Stochastic Processes & Optimization in Machine Learning

Vasilis Maglaris, Spring 2024

Course Outline

Basic definitions of Machine Learning (ML) & Artificial Intelligence (AI). Training, Validation & Testing Datasets. Review of Opimization Algorithms in ML: Supervised, Unsupervised, Reinforcement Learning. Discriminative & Generative Models, the Chat Generative Pre-trained Transformer - ChatGPT hype. Linear & Logistic Regression

Neural Networks, Hebb's rule, parameter setting via Supervised Learning, Rosenblatt's Perceptron, Back-Propagation Algorithm

Unsupervised Learning: K-Means Clustering, Principal Components Analysis (PCA), Self-Organizing Maps (SOM), Autoencoders

Machine Learning & Statistical Mechanics Concepts: Markov Chains & state classification. Chapman-Kolmogorov equations, asymptotic behavior, irreducibility, recurrence, ergodicity, invariant probabilities

Markov Chain Monte Carlo (MCMC) methods, Metropolis-Hastings algorithm, Simulated Annealing, Gibbs Sampling. Generative models, Boltzmann Machine, Restricted Boltzmann Machine (RBM), Deep Belief Networks (DBN)

Reinforcement Learning and Dynamic Programming: Markov Decision Processes, Bellman's Optimality Criterion, optimization algorithms – Value Iteration & Policy Improvement. Approximate dynamic programming methods, Temporal-Difference (TD) & Q-Learning

Reinforcement Learning & Internet Routing: Bellman – Ford algorithm, Border Gateway Protocols (BGP)

Kernel Classification Methods: Pattern Separability - Cover's Theorem. Applications in Radial-Basis Function (RBF) Networks, Hybrid Learning and Support Vector Machines (SVM)

Non-parametric Classifiers, based on K-Nearest Neighbors (KNN) labelled classification

Statistical evaluation of binary classification, Confusion Matrix, Receiver Operating Characteristics (ROC) & Area Under the Curve (AUC). Parametric Probabilistic Classification, Bayes rule, approximate methods, Naive Bayes Classifiers

Decision Trees: Binary Splitting, Classification and Regression Trees (CART), Gini Index, Random Forests, Bagging (Bootstrap & aggregating) algorithms

Recurrent Neural Nets (RNN) & Neurodynamic Models: Associative Memory - Content Addressable Memory (CAM), Hopfield Networks. Sequence modeling RNNs based on time/character series & speech processing datasets, Long-Short Term Memory (LSTM) nets

Explainability in ML/AI - XAI (eXplainable AI): Definitions, Intrinsic & Model Agnostic XAI methods, PI (Permutation Feature Importance), SHAP (Shapley Additive exPlanations), LIME (Local Interpretable Model Agnostic Explanation)

Laboratory Exercises

Students are required to practice via Python-based exercises

Instructor

Vasilis Maglaris (Professor Emeritus of the School of Electrical & Computer Engineering, NTUA)

Support

Mary Grammatikou (Teaching Fellow of the School of Electrical & Computer Engineering, NTUA)

Laboratory Assistants

Nikos Kostopoulos, *Dimitris Pantazatos* (Research Associates at the Network Management & Optimal Design Lab - NETMODE, School of Electrical & Computer Engineering, NTUA)

Suggested Bibliography

- Simon Haykin, "**Neural Networks and Learning Machines**", 3rd Edition, Pearson Education, 2009

- Bernhard Mehlig, "**Machine learning with neural networks**", Cambridge Univ. Press, 2021 https://arxiv.org/pdf/1901.05639.pdf

- Kevin P. Murphy, "**Machine Learning: A Probabilistic Perspective**", MIT Press, 2012

- Ian Goodfellow, Yoshua Bengio and Aaron Courville, "**Deep Learning**", MIT Press, 2016 https://www.deeplearningbook.org/

- Daniel Jurafsky and James H. Martin, "**Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition**", 3rd Edition draft, 2018

- Andrew Ng, "**CS229 Lecture Notes**", Stanford University, Fall 2018 https://see.stanford.edu/materials/aimlcs229/cs229-notes1.pdf

- James Gareth, Daniela Witten, Trevor Hastie and Robert Tibshirani, "**An Introduction to Statistical Learning with Applications in R**", 2nd Edition, Springer, 2021 https://hastie.su.domains/ISLR2/ISLRv2_website.pdf

- Richard Sutton and Andrew Barto, "**Reinforcement Learning: An Introduction**", 2nd Edition, MIT Press, 2018

- Dimitri P. Bertsekas and John Tsitsiklis, "**Neuro-Dynamic Programming**", Athena Scientific, Belmont MA, 1996

- Christopher Bishop, "**Pattern Recognition and Machine Learning**", Springer, 2006

- Tom Mitchell, "**Machine Learning**", McGraw Hill, 1997 http://www.cs.cmu.edu/~tom/mlbook.html

- Charu C. Aggarwal, "**Outlier Analysis**", Springer 2013 https://link.springer.com/chapter/10.1007/978-1-4614-6396-2_1

- Christoph Molnar, "**Interpretable Machine Learning**", 2nd Edition, Munich, 2022 https://christophm.github.io/interpretable-ml-book/

- Leonida Gianfagna and Antonio Di Cecco, "**Explainable AI with Python**", Springer, 2021 **https://link.springer.com/book/10.1007/978-3-030-68640-6**

- Robert Hogg, Joseph McKean and Allen Craig, "**Introduction to Mathematical Statistics**", 8th Edition, Pearson Education, 2020

- Frank Kelly, "**Reversibility and Stochastic Networks**", Wiley, 1979 http://www.statslab.cam.ac.uk/~frank/BOOKS/book/whole.pdf

- Sheldon Ross, "**Applied Probability Models with Optimization Applications**", Dover, 1992