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

**Semester C 2023**

**Logbook (Activities and GitHub submissions)**

**Student Name and ID: Sravani Goka**

**Project Title: USE OF MACHINE LEARNING IN PREDICTIVE MODELING FOR BREAST CANCER DIAGNOSIS**

**Supervisor: Peter Scicluna**

**Student GitHub URL:**<https://github.com/Sravanigooka11/final-project>

**Number of versions of the code submitted on GitHub: 7**

**User documentation has been submitted on GitHub: YES**

**Student GitHub URL has been shared with markers: YES**

**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 | 14th may | Lecture 1 |  | The first lesson covered the overview of the module and allowed me to choose my project and dataset. |
| 2 | 21st may | Lecture 2 |  | Project and data management plan: from this, I was able to see my project clearly and gathered all the necessary information. |
| 3 | 4th June | Supervision 1 |  | The questions, clear information about how to select the best dataset, and deadlines were discussed. |
| 4 | 3rd June | Lecture 3 |  | Here I have got the related information about the data ethics quiz and all the theory and preparation modules |
| 5 | 9th June | Supervision 2 |  | I had a clear note regarding the project completion and submission dates as well as a detailed explanation of the PDP strategy here. |
| 6 | 17th June | Lecture 4 |  | Presentation materials and recording on data ethics available here For reference, I've started reading through my project's thesis and journals. I also obtained precise information on the index and prospective. |
| 7 | 14th June | PDM |  | Project development presentation and data management |
| 8 | 3rd July | Data ethics quiz |  | Attended data ethics quiz |
| 9 | 26th june | Supervision 3 |  | Data ethics quiz and doubts clarification |
| 10 | 15th june | Lecture 5 |  | I got the clear information about the project guidelines and submission and also the plagiarism check about the Turnitin check and the total information about the viva and guidelines |

**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).** |
| 30th july | Update my code. ipynb | My code till date |
| 5st August | Calling and loading data | Modified changes |
| 11rd August | Syntax change in code.ipynb | Total change in code for best algorithm check |
| 24th August | Update my code.ipynb | Total change in code for best algorithm check |
| 26th August | Calling and loading the data | Checking with the different modes of image xrays |
| 27th August | Breast\_cancer\_project1.ipynb | Change in the code according to the new versions |
| 28th August | Breast\_cancer\_project1.ipynb | Modifying and finalising the code with explanation and headings |