Pattern Recognition and Neural Networks - Video course

COURSE OUTLINE

Introduction to pattern recognition, introduction to classifier design and supervised learning from data, classification and regression, basics of Bayesian decision theory, Bayes and nearest neighbour classifiers, parametric and non-parametric estimation of density functions, linear discriminant functions, Perceptron, linear least-squares regression, LMS algorithm.

Fisher linear discriminant, introduction to statistical learning theory and empirical risk minimization, non-linear methods for classification and regression, artificial neural networks for pattern classification and regression, multilayer feedforward networks, backpropagation, RBF networks, Optimal separating hyperplanes, Supoort Vector Machines and some variants, Assessing generalization abilities of a classifier, Bias-variance trade-off, crossvalidation, bagging and boosting, AdaBoost algorithm, brief discussion of feature selection and dimensionality reduction methods.

The course is designed for graduate students (i.e. first year ME or research students). The course is intended to give the students a fairly comprehensive view of fundamentals of classification and regression. However, not all topics are covered.

For example, we do not discuss Decision tree classifiers. Also, the course deals with neural networks models only from the point of view of classification and regression. For example, no recurrent neural network models (e.g., Boltzman machine) are included. The main reason for leaving out some topics is to keep the course content suitable for a one semester course.

COURSE DETAIL

Module1 - Overview of Pattern classification and regression
Lecture 1 - Introduction to Statistical Pattern Recognition
Lecture 2 - Overview of Pattern Classifiers
Module2 - Bayesian decision making and Bayes Classifier
Lecture 3 - The Bayes Classifier for minimizing Risk
Lecture 4 - Estimating Bayes Error; Minimax and Neymann-

NPTEL

http://nptel.iitm.ac.in

Electronics & Communication Engineering

Pre-requisites:

1. Probability Theory Some knowledge of optimization methods.

Additional Reading:

1. C.M. Bishop, "Pattern Recognition and Machine Learning", Springer, 2006.

Coordinators:

Prof. P.S. Sastry
Department of Electrical EngineeringIISc Bangalore

Pearson classifiers

Module3 - Parametric Estimation of Densities

Lecture 5 - Implementing Bayes Classifier; Estimation of Class Conditional Densities
Lecture 6 - Maximum Likelihood estimation of different densities
Lecture 7 - Bayesian estimation of parameters of density functions, MAP estimates
Lecture 8 - Bayesian Estimation examples; the exponential family of densities and ML estimates
Lecture 9 - Sufficient Statistics; Recursive formulation of ML and Bayesian estimates

Module4 - Mixture Densities and EM Algorithm

Lecture 10 - Mixture Densities, ML estimation and EM algorithm
Lecture 11 - Convergence of EM algorithm; overview of Nonparametric density estimation

Module5 - Nonparametric density estimation

Lecture 11 - Convergence of EM algorithm; overview of Nonparametric density estimation
Lecture 12 - Nonparametric estimation, Parzen Windows, nearest neighbour methods

Module6 - Linear models for classification and regression

Lecture 13 - Linear Discriminant Functions; Perceptron -- Learning Algorithm and convergence proof
Lecture 14 - Linear Least Squares Regression; LMS algorithm
Lecture 15 - AdaLinE and LMS algorithm; General nonliner least-squares regression
Lecture 16 - Logistic Regression; Statistics of least squares method; Regularized Least Squares
Lecture 17 - Fisher Linear Discriminant
Lecture 18 - Linear Discriminant functions for multi-class case; multi-class logistic regression

Module7 - Overview of statistical learning theory, Empirical Risk Minimization and VC-Dimension

Lecture 19 - Learning and Generalization; PAC learning framework
Lecture 20 - Overview of Statistical Learning Theory; Empirical Risk Minimization
Lecture 21 - Consistency of Empirical Risk Minimization
Lecture 22 - Consistency of Empirical Risk Minimization; VC-Dimension
Lecture 23 - Complexity of Learning problems and VC-Dimension
Lecture 24 - VC-Dimension Examples; VC-Dimension of hyperplanes

Module8 - Artificial Neural Networks for Classification and regression

Lecture 25 - Overview of Artificial Neural Networks
Lecture 26 - Multilayer Feedforward Neural networks with Sigmoidal activation functions;
Lecture 27 - Backpropagation Algorithm; Representational abilities of feedforward networks
Lecture 28 - Feedforward networks for Classification and

Regression; Backpropagation in Practice
Lecture 29 - Radial Basis Function Networks; Gaussian RBF networks
Lecture 30 - Learning Weights in RBF networks; K-means clustering algorithm

Module9 - Support Vector Machines and Kernel based methods

Lecture 31 - Support Vector Machines -- Introduction, obtaining the optimal hyperplane
Lecture 32 - SVM formulation with slack variables; nonlinear SVM classifiers
Lecture 33 - Kernel Functions for nonlinear SVMs; Mercer and positive definite Kernels
Lecture 34 - Support Vector Regression and ε-insensitive Loss function, examples of SVM learning
Lecture 35 - Overview of SMO and other algorithms for SVM; ν-SVM and ν-SVR; SVM as a risk minimizer
Lecture 36 - Positive Definite Kernels; RKHS; Representer Theorem

Module10 - Feature Selection, Model assessment and cross-validation

Lecture 37 - Feature Selection and Dimensionality Reduction; Principal Component Analysis
Lecture 38 - No Free Lunch Theorem; Model selection and model estimation; Bias-variance trade-off
Lecture 39 - Assessing Learnt classifiers; Cross Validation;

Module11 - Boosting and Classifier ensembles

Lecture 40 - Bootstrap, Bagging and Boosting; Classifier Ensembles; AdaBoost
Lecture 41 - Risk minimization view of AdaBoost

References:

1. R.O.Duda, P.E.Hart and D.G.Stork, Pattern Classification, John Wiley, 2002.

2. C.M.Bishop, Neural Networks and Pattern Recognition, Oxford University Press (Indian Edition), 2003.