**Data Science Project 7PAM2002-0509-2023**

**Semester C 2023**

**Logbook (Activities and GitHub submissions)**

**Student Name and ID: NIVEDITHA CHERUKURI (21082045)**

**Project Title: Predicting Chronic Kidney Disease with Machine Learning**

**Supervisor: Luigi Alfonsi**

**Student GitHub URL:** [**https://github.com/niveditha2000**](https://github.com/niveditha2000)

**Number of versions of the code submitted on GitHub: 1**

**User documentation has been submitted on GitHub: YES / NO**

**Student GitHub URL has been shared with markers: YES / NO**

**Log of Activities**

**Must record attendance at lectures and supervisions**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Week** | **Date** | **Activity**  **incl. lectures & supervisions** | **Reason if not attend lecture or supervision** | **Weekly project progress.**  **How lecture/supervision was helpful to your project.** |
| 1 | 14-May-2024 | Lecture 1 |  | Choosed my project topic and conducted initial literature review on that in which my lecture was quite helpful. |
| 2 | 21-May-2024 | Lecture 2 |  | Started drafting my PDM and performed two commits in github. |
| 3 | 28-May-2024 | Supervision 1 |  | Changed PDM according to instructions and suggestions |
| 4 | 03-June-2024 | Lecture 3 |  | Checked whether my dataset is according to data ethics and performed more coding. |
| 5 | 12-June-2024 | Supervision 2 |  | Started working on the dataset, preprocessing and literature review |
| 6 | 17-June-2024 | Lecture 4 |  | Data Ethics and coding progress along with literature review |
| 7 | 03-July-2024 | Supervision  3 |  | Progress review |
| 8 | 15-July-2024 | Lecture 5 |  | Result analysis and preparation of FPR and Viva. |
| 9 | 26-July-2024 | Supervison  4 |  | Code review and discussion on references and citations. |
| 10 | 31-July-2024 | Supervision  5 |  | Discussion on literature review and code complications |
| 11 | 01-Aug-2024 | FPR |  | Started drafting my FPR with proper abstract and introduction |
| 12 | 13-Aug-2024 | Code & FPR |  | Asked my supervisor about doubts in code and changed the code according to, continued drafting FPR |
| 13 | 20-Aug-2024 | Code |  | Replaced plotly confusion matrix with matplotlib |
| 14 | 22-Aug-2024 | FPR |  | Completed my FPR and discussed with supervisor |
| 15 | 23-Aug-2024 | Supervision  6 |  | Online supervision with mock viva and FPR reviewss |

**Log of GitHub Submissions**

**Record the versions of code and user documentation submitted on GitHub**

|  |  |  |
| --- | --- | --- |
| **Date** | **Filename and version submitted to GitHub** | **Description of code and/or documentation submitted (what has been added since the previous version).** |
| 21-05-2024 | README.md | Initial commit |
| 26-05-2024 | .ipynb\_checkpoints | Added source code file to workspace |
| 27-05-2024 | Kidney\_disease.csv | Added dataset file to workspace |
| 27-05-2024 | Chronic\_kidney\_disease\_status\_prediction.ipynb | Started data preprocessing on dataset |
| 30-05-2024 | Chronic\_kidney\_disease\_status\_prediction.ipynb | Changed dataset coloumn names for more readability |
| 03-06-2024 | Descriptions.txt | Added coloumn descriptions. |
| 10-06-2024 | Chronic\_kidney\_disease\_status\_prediction.ipynb | Checking datatypes and extracting numerical and categorical features. |
| 19-06-2024 | Chronic\_kidney\_disease\_status\_prediction.ipynb | Replacing incorrect values and Examining the distribution of features. |
| 01-07-2024 | Chronic\_kidney\_disease\_status\_prediction.ipynb | Checking correlations and filling missing values with random values. |
| 15-07-2024 | Chronic\_kidney\_disease\_status\_prediction.ipynb | Performing label encoding and feature significance |
| 24-07-2024 | Chronic\_kidney\_disease\_status\_prediction.ipynb | Model building and printing accuracy scores. |
| 30-07-2024 | Chronic\_kidney\_disease\_status\_prediction.ipynb | Added plot labels and removed unnecessary code. |