Chemometrics & Machine Learning for Metabolomics — Hands-on

Take your metabolomics statistics beyond basic PCA and PLS-DA. This module focuses on chemometric thinking and practical machine learning workflows for metabolomics and lipidomics: feature engineering, supervised and non linear models, robust validation and interpretation so that you can build defensible classifiers and prediction models for real research questions.

Chemometrics & Machine Learning for Metabolomics

Help Desk · WhatsApp

Session Index

Session 1

Fee: Rs 8800 Apply Now

Chemometric Thinking & Feature Engineering

Chemometric view of metabolomics data matrices

samples vs variables perspective **blocks, batches and design structure** **unsupervised vs supervised goals**

Pre processing and scaling choices for ML workflows

log, Pareto, unit variance concepts | centered vs uncentered data | impact of scaling on models

Feature engineering and basic feature filtering strategies

variance and missingness filters | biologically motivated groupings | simple ratios and indices (concepts)

Session 2

Fee: Rs 11800 Apply Now

Supervised Models & Performance Assessment

Supervised learning tasks in metabolomics (concepts)

classification vs regression examples | multiclass and ordinal setups | balanced vs imbalanced designs

PCA, PLS, PLS DA and regularised linear models (high level)

chemometric roots of PLS | Lasso, Ridge and Elastic Net ideas | when simple models are enough

Cross validation, test sets and basic performance metrics

k fold and repeated CV concepts | ROC AUC, accuracy, RMSE | avoiding optimistic bias

Session 3

Fee: Rs 14800 Apply Now

Non Linear Models, Robustness & Interpretation

Tree based and other non linear model concepts

Random Forest and gradient boosting ideas | support vector machines (high level) | pros and cons for metabolomics

Model robustness, leakage checks and repeated CV thinking

feature selection within CV concept | permutation style sanity checks | stability of feature importance

Interpreting models and presenting results to biologists

variable importance ranking ideas | partial dependence concept | link to pathways and hypotheses

Session 4

Fee: Rs 18800 Apply Now

Mini Capstone: Metabolomics ML Workflow & Report

Designing an end to end ML workflow for a toy dataset

Theory + Practical

Comparing a simple linear model with a non linear alternative

performance vs interpretability | stability across CV runs | practical recommendation for users

Deliverables: ML workflow summary & result bundle

workflow outline (PDF/HTML) | metrics and feature importance table (CSV/TSV) | short interpretation note for collaborators