FOUNDATIONS OF MACHINE LEARNING (AI60203)

Announcements

Course Logistics

Course Logistics

Teachers: Mahesh Mohan M R
References
• Introduction to Machine Learning, Ethem Alpaydin, MIT Press (available online)
• Machine Learning, Tom M Mitchell, Mc Graw Hill Education (available online)
Theory
• Monday – 11:00-11:55 AM, Tuesday – 8:00-9:55 AM
• Evaluation: Class Test-1, Midsem, Class Test-2, and Endsem
Mini-Project
• Open ended (will be announced)
Office Hours: Fridays 4-5 PM based on appointment

Course_logistics

Introducton and Motivation

Intro to ML

Motivations of ML. Why ML is trending now?
Classification of ML: Supervised, Unsupervised and Reinforcement Learning
Algorithm and its ineffectiveness for problems intuitive to humans (e.g., digit classification)
Basics of Supervised Learning: Labelled Data, Model, Loss, and Parameter Optimization

Introduction

Reading Exercise: Intro to ML by Ethem Alpaydin (first reference): Sec 1.1, Secs. 1.2.1 - 1.2.3.

K Nearest Neighbor

Instance based vs Model based Learning
Local vs Global Approximation of Target
K Nearest Neighbor for Classification and Regression
Effect of K (less K -- > sensitive to noise and large K --> sensitive to possibly irrelevant inputs)
Weighted KNN for Classification and Regression
Issues of KNN
Measurement Scales (Solution: min-max or z-score normalization)
Curse of Dimensionality
Expensive and Storage Need

KNN

KNN Study

Linear and Logistic Regression

Motivation of Linear Models
Prerequisites of Linear Algebra
Vector Inner Product (for linear models with a single scalar output)
Matrix-vector multiplication (for linear models with multidimensional outputs)
Inner Product perspective --> for a single op value
Outer product perspective --> for the entire op
Perceptron
History
Pre-inner product Interpretation
Post-inner product Interpretation
Optimization
Gradient descent
Batch Gradient Descent
Stochastic Gradient Descent
Minibatch Gradient Descent
Closed Form Solution
Prerequisites of Probability
Probability and Random Variables
Probability Distribution (Joint, Marginal, and Conditional)
Bayes' Theorem
Independence
Mean, Variance, and Covariance
Logistic Regression
Linear vs Logistic Regression

linear_reg

Logistic Regression

Naive Bayes

Naive Bayes' Model
Main assumption: Independence. Why?
Naive Bayes' Method for Categorical Inputs
Naive Bayes' Method for Continuous Inputs
Numerical Stability
Merits and Demerits

naive_bayes

Decision Tree

Decision Tree for Classification and Regression

Motivation for Decision Tree: Interpretability
Elements of Decision Tree: Root, Nodes, and Leaves
When to Split a Node?
Concept of Impurity -- Which feature to consider in a given node
Entropy Measure and other measures
Dealing with categorical as well as numerical features
Rule Extraction from Decision Trees

Regression Trees
Pruning to address overfitting

decision_tree

Bias Variance Tradeoff

Bias-variance Tradeoff
Graphical Illustration
Mathematical derivation
Addressing Bias-variance issues
Underfitting and Overfitting in various models
Training and Validation Sets
K-fold Cross validation (to eliminate the drawback of the above)

Bias Variance Tradeoff

Proof

SVMs

Motivation of SVMs
Hard-margin Hyperplane SVMs
Soft-margin Hyperplane SVMs
Kernel SVMs
Optimization Problems for different SVMs

SVM

Feed Forward NN

1. XOR problem: Solution with Multi-layer Neural Networks
2. Biological Motivation of Feedforward Neural Network
3. Machine Learning Vs Deep Learning
   - Case study: Face detection
   - Advantages of Deep Learning over Machine Learning
   - Advantages of Machine Learning over Deep Learning
4. Feed Forward Neural network
   - Overview
     - Terminologies (input, output, and hidden layers, and learnable parameters)
     - Impact of Bias (similar to the threshold in Perceptron)
     - Extracting low-level to high-level features, and Connectionism
   - Forward Pass
     - Matrix multiplication based mapping
     - The need for Non-linearities between layers
   - Backpropagation for optimizing weights
     - Vanishing/Exploding Gradient
     - Gradients for MSE, Linear Non-linear Activations, and Linear Layer
     - Reuse of component Gradients for different parameters

Feedforward NN

Convolutional and Recurrent NN

1. Biological Motivation

2. CNNs from MLPs
3. Different Layers of CNNs
4. Applications of CNNs for Classification and Regression
5. Standard CNN Architectures
   - AlexNet
   - VGGNet
   - GoogleNet
   - ResNet
6. Feedforward vs Recurrent Networks
7. Motivation for RNNs
8. Different RNN types and Applications
   - One-to-Many
   - Many-to-One
   - Many-to-Many
9. RNN model and parameters
10. Optimizing RNNs (Backpropagation through time)
11. Applications of RNNs

CNN

CNN Visualization

RNN

Class Test 1 Evaluation

CT1

Ensemble Learning

Motivation of Ensemble Learning
Insights on How Bagging and Boosting reduce Variance and Bias, respectively.
Random Forests
Adaboost

Ensemble

Dimensionality Reduction

Role of Dimensionality Reduction in Machine Learning
Feature Selection Vs Feature Extraction
Feature Selection
Forward Search
Backward Search
Feature Extraction
Principal Component Analysis (PCA)
Linear Discriminant Analysis (LDA)

DimRed

Unsupervised Learning: Clustering

Clustering -- Unsupervised Learning
Flat Clustering
K Nearest Neighbor --> Compactness

Spectral Clustering --> Connectivity
Mixture of Gaussian --> Soft Clustering
Hierarchical Clustering
Divisive -- > top-down
Agglomerative --> bottom-up

Clustering

Course Project

Course Project Slide

Course Project Google Collab

Results

Class Test 2 Evaluations

Class Test 2

Moodle Docs for this page

You are logged in as Dr. Mahesh Mohan MR (Log out)
Home

Data retention summary