



Jansons Institute of Technology

Karumathampatti, Coimbatore - 641 659

Approved by AICTE and Affiliated to Anna University

An ISO 9001:2015 certified institution

Industry Readiness Course - Syllabus

Department of Computer Science and Engineering

Genomic Data Mining

2023 - 2024 (EVEN Semester)

Learning Objectives

- Basics for statistical learning, with an emphasis on its application to genomic data.
- Cover methods on classification, resampling methods, linear models with regularization (e.g. LASSO), additive models, classification and regression trees, random forests, support vector machines and basics of unsupervised learning.
- An emphasis will be on its applications to cutting edge genetics and genomics problems.
- Areas of genomic applications will include (but not limited to) variant annotation, genetic association analysis, variant calling and filtering from next generation sequencing.

Learning Outcomes

- Sequel of statistical learning course for high dimensional statistical learning.
- Cover a broad class of methods in statistical learning and discuss their applications in cutting edge genomics research.
- Introduced to current research topics in genomics where data mining methods play a critical role.

MODULE 1

Introduction to statistical Learning - Overview of Concurrent Genomic Research - Linear Models - Classification - Linear Discriminant Analysis (LDA) - Applications of LDA in Genomics

MODULE 2


Resampling Methods I - Validation Method - Resampling Methods II - Bootstrap - Linear Model Selection I - Shrinkage Method - Linear Model Selection II - Dimension Reduction and High Dimensional Inference - Variable Selection and Prediction in Genomics - Nonlinear Models I - Spline Regression - Nonlinear Models II - Generalized Additive Models - Nonlinear Models in Genomics



MODULE 3

Tree-based Methods – Decision Trees - Overview of Tree Based Methods in Genomics - Support Vector Machines I - Exemplar Application of SVM in Genomics – Variant Filtering from Next Generation Sequencing - Unsupervised Learning I – Principal Component Analysis - Unsupervised Learning II – Clustering Method

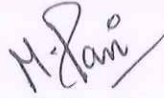
Course Designed By



Dr A Velayudham

Professor / CSE

Jansons Institute of Technology



Dr.M Pavithra

AP / CSE

Jansons Institute of Technology

Approved By

Mr.S.M.Srihari shankar
AP/ CSE



Industry Readiness Coordinator



Approved By

Dr A Velayudham,
Prof & Head /CSE



Head of Department

Approved By



Principal

PRINCIPAL
JANSONS INSTITUTE OF TECHNOLOGY
(Autonomous)
KARUMATHAMPATTI, COIMBATORE - 641659