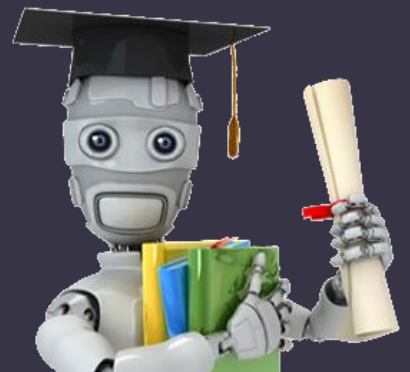
Introduction to Machine Learning

Instituto de Matemática Aplicada San Luis
Universidad Nacional de San Luis
July 13, 2017



PyData
San Luis

$ whoami

- Juan Martín Loyola
- BS. Computer Science
- Doctoral Student
- jmloyola@outlook.com

Disclaimer

- I'm not an expert in Machine Learning.
- I recently started learning Python

Before starting…

- Install Python (preferably > 3.6.0)
  - Python https://www.python.org/downloads/
  - Anaconda https://www.continuum.io/downloads
- Requirements:
  - Jupyter / iPython
  - Numpy
  - Matplotlib
- Download CIFAR-10 dataset (~165MB)
  - Run the shell script */Notebook/complementary_code/datasets/get_datasets.sh*
  - http://www.cs.toronto.edu/~kriz/cifar-10-python.tar.gz

Definition

- A field of study that gives computers the ability to learn without being explicitly programmed.
- A computer program is said to learn from experience **E** with respect to some class of tasks **T** and performance measure **P**, if its performance at tasks in **T**, as measured by **P**, improves with experience **E**.

Examples

- **A handwriting recognition learning problem:**
  - Task T: recognizing and classifying handwritten words within images
  - Performance measure *P:* percent of words correctly classified
  - Training experience E: a database of handwritten words with given classifications

- **A robot driving learning problem:**
  - Task T: driving on public four-lane highways using vision sensors
  - Performance measure P: average distance traveled before an error (as judged by human overseer)
  - Training experience E: a sequence of images and steering commands recorded while observing a human driver

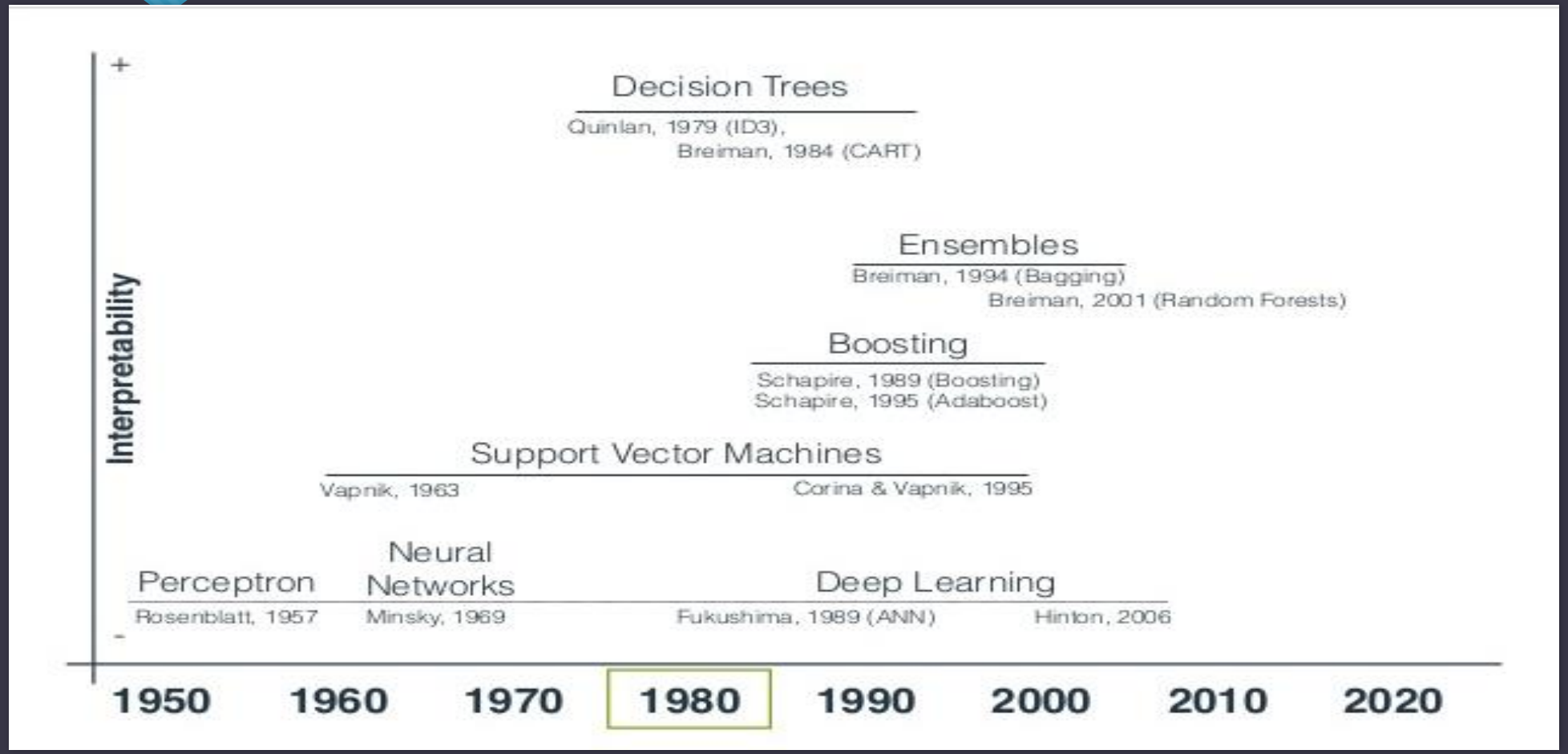
History

- Disciplines that influence Machine Learning:
  - Artificial Intelligence
  - Computational complexity theory
  - Information theory
  - Philosophy
  - Psychology and neurobiology
  - Bayesian methods
  - Statistics

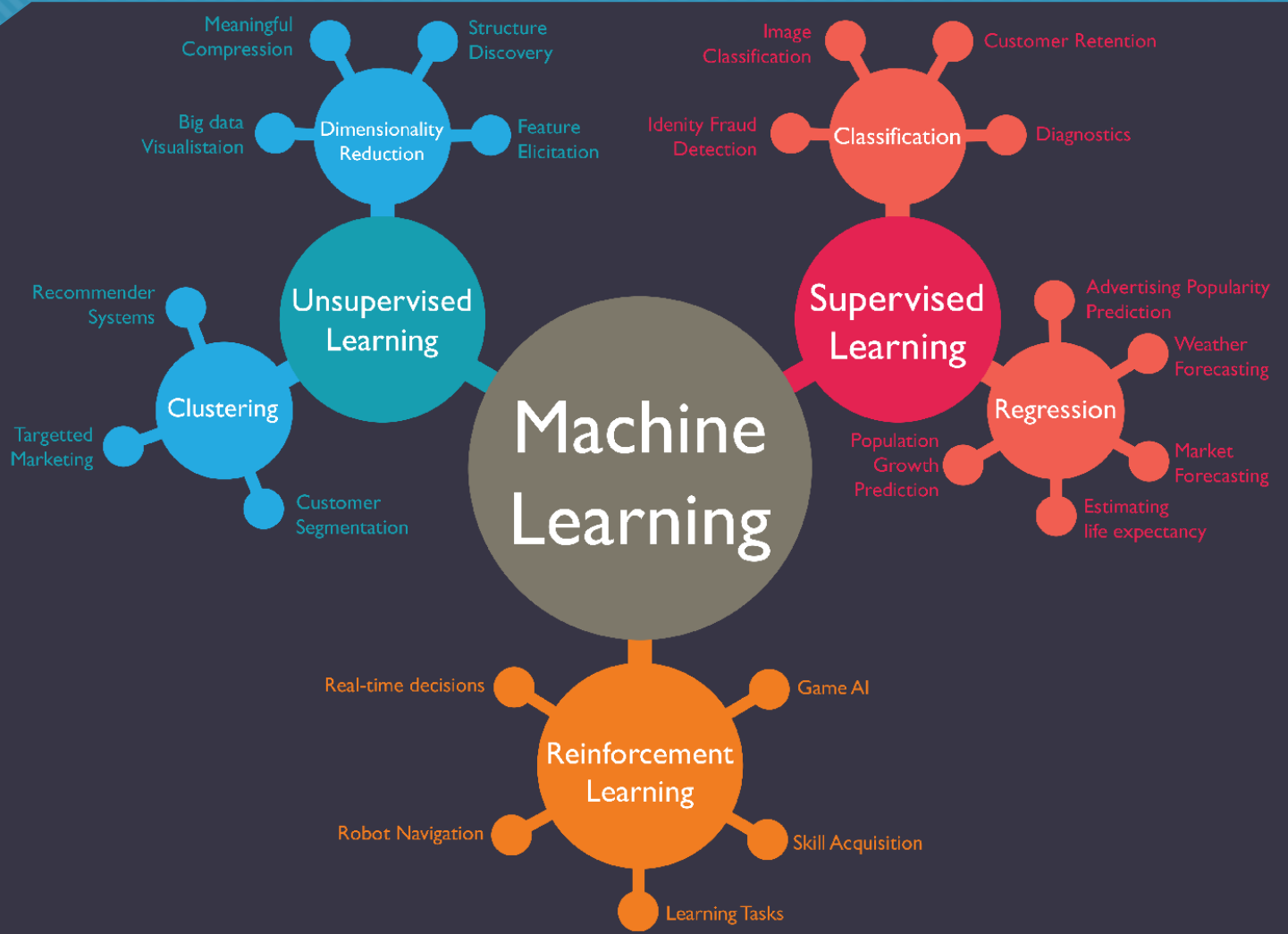
History

- 1812: Bayes' Theorem
- 1913: Markov Chains
- 1950: Turing's Learning Machine
- 1951: First Neural Network Machine
- 1952: Machines Playing Checkers
- 1957: Perceptron
- 1967: Nearest Neighbor
- 1969: Limitations of Neural Networks
- 1970: Automatic Differentiation (Backpropagation)
- 1982: Recurrent Neural Network
- 1989: Reinforcement Learning
- 1995: Random Forest Algorithm
- 1995: Support Vector Machines
- 1997: IBM Deep Blue Beats Kasparov
- 2011: Beating Humans in Jeopardy
- 2016: Beating Humans in Go

History


+
Interpretability
-
Decision Trees
Quinlan, 1979 (ID3),
Breiman, 1984 (CART)
Ensembles
Breiman, 1994 (Bagging)
Breiman, 2001 (Random Forests)
Boosting
Schapire, 1989 (Boosting)
Schapire, 1995 (Adaboost)
Support Vector Machines
Vapnik, 1963
Corina & Vapnik, 1995
Perceptron
Rosenblatt, 1957
Neural Networks
Minsky, 1969
Deep Learning
Fukushima, 1989 (ANN)
Hinton, 2006
1950
1960
1970
1980
1990
2000
2010
2020

Problems Machine Learning tries to solve

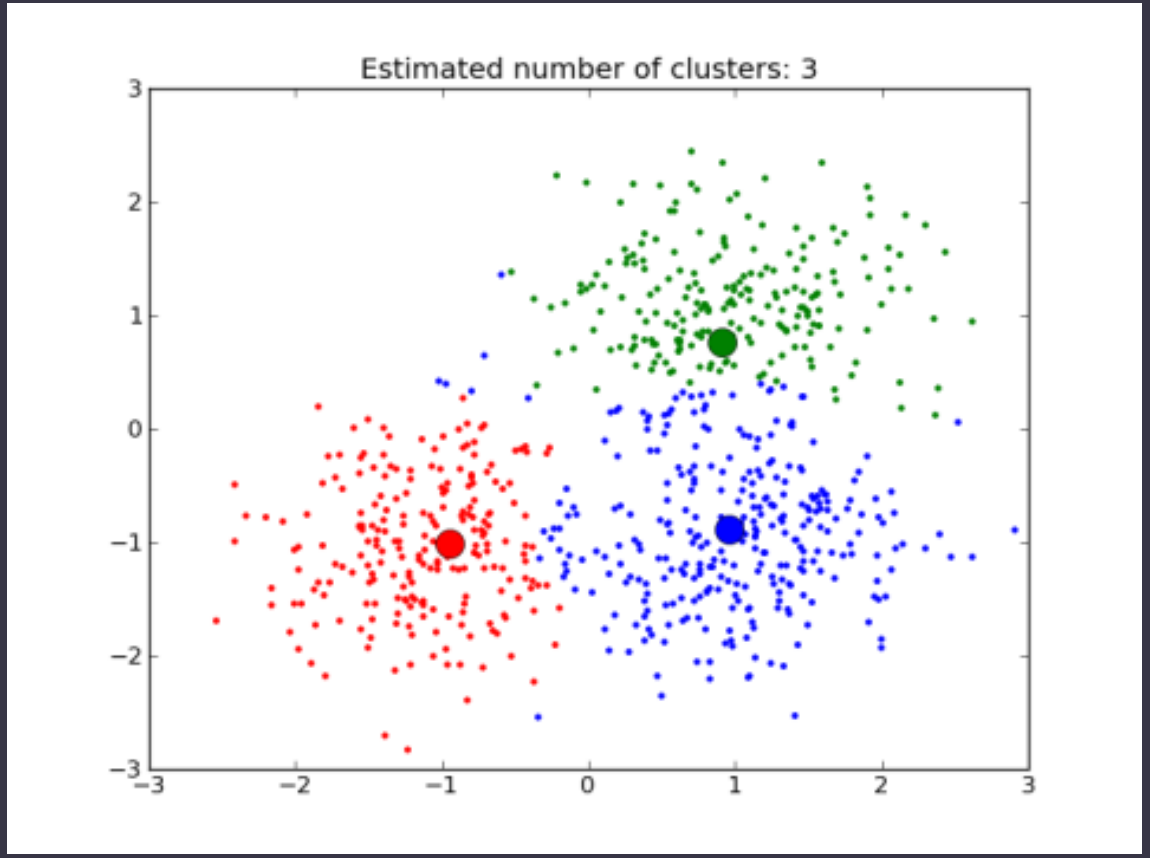


Supervised Learning

- Task of inferring a function from labeled training data.
- Each example is a pair consisting of an input object and a desired output value.
- A supervised learning algorithm analyzes the training data and produces an inferred function, which can be used for mapping new examples.

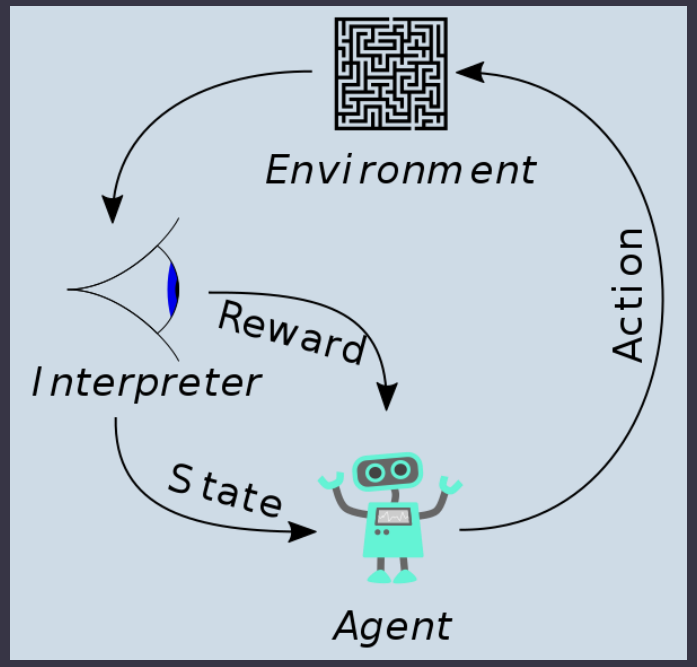
Unsupervised Learning

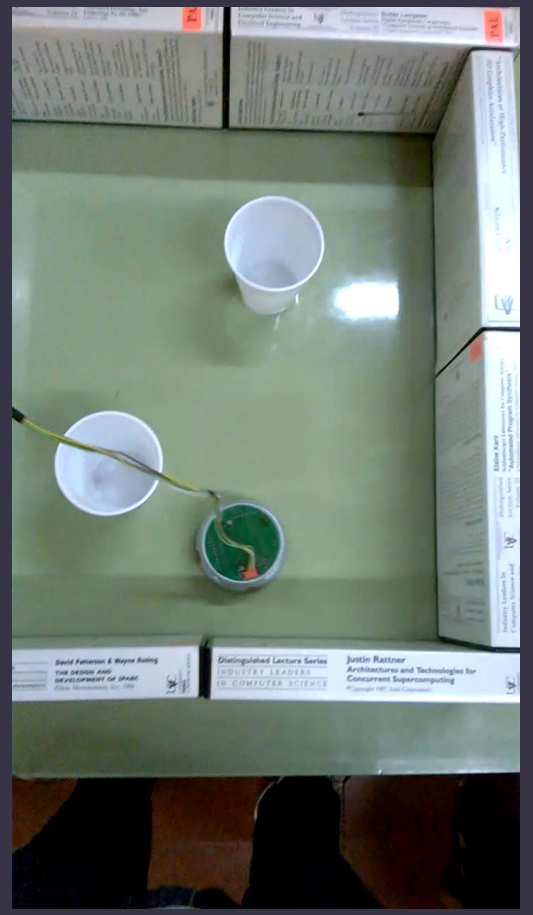
- Task of inferring a function to describe hidden structure from "unlabeled" data.
- Since the examples given to the learner are unlabeled, there is no evaluation of the accuracy of the structure that is output by the relevant algorithm



Reinforcement Learning

- Concerned with how agents ought to take actions in an environment so as to maximize some notion of cumulative reward.

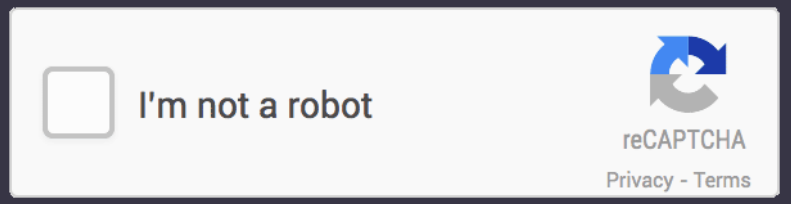


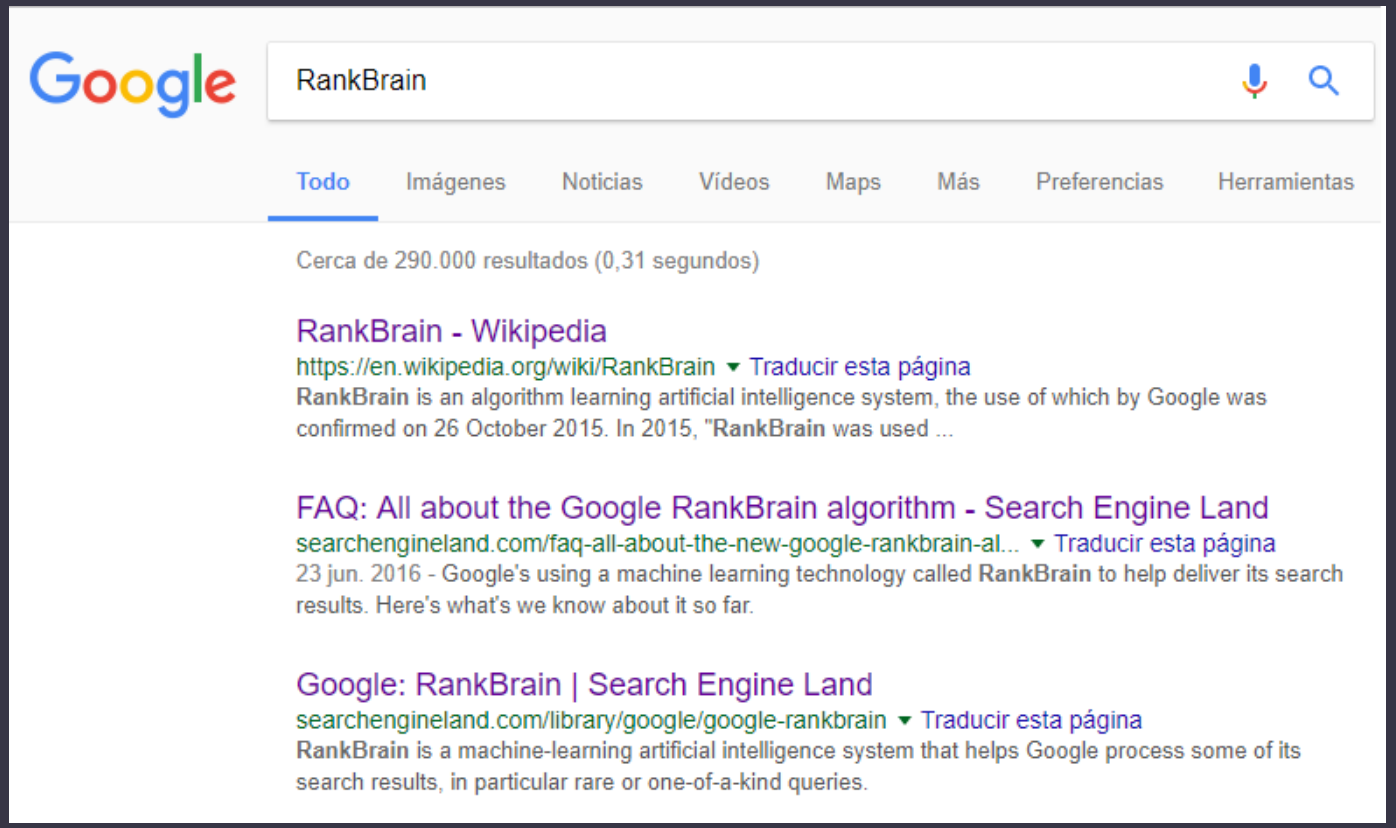


LEARNING = REPRESENTATION + EVALUATION + OPTIMIZATION

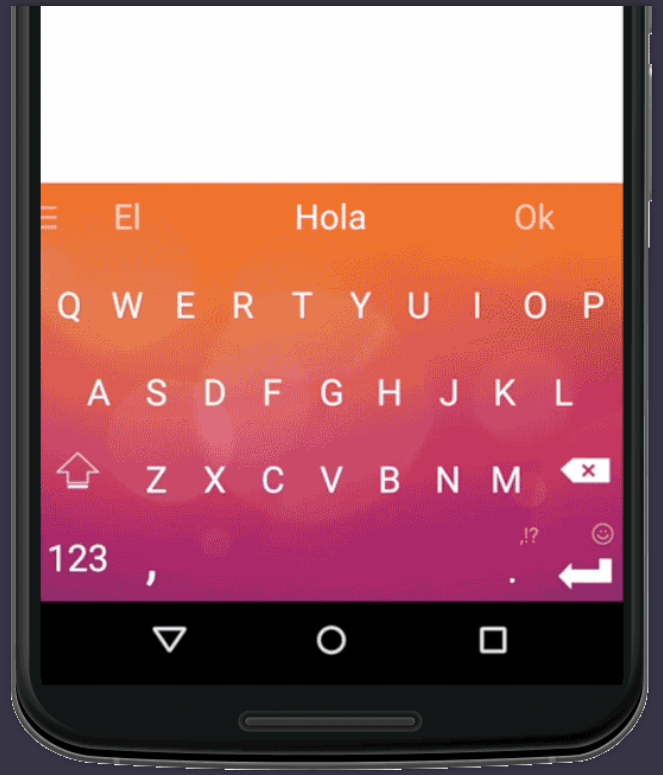
Representation	Evaluation	Optimization
Instances	Accuracy/Error rate	Combinatorial optimization
K-nearest neighbor	Precision and recall	Greedy search
Support vector machines	Squared error	Beam search
Hyperplanes	Likelihood	Branch-and-bound
Naive Bayes	Posterior probability	Continuous optimization
Logistic regression	Information gain	Unconstrained
Decision trees	K-L divergence	Gradient descent
Sets of rules	Cost/Utility	Conjugate gradient
Propositional rules	Margin	Quasi-Newton methods
Logic programs		Constrained
Neural networks		Linear programming
Graphical models		Quadratic programming
Bayesian networks		
Conditional random fields		

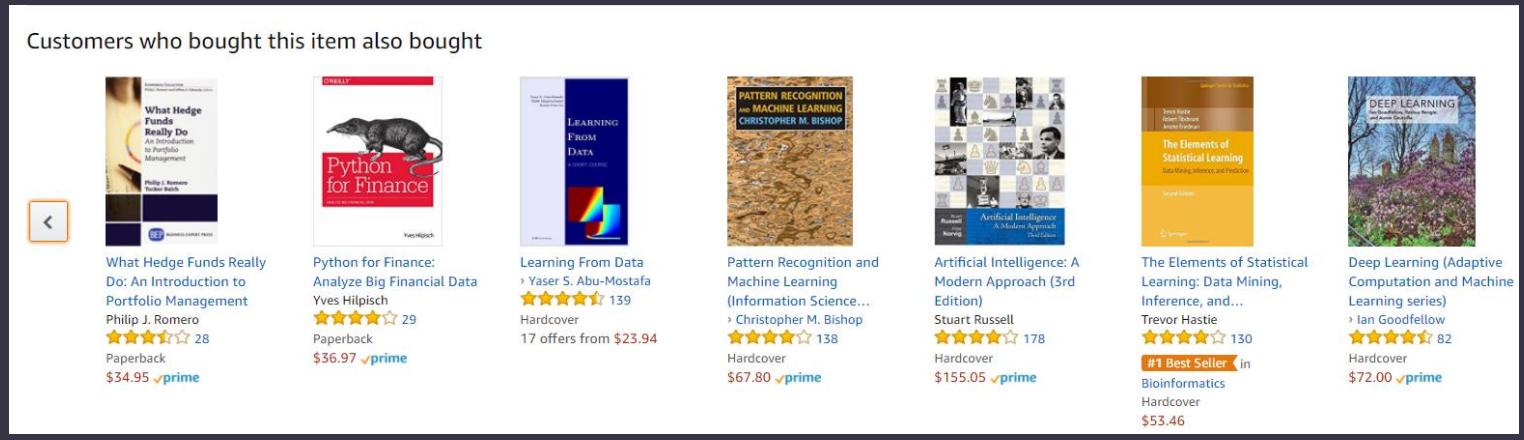
Applications

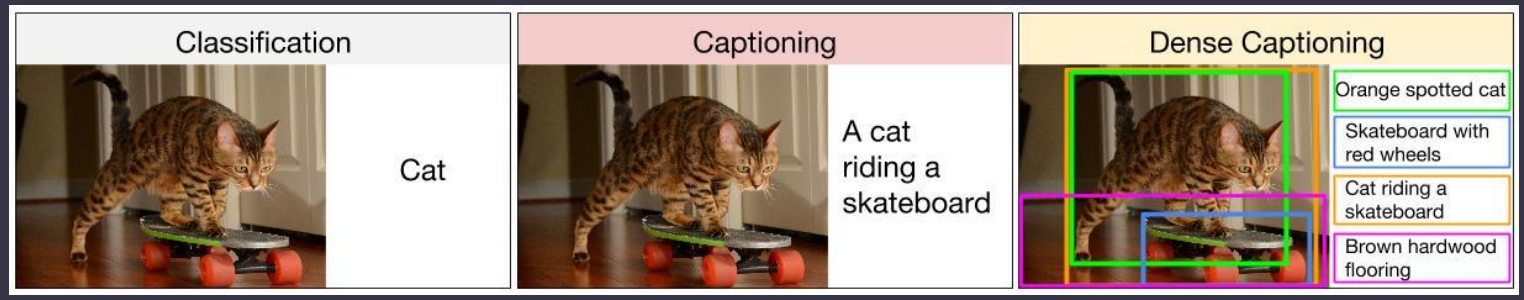




Applications





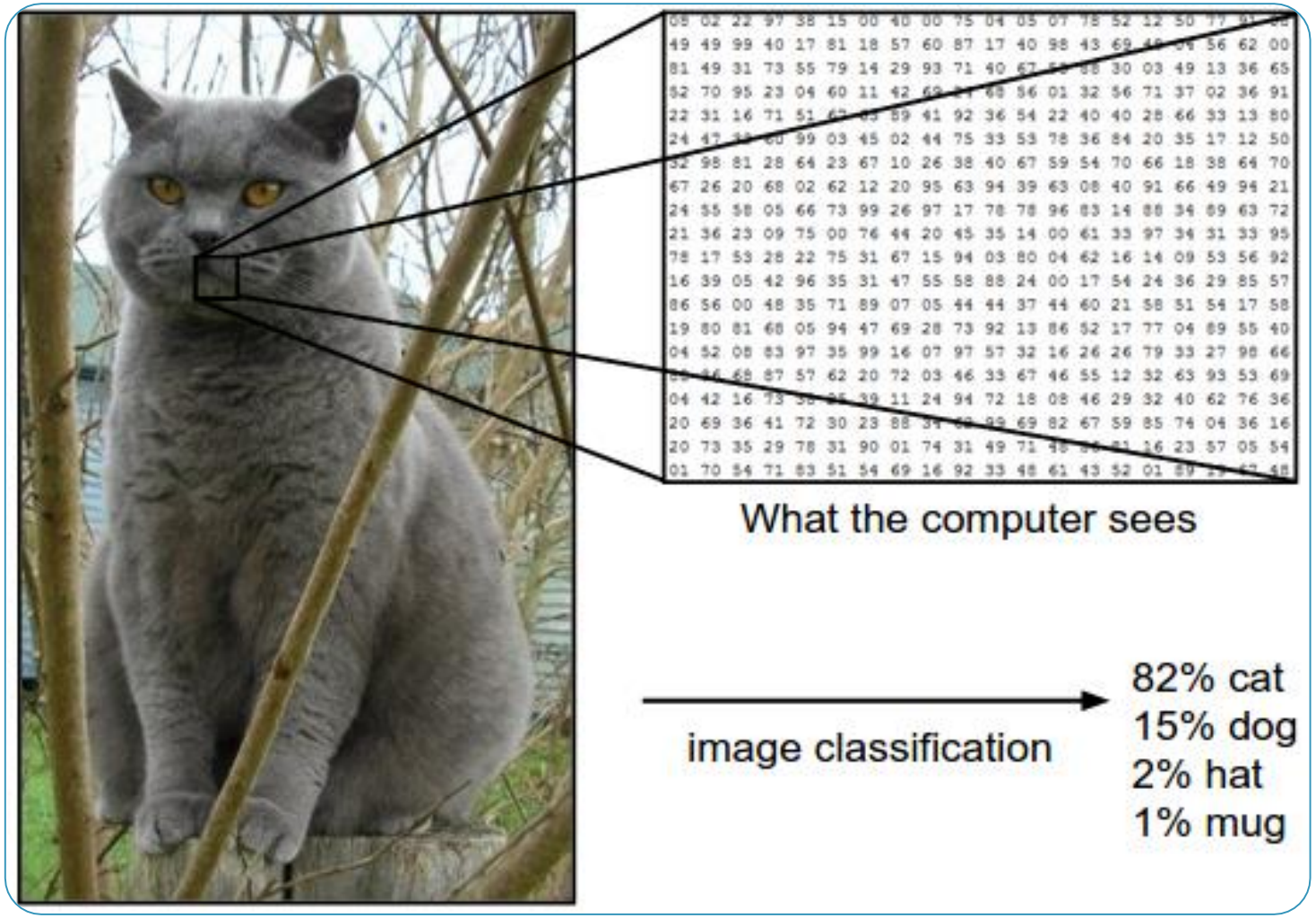


Applications

And many more…

- Prediction of credit-worthy applicants.
- Detection of credit card fraud.
- Recognition of handwritten digit and letter at the post office.
- Selecting relevant ads to show.
- Finding "People You Might Know" in Facebook (friend suggestions).
- Knowledge tracing and dropout detection for online courses.
- Playing chess, go and jeopardy against top players.
- Self-driving car.
- Recommendation systems (Netflix, YouTube).

Image Classification Problem



The intent of the classification process is to categorize all pixels in a digital image into one of several classes.

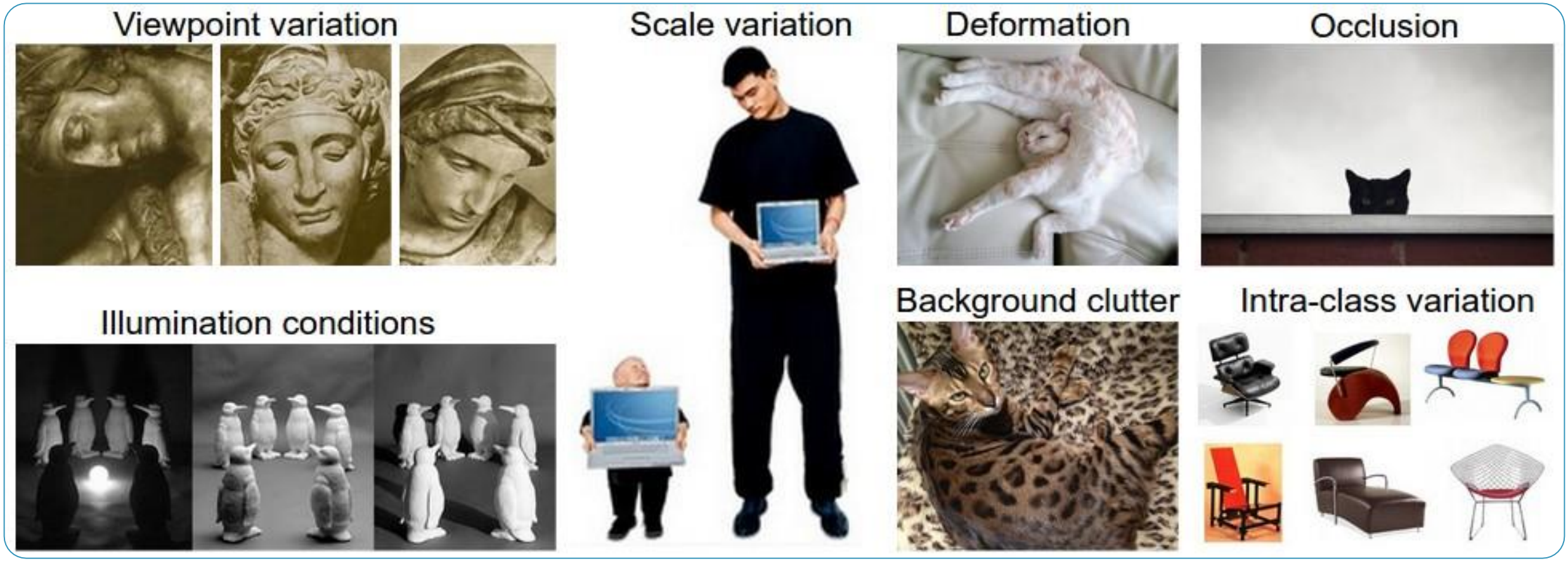
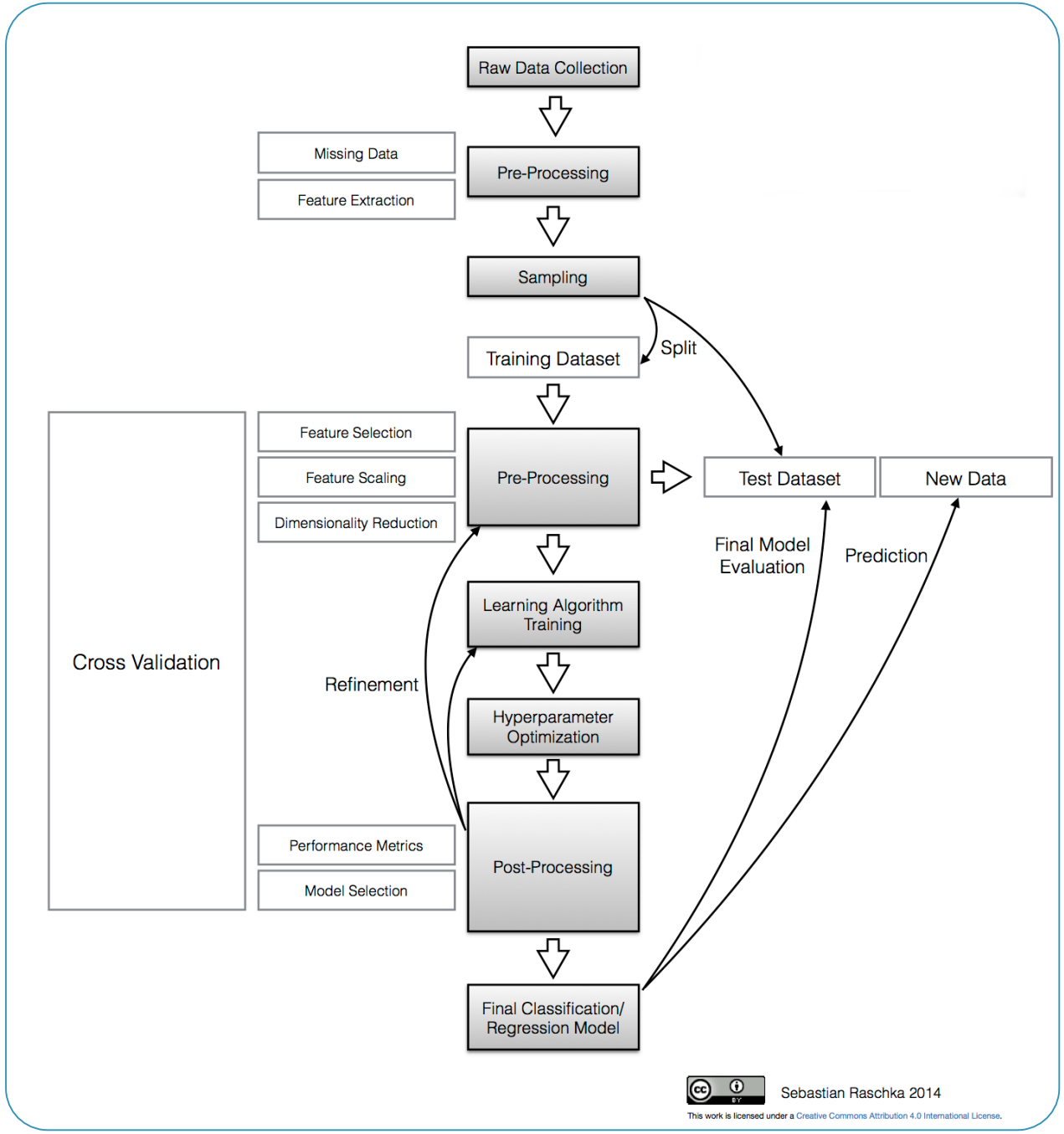
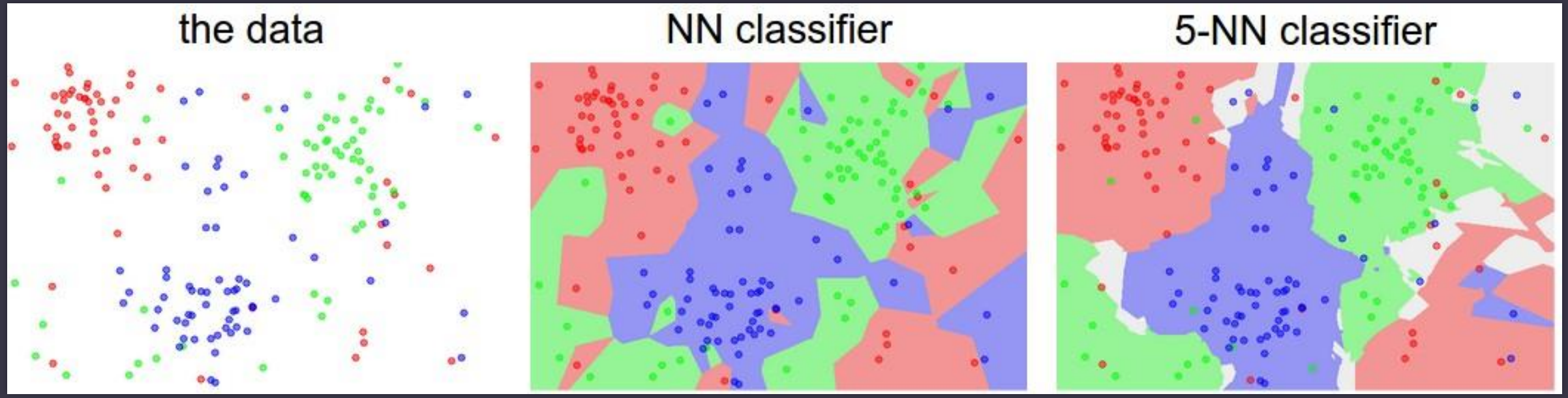


Image Classification Challenges

Supervised Learning Framework



Nearest Neighbor Classifier



Distance Metrics

- L1 (Manhattan) distance.

$$d_1(I_1, I_2) = \sum_p |I_1^p - I_2^p|$$

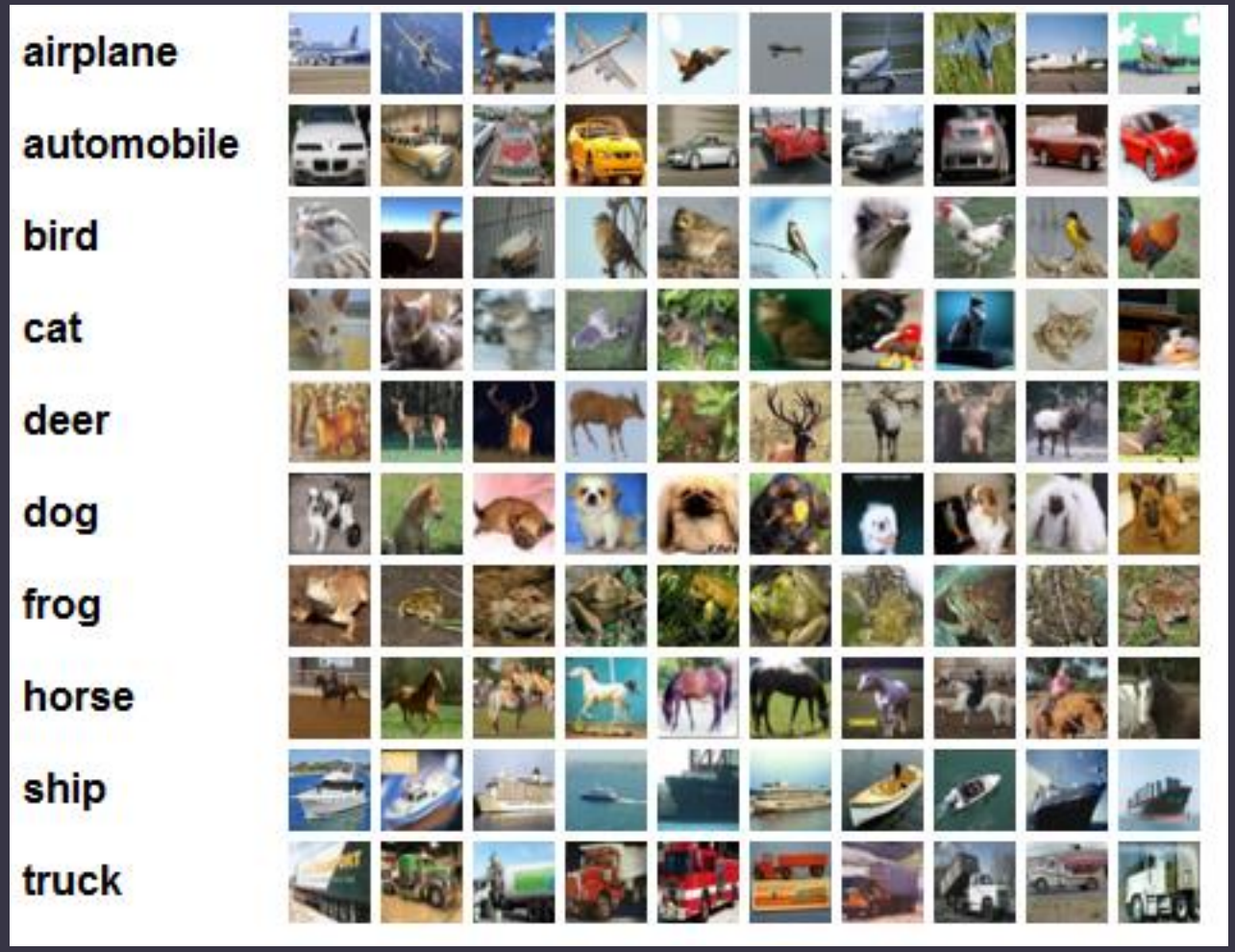
- L2 (Euclidean) distance.

$$d_2(I_1, I_2) = \sqrt{\sum_p (I_1^p - I_2^p)^2}$$

PyData San Luis

CIFAR-10 Dataset

- 10 labels
- 50000 training images
- 10000 test images



Supervised Learning Framework

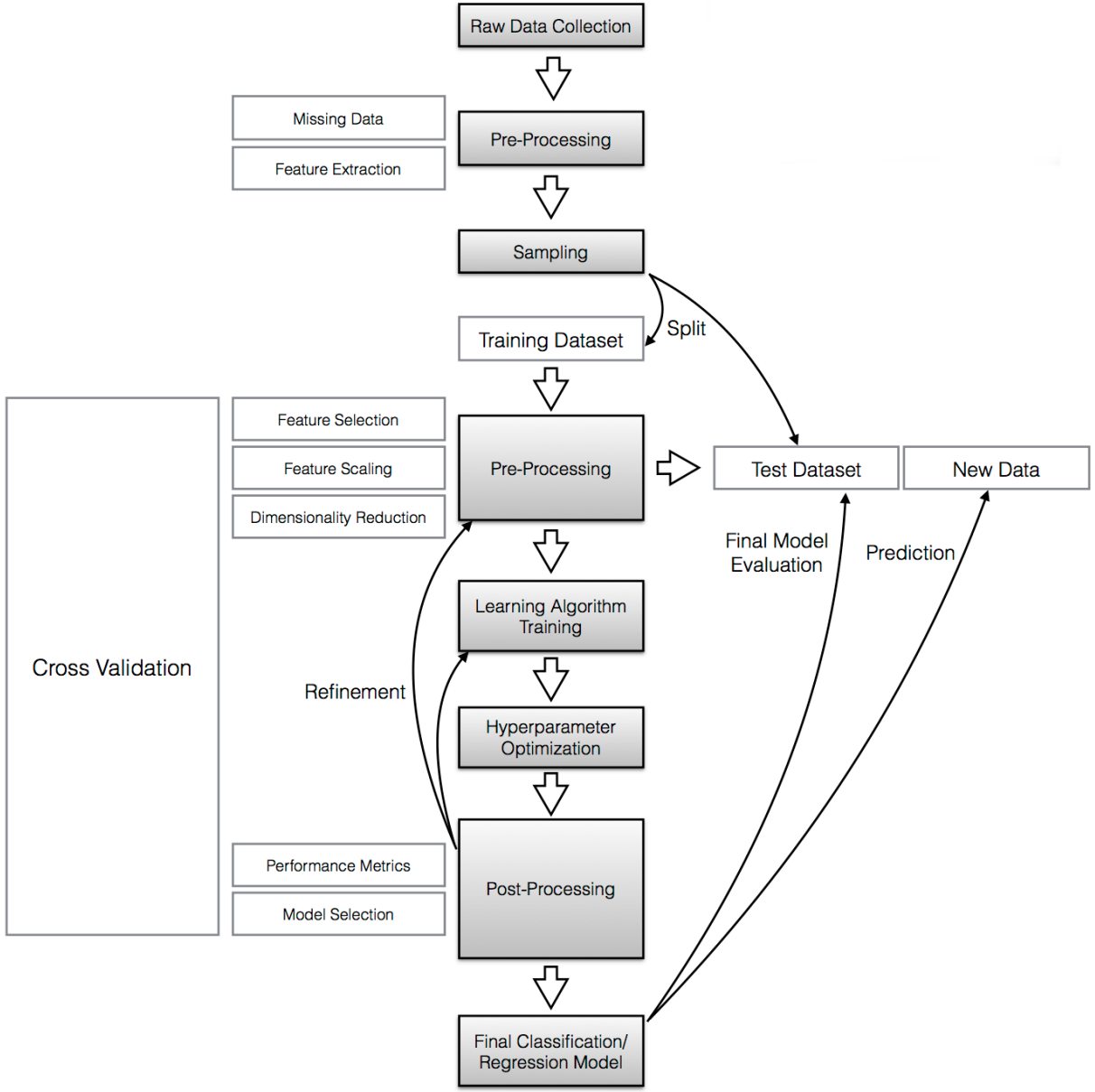


Image Classification Assignment

https://github.com/PyDataSanLuis/Encuentros

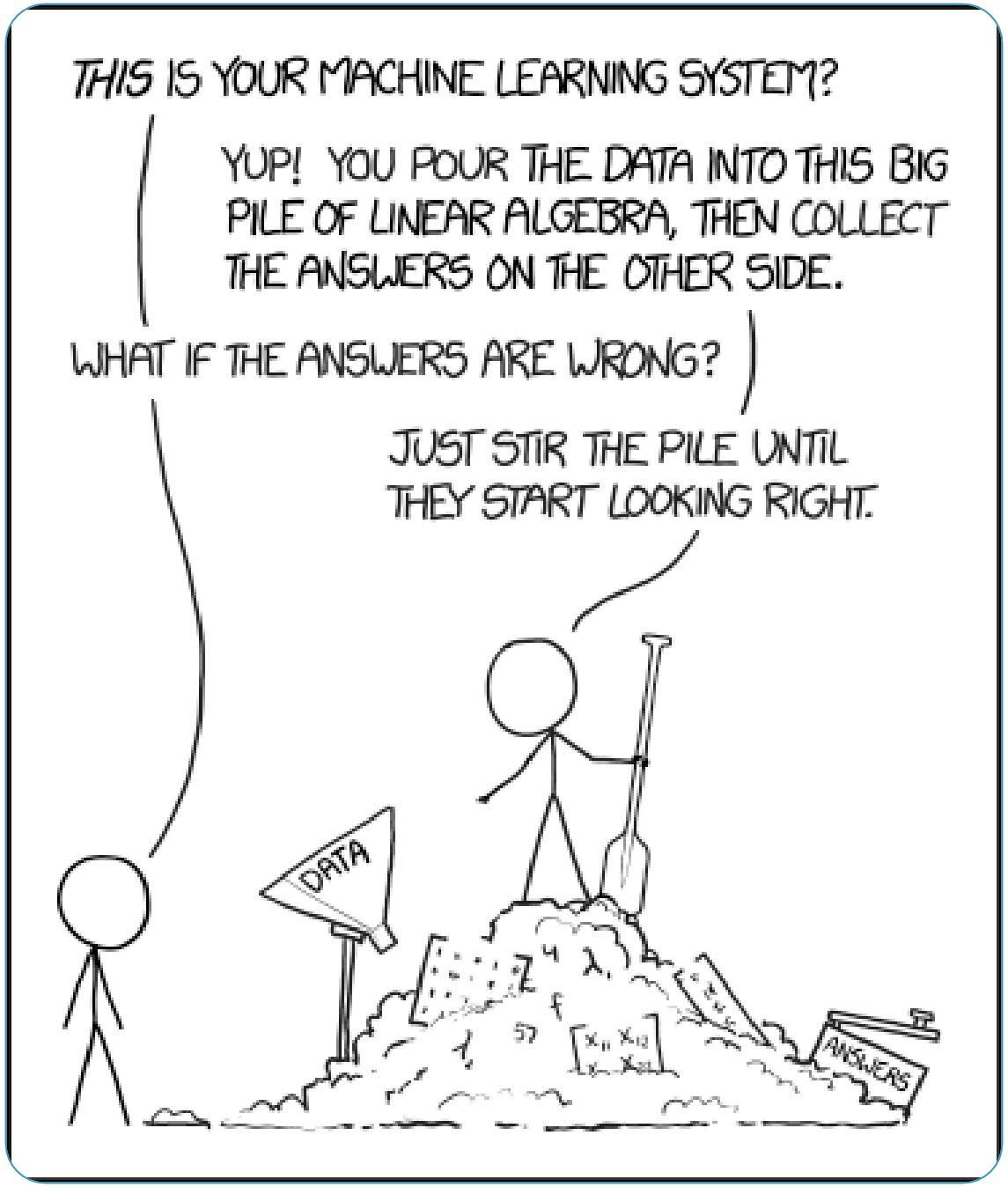
Frameworks

- Scikit-learn http://scikit-learn.org/stable/
- WEKA http://www.cs.waikato.ac.nz/ml/weka/
- TensorFlow https://www.tensorflow.org/
- Theano http://deeplearning.net/software/theano/
- PyMC3 http://pymc-devs.github.io/pymc3/

- R packages:
  - nnet
  - randomForest
  - caret
  - kernlab
  - tree

Further Reading

- "Machine Learning". Online Course in Coursera from Andrew Ng. https://www.coursera.org/learn/machine-learning
- "Convolutional Neural Networks for Visual Recognition". Stanford's Course from Fei-Fei Li. http://cs231n.github.io/
- "Machine Learning". Book from Tom Mitchell (1997).
- "Pattern Recognition and Machine Learning". Book from Michael Bishop (2006).
- "The Elements of Statistical Learning: Data Mining, Inference, and Prediction". Book from Trevor Hastie, Robert Tibshirani and Jerome Friedman (2009).


THIS IS YOUR MACHINE LEARNING SYSTEM?
YUP! YOU POUR THE DATA INTO THIS BIG PILE OF LINEAR ALGEBRA, THEN COLLECT THE ANSWERS ON THE OTHER SIDE.
WHAT IF THE ANSWERS ARE WRONG?
JUST STIR THE PILE UNTIL THEY START LOOKING RIGHT.
DATA
ANSWERS

Questions?