100 beats per minute
RESPIRATORY RATE	12–18 breaths per minute
BLOOD OXYGEN	95–100%
BLOOD PRESSURE	120/80 mm Hg

The measurements and readings of these parameters can assist in detecting irregularities or modifications in a person's well-being, which may suggest an underlying ailment or disorder.

2.4 Machine Learning (ML) and Deep Learning (DL) Prediction of CVD

CVD is the leading cause of death worldwide, with a greater prevalence among those from poor socioeconomic backgrounds (Mishra and Monica, 2019). There is a necessity for digitised healthcare facilities that can monitor vital signs and propose therapy based on medical data to diagnose CVD early and automatically. The information might come from previously acquired data or information from an existing or current source. This data is learned and studied in the process of ML or DL algorithms (as a subset of AI) for insights into trends, patterns, and analysis, or to develop projections for the future utilising what is learned; this procedure is known as machine learning (ML) or deep Learning (DL) (Tripoliti et al., 2017; Swathy and Saruladha, 2021; Pal et al., 2022). This approach of predicting CVD may give treatments such as early detection, tracking, and control of persons living with CVD, as well as general illness therapy. There is always room for improvement, thus additional research into techniques to enhance the ML and DL prediction of CVD is welcomed. This is the primary goal of this research, which will investigate and use an extensive dataset containing the five major human vital parameters as factors to determine which of both models, ML and DL, has the better overall performance metrics by means of a comparative analysis of the two models and algorithms.

(Mohd Faizal et al., 2021) analyses how ML outperforms standard statistical approaches in predicting and evaluating CVD. Even though their mathematical concepts are identical, ML provides quicker feature selection and can analyse non-linear interactions and missing data. ML can find additional factors and generate better predictions than statistical approaches that concentrate on correlations between restricted variables. ML may be used to choose features, classify them, or both. In CVD investigations, Table 2.3 compares conventional risk score methodologies to AI-based risk prediction approaches.

Table 2.3: Conventional risk prediction approach vs. AI risk prediction approach in CVD.

Feature	Conventional Risk Score Approach	AI-based Risk Prediction Approach
Hypothesis	Yes	No
Approach	Estimates and explains data	Practical prediction from data
Measurement	Coefficients, goodness-of-fit	Precision, recall, F-measure, accuracy, area under curve
Learning ability	No	Yes
Data size	Suitable-size data that fits a certain hypothesis	Big and complex data
Data type	Single-type of data, structured data	Supports multi-modality data, structured and unstructured data
Model	Simple parametric model	Complex non-parametric model
Output	Validate the hypothesis, causality	Predict new data, identify hidden patterns
Limitation	Required more assumptions, low-dimensionality data	Overfitting, data privacy and security issues
Risk factors	Clinical/demographic factors only e.g. age, gender, ethnicity, smoking, hypertension etc.	Multimodality e.g. age, gender, ethnicity, ECG variations, differential expressed genes, features from images etc.

Source: Mohd Faizal et al., (2021)

Tran et al. (2019) conducted a study on ML and AI applications in strokes and heart disease care, highlighting their potential in diagnosis, mortality prediction, and improving health behaviors. The research revealed significant growth in areas like big data analysis, robotic prosthesis, robotic-assisted stroke therapy, and minimally invasive surgery. Other publications, such as Baashar et al. (2022) and Barry et al. (2023), also recognize the immense promise of ML and DL in cardiovascular disease (CVD) treatments. These technologies show potential in enhancing heart failure treatment, echocardiography interpretation, evaluation reliability, and establishing relationships between medical data. The use of algorithms for identifying, classifying, and predicting may lead to advancements in managing CVD patients and improving their care.

2.5 Evaluation of ML and DL Models, Datasets, and their Performance Metrics

ML and DL can predict CVD outcomes, according to recent studies. These approaches may find CVD risk factors in large datasets. This study tests ML and DL CVD prediction models using a large lifestyle and physiological dataset. A solid prediction model will help physicians improve patient outcomes.

Verma, Srivastava, and Negi (2016) developed a hybrid CVD diagnostic approach operating correlation-based feature subset (CFS) selection, particle swarm optimisation (PSO), and KNN clustering to identify risk factors. Research simulates CVD cases using multi-layer perception (MLP), and Multinomial Logistic Regression (MLR). The approach was tested on 26 attributes and 335 occurrences of medical data gathered at the Department of Cardiology, Indira Gandhi Medical College, Shimla, India. The result shows MLR predicted best with accuracy of 88%.

Acharya et al. (2018) propose an 11-layer CNN for congestive heart failure (CHF) diagnosis. This unique method pre-processes electrocardiogram (ECG) data minimally and categorises it without artificial characteristics. The model was trained and tested on four different datasets (A, B, C, and D), and Set B achieved outstanding accuracy of 98.97% with specificity and sensitivity levels at 99.01% and 98.87%, respectively.

Lorenzoni et al. (2019) explores eight machine learning techniques (MLTs) for predicting heart failure hospitalisation using data from the present Gestione Integrata dello Scompenso Cardiaco (GISC) investigation in Puglia, Southern Italy. Logistic regression, classification and regression trees, random forests, adaboost, logitboost, SVMs, and neural networks were employed. 380 long-term support programme heart failure patients utilising an online portal to exchange data with healthcare professionals were investigated. MLTs predicted hospitalisation using patient demographic, medical history, and clinical data. BMI, age, heart rate, BNP, pulmonary pressure, serum creatinine, mean years between clinical examinations, and gender were among these variables. The generalised linear model net (GLMN) predicted hospitalisation the best, with an average accuracy of 0.81, an AUC of 0.80, a positive predictive value of 0.87, and a negative predictive value of 0.75. MLTs can predict heart failure hospital admissions using healthcare information from patient interaction with the healthcare system, but the publication recommends further study to enhance accuracy and precision.

Nahiduzzaman et al. (2019) compare MLP and SVM for heart disease identification. SVM beats MLP in two-class classification, 92.45% versus 90.57%. MLP outperforms SVM in five-class classification, 68.86% to 59.01%. SVM or MLP use class count for classification, according to study. Due to its accuracy and efficiency in heart disease identification, MLP is recommended for five-class issues and SVM for two-class problems. Alternative classifiers and real-time clinical data from various healthcare institutions may increase heart issue screening accuracy, according to studies.

Rim et al. (2020) examined medical signals for two years using DL. Electromyography (EMG), Electrocardiography (ECG), Electroencephalography (EEG), and Electrooculography (EOG), as well as combinations of these signals, are analysed. The researchers analysed 147 DL studies that predicted hand motion, CVD, mental disorders, emotional states, sleep phases, age, and gender. CNN anticipates physiological signals with precision. Transfer learning is also utilised by certain contributions, which instead of training from inception utilise pre-trained weights from state-of-the-art models such as AlexNet, VGG, ResNet, Inception, or DenseNet. The research indicates that DL methods could enhance medical signal processing.

Swathy and Saruladha (2021) examined classification, data mining (DM), ML, and DL CVD prediction models. Each strategy has performance metrics, information, and tools in the survey. DM extracts valuable data without assumptions using many methods. Neural networks outperform decision tree (DT) and Naive Bayes classifiers (NBC). Multi-Layer Perceptron Neural Networks (MLP) with Back Propagation outperformed others in 15 features.

Pal et al. (2022) diagnosed CVD using two trustworthy ML algorithms, multi-layer perceptron (MLP) and K-nearest neighbour (K-NN). Model training and testing used 13 main features from 303 samples with 76 characteristics. The model's goal is to tell if a patient has CVD (1) or not (0). Using MLP and K-NN classification methods, the model was trained and tested, and its accuracy and AUC score were used to measure and compare its performance. The results show that the MLP model is more accurate than the K-NN models. It has an accuracy rate of 0.82 and an AUC score of 0.86, while the K-NN models have accuracy rates of 0.73 and 0.86, respectively.

Lee et al. (2022) designed and validated statistical and DL models to predict hypertensive CVD mortality and hospitalisation. In the Korean National Health Insurance Service (NHID) database, 2,037,027 hypertensives were hospitalised or died from CVD (myocardial infarction or stroke) within a year of their last visit. NHID data analysed hypertension-related CVD hospitalisation and death. The national health screening database gathered age, sex, diabetes status, statin medication usage, BMI, SBP, DBP, FPG, total cholesterol, income, smoking status, and physical

Paper ID: SR23809044938 DOI: 10.21275/SR23809044938

activity. To tackle case-control ratios, researchers used synthetic minority oversampling technique (SMOTE), a ML method for datasets with severe class distribution imbalances. Logistic regression (LR) and deep neural networks (DNN) predict CVD hospitalisation and death. DNN predicted CVD hospitalisation and death better than LR. The DL model had higher accuracy (0.863 vs. 0.655 for hospitalisation and 0.925 vs. 0.780 for death). The DNN model contained 64 nodes in a single hidden layer. Nonlinearity was applied to the hidden layer using the Rectified Linear Unit (ReLU) activation function, and the dropout ratio was set to 0.2 to prevent overfitting. The DNN model can predict CVD hospitalisation and death in hypertensive patients, helping manage resources and prevent high-risk cases.

ML helps data-driven CVD prediction (Dritsas and Trigka, 2023). Their study employed supervised ML to create excellent CVD prediction models, underlining the SMOTE data augmentation technique's advantages for class imbalance classification challenges. Using a Kaggle CVD dataset, age, gender, BMI, Sys BP, Dias BP, glucose, smoking, alcohol consumption, physical activity, and CVD are input into ML models to determine their CVD prediction value. Binary classification problems with uniform class probability distributions train and test these models. ML models with and without SMOTE compare accuracy, recall, precision, and AUC. The stacking model with SMOTE and 10-fold cross-validation has the greatest accuracy, recall, precision, and AUC of 0.87, 0.88, 0.87 and 0.98 respectively.

Ashfaq et al. (2022) suggested an IoMT-based remote patient tracking systems that uses AI and edge computing to constantly monitor vital signs of CVD patients and communicate their wellness status to authorised users. The MIMIC-III clinical database supplied data on heart rate (BPM), blood oxygen saturation (SpO2) and body temperature (°C) used for the ML model. K-Nearest Neighbours (KNNs) algorithm gave the greatest test set accuracy of 0.96. Precision, recall, and F1 score suggest KNN performs best at 0.86, 0.99, and 0.92 respectively. The smart, IoT, and ML solution seeks to reduce CVD patients' high death rate and enhance their quality of life. The authors suggested adding ECG, blood pressure, and breathing sensors to this project.

Finally, Khan et al. (2023) used DT, RF, LR, NBC, and SVM to classify and predict CVD patients in Lady Reading Hospital and the Khyber Teaching Hospital in Khyber Pakhtunkhwa (KPK) Province, Pakistan. Two hospitals' patients were randomly selected for exploring and experimental research. The CVD dataset includes age, gender, height, weight, cholesterol, glucose, smoking and drinking habits, physical activity, and BMI. CVD was categorised as "presence" or "absence". RF approach was best for CVD classification and prediction, with 0.85 accuracy, 0.92 precision, and 0.88 ROC curve. The approach might be used worldwide to classify and predict CVD.

3. Research Methodology

3.1 Machine Learning Pipeline

A comprehensive ML pipeline includes processes such as raw data collection, training, testing, prediction, and even retraining and algorithm for reuse. It is a strategy that seeks to improve the efficiency of ML models and the procedures involved in their creation. Typically, a ML pipeline is composed of five stacked components that are continually monitored by a monitoring pattern.

According to Stodt, Stodt, and Reich (2023), the ML algorithm consists of a series of steps, including data collection, followed by data pre-processing, model training that reuses historical data in parallel with other stages, ML usage in which the trained model is applied to cleaned-up data, and finally, the result, which signifies the end of the pipeline and displays the performance index result as the strength of ML model.

Figure 3.1: Machine Learning Pipeline (Source: Ashfaq et al., 2022)

3.2 Deep Learning (DL) Pipeline

The DL model infrastructure consists of data preparation, model structure design, train-the-model, test-the-model, and model refining. Attallah and Samir (2022) proposed a covid-19 diagnosis-based DL pipeline that prepares input data and generates training, validation, and testing datasets. The model framework is created by selecting the model type, number of layers, and activation functions. Using input data from the network layer, the trained model minimises loss. The model is then evaluated on a separate validation dataset utilising precision, recall, precision, etc. In DL, it is advantageous to enhance the model's structure or training parameters before reconfiguring the algorithm to satisfy performance requirements. The recommended stages for DL implementation are data preparation, model creation, training, evaluating, and fine-tuning.

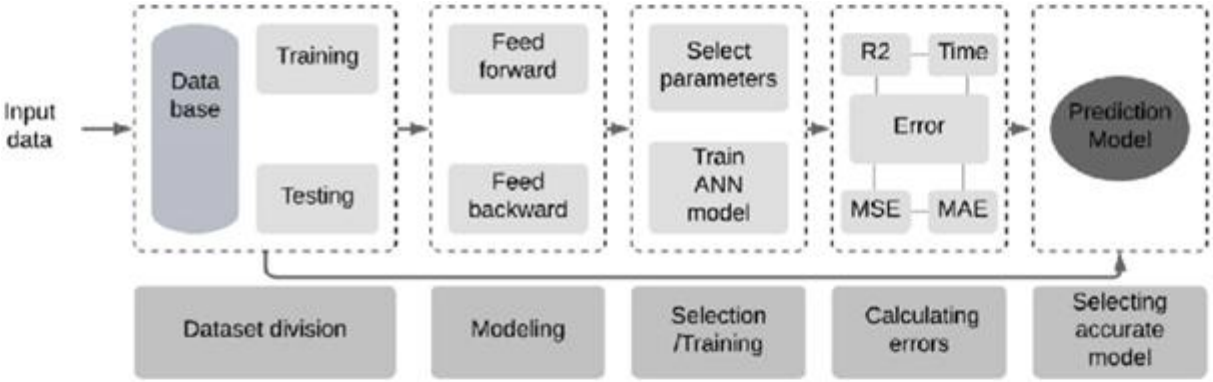


Figure 3.2: A standard pipeline for prediction using Deep Learning Pipeline. (Source: Qadir et al., 2021)

3.3 Machine Learning (ML) and Deep Learning (DL) Model

DL is a kind of ML that employs artificial neural networks and has been shown to be more successful than ML and other traditional data analysis techniques in many cases (Janiesch, Zschech, and Heinrich, 2021). The present research recommends employing both ML and DL algorithms. The two models will be built and performed in a Python environment. The objective is to investigate, via performance result, how the variables affect the prognosis of CVD and to compare the resulting models and algorithms to determine the optimal model and algorithm for CVD prediction.

3.3.1 Data Collection

Data collection is the systematic collecting and measurement of data for organised assessment, hypothesis testing, and research. Careful data collection allows researchers to make insightful observations, answer research questions, understand human experience, and critically assess present societal norms and practises.

Current investigations will utilise the clinical database MIMIC-III. According to Johnson, Pollard, and Mark (2016), the MIMIC-III database is an invaluable asset for academics due to the vast amount of deidentified clinical data collected from over 50,001 critically ill patients managed at Beth Israel Deaconess Medical Centre between 2001 and 2012. It involves the patient's demographic information, vital signs, lab test results, treatments, medications, attendant written notes, imaging reports, and post-hospital outcomes. Both Google's Big Query cloud and Amazon's AWS cloud host the database. This research will utilise Amazon S3 to gain access to and extract the required data for CVD forecasting.

3.3.2 Data Pre-Processing

Data pre-processing, as defined by (Chaki and Ucar, 2023), is the modification of the initial data before it is utilised for analysis or fed to a model or algorithms. The term "data pre-processing" is used to describe the steps used to clean, prepare, and alter raw data. The purpose of data pre-processing is to clean up the raw data so that useful information may be extracted about the target attributes and included into accurate models (Mishra et al., 2020). Pre-processing techniques have distinct aims depending on the nature of the artefacts that must be removed. According to (Kilkenny and Robinson, 2018), "garbage in-garbage out" suggests that sloppy data input might lead to unreliable results, highlighting the need of pre-processing data to ensure accuracy. A lack of precision in the data obtained might result in ineffective data analytics, applications, or procedures.

The following procedures will be used to prepare the data for this study:

3.3.2.1 Create Local SQL Server Database

The SQL database, also called RDBMS (Relational Database Management System), is where the SQL language is used to pull relevant data from relational databases (Khan et al., 2023). Amazon Web Services (AWS) has a SQL-based query tool called Athena that is common. Since the relational database had been hosted by AWS, therefore Athena utilised it to query it and get the necessary data locally. This local dataset had the five necessary human physiological factors, which were described in the previous chapters: pulse rate (BPM), blood pressure (mmHg), blood oxygen saturation (SpO_2), respiration rate (RR), and body temperature (°C). The information was then saved to a CSV file from the Athena interface. The AWS Athena interface is used to get the relevant information from the MIMIC-III database. The figures below show the SQL query and the results.

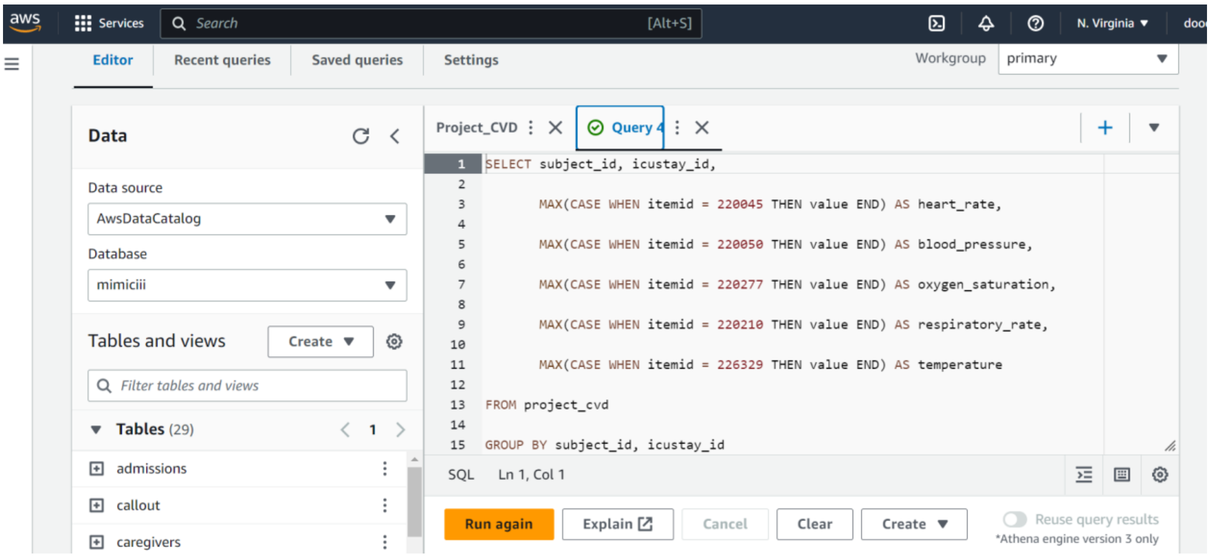


Figure 3.4: showing the SQL query used to extract the relevant dataset

3.3.2.2 Data Cleaning

Data cleansing, as emphasized by Woo et al. (2019), is a pivotal procedure that rectifies errors and inconsistencies in datasets, ensuring accuracy and reliability in insights and predictions. While high-quality data is fundamental for robust applications, discrepancies and gaps stemming from fractured sensors or human errors can hamper Machine Learning (ML) system performance (Lee et al., 2021). In the context of this study, where data gaps are evident, data cleansing becomes imperative. The process involves eliminating duplicates, rectifying data structure, handling anomalies, and addressing missing values through deletion or imputation, as highlighted by Ilyas and Rekatsins (2022). To counter the negative impact of null or inconsistent data on ML algorithms, the Python library pandas, specifically the Jupiter version, is employed to ensure data consistency and accuracy by identifying null values. For imputing missing data, the median of each column is utilized, thereby enhancing the dataset's integrity and suitability for analysis and prediction.

3.3.2.3 Dealing with Outliers

Data points that are outliers are those that are significantly different from the mean and standard deviation of the data set (Kwak and Kim, 2017). While outliers may be the result of human mistake in data input or in the way observations were recorded, they may also represent genuine data that sheds light on the data or indicates a pattern and hence should not be discarded (Gress, Denvir, & Shapiro, 2018). The data used in the present study comes from a trustworthy source, and no outliers will be removed throughout the training and testing phases of the ML and DL pipeline.

3.3.2.4 Data Transformation

The study employs data transformation, including mathematical alterations and feature scaling, to convert raw data into a usable dataset with labelled variables, aiming for accurate interpretation and pattern recognition (Lachlan, 2017; Singh & Singh, 2019).

3.3.3 Exploratory Data Analysis

Exploratory data analysis (EDA) should be a standard part of every study's methodology. The goal of EDA is to help researchers in finding trends in natural phenomena through studying data distribution, outliers, and abnormalities, and providing tools for hypothesis testing by means of visual analysis and comprehension of graphical data (Komorowski et al., 2016; Nicodemo and Satorra, 2020).

3.3.3.1 Univariate Analysis

The goal of univariate analysis is to provide an overview or account of the statistical distribution of just one variable across a given sample (Sandilands, 2014; Canova, Cortinovis, and Ambrogi, 2017). The density plot and overall statistics for each variable are shown for the current study using univariate analysis.

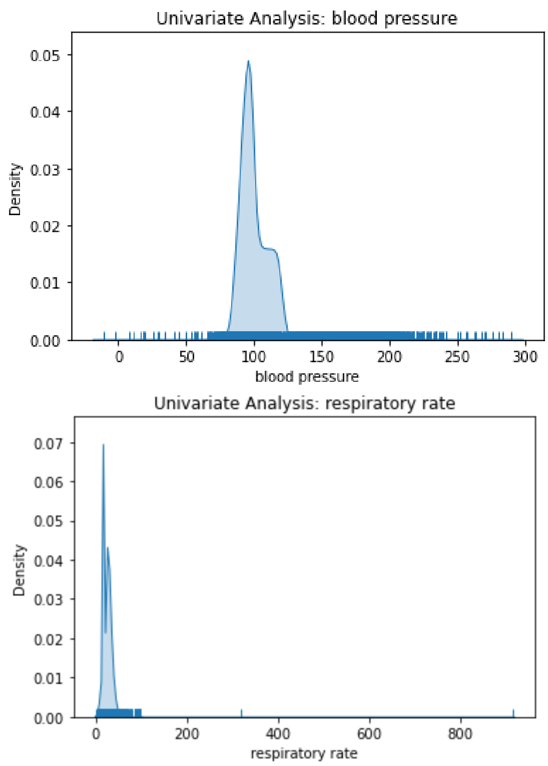


Figure 3.8: Showing the density plot of each feature.

3.3.3.2 Multivariate Analysis

The statistical method known as multivariate analysis allows for the simultaneous examination of numerous outcome variables. Research across several dimensions is possible with this strategy, as are analyses that consider the impact of all relevant factors on the results of interest (Mengual-Macenlle et al., 2015).

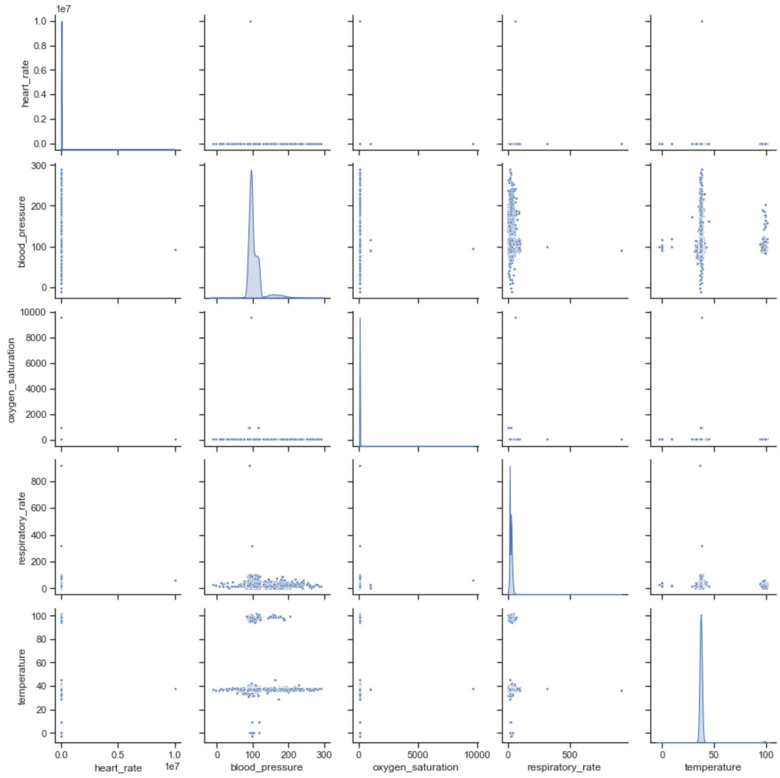


Figure 3.10: showing multivariate analysis using the scatter matrix

As can be seen in the image above, there is no obvious linear connection between the variables, hence they do not correlate with one another. This will be confirmed by further correlation procedures.

3.3.3.3 Correlation Analysis

Correlation analysis uses statistical methods to determine whether two sets of data are related. When two or more variables are highly linked, multicollinearity may cause problems for the ML model and bad results. According to (Kim, 2019), Multicollinearity arises when independent variables in a multiple regression model have a strong linear relationship, which may lead to misleading findings. 0.6 or higher correlation coefficients indicate a meaningful association between two or more variables, whereas 0.0 shows no relationship. Since there is no correlation, every variable will be explored in current study.

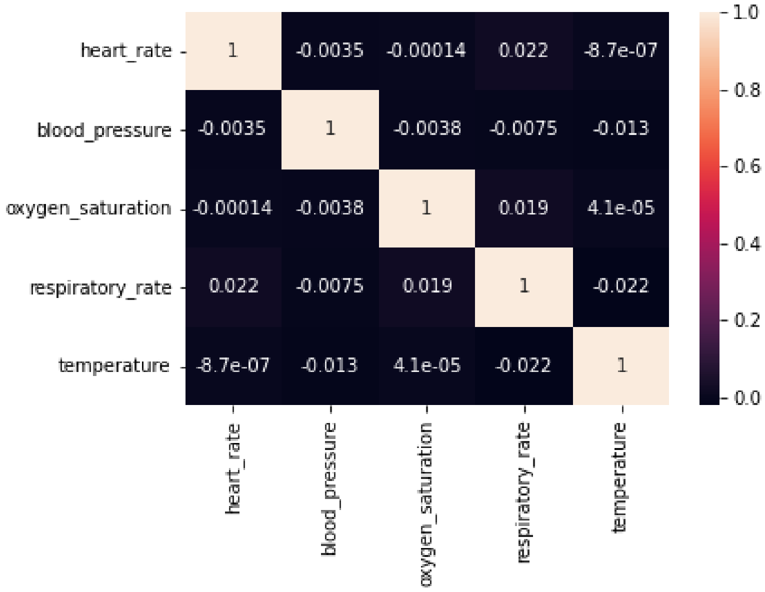


Figure 3.11: Showing the correlation analysis of the Dataset

3.4 Model Building

Different approaches to model construction are used depending on the intended use of the model or the availability of the dataset. Model construction, as defined by (Janiesch, Zschech, & Heinrich, 2021), includes different procedures for constructing ML and DL models. The present study outlines the process of model construction as follows:

3.4.1 Model Selection

According to Ghosh and Dasgupta (2022), a crucial step in using a ML and DL algorithm is selecting the appropriate model to use. There may be more than one algorithm that may be used to solve a given issue, which algorithm is ultimately chosen depends on several parameters, including the nature of the data, the desired degree of complexity, the available resources, and the statistical cost function.

3.4.1.1 Machine Learning Algorithms

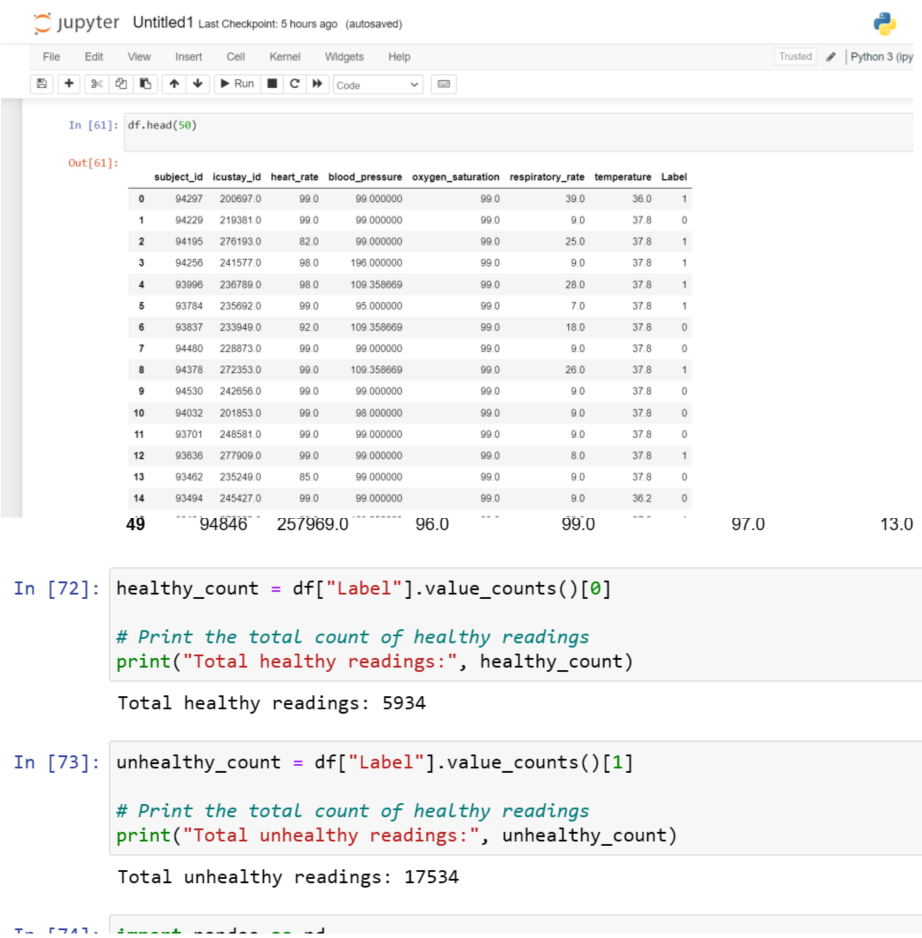
ML algorithms are mathematical models that uncover hidden trends in data, with RF and KNN, described as efficient and simple models for classification, being utilized in the study (Wei, 2019; Abraham et al., 2020; Breiman, 2001; Fawagreh et al., 2014; Schonlau and Zou, 2020; Zhang, 2016; Taunk et al., 2019).

3.4.1.2 Deep Learning Algorithms

The study will utilize Convolutional Neural Networks (CNN) and Multilayer Perceptron (MLP) methods, incorporating weight sharing, convolutional, pooling, and fully connected layers in CNN, along with backpropagation for training MLP, for effective deep learning in various fields like medicine, computer vision, and cybersecurity (Sarker, 2021; Albawi et al., 2017; Indolia et al., 2018; Nahiduzzaman et al., 2019).

3.4.2 Labelling the Data

The MIMIC-III database lacks a constant target category for ML models; hence, data is labeled to enable context for training, focusing on "Healthy" (0) and "Unhealthy" (1) binary classification based on accepted physiological ranges (Zhao et al., 2023), where parameters within normal limits are labeled "healthy" (0), and those outside are labeled "unhealthy" (1) using predefined ranges for various bodily factors (blood pressure, respiratory rate, heart rate, body temperature, and oxygen saturation) (a) 90/60-120/80 mmHg, (b) 12-18 breaths/min, (c) 60-100 beats/min, (d) 36.0-38.3°C, and (e) 95-100 saturation; Python is utilized for data tagging.


jupyter Untitled1 Last Checkpoint: 5 hours ago (autosaved)

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3 (ipy

```
In [61]: df.head(50)
```

Out[61]:

	subject_id	icustay_id	heart_rate	blood_pressure	oxygen_saturation	respiratory_rate	temperature	Label
0	94297	200697.0	99.0	99.000000	99.0	39.0	36.0	1
1	94229	219381.0	99.0	99.000000	99.0	9.0	37.8	0
2	94195	276193.0	82.0	99.000000	99.0	25.0	37.8	1
3	94256	241577.0	98.0	196.000000	99.0	9.0	37.8	1
4	93996	236789.0	98.0	109.358669	99.0	28.0	37.8	1
5	93784	235692.0	99.0	95.000000	99.0	7.0	37.8	1
6	93837	233949.0	92.0	109.358669	99.0	18.0	37.8	0
7	94480	228873.0	99.0	99.000000	99.0	9.0	37.8	0
8	94378	272353.0	99.0	109.358669	99.0	26.0	37.8	1
9	94530	242656.0	99.0	99.000000	99.0	9.0	37.8	0
10	94032	201853.0	99.0	98.000000	99.0	9.0	37.8	0
11	93701	248581.0	99.0	99.000000	99.0	9.0	37.8	0
12	93636	277909.0	99.0	99.000000	99.0	8.0	37.8	1
13	93462	235249.0	85.0	99.000000	99.0	9.0	37.8	0
14	93494	245427.0	99.0	99.000000	99.0	9.0	36.2	0

49 94846 257969.0 96.0 99.0 97.0 13.0

```
In [72]: healthy_count = df["Label"].value_counts()[0]

         # Print the total count of healthy readings
         print("Total healthy readings:", healthy_count)

         Total healthy readings: 5934

In [73]: unhealthy_count = df["Label"].value_counts()[1]

         # Print the total count of healthy readings
         print("Total unhealthy readings:", unhealthy_count)

         Total unhealthy readings: 17534
```

Figure 3.12: Showing the labelled dataset along Healthy (0) and Unhealthy (1) labels

Table 3.3: Summary of the Labelled Data

Summary	Healthy (0)	Unhealthy (1)
Total number of features	7	7
Total count	5934	17534
Dataset size 23468		

3.4.3 Train- Test Split

Data splitting involves separating a dataset into training and test subsets to evaluate ML and DL model predictive ability, with an 80-20 proportion for training and testing, as outlined by (Mücke et al., 2021).

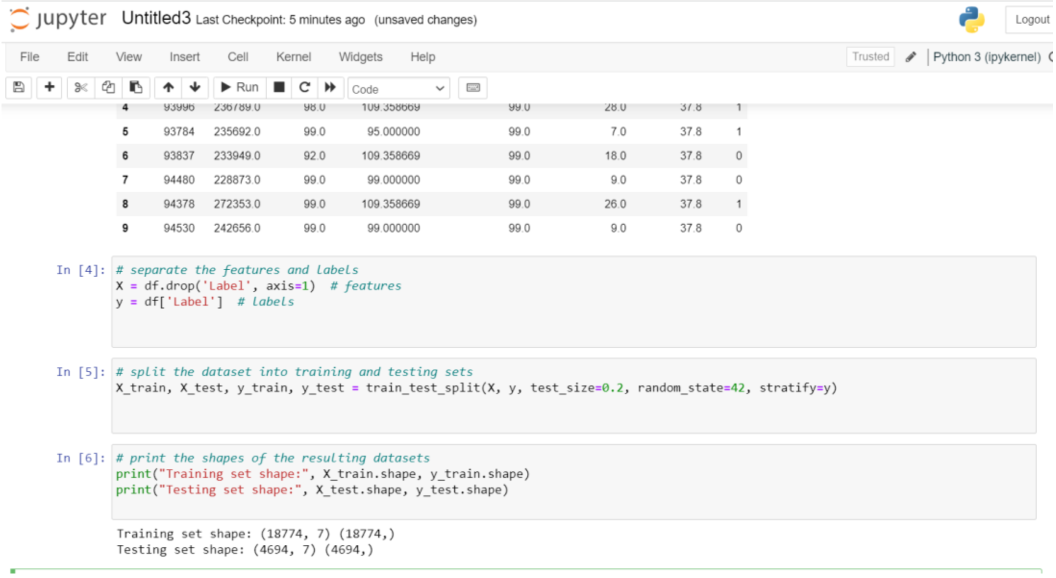


Figure 3.9: showing the Train Test Split of ratio 80:20 outcome of the dataset

Table 3.3: Summary of Train-Test Split

Summary	Dataset
Dataset size	23468
Total number of features	7
Training Dataset	18774
Test Dataset	4694

3.4.4 Model Hyperparameter Optimization

Efficient optimization algorithms are crucial for enhancing the effectiveness of machine learning (ML) across various approaches, as hyperparameter tuning plays a vital role in regulating training algorithms and improving ML model performance, especially for deep neural networks and computationally intensive models (Wu et al., 2019; Feurer & Hutter, 2019), with predictive ML models fine-tuned using strategies like "StandardScaler" and "GridSearchCV", while deep learning (DL) algorithms' performance can be refined by adjusting hyperparameters like layer count, neuron count, learning rate, regularization strength, and activation function.

3.4.5 Model Performance and Validation Metrics

In the context of binary classification, the research will use performance measures such as recall, precision, accuracy, and F1 score, along with the Confusion Matrix, to evaluate the effectiveness of the Machine Learning model (Erickson and Kitamura, 2021; Kulkarni, Chong, & Batarseh, 2020).

3.5 Chapter Summary

The chapter explores ML and DL models, pipelines, and algorithms, discussing data processing, thorough model creation, and evaluation steps crucial for trustworthy results, with the following chapter focusing on in-depth analysis of predictive abilities for enhanced subject understanding and future research insights.

4. Results and Analysis

4.1 Introduction

This chapter demonstrates the analysis established from the present study using the specified research methodology, as well as the methods used to achieve the findings of the corresponding ML and DL model predictions. Using the model performance and validation metrics, further comparisons will be made between the two models.

4.2 Machine Learning (ML) and Deep Learning (DL) Models Prediction and Results

The performance of KNN and RF, two ML models, and MLP and CNN, two DL models, are studied. Results will demonstrate the precision of the forecast and be compared to other metrics of measurement to choose the top performer. Appropriate measuring measures are used to assess the analysis and forecast outcomes. The following metrics will be used to measure the success of machine learning and deep learning projects.

- Accuracy: a quantitative representation of the model's success in making predictions. It is defined as the ratio of properly labelled data items to the total number of observations, as stated by AlZoman and Alenazi (2021). The corresponding equation (1) is shown in this section.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \dots\dots\dots 1$$

where, TP is the True Positive. Where, the positive class predicted by the model is correct.
TN is the True Negative. Where, the Negative class predicted by the model is correct.
FP is the False Positive. Where, the Positive class predicted by the model is incorrect.
FN is the False Negative. Where, the Negative class predicted by the model is incorrect.

- Testing accuracy: demonstrates the output of the model's measurement on a distinct testing dataset.

- Precision: this indicates the proportion of True Positives correctly predicted out of the total number of predicted positives. This is illustrated by the equation 2 below.

$$Precision = \frac{TP}{TP + FP} \dots \dots \dots 2$$

- Recall: this indicates the proportion of actual Positives that were accurately predicted among the actual Positives values. This is depicted by the equation 3 shown below.

$$Recall = \frac{TP}{TP + FN} \dots \dots \dots 3$$

- F1 score: this provides balance to the model's performance by taking the mean of both recall and precision into account. This is illustrated by the equation shown below.

$$F1\ score = \frac{2 \times Precision \times Recall}{Precision + Recall} \dots \dots \dots ..4$$

- Receiver Operating Characteristics of the Area, or ROC AUC by considering various data value ranges, the measurement of the True Positive rate or Recall versus the False Positive is shown under the curve. AUC, or area under the ROC curve, is a binary classification metric that, according to Melo (2013), assesses the performance of classifiers by calculating the area under the curve, with values ranging from 0.5 to 1.0, where 0.5 denotes a random classifier and 1.0 denotes an ideal classifier with no classification errors.
- Confusion matrix: According to Tiwari (2022), the confusion matrix is a structured brief of forecasting that displays the proportion of right and wrong guesses for each class, making it simple to determine which classes the model might be confusing with others. This summary also aids in evaluating the performance of the model for each class.

4.2.1 Result of K-Nearest Neighbours (KNN) Prediction

By using its techniques and model, KNN is used to train and test the dataset's predictive power. Hyperparameter optimisations like "StandardScaler" and "GridSearchCV" are used to enhance the outcome. Table 4.1 displays the outcome of a KNN model that has been optimised.

Table 4.1: Optimized KNN Prediction Performance

Performance metrics	Score
Accuracy	0.97
Precision	0.98
Recall	0.96
F1- score	0.97
ROC AUC	0.97

The performance of the KNN model across the measurement metrics is shown in Table 4.1.

With 97% accuracy, the model classified the data effectively. A precision of 0.98 means that the model successfully identified 98% of the relevant examples, and a recall of 0.96 means that the model accurately recognized 96% of the cases that should have been discovered. With a ROC AUC of 0.97 the model has a strong ability to differentiate between the two classes at 97%.

The confusion matrix is used to conduct further performance assessment, as depicted.

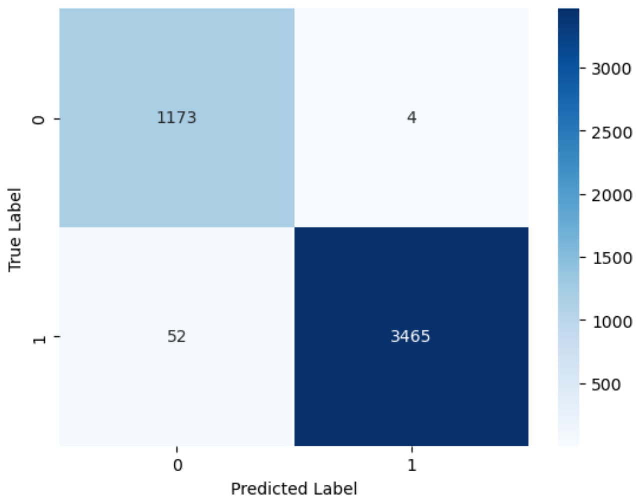


Figure 4.1: showing the confusion matrix output for KNN model.

According to the outcome of Figure 4.1, the healthy class is represented by (class 0), while the unhealthy class is shown by (class 1). The model is predicted as follows using the real labels and the predicted labels:

True Negative (TN): Of the examples that were healthy, the model correctly identified 1173 instances as being healthy (class 0). This indicates that the model had a recall of 99.7% and was able to identify 1173 out of 1177 real healthy instances.

True Positive (TP): Of the examples that were unhealthy, the model correctly identified 3465 occurrences as unhealthy (class 1) in total. This results in a recall of 98.5% since the model was able to identify 3465 out of 3517 unhealthy class.

False Positive (FP): The model incorrectly identified 4 cases as unhealthy (class 1) when they belong to the healthy class (class 0). This results to a precision outcome of 98% given the calculation from equation 2, the value is inputted as (3465/ (3465+4)).

False Negative (FN): shows the value 52, meaning that 52 cases were classified as healthy (class 0), when they belong to the unhealthy class (class 1). Using the same equation 2 to calculate the precision at values (3465/ (3465+52)), the result shows 98.5% precision of identifying healthy class.

4.2.2 Result of RF Prediction

RF is another high performing ML model used for classification purposes. RF algorithm and model is applied on the same dataset and the result is given below.

Table 4.2: RF Prediction Performance

Performance metrics	Score
Accuracy	0.98
Precision	0.99
Recall	0.98
F1- score	0.97
ROC AUC	0.97

Table 4.2 indicates that the RF model has a high accuracy of 98%, which indicates that it correctly classified 98% of all cases. The model's precision of 0.99, means that out of all the predicted cases, 99% were accurate. The recall of the

International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2022): 7.942

model is 0.98, indicating that it correctly identified 98% of the actual instances. The F1-score, which combines precision and recall, is 97% (0.97), indicating that the model performs well in both categories. At 97%, the ROC AUC score is also high, indicating that the model has a high capacity to distinguish between the class 0 and class 1.
The model will further be evaluated by the confusion matrix.

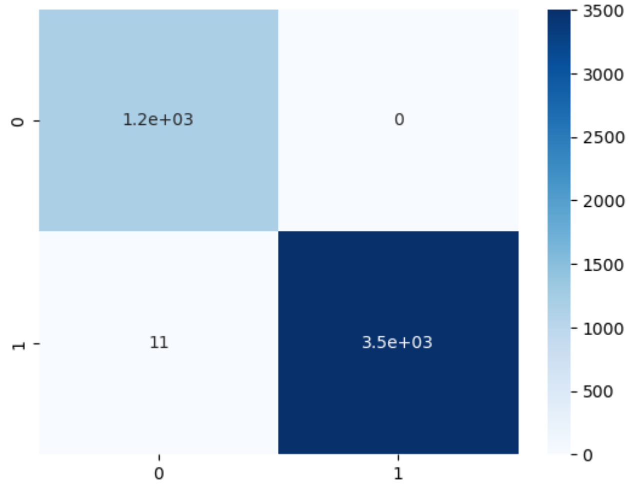


Figure 4.2: showing the confusion matrix of the RF model.

The RF model identified class 0 as healthy and class 1 as unhealthy, as shown in Figure 4.2. Using the true and predicted labels, the following represents the model's performance:

True Positive (TP): The model accurately predicted 3503 instances as unhealthy (class 1) out of all the genuinely unhealthy instances.

True Negative (TN): The model accurately predicted 1203 instances as healthy (class 0) out of all instances that were healthy.

False Negative (FN): The model incorrectly predicted 11 instances as healthy (class 0) when they were unhealthy (class 1), resulting in a precision of 99% using the values for equation 2, (3503/ (3503+11)).

False Positive at 0 means that the RF model predicted all the unhealthy class (class 1) as actual unhealthy, proving a 100% precision.

4.2.3 Result of Multi-Layer Perceptron (MLP) Prediction

Using the same dataset from current research, MLP is used to make prediction based on the input data. The result performance metrics is evaluated.

Table 4.3: MLP Prediction Performance

Performance metrics	Score
Accuracy	0.99
Testing Accuracy	0.97
Precision	0.99
Recall	0.99
F1- score	0.98
ROC AUC	0.98

The accuracy of 0.99 in Table 4.3 for the MLP model indicates that 99% of the predicted labels were accurate, while the testing accuracy of 0.97 indicates that the model also performed well on distinct dataset known as test data. A precision of 0.99 indicates that the model predicted 99% of the observations accurately. A recall of 0.99 means that the model correctly identifies 99% of the actual occurrences of the observation. The F1-score of 0.98 shows the balance or harmonic mean between the model's precision and recall which proves that the performance of the model is high at 98%. The ROC AUC value of 0.98 shows the overall capacity of the model to distinguish between class 1 and class 0 at 98%.

The confusion matrix output is displayed below.

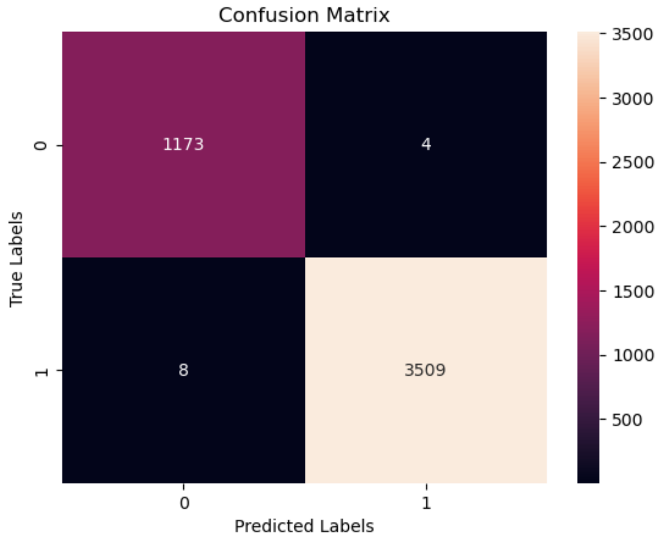


Figure 4.3: showing the confusion matrix of the MLP model

Figure 4.3 shows that 1173 true negatives (TN) mean that the model correctly predicted 1173 healthy data points (class 0) as true healthy, while 8 of false negatives (FN) shows that the model predicted 8 unhealthy class as healthy. 4 false positives (FP) means that the model predicted 4 healthy data points were unhealthy. However, 3509 true positives (TP) shows that the model predicted unhealthy class correctly, showing a precision of 99%. Briefly, MLP model performs well.

4.2.4 Result of Convolutional Neural Network (CNN) Prediction

CNN, another DL model is used to test for the prediction of the CVD using the same dataset of the current research.

```
587/587 [==============================] - 1s 3ms/step - loss: 0.2631 - accuracy: 0.9292
Epoch 9/10
587/587 [==============================] - 2s 3ms/step - loss: 0.2661 - accuracy: 0.9224
Epoch 10/10
587/587 [==============================] - 1s 3ms/step - loss: 0.2619 - accuracy: 0.9297
147/147 [==============================] - 0s 2ms/step
Accuracy: 0.9352364720920324
Confusion Matrix:
 [[1144  109]
 [ 195 3246]]
Classification Report:
              precision    recall  f1-score   support

           0       0.85      0.91      0.88      1253
           1       0.97      0.94      0.96      3441

    accuracy                           0.94      4694
   macro avg       0.91      0.93      0.92      4694
weighted avg       0.94      0.94      0.94      4694
```

Figure 4.4: shows the CNN model classification performance report.

According to Figure 4.4, the CNN model accurately identified 93.5% of the cases in the dataset. The precision for class 0 is 0.85, meaning that 85% of the occurrences which were predicted to be in class 0 were adequately classified. 91% of actual instances of class 0 were correctly predicted as such, according to the recall value of 0.91. The F1-score, or the average of precision and recall, is 0.88 (88%).

97% of the occurrences that were predicted to belong to class 1 were correct according to the precision for class 1. With a recall score of 0.94, 94% of real instances of class 1 were correctly identified as being in that class. Also, the average of precision and recall is represented by the F1-score, which is 0.96 (96%).

Further evaluation of MLP is explored through the confusion matrix below:

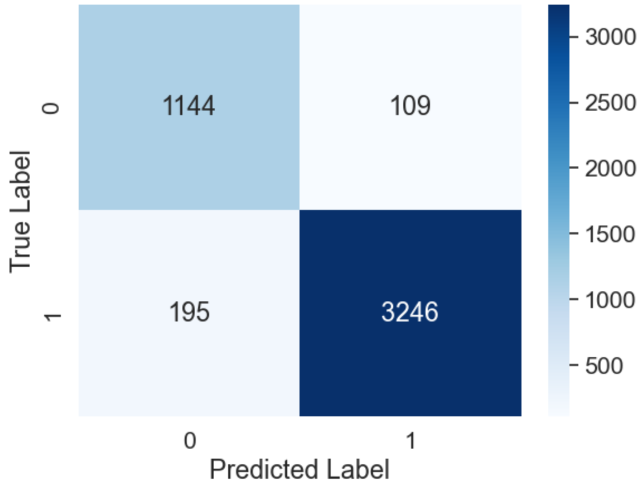


Figure 4.5: shows the confusion matrix of the CNN model.

Figure 4.5 demonstrates that the model made predictions on a total of 4694 instances, which were separated into two classifications. The row of the confusion matrix represents the true or actual classes, whereas the columns represent the predicted classes.

The first row and column (1144) indicate the number of instances that were correctly classified as "healthy" (0). The second row and second column (3246) indicate the total number of instances that were correctly assigned to the "unhealthy" class (1). They are known as genuine positives. The first row and second column (109) represent the number of instances that were misclassified as "unhealthy" when they belonged to the "healthy" class. This is known as a false positive.

The second row and first column (195) represent the number of instances that were misclassified as "healthy" when they belonged to the "unhealthy" class. This is known as a false negative.

Overall, the model correctly classified 4390 out of 4694 instances, or approximately 93.4% accuracy. The remaining 304 instances were incorrectly classified, resulting in an approximate 6.5% error rate.

4.3 Comparative Analysis of ML and DL Models

The table below displays the Comparative Analysis, which assesses the results of each model's individual output on the same input data, makes interpretations based on their similarities and differences, and compares their strengths, weaknesses, and overall performance potential.

Table 4.4: Comparative Analysis of ML and Dl Models

Algorithms	Accuracy	Precision	Recall	F1-score	ROC AUC
KNN	0.97	0.98	0.96	0.97	0.97
RF	0.98	0.99	0.98	0.97	0.97
MLP	0.99	0.99	0.99	0.98	0.98
CNN	0.93	0.92	0.93	0.92	0.92

A comparison of the four models based on different evaluation metrics from Table 4.4 shows the following results:

Accuracy: MLP has the highest accuracy at 0.99, followed by RF with 0.98 accuracy, KNN with 0.97 accuracy, and CNN with 0.935 accuracy.

With the same precision of 0.99, RF and MLP have the highest precision, indicating a high proportion of correctly identified positive cases. The precision of KNN is 0.98, which is also excellent. CNN's precision for class 0 is 0.85 while its precision for class 1 is 0.97.

With a recall of 0.99, MLP and RF have the highest ability to identify positive cases. The recall of KNN is 0.96, which is also excellent. CNN has a lesser class 0 recall of 0.91 and a higher class 1 recall of 0.94.

With an F1-score of 0.98, MLP has the highest F1-score, followed by RF with 0.97, KNN with 0.97, and CNN with a weighted average F1-score of 0.94. The F1-score is a strong indicator of a model's overall performance because it takes precision and recall into account.

MLP, RF, and KNN have exceptionally high ROC AUC ratings of 0.98 or 0.97, indicating a robust ability to distinguish between positive and negative classifications. CNN has a respectable ROC AUC score of 0.93.

Overall, MLP and RF exhibit the highest levels of precision, recall, F1-score, and ROC AUC. CNN has lower accuracy but performs reasonably well, particularly in terms of precision and recall for class 1.

4.4 Comparison of Result with Existing Studies

The current study's findings indicate that the MLP and RF models outperformed the CNN and KNN models, both in agreement with and contrast to previous research. According to (Swathy and Saruladha, 2021), MLP outperforms other ML models on the datasets (Cleveland database, UCI repository, publicly available heart, Disease dataset, Kaggle). In an additional existing work, (Nahiduzzaman et al., 2019) demonstrate that MLP obtained 0.91 accuracy with The Cleveland database from the UCI machine learning repository. In the University of California Irvine repository dataset, (Pal et al., 2022) confirm that the MLP performs better than KNN with an overall accuracy of 0.82.

In contrast to previous research by Ashfaq et al., (2022), KNN performs better on the MIMIC-III clinical database, however, current research shoes that MLP and RF outperform KNN, although more features/parameters were considered in the current study.

Table 4.5: showing the comparative analysis of existing work and current work on the same dataset.

Research	Dataset	Algorithms	Accuracy	Precision	Recall	F1-score
Ashfaq et al., (2022)	MIMIC-III clinical database	KNN	0.96	0.86	0.99	0.92
Current Research	MIMIC-III clinical database	MLP	0.99	0.97	0.99	0.98
Current Research	MIMIC-III clinical database	RF	0.98	0.99	0.98	0.97

Table 4.5 confirms the results of the existing work and compares it with the best performing model of the current research. The results indicates that present study achieved higher performance evaluation with the MLP and RF models against the KNN model of Ashfaq et al., (2022). However, current research has employed more feature inclusion in its analysis.

4.5 Chapter Summary

In this chapter, the capacity of ML and DL models to accurately predict CVD was compared by analysing the performance metrics of ML and DL models using the MIMIC-III clinical dataset. A comparative analysis of the results was conducted to determine which of the analysed models performed the best. Existing research results were also compared with the present study.

5. Critical Evaluation, Conclusion, and Recommendations

5.1 Introduction

This chapter presents a self-reflective evaluation of the research study's findings, including a summary, results, recommendations, and critical evaluation.

5.2 Critical Evaluation

The research aims to enhance cardiovascular disease (CVD) diagnosis, thereby advancing CVD mortality intervention and patient management; accessing the MIMIC-III database posed challenges, addressed through data cleaning, while literature review and existing studies informed ML and DL model selection, with MLP achieving superior accuracy, precision, recall, F1-score, and ROC AUC for CVD prediction, aligned with recommendations from Nahiduzzaman et al. (2019), Swathy and Saruladha (2017), and Pal et al. (2022); KNN's 97% performance surpasses RF's 94%, even as the current study employs 5 features instead of 3; research objectives are achieved through improved prediction performance of ML and DL models.

5.3 Ethical, Legal, Social, Professional and Security Issues

The dataset was obtained through procedures adhering to ethical, legal, social, and security considerations, following guidelines such as data protection statutes, HIPAA rules, and GDPR, while upholding professional standards and proper citation of resources (Sources: Ethical guidelines; Legal regulations; Data protection statutes; GDPR; HIPAA; Professional standards).

5.4 Conclusion, Limitation and Recommendations

This study compares ML and DL models for CVD prediction using human physiological parameters, addressing gaps in prior research. DL models incorporating five vital signs outperformed prior ML-only studies with three signs. MLP showed superior prediction accuracy. Limitations included dataset availability and quality from MIMIC-III. Previous research (Nahiduzzaman et al., 2019; Rim et al., 2020; Lee et al., 2022; Ashfaq et al., 2022) demonstrated successes in CVD prediction and DL analysis. While KNN excelled in ML, all four algorithms performed well, indicating varied performance across healthcare contexts, necessitating further investigation (Nahiduzzaman et al., 2019; Rim et al., 2020; Lee et al., 2022; Ashfaq et al., 2022).

5.5 Suggested Scope for Future Research

The observation of existing study suggests future efforts. More high-quality data on measured human physiological body signals are required. Quality data will enhance future study results and aid in the development of new CVD prediction models. More measured body vitals may be added to provide more informed findings in the prediction of CVD and other associated illnesses, particularly since the medical field is vast and numerous diseases might have identical symptoms. An electrocardiogram may be used to assess cardiac impulses and record them for analysis and decision-making in the prediction of CVD. Further research may add data mining approaches into current models to increase the accuracy and efficiency of CVD prediction by extracting relevant information from massive data sets and text data. This combination may lead to the creation of more accurate prediction systems. Above all, an Internet of things (IomT) prototype system that will use both AI and ML/DL will be improved by integrating the high performing model into sensor or wearable devices. This sensor can create predictions based on clever monitoring of the patient's vital signs to make rapid or trustworthy judgements.

References

[1] Abegaz, T.M., Shehab, A., Gebreyohannes, E.A., Bhagavathula, A.S. and Elnour, A.A. (2017). Nonadherence to antihypertensive drugs. *Medicine*, [online] 96(4), p.e5641. doi:https://doi.org/10.1097/md.0000000000005641.

[2] Abraham, G.K., Jayanthi, V.S. and Bhaskaran, P. (2020). *10 - Convolutional neural network for biomedical applications*. [online] ScienceDirect. Available at: https://www.sciencedirect.com/science/article/pii/B9780128206041000108.

[3] Acharya, U.R., Fujita, H., Oh, S.L., Hagiwara, Y., Tan, J.H., Adam, M. and Tan, R.S. (2018). Deep convolutional neural network for the automated diagnosis of congestive heart failure using ECG signals. *Applied Intelligence*, 49(1), pp.16–27. doi:https://doi.org/10.1007/s10489-018-1179-1.

[4] Ada's Medical Knowledge Team (2022). *Cardiovascular Disease Risk Factors «Ada*. [online] Ada. Available at: https://ada.com/cardiovascular-disease-risk-factors/.

[5] Al Bataineh, A. and Manacek, S. (2022). MLP-PSO Hybrid Algorithm for Heart Disease Prediction. *Journal of Personalized Medicine*, 12(8), p.1208. doi:https://doi.org/10.3390/jpm12081208.

[6] Albawi, S., Mohammed, T.A. and Al-Zawi, S. (2017). Understanding of a Convolutional Neural Network. *2017 International Conference on Engineering and Technology (ICET)*, [online] pp.1–6. doi:https://doi.org/10.1109/icengtechnol.2017.8308186.

[7] Ali R , et al. (2021). *Ethnic disparities in the major causes of mortality and their risk factors – a rapid review*. [online] GOV.UK. Available at: https://www.gov.uk/government/publications/the-report-of-the-commission-on-race-and-ethnic-disparities-supporting-research/ethnic-disparities-in-the-major-causes-of-mortality-and-their-risk-factors-by-dr-raghib-ali-et-al.

[8] AlZoman, R.M. and Alenazi, M.J.F. (2021). A Comparative Study of Traffic Classification Techniques for Smart City Networks. *Sensors*, 21(14), p.4677. doi:https://doi.org/10.3390/s21144677.

[9] Amegah, A.K. (2018). Tackling the Growing Burden of Cardiovascular Diseases in Sub-Saharan Africa. *Circulation*, 138(22), pp.2449–2451. doi:https://doi.org/10.1161/circulationaha.118.037367.

[10] Amini, M., Zayeri, F. and Salehi, M. (2021). Trend analysis of cardiovascular disease mortality, incidence, and mortality-to-incidence ratio: results from global burden of disease study 2017. *BMC Public Health*, 21(1). doi:https://doi.org/10.1186/s12889-021-10429-0.

[11] Anand, S., Bradshaw, C. and Prabhakaran, D. (2020). Prevention and management of CVD in LMICs: why do ethnicity, culture, and context matter? *BMC Medicine*, 18(1). doi:https://doi.org/10.1186/s12916-019-1480-9.

[12] Ashfaq, Z., Mumtaz, R., Rafay, A., Zaidi, S.M.H., Saleem, H., Mumtaz, S., Shahid, A., Poorter, E.D. and Moerman, I. (2022). Embedded AI-Based Digi-Healthcare. *Applied Sciences*, [online] 12(1), p.519. doi:https://doi.org/10.3390/app12010519.

[13] Attallah, O. and Samir, A. (2022). A wavelet-based deep learning pipeline for efficient COVID-19 diagnosis via CT slices. *Applied Soft Computing*, 128, p.109401. doi:https://doi.org/10.1016/j.asoc.2022.109401.

[14] Balakumar, P., Maung-U, K. and Jagadeesh, G. (2016). Prevalence and prevention of cardiovascular disease and diabetes mellitus. *Pharmacological Research*, 113, pp.600–609. doi:https://doi.org/10.1016/j.phrs.2016.09.040.

[15] Barry, T., Farina, J.M., Chao, C.-J., Ayoub, C., Jeong, J., Patel, B.N., Banerjee, I. and Arsanjani, R. (2023). The Role of Artificial Intelligence in Echocardiography. *Journal of Imaging*, [online] 9(2), p.50. doi:https://doi.org/10.3390/jimaging9020050.

[16] Bowry, A.D.K., Lewey, J., Dugani, S.B. and Choudhry, N.K. (2015). The Burden of Cardiovascular Disease in Low- and Middle-Income Countries: Epidemiology and Management. *Canadian Journal of Cardiology*, [online] 31(9), pp.1151–1159. doi:https://doi.org/10.1016/j.cjca.2015.06.028.

[17] Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), pp.5–32. doi:https://doi.org/10.1023/a:1010933404324.

[18] British Heart Foundation (2019a). *Cardiovascular heart disease*. [online] British Heart Foundation. Available at: https://www.bhf.org.uk/informationsupport/conditions/cardiovascular-heart-disease#:~:text=Cardiovascular%20disease%20(CVD)%2C%20also.

[19] Canova, S., Cortinovis, D.L. and Ambrogi, F. (2017). How to describe univariate data. *Journal of Thoracic Disease*, 9(6), pp.1741–1743. doi:https://doi.org/10.21037/jtd.2017.05.80.

[20] CDC (2022). *Heart disease Facts*. [online] Centers for Disease Control and Prevention. Available at: https://www.cdc.gov/heartdisease/facts.htm.

[21] Chaki, J. and Ucar, A. (2023). *Current Applications of Deep Learning in Cancer Diagnostics*. [online] *Google Books*. CRC Press. Available at: https://books.google.co.uk/books?hl=en&lr=&id=8AmmEAAAQBAJ&oi=fnd&pg=PA19&dq=data+preprocessing&ots=MrfVupCpuy&sig=O4sbCpvBiJq-1qIMIYztobDNRAU#v=onepage&q=data%20preprocessing&f=false [Accessed 12 Mar. 2023].

[22] Cleveland Clinic (2021). *Cardiovascular Disease: Symptoms, Types, Causes, Management & Prevention*. [online] Cleveland Clinic. Available at: https://my.clevelandclinic.org/health/diseases/21493-cardiovascular-disease.

[23] Dookie, K., Sanjiban, S.R., Tim, L., Ravinesh, D. and Pijush, S. (2018). *Handbook of Research on Predictive Modeling and Optimization Methods in Science and Engineering*. [online] *Google Books*. IGI Global. Available at: https://www.google.co.uk/books/edition/Handbook_of_Research_on_Predictive_Model/xdZdDwAAQBAJ?hl=en&gbpv=1&dq=predictive+research+designs&printsec=frontcover [Accessed 21 Feb. 2023].

[24] Dritsas, E. and Trigka, M. (2023). Efficient Data-Driven Machine Learning Models for Cardiovascular Diseases Risk Prediction. *Sensors*, 23(3), p.1161. doi:https://doi.org/10.3390/s23031161.

[25] Elman, C., Gerring, J. and Mahoney, J. (2020). *The Production of Knowledge: Enhancing Progress in*

Social Science. [online] *Google Books*. Cambridge University Press. Available at: https://books.google.co.uk/books?hl=en&lr=&id=vlTMDwAAQBAJ&oi=fnd&pg=PA17&dq=exploratory+research+designs&ots=lTtGf4Vd_n&sig=hLR3njycqgk6GK74fmzVvvphjmw#v=onepage&q=exploratory%20research%20designs&f=false [Accessed 21 Feb. 2023].

[26] Erickson, B.J. and Kitamura, F. (2021). Magician's Corner: 9. Performance Metrics for Machine Learning Models. *Radiology: Artificial Intelligence*, 3(3), p.e200126. doi:https://doi.org/10.1148/ryai.2021200126.

[27] Fawagreh, K., Gaber, M.M. and Elyan, E. (2014). Random forests: from early developments to recent advancements. *Systems Science & Control Engineering*, 2(1), pp.602–609. doi:https://doi.org/10.1080/21642583.2014.956265.

[28] Feurer, M. and Hutter, F. (2019). Hyperparameter Optimization. *Automated Machine Learning*, pp.3–33. doi:https://doi.org/10.1007/978-3-030-05318-5_1.

[29] Flowers, E., Froelicher, E.S. and Aouizerat, B.E. (2011). Gene-environment interactions in cardiovascular disease. *European Journal of Cardiovascular Nursing*, 11(4), pp.472–478. doi:https://doi.org/10.1016/j.ejcnurse.2011.06.001.

[30] Ghosh, S. and Dasgupta, R. (2022). Model Selection for Machine Learning. *Machine Learning in Biological Sciences*, pp.51–57. doi:https://doi.org/10.1007/978-981-16-8881-2_5.

[31] Goldstein, C.M., Gathright, E.C. and Garcia, S. (2017). Relationship between depression and medication adherence in cardiovascular disease: the perfect challenge for the integrated care team. *Patient preference and adherence*, [online] 11, pp.547–559. doi:https://doi.org/10.2147/PPA.S127277.

[32] Gress, T.W., Denvir, J. and Shapiro, J.I. (2018). Effect of Removing Outliers on Statistical Inference: Implications to Interpretation of Experimental Data in Medical Research. *Marshall Journal of Medicine*, 4(2). doi:https://doi.org/10.18590/mjm.2018.vol4.iss2.9.

[33] Hamid, S., Groot, W. and Pavlova, M. (2019). Trends in cardiovascular diseases and associated risks in sub-Saharan Africa: a review of the evidence for Ghana, Nigeria, South Africa, Sudan and Tanzania. *The Aging Male*, 22(3), pp.169–176. doi:https://doi.org/10.1080/13685538.2019.1582621

[34] Hassani, H., Silva, E.S., Unger, S., TajMazinani, M. and Mac Feely, S. (2020). Artificial Intelligence (AI) or Intelligence Augmentation (IA): What Is the Future? *AI*, [online] 1(2), pp.143–155. doi:https://doi.org/10.3390/ai1020008.

[35] Health and Safety Executives (HSE) (2017). *HSE - ART tool: Risk factors*. [online] Hse.gov.uk. Available at: https://www.hse.gov.uk/msd/uld/art/riskfactors.htm.

[36] Hsu, W., Warren, J.R. and Riddle, P.J. (2022). Medication adherence prediction through temporal modelling in cardiovascular disease management. *BMC Medical Informatics and Decision Making*, 22(1). doi:https://doi.org/10.1186/s12911-022-02052-9.

[37] Ilyas, I.F. and Rekatsinas, T. (2022). Machine Learning and Data Cleaning: Which Serves the Other? *Journal of Data and Information Quality*. doi:https://doi.org/10.1145/3506712.

[38] Indolia, S., Goswami, A.K., Mishra, S.P. and Asopa, P. (2018). Conceptual Understanding of Convolutional Neural Network- A Deep Learning Approach. *Procedia Computer Science*, [online] 132, pp.679–688. doi:https://doi.org/10.1016/j.procs.2018.05.069.

[39] Ivanov, S. and Webster, C. (2018). What Should Robots Do? A Comparative Analysis of Industry Professionals, Educators and Tourists. *Information and Communication Technologies in Tourism 2019*, [online] pp.249–262. doi:https://doi.org/10.1007/978-3-030-05940-8_20.

[40] Jacob, V., Reynolds, J.A., Chattopadhyay, S.K., Hopkins, D.P., Therrien, N.L., Jones, C.D., Durthaler, J.M., Rask, K.J., Cuellar, A.E., Clymer, J.M. and Kottke, T.E. (2022). Pharmacist Interventions for Medication Adherence: Community Guide Economic Reviews for Cardiovascular Disease. *American Journal of Preventive Medicine*, 62(3), pp.e202–e222. doi:https://doi.org/10.1016/j.amepre.2021.08.021.

[41] Jagannathan, R., Patel, S.A., Ali, M.K. and Narayan, K.M.V. (2019). Global Updates on Cardiovascular Disease Mortality Trends and Attribution of Traditional Risk Factors. *Current Diabetes Reports*, 19(7). doi:https://doi.org/10.1007/s11892-019-1161-2.

[42] Janiesch, C., Zschech, P. and Heinrich, K. (2021). Machine learning and deep learning. *Electronic Markets*, [online] 31. doi:https://doi.org/10.1007/s12525-021-00475-2.

[43] Johnson, A., Pollard, T. and Mark, R. (2016). *MIMIC Database v1.0.0*. [online] physionet.org. Available at: https://physionet.org/content/mimicdb/1.0.0/.

[44] Karatzia, L., Aung, N. and Aksentijevic, D. (2022). Artificial intelligence in cardiology: Hope for the future and power for the present. *Frontiers in Cardiovascular Medicine*, 9. doi:https://doi.org/10.3389/fcvm.2022.945726.

[45] Kebe, M., Gadhafi, R., Mohammad, B., Sanduleanu, M., Saleh, H. and Al-Qutayri, M. (2020). Human Vital Signs Detection Methods and Potential Using Radars: A Review. *Sensors*, 20(5), p.1454. doi:https://doi.org/10.3390/s20051454.

[46] Khan, A., Qureshi, M., Daniyal, M. and Tawiah, K. (2023a). A Novel Study on Machine Learning Algorithm-Based Cardiovascular Disease Prediction. *Health & Social Care in the Community*, [online] 2023, p.e1406060. doi:https://doi.org/10.1155/2023/1406060.

[47] Khan, M.M., Mehnaz, S., Shaha, A., Nayem, M. and Bourouis, S. (2021). IoT-Based Smart Health Monitoring System for COVID-19 Patients. *Computational and Mathematical Methods in Medicine*, 2021, pp.1–11. doi:https://doi.org/10.1155/2021/8591036.

[48] Khan, M.Z., Zaman, F.U., Adnan, M., Imroz, A., Rauf, M.A. and Phul, Z. (2023b). Comparative Case Study: An Evaluation of Performance Computation between SQL and NoSQL Database. *Journal of*

Software Engineering, [online] 1(2), pp.14–23. Available at: http://sjhse.smiu.edu.pk/sjhse/index.php/SJHSE/article/view/42 [Accessed 12 Mar. 2023].

[49] Kilkenny, M.F. and Robinson, K.M. (2018). Data quality: 'Garbage in – garbage out'. *Health Information Management Journal*, 47(3), pp.103–105. doi:https://doi.org/10.1177/1833358318774357.

[50] Kim, J.H. (2019). Multicollinearity and misleading statistical results. *Korean Journal of Anesthesiology*, 72(6). doi:https://doi.org/10.4097/kja.19087.

[51] Komorowski, M., Marshall, D.C., Salciccioli, J.D. and Crutain, Y. (2016). Exploratory Data Analysis. *Secondary Analysis of Electronic Health Records*, pp.185–203. doi:https://doi.org/10.1007/978-3-319-43742-2_15.

[52] Kulkarni, A., Chong, D. and Batarseh, F.A. (2020). *5 - Foundations of data imbalance and solutions for a data democracy*. [online] ScienceDirect. Available at: https://www.sciencedirect.com/science/article/abs/pii/B9780128183663000058 [Accessed 2 May 2023].

[53] Kwak, S.K. and Kim, J.H. (2017). Statistical data preparation: management of missing values and outliers. *Korean Journal of Anesthesiology*, 70(4), p.407. doi:https://doi.org/10.4097/kjae.2017.70.4.407.

[54] Lachlan, K.A. (2017). *The SAGE Encyclopedia of Communication Research Methods*. [online] *methods.sagepub.com*. SAGE Publications, Inc. Available at: https://methods.sagepub.com/reference/the-sage-encyclopedia-of-communication-research-methods/i4037.xml [Accessed 3 Apr. 2023].

[55] Lee, G.Y., Alzamil, L., Doskenov, B. and Termehchy, A. (2021). A Survey on Data Cleaning Methods for Improved Machine Learning Model Performance. *arXiv:2109.07127 [cs]*. [online] Available at: https://arxiv.org/abs/2109.07127.

[56] Lee, S.-J., Lee, S.-H., Choi, H.-I., Lee, J.-Y., Jeong, Y.-W., Kang, D.-R. and Sung, K.-C. (2022). Deep Learning Improves Prediction of Cardiovascular Disease-Related Mortality and Admission in Patients with Hypertension: Analysis of the Korean National Health Information Database. *Journal of Clinical Medicine*, [online] 11(22), p.6677. doi:https://doi.org/10.3390/jcm11226677.

[57] Lin, W.-H., Zhang, H. and Zhang, Y.-T. (2013). *Investigation on Cardiovascular Risk Prediction Using Physiological Parameters*. [online] Computational and Mathematical Methods in Medicine. Available at: https://www.hindawi.com/journals/cmmm/2013/272691/.

[58] Lindsey, H. (2022). *What Are Vital Signs, and Why Are They Important?* [online] Healthline. Available at: https://www.healthline.com/health/what-are-vital-signs.

[59] Lorenzoni, G., Sabato, S.S., Lanera, C., Bottigliengo, D., Minto, C., Ocagli, H., De Paolis, P., Gregori, D., Iliceto, S. and Pisanò, F. (2019). Comparison of Machine Learning Techniques for Prediction of Hospitalization in Heart Failure Patients. *Journal of Clinical Medicine*, [online] 8(9), p.1298. doi:https://doi.org/10.3390/jcm8091298.

[60] Malo, S., Aguilar-Palacio, I., Feja, C., Lallana, M.J., Rabanaque, M.J., Armesto, J. and Menditto, E. (2017). Different approaches to the assessment of adherence and persistence with cardiovascular-disease preventive medications. *Current Medical Research and Opinion*, 33(7), pp.1329–1336. doi:https://doi.org/10.1080/03007995.2017.1321534

[61] MedlinePlus (2015). *Vital signs: MedlinePlus Medical Encyclopedia*. [online] Medlineplus.gov. Available at: https://medlineplus.gov/ency/article/002341.htm.

[62] Melo, F. (2013). Area under the ROC Curve. *Encyclopedia of Systems Biology*, pp.38–39. doi:https://doi.org/10.1007/978-1-4419-9863-7_209.

[63] Mengual-Macenlle, N., Marcos, P.J., Golpe, R. and González-Rivas, D. (2015). Multivariate analysis in thoracic research. *Journal of Thoracic Disease*, [online] 7(3), pp.E2–E6. doi:https://doi.org/10.3978/j.issn.2072-1439.2015.01.43.

[64] Mensah, G.A., Roth, G.A. and Fuster, V. (2019). The Global Burden of Cardiovascular Diseases and Risk Factors. *Journal of the American College of Cardiology*, [online] 74(20), pp.2529–2532. doi:https://doi.org/10.1016/j.jacc.2019.10.009.

[65] Mishra, P., Biancolillo, A., Roger, J.M., Marini, F. and Rutledge, D.N. (2020). New data preprocessing trends based on ensemble of multiple preprocessing techniques. *TrAC Trends in Analytical Chemistry*, 132, p.116045. doi:https://doi.org/10.1016/j.trac.2020.116045.

[66] Mishra, R. and Monica (2019). Determinants of cardiovascular disease and sequential decision-making for treatment among women: A Heckman's approach. *SSM - Population Health*, 7, p.100365. doi:https://doi.org/10.1016/j.ssmph.2019.100365.

[67] Mohd Faizal, A.S., Thevarajah, T.M., Khor, S.M. and Chang, S.-W. (2021). A review of risk prediction models in cardiovascular disease: conventional approach vs. artificial intelligent approach. *Computer Methods and Programs in Biomedicine*, [online] 207, p.106190. doi:https://doi.org/10.1016/j.cmpb.2021.106190.

[68] Movsisyan, N.K., Vinciguerra, M., Medina-Inojosa, J.R. and Lopez-Jimenez, F. (2020). Cardiovascular Diseases in Central and Eastern Europe: A Call for More Surveillance and Evidence-Based Health Promotion. *Annals of Global Health*, 86(1). doi:https://doi.org/10.5334/aogh.2713.

[69] Mücke, N., Reiss, E., Rungenhagen, J. and Klein, M. (2021). Data splitting improves statistical performance in overparametrized regimes. *arXiv:2110.10956 [cs, stat]*. [online] Available at: https://arxiv.org/abs/2110.10956 [Accessed 14 Mar. 2023].

[70] Muthukumar, B., Dhanagopal, R. and Ramesh, R. (2019). KYP modeling architecture for cardiovascular diseases and treatments in healthcare institutions. *Journal of Ambient Intelligence and Humanized Computing*, 12(3), pp.3725–3734. doi:https://doi.org/10.1007/s12652-019-01653-z.

[71] Nahiduzzaman, Md., Nayeem, Md.J., Ahmed, Md.T. and Zaman, Md.S.U. (2019). Prediction of Heart

Disease Using Multi-Layer Perceptron Neural Network and Support Vector Machine. *2019 4th International Conference on Electrical Information and Communication Technology (EICT)*. doi:https://doi.org/10.1109/eict48899.2019.9068755.

[72] National Cancer Institute (2019). *NCI Dictionary of Cancer Terms*. [online] National Cancer Institute. Available at: https://www.cancer.gov/publications/dictionaries/cancer-terms/def/risk-factor.

[73] National Heart, Lung, and Blood Institute (2022). *Coronary Heart Disease - What Is Coronary Heart Disease? | NHLBI, NIH*. [online] www.nhlbi.nih.gov. Available at: https://www.nhlbi.nih.gov/health/coronary-heart-disease.

[74] National Intitute of Health and Care Excellence (NICE) (2023). *CKS is only available in the UK*. [online] NICE. Available at: https://cks.nice.org.uk/topics/cvd-risk-assessment-management/.

[75] NHS (2022). *Cardiovascular disease*. [online] NHS. Available at: https://www.nhs.uk/conditions/cardiovascular-disease/.

[76] Nicodemo, C. and Satorra, A. (2020). Exploratory data analysis on large data sets: The example of salary variation in Spanish Social Security Data. *BRQ Business Research Quarterly*, p.234094442095733. doi:https://doi.org/10.1177/2340944420957335.

[77] Nnate, D.A., Eleazu, C.O. and Abaraogu, U.O. (2021). Ischemic Heart Disease in Nigeria: Exploring the Challenges, Current Status, and Impact of Lifestyle Interventions on Its Primary Healthcare System. *International Journal of Environmental Research and Public Health*, 19(1), p.211. doi:https://doi.org/10.3390/ijerph19010211.

[78] OECD and Organization, W.H. (2022). *Mortality from cardiovascular disease*. [online] OECD iLibrary. Available at: https://www.oecd-ilibrary.org/social-issues-migration-health/health-at-a-glance-asia-pacific-2022_138f8d8f-en [Accessed 2 Mar. 2023].

[79] Pal, M., Parija, S., Panda, G., Dhama, K. and Mohapatra, R.K. (2022). Risk prediction of cardiovascular disease using machine learning classifiers. *Open Medicine*, 17(1), pp.1100–1113. doi:https://doi.org/10.1515/med-2022-0508.

[80] Porumb, M., Iadanza, E., Massaro, S. and Pecchia, L. (2020). A convolutional neural network approach to detect congestive heart failure. *Biomedical Signal Processing and Control*, [online] 55, p.101597. doi:https://doi.org/10.1016/j.bspc.2019.101597.

[81] Qadir, Z., Khan, S.I., Khalaji, E., Munawar, H.S., Al-Turjman, F., Mahmud, M.A.P., Kouzani, A.Z. and Le, K. (2021). Predicting the energy output of hybrid PV–wind renewable energy system using feature selection technique for smart grids. *Energy Reports*. doi:https://doi.org/10.1016/j.egyr.2021.01.018.

[82] Qureshi, N.Q., Mufarrih, S.H., Bloomfield, G.S., Tariq, W., Almas, A., Mokdad, A.H., Bartlett, J., Nisar, I., Siddiqi, S., Bhutta, Z., Mark, D., Douglas, P.S. and Samad, Z. (2021). Disparities in Cardiovascular Research Output and Disease Outcomes among High-, Middle- and Low-Income Countries – An Analysis of Global Cardiovascular Publications over the Last Decade (2008–2017). *Global Heart*, 16(1), p.4. doi:https://doi.org/10.5334/gh.815.

[83] Reber, E., Gomes, F., Vasiloglou, M.F., Schuetz, P. and Stanga, Z. (2019). Nutritional Risk Screening and Assessment. *Journal of Clinical Medicine*, [online] 8(7), p.1065. doi:https://doi.org/10.3390/jcm8071065.

[84] Rim, B., Sung, N.-J., Min, S. and Hong, M. (2020). Deep Learning in Physiological Signal Data: A Survey. *Sensors*, 20(4), p.969. doi:https://doi.org/10.3390/s20040969.

[85] Roth, G.A., Abate, D., Abate, K.H., Abay, S.M., Abbafati, C., Abbasi, N., Abbastabar, H., Abd-Allah, F., Abdela, J., Abdelalim, A., Abdollahpour, I., Abdulkader, R.S., Abebe, H.T., Abebe, M., Abebe, Z., Abejie, A.N., Abera, S.F., Abil, O.Z., Abraha, H.N. and Abrham, A.R. (2018). Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet*, [online] 392(10159), pp.1736–1788. doi:https://doi.org/10.1016/s0140-6736(18)32203-7.

[86] Roth, G.A., Mensah, G.A., Johnson, C.O., Addolorato, G., Ammirati, E., Baddour, L.M., Barengo, N.C., Beaton, A.Z., Benjamin, E.J., Benziger, C.P., Bonny, A., Brauer, M., Brodmann, M., Cahill, T.J., Carapetis, J., Catapano, A.L., Chugh, S.S., Cooper, L.T., Coresh, J. and Criqui, M. (2020). Global Burden of Cardiovascular Diseases and Risk Factors, 1990-2019: Update From the GBD 2019 Study. *Journal of the American College of Cardiology*, 76(25), pp.2982–3021. doi:https://doi.org/10.1016/j.jacc.2020.11.010.

[87] Sandilands, D. (2014). Univariate Analysis. *Encyclopedia of Quality of Life and Well-Being Research*, pp.6815–6817. doi:https://doi.org/10.1007/978-94-007-0753-5_3108.

[88] Sapra, A., Malik, A. and Bhandari, P. (2021). *Vital sign assessment*. [online] PubMed. Available at: https://www.ncbi.nlm.nih.gov/books/NBK553213/.

[89] Sarker, I.H. (2021). Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions. *SN Computer Science*, 2(6). doi:https://doi.org/10.1007/s42979-021-00815-1.

[90] Schonlau, M. and Zou, R.Y. (2020). The random forest algorithm for statistical learning. *The Stata Journal: Promoting communications on statistics and Stata*, 20(1), pp.3–29. doi:https://doi.org/10.1177/1536867x20909688.

[91] Siedlecki, S.L. (2020). Understanding Descriptive Research Designs and Methods. *Clinical Nurse Specialist*, [online] 34(1), pp.8–12. doi:https://doi.org/10.1097/nur.0000000000000493.

[92] Singh, D. and Singh, B. (2019). Investigating the impact of data normalization on classification performance. *Applied Soft Computing*, p.105524. doi:https://doi.org/10.1016/j.asoc.2019.105524.

[93] Stewart, J., Manmathan, G. and Wilkinson, P. (2017). Primary prevention of cardiovascular disease: A review of contemporary guidance and literature. *JRSM Cardiovascular Disease*, [online] 6(1), p.204800401668721. doi:https://doi.org/10.1177/2048004016687211.

[94] Stodt, F., Stodt, J. and Reich, C. (2023). Blockchain Secured Dynamic Machine Learning Pipeline for Manufacturing. *Applied Sciences*, 13(2), p.782. doi:https://doi.org/10.3390/app13020782.

[95] Swathy, M. and Saruladha, K. (2021). A comparative study of classification and prediction of Cardio-Vascular Diseases (CVD) using Machine Learning and Deep Learning techniques. *ICT Express*. doi:https://doi.org/10.1016/j.icte.2021.08.021.

[96] Taunk, K., De, S., Verma, S. and Swetapadma, A. (2019). *A Brief Review of Nearest Neighbor Algorithm for Learning and Classification*. [online] IEEE Xplore. doi:https://doi.org/10.1109/ICCS45141.2019.9065747.

[97] Tiwari, A. (2022). *Chapter 2 - Supervised learning: From theory to applications*. [online] ScienceDirect. Available at: https://www.sciencedirect.com/science/article/abs/pii/B9780128240540000265.

[98] Townsend, N., Kazakiewicz, D., Lucy Wright, F., Timmis, A., Huculeci, R., Torbica, A., Gale, C.P., Achenbach, S., Weidinger, F. and Vardas, P. (2021). Epidemiology of cardiovascular disease in Europe. *Nature Reviews Cardiology*, 19. doi:https://doi.org/10.1038/s41569-021-00607-3.

[99] Tran, B.X., Latkin, C.A., Vu, G.T., Nguyen, H.L.T., Nghiem, S., Tan, M.-X., Lim, Z.-K., Ho, C.S.H. and Ho, R.C.M. (2019). The Current Research Landscape of the Application of Artificial Intelligence in Managing Cerebrovascular and Heart Diseases: A Bibliometric and Content Analysis. *International Journal of Environmental Research and Public Health*, 16(15), p.2699. doi:https://doi.org/10.3390/ijerph16152699.

[100] Tripoliti, E.E., Papadopoulos, T.G., Karanasiou, G.S., Naka, K.K. and Fotiadis, D.I. (2017). Heart Failure: Diagnosis, Severity Estimation and Prediction of Adverse Events Through Machine Learning Techniques. *Computational and Structural Biotechnology Journal*, 15, pp.26–47. doi:https://doi.org/10.1016/j.csbj.2016.11.001.

[101] Tsao, C.W., Aday, A.W., Almarzooq, Z.I., Alonso, A., Beaton, A.Z., Bittencourt, M.S., Boehme, A.K., Buxton, A.E., Carson, A.P., Commodore-Mensah, Y., Elkind, M.S.V., Evenson, K.R., Eze-Nliam, C., Ferguson, J.F., Generoso, G., Ho, J.E., Kalani, R., Khan, S.S., Kissela, B.M. and Knutson, K.L. (2022). Heart Disease and Stroke Statistics—2022 Update: A Report From the American Heart Association. *Circulation*, [online] 145(8). doi:https://doi.org/10.1161/cir.0000000000001052.

[102] Verma, L., Srivastava, S. and Negi, P.C. (2016). A Hybrid Data Mining Model to Predict Coronary Artery Disease Cases Using Non-Invasive Clinical Data. *Journal of Medical Systems*, 40(7). doi:https://doi.org/10.1007/s10916-016-0536-z.

[103] Wei, L.K. (2019). *Computational Epigenetics and Diseases*. [online] ScienceDirect. Available at: https://www.sciencedirect.com/book/9780128145135/computational-epigenetics-and-diseases [Accessed 21 Mar. 2023].

[104] Wein, H. (2017). *Understanding Health Risks*. [online] NIH News in Health. Available at: https://newsinhealth.nih.gov/2016/10/understanding-health-risks.

[105] WHO (2022). *Cardiovascular Diseases*. [online] World Health Organization. Available at: https://www.who.int/health-topics/cardiovascular-diseases#tab=tab_1.

[106] Windisch, P., Schröder, C., Förster, R., Cihoric, N. and Zwahlen, D.R. (2023). Accuracy of the Apple Watch Oxygen Saturation Measurement in Adults: A Systematic Review. *Cureus*. doi:https://doi.org/10.7759/cureus.35355.

[107] Woo, H., Kim, K., Cha, K., Lee, J.-Y., Mun, H., Cho, S.J., Chung, J.I., Pyo, J.H., Lee, K.-C. and Kang, M. (2019). Application of Efficient Data Cleaning Using Text Clustering for Semistructured Medical Reports to Large-Scale Stool Examination Reports: Methodology Study. *Journal of Medical Internet Research*, 21(1), p.e10013. doi:https://doi.org/10.2196/10013.

[108] World Health Organization: WHO (2022). *Cardiovascular diseases*. [online] Who.int. Available at: https://www.who.int/health-topics/cardiovascular-diseases/#tab=tab_1.

[109] Wu, J., Chen, X.-Y., Zhang, H., Xiong, L.-D., Lei, H. and Deng, S.-H. (2019). Hyperparameter Optimization for Machine Learning Models Based on Bayesian Optimizationb. *Journal of Electronic Science and Technology*, [online] 17(1), pp.26–40. doi:https://doi.org/10.11989/JEST.1674-862X.80904120.

[110] Zaidi, S. and Brueckner, M. (2017). Genetics and Genomics of Congenital Heart Disease. *Circulation Research*, [online] 120(6), pp.923–940. doi:https://doi.org/10.1161/CIRCRESAHA.116.309140.

[111] Zhang, Z. (2016). Introduction to machine learning: k-nearest neighbors. *Annals of Translational Medicine*, [online] 4(11), pp.218–218. doi:https://doi.org/10.21037/atm.2016.03.37.

[112] Zhao, D. (2021). Epidemiological Features of Cardiovascular Disease in Asia. *JACC: Asia*, 1(1), pp.1–13. doi:https://doi.org/10.1016/j.jacasi.2021.04.007.

[113] Zhao, X., An, Y., Xu, N. and Geng, X. (2023). Continuous label distribution learning. *Pattern Recognition*, 133, p.109056. doi:https://doi.org/10.1016/j.patcog.2022.109056.