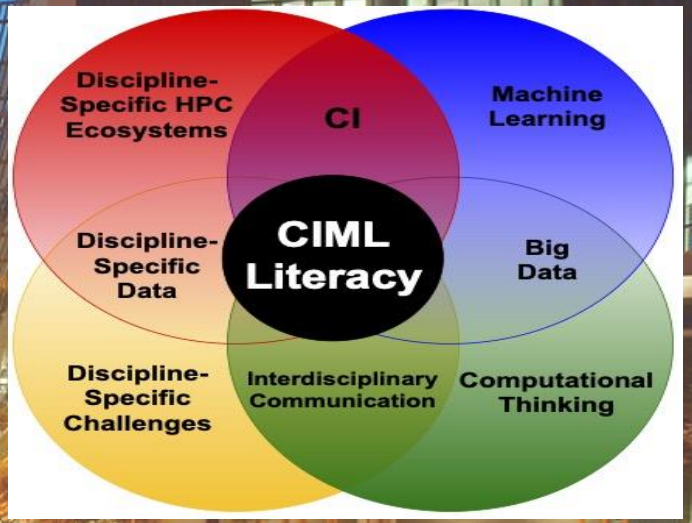
Machine Learning/Deep Learning Schedule Overview

July 2022

Paul Rodriguez, PhD
(SDSC)



What is Machine Learning?

We often say something like:

Programs that learn from data (as opposed to being given rules)

Statistical Learning (as opposed to Statistical Inferencing)

Or, we often just use certain terms:

Supervised/Unsupervised; Classification/Regression;
Overfitting/Regularization; Data Prep; Hyperparameters; etc..

Or, we often talk about algorithm functions like:

fit(); predict(); evaluate(); etc…

Other frameworks with parameters, some learning, and predictions

- Probabilistic Reasoning
- Reinforcement Learning
- Simulations of Physical Systems
- Agent modeling
- Network Analysis
- Time Series Regression Analysis
- Databases and Business Analytics

HPC with ML and Deep Networks are sometimes part and parcel of these

The HPC & Machine Learning landscape

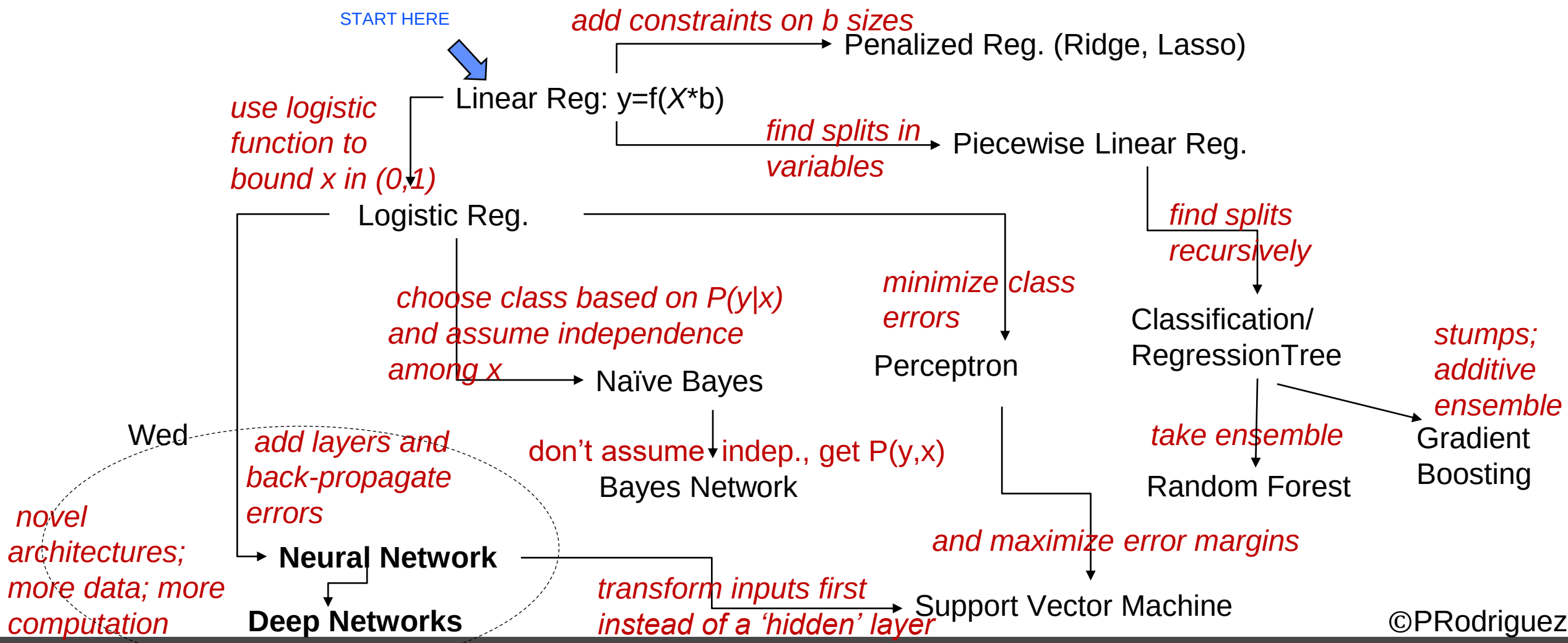
Why use HPC?

Big data and/or Big computation

Possibilities for parallelization; eg distributing data and/or computation

For ML models parallelization often depends on the algorithm to fit parameters

ML Model Space Map – in a nutshell



©PRodriguez

Schedule overview

- **Today - Scaling**
  - R on HPC with Demo, some references to other packages
  - Spark with Tutorials

- **Wednesday – Deep Learning**
  - Intro to NN/CNN/Deep Learning
  - Intro DL, MNIST and Parameter Tuning Exercises
  - DL Layers and Models
  - DL Transfer Learning with Exercises
  - DL Functional API, Special Connections with Exercise (skip connections)