TECHNO INDIA
SILIGURI INSTITUTE
OF TECHNOLOGY
A Satyam Roychowdhury Initiative

SILIGURI INSTITUTE OF TECHNOLOGY
COMPUTER SCIENCE AND ENGINEERING

COURSE FILE
7TH SEM, 4TH YEAR

MACHINE LEARNING (PEC-CS701E)

Course Title	: **Machine Learning**
Course Code	: **PEC-CS701E**
L-T-P-S Structure	: 3L-0T-0P-0S
Credits	: 3
Pre-requisite	: **Artificial Intelligence, Data Warehousing**
Course Coordinator	: Sucharita Das
Team of Instructors	: NA
Teaching Associates (For LAB only)	: NA

SYLLABUS: *Machine Learning*

Code: **PCC-CS701E**
Contacts: 3L
PEC-CS701E

Introduction:
- What is Machine learning? How is it different from traditional programming?
- Why do we need Machine Learning? History of Machine Learning
- Machine Learning at Present, Features of Machine Learning
- Types of machine learning
- Supervised Learning
- Unsupervised Learning
- Reinforcement Learning

Unit 1: Supervised Learning (Regression/Classification)
Basic methods:
- Distance-based method (K-Nearest-Neighbours)
- Decision Trees
- Naive Bayes

Linear models:
- Linear Regression
- Logistic Regression
- Generalized Linear Models.
- Support Vector Machines, Nonlinearity and Kernel Methods.
- Beyond Binary Classification: Multi-class/Structured Outputs, Ranking.

Unit 2: Unsupervised Learning
Clustering:
- K-means.
- Kernel K-means

Dimensionality Reduction:
- PCA

- kernel PCA.

• Matrix Factorization
• Matrix Completion.
• Generative Models (mixture models and latent factor models).

Unit 3:
• Evaluating Machine Learning algorithms and Model Selection
• Introduction to Statistical Learning Theory.
• Ensemble Methods:
• Boosting
• Bagging
• Random Forests

Unit 4:
• Sparse Modelling and Estimation
• Modelling Sequence/Time-Series Data
• Deep Learning
• Feature Representation Learning.

Unit 5: Scalable Machine Learning
• Semi-supervised Learning.
• Active Learning.
• Reinforcement Learning.
• Inference in Graphical Models.
• Introduction to Bayesian Learning and Inference.

Unit 6:
• Recent trends in various learning techniques of machine learning and classification methods.

References:

1. Kevin Murphy, Machine Learning: A Probabilistic Perspective, MIT Press, 2012.
2. Tom. M. Mitchell , Machine learning, TMH.
3. Trevor Hastie, Robert Tibshirani, Jerome Friedman, The Elements of Statistical Learning, Springer 2009 (freely available online)
4. Christopher Bishop, Pattern Recognition and Machine Learning, Springer, 2007.
5. Dr. Rajiv Chopra, Machine Learning, Khanna Publishing House, 2018.

Course Hand out

ClassSchedule:

Dept./Day	Tuesday	Wednesday	Friday
CSE	11:40AM–12:30PM	12:30PM-1:20P.M.	11:40AM–12:30PM

HoursofMeetingStudents:-Anyday(between4:30PM to5:30PM)(ifrequired)

Course Objective: 1. To learn the concept of how to learn patterns and concepts from data without being explicitly programmed.
2 .To design and analysis various machine learning algorithms and techniques with a modern outlook focusing on recent advances.
3. Explore supervised and unsupervised learning paradigms of

machine learning.
4. To explore Deep Learning technique and various feature extraction strategies.

COURSE OUTCOMES (COs):

CO No	Course Outcome (CO)	Blooms Taxonomy Level (BTL)	Target %
PCC-CS701E.1	1. To formulate a machine learning problem.	BT Level 1	Students will attain 65% marks
PCC-CS701E.2	2. Select an appropriate pattern analysis tool for analysing data in a given feature space.	BT Level 4	Students will attain 65% marks
PCC-CS701E.3	3. Apply pattern recognition and machine learning techniques such as classification and feature selection to practical applications and detect patterns in the data.	BT Level 3	Students will attain 65% marks
PCC-CS701E.4	4. Design and program efficient algorithms related to recent machine learning techniques, train models, conduct experiments, and develop real-world ML-based applications and products.	BT Level 6	Students will attain 65% marks

PROGRAM OUTCOMES(POs):

PO Number	Description
1. **Engineering Knowledge**	Apply the knowledge of mathematics, science, engineering fundamentals, and an engineering specialization to the solution of complex engineering problems
2. **Problem Analysis**	Identify, formulate, review research literature, and analyze complex engineering problems reaching substantiated conclusions using first principles of mathematics, natural sciences, and engineering sciences.
3.	Design solutions for complex engineering problems and design system components or processes that meet the specified needs with appropriate consideration for the public health and safety, and the cultural, societal, and environmental considerations.

PO Number	Description
Design/ development of solutions	
4. **Conduct investigations of complex problems**	Use research-based knowledge and research methods including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.
5. **Modern tool usage**	Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools including prediction and modeling to complex engineering activities with an understanding of the limitations.
6. **The engineer and society**	Apply reasoning informed by the contextual knowledge to assess societal, health, safety, legal and cultural issues and the consequent responsibilities relevant to the professional engineering practice.
7. **Environment and sustainability**	Understand the impact of the professional engineering solutions in societal and environmental contexts, and demonstrate the knowledge of, and need for sustainable development.
8. **Ethics**	Apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice.
9. **Individual and team work**	Function effectively as an individual, and as a member or leader in diverse teams, and in multidisciplinary settings.
10. **Communication**	Communicate effectively on complex engineering activities with the engineering community and with society at large, such as, being able to comprehend and write effective reports and design documentation, make effective presentations, and give and receive clear instructions.
11. **Project management and finance**	Demonstrate knowledge and understanding of the engineering and management principles and apply these to one's own work, as a member and leader in a team, to manage projects and in multidisciplinary environments.
12. **Lifelong learning**	Recognize the need for, and have the preparation and ability to engage in independent and life-long learning in the broadest context of technological change.

Mapping of Course Outcomes and Program Outcomes: (Sample Attached)

Course Outcomes	Program Outcomes												PSOs	
	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12	1.	2.
PCC-CS701E.1	2	3	3	-	1	1	2	1	-	-	-	2	2	-
PCC-CS701E.2	2	2	1	1	1	-	-	-	1	1	-	2	2	1
PCC-CS701E.3	2	3	1	2	2	1	-	-	1	1	-	2	2	-
PCC-CS701E.4	3	3	-	1	3	1	1	2	2	1	-	2	2	1
PCC-CS701E	**2**	**3**	**2**	**1**	**2**	**1**	**1**	**2**	**1**	**1**	**-**	**2**	**2**	**1**

1 = courses in which the student will be exposed to a topic
2 = courses in which students will gain competency in that area
3= courses in which students will master that skill

COURSE DELIVERY PLAN:

Week	Sess. No.	CO	Topic (s)	**Book No [CH No][Page No]**	Teaching-Learning Methods	Planned Date	Execution Date
1	1	CO1	Discussion on course outcome and program outcome. What is Machine learning? How is it different from traditional programming? • Why do we need Machine Learning? History of Machine Learning • Machine Learning at Present, Features	2[1][11-15]	T: Chalk & Talk L: Observes understands	15/7/2023	
1	2	CO1	Types of machine learning • Supervised Learning • Unsupervised Learning • Reinforcement Learning	2[1][29-33]	T:Questioning /Discussion L: Answering questions, Participates	19/7/2023	
1	3	CO2	**Unit 1: Supervised Learning (Regression/Classification)** • Distance-based method (K-Nearest-Neighbours)	2[1][37-39]	**T:** Chalk & Talk L: Observes understands	20/7/2023	

2	4	CO2	• Decision Trees • Naive Bayes	2[3][52-59]	T: Lecturing L: Observes understands	22/7/2023	
2	5	CO2	**Linear models:** • Linear Regression	2[3][66-75]	**T:** Chalk & Talk L: Observes understands	26/7/2023	
2	6	CO2	• Logistic Regression • Generalized Linear Models.	2[4][86-94]	T: Lecturing L: Observes understands Video synthesis	27/7/2023	
3	7	CO2	• Support Vector Machines, Nonlinearity and Kernel Methods.	2[1][11-15]	T: Lecturing L: Observes understands	29/7/2023	
3	8	CO2	• Beyond Binary Classification: Multi-class/Structured Outputs, Ranking.	2[1][11-15]	**T:** Chalk & Talk L: Observes understands	2/8/2023	
3	9	CO2	**Unit 2: Unsupervised Learning** **Clustering:** • K-means.	2[8][230-234]	**T:** Chalk & Talk L: Observes understands	3/8/2023	
4	10	CO2	Kernel K-means	2[8][230-234]	**T:** Chalk & Talk L: Observes understands	5/8/2023	
4	11	CO3	**Dimensionality Reduction:** • PCA • kernel PCA.	2[7][201-205]	**T:** Chalk & Talk L: Observes understands	10/8/2023	
4	12	CO3	Matrix Factorization • Matrix Completion.	2[7][207-215]	**T:** Chalk & Talk L: Observes understands	12/8/2023	
5	13	CO3	• Generative Models (mixture models and latent factor models).	2[7][217-229]	**T:** Chalk & Talk L: Observes understands, Problem solving	16/8/2023	
5	*14*	CO3	**Unit 3:** • Evaluating Machine Learning algorithms and Model Selection	2[9][249-255]	**T:** Chalk & Talk L: Observes understands	17/8/2023	
5	*15*	CO3	• Introduction to Statistical Learning Theory.	2[9][249-255]	**T:** Chalk & Talk	19/8/2023	

			•		L: Observes understands		
6	*16*	CO3	Ensemble Methods:	2[10][274-280]	**T:** Chalk & Talk L: Observes understands	23/8/2022	
6	*17*	CO3	• Boosting • Bagging	2[10][283-290]	T: Lecturing L: Observes understands	24/8/2023	
6	*18*	CO3	• Random Forests	2[9][249-255]	**T:** Chalk & Talk L: Observes understands	30/8/2023	
7	*19*	CO4	**Unit 4:** • Sparse Modelling and Estimation	2[10][293-299]	**T:** Chalk & Talk L: Observes understands	31/8/2023	
7	*20*	CO4	• Modelling Sequence/Time-Series Data	5[7] [187-199]	T: Lecturing L: Observes understands	2/9/2023	
8	*21*	CO4	• Deep Learning	5[8] [205-207]	T: Lecturing L: Problem based learning	6/9/2023	
8	22	CO3	• Feature Representation Learning.	5[8] [215-221]	T **T:** Chalk & Talk L: Observes understands	7/9/2023	
8	23	CO3	**Unit 5:** Scalable Machine Learning • Semi-supervised Learning.	5[8] [231-240]	T: Lecturing L: Observes understands	9/9/2023	
9	24	CO4	• Active Learning.	5[8] [245-251]	T: Chalk & Talk L: Observes understands	13/9/2023	
9	25	CO4	• Reinforcement Learning.	2[13][267-272]	T: Chalk & Talk L: Observes understands	14/9/2023	
10	26	CO4	• Inference in Graphical Models.	2[13][267-272]	T: Chalk & Talk L: Observes understands	16/9/2023	
10	27	CO4	• Introduction to Bayesian Learning and Inference.	2[13][274-280]	T: Chalk & Talk L: Observes understands	20/9/2023	
11	28	CO4	• Recent trends in various learning techniques of machine learning and classification methods.	2[13][281-288]	T: Chalk & Talk L: Observes understands	21/9/2023	
11	*29*		Discussion on Previous Question Paper on WBUT			23/10/2023-27/10/2023	

LIST OF TUTORIALS:OPTIONAL

Tutorial session no	Topics	CO-Mapping
	NA	

WEEKLY HOMEWORK ASSIGNMENTS/ PROBLEM SETS/OPEN ENDEDED PROBLEM-SOLVING EXERCISES etc.

Week	Assignment/Quiz	Topic	Details	CO
2	A01	**Supervised Learning**	Distance-based method (K-Nearest-Neighbours)	CO1
4	A02	**Linear models**	Logistic Regression • Generalized Linear Models.	CO2
6	A03	**Dimensionality Reduction:**	Generative Models (mixture models and latent factor models).	CO3

COURSE TIME TABLE

REMEDIAL CLASSES:
Supplement course handout, which may perhaps include special lectures and discussions that would be planned, and schedule notified accordingly.

DELIVERY DETAILS OF CONTENT BEYOND SYLLABUS:
Content beyond syllabus covered (if any) should be delivered to all students that would be planned, and schedule notified accordingly.

S.No	Advanced Topics, Additional Reading, Research papers and any	CO	POs & PSOs	ALM	References/MOOCS
1	Vanishing gradients and fancy RNNs.	CO2	PO1,PO2 & PSO1	PPT & Lab	
2	Typical NLP task called 'machine translation',	CO3	PO1 & PSO1	PPT& Lab	
3	Performance of NNs across NLP tasks	CO3	PO1,PO2 & PSO1	PPT& Lab	

EVALUATION: AS PER MAKAUT GUIDELINES

Schedule for Continuous Assessment (CA):

CA	Assessment By	Schedule
CA-I	Presentation, Quiz, Group Discussion (25 Marks)	11th July – 14th August 2023
CA-II	Report writing (25 Marks)	11th July – 14th September 2023

CA-III	Class test in pen and paper mode to be conducted at the College Level. (25 Marks)	3rd October – 6th October, 2023
CA-IV	Centralized online test to be arranged by the University (25 Marks)	As Notified by MAKAUT
PCA1	Rubrics based Evaluation and Viva –Voce (40 Marks)	11th July – 14th September 2023
PCA2	Rubrics based Evaluation and Viva –Voce (40 Marks)	6th November – 10th November, 2023

ATTENDANCE POLICY

Every student is expected to be responsible for regularity of his/her attendance in class rooms and laboratories, to appear in scheduled tests and examinations and fulfil all other tasks assigned to him/her in every course. For Promotion, a Minimum of 50% of internal marks must be obtained. In every course, student has to maintain a minimum of 75% attendance to be eligible for appearing in Semester end examination of the course, for cases of medical issues and other unavoidable circumstances the students will be condoned if their attendance is between 60% to 75% in every course, subjected to submission of medical certificates, medical case file and other needful documental proof to the concerned departments.

DETENTION POLICY

In any course, a student has to maintain a minimum of75% attendance and must secure a minimum of 50% marks in In-Semester Examinations to be eligible for appearing to the Semester End Examination, failing to fulfill these conditions will deem such student to have been detained in that course.

PLAGIARISM POLICY

Use of unfair means in any of the evaluation components will be dealt with strictly, and the case will be reported to the examination committee.

COURSE TEAM MEMBERS, CHAMBER CONSULTATION HOURS AND CHAMBER VENUE DETAILS:

Each instructor will specify his / her chamber consultation hours during which the student can contact him / her in his / her chamber for consultation.

S.No.	Name of Faculty	Chamber Consultation Day (s)	Chamber Consultation Timings for each day	Chamber Consultation Room No:	Signature of Course faculty
1	SUCHARITA DAS	As per appointment	As per appointment	Faculty Room1	

GENERAL INSTRUCTIONS
Students should come prepared for classes and carry the text book(s) or material(s) as prescribed by the Course Faculty to the class.

NOTICES
All notices will be communicated through the institution email.
All notices concerning the course will be displayed on the respective Notice Boards.

Session Documents:

SESSION WISE TEACHING – LEARNING PLAN (17TH - 21ST JULY, 2023)

Course Title : Machine Learning (B. Tech ,7th Semester CSE)
Course Code : PEC-CS701E
L-T-P-S Structure : 3-0-0-0
Credits : 3
Course Coordinator : Mrs. SUCHARITA DAS

==========

Session No. 1 (17/07/2023)

Session Outcome:

1. Introduction:

Time (Min)	Topic	BTL	Teaching Learning Method	Active learning method
05	Record of Attendance	--	NA	
25	Introduction: • What is Machine learning? How is it different from traditional programming? •	1	Chalk & talk	Concept Mapping
15	Why do we need Machine Learning?	2	Chalk & talk	Concept Mapping
5	Summary and Conclusion	--	NA	Brainstorming

Session No. 2(18/07/2023)

Session Outcome:

2. Features of Machine Learning

Time (Min)	Topic	BTL	Teaching Learning Method	Active learning method
05	Record of Attendance	--	NA	

10	History of Machine Learning • Machine Learning at Present	1	Chalk & talk	Concept Mapping
30	Features of Machine Learning • Types of machine learning • Supervised Learning •	2	Chalk & talk	Concept Mapping
5	Summary and Conclusion	--	NA	Brainstorming

Session No. 3(20/07/2023)

Session Outcome:

3. Supervised Learning and Unsupervised Learning

Time (Min)	**Topic**	**BTL**	**Teaching Learning Method**	**Active learning method**
05	Record of Attendance	--	NA	
10	Introduction to Unsupervised Learning	1	Chalk & talk	Concept Mapping
30	Unsupervised Learning • Reinforcement Learning	2	Chalk & talk	Concept Mapping
5	Summary and Conclusion	--	NA	Brainstorming

Assignments:

Siliguri Institute of Technology

Computer Science & Engineering Department

Year: 4TH Semester: 7TH Section: A&B

Paper Name: Machine Learning

Paper Code: PEC-CS701E

Last Date of Submission: 30th August, 2023

Full Marks: 25

Assignment policy:

- Assignment must be written in A4 sheet along with the front sheet then uploaded in **softcopy as .pdf** file in the given **link** as well as submitted in **hardcopy** (A4 sheet) within the due date mentioned above.

- ***Link for online submission***: https://drive.google.com/drive/folders/1qEcGb9cSEiLsGmCzFCwNIuf5kGOnTXiY?usp=sharing

- All assignments must be done individually.
- **No late submissions will be allowed.**
- **Each question will carry 5 marks.**

Assignment – I

1. a) Why k-NN algorithm does more computation on test time rather than train time.
   b) Why should we not use the KNN algorithm for large datasets? 2+3=5

2 a) Write down the steps of **k-NN algorithm .** 2+3=5
   b) What are the advantages and disadvantages of **k-NN algorithm.**

3. a) Which should be preferred among Gini impurity and Entropy? Why? 3+2=5
   b) List down the attribute selection measures used by the ID3 algorithm to construct a Decision Tree.

4. a) What is the Decision Tree Algorithm? 3+2=5
   b) How does a Decision Tree handle missing attribute values?

5. a) Explain the relationship between Machine Learning, Artificial Intelligence, and Deep Learning using diagram. 3+2=5
   b) What are the different types of machine learning algorithms with examples.

Sr. No.	Roll No./ NAME	CA1(out of 25)	CA2(out of 25)	CA3(out of 25)
1	11900120008 - ANSHIKA JAISWAL	25	24	23.5
2	11900120027 - NABIN DAS	25	24	24
3	11900120050 - rudra prasad das	21	24	24.5
4	11900120059 - ROUNAK SHAW	25	24	19
5	11900120061 - AVIJEET PAUL	25	24	24.5
6	11900120062 - DIKSHITA SHAH	25	24	23
7	11900120064 - ANUSHKA GHOSH CHOWDHURY	25	24	24
8	11900120071 - Preetam Das	25	24	18.5
9	11900120082 - Gaddar Kumar Chaudhary	25	24	19
10	11900120084 - ABHINASH KUMAR TIWARI	25	24	23
11	11900120088 - Swaroop Raj Lama	25	24	22.5
12	11900120090 - SHWETA BHAGAT	25	24	24
13	11900121183 - SAJEEB GHOSH	25	24	21

Syllabus Progress Report of Machine Learning upto 21/11/2023

Course Name:	Machine Learning	Course Code:	PEC-CS701E
Teacher Name:	Sucharita Das	Teacher Employee Code:	6005
Department Name:	Computer Science and Engineering	Semester:	7th sem
LTP Structure:	3L+0T+0P		
Total No. of classes Allotted:	36		
Total No. of clases conducted:	22		
% of syllabus covered:	97%		

Module No.	Topics Covered	Methodology Used	No. of Lectures conducted
1	Introduction: • What is Machine learning? How is it different from traditional programming? • Why do we need Machine Learning?	Google ClassRoom and Handouts	1
	History of Machine Learning • Machine Learning at Present, Features of Machine Learning • Types of machine learning • Supervised Learning •		1
	Unsupervised Learning • Reinforcement Learning		1
	Supervised Learning (Regression/Classification)		1
	Distance-based method(K-Nearest-Neighbours)		1
	Decision Trees with example		1
	Naive Bayes with explanation and examples		1
	Linear Regression algorithm with examples.		1

Module No.	Topics Covered	Methodology Used	No. of Lectures conducted
1	Logistic regression	Google ClassRoom and Handouts	1
2	K-means, Kernel K-means		1
	Generalized Linear Models.· Support Vector Machines, Nonlinearity and Kernel Methods, Beyond Binary Classification:Multi-class/Structured Outputs, Ranking.		1
3	Evaluating Machine Learning algorithms and Model Selection ,Introduction to Statistical Learning Theory. Ensemble,Methods: Boosting,Bagging ,Random Forests		1
4	Deep Learning:Feature Representation Learning,Semi-supervised,Learning		1

5	Active Learning. Reinforcement Learning. Inference in Graphical Models. Introduction to Bayesian Learning and Inference.		1
4	Sparse Modelling and Estimation • Modelling Sequence/Time-Series Data • Deep Learning		1
5	• Feature Representation Learning.• Semi-supervised Learning. • Active Learning. • Reinforcement Learning.		1
5	• Inference in Graphical Models. • Introduction to Bayesian Learning and Inference.		1
6	• Recent trends in various learning techniques of machine learning and classification methods.		1
Signature of the Faculty Member			**Signature of HOD**