



Machine Learning (IS ZC464)

Session 8: Review

Session 1: Introduction

- Learning – experience from past
- Learning -- humans vs. machines
- Why machine learning?
- Machine learning and Artificial intelligence
- Intelligent systems
- Examples – vehicle recognition, fruit recognition and face recognition: attributes and their significance
- Traditional vs. Machine learning
- etc.

Session 2: Regression

- Training and testing in learning systems
- Prediction – training and testing
- (X,Y) – input output pair
- Understanding error
- Function approximation : given training (x,y) pairs
- Best hypothesis-minimum error tradeoff
- Learning parameters defining a hypothesis
- Line fitting – slope and intercept
- Squared mean error
- etc.

Session 3: Classification

- Decision boundaries and regions
- Binary classification
- Linear separability
- Discriminant functions
- Identifying the decision boundary on 2D plane
- Classifying a test vector using $h(x) = x_2 - mx_1 - c$
- etc.

Session 4: Classification Algorithms



- K-Nearest Neighbors
- Lazy learning
- Distance based weighted k-NN algorithm
- Bayes' Theorem
- Naïve Bayes' Classifier
- Problem solving- classifying a test pattern
- etc.

Session 5: Decision tree and data visualization



- Decision tree classifier
- Decision trees and its limitations
- Attributes and their significance in decision tree based classification
- Leaf and non leaf nodes of a decision tree
- Information content – entropy of the given set of observations
- Splitting the observations based on decision key and attribute selected for splitting
- Gain and remainder computations
- Data visualization
- Scatter plots
- Parallel coordinate graphs
- Pearson Correlation coefficient
- etc.

Session 6: Feature Engineering

- What is feature engineering
- Machine learning model with feature engineering
- Dimensionality reduction
- Feature Extraction techniques
- Bag of words
- TF-IDF, n-grams, stemming
- Image features, glimpse of transform based features
- Feature selection algorithms etc.

Session 7: Evaluation of Classification models



Overfitting

Bias and variance

Partitioning the data

Hold out, Cross validation and bootstrapping methods

Class-wise accuracy

true/false positives/negatives

Precision and recall

sensitivity analysis

ROC curves

Confusion matrix etc.

Mid Term Syllabus

- Lecture sessions 1- 8

Queries?
