

<b>Program</b>	Post Graduate Diploma in Data Science
<b>Semester</b>	1
<b>Subject Code and Name</b>	1618003 Fundamentals of Machine Learning
<b>Credit</b>	5

### Objectives

- To learn basic concepts of supervised machine learning methods.
- To learn mathematical concepts, and algorithms used in machine learning

<b>Unit No.</b>	<b>Topic(s)</b>	<b>No. of Hours</b>
1.	<b>Introduction</b> Learning Problems, designing a learning system, Issues with machine learning. Concept Learning, Version Spaces and Candidate Eliminations, Inductive bias, Supervised, Unsupervised and Semi supervised Learning	4
2.	<b>Supervised Learning</b> Decision Tree Representation, Appropriate problems for Decision tree learning, Algorithm, Hypothesis space search in Decision tree learning, inductive bias in Decision tree learning, Issues in Decision tree learning, Overfitting, Underfitting, Pre-pruning, Post-pruning	10
3.	<b>Artificial Neural Networks</b> Neural Network Representation, Appropriate problems for Neural Network Learning, perceptron, Cost function, Gradient descent, Hypothesis formation, Multilayer Networks and Back Propagation Algorithms, Remarks on Back Propagation Algorithms Case Study: face Recognition	10
4.	<b>Bayesian Learning</b> Bayes Theorem, Bayes Theorem and Concept Learning, Maximum Likelihood and Least squared Error Hypothesis, Maximum likelihood hypothesis for Predicting probabilities, Minimum Description Length Principle, Bayes Optimal Classifier, Gibbs Algorithm, Naïve Bayes Classifier , Bayesian Belief Network, EM Algorithm Case Study: Learning to classify text	10
5.	<b>Overview of typical application areas such as Recommender System etc</b>	6

### Reference Books

1. Machine Learning  
by Tom M Mitchell  
McGraw Hill, ISBN: 0070428077
2. Machine Learning in Action  
by Peter Harrington  
Manning Publications Co., ISBN 9781617290183
3. The Elements of Statistical Learning: Data Mining, Inference, and Prediction  
by Trevor Hastie, Robert Tibshirani, Jerome H. Friedman  
Springer Series in Statistics, Second edition, ISBN 978-0-387-84858-7
4. Real-World Machine Learning

## **Outcomes**

After completion of subject, students would be able to:

- understand the concept of Machine learning and range of problems that can be solved by machine learning.
- compare different types of learning algorithms and apply machine learning concepts in real life problems.

**Suggested list of Practical (at least 10 practical are to be performed by students. These practical should cover majority of all topics of syllabus.)**

**This is the suggested list of practical but it may not be limited only to this list.**

1. Import the real world dataset from any of the public dataset repository. Apply various data functions to clean the data (i.e. handling missing values, normalization).
2. For cleaned dataset, construct the decision tree, visualize it and use it for classification of the new samples.
3. Apply various decision tree pruning methods by tuning the parameters of functions. Make a comparison chart based on it for five different real world datasets.
4. Measure the precision, recall, accuracy, and F-score of the decision tree classifier on any text classification using 10-fold cross-validation. Also apply tree pruning option. Compare the results without pruning and with pruning.
5. For cleaned dataset, construct the random forest, visualize it and use it for classification of the new samples.
6. On cleaned dataset, apply Naïve Bayes classification and compare its result with decision tree and random forest.
7. Develop a code to classify spam mail with Naïve Bayes.
8. Estimate the accuracy of Naïve Bayes algorithm using 10-fold cross validation on the house-votes-84 data set.
9. Develop a feed forward neural network with backpropagation function to improve a handwritten character recognition system.
10. Develop a neural network, use learning functions and tune the parameters to reduce the mean square error for recognizing the face.

\*\*\*\*\*