NATIONAL INSTITUTE OF TECHNOLOGY, TIRUCHIRAPPALLI

DEPARTMENT OF <u>COMPUTER APPLICATIONS</u>

COURSE PLAN – PART I			
Name of the programme and specialization	**Bachelor of Technology**		
Course Title	**Data Analytics**		
Course Code	**CAMI16**	No. of Credits	**3**
Course Code of Pre-requisite subject(s)	**-**		
Session	**July / ~~January~~ 2021**	Section (if, applicable)	
Name of Faculty	**Dr. Jitendra Kumar**	Department	**Computer Applications**
Email	**jitendra@nitt.edu**	Telephone No.	**0431-2503734**
Name of PAC Chairman	**Dr. P. J. A. Alphonse**		
E-mail	**alphonse@nitt.edu**	Telephone No.	**0431-2503730**
Course Type	**Open Elective Course**		

Syllabus (approved in BoS)
Introduction: Data Analytics – Data Collection – Integration – Management – Modelling – Analysis – Visualization – Prediction and Informed decision making. General Linear Regression Model, Estimation for BETA, Error Estimation, Residual Analysis.   Test of significance – ANOVA, 't' test, Forward, Backward, Sequential, Stepwise, and all possible subsets, Dummy Regression, Logistic Regression, and Multi-collinearity.   Discriminant Analysis –Two group problem, Variable contribution, Violation of assumptions, Discrete and Logistic Discrimination, The k-group problem, multiple groups, Interpretation of Multiple group Discriminant Analysis solutions.   Principal Component Analysis – Extracting Principal Components, Graphing of Principal Components, some sampling Distribution results, Component scores, large sample Inferences, Monitoring Quality with principal Components.   Factor Analysis – Orthogonal Factor Model, Communalities, Factor Solutions, and rotation.   Machine Learning - Supervised learning (rules, trees, forests, nearest neighbor, regression) – optimization (gradient descent and variants) – unsupervised learning   REFERENCES:

1. Richard A. Johnson and Dean W. Wichern, "Applied Multivariate Statistical Analysis", fifth edition, Pearson Education, 2002.
2. Hastie, Trevour, et al. "The element of statistical learning", Springer, 2009.
3. Montgomery, Douglas C., and George C. Runger, "Applied Statistics and probability for engineers", John Wiley & sons, 2010

COURSE OBJECTIVE(S)

To understand the data analytics approaches and apply them for decision making problems.

COURSE OUTCOMES (CO)

Course Outcomes	Aligned Programme Outcomes (PO)
Students will be able to: 1. Understand the basics of statistical modeling for data analytics	**PO I, II, III**
2. Understand the use of machine learning in data analytics	**PO I, II, IV**
3. Solve real-world decision-making problems using various data analytics techniques	**PO IV, V**

COURSE PLAN – PART II

COURSE OVERVIEW

This course introduces concepts of data analytics such as regression analysis, discriminant analysis, principal component analysis, factor analysis, etc. It also sheds light on the importance of statistical testing in the domain of data analytics. The course also develops a foundation of machine learning algorithms and their usage in different application domains in the analytics.

COURSE TEACHING AND LEARNING ACTIVITIES

S. No.	Week/ Contact Hours	Topic	Mode of Delivery
1	Week 1 (3 Classes)	Introduction to Data Analytics, Data	PPT, Pen and Paper (Online)
2	Week 2 (3 Classes)	Basic Statistical Description of Data, General Linear Regression Model, Parameter Estimation, Residual Analysis	PPT, Pen and Paper (Online)
3	Week 3 (3 Classes)	Test of Significance, ANOVA, 't' test	PPT, Pen and Paper (Online)

4	Week 4 (3 Classes)	Linear Regression Analysis and Variants, Logistic Regression and Multi-Colinearity	PPT, Pen and Paper (Online)
5	Week 5 (3 Classes)	Introduction to MATLAB	PPT, Pen and Paper (Online)
6	Week 6 (3 Classes)	Discriminant Analysis – Two Group Problem, Variable Contribution, Violation of Assumptions	PPT, Pen and Paper (Online)
7	Week 7 (3 Classes)	Discrete and Logistic Discrimination, K-group Problem, Multiple Groups	PPT, Pen and Paper (Online)
8	Week 8 (3 Classes)	Principal Component Analysis – Extracting Principal Components, Graphing of Principal Components	PPT, Pen and Paper (Online)
9	Week 9 (3 Classes)	Sampling Distribution Results, Component Scores, Large Sample Inferences, Monitoring Quality with Principal Components	PPT, Pen and Paper (Online)
10	Week 10 (3 Classes)	Factor Analysis	PPT, Pen and Paper (Online)
11	Week 11 (3 Classes)	Supervised Learning-I (Rules, Nearest Neighbors)	PPT, Pen and Paper (Online)
12	Week 12 (3 Classes)	Supervised Learning-II (Trees, Forests, Regression)	PPT, Pen and Paper (Online)
13	Week 13 (3 Classes)	Optimization: Gradient Descent and Variants	PPT, Pen and Paper (Online)
14	Week 14 (3 Classes)	Unsupervised Learning	PPT, Pen and Paper (Online)

COURSE ASSESSMENT METHODS

S. No.	Mode of Assessment	Week/Date	Duration	% Weightage
1	First Assessment	6th week	60 Minutes	20
2	Second Assessment	10th week	60 Minutes	20
3	Assignment	9th week	-	30
*	Compensation Assessment	-	60 Minutes	20
4	Final (End Term) Assessment	As per academic schedule	120 Minutes	30
Total Marks				100

*mandatory

COURSE EXIT SURVEY (mention the ways in which the feedback about the course shall be assessed)

The students through the class representattive may give their feedback at any time to the course coordinator which will be duly adddressed.
The students may give their feedback during class committee meetings.

COURSE POLICY (including compensation assessment to be specified)

MODE OF CORRESPONDENCE
By Email: jitendra@nitt.edu

COMPENSATION ASSESSMENT POLICY
One compensation assessment will be conducted for absentees in assessments (other than final assessment) only after the submission of medical or On-Duty certificates signed by competent authority.

ATTENDANCE POLICY (A uniform attendance policy as specified below shall be followed)

➢ **At least 75% attendance in each course is mandatory.**

➢ **A maximum of 10% shall be allowed under On Duty (OD) category.**

➢ Students with **less than 65% of attendance** shall be prevented from writing the

final assessment and **shall be awarded 'V' grade.**

ACADEMIC DISHONESTY & PLAGIARISM

➢ Possessing a mobile phone, carrying bits of paper, talking to other students, copying from others during an assessment will be treated as punishable dishonesty.

➢ Zero mark to be awarded for the offenders. For copying from another student, both students get the same penalty of zero mark.

➢ The departmental disciplinary committee including the course faculty member, PAC chairperson and the HoD, as members shall verify the facts of the malpractice and award the punishment if the student is found guilty. The report shall be submitted to the Academic office.
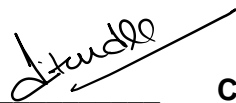
The above policy against academic dishonesty shall be applicable for all the programmes.

ADDITIONAL INFORMATION

NIL

FOR APPROVAL

Course Faculty _____   **CC-Chairperson** _____   **HOD** _____