51.504 Machine Learning (2020)
Lecture Note 03: Classification

Jon Kartago Lamida

2020-10-01

Examples

- Tumor Classification
- Spam Filters

Methodology

Machine Learning > Supervised Learning > Classification

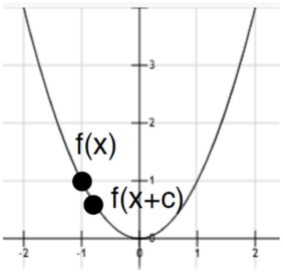
- **Task**. $h : \mathbb{R}^d \to \{-1, +1\}$ such that $y \approx h(x; \theta)$
- **Experience**. Training data $(x^{(1)}, y^{(1)}), \ldots, (x^{(n)}, y^{(n)})$
- **Performance**. Prediction error on test data

Gradient Descent

The slope at a point is called the *derivative* at that point

Intuition: Measure the slope between two points that are really close together

$$\frac{f(x+c) - f(x)}{c}$$



Limit as c goes to zero

Linear Classifier

Methodology

Model

Test Loss

Decision Region

Linearly Separable

Not Linearly Separable

Perceptron Algorithm

Linear classifiers are often also called perceptrons.

Zero-One Loss

Training Loss

Hypothesis Function

The hypothesis function is given by

$$h_\theta\left(x^{(i)}\right) = sgn\left(\left\langle \theta, x(i) \right\rangle\right),$$

where

$$sgn(z) = \begin{cases} 1 & if z \geq 0 \\ -1 & if z < \end{cases}$$

Loss Function

Perceptron Algorithm

Error Reduction

Perceptron Summary

Convergence

Example

Disadvantages

Logistic Regression

Almost Linearly Separable

Sigmoid Neuron

Probabilistic Model

Sigmoid Function

Solving for Hyperlapse

Decision Boundary

Sigmoid Function Formulas

LOG-Likelihood Function

Gradient

Gradient Ascent

Multiclass Classification