

Semester	TWO
Paper Number	8
Paper Code	MDTS 4214
Paper Title	Predictive Analytics
No. of Credits	6
Course Description	CORE Composite PaperOne Module No. of classes assigned Theory: 4 classes per week Practical: 3 classes per week
Course Objective	At the end of the course, the students should be able to <ul style="list-style-type: none"> ○ Develop the concept of regression and classification ○ Understand the different model selection methods ○ Apply different dimension reduction techniques to real life data ○ Analyse current data and make future predictions
Syllabus	<p><i>UNIT 1: Introduction</i></p> <p>Diagnostic versus prognostic models. Regression versus classification problems. Bias-variance trade off. [4]</p> <p><i>UNIT 2: Linear Regression</i></p> <p>Least square method, simple linear regression, multiple linear regression with quantitative & qualitative predictors. Dummy variables, regression diagnostics (Outlier detection, leverage, Influential point, Cook's distance, Model selection via AIC and BIC, adjusted R-Square). K-nearest neighbour regression. [10]</p> <p><i>UNIT 3: Classification</i></p> <p>Logistic regression, multiple logistic regression, multi-category logistic regression (model, parameter estimation and prediction). Multiclass discriminant analysis. Decision trees (CART and CHAID)</p>

	<p>[15]</p> <p>UNIT 4: Model selection and Regularization</p> <p>Subset selection method (forward and backward stepwise selection), Shrinkage methods: Penalized likelihood and Bayesian linear regression; Ridge regression, the LASSO and Elastic NET. Applications of dimension reduction techniques. [15]</p> <p>UNIT 5: Generalized linear model</p> <p>Components of GLM, link functions (logit, probit, log link). Fitting of GLM (parameter estimation and prediction). Contingency tables, odds ratio and log linear models. Generalized linear mixed models (Inference for model parameters and predictions). [8]</p>
Practical	Based on theory topics
Reading/Reference Lists	<ol style="list-style-type: none"> 1. James, Witten, Hastie and Tibshirani: <i>An Introduction to Statistical Learning</i>. Second edition, Springer. 2. Hastie, Tibshirani, Friedman: <i>The Elements of Statistical Learning, Data Mining, Inference and Prediction</i>. Second Edition, Springer Series in Statistics. 3. McCullagh, P & Nelder, J.A.(1995), <i>Generalized Linear Models</i>. Chapman and Hall. 4. Agresti, A. (2007): <i>An Introduction to Categorical data analysis</i>. Wiley

Evaluation	Theory CIA: 10 End Sem Exam: 50 Total : 60	Practical Continuous Assessment: 30 End Sem Viva: 10 Total: 40
Paper Structure for End Semester Theory	Short questions: 5 marks each	Long questions: 10 marks each
	2 out of 4	4 out of 6