

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

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

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

<ul style="list-style-type: none"> • 	5.4 Principal component analysis (PCA) 5.5 Linear discriminant analysis (LDA)		
<ul style="list-style-type: none"> • Evaluate classification accuracy. • Construct and interpret confusion matrices. • Analyze misclassification costs. • Understand precision, recall, F1-score, and ROC curves. • Conduct cross-validation for performance assessment. • 	Measures for Performance Evaluation 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"> • Practical sessions on evaluating ML models. • Use of visualization tools like ROC plots. • Case studies on model performance assessment.
<ul style="list-style-type: none"> • Define deep learning and neural networks. • Understand feed-forward and backpropagation concepts. • Implement activation functions (Sigmoid, Tanh, ReLU, Softmax). • Learn about CNN and RNN architectures. • Explore ML applications in security (anomaly detection, fraud detection, etc.) • 	Deep Learning Basic 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"> • Hands-on coding in TensorFlow and Keras. • Case studies on security-related ML applications. • Interactive discussions on deep learning architectures. •

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
Total		60

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
