Data Glacier

Your Deep Learning Partner

FA

Project lifecycle along with the deadline

Week 7: 13 – 19 November 2022

Deliverables

1- Problem description (14 November 2022)
2- Business understanding (15 November 2022)
3- Project lifecycle along with the deadline (16 November 2022)
4- Data Intake report (17 November 2022)
5- GitHub Repo link (18 November 2022)
6- Deadline for this part (19 November 2022)

Week 8: 20 – 26 November 2022

Deliverables

1- Problem description (20 November 2022)
2- Data understanding (21 November 2022)
3- What type of data you have got for analysis (22 November 2022)
4- What are the problems in the data (number of NA values, outliers, skewed, etc.) (23 November 2022)
5- What approaches you are trying to apply to your data set to overcome problems like NA value, outlier, etc., and why? (24 November 2022)
6- GitHub Repo link (25 November 2022)
7- Deadline for this part (26 November 2022)

Week 9: 27 November – 2 December 2022

Data Cleansing and Transformation

Submit a pdf document and ipynb notebook which should contain the following details:

1- Data cleansing and transformation have been done on the data. (28 - 29 November 2022)
2- Try at least 2 techniques to clean the data (for NA values: mean/median/mode/Model-based approach to handle NA value/WOE and like this try different techniques to identify and handle outliers as well) (30 November, 1 – 2 December 2022)
3- Deadline for this part (2 December 2022)

Week 10: 3 – 9 December 2022

Deliverables

Submit a pdf document and EDA ipynb file which should contain the following details:

Team member's details: Group Name (give a name to your group), Name, Email, Country, College/Company, Specialization (Data Science, NLP, Data Analyst)

1- Problem description (3-4 December 2022)
2- GitHub Repo link (4 December 2022)
3- EDA performed on the data (5,6 December 2022)
4- Final Recommendation (7,8 December 2022)
5- Deadline for this part (9 December 2022)

Week 11: 10 – 16 December 2022

EDA Presentation and proposed modeling technique

Team member's details: Group Name (give a name to your group), Name, Email, Country, College/Company, Specialization (Data Science, NLP, Data Analyst)

1- Problem description (11 December 2022)
2- GitHub Repo link (12 December 2022)
3- EDA presentation for business users (13 - 15 December 2022)
4- Deadline for this part (16 December 2022)

The last slide of EDA should be dedicated to the technical user and should contain recommended models for this data set.

Week 12: 17 -23 December 2022

Model Selection and Model Building/Dashboard

1- Select your base model and then explore 1 model of each family if its a classification problem then 1 model for Linear models, 1- Model for Ensemble, 1-Model for boosting, and other models if you have time (like stacking) (18 - 22 December 2022)
2- Deadline for this part (23 December 2022)

Week 13: 24 – 30 December 2022

Final Project Report and Code

1- Provide the link to your code and report.
2- select the solution which is best and as per the requirement.
3- PowerPoint presentation is a must.
4- Deadline for this part (30 December 2022)