Machine Learning II		Semester	7
Course Code	**BAI702**	CIE Marks	50
Teaching Hours/Week (L:T:P: S)	3:0:2:0	SEE Marks	50
Total Hours of Pedagogy	40 hours Theory + 8-10 Lab slots	Total Marks	100
Credits	04	Exam Hours	3
Examination nature (SEE)	Theory/practical		

Course objectives:

- To introduce concept learning and General to specific learning
- To learn set of rules using Sequential Covering approach
- To make decisions on decision by committee
- To understand similarities using unsupervised learning .
- To understand Markov Chain Monte Carlo (MCMC) and Graphical Methods

Teaching-Learning Process (General Instructions)

These are sample Strategies; that teachers can use to accelerate the attainment of the various course outcomes.

1. Lecturer method (L) need not to be only a traditional lecture method, but alternative effective teaching methods could be adopted to attain the outcomes.
2. Use of Video/Animation to explain functioning of various concepts.
3. Encourage collaborative (Group Learning) Learning in the class.
4. Ask at least three HOT (Higher order Thinking) questions in the class, which promotes critical thinking.
5. Adopt Problem Based Learning (PBL), which fosters students' Analytical skills, develop design thinking skills such as the ability to design, evaluate, generalize, and analyze information rather than simply recall it.
6. Introduce Topics in manifold representations.
7. Show the different ways to solve the same problem with different logic and encourage the students to come up with their own creative ways to solve them.
8. Discuss how every concept can be applied to the real world - and when that's possible, it helps improve the students' understanding

MODULE-1

Introduction: Well-Posed Learning Problems, Designing a Learning System, Perspectives and Issues in Machine Learning.

Concept Learning and the General-to-Specific Ordering: A Concept Learning Task, Concept Learning as Search, Find-S: Finding a Maximally Specific Hypothesis, Version Spaces and the Candidate-Elimination Algorithm, Remarks on Version Spaces and Candidate-Elimination, Inductive Bias.

Text Book 1 : Ch 1 & 2

MODULE-2

Learning Sets of Rules: Sequential Covering Algorithms, Learning Rule Sets: Example-Based Methods, Learning First-Order Rules, FOIL: A First-Order Inductive Learner.

Analytical Learning: Perfect Domain Theories: Explanation-Based Learning, Explanation-Based Learning of Search Control Knowledge, Inductive-Analytical Approaches to Learning.

Text Book 1 : Ch 10 & 11

MODULE-3

Decision by Committee: Ensemble Learning: Boosting: Adaboost , Stumping, Bagging: Subagging, Random Forests, Comparison With Boosting, Different Ways To Combine Classifiers.

Unsupervised Learning: The K-MEANS algorithm : Dealing with Noise ,The k-Means Neural Network , Normalisation ,A Better Weight Update Rule ,Using Competitive Learning for Clustering.

Text Book 2: Chap 13 and 14.1

MODULE-4

Unsupervised Learning: Vector Quantisation**,** the self-organising feature map , The SOM Algorithm, Neighbourhood Connections, Self-Organisation, Network Dimensionality and Boundary Conditions, Examples of Using the SOM. **Markov Chain Monte Carlo (MCMC) Methods:** Sampling : Random Numbers ,Gaussian Random Numbers ,Monte Carlo Or Bust ,The Proposal Distribution , Markov Chain Monte Carlo. **Text Book 2: Chap 14.2, 14.3, 15**
MODULE-5
Graphical Models: Bayesian Networks : Approximate Inference , Making Bayesian Networks , Markov Random Fields , Hidden Markov Models (Hmms), The Forward Algorithm , The Viterbi Algorithm , The Baum–Welch Or Forward–Backward Algorithm , Tracking Methods , The Kalman Filter, The Particle Filter. **Text Book 2 : Chap 16**

PRACTICAL COMPONENT OF IPCC *(May cover all / major modules)*

Sl.NO	Experiments
1	Read a dataset from the user and i. Use the Find-S algorithm to find the most specific hypothesis that is consistent with the positive examples. Ii. What is the final hypothesis after processing all the positive examples? Using the same dataset, apply the Candidate Elimination algorithm. Determine the final version space after processing all examples (both positive and negative). What are the most specific and most general hypotheses in the version space?
2	Read a dataset and use an **example-based method** (such as **RIPPER** or **CN2**) to generate a set of classification rules . Apply the **FOIL algorithm** (First-Order Inductive Learner) to learn first-order rules for predicting.
3	Read a supervised dataset and use bagging and boosting technique to classify the dataset. Indicate the performance of the model.
4	Read an unsupervised dataset and group the dataset based on similarity based on k-means clustering .
5	Read a dataset and perform unsupervised learning using SOM algorithm.
6	Write a function to generate **uniform random numbers** in the interval [0, 1]. Use this function to generate 10 random samples and evaluate f(x) for each sample. What are the sampled function values? Using the samples generated in the previous step, estimate the integral I using the **Monte Carlo method**.
7	Read a dataset and indicate the likelihood of an event occurring using Bayesian Networks.
8	Refer to the dataset in question 7 and indicate inferences based on the sequence of steps .

Course outcomes (Course Skill Set):

At the end of the course, the student will be able to:

1. Apply concept learning and General to specific learning
2. Design models to classify supervised data .
3. To analyze methods to identify similarities using unsupervised learning .
4. To understand Markov Chain Monte Carlo (MCMC) and Graphical Methods.

Assessment Details (both CIE and SEE)

The weightage of Continuous Internal Evaluation (CIE) is 50% and for Semester End Exam (SEE) is 50%. The minimum passing mark for the CIE is 40% of the maximum marks (20 marks out of 50) and for the SEE minimum passing mark is 35% of the maximum marks (18 out of 50 marks). A student shall be deemed to have satisfied the academic requirements and earned the credits allotted to each subject/ course if the student secures a minimum of 40% (40 marks out of 100) in the sum total of the CIE (Continuous Internal Evaluation) and SEE (Semester End Examination) taken together.

CIE for the theory component of the IPCC (maximum marks 50)

- IPCC means practical portion integrated with the theory of the course.
- CIE marks for the theory component are **25 marks** and that for the practical component is **25 marks**.
- 25 marks for the theory component are split into **15 marks** for two Internal Assessment Tests (Two Tests, each of 15 Marks with 01-hour duration, are to be conducted) and **10 marks** for other assessment methods mentioned in 22OB4.2. The first test at the end of 40-50% coverage of the syllabus and the second test after covering 85-90% of the syllabus.
- Scaled-down marks of the sum of two tests and other assessment methods will be CIE marks for the theory component of IPCC (that is for **25 marks)**.
- The student has to secure 40% of 25 marks to qualify in the CIE of the theory component of IPCC.

CIE for the practical component of the IPCC

- **15 marks** for the conduction of the experiment and preparation of laboratory record, and **10 marks** for the test to be conducted after the completion of all the laboratory sessions.
- On completion of every experiment/program in the laboratory, the students shall be evaluated including viva-voce and marks shall be awarded on the same day.
- The CIE marks awarded in the case of the Practical component shall be based on the continuous evaluation of the laboratory report. Each experiment report can be evaluated for 10 marks. Marks of all experiments' write-ups are added and scaled down to **15 marks**.
- The laboratory test **(duration 02/03 hours)** after completion of all the experiments shall be conducted for 50 marks and scaled down to **10 marks.**
- Scaled-down marks of write-up evaluations and tests added will be CIE marks for the laboratory component of IPCC for **25 marks**.
- The student has to secure 40% of 25 marks to qualify in the CIE of the practical component of the IPCC.

SEE for IPCC

Theory SEE will be conducted by University as per the scheduled timetable, with common question papers for the course (**duration 03 hours**)

1. The question paper will have ten questions. Each question is set for 20 marks.
2. There will be 2 questions from each module. Each of the two questions under a module (with a maximum of 3 sub-questions), **should have a mix of topics** under that module.
3. The students have to answer 5 full questions, selecting one full question from each module.
4. Marks scoredby the student shall be proportionally scaled down to 50 Marks

The theory portion of the IPCC shall be for both CIE and SEE, whereas the practical portion will have a CIE component only. Questions mentioned in the SEE paper may include questions from the practical component.

Suggested Learning Resources:

Books

1. Tom Mitchell, ―Machine Learning, McGraw Hill, 3rd Edition, 1997.
2. Stephen Marsland, "Machine Learning - An Algorithmic Perspective", Second Edition, CRC Press - Taylor and Francis Group, 2015.

Web links and Video Lectures (e-Resources):

https://archive.nptel.ac.in/courses/106/106/106106139
https://www.youtube.com/watch?v=i_LwzRVP7bg
https://www.youtube.com/watch?v=NWONeJKn6kc

Activity Based Learning (Suggested Activities in Class)/ Practical Based learning

Miniproject in the topics of machine learning.