NBTC – ITU Training
on Building IoT solutions for e-applications

Session 3: Machine learning | A Big picture

CELEBRATING 25 YEARS OF ACHIEVEMENTS

A MOTIVATING EXAMPLE

How a Japanese cucumber farmer is using deep learning and TensorFlow

Wednesday, August 31, 2016

Posted by Kaz Sato, Developer Advocate, Google Cloud Platform

https://cloud.google.com/blog/big-data/2016/08/how-a-japanese-cucumber-farmer-is-using-deep-learning-and-tensorflow

WHAT IS MACHINE LEARNING [1] ?

- GIVING COMPUTERS THE ABILITY TO LEARN FROM DATA
- PREDICTING vs. EXPLAINING
- CAN BE SUPERVISED, UNSUPERVISED, REINFORCEMENT LEARNING

MACHINE LEARNING PRACTITIONER CARICATURE?

https://xkcd.com/1838/

CELEBRATING 25 YEARS OF ACHIEVEMENTS

WHAT IS MACHINE LEARNING [2] ?

SUPERVISED LEARNING

- Labeled data
- Direct feedback
- Predict outcome/future

UNSUPERVISED LEARNING

- No labels/targets
- No feedback
- Find hidden structure in data

REINFORCEMENT LEARNING

- Decision process
- Reward system
- Learn series of action

S. Raschka & V. Mirjalili, Python Machine Learning, Second Edition, Packt, 2017

SUPERVISED LEARNING

scikits learn
machine learning in Python

The problem solved in supervised learning

Supervised learning consists in learning the link between two datasets: the observed data `X` and an external variable `y` that we are trying to predict, usually called "target" or "labels". Most often, `y` is a 1D array of length `n_samples`.

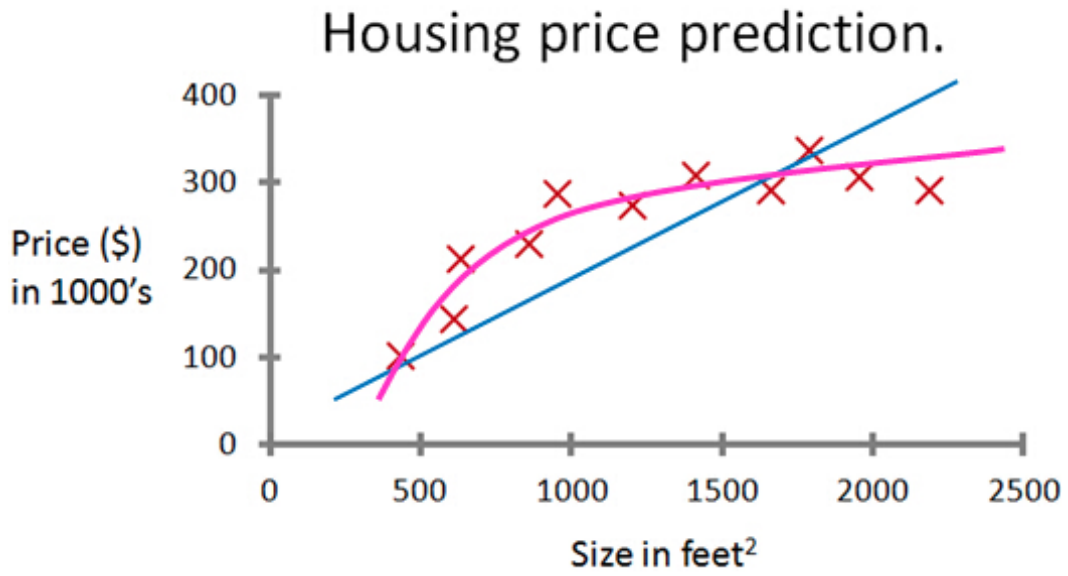
Vocabulary: classification and regression

If the prediction task is to classify the observations in a set of finite labels, in other words to "name" the objects observed, the task is said to be a **classification** task. On the other hand, if the goal is to predict a continuous target variable, it is said to be a **regression** task.

http://scikit-learn.org/stable/tutorial/statistical_inference/supervised_learning.html

SUPERVISED LEARNING - REGRESSION

TARGET
CONTINUOUS

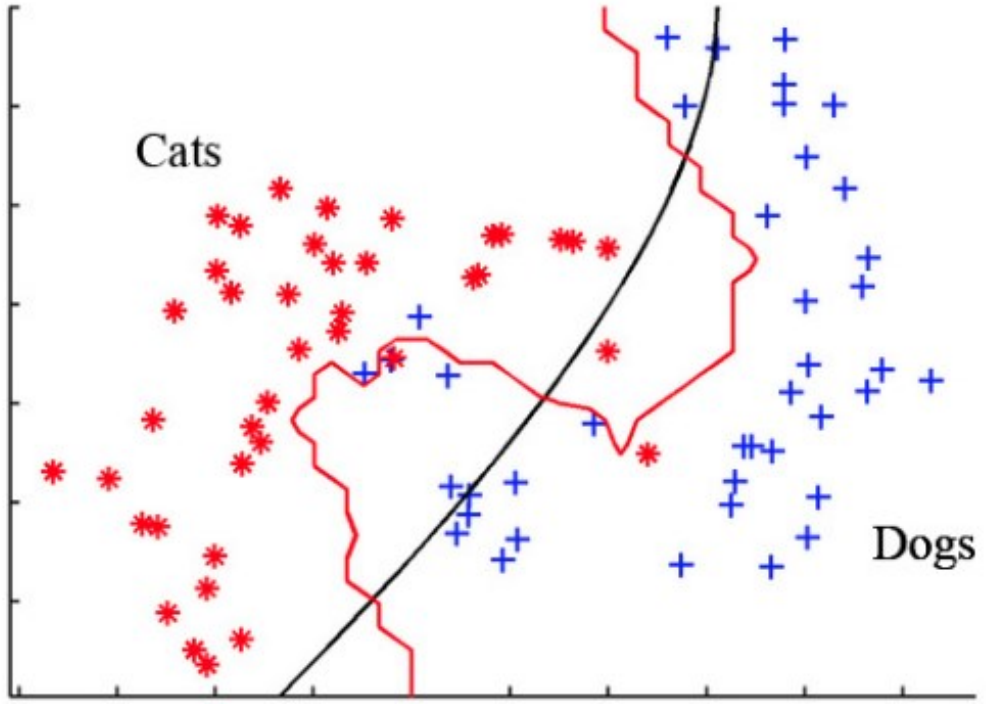


FEATURE(S)

SUPERVISED LEARNING - CLASSIFICATION

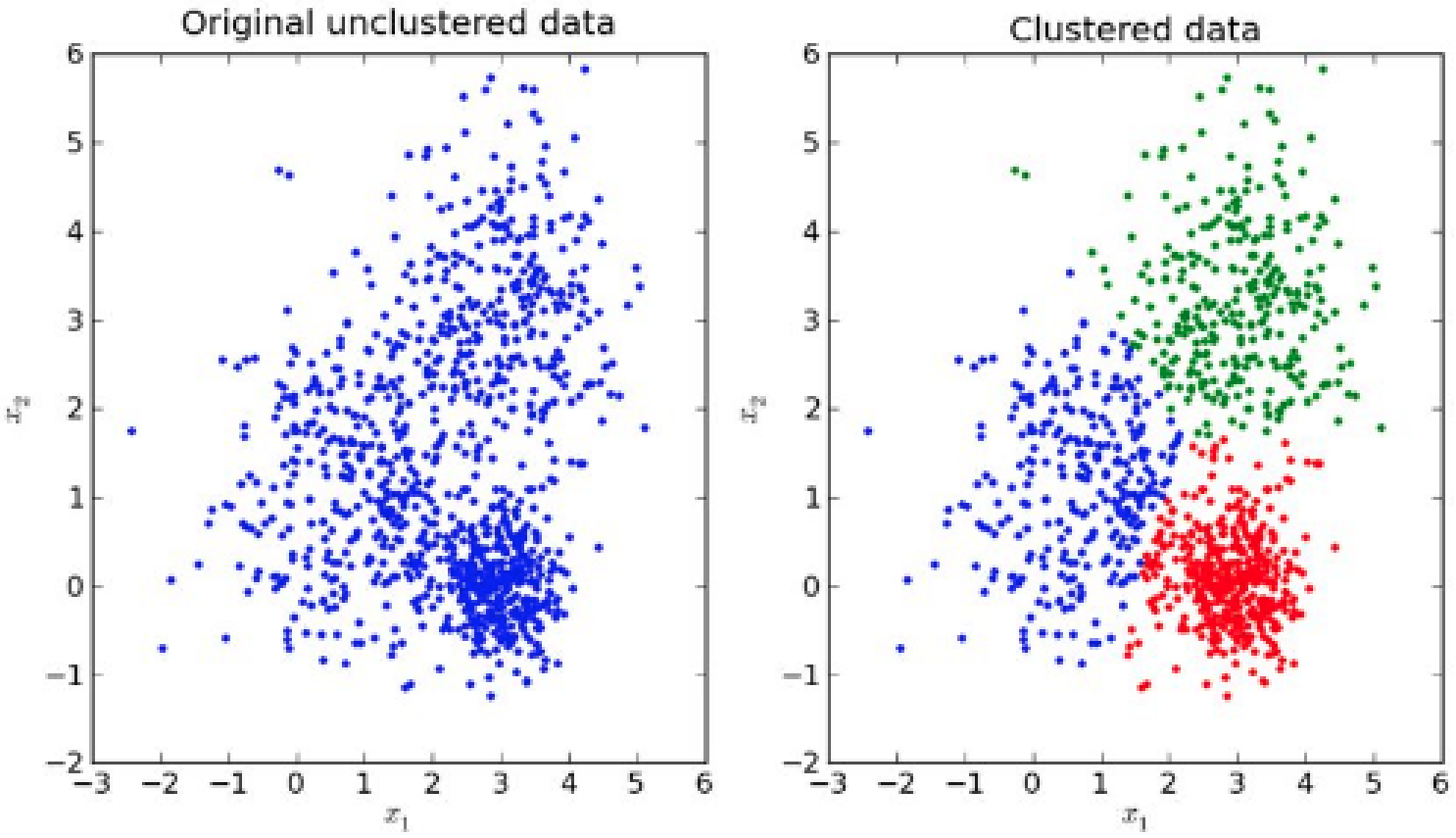
TARGET DISCRETE: DOGS vs CATS

FEATURE 2



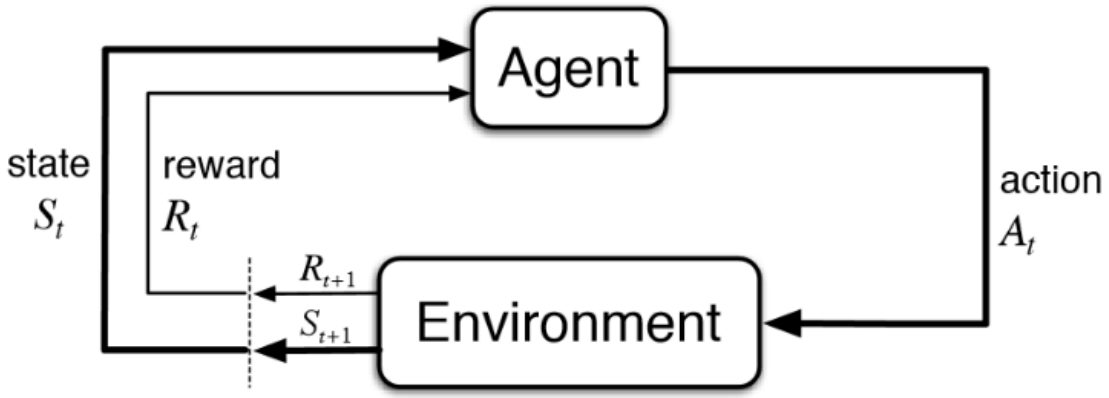
FEATURE 1

UNSUPERVISED LEARNING | DISCOVERING HIDDEN STRUCTURE



http://scikit-learn.org/stable/unsupervised_learning.html#unsupervised-learning

REINFORCEMENT LEARNING



EXAMPLE: ROBOT LEARNING HOW TO PLAY TABLE TENNIS

https://www.youtube.com/watch?v=SH3bADiB7uQ

THE LEARNING PROCESS [1]

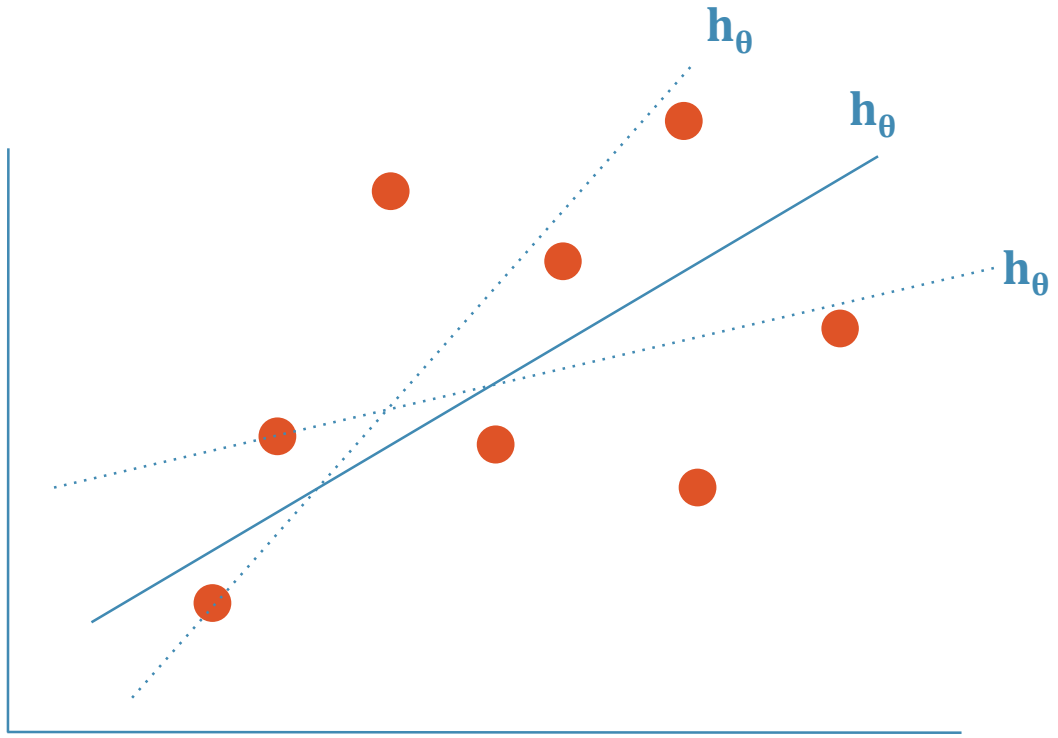
ZOOM ON SUPERVISED LEARNING & REGRESSION

- DATA (TRAINING/TEST)
- HYPOTHESIS FUNCTION

$$h_\theta(X) = \theta_0 + \theta_1 X$$

- COST FUNCTION TO MINIMIZE

$$J(\theta_0, \theta_1) = \sum(h\theta(x(i)) - y(i))2$$



THE LEARNING PROCESS [2]

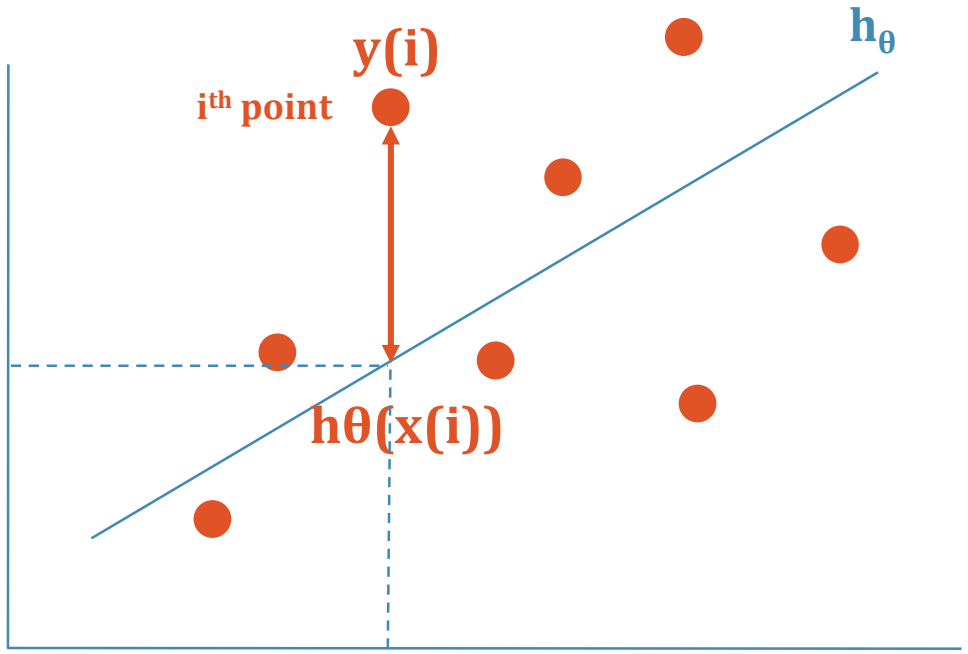
ZOOM ON SUPERVISED LEARNING & REGRESSION

- DATA (TRAINING/TEST)
- HYPOTHESIS FUNCTION

$$h_\theta(X) = \theta_0 + \theta_1 X$$

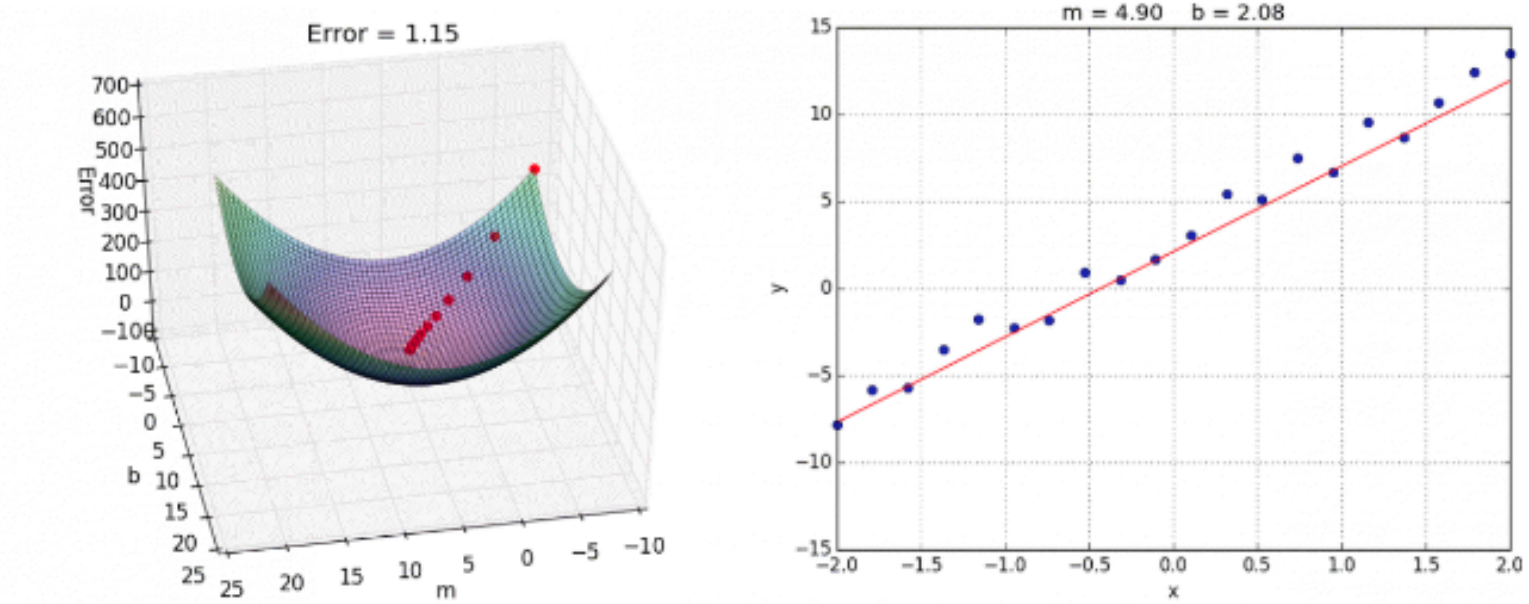
- COST FUNCTION TO MINIMIZE

$$J(\theta_0, \theta_1) = \sum(h\theta(x(i)) - y(i))2$$



THE LEARNING PROCESS [3]

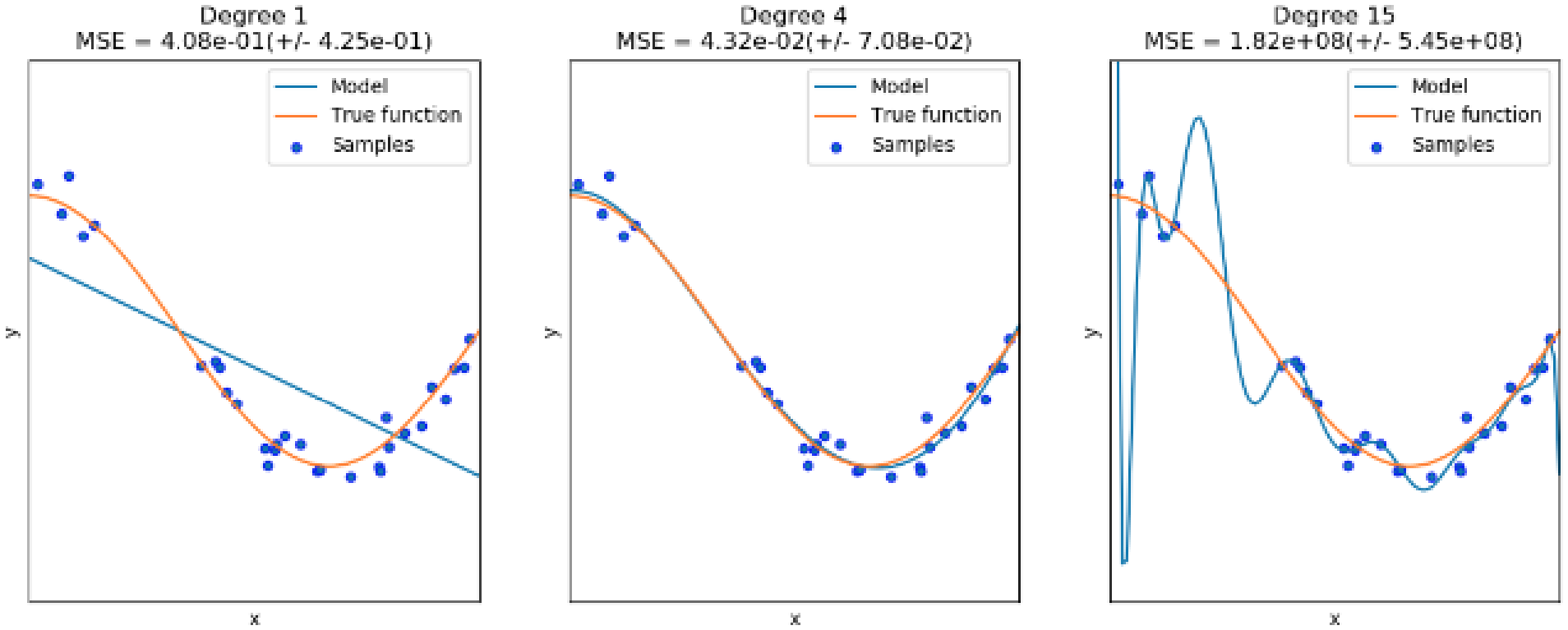
ZOOM ON SUPERVISED LEARNING & REGRESSION



20 *steps of gradient descent with learning rate of* 0.01

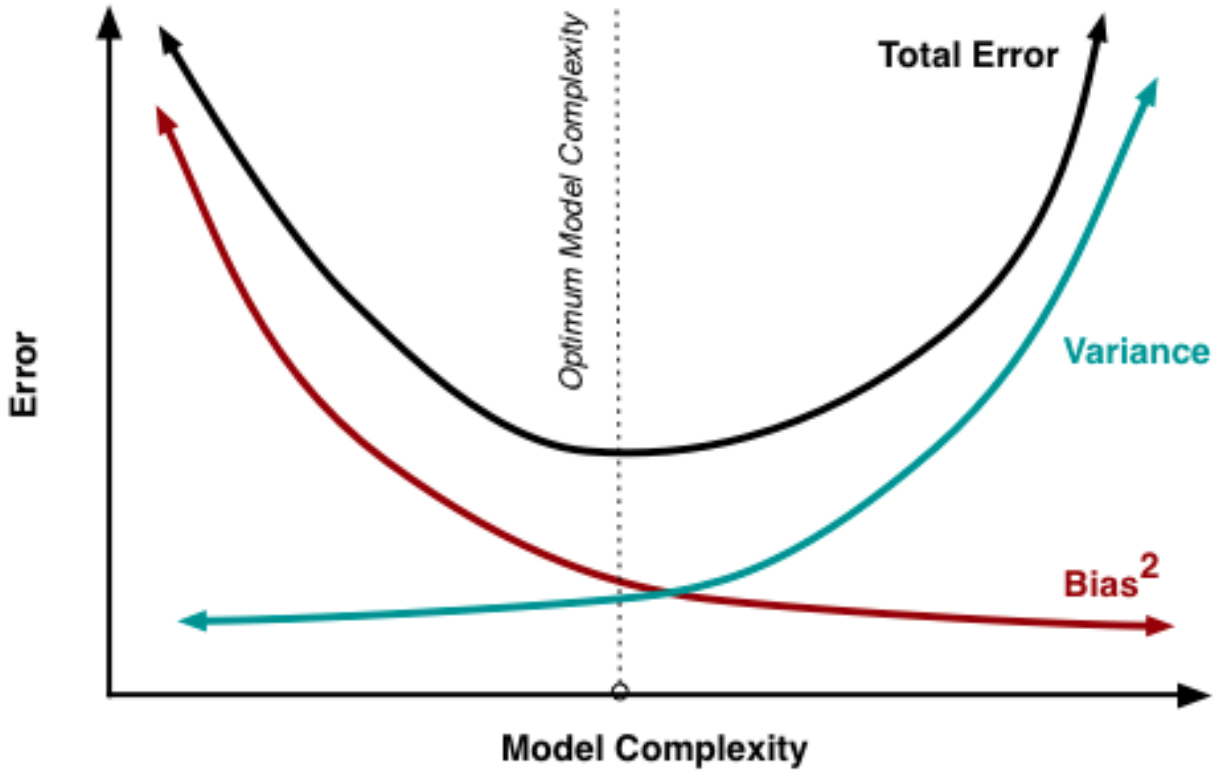
https://alykhantejani.github.io/a-brief-introduction-to-gradient-descent/

BIAS/VARIANCE & OVER/UNDERFITTING



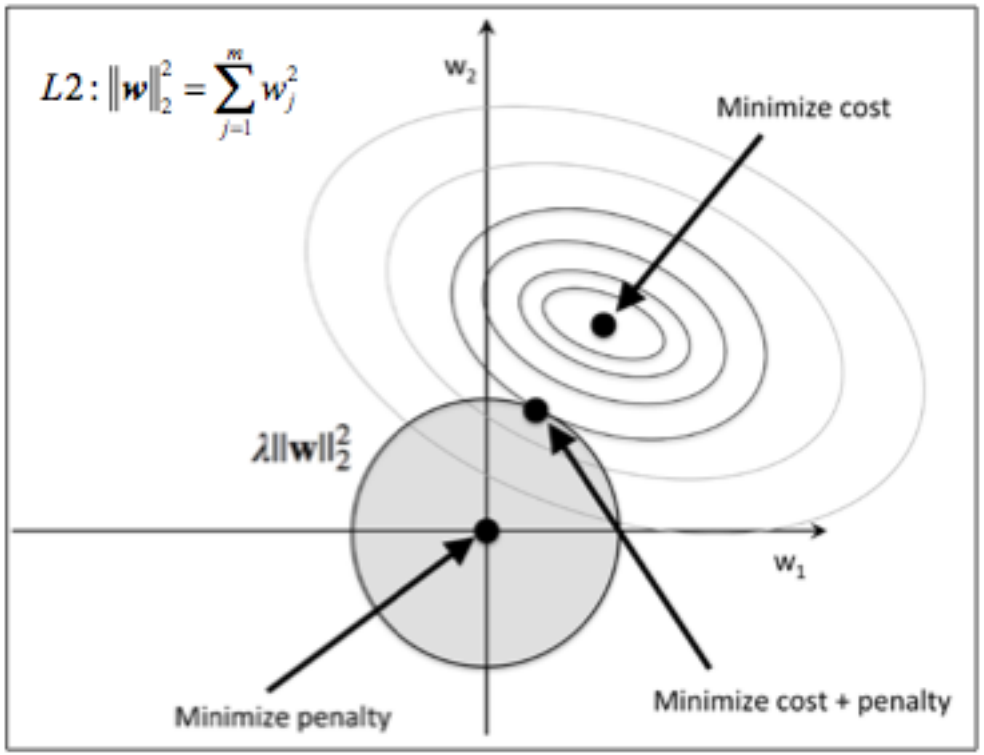
http://scikit-learn.org/stable/auto_examples/model_selection/plot_underfitting_overfitting.html

DIAGNOSING BIAS/VARIANCE

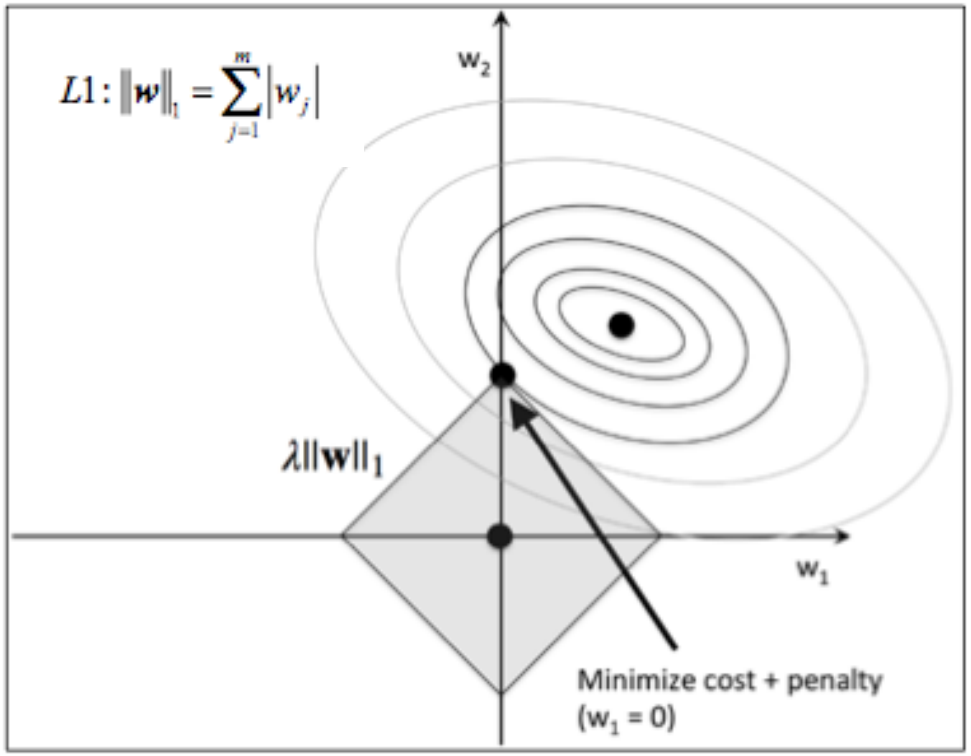


https://www.researchgate.net/post/How_does_model_complexity_impact_the_bias-variance_tradeoff

REGULARIZATION | PENALIZING COST FUNCTION



RIDGE REGRESSION



LASSO REGRESSION

S. Raschka & V. Mirjalili, Python Machine Learning, Second Edition, Packt, 2017

CELEBRATING 25 YEARS OF ACHIEVEMENTS

BUT THIS IS ONLY ONE TYPE OF ALGORITHM!
WHAT ELSE?

https://machinelearningmastery.com/a-tour-of-machine-learning-algorithms/

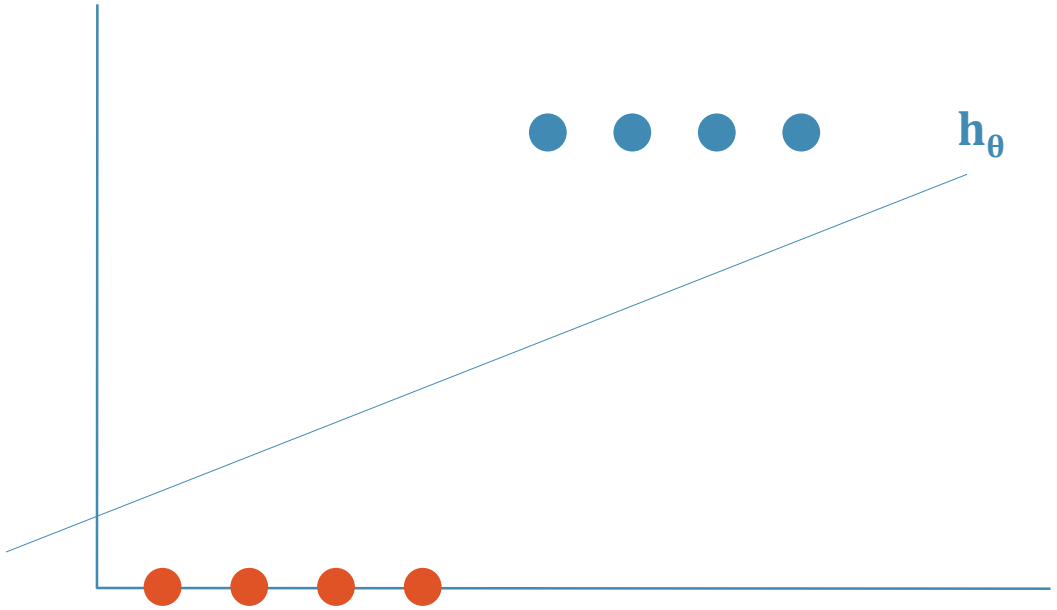
CELEBRATING 25 YEARS OF ACHIEVEMENTS

CLASSIFICATION | LOGISTIC REGRESSION [1]

- DATA (TRAINING/TEST)
- HYPOTHESIS FUNCTION

$$h_\theta(X) = \theta_0 + \theta_1 X \quad ?$$

- COST FUNCTION TO MINIMIZE

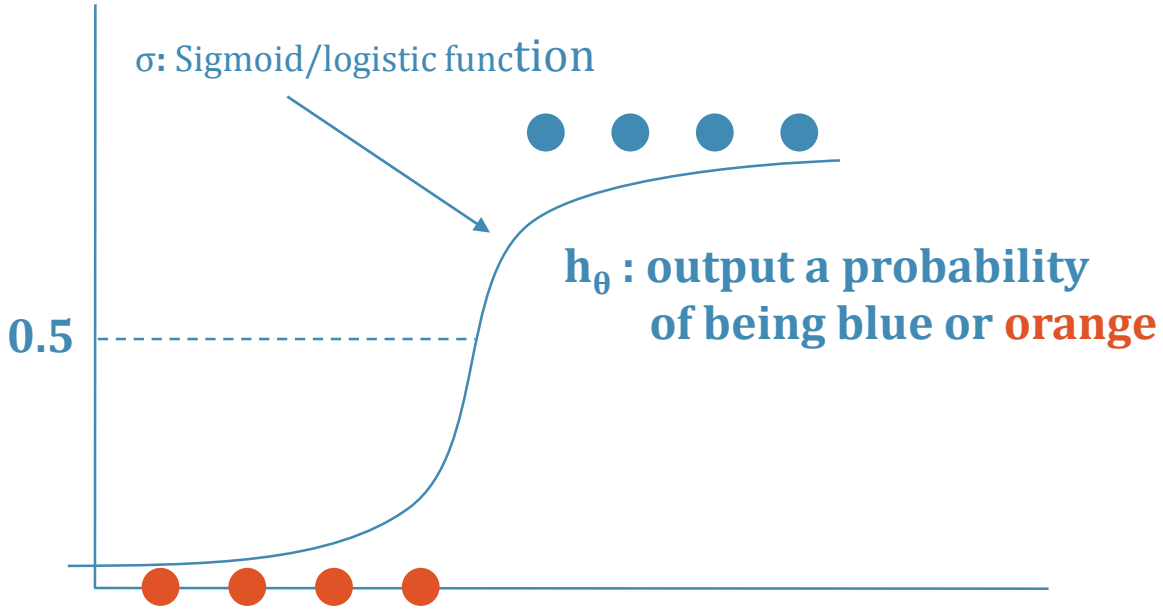


CLASSIFICATION | LOGISTIC REGRESSION [2]

- DATA (TRAINING/TEST)
- HYPOTHESIS FUNCTION

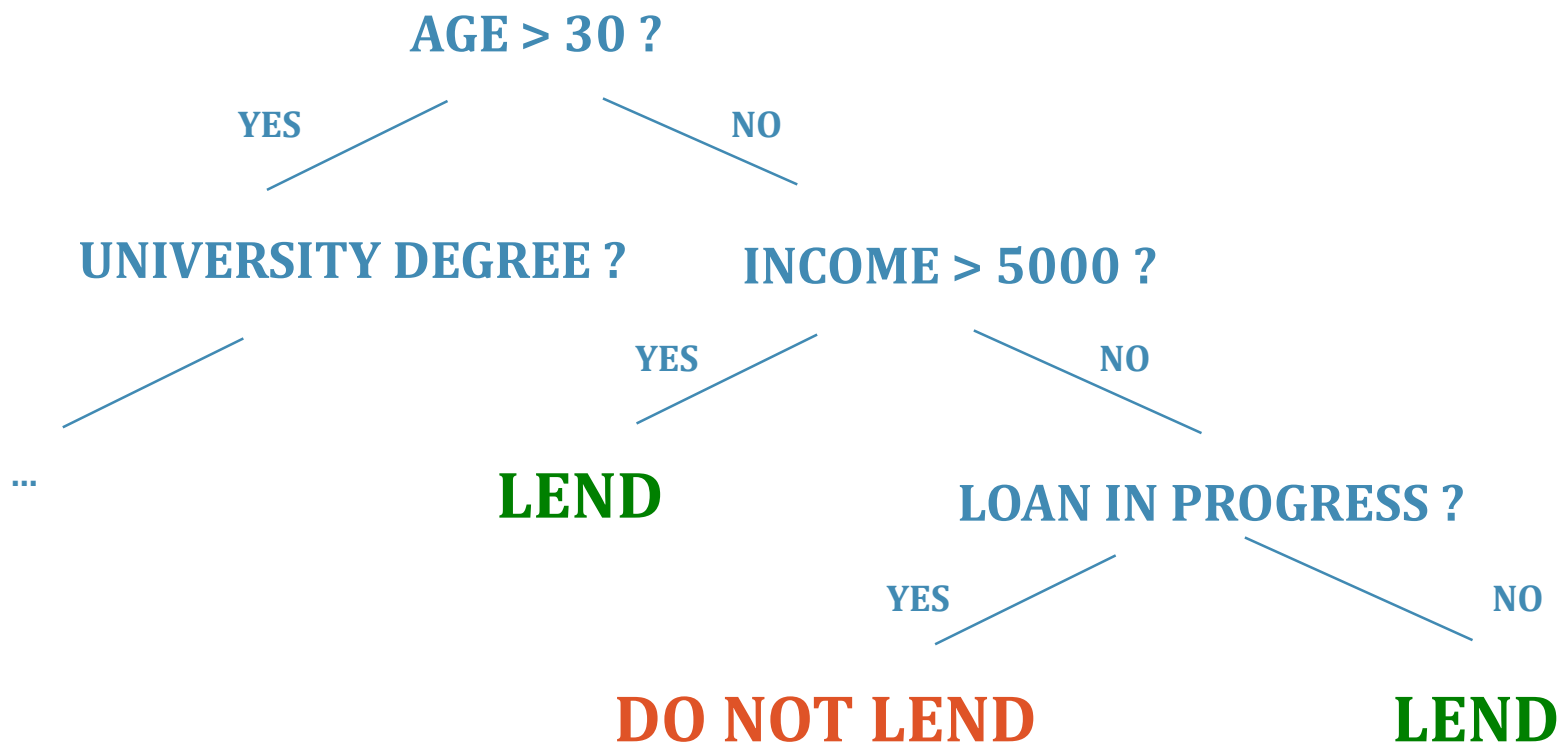
$$h_\theta(X) = \sigma(\theta_0 + \theta_1 X)$$

- COST FUNCTION TO MINIMIZE

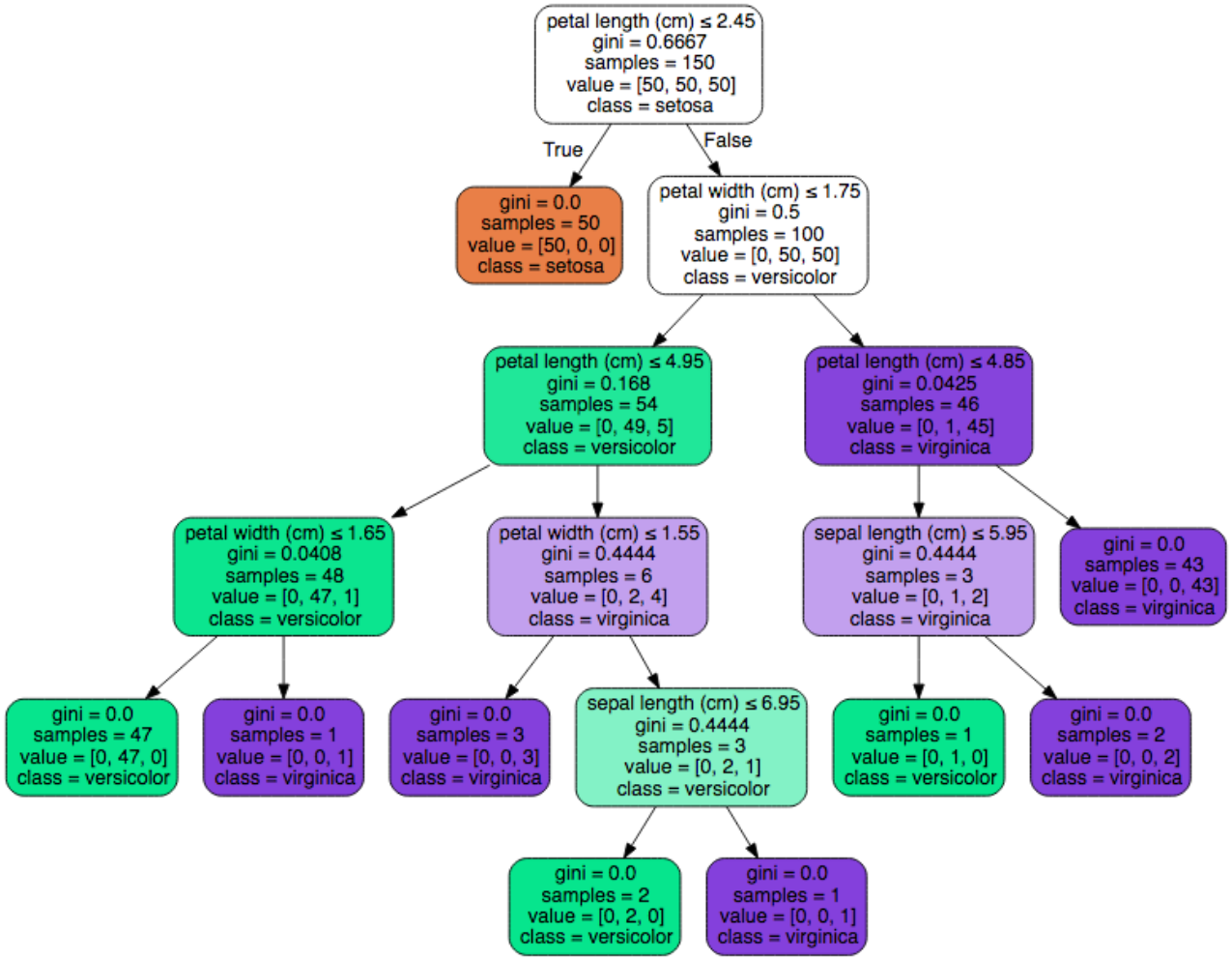


CLASSIFICATION | DECISION TREE

A TOY EXAMPLE: ARRANGING A LOAN [the 20 questions game]



DECISION TREE | IRIS DATASET



IRIS dataset

Iris Versicolor

Iris Virginica

Iris Setosa

http://scikit-learn.org/stable/modules/tree.html#classification

DECISION TREE | ENSEMBLE METHODS

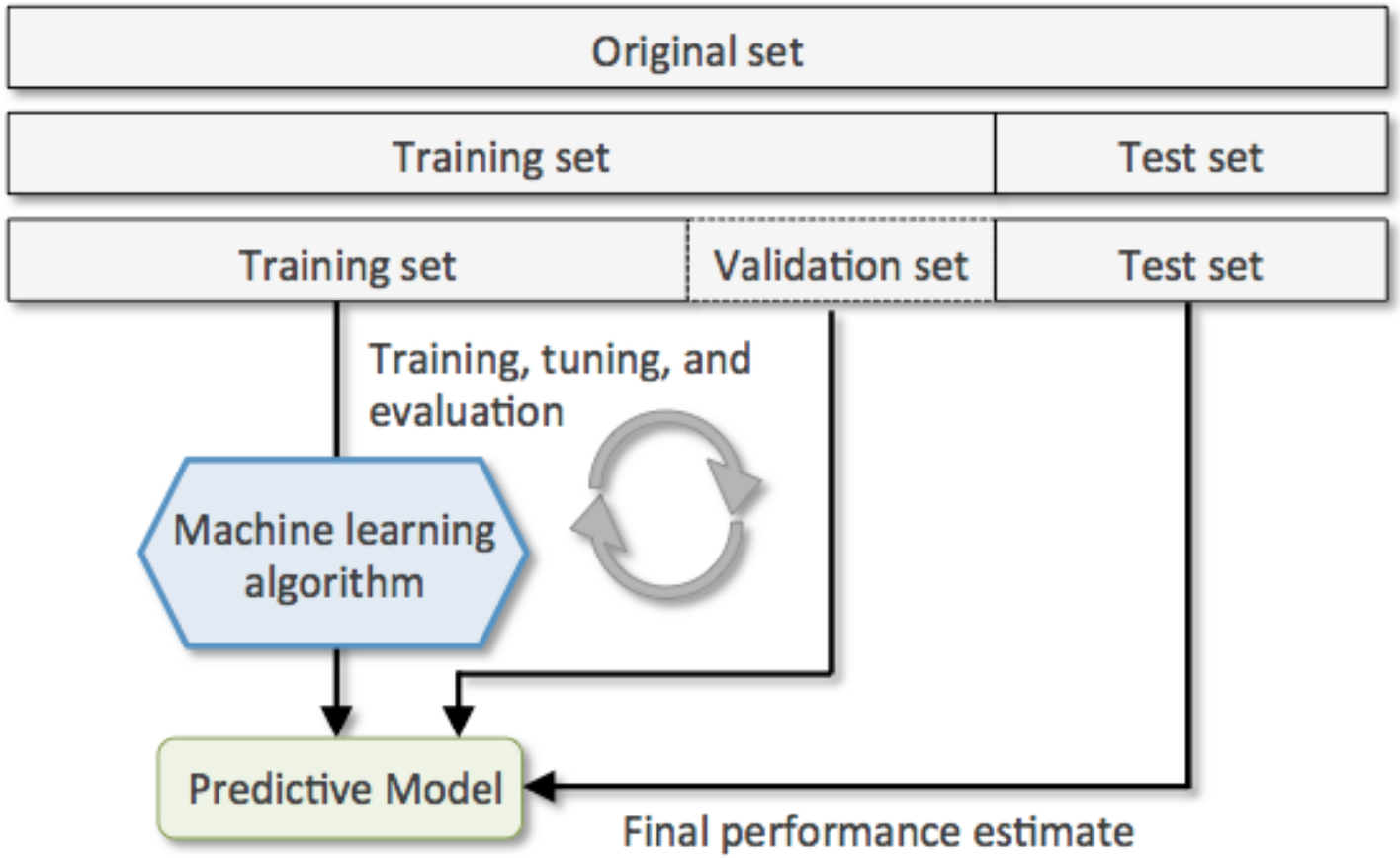
PREDICTING AS A TEAM RATHER THAN SOLO

1. SEVERAL SUBSETS OF FEATURES
2. SEVERAL SUBSETS OF SAMPLES
3. INDIVIDUAL CLASSIFIER FOR EACH SUBSETS
4. AGGREGATE PREDICTIONS

See labs exercise

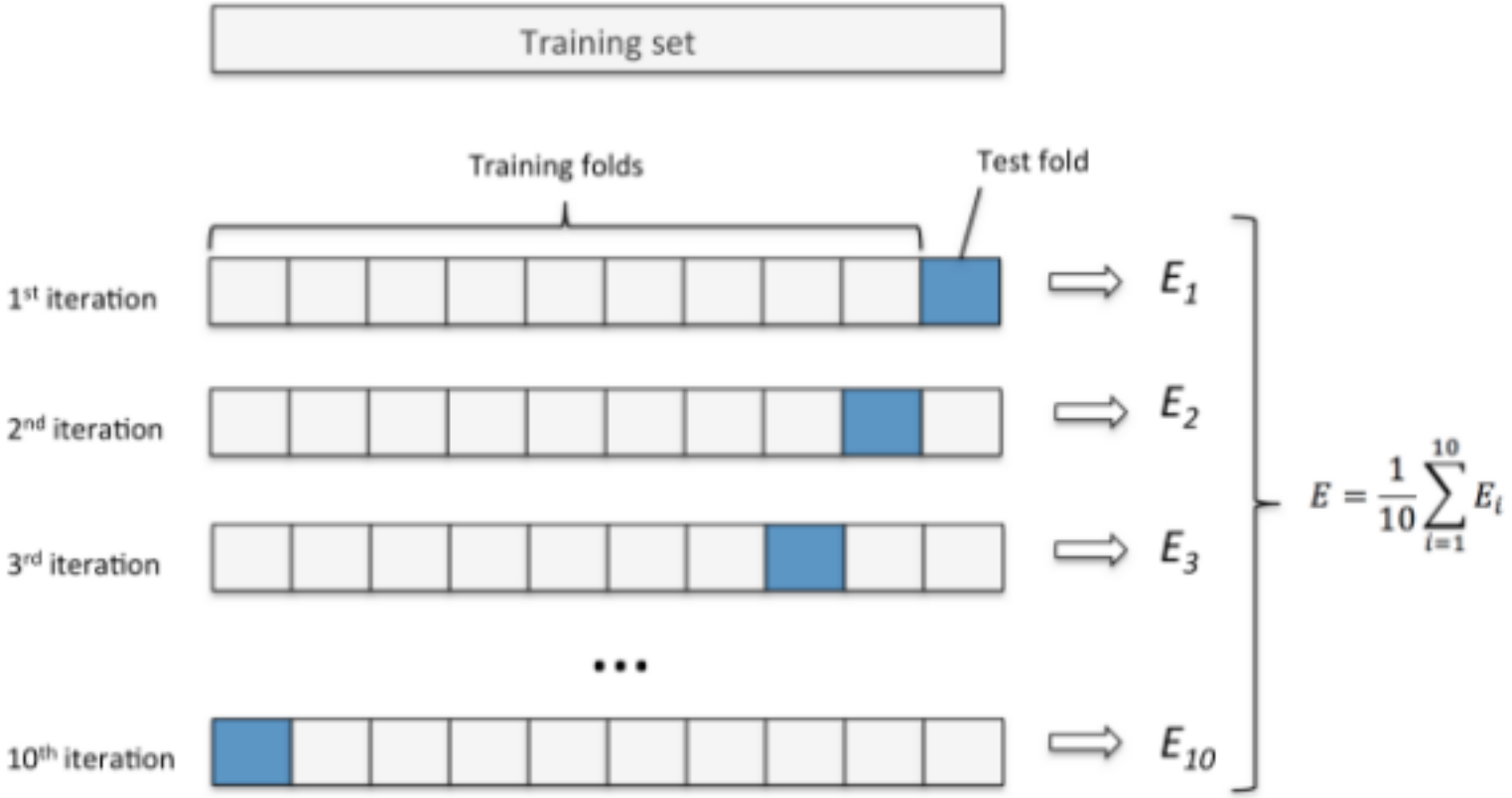
TUNING & EVALUATING MODELS

TYPICAL PIPELINE | HOLD OUT METHOD



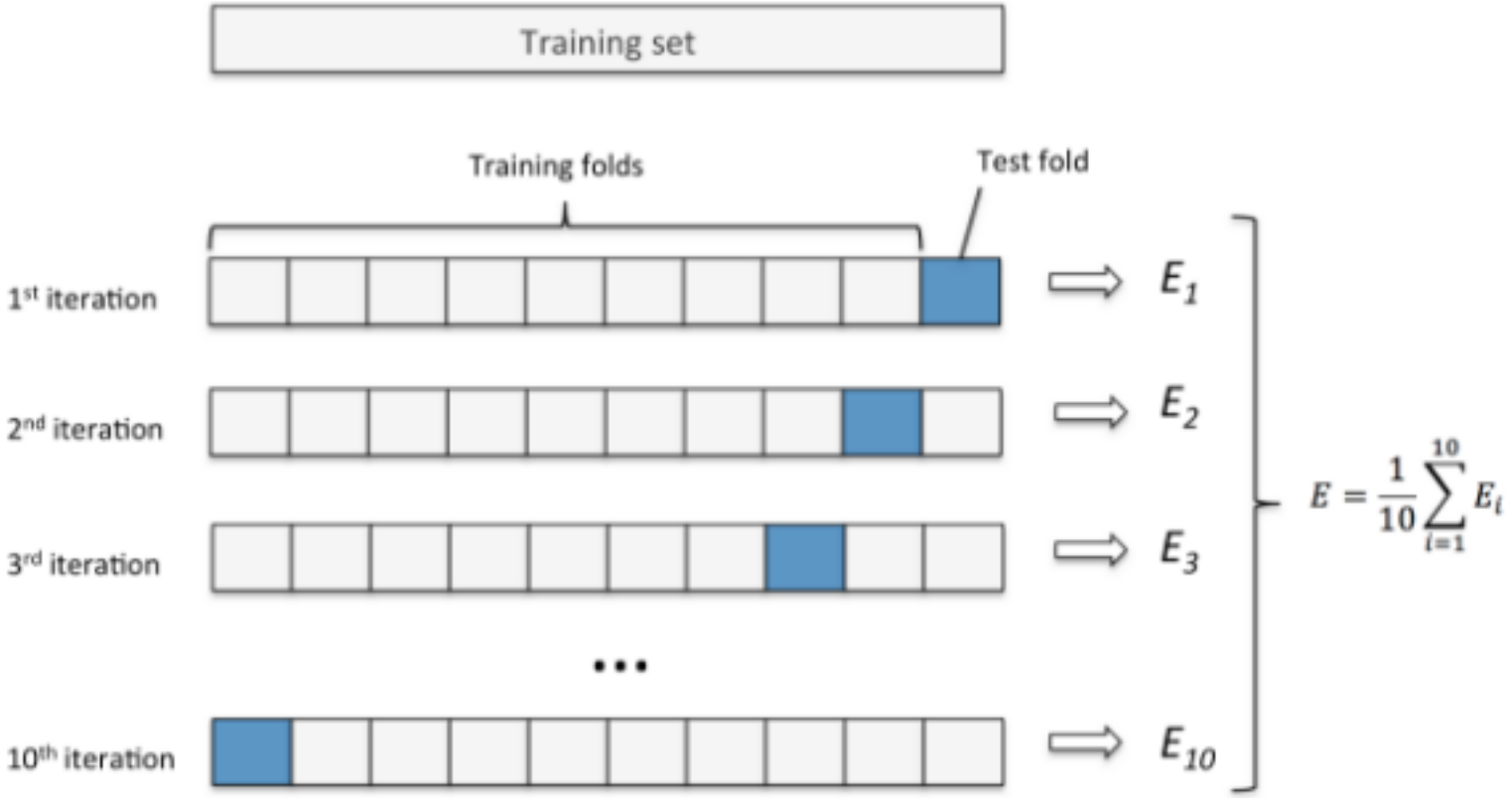
http://www.cs.nthu.edu.tw/

TYPICAL PIPELINE | K-FOLD CROSS VALIDATION



http://www.cs.nthu.edu.tw/

TYPICAL PIPELINE | K-FOLD CROSS VALIDATION



http://www.cs.nthu.edu.tw/

TYPICAL PIPELINE | GRID SEARCH

class `sklearn.ensemble.` **`RandomForestClassifier`** (*n_estimators=10, criterion='gini', max_depth=None, min_samples_split=2, min_samples_leaf=1, min_weight_fraction_leaf=0.0, max_features='auto', max_leaf_nodes=None, min_impurity_decrease=0.0, min_impurity_split=None, bootstrap=True, oob_score=False, n_jobs=1, random_state=None, verbose=0, warm_start=False, class_weight=None*) [source]

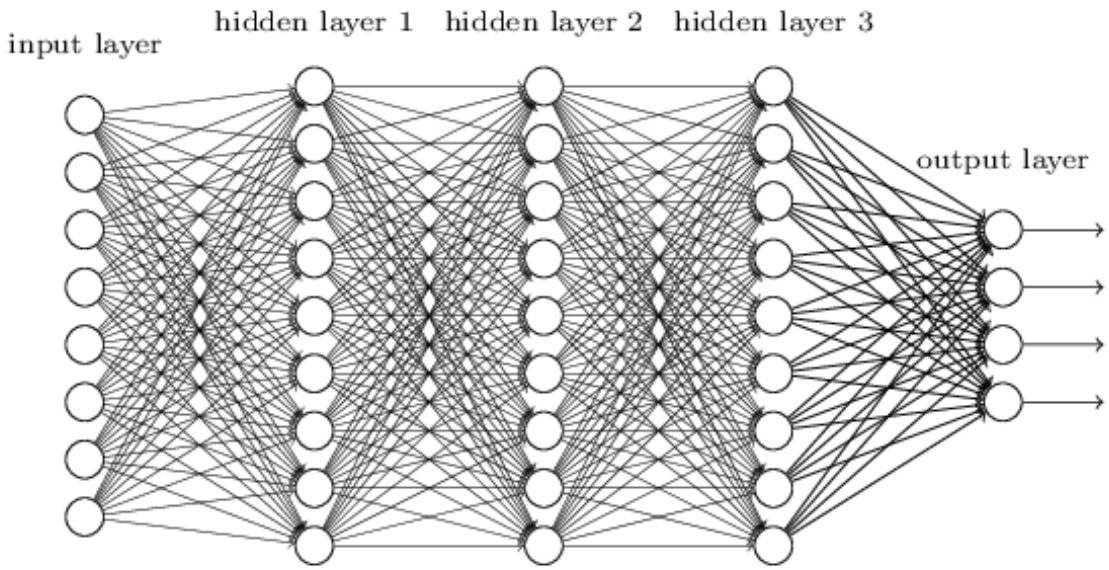
THE IMPORTANCE OF FEATURES

- MIGHT BE 1,000,000 INPUT FEATURES
- DIMENSIONALITY REDUCTION SOMETIMES RELEVANT
- FEATURE ENGINEERING IS SOMETIMES KEY (see lab exercise)
- THE CASE OF DEEP LEARNING AND COMPUTER VISION

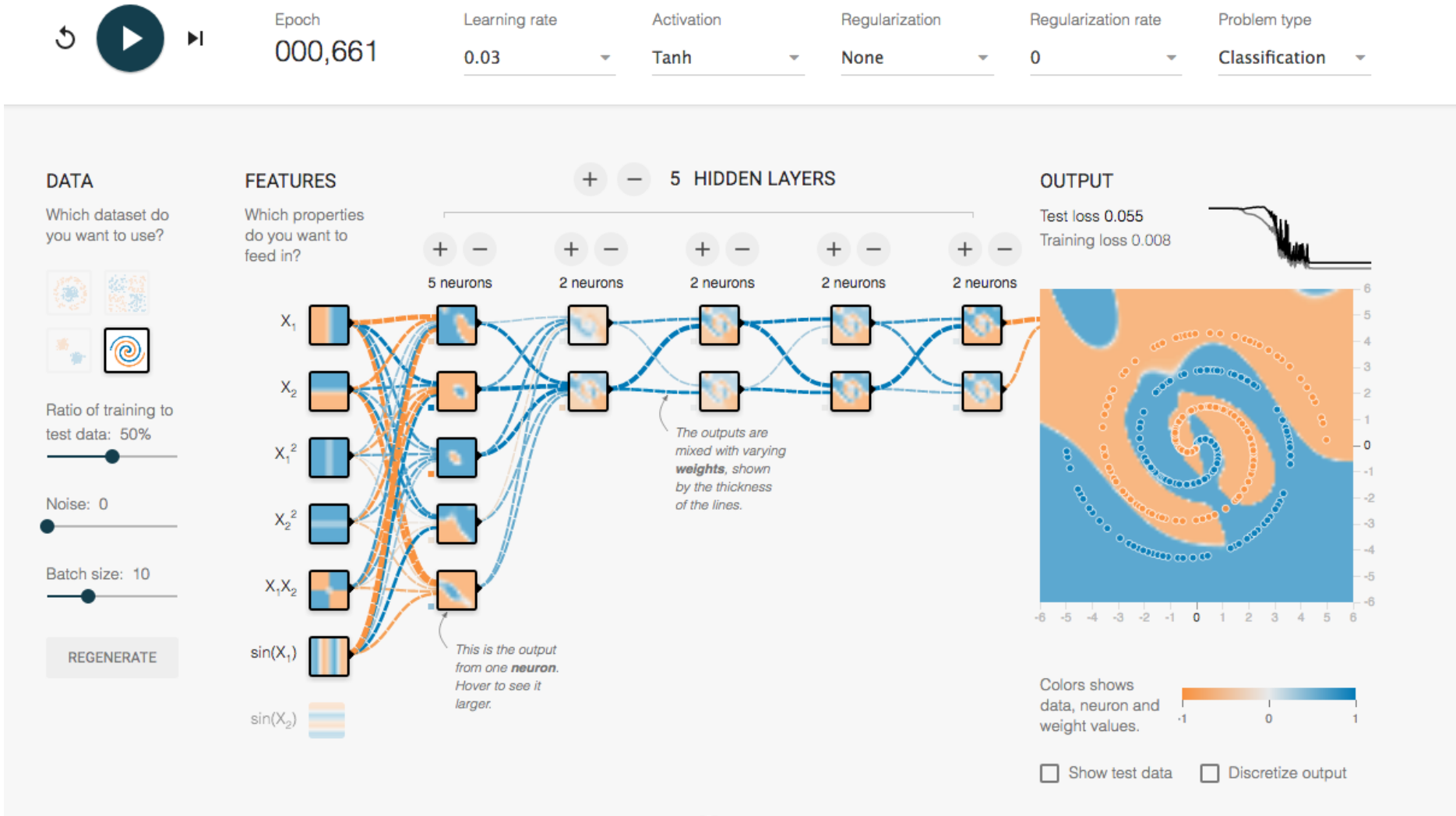
DEEP LEARNING & NEURAL NETWORKS

DEEP LEARNING & NEURAL NETWORKS

- A LOGISTIC REGRESSION IS ALREADY A NEURON (BUT NEED NON-LINEARITY)
- BY INCREASING # OF NODES/NEURON and CONNECTIVITY -> NON-LINEARITY



http://neuralnetworksanddeeplearning.com/chap6.html



http://playground.tensorflow.org/

DEEP LEARNING FOR IMAGE RECOGNITION

THE CASE OF "MNIST" HANDWRITTEN DIGIT DATASET RECOGNITION

TRAINING SET OF 60,000 EXAMPLES, TEST SET OF 10,000 EXAMPLES OF 28 x 28 IMAGES

CLASSIFIERS BENCHMARK:

http://yann.lecun.com/exdb/mnist/

CELEBRATING 25 YEARS OF ACHIEVEMENTS

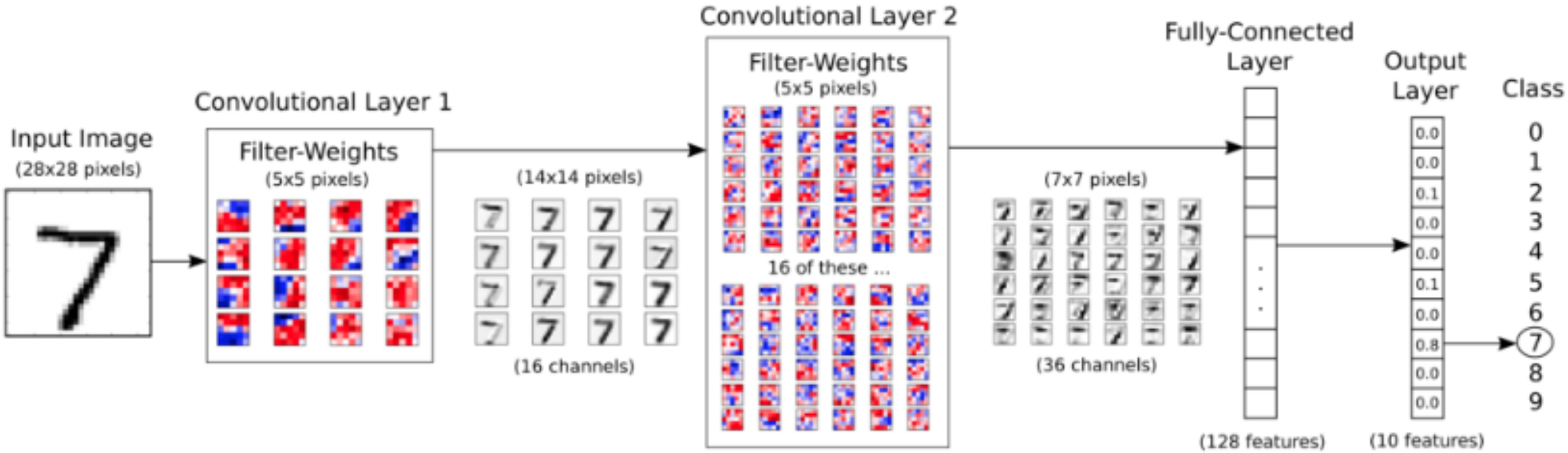
DEEP LEARNING FOR IMAGE RECOGNITION

CONVOLUTIONAL NETS MOTIVATION(S)

- 28 x 28 PIXELS IMAGE = 784 FEATURES, 1 BY PIXEL, VALUE FROM 0-255 (LEVEL OF GRAY)
- WHAT IF 1000 x 1000 x 3 (R,G,B) + 4 HIDDEN LAYERS + … ?
- WHAT IF IMAGE IS TRANSLATED, ROTATED, AT DIFFERENT SCALES ?
- NEED FOR HIGHER ABSTRACTION
- ONE OF THE GREATEST SUCCESS STORY OF BIOLOGICALLY INSPIRED AI
  https://en.wikipedia.org/wiki/David_H._Hubel

DEEP LEARNING FOR IMAGE RECOGNITION

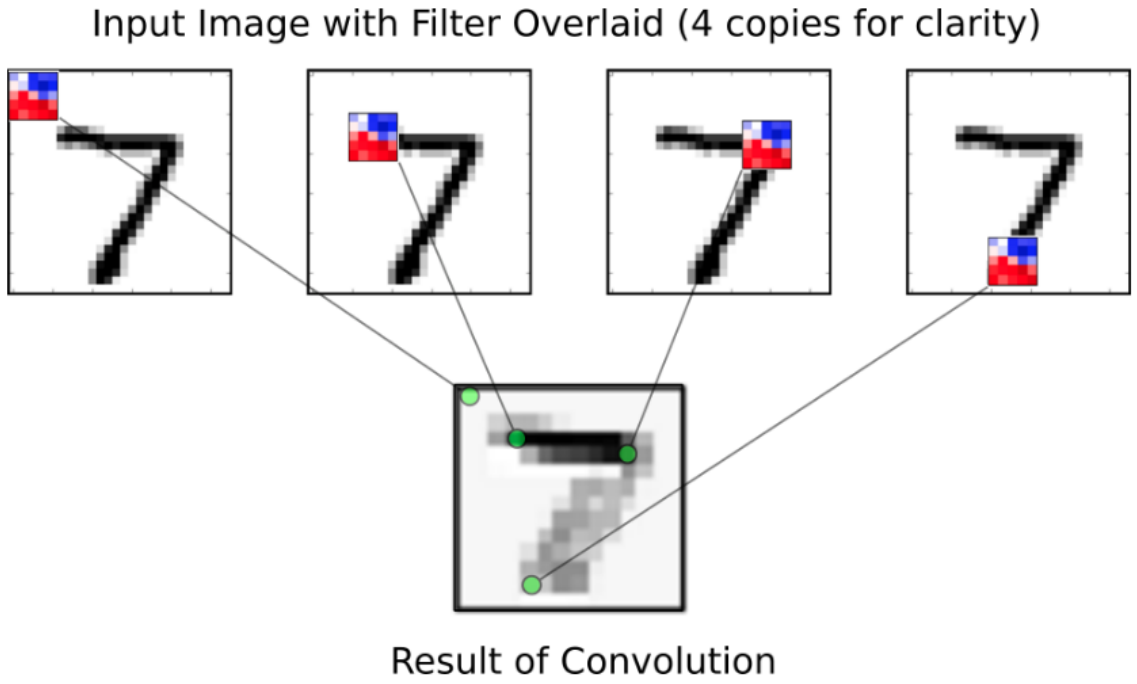
CONVOLUTIONAL NETS EXAMPLE [1]



https://github.com/Hvass-Labs/TensorFlow-Tutorials/blob/master/02_Convolutional_Neural_Network.ipynb

DEEP LEARNING FOR IMAGE RECOGNITION

CONVOLUTIONAL NETS EXAMPLE [2]



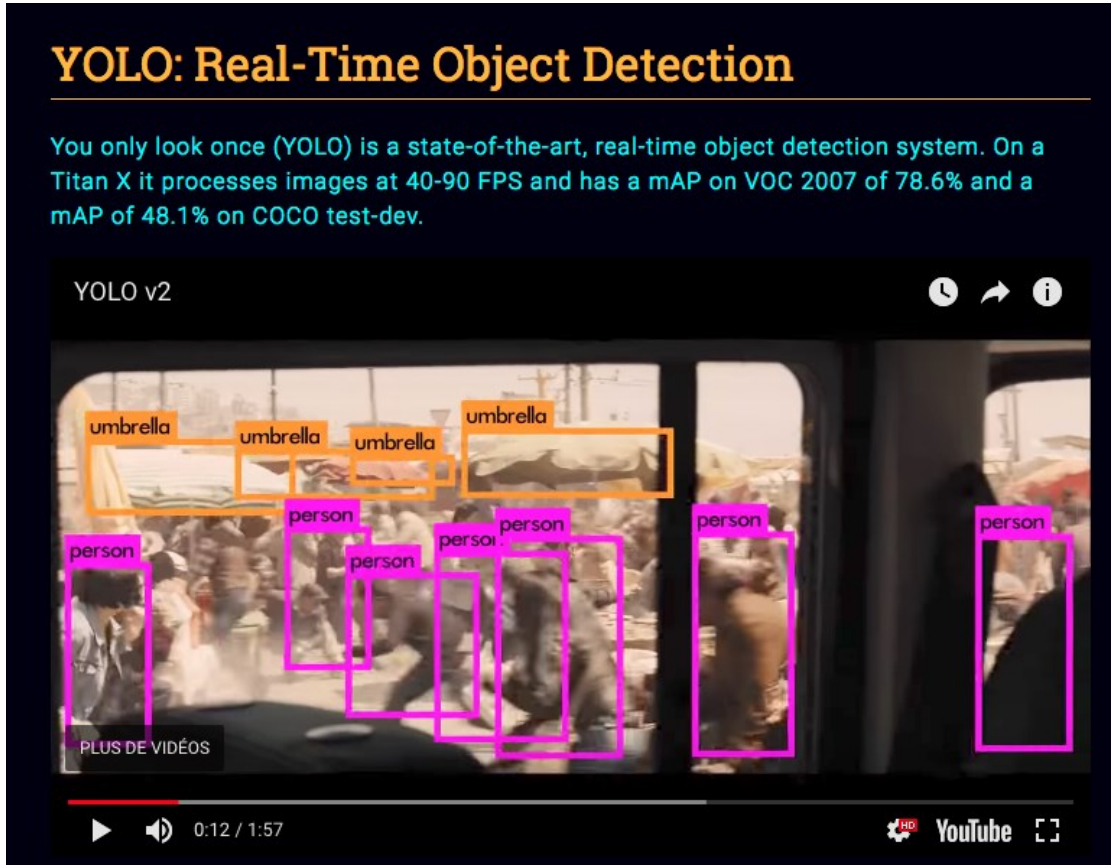
https://github.com/Hvass-Labs/TensorFlow-Tutorials/blob/master/02_Convolutional_Neural_Network.ipynb

DEEP LEARNING | TRANSFER LEARNING

RE-USE (USE WEIGHTS & TOPOLOGY) OF STATE OF THE ART MODELS

- LeNet-5
- AlexNet
- VGG
- ResNet
- Inception

STATE OF THE ART IN OBJECT DETECTION



https://pjreddie.com/darknet/yolo/

CELEBRATING 25 YEARS OF ACHIEVEMENTS

DEEP LEARNING | TRANSFER LEARNING

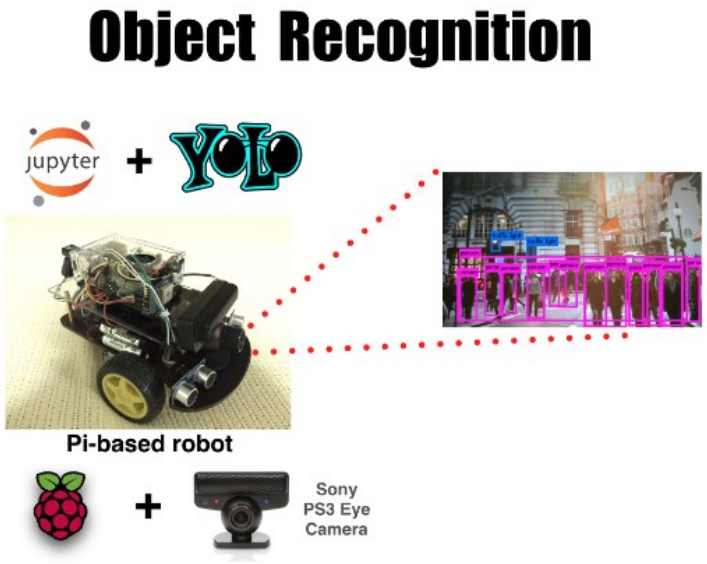
RE-USE (USE WEIGHTS & TOPOLOGY) OF STATE OF THE ART MODELS

- LeNet-5
- AlexNet
- VGG
- ResNet
- Inception

A WORD ON CPU vs GPU

- FORWARD/BACKWARD PROPAGATION CAN BE // AND VECTORIZED
- GAMER'S GPUs CAN BE FULLY HARNESSED
- BACKWARD PROPAGATION WITH MILLION OF PARAMETERS REQUIRES CONSIDERABLE RESOURCES
- FORWARD PROPAGATION IS FAST

BTW, WHAT THE RELATION WITH IOT ? [1]



https://github.com/leehaesung/YOLO-Powered_Robot_Vision

Deep learning on the Raspberry Pi with OpenCV

by Adrian Rosebrock on October 2, 2017 in Deep Learning, Machine Learning

https://www.pyimagesearch.com/2017/10/02/deep-learning-on-the-raspberry-pi-with-opencv/

BTW, WHAT THE RELATION WITH IOT ? [2]

- OTHER DOMAIN SUCH AS SEQUENTIAL MODELS ARE RELEVANT AS WELL (TIME SERIES PREDICTION, …)
- WILL NOT RUN ON A IoT NODE BUT SOMEWHERE IN YOUR TOPOLOGY …
- THOUGH SIMPLER MODEL SUCH AS LOGISTIC REG., DECISION TREE, … MIGHT BE EMBEDDED IN IoT DEVICES ?
- USE CASES ?

PYTHON LIBRARIES FOR ML & DL

- Scikit-learn: http://scikit-learn.org/
- TensorFlow: https://www.tensorflow.org/
- Theano: http://deeplearning.net/software/theano/
- Keras [wrapper]: https://keras.io/
- PyTorch: http://pytorch.org/
- …

LEARNING RESOURCES

- **Coursera/Stanford:** https://www.coursera.org/learn/machine-learning
- **Deeplearning.ai:** **https://www.deeplearning.ai/**
- **FastAI:** **http://www.fast.ai/**
- **UDACITY:** **https://www.udacity.com/courses/all**
- **…**

CELEBRATING 25 YEARS OF ACHIEVEMENTS

THANK YOU

CELEBRATING 25 YEARS OF ACHIEVEMENTS