

Group Presentation Personal Work Explanation

Author: Peizhao Wang

UCL Student Number: 21039541

Work within the presentation:

- Finding the dataset we used in the presentation.
- Early data exploration with visualization.
- Data preprocessing.
- Finishing the cluster model and PCA with Yuning.
- Finding the best number of PCA components.
- Finding the best number of clusters with highest silhouette score.
- K-means clustering test & result visualization with Yuning.

Extended work (not included in presentation):

- Multiple regression model with considering 14 unique feature variables to predict house price, residual analysis, obtaining R-squared value of 0.717.
- Multi-variable Polynomial Regression model with considering 14 unique feature variables to predict house price, residual analysis, obtaining almost perfect R-squared value of 1.
- Fully-connected Neural Network model with considering 14 unique feature variables to predict house price, residual analysis, obtaining R-squared value of 0.727.
- Fully-connected Neural Network model with considering 4 unique feature variables in 5 years of data to predict house price, residual analysis, obtaining R-squared value of 0.9357. Obtaining MSE of 0.01 in normalized view.
- Support Vector Machine model with considering 4 unique feature variables in 5 years of data to predict house price, residual analysis, obtaining R-squared value of 0.923. Obtaining MSE of 0.01 in normalized view.