Final Project

MONTGOMERY COUNTY HOUSE SALES PREDICTION

Introduction:

The "Montgomery County House Sales Price Prediction " is a project prepared after attending a three-month course organized for grad students in partial fulfillment of the course DATA602-Introduction to Machine Language for Data Science program at the University of Maryland Baltimore County.

Project Overview

As a first experience, I preferred to engage in the house sale regression analysis algorithm and predict to focus on some features in the dataset.

The data was obtained from the state of Maryland open data source

Project Intro and Objective:

The goal of the project is to predict the future house sales in Montgomery county. the pattern of the Employee Income and Spending habit using unsupervised learning clustering K means algorithm **.**

Methods Used

In this section, linear regression, ridge regression, and lasso regression will be compared and see how well each model makes a prediction.

Data Size : 3.5MB

Technologies

- **Python**
- **Jupiter Notebook /Panda**
- **Scikit-learn library**

In this project K Means was used to study to classify and form clusters . to this effect, Elbow method was used to find out the size of "K" equal to 5 to place similar cluster together and data points in the different clusters are farther apart.

Needs of this project:

- **data exploration/descriptive statistics**
- **data processing/cleaning**
- **statistical modeling**

2 Conclusion

Based on the analysis conducted to predict the price of house sale in Montgomery County for classification problem, the best-performing model was Lasso which shrink the data and reduced

the outlier to perform 0.81 accuracy rate the best-performing model to resolve the regression problem.

3 Future Work:

This project is my first experience. I planned to work on same project with additional data models and optimize the performance.