

Summer Training (2021-22) : “Machine Learning & its Applications”

A. Name of the Institute	Centre for Development of Advanced Computing, Mohali
B. Name/title of the Course	Summer Training Programme in “ Machine Learning & its Applications ”
C. Duration of Module	6 weeks
D. Eligibility Criteria for Participants	Engineering Students (Electronics/Computer Science/IT/Telecommunications or equivalent) with working knowledge of computers. Familiarity with computer programming, basic linear algebra (matrices, vectors, matrix-vector multiplication), and basic probability (random variables, basic properties of probability) is assumed.
E. Aims & Objectives of the Course	This course provides a broad introduction to machine learning, data mining, and statistical pattern recognition. Topics include: (i) Supervised learning (parametric/nonparametric algorithms, support vector machines, kernels, neural networks). (ii) Unsupervised learning (iii) Best practices in machine learning (bias/variance theory; innovation process in machine learning and AI). The course will also draw from numerous case studies and applications, so that you'll also learn how to apply learning algorithms to real world problems like computer vision, medical informatics, database mining, and other areas.
F. Min Number of participants	30 (Thirty)
G. Hardware Requirements (in case of offline classes)	<ul style="list-style-type: none"> • No. of Systems = 30 (one system per student) with following configuration : <ul style="list-style-type: none"> ○ Multi-core 64-bit CPU System with Minimum 4GB RAM and Minimum 20 GB Hard disk space. • Projector to present the lectures.

COURSE CONTENTS :

1. Introduction to Machine Learning

- What is Machine Learning
- Supervised Learning Introduction
- Unsupervised Learning Introduction
- Classification, Regression and Clustering

2. Linear Algebra

- Matrices and Vectors
- Addition and Scalar Multiplication
- Matrix Vector Multiplication
- Matrix Matrix Multiplication
- Matrix Multiplication Properties
- Inverse and Transpose

3. Introduction to Python Programming

- Overview of language, Features, Installation
- Python Basics: Variable, Data types, Strings, Operators & Expressions etc.
- Control Flow: Loops & Selection Statements, Input & Output, etc.
- Data Structures: Lists, Dictionaries, Tuples etc.
- Functions & Modules (List of Built in Functions: lambda, filter, map, sum, max, min, enumerate, reversed, sorted)
- Pandas: Getting started with pandas
- Python Standard Library & Python External Packages
- Data Visualization
- Object oriented concepts: Classes, Methods, Inheritance, Overriding, etc
- File Handling, Exception Handling

4. Regression

- Predicting house prices: A case study in regression
- Linear Regression Introduction
- Model Representation
- Regression ML block diagram
- Cost Function
- Gradient Descent
- Gradient Descent for Linear Regression
- Vectorized Implementation
- Polynomial Regression Introduction
- Multiple Features
- Gradient Descent For Multiple Variables
- Feature Scaling, Learning Rate
- Normal Equation, Normal Equation Noninvertibility
- Lab: Predicting house price with linear regression

Summer Training (2021-22) : “Machine Learning & its Applications”

5. Applying Machine Learning and performance measures

- Evaluating a Hypothesis
- Model Selection and Train/Validation/Test Sets
- Loading & exploring house sale data
- Splitting the data into training and test sets
- cross validation
- Diagnosing Bias vs. Variance
- Learning Curves
- Error Analysis, Error Metrics
- Lab: Predicting house price with linear regression

6. Classification

- Logistic Regression Introduction
- Representation Model
- Decision Boundaries
- Cost Function
- Gradient Descent
- Optimization Objectives
- Training and evaluating a classifier
- False positives, false negatives, and confusion matrices
- Classification ML block diagram
- Multiclass Classification: One-vs-all
- *Lab*: Classification with logistic regression

7. Classification : Neural Networks

- Non-linear Hypotheses
- Neurons and the Brain
- Model Representation
- Examples and Intuitions
- Multiclass Classification
- Cost Function
- Backpropagation Algorithm
- Gradient Checking
- Random Initialization
- *Lab*: Classification with Neural network

8. Clustering

- Unsupervised Learning: Introduction
- Basic issues in clustering
- clustering with K-Means Algorithm
- Optimization Objective
- Random Initialization
- What is your optimal K
- Lab Session

9. Experiments with Weka

- Introduction & Installation
- Data preprocessing
- Data Mining with Weka

10. Application Example: PhotoOCR

- Problem Description and Pipeline
- Sliding Windows

11. Project

Justification for Training Programme

Machine learning is the science of getting computers to act without being explicitly programmed. In the past decade, machine learning has given us self-driving cars, practical speech recognition, effective web search, and a vastly improved understanding of the human genome. Machine learning is so pervasive today that you probably use it dozens of times a day without knowing it. Many researchers also think it is the best way to make progress towards human-level AI. In this class, you will learn about the most effective machine learning techniques, and gain practice implementing them and getting them to work for yourself. More importantly, you'll learn about not only the theoretical underpinnings of learning, but also gain the practical know-how needed to quickly and powerfully apply these techniques to new problems

This course provides a broad introduction to machine learning, data mining, and statistical pattern recognition. Topics include: (i) Supervised learning (parametric/non-parametric algorithms, support vector machines, kernels, neural networks). (ii) Unsupervised learning (clustering, recommender systems). (iii) Best practices in machine learning (bias/variance theory; innovation process in machine learning and AI). The course will also draw from numerous case studies and applications, so that you'll also learn how to apply learning algorithms to real world problems like computer vision, audio, database mining, and other areas. At the end of the course, candidate should be able to conceptualize & design the approach for the data analysis problems.