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

**Semester C 2023**

**Logbook (Activities and GitHub submissions)**

**Student Name and ID:** Harish Kumar Kodidela (21060336)

**Project Title:** Prediction of Precipitation using Machine Learning Approaches over Bangalore Region of India

**Supervisor:** Dr. Pushp Raj Tiwari

**Student GitHub URL:** <https://github.com/harishkodidela/Data-Science-Project.git>

**Number of versions of the code submitted on GitHub: 13**

**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 | 14/05/2024 | Lecture 1 |  | Read the project handbook. Discussed to attend all the lectures and supervisions. Choose and send emails to supervisor about the project topic and dataset and get agreed with the supervisor. Save work in different places. |
| 2 | 21/05/2024 | Lecture 2 |  | We talked about the project and data management plan and how it should be presented. We also decided to start the literature review and to review papers and doing timely updates to supervisor regarding work. |
| 3 | 03/06/2024 | Supervision 1 |  | We talked about how to plan project and how to come up with a good question for study. We talked about the literature study, discussion about dataset changes and how to start initial work and PDM plan. |
| 4 | 03/06/2024 | Lecture 3 |  | Discussed about Data Ethics importance in project and having quiz on the this on July 3rd. |
| 5 | 14/06/2024 | Supervision 2 |  | Presented the Project and Data Management Plan to supervisor and submission of PDM report on canvas before deadline, worked on data set. |
| 6 | 17/06/2024 | Lecture 4 |  | Data Ethics continuation class discussion about UH ethics policies and regulations and UK government data ethics. Required reading which is essential for quiz coming on July 3rd. |
| 7 | 03/07/2024 | Data Ethics Quiz |  | Completed Data Ethics Quiz |
| 8 | 15/07/2024 | Supervision 3 |  | Completed all pre-processing sections of code and discussed ML model implementations and how to work out models with best metrics. |
| 9 | 15/07/2024 | Lecture 5 |  | Discussion on results analysing and final project report. |
| 10 | 19/07/2024 | Supervision 4 |  | Model Training & Testing done, working on some visualization of dataset. Discussion on geo spatial analysis and pdf form of data to work out. |
| 11 | 26/07/2024 | Supervision 5 |  | Discussion on some changes in EDA and model training. |
| 12 | 09/08/2024 | Supervision 6 |  | Changes in Gamma fit distribution, working on model score improvement and reviewing the code with necessary changes required. |
| 13 | 13/08/2024 | Supervision 7 |  | Reviewing the presentation having plots of all outputs and model implementation with changes suggested. |
| 14 | 19/08/2024 | Supervision 8 |  | Final presentation discussion with all code, plots and time series analysis of predicted values with all necessary requirements. |
| 15 | 23/08/2024 | Supervision 9 |  | Report review with all requirements followed from project handbook and logbook update. |

**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).** |
| 16/06/2024 | README.md (version - b95149) | Initial file commit for data science project repository. |
| 20/06/2024 | Imd\_data\_csv.ipynb  (version - 756e1f7) | Code for fetching dataset from IMD library of python package |
| 27/06/2024 | bengaluru1972\_2022.csv  (version - 234f712) | Dataset for the project having data from imd website. |
| 01/07/2024 | Precipitation\_Pred).ipynb  (version - b4a2b74) | Code for project, done with initial part and working on pre-processing section. |
| 10/07/2024 | Precipitaion\_Pred.ipynb  (version – f9bce6e) | Done with initial and pre-processing sections of code |
| 15/07/2024 | Precipitation\_Pred.ipynb  (version - 1c9c2ed) | Model training and testing done |
| 25/07/2024 | Precipitaion\_Pred.ipynb  (version - 33ecb9a) | Changes in models to get better scores trying with other models and implementation. |
| 11/08/2024 | Precipitation\_Pred.ipynb  (version - 012f064) | Changes in numeric features and working on results. |
| 21/08/2024 | Precipitation\_Pred.ipynb  (version - 1139cf7) | Worked on time series analysis plotting for actual and predicted values to get better visuals on results. |
| 26/08/2024 | Precipitation\_Pred.ipynb  (version - 8218bd1) | Working on comments and code organization, putting headings and subheadings for better understand. |
| 28/08/2024 | geospatial\_analysis.ipynb  (version - 7dc01fa) | Work on geo spatial analysis with rainfall and temperature data plotting in geo spatial map. |
| 28/08/2024 | Precipitation\_Pred.ipynb  (version - 28e430d) | Final checks on code with all sections and flow of code maintained well. |