CLASS – XII : SEMESTER – IY

SUBJECT: ARTIFICIAL INTELLIGENCE AND DATA SCIENCE (AIDS)-THEORY

FULL MARKS: 35 **CONTACT HOURS: 60 Hours**

UNIT NO.	SUB UNIT	TOPICS	CONTACT HOURS	MARKS
Unit -6 Supervised Learning (10)	**6**	**Probabilistic classifier:** Basics of Bayesian Learning, Conditional independence, Naive Bayes classifier. Applications of Naive Bayes Classifier to sentiment classification task, add-one smoothing. **Instance based learning or lazy learning:** K-nearest neighbor classifier, curse of dimensionality	**20**	**10**
Unit -7 Unsupervised Learning (5)	**7**	What is unsupervised learning? Difference between supervised and unsupervised learning, What is clustering? Difference between clustering and classification, Why do we use clustering in an unsupervised learning technique? , Some examples of real world application of clustering, K-means clustering algorithm and its drawback. Simple use cases	**12**	**5**

Unit -8 Artificial Neural Network (17)	**8**	Biological motivation for Artificial Neural Networks(ANN), A simple mathematical model of a neuron (McCulloch and Pitts(1943)), Concept of activation function: threshold function and Sigmoid function, Perceptron as a linear classifier, perceptron training rule, Implementation of basic Boolean functions of two inputs using threshold perceptron, Equation of a linear separator in the input space, Representational power of perceptrons, Training unthresholded perceptron using Delta rule(with derivation) , What is the need for hidden layers ? - XOR example. Why do we need non-linearity in ANN?, Network structures: feed forward networks and recurrent networks **(basic concept only).** Training multiplayer feed-forward neural networks using Back propagation algorithm **(Concepts only and no derivation),** Generalization, overfitting, and stopping criterion, overcoming the overfitting problem using a set of validation data. An Illustrative example of an ANN architecture for handwritten digit recognition **(Only input representation, output representation and a block diagram of the network),** Need for automatic feature learning, difference between the conventional feed-forward neural networks and CNN, role of convolution layer and fully connected layer in CNN, function of pooling layer in CNN. An example of 2D convolution, a block diagram recognition task	**26**	**17**

Unit -9 Ethics in Large Language Models (LLM)	**9**	Brief introduction to ethics in Large Language Models(LLM), for example, GPT, ChatGPT.	**2**	**3**

NB : Additional 10 hours for Remedial and/or Tutorial classes

CLASS: XII

SUBJECT: ARTIFICIAL INTELLIGENCE AND DATA SCIENCE (AIDS)-PRACTICAL

FULL MARKS: 30 **CONTACT HOURS: 60 HOURS**

SL NO	TOPICS	CONTACT HOURS	MARKS
1. Data Visualization Techniques [Marks 5]			
1	Introduction to plotly library in python and plotting different types of plot using the library refer this *(https://plotly.com/python/plotly-express/)* 1D Histogram of four attributes of the IRIS dataset, 2D Histogram(considering the IRIS dataset, plot 2D histogram of petal length and width), Box Plots (Considering the IRIS dataset, show the Box plots of attributes for IRIS attributes and species) , Plot the Pie chart, showing the distribution of IRIS flowers (use IRIS dataset), Scatter Plots for each pair of attributes of the IRIS dataset, Heatmap, Word clouds, Geospatial maps.	8	5
2. Introduction to Machine Learning [6 Marks]			

2a	Introduction to python libraries like scipy and statsmodel to various basic codes Revisit matrix operations using scipy (basic matrix operations of addition, subtraction, multiplication, transpose), Using Scipy for advanced matrix operations - inverse. Write a python program to find the values of the parameters for the regression model with multiple features (using only formula for normal method).	6	3
2b	Generation of random (x, y) pairs where y = f(x) + d (d varies from -r to +r , a random value), f being a linear function, Linear regression or line fitting of the data, Optimizing the function using gradient descent, Plotting the steps using matplotlib	6	3
3. Supervised Learning [10 Marks]			
3a	● Building linear regression-based threshold classifier and testing the model on Diabetes Data set downloadable from UCI Machine Learning Repository ● Building Logistic regression model for binary classification of Diabetes Data. Vary learning rate and verify the impact of learning rate on classification performance. ● Introduction to the IRIS dataset, building a logistic regression for multi-class classification and testing the model on the IRIS dataset downloadable from UCI Machine Learning Repository ● Building K-nearest neighbor classifier and testing on the IRIS dataset downloadable from UCI Machine Learning Repository (Use Scikit-learn open source data analysis library for implementing the models)	18	7
3b	Building a naive Bayes classifier for sentiment analysis (Use Scikit-learn open source data analysis library)	6	3

4. Unsupervised Learning [3 Marks]			
4a	Using Scikit-learn library to use the K-means algorithm for clustering IRIS data and its visualization	6	3
5. Artificial Neural Network [6 marks]			
5a	• Using MLP from Scikit learn library, develop a handwritten digit recognition model using MLP and MNIST dataset • Using CNN from keras library, develop a handwritten digit recognition model using CNN and MNIST dataset • Compare the performance of the MLP based model and the CNN based model for the handwritten digit recognition task	10	6

NB : Additional 10 hours for Remedial and/or Tutorial classes