UNIVERSITY OF MARYLAND

MSML610: Advanced Machine Learning

Class mechanics

Instructor: GP Saggese, PhD - gsaggese@umd.edu

MSML610

- **MSML610**
- Class map

Books of the class

- The goal is to make the slides self-sufficient
- For each lesson, I recommend a few books I liked about that topic
  - Simple -> Medium -> Hardcore
- Simple
  - Burkov: "Machine Learning Engineering" (2020)
  - Burkov: "The Hundred-Page Machine Learning Book" (2019)
- Medium
  - Abu-Mostafa et al.: "Learning From Data" (2012)
  - Martin: "Bayesian Analysis with Python" (2nd ed, 2021)
  - Russell et al.: "Artificial Intelligence: A Modern Approach" (4th ed, 2020)
- Hardcore
  - Hastie et al.: "The Elements of Statistical Learning" (2nd ed, 2009)
  - Koller et al.: "Probabilistic Graphical Models: Principles and Techniques" (2009)
  - Murphy: Machine Learning: "A Probabilistic Perspective" (2012)
  - Sutton et al.: "Reinforcement Learning: An Introduction" (2nd ed, 2018)

Invariants of a class lecture

- Focus on intuition over math (unless necessary)
- Emphasize realistic assumptions and numerical methods
  - Analytical solutions are so 1800s
- Lessons include an interactive Jupyter notebook tutorial to foster intuition and hands-on approach
  - Tutorials are mainly done at home
  - Videos of each tutorial will be added over time
- Lessons alternate between:
  - Slides
  - Whiteboard
  - Tutorials
- 2:45 hours per class lessons
  - 50 slides
  - 10 break
  - 50 slides
  - 10 break
  - 45 slides (Topic refresher!)

Grading

- Class Participation (10%)
  - Attendance
  - Contributions to discussions and engagement
- Quizzes (40%)
  - Multi-choice quizzes
  - 4-5 quizzes to make sure everyone studies during the semester and doesn't cram
- Final Project (50%)
  - A comprehensive application of course concepts
  - Python project selected from a list of topics

Class projects

- There is a list of topics you can pick from, e.g.,
  - LLMs
  - Deep learning
  - Big data
  - Statistical learning
  - ...
- Different levels of difficulty
- Each project is individual
- Content of each project `XYZ`
  - Study / describe a technology
  - Implement a use case using the technology
  - Create Jupyter notebooks to demo your project
  - Commit code to GitHub and contribute to open-source repo
  - Write a blog entry

Links

- Syllabus
- FAQs
- Project specs
- Announcements on ELMS

Class map

- MSML610
- **Class map**

1. Intro

- A map of machine learning
- What is Artificial Intelligence
    - AI
    - ML
    - AI vs ML vs Deep-learning
    - The foundation of AI
    - Brief history of AI
    - AI state of the art
    - Risks and benefits of AI

2. Techniques

- Paradigms
- Techniques
  - Machine learning in practice
  - How to do research
    - Simple is better
    - Research methodology
  - Pipeline organization
  - Input processing
  - Learning algorithms
    - Gradient descent
    - Stochastic gradient descent
  - Performance metrics
    - Precision and recall
  - Model selection
  - Aggregation
    - Bagging
    - Boosting
    - Stacking

3. Knowledge Representation

- Knowledge Representation
  - Basics of Knowledge Representation
  - Examples of Logic
  - Logical Agents
  - Ontologies
  - Reasoning in Ontologies
- Propositional logic
- First-order Logic
- Non-classical Logics
- Description Logics
  - Semantic Web
- Advanced topics

4. Machine Learning Models

- Models
    - Naive Bayes
    - Decision trees
    - Random forests
    - Linear models
    - Perceptron
    - Logistic regression
    - LDA, QDA
    - Kernel methods
    - Support vector machines
    - Similarity-based models
    - Clustering
    - Anomaly detection

5. Machine Learning Theories

- Is learning possible?
  - Training vs Testing
  - Growth function
  - The VC dimension
  - Regularization
    - Bias vs variance
    - Learning curves
    - Learn-validation approach

6. Bayesian statistics

- Quantifying uncertainty
- Probabilistic reasoning

7. Probabilistic programming

- Probabilistic programming
  - Probability theory
  - Single-parameter inference
  - How to choose priors
  - Communicating a Bayesian analysis
  - Probabilistic programming
  - Posterior-based decisions
  - Gaussians all the way down
  - Posterior predictive checks
  - Robust inference
  - Groups comparison
  - Hierarchical models
  - Simple linear model
  - Variable variance
  - Hierarchical linear regression
  - Multiple linear regression
  - Comparing models
    - Posterior predictive checks
  - The balance between simplicity and accuracy
  - Measures of predictive accuracy
    - Information criteria
    - Cross-validation

8. Reasoning over time

- Reasoning over time
- HMMs
- Markov random fields
- Markov logic network
- State space models and Kalman filter
  - g-h filter
  - Discrete Bayes filter
- Dynamic Bayesian networks
- State space model
- Variational Inference
  - Expectation-Maximization (EM) Algorithm

9. Causal inference

- Causal AI
  - Why Causal AI?
  - Concepts in Causal AI
  - Variables
  - Paths
  - The Ladder of Causation
  - Correlation vs causation models
- Business processes around data modeling
  - Modeling processes
  - Roles

10. Timeseries forecasting

- Time Series
  - Basic definition
  - Time series operators
  - Time series decomposition
- Classical Methods
  - Simple models for stochastic process
  - Autoregressive models
  - Moving average models
  - ARMA(p, q) process
  - ARIMA model
  - ARCH model

11. Probabilistic deep learning

- Neural networks
  - Biological inspiration
  - Neural networks
- Advanced Neural Network Architectures
  - Convolutional networks
  - Recurrent Neural Networks (RNNs)
  - Deep learning learning algorithms
  - Deep learning architectures
- Fundamentals of Deep Learning
- Training Deep Neural Networks
- Interpretability and Explainability
- Deep Generative Models
- Bayesian Deep Learning
- Deep Probabilistic Models
- Uncertainty Quantification
- Probabilistic Programming and Inference
- Modern Research Frontiers
- Bonus Topics

12. Reinforcement learning

- Sequential decision problems
  - Utilities over time
  - Algorithms for MDPs
- Reinforcement learning
  - Passive reinforcement learning
  - Active reinforcement learning
  - Generalization in reinforcement learning
  - Policy search
- Fundamentals
- Classical Methods
- Exploration Strategies
- Policy Gradient Methods
- Value Function Approximation
- Deep Reinforcement Learning
- Model-Based Reinforcement Learning
- Advanced Topics
- Applications

Refresher: Probability

- Probability
    - Probability definition
    - Probability measure
    - Independent events
    - Conditional probability
    - Law of total probability
    - Bayes theorem
- Random variables
    - Random variables
    - CDF, PMF, PDF of Random Variables
    - Joint distributions
    - Marginal distributions
    - Independent RVs
    - Conditional PDF RVs
- Mathematical expectation of RVs
    - Mean
    - Variance and covariance
    - Statistics of RVs
- Probability inequalities
- Statistical Inference
    - Definitions

Refresher probability distributions

- Interesting RVs
  - Bernoulli
  - Binomial
  - Gaussian
  - Log-Normal
  - Poisson
  - Chi-square
  - Student's t-distribution
- Probability inequalities

Refresher linear algebra

- Linear algebra
  - Vector and vector spaces
  - Affine spaces
  - Vectors and matrices
  - Linear functions
  - Connections between Machine Learning and Linear Algebra

Refresher information theory

- Information theory
  - Entropy
    - Kullback-Leibler divergence
  - Connections between Information Theory and ML

Refresher game theory

- Game theory
  - Connections between Machine Learning and Game Theory

Refresher: numerical optimization

- Optimization / numerical methods

Refresher: stochastic processes

- Stochastic processes