CODE	COURSE NAME	CATEGORY	L	T	P	CREDIT
EET474	MACHINE LEARNING	PEC	2	1	0	3

Preamble:. This course will enable students to:

1) Develop an appreciation for what is involved in learning models from data.
2) Understand a wide variety of learning algorithms.
3) Understand how to evaluate models generated from data.
4) Apply the algorithms to a real-world problem, optimize the models learned and report on the expected accuracy that can be achieved by applying the models.

Prerequisite: Nil

Course Outcomes: After the completion of the course the student will be able to:

CO1	Understand various basic learning techniques
CO2	Perform dimensionality reduction for multivariate problems
CO3	Implement machine learning solutions to classification, regression, and clustering problems
CO4	Use Perceptron modelling based learning techniques and Support Vector Machines to design solutions
CO5	Design and analyse machine learning experiments for real-life problems

Mapping of course outcomes with program outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7	PO 8	PO 9	PO 10	PO 11	PO 12
CO 1	3	3	-	-	-	-	-	-	-	-	-	3
CO 2	3	3	2	-	-	-	-	-	-	-	-	3
CO 3	3	3	3	-	-	-	-	-	-	-	-	3
CO 4	3	3	-	-	-	-	-	-	-	-	-	3
CO 5	3	3	2	-	-	-	-	-	-	-	-	3

Assessment Pattern:

Bloom's Category	Continuous Assessment Tests		End Semester Examination
	1	2	
Remember (K1)	10	10	20
Understand (K2)	10	10	20
Apply (K3)	20	20	50
Analyse (K4)	10	10	10
Evaluate (K5)			
Create (K6)			

Total Marks	CIE marks	ESE marks	ESE Duration
150	50	100	03 Hrs

End Semester Examination Pattern: There will be two parts; Part A and Part B.
Part A contains 10 questions with 2 questions from each module, having 3 marks for each question. Students should answer all questions.
Part B contains 2 questions from each module of which student should answer any one. Each question carries 14 marks and can have maximum 2 sub-divisions.

Course Level Assessment Questions

Course Outcome 1 (CO1):

1. Distinguish between overfitting and underfitting. How it can affect model generalization?

2. Explain bias- variance dilemma.

3. Distinguish between classification and regression with an example.

Course Outcome 2 (CO2)

1. Define VC dimension. Show that an axis aligned rectangle can shatter 4 points in 2 dimensions.

2. Compare Simple Regression, Multiple Regression and Multivariate Regression.

3. Describe any two techniques used for Ensemble Learning.

Course Outcome 3(CO3):

1. Given a linearly separable dataset with one group containing 5 instances and a second group containing 20 instances, is k-means clustering with k =2 guaranteed to find these two clusters? Explain why or why not.

2. Explain Basic decision tree learning algorithm for classification problems

3. Draw the decision tree structure for X1 XOR X2

Course Outcome 4 (CO4):

1. What is kernel trick? Why does the kernel trick allow us to solve SVMs with high dimensional feature spaces, without significantly increasing the running time?

2. Can you represent the following Boolean function with a single binary perceptron? If yes, show the weights. If not, explain why not in 1-2 sentences.

A	B	f(A,B)
1	1	0
0	0	0
1	0	1
0	1	0

3. Formulate the SVM regression problem using insensitive loss.

Course Outcome 5 (CO5):

1) Suppose that the datamining task is to cluster the following seven points (with (x,y) representing location) into two clusters A1(1,1), A2(1.5,2), A3(3,4), A4(5,7), A5(3.5,5), A6(4.5,5), A7(3.5,4.5) The distance function is City block distance. Suppose initially we assign A1,A5 as the centre for each cluster respectively. Using the K-means algorithm to find the three clusters and their centres after two round of execution.

2) Explain the concept of Reinforcement Learning with a practical example.

3) Draw the structure of CNN, and explain the classification process with an example.

Model Question Paper PAGES: 3

QP CODE:

Reg.No:____________

Name:______________

APJ ABDULKALAM TECHNOLOGICAL UNIVERSITY
EIGHTH SEMESTER B.TECH DEGREE
EXAMINATION MONTH & YEAR

Course Code**: EET474**

Course Name: **MACHINE LEARNING**

Max. Marks: 100 **Duration: 3 Hours**

PART A

Answer all Questions. Each question carries 3 Marks

1	Explain false negative, false positive, true negative and true positive with a simple example.
2	While using all features of a data set, if we achieve 100% accuracy on my training set, but ~70% on validation set, discuss whether we might see an underfitting, overfitting or perfect model? Please justify.
3	Differentiate a Perceptron and Logistic Regression?
4	Explain the difference between L1 and L2 regularization.
5	Can we design a neural network without an activation function? Justify your answer.
6	Is Occam's Razor an inductive bias scenario? State reasons with examples.
7	What are the standard use cases for Bayesian belief networks? What is its basic difference with respect to Hidden Markov Models?
8	We have designed an RBF kernel in SVM with high Gamma value. What does this signify?
9	In a binary classification problem, there are 3 models each with 70% accuracy. If we want to ensemble these models using majority voting method, what will be the maximum accuracy we can get?
10	What are the basic elements of reinforcement learning?

PART B

Answer any one full question from each module. Each question carries 14 Marks

Module 1

11 a) Discuss the influence of model complexity on underfitting and overfitting? (7 Marks)

b) How do we measure the power of a classifier? What is the VC dimension for a linear classifier? (7 Marks)

12 a) List out the critical assumptions for applying linear regression, with emphasis to Heteroscedasticity. How can we improve the accuracy of a linear regression model? (9 Marks)

b) Discuss L1 and L2 regularization? (5 Marks)

Module 2

13 a) Explain Naïve Bayes Classifier (10 Marks)

b) Discuss the inconsistencies in Bayesian inference (4 Marks)

14 a) What are the various multivariate learning techniques? Discuss with use cases and applications (7 Marks)

b) Suppose we have 3 cards identical in form except that both sides of the first card are colored red, both sides of the second card are colored black, and one side of the third card is colored red and the other side is colored black. The 3 cards are mixed up in a hat, and 1 card is randomly selected and put down on the ground. If the upper side of the chosen card is colored red, what is the probability that the other side is colored black? (7 Marks)

Module 3

15 a) Consider the following data where x and y are the 2 input variables and Class is the dependent variable. (10 Marks)

x	y	Class
−1	1	−
0	1	+
0	2	−
1	−1	−
1	0	+
1	2	+
2	2	−
2	3	+

Draw the scatter plot for this dataset in a two dimensional space. Assuming a Euclidian distance of in 3-NN, to which class will the new point of x=1 and y=1 belong to?

b) Write four termination conditions for k-means clustering algorithm (4 Marks)

16 a) Describe the expectation-maximization algorithm? (9 Marks)

b) Write short note on Random Forest Decision tree . (5 Marks)

Module 4

17 a) Write the pseudo code for back propagation algorithm and explain? (10 Marks)

b) Differentiate CNN from RNN with respect to its use cases. (4 Marks)

18 a) Discuss the geometric intuition behind SVMs.

Discuss soft margin and hard margin SVMs (10 Marks)

b) When do you apply "Kernel Trick"? (4 Marks)

Module 5

19 a) In an election, N candidates are competing against each other and people are voting for either of the candidates. Voters don't communicate with each other while casting their votes. Which ensemble method works similar to above-discussed election procedure? (11 Marks)

b) Illustrate K-Arm bandit algorithm with an example (3 Marks)

20 a) Discuss problem characteristics in the Reinforcement Learning method (5 Marks)

b) With an example, demonstrate the Q-Function and Q-Learning algorithm, assuming deterministic reward and action. (9 Marks)

Syllabus

Module – 1

Introduction: What Is Machine Learning? Examples of Machine Learning Applications, Learning Associations, Classification, Regression, Unsupervised Learning, Reinforcement Learning

Supervised Learning: Learning a Class from Examples, Vapnik-Chervonenkis (VC) Dimension, Noise, Learning Multiple Classes, Regression, Model Selection and Generalization

Parametric Methods: Maximum Likelihood Estimation, Evaluating an Estimator: Bias and Variance, Parametric Classification, Regression, Tuning Model Complexity and Model Validation: Bias/Variance Dilemma

Module – 2

Bayesian Learning: Introduction to conditional probability and conditional expectations, Bayes theorem, Bayes theorem and concept learning, ML and LS error hypothesis, ML for predicting probabilities, Naive Bayes classifier, Bayesian belief networks,

Multivariate Data, Multivariate Classification, Multivariate Regression

Module – 3

Clustering: Introduction, Mixture Densities, k-Means Clustering, Expectation-Maximization Algorithm, Other methods of clustering.

Nonparametric Methods: Nonparametric Density Estimation, Histogram Estimator, Kernel Estimator, k-Nearest Neighbor Estimator

Decision Tree Based Learning: Decision tree representation, Appropriate problems for decision tree learning, Basic decision tree learning algorithm, hypothesis space search in decision tree learning, Inductive bias in decision tree learning, Issues in decision tree learning

Module – 4

Neural Networks: Neural Networks as a Paradigm for Parallel Processing, Feed Forward Networks, Backpropagation Algorithm, Fundamentals of Deep Learning, Basic Deep Learning Architectures

Local Models: Competitive Learning, Radial Basis Functions, Incorporating Rule-Based Knowledge

Kernel Machines: SVM Formulations, Optimal Separating Hyperplane, The Nonseparable Case: Soft Margin Hyperplane, ν-SVM, Kernel Types, Kernel Machines for Regression

Module – 5

Combining Multiple Learners: Rationale, Generating Diverse Learners, Model Combination Schemes, Voting, Error-Correcting Output Codes, Bagging, Boosting

Reinforcement Learning: The State Space Theory, K-Armed Bandit, Elements of Reinforcement Learning, Q Learning

Text Books:

- Pattern Recognition and Machine Learning. Christopher Bishop. Springer. 2006. [CB-2006]
- Machine Learning. Tom Mitchell, McGraw-hill, 1997

Reference Books:

- Understanding Machine Learning. Shai Shalev-Shwartz and Shai Ben-David. Cambridge University Press. 2017. [SS-2017]
- Haykin, Simon. Neural networks and learning machines, 3/E. Pearson Education India, 2010.
- The Elements of Statistical Learning. Trevor Hastie, Robert Tibshirani and Jerome Friedman. Second Edition. 2009. [TH-2009]
- Foundations of Data Science. Avrim Blum, John Hopcroft and Ravindran Kannan. January 2017. [AB-2017]

Course Contents and Lecture Schedule:

No	Topic	No. of Lectures
1	**Module 1**	**(7 hours)**
1.1	What Is Machine Learning? Examples of Machine Learning Applications, Learning Associations, Classification, Regression, Unsupervised Learning, Reinforcement Learning	2
1.2	Supervised Learning: Learning a Class from Examples, Vapnik-Chervonenkis (VC) Dimension, Noise, Learning Multiple Classes, Regression, Model Selection and Generalization	2
1.3	Parametric Methods: Maximum Likelihood Estimation, Evaluating an Estimator: Bias and Variance, Parametric Classification, Regression, Tuning Model Complexity and Model Validation: Bias/Variance Dilemma	3
2	**Module 2**	**(7 hours)**
2.1	Bayesian Learning: Introduction to conditional probability and conditional expectations, Bayes theorem, Bayes theorem and concept learning, ML and LS error hypothesis,	3
2.2	ML for predicting probabilities, Naive Bayes classifier, Bayesian belief networks,	2
2.3	Multivariate Data, Multivariate Classification, Multivariate Regression	2
3	**Module 3**	**(7 hours)**
3.1	Clustering: Introduction, Mixture Densities, k-Means Clustering, Expectation-Maximization Algorithm, Other methods of clustering.	2
3.2	Nonparametric Methods: Nonparametric Density Estimation, Histogram Estimator, Kernel Estimator, k-Nearest Neighbor Estimator	2
3.3	Decision Tree Based Learning: Decision tree representation, Appropriate problems for decision tree learning, Basic decision tree learning algorithm, hypothesis space search in decision tree learning, Inductive bias in decision tree learning, Issues in decision tree learning	3
4	**Module 4**	**(7 hours)**
4.1	Neural Networks: Neural Networks as a Paradigm for Parallel Processing, Feed Forward Networks, Backpropagation Algorithm, Fundamentals of Deep Learning, Basic Deep Learning Architectures	2
4.2	Local Models: Competitive Learning, Radial Basis Functions, Incorporating Rule-Based Knowledge	2
4.3	Kernel Machines: SVM Formulations, Optimal Separating Hyperplane, The Nonseparable Case: Soft Margin Hyperplane, ν-SVM, Kernel Types, Kernel Machines for Regression	3
5	**Module 5**	**(7 hours)**
5.1	Combining Multiple Learners: Rationale, Generating Diverse Learners	2
5.2	Model Combination Schemes, Voting, Error-Correcting Output Codes, Bagging, Boosting	2
5.3	Reinforcement Learning: The State Space Theory, K-Armed Bandit, Elements of Reinforcement Learning, Q Learning	3

-------------- x ------------