Pattern Recognition (2016 Spring at Nanjing University)

Chong Liu
chongliu@ucsb.edu
7595382

February 2019

1 Introduction

A graduate level course provided for master and PhD students at Nanjing University, Nanjing, China.

2 Instructor

Prof. Jianxin Wu, Ph.D. at Georgia Tech, 2009 (`https://cs.nju.edu.cn/wujx/`)

3 My Score

92 / 100

4 Syllabus

1. **Introduction to Pattern Recognition**
   An example: autonomous driving
   Pattern recognition and machine learning
   Structure of this lecture
2. **Mathematical background**
   Linear algebra
   Probability
   Optimization and matrix calculus
   Complexity of algorithms
3. **Overview of a pattern recognition system**
   Face recognition
   A simple nearest neighbor classifier
   The ugly details
   Making assumptions and simplifications
   A framework

4. **Evaluation**
   Accuracy and error in the simple case
   Minimizing the cost or loss
   Evaluation in imbalanced problems
   Can we reach 100% accuracy?
   Confidence in the evaluation results
5. **Principal component analysis**
   Motivation
   PCA to zero-dimensional subspace
   PCA to one-dimensional subspace
   PCA to more dimensions
   The complete PCA algorithm
   Variance analysis
   When to use or not to use PCA?
   The whitening transform
   Eigen-decomposition vs. SVD
6. **Fishers linear discriminant**
   FLD for binary classification
   FLD for more classes
7. **Distance metrics and data transformations**
   Distance metrics and similarity measures
   Data transformation and normalization
8. **Support vector machines**
   The key SVM idea
   Visualizing and calculating the margin
   Maximizing the margin
   The optimization and the solution
   Nonlinear and multiclass extensions
   Kernel SVM
9. **Probabilistic methods** The probabilistic way of thinking
   Choices
   Parametric estimation
   Nonparametric estimation
   Making decisions
10. **Information theory and decision tree** Prefix code and Huffman tree
    Basics of information theory
    Information theory for continuous distributions
    Information theory in ML and PR
    Decision trees

11. **Sparse and misaligned data**

    Sparse machine learning

    Dynamic time warping

12. **Hidden Markov model**

    Sequential data and the Markov property

    Three basic problems in HMM learning α, β, and the evaluation problem

    γ, δ, ψ, and the decoding problem

    ξ and learning HMM parameters

13. **Convolutional neural networks**

    Preliminaries

    CNN overview

    Layer input, output and notations

    The ReLU layer

    The convolution layer

    The pooling layer

    A case study: the VGG-16 net

    Hands-on CNN experiences