Devi Ahilya University, Indore, India Institute of Engineering & Technology		BE IV Year Computer Engineering					
Subject Code & Name 4CO205 Machine Learning	**Periods Hours/ Week** Lectures- 4 Tutorials-0 Practical- 0	**Marks**	TH	CW	SW	Pr	Total
		Max	100	50	-	-	150
Duration of Theory Paper:3 Hours		Min	35	25	-	-	60

Course Objectives: The field of machine learning is concerned with the question of how to build computer programs able to construct new knowledge or to improve already possessed knowledge by using input information. The goal of this course is to introduce the theoretical foundations of machine learning, to provide practical experience of applying machine learning techniques and to investigate new problems where machine learning techniques can do better.

Prerequisite: Basic knowledge of a programming language is required. Basic knowledge of probabilities and statistics is required

Course of Contents

Unit-I

Introduction: Definition of learning systems. Goals and applications of machine learning. Aspects of developing a learning system: training data, concept representation, function approximation.

Inductive Classification: The concept learning task. Concept learning as search through a hypothesis space. General-to-specific ordering of hypotheses. Finding maximally specific hypotheses. Version spaces and the candidate elimination algorithm.

Unit-II

Decision Tree Learning: Representing concepts as decision trees. Recursive induction of decision trees. Picking the best splitting attribute: entropy and information gain. Searching for simple trees and computational complexity.

Experimental Evaluation of Learning Algorithms: Measuring the accuracy of learned hypotheses. Comparing learning algorithms: cross-validation, learning curves, and statistical hypothesis testing.

Unit-III

Computational Learning Theory: Models of learnability: learning in the limit; probably approximately correct (PAC) learning. Sample complexity: quantifying the number of examples needed to PAC learn. Computational complexity of training. Sample complexity for finite hypothesis spaces. PAC results for learning conjunctions, kDNF, and kCNF. Sample complexity for infinite hypothesis spaces, Vapnik-Chervonenkis dimension.

Unit-IV

Language Learning: Classification problems in language: word-sense disambiguation, sequence labeling. Formal Language learning, introduction to Hidden Markov models (HMM's).

Support Vector Machines: Maximum margin linear separators. Quadratic programming solution to finding maximum margin separators. Kernels for learning non-linear functions.

Unit-V

Genetic Algorithms (GAs): Motivation, Representing Hypotheses, Genetic operators, fitness Function and Selection. How do GAs work? Machine Learning: The Michigan Approach, The Pitt Approach, An evolution program: the GIL system. Evolutionary Programming and Genetic Programming.

Books Recommended:

[1] Tom Mitchell, *Machine Learning*, McGraw-Hill, 1997.

[2] Richard O. Duda, Peter E. Hart & David G. Stork, *Pattern Classification,* Wiley & Sons, 2001.

[3] Ethem Alpaydin, *Introduction to Machine Learning*, MIT Press, 2004.

[4] David E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning,* Kluwer Academic Publishers, Boston, MA, 1989.

[5] Zbigniew Michalewicz, *Genetic Algorithms + Data Structures = Evolution Programs*, Springer, 1999.