I Semester M. Tech. CSE (BDA)

Sr. No	Course Code	Course Title	TOTAL Hours			Credits	Maximum Marks		Total	Minimum Passing Marks	No. of Hrs for ESE
			L	T	P		Continual Assessment	End Sem Examination			
1	25BD101T	Fundamentals of Data Analytics	3	-	-	3	40	60	100	50	3
2	25BD102T	Artificial Intelligence and Machine Learning	3	-	-	3	40	60	100	50	3
3	25BD103T	Data Engineering Foundation	3	-	-	3	40	60	100	50	3
4	25BD104T	Program Elective-I	3	-	-	3	40	60	100	50	3
5	25BD105P	Technical Seminar-I	-	-	8	4	100	-	100	50	-
6	25BD106P	Mini Project-I	-	-	8	4	100	-	100	50	-
Total			12	-	16	20	360	240	600	300	-

Program Elective Basket

Program Elective-I	
25BD104T (i)	Data Science for Engineers
25BD104T (ii)	Big Data Computing

FIRST SEMESTER

Course Code	Course Name	Th	Tu	Pr	Credits	Evaluation		
						CA	ESE	Total
25BD101T	Fundamentals of Data Analytics	3	-	-	3	40	60	100

Course Objectives	Course Outcomes
This course is intended 1. To make the students understand the fundamentals of excel for data analysis. 2. To equip the students with skills in Excel basics and intermediate features for data processing and analysis. 3. To facilitate the students to advanced Excel tools for data transformation and visualization. 4. Develop proficiency in organizing, managing, and visualizing data using Excel. 5. Solve engineering problems through data-driven decision making.	**At the end of the Course, the Student will be able to:** 1. Apply statistical methods in engineering data analysis using Excel. 2. Apply Excel basics and intermediate features for data processing and analysis. 3. Use advanced Excel tools and functions for data transformation and visualization. 4. Competence in interpreting and presenting data through charts and graphs. 5. Create and present data-driven solutions using Excel.

Unit I **[9Hrs]**

EXCEL FUNDAMENTALS FOR DATA PREPARATION AND VISUALIZATION:

Introduction to Excel, Absolute verses Relative Referencing, Merging and Looking Up Data, Decision making with IF Function, Preparing Data for Analysis, Pivot Tables Overview, creating a Pivot Table, Using Slicer to Filter Information, Visualizing Data with Charts.

Unit II **[9Hrs]**

ADVANCED EXCEL FUNCTIONS AND WHATIF ANALYSIS FOR DATA FORECASTING:

Forecasting sheet, forecasting function, conditional formatting, INDIRECT function, OFFSET Function, Histograms, Regression, WhatIf Analysis, Advanced Excel Outline, Function

Unit III **[9Hrs]**

ADVANCED EXCEL TECHNIQUES: LOOKUP, FINANCIAL, AND STATISTICAL FUNCTIONS:

Advanced Excel, Lookup Functions, Financial Functions and Terminology, Personal Financial Function, Principal and Interest Payment, Statistical Functions, Description and Forecasting, Connecting to External Data, Working with Tables, Pivot and Charts

Unit IV **[9Hrs]**

DATA ANALYSIS AND VISUALIZATION IN ADVANCED EXCEL:

Data Analysis, WhatIf, Graphs and Charts, Area Charts, Surface Charts, Radar Charts, Bubble Charts, Web App, Advanced Pivot Table Introduction, Grouping and Ungrouping, Reporting Layouts, Formatting Error Values and Empty Cells

Unit V **[9Hrs]**

ADVANCED PIVOTTABLE TECHNIQUES AND DASHBOARD CREATION:

Formatting, Custom Number Formatting, Value Field Setting, Sorting and Filtering, Interacting with Pivot Tables, Creating Calculated Field and Item, Solved order and List Formulas, GETPIVOTDATA, Pivot Charts, Graphical and Conditional Formatting, Dashboard

Text Books

S. No	Title	Authors	Edition	Publisher
1	Data Analytics: Principles, Tools and Practice	Dr. Gaurav Aroraa, Chitra lele	2nd Edition	BPB Publication
2	Data Analytics	Anil Maheshwari	3rd Edition	McGraw Hill

Reference Books

S. No	Title	Authors	Edition	Publisher
1	Excel 2023 for Dummies	Greg Harvey	5th	Wiley

		July 2025	1	**Applicable for 2025-26**
Chairman - BoS	**Dean – Academics**	**Date of Release**	**Version**	

FIRST SEMESTER

Course Code	Course Name	Th	Tu	Pr	Credits	Evaluation		
						CA	ESE	Total
25BD102T	Artificial Intelligence and Machine Learning	3	-	-	3	40	60	100

Course Objectives	Course Outcomes
This course is intended 1. Explore topics in Statistical Methods useful for Machine learning Algorithms. 2. Understand the Predictive Modelling Techniques. 3. Demonstrate various Machine Learning Algorithms with Various Python Supported Libraries. 4. Implement Recommender Systems with Collaborative Filtering Techniques. 5. Explore Neural Networks and Deep learning Methods.	**At the end of the Course, the Student will be able to:** 1. Describe Statistical Methods and Predictive Modelling techniques. 2. Discuss Various Machine Learning Algorithms with Various case studies. 3. Construct the Recommender Systems in the context of Collaborative Filtering Techniques. 4. Evaluate ML models and apply various best practices to get optimized results. 5. Demonstrate Deep Learning models and neural networks usage with TensorFlow.

Unit I **[9Hrs]**

INTRODUCTION TO STATISTICS AND PREDICTIVE MODELS: Python Basics, Running Python Scripts, Types of Data, Central Tendency: Mean, Median, and Mode, Variation and Standard Deviation, Probability Density Function, Probability Mass Function, Linear Regression, Polynomial Regression, Multi-Level Model

Unit II **[9Hrs]**

MACHINE LEARNING ALGORITHMS: Introduction to ML, Supervised and Unsupervised Learning, Train/Test, Classification and Regression, Bayesian Methods, Implementing a Spam Classifier with Naive Bayes, K-Means Clustering, Decision Trees, Ensemble Learning, Support Vector Machines (SVM)

Unit III **[9Hrs]**

RECOMMENDER SYSTEMS: Collaborative Filtering, Making Movie Recommendations to People, K-Nearest Neighbors, Using KNN to predict a rating for a movie, Dimensionality Reduction and Principal Component Analysis, Data Warehousing ETL and ELT, Dealing With Real-World Data, Bias/Variance Trade-off, K-Fold Cross Validation to Avoid Overfitting, Cleaning Web Log Data

Unit IV **[9Hrs]**

APACHE SPARK ML ON BIG DATA AND ML BEST PRACTICES: Installing Spark, Spark Introduction, Spark and the Resilient Distributed Dataset (RDD), Introducing MLLib, Decision Trees in Spark, K-Means Clustering in Spark, TF/IDF, Searching Wikipedia with Spark, Using the Spark 2.0 Data Frame API for MLLib.

Unit V **[9Hrs]**

INTRODUCTION TO DEEP LEARNING AND NEURAL NETWORKS: Deep Learning Overview, The History of Artificial Neural Networks, Deep Learning in the TensorFlow Playground, Deep Learning Details, Keras Framework, Using Keras to Predict Political Affiliations, Convolutional Neural Networks (CNN's), Using CNN's for handwriting recognition, Recurrent Neural Networks (RNN's), Using a RNN for sentiment analysis.

Text Books

S. No	Title	Authors	Edition	Publisher
1	Hands-On Data Science and Python Machine Learning	Kane	1st	Packt Publishing.
2	Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems	Géron. A	3rd	O'Reilly Media.

FIRST SEMESTER

Course Code	Course Name	Th	Tu	Pr	Credits	Evaluation		
						CA	ESE	Total
25BD103T	Data Engineering Foundation	3	-	-	3	40	60	100

Course Objectives	Course Outcomes
This course is intended to: 1. Understand the fundamentals of Databases and SQL 2. Learn manipulating data with SQL techniques. 3. Gain proficiency in advanced SQL statements. 4. Understand the document base NoSQL MongoDB 5. Explore the applications of MongoDB CRUD Operations.	**At the end of the Course, the Student will be able to:** 1. Understand the fundamentals of Databases and SQL 2. Apply the learned SQL techniques to manipulate data effectively. 3. Analyse complex SQL statements to demonstrate proficiency in writing and understanding them. 4. Evaluate the principles of document-based NoSQL databases, specifically MongoDB, and apply them appropriately. 5. Create applications implementing MongoDB CRUD operations.

Unit I **[9Hrs]**

INTRODUCTION TO DATABASES: Databases Introduction, Tables and Relations, Primary and Foreign Keys and Other Constraints, The ANSI SQL Standard DML, DDL and More Common Query Tools, The SELECT Statement, Limiting Output Columns, Formatting and Sorting Output, Column Aliases, The WHERE Clause, Creating a Filter Condition, Applying Multiple Filter Conditions

Unit II **[9Hrs]**

Manipulating Data with SQL: Counting Records, Common Aggregate Functions: SUM, AVG, MIN, MAX, Do's and Don'ts When Consolidating, Unions and Other Multiset Consolidations, Grouping Your Data, The HAVING Clause, Table Aliases, Table Joins, Complex Multi Table Joins.

Unit III **[9Hrs]**

Advanced SQL Techniques: Filtering Using Subqueries, The EXISTS Clause, Derived Tables, The INSERT, UPDATE, DELETE Statements, Transactions, Initiating a Transaction, The COMMIT and ROLLBACK Commands, Creating Tables, Creating and Using Views, Developing and Calling Stored Procedures, Implementing Triggers.

Unit IV **[9Hrs]**

INTRODUCTION TO MONGODB : MongoDB Overview, Installing MongoDB, Mongo Shell, Commands and Queries, Adding MongoDB to your Application, Local Node.js Server, Package Manager Preview, RESTful API, Creating Express Application and Setting up Cloud Host for MongoDB, Root Setup, Setting up a Controller, Schema Configuration and Creating a Schema, Custom Validators and Enums

Unit V **[9Hrs]**

MONGODB CRUD OPERATIONS: Basic Query, Semantic Naming Conventions, Dynamic Inserts, Array Query Operators, Dynamic Update and Delete, Adding New Collections, Embedded Documents, Sharding and Replicas, Unit Testing, Pagination

Text Books

S. No	Title	Authors	Edition	Publisher
1	Querying MySQL: Make your MySQL database analytics accessible with SQL operations, data extraction, and custom queries	Adam Aspin	2nd	BPB Publication
2	SQL in 10 Minutes, Sams Teach Yourself in 10 Minutes	Ben Forta	4th	Pearson

		July 2025	1	**Applicable for 2025-26**
Chairman - BoS	**Dean – Academics**	**Date of Release**	**Version**	

FIRST SEMESTER

Course Code	Course Name	Th	Tu	Pr	Credits	Evaluation		
						CA	ESE	Total
25BD104T(i)	PE - I Data Science for Engineers	3	-	-	3	40	60	100

Course Objectives	Course Outcomes
This course is intended 1. Introduce R as a programming language 2. Introduce the mathematical foundations required for data science 3. Introduce the first level data science algorithms 4. Introduce a data analytics problem solving framework 5. Introduce a practical capstone case study	**At the end of the Course, the Student will be able to:** 1. Describe a flow process for data science problems (Remembering) 2. Classify data science problems into standard typology (Comprehension) 3. Develop R codes for data science solutions (Application) 4. Correlate results to the solution approach followed (Analysis) 5. Assess the solution approach (Evaluation) 6. Construct use cases to validate approach and identify modifications required (Creating)

NPTEL Course: - **Data Science for Engineers**
By Prof. Ragunathan Rengasamy, Prof. Shankar Narasimhan | IIT Madras

Week 1: Course philosophy and introduction to R
Week 2: Linear algebra for data science
1. Algebraic view - vectors, matrices, product of matrix & vector, rank, null space, solution of over-determined set of equations and pseudo-inverse)
2. Geometric view - vectors, distance, projections, eigenvalue decomposition

Week 3: Statistics (descriptive statistics, notion of probability, distributions, mean, variance, covariance, covariance matrix, understanding univariate and multivariate normal distributions, introduction to hypothesis testing, confidence interval for estimates)
Week 4: Optimization
Week 5: 1. Optimization
2. Typology of data science problems and a solution framework
Week 6: 1. Simple linear regression and verifying assumptions used in linear regression
2. Multivariate linear regression, model assessment, assessing importance of different variables, subset selection
Week 7: Classification using logistic regression

Week 8: Classification using KNN and k-means clustering

Week 9 – Week 12 : Project related activities based on the above course

Text Books

S. No	Title	Authors	Edition	Publisher
1	Data Analytics: Principles, Tools and Practice	Dr. Gaurav Aroraa, Chitra lele	2nd Edition	BPB Publication
2	Data Analytics	Anil Maheshwari	3rd Edition	McGraw Hill

FIRST SEMESTER

Course Code	Course Name	Th	Tu	Pr	Credits	Evaluation		
						CA	ESE	Total
25BD104T(ii)	PE - I Big Data Computing	3	-	-	3	40	60	100

Course Objectives	Course Outcomes
This course is intended 1. Understand the Fundamentals of Big Data 2. Explore Big Data Frameworks and Technologies 3. Develop Skills in Data Processing and Analysis 4. Design Scalable and Fault-Tolerant Applications 5. Implement Real-Time Big Data Solutions	**At the end of the Course, the Student will be able to:** 1. Explain the characteristics and challenges of Big Data systems. 2. Demonstrate the use of Hadoop ecosystem components like HDFS, MapReduce, Hive, and Pig. 3. Apply Spark and other in-memory technologies for data analytics. 4. Design and implement scalable data pipelines using appropriate storage (HBase, MongoDB, Cassandra). 5. Analyze structured and unstructured data using distributed data processing techniques.

NPTEL Course: - **Big Data Computing**

By Prof. Rajiv Misra | IIT Patna

Week 1 : Introduction to Big Data

Week 2 : Introduction to Enabling Technologies for Big Data

Week 3 : Introduction to Big Data Platforms

Week 4 : Introduction to Big Data Storage Platforms for Large Scale Data Storage

Week 5 : Introduction to Big Data Streaming Platforms for Fast Data

Week 6 : Introduction to Big Data Applications (Machine Learning)

Week 7 : Introduction of Big data Machine learning with Spark

Week 8 : Introduction to Big Data Applications (Graph Processing)

Week 9 – Week 12: Project related activities based on the above course

Text Books

S. No	Title	Authors	Edition	Publisher
1	Analytics in a Big Data World: The Essential Guide to Data Science and its Applications	Bart Baesens	2nd Edition	Wiley

FIRST SEMESTER

Course Code	Course Name	Th	Tu	Pr	Credits	Evaluation		
						CA	ESE	Total
25BD105P	Technical Seminar-I	-	-	8	4	100	-	100

Course Objectives	Course Outcomes
This course is intended To provide an opportunity to the students to explore and deepen their knowledge in emerging technologies.	**Students will be able to** ● Enhance knowledge and critical thinking ● Improve presentation skills ● Learn new research skills and problem-solving technique. ● Increase their confidence

Student need to present seminar on various emerging technologies to explore and deepen their knowledge in a specific technical field, trends, and industry advancements, enabling them to enhance their expertise, critical thinking, and problem-solving skills within the chosen domain.

FIRST SEMESTER

Course Code	Course Name	Th	Tu	Pr	Credits	Evaluation		
						CA	ESE	Total
25BD106P	Mini Project-I	-	-	8	4	100	-	100

Course Objectives	Course Outcomes
This course is intended To provide an opportunity for the students to apply the knowledge, develop the skills and provide hands-on experience on a practical based project.	**Students will be able to** 1. Apply theoretical knowledge to address real-world problems, showcasing research, problem-solving, and critical thinking skills. 2. Effectively communicate project findings and demonstrate proficiency in project management and interdisciplinary learning. 3. Develop practical experience, ethical considerations, and the ability to adapt to challenges in a hands-on learning environment. 4. Develop the ability to document the progress, methodologies and results effectively in the form of project report.

Student need to build a project by ssuccessfully applying their academic knowledge to solve practical problems, demonstrating research, critical thinking, communication skills in a real-world context and prepare the project report.