Course Plan

CHITKARA UNIVERSITY

A. Course Handout

Institute/School/College Name	Chitkara University Institute of Engineering & Technology
Department/Centre Name	Department of Computer Science & Engineering
Programme Name	Bachelor of Engineering- Computer Science & Engineering (Artificial Intelligence)

Course Name	Supervised and Unsupervised Learning	**Session**	2024-25
Course Code	22AI010	**Semester/Batch**	4th/2023
Lecture/Tutorial (Per Week)	2-0-4	**Course Credits**	4
Course Coordinator Name	Dr. Harshvardhan		

1. Objective of the Course:

The main objectives of the courses are:
- Use Supervised and Unsupervised Learning algorithms to make predictions.
- Use Regression and Classification algorithms for predicting continuous and discrete values.
- Learn best practices for real world data sets.
- Learn how to use sklearn library for pre-processing the data.
- Understanding important concepts which helps to improve model more.
- Understand the full product workflow for the machine learning lifecycle.

2. Course Learning Outcome

After completing the course, the student will be able to: -

CLO	Course Outcome	POs	CL	KC	Sessions
CLO01	Understand the fundamental concepts and principles of machine learning, including the types of problems it can solve and the tools and techniques used to implement ML algorithms.	PO1, PO2, PO3, PO4, PO5, PO6	K2	Factual, Conceptual	**04**
CLO02	Develop the skills of pre-process and clean the data for ML tasks, including handling missing values or corrupted data, selecting relevant features, and scaling and normalizing the data.	PO1, PO2, PO3, PO4, PO5, PO6	K3	Fundamental, Conceptual	**10**
CLO03	Become proficient in using popular ML libraries and frameworks such as sklearn to build ML models.	PO1, PO2, PO3, PO4, PO5, PO6	K3	Conceptual, Procedural	**16**
CLO04	Understand how to evaluate the performance of ML models, including common metrics and methods for model selection and hyper-parameter tuning	PO1, PO2, PO3, PO4, PO5, PO6	K3	Conceptual, Procedural	**16**
CL005	Understand the ethical and societal implications of ML, including issues of bias, privacy, and transparency.	PO1, PO2, PO3, PO4, PO5, PO6, PO7, PO8, PO9	K3	Conceptual, Procedural	**14**
Total Contact Hours					**60**

CLO-PO-PSO Mapping grid |Program outcomes (POs) and Program Specific Outcomes (PSOs) are available as a part of Academic Program Guide

Course Learning Outcomes	P01	P02	P03	P04	P05	P06	P07	P08	P09	P10	P11	P12	PSO1	PSO2	PSO3
CLO1	H	M	L	M	L	L							M		M
CLO2	H	L	M	M	H	M								M	H
CLO3	H	L	L	M	H	M								H	M
CLO4	H	M	M	M	H	M							H		
CLO5	H	H	M	M	M	H	L	H	M				H	L	H

*H=High, M=Medium, L=Low

3. **ERISE Grid Mapping**

Feature Enablement	Level (1-5, 5 being highest)
Entrepreneurship	3
Research	5
Innovation	4
Skills	4
Employability	4

4. **Recommended Books (Reference Books/Textbooks):**

B01: "Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow" by Aurélien (2nd edition)
B02: Deep Learning" by Ian Goodfellow, Yoshua Bengio, and Aaron Courville (2nd edition)
B03: Introduction to Machine Learning" by Alpaydin (4th edition)
B04: The Elements of Statistical Learning" by Trevor Hastie, Robert Tibshirani (2nd edition)
B05: Pattern Classification" by Duda and Hart (2nd edition)

5. **Other readings & relevant websites:**

S. No.	Link of Journals, Magazines, websites and Research Papers
1	https://www.analyticsvidhya.com/machine-learning/
2	https://towardsdatascience.com/machine-learning/home
3	https://www.geeksforgeeks.org/machine-learning/
4	https://www.javatpoint.com/machine-learning
5	https://medium.com/@randylaosat/a-beginners-guide-to-machine-learning-dfadc19f6caf

6. **Recommended Tools and Platforms**

Anaconda, Jupyter Notebook, Pycharm, VSCode, Google Colab

7. **Course Plan:**
   a. **Lecture Plan**

Lecture Number	Topics	Recommended Books
1-3	Introduction to Machine Learning: Basic Terminology, Types of Machine Learning, Supervised, Unsupervised, and Semi-Supervised Learning.	B02, B03
4-7	Revision - Python, Numpy, Pandas, Matplotlib (Prerequisite)	B02, B03

8-9	Linear Regression - Linear Regression Ideal, Best Line, Concept, Best Line Metric, Best Line Math Metric	B01, B03
10-12	Linear Regression - Loss functions, Loss function minimize, Square Function, MSE/RMSE, Assumptions of Linear Regression	B02, B03
13-14	Introduction to Gradient Descent, Stochastic Gradient Descent, Mini Batch Gradient Descent.	B04, B05
15-17	Multiple Linear Regression, Polynomial Regression using case studies.	B02, B03
Formative Assessment-1 (1-17 lectures)		
18-20	Classification - Evaluation Metric (Confusion metric, Precision, Recall, F1-Score), Loss function vs Cost Function.	B02, B04
21-24	Classification - Logistic Regression, Step function, Sigmoid function, Loss Function, Optimization technique for loss function	B05, B03
Sessional Test-1 (1-24 lectures)		
25-26	Feature Selection, Bias-Variance Trade offs, Regularization techniques	B02, B03
27-30	Augmentations, K-Fold Cross Validation, Grid Search Cv/Random Search	B01, B02
31-32	Classification and Regression with KNN.	B03, B04
33-36	KNN Complexity and Overfitting vs Underfitting.	B03, B05
Formative Assessment-2 (18-36 lectures)		
37-38	Sklearn Documentation and cross_val_score	B02, B03
39-41	Imbalanced Dataset - Oversampling, Downsampling, Weighting classes	B01, B03
42-45	How to Handle Categorical and missing values in a dataset, Feature Importance, Collinearity, Hyper-parameter Tuning	B02, B03
46-47	Dimensionality Reduction - PCA, TSNE	B02, B04
48-50	Dimensionality Reduction - Applications and use case studies	B01, B03
Sessional Test-2 (25-50 lectures)		
51-54	PCA and SVD it's working, Eigen Values, Eigen vectors, Covariance matrix.	B02, B03
55-56	Unsupervised Learning Algorithm - K-Means Clustering, Elbow Method	B03, B05
56-60	Classification and Regression with KNN, SVM, Decision Trees, Random Forest	B02, B03
END-TERM EXAM (FULL SYLLABUS)		

8. **Delivery/Instructional Resources**

Lecture No.	Topics	PPT *(Link of ppts on the central server)*	Industry Expert Session *(If yes: link of ppts on the central server)*	Web References	Audio-Video
1-3	Introduction to Machine Learning: Basic Terminology, Types of Machine Learning, Supervise Learning, Unsupervised Learning, Semi-Supervised Learning.	https://www.salesforce.com/eu/blog/2020/06/real-world-examples-of-machine-learning.html			www.youtube.com/watch?v=jGwO_UgTS7I&list=PLoROMvodv4rMiGQp3WXShtMGgzqpfVfbU
4-7	Revision - Python, Numpy, Pandas, Matplotlib (Pre-requisite)			https://towardsdatascience.com/the-data-science-trilogy-numpy-pandas-and-matplotlib-basics-42192b89e26	
8-9	Linear Regression - Linear Regression Ideal, Best Line, Concept, Best Line Metric, Best Line Math Metric			https://www.analyticsvidhya.com/blog/2021/10/everything-you-need-to-know-about-linear-regression/	
10-12	Linear Regression - Loss functions, Loss function minimize, Square Function, MSE/RMSE, Assumptions of Linear Regression			https://towardsdatascience.com/evaluation-metrics-model-selection-in-linear-regression-73c7573208be	
13-14	Introduction to Gradient Descent, Stochastic Gradient Descent, Mini Batch Gradient Descent.			https://www.javatpoint.com/gradient-descent-in-machine-learning	
15-17	Classification - Logistic Regression, Step function, Sigmoid function, Loss Function, Optimization technique for loss function			https://www.javatpoint.com/gradient-descent-in-machine-learning	
18-20	Classification - Evaluation Metric (Confusion metric, Precision, Recall, F1-Score), Loss function vs Cost Function.			https://medium.com/@nandinisekar27/logistic-regression-overview-and-its-error-metrics-8b54512ca39f	
21-24	Multiple Linear Regression, Polynomial Regression using case studies.			https://towardsdatascience.com/polynomial-regression-bbe8b9d97491	

25-26	Feature Selection, Bias-Variance Trade offs, Regularization techniques			https://towardsdatascience.com/regularization-the-path-to-bias-variance-trade-off-b7a7088b4577	
27-30	Augmentations, K-Fold Cross Validation, Grid Search CV/Random Search			https://towardsdatascience.com/regularization-the-path-to-bias-variance-trade-off-b7a7088b4577	
31-32	Classification and Regression with KNN.			https://www.javatpoint.com/k-nearest-neighbor-algorithm-for-machine-learning	
33-36	KNN Complexity and Overfitting vs Underfitting.			https://www.geeksforgeeks.org/underfitting-and-overfitting-in-machine-learning/	
37-38	Sklearn Documentation and cross_val_score			https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.cross_val_score.html	
39-41	Imbalanced Dataset - Oversampling, Downsampling, Weighting classes			https://machinelearningmastery.com/random-oversampling-and-undersampling-for-imbalanced-classification/	
42-45	How to Handle Categorical and missing values in a dataset, Feature Importance, Collinearity, Hyper-parameter Tuning			https://towardsdatascience.com/supervised-machine-learning-feature-engineering-and-hyper-parameter-tuning-a3da583dd7b9	
46-47	Dimensionality Reduction - PCA, TSNE			https://www.geeksforgeeks.org/difference-between-pca-vs-t-sne/	
48-50	Dimensionality Reduction - Applications and use case studies			https://www.geeksforgeeks.org/dimensionality-reduction/	
51-54	PCA and SVD it's working, Eigen Values, Eigen vectors, Covariance matrix.			https://medium.com/@notsokarda/pca-vs-svd-simplified-32c5c753998	
55-56	Unsupervised Learning Algorithm - K-Means Clustering, Elbow Method				https://youtu.be/EItlUEPCIzM?si=IxUiZm2PsZFbDebU

56-60	Classification and Regression with KNN, SVM, Decision Trees, Random Forest				https://www.youtube.com/watch?v=yN7ypxC7838

9. **Action plan for different types of learners**

Slow Learners	Average Learners	Fast Learners
● Multiple Remedial Extra Classes ● Encouragement for improvement using Peer Tutoring	● Doubt-sessions ● Pre-coded algorithms to illustrate concepts and notions ● E-notes and E-exercises to read in addition to pedagogic material	● More Practice assignments on real life problems ● Engaging students to hold hands of slow learners by creating a Peer Tutoring Group ● Participation in Hackathons, competitions.

10. **Evaluation Scheme & Components:**

Evaluation Component	Type of Component	No. of Assessments	Weightage of Component
Component 1	Formative Assessments (FAs)	02*	20%
Component 2	Sessional Tests (STs)	02**	30%
Component 3	End Term Examinations	01	50%
Total			**100%**

*Out of 02 FAs, the ERP system automatically picks the best 01 FA marks for evaluation of the FAs as final marks.
**ST1 and ST2 average will be taken for the total ST marks.

11. **Details of Evaluation Components:**

Evaluation Component	Description	Syllabus Covered (%)	Timeline of Examination	Weightage (%)
Component 01	Formative Assessment 01	Up to 28%	Week 6	20%
	Formative Assessment 02	29%-50%	Week 14	
Component 02	ST 01	Up to 40%	Week 7	30%
	ST 02	41% - 80%	Week 15	
Component 03	End Term Examination*	100%	To be notified by Dean Examination	50%

* As per Academic Guidelines minimum 75% attendance is required to become eligible for appearing in the End Semester Examination

12. Evaluation Components

Type of Assessment	Timeline of Conduct	Total Marks	Question Paper Format			
			1 Mark MCQ	2 Mark MCQ/2 Marks Question	5 Mark Question	10 Mark Question
Formative Assessment 1	Week 6	20	20	0	0	0
Formative Assessment 2	Week 7	20	20	0	0	0
Sessional Test 1	Week 14	30	20	5	0	0
Sessional Test 2	Week 15	30	20	5	0	0
End Term Examination		50	10	5	4	1

B. Syllabus of the Course:

Subject: Supervised and Unsupervised Learning	Subject Code: 22AI010

S. No.	Topic (s)	No. of Lectures	Weightage %
1-3	Introduction to Machine Learning: Basic Terminology, Types of Machine Learning, Supervise Learning, Unsupervised Learning, Semi-Supervised Learning.	3	10%
4-10	Revision - Python, Numpy, Pandas, Matplotlib (Pre-requisite)	7	5%
11-20	Linear Regression - Loss functions, Loss function minimize, Square Function, MSE/RMSE, Assumptions of Linear Regression, Introduction to Gradient Descent, Stochastic Gradient Descent, Mini Batch Gradient Descent. Classification - Logistic Regression, Step function, Sigmoid function, Loss Function, Optimization technique for loss function, Classification - Evaluation Metric (Confusion metric, Precision, Recall, F1-Score), Loss function vs Cost Function.	10	20%
21-30	Multiple Linear Regression, Polynomial Regression using case studies. Feature Selection, Bias-Variance Trade offs, Regularization techniques, Augmentations-Fold Cross Validation, Grid Search Cv/Random Search	10	20%
31-43	Classification and Regression with KNN. KNN Complexity and Over fitting vs Under fitting. Sklearn Documentation and cross_val_score Imbalanced Dataset - Oversampling, Downsampling, Weighting classes	13	20%
44-54	How to Handle Categorical and missing values in a dataset, Feature Importance, Collinearity, Hyper-parameter Tuning, Dimensionality Reduction - PCA, TSNE, Dimensionality Reduction - Applications and use case studies. PCA and SVD it's working, Eigen Values, Eigen vectors, Covariance matrix.	11	20%

55-60	Unsupervised Learning Algorithm - K-Means Clustering, Elbow Method, Classification and Regression with KNN, SVM, Decision Trees, Random Forest	6	5%

This Document is approved by:

Designation	Name	Signature
Course Coordinator	Dr. Harshvardhan	
Head Academic Delivery	Dr. Kamal Deep Garg	
Dean (CSE-AI)	Dr. Sushil Kumar Narang	
Date	January 07, 2025	