**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)  This lecture tells about the ethics need to follow for choosing the dataset, kind of dataset need to be selected  About data ethics, research question about the project |
| 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)  In this lecture got to know about the UH Ethical Policy and Procedures, Codes of Conduct need to followed during the project |
| 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 | 16/07/24 | Lecture 5 |  | gives important information about the report and  tips of how to prepare your final report and viva |
| 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 |
| 02/07/24 | ds\_final\_proj\_stock\_market\_colab.ipynb | Build Lstm model from scratch |
| 07/07/24 | ds\_final\_proj\_stock\_market\_colab.ipynb | Updated and tested the model  Train the model |
| 15/07/24 | ds\_final\_proj\_stock\_market\_colab.ipynb  [project\_report.docx](https://github.com/RaghavendhraRao/stock_market_prediction_ISE_DSFP/blob/main/project_report.docx) | Checking the model  Testing the model with different learning rate, checking RMSE values  Also updated the report work on model (LSTM) |
| 16/07/24 | Project Logbook Sem C Data Science 2024.docx  ds\_final\_proj\_stock\_market\_colab.ipynb  [stock\_market\_prediction\_ds.py](https://github.com/RaghavendhraRao/stock_market_prediction_ISE_DSFP/blob/main/stock_market_prediction_ds.py) | Updated project log-book  Updated code on python file and colab |
| 26/07/24 | ds\_final\_proj\_stock\_market\_colab.ipynb | LSTM model tuned by updating “Learning\_rate”  Increased the learning-rate to 0.01 and tested with epochs 100 |
| 29/07/24 | [project\_report.docx](https://github.com/RaghavendhraRao/stock_market_prediction_ISE_DSFP/blob/main/project_report.docx) | Worked Project report |
| 31/07/24 | [project\_report.docx](https://github.com/RaghavendhraRao/stock_market_prediction_ISE_DSFP/blob/main/project_report.docx) | Worked Project report |
| 02/08/24 | [project\_report.docx](https://github.com/RaghavendhraRao/stock_market_prediction_ISE_DSFP/blob/main/project_report.docx) | Worked Project report and updated all work |
| 06/08/24 | ds\_final\_proj\_stock\_market\_colab.ipynb  [project\_report.docx](https://github.com/RaghavendhraRao/stock_market_prediction_ISE_DSFP/blob/main/project_report.docx) | Worked on code , checking the results of the model, tuning the parameters.  Worked Project report |
| 10/08/24 | ds\_final\_proj\_stock\_market\_colab.ipynb  [project\_report.docx](https://github.com/RaghavendhraRao/stock_market_prediction_ISE_DSFP/blob/main/project_report.docx) | Worked on model , feature extraction of the model and trained the model.  Worked Project report and updated the work on report. |
| 12/08/24 | ds\_final\_proj\_stock\_market\_colab.ipynb | Created the ARIMA model and check the modle results with LSTM model |
| 16/08/24 | ds\_final\_proj\_stock\_market\_colab.ipynb | Worked on developing the ARIMA model and also updated the LSTM model for better accuracy |
| 18/08/24 | ds\_final\_proj\_stock\_market\_colab.ipynb | Increased the accuracy for both models along with learning rate and epochs |
| 19/08/24 | ds\_final\_proj\_stock\_market\_colab.ipynb | developing the ARIMA and LSTM models.  Increased the accuracy for both models along with learning rate and epochs |
| 20/08/24 | ds\_final\_proj\_stock\_market\_colab.ipynb  [project\_report.docx](https://github.com/RaghavendhraRao/stock_market_prediction_ISE_DSFP/blob/main/project_report.docx) | Checking the model results and analysing the code and results  And worked on report (Didn’t commit) |
| 23/08/24 | ds\_final\_proj\_stock\_market\_colab.ipynb  [project\_report.docx](https://github.com/RaghavendhraRao/stock_market_prediction_ISE_DSFP/blob/main/project_report.docx) | Changes in the code, checking results,  Analyzing the results through plots and error metrics.  And worked on report (Didn’t commit) |
| 24/08/24 | ds\_final\_proj\_stock\_market\_colab.ipynb  [project\_report.docx](https://github.com/RaghavendhraRao/stock_market_prediction_ISE_DSFP/blob/main/project_report.docx) | Added and updated the code  Created a new model called ‘Prophet’  Comparing the results among the models  Choosing the best model  Worked on report and updated the format in the report. |
|  |  |  |
|  |  |  |