Machine Learning

Lecture Plan

Lecture 1: Announcement of methods of internal/class assessment and marking scheme, Syllabus, study material and resources, Motivation, Introduction to machine learning.

Lecture 2: Review of Linear Algebra: Concept of feature vector representation, Norms of the vectors, vector and matrix operations, vector space, basis vectors, linear independence, orthogonality.

Lecture 3: Vector sub-spaces, projection, Eigen vectors and Eigen values, properties of symmetric matrices.

Lecture 4: Review of Probability Theory: Definitions, independent events, joint probability, marginal probability, conditional probability, sum rule, product rule, Bayes' theorem, concept of probability distribution, likelihood.

Lecture 5: Random process and random variable: Definitions, continuous and discrete random variables, expectation, variance, covariance.

Lecture 6: Classification and Regression: curve fitting, model selection, curse of dimensionality, loss function.

Lecture 7: Evaluation of ML models: The train/test/validation split, under-fitting, overfitting, generalization, Bias vs Variance, validation curves.

Lecture 8: Metrics: Confusion matrix, Accuracy, Precision, Recall, Specificity, F1 score Precision-Recall or PR curve, ROC (Receiver Operating Characteristics) curve, PR vs ROC curve.

Lecture 9: Information Theory: Concept of information, Entropy, Information gain, relative and mutual information.

Lecture 10: Classification using Decision Trees: Iterative Dichotomiser 3 (ID3), Greedy decision tree learning, selecting best feature for split, classification error, prediction with decision trees.

Lecture 11: Decision trees with real valued features, threshold split in 1-D, threshold split in 2-D, finding optimal threshold split.

Lecture 12: Overfitting in decision tress: Principle of *Occam's Razor,* complex and simpler decision trees, early stopping, Decision tree pruning.

Lecture 13: CART, C4.5

Lecture 14: K_n –nearest neighbor density estimation, K-nearest neighbor classifier (K-NN).

Lecture 15: Naïve-Bayes Classifier

Lecture 16: Linear discriminant functions, logistic discrimination, Linear separability, generalized linear discriminants.

Lecture 17: Least-square techniques, gradient descent algorithms.

Lecture 18: Supervise Learning-Linear Regression, linear regression with one variable, Derivative of cost function, gradient descent algorithm.

Lecture 19: Logistic regression: Classification, learning parameters, cost function for logistic regression, gradient descent algorithm in logistic regression.

Lecture 20: Artificial Neural Network (ANN): Introduction, Perceptron model, applications of linear model.

Lecture 21: Perceptron learning, perceptron convergence theorem, limitations of perceptron.

Lecture 22: Fisher's linear Discriminant-Linear discriminant analysis (LDA).

Lecture 23: Multi-layer perceptron:

Lecture 24: Error back-propagation

Lecture 25: Support Vector Machine (SVM)

Lecture 26: Unsupervised Learning: similarity measures, k-means clustering, k-means as coordinate descent algorithm, k-mean++

Lecture 27: Convergence of k-means algorithm, limitations: uncertainty in cluster assignment, failure modes of k-means, mixture models.

Lecture 28: Gaussian mixture models (GMM), Maximum likelihood estimation (MLE), Expectation Maximization, Inferring soft assignments with expectation maximization (EM), Convergence and overfitting of MLE

Lecture 29: Hierarchal Clustering

Lecture 30: Data representation and Dimension Reduction: change of basis vectors, principle component analysis (PCA).

Lecture 31: Factor analysis, Manifold Learning

Lecture 32: Reinforcement learning: Introduction, difference with supervised learning, Evaluative feedback: n-armed bandit problem, action-value methods, softmax action selection

Lecture 33: The reinforcement learning problem: Agent-Environment interface, goals and rewards, returns, unified notation for episodic and continuing tasks, Markov property, Markov decision processes, value functions, optimal value functions.

Lecture 34: Dynamic programming: policy evaluation, policy improvement, policy iteration.

Lecture 35: Temporal-Difference Learning: TD prediction and advantages, Optimality of TD(0), Sarsa: On-Policy TD Control.

Lecture 36: Q-Learning: Off-Policy TD Control, Actor-Critic Methods.