INTRODUCTION TO MACHINE LEARNING - IITKGP

PROF. SUDESHNA SARKAR
Department of Computer Science and Engineering
IIT Kharagpur

INTENDED AUDIENCE : Elective course for UG, PG, BE, ME, MS, M.Sc, PhD
PRE-REQUISITES : Basic programming skills (in Python), algorithm design, basics of probability & statistics
INDUSTRIES APPLICABLE TO : Data science companies and many other industries value machine learning skills.

COURSE OUTLINE :

This course provides a concise introduction to the fundamental concepts in machine learning and popular machine learning algorithms. We will cover the standard and most popular supervised learning algorithms including linear regression, logistic regression, decision trees, k-nearest neighbour, an introduction to Bayesian learning and the naïve Bayes algorithm, support vector machines and kernels and neural networks with an introduction to Deep Learning. We will also cover the basic clustering algorithms. Feature reduction methods will also be discussed. We will introduce the basics of computational learning theory. In the course we will discuss various issues related to the application of machine learning algorithms. We will discuss hypothesis space, overfitting, bias and variance, tradeoffs between representational power and learnability, evaluation strategies and cross-validation. The course will be accompanied by hands-on problem solving with programming in Python and some tutorial sessions.

ABOUT INSTRUCTOR :

Prof. Sudeshna Sarkar is a Professor and currently the Head of the Department of Computer Science and Engineering at IIT Kharagpur. She completed her B.Tech. in 1989 from IIT Kharagpur, MS from University of California, Berkeley, and PhD from IIT Kharagpur in 1995. She served briefly as faculty at IIT Guwahati and at IIT Kanpur before joining IIT Kharagpur in 1998. Her research interests are in Machine Learning, Natural Language Processing, Data and Text Mining.

COURSE PLAN :

Week 01 : Introduction: Basic definitions, types of learning, hypothesis space and inductive bias, evaluation, cross-validation.

Week 02 : Linear regression, Decision trees, overfitting.

Week 03 : Instance based learning, Feature reduction, Collaborative filtering based recommendation. Probability

Week 04 : Probability and Bayes learning

Week 05 : Logistic Regression, Support Vector Machine, Kernel function and Kernel SVM.

Week 06 : Neural network: Perceptron, multilayer network, backpropagation, introduction to deep neural network.

Week 07 : Computational learning theory, PAC learning model, Sample complexity, VC Dimension, Ensemble learning.

Week 08 : Clustering: k-means, adaptive hierarchical clustering, Gaussian mixture model.