

Level : M. Sc.

Year : I

Program : MSNCS

Part : I

**Course Objectives:**

The objective of this course is to provide a fundamental understanding of Machine Learning (ML), Deep Learning, and Data Analytics. This course also explores the understanding of Supervised and unsupervised learning techniques, probability-based learning techniques, performance evaluation of ML algorithms, and applications of ML such as in information /cyber security, etc.

Learning Outcomes	Chapter Contents	Credit Hours	Teaching Methods
<ul style="list-style-type: none"> <li>Understand the history, definition, and types of machine learning.</li> <li>Review basic statistical concepts relevant to machine learning.</li> <li>Comprehend fundamental machine learning terminology.</li> <li>Differentiate between training, validation, and test data.</li> <li>Understand key concepts like generalization tradeoff, bias-variance tradeoff, and learning curves.</li> </ul>	<p><b>1. Basics of Machine Learning (6 hrs)</b></p> <p>1.1 History of machine learning, definition of learning, types of learning, and importance of machine learning</p> <p>1.2 Review of Statistics: Min., Max., Mean, Mode, Median, Standard deviation, MSE</p> <p>1.3 Basics of machine learning terminology: class, pattern, feature, training, validation and test data</p> <p>1.4 Feasibility of learning – error and noise – training versus testing</p> <p>1.5 Generalization tradeoff – bias and variance – learning curve</p> <p>1.6 Overfitting and Underfitting</p>	6	<ul style="list-style-type: none"> <li>Lectures with real-world examples.</li> <li>Hands-on exercises using datasets for statistical calculations.</li> <li>Interactive discussions and case studies.</li> <li>Visual illustrations of overfitting and underfitting.</li> </ul>

<ul style="list-style-type: none"> <li>Identify overfitting and underfitting problems.</li> </ul>			
<ul style="list-style-type: none"> <li>Understand the process of data analytics and its key steps.</li> <li>Differentiate between various data types and attributes.</li> <li>Learn data pre-processing techniques for machine learning.</li> <li>Utilize data visualization methods for exploration.</li> <li>Understand architectural design patterns for handling Big Data.</li> <li>Identify different types of analytics (descriptive, diagnostic, predictive, prescriptive).</li> </ul>	<p>Unit 2: Data Analytics Process</p> <p>2.1 Process of data analytics</p> <p>2.2 Data types and attributes</p> <p>2.3 Data pre-processing</p> <p>2.4 Visualization and exploring data</p> <p>2.6 Architectural design patterns and stack for handling Big Data</p> <p>2.5 Descriptive, diagnostic, predictive, prescriptive analytics</p>	9	<ul style="list-style-type: none"> <li>Hands-on labs on data pre-processing and visualization.</li> <li>Group discussions on Big Data handling techniques.</li> <li>Practical exercises using Python libraries (Pandas, Matplotlib, Seaborn).</li> </ul>
<ul style="list-style-type: none"> <li>Understand the concept of supervised learning and classification problems.</li> <li>Learn about classifiers and discriminant functions.</li> <li>Implement linear supervised learning models such as linear regression and perceptron.</li> <li>Comprehend neural network structures and decision tree models.</li> </ul>	<p><b>Chapter #3: Supervised Learning</b></p> <p>3.1 Definition and classification problem</p> <p>3.2 Classifiers and discriminant functions</p> <p>3.3 Linear supervised learning models: linear regression, Perceptron</p> <p>3.4 Learning neural network structures</p>	9	<ul style="list-style-type: none"> <li>Coding exercises using Python and Scikit-Learn.</li> <li>Hands-on implementation of classifiers.</li> <li>Comparative analysis of different supervised learning models.</li> <li></li> </ul>

<ul style="list-style-type: none"> <li>• Explore support vector machines and their applications.</li> </ul>	<p>3.5 Decision tree representation model, basic decision tree algorithm, and application</p> <p>3.6 Support vector machines and applications</p>		
<ul style="list-style-type: none"> <li>• Understand Bayes' probability theory and conditional probability.</li> <li>• Analyze decision surfaces and classification using Bayes decision theory.</li> <li>• Explore Bayesian belief networks and their applications.</li> <li>• Implement the gradient descent method for optimization.</li> <li>• Understand K-nearest neighbor (KNN) algorithm.</li> <li>•</li> </ul>	<p><b>Unit 4: Bayesian Decision based learning (9 hrs)</b></p> <p>4.1 Bayes probability theory and conditional probability</p> <p>4.2 Decision surfaces and classifying with Bayes decision theory</p> <p>4.3 Bayesian belief network and applications</p> <p>4.4 Gradient descent method</p> <p>4.5 K-nearest neighbor</p>	9	<ul style="list-style-type: none"> <li>• Mathematical derivations and problem-solving sessions.</li> <li>• Algorithmic implementations using Python.</li> <li>• Real-life applications of Bayesian methods.</li> <li>•</li> </ul>
<ul style="list-style-type: none"> <li>• Understand the concept of clustering and different clustering algorithms.</li> <li>• Implement K-means, hierarchical, and other clustering techniques.</li> <li>• Comprehend the importance of dimensionality reduction.</li> <li>• Apply Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA).</li> </ul>	<p><b>Un-Supervised learning and dimensionality reduction</b></p> <p>5.1 Introduction to clustering, criterion function for clustering</p> <p>5.2 Algorithms for clustering; K-means, hierarchical, and other methods</p> <p>5.3 Dimensionality reduction techniques and need</p>	9Hrs	<ul style="list-style-type: none"> <li>• Interactive demonstrations of clustering techniques.</li> <li>• Implementation of dimensionality reduction in Python.</li> <li>• Case studies on real-world datasets.</li> <li>•</li> </ul>

<ul style="list-style-type: none"> <li>•</li> </ul>	5.4 Principal component analysis (PCA)  5.5 Linear discriminant analysis (LDA)		
<ul style="list-style-type: none"> <li>• Evaluate classification accuracy.</li> <li>• Construct and interpret confusion matrices.</li> <li>• Analyze misclassification costs.</li> <li>• Understand precision, recall, F1-score, and ROC curves.</li> <li>• Conduct cross-validation for performance assessment.</li> <li>•</li> </ul>	<b>Measures for Performance Evaluation</b>  6.1 Classification accuracy 6.2 Confusion matrix 6.3 Misclassification costs 6.4 Sensitivity and specificity, recall, precision, and F1-score 6.5 ROC curve, box plot, confidence interval 6.6 Cross-validation	9Hrs	<ul style="list-style-type: none"> <li>• Practical sessions on evaluating ML models.</li> <li>• Use of visualization tools like ROC plots.</li> <li>• Case studies on model performance assessment.</li> </ul>
<ul style="list-style-type: none"> <li>• Define deep learning and neural networks.</li> <li>• Understand feed-forward and backpropagation concepts.</li> <li>• Implement activation functions (Sigmoid, Tanh, ReLU, Softmax).</li> <li>• Learn about CNN and RNN architectures.</li> <li>• Explore ML applications in security (anomaly detection, fraud detection, etc.)</li> <li>•</li> </ul>	<b>Deep Learning Basic</b> 7.1 Definition of deep networks 7.2 Feed-Forward and backpropagation 7.3 Activation functions sigmoid, Tanh, ReLU and Softmax 7.4 Convolution neural networks: CNN architectures 7.5 Recurrent neural networks: RNN architectures 7.6 ML applications in Security 7.6.1 Anomaly detection /intrusion detection 7.6.2 Malware/phishing / fraud detection	9Hrs	<ul style="list-style-type: none"> <li>• Hands-on coding in TensorFlow and Keras.</li> <li>• Case studies on security-related ML applications.</li> <li>• Interactive discussions on deep learning architectures.</li> <li>•</li> </ul>

## Practical

Practical work should be done covering all the topics listed above and a small project work should be carried out using the concepts learned in this course using software like Matlab and Python.

## Evaluation Scheme:

The question will cover all chapters of the syllabus. The evaluation scheme will be as indicated in the table:

Units	Chapters	Marks
1	Basics of Machine Learning	7
2	Data Analytics Process	9
3	Supervised Learning	9
4	Bayesian Decision-based learning	9
5	Un-supervised learning and dimensionality reduction	9
6	Measures for Performance Evaluation	8
7	Deep Learning Basic	9
<b>Total</b>		<b>60</b>

---

## References:

1. Pablo Duboue, *The Art of Feature Engineering: Essentials for Machine Learning*, Cambridge University Press, First Edition, 2020
2. Christopher M. Bishop, *Pattern Recognition and Machine Learning*, Springer, First Edition, 2011
3. Kevin P. Murphy, *Machine Learning: A Probabilistic Perspective*, MIT Press, First Edition, 2012
4. Oliver Theobald, *Machine Learning For Absolute Beginners*, Kindle Edition, 2017
5. Geron Aurelien, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*, O'Reilly Media, Inc. 2019
7. Ian Goodfellow, Yoshua Bengio, Aaron Courville. *Deep Learning*, MIT Press. 2016

---