

Day 47

Jan. 17, 2022

Intro to Machine Learning from kassle

## How Models work

Models work by learning data and from this they make predictions.

→ Similar to what we can say as intuition / experience.

→ They make predictions from experience.

# Machine learning model - Decision tree

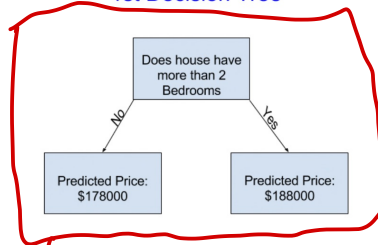
The decision tree is an example of a ML model

## Sample Decision Tree

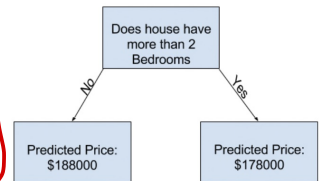
### Improving the Decision Tree

Which of the following two decision trees is more likely to result from fitting the real estate training data?

1st Decision Tree



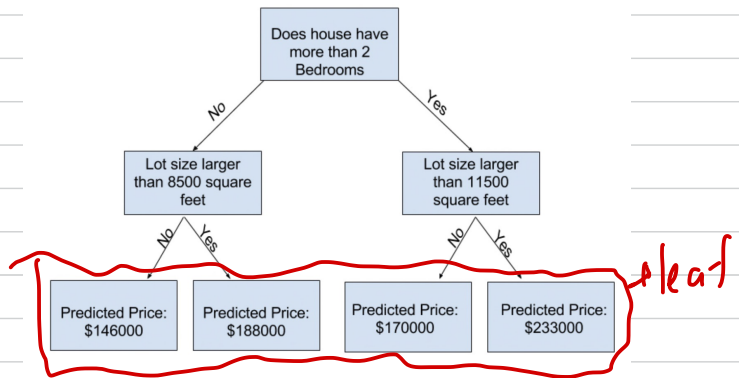
2nd Decision Tree



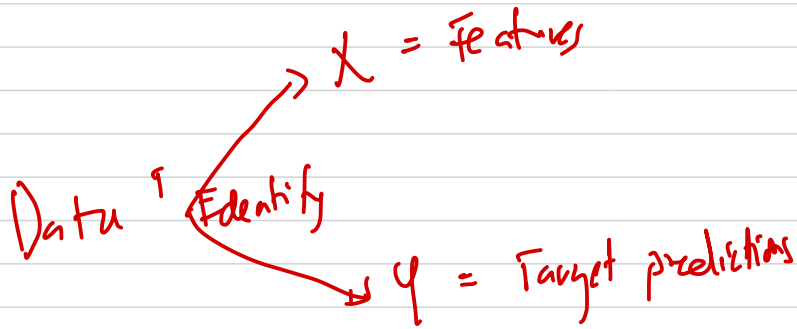
Butter model because  
in reality as no of  
bedrooms ↑, price  
also increases.

More branches can be included in a decision tree to make/improve decisions.

The house example, factors like sizes, locations, type of house can be included.

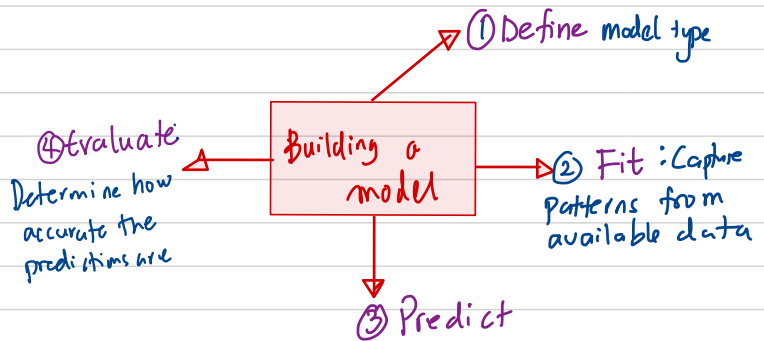


The point where the final price is decided



Scikit learn will be used to  
predict model.

## Building a model - DFPE



Several ways to test a model accuracy. Common one is

Mean Absolute Error (MAE) =

Average  $| \text{Predicted} - \text{Actual} |$

↑  
absolute

## The Problem of "In-Sample" Scores

• Don't use the same sample for "building" & "Evaluating" the model

o In-sample score : measure of a model's accuracy based on the data used to train the model

o Quite misleading because it may not accurately reflect the model's performance on new, unseen data.

o Use validation data or data that was not used to train the model, to test its accuracy.