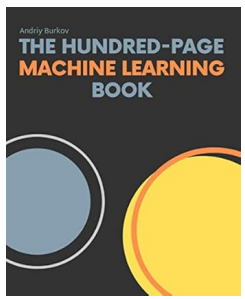
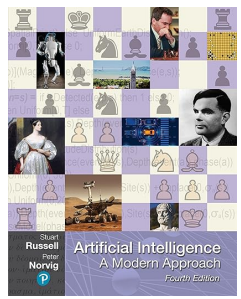
Lesson 02.1: A Map of Machine Learning

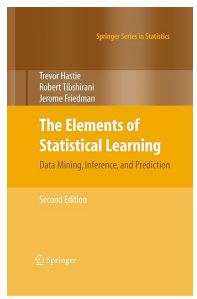
Instructor: Dr. GP Saggese - gsaggese@umd.edu

References:

- Burkov: *"The Hundred-Page Machine Learning Book"* (2019)
- Russell et al.: *"Artificial Intelligence: A Modern Approach"* (4th ed, 2020)
- Hastie et al.: *"The Elements of Statistical Learning"* (2nd ed, 2009)





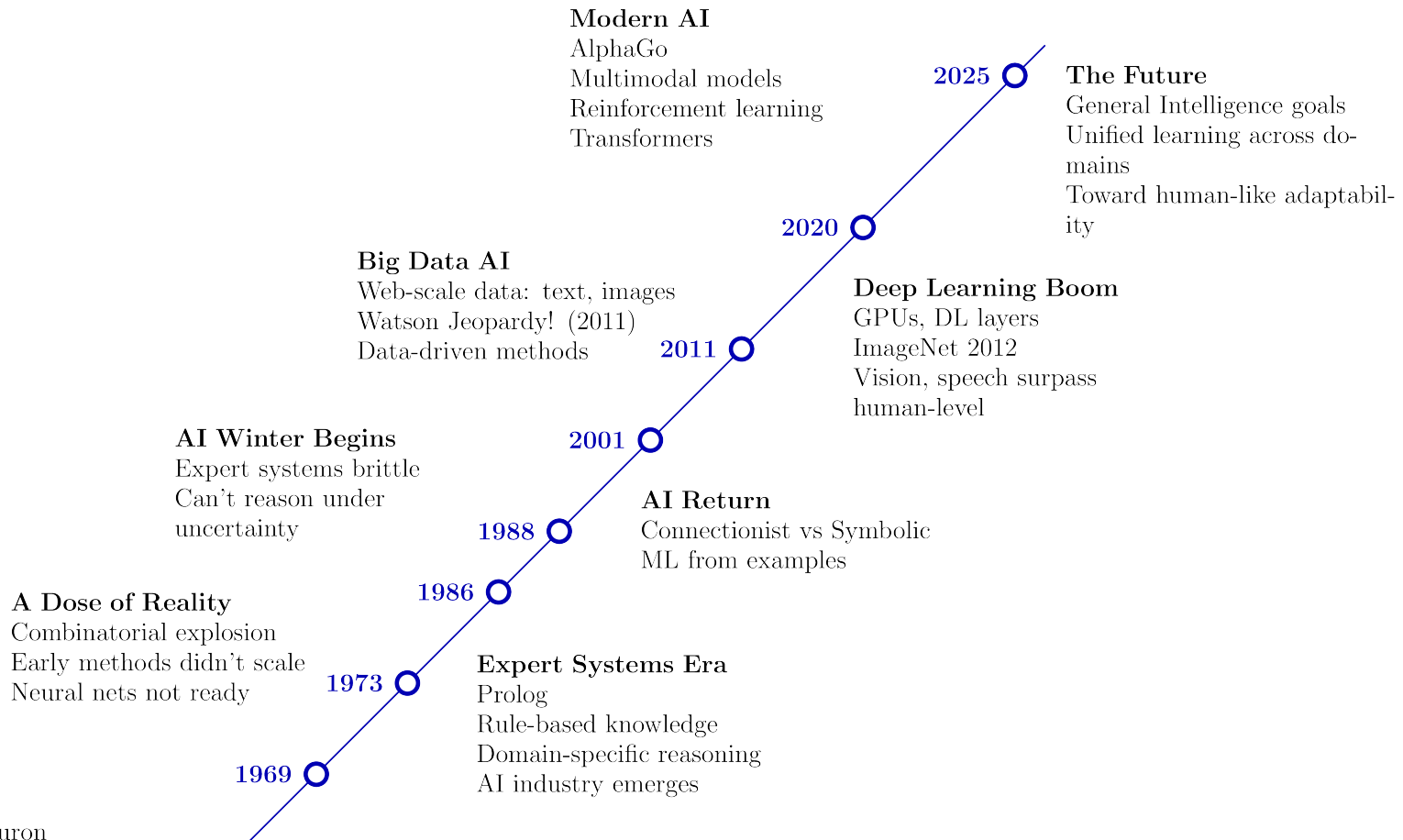


SCIENCE ACADEMY

- ***A Map of Machine Learning***

A Map of Machine Learning

- Machine Learning is a field with many branches
  - Paradigms
  - Theory
  - Models
  - Techniques
  - ...



Machine Learning Paradigms

- How do you set up the learning problem?
  - **Supervised learning**
    - The dataset includes inputs with corresponding outputs
    - Develop an input-output relationship
  - **Unsupervised learning**
    - The data is unlabeled, discover structure within the data
    - E.g., anomaly detection, clustering
  - **Reinforcement learning**
    - The correct answer is not immediately available
    - Evaluate actions based on final outcomes
  - **Active learning**
    - Not all examples are available initially
    - Request outputs for specific inputs
  - ...

Machine Learning Theory

- **VC theory**
  - Measure model capacity and generalize based on hypothesis space complexity
- **Bias-variance decomposition**
  - Prediction error is the sum of:
    - Bias: Error from simplistic model assumptions
    - Variance: Error due to sensitivity to training data fluctuations
- **Computation complexity**
  - Related to information theory and compression
  - E.g., Minimum Description Length (MDL) measures model complexity via efficient model and data description
- **Bayesian approach**
  - Treat ML as probability
  - Combine prior knowledge with observed data to update belief about a model
- **Problem in ML theory**
  - Assumptions may not align with practical problems

Machine Learning Models

- What is the **form of the model** and **how to fit / predict** from the data?
  - Linear models
  - Generalized linear models
    - E.g., logistic, Poisson regression
  - Support Vector Machines (SVM)
  - Nearest neighbors
    - E.g., k-means clustering, KNN
  - Gaussian processes
  - Graphical models
    - Model joint distributions with graphs
    - E.g., hidden Markov models (HMM), Kalman filters, Bayesian networks
  - Neural networks
  - ...

Machine Learning Techniques

- What are the stages of a ML pipeline?
  - **Input processing**
    - Data cleaning
    - Dimensionality reduction
    - Feature engineering
  - **Model building**
    - Models
    - Learning algorithms
  - **Performance evaluation**
    - Cross-validation
    - Bias-variance curves
    - Learning curves
  - **Regularization**
  - **Aggregation**
    - Boosting
    - Bagging
    - Stacking

Machine Learning Adages

- *"An explanation of the data should be as simple as possible, but not simpler"* (Einstein)
- *"The simplest model that fits the data is also the most plausible"* (Occam's razor)
- *"Garbage in, garbage out"* (Fuechse, 1957)
- *"All models are wrong, but some are useful"* (Box, 1976)
- *"If you torture the data long enough it will confess whatever you want"* (Coase, 1982)
- *"Data is the new oil"* (Humby, 2006)
- *"More data beats clever algorithms"* (Norvig, ~2006)
- *"The unreasonable effectiveness of data"* (Halevy, Norvig, Pereira, 2009)
- *"The bitter lesson"* (Sutton, 2019)