Machine Learning

Abdelhak Mahmoudi
abdelhak.mahmoudi@um5.ac.ma

INPT- 2020

Content

1. The Big Picture
2. Supervised Learning
   - Linear Regression, Logistic Regression, Support Vector Machines, Trees, Random Forests, Boosting, Artificial Neural Networks
3. Unsupervised Learning
   - Principal Component Analysis, K-means, Mean Shift

Supervised Learning

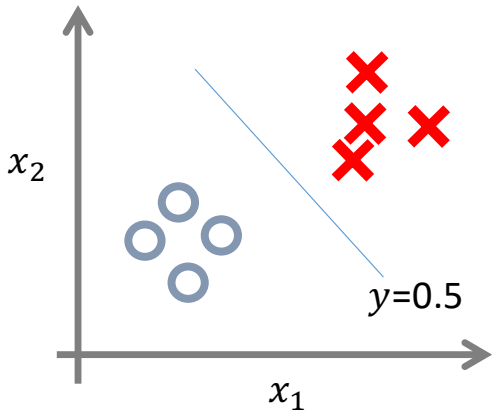
- Linear Regression
- **Logistic Regression**
- Support Vector Machines
- Trees (Decision and Regression)
- Random Forests
- Boosting
- Artificial Neural Networks

Logistic Regression

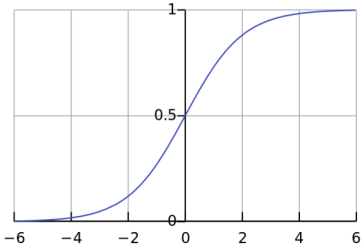
- The output y is discrete
- Classify X with a line $y = g(w_0 + w_1 x_1 + w_2 x_2)$
- The best line is the one with minimum loss

$$\mathrm{L}(w) = \frac{1}{m}\sum_{i=1}^{m}\left[\hat{y}^{(i)}\log(y^{(i)}) + (1-\hat{y}^{(i)})\log(1-y^{(i)})\right]$$

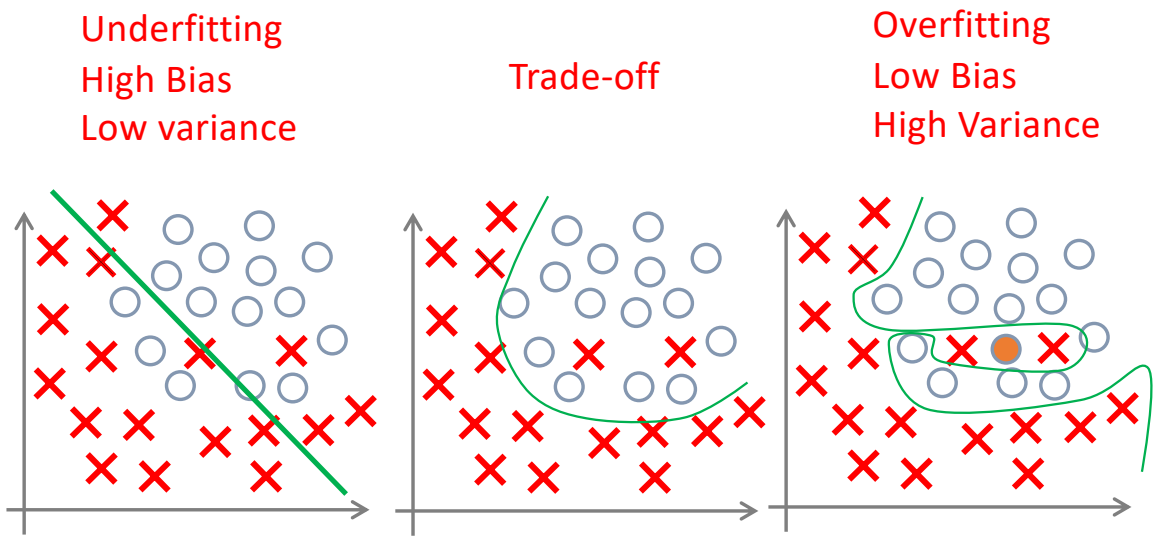
- Solved with gradient descent



$$g(z) = \frac{1}{1+e^{-z}}$$



Overfitting vs. Underfitting



Linear and Logistic Regression

- Hyper-Parameters Tuning
  - λ : regularization hyper-parameter
  - d: degree of polynomial