Revision of Existing Course
Department of Computer Science and Engineering
Indian Institute of Technology, Kanpur

Course number: CS771
Course title: Introduction to Machine Learning
Course prerequisites: None
Course credits: [12] (3-0-3-0)
Course duration: Full semester
Course type: UG + PG department elective
Proposing instructor: Purushottam Kar
Other faculty members interested in teaching the course: Piyush Rai (CSE)
Other departments interested in the proposed course: None

Course description:

a. *Objectives*: Machine Learning is the discipline of designing adaptive systems to identify and exploit data patterns to present useful predictive and analytical outputs. This course will introduce the basic building blocks of ML and their foundations. The course will also offer hands-on experience in building small-scale ML models.
b. *Logistics*: The course will serve as a DE for CSE PG students (MT, MS, PhD), a basket DE for CSE UG students (BT, BT-MT), a basket DE for students selected into the CSE double major program, and a compulsory course for students selected into the "*Machine Learning Applications*" minor offered by CSE. The course may be offered one or more times every academic year depending on demand and availability of resources.
c. *Content*: The course will have two parts:
    i. **Foundations**: this content will be offered in the form of lectures, demonstrations, etc. A weekly breakup is presented below.
    ii. **Applications**: hands-on experience will be offered with individual instructors curating student experience by selecting one or more components from the following or developing others:
        - One or more projects
        - In-person or virtual labs
        - Programming assignments
        - Reading exercises and seminars
d. *Evaluation*: Every offering of the course will specify, as a part of the "*First Course Handout*", evaluation points (exams, quizzes, projects, reading exercises, seminars, assignments, labs, etc) and weightages for the foundations and applications portions.

Breakup of lecture content: there will be an equivalent of 40 lectures of 50 minutes each. The numbers in square brackets [] against each topic indicate the number of lectures for that topic.

Preliminaries

i. Multivariate calculus [2]: gradient, Hessian, Jacobian, chain rule
ii. Linear algebra [2]: determinants, eigenvalues/vectors, SVD
iii. Probability theory [2]: conditional and marginal probability, Bayes rule

Supervised Learning

iv. Local/proximity-based methods [2]: nearest-neighbors, decision trees
v. Learning by function approximation [8]: linear models (SVMs, ridge regression), non-linear models (kernel methods, neural networks)
vi. Learning by probabilistic modeling [3]: discriminative models (logistic regression, generalized linear models), generative models (naïve Bayes)

Unsupervised Learning

vii. Discriminative Models [4]: clustering (e.g. k-means), dimensionality reduction (e.g. PCA)
viii. Generative Models [4]: latent variables (expectation maximization), applications (e.g. Gaussian mixture models, probabilistic PCA)

Practical Aspects [3]

ix. Over-fitting and generalization, bias-variance tradeoffs, model and feature selection
x. Optimization for machine learning: (stochastic/mini-batch) gradient descent

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

xi. Deep learning: CNN, RNN, LSTM, autoencoders
xii. Structured output prediction: multi-label classification, sequence tagging, ranking
xiii. Ensemble methods: boosting, bagging, random forests
xiv. Recommendation systems: ranking methods, collaborative filtering via matrix completion
xv. Reinforcement learning and applications
xvi. Kernel extensions for PCA, clustering, spectral clustering, manifold learning
xvii. Probability density estimation and anomaly detection
xviii. Time-series analysis and modeling sequence data
xix. Sparse modeling and estimation
xx. Online learning algorithms: perceptron, Widrow-Hoff, explore-exploit
xxi. Statistical learning theory: PAC learning, VC dimension, generalization bounds
xxii. A selection from some other advanced topics such as semi-supervised learning, active learning, inference in graphical models, Bayesian learning and inference.

Short summary for inclusion in the Courses of Study booklet: the course is intended to offer an introduction to basic techniques used to design machine applications using theoretical analyses and hands-on exercises.

Textbook: There will be no textbook for this course. However, 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 may be referred to in part/full. A few excellent reference texts are listed below.

- Christopher Bishop, Pattern Recognition and Machine Learning, Springer, 2007.
- Hal Daumé III, A Course in Machine Learning, 2015.
- Marc Peter Deisenroth, A. Aldo Faisal, and Cheng Soon Ong. Maths for Machine Learning, Cambridge University Press, 2023.
- Trevor Hastie, Robert Tibshirani, Jerome Friedman, The Elements of Statistical Learning, Springer, 2009.
- John Hopcroft, Ravindran Kannan, Foundations of Data Science, 2014.
- Mehryar Mohri, Afshin Rostamizadeh, Ameet Talwalkar. Foundations of Machine Learning, The MIT Press, 2012.
- Kevin P. Murphy. Probabilistic Machine Learning: An Introduction, The MIT Press, 2022.
- Simon J.D. Prince. Understanding Deep Learning, MIT Press, 2023.
- David G. Stork, Peter E. Hart, and Richard O. Duda. Pattern Classification, Wiley-Blackwell, 2000.
- Aston Zhang, Zack C. Lipton, Mu Li, Alex J. Smola. Dive into Deep Learning, Cambridge University Press, 2023.

Course proposer:
Date:

Convener DPGC:
Date:

The course is approved/not approved

Chairperson, SPGC
Date: