Indian Institute of Technology, Kanpur

Department of Computer Science and Engineering

Revision of Existing Course CS771

Title: Introduction to Machine Learning

Course No: CS771

Units: 3-0-0-0-9

Pre-requisites: Instructor's consent (no course prerequisite).
Desirable: ESO207/equivalent, MSO201/equivalent, familiarity in a programming environment such as MATLAB/Octave/Python/R.

Proposed by: Purushottam Kar

Estimated Enrollment: 150-200.

Faculty members interested in teaching the course: Purushottam Kar (CSE), Vinay Namboodiri (CSE), Piyush Rai (CSE), Nisheeth Srivastava (CSE)

Departments which may be interested: CSE, EE, MTH, IME, ECO

Level of the course: PG (6xx level).

Short Description: Machine Learning is the discipline of designing algorithms that allow machines (e.g., a computer) to learn patterns and concepts from data without being explicitly programmed. This course will be an introduction to the design (and some analysis) of machine learning algorithms, with a modern outlook focusing on recent advances, and examples of real-world applications of machine learning algorithms.

Topics (total [40 lecs] of 50 minutes each)

- **Preliminaries**
  1. Multivariate calculus [**2 lecs**]: gradient, Hessian, Jacobian, chain rule
  2. Linear algebra [**2 lecs**]: determinants, eigenvalues/vectors, SVD
  3. Probability theory [**2 lecs**]: conditional probability, marginal probability, Bayes rule
- **Supervised Learning**
  4. Local/proximity-based methods [**2 lecs**]: nearest-neighbors, decision trees
  5. Learning by function approximation:
     - Linear models [**4 lecs**]: (multiclass) support vector machines, ridge regression
     - Non-linear models [**4 lecs**]: kernel methods, neural networks (feedforward)
  6. Learning by probabilistic modeling:
     - Discriminative methods [**2 lecs**]: (multiclass) logistic regression, generalized linear models
     - Generative methods [**1 lec**]: naïve Bayes
- **Unsupervised Learning**
  7. Discriminative Models [**4 lecs**]: k-means (clustering), PCA (dimensionality reduction)

8. Generative Models:
   - Latent variable models [**2 lecs**]: expectation maximization for learning latent variable models
   - Applications [**2 lecs**]: Gaussian mixture models, probabilistic PCA

- **Practical Aspects** [**3 lecs**]

9. Concepts of over-fitting and generalization, bias-variance tradeoffs
10. Model and feature selection using above concepts
11. Optimization for machine learning: (stochastic/mini-batch) gradient descent

- **Additional Topics [10 lecs]** (a subset to be covered depending on interest)

12. Deep learning: CNN, RNN, LSTM, autoencoders
13. Structured output prediction: multi-label classification, sequence tagging, ranking
14. Ensemble methods: boosting, bagging, random forests
15. Recommendation systems: ranking methods, collaborative filtering via matrix completion
16. Reinforcement learning and applications
17. Kernel extensions for PCA, clustering, spectral clustering, manifold learning
18. Probability density estimation and anomaly detection
19. Time-series analysis and modeling sequence data
20. Sparse modeling and estimation
21. Online learning algorithms: perceptron, Widrow-Hoff, explore-exploit
22. Statistical learning theory: PAC learning, VC dimension, generalization bounds
23. A selection from some other advanced topics such as semi-supervised learning, active learning, inference in graphical models, Bayesian learning and inference.

References: There will be no textbook for this course. Lecture notes, monographs and research papers at major machine learning venues including conferences such as NeurIPS, ICML, COLT, AISTATS, and journals such as Journal of Machine Learning Research, Machine Learning Journal, and IEEE Transactions on Information Theory will form the reference base for the course. Some recommended (though non compulsory) reference texts are listed below.

1. Christopher Bishop, Pattern Recognition and Machine Learning, Springer, 2007.
2. Hal Daumé III, A Course in Machine Learning, 2015 (freely available online).
3. Trevor Hastie, Robert Tibshirani, Jerome Friedman, The Elements of Statistical Learning, Springer 2009.
4. John Hopcroft, Ravindran Kannan, Foundations of Data Science, 2014 (freely available online).
5. Mehryar Mohri, Afshin Rostamizadeh, and Ameet Talwalkar. Foundations of Machine Learning, The MIT Press, 2012.
6. Kevin Murphy, Machine Learning: A Probabilistic Perspective, The MIT Press, 2012.

Signature of Course Proposer

Signature of DPGC Convenor