Machine Learning

Rationale

The rapid growth of computer power and the needs for information technology have made Machine Learning an essential part of systems that must interpret data by classifying or clustering. This course gives a a thorough grounding in the methodologies, technologies, mathematics and algorithms currently needed by people who do research in machine learning.

Description

Machine learning is a branch of artificial intelligence, concerned with the construction and study of systems that can learn from data. Data may be numeric or symbolic and typically has the form of an N-tuple. The anthropomorphic term learning in the machine learning context means being able to predict some unobserved components of an N-tuple given some observed components of the N-tuple. This course provides a detailed explanation of many of the techniques used in machine learning and statistical pattern recognition.

Topic List

Topics may include but are not limited to:

- Bayesian Classification
  - Class conditional probabilities
  - Prior Probabilities
  - Gain Matrix
  - Maximizing Expected Gain
  - Minimax Classification
- Parametric Probability Models
- Non-Parametric Probability Models

- Making Decisions in Context
  - Conditional Independence
  - Hidden Markov Models
  - Forward Backward Algorithm
- Graphical Models
  - Semi-graphoids
  - Graphoids
  - Bayesian Nets
- Decision Trees
- Nearest Neighbor
- Linear Regression
- Logistic Regression
- Principal Component Analysis
- Neural Networks
  - The Perceptron Algorithm
  - The Back Propagation Algorithm
  - Deep Learning
- Linear Decision Rules
  - Fisher Linear Decision Rule
  - Support Vector Machines
  - Kernel Methods
- Ensemble Learning
- Evolutionary Learning
- Clustering

  - K-Means Clustering
  - Expectation Maximization
  - Linear Manifold Clustering
  - Gaussian Mixture Models
  - Clustering Evaluation Measures
- Experimental Protocols
  - Training Sets
  - Test Sets
  - Cross-Validation
  - Performance Characterization

Learning Goals

The student must be able to demonstrate a working knowledge of the theoretical foundations and software of machine learning represented by the topics of

- Bayesian Classification
- Non-parametric Probability Models
- Clustering
- Dimensionality Reduction
- Performance Characterization

Assessment

Written exams and course projects will be assigned to make sure students are capable of identifying suitable algorithms for making certain types of predictions, designing experimental protocols to evaluate the performance of those proposed algorithms, and implement experiments on the algorithms and evaluations. 40% Important machine learning knowledge to be assessed

by a final project includes but not limited to: Classification, Regression, Clustering, Dimensionality Reduction, and Performance Characterization. 60%