Heart Disease Prediction System Using Machine Learning

Ranjit Shrestha[1] and Jyotir Moy Chatterjee[2]
[1] UG Student, Lord Buddha Education Foundation, Kathmandu, Nepal
[2] Assistant Professor (IT), Lord Buddha Education Foundation, Kathmandu, Nepal

Abstract

The major killer cause of human death is Heart Disease (HD). Many people die due to this disease. Lots of researchers have been discovering new technologies to prognosticate the disease early before it's too late for helping healthcare as well as people. These processes are still under research phase. Machine Learning (ML) is faster-emerging technology of Artificial Intelligence (AI) that contributes various algorithms for HD. Based on the proposed problem, ML provides different classification algorithms to divine the probability of patient having HD. For predicting HD, a lot of research scholars contributes their effort in this work using various techniques and algorithms such as Decision Tree (DT), Naïve Bayes (NB), Support Vector Machine (SVM), KNN (K-Nearest Neighbor), Neural Network (NN), etc. In order to give some effort on this work, we are going to develop a Web-based Heart Disease Prediction System (HDPS) by applying DT and NB ML algorithms. We are using the UCI repository HD dataset to train a model by comparing DT and NB algorithm for HDPS Web application. The dataset contains 303 instances with 14 attributes that help to train a prediction model that will be deployed into a web application for prediction. The main aim of this project is to build an efficient prediction model and deploy for prediction of disease. An HDP Model is built by using NB algorithm that provides 88.163% accuracy among others. A web-based HDPS application is developed through the waterfall model. Each phase is efficiently done. The project is successfully created with help of requirement analysis and project plan, system design, database design, testing plan, identifying features and functionalities, and system validation and deployment. The limitation of this project is to have only predicted the presence of heart disease but not identify which type of HD does have at patient. In future work, we can enhance the project by appending more detail prediction of HD at patient and incorporate with smart wear devices that integrate to Hospital Emergency System.

Keywords: Machine Learning (ML), Decision Tree (DT), Naïve Bayes (NB), Heart Disease, Classification.

1. INTRODUCTION

ML is an emerging application of AI that uses different analytics and statistical techniques in order to improve the performance of particular machine learning from old data. It enables a particular machine to learn from database and enhance the performance by experience. It helps to build an intelligent machine to solve the specific problem. ML solved a various complex problem that doesn't solve by statistic algorithms. ML provides dynamic algorithms which are being without explicitly program in order to build an intelligent machine (Kautish et al, 2008, 2012, 2013, 2020) that can easy various difficult problems. ML solved the different type of problem which is categorized into three part. Supervised, unsupervised, and reinforcement type problems. In supervised, there is two types of problems such as classification and regression problems. In unsupervised, there is clustering type problem can solved ML algorithms. ML assigned different algorithms based on their type of problem. ML project is done by the following steps:

- Defines a problem statement.
- Classifying the problem into ML problems.

- Selecting suitable ML algorithms based on their type of problems.
- Collecting and cleaning the data.
- Training a Model from data.
- Test the Model from test data
- Evaluate a model from their accuracy.

This work is closely related to the supervised problem of ML. However, many researchers have also solved this problem from unsupervised and NN. NN is a subset of ML that solved the complex problem which does not solve by normal ML. this project will solve by two classification algorithms of supervised ML such as DT and NB. This system will do by Python programming language which is handled entire development of this system using its ML's libraries. At the initial phase, a Heart Disease Prediction Model (HDPM) will build by one of both mentioned algorithms through the comparison of them. This process is done through the ML Process in order to build a model. After that, the model will deploy into web application by implanting Python Flask server-side libraries through the Waterfall Model. The proposed system used by Doctors that can access the system in order to decide whether the patient having HD or not. This system provides the level of HD presence such as no HD, having HD, and most likely having HD. This system has one admin user that manage and control the overall system and data of doctors and patients reports.

1.1 Background to the project

In the last fifteen years, HDs become still endured the leading causes of death. (WHO, 2019). In United States, Millions of human are having a HD every year so that the HD takes placed the biggest killer of people in the world. According to analyzation of WHO, twelve million people are death due to HD in worldwide. One person dies almost every 34 seconds from HD. (Patel Jaymin, 2016) Diagnosis of HDs is an essential task and yet intricate task to perform accurately and efficiently in the hospital and clinic. These things are motivated to build a Web-based HDPS application using ML algorithms. This proposed system can reserve problem by accurately predicting the presence of HD in the patient. ML is an emerging technology of AI that solved the various type of classification problem by producing accurate output. ML algorithms are applied to forecast the HD of the patient. This proposed system can be used either NB or DT algorithms by comparing their accurate result and trained time. A HDPM is built by employing one of the algorithms and deploying this model into a web application. The outcome of this project is to deliver a web-based application named HDPS that successfully and accurately predict the presence of HD of patients. This system is useful to support the decision making of doctors and healthcare members in hospitals.

1.2 Problem context

HD defines a condition that affects a heart. HD contains differences diseases such as Coronary Artery Disease (CAD), Congenital HD, Mitral Value Prolapse, Arrhythmia, Pulmonary Stenosis, Dilated Cardiomyopathy, Heart Failure, Hypertrophic Cardiomyopathy, and Myocardial Infarction. One of them, Cardiovascular Disease (CVD) is one of the main diseases of the heart that refers to the condition of obstructed blood vessels that can be happened a stroke and heart attack. Another form of HD can be rhythm, heart's muscle, etc. (Mayo Clinic, 2019)

CVDs are one of the major cause of people death globally. Many people have died from CVDs compare to other cause. In 2016, due to CVDs, an estimated 17.9 million human died. It's illustrating 31% of human deaths all over the world. Stroke and heart attack have occupied 85% of these deaths. (World Health Organization, 2019)
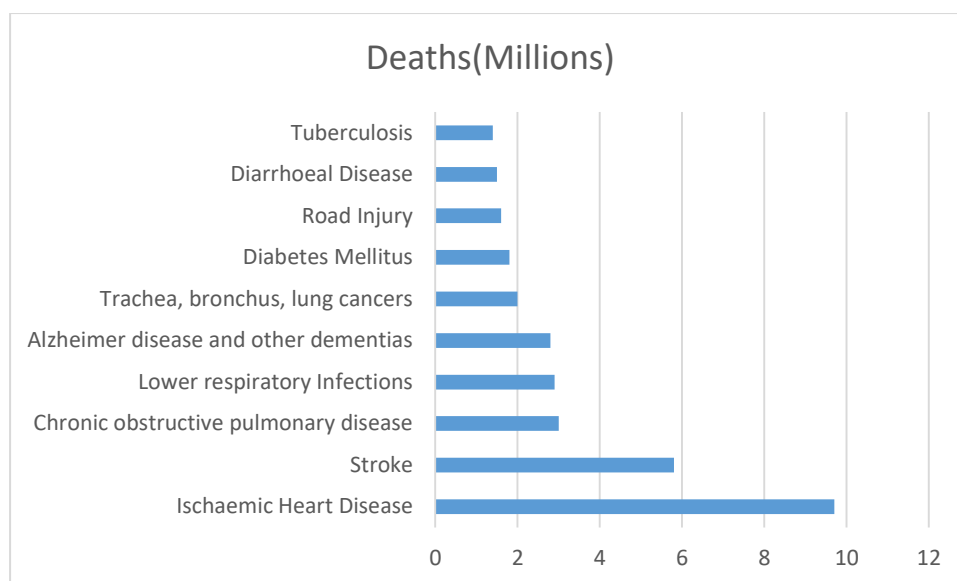
Figure 1: Top 10 Global causes of deaths (WHO, 2019)

In Figure-1.1, the WHO analysis the data about causes of deaths in between 2000-2016 and result clearly shows that the causes of heart disease's death is higher than other causes of death.

In 2017, the latest fact data of Word Health Organization (WHO) published that Nepal has reached 18.72% or 30,559 deaths from Coronary HD. The rate of age fixed death is 158.35 out of 100,000 population and world rank is #41. (World Life Expectancy, 2019)

According to The Heart Foundation; 13% of men and 10% of women are died due to HD in Australia. In 2017, Whilst HD had 18,590 deaths. So that HD was a one four death of cause factor in 2017. (The Heart Foundation., 2019)

So, Nepal government also needs to use this system to aware the patient before being critical situation. This system provide accurate result that help to less worry about the doctor's negligence.

1.3 Rationale

With the consideration of WHO statistical facts, the most powerful causes of death globally are a HD. It seemed to the negligence of patients as well as doctors to increase a HD patient. Some of the difficulties to execute the doctor's decision and lack of application to clearly diagnosis of HD become the cause of human death.

Regarding the above issues, we are proposing a web-based HDPS that is one of the best solutions to efficiently and accurately predict the HD patients. The proposed system eliminates the various testing of HD and supports the decision making of doctors. This system can accept a singleton query and display the clear output of the presence of HD level. This system is useful for any hospital and clinic to evaluate the patient getting a HD. It is reduced the number of tests and provide an efficient output of patient HD. It supports to make the decision of doctors that consulates with their patients easily.

The proposed system will support the healthcare systems as well as health-related application to expand their services with efficiently and accurately providing results. It mitigates the time to checkup of doctors.

1.4 Target users

The proposed system will cover a particular field of hospital and healthcare center and the target users of this system are also specific like heart-related doctors who have directly used this system in hospitals and healthcare centers.

1.5 Aim and objectives

Developing a Web-based HDPS using ML method for doctors as well as for the hospitals.

- To study about HDs prediction using various algorithms and identify the important attributes of HD
- To study and research on NB and DT algorithms for comparing accuracy.
- To evaluate and identify the best out of two algorithms
- To build a HD Prediction Model by utilizing UCI Datasets and deploy the Model into a web app.

1.6 Limitation of project

A web-based HDPS application is reserved for the specific task of HD due to researched knowledge of the domain and technical. Considering the time to accomplish this project, it will conduct the following tasks only:

- HDPS can only find the presence of HD of patients. In future, we will enhance the model for predicting specific type of HD.
- HD prediction model can be trained only 303 data of HD patient due to difficult of collecting Nepalese heart patient data but in future, we will collect large data and train model with their high accuracy.
- HDPS can be run with the internet and can be open in the only browser. In the future, we will develop this system as an offline based application.
- HDPS can only reserve heat disease prediction. In the future, we will integrate a healthcare system.

1.7 Nature of Challenges

HD is the domain of health and medical science field which distinct from the IT field, as well as ML, is an emerging field of AI that is a new technology to open new horizons. ML projects are always technology-driven that comes to the actual product is not easy. These are some high-level challenges will be faced during the developing the project.

The followings are addressing the challenges while developing this project as listed below:

- In the survey of this project, the interview of HD doctors will difficult to conduct due to their busy time schedule. Thereby it will make hard to identify the right attribute of the HD.
- The main problem will be built a classifier model that will predict if it is getting HD or not it means how to train the classifier model so that the system will get exact if the patient is possessing HD or not that is the main challenge of this project.
- ML is a new technology that provides a solution to the contemporary problem. It is very young technology environment that every big company such as Google, Facebook, and Microsoft are working ML project in their own environment like Tensorflow (released on 2017), PyTorch (released on 2017), and Microsoft Cognitive Toolkits respectively. Thereby, it is still research phased. It means it may not be production-ready or be borderline production ready. So, this purposed project will be faced with the problem during the deployment.

2. LITERATURE REVIEW

2.1 Heart Disease (HD)

HD is defined a range of conditions that affect your heart. It is describing any disorder of the heart. The umbrella of HD consists of different type of HD such as blood

vessel diseases (coronary artery disease, and arrhythmias) and heart defects when you're born with congenital heart defects, among others. (Mayo Clinic, 2019).

The term "Heart Disease" is always used interchangeably with the term "Cardiovascular Disease (CVD)". CVD generally refers to conditions that involve blocked or narrowed blood vessels that can lead to a heart attack, stroke or chest pain (angina). (Mayo Clinic, 2019).

2.2 Heart Disease Prediction

(Chala Beyene, 2018) Proposed a methodology to foretell the occurrence of HD to overcome the problem of diagnosis of HD. It improved the existence methodology by choosing Naïve Bayes, J48, and SVM for predicting the occurrence of HD for early automatic diagnosis in short time in order to support the qualities of services and reduce costs to save the life of individuals. This methodology uses various attributes of HD in order to identify whether a patent has HD or not. The comparison of analysis in the dataset is used WEKA software.

(P.Sai Chandrasekhar Reddy, 2017) Recommended ANN algorithms for HD prediction system in DM. The main aim of this predicting system is to reduce cost of a diagnosis like different type of test was done to make a decision for diagnosis of HD. so, they have proposed a new system to prophesy the condition of the patient based on their parameters such as age, blood pressure, heartbeat rate, cholesterol, etc. and evaluate if a patient has HD or not. The proposed system is provided its accuracy in java.

(Dwivedi, 2016) Focused to evaluate the performance of different ML algorithms for HD prediction. The comparison between different algorithms such as Naïve Bayes, KNN, Logistic Regression and Classification tree in order to identify the high performance for predicting the HD.

3. RESEARCH METHODS

Research method is one of the processes used to gather data and information for the purpose of making a decision in the research project. In another words, research method is a set of structuration that used to perform that research plan (guides, 2018). Research methods could have various types such as observation, questionnaire, interview, focus groups, experiments, secondary data analysis, etc. each research tools have their own advantage and disadvantage. Selecting appropriate research tools for collecting the target data, various parameters need to be checked that research objectives achieved by the chosen research tool or not. In our research project, we have chosen questionnaire research tool which is appropriate for collecting information that helps to achieve our research objective. The main goal of selecting this research method is to seek feedback and professional advice from the potential domain, technical expertise, and the general public.

3.1 Justification of chosen Research Method

In order to achieve our research objectives, a questionnaire is an appropriate tool for research method. The following are the justification of choosing questionnaire for our research project as below:

- Questionnaire collects a large volume of data from the mass of people that can be scientifically analyzed data and achieved the right option to make the decision for the research project.
- Questionnaire can be conducted by distributing a group of people and from the online that helps to easily collect data and information.
- Questionnaire allows no pressure to respond immediately to participators. It is a platform that participators can feel free to submit their response.

- Questionnaire is a cost-effective way in order to gather information from participators rather than other research tools.
- Questionnaire is covered a wide area that helps to finds information from all type of participators such as general public, heart surgeon, heart patient, and technical professional that are our main participators of this research project.

4. SYSTEM ARCHITECTURE

4.1 Introduction

The system architecture (Kautish et al, 2016, 2018, 2019) is like a blueprint of any object. It is a conceptual model to integrate between business logic and physical system in an organized way. It demonstrates the structure, view, behavior, features, and functionalities of the system. It is the way of portraying the desired system in visualizing a way to well understand for people. The System architecture is the foundational orchestrate of a system that incorporated in its elements, their relationships of elements, and the science of its design (MITRE, 2019).

HDPS is a web-based application that runs on the browser. This system is embodied in a web application. The web application architecture of the HDPS is to define the communication between applications, and database on the web. It also helps to know about other third-party application required like python's packages. It represents the represent the relationship between them and visualizes how they work together simultaneously.

4.2 Abstract Architecture

Abstract Architecture of HDPS has demonstrated the comprehensive structure of the system that easily understands all components and their relationship. The way to represent the system as a structural way can beneficial to understand the system for technical and non-technical users. The abstract architecture consists of a system, database, and interface design to visualize the better way of viewing. It describes the overall design of the system to easily understand the system elements, feature and functionalities, and behavior of the system.

4.3.1 System Design

System design is the systematic process of defining the components of the purposed system that consists of model, architecture, and interface of different elements. It describes the operation of a system that demonstrates data flow structure and a link between the database tables (Odhiambo, 2018). In this system design, we are going to design a procedure programming system design related to our project. They can be a Context diagram, DFD Level-1, DFD Level-2. It is an approach to design the system to organize a way to easily understand the whole system.

4.3.2 SSADM

SSADM is abbreviated as Structured Systems Analysis and Design Method that is traditional software designing method that aids to design information system. This methodology was developed by CCT (Central Computer and Telecommunications Agency) in the early 1980s in the UK (Edrawsoft, 2019). It mainly focuses on a procedural programming system to design the system. It follows a sequential or linear model to complete the system design.
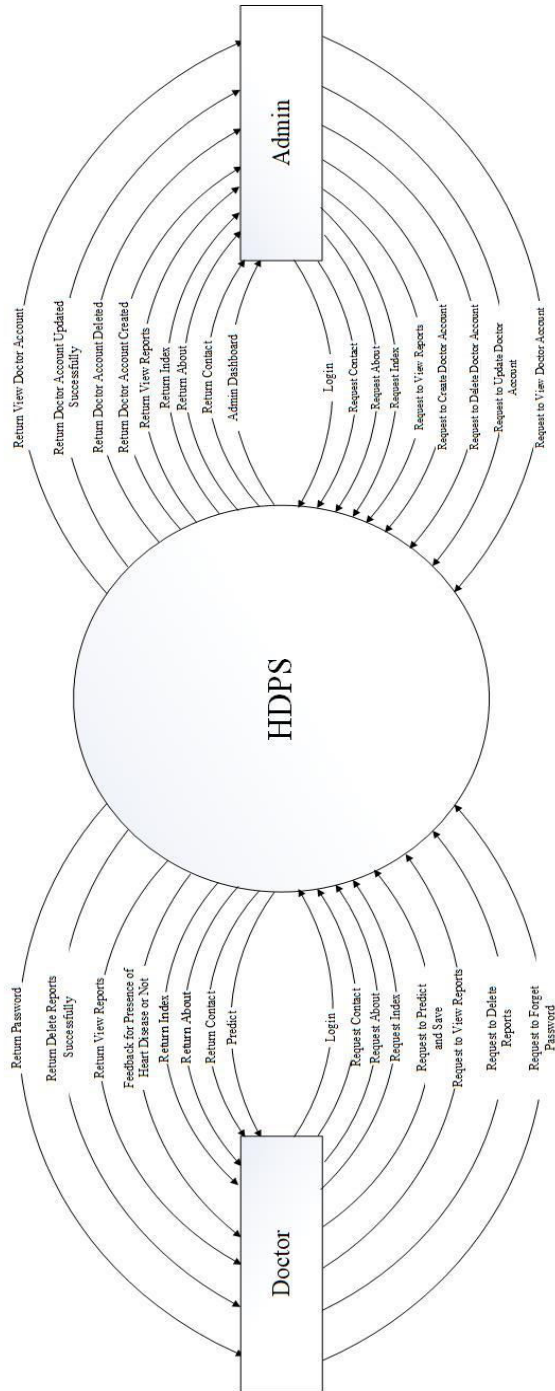
4.2.2.1 Context Diagram or DFD Level-0

Figure 2: Context Diagram of HPDS

4.2.2.2 DFD Level-1

Figure 3: DFD Leve-1 of HDPS
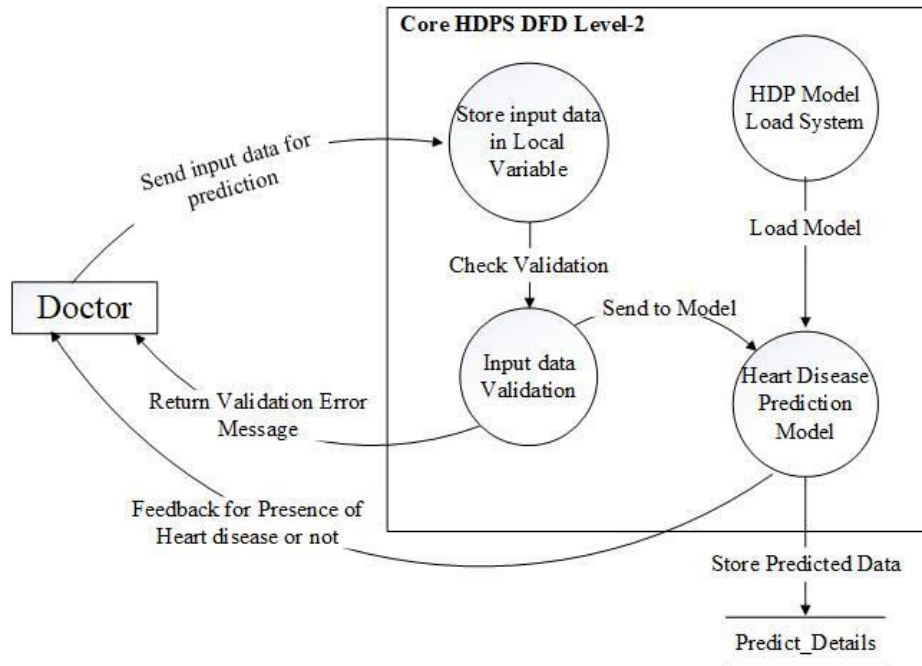
4.2.2.3   DFD Level-2

Figure 4: DFD Level-2 of HDPS
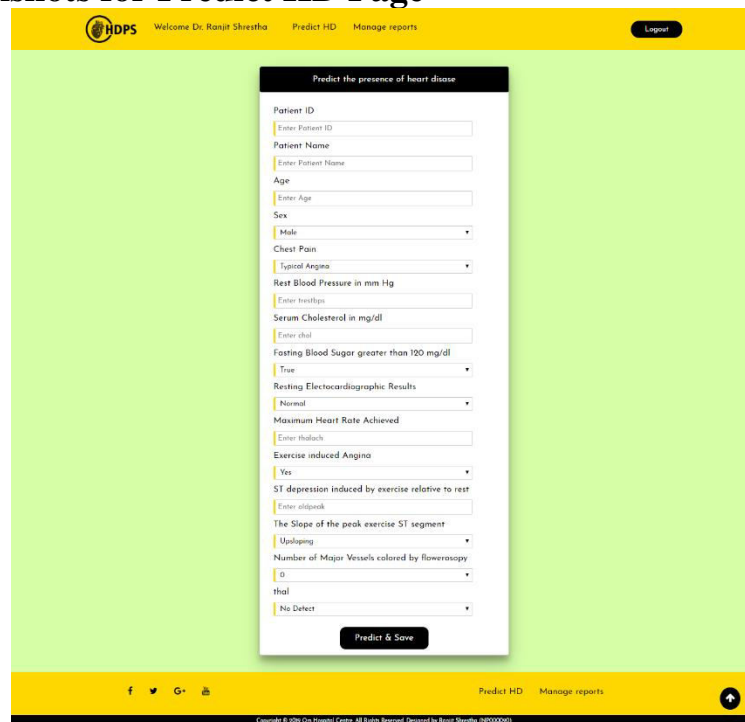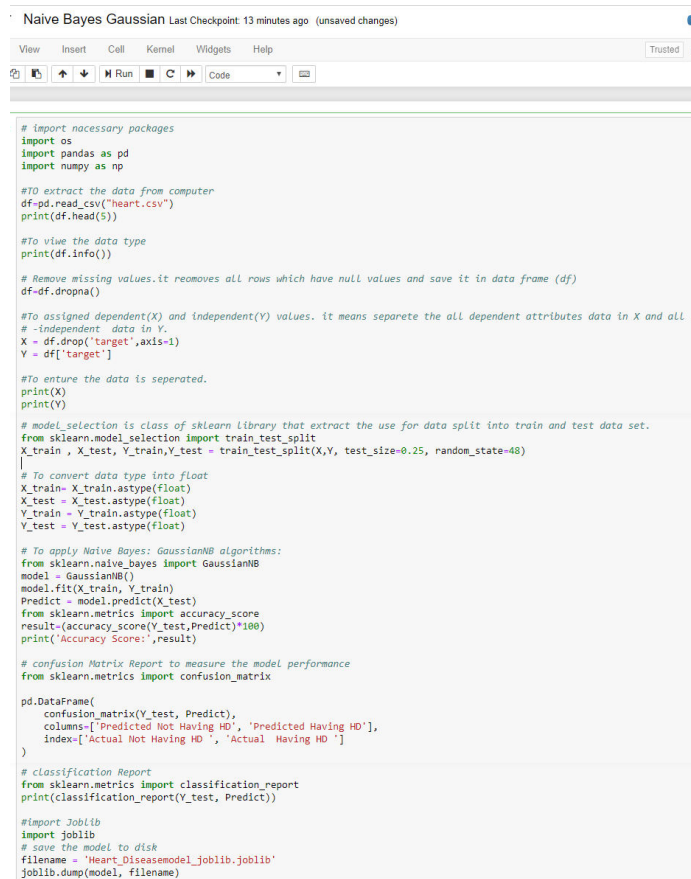
4.3 Screenshots for Predict HD Page

Figure 5: Predict HD page

4.3.1   Description

The above page shows the Predict HD page that appears only access by doctors. This page demonstrates the form of prediction HD that contains details about the patient's data. It has input validation to restrict the garbage input that displays error message within this form. Must of input data are dropdown list due to restricting the garbage input. When the doctor successfully predicts the HD, it will automatically open the success page that

displays the result of the patient. Furthermore, this function also stores the patient's report into the database table.

4.4 Sample Codes of Heart Disease Prediction Model Development

5. CONCLUSION AND REFLECTIONS

5.1 Critical Evaluation

Heart Disease is a killer disease of death in the world. According to WHO and other statistical facts, HD is the most dangerous disease that is the cause of death of a human. In 2017, the latest fact data of the World Health Organization (WHO) published that Nepal has reached 18.72% or 30,559 deaths from Coronary HD. The rate of age fixed death is 158.35 out of 100,000 population and world rank is #41. (World Life Expectancy, 2019).

After identifying the problem statement, we strived to find a way of solving the problem through ML. After that, we conducted requirement analysis and planning of the system. We made the line of research boundary to complete this project by planning the system requirement and planning. System requirement and planning consists of aim, objective, deliverables, and target audience of the system. We conducted the literature review by studying various journal papers and articles to understand the way of solving the problem as well as collecting important attributes of HD that lead role of the HD in patient. We found a various way to resolve the problem but dilemma to select the best algorithm. Eventually, we chose two algorithms (DT and NB) among them which have high accuracy in HD datasets. The dataset of HD is retrieved from UCI repertory which has important 13 attributes among 76 attributes that play the role of HD. it has 303 records that a small dataset. It may difficult to achieve a high accuracy to predict an HD.

In technical research, we found a Python programming language that is suitable for the HDPS by comparing two other languages such as R and Java. Python is only one

language to handle the both function such as building and deploying a Prediction Model and developing a web application. For building an HDP Model, we chose Jupyter notebook for interactive development environment. We selected some essential Python libraries that supports to build an HDP Model such as NumPy, Pandas, Matplotlib, and Sk-learn. For developing web application, we select Python Flask Server-side scripting language to deploy the HDP Model and develop a web application. Furthermore, we selected Windows OS, SQL Database, Sublime text editor, Python WSGI web server, and Google Chrome for web browser by critical evaluating.

For developing the HDPS, we chose the waterfall methodology by comparing two other methodologies such as RAD and Prototype Model. We planned the all development tasks in this methodology. After that, we conducted the feasibility study to ensure the system can meet the business and user's requirement by conducting the questionnaire survey. We designed five section of question into four participators such as General Public, Heart Patient, Heart Surgeon, and IT Professional. In section-1, we designed Demography question for all participators, and General question, HD related question for patient, HD related question for Doctor, and System related question for IT professional in section-2, section-3, section-4, and section-5 respectively. This design ensures to reach right question for right participators. we received a total of 101 responded out of that 80,1,5,15 responded people are General public, Heart Patient, Heart Surgeon, and IT Professional respectively. The result of survey led the system can meet the requirement of users. We proceed this project toward the next phase like system design.

In system design phase, we discussed on the system and abstract architecture that includes system design by choosing appropriate methodology. We selected SSADM methodology to design system because we planned to use procedure programming concept. In this SSADM, we designed Context diagram (DFD Level-0), DFD Level-1 and DFD Level -2 to portray the architecture of the system. We strived to demonstrate the flow of data, process and functionalities of modules in these diagrams. Moreover, we conducted a database design that includes ERD, Database table structure like Data dictionary. For interface design, we conducted storyboard designed by using a Balsamiq tool. Each and every interface design of the system portray clear picture of the system.

In the project plan, we planned a release version of the system based on the features of HDPS and test plan of the HDPS. We discussed details about features in each version. In the test plan, we planned to conduct Unit, integration, system, usability, and user acceptance testing. For the implementation of the HDPS, we faced the problem of accuracy of algorithms due to the result has three-level, that's why data is distributed and give low accuracy. So, we chose Kaggle dataset of HD which have two results. Again, we faced the problem of low accuracy of DT and NB. To solve this problem, we explored other ML algorithms like SVM, NN, etc. Finally, we got about 88 % accuracy of Gaussian NB algorithm and built an HDP Model for the system. The lesion of building a choosing an algorithm is *"Any algorithm is not the best algorithm for ML. It depends on the dataset. We conducted hit and trial method to find out best fit algorithm for the chosen dataset"*. After that, we conducted coding for deploying an HDP Model and developing a web application based on the storyboard.

At last but not least, we conducted the system validation that includes testing of system such as unit, integration, system, usability and user acceptance testing based on the testing plan of the system. In unit testing, we conducted four testing such as Create Doctor Account, Manage Doctor Account, Login, Forget Password, and Predict Heart Disease. We conducted three UAT. The overall testing was a success but there was unit testing had a problem. These were hard to solve the displaying error message of the system. In the future, we can solve that problem.

5.2 Conclusion

Eventually, we are able to achieve all the proposed objectives of this project. We developed a web application of HDPS that has all features and functionalities what we planned. The research of this project is successfully met in the system. This system resolves the real-environment problem. It is successfully predicted the presence of HD in patient. It is also store and manage the prediction report of heart patient by doctor account. Admin user can handle create doctor account, mange doctor account and view the report of the patient. The overall system can solve the problem statement of the project and make a novel tool to predict an HD at hospital.

We were able to research or investigate our topic by conducting literature review as domain and technical research. At domain research, we achieved 13 important attributes of HD that every researcher had done their research to predict the HD. we also achieved an appropriate two algorithm to provide high accuracy in heart disease dataset of the heart patient such as DT and NB and finally NB provides 88% accuracy among other algorithms. But the research of dataset of heart may failed due to distributed result of HD dataset. That's why we change the dataset from Kaggle. At technical and other research, we were successfully meet finding to final project. Programming language, database, methodology is suitable to conduct best performance to accomplish this project.

In our research of the topic, we were difficult to find details about attributes of the HD. The improper attributes name, their impact, and their sub-attributes were less information in journal paper and dataset description. Due to lack of spare time of doctor, we were unable to collect more data about HD. lack of time constraint, we were unable to explore other features to include the system and explore different heart disease type prediction.

5.3 Future Work

Future enhance of the HDPS is to predict a specific HD type such Heart attracts, CVD, CAD, etc. the potential of the HDPS in a different area are hospital, Clinic, smartphone, smart wear, hospital/police emergency system and integrate with fitness mobile application. We will integrate this model in hospital and clinic system to predict heart disease. We will implement this HDP Model into smart wears to detect essential attributes of HD and suggest to the precaution of HD. we will also apply this model into a mobile app to easily test ourselves HD. we will integrate smart wear to the hospital and police emergency system to save the life of the patient at the emergency condition.

REFERENCES

Segue Technologies, 2015. *The Benefits of Adhering to Software Development Methodology Concepts.* [Online]
Available at: https://www.seguetech.com/benefits-adhering-software-development-methodology-concepts/
[Accessed 28 01 2019].

Abhishek, t., 2013. Heart Disease Prediction System Using Data Mining Techniques. *Oriental Scientific Publishing Co., India,* 6(4), pp. 457-466.

Afroz, R. R. a. F., 2013. Comparison of Various Classification Techniques Using Different Data Mining Tools for Diabetes Diagnosis. *Journal of Software Engineering and Applications,* 6(3), pp. 85-97.

Aparna K., D. R. N. C. S. P. I. S. S. D. K. V., 2014. Disease Prediction in Data Mining Techniques. *InternatIonal Journal of Computer SCIenCe and teChnology,* 5(2), pp. 246-249.

Arsenault, R., 2019. *Sublime Text.* [Online]
Available at: https://www.utest.com/tools/sublime-text
[Accessed 30 01 2019].

Asthana, S., 2018. *Essential libraries for Machine Learning in Python.* [Online]
Available at: https://medium.freecodecamp.org/essential-libraries-for-machine-learning-in-python-82a9ada57aeb
[Accessed 08 02 2019].

Bahrami, B. a. S. M., 2015. Prediction and Diagnosis of Heart Disease by Data Mining Techniques. *Journal of Multidisciplinary Engineering Science and Technology (JMEST),* 2(2), pp. 164-168.

Chala Beyene, P. K., 2018. "Survey on Prediction and Analysis the Occurrence of Heart Disease Using Data Mining Techniques. *International Journal of Pure and Applied Mathematics,* 118(8), pp. 165-174.

Dangare Chaitrali S., P. A. S. S., 2012. Improved Study of Heart Disease Prediction System using Data Mining Classification Technique. *International Journal of Computer Applications (0975 – 888) ,* 47(10), pp. 44-48.

db-engines, 2019. *System Properties Comparison Microsoft SQL Server vs. MySQL vs. SQLite.* [Online]
Available at: https://db-engines.com/en/system/Microsoft+SQL+Server%3BMySQL%3BSQLite
[Accessed 03 02 2019].

Dwivedi, A. k., 2016. Evaluate the performance of different machine learning techniques for prediction of heart disease using ten-fold cross-validation. *Springer.*

Edrawsoft, 2019. *SSADM Diagram Software - Structured Systems Analysis and Design Methodology.* [Online]
Available at: https://www.edrawsoft.com/SSADM.php
[Accessed 11 06 2019].

experienceux, 2018. *What is usability testing?.* [Online]
Available at: https://www.experienceux.co.uk/faqs/what-is-usability-testing/
[Accessed 17 12 2018].

fullstackpython, 2019. *Full Stack Python.* [Online]
Available at: https://www.fullstackpython.com/web-servers.html
[Accessed 05 02 2019].

Gupta Y, Ahmed R K A, Kautish S K; Application of data mining and knowledge management in special reference to medical informatics: a review; International Journal of Medical Laboratory Research 2017, 2(2): 60-76

guides, 2018. *Research Methods Guide: Research Design & Method.* [Online]
Available at: https://guides.lib.vt.edu/researchmethods/design-method
[Accessed 09 02 2019].

heart, 2015. *Transesophageal Echocardiography (TEE).* [Online]
Available at: https://www.heart.org/en/health-topics/heart-attack/diagnosing-a-heart-attack/transesophageal-echocardiography-tee
[Accessed 25 01 2019].

HN, N. P., 2017. *Best Python libraries for Machine Learning and Data Science : (Part - 1).* [Online]
Available at: https://towardsdatascience.com/best-python-libraries-for-machine-learning-and-data-science-part-1-f18242424c38
[Accessed 07 02 2019].

Joshi, S. a. N. M., 2015. Prediction of Heart Disease Using Classification Based Data Mining Techniques. *Computational Intelligence in Data Mining-Volume 2,* pp. 503-511.

Jyoti, N. B. K., 2012. An Analysis of Heart Disease Prediction using Different Data Mining Techniques. *International Journal of Engineering Research & Technology(IJERT),* 1(8), pp. 1-4.

Kaur Beant, S. W., 2014. Review on Heart Disease Prediction System using Data Mining Techniques. *International Journal on Recent and Innovation Trends in Computing and Communication ,* 2(10), p. 3003 – 3008 .

Kaur, H. and Kautish, S., (2016). "An Implementation of Wireless Sensor Network Using Voronoi _
    PSO (Particle Swarm Optimization)", International Journal for Research in Applied Science & Engineering Technology (IJRASET), Volume 4, Issue XI, November 2016, pp.361-368

Kautish,S. (2008), "Online Banking : A Paradigm Shift", E-Business, Vol. 8, No.10, pp. 5459.

Kautish S., Thapliyal M P, "Concept of Decision Support Systems in relation withKnowledge Management – Fundamentals, theories, frameworks and practices", International Journal of Application or Innovation in Engineering & Management (IJAIEM) Volume 1, Issue 2, October 2012 ISSN 2319 – 4847

Kautish S. 2013. Knowledge sharing: A contemporary review of literature in context to information systems designing. Academia 3(1). The South Asian Academic Research Journal: 101-113.

Kautish, S., & Thapliyal, M. P. (2013). Design of new architecture for model management systems using knowledge sharing concept. International Journal of Computer Applications, 62(11), 27–30.

Kumar, A., Rajpurohit, V.S. and Kautish, S., 2020. A Study on Technology-LED Solutions for Fruit Grading to Address Post- Harvest Handling Issues of Horticultural Crops. In Modern Techniques for Agricultural Disease Management and Crop Yield Prediction (pp. 203-221). IGI Global.

Kaur, R., & Kautish, S. (2019). Multimodal Sentiment Analysis: A Survey and Comparison. International Journal of Service Science, Management, Engineering, and Technology (IJSSMET), 10(2), 38-58.

M. Marimuthu, M. A. K. H. K. M. V., 2018. A Review on Heart Disease Prediction using Machine Learning and Data Analytics Approach. *International Journal of Computer Applicationss (0975 – 8887) ,* 181 (18), pp. 20-25.

MARSH, J., n.d. *The Importance of Unit Testing in Software Development.* [Online]
Available at: https://www.upwork.com/hiring/development/unit-testing-in-software-development/
[Accessed 13 12 2018].

Mayo Clinic, 2019. *Heart Disease.* [Online]
Available at: https://www.mayoclinic.org/diseases-conditions/heart-disease/symptoms-

causes/syc-20353118
[Accessed 15 01 2019].

Mayo Clinic, 2019. *Heart Disease.* [Online]
Available at: https://www.mayoclinic.org/diseases-conditions/heart-disease/symptoms-causes/syc-20353118
[Accessed 07 03 2019].

Methaila Aditya, K. P. A. H. K. P., 2014. EARLY HEART DISEASE PREDICTION USING DATA MINING TECHNIQUES. *Computer Science & Information Technology (CS & IT),* p. 53–59.

Methaila, A. K. P. A. H. a. K. P., 2014. Early heart disease prediction using data mining techniques.. *Computer Science & Information Technology Journal,* pp. 53-59.

Miller, N., 2018. *Waterfall vs. RAD – Which One Is Best For You?.* [Online]
Available at: https://kissflow.com/rad/rad-waterfall-agile-which-one-is-best-for-you/
[Accessed 06 02 2019].

Miss. Chaitrali S. Dangare, D. M. S. S. A., 2012. A data mining approach for prediction of heart disease using neural networks. *international journal of computer engineering and technology,* 3(3), pp. 30-40.

MITRE, 2019. *System Architecture.* [Online]
Available at: https://www.mitre.org/publications/systems-engineering-guide/se-lifecycle-building-blocks/system-architecture
[Accessed 3 06 2019].

Mohandoss S., V. S. S. R. a. R. S., 2017. a Hybrid Heart Disease Prediction System using Evolutionary Learning Algorithms. *Asian Journal of Information Technology,* 16(7), pp. 639-644.

Monali Dey, S. S. R., 2014. Study and Analysis of Data mining Algorithms for Healthcare Decision Support System,. *International Journal of Computer Science and Information Technologies,* 5(1), pp. 470-477.

Ms. Shinde Swati B., P. P. A., 2015. Decision Support System on Prediction of Heart Disease Using Data Mining Techniques. *International Journal of Engineering Research and General Science,* 3(2), pp. 1453-1458.

Niraula, P. and Kautish, S., Study of The Digital Transformation Adoption in The Insurance Sector of Nepal. LBEF Research Journal of Science, Technology and Management, 1(1), pp.43-60.

NGBAGARO, L., 2016. *Systems development methodology.* [Online]
Available at: https://www.supinfo.com/articles/single/2880-systems-development-methodology
[Accessed 09 12 2018].

nginx, 2017. [Online]
Available at: https://www.nginx.com/resources/wiki/
[Accessed 05 02 2019].

Odhiambo, D., 2018. [Online]
Available at: https://medium.com/the-andela-way/system-design-in-software-

development-f360ce6fcbb9
[Accessed 05 06 2019].

P.Sai Chandrasekhar Reddy, P. P. S., 2017. Heart Disease Prediction using ANN Algorithm in Data Mining. *International Journal of Computer Science and Mobile Computing,* pp. 168-172.

Palaniappan Sellappan, A. R., 2008 . Intelligent Heart Disease Prediction System Using Data Mining Techniques. *International Journal of Computer Science and Network Security,* 8(8), pp. 343-350.

Patel Jaymin, P. U. T. D. P. S., 2016. Heart Disease Prediction using Machine Learning and Data Mining Technique. *IJCSC,* Volume 7, pp. 129-137.

Patel Jaymin, P. U. T. D. P. S., 2016. Heart Disease Prediction Using Machine Learning and Data Mining Tehnique. *IJCSC 0973-7391,* 7(1), pp. 129-137.

Patel Jaymin, U. P. D. P. S., 2015. heart disease prediction using machine learning and data mining technique. *IJCSC (0973-7391),* 7(1), pp. 129-137.

Patel, J. T. D. a. P. S., 2015. Heart disease prediction using machine learning and data mining technique.. *IJCSC,* 7(1), pp. 129-137.

Patil, K., 2017. *WHY IS INTEGRATION TESTING IMPORTANT IN SOFTWARE TESTING LIFE CYCLE?.* [Online]
Available at: https://bitwaretechnologies.com/integration-testing-important-software-testing-life-cycle/
[Accessed 15 12 2018].

Rani, S. and Kautish, S., 2018, June. Association Clustering and Time Series Based Data Mining in Continuous Data for Diabetes Prediction. In 2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS) (pp. 1209-1214). IEEE.

Rani, S., Kautish, S. (2018). Application of data mining techniques for prediction of diabetes-A review. International Journal of Scientific Research in Computer Science, Engineering and Information Technology, 3(3), 1996-2004.

Rouse, M., 2018. *Microsoft Windows.* [Online]
Available at: https://searchwindowsserver.techtarget.com/definition/Windows
[Accessed 04 02 2019].

Singh, C. and Kautish, S.K., Page Ranking Algorithms for Web Mining: A Review. International Journal of Computer Applications, 975, p.8887.

Saxena, P.D. and Kautish, S.K. (2015). Risk Assessment of Flood for Various Spatial and Temporal Mining Scales Using Flood Vulnerability Index [FVI] – A Case Study of Madhubani Kosi River Flood Affected District in North Bihar, India. International Journals Digital Communication and Analog Signals, [online] 1(2), pp.20–29.

Santhanam, T. a. E. E., 2015. Heart Disease Prediction Using Hybrid. *Indian Journal of Science and Technology,* 8(9), p. 797–803.

Sharma, L., 2016. *WaterFall Model.* [Online]
Available at: http://toolsqa.com/software-testing/waterfall-model/
[Accessed 11 12 2018].

Shinde Swati B., P. A., 2015. Decision Support System on Prediction of Heart Disease Using Data Mining Techniques. *International Journal of Engineering Research and General Science,* 3(2), pp. 1453-1458.

Techopedia, 2019. *User Acceptance Testing (UAT).* [Online]
Available at: https://www.techopedia.com/definition/3887/user-acceptance-testing-uat
[Accessed 11 8 2019].

The Heart Foundation., 2019. *Heart disease fact sheet | The Heart Foundation..* [Online]
Available at: https://www.heartfoundation.org.au/about-us/what-we-do/heart-disease-in-australia/heart-disease-fact-sheet.
[Accessed 02 01 2019].

Tutorialspoint, 2019. *What is Unit Testing?.* [Online]
Available at:
https://www.tutorialspoint.com/software_testing_dictionary/unit_testing.htm#
[Accessed 10 8 2019].

Tyler, K., 2019. *R vs. Java vs. Python: Which Is Right for Your Project?.* [Online]
Available at: https://www.upwork.com/hiring/data/r-vs-java-vs-python-which-is-best/
[Accessed 08 03 2019].

UCI repository, 2019. *Heart Disease Data Set.* [Online]
Available at: https://archive.ics.uci.edu/ml/datasets/Heart+Disease
[Accessed 06 02 2019].

Vasconcellos, P. H., 2018. *Top 5 Python IDEs For Data Science.* [Online]
Available at: https://www.datacamp.com/community/tutorials/data-science-python-ide
[Accessed 08 03 2019].

Visual-Paradigm, 2019. *What is Entity Relationship Diagram (ERD)?.* [Online]
Available at: https://www.visual-paradigm.com/guide/data-modeling/what-is-entity-relationship-diagram/
[Accessed 10 06 2019].

WHO, 2019. *The top 10 causes of death.* [Online]
Available at: https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death
[Accessed 18 02 2019].

Woody, A., 2015. *Why is a software development methodology so important?.* [Online]
Available at: https://www.quora.com/Why-is-a-software-development-methodology-so-important
[Accessed 29 01 2019].

World Health Organization(WHO), 2018. *The top 10 causes of death.* [Online]
Available at: https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death
[Accessed 03 01 2019].

World Health Organization, 2019. *cardiovascular-diseases-(cvds).* [Online]
Available at: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)
[Accessed 16 01 2019].

World Life Expectancy, 2019. *Coronary Heart Disease in Nepal.* [Online]
Available at: https://www.worldlifeexpectancy.com/nepal-coronary-heart-disease
[Accessed 17 01 2019].

WSGI, 2015. *Introduction.* [Online]
Available at: http://wsgi.tutorial.codepoint.net/
[Accessed 06 02 2019].

Ziasabounchi, N. a. A. I., 2014. ANFIS based classification model for heart disease
prediction.. *International Journal of Electrical & Computer Sciences IJECS-IJENS,*
14(02), pp. 7-12.