

Machine Learning for Macroeconomics

Jesús Fernández-Villaverde¹ and Galo Nuño²

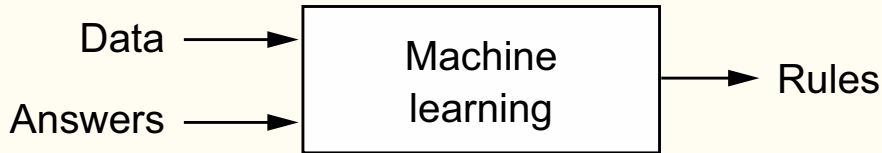
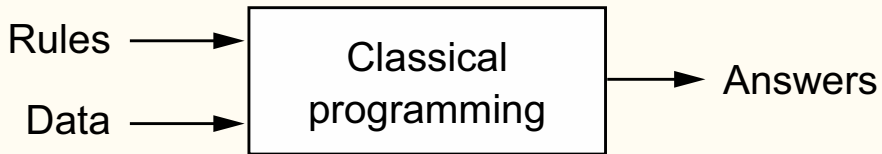
August 7, 2023

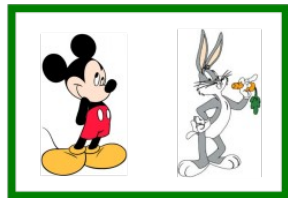
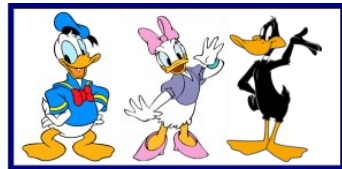
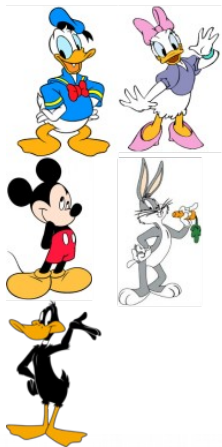
¹University of Pennsylvania

²Banco de España

What is machine learning?, I

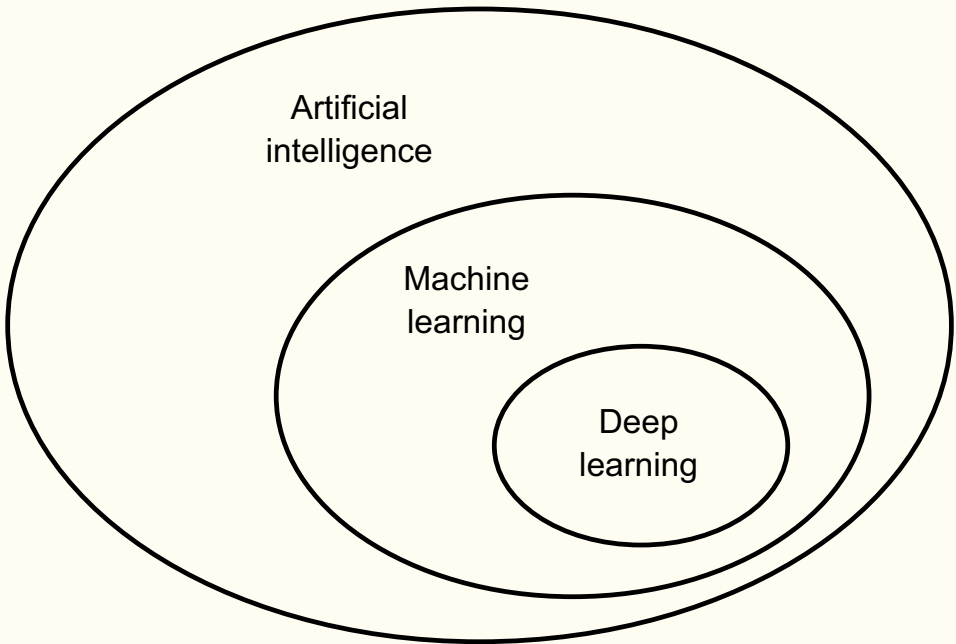
- Wide set of algorithms to detect and learn from patterns in the data and use them for decision making or to forecast future realizations of random variables.
- Focus on recursive processing of information to improve performance over time.
- In fact, this is clearer to see in its name in other languages: [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.





What is machine learning?, II

- Opposition with traditional scientific computation (both standard numerical analysis and Monte Carlo).
- 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).

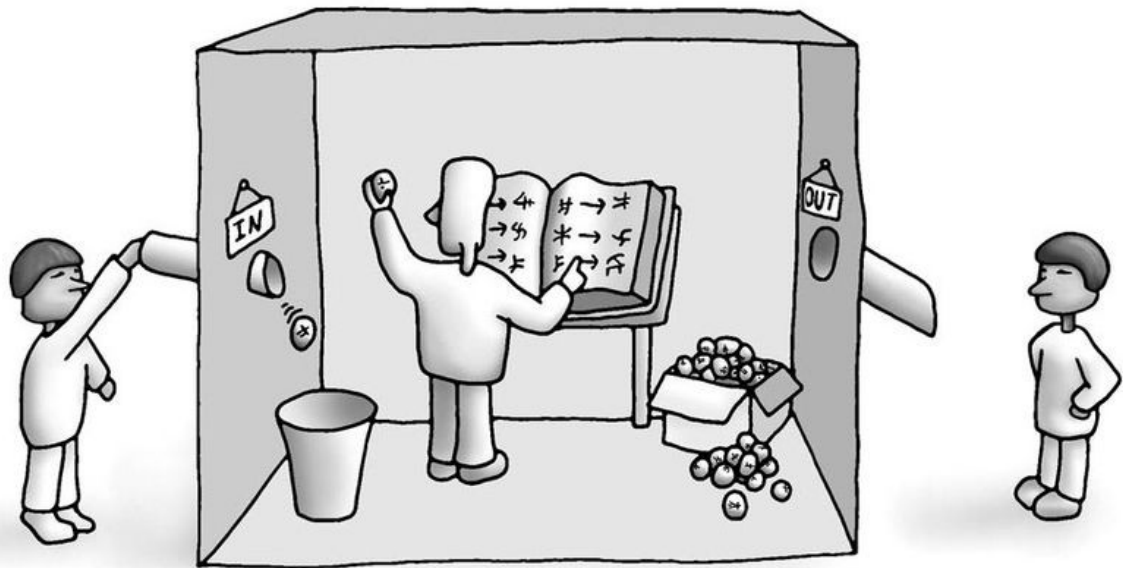


Artificial
intelligence

Machine
learning

Deep
learning

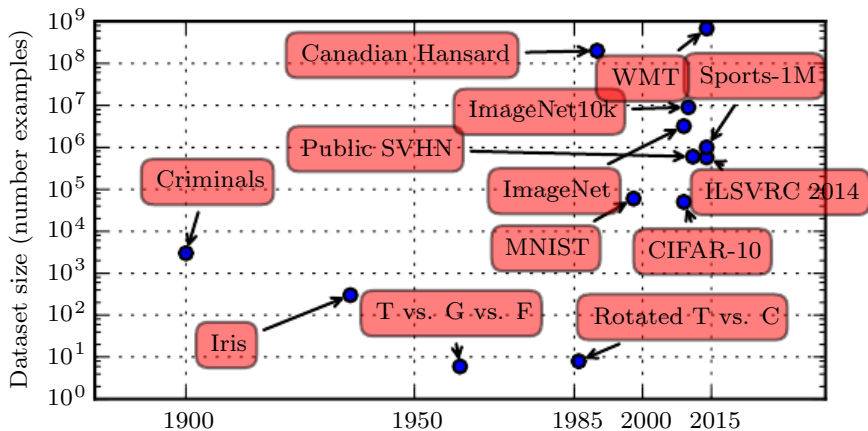




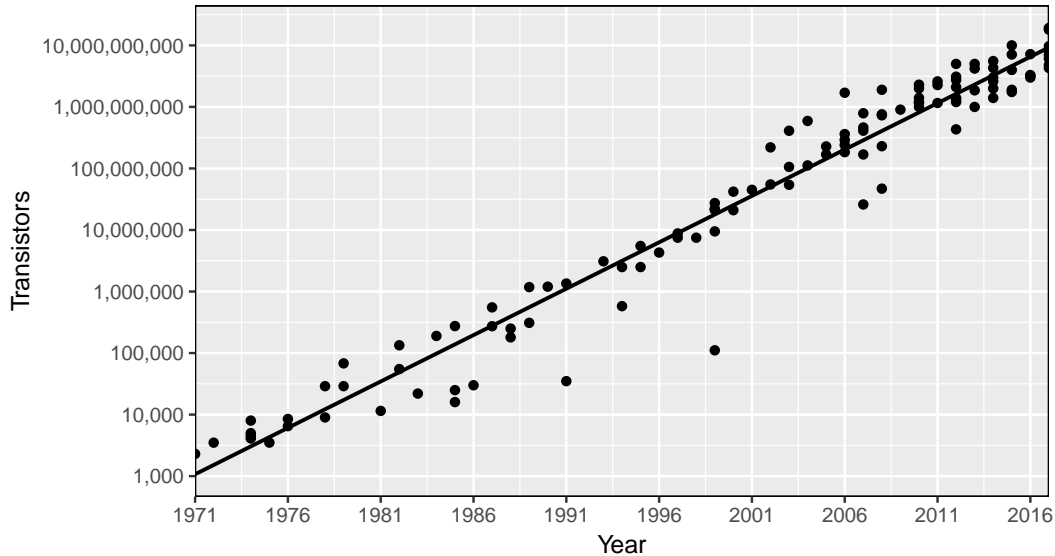
Why now?

- 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 \Rightarrow 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 among predictors. In simple terms, a Naïve Bayes Classifier assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature.	naiveBayes	naiveBayes(x, data = X)
Support Vector Machines	A non-parametric method used for classification and regression. In both cases, the input vectors of the training examples in the feature space. The output depends on whether a SVM is used for classification or regression.	svm	svm(Svm, data = 1:n, y = 0:prob + FALSE, useLib = TRUE)
Linear Regression	Model the linear relationship between a single dependent variable Y and one or more explanatory variables X. Independent variables is denoted as X.	lm	lmfit ~ regnd, dataset)
Logistic Regression	Used to predict a binary outcome (0/1, Yes/No, True/False) given a set of independent variables.	glm	glm(f ~, family = binomial(link = logit), data = X)
Tree Based Models	The idea is to recursively divide (split) the training dataset based on the input features until an optimality criterion is met for the target variable. "Data bucket" needs to be created.	cart	cart(x = data, y = factor(1:2))
Ensemble Neural Networks	Neural networks are built from unified connections. Hierarchical flow of more inputs, an activation function and an output. An ANN model is built up by combining principles of neural networks.	neuralnet	neuralnet::neuralnet
Support Vector Machine	A data classification method that separates data using hyperplanes.	svm	svm(x = data, y = factor(1:2))
Principal Component Analysis	A procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of uncorrelated variables called principal components.	prcomp	prcomp(x = data, y = factor(1:2))
K-Means Clustering	One of several clustering algorithms in which each observation belongs to the cluster with the nearest mean.	kmeans	kmeans(x = data, y = factor(1:2))
Hierarchical Clustering	An approach which builds a hierarchy from the bottom up. It requires the number of clusters to be specified beforehand.	hclust	hclust(x = data, y = factor(1:2))

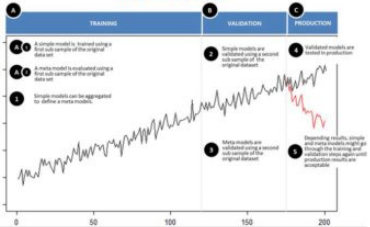
Meta-Algorithm, Time Series & Model Validation

ALGORITHM	DESCRIPTION	R PACKAGE/FUNCTION	SAMPLE CODE
Regularization (L1/L2)	Regularization adds a penalty on the different parameters of the model. The idea is to penalize the model so that it does not fit the noise of the training data and will improve the generalization of the model.	glmnet	glmnet(x = data, y = factor(1:2))
Boosting	Aggregates of weakly performing models. The idea is to combine many weak models to create a strong model. The idea is to combine many weak models to create a strong model.	gbm	gbm(x = data, y = factor(1:2))
Bagging	Bagging is a technique that reduces the variance of a model. The idea is to combine many weak models to create a strong model. The idea is to combine many weak models to create a strong model.	randomForest	randomForest(x = data, y = factor(1:2))
Pruning	Pruning is a technique that reduces the size of decision trees by removing branches of the tree that provide little power to classify instances. Pruning reduces the complexity of the final classifier and hence improves prediction accuracy by reducing overfitting.	prune	prune(x = data, y = factor(1:2))
Random Forest	An ensemble learning method for classification, regression and other tasks. This method uses random sampling with replacement from the training set and building a collection of decision trees at training time and outputting the class that is the mode of the results (classification) or mean prediction (regression).	randomForest	randomForest(x = data, y = factor(1:2))
Grid Search	Random sampling of observations for training and testing a model on a set of parameters. The idea is to find the best model by testing different combinations of parameters.	caret	caret::train(x = data, y = factor(1:2))
Performance Metrics	Measures the performance of a model. The idea is to find the best model by testing different combinations of parameters.	caret	caret::perf(x = data, y = factor(1:2))
Cross Validation	Cross validation compares the test performance of different model realizations with different sets or values of parameters.	caret	caret::train(x = data, y = factor(1:2))
Learning Curves	Learning curves show a model's training and test errors, as the chosen performance metric, depending on the training set size.	caret	caret::learning_curves(x = data, y = factor(1:2))

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

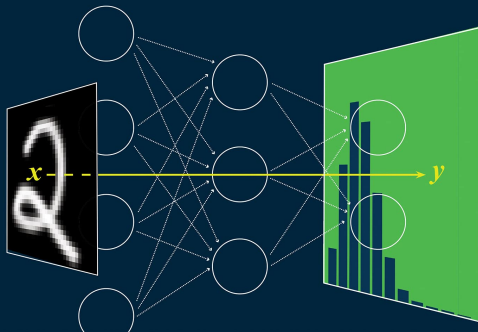
- Recent boom in economics:
 1. New solution methods for economic models: my own work on deep learning.
 2. Alternative to older bounded rationality models: reinforcement learning.
 3. Data processing: [Blumenstock *et al.* \(2017\)](#).
 4. Alternative empirical models: deep IVs by [Hartford *et al.* \(2017\)](#) 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).

- 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).

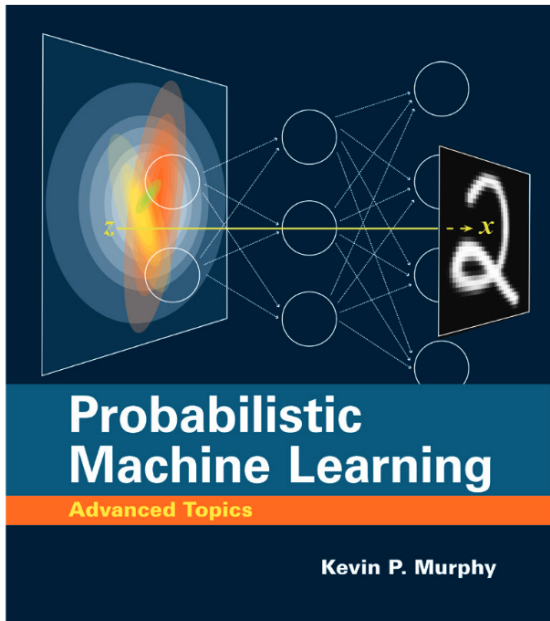
References



Probabilistic Machine Learning

An Introduction

Kevin P. Murphy



LEARN ABOUT AI FROM THE BEST IN THE FIELD.



machine learning street talk



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



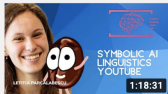
#72 Prof. KEN STANLEY 2.0 - On Art and Subjectivity...

3.6K views • 1 month ago



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

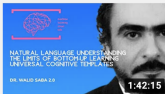
4.9K views • 1 month ago



#70 - LETITIA PARCALABESCU - Symbolic...



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



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



#62 Dr. GUY EMERSON [Unplugged]



#67 Prof. KARL FRISTON 2.0 [Unplugged]

**Statistics for
Engineering and
Information Science**

Vladimir N. Vapnik

**The Nature
of Statistical
Learning Theory**

Second Edition



Springer

