

Can we predict house pricing?

Workshop Lombok {Dev}

Menurut temen2 harga rumah itu pengaruh nya apa ?

- luas tanah
- lokasi
- jumlah kamar
- luas bangunan
- dll

**Apakah data berbicara
seperti itu juga?**

Let's explore it

Sudah dapat intuisi dari data tersebut ?

- Korelasi adalah sesuatu yang sangat diperhatikan
- Distribusi sebuah variabel sedikit memberikan insight mengenai current status

Machine Learning

Sub Bagian Machine Learning

- Supervised Machine Learning
- Unsupervised Machine Learning

Variable

- Regressi
 - Untuk memprediksi continuous number, contoh: gaji, harga, tinggi badan, harga rumah
- Klasifikasi
 - Untuk memprediksi categorical value, contoh : spam/ tidak spam, fraud/tidak fraud, jomblo/tidak jomblo, suka/tidak suka

Current Case ?

of course Continious Number

Flow ?

**3 Menit, bayangkan dulu
kita mau belajar apa aja
tahap nya waktu SD?**

Ada 3 Tahap dalam membuat machine learning model

- Preparation Train Data
- Train the Model
- Test the Model

Preparation Test Data

- Load the Data
- fill the null value (for now, just fill it with 0)
- change the categorical value to number
- Split the data into 2 part, train data and test data

Train Model

- Choose Algorithm for regression:
 1. LinearRegression
 2. SVR
 3. DecisionTreeRegressor
 4. XGBoostRegressor

**Don't forget to do the
Cross Validation to avoid
overfitting**

Test The Model

- Test data must be not part of Train Data
- Test data must be randomly selected
- Predict the result using the model and test data
- Use metrics, Mean Absolute Error and Mean Squared Error to know our performance

Extra !!!!

Upload your result to
kaggle.com

BEAT MY SCORE 0.13950