

Use of Machine Learning Model for Improving Cardiovascular Condition Using Cloud Computing

MSc Research Project
MSc in Cloud Computing

Tanuj Kumar
Student ID: X20228821

School of Computing
National College of Ireland

Supervisor: Dr. Shivani Jaswal

National College of Ireland
MSc Project Submission Sheet
School of Computing



StudentTanuj Kumar.....
Name:
Student ID:X20228821.....
Program:MSc in Cloud Computing..... **Year:** 2023.....
Module:MSc Research Project.....
Supervisor:Dr. Shivani Jaswal.....
Submission
Due Date:25th May2023.....
Project Use of Machine Learning Model for Improving Cardiovascular Condition
Title: Using Cloud Computing
Word Count5699..... **Page Count**.....19.....

I hereby certify that the information contained in this (my submission) is information about research I conducted for this project. All information other than my contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use another author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:Tanuj Kumar.....
Date:25th April 2023.....

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	<input type="checkbox"/>
Attach a Moodle submission receipt of the online project submission , to each project (including multiple copies).	<input type="checkbox"/>
You must ensure that you retain a HARD COPY of the project , both for your reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on a computer.	<input type="checkbox"/>

Assignments that are submitted to the Programme Coordinator's Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	

Table of Contents

1	Introduction.....	3
1.1	Background.....	3
1.2	Research Aim.....	4
1.3	Research objectives.....	4
1.4	Research Question.....	5
1.5	Ethics declaration.....	5
2	Related Work (Literature review).....	5
2.1	Machine learning (ML) and deep learning (DL).....	5
2.2	Machine Learning for Improvement of Health care Services.....	6
2.2.1	Patient Intake.....	7
2.2.2	Radiology.....	7
2.2.3	Hematology.....	7
2.2.4	Neurology.....	7
2.2.5	Oncology.....	8
2.2.6	Cell Biology.....	8
2.2.7	Cardiology.....	8
2.3	Flask as the API for Machine Learning model deployment.....	8
2.4	Web App Technologies for the Machine learning model deployment.....	9
2.4.1	Flask.....	9
2.4.2	Stramlit.....	9
2.4.3	Django.....	9
2.5	Container image method for Machine learning model deployment.....	9
3	Research Methodology.....	11
3.1	Theoretical setup.....	11
4	Design Specification.....	13
4.1	Experiment setup.....	13
	Implementation.....	14
5	Evaluation.....	16
5.1	Experiment.....	16
5.2	Experiment.....	18
5.3	Experiment.....	18

5.4	Discussion	18
6	Conclusion and Future Work	19
	References	20

List of Figures

Figure 1:	Web application flowchart for machine learning model	8
Figure 2:	Flow chart for experiment	13
Figure 3:	Mathematical representation of logistic regression	13
Figure 4:	Correlation heatmap	16
Figure 5:	Pair plot of all the variables	17
Figure 6:	Relation between the different variables	18
Figure 7:	Relation between the different variables	18

Machine Learning Model To Improve the Cardiovascular Condition By Cloud Computing

Abstract

There are unsupervised, supervised, and reinforcement learning are three categories into which the machine learning algorithm can be divided. By taking into consideration the many variables in the dataset, the research is integrated to predict cardiac disease. The research involves the application of logistic regression and other heart disease prediction techniques. The study falls under the experimental research category and is connected to the deployment area. The logistic Regression algorithm is taken into consideration to predict the outcome. The research generates that the model can be utilized in web applications using Flask in the future based on the feature score. The user must enter their information to be analyzed by a machine learning model to forecast the status of heart disease. The model then returns the predicted status.

1 Introduction

1.1 Background

There are several ways that machine learning can be applied in the healthcare industry. Identifying the condition early on, it can also be used to enhance patient care. Machine learning analyses numerous events using historical data to make precise predictions. In this project, the heart disease prediction model is built using the machine learning algorithm, flask, HTML, and CSS. As heart disease is one of the most common causes of death worldwide. It is possible to significantly improve the chances of properly diagnosing and treating heart disease. Machine learning algorithms can be used to forecast heart disease

based on several risk factors. In this report, we'll go through how a web application that predicts heart sickness was made using Flask and machine learning models. The web application was made with the Flask Python web framework. User input is gathered on the front end, and machine learning models that have been trained to produce predictions are used in the back end. HTML, CSS, and JavaScript are used to develop the front end of the web application. The logistic regression algorithm can be used to predict heart disease. It also includes a form where various information about heart disease risk factors, such as age, sex, type of chest discomfort, resting

blood pressure, serum cholesterol levels, and more, is gathered. The data input on the form is sent to the backend

when the submit button is clicked. Finding the causes of heart and circulatory disease and predicting it early on will give patients more time to receive therapy, which is the aim of the research.

1.2 Research aim

The study has main aim is to make a machine-learning model to enhance cloud-based medical care.

1.3 Research objectives

- To build a logistic regression model for improving healthcare services in a cloud computing environment.
- To assess the function of cloud computing and machine learning in healthcare services.

1.4 Research Question

How the machine learning algorithm can be used for the prediction of heart disease to improve medical care?

1.5 Ethics declaration

The required ethical standards are upheld to conclude the research and preserve the integrity of the work. The study uses only verified sources to gather data; everything in it has been independently verified. Secondary sources included in the study include Kaggle, academic publications and journals, research articles, and company websites. The reference appears after every piece of information. Sensitive data will not be a part of the study. The research solely employs a syntax of programming languages to carry out programming, and the researchers themselves have devised every aspect of the development process. No patient-specific information is included in the study.

2 Related Work

2.1 Machine learning (ML) and Deep learning (DL)

Machine learning can be described as a system capable of learning based on problem-specific data to develop an analytical model for different purposes. Whereas DL is defined as the concept of machine learning that follows the concept of artificial neural networks. With the

help of historical data, machine learning can learn interesting and meaningful patterns and relationships. ML is integrated with the different types of data analysis which are descriptive, predictive, cognitive, and prescriptive analysis. Nowadays, the industry is incorporated with intelligent systems for the automation of the daily routine tasks that systems work with the cognitive capacity like a human. The ML can also be useful for the decision-making approach to increase employee retention, productivity, and engagement. The systems that are integrated with the machine learning or AI environment can be used for the problem-solving approach. In the last few decades, ML achieves remarkable advancement with the help of sophisticated algorithms of learning and they can be considered pre-processing techniques. In the evolution of machine learning capabilities, the ANN is considered the biggest advancement because it is associated with a higher learning capacity. Several techniques, including the KNN (K-Nearest Neighbor Algorithm),

Genetic Algorithms (GA), Decision Trees (DT), and Naive Bayes are used to categorize the severity of the condition.

From the perspective of Thomas, 2022, The application of cutting-edge technologies can enhance a variety of healthcare applications. Automation and innovation are adaptable because machine learning is seen as a component of artificial intelligence. ML can be used to manage patient information, create fresh, cutting-edge medical techniques, and improve the management of chronic diseases. Healthcare can be enhanced in several areas with machine learning, including disease treatment and prediction, diagnostic and medical imaging services, medication discovery, patient record management, and strong supply chains.

Panch et al., 2018 To make better decisions, explain how machine learning is used in the healthcare system. It is necessary to have a strong management system in place to ensure a better flow of healthcare services. The organization can create better policies because the usage of ml makes information processing efficient. ML is regarded as an effective tool for decision-making, it can be said.

Uddin et al., 2019, According to the needs of the organization, multiple algorithms can be utilized for the model's development. By taking into account various pieces of data, the author proposes various methods that can be utilized to predict diseases. Additionally, the author states that the naive Bayes classification and support vector machine algorithms are the two most used methods for predicting diseases. The study describes the use of algorithms and incorporates tagged data from several individuals. The algorithms used are Naive Bayes, Support Vector Machine, Decision Tree, and Random Forest. Ramalingam et al., 2018 say

that numerous algorithms can be applied to the prediction of cardiac disease. Principal component analysis can be applied to the dataset to improve its accuracy. The Naive Bayes, Decision Tree, K-Nearest Neighbor, Random Forest, Ensemble approach, and Support Vector Machine techniques are among those that the author defines as being useful for predicting heart disease. These methods can be used to forecast cardiovascular and heart disease.

2.2 Machine Learning for Improvement of Health care services

Alsuliman et al., 2020 define the term intelligence to develop an understanding of artificial intelligence. Intelligence can be defined as the ability to problem-solving, make consistently design, and creation of products or services in different business domains.

Whereas the term artificial intelligence means the ability of a system the interpretation of external data in the correct format to achieve the pre-defined task or goals. Machine learning is a part or main component of artificial intelligence which can be used for analyzing statistical data by creating different types of models. The algorithms of machine learning can be differentiated into different domains which are the supervised, unsupervised, and reinforcement machine learning methods. Machine learning can be used in the following domains to improve healthcare services. The reason behind defining these points is to develop an understanding of machine learning applications in medical science. With the help of these, the reader can understand the potential of machine learning in the healthcare sector.

2.2.1 Patient Intake

To collect patient data properly, advanced technologies can be used. The system must associate with the predefined questions. So the system can get the information with the help of the user and it leads to an easy process because everyone from the healthcare department can access the information.

2.2.2 Radiology

Different sections can be considered to predict the requirement of radiology such as the segmentation for images, labeling for the different elements of images, diagnostic detection, and dictating the final report of radiology.

2.2.3 Hematology

In this section, identification of cell diagnostic can be performed by considering the three major areas which are the analysis of bone marrow, analysis of peripheral blood, and immunophenotyping with the help of flow cytometry. The author defines the use of ADTree (alternating decision tree) for such prediction.

2.2.4 Neurology

Machine learning can also be used in the field of neurology for performing its diagnostic tools such as electroencephalography and provides prediction with the help of signals.

2.2.5 Oncology

AI and ML are considered effective approaches to fighting cancer disease because they are associated with gene therapy, engineered biotherapy, and molecule inhibitors.

2.2.6 Cell Biology

The image processing algorithms can be used for analyzing the high number of cells to screen the different images because machine learning is associated with manual adjustments.

2.2.7 Cardiology

It can also be used in performing heart surgery by analyzing electrical signals (Alsuliman et al., 2020).

2.3 Flask as the API for Machine Learning model deployment

The author defines the development of web applications to predict diabetes. The flask is nothing but interference between the machine learning model and web application so it can be stated that it is a useful approach for performing the functionality of prediction. For the best practice of web application, there is a requirement for developing the smart solution in the form of a web app that can be performed with the help of the micro-framework of Flask by integration with the best model. As it is known that deployment is necessary for the working machine learning model so it is necessary to develop an understanding of model deployment efficiently.

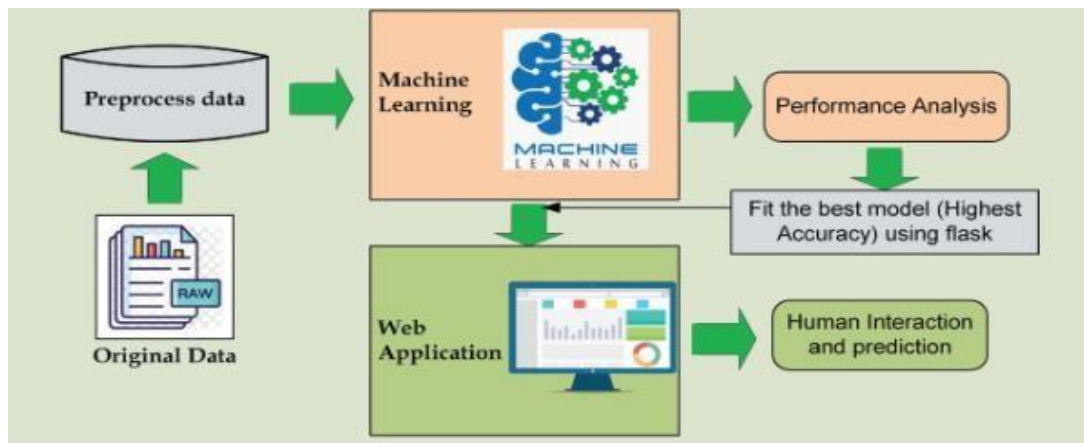


Figure 1: Web application flowchart for machine learning model

(Source: Ahmed et al., 2021)

2.4 Web App Technologies for the Machine learning model deployment

different frameworks can be used for the deployment of machine learning models which are listed below.

2.4.1 Flask

Flask is defined as the micro-framework to develop a website on a small scale. The flask allows the user to host the web application on the local server or cloud it is associated with different features such as the Jinja2 template for search engines, RESTful requests, and engine compatibility with Google, and it can easily deploy with production.

2.4.2 Streamlit

Streamlit is an open-source framework that does not require another language and it also defines as a tool for creating interfaces for the user. It is compatible with different machine learning libraries which are Keras, Scikit-learn, NumPy, matplotlib, Pandas, and latex. Streamlit directly calls the HTML code for the Python file.

2.4.3 Django

It is also a Python-based framework that helps create a professional environment during the development of web apps. It is associated with different features which are high scalability, association with SEO, web framework, versatile nature, and rapid development (Dani et al., 2022).

2.5 Container image method for Machine learning model deployment

There are the following stages required for the deployment of the model in the form docker image.

- The first stage requires the building of a model with the help of training and testing data.
- After that, there is a requirement for creating the API with Flask, Django, and Streamlit.
- The stage requires the containerizing of flask API with the help of docker.
- The fourth stage refers to the testing of the Flask application on the local server.
- Then the model will be deployed with the help of AWS Elastic Beanstalk.
- As the model is deployed successfully then it will be able to work (Openja et al., 2022).

Table 1: Summary of literature

Author	Research Area	Summary
Qayyum et al., 2020	Health care services	The healthcare industry allows a range of machine learning applications that can be used for analysis, data-driven approaches, and innovation in order to help patients.
Sarker, 2021	Machine learning	Machine learning is a process of learning the patterns of the business process, workflow, events, and individuals.
Wulf et al., 2021	Cloud computing environment	The cloud computing environment or service is defined as the availability of the resources such as computing ability and data storage.
Thomas, 2022	Improvement in healthcare services	ML can be used for handling the records of patients, developing new and innovative medical procedures, and better treatment of chronic diseases.
Uddin et al., 2019	Machine learning algorithms	The author provides the different algorithms that can be used for disease prediction by considering the different information
Aceto et al., 2020	cloud computing environment	Digital transformation at industry can be achieved with the help of the cloud computing environment.
Stančin and Jović, 2019	ML procedure	This section contains the life cycle of ML project from data extraction to model deployment.
Panesar, 2019	Model deployment at cloud	The author defines the different stages that are requires for the deployment of model.

3 Research Methodology

3.1 Theoretical setup

The research is integrated with the scientific procedure of machine learning. The machine is described as the subset of artificial intelligence that works with the common concepts of statistics and computer science. We are adapting the different machine learning algorithms which major one being Logistic regression. The research philosophy defines the knowledge, nature of the study, and different assumptions. During the execution of ML algorithms, there are different assumptions are considered for the experimental analysis. To perform the logistic regression, it can be considered that the dependent variable is considered dichotomous or binary, with no multicollinearity, presence of linearity, large sample size, and no outliers at an extreme level. Apart from the logistic regression, there are different algorithms considered which are the support vector machine, logistic regression, random forest algorithm, and gradient boost algorithms. All the algorithms of machine learning are associated with their assumption. This is the main reason behind the consideration of positivist philosophy. The research will determine the various methods and predicts the results so an analytical approach is taken for this. With the help of the two machine learning algorithms, there are two algorithms have been adapted that work for predictive analysis. Predictive data analysis is associated with the algorithms of machine learning and deep learning. The model of ML requires historical data that is trained with a specific period that works for delivering the business needs. The machine learning model can work for both regression and classification tasks. To complete the research, there is a range of algorithms is used that perform the data modeling to perform the predictive analytics As the previous related research is included in the literature review section with the help of the research strategy, so specific dimensions can be adapted. The research is associated with the complete life cycle of machine learning from business understanding to the deployment of the machine learning model on Amazon web service. There are different stages are considered such as the processing of data, data visualization, data modeling, and model deployment. It is well known that machine learning algorithms are works on historical data. For the analysis of the historical data, there is a requirement of statistical analysis that requires the concept of programming from the field of computer science and computation with mathematics. The research method includes the experiment, action research, survey, experiment, and grounded theory. There are different quantitative research methods are

used such as correlation, descriptive statistics, and so on. To carry out the research, secondary data is gathered. Research papers, journals, and websites are used to gather the information. Additionally, Kaggle is used to collect the dataset, and Python programming is used to conduct the qualitative analysis.

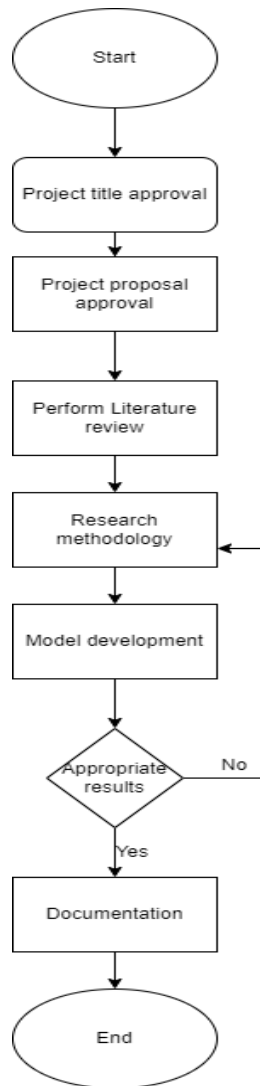


Figure 2: Flow chart for experiment

4 Design Specification

4.1 Experiment setup

- **Amazon Web Service:** It is further referred to as AWS and is a platform for cloud computing created by the Amazon Group. AWS offers a variety of services for tasks involving artificial intelligence and machine learning, such as developing, training, and deploying machine learning models. Projects using machine learning along with artificial intelligence can make use of a variety of advantages made possible by AWS.
 1. Logical capabilities for applications with APIs.
 2. Faster deployment of AI solutions for faster business processes.
 3. AI-written content investigation in the application.
 4. Building the bot by connecting from the different channels.
- **Google Colab:** It is a platform that enables users to write and run Python code in the browser while taking into account the following factors or advantages. For running the Python script, Google Colab offers an interactive web page.
 1. No need for the configuration.
 2. It gives free access to GPU.
 3. It has easy sharing (Google, 2022).
- **Logistic regression:** The definition of logistic estimate is the estimation of the possibility of a thing happening. Due to the probability values being inside the aforementioned range, the logistic regression's output is restricted to the range of 0 to 1. The statistical model known as the logit model, sometimes known as the logistic regression model, is defined. The formula for the logistic regression is:

$$\text{Logit}(\pi) = 1/(1 + \exp(-\pi))$$

$$\ln(\pi/(1-\pi)) = \text{Beta}_0 + \text{Beta}_1 * X_1 + \dots + B_k * K_k$$

Figure 3: Logistic regression formula

(Source: IBM, 2022)

- **FLASK:** Flask can be categorized as the microframework because it does not require any kind of external library or tools. It does not associate with the abstraction layer of the

database because it is already associated with common functions. The main aim behind introducing the flask is to develop the application efficiently. It can also be defined as the platform for creating the API.

- **HTML:** As the model is deployed with Amazon web service, then there is a requirement for presenting the model on the web page. HTML is a standard markup language that is used for displaying content on the web browser. It is used for the front-end section and it also provides the form section so, the user can choose the input variables.
- **CSS:** CSS stands for the cascading style sheet which works for creating stylish web pages or it can be stated that It is considered a useful option for a better user interface.

Implementation

- **Data Collection:** The data was gathered through the Kaggle data science community site. Numerous variables, including physical discomfort, headaches, and others, are included in the dataset. The entire machine learning model is created using the dataset, making it the main stage. To build a relationship, a column in the dataset that is present treats the events as such. The programmer can determine how the various events are related to one another.
- **Data processing:** In this step, the data is processed and manipulated to produce the intended outcomes. This task can be completed quickly because a variety of libraries are supported by the Python language. NumPy and pandas are two of the most utilized Python libraries. These libraries offer numerical computation along with the dataset because they are linked to many statistical functions.
- **Data visualization:** It is the process of presenting data and information graphically. The conclusions that arise from the evidence can be quickly detected if it is presented effectively. The ability to share information easily, the interactive depiction of data, and the ability to see links and trends are just a few advantages of data visualization. For the graphical display of the dataset, we'll utilize matplotlib. It is regarded as the NumPy library's graphical extension.
- **Data modeling:** Machine learning techniques that are used for model training and data modeling are both necessary. Because of the way the dataset is divided, the majority of it can be used to train prediction models. Using logistic regression, we can forecast the chance of a variable occurring. There is a need for a sklearn library that supports the

variety of packages to create the machine learning model(Vellido, 2020). SVM is another considerable algorithm for both regression and classification problems. It works by evaluating the hyperplane then it best separates the positive and negative data points in the data. SVM can handle non-linear relationships between the input variables and the result, and it is also effective in high-dimensional datasets. However, SVM can be computationally intensive and may require careful tuning of the parameters to achieve optimal results (Abdullah and Abdulazeez, 2021). Random forest classifiers consider the ensemble methods that combine multiple decisions based on decision trees to make predictions. The model is trained on random subsets of the data and features, which helps to reduce overfitting and improve accuracy. Random forests are robust and can handle missing data and outliers. However, they can be complex to interpret, and the computation time can be high for large datasets (Speiser et al., 2019).

- **Model evaluation:** Model evaluation is a procedure for determining a machine learning model's performance in terms of its shortcomings and strengths using various evaluation criteria. The cross-validation function will be used to assess the machine learning models and determine whether it has successfully decreased the overfitting of the dataset's data. Machine learning algorithms have the main three evaluation matrix.
 1. R2 score or R squared
 2. Root mean square error value (RMSE) / mean square error value (MSE)
 3. Mean absolute error value (MAE) (Hodson, 2022)
- **Deployment:** It may be said that AWS offers a complete range of machine learning services, making it a suitable option for the model's deployment. The platform offers the ability to process data, develop and test algorithms, and assess machine learning models. AWS supports the cross-validate model and the evaluated model, two different methods of evaluation. For putting the model into the cloud, the aforementioned tools are used.
- **Flask:** As the model is deployed with a machine learning algorithm, then we proceed with the model in the application that was developed with the help of Flask. In the first stage, we open the machine learning model and then create the application route. In the last stage, we created a function that returns the output in the form of an HTML page. The application contains the input from the HTML page then it returns the output in the same HTML page as the part of the prediction by analyzing the historic dataset. It is a Python-based framework that allows the developer for creating the API (Flask, 2021).

5 Evaluation

5.1 Experiment

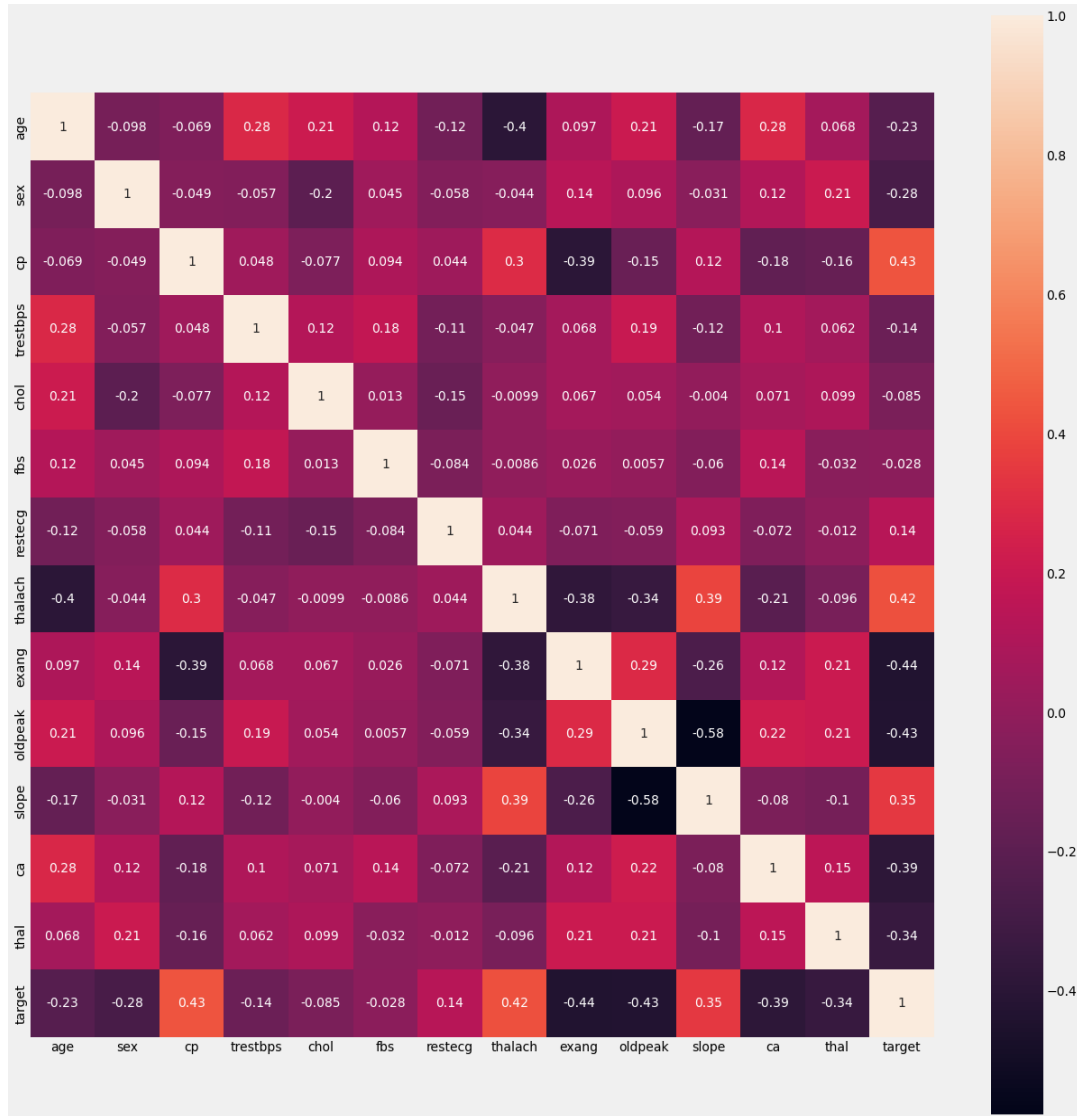


Figure 4: Correlation heatmap

The above-mentioned figure is the depiction of the correlation heatmap that defines the relation between the variables. It is considered a useful approach to selecting the best features with the help of the Pearson correlation coefficient. The values of the coefficient lie between -1 to 1. If the values are tending towards the +1 then it can be stated that the events are highly correlated and lead for similar effect in occurring of same event (Brittain et al., 2018).

5.2 Experiment

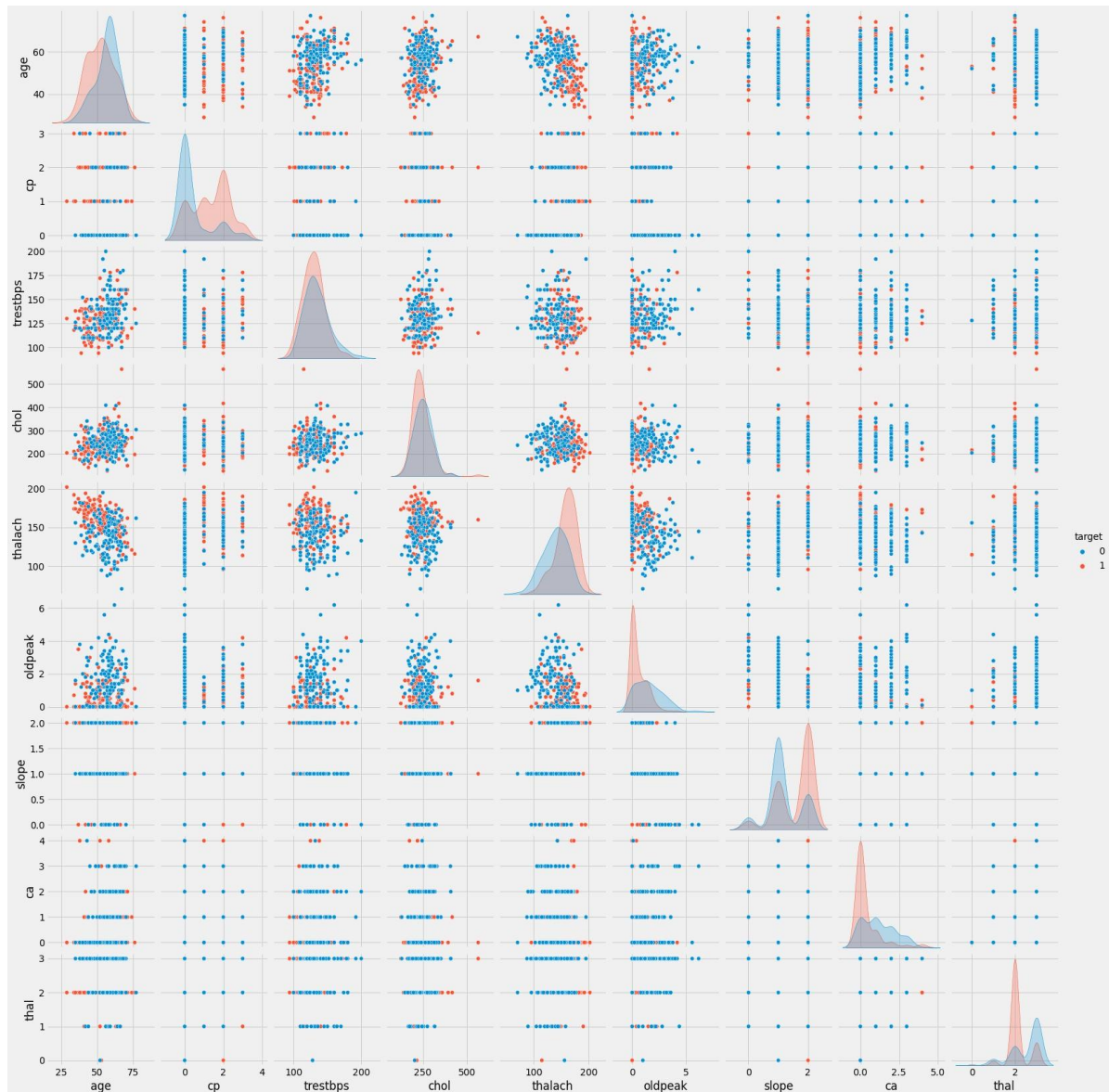


Figure 5: Pair plot of all the variables

The above figure is designed with the help of the seaborn library and it can be called the pair plot. It is a useful approach for plotting the relationship between the different events based on the pairs. It is a useful approach for summarizing the different visualization and it also helps to develop the understanding of data points. The Seaborn library allows the user to plot the entire data set in a few lines of code with different programming languages or analytical software. We developed it using the python language (Anisha et al., 2021).

5.2 Experiment

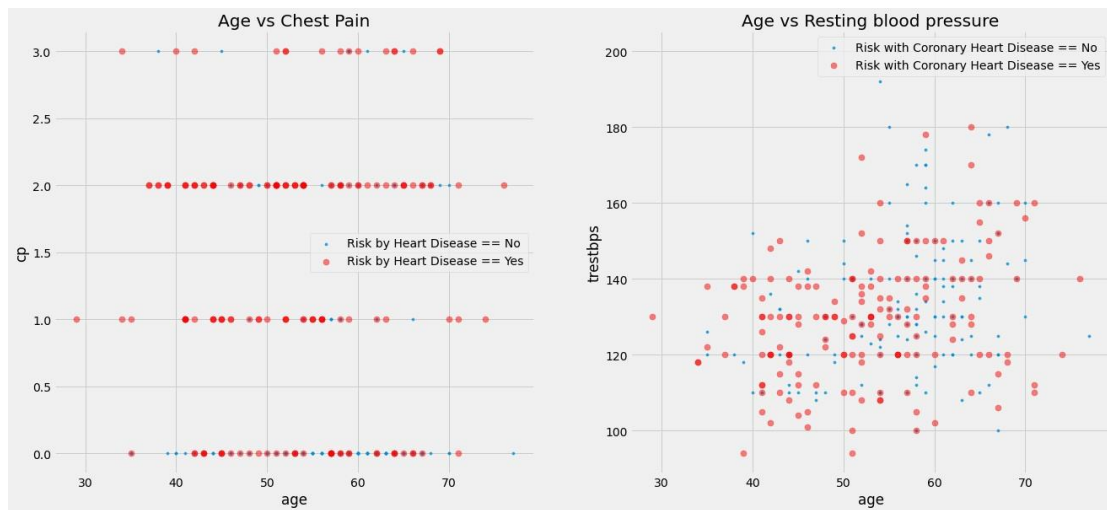


Figure 6: Relation between the different variable

5.3 Experiment

From the above correlation heatmaps and plotting the relation between the variables. Find the variables that are not useful in prediction.



Figure 7: Relation between the different variable

5.4 Discussion

This section gives a summary of the different experiments. Major of the experiments is performed for performing the data analysis to achieve descriptive and predictive analysis. Data visualization is considered a useful approach for defining the relationships among the different variables. At first correlation, the heatmap is plotted to visualize the relationship among the variables. It is used to pick the best variables and remove the un-useful variables. The Pair plot is used for an outlook of the different visualization and understanding of the data points. And from the above graph, a relation can be identified among the different variables.

We are finding the relationship between the variables using the correlation heatmap and by pair plotting the variables in the dataset of the heart disease prediction.

6 Conclusion and Future Work

From the above research, it can be concluded that machine learning and artificial intelligence can be used for improving the healthcare system. As is known that machine learning is associated with different types of analytics such as descriptive, predictive, prescriptive, and cognitive so it can be used for the prediction of heart disease based on the available physical characteristics of the user. The research is performed based on experiments that are performed with the help of machine learning algorithms. There are different stages are considered for the ML life cycle which is from business understanding to building the machine learning model on the cloud with the help of the AWS platform. There is a range of machine learning algorithms considered which major one is logistic regression. Apart from machine learning, there is a range of tools used to build the appropriate system which are Flask, HTML, and CSS. To deploy the model, we created a docker image and then push into AWS ECR then the code is successfully executed and it can provide the prediction based on the user physical characteristics.

As it is known that machine learning is associated with a different type of analytics so the research can also be used for prescriptive and cognitive modeling. With the help of the different types of analytics, the system which is developed in the research will be able to provide the reason and prevention technique for patients of heart disease. The task of prescriptive and cognitive modeling requires the additional dataset so the machine learning model can provide the prevention to the user. In terms of technology, the research can integrate with the better UI application by containing more features so different techniques can be used for the system such as the React.js, Node.js, and additional features of CSS. The research can also be used in the future for study purposes in different domains such as the

medical and technology. This report contains in-depth knowledge of medical functionality, machine learning algorithms, a technique for cloud deployment, and so on. The research will be integrated with the unsupervised learning method so models can learn by themselves.

Apart from the range of benefits and future potentials, the research is associated with some limitations. Eliminating that limitation can lead to the development of a robust system. There are two major issues identified during the research work which are the limited scope with the word count so we are not able to deep dive into the literature review section because it is considered the base approach for developing the understanding of business. Whereas another limitation is associated with the dataset. A dataset with a range of features and entries leads to the effective and accurate analysis.

References

Abdullah, D.M. and Abdulazeez, A.M., 2021. Machine learning applications based on SVM classification a review. *Qubahan Academic Journal*, 1(2), pp.81-90.

Aceto, G., Persico, V. and Pescapé, A., 2020. Industry 4.0 and health: Internet of things, big data, and cloud computing for Healthcare 4.0. *Journal of Industrial Information Integration*, 18, p.100129.

Ahmed, N., Ahammed, R., Islam, M.M., Uddin, M.A., Akhter, A., Talukder, M.A. and Paul, B.K., 2021. Machine learning-based diabetes prediction and development of smart web application. *International Journal of Cognitive Computing in Engineering*, 2, pp.229-241.

Alsuliman, T., Humaidan, D. and Sliman, L., 2020. Machine learning and artificial intelligence in the service of medicine: Necessity or potentiality? *Current research in translational medicine*, 68(4), pp.245-251.

Anisha, P.R., Reddy, C.K.K., Apoorva, K. and Mangipudi, C.M., 2021, April. Early Diagnosis of Breast Cancer Prediction using Random Forest Classifier. In *IOP Conference Series: Materials Science and Engineering* (Vol. 1116, No. 1, p. 012187). IOP Publishing.

Brittain, J., Cendon, M., Nizzi, J. and Pleis, J., 2018. Data scientist's analysis toolbox: Comparison of Python, R, and SAS Performance. *SMU Data Science Review*, 1(2), p.7.

Dani, H., Bhole, P., Waghmare, H., Munginwar, K. and Patil, A., 2022. Review on Frameworks Used for Deployment of Machine Learning Model.

Flask, D., 2021. Flask.

Google. 2022. *Google Colaboratory*. Google.com. [online] Available at: https://colab.research.google.com/#scrollTo=GJBs_flRovLc [Accessed 13 March 2022].

Hodson, T.O., 2022. Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not. *Geoscientific Model Development*, 15(14), pp.5481-5487.

Nazihovna, Y.G., 2022. CREATING A PLATFORM USING HTML, CSS, AND JAVASCRIPT METHODS AND STRENGTHENING EDUCATION WITH THIS STEAM. *Confrencea*, 5(5), pp.17-38.

Openja, M., Majidi, F., Khomh, F., Chembakottu, B. and Li, H., 2022, June. Studying the practices of deploying machine learning projects on docker. In *Proceedings of the International Conference on Evaluation and Assessment in Software Engineering 2022* (pp. 190-200).

Panch, T., Szolovits, P. and Atun, R., 2018. Artificial intelligence, machine learning, and health systems. *Journal of global health*, 8(2).

Panesar, A., 2019. *Machine learning and AI for healthcare* (pp. 1-73). Coventry, UK: Apress.

Qayyum, A., Qadir, J., Bilal, M. and Al-Fuqaha, A., 2020. Secure and robust machine learning for healthcare: A survey. *IEEE Reviews in Biomedical Engineering*, 14, pp.156-180.

Ramalingam, V.V., Dandapath, A. and Raja, M.K., 2018. Heart disease prediction using machine learning techniques: a survey. *International Journal of Engineering & Technology*, 7(2.8), pp.684-687.

Sarker, I.H., 2021. Machine learning: Algorithms, real-world applications, and research directions. *SN Computer Science*, 2(3), pp.1-21.

Speiser, J.L., Miller, M.E., Tooze, J. and Ip, E., 2019. A comparison of random forest variable selection methods for classification prediction modeling. *Expert systems with applications*, 134, pp.93-101.

Stančín, I. and Jović, A., 2019, May. An overview and comparison of free Python libraries

for data mining and big data analysis. In 2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO) (pp. 977-982). IEEE.

Tableau. 2022. *What Is Data Visualization? Definition, Examples, And Learning Resources*. [online] Available at: <https://www.tableau.com/learn/articles/data-visualization> [Accessed 13 March 2022].

Thomas, M., 2022. *16 Machine Learning in Healthcare Examples*. [online] Available at: <https://builtin.com/artificial-intelligence/machine-learning-healthcare> [Accessed 13 March 2022].

Thomas, M., 2022. *16 Machine Learning in Healthcare Examples*. [online] Available at: <https://builtin.com/artificial-intelligence/machine-learning-healthcare> [Accessed 23 Dec. 2022].

Uddin, S., Khan, A., Hossain, M.E. and Moni, M.A., 2019. Comparing different supervised machine learning algorithms for disease prediction. *BMC medical informatics and decision making*, 19(1), pp.1-16.

Vellido, A., 2020. The importance of interpretability and visualization in machine learning for applications in medicine and health care. *Neural computing and applications*, 32(24), pp.18069-18083.

Wulf, F., Lindner, T., Strahringer, S. and Westner, M., 2021. IaaS, PaaS, or SaaS? The Why of Cloud Computing Delivery Model Selection: Vignettes on the Post-Adoption of Cloud Computing. In *Proceedings of the 54th Hawaii International Conference on System Sciences*, 2021 (pp. 6285-6294).