AIPE3006 PREDICTIVE ANALYSIS AND STATISTICAL MODELLING (3-0-0)

COURSE OBJECTIVES:

- To understand the different approaches to formulate statistical prediction and inference problems.
- To build a solid foundation for the statistical theory for predictive modelling and inference.
- To understand the pros and cons of statistical models for predictive modeling.
- To perform data analysis (including exploratory data analysis and visualization) with R and to explain the relevance of the statistical methods chosen.
- To communicate statistical results without use of statistical jargon.

MODULE - I (08 HOURS)

INTRODUCTION TO PREDICTIVE ANALYSIS & LINEAR REGRESSION

What and Why Analytics, Introduction to Tools and Environment, Application of Modeling in Business, Databases & Types of data and variables, Data Modeling Techniques, Missing imputations etc. Need for Business Modeling, Regression — Concepts, Blue property-assumptions-Least Square Estimation, Variable Rationalization, and Model Building etc.

MODULE -II (08 HOURS)

LOGISTIC REGRESSION

Model Theory, Model fit Statistics, Model Conclusion, Analytics applications to various Business Domains etc. Regression Vs Segmentation — Supervised and Unsupervised Learning, Tree Building — Regression, Classification, over fitting, Pruning and complexity, Multiple Decision Trees etc.

MODULE -III (08 HOURS)

OBJECTIVE SEGMENTATION

Regression Vs Segmentation — Supervised and Unsupervised Learning, Tree Building — Regression, Classification, over-fitting, Pruning and complexity, Multiple Decision Trees etc. Develop Knowledge, Skill and Competences, Introduction to Knowledge skills & competences, Training & Development, Learning & Development, Policies and Record keeping. etc.

MODULE - IV (08 HOURS)

TIME SERIES ANALYSIS AND FORECASTING

Time series decomposition, smoothing techniques, Feature Extraction: ARIMA models, Measures of Forecast Accuracy, STL approach, Extract features from generated model as Height. Average, Energy etc., and analyze for prediction

MODULE - V (08 HOURS)

STATISTICAL METHODS IN THE FREQUENCY DOMAIN

Introduction, Spectral Matrices and Likelihood Functions, Regression for Jointly Stationary Series, Regression with Deterministic Inputs, Random Coefficient Regression, Analysis of Designed Experiments, Discrimination and Cluster Analysis, Principal Components and Factor Analysis.

COURSE OUTCOMES:

After the completion of the course, students will be able to:

- Understand the role of predictive analysis in decision-making.
- Apply linear and logistic regression techniques to build and evaluate predictive models for business problems.
- Analyze customer or business data using segmentation techniques to derive objective insights and support strategic decisions.
- Develop time series forecasting models using techniques like ARIMA, STL, and smoothing to predict future trends and measure forecast accuracy.
- Evaluate complex datasets using frequency domain statistical methods, including spectral analysis, PCA, and cluster analysis for deeper business insights.

TEXTBOOKS:

1. Time Series Analysis and Its Applications: With R Examples by Robert H. Shumway & David S. Stoffer.
2. Business Analytics: Data Analysis & Decision Making by S. Christian Albright & Wayne L. Winston.

REFERENCE BOOKS:

1. An Introduction to Statistical Learning: with Applications in R (Second Edition) by Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani.
2. Elements of Statistical Learning (12th printing) by Trevor Hastie, Robert Tibshirani and Jerome Friedman.

WEBLINKS AND VIDEO LECTURES (E-RESOURCES):

1. https://onlinecourses.swayam2.ac.in/imb23_mg42/preview
2. https://www.coursera.org/learn/predictive-modeling-analytics