DJS-Compute

Machine Learning

Assignment-4

Topic: Linear and Logistic Regression

General Instructions:

- The deadline for completing the tasks is 30th Dec, 2022. A discussion session, addressing any doubts and task related topics, will be conducted in the following days.
- You can refer to additional online resources if needed .

LINEAR REGRESSION

Reference material:

- Supervised vs Unsupervised vs Reinforcement
- https://machine-learning.paperspace.com/wiki/supervised-unsupervised-and-reinforcement-learning
- Linear Regression Master Notebook (Kaggle)
- Evaluation Metrics for Linear Regression

Task:

Dataset Link:
Dataset

Colab Link:
https://colab.research.google.com/drive/1pUYCVcl6rOoa0e91LuEm9NVPJLESo_1E?usp=sharing

LOGISTIC REGRESSION

Reference material:

- Introduction to Logistic Regression
- Logistic Reg. a bit more in depth with the code
- Performance Metrics - Precision and Recall

Task-1:

Dataset Link:
heart_study.csv

Colab Link:
https://colab.research.google.com/drive/1PmUPdwAdbKwzUNGSatCDfBRAvBDgUNaW?usp=sharing

Note: Make sure to answer each question with relevant code or text explanation **as per your understanding.**

Task-2:

Dataset:
The following dataset gives information about whether a person had clicked on an ad or not given Daily Time Spent on Site, Age, Area Income, Daily Internet Usage, Ad Topic Line, City, Male, Country and Timestamp.

https://drive.google.com/file/d/1rSdGgfiwwAhkW25MVAz_tezyYrbDV578/view?usp=sharing

1. Perform Data Cleaning
2. Select Appropriate Features for prediction
3. Plot relevant visualizations
4. Apply logistic regression on 70% of the original dataset
5. Plot the confusion matrix for the remaining 30% of the dataset (test set).