MA848 Advanced Statistical Methods for Data Analytics 4 0 0 4

Unit I: Data collection, sampling and estimation

Data collection – types of data - Sampling methods – data generation methods- Bootstrap sampling – Jackknife sampling – bias and variance- simulation – confidence levels - sample size determination – descriptive statistics- sampling and sampling distributions- maximum likelihood estimation- Bayesian estimation.

Unit II: Fitting of data and inferential statistics

Hypothesis testing -Fitting of distribution to data – Binomial – Poisson – uniform – exponential –Normal distributions — one- way, two-way- analysis of variance – Multiple range test for one way ANOVA for grouping of means – Bayesian belief networks.

Unit III: Dimensionality Reduction Methods and supervised learning methods

Multivariate statistics – multivariate normal distribution – multivariate regression analysis – testing the significance of regression coefficients - Principal component analysis – Linear discriminant analysis- Fisher's discriminant analysis – Statistical decision making – Bayesian classification.

Unit IV: Markov model and Hidden Markov model

Markov process – Markov chains – transition probabilities- steady state probabilities- classification of states – hidden Markov models – computation, evaluation and decoding in hidden Markov models.

Unit V: Clustering and regression based classification

Hierarchical clustering – agglomerative clustering algorithm – single linkage algorithm – complete linkage algorithm – average linkage algorithm - Partitional clustering – Forgy's algorithm – k-means algorithms. regression and classification Trees-decision trees – CART methods - histograms – nearest neighbor classification techniques –statistical distribution-based grouping.

Text Book/References:

1. Richard O. Duda, Peter E. Hart and David G. Stork, "Pattern Classification", Second Edition, 2003, John wily & sons.
2. Earl Gose, Richard Johnsonbaugh and Steve Jost, "Pattern Recognition and Image Analysis", 2002, Prentice Hall of India.
3. Ravichandran . J. "Probability and Statistics for Engineers", First edition, Wiley, 2012.
4. Hastie, T., Tibshirani . R., and Friedman, J. The elements of statistical learning. Vol. 2. No. 1. New York: Springer, 2009.