PVP SIDDHARTHA INSTITUTE OF TECHNOLOGY

**LESSON PLAN**

**Academic Year :** 2024 - 2025

## Year & Semester : III B.Tech / I Sem

**Branch :** Computer Science & Engineering

**Subject Code & Name :** 20CS4501A & Data Science

**Name of Faculty :** A.Madhuri

|  |  |  |
| --- | --- | --- |
| **S. No.** | **Course Outcomes** | **Blooms Level** |
| **1** | Understand the life cycle process of data science. | **L2** |
| **2** | Apply different data pre-processing techniques for improving data quality. | **L3** |
| **3** | Apply Statistical methods to evaluate the data. | **L3** |
| **4** | Apply Statistical Learning techniques for model building. | **L3** |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **UNIT NO** | **TOPIC OF SYLLABUS**  **TO BE COVERED** | **LEARNING OUTCOMES** | **COURSE OUT COMES AND COGNITIVE LEVELS** | **TEACHING MODE**  **(WB/PPT)** | **HOURS REQUIRED** | | **CUMULATIVE HOURS** | **Expected date of Topic to be covered** | **REVIEW/**  **REMARKS**  **(BY HOD)** |
| **L** | **T** |
|  | **Learning Objectives of the Course, Importance of the Course, and Course Outcomes** | | | WB/PPT | 1 |  | 1 |  |  |
| 1 | **Introduction to Data Science**  What is Data Science? | What is Data?  What is Information?  What is Data Science?  Use cases of Data Science, Applications of Data Science | CO1-L2 | WB/PPT | 1 |  | 2 |  |  |
| 1 | **Phases of Data Science:** Data Acquisition, Cleansing | Understand the phases of data science or data science life cycle with examples   * Problem Understanding * Data Collection * Data Cleaning and preparation | CO1-L2 | WB/PPT | 1 |  | 3 |  |  |
| 1 | **Phases of Data Science:**  Exploratory Data Analysis, Data Preparation, Data Modeling | * Exploratory Data Analysis * Feature Engineering * Model Building * Model Evaluation * Data Visualization | CO1-L2 | WB/PPT | 1 |  | 4 |  |  |
| 1 | **Engineering Aspects of Data Science:**  Business Understanding, Data Understanding, Data Preparation,  Model Building,  Model Evaluation | Understand the Engineering Data Science Systems   * What is Systems Thinking? Examples * Components in Engineering Data Science Systems * CRISP-DM | CO1-L2 | WB/PPT | 1 |  | 5 |  |  |
| 1 | Hyper Parameter Optimization and Deployment | Understand Hyper Parameter Optimization   * Importance * Optimizer Hyperparameters * Model Hyperparameters * Types of Hyperparameters * Deployment   Sample Data Science Projects | CO1-L2 | WB/PPT | 1 |  | 6 |  |  |
| 2 | **Data Preprocessing:**  Introduction, Data Quality | Why preprocess the data?  Why is data dirty?  Why data preprocessing is important?  Understand the multi-dimensional measure of data quality, Major Tasks in data processing | CO1-L2 | WB/PPT | 1 |  | 7 |  |  |
| 2 | Data Cleaning- Missing Values, Noisy data | Understand about handling missing data and noisy data | CO1-L2 | WB/PPT | 1 |  | 8 |  |  |
| Apply Data Smoothing Techniques-   * Binning * Regression * Outlier Analysis * Examples   Data Cleaning as a process | CO2-L3 | WB/PPT | 2 |  | 10 |  |  |
| 2 | Data Integration | Apply data integration techniques for handling data with examples | CO2-L3 | WB/PPT | 2 |  | 12 |  |  |
| 2 | Data Reduction | Apply data reduction techniques on data   * Wavelet Transforms * Principal Components Analysis * Attribute Subset Selection * Histograms * Clustering * Sampling   Examples | CO2-L3 | WB/PPT | 3 |  | 15 |  |  |
| 2 | Data Transformation and Discretization | Apply data transformation techniques on data   * Smoothing * Attribute construction * Aggregation * Normalization   Examples | CO2-L3 | WB/PPT | 2 |  | 18 |  |  |
|  |  | Quiz – Unit-I & II |  |  | 1 |  | 19 |  |  |
| 3 | **Random Variables and Probability Distributions** | Understand about Random variables: discrete and continuous-Mathematical Expectation | CO1-L2 | WB/PPT | 1 |  | 20 |  |  |
| Understand Probability Density Function (PDF), Probability Mass Function (PMF), and Cumulative Density Function (CDF). | CO1-L2 | WB/PPT | 1 |  | 21 |  |  |
| 3 | Discrete distributions | Uniform, Binomial distribution, Bernoulli and Poisson distribution – Uses and applications | CO1-L2  CO3-L3 | WB/PPT | 3 |  | 23 |  |  |
| 3 | Continuous Distributions | Normal distribution, Standard Normal distribution, Student's T distribution, Chi-squared distribution – Uses and applications. | CO1-L2  CO3-L3 | WB/PPT | 3 |  | 26 |  |  |
| 3 | Sampling Strategies | Introduction, Simple Random sampling, Systematic sampling, Stratified sampling, Cluster sampling - applications | CO1-L2  CO3-L3 | WB/PPT | 2 |  | 28 |  |  |
| 4 | **Linear methods for Regression:**  Linear Regression Models | Understand about Linear Regression  Apply Linear Regression to build a model on datasets | CO1-L2  CO4-L3 | WB/PPT | 2 |  | 30 |  |  |
| 4 | Least Squares | Apply Least Squares to minimize the residual sum of squares | CO4-L3 | WB/PPT | 1 |  | 31 |  |  |
| 4 | Multiple Regression | Understand Multiple Regression  Apply Multiple Regression to build a model on datasets | CO1-L2  CO4-L3 | WB/PPT | 2 |  | 33 |  |  |
| 4 | Linear Discriminant Analysis | Understand Linear Discriminant Analysis Apply Linear Discriminant Analysis to build a model on datasets | CO1-L2  CO4-L3 | WB/PPT | 1 |  | 34 |  |  |
|  |  | Flip class - Linear Discriminant Analysis |  |  | 1 |  | 35 |  |  |
| 4 | Logistic Regression | Understand Logistic Regression  Apply Logistic Regression to build a model on datasets | CO1-L2  CO4-L3 | WB/PPT | 2 |  | 37 |  |  |
| 5 | **Model Assessment and Selection:** Bias, Variance and Model Complexity | Understand Model complexity using Bias and Variance | CO1-L2 | WB/PPT | 1 |  | 38 |  |  |
| 5 | Bias–Variance Decomposition | Understand Bias- Variance Tradeoff | CO1-L2 | WB/PPT | 1 |  | 39 |  |  |
| 5 | Model Parameters | Understand number of parameters and MLD of a model | CO1-L2 | WB/PPT | 2 |  | 41 |  |  |
| 5 | Optimism of the Training Error Rate | Apply fine-tuning concept on model parameters to Optimism of the Training Error Rate | C04-L3 | WB/PPT | 2 |  | 43 |  |  |
| 5 | Cross-Validation | Apply K-Fold Cross-Validation method for estimating prediction  error | CO4-L3 | WB/PPT | 2 |  | 45 |  |  |

**Legend**: Teaching Mode WB: White Board/ PPT: Power Point Presentation, Flip Class, Quiz

W

Signature of the faculty Signature of HOD