**Faculty of Computing and Technology**

**University of Kelaniya**

**Bachelor of Science Honours in Computer Science Degree**

**CSCI 43018 – Project**

**Academic Year 2021/2022**

**Project Diary**

**Project Title: Heart Disease prediction using Machine learning and Deep learning techniques.**

**Name of the Supervisor: Ms. P.H.A.H.K. Yasodara**

**Student Name: P.M.B.D. SAMARAKOON**

**Student Index Number:** CS/2018/037

Meeting No: ……01….. Meeting Date: 2023-12-03

|  |  |  |
| --- | --- | --- |
| Tasks assigned in the previous meeting | If **completed** briefly describe the way you reached/solved the task/issue | If **not completed** clearly state the reason |
| 1.  2.  3.  4.  5. |  |  |
| Points/tasks/issues discussed at the current meeting | * Discussed the selected topic, “ Heart Disease prediction using Machine learning and Deep learning techniques” with the supervisor. Also discussed the pros and cons of the approach. * Overview of the project objectives and scope. * Initial literature review and dataset search. * Setting up the project plan and timeline. * Identifying initial tasks and responsibilities. | |
| Targets/tasks assign to complete before the next meeting | * Should found a suitable data set * Search and identify suitable datasets. * Prepare the development environment. * Outline the initial project structure. * Schedule the next meeting. | |

\*\*Please make sure to contact/meet your supervisor/s **at least** once a month

Signature of the supervisor: ………………… Date: ……………

**Self-diary**

Project Title: Heart Disease prediction using Machine learning and Deep learning techniques

Month: 2023/12 Index Number: CS/2018/037

|  |  |  |
| --- | --- | --- |
| **Week** | **Date** | **Description of work carried out, Problems found and the way solved, etc.** |
| **Week 1** | 12/03 |  |
| 12/04 | Started searching for a suitable dataset on Google. |
| 12/05 | Found an initial dataset but deemed it unsuitable for the project. |
| 12/06 | Identified a more suitable dataset for the project. |
| 12/07 | Downloaded and prepared the selected dataset for analysis |
| 12/08 |  |
| 12/09 |  |
|  | | |
| **Week 2** | 12/10 |  |
| 12/11 | Summarized key findings from the reviewed papers. |
| 12/12 | Go through the data set and studied it |
| 12/13 | Started the project |
| 12/14 | Studied the methods |
| 12/15 |  |
| 12/16 |  |
|  | | |
| **Week 3** | 12/17 |  |
| 12/18 |  |
| 12/19 | Began initial data preprocessing, handling missing values. |
| 12/20 | Normalized features in the dataset. |
| 12/21 | Set up the development environment with necessary libraries. |
| 12/22 | Studied a ML models |
| 12/23 | Studied a ML models |
|  | | |
| **Week 4** | 12/24 |  |
| 12/25 |  |
| 12/26 | Continued data preprocessing, focusing on feature engineering. |
| 12/27 | Studied the structure and contents of the dataset. |
| 12/28 | Insert dataset and do some analyze |
| 12/29 | Started the project implementation. |
| 12/30 |  |
|  | | |
| **Week 5** | 12/31 |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

Meeting No: ……02…… Meeting Date: 2024-01-11

|  |  |  |
| --- | --- | --- |
| Tasks assigned in the previous meeting | If **completed** briefly describe the way you reached/solved the task/issue | If **not completed** clearly state the reason |
| 1. Find a dataset  2. Study the model  3. Study ML models  4. Conduct a detailed literature review on heart disease prediction using ML and DL.  5. Search and identify suitable datasets.  6. Prepare the development environment.  7. Schedule the next meeting. | * Found a suitable data set from the Kaggle. Go through some ML models * Completed by reviewing and summarizing key findings from research papers. * Found and prepared a suitable dataset for analysis. * Set up with necessary libraries and tools. |  |
| Points/tasks/issues discussed at the current meeting | * Review of literature and datasets. * Initial data preprocessing and feature engineering. * Setting up machine learning models. * Discussion of initial results and challenges faced. | |
| Targets/tasks assign to complete before the next meeting | * Complete data preprocessing and feature engineering. * Implement Logistic Regression and Decision Tree classifiers. * Evaluate the initial models' performance. * Research additional ML models like Random Forest and SVM. | |

\*\*Please make sure to contact/meet your supervisor/s **at least** once a month

Signature of the supervisor: ………………… Date: ………………

**Self-diary**

Project Title: Heart Disease prediction using Machine learning and Deep learning techniques

Month: 2024/01 Index Number: CS/2018/037

|  |  |  |
| --- | --- | --- |
| **Week** | **Date** | **Description of work carried out, Problems found and the way solved, etc.** |
| **Week 1** |  |  |
|  |  |
| 01/02 | Completed initial Logistic Regression implementation. |
| 01/03 | Researched various machine learning methods relevant to the project. |
| 01/04 | Continued research on machine learning models. |
|  |  |
|  |  |
|  | | |
| **Week 2** |  |  |
|  |  |
|  |  |
| 01/10 | Began implementing Logistic Regression using scikit-learn. |
| 01/11 | Studied Random forest |
| 01/12 | Studied SVM |
|  |  |
|  | | |
| **Week 3** |  |  |
|  |  |
|  |  |
| 01/17 | Implemented Decision Tree classifier using scikit-learn. |
| 01/18 | Studied Decision Tree algorithms for implementation. |
| 01/19 | Faced issues with missing values; resolved by imputing median values. |
| 01/20 | Started clean the data set |
|  | | |
| **Week 4** |  |  |
|  |  |
| 01/23 | Started clean the data set |
|  |  |
|  |  |
| 01/26 | Watched related videos and study more |
| 01/27 | Evaluated the performance of Logistic Regression and Decision Tree, achieving 78.5% and 81.0% accuracy, respectively. |
|  | | |
| **Week 5** | 01/28 |  |
| 01/29 |  |
| 01/30 |  |
| 01/31 | Reviewed the performance metrics of initial models. |
|  |  |
|  |  |
|  |  |

Meeting No: ……03………. Meeting Date: 2024-02-15

|  |  |  |
| --- | --- | --- |
| Tasks assigned in the previous meeting | If **completed** briefly describe the way you reached/solved the task/issue | If **not completed** clearly state the reason |
| 1.. fix the issues with the dataset and model issues  2. Complete data preprocessing and feature engineering.  3. implement Logistic Regression and Decision Tree classifiers. | * Completed with normalization and feature engineering techniques. * Successfully implemented using scikit-learn. * Achieved 78.5% and 81.0% accuracy for Logistic Regression and Decision Tree respectively. |  |
| Points/tasks/issues discussed at the current meeting | * Review of completed tasks and model performance. * Discussion on challenges faced during model implementation. | |
| Targets/tasks assign to complete before the next meeting | * Implement and evaluate Random Forest and SVM models. * Finalize CNN architecture and start implementation. | |

\*\*Please make sure to contact/meet your supervisor/s **at least** once a month

Signature of the supervisor: ………………… Date: ………………

**Self-diary**

Project Title: Heart Disease prediction using Machine learning and Deep learning techniques

Month: 2024/02 Index Number: CS/2018/037

|  |  |  |
| --- | --- | --- |
| **Week** | **Date** | **Description of work carried out, Problems found and the way solved, etc.** |
| **Week 1** |  |  |
|  |  |
|  |  |
|  |  |
|  |  |
| 02/06 | Researched Random Forest and SVM for further implementation. |
| 02/07 | Began implementing Random Forest classifier. |
|  | | |
| **Week 2** |  |  |
|  |  |
|  |  |
|  |  |
| 02/12 | Faced challenges with hyperparameter tuning; started using GridSearchCV. |
| 02/13 | Completed implementation of Random Forest classifier |
| 02/14 | Implemented SVM classifier. |
|  | | |
| **Week 3** |  |  |
|  |  |
|  |  |
|  |  |
| 02/19 | Evaluated SVM, achieving 82.5% accuracy. |
| 02/20 | compared the performance of all implemented machine learning models. |
|  |  |
|  | | |
| **Week 4** |  |  |
|  |  |
|  |  |
|  |  |
| 02/25 | Started setting up the environment for CNN implementation. |
| 02/26 | Faced data reshaping issues for CNN; resolved by adjusting input dimensions. |
|  |  |
|  | | |
| **Week 5** |  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

Meeting No: ……04……. Meeting Date: 2024-03-28

|  |  |  |
| --- | --- | --- |
| Tasks assigned in the previous meeting | If **completed** briefly describe the way you reached/solved the task/issue | If **not completed** clearly state the reason |
| 1. Evaluate the initial models' performance.  2. Research additional ML models like Random Forest and SVM.  3. Plan for CNN implementation. | * Completed with normalization and feature engineering techniques. * Successfully implemented using scikit-learn. * Achieved 78.5% and 81.0% accuracy for Logistic Regression and Decision Tree respectively. |  |
| Points/tasks/issues discussed at the current meeting | * Review of completed tasks and model performance. * Discussion on challenges faced during model implementation. | |
| Targets/tasks assign to complete before the next meeting | * Implement and evaluate Random Forest and SVM models. * Finalize CNN architecture and start implementation. * Plan for CNN implementation. | |

\*\*Please make sure to contact/meet your supervisor/s **at least** once a month

Signature of the supervisor: ………………… Date: ………………

**Self-diary**

Project Title: Heart Disease prediction using Machine learning and Deep learning techniques

Month: 2024/03 Index Number CS/2018/037

|  |  |  |
| --- | --- | --- |
| **Week** | **Date** | **Description of work carried out, Problems found and the way solved, etc.** |
| **Week 1** |  |  |
|  |  |
|  |  |
|  |  |
| 03/05 | Continued work on CNN implementation. |
| 03/06 | Completed initial training of CNN. |
|  |  |
|  | | |
| **Week 2** |  |  |
|  |  |
|  |  |
|  |  |
| 03/11 | Documented final CNN model and results. |
|  |  |
| 03/13 | Integrated all models for comprehensive comparison. |
|  | | |
| **Week 3** |  |  |
|  |  |
|  |  |
|  |  |
|  |  |
| 03/19 | Finalized the report with all necessary improvements. |
| 03/20 | Rehearsed the presentation for the final session. |
|  | | |
| **Week 4** |  |  |
|  |  |
|  |  |
|  |  |
|  |  |
| 03/26 | Followed up on any remaining feedback. |
| 03/27 | Addressed final feedback and ensured all project components were completed and submitted. |
|  | | |
| **Week 5** |  |  |
|  |  |
|  |  |
|  |  |
|  |  |

Meeting No: ……05……. Meeting Date: 2024-05-10

|  |  |  |
| --- | --- | --- |
| Tasks assigned in the previous meeting | If **completed** briefly describe the way you reached/solved the task/issue | If **not completed** clearly state the reason |
| 1. Continue hyperparameter tuning for all models.  2. Complete CNN implementation..  3. Integrate all models for final evaluation. | * Continued and completed hyperparameter tuning for all models. * Continued and completed hyperparameter tuning for all models. |  |
| Points/tasks/issues discussed at the current meeting | * Review of CNN performance. * Integration of all models for final evaluation. * Discussion on the draft report and presentation. * Feedback on initial results. | |
| Targets/tasks assign to complete before the next meeting | * Finalize the integration of all models. * Complete the final evaluation of models. * Schedule the next meeting. * Add another DL model * Start to write thesis | |

\*\*Please make sure to contact/meet your supervisor/s **at least** once a month

Signature of the supervisor: ………………… Date: ………………

**Self-diary**

Project Title: Heart Disease prediction using Machine learning and Deep learning techniques

Month: 2024/05 Index Number CS/2018/037

|  |  |  |
| --- | --- | --- |
| **Week** | **Date** | **Description of work carried out, Problems found and the way solved, etc.** |
| **Week 1** |  |  |
|  |  |
|  |  |
|  |  |
| 05/16 | Finalize the integration of all models. |
| 05/17 | Give comparison of all models. |
|  |  |
|  | | |
| **Week 2** |  |  |
|  |  |
|  |  |
|  |  |
|  |  |
| 05/24 | Complete the final evaluation of models. |
| 05/25 | Add keras Model |
|  | | |
| **Week 3** |  |  |
|  |  |
|  |  |
|  |  |
|  |  |
| 05/31 | Studied the Keras model |
|  |  |
|  | | |
| **Week 4** |  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  | | |
| **Week 5** |  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

**Self-diary**

Project Title: Heart Disease prediction using Machine learning and Deep learning techniques

Month: 2024/06 Index Number: CS/2018/037

|  |  |  |
| --- | --- | --- |
| **Week** | **Date** | **Description of work carried out, Problems found and the way solved, etc.** |
| **Week 1** |  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
| 06/07 | Implement the Keras model |
|  | | |
| **Week 2** |  |  |
|  |  |
|  |  |
|  |  |
|  |  |
| 06/13 | Finished with keras model |
| 06/14 | Get comparison of CNN and Keras |
|  | | |
| **Week 3** |  |  |
|  |  |
|  |  |
| 06/18 | Make structures for thesis |
| 06/19 | Write a introduction and literature survey |
| 06/20 | Write a introduction and literature survey |
|  |  |
|  | | |
| **Week 4** |  |  |
|  |  |
|  |  |
|  |  |
| 06/25 | Completed introduction and literature survey |
| 06/26 | Tried to get more accuracy |
|  |  |
|  | | |
| **Week 5** |  |  |
|  |  |
|  |  |
| 06/30 | Tried to get more accuracy and low log loss |
|  |  |
|  |  |
|  |  |

Meeting No: ……06………. Meeting Date: 2024-07-03

|  |  |  |
| --- | --- | --- |
| Tasks assigned in the previous meeting | If **completed** briefly describe the way you reached/solved the task/issue | If **not completed** clearly state the reason |
| 1. Finalize the integration of all models.  2. Complete the final evaluation of models.  3. continue the thesis | * Integrated all models for comprehensive comparison. |  |
| Points/tasks/issues discussed at the current meeting | * Review of final evaluation results. * Discussion on improving model accuracy. * Future work and potential improvements. * Run-time and performance of the project. | |
| Targets/tasks assign to complete before the next meeting | * Explore additional techniques to improve model accuracy. * Document future work and potential improvements. * Ensure the project runs smoothly and address any issues. * Schedule the next meeting. | |

\*\*Please make sure to contact/meet your supervisor/s **at least** once a month

Signature of the supervisor: ………………… Date: ………………

**Self-diary**

Project Title: Heart Disease prediction using Machine learning and Deep learning techniques

Month: 2024/07 Index Number: CS/2018/037

|  |  |  |
| --- | --- | --- |
| **Week** | **Date** | **Description of work carried out, Problems found and the way solved, etc.** |
| **Week 1** |  |  |
|  |  |
|  |  |
|  |  |
|  |  |
| 07/06 | Completed the final evaluation of models and documented results. |
| 07/07 | |  | | --- | |  |  |  | | --- | | Worked on improving model accuracy using ensemble techniques. | |
|  | | |
| **Week 2** |  |  |
|  |  |
|  |  |
|  |  |
| 07/14 | Started write Objectives and Requirements Specification |
| 07/13 | Add more to Objectives and Requirements Specification |
| 07/15 | Documented potential future work and improvements. |
|  | | |
| **Week 3** |  |  |
|  |  |
|  |  |
|  |  |
|  |  |
| 07/19 | Started write Methodology part |
| 07/20 | |  | | --- | |  |  |  | | --- | | Ensured the project runs smoothly and addressed any performance issues. | |
|  | | |
| **Week 4** |  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
| 07/27 | Explore additional techniques to improve model accuracy |
|  | | |
| **Week 5** |  |  |
|  |  |
|  |  |
|  |  |

Meeting No: ……07………. Meeting Date: 2024-08-03

|  |  |  |
| --- | --- | --- |
| Tasks assigned in the previous meeting | If **completed** briefly describe the way you reached/solved the task/issue | If **not completed** clearly state the reason |
| 1. Finalize the integration of all models.  2. Finalize the report and presentation.  3. continue the thesis | * Conducted final evaluation and documented results. * Integrated all models for comprehensive comparison. |  |
| Points/tasks/issues discussed at the current meeting | * Review of final evaluation results. * Final adjustments to thesis. * Some points should be correct in thesis | |
| Targets/tasks assign to complete before the next meeting | * Submit the final thesis. * Follow up on any remaining feedback. | |

\*\*Please make sure to contact/meet your supervisor/s **at least** once a month

Signature of the supervisor: ………………… Date: ………………

**Self-diary**

Project Title: Heart Disease prediction using Machine learning and Deep learning techniques

Month: 2024/08 Index Number: CS/2018/037

|  |  |  |
| --- | --- | --- |
| **Week** | **Date** | **Description of work carried out, Problems found and the way solved, etc.** |
| **Week 1** |  |  |
| 08/02 | Prepared the thesis. Finished it |
| 08/03 | Add some more points to literature survey and introduction |
| 08/04 | Add more to methodology |
|  |  |
|  | | |
| **Week 2** | 08/05 | Add implementation part, testing and model |
| 08/06 | Add result part. Completed the thesis |
| 08/07 | Do corrections in thesis |
|  |  |
|  |  |
|  |  |
|  |  |
|  | | |
| **Week 3** |  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  | | |
| **Week 4** |  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  | | |
| **Week 5** |  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |