# Regression Overview

Linear and Logistic regression discussion

Linear versus Logistic Regression

- Linear Regression
- Continuous function output
  - E.g.: How much will this cost
- Output based on a derived formula
- Formula MUST be linear
- When there is only feature it is called *Uni-variate* Linear Regression
- If there are multiple features, it is called *Multi*ple Linear Regression.

- Logistic Regression
- Discrete function output
  - E.g.: What class / Over or under a certain price
- Formula MAY be linear (but can vary)
- Output usually sigmoid (yes/no, a/b)... multiple values allowed, but rare
- Linear function based on one or more predictive factors

Visual: Linear versus Logistic

Linear Regression: Cost vs data
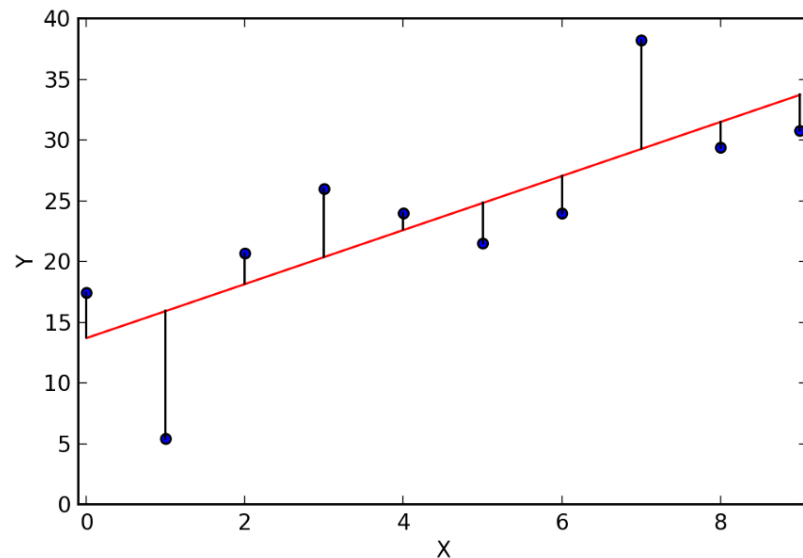
Logistic regression model
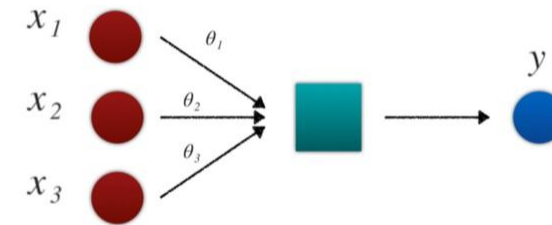
Image Credit : toshistats.net

Logistic Regression versus K nearest neighbors

- Logistic Regression
- Some training required (to find the line)
- Uses linear distance (error squared of y distance) to plot the line
- Line ends up 2-dimensional
- Prediction is probability of classifier, sigmoid output

- K nearest neighbors
- No training required
- Uses distance formula from boundary
- Separation can be n-dimensional, so linear across multiple dimensions (hyperplane)
- Better with more data/classifier
- Prediction is value of classifier

Visual: Logistic versus KNN
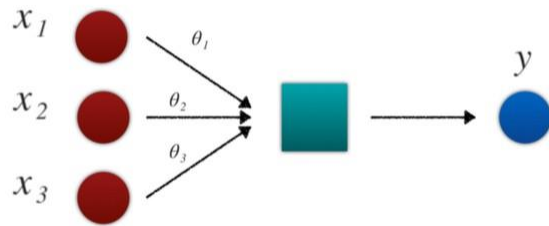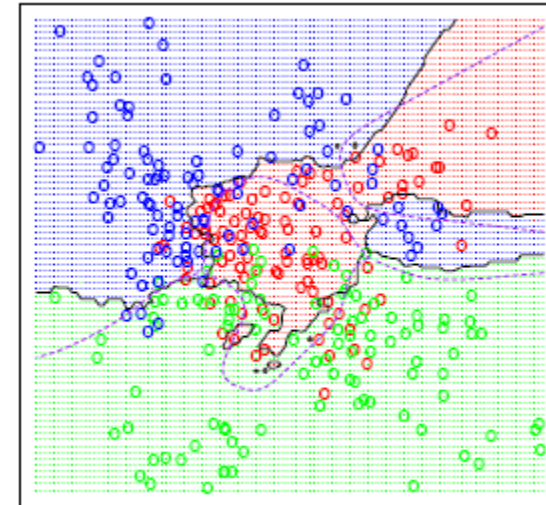
Logistic regression model

Image Credit : toshistats.net

https://towardsdatascience.com/machine-learning-part-3-logistics-regression-9d890928680f

k-Nearest-Neighbors

https://towardsdatascience.com/introduction-to-k-nearest-neighbors-3b534bb11d26