

NEP and Learning Outcomes-based Curriculum Framework (LOCF)

For
PG Diploma in Business Analytics
(To be effective from the Academic Session 2022-23)



Department of Management
Gurugram University, Gurugram
(A State Govt. University Established Under Haryana Act 17 of 2017)

Table of Content

S.No.	Particulars	Page No.
1	Background	1
2	Programme Outcomes	2
3	Programme Specific Outcomes	3
4	Postgraduate Attributes	4
5	Qualification descriptors	5
6	Teaching Pedagogy	5
7	Scheme of Programme/Details of Courses	6-7
8	Learning outcome index	8
9	Course Outcomes and Mapping Matrix	8
10	Syllabus	9-33

Background

In the modern age, technology is making its way into all courses and sectors. Information is required to be collected, analysed, sorted and arranged with the use of technology. This comes under the Business Analytics specialization. Hence, it is one of the fastest rising specialisations among students today. A career in business analytics is highly rewarding with great future prospects.

PGDM Business Analytics is a professional course which is gaining huge popularity amongst the students pursuing management after graduation. PGDM Business Analytics puts the students in the forefront of technological advancements around the world through its course.

Easy Adaptability: PGDM Business Analytics is required in a wide range of industries and sectors. From the top MNC's to even the Hospitality and Entertainment industry, it includes all. It is relatively easier to switch and adapt between industries with this as your specialisation. Past experience and professional background is not a restriction on these choices as the degree is quite wide and far reaching. **A Dynamic Career:** This career makes way for professional development throughout their career. The nature of the job urges one to keep updated with various departments and sectors, which is a continuous and enriching learning process.

Top Recruiters: Graduates with this degree get exposed to top recruits like Amazon, JSW Steel, HCL, HP, Dell, Flipkart, GODREJ, IBM, Hitachi, etc. These companies not only offer decent starting packages but also provide recognition to the employees. **High Packages:** PGDM Business Analytics professionals can easily bag high paying jobs in either private or government sectors. The average salary offered to these graduates ranges is very high.

The benefits of choosing PGDM Business Analytics program are:

- **Trend:** Technology has become an integral part of every industry around the world. PGDM Business analytics is a trending course particularly in today's business scenario.
- **Future Scope:** Data is the new oil. It is the future of everything according to the experts and the role of Business Analytics is about making use of the data to make business decisions.
- **Different Course:** PGDM Business Analytics is different from other Management courses. It is quite versatile and deals in technology, numbers, analysis and decision making.
- **Professional Prospects:** PGDM Business Analytics professionals are high in demand all around the world due to modern day business working. The field is growing at a fast pace as the technological advancements are taking place around the world.

1.

Programme Outcomes

On completing PG Diploma in Business Analytics Programme, the students shall be able to realize following Programme outcomes:

PO	Description
PO-1	Self-Directed Learning: Develop the ability to work independently as well as effectively in the changing environment.
PO-2	Problem Solving: Visualize, conceptualize, articulate and solve complex problems through experimentation and observation using theoretical framework of social science disciplines.
PO-3	Critical Thinking and Scientific Enquiry: Critically analyse everyday problems faced by the society, evaluate specific policy proposals, compare arguments with different conclusions to a specific societal issue and assess the role played by assumptions in such arguments. Develop the capability of defining problems, formulate hypothesis, collect relevant data, develop empirical evidence and interpret the results of such analyses.
PO-4	Usage of Analytical Tools: Develop the ability to apply appropriate quantitative/qualitative techniques used in social science disciplines along with ICT, software's etc.
PO-5	Specialization and Employability: Develop deeper understanding, creativity, originality, analytical and critical skills in chosen specialized areas of social science disciplines leading to employability.
PO-6	Life Skills: Articulate and apply ethics, values and ideals that demonstrate awareness of current societal challenges leading to better quality of life.
PO-7	Leadership: Build skills to work as part of a team and lead others, setting directions and formulating inspiring vision.
PO-8	Communication: Communicate conclusions, interpretations and implications clearly, concisely and effectively, both orally and in writing for different types of audiences.

2.

Programme Specific Outcomes

On completing PG Diploma in Business Analytics Programme, the students shall be able to realize following outcomes:

PSO	Description
PSO-1	Business Environment and Domain Knowledge (BEDK): Economic, legal and social environment of Indian business. Graduates are able to improve their awareness and knowledge about functioning of local and global business environment and society. This helps in recognizing the functioning of businesses, identifying potential business opportunities, involvement of business enterprises and exploring the entrepreneurial opportunities.
PSO-2	Critical thinking, Business Analysis, Problem Solving and Innovative Solutions (CBPI): Competencies in quantitative and qualitative techniques. Graduates are expected to develop skills on analysing the business data, application of relevant analysis, and problem solving in other functional areas such as marketing, business strategy and human resources.
PSO-3	Global Exposure and Cross-Cultural Understanding (GECCU): Demonstrate a global outlook with the ability to identify aspects of the global business and Cross-Cultural Understanding
PSO-4	Social Responsiveness and Ethics (SRE): Developing responsiveness to contextual social issues / problems and exploring solutions, understanding business ethics and resolving ethical dilemmas. Graduates are expected to identify the contemporary social problems, exploring the opportunities for social entrepreneurship, designing business solutions and demonstrate ethical standards in organizational decision making. Demonstrate awareness of ethical issues and on distinguish ethical and unethical behaviours.
PSO-5	Effective Communication (EC): Usage of various forms of business communication, supported by effective use of appropriate technology, logical reasoning, articulation of ideas. Graduates are expected to develop effective oral and written communication especially in business applications, with the use of appropriate technology (business presentations, digital communication, social network platforms and so on).
PSO-6	Leadership and Teamwork (LT): Understanding leadership roles at various levels of the organization and leading teams. Graduates are expected to collaborate and lead teams across organizational boundaries and demonstrate leadership qualities, maximize the usage of diverse skills of team members in the related context.

3.

Postgraduate Attributes

- Disciplinary Knowledge
- Creative and Critical Thinking
- Reflective Thinking
- Problem Solving
- Analytical Reasoning
- Communication Skills
- Research Skills
- Life Skills
- Multicultural Competence
- Moral and Ethical Values
- Life-long Learning
- Global Competence

4.

Qualification Descriptors

The eligibility criteria is that the candidate should have passed its Bachelor's Degree in any discipline from a UGC recognised university or college. Candidates must have scored 50% marks in Bachelor's Degree in the any field. Students must have a minimum of three-year graduate degree with a minimum 50% aggregate marks from any University recognized by AIU /AICTE are eligible. Any graduate can pursue Business Analytics. However, a Science or Commerce background is generally more helpful as it provides a link to maths and statistics which is involved in Business Analytics.

5.

Teaching Pedagogy

1. The course contents will be delivered through lectures, presentations, case analysis, discussions, assignments, and audio-visual tools.
2. Spreadsheet is the recommended software for doing basic calculations in subjects applicable, hence shall be used for teaching, practice, problem solving and assignments.
3. It is mandatory to include at least 2 case studies for the delivery of course. The list of cases and specific references including recent articles will be announced in the class at the time of launching of the course.

6.

Scheme of Programme**Semester 1**

Course Code	Course Title	L	T	P	Credits	External Marks (Theory)	Internal Assessment Marks/Workshop	External/Practical Marks
221DBACC1	Statistics with R (BL)	2	0	4	4	70	30	
221DBACC2	Python Programming	2	0	4	4	70	30	
221DBASEC3	Structured Query Language	2	0	4	4	70	30	
221DBASEC4	MOOC Course (BL) - Note 3	0	3	0	3			100*
221DBAAEC5	SaS and Tableau	2	0	4	4	35	30	35
221DBAVAC6	Project Report	0	3	0	3		30	70
Total Credits					22			

- Notes:** 1. After successfully completed first semester and 22 credits, Post Graduate certificate in Business Analytics will be awarded to the students.
2. BL is blended learning mode
3. A student can undertake a project or can complete any additional subject (s) for 3 credits related to Business Analytics from an open source and can submit the proof to the examination branch of the university after successful completion of the course.
4. *Each student is required to pass MOOC available on SWAYAM portal or any other online educational platform of repute of 3 credits (Option will be given by the Course Coordinator). The student is required to submit the passing certificate of the same to the department. The Evaluation of MOOC Courses will be 50% certificate earned and 50% for External viva.
5. Project Report Evaluation: Project report evaluation has 30 marks for internal evaluation which will be done by Internal Guide /Mentor and 70 marks will be based on External viva before the committee of three members constituted by Dean/Chairperson of the Department.

Semester 2

Course Code	Course Title	L	T	P	Credits	External Marks (Theory)	Internal Assessment Marks/Workshop	External/Practical Marks
222DBACC1	Big Data Analytics (BL)	2	0	4	4	70	30	
222DBACC2	Social Media Analytics	2	0	4	4	70	30	
222DBAAEC3	Natural Language Processing	2	0	4	4	35	30	35
222DBASEC4	Machine Learning and Artificial Intelligence	2	0	4	4	70	30	
222DBASEC5	MOOC Course (BL) - Note 3	0	3	0	3			100*
222DBAVAC6	Project Report	0	3	0	3		30	70
Total Credits					22			

Notes: 1. After successfully completed second semester and a total of 44 credits, Post Graduate Diploma in Business Analytics will be awarded to the students.

2. BL is blended learning mode

3. A student can undertake a project or can complete any additional subject (s) for 3 credits related to Business Analytics from an open source and can submit the proof to the examination branch of the university after successful completion of the course.

After successfully completed second semester and a total of 44 credits, Post Graduate Diploma in Business Analytics will be awarded to the students.

4.*Each student is required to pass MOOC available on SWAYAM portal or any other online educational platform of repute of 3 credits (Option will be given by the Course Coordinator). The student is required to submit the passing certificate of the same to the department. The Evaluation of MOOC Courses will be 50% certificate earned and 50% for External viva.

5. Project Report Evaluation: Project report evaluation have 30 marks for Internal evaluation which will be done by Internal Guide /Mentor and 70 marks will be based on External viva before the committee of three members constituted by Dean/Chairperson of the Department.

Details of courses

Semester	Core Courses (CC)		Skill Enhancement Courses (SEC)		Ability Enhancement Courses (AEC)		Value Addition Courses (VAC)		Total Credits
	No. of Courses	Total Credits	No. of Courses	Total Credits	No. of Courses	Total Credits	No. of Courses	Total Credits	
I	2	8	2	7	1	4	1	3	22
II	2	8	2	7	1	4	1	3	22
Total	Core Credits	16	Skill Enhancement Credits	14	Ability Enhancement Courses (AEC)	8	Value Addition Courses (VAC)	6	44
%age	Core Credits	36	Skill Enhancement Credits	32	Ability Enhancement Courses (AEC)	18	Value Addition Courses (VAC)	14	100.00

	After 1 years		
	No of courses	No of credits per course	Total no of credits
Core Courses	4	4	16
Ability Enhancement Courses	2	4	8
Skill enhancement	2	4	8
Value Added Courses	2	6	12
TOTAL CREDITS			44

Learning Outcome Index

Semester	PSO →	PSO1	PSO2	PSO3	PSO4	PSO5	PSO6	PSO7	PSO8
	Course code ↓								
Semester-1	221DBACC1	✓	✓				✓		✓
	221DBACC2	✓		✓		✓			
	221DBASEC3	✓	✓		✓		✓		✓
	221DBASEC4	✓	✓		✓	✓	✓	✓	
	221DBAAEC5			✓			✓	✓	
	221DBAVAC6			✓					✓
Semester-2	222DBACC1	✓	✓		✓				
	222DBACC2	✓		✓			✓		
	222DBAAEC3		✓	✓			✓	✓	
	222DBASEC4	✓					✓		✓
	222DBASEC5			✓			✓	✓	
	222DBAVAC6				✓	✓			

8. Course Outcomes and Mapping Matrix:

Each Course of the Diploma Programme results in few Course/Learning Outcomes (COs) which are broadly mapped or associated with POs as well as PSOs.

Mapping is a process of representing the correlation between COs and POs, COs and PSOs in the scale of 1 to 3 as follows.

Scale of Mapping between COs and POs & COs and PSOs

Scale 1	If the contents of course have low correlation (i.e. in agreement with the particular PO to a small extent) with the particular Programme outcome.
Scale 2	If the contents of course have medium correlation (i.e. in agreement with the particular PO to a reasonable extent) with the particular Programme outcome.
Scale 3	If the contents of course have strong correlation (i.e. in agreement with the particular PO to a large extent) with the particular Programme outcome.

Syllabus

Statistics with R Course Code: 221DBACC1

L T P
2 0 4

External Marks: 70
Internal Marks: 30
Time Allowed: 3 Hrs

Employability Level: Professional Core

Core Courses	Ability Enhancement Courses	Skill enhancement	Value Added Courses
✓			

Introduction to the Course: The following module comprises of R programming basics and application of several Statistical Techniques using it. The module aims to provide exposure in terms of Statistical Analysis, Hypothesis Testing, Regression and Correlation using R programming language. The objective of this module to make students exercise the fundamentals of statistical analysis in R environment. They would be able to analysis data for the purpose of exploration using Descriptive and Inferential Statistics. Students will understand Probability and Sampling Distributions and learn the creative application of Linear Regression in multivariate context for predictive purpose.

Course Outcomes: On completion of this course, the students will be able to

CO1: Install Code and Use R Programming Language in R Studio IDE to perform basic tasks on Vectors, Matrices and Data frames.

CO2: Describe key terminologies, concepts and techniques employed in Statistical Analysis.

CO3: Define Calculate, Implement Probability and Probability Distributions to solve a wide variety of problems.

CO4: Conduct and interpret a variety of Hypothesis Tests to aid Decision Making.

Detailed Syllabus:

Unit -I

R and R Studio, Logical Arguments, Missing Values, Characters, Factors and Numeric, Help in R, Vector to Matrix, Matrix Access, Data Frames, Data Frame Access, Basic Data Manipulation Techniques, Usage of various apply functions – apply, lapply, sapply and tapply, Outliers treatment.

Unit II

Types of Data, Nominal, Ordinal, Scale and Ratio, Measures of Central Tendency, Mean, Mode and Median, Bar Chart, Pie Chart and Box Plot, Measures of Variability, Range, Inter-Quartile- Range, Standard Deviation, Skewness and Kurtosis, Histogram, Stem and Leaf Diagram, Standard Error of Mean and Confidence Intervals. Experiment, Sample Space and Events, Classical Probability, General Rules Of Addition, Conditional Probability, General Rules For Multiplication, Independent Events, Bayes' Theorem, Discrete Probability Distributions: Binomial, Poisson, Continuous Probability Distribution, Normal Distribution & t-distribution, Sampling Distribution and Central Limit Theorem.

Unit III

Population and Sample, Null and Alternate Hypothesis, Level of Significance, Type I and Type II Errors, One Sample t Test, Confidence Intervals, One Sample Proportion Test, Paired Sample t Test, Independent Samples t Test, Two Sample Proportion Tests, One Way Analysis of Variance and Chi Square Test.

Unit IV

Analysis of Relationship, Positive and Negative Correlation, Perfect Correlation, Correlation Matrix, Scatter Plots, Simple Linear Regression, R Square, Adjusted R Square, Testing of Slope, Standard Error of Estimate, Overall Model Fitness, Assumptions of Linear Regression, Multiple Regression, Coefficients of Partial Determination, Durbin Watson Statistics, Variance Inflation Factor.

Suggested Readings:

1. Ken Black, 2013, Business Statistics, New Delhi, Wiley.
2. Lee, Cheng. et al., 2013, Statistics for Business and Financial Economics, New York: Heidelberg Dordrecht.
3. Anderson, David R., Thomas A. Williams and Dennis J. Sweeney, 2012, Statistics for Business and Economics, New Delhi: South Western.
4. Waller, Derek, 2008, Statistics for Business, London: BH Publications.
5. Levin, Richard I. and David S. Rubin, 1994, Statistics for Management, New Delhi: Prentice Hall.

Instructions for External Examiner: The question paper shall be divided in two sections as follows:

Section A	Seven (7) short answer type questions from whole of the syllabus carrying two marks each, This section will be compulsory	7*2=14 marks
Section B	8 questions (2 questions from each unit). The students will be required to attempt four questions selecting one question of 14 marks from each unit.	14*4= 56 marks
	Total Marks	70 marks

Instructions for Internal Examiner: The internal assessment should be spread evenly throughout the semester and must include at least 3 independent components including a mid-term exam. Below are the suggested components for 30 marks. A teacher has a choice to change these components as per the need except the mid-term exam.

S. No.	Course Assessment Components	Marks/Weightage (%)
1	Assessment 1 : Class Participation(CP) And Individual Assessment	10
2	Assessment 2 : Mid Term Exam (MTE)	10
3	Assessment 3 : Case Analysis / Presentation (CAP)/ Group Project (GP) / Role Play / Live Projects/ Simulation / Worksheet Assessment	10
	Internal Assessment (IA) (1+2+3)	30 (30%)
	End-Term Examination (EE)	70 (70%)
Total Marks (IA+EE)		100

Mapping Matrix of Course: 221DBACCI

Table 1: CO-PO Matrix for the Course: 221DBACCI

COURSE OUTCOMES	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8
CO1	3	2	3	2	3	3	2	3
CO2	3	3	3	3	3	3	3	3
CO3	3	3	3	3	3	3	2	2
CO4	3	3	3	3	3	3	3	2
Average	3	2.75	3	2.75	3	3	2.5	2.5

Table 2: CO-PSO Matrix for the Course: 221MGCC1

CO	PSO1	PSO2	PSO3	PSO4	PSO5	PSO6
CO1	3	3	2	2	3	3
CO2	3	3	2	2	2	3
CO3	3	3	3	3	3	3
CO4	3	3	3	3	3	3
Average	3	3	2.5	2.5	2.75	3

Python Programming
Course Code: 221DBACC2

L T P
2 0 4

External Marks: 70
Internal Marks: 30
Time Allowed: 3 Hrs

Employability Level: Professional Core

Core Courses	Ability Enhancement Courses	Skill enhancement	Value Added Courses
✓			

Introduction to the Course: Python Programming module is intended for students who wish to learn the Python programming language. This module is highly important so as to proceed with this programme. The module comprises of Programming basics with regards to Python Language such as Data Types, Operators, Functions, Classes and Exception Handling. This module will help students gain much needed knowledge pertaining to Python Programming, so as to prepare them for the advanced modules such as ML. Python scripting is user-friendly and is the most used language in industry when it comes to designing and scripting applications with respect to Emerging Technologies.

Course Outcomes: On completion of this course, the students will be able to

CO1: To understand why Python is a useful scripting language.

CO2: To learn how to use lists, tuples, and dictionaries in Python programs.

CO3: To learn how to write loops and decision statements in Python.

CO4: To learn how to write functions and pass arguments in Python.

Detailed Syllabus:

Unit-I

History of Python, Need of Python Programming, Applications Basics of Python Programming Using the REPL(Shell), Running Python Scripts, Variables, Assignment, Keywords, Input- Output, Indentation. Types - Integers, Strings, Booleans; Operators- Arithmetic Operators, Comparison (Relational) Operators, Assignment Operators, Logical Operators, Bitwise Operators, Membership Operators, Identity Operators, Expressions.

Unit-II

Lists, Operations, Slicing, Methods, Tuples, Sets, Dictionaries, Sequences, Comprehensions, Conditional blocks using If, Else and El-if, For Loop, For loop using Ranges, String, list and Dictionaries, While Loop, Loop Manipulation using Pass, Continue, Break and Else, Conditional and Loops Block.

Unit-III

Defining Functions, Calling Functions, Passing Arguments, Keyword Arguments, Default Arguments, Variable-length arguments, Anonymous Functions, Function Returning Values, Scope of the Variables in a Function - Global and Local Variables. Creating modules, Name Spacing, Introduction to PIP, Installing Packages via PIP, Using Python Packages.

Unit-IV

Classes, Self-Variable, Methods, Constructor Method, Inheritance, Overriding Methods, Data Hiding, Difference between an Error and Exception, Handling Exception, Try Except Block, Raising Exceptions, and User Defined Exceptions.

Suggested Readings:

1. R.Nageswara Rao, 2018, Core Python Programming, Dreamtech.
2. John Hearty, 2016, Advanced Machine Learning with Python, Packt.
3. Jake VanderPlas, 2016, Python Data Science Handbook: Essential Tools for Working with Data, O'Reilly.
4. Mark Lutz, 2010, Programming Python, O'Reilly.
5. Tim Hall and J-P Stacey, 2009, Python 3 for Absolute Beginners, Apress.

Instructions for External Examiner: The question paper shall be divided in two sections as follows:

Section A	Seven (7) short answer type questions from whole of the syllabus carrying two marks each, This section will be compulsory	7*2=14 marks
Section B	8 questions (2 questions from each unit). The students will be required to attempt four questions selecting one question of 14 marks from each unit.	14*4= 56 marks
	Total Marks	70 marks

Instructions for Internal Examiner: The internal assessment should be spread evenly throughout the semester and must include at least 3 independent components including a mid-term exam. Below are the suggested components for 30 marks. A teacher has a choice to change these components as per the need except the mid-term exam.

S. No.	Course Assessment Components	Marks/Weightage (%)
1	Assessment 1 : Class Participation(CP) And Individual Assessment	10
2	Assessment 2 : Mid Term Exam (MTE)	10
3	Assessment 3 : Case Analysis / Presentation (CAP)/ Group Project (GP) / Role Play / Live Projects/ Simulation / Worksheet Assessment	10
	Internal Assessment (IA) (1+2+3)	30 (30%)
	End-Term Examination (EE)	70 (70%)
Total Marks (IA+EE)		100

Mapping Matrix of Course : 221DBACC2**Table 1: CO-PO Matrix for the Course: 221DBACC2**

COURSE OUTCOMES	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8
CO1	3	2	3	2	3	3	2	3
CO2	3	2	3	3	3	3	3	3
CO3	3	3	3	3	3	3	2	2
CO4	3	2	3	3	3	3	3	2
Average	3	2.25	3	2.75	3	3	2.5	2.5

Table 2: CO-PSO Matrix for the Course: 221DBACC2

CO	PSO1	PSO2	PSO3	PSO4	PSO5	PSO6
CO1	3	2	2	2	3	3
CO2	3	3	2	2	2	3
CO3	3	3	3	3	3	3
CO4	3	3	3	3	3	3
Average	3	2.75	2.5	2.5	2.75	3

Structured Query Language

Course Code: 221DBASEC3

L T P
2 0 4

External Marks: 70
Internal Marks: 30
Time Allowed: 3 Hrs

Employability Level: Skill enhancement

Core Courses	Ability Enhancement Courses	Skill enhancement	Value Added Courses
		✓	

Introduction to the Course: In this course, the students will learn the basics of the SQL/No SQL and the Relational Databases. They will learn about the Relational Model and Relational Model concepts and constraints. The students will get exposure to key concepts with regards to SQL Language and DBMS such as Normalization, Transaction Processing along-side an exposure to No SQL programming. This module will help students gain much needed knowledge pertaining to Relational Database Management Systems, Data Models, SQL query processing, Normalization along with an introduction to No SQL Database systems using Mongo DB.

Course Outcomes: On completion of this course, the students will be able to

CO1: To understand the basic concepts and the applications of Database Systems.

CO2: To master the basics of SQL and construct queries using SQL.

CO3: To become familiar with the basic issues of Transaction Processing and Concurrency Control.

CO4: To become familiar with NO SQL Programming Language.

Detailed Syllabus:

Unit-I

Introduction-Database System Applications, Purpose of Database Systems, Views of Data, Data Abstraction, Instances and Schemas, Data Models, Database Languages, DDL, DML, Database Architecture, Database Users and Administrators, Database Design, ER Diagrams, Entities, Attributes and Entity Sets, Relationships and Relationship sets, Integrity Constraints, Views.

Unit-II

Relational Algebra and Calculus, Selection and Projection, Set Operations, Renaming, Joins, Division, Relational calculus, Tuple Relational Calculus, Domain Relational Calculus, Forms of Basic SQL Query, Nested Queries, Comparison Operators, Aggregate Operators, NULL values, Logical connectives, AND, OR and NOT, Outer Joins, Triggers. Problems Caused by Redundancy, Decompositions, Functional Dependencies, Normal Forms, First, Second, Third Normal forms, BCNF, Properties of Decompositions, Loss less Join Decomposition, Dependency Preserving Decomposition, Multi Valued Dependencies, Fourth Normal Form, Join Dependencies, Fifth Normal Form.

Unit-III

Transaction Management, Transaction Concept, Transaction State, Implementation of Atomicity and Durability, Concurrent, Executions, Serializability, Recoverability, Implementation of Isolation, testing for serializability, Concurrency Control, Lock, Timestamp Based Protocols, Validation Based Protocols, Recovery, Failure Classification, Storage Structure, Atomicity, Log Based Recovery, Remote Backup Systems.

Unit-IV

Overview of No SQL, Types of No SQL Databases, No SQL Storage Architecture, CRUD Operations in MongoDB, Querying, Modifying and Managing No SQL Databases, Indexing and Ordering, Migrating from RDBMS to No SQL, No SQL in Cloud, Database Administration.

Suggested Readings:

1. Guy Harrison, 2015, Next Generation Databases: No SQL and Big Data, Apress.
2. Ramez Elmasri, ShamkatB.Navathe, 2013, Database Systems, Pearson.
3. Pramod J. Sadalage, Martin Fowler, 2012, No SQL Distilled, Addison Wesley.
4. Silberschatz, H.F. Korth, S.Sudarshan, 2006, Database System Concepts, McGraw Hill.
5. Raghurama Krishnan, Johannes Gehrke, 2003, Database Management Systems, McGraw Hill.

Instructions for External Examiner: The question paper shall be divided in two sections as follows:

Section A	Seven (7) short answer type questions from whole of the syllabus carrying two marks each, This section will be compulsory	7*2=14 marks
Section B	8 questions (2 questions from each unit). The students will be required to attempt four questions selecting one question of 14 marks from each unit.	14*4= 56 marks
	Total Marks	70 marks

Instructions for Internal Examiner: The internal assessment should be spread evenly throughout the semester and must include at least 3 independent components including a mid-term exam. Below are the suggested components for 30 marks. A teacher has a choice to change these components as per the need except the mid-term exam.

S. No.	Course Assessment Components	Marks/Weightage (%)
1	Assessment 1 : Class Participation(CP) And Individual Assessment	10
2	Assessment 2 : Mid Term Exam (MTE)	10
3	Assessment 3 : Case Analysis / Presentation (CAP)/ Group Project (GP) / Role Play / Live Projects/ Simulation / Worksheet Assessment	10
	Internal Assessment (IA) (1+2+3)	30 (30%)
	End-Term Examination (EE)	70 (70%)
Total Marks (IA+EE)		100

Mapping Matrix of Course: 221DBASEC3**Table 1: CO-PO Matrix for the Course: 221DBASEC3**

COURSE OUTCOMES	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8
CO1	3	2	3	2	3	3	2	3
CO2	3	3	3	3	3	3	3	3
CO3	3	3	3	3	3	3	2	2
CO4	3	3	3	3	3	3	3	2
Average	3	2.75	3	2.75	3	3	2.5	2.5

Table 2: CO-PSO Matrix for the Course: 221DBASEC3

CO	PSO1	PSO2	PSO3	PSO4	PSO5	PSO6
CO1	3	3	2	2	3	3
CO2	3	3	2	2	2	3
CO3	3	2	3	3	3	3
CO4	3	3	3	3	3	3
Average	3	2.75	2.5	2.5	2.75	3

SaS and Tableau

Course Code: 221DBAAEC5

L T P

2 0 4

External Marks: 35
External Practical :35
Internal Marks: 30
Time Allowed: 3 Hrs

Employability Level: Ability Enhancement Courses

Core Courses	Ability Enhancement Courses	Skill enhancement	Value Added Courses
	✓		

Introduction to the Course: This module forms an introduction as well an in-depth study in the discipline of SaS programming with regards to Statistical Analysis along with Data Visualization through Tableau. The module comprises of basic introduction and programming in SaS, using Procedures within SaS and Data Visualization using Tableau application. This course will provide students and exposure towards SaS, Tableau and its usability in the field of analytics. The course comprises of an introduction to SaS, its procedures, visualizations along with Tableau application usage and visualization basics.

Course Outcomes: On completion of this course, the students will be able to

CO1: Deploy SaS in a virtual environment and import data for analysis.

CO2: Prepare and manipulate datasets for analysis in SaS.

CO3: Perform exploratory data analysis within SaS environment using various procedures.

CO4: Understand Tableau Interface, Panes and Implement Visualization Techniques.

Detailed Syllabus:

Unit-I

Overview of SaS university edition, Deploying SaS Studio on virtual platform, File Management, SaS libraries, importing data, Structure of Data and Data Types, Program Syntax, saving data, PROC IMPORT and PROC CONTENTS, Displaying Data and Generating Logs, List Input. Variables and Syntax Rules, Data Set Options, Operators, In-File Statement, Input Styles, Select Statements, Leave and Continue, Decision Making via SaS, Where Statement, Looping Constructs, SaS Functions, Arrays and Array Processing, Modifying and Combining Data Sets.

Unit-II

Proc Data, Proc Sort, Proc Means Sort, Proc Means, Proc Univariate, Proc Freq, Proc Plot, Proc Sgplot, Proc Summary, Proc Contents, Proc Append, Proc Copy, Proc SQL, Proc Delete, Proc Format, Proc Import, Proc Export, Proc Transpose, Proc GChart, Proc GPlot, Proc Report.

Unit-III

Tableau Software Ecosystem, Toolbar Icons, Data Window and Aggregation, Tableau Data Source, Data Extract, Connect to Data, Measure Names, Number of Records & Measures, Heat Maps, Tree maps, Bar Chart, Line Chart, Area Fill Charts, Pie Chart, Scatter Plot, Circle View, Bullet Graph, Packed Bubble, Histogram, Boxplot and Gantt Chart, Sorting Data, Enhancing Views with Filters, Sets, Groups & Hierarchies.

Unit-IV

Cross-tabulation, Dashboard Designing, Dashboard Actions, Joining Database, Functions in Tableau, Aggregate Functions, Numeric Functions, Date Functions, Stories, Advanced Mapping, Advanced Parameters, Tableau Best Practices, Combining Multiple Dashboards into Stories, Publishing Stories and Dashboards.

Recommended Readings:

1. Ron Cody, 2018, An Introduction to SaS University Edition, SaS Institute.
2. Ron Cody, 2018, Learning SaS by Example, SaS Institute.
3. Deepti Gupta. 2018, Applied Analytics through Case Studies Using SaS, Apress.
4. Joshua N. Milligan, 2015, Learning Tableau, Packt.
5. Ben Jones, 2014, Communication Data with Tableau: Designing, Developing and Delivering Data Visualization, O'Reilly.

Instructions for External Examiner: The question paper shall be divided in two sections as follows:

Section A	Seven (7) short answer type questions from whole of the syllabus carrying one marks each, This section will be compulsory	7*1=7 marks
Section B	8 questions (2 questions from each unit). The students will be required to attempt four questions selecting one question of 7 marks from each unit.	4*7= 28 marks
	Total Marks	35 marks

Instructions for Internal Examiner: The internal assessment should be spread evenly throughout the semester and must include at least 3 independent components including a mid-term exam. Below are the suggested components for 30 marks. A teacher has a choice to change these components as per the need except the mid-term exam.

S. No.	Course Assessment Components	Marks/Weightage (%)
1	Assessment 1 : Class Participation(CP) And Individual Assessment	10
2	Assessment 2 : Mid Term Exam (MTE)	10
3	Assessment 3 : Case Analysis / Presentation (CAP)/ Group Project (GP) / Role Play / Live Projects/ Simulation / Worksheet Assessment	10
	Internal Assessment (IA) (1+2+3)	30 (30%)
	End-Term Examination (EE)	35(35%)
Total Marks (IA+EE)		100

Mapping Matrix of Course: 221DBAAEC5**Table 1: CO-PO Matrix for the Course: 221DBAAEC5**

COURSE OUTCOMES	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8
CO1	3	3	3	2	3	3	2	3
CO2	3	3	3	3	3	3	3	3
CO3	3	3	3	3	3	3	3	2
CO4	3	3	3	3	3	3	3	2
Average	3	3	3	2.75	3	3	2.75	2.5

Table 2: CO-PSO Matrix for the Course: 221DBAAEC5

CO	PSO1	PSO2	PSO3	PSO4	PSO5	PSO6
CO1	3	3	2	2	3	3
CO2	3	3	2	2	2	3
CO3	3	2	3	3	3	3
CO4	3	3	3	3	3	3
Average	3	2.75	2.5	2.5	2.75	3

Semester 2

Semester 2

Course Code	Course Title	Course ID	L	T	P	Credits
222DBACC1	Big Data Analytics (BL)		2	0	4	4
222DBACC2	Social Media Analytics		2	0	4	4
Ability Enhancement Course(s)						
222DBAAEC3	Natural Language Processing		2	0	4	4
Skill Enhancement Course(s)						
222DBASEC4	Machine Learning and Artificial Intelligence		2	0	4	4
222DBASEC5	MOOC Course (Note 3) (BL)		0	3	0	3
Value Addition Course (s)						
222DBAVAC6	Project Report		0	3	0	3
Total Credits						22

Big Data Analytics
Course Code: 222DBACC1

L T P
2 0 4

External Marks: 70
Internal Marks: 30
Time Allowed: 3 Hrs

Employability Level: Core Courses

Core Courses	Ability Enhancement Courses	Skill enhancement	Value Added Courses
✓			

Introduction to the course: This module forms an introduction as well an in-depth study in the field of Big Data and Hadoop. It comprises of the fundamentals as well as advanced topics needed to progress in this technology. The students will learn about the applications, usage and several use case scenarios pertaining to Big Data- Hadoop where they can use the knowledge and progress ahead. This course will help students gain knowledge and understanding about Big Data Technology, Hadoop Ecosystem and various tools related to it. The students will learn about the HDFS File System, Map Reduce Framework, Analysing data using Hbase and Hive along with the Integration of R with Hadoop.

Course Outcomes: On completion of this course, the students will be able to

CO1: Understand the fundamentals of Big Data and its Applications in various Domains.

CO2: Conceptualize and Incorporate the Technologies behind Big Data.

CO3: Understand HDFS File Structure, Map Reduce Framework, the architectures related to them

CO4: Integrate R with Hadoop and solve analytical problems.

Detailed syllabus:

Unit-I

What Is Big Data? History of Data Management, Evolution of Big Data, Structuring of Big Data, Elements of Big Data, Application of Big Data in the Business Context, Careers in Big Data. Business Applications of Big Data: The Significance of Social Network Data, Financial Fraud and Big Data, Fraud Detection in Insurance, Use of Big Data in the Retail Industry. Distributed and Parallel Computing for Big Data, Understanding Hadoop, Cloud Computing, Grid Computing and In-Memory Technology for Big Data. VMWare Installation of Hadoop, Linux and its Shell Commands, Different Hadoop Distributions and their advantages, Hortonworks, Cloudera, MapR.

Unit-II

The Hadoop Ecosystem, Storing Data with HDFS, Design of HDFS, HDFS Concepts, Command Line Interface to HDFS, Hadoop File Systems, Java Interface to Hadoop, Anatomy of a file read, Anatomy of a file write, Replica placement and Coherency Model. Parallel Copying with distcp, keeping an HDFS Cluster Balanced.

Unit-III

Origins of Map Reduce, How Map Reduce Works, Optimization Techniques for Map Reduce Jobs, Applications of Map Reduce, Java Map Reduce classes (new API), Data flow, combiner functions, running a distributed Map Reduce Job. Configuration API, setting up the development environment, Managing Configuration.

Unit-IV

Understanding R-Hadoop, Integration Procedure, Packages needed for R under Hadoop Ecosystem, Text Mining for Deriving Useful Information using R within Hadoop, Introduction to Hive & Hbase, Hive and Hbase Architecture, Understanding Queries, Mining Big Data with Hive & Hbase.

Suggested Readings:

1. Arshdeep Bahga, 2016, Big Data Science & Analytics: A Hands-On Approach, VPT.
2. Tom White, 2012, Hadoop: The Definitive Guide, O'Reilly.
3. Adam Shook and Donald Miner, 2012, Map Reduce Design Patterns: Building Effective
4. Algorithms and Analytics for Hadoop and Other Systems, O'Reilly.
5. Dean Wampler, Edward Capriolo & Jason Rutherglen, 2012, Programming Hive, O'Reilly.
6. Lars George, 2011, HBase - The Definitive Guide: Random Access to Your Planet- Size Data, O'Reilly.

Instructions for External Examiner: The question paper shall be divided in two sections as follows:

Section A	Seven (7) short answer type questions from whole of the syllabus carrying two marks each, This section will be compulsory	7*2=14 marks
Section B	8 questions (2 questions from each unit). The students will be required to attempt four questions selecting one question of 14 marks from each unit.	14*4= 56 marks
	Total Marks	70 marks

Instructions for Internal Examiner: The internal assessment should be spread evenly throughout the semester and must include at least 3 independent components including a mid-term exam. Below are the suggested components for 30 marks. A teacher has a choice to change these components as per the need except the mid-term exam.

S. No.	Course Assessment Components	Marks/Weightage (%)
1	Assessment 1 : Class Participation(CP) And Individual Assessment	10
2	Assessment 2 : Mid Term Exam (MTE)	10
3	Assessment 3 : Case Analysis / Presentation (CAP)/ Group Project (GP) / Role Play / Live Projects/ Simulation / Worksheet Assessment	10
	Internal Assessment (IA) (1+2+3)	30 (30%)
	End-Term Examination (EE)	70 (70%)
Total Marks (IA+EE)		100

Mapping Matrix of Course: 222DBACC1**Table 1: CO-PO Matrix for the Course: 222DBACC1**

COURSE OUTCOMES	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8
CO1	3	2	3	3	3	3	2	3
CO2	3	3	3	3	3	3	3	3
CO3	3	3	3	3	3	3	3	2
CO4	3	3	3	3	3	3	3	2
Average	3	2.75	3	3	3	3	2.75	2.5

Table 2: CO-PSO Matrix for the Course: 222DBACC1

CO	PSO1	PSO2	PSO3	PSO4	PSO5	PSO6
CO1	3	3	2	2	3	3
CO2	3	3	3	2	2	3
CO3	3	3	3	3	3	3
CO4	3	3	3	3	3	3
Average	3	3	2.75	2.5	2.75	3

Social Media Analytics
Course Code: 222DBACC2

L T P
2 0 4

External Marks: 70
Internal Marks: 30
Time Allowed: 3 Hrs

Employability Level: Core Courses

Core Courses	Ability Enhancement Courses	Skill enhancement	Value Added Courses
✓			

Introduction to the Course: This module comprises of advanced disciplines and units pertaining to Analytics, Data Acquisition via web, Understanding and analysing Unstructured Data along with the future implementations and growth in the discipline of Analytics. This course aims at giving exposure on the advanced aspects with regards to Analytics. The course comprises of Social Media, Mobile, Text Analytics along with Web Scrapping and the future advancements in the field of Analytics.

Course Outcomes: On completion of this course, the students will be able to

CO1: Apply and use Social Media Analytics for the betterment of the business.

CO2: Use Mobile Analytics for solving complex business problems and to stop churn.

CO3: Evaluate the business problem and apply analytics techniques for better output.

CO4: Analyse and understand patterns and techniques in Social Media & Mobile Analytics to solve complex problems.

Detailed Syllabus:

Unit -I

Social Media, On-Line Social Network, Off-Line Social Network, Metrics and Measurement, Dashboard, Target Audience, Desired Action, Content, Market Research Online Communities, Cluster Analysis, Conjoint Analysis, Multidimensional Scaling, Social Media Listening, Social Media Scoring, Social Media Modelling. Understanding Mobile Analytics Concepts, difference between Mobile Analytics and Site Analytics, Natural language Processing with Mobile Analytics, Text Mining for Mobile Analytics, Mobile Analytics Tools, Churn Analytics.

Unit -II

Text Data, Sources of Text Data, Information Clusters, Patterns, Trends, Tagging, Natural Learning Process, Lexical Analysis, Social Network Nodes, Linkage Structure, Node Labelling, Content-Based Classification, Word Stemming, Stemming Algorithms, Polarity of the Attitude, Psychological Profiling, Sentiment Analysis.

Unit -III

Web Scrapping of unstructured data, Gathering data from HTTP and HTTPS format, Web Scrapping from XML and JSON file, Regular expressions, Extraction Strategies, Term Document Matrix, Data Cleansing, Data Manipulation and Data Transformation after Scrapping.

Unit -IV

Introduction to Big Data, Predictive Analysis for Business, Social Information Processing and Distributed Computing, Advances in Machine Learning, Traditional Data Models Evolve, Analytics to Solve Social Problems, Location Based Data Explosion, Data Privacy Backlash, Internet of

Things, Artificial Intelligence.

Suggested Readings:

1. Galit Shamuelli, 2017, Data Mining for Business Analytics: Concepts, Techniques and Applications with R, Wiley.
2. Luis Torgo, 2017, Data Mining with R: Learning Case Studies, Chapman.
3. Zaki & Meira, 2014, Data Mining and Analysis Fundamental Concepts and Algorithms, Cambridge.
4. Han, Kamber & Pei, 2013, Data Mining: Concepts and Techniques, Morgan Kaufmann.
5. Han, Jiawei and Kamber, Micheline, 2012, Data Mining: Concepts and Techniques, Morgan Kaufman.

Instructions for External Examiner: The question paper shall be divided in two sections as follows:

Section A	Seven (7) short answer type questions from whole of the syllabus carrying two marks each, This section will be compulsory	7*2=14 marks
Section B	8 questions (2 questions from each unit). The students will be required to attempt four questions selecting one question of 14 marks from each unit.	14*4= 56 marks
	Total Marks	70 marks

Instructions for Internal Examiner: The internal assessment should be spread evenly throughout the semester and must include at least 3 independent components including a mid-term exam. Below are the suggested components for 30 marks. A teacher has a choice to change these components as per the need except the mid-term exam.

S. No.	Course Assessment Components	Marks/Weightage (%)
1	Assessment 1 : Class Participation(CP) And Individual Assessment	10
2	Assessment 2 : Mid Term Exam (MTE)	10
3	Assessment 3 : Case Analysis / Presentation (CAP)/ Group Project (GP) / Role Play / Live Projects/ Simulation / Worksheet Assessment	10
	Internal Assessment (IA) (1+2+3)	30 (30%)
	End-Term Examination (EE)	70 (70%)
Total Marks (IA+EE)		100

Mapping Matrix of Course: 222DBACC2**Table 1: CO-PO Matrix for the Course: 222DBACC2**

COURSE OUTCOMES	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8
CO1	3	2	3	2	3	3	2	3
CO2	3	3	3	3	3	3	3	3
CO3	3	3	3	3	3	3	2	2
CO4	3	3	3	3	3	3	3	2
Average	3	2.75	3	2.75	3	3	2.5	2.5

Table 2: CO-PSO Matrix for the Course: 222DBACC2

CO	PSO1	PSO2	PSO3	PSO4	PSO5	PSO6
CO1	3	3	2	2	3	3
CO2	3	3	2	2	2	3
CO3	3	3	3	3	3	3
CO4	3	3	3	3	3	3
Average	3	3	2.5	2.5	2.75	3

Natural Language Processing

Course Code: 222DBAAEC3

L T P
2 0 4

External Marks: 35
External Practical :35
Internal Marks: 30
Time Allowed: 3 Hrs

Employability Level: Ability Enhancement Courses

Core Courses	Ability Enhancement Courses	Skill enhancement	Value Added Courses
	✓		

Introduction to the course: The following module aims to provide an exposure to students based on Natural Language Processing Techniques, Algorithms and Analytics based on Textual Data. The module comprises of NLP basics, Feature Extraction, Relationship Extraction, Text Categorization, Clustering and Topic Modelling Algorithms and Practices. The course introduces the concepts of Text Analytics, Unstructured Information Analysis for better decision making by deriving valuable insights. The course will help the students understand the roots behind Text Mining which evolved from Machine Learning, Natural Language Processing and Statistics. Upon completion, students are expected to be able to describe basic concepts and methods of Text Mining, Information Extraction, Text Classification and Clustering, Topic Modelling.

Course Outcomes: On completion of this course, the students will be able to

CO1: Understand approaches to Syntax and Semantics in NLP.

CO2: Understand various methods for Statistical approaches to Machine Translation.

CO3: Build Models which extract information from Textual Unstructured Data.

CO4: Understand and implement Topic Modelling and Probabilistic Models for Information Extraction.

Detailed Syllabus:

Unit-I

Basics of Text Mining, Natural Language Content Analysis, Core Text Mining Operations, Associations, Using Background Knowledge for Text Mining, Domain Ontologies, Domain Lexicons. Text Mining Pre-processing Techniques, Task Oriented Approaches, NLP Tasks, Tokenization, Part-of-Speech Tagging, Syntactical Parsing and Shallow Parsing. Finding Implicit Features, Finding Opinion Phrases and their Polarity, Context-Specific Word Semantic Orientation, Analysis of Word and Document Frequency, tf-idf, Zipf's Law, bind tf_idf Function, Subsequence Kernels for Relation Extraction, Capturing Relation Patterns with a String Kernel.

Unit-II

Applications of Text Categorization, Document Representation, Knowledge Engineering Approach to Text Categorization, Machine Learning Approach to Text Categorization, Evaluation of Text Classifiers. Clustering Tasks in Text Analysis, Clustering Algorithms and Clustering of Textual Data.

Unit-III

Tokenizing by N-gram, Counting and Filtering N-gram, Analysing Bigrams to provide Context in Sentiment Analysis, visualizing a Network of Bigrams using ggraph, Counting and Correlating Pairs of Words with the widyr Package, Counting and Correlating among Sections, Examining Pairwise Correlation.

Unit-IV

Latent Dirichlet Allocation, Word Topic Probabilities, Per-Document Classification, By-words Assignments, Alternative LDA Implementations. Hidden Markov models, Stochastic Context Free Grammar, Conditional Random fields, Parallel Learning Algorithms.

Suggested Readings:

1. Julia Silge, David Robinson, 2018, Text Mining with R-A Tidy Approach, O'Reilly
2. Matthew L. Jockers, 2014, Text Analysis with R for Students of Literature, Springer.
3. James Pustejovsky, Amber Stubbs, 2012, Natural Language Annotation for Machine Learning, O'Reilly.
4. Steve R. Poteet, 2007, Natural Language Processing with Text Mining, Springer.
5. James Sanger, Ronen Feldman, 2002, The Text Mining Handbook: Advanced Approaches in Analysing Unstructured Data, Cambridge.

Instructions for External Examiner: The question paper shall be divided in two sections as follows:

Section A	Seven (7) short answer type questions from whole of the syllabus carrying one marks each, This section will be compulsory	7*1=7 marks
Section B	8 questions (2 questions from each unit). The students will be required to attempt four questions selecting one question of 7 marks from each unit.	4*7= 28 marks
	Total Marks	35 marks

Instructions for Internal Examiner: The internal assessment should be spread evenly throughout the semester and must include at least 3 independent components including a mid-term exam. Below are the suggested components for 30 marks. A teacher has a choice to change these components as per the need except the mid-term exam.

S. No.	Course Assessment Components	Marks/Weightage (%)
1	Assessment 1 : Class Participation(CP) And Individual Assessment	10
2	Assessment 2 : Mid Term Exam (MTE)	10
3	Assessment 3 : Case Analysis / Presentation (CAP)/ Group Project (GP) / Role Play / Live Projects/ Simulation / Worksheet Assessment	10

	Internal Assessment (IA) (1+2+3)	30 (30%)
	End-Term Examination (EE)	35(35%)
	External Practical	35(35%)
Total Marks (IA+EE)		100

Mapping Matrix of Course: 222DBAAEC3

Table 1: CO-PO Matrix for the Course: 222DBAAEC3

COURSE OUTCOMES	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8
CO1	3	3	3	2	3	3	2	3
CO2	3	3	3	3	3	3	3	3
CO3	3	3	3	3	3	3	2	2
CO4	3	3	3	3	3	3	3	2
Average	3	3	3	2.75	3	3	2.5	2.5

Table 2: CO-PSO Matrix for the Course: 222DBAAEC3

CO	PSO1	PSO2	PSO3	PSO4	PSO5	PSO6
CO1	3	3	2	3	3	3
CO2	3	3	2	3	2	3
CO3	3	3	3	3	3	3
CO4	3	3	3	3	3	3
Average	3	3	2.5	3	2.75	3

Machine Learning and Artificial Intelligence

Course Code: 222DBASEC4

L T P
2 0 4

External Marks: 70

Internal Marks: 30

Time Allowed: 3 Hrs

Employability Level: Skill enhancement

Core Courses	Ability Enhancement Courses	Skill enhancement	Value Added Courses
		✓	

Introduction to the Course: This module comprises of conceptual, statistical and Machine Learning algorithms. It covers all the necessary models pertaining to Machine Learning such as Regression, Decision Trees, Support Vectors, Clustering, Association Analysis and Dimensionality Reduction. After this course students will gain critical knowledge and understanding about major Data Mining procedures like Decision Tree, Cluster Analysis, Neural Networks, Support Vector Machine, Bayesian Networks and Machine Learning fundamentals. Students will be able to apply and practice this gained knowledge in variety of Business Scenarios.

Course Outcomes: On completion of this course, the students will be able to

CO1: Understand and employ a wide variety of Statistical and Machine Learning Algorithms.

CO2: Identify the characteristics of Datasets, Problem Statement and develop Machine Learning programs with reference to known Computing Techniques.

CO3: Understand the Model Performance Evaluation and select the best one based on the solution.

CO4: Implement Machine Learning techniques and the Programming Framework to obtain acceptable decisions for the Real-World problems.

Detailed syllabus:

Unit - I

Classification & Regression, working of a Decision Tree, Attribute Selection Measures, Information Gain, Gain Ratio, Gini Index, Building Decision Trees, CART, C5.0, and CHAID Trees, Prediction by Decision Tree, Advantages and Disadvantages of Decision Trees, Model Overfitting, Building Decision Trees in R.

Unit - II

Cluster Analysis versus Factor Analysis, Overview of Basic Clustering Methods, Agglomerative Hierarchical Clustering, Within-Group Linkage, Nearest Neighbour or Single, Linkage, Furthest Neighbour or Complete Linkage, Centroid Clustering, Ward's Method, K- Means Algorithm, Dendrogram, Profiling of Cluster, Cluster Evaluation. Decision Boundaries for Support Vector Machine, Maximum Margin Hyperplanes, Structural Risk Minimization, Linear SVM-Separable Case, Linear SVM-Non-Separable Case, Kernel Function, Kernel Trick, Kernel Hilbert Space, Model Evaluation.

Unit - III

Market Basket Analysis and Association Analysis, Market Basket Data, Stores, Customers, Orders, Items, Order Characteristics, Product Popularity, Tracking Marketing Interventions, Association Rules, Support, Confidence, Lift, Chi-Square Value, Sequential Pattern Analysis.

Unit - IV

Current Trends in AI, Intelligent Agents, Environments, Problem Solving Agents, Searching Techniques, Knowledge and Reasoning in AI, Forms of Learning, Structure of a Neural Network, Analogy with Biological Neural Network, Activation Functions, Gradient Descent, Model Accuracy.

Suggested Readings:

1. Kevin Knight, Elaine Rich, B.Nair, 2017, Artificial Intelligence, McGraw.
2. Han, Jiawei and Kamber, Micheline, 2012, Data Mining: Concepts and Techniques, Morgan Kaufman Publishers.
3. AnandRajaraman, 2011, Mining of Massive Datasets, Cambridge University Press.
4. Mitchell, 2013, Machine Learning, McGraw Hill.
5. Stuart Russell, Peter Norvig, 2004, Artificial Intelligence – A Modern Approach, Pearson.

Instructions for External Examiner: The question paper shall be divided in two sections as follows:

Section A	Seven (7) short answer type questions from whole of the syllabus carrying two marks each, This section will be compulsory	7*2=14 marks
Section B	8 questions (2 questions from each unit). The students will be required to attempt four questions selecting one question of 14 marks from each unit.	14*4= 56 marks
	Total Marks	70 marks

Instructions for Internal Examiner: The internal assessment should be spread evenly throughout the semester and must include at least 3 independent components including a mid-term exam. Below are the suggested components for 30 marks. A teacher has a choice to change these components as per the need except the mid-term exam.

S. No.	Course Assessment Components	Marks/Weightage (%)
1	Assessment 1 : Class Participation(CP) And Individual Assessment	10
2	Assessment 2 : Mid Term Exam (MTE)	10
3	Assessment 3 : Case Analysis / Presentation (CAP)/ Group Project (GP) / Role Play / Live Projects/ Simulation / Worksheet Assessment	10
	Internal Assessment (IA) (1+2+3)	30 (30%)

	End-Term Examination (EE)	70 (70%)
Total Marks (IA+EE)		100

Mapping Matrix of Course: 222DBASEC4

Table 1: CO-PO Matrix for the Course: 222DBASEC4

COURSE OUTCOMES	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8
CO1	3	2	3	3	3	3	2	3
CO2	3	3	3	3	3	3	3	3
CO3	3	3	3	3	3	3	3	2
CO4	3	3	3	3	3	3	3	2
Average	3	2.75	3	3	3	3	2.75	2.5

Table 2: CO-PSO Matrix for the Course: 222DBASEC4

CO	PSO1	PSO2	PSO3	PSO4	PSO5	PSO6
CO1	3	3	3	2	3	3
CO2	2	3	3	2	3	3
CO3	3	3	3	3	3	3
CO4	3	3	3	3	3	3
Average	2.75	3	3	2.5	3	3