

**Programme Name/s** : **Computer Technology/ Computer Engineering/ Computer Software Technology/ Computer Science & Engineering/ Computer Hardware & Maintenance/ Information Technology/ Computer Science & Information Technology/ Computer Science/**

**Programme Code** : **CM/ CO/ CST/ CW/ HA/ IF/ IH/ SE**

**Semester** : **Sixth**

**Course Title** : **MACHINE LEARNING**

**Course Code** : **316316**

### I. RATIONALE

A machine learning concepts equip students with fundamental knowledge and practical skills. Machine learning algorithms are important for data analysis, AI and Data Science. It emphasizes real-world applications such as data analytics, predictive analytics and problem-solving, which are in high demand across industries like healthcare, finance and technology.

### II. INDUSTRY / EMPLOYER EXPECTED OUTCOME

The aim of this course is to help the students to attain the following industry identified outcome through various teaching learning experiences: Apply machine learning effectively across various domains.

### III. COURSE LEVEL LEARNING OUTCOMES (COS)

Students will be able to achieve & demonstrate the following COs on completion of course based learning

- CO1 - Explain the role of machine learning in AI and data science.
- CO2 - Implement data preprocessing.
- CO3 - Implement feature engineering techniques to prepare data for machine learning models.
- CO4 - Apply supervised learning models to train and evaluate.
- CO5 - Apply unsupervised learning models to train and evaluate.

### IV. TEACHING-LEARNING & ASSESSMENT SCHEME

Course Code	Course Title	Abbr	Course Category/s	Learning Scheme						Credits	Assessment Scheme											
				Actual Contact Hrs./Week			SLH	NLH	Paper Duration		Theory				Based on LL & TL				Based on SL			Total Marks
				CL	TL	LL					Total		Practical		SLA							
							FA-TH	SA-TH			Max	Min	FA-PR	SA-PR	Max	Min	Max	Min				
316316	MACHINE LEARNING	MAL	DSE	3	-	2	1	6	3	3	30	70	100	40	25	10	25#	10	25	10	175	

## Total IKS Hrs for Sem. : 0 Hrs

Abbreviations: CL- Classroom Learning , TL- Tutorial Learning, LL-Laboratory Learning, SLH-Self Learning Hours, NLH-Notional Learning Hours, FA - Formative

Assessment, SA -Summative assessment, IKS - Indian Knowledge System, SLA - Self Learning Assessment

Legends: @ Internal Assessment, # External Assessment, \*\* On Line Examination , @\$ Internal Online Examination

Note :

1. FA-TH represents average of two class tests of 30 marks each conducted during the semester.
2. If candidate is not securing minimum passing marks in FA-PR of any course then the candidate shall be declared as "Detained" in that semester.
3. If candidate is not securing minimum passing marks in SLA of any course then the candidate shall be declared as fail and will have to repeat and resubmit SLA work.
4. Notional Learning hours for the semester are (CL+LL+TL+SL)hrs.\* 15 Weeks
5. 1 credit is equivalent to 30 Notional hrs.
6. \* Self learning hours shall not be reflected in the Time Table.
7. \* Self learning includes micro project / assignment / other activities.

## V. THEORY LEARNING OUTCOMES AND ALIGNED COURSE CONTENT

Sr.No	Theory Learning Outcomes (TLO's) aligned to CO's.	Learning content mapped with Theory Learning Outcomes (TLO's) and CO's.	Suggested Learning Pedagogies.
1	<p>TLO 1.1 Describe machine learning concept.</p> <p>TLO 1.2 Compare Traditional vs ML based Programming.</p> <p>TLO 1.3 Distinguish Supervised , unsupervised and Reinforcement learning.</p> <p>TLO 1.4 Explain the Challenges for Machine learning.</p> <p>TLO 1.5 Explain the features of python libraries used for Machine Libraries.</p>	<p><b>Unit - I Introduction to Machine Learning</b></p> <p>1.1 Basics of ML - Define Machine Learning, Traditional programming vs ML-based approaches, Role and application of ML in artificial intelligence and data science</p> <p>1.2 Types of ML (Supervised, Unsupervised, Reinforcement Learning)-Supervised Learning: Definition, working principle, examples (classification, regression), Unsupervised Learning: Definition, working principle, examples, Clustering, dimensionality reduction techniques, Reinforcement Learning: Concept of agents, rewards, and policy learning, Comparison of different ML types with real-world applications</p> <p>1.3 Applications of ML- Real-world use cases of ML in various domains such as healthcare, finance, e-commerce, etc, challenges in Machine Learning</p> <p>1.4 Introduction to Python for ML-Basics of Python programming relevant to ML, Overview required libraries: NumPy, Pandas, Matplotlib, Scikit-learn, Writing and executing simple ML scripts in Python</p>	<p>Lecture Using</p> <p>Chalk-Board</p> <p>Presentations</p>
2	<p>TLO 2.1 Illustrate the process of cleaning the data.</p> <p>TLO 2.2 Explain the functioning of algorithm used for handling missing values.</p> <p>TLO 2.3 Select the appropriate dataset splitting technique.</p>	<p><b>Unit - II Data Preprocessing</b></p> <p>2.1 Data Cleaning :Introduction to Data Cleaning, Identifying and Handling Noisy Data, Removing Duplicates and Inconsistencies, Standardizing and Normalizing Data, Handling Outliers</p> <p>2.2 Handling Missing Values -Types of Missing Data, Identifying Missing Values, Techniques to Handle Missing Data, Removing Missing Data, Mean, Median, and Mode Imputation, Predictive Imputation (Regression, KNN), Using Algorithms that Support Missing Values</p> <p>2.3 Splitting Dataset for Training and Testing -Importance of Dataset Splitting, Train-Test Split Ratio Selection, Cross-Validation Techniques, K-Fold Cross Validation, Leave-One-Out Cross Validation, Stratified Sampling vs Random Sampling</p>	<p>Lecture Using</p> <p>Chalk-Board</p> <p>Presentations</p> <p>Video</p> <p>Demonstrations</p>

Sr.No	Theory Learning Outcomes (TLO's) aligned to CO's.	Learning content mapped with Theory Learning Outcomes (TLO's) and CO's.	Suggested Learning Pedagogies.
3	TLO 3.1 Explain different feature selection methods. TLO 3.2 Illustrate the different feature extraction methods.	<b>Unit - III Feature Selection</b> 3.1 Feature Scaling and Selection-Importance of Feature Scaling, Normalization vs Standardization, Feature Selection Methods, Filter Methods (Correlation, Chi-Square), Wrapper Methods (Forward, Backward Selection), Embedded Methods (Lasso, Decision Trees) 3.2 Feature Extraction Techniques - Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA) 3.3 Mutual Information-Based Feature Selection, ANOVA (Analysis of Variance) for Feature Selection, Recursive Feature Elimination (RFE) with Cross-Validation (RFECV), Feature Importance from Tree-Based Models (Beyond Decision Trees), XGBoost, LightGBM, Random Forest provide built-in feature importance scores (Gini importance, SHAP values).	Lecture Using Chalk-Board Presentations Demonstration
4	TLO 4.1 Distinguish the working of Decision Tree, KNN and SVM classification algorithm. TLO 4.2 Explain the working of specified Regression algorithm. TLO 4.3 Explain the process of model performance evaluation.	<b>Unit - IV Supervised Learning</b> 4.1 Classification Algorithms: Decision Trees, KNN(K-Nearest Neighbors), SVM(Support Vector Machine) 4.2 Regression Algorithms: Linear Regression, Logistic Regression, Ridge Regression 4.3 Model Performance Evaluation: Confusion Matrix, Accuracy, Precision, Recall	Lecture Using Chalk-Board Presentations Demonstration
5	TLO 5.1 Explain the concept of clustering in data analysis and its significance. TLO 5.2 Compare K-Means and Hierarchical Clustering and choose the appropriate technique based on the problem. TLO 5.3 Describe the importance of Dimensionality reduction. TLO 5.4 Illustrate the process of dimensionality reduction using PCA.	<b>Unit - V Unsupervised Learning</b> 5.1 Clustering Techniques:-Define Clustering, Importance of clustering in data analysis, Applications of Clustering 5.2 K-Means Clustering: Definition and working principle, Steps involved in the K-Means algorithm, Advantages of K-Means, Disadvantages of K-Means, Hierarchical Clustering: Definition and types, Steps in Hierarchical Clustering, Advantages of Hierarchical Clustering, Disadvantages of Hierarchical Clustering, Comparing K-Means and Hierarchical Clustering 5.3 Dimensionality Reduction: Importance of Dimensionality Reduction 5.4 PCA -Definition and fundamental principles of PCA, Eigenvectors and eigenvalues, Steps in PCA, Explained Variance, Choosing the Optimal Dimensionality, Advantages and Disadvantages of PCA, Applications of PCA	Lecture Using Chalk-Board Presentations Demonstration

## VI. LABORATORY LEARNING OUTCOME AND ALIGNED PRACTICAL / TUTORIAL EXPERIENCES.

Practical / Tutorial / Laboratory Learning Outcome (LLO)	Sr No	Laboratory Experiment / Practical Titles / Tutorial Titles	Number of hrs.	Relevant COs
LLO 1.1 Install required platform to use Scikit-learn library.	1	*Installation of IDE with necessary libraries	2	CO1
LLO 2.1 Write a program for handling missing values. LLO 2.2 Normalize data to make models work. LLO 2.3 Standardize data to make models work. LLO 2.4 Transform categorical data into numerical form using encoding methods. LLO 2.5 Split the dataset into train and test.	2	*Implement program for Data Preprocessing Techniques	2	CO1 CO2
LLO 3.1 Write a program to read dataset and differentiate attributes in various category.	3	Implement program to read dataset ( Text, CSV, JSON ,XML)	2	CO2
LLO 4.1 Write a program to implement the Decision Tree model. LLO 4.2 Write a program to implement the K-Nearest Neighbor model. LLO 4.3 Evaluate classification performance using accuracy, precision, recall, and F1-score .	4	Implement the classification algorithms on previously prepared dataset	4	CO2 CO3 CO4

Practical / Tutorial / Laboratory Learning Outcome (LLO)	Sr No	Laboratory Experiment / Practical Titles / Tutorial Titles	Number of hrs.	Relevant COs
LLO 5.1 Write a program to implement the Linear Regression by using the suitable dataset. LLO 5.2 Write a program to implement the logistic regression by using the suitable dataset. LLO 5.3 Write a program to implement the Ridge Regression by using the suitable dataset. LLO 5.4 Evaluate model performance using metrics.	5	*Implement the regression model by using the suitable dataset	4	CO2 CO4
LLO 6.1 Write a program to use clustering algorithms to find patterns in data. LLO 6.2 Visualize the results using Matplotlib / Seaborn.	6	*Implement program to use clustering algorithms to find patterns in data	2	CO3 CO4
LLO 7.1 Write a program to identify the most important features that contribute to the model accuracy.	7	Implement program to identify the most important features that contribute to the model accuracy	2	CO2 CO3
LLO 8.1 Write a program to use k-Nearest Neighbors (KNN) model for Classification on give dataset. LLO 8.2 Experiment with different values of K and measure model performance.	8	Implement program to use k-Nearest Neighbors (KNN) model for Classification on given dataset	2	CO4
LLO 9.1 Write a program to Train an SVM model on dataset.	9	*Implement program to train an SVM model on given dataset	2	CO4
LLO 10.1 Write a program to use logistic regression model to classify binary outcomes.	10	Implement program to use logistic regression model to classify binary outcomes	2	CO3
LLO 11.1 Write a program to use PCA technique to reduce the number of features while retaining important information.	11	*Implement program to use PCA technique to reduce the number of features while retaining important information	2	CO4
LLO 12.1 Write a program to use any machine learning model on given dataset. LLO 12.2 Preprocess, analyze, and build models to get the meaningful insights.	12	*Implement program to use machine learning model on given dataset	2	CO4 CO5
LLO 13.1 Write a program to use Pandas, Matplotlib to analyze data of previously created model . LLO 13.2 Plot histograms, box plots, scatter plots, and correlation heatmaps to understand relationships between variables.	13	Implement program to use Pandas, Matplotlib to analyze data of previously created model	2	CO4 CO5
LLO 14.1 Write a program to use any machine learning model on given dataset. LLO 14.2 Predict results based on various features using a regression model.	14	Implement program to use machine learning model on Boston Housing Dataset (available in Scikit-learn)	2	CO4 CO5
LLO 15.1 Write a program to make groups based on their features using K-Means Clustering.	15	Implement a program to segment customers into different groups based on their purchasing behavior features using K-Means Clustering	4	CO5

### Note : Out of above suggestive LLOs -

- \*\* Marked Practicals (LLOs) Are mandatory.
- Minimum 80% of above list of lab experiment are to be performed.
- Judicial mix of LLOs are to be performed to achieve desired outcomes.

## VII. SUGGESTED MICRO PROJECT / ASSIGNMENT/ ACTIVITIES FOR SPECIFIC LEARNING / SKILLS DEVELOPMENT (SELF LEARNING)

### Micro project

- The micro project should be based on one of the following area as suggested by the teacher: Industry application/Internet-based/workshop-based/laboratory-based/field-based.
  1. Develop a model for Waiter's Tip Prediction using Machine Learning algorithm , predict what amount of tip a person will give based on his/her visit to the restaurant using some features related to the same.
  2. Implement a machine learning model which can predict the stock amount for the different products which are sold in different stores.
  3. Implement a model for Stock Price Prediction is the task of forecasting future stock prices based on historical data and various market indicators.
  4. Implement model to Predict Employee Attrition: Analyze HR datasets to determine factors affecting employee retention.
  5. Implement model to perform Human Scream Detection and Analysis for Controlling Crime Rate using Machine Learning.

### **Assignment**

- Solve assignment covering all COs given by teacher.

### **Other Courses**

- Students are encouraged to register themselves in various Swayam/ MOOC's/Infosys Springboard courses to enhance their learning. Various courses available on machine learning courses with certificates for Free such as Generative AI / AI Python for beginners/ Google AI for Anyone / Linear regression/ AWS Deep Learning Containers / CS229: Machine Learning By Stanford/ Great Learning Academy.
-

## Note :

- Above is just a suggestive list of microprojects and assignments; faculty must prepare their own bank of microprojects, assignments, and activities in a similar way.
- The faculty must allocate judicious mix of tasks, considering the weaknesses and / strengths of the student in acquiring the desired skills.
- If a microproject is assigned, it is expected to be completed as a group activity.
- SLA marks shall be awarded as per the continuous assessment record.
- For courses with no SLA component the list of suggestive microprojects / assignments/ activities are optional, faculty may encourage students to perform these tasks for enhanced learning experiences.
- If the course does not have associated SLA component, above suggestive listings is applicable to Tutorials and maybe considered for FA-PR evaluations.

## VIII. LABORATORY EQUIPMENT / INSTRUMENTS / TOOLS / SOFTWARE REQUIRED

Sr.No	Equipment Name with Broad Specifications	Relevant LLO Number
1	Computer system with a minimum 8GB RAM, Python/Jupyter Notebook/ Google Colab Free/unpaid Account/ Anaconda	All
2	Other Useful Tools: VS Code / PyCharm " Alternative Python development environments. Git & GitHub " For version control and project collaboration. Kaggle " Free datasets and cloud computing for ML experiments.	All

## IX. SUGGESTED WEIGHTAGE TO LEARNING EFFORTS & ASSESSMENT PURPOSE (Specification Table)

Sr.No	Unit	Unit Title	Aligned COs	Learning Hours	R-Level	U-Level	A-Level	Total Marks
1	I	Introduction to Machine Learning	CO1	6	2	6	4	12
2	II	Data Preprocessing	CO2	10	4	8	4	16
3	III	Feature Selection	CO3	9	2	4	4	10
4	IV	Supervised Learning	CO4	10	4	8	4	16
5	V	Unsupervised Learning	CO5	10	4	8	4	16
<b>Grand Total</b>				<b>45</b>	<b>16</b>	<b>34</b>	<b>20</b>	<b>70</b>

## X. ASSESSMENT METHODOLOGIES/TOOLS

## Formative assessment (Assessment for Learning)

- Continuous assessment based on process and product related performance indicators. Each practical will be assessed considering 60% weightage to process 40% weightage to product. A continuous assessment-based term work.

## Summative Assessment (Assessment of Learning)

- End Semester Examination, Lab Performance, Viva-voce.

## XI. SUGGESTED COS - POS MATRIX FORM

Course Outcomes (COs)	Programme Outcomes (POs)							Programme Specific Outcomes* (PSOs)		
	PO-1 Basic and Discipline Specific Knowledge	PO-2 Problem Analysis	PO-3 Design/ Development of Solutions	PO-4 Engineering Tools	PO-5 Engineering Practices for Society, Sustainability and Environment	PO-6 Project Management	PO-7 Life Long Learning	PSO- 1	PSO- 2	PSO- 3
CO1	3	2	2	3	1	1	3			
CO2	3	2	2	3	1	1	3			
CO3	3	2	2	3	1	1	3			
CO4	3	3	3	3	1	3	3			
CO5	3	3	3	3	1	3	3			

Legends :- High:03, Medium:02,Low:01, No Mapping: -  
 \*PSOs are to be formulated at institute level

## XII. SUGGESTED LEARNING MATERIALS / BOOKS

Sr.No	Author	Title	Publisher with ISBN Number
1	Yuxi Liu	Python machine learning by examples	Packt publication ISBN:978-178355-311-2
2	Saikat Dutt,Subramanian ChandramouliAmit Kumar Das	Machine Learning	Person ISBN 978-93-530-6669-7,eISBN: 978-93-895-8813-2
3	Tom M Mitchell	Machine Learning	McGraw Hill, First edition ISBN: 13: 978-0070428072
4	Aurélien Géron	Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems	O'Reilly Media ISBN-10 1492032646
5	Sebastian Raschka, Vahid Mirjalili	Python Machine Learning	Packt Publishing ISBN: 978-1800567703

## XIII . LEARNING WEBSITES & PORTALS

Sr.No	Link / Portal	Description
1	<a href="https://machinelearningmastery.com/how-machine-learning-algorithms-work/">https://machinelearningmastery.com/how-machine-learning-algorithms-work/</a>	ML with Python

Sr.No	Link / Portal	Description
2	<a href="https://www.geeksforgeeks.org/machine-learning/">https://www.geeksforgeeks.org/machine-learning/</a>	ML Introduction
3	<a href="https://www.w3schools.com/python/python_ml_getting_started.asp">https://www.w3schools.com/python/python_ml_getting_started.asp</a>	ML Python Basics
4	<a href="https://www.coursera.org/learn/machine-learning-with-python">https://www.coursera.org/learn/machine-learning-with-python</a>	ML with Python
5	<a href="https://www.kdnuggets.com/5-free-courses-to-master-machine-learning">https://www.kdnuggets.com/5-free-courses-to-master-machine-learning</a>	Free ML Courses
6	<a href="https://developers.google.com/machine-learning/crash-course">https://developers.google.com/machine-learning/crash-course</a>	Google ML Crash Course - Google ML Guide
7	<a href="https://www.youtube.com/@machinelearning-sudeshnasa3607">https://www.youtube.com/@machinelearning-sudeshnasa3607</a>	ML Tutorials Video(NPTEL)
8	<a href="https://www.youtube.com/watch?v=ukzFI9rgwfU">https://www.youtube.com/watch?v=ukzFI9rgwfU</a>	ML Overview
9	<a href="https://ekumbh.aicte-india.org/allbook.php">https://ekumbh.aicte-india.org/allbook.php</a>	Python Programming
10	<a href="https://arxiv.org/abs/2006.10092">https://arxiv.org/abs/2006.10092</a>	Housing Market Prediction Problem using Different Machine Learning Algorithms: A Case Study

### Note :

- Teachers are requested to check the creative common license status/financial implications of the suggested online educational resources before use by the students