EAS 595 Assignment 2

Date: March 11, 2022

Total: 50 Points

Deadline: March 20, 2022, 11.59PM

You are free to use any computing resources available at your disposal. Please check out free resources such as Google Collab. The assignment is expected to be submitted in Python. There is no restriction for python libraries. Doubts should be cleared up two days before the deadline for submission. Plagiarism and copying material are not allowed. You are allowed to use various sources for reference but mention the references used.

Submission Details – Submit one jupyter notebook (.ipynb) that includes all the 3 questions and solutions for the same. Create a 5 minute demo video explaining the code. Create a Zip file along with the notebook, data, and video. Submit the zip file. Naming Convention – ubitname_a2.zip (ex: sachinge_a2.zip)

Dataset – You can use your own dataset (no restriction on the dataset)

https://www.kaggle.com/camnugent/california-housing-prices?select=housing.csv

Q1) 20 Points

Implement Linear regression using closed-form approach. You are allowed to use ML libraries such as sklearn or scikit learn.

Closed Form Equation $\Theta = (X^T X)^{-1} X^T y$

Follow Q3 in Quiz 3 and implement an algorithm to calculate the theta values and form the linear equation.

a. Simple Linear regression using 1 feature
   $Y = \Theta_0 + \Theta_1 x$ (9 points)
b. Multiple Linear Regression using 2 – 4 features of your choice from the dataset
   $Y = \Theta_0 + \Theta_i x_i$ (9 points)
c. Report evaluation metrics such as Rsquared error, Mean square error, Mean absolute error. (2 points)

Q2) 20 Points

Implement Linear regression using Stochastic Gradient Descent. You are allowed to use ML libraries such as sklearn or scikit learn. Report evaluation metrics such as R-squared error, Mean square error, Mean absolute error

Follow Q1 in Quiz 3 and implement an algorithm to calculate the theta values and form the linear equation.

$\Theta = \Theta$ – learning_rate.error.X

You can make assumptions for learning_rate and initial theta values

a. Simple Linear regression using 1 feature
   $Y = \Theta_0 + \Theta_1 x$ (9 points)
b. Multiple Linear Regression using 2 – 4 features of your choice from the dataset
   $Y = \Theta_0 + \Theta_i x_i$ (9 points)
c. Report evaluation metrics such as Rsquared error, Mean square error, Mean absolute error. (2 points)

Q3) 10 points

Implement a Multi-Layer perceptron with 2 hidden layers for the entire dataset. You are allowed to use ML libraries such as sklearn or scikit learn. Report evaluation metrics such as R-squared error, Mean square error, Mean absolute error. The idea is to perform regression using neural networks.