



Machine Learning for Economists

Jesús Fernández-Villaverde¹ and Galo Nuño²

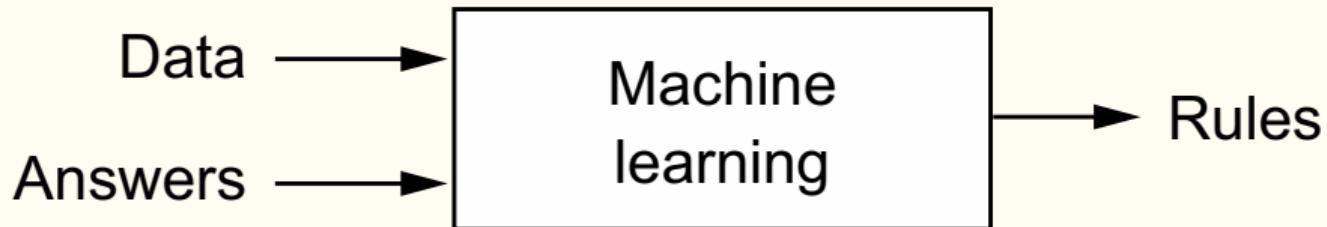
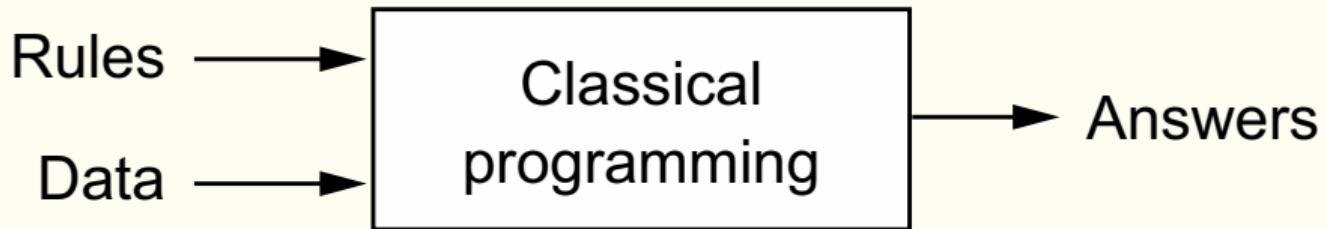
January 5, 2025

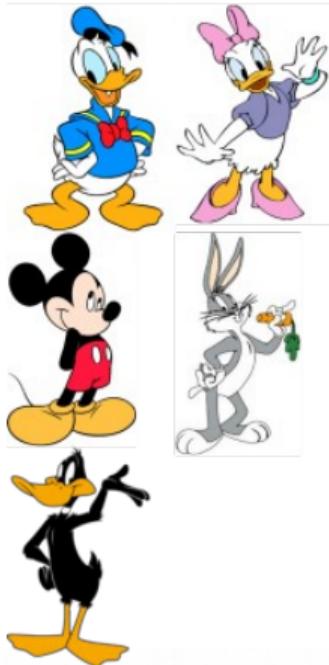
¹University of Pennsylvania

²Banco de España

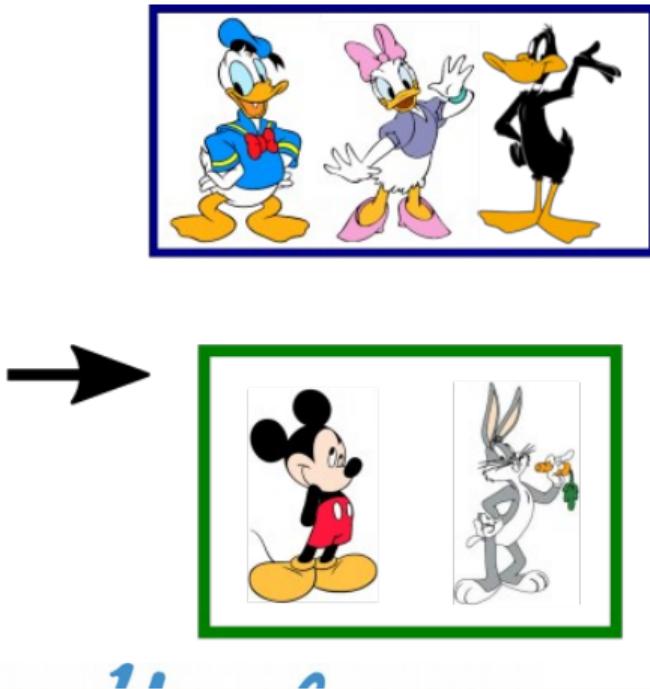
What is machine learning?, I

- Wide set of algorithms to detect and learn from patterns in the data and use them for decision making or to forecast future realizations of random variables.
- Focus on recursive processing of information to improve performance over time.
- In fact, this is clearer to see in its name in other languages: [Apprentissage automatique](#) or [aprendizaje automático](#).
- Even in English: [Statistical learning](#).
- More formally: we use rich data to select appropriate functions in a dense functional space.



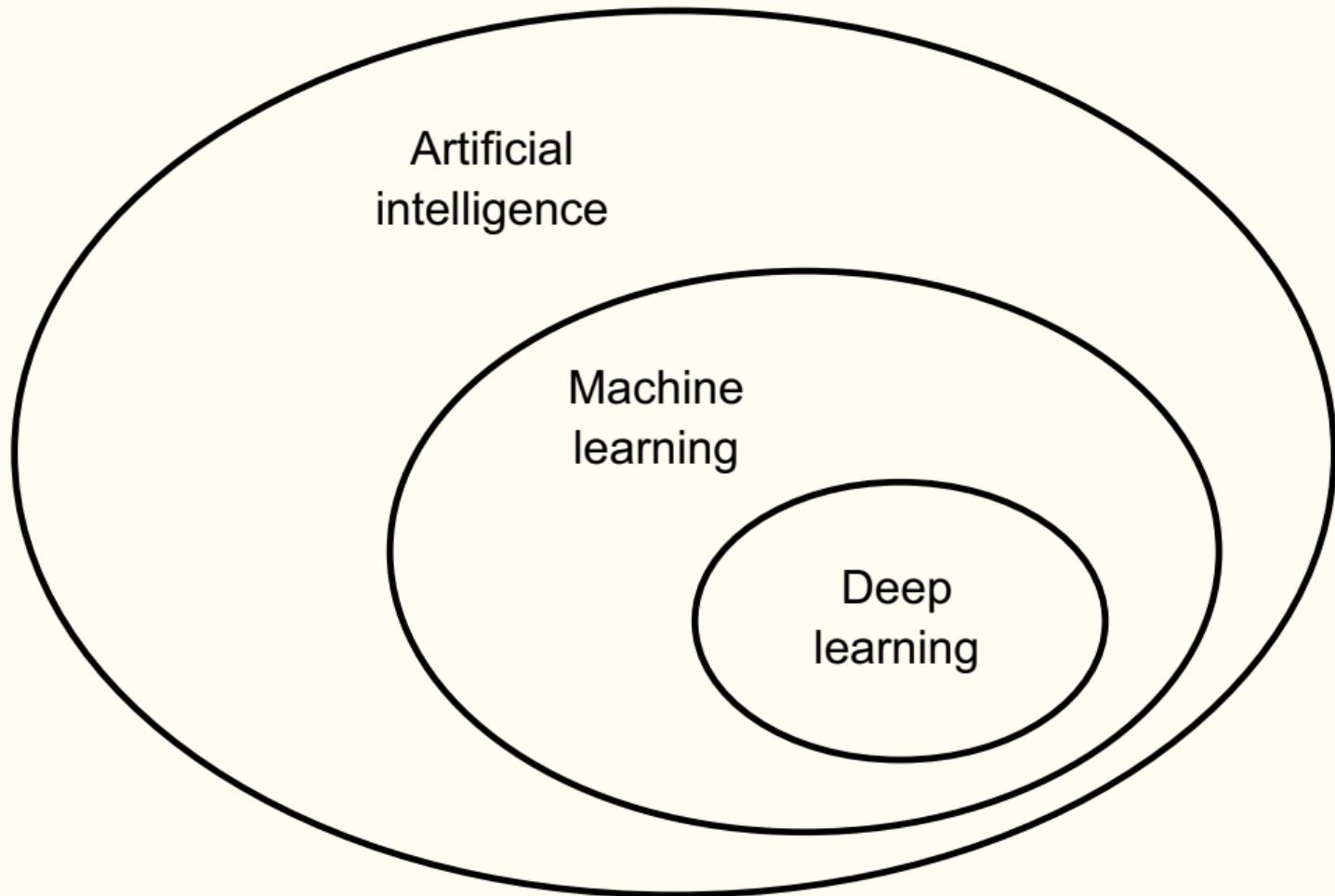


→ **Unsupervised
Learning**

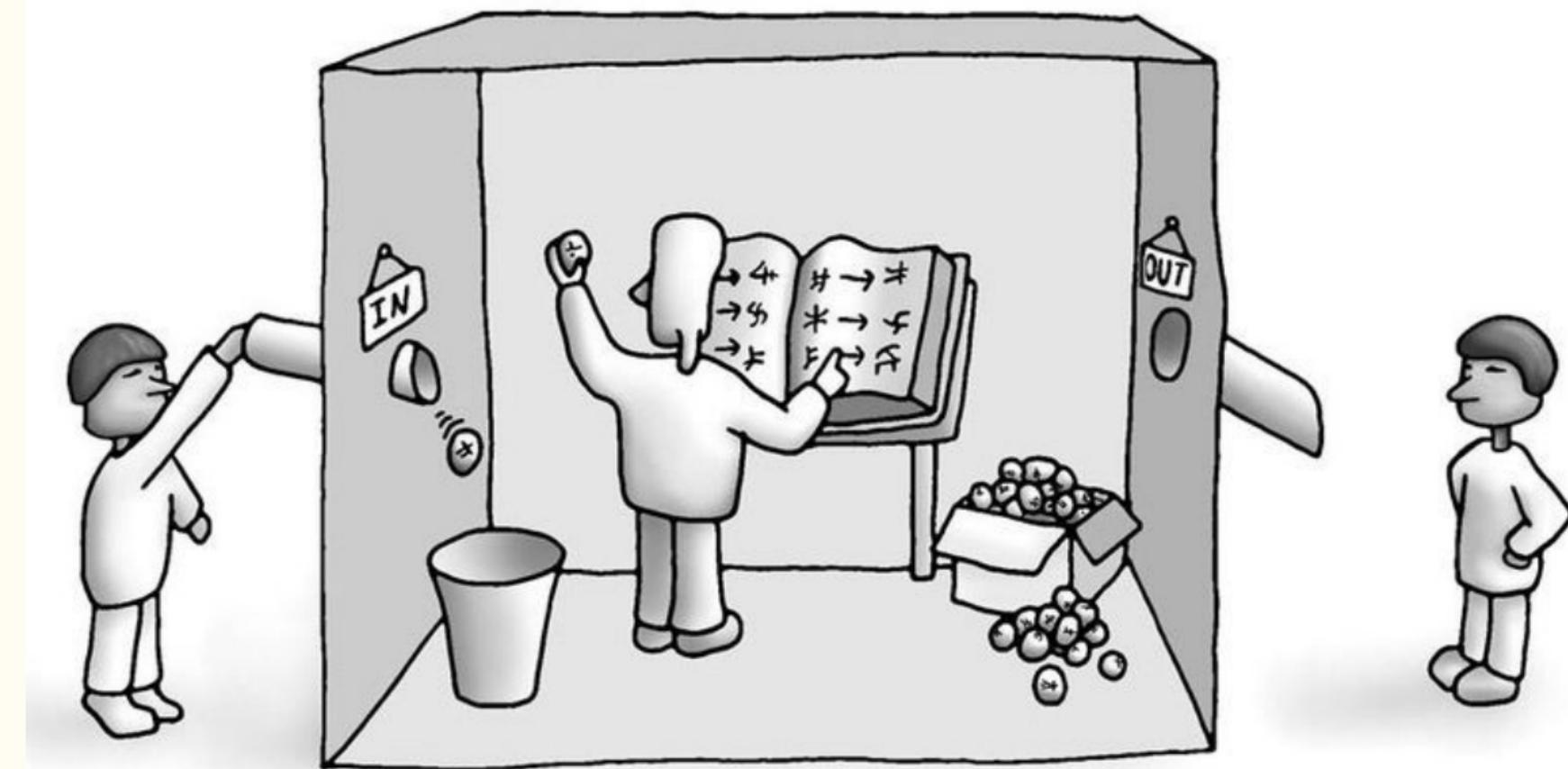


What is machine learning?, II

- Opposition with traditional scientific computation (both standard numerical analysis and Monte Carlos).
- Opposition with symbolic reasoning, expert systems, and cognitive model approaches in artificial intelligence.
- Think about the example of how to program a computer to play chess.
- Operational definition of learning (i.e., Turing test and Chinese room).



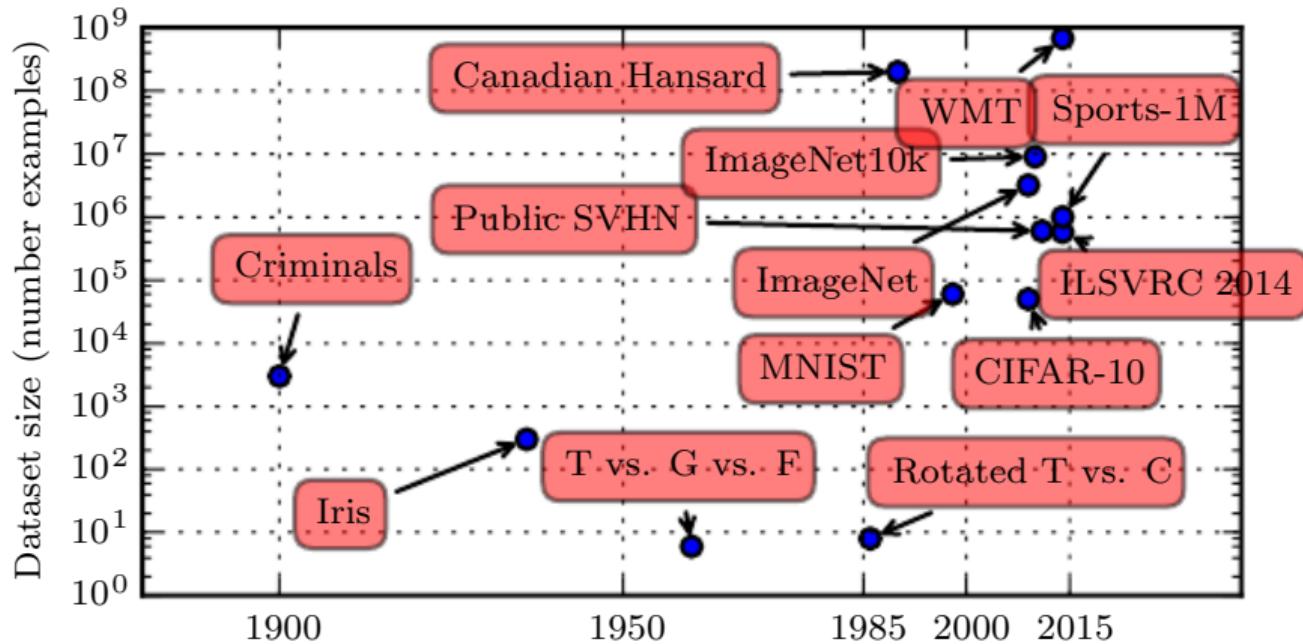




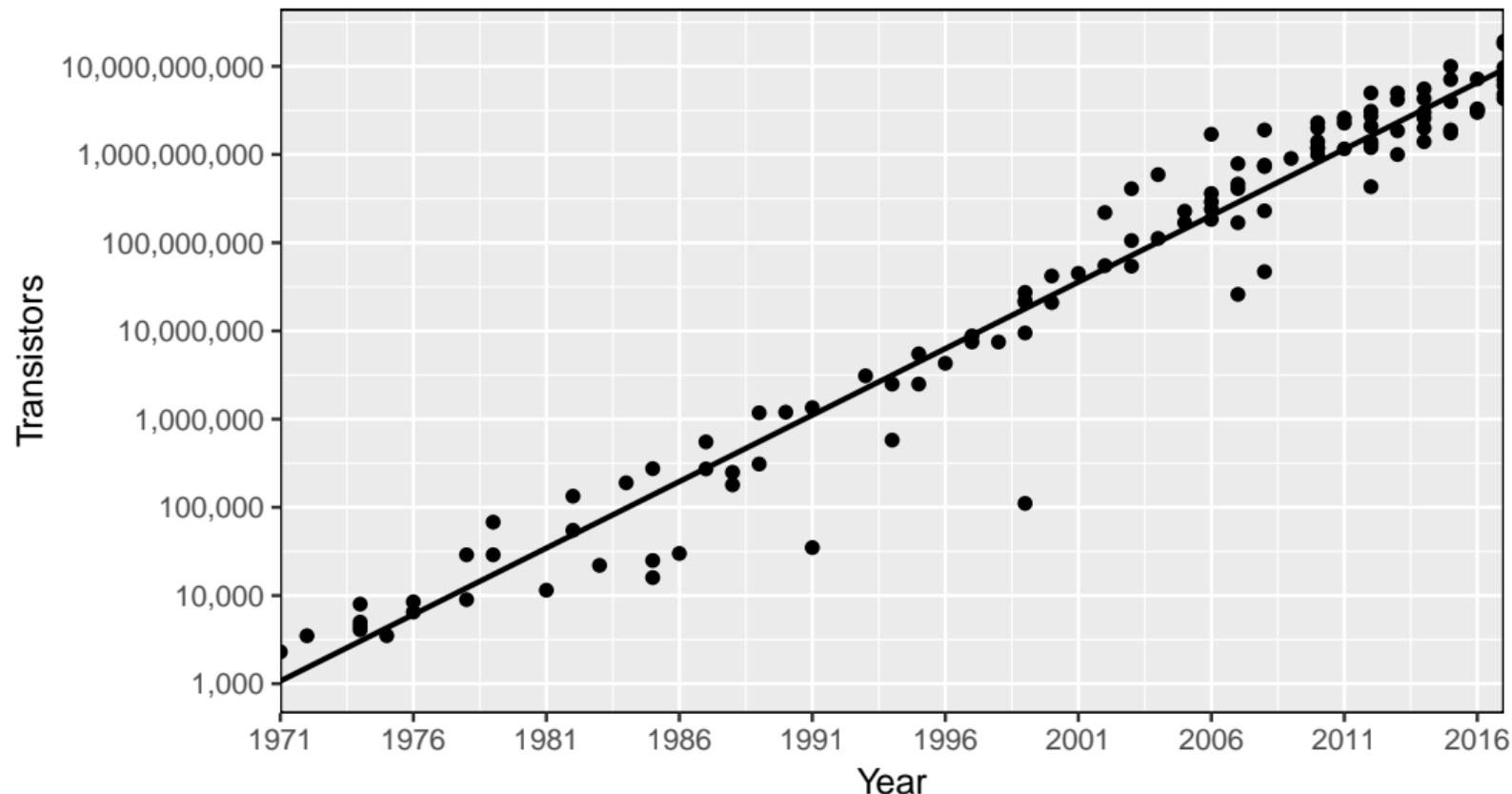
Why now?

- Many of the ideas of machine learning (e.g., basic neural network by McCulloch and Pitts, 1943, and perceptron by Rosenblatt, 1958) are decades old.
- Previous waves of excitement: late 1980s and early 1990s. Those decades were followed by a backlash.
- Four forces behind the revival:
 1. Big data.
 2. Long tails.
 3. Cheap computational power.
 4. Algorithmic advances.
- Likely that these four forces will become stronger over time.
- Exponential growth in industry → plenty of packages for Python, R, and other languages.

Data sizes



Number of transistors





Machine Learning Modelling in R : : CHEAT SHEET

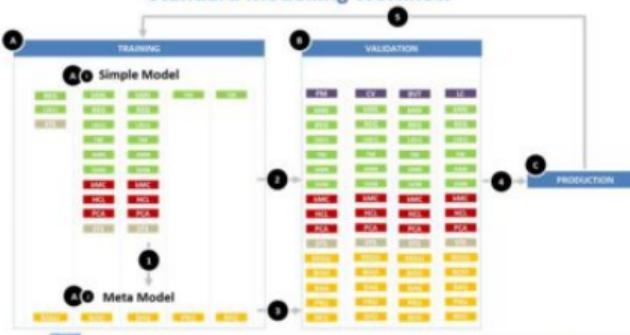
Supervised & Unsupervised Learning

ALGORITHM	DESCRIPTION	R PACKAGE / FUNCTION	SAMPLE CODE
Naïve Bayes classifier	A classification technique based on Bayes' Theorem with an assumption of independence between features. In other terms, a Naïve Bayes classifier assumes that the presence or absence of a particular feature is unrelated to any other feature.	e1071::naivebayes	naivebayes(class ~ ., data = iris)
Linear Regression	An n-parameter linear regression model is fitted to the training data. The n parameters describe the relationship between the dependent variable and the independent variables.	glm::lm	glm(lifeExp ~ 1 + L + I + P, data = world)
Logistic Regression	Model the linear relationship between a scalar dependent variable and one or more explanatory variables (or independent variables). Example:	stats::glm	glm(lifeExp ~ speed, data = world)
Tree-Based Models	The idea is to constructively divide (branch) the training data into smaller and smaller subgroups with respect to the target variable (iris ~ setosa, versicolor, virginica).	randomForest::randomForest	randomForest(lifeExp ~ ., data = world, ntree = 500)
Artificial Neural Network	Neural networks are built from unlabelled perceptions. Perceptions have one or more inputs, an activation function and one or more outputs. They are the building blocks in structured learners.	caret::train	train(lifeExp ~ age + Number + Star_L, data = world)
Support Vector Machine	A data classification method that separates data using hyperplanes.	e1071::svm	svm(lifeExp ~ ., kernel = "linear", C = 100)
Principal Component Analysis	An algorithm that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components.	stats::prcomp	prcomp(lifeExp ~ ., data = world, center = TRUE)
K-Means Clustering	Alters at partitioning clustering algorithm which each observation belongs to the cluster with the nearest mean	stats::kmeans	kmeans(lifeExp ~ ., data = world, centers = 3)
Hierarchical Clustering	An approach which builds a hierarchy from the bottom up and doesn't require the number of clusters to be specified beforehand.	stats::hclust	hclust(lifeExp ~ ., method = "complete", members = NA)

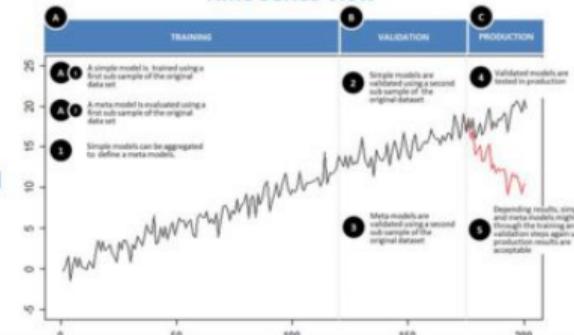
Meta-Algorithm, Time Series & Model Validation

ALGORITHM	DESCRIPTION	R PACKAGE / FUNCTION	SAMPLE CODE
Bagging	Regularization with a penalty on the different parameters of a model to reduce the freedom of the model. Hence, the model becomes more general and less prone to overfitting and improves the generalization abilities of the model.	glmnet::glmnet	glmnet(lifeExp ~ ., family = "gaussian", alpha = 0.5)
Pruning	A process of tree pruning, e.g., the reweighting of the training data set, to remove branches of the tree that has primarily been applied to the latter, in order to improve predictive ability.	tree::prune	prune(tree, cp = 0.05)
Random Forest	Bagging is a technique to increase the power of a predictive model by averaging the predictions of multiple models (bagging replacement) of the training data set, and using each model's prediction to make the final prediction and use the predictions for the original test.	randomForest::randomForest	randomForest(lifeExp ~ ., data = world, ntree = 500)
Leave-one-out	Random sampling of observations for training and testing is done on one observation at a time. This is a very slow process and other tests, that split the data into training and testing sets, are preferred.	boot::boot	boot(lifeExp ~ ., data = world, R = 1000)
Performance metrics	Metrics to evaluate the performance of a model. These metrics are often used in competitions.	mlbench::accuracy	accuracy(lifeExp ~ ., data = world)
Statistical Test	Statistical analysis, hypothesis testing, descriptive statistics, probability distributions, etc.	stats::t.test	t.test(lifeExp ~ ., data = world)
Cross-validation	Cross-validation compares the test performance of different models with different sets or values of parameters.	mlbench::crossValidation	crossValidation(lifeExp ~ ., data = world)
Learning Curves	Learning curves plot a model's training and test scores, or the chosen performance metric, depending on the training set size.	mlbench::learningCurve	learningCurve(lifeExp ~ ., data = world, nfolds = 10, nfold_cv = 10, nrep = 10, ntrain = 100, ntest = 100)

Standard Modelling Workflow



Time Series View



Relation with other fields

- Link with computer science, statistical learning, data science, data mining, predictive analytics, and optimization: frontiers are often fuzzy.
- Many similarities with econometrics and statistical learning, but emphasis is somewhat different:
 1. No unified approach.
 2. Practical algorithms vs. theoretical properties (scalability vs. asymptotic properties).
 3. Traditional statistical inference is de-emphasized.
 4. More interest in forecasting than in causality assertions (cross-validation, regularization).

The many uses of machine learning in macroeconomics

The many uses of machine learning in macroeconomics

- Recent boom in economics:
 1. New solution methods for economic models: my own work on deep learning.
 2. Alternative to older bounded rationality models: reinforcement learning.
 3. Data processing: [Blumenstock et al. \(2017\)](#).
 4. Alternative empirical models: deep IVs by [Hartford et al. \(2017\)](#) and text analysis.
 5. Large language models: [Korinek \(2023\)](#).
- However, important to distinguish signal from noise.
- Machine learning is a catch-all name for a large family of methods.
- Some of them are old-fashioned methods in statistics and econometrics presented under alternative names.

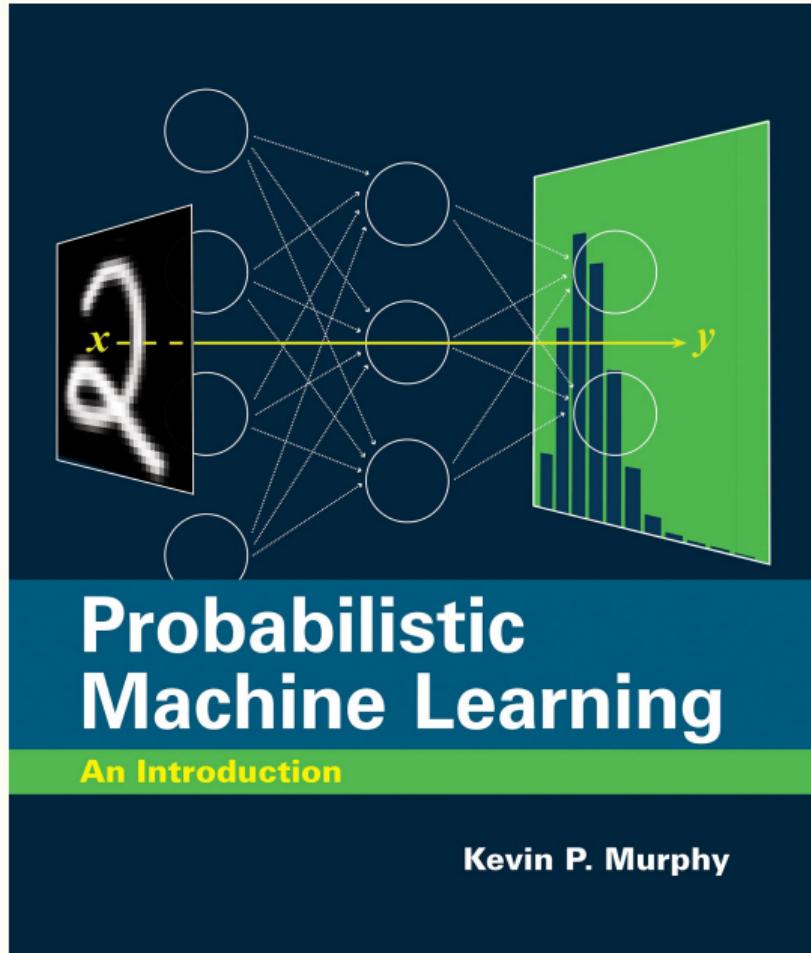
Course outline, I

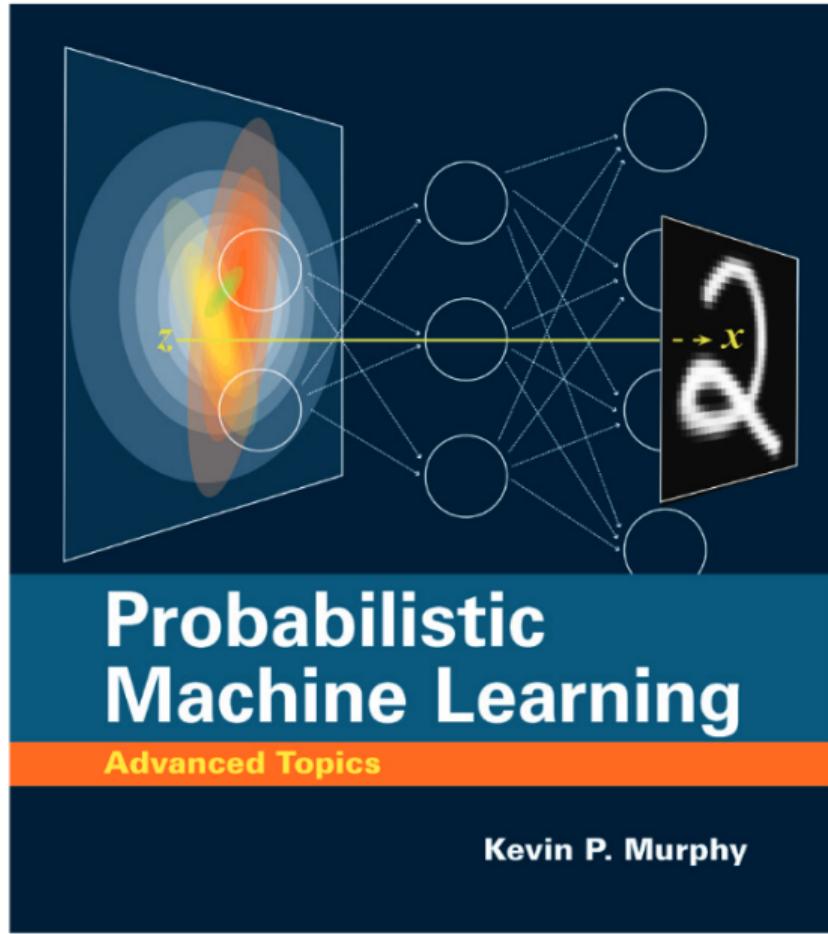
- Block 1: Coding Machine Learning Algorithms.
- Block 2: Challenges Solving Economic Models.
- Block 3: Introduction to Deep Learning.
- Block 4: Optimization in Deep Learning.
- Block 5: Deep Learning for Solving Economic Models.
- Block 6: Advanced Topics in Deep Learning.
- Block 7: Symmetry in Dynamic Programming (if time allows).
- Block 8: Transversality and Stationarity with Deep Learning (if time allows).

Course outline, II

- Block 9: Reinforcement Learning.
- Block 10: Machine Learning for Data Analysis (if time allows).
- Block 11: Text Analysis (if time allows).
- Block 12: Structural Estimation with Unstructured Data (if time allows).

References





LEARN ABOUT AI FROM THE BEST IN THE FIELD.



Machine Learning Street Talk

31.5K subscribers

SUBSCRIBE

HOME

VIDEOS

PLAYLISTS

COMMUNITY

STORE

CHANNELS

ABOUT



Uploads

SORT BY



#75 - Emergence [Special Edition] with Dr. DANIELE...

5.8K views • 3 weeks ago



#74 Dr. ANDREW LAMPINEN - Symbolic behaviour in AI...

4K views • 1 month ago



#73 - YASAMAN RAZEGHI & Prof. SAMEER SINGH - NLP...

2.5K views • 1 month ago



#72 Prof. KEN STANLEY 2.0 - On Art and Subjectivity...

3.6K views • 1 month ago



#71 - ZAK JOST (Graph Neural Networks + Geometr...

4.9K views • 1 month ago



#70 - LETITIA PARCALABESCU - Symbolic...

1:18:31



#69 DR. THOMAS LUX - Interpolation of Sparse High...

50:39



#68 DR. WALID SABA 2.0 - Natural Language...

1:42:15



#62 Dr. GUY EMERSON [Unplugged]

1:29:15



#67 Prof. KARL FRISTON 2.0 [Unplugged]

1:42:11

**Statistics for
Engineering and
Information Science**

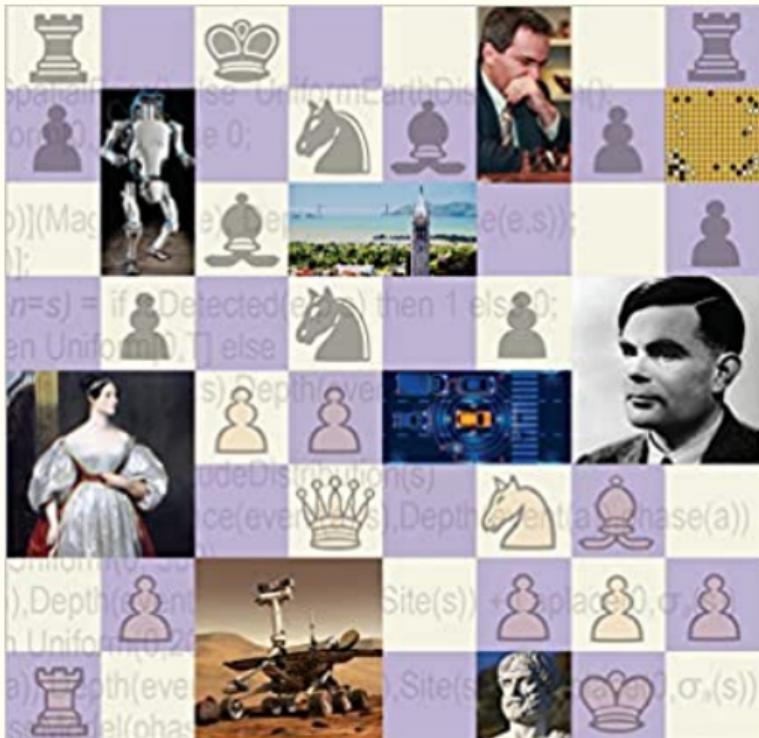
Vladimir N. Vapnik

**The Nature
of Statistical
Learning Theory**

Second Edition



Springer



Artificial Intelligence

A Modern Approach

Fourth Edition



Stuart
Russell
Peter
Norvig