Instance Based vs Model Based

Tuesday, December 3, 2019   7:21 PM

Note Written by: Zulfadli Zainal

Main focus: How machine learning algorithm generalizing the data.

2 main approach in generalization: Instance Based vs Model Based

Instance Based

Similarly said = Learned by Heart

Algo will compare features and "generalize' the outcome based on similarity comparison or etc, etc.

 Eg;

"Instead of just flagging emails that are identical to known spam emails, your spam filter could be programmed to also flag emails that are very similar to known spam emails. This requires a measure of similarity between two emails. A (very basic) similarity measure between two emails could be to count the number of words they have in common. The system would flag an email as spam if it has many words in common with a known spam email."
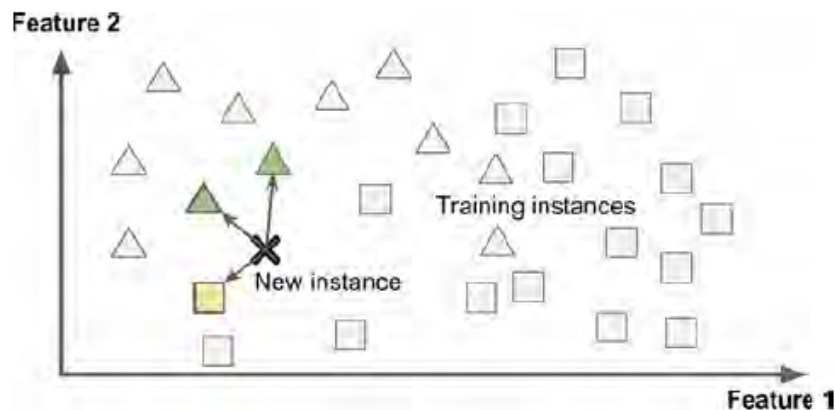
Model Based

Another way to generalize: Make model -> And use model to make predictions.
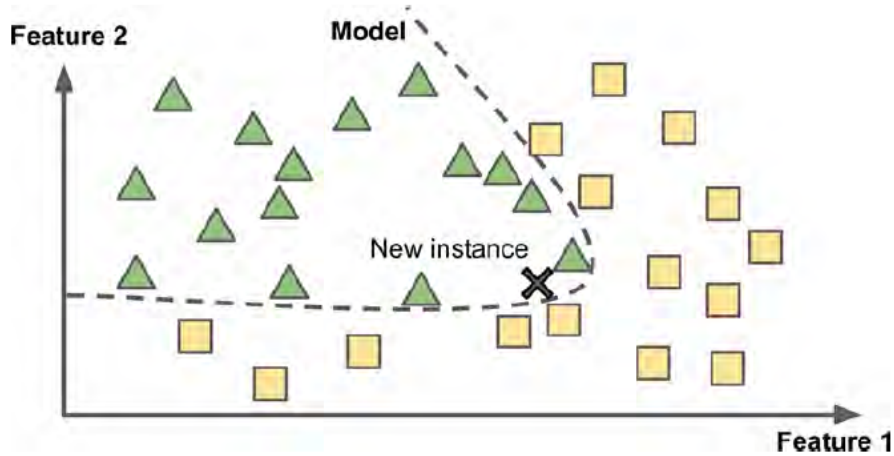

Example of model creation:

1. From data, visualize.

Table 1-1. Does money make people happier?

Country	GDP per capita (USD)	Life satisfaction
Hungary	12,240	4.9
Korea	27,195	5.8
France	37,675	6.5
Australia	50,962	7.3
United States	55,805	7.2

Let's plot the data for a few random countries

2. Do you see a trend here? There does seem to be a trend here! Although its noisy, its seems like people more happy linearly if country GDP is higher.
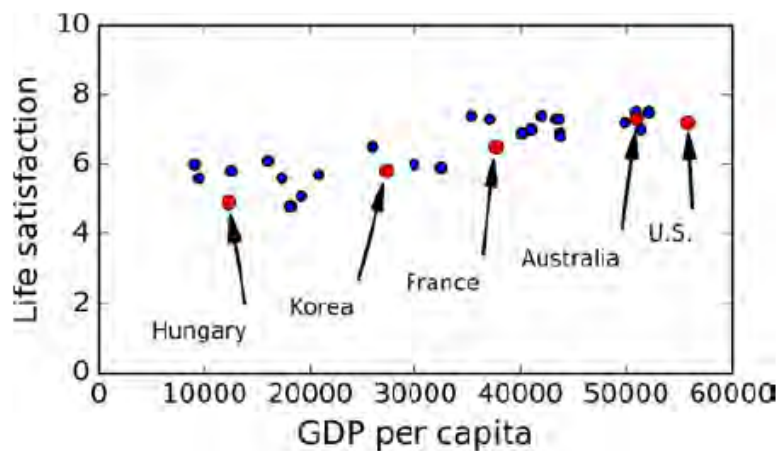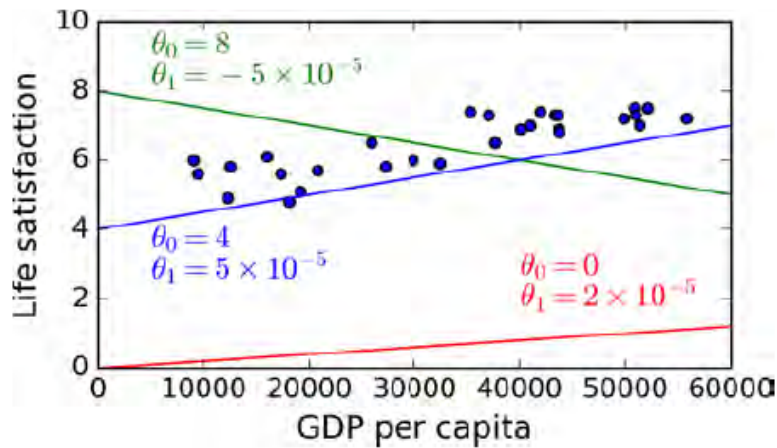
3. Trend looks very linear. Choose model based on this. In this case linear model.
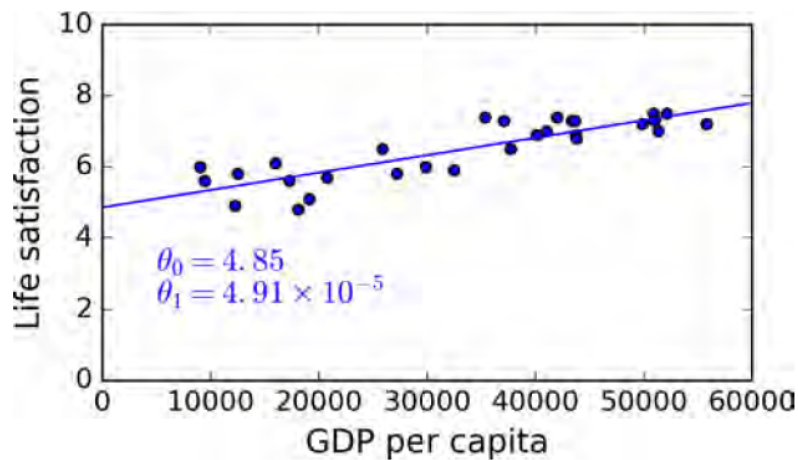
Equation 1-1. A simple linear model

$$life\_satisfaction = \theta_0 + \theta_1 \times GDP\_per\_capita$$

4. This model has two model parameters, θ0 and θ1. By tweaking these parameters, you can make your model represent any linear function.

5. Need to find θ0 and θ1 that closely fit the model.

6. Now the model can be used directly.