

MACHINE LEARNING
[As per Choice Based Credit System (CBCS) scheme]
(Effective from the academic year 2016 -2017)
SEMESTER – VII

Subject Code	15CS73	IA Marks	20
Number of Lecture Hours/Week	03	Exam Marks	80
Total Number of Lecture Hours	50	Exam Hours	03

CREDITS – 04

Course Objectives: This course will enable students to

- Define machine learning and problems relevant to machine learning.
- Differentiate supervised, unsupervised and reinforcement learning
- Apply neural networks, Bayes classifier and k nearest neighbor, for problems appear in machine learning.
- Perform statistical analysis of machine learning techniques.

Module – 1	Teaching Hours
-------------------	-----------------------

<p>Introduction: Well posed learning problems, Designing a Learning system, Perspective and Issues in Machine Learning.</p> <p>Concept Learning: Concept learning task, Concept learning as search, Find-S algorithm, Version space, Candidate Elimination algorithm, Inductive Bias.</p> <p>Text Book1, Sections: 1.1 – 1.3, 2.1-2.5, 2.7</p>	10 Hours
---	-----------------

Module – 2	
-------------------	--

<p>Decision Tree 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.</p> <p>Text Book1, Sections: 3.1-3.7</p>	10 Hours
---	-----------------

Module – 3	
-------------------	--

<p>Artificial Neural Networks: Introduction, Neural Network representation, Appropriate problems, Perceptrons, Backpropagation algorithm.</p> <p>Text book 1, Sections: 4.1 – 4.6</p>	08 Hours
---	-----------------

Module – 4	
-------------------	--

<p>Bayesian Learning: Introduction, Bayes theorem, Bayes theorem and concept learning, ML and LS error hypothesis, ML for predicting probabilities, MDL principle, Naive Bayes classifier, Bayesian belief networks, EM algorithm</p> <p>Text book 1, Sections: 6.1 – 6.6, 6.9, 6.11, 6.12</p>	10 Hours
--	-----------------

Module – 5	
-------------------	--

<p>Evaluating Hypothesis: Motivation, Estimating hypothesis accuracy, Basics of sampling theorem, General approach for deriving confidence intervals, Difference in error of two hypothesis, Comparing learning algorithms.</p> <p>Instance Based Learning: Introduction, k-nearest neighbor learning, locally weighted regression, radial basis function, cased-based reasoning,</p> <p>Reinforcement Learning: Introduction, Learning Task, Q Learning</p> <p>Text book 1, Sections: 5.1-5.6, 8.1-8.5, 13.1-13.3</p>	12 Hours
--	-----------------

Course Outcomes:After studying this course, students will be able to

- Identify the problems for machine learning. And select the either supervised,

unsupervised or reinforcement learning.

- Explain theory of probability and statistics related to machine learning
- Investigate concept learning, ANN, Bayes classifier, k nearest neighbor, Q,

Question paper pattern:

The question paper will have ten questions.

There will be 2 questions from each module.

Each question will have questions covering all the topics under a module.

The students will have to answer 5 full questions, selecting one full question from each module.

Text Books:

1. Tom M. Mitchell, Machine Learning, India Edition 2013, McGraw Hill Education.

Reference Books:

1. Trevor Hastie, Robert Tibshirani, Jerome Friedman, h The Elements of Statistical Learning, 2nd edition, springer series in statistics.
2. Ethem Alpaydm, Introduction to machine learning, second edition, MIT press.

MACHINE LEARNING LABORATORY
[As per Choice Based Credit System (CBCS) scheme]
(Effective from the academic year 2016 -2017)
SEMESTER – VII

Subject Code	15CSL76	IA Marks	20
Number of Lecture Hours/Week	01I + 02P	Exam Marks	80
Total Number of Lecture Hours	40	Exam Hours	03

CREDITS – 02

Course objectives: This course will enable students to

1. Make use of Data sets in implementing the machine learning algorithms
2. Implement the machine learning concepts and algorithms in any suitable language of choice.

Description (If any):

1. The programs can be implemented in either JAVA or Python.
2. For Problems 1 to 6 and 10, programs are to be developed without using the built-in classes or APIs of Java/Python.
3. Data sets can be taken from standard repositories (<https://archive.ics.uci.edu/ml/datasets.html>) or constructed by the students.

Lab Experiments:

1. Implement and demonstrate the **FIND-S algorithm** for finding the most specific hypothesis based on a given set of training data samples. Read the training data from a .CSV file.
2. For a given set of training data examples stored in a .CSV file, implement and demonstrate the **Candidate-Elimination algorithm** to output a description of the set of all hypotheses consistent with the training examples.
3. Write a program to demonstrate the working of the decision tree based **ID3 algorithm**. Use an appropriate data set for building the decision tree and apply this knowledge to classify a new sample.
4. Build an Artificial Neural Network by implementing the **Backpropagation algorithm** and test the same using appropriate data sets.
5. Write a program to implement the **naïve Bayesian classifier** for a sample training data set stored as a .CSV file. Compute the accuracy of the classifier, considering few test data sets.
6. Assuming a set of documents that need to be classified, use the **naïve Bayesian Classifier** model to perform this task. Built-in Java classes/API can be used to write the program. Calculate the accuracy, precision, and recall for your data set.
7. Write a program to construct a **Bayesian network** considering medical data. Use this model to demonstrate the diagnosis of heart patients using standard Heart Disease Data Set. You can use Java/Python ML library classes/API.
8. Apply **EM algorithm** to cluster a set of data stored in a .CSV file. Use the same data set for clustering using **k-Means algorithm**. Compare the results of these two algorithms and comment on the quality of clustering. You can add Java/Python ML library classes/API in the program.
9. Write a program to implement **k-Nearest Neighbour algorithm** to classify the iris data set. Print both correct and wrong predictions. Java/Python ML library classes can be used for this problem.
10. Implement the non-parametric **Locally Weighted Regression algorithm** in order to fit data points. Select appropriate data set for your experiment and draw graphs.

Study Experiment / Project:
NIL
Course outcomes: The students should be able to:
<ol style="list-style-type: none">1. Understand the implementation procedures for the machine learning algorithms.2. Design Java/Python programs for various Learning algorithms.3. Apply appropriate data sets to the Machine Learning algorithms.4. Identify and apply Machine Learning algorithms to solve real world problems.
Conduction of Practical Examination:
<ul style="list-style-type: none">• All laboratory experiments are to be included for practical examination.• Students are allowed to pick one experiment from the lot.• Strictly follow the instructions as printed on the cover page of answer script• Marks distribution: Procedure + Conduction + Viva: 20 + 50 +10 (80)
Change of experiment is allowed only once and marks allotted to the procedure part to be made zero.