



THE UNIVERSITY
of ADELAIDE

CRICOS PROVIDER 00123M

Introduction to Statistic Machine Learning Review

Lingqiao Liu
University of Adelaide

adelaide.edu.au

seek LIGHT

Overview of Machine Learning

- Types of machine learning systems
- Basic math skills
 - The same set of skills you will need to use in the exam

Classification, KNN, Overfitting

- What is the classification system?
 - Describe steps in building a classification system
- Nearest neighbour classifier
 - 1 nearest neighbour
 - K nearest neighbours
 - The effect of K
- Model selection problem
 - We introduce the model selection problem from the example of choosing k in KNN classifiers.
 - Concept of overfitting and generalization
 - Validation set
 - K-fold cross validation: special case, leave-one-out cross validation

Linear Classifier, Linear SVM

- Linear discriminant function
 - Know basic concepts, like separating hyperplanes
 - Linear and non-linear classifiers
- Basic idea of linear SVM
 - Concepts: support vectors, margin
- Hard margin and soft margin SVM
 - What's the difference?
 - Motivation of soft-margin SVM
- Primal and Dual problems of linear SVMs
 - Formulation, relationship between variables in primal problem and dual problem, meaning of each term (objective terms and constraints)
 - How to derive dual from primal problem

Regression

- What is regression problem?
- Linear Regression
 - Regression to scalar value and vector values.
 - Close-form solution
- Regularized linear regression
 - P-norm
 - L2 regularized linear regression, or ridge regression and L1 regularized linear regression or Lasso
 - Benefit of ridge regression, its close-form solution
 - Benefit of Lasso
- Support vector regression
 - motivation and intuitive idea
 - The primal problem and dual problem (optional)

Ensemble methods

- Basic concepts
 - Why ensemble methods, what is ensemble methods
 - General idea or workflow
- Bagging
 - Algorithm
- Random forest
 - Decision tree (optional)
 - How does random forest randomize decision trees to make a random forest
- Adaboost
 - Concepts: weak_learner, when does it work
 - Algorithms: the update of each components.

PCA and LDA

- Concept of dimensionality reduction
 - Benefit, why it is possible, applications
- PCA
 - Motivation and understanding of PCA
 - How PCA is derived, i.e., the relationship between PCA and covariance matrix
 - How to perform PCA
 - Eigen-face model: how to solve the issue of calculating eigen vectors for high-dimensional data
 - Roles of eigen-vectors: the face reconstruction experiment
- LDA
 - Motivation and intuitive idea of LDA (binary-class case)
 - Solution of LDA and multi-class case (optional)

Unsupervised learning

- K-means clustering
 - Steps, objective function
 - Advantages and disadvantages
- GMM model
 - Advantages over k-means
 - Interpret GMM from the viewpoint of clustering, e.g. class membership.
 - EM algorithm [optional]

Kernel Method

- Basic concepts
 - Benefit of using kernel
 - How to prove one function is a valid kernel function
 - Commonly used kernels
- Kernelize algorithms
 - Kernel SVM
 - How to kernelize algorithms: Euclidean distance,
 - Kernel k-means
 - Kernel PCA [optional]
 - Kernel regression: representing w by weighted combination of features

Neural Networks and Deep Learning

- Multi-layer perceptron
 - Structure and benefit
- Convolutional Neural networks
 - Structure and benefit
 - Convolution operator
 - Pooling operator
 - How many parameters, how many activations
- Optimization in deep learning: Stochastic Gradient descent (SGD)
 - Relationship between gradient descent and SGD, why should we use SGD
 - Concepts like learning rate and batch size.

Semi-supervised Learning

- Concepts and basic setting
- Pseudo-labelling
 - Assumption, Algorithm
 - Advantage and disadvantage
- Co-training
 - Basic idea
 - Advantage and disadvantage
- S3SVM and Graph-based Semi-supervised learning
 - Assumption
 - loss functions
- Deep semi-supervised learning (optional)

Generative model

- Autoregressive model
 - Theoretical foundation and key idea
 - How to generate (sample) from an auto-regressive model
 - How to train an auto-regressive model
- Generative Adversarial Networks (GAN)
 - Basic idea
 - Components in GAN and their roles
 - Loss function
 - Applications (optional)

Exam

- Open-book
 - Calculator - Standard, Drawing instruments or rulers, Lecture notes, Paper dictionary, Paper english dictionary, Paper translation dictionary are permitted .
- Length
 - 150 minutes
 - Extra 30 minutes uploading time for offshore students
- Difficulty level
 - 10-15 marks: need a good understanding of ML concepts and a little bit math skill
 - 20-25 marks: need a good understanding of ML concepts
 - The rest: should be easy