

Overview

1 Regression

Set of statistical processes for estimating the relationship between a dependent variable (also called the “outcome variable”) and one or more independent variables (also called “predictors”, “covariates” or “features”). Usually related to continuous variables.

Algorithms:

- **Linear Regression:** establishes linear function from features to the outcome variable
- **Linear Regression with Basis Function Expansion:** generalizes hypothesis class of linear regression to include any function while applying similar techniques

2 Classification

The problem of identifying to which category an observation belongs, based on training data whose categories are known. Can also be thought of as partitioning the feature space into categories.

Algorithms:

- **k -Nearest Neighbors:** classifies data points based on the most common category from the k nearest training points
- **Decision Trees:** partitions feature space by creating a tree with decisions at each node and labels at the leafs
- **Logistic Regression:** fits a logistic function to the data to obtain a hyperplane that divides the feature space
- **Support Vector Machines:** finds hyperplane that divides the feature space with the largest margin to data points
- **Naive Bayes:** models probabilities of attributes to find probability of label using Bayes rule

3 Ensemble Learning

- **Bagging:**
- **Boosting:**

4 Feature Transformation

- **Feature selection:**
- **PCA:**

5 Unsupervised Learning

- **Clustering:** find compact representation of data by partitioning it into groups or clusters

6 Optimization

- **Gradient Descent:**
- **Stochastic Gradient Descent:**
- **Genetic Algorithms:**

7 Miscellaneous

- **“Bayesian Learning”:** usage of Bayes rule to analyze which hypothesis from a hypothesis class is the most likely; probabilities of hypotheses and their relationship with the hypothesis “length”; and the derivation of the Bayes Optimal Classifier
- **Bayes Nets:** graphs of conditional dependencies between random variables
- **Distance functions:**

8 To Do

- Distance functions
- Bagging
- Boosting
- Bayes Nets
- Feature selection
- PCA
- Gradient Descent
- Stochastic Gradient Descent

- Hill Climbing
- Simulated Annealing
- Genetic Algorithms
- Hidden Markov Models
- Markov Decision Process
- Reinforcement Learning
- Artificial Neural Networks
- Evaluation Metrics