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

**Semester C 2023**

**Logbook (Activities and GitHub submissions)**

**Student Name and ID:** Raghavendhra Rao Devineni, 21072747

**Project Title:** Stock Market Prediction**s**

**Supervisor:** Vandana Das

**Student GitHub URL:** [**https://github.com/RaghavendhraRao/stock\_market\_prediction\_ISE\_DSFP**](https://github.com/RaghavendhraRao/stock_market_prediction_ISE_DSFP)

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

**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 | 14/05/24 | Lecture 1 |  | **Module Overview and Choosing a Project** This lecture informed me about the project module and the type of work I will be doing during that time. In addition, the projects that are chosen must be appropriate for the current course. |
| 2 | 21/05/24 | Lecture 2 |  | **Project and Data Management Plan**  I gained knowledge about the PDM (project and data management plan) in this lecture, which aids in project design (i.e., how to work on a project while sticking to ethical principles). Additionally, by following the plan from beginning to end, I created a Gantt chart that makes it easier to monitor the work that has been performed and that still needs to be worked on. helps in submitting the work by the given date. |
| 3 | 28/05/24 | Supervisor session |  | **Supervisor session 1**   * Project discussion. Discussed on the options to choose the project. * Discussed about the area of interest, dataset resources. * Finalized the project. |
| 4 | 03/06/24 | Lecture 3 |  | Data Ethics (part 1) |
| 5 | 06/06/24 | Supervisor session |  | **Supervisor session 2**   * Discussed about the Literature papers relevant to the project. * Discussed about the dataset for the project (choosing dataset must follow GDPR & UH ethical requirements). * Also plan of action for the “Presentation of Project and Data Management Plan”. |
| 6 | 12/06/24 | Project Presentation |  | Given project presentation on the project.  Presented a plan to work on the project.  Explained about the datasets, Machine learning models, Literature papers for the project. |
| 7 | 17/06/24 | Lecture 4 |  | Data Ethics (part 2) |
| 8 | 17/06/24 | Supervisor session |  | **Supervisor session 3**   * Discussed about the work done so far also, tracking the Gantt chart to check the work progress. * Suggestions given by the professor on reading & understanding different literature papers. * Discussed about views on “Research Questions” for project. |
| 9 | 26/06/24 | Supervisor session |  | **Supervisor session 4**   * Discussed about the current work progress and future work to be done. * Discussed about the Methodology, that I have worked so far. * Also Discussed about the “mock viva” preparations. |
| 10 | 03/07/24 | Project ethical quiz |  | Present for the quiz |
| 11 | 08/07/24 | Supervisor session |  | **Supervisor session 5**   * Updated on the literature paper * Given some suggestions on the report work (literature review, methodology...!) * Given update on the code progress |
| 12 | 15/07/24 | Mock viva |  | **Supervisor session 6**   * I had a mock viva regarding the project * Explained about the project so far done, answered couple of question asked by the supervisor. * Received feedback to improve in the report and the project code. |
| 13 |  | Lecture |  |  |
| 14 |  |  |  |  |
| 15 |  |  |  |  |
| 16 |  |  |  |  |

**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).** |
| 09/06/24 | README.md  stock\_market\_prediction\_ds.py (Version -1)  Project Logbook Sem C Data Science 2024.docx | Updated the Readme file about installing packages, read & load dataset.  Created a basic code , reading and visualizing the data. |
| 10/06/24 | stock\_market\_prediction\_ds.py | Worked on preprocessing the data |
| 10/06/24 | stock\_market\_prediction\_ds.py | Visualizing each column and understanding the data & its relations between columns. |
| 16/06/24 | README.md  stock\_market\_prediction\_ds.py (Version -2)  Project Logbook Sem C Data Science 2024.docx | **Created a new project repository** (moved all the files to new repository).  Updated the readme file.  Updated code (reading and changing date datatype, identifying data relations) |
| 17/06/24 | stock\_market\_prediction\_ds.py (Version -3) | Restructure the code.  Defined a function to read and load data. |
| 22/06/24 | stock\_market\_prediction\_ds.py (Version -5) | Preprocessing the data, checking null values, before splitting the data for model training. |
| 23/06/24 | stock\_market\_prediction\_ds.py (Version -6) | Identifying the target features(dependent variable) to be tested, independent variable(training for model). |
| 24/06/24 | stock\_market\_prediction\_ds.py (Version -6) | Splitting the data into train, testing, validation |
|  |  |  |