Foundations of Machine Learning
AI2000 and AI5000

FoML-38
Conclusion

Dr. Konda Reddy Mopuri
Department of AI, IIT Hyderabad
July-Nov 2025

భారతీయ సాంకేతిక విజ్ఞాన సంస్థ హైదరాబాద్
भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Data-driven Intelligence
& Learning Lab

So far in FoML

- Intro to ML and Probability refresher
- MLE, MAP, and fully Bayesian treatment
- Supervised learning
    a. Linear Regression with basis functions
    b. Bias-Variance Decomposition
    c. Decision Theory - three broad classification strategies
    d. Neural Networks
- Unsupervised learning
    a. K-Means, Hierarchical, and GMM for clustering
- Kernelizing linear Models
    a. Dual representation, Kernel trick, SVM (max-margin classifier)
- Tree-based Methods
- Model combination - Bagging, Boosting

For today

- Course recap and conclusion

Recap

First principles to Practical systems

- Starting with mathematics
  - (probability, linear algebra, and optimization)
- Progress through core paradigms of learning

Probabilistic Core

- Began with ground our models in probability theory
  - (foml-02 and 03)
- Understand and implement fundamental estimation techniques
  - MLP (foml-04)
  - MAP (foml-05)
  - Fully-Bayesian Treatment (foml-06 and 11)

Supervised learning

1. Regression (foml-06 to 11)
   a. Minimizing the squared error (foml-06)
   b. Controlling the complexity via regularization (Ridge/Lasso in foml-09)
   c. Bias-Variance decomposition (foml-10)

Supervised learning

1. Regression
2. Classification (foml-12 to 21)
    a. We analysed the models through three lenses
    b. Generative models (foml-13 and 14)
    c. Discriminant functions (foml-15 to 17)
    d. Probabilistic discriminative models (foml-19 to 21)

Learning Basis functions

- Neural networks (foml-22 to 24)
  - A sophisticated extension of linear models
  - Data-driven features
  - Gradient descent → backpropagation

Unsupervised Learning

- Explored structure discovery (foml-25 to 28)
  - K-Means, GMM, Hierarchical clustering
- Dimensionality reduction (foml-29 and 30)
  - PCA

Advanced conventional methods

- Kernel methods (foml- 31 & 32)
  - Map data into high-dim feature spaces → max-margin classifier (foml-33 to 35)
- Ensemble methods (foml - 36 & 37)
  - Bagging, Boosting (e.g., Random Forests)

Conventional ML: when classics beat the hype

Deep Learning

- Dominates the fields such as CV, NLP
- However, understanding the foml is essential
    - Because they remain the choice in some of the practical scenarios

Low-data Regime

- DL models are data-hungry
  - May need millions of training data samples
- Tree-based methods handle high-variance situations better
  - Train efficiently

Tabular/structured data

- In business, finance, healthcare, data is typically structured (rows & columns)
  - DL often fails to outperform boosting algorithms
  - Particularly when the data's underlying structure is not sequential or hierarchical

High need for Interpretability

- Fields such as clinical trials, finance, insurance, etc., black-box models are unacceptable
- Classical models provide intrinsic transparency
  - Linear (predictor to o/p relation), shallow DTs (rules for prediction)

Feature engineering is Key

- When experts can create powerful 'features' (based on the domain knowledge)
  - → simple linear models or SVM can be more efficient and robust than DL models that learn from scratch

Resources and latency constraints

- Classical models have
  - Lesser memory footprint
  - Faster inference time
- Ideal for real-time systems
  - Edge devices without much of resources
- DL models need specialised, expensive hardware

Thank you!

భారతీయ సాంకేతిక విజ్ఞాన సంస్థ హైదరాబాద్
भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

DiL
Data-driven Intelligence
& Learning Lab