

MLP 3 - Real estate evaluation web application.

Industry

Real estate | Investment

Skills

Python | EDA | Linear regression | Machine learning model creation and design | Machine learning model deployment | Model testing | Docker | Machine learning | FLASK | API | Postman | JSON

Problem statements

A REIT company wanted to predict the evaluation of properties to identify new investment opportunities for the real estate investment portfolio.

Data Description

Each feature name corresponds to a column in the dataset's input features (data). The printed output will show the names of the features, typically including information such as:

- CRIM: Per capita crime rate by town.
- ZN: Proportion of residential land zoned for lots over 25,000 sq.ft.
- INDUS: Proportion of non-retail business acres per town.
- CHAS: Charles River dummy variable (= 1 if tract bounds river; 0 otherwise).
- NOX: Nitric oxides concentration (parts per 10 million).
- RM: Average number of rooms per dwelling.
- AGE: Proportion of owner-occupied units built prior to 1940
- DIS: Weighted distances to five Boston employment centers.
- RAD: Index of accessibility to radial highways.
- TAX: Full-value property tax rate per \$10,000.
- PTRATIO: Pupil-teacher ratio by town.
- B: $1000(B_k - 0.63)^2$ where B_k is the proportion of Black people by town.
- LSTAT: Percentage of lower status of the population.
- MEDV: Median value of owner-occupied homes in \$1000s.

Methods

- Used linear regression.
- Normalized data to train ML model
- Denormalized again after model trained

Results

- Accurate model
- R-squared score of 0.711 is generally considered a reasonably good score for a linear regression model.
- An adjusted R-squared value of 0.684 is generally considered a good score for a linear regression model.