Wednesday, February 20, 2019

Machine Learning

Chapter 1

Learning => supervised, semi-supervised, unsupervised and reinforcement.

- Supervised Learning
  - Goal: Use the dataset to train a model which can predict the label based on a given feature.
  - Dataset <- labeled examples, $\{(\mathbf{x}_i, y_i)\}_{i=1}^{N}$.
  - Each x is a feature vector //input, which describe the example, each y is a label// output
  - Usually when writing programs D::=number of features, N::=number of the sample.

- Unsupervised learning
  - The same thing using unlabeled dataset. $\{\mathbf{x}_i\}_{i=1}^{N}$.
  - Either transform x into another vector, or into a value that can be used.

- Semi-Supervised Learning
  - Using both labeled and unlabeled data. Hope that unlabeled can help to train a better model.

- Reinforcement Learning
  - Learn policy(function) in an environment? Feel like something super filed of supervised/semi/unsupervised learning?

A demo of supervised learning

- Data:
  - In pairs (x,y) // x could be anything, but y are usually real numbers, for predicting (for computer convenience) or labels (spam, not_spam, etc.), or vectors ◦
- Algorithm:
  - Support Vector Machine
    - Requires the positive labels to be +1 and negative labels to be -1.
- Prediction function y=wx-b
  - w ::= a real-valued vector of the same dimensionality as x.
  - b ::= adjustment
  - The goal of SVM : To find optimal w and b.
- Also try to maximize the MARGIN (to make the boundary conditions more clear).

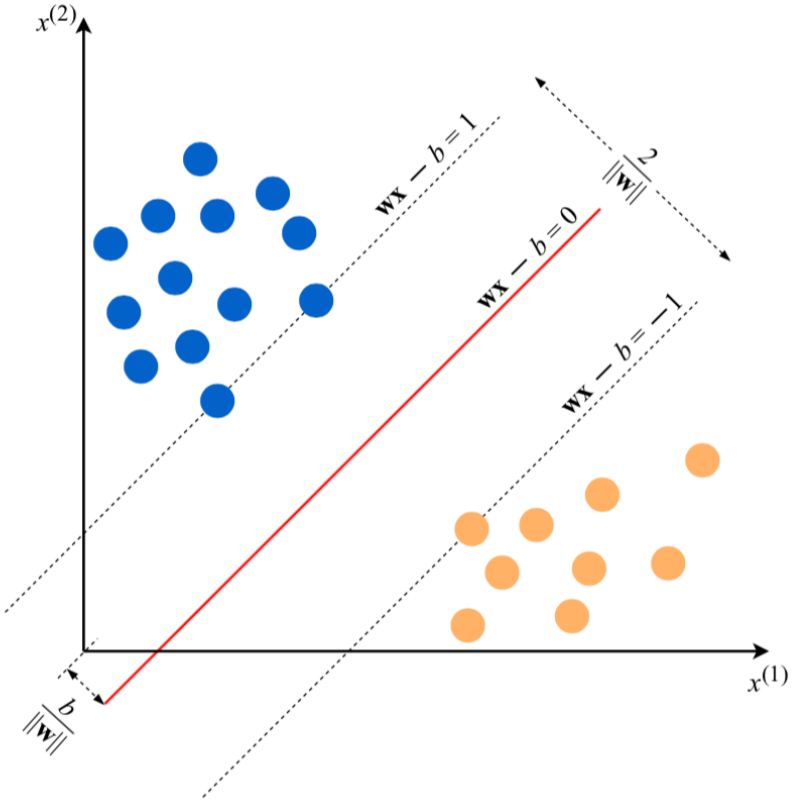

Figure 1: An example of an SVM model for two-dimensional feature vectors.