Name of the Programme: M.Sc. in Artificial Intelligence
Course Code: CSI-528
Title of the Course: Regression Analytics and Predictive Models
Number of Credits: 4 (2L+2T+0P)
Effective from AY: 2023-24

Prerequisites for the course	Probability Theory and Distributions	
Objectives	● Develop an understanding of regression analysis and model building. ● Provide the ability to develop relationship between variables ● Investigate possible diagnostics in regression techniques ● Formulate feasible solutions using a regression model for real-life problems.	
Theory:	**Simple Regression Analysis** Introduction to a linear and nonlinear model. Ordinary Least Square methods. Simple linear regression model, using simple regression to describe a linear relationship. Fitting a linear trend to time series data, Validating simple regression model using t, F and p test. Developing confidence interval. Precautions in interpreting regression results.	4 hours
	Multiple Regression Analysis Concept of Multiple regression model to describe a linear relationship, Assessing the fit of the regression line, inferences from multiple regression analysis, problem of overfitting of a model, comparing two regression model, prediction with multiple regression equation.	4hours
	Fitting Curves and Model Adequacy Checking Introduction, fitting curvilinear relationship, residual analysis, PRESS statistics, detection and treatment of outliers, lack of fit of the regression model, test of lack of fit, Problem of autocorrelation and heteroscedasticity. Estimation of pure errors from near neighbors.	4 hours
	Transformation techniques Introduction, variance stabilizing transformations, transformations to linearize the model, Box Cox methods, transformations on the repressors variables, Generalized and weighted least squares, Some practical applications.	4 hours
	Multicollinearity Introduction, sources of multicollinearity, effects of multicollinearity. Multicollinearity diagnostics: examination of correlation matrix, variance Inflation factors (VIF), Eigen system analysis of X1X. Methods of dealing with Multicollinearity: collecting additional data, model re-specification, and ridge regression.	4 hours
	Generalized Linear Models Generalized linear model: link functions and linear predictors, parameter estimation and inference in the GLM, prediction and estimation with the GLM, Residual Analysis, and concept of over dispersion.	4 hours
	Model building and Nonlinear Regression Variable selection, model building, model misspecification. Model validation techniques: Analysis of model coefficients, and predicted values, data splitting method. Nonlinear regression model, nonlinear least squares, transformation to linear model, parameter estimation in nonlinear system, statistical inference in nonlinear regression.	6 hours

Assignments and a Mini Project to be discussed during the Tutorial slots:	1. Linear Regression 2. Minimum Least Square Method 3. Calculating coefficients values 4. Ascombe's Quartet 5. Regression Equations- x on y & y on x 6. Predicting mom's height based on daughter's height 7. Regression-Solved problem-2 8. Probable Error- Calculating correlation coefficient of POPULATION 9. Predictive modelling project for credit card fraud detection 10. Predictive modelling project for customer value prediction 11. Predictive modelling project for stock market forecasting 12. Predictive modelling project for corporate bankruptcy prediction	12 * 2 = 24 hours + 6 hours for Mini Project discussions
Pedagogy	Lectures/ tutorials/assignments/self-study	
References/ Readings	1. Douglas C. Montgomery, Elizabeth A. Peck, G. Geoffrey Vining, Introduction to Linear Regression Analysis, Third Ed., Wiley India Pvt. Ltd., 2016. Norman R. 2. Draper, Harry Smith; Applied Regression Analysis, WILEY India Pvt. Ltd. New Delhi; Third Edition, 2015. 3. Johnson, R A., Wichern, D. W., Applied Multivariate Statistical Analysis, Sixth Ed., PHI learning Pvt., Ltd., 2013. 4. Iain Pardoe, Applied Regression Modeling, John Wiley and Sons, Inc, 2012.	
Course Outcomes	1.Develop in-depth understanding of the linear and nonlinear regression model. 2.Demonstrate the knowledge of regression modeling and model selection techniques. 3.Examine the relationships between dependent and independent variables. 4.Estimate the parameters and fit a model.	

(Back to top)