

**UNIT - I****Chapter 1 : Introduction to Machine Learning****1-1 to 1-19**

1.1	Overview of Human Learning and Machine Learning .....	1-1
1.1.1	How does Machine Learning Work? .....	1-3
1.1.2	Key Terms Associated with Machine Learning .....	1-5
1.2	Types of Machine Learning .....	1-5
1.2.1	Supervised Learning .....	1-5
1.2.2	Unsupervised Learning .....	1-6
1.2.3	Reinforcement Learning .....	1-7
1.2.3(A)	Comparison between Supervised and Unsupervised Learning .....	1-8
1.2.4	How to Choose the Right Machine Learning Algorithm? .....	1-8
1.3	Issues in Machine Learning .....	1-10
1.4	Application of Machine Learning .....	1-11
1.5	Tools and Technology for Machine Learning .....	1-18

**UNIT - II****Chapter 2 : Preparing to Model****2-1 to 2-17**

2.1	Steps in Developing a Machine Learning Application (Machine Learning Activities) .....	2-1
2.2	Data Formats (Types of data in Machine Learning - Structures of Data) .....	2-6
2.3	DIKW Pyramid .....	2-7
2.4	Categories of Data Analytics (Selecting a Machine Learning Model) .....	2-8
2.4.1	Comparison between Categories of Data Analytics .....	2-10
2.5	Data Quality and Remediation .....	2-10
2.5.1	Common Data Quality Issues .....	2-10
2.5.2	Remediating (Fixing) Data Quality Issues .....	2-12

2.6	Data Pre-Processing (Dimensionality Reduction Techniques) .....	2-14
2.6.1	Types of Dimensionality Reduction Techniques .....	2-16

**UNIT - III****Chapter 3 : Modelling and Evaluation****3-1 to 3-3**

3.1	Selecting a Model : Predictive/Descriptive .....	3-1
3.2	Training a Model for supervised learning .....	3-1
3.3	Evaluating Performance of a Model .....	3-1
3.4	Improving Performance of a Model .....	3-1
3.5	Model Representation and Interpretability .....	3-1
3.5.1	Overfitting .....	3-2
3.5.2	Underfitting .....	3-3

**UNIT - IV****Chapter 4 : Basics of Feature Engineering****4-1 to 4-32**

4.1	Feature and Feature Engineering .....	4-1
4.1.1	Data .....	4-1
4.1.2	Tasks .....	4-1
4.1.3	Models .....	4-2
4.1.4	Features .....	4-3
4.1.5	Feature Engineering .....	4-3
4.1.6	Data Engineering -vs- Feature Engineering .....	4-5
4.2	Feature Transformation .....	4-6
4.2.1	Feature Construction .....	4-6
4.2.2	Quantization or Binning .....	4-7
4.2.3	Log Transform .....	4-8
4.2.4	Feature Scaling or Normalisation .....	4-9
4.2.5	Min-Max Scaling .....	4-9
4.2.6	Standardisation (Variance Scaling) .....	4-10



4.2.7	Encoding Categorical Variables .....	4-12	5.2.2	Difference of Means.....	5-8
4.2.8	One-Hot Encoding.....	4-13	5.2.3	Type I and Type II Errors.....	5-9
4.2.9	Dummy Coding.....	4-14	5.2.3(A)	Comparison between Type I and Type II errors ..	5-11
4.2.10	Feature Hashing.....	4-14	5.2.4	Power and Sample Size.....	5-11
4.2.11	Handling Textual Features .....	4-16	5.3	Concepts of Probability.....	5-12
4.2.12	Feature Extraction.....	4-17	5.3.1	Fundamental Rules of Probability.....	5-12
4.3	Principal Component Analysis (PCA) .....	4-17	5.4	Random Variables .....	5-13
4.4	Independent Component Analysis (ICA).....	4-19	5.4.1	Discrete Random Variables .....	5-14
4.4.1	How ICA Works?.....	4-21	5.4.2	Continuous Random Variables.....	5-14
4.4.2	Comparison between PCA and ICA .....	4-22	5.4.3	Multiple Random Variables .....	5-15
4.5	Singular Value Decomposition (SVD) .....	4-23	5.5	Common Discrete Distributions .....	5-15
4.5.1	How SVD Works? .....	4-24	5.5.1	The Binomial Distribution.....	5-15
4.6	Feature(Subset) Selection .....	4-25	5.5.2	The Bernoulli Distribution .....	5-17
4.7	Issues in High-Dimensional Data .....	4-25	5.5.3	The Multinomial Distribution .....	5-18
4.8	Key Drivers of Feature Selection.....	4-25	5.5.4	Multinoulli Distribution.....	5-20
4.8.1	Measures of Feature Relevance .....	4-26	5.5.5	The Poisson distribution.....	5-20
4.8.2	Measures of Feature Redundancy .....	4-27	5.5.6	The Empirical Distribution .....	5-22
4.9	Overall Feature Selection Process .....	4-28	5.6	Common Continuous Distributions .....	5-23
4.10	Feature Selection Approaches.....	4-29	5.6.1	Uniform Distribution .....	5-23

## UNIT - V

### **Chapter 5 : Overview of Probability                    5-1 to 5-33**

5.1	Statistical Tools in Machine Learning.....	5-1
5.1.1	Sampling Distributions .....	5-2
5.1.2	General Statistics.....	5-3
5.1.2(A)	Standard Deviation .....	5-3
5.1.2(B)	Variance.....	5-4
5.1.2(C)	Covariance.....	5-4
5.2	Hypothesis Testing.....	5-6
5.2.1	How Hypothesis Testing Works? .....	5-7

5.6.2	Gaussian (Normal) Distribution .....	5-25
5.6.3	The Laplace Distribution.....	5-26
5.7	Central Limit Theorem (CLT) .....	5-27
5.8	Monte Carlo Methods (Approximation) .....	5-30

## UNIT - VI

### **Chapter 6 : Bayesian Concept Learning                    6-1 to 6-8**

6.1	Importance of Bayesian Methods.....	6-1
6.2	Bayes' Algorithm (Theorem) .....	6-2
6.3	Bayes' Theorem and Concept Learning.....	6-4
6.4	Bayesian Belief Network.....	6-5

**UNIT - VII****Chapter 7 : Supervised Learning : Classification and Regression 7-1 to 7-94**

7.1	Supervised Learning .....	7-1	7.5.2	Smoothing .....	7-59
7.2	Classification Model.....	7-1	7.5.3	Advantages of Naïve Bayes Classifier .....	7-61
7.2.1	Decision Trees .....	7-2	7.5.4	Disadvantages of Naïve Bayes Classifier .....	7-61
7.2.2	Key Terms and Concepts.....	7-3	7.6	Support Vector Machines (SVM) .....	7-62
7.2.2(A)	Entropy .....	7-3	7.6.1	Maximum Margin Linear Separators .....	7-63
7.2.2(B)	Information Gain .....	7-6	7.6.2	Quadratic Programming Solution to Finding Maximum Margin Separators .....	7-64
7.2.2(C)	Gain Ratio .....	7-11	7.6.3	Kernels for Learning Non-Linear Functions (Kernel Trick).....	7-66
7.2.2(D)	Gini Index.....	7-15	7.6.4	Comparison between Logistic Regression and SVM.....	7-67
7.3	Decision Tree Algorithms .....	7-29	7.7	k-Nearest Neighbours (kNN) Classification Algorithm .....	7-68
7.3.1	The General Algorithm.....	7-29	7.8	Diagnostics (Evaluation Measures) of Classifiers.....	7-68
7.3.2	ID3 Algorithm .....	7-29	7.8.1	ROC Curve .....	7-71
7.3.3	C4.5 Algorithm .....	7-34	7.8.2	Area Under the Curve (AUC).....	7-72
7.3.4	CART Algorithm.....	7-34	7.9	Additional Classification Methods.....	7-73
7.3.5	Evaluating a Decision Tree.....	7-35	7.9.1	Bagging .....	7-73
7.4	Rule-based Classification .....	7-35	7.9.2	Boosting.....	7-75
7.4.1	Basic Components of a Rule-based Classifier .....	7-36	7.9.3	Random Forests .....	7-76
7.4.2	Rule Coverage and Accuracy.....	7-39	7.10	Regression Analysis.....	7-77
7.4.3	Desirable Characteristics of a Rule-based Classifier.....	7-43	7.10.1	Linear Regression.....	7-78
7.4.4	Managing Rule Conflicts .....	7-46	7.10.1(A)	Use Cases (or Applications of) for Linear Regression .....	7-83
7.4.4(A)	Ordered Ruleset.....	7-47	7.10.2	Logistic Regression.....	7-84
7.4.4(B)	Unordered Ruleset.....	7-48	7.10.2(A)	Use Cases (or Applications of) for Logistic Regression.....	7-85
7.4.5	Default Rule .....	7-49	7.11	Reasons to Choose and Cautions.....	7-86
7.4.6	Building Rule-based Classifier.....	7-49	7.12	Additional Regression Models .....	7-86
7.5	Naïve Bayes (Classification by Bayesian Belief Networks).....	7-51			
7.5.1	Naïve Bayes Classifier.....	7-51			

**UNIT - VIII****Chapter 8 : Unsupervised Learning 8-1 to 8-50**

8.1	Supervised vs. Unsupervised Learning.....	8-1
8.2	Clustering.....	8-1
8.2.1	Properties of a Cluster.....	8-3
8.2.2	Types of Clustering .....	8-3
8.2.3	Use Cases (Applications) of Clustering .....	8-4
8.2.4	K-means.....	8-6
8.2.5	Determining the Number of Clusters.....	8-15
8.2.6	Diagnostics.....	8-19
8.2.7	Reasons to Choose and Cautions (Drawbacks / Challenges) .....	8-20
8.3	k-Nearest Neighbours (kNN) Classification Algorithm .....	8-21
8.4	Association Rules.....	8-23
8.4.1	Key Terms and Properties of Association Rules..	8-24
8.4.1(A)	Itemset .....	8-24
8.4.1(B)	Support .....	8-25
8.4.1(C)	Confidence.....	8-27
8.4.1(D)	Lift .....	8-28
8.4.1(E)	Leverage.....	8-31
8.4.1(F)	Subsets .....	8-32
8.4.2	Evaluation of Candidate Rules.....	8-33
8.4.3	Apriori Algorithm .....	8-33
8.4.3(A)	How Apriori Algorithm Works.....	8-34

8.4.3(B)	Case Study – Transactions in Grocery Store.....	8-39
8.4.3(C)	Validation and Testing.....	8-44
8.4.3(D)	Diagnostics.....	8-44

**UNIT - IX****Chapter 9 : Neural Network 9-1 to 9-32**

9.1	Introduction to Neural Networks .....	9-1
9.1.1	Fundamental Concept.....	9-1
9.1.2	Biological Neuron .....	9-3
9.1.3	Evolution of Neural Networks (Hebb's Rule) .....	9-4
9.2	Artificial Neural Networks (ANN) .....	9-5
9.2.1	The Perceptron.....	9-10
9.2.2	The Bias Input .....	9-11
9.2.3	Multilayer Perceptron .....	9-12
9.2.4	Shallow and Deep Neural Network.....	9-13
9.3	Neural Network (NN) Architecture .....	9-14
9.4	Activation Functions .....	9-16
9.4.1	Types of Activation Functions.....	9-17
9.5	Learning Process.....	9-21
9.6	Backpropagation.....	9-22
9.6.1	How Backpropagation Works? .....	9-24
9.6.2	Forward Propagation .....	9-27
9.6.3	Error Derivative .....	9-28
9.6.4	Additional Derivatives .....	9-29
9.6.5	Backpropagation.....	9-30