

Course code	Course Name	L-T-P Credits	Year of Introduction
CS467	MACHINE LEARNING	3-0-0-3	2016

**Course Objectives:**

- To introduce the prominent methods for machine learning
- To study the basics of supervised and unsupervised learning
- To study the basics of connectionist and other architectures

**Syllabus:**

Introduction to Machine Learning, Learning in Artificial Neural Networks, Decision trees, HMM, SVM, and other Supervised and Unsupervised learning methods.

**Expected Outcome:**

The Students will be able to :

- differentiate various learning approaches, and to interpret the concepts of supervised learning
- compare the different dimensionality reduction techniques
- apply theoretical foundations of decision trees to identify best split and Bayesian classifier to label data points
- illustrate the working of classifier models like SVM, Neural Networks and identify classifier model for typical machine learning applications
- identify the state sequence and evaluate a sequence emission probability from a given HMM
- illustrate and apply clustering algorithms and identify its applicability in real life problems

**References:**

1. Christopher M. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006.
2. Ethem Alpaydm, *Introduction to Machine Learning (Adaptive Computation and Machine Learning)*, MIT Press, 2004.
3. Margaret H. Dunham. *Data Mining: introductory and Advanced Topics*, Pearson, 2006
4. Mitchell. T, *Machine Learning*, McGraw Hill.
5. Ryszard S. Michalski, Jaime G. Carbonell, and Tom M. Mitchell, *Machine Learning : An Artificial Intelligence Approach*, Tioga Publishing Company.

**Course Plan**

Module	Contents	Hours	End Sem. Exam Marks %
I	Introduction to Machine Learning, Examples of Machine Learning applications - Learning associations, Classification, Regression, Unsupervised Learning, Reinforcement Learning. Supervised learning- Input representation, Hypothesis class, Version space, Vapnik-Chervonenkis (VC) Dimension	6	15

II	Probably Approximately Learning (PAC), Noise, Learning Multiple classes, Model Selection and Generalization, Dimensionality reduction- Subset selection, Principle Component Analysis	8	15
<b>FIRST INTERNAL EXAM</b>			
III	Classification- Cross validation and re-sampling methods- K-fold cross validation, Boot strapping, Measuring classifier performance- Precision, recall, ROC curves. Bayes Theorem, Bayesian classifier, Maximum Likelihood estimation, Density functions, Regression	8	20
IV	Decision Trees- Entropy, Information Gain, Tree construction, ID3, Issues in Decision Tree learning- Avoiding Over-fitting, Reduced Error Pruning, The problem of Missing Attributes, Gain Ratio, Classification by Regression (CART), Neural Networks- The Perceptron, Activation Functions, Training Feed Forward Network by Back Propagation.	6	15
<b>SECOND INTERNAL EXAM</b>			
V	Kernel Machines- Support Vector Machine- Optimal Separating hyper plane, Soft-margin hyperplane, Kernel trick, Kernel functions. Discrete Markov Processes, Hidden Markov models, Three basic problems of HMMs- Evaluation problem, finding state sequence, Learning model parameters. Combining multiple learners, Ways to achieve diversity, Model combination schemes, Voting, Bagging, Booting	8	20
VI	Unsupervised Learning - Clustering Methods - K-means, Expectation-Maximization Algorithm, Hierarchical Clustering Methods , Density based clustering	6	15
<b>END SEMESTER EXAM</b>			

### Question Paper Pattern

1. There will be **FOUR** parts in the question paper – **A, B, C, D**
2. **Part A**
  - a. **Total marks : 40**
  - b. **TEN** questions, each have **4 marks**, covering **all the SIX modules** (**THREE** questions from **modules I & II**; **THREE** questions from **modules III & IV**; **FOUR** questions from **modules V & VI**).  
*All the TEN* questions have to be answered.
3. **Part B**
  - a. **Total marks : 18**
  - b. **THREE** questions, each having **9 marks**. One question is from **module I**; one question is from **module II**; one question *uniformly* covers **modules I & II**.
  - c. **Any TWO** questions have to be answered.
  - d. Each question can have *maximum THREE* subparts.

4. Part C

a. Total marks : 18

b. *THREE* questions, each having **9 marks**. One question is from **module III**; one question is from **module IV**; one question *uniformly* covers **modules III & IV**.

c. *Any TWO* questions have to be answered.

d. Each question can have *maximum THREE* subparts.

5. Part D

a. Total marks : 24

b. *THREE* questions, each having **12 marks**. One question is from **module V**; one question is from **module VI**; one question *uniformly* covers **modules V & VI**.

c. *Any TWO* questions have to be answered.

d. Each question can have *maximum THREE* subparts.

6. There will be **AT LEAST 60%** analytical/numerical questions in all possible combinations of question choices.

