FACULTY OF ENGINEERING AND TECHNOLOGY

Effective from Academic Batch: 2022-23

Programme: **M.TECH. ARTIFICIAL INTELLIGENCE**

Semester: **II**

Course Code: **2023102055**

Course Title: **Advanced Data Science**

Course Group: Program Elective-III

Course Objectives:
To familiarize the scope, process and advantages of business analytics, to introduce the forecasting models and techniques used in analytics, to expose the formulation and decision strategies used in business analytics

Teaching & Examination Scheme:

Contact hours per week			Course Credits	Examination Marks (Maximum / Passing)				
				Theory		J/V/P*		Total
Lecture	Tutorial 1	Practical 1		Internal	External	Internal	External	
3	0	2	4	50/20	50/20	25/10	25/10	150/60

* **J**: Jury; **V**: Viva; **P**: Practical

Detailed Syllabus:

Sr.	Contents	Hours
1	Introduction to Business analytics: The science of Data-Driven Decision Making, Descriptive-predictive-prescriptive analytics, Bigdata Analytics, Web Analytics, Social media Analytics. Framework for data driven-decision making, Challenges in data driven-decision making and Future, Business Analytics in Practice	4
2	Hypothesis Testing: Introduction to Hypothesis Testing, Setting Up a Hypothesis Test, One-Tailed and Two-tailed Test, Type I Error, Type II Error, and Power of The Hypothesis Test, Hypothesis Testing for Population mean with Known Variance: Z-Test, Hypothesis Testing for Population Proportion: Z-Test for Proportion, Hypothesis Test for Population mean under Unknown Population Variance, Paired Sample t-Test, Hypothesis Test for Difference in Population Proportion under Large Samples: Two-Sample Z-Test for Proportions, Effect Size: Cohen's D, Hypothesis Test for Equality of Population Variances, Non-Parametric Tests: Chi-Square Tests	6

Opp. Shastri Maidan, Beside BVM College, Vallabh Vidyanagar, Dist: Anand, Gujarat - 388120
(O): 02692-238001 | Email: adminoffice@cvmu.edu.in | www.cvmu.edu.in

3	Analysis of Variance: Introduction to Analysis of Variance (ANOVA), Multiple t-Tests for Comparing Several Means, One-way Analysis of Variance (ANOVA), Two-Way Analysis of Variance (ANOVA)	**4**
4	Multiple Linear Regression: Introduction, Ordinary Least Squares Estimation for Multiple Linear Regression, Multiple Linear Regression Model Building, Part (Semi-Partial) Correlation and Regression Model Building, Interpretation of MLR Coefficients -Partial Regression Coefficient, Standardized Regression Co-efficient, Regression Models with Qualitative Variables, Validation of Multiple Regression Model, Co-efficient of Multiple Determination (R-Square) and Adjusted R-Square, Statistical Significance of Individual Variables in MLR – t-Test, Validation of Overall Regression Model: F-Test, Validation of Portions of a MLR Model – Partial F-Test, Residual Analysis in Multiple Linear Regression, Multi-Collinearity and Variance Inflation Factor, Auto-correlation, Variable Selection in Regression Model Building, Avoiding Overfitting: Mallows's Cp	**7**
5	Logistic Regression: Introduction – Classification Problems, Introduction to Binary Logistic Regression, Estimation of Parameters in Logistic Regression, Interpretation of Logistic Regression Parameters, Logistic Regression Model Diagnostics, Classification Table, Sensitivity, and Specificity, Optimal Cut-Off Probability, Variable Selection in Logistic Regression, Application of Logistic Regression in Credit Rating, Gain Chart and Lift Chart	**7**
6	Forecasting Techniques: Introduction to Forecasting, Time-Series Data and Components of Time-Series Data, Forecasting Techniques and Forecasting Accuracy, Moving Average Method, Single Exponential Smoothing (ES), Double Exponential Smoothing – Holt's Method, Triple Exponential Smoothing (Holt-Winter Model), Regression Model for Forecasting, Auto-Regressive (AR),Moving Average (MA) and ARMA Models, Auto-Regressive Integrated Moving Average (ARIMA) Process, Power of Forecasting Model: Theil's Coefficient	**7**
7	Clustering: Introduction to Clustering, Distance and Dissimilarity Measures used in Clustering, Quality and Optimal Number of Clusters, Clustering Algorithms, K-Means Clustering, Hierarchical Clustering	**5**

List of Practical's / Tutorials:

1	Tutorials on Hypothesis testing
2	Tutorials on Analysis of Variance
3	Study and perform Multilinear regression on given datasets
4	Perform Logistic Regression on given datasets
5	Apply Clustering techniques for datasets
6	Apply Clustering techniques for datasets
7	Apply forecasting techniques for various time series datasets
8	Apply forecasting techniques for various time series datasets

Opp. Shastri Maidan, Beside BVM College, Vallabh Vidyanagar, Dist: Anand, Gujarat - 388120
(O): 02692-238001 | Email: adminoffice@cvmu.edu.in | www.cvmu.edu.in

Reference Books:

1	Business analytics: The science of Data Driven Decision Making by u Dinesh Kumar, Willey
2	Essential of Business Analytics by Jeffrey D. Camm, James J. Cochran, Michael J. Fry, Jeffrey W. Ohlman, David R. Anderson
3	Business analytics Principles, Concepts, and Applications by Marc J. Schniederjans, Dara G. Schniederjans, Christopher M. Starkey, Pearson FT Press

Supplementary learning Material:	
1	Open source Tools (R Studio, Candela, Python, Qlik, etc..)
2	NPTEL Course - Data Science for Engineers https://nptel.ac.in/courses/106/106/106106179/
3	NPTEL Course - Python for Data Science https://nptel.ac.in/courses/106/106/106106212/
4	towardsdatascience.com

Pedagogy:

- Direct classroom teaching
- Audio Visual presentations/demonstrations
- Assignments/Quiz
- Continuous assessment
- Interactive methods
- Seminar/Poster Presentation
- Industrial/ Field visits
- Course Projects

Internal Evaluation:

The internal evaluation comprised of written exam (40% weightage) along with combination of various components such as Certification courses, Assignments, Mini Project, Simulation, Model making, Case study, Group activity, Seminar, Poster Presentation, Unit test, Quiz, Class Participation, Attendance, Achievements etc. where individual component weightage should not exceed 20%.

Suggested Specification table with Marks (Theory) (Revised Bloom's Taxonomy):

Distribution of Theory Marks in %						**R**: Remembering; **U**: Understanding; **A**: Applying; **N**: Analyzing; **E**: Evaluating; **C**: Creating
R	**U**	**A**	**N**	**E**	**C**	
15%	25%	20%	30%	10%	--	

Note: This specification table shall be treated as a general guideline for students and teachers. The actual distribution of marks in the question paper may vary slightly from above table.

Course Outcomes (CO):

Sr.	Course Outcome Statements	%weightage
CO-1	Define the scope, process and advantages of business analytics	**30**
CO-2	Understand and apply the Concepts of Descriptive and Predictive analytics	**40**
CO-3	Understand, analyze and apply concepts of hypothesis testing, analysis of variance and forecasting	**30**

Opp. Shastri Maidan, Beside BVM College, Vallabh Vidyanagar, Dist: Anand, Gujarat - 388120
(O): 02692-238001 | Email: adminoffice@cvmu.edu.in | www.cvmu.edu.in

Curriculum Revision:	
Version:	2.0
Drafted on (Month-Year):	June-2022
Last Reviewed on (Month-Year):	-
Next Review on (Month-Year):	June-2025

Opp. Shastri Maidan, Beside BVM College, Vallabh Vidyanagar, Dist: Anand, Gujarat - 388120
(O): 02692-238001 | Email: adminoffice@cvmu.edu.in | www.cvmu.edu.in