

## UNIT I

### Chapter 1 : Introduction to Machine Learning 1-1 to 1-11

**Syllabus :** Classic and adaptive machines, Machine learning matters, Beyond machine learning-deep learning and bio inspired adaptive systems, Machine learning and Big data. Important Elements of Machine Learning-Data formats, Learnability

|       |                                                                              |      |
|-------|------------------------------------------------------------------------------|------|
| 1.1   | Classic and Adaptive Machines .....                                          | 1-1  |
| 1.2   | Machine Learning Matters.....                                                | 1-2  |
| 1.2.1 | Role of Machine Learning Algorithms in Spam Filtering .....                  | 1-3  |
| 1.2.2 | Role of Machine Learning Algorithms in Natural Language Processing.....      | 1-4  |
| 1.3   | Beyond Machine Learning-Deep Learning and Bio-Inspired Adaptive Systems..... | 1-4  |
| 1.3.1 | Deep Learning .....                                                          | 1-4  |
| 1.3.2 | Difference between Machine Learning and Deep Learning .....                  | 1-5  |
| 1.3.3 | Bio-Inspired Adaptive Systems .....                                          | 1-6  |
| 1.4   | Machine Learning and Big Data .....                                          | 1-6  |
| 1.5   | Important Elements of Machine Learning .....                                 | 1-7  |
| 1.5.1 | Data Formats .....                                                           | 1-7  |
| 1.5.2 | Parametric Learning.....                                                     | 1-8  |
| 1.5.3 | Non-parametric Learning .....                                                | 1-8  |
| 1.5.4 | Multiclass Strategies .....                                                  | 1-9  |
| 1.6   | Learnability .....                                                           | 1-9  |
| 1.6.1 | Underfitting and Overfitting .....                                           | 1-10 |
| 1.6.2 | Error Measures .....                                                         | 1-10 |
| 1.6.3 | PAC Learning .....                                                           | 1-11 |

## UNIT II

### Chapter 2 : Feature Selection 2-1 to 2-34

**Syllabus :** Scikit-learn Dataset, Creating training and test sets, managing categorical data, Managing missing features, Data scaling and normalization, Feature Selection and Filtering, Principle Component Analysis(PCA) - Non negative matrix factorization, Sparse PCA, Kernel PCA, Atom Extraction and Dictionary Learning.

|       |                                                                                                                                                                 |      |
|-------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| 2.1   | Introduction to Feature Selection .....                                                                                                                         | 2-1  |
| 2.2   | About Scikit-learn Library .....                                                                                                                                | 2-1  |
| 2.2.1 | Scikit - Learn Datasets [Ref URL : <a href="https://scikit-learn.org/stable/datasets/index.html">https://scikit-learn.org/stable/datasets/index.html</a> ]..... | 2-2  |
| 2.2.2 | The Toy Datasets .....                                                                                                                                          | 2-2  |
| 2.2.3 | Real World Datasets .....                                                                                                                                       | 2-8  |
| 2.2.4 | Generated Datasets.....                                                                                                                                         | 2-10 |
| 2.3   | Creating Training and Test Sets.....                                                                                                                            | 2-10 |
| 2.4   | Types of Data .....                                                                                                                                             | 2-11 |
| 2.4.1 | Managing Categorical Data.....                                                                                                                                  | 2-14 |
| 2.5   | Managing Missing Features .....                                                                                                                                 | 2-17 |
| 2.6   | Data Scaling and Normalization .....                                                                                                                            | 2-17 |



|        |                                                                         |      |
|--------|-------------------------------------------------------------------------|------|
| 2.6.1  | Data Scaling .....                                                      | 2-17 |
| 2.6.2  | Data Normalization .....                                                | 2-19 |
| 2.7    | Feature Selection and Filtering .....                                   | 2-19 |
| 2.7.1  | Filter Methods .....                                                    | 2-20 |
| 2.7.2  | Wrapper Methods .....                                                   | 2-20 |
| 2.7.3  | Embedded Methods .....                                                  | 2-20 |
| 2.8    | Principle Component Analysis (PCA) .....                                | 2-20 |
| 2.8.1  | Pre-requisites for PCA (Mathematics Background) .....                   | 2-20 |
| 2.8.2  | Eigen Values and Eigen Vectors .....                                    | 2-23 |
| 2.8.3  | Steps for Computing PCA .....                                           | 2-26 |
| 2.9    | Non-negative Matrix Factorization .....                                 | 2-29 |
| 2.10   | Sparse PCA .....                                                        | 2-30 |
| 2.10.1 | Sparse Principal Component Analysis using the Lasso (Elastic Net) ..... | 2-31 |
| 2.11   | Kernel PCA .....                                                        | 2-31 |
| 2.12   | Atom Extraction and Dictionary Learning .....                           | 2-33 |

### UNIT III

## Chapter 3 : Regression

3-1 to 3-36

**Syllabus : Linear Regression :** Linear Models, A bi-dimensional example, Linear Regression and higher dimensionality, Ridge, Lasso and ElasticNet, Robust regression with random sample consensus, Polynomial regression, Isotonic regression.

**Logistic Regression :** Linear classification, Logistic regression, Implementation and Optimizations, Stochastic gradient descent algorithms, Finding the optimal hyper-parameters through grid search, Classification metric, ROC Curve.

|          |                                                      |      |
|----------|------------------------------------------------------|------|
| 3.1      | Introduction to Regression Analysis .....            | 3-1  |
| 3.1.1    | Steps in Conducting a Regression Analysis .....      | 3-1  |
| 3.2      | Linear Regression : A Bi-dimensional Example .....   | 3-6  |
| 3.2.1    | The Equation of Line .....                           | 3-8  |
| 3.2.2    | Simple Linear Regression Model .....                 | 3-9  |
| 3.2.2(A) | Least Squares Method .....                           | 3-10 |
| 3.3      | Step by Step Procedure for Linear Regression .....   | 3-10 |
| 3.3.1    | Is the Regression Line Really Good ? .....           | 3-15 |
| 3.4      | Linear and Higher Dimensionality .....               | 3-18 |
| 3.5      | Ridge Regression .....                               | 3-19 |
| 3.5.1    | Lasso Regression .....                               | 3-20 |
| 3.5.2    | ElasticNet Regression .....                          | 3-21 |
| 3.6      | Robust Regression with Random Sample Consensus ..... | 3-21 |
| 3.6.1    | Advantages and Disadvantages of RANSAC .....         | 3-21 |
| 3.7      | Polynomial Regression .....                          | 3-22 |
| 3.7.1    | The Bias vs. Variance Trade-off .....                | 3-25 |
| 3.8      | Isotonic Regression .....                            | 3-25 |



|        |                                                               |      |
|--------|---------------------------------------------------------------|------|
| 3.9    | Linear Classification.....                                    | 3-26 |
| 3.10   | Logistic Regression.....                                      | 3-27 |
| 3.11   | Implementation and Optimizations .....                        | 3-30 |
| 3.12   | Stochastic Gradient Descent Algorithms .....                  | 3-31 |
| 3.12.1 | Introduction to Gradient Descent.....                         | 3-31 |
| 3.12.2 | Stochastic Gradient Descent.....                              | 3-33 |
| 3.13   | Finding the Optimal Hyper-parameters through Grid Search..... | 3-33 |
| 3.14   | Classification Metric.....                                    | 3-34 |
| 3.15   | ROC Curve .....                                               | 3-36 |

## UNIT IV

### Chapter 4 : Naïve Bayes and Support Vector Machine

4-1 to 4-22

**Syllabus :** Bayes' Theorem, Naïve Bayes' Classifiers, Naïve Bayes in Scikit-learn - Bernoulli Naïve Bayes, Multinomial Naïve Bayes, and Gaussian Naïve Bayes. Support Vector Machine (SVM) : Linear Support Vector Machines, Scikit-learn implementation- Linear Classification, Kernel based classification, Non-linear Examples. Controlled Support Vector Machines, Support Vector Regression.

|        |                                                          |      |
|--------|----------------------------------------------------------|------|
| 4.1    | Introduction to Naïve Bayes .....                        | 4-1  |
| 4.1.1  | Bayes Theorem .....                                      | 4-1  |
| 4.1.2  | Maximum A Posteriori (MAP) Hypothesis.....               | 4-2  |
| 4.1.3  | Maximum Likelihood .....                                 | 4-2  |
| 4.2    | Naïve Bayes Classifier.....                              | 4-2  |
| 4.2.1  | Example of Naïve Bayesian Classifier.....                | 4-3  |
| 4.2.2  | Properties of Bayes Classifiers .....                    | 4-5  |
| 4.3    | Bayes in Scikit-Learn .....                              | 4-5  |
| 4.4    | Bernoulli Naïve Bayes.....                               | 4-5  |
| 4.5    | Multinomial Naïve Bayes.....                             | 4-7  |
| 4.6    | Gaussian Naïve Bayes .....                               | 4-8  |
| 4.7    | Application of Naïve Bayes .....                         | 4-9  |
| 4.8    | Introduction to Support Vector Machine (SVM) .....       | 4-10 |
| 4.8.1  | Tuning Hyperparameters .....                             | 4-12 |
| 4.8.2  | Problems Faced by SVM when used with Real Datasets ..... | 4-13 |
| 4.9    | Linear Support Vector Machines .....                     | 4-13 |
| 4.10   | Scikit-Learn Implementation.....                         | 4-16 |
| 4.10.1 | Linear Based Classification .....                        | 4-16 |
| 4.10.2 | Kernel Based Classification .....                        | 4-19 |
| 4.10.3 | Non Linear Examples.....                                 | 4-20 |
| 4.11   | Controlled Support Vector Machines.....                  | 4-21 |
| 4.12   | Support Vector Regression .....                          | 4-22 |

**UNIT V****Chapter 5 : Decision Trees and Ensemble Learning****5-1 to 5-30**

**Syllabus :** Decision Trees : Impurity measures, Feature Importance. Decision Tree Classification with Scikit-learn, Ensemble Learning- Random Forest, AdaBoost, Gradient Tree Boosting, Voting Classifier. Clustering Fundamentals: Basics, K-means: Finding optimal number of clusters, DBSCAN, Spectral Clustering. Evaluation methods based on Ground Truth- Homogeneity, Completeness, Adjusted Rand Index. Introduction to Meta Classifier: Concepts of Weak and eager learner, Ensemble methods, Bagging, Boosting, Random Forests.

|          |                                                                                               |      |
|----------|-----------------------------------------------------------------------------------------------|------|
| 5.1      | Decision Trees.....                                                                           | 5-1  |
| 5.1.1    | Impurity measures .....                                                                       | 5-2  |
| 5.1.2    | Feature Importance.....                                                                       | 5-4  |
| 5.1.3    | Decision Tree Classification with Scikit-learn .....                                          | 5-5  |
| 5.1.4    | Ensemble Learning.....                                                                        | 5-6  |
| 5.1.4(A) | Random Forest.....                                                                            | 5-7  |
| 5.1.4(B) | Ada Boost.....                                                                                | 5-8  |
| 5.1.4(C) | Gradient Tree Boosting.....                                                                   | 5-9  |
| 5.1.4(D) | Voting Classifier.....                                                                        | 5-10 |
| 5.2      | Clustering Fundamentals .....                                                                 | 5-12 |
| 5.2.1    | Basics of Clustering .....                                                                    | 5-12 |
| 5.2.2    | K-means .....                                                                                 | 5-13 |
| 5.2.2(A) | K-means Algorithm .....                                                                       | 5-14 |
| 5.2.2(B) | Clustering with Scikit-learn.....                                                             | 5-15 |
| 5.2.2(C) | Finding Optimal Number of Clusters .....                                                      | 5-17 |
| 5.2.3    | DBSCAN.....                                                                                   | 5-22 |
| 5.2.4    | Spectral Clustering.....                                                                      | 5-26 |
| 5.3      | Evaluation Methods Based on Ground Truth- Homogeneity, Completeness, Adjusted Rand Index..... | 5-26 |
| 5.4      | Introduction to Meta Classifier.....                                                          | 5-29 |
| 5.4.1    | Concepts of Weak/Lazy and Eager Learner.....                                                  | 5-29 |

**UNIT VI****Chapter 6 : Clustering Techniques****6-1 to 6-18**

**Syllabus :** Hierarchical Clustering, Expectation maximization clustering, Agglomerative Clustering - Dendrograms, Agglomerative clustering In Scikit- learn, Connectivity Constraints. Introduction to Recommendation Systems : Naïve User based systems, Content based Systems, Model free collaborative filtering-singular value decomposition, alternating least squares. Fundamentals of Deep Networks : Defining Deep learning, common architectural principles of deep networks, the building blocks of deep networks.

|       |                                                |     |
|-------|------------------------------------------------|-----|
| 6.1   | Hierarchical Clustering.....                   | 6-1 |
| 6.2   | Expectation Maximization Clustering.....       | 6-2 |
| 6.3   | Agglomerative Clustering.....                  | 6-4 |
| 6.3.1 | Affinity .....                                 | 6-4 |
| 6.3.2 | Strategy to Aggregate Different Clusters.....  | 6-4 |
| 6.3.3 | Dendrograms.....                               | 6-6 |
| 6.3.4 | Agglomerative Clustering in Scikit-learn ..... | 6-8 |



|          |                                                                                                          |                   |
|----------|----------------------------------------------------------------------------------------------------------|-------------------|
| 6.3.5    | Connectivity Constraints .....                                                                           | 6-10              |
| 6.4      | Introduction to Recommendation Systems .....                                                             | 6-11              |
| 6.4.1    | Naïve User based Systems.....                                                                            | 6-11              |
| 6.4.2    | Content Based Systems.....                                                                               | 6-12              |
| 6.4.3    | Model Free Collaborative Filtering .....                                                                 | 6-12              |
| 6.4.4    | Singular Value Decomposition .....                                                                       | 6-15              |
| 6.4.5    | Alternating Least Squares.....                                                                           | 6-16              |
| 6.5      | Fundamentals of Deep Networks.....                                                                       | 6-16              |
| 6.5.1    | Defining Deep Learning .....                                                                             | 6-17              |
| 6.5.2    | Common Architectural Principles of Deep Networks .....                                                   | 6-17              |
| 6.5.2(A) | Multilayer Perceptron with Back-propagation : First Deep Learning Model .....                            | 6-17              |
| 6.5.2(B) | Deep Belief Networks (Hinton et al. 2006).....                                                           | 6-18              |
| 6.5.2(C) | Five Popular and Widely-used Deep Learning Architectures.....                                            | 6-18              |
| 6.5.2(D) | Building Blocks of Deep Networks .....                                                                   | 6-18              |
| •        | <b>Appendix A : Solved University Question Papers of In Sem (March 2019) and End Sem (May 2019).....</b> | <b>A-1 to A-9</b> |

