

BLM-3120 Information Retrieval and Web Search Engines Term Project Proposal

Subject: House Price Assignment

Group: Engin Deniz Çağlar - 17011080 & Umut Sümer - 19011618

Dataset: <https://www.kaggle.com/shivachandel/kc-house-data>

1. Why?

Making a decision when buying high cost products has always been a struggle for customers. Houses are unarguably one of them and varying prices in the market makes it even harder. That's why we aim to analyse house prices using machine learning algorithms.

2. What?

It recommends an price range for a house with specific features, according to the prices and the features we supplied to the algorithms.

3. How?

We will use a dataset from Kaggle with features and prices of houses. There are 21613 rows and 21 columns in the dataset. Columns of the dataset correspond to most significant features like bedrooms, bathrooms, floors, ratings or square footage of a room.

We will train different machine learning algorithms with these features. Eventually the algorithms will tell us what is the best price range for the given features of a house.

We will use Support Vector Classifier, Decision Tree, Random Forest, K-Nearest Neighbors and XGBoost algorithms.

4. When?

Week 9-10 : We will implement the algorithms to the dataset and compare the results of different ML algorithms.

Week 11 : We will build a GUI using the algorithm with the best accuracy score.

Week 12 : We will do various tests and check the outputs.