ARTIFICIAL INTELLIGENCE

A BIRD'S EYE VIEW

RAMESH KRISHNAN BALAKRISHNAN

18TH MAY 2017

ARTIFICIAL INTELLIGENCE

Knowledge Representation
- to interact with people in real world, computers need to understand what are objects, people, language

Pattern Matching
- Compare what it sees with a pattern.

Reasoning
- Play chess, solve algebra problems, diagnose disease

Generalized Intelligence (including Emotional, Creativity, Moral Reasoning, Intuition etc..)

Natural language processing
- speak, understand, create sentences, understand context.

Perception
- How to see things, hear, feel things in world to get to human intelligence and even to smell.

Planning (Including Navigation)
- go from place A to B, where are doors, what are paths, where its safe to go, optimal route/path

Artificial Intelligence is the broader concept of machines being able to carry out tasks in a way that we would consider "smart".
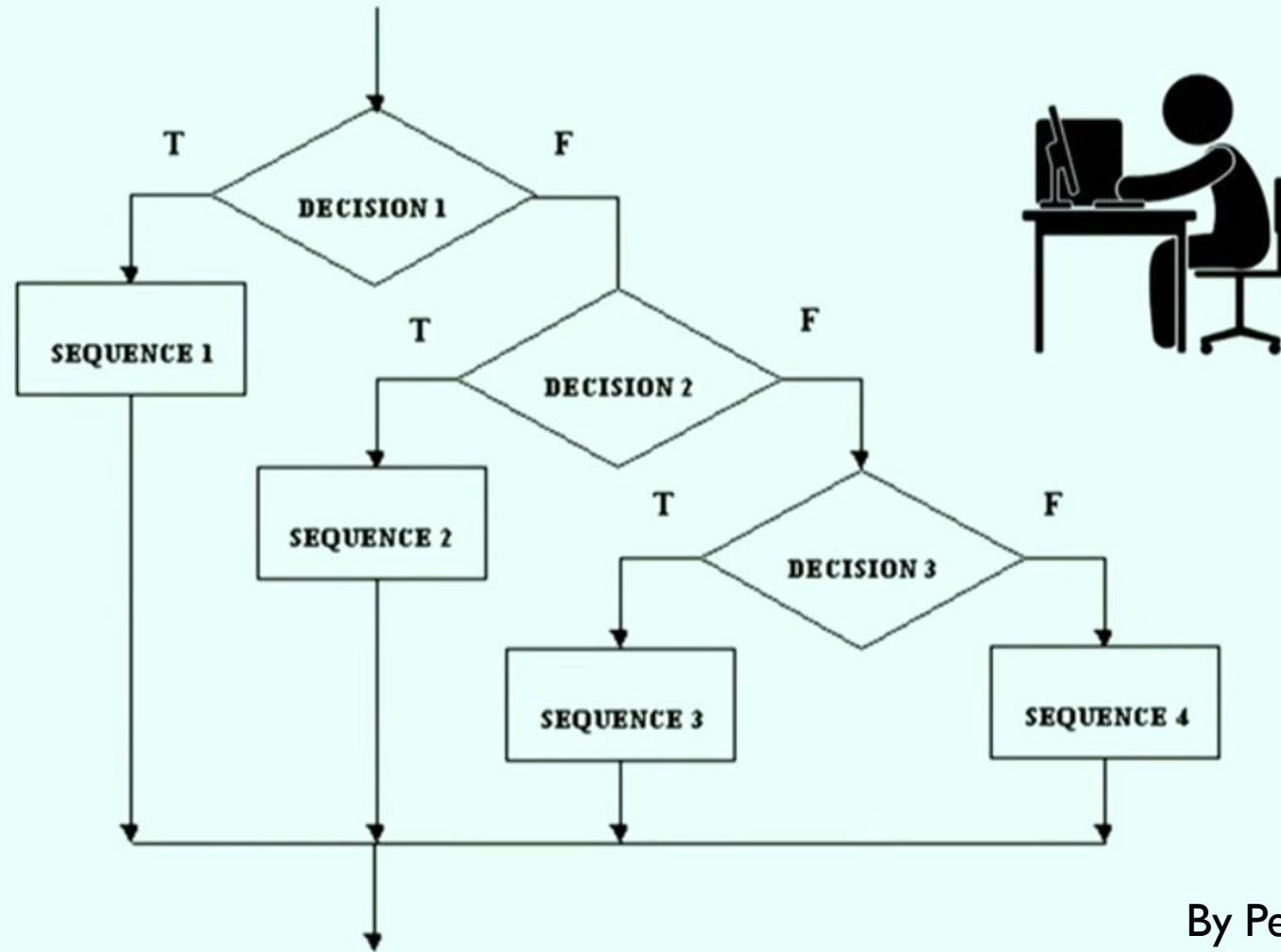
Machine Learning is a current application of AI based around the idea that we should really just be able to give machines access to data and let them learn for themselves.

Why Machine Learning Explosion now?
- More data available
- Better processing power (CPU/GPUs)
- Improvement in techniques

http://a16z.com/2016/06/10/ai-deep-learning-machines/

PROGRAMMING AS WE KNOW

By Peter Norving in EmTech 2016

MACHINE LEARNING

- Probabilistic logic than Boolean

- Program by example – from data

- Can't put a break point and debug

- New concepts - Overfitting/Underfitting, Normalization, Accuracy

- Involves Statistics, Probability, Linear Algebra, Calculus

Broadly classified into:

- Supervised Learning
    - Regression – Continuous and Numeric
   - Classification - Categorical
- Unsupervised Learning
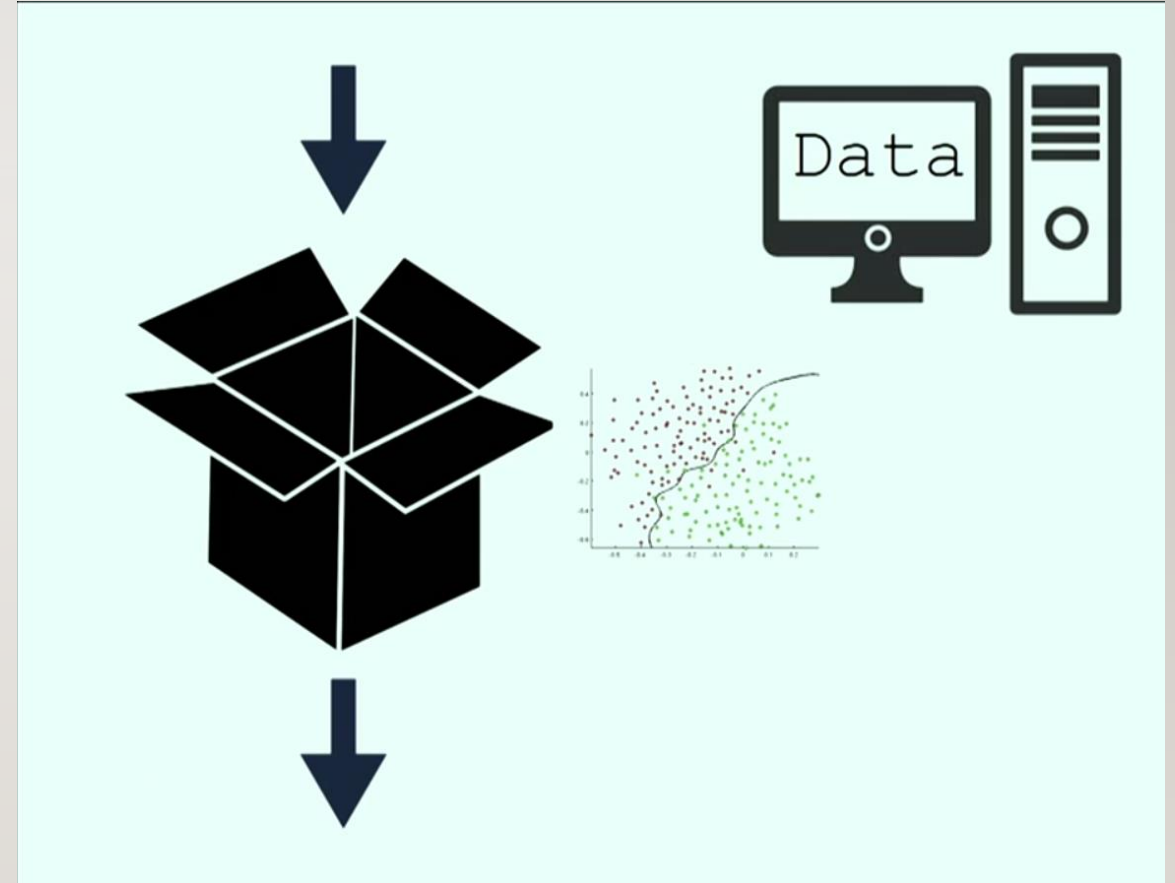
By Peter Norving in EmTech 2016

ML: DEFINITIONS AND QUOTES

- Arthur Samuel: Machine learning: Field of study that gives computers the ability to learn without being explicitly programmed.

- Tom Mitchell – Computer program is said to learn from Experience E with respect to some task T and some performance measure P, if its performance on T, as measure by P, improves with experience E.

- Jeff Bezos - Over the past decades computers have broadly automated tasks that programmers could describe with clear rules and algorithms. Modern machine learning techniques now allow us to do the same for tasks where describing the precise rules is much harder.

- AI is new electricity – Andrew Ng

- Google was Search first, mobile first later, Now Sundar Pichai says – Google is AI first

MACHINE LEARNING METHODS

scikit-learn algorithm cheat-sheet
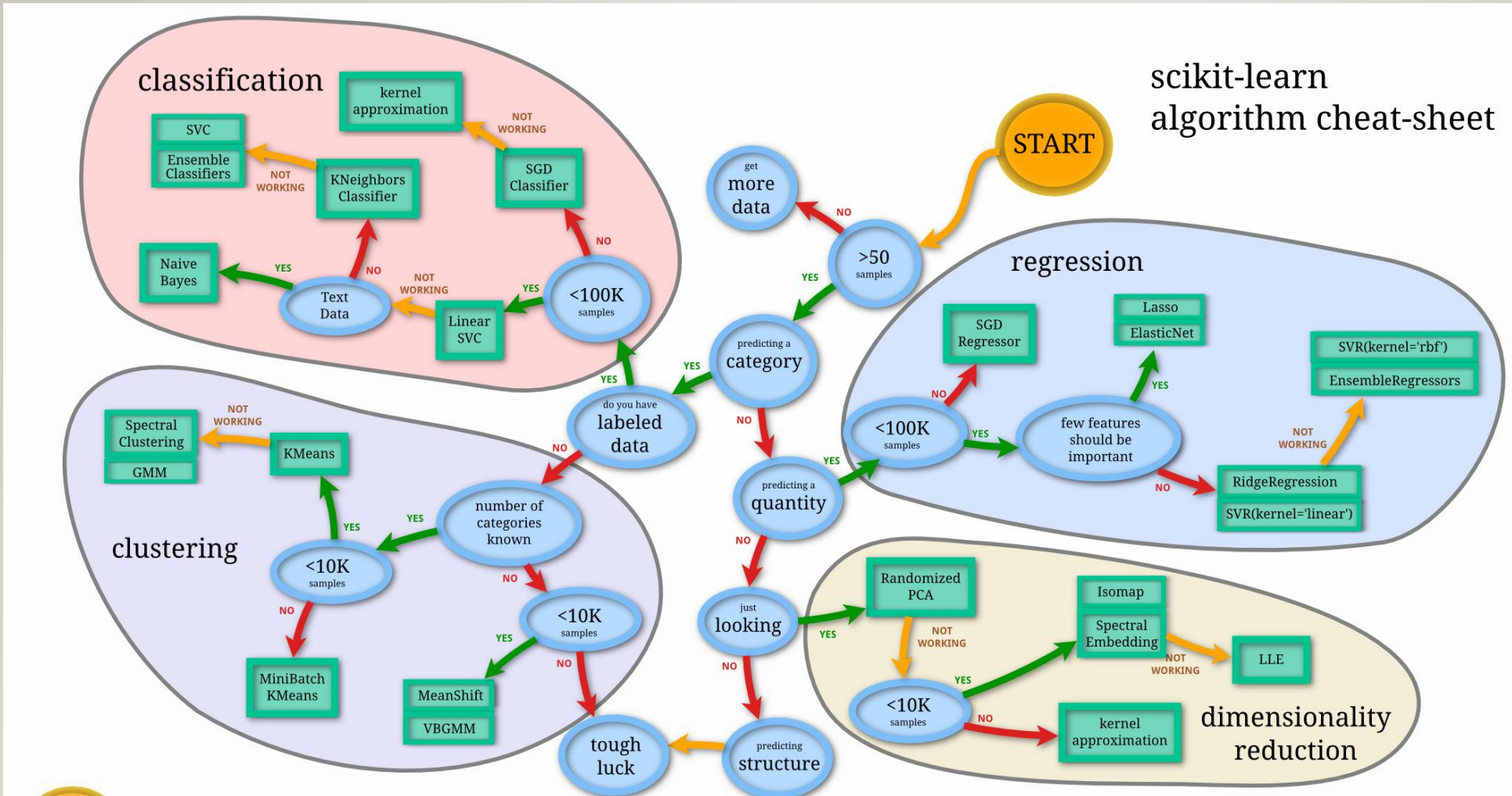
Languages & Tools:
- R
- Python – Anaconda
- Matlab
- Octave

Platforms:
- Amazon machine
- Learning
- Microsoft Azure ML
- IBM Data Science
- H2o.ai

http://scikit-learn.org/stable/tutorial/machine_learning_map/

EXAMPLES

- Regression:
  - House rent prediction

- Classification:
  - Breast cancer – malignant or not
  - Classifying emails as spam or not

- Netflix Recommendations

- Unsupervised Learning:
  - Google News, Market Segmentation

- NLP
  - Google Translate
  - Chatbots – NLP

- Computer Vision
  - Captioning of Images – Computer Vision + NLP (Not HotDog App ☺)

Amazon uses ML for:
Demand forecasting, product and deals recommendations, Merchandising placements, fraud detection, translations, and much more.

DEMO

- Linear Regression using Stochastic Gradient Descent

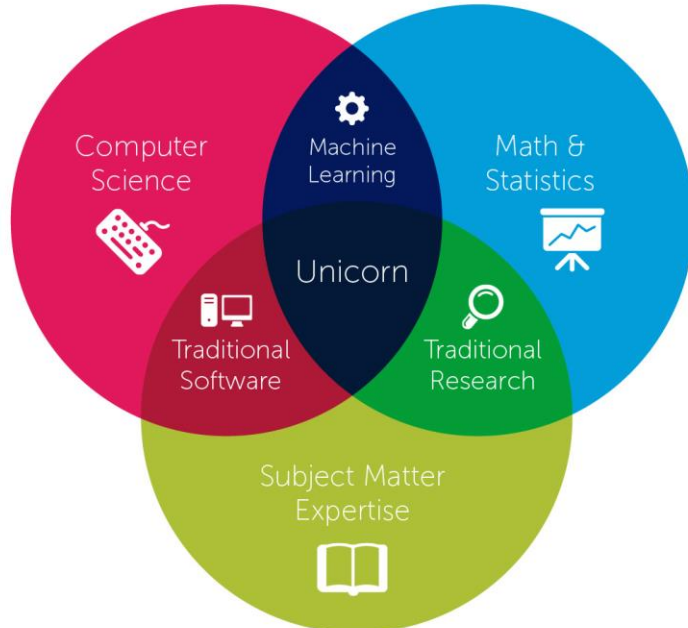- Recommender Systems using Collaborative filtering

EVERYTHING is a Recommendation
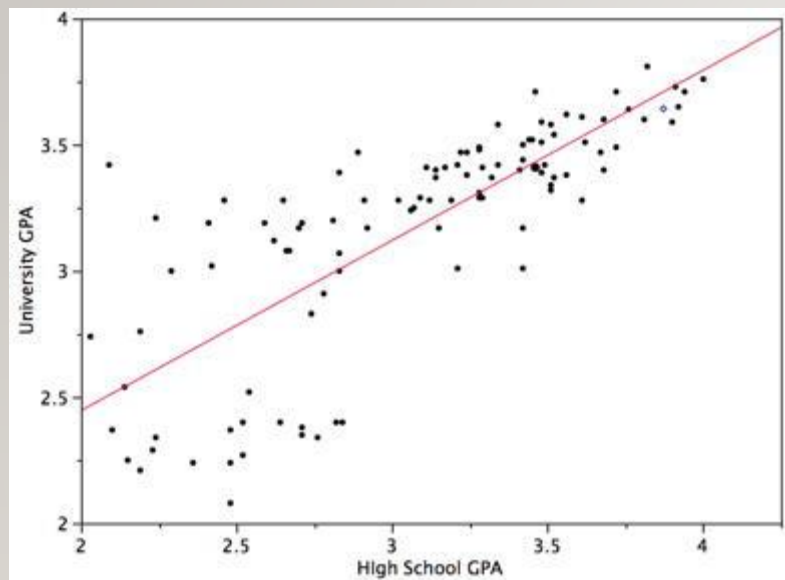
Images from Google Images search
Excels used for demonstration of SGD(graddesc.xlsm) and Collaborative filtering (collab_filter.xlsx) is from below github link.
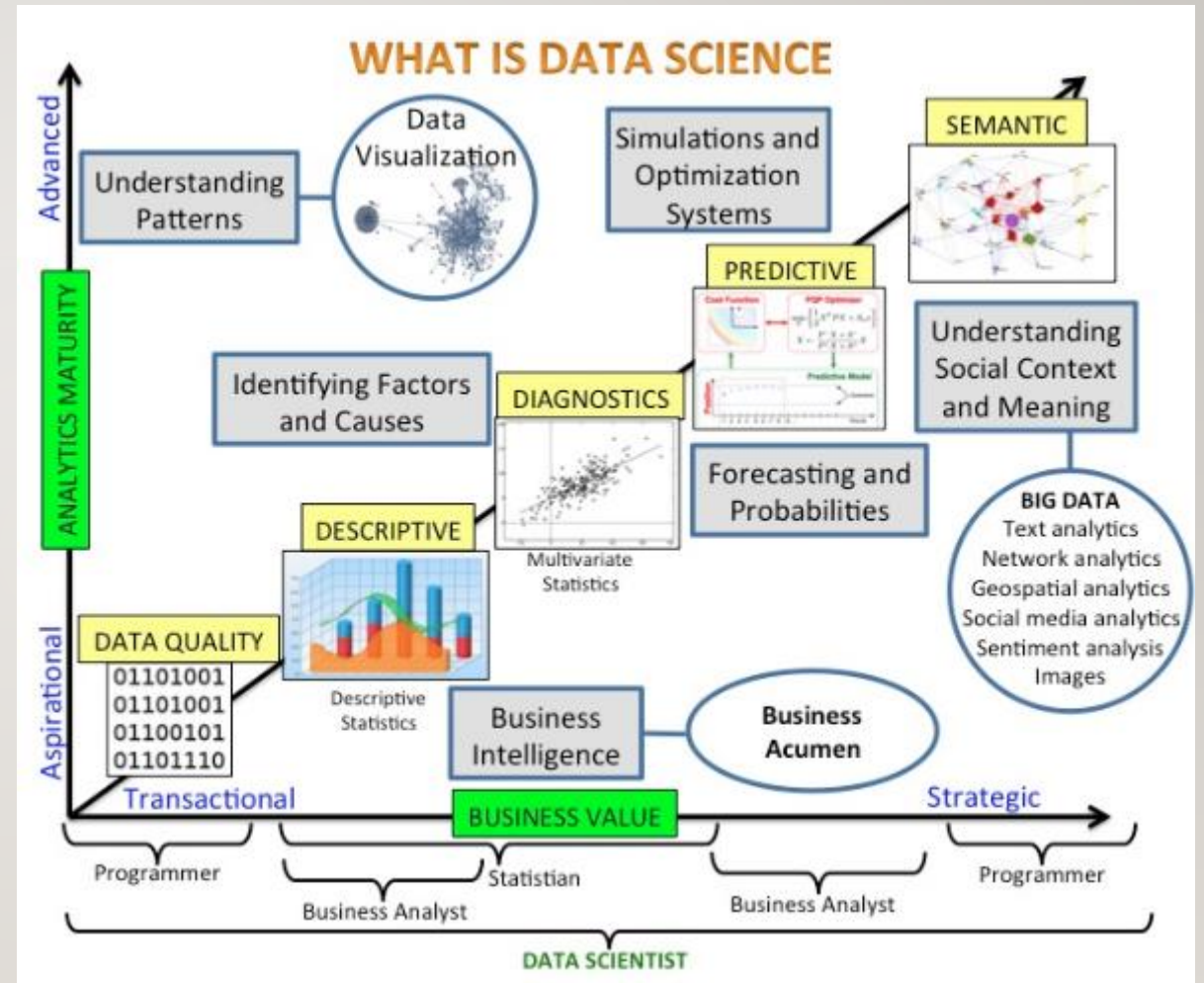 Part of fast.ai course https://github.com/fastai/courses/tree/master/deeplearning1/excel

DATA SCIENCE

Demo: Titanic: ML for Disaster from Kaggle

http://www.kdnuggets.com/2016/03/data-science-process.html
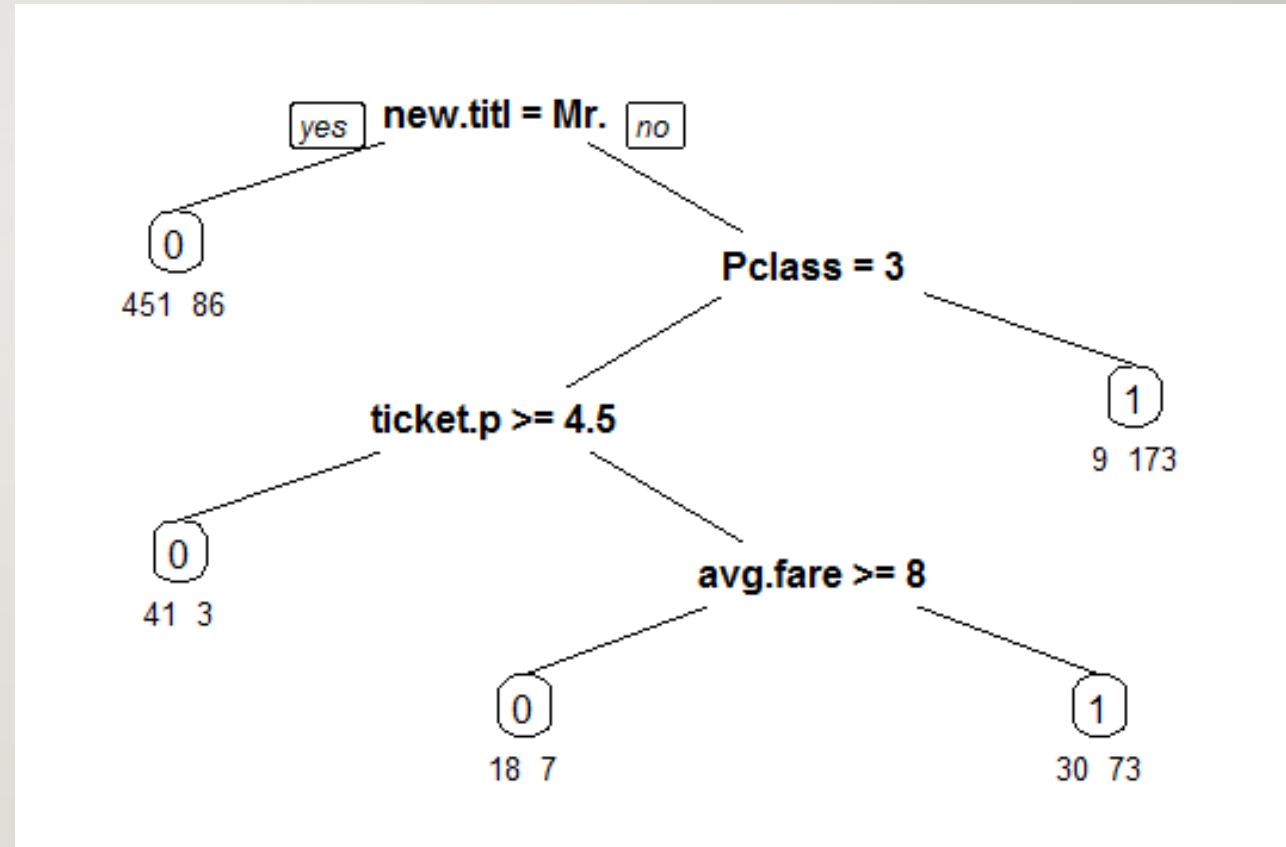
DEMO: TITANIC: ML FOR DISASTER FROM KAGGLE

- Training Set

- Test set

- Cross-Validation set - Overfitting

- Data Cleanup/Understanding

- Feature Identification/Engineering

- Modelling:
  - Decision Tree
  - Random Forest

- Accuracy

https://www.kaggle.com/c/titanic
https://www.youtube.com/watch?v=32o0DnuRjfg

CONFLICTING VIEWS ON AI

Boom and Bust cycles – So busts were called AI Winters where funding got exhausted.

Emtech 2016 talk by Oren Etzioni, AI for the Common Good

ALPHA GO

DEEP LEARNING

- GPUs eating Linear Algebra. LA eating Deep Learning.

  - DL eating Machine Learning. ML eating AI.

    - AI eating software. Software eating the world.

- Deep learning is a subfield of machine learning concerned with algorithms inspired by the structure and function of the brain called artificial neural networks."— Jason Brownlee

- Popular Neural network techniques:

  - CNNs –Convoluted Neural networks – Used in Image Classification (Dog vs Cat, Image captioning

- RNNs - Recurrent Neural Networks. Used in NLP.

Swift keyboard in iOS/Android) uses this for predicting

 next character, word)

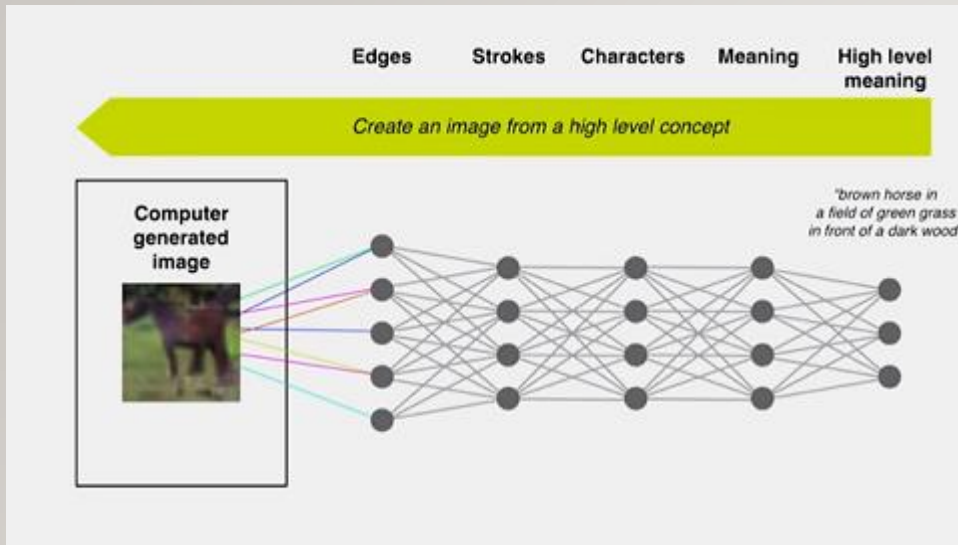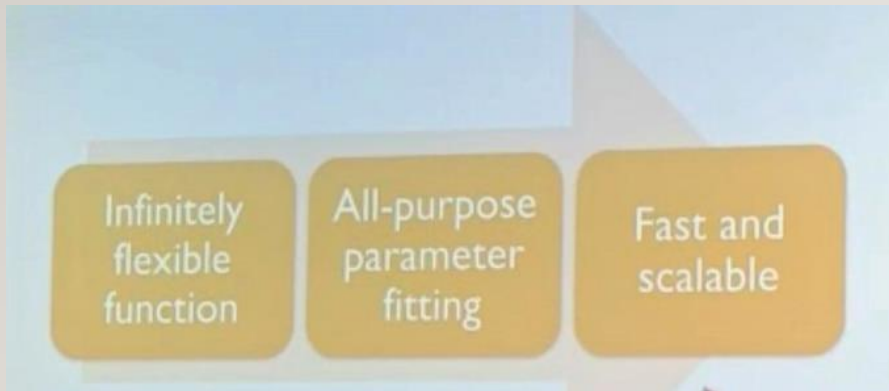THIS IS YOUR MACHINE LEARNING SYSTEM?
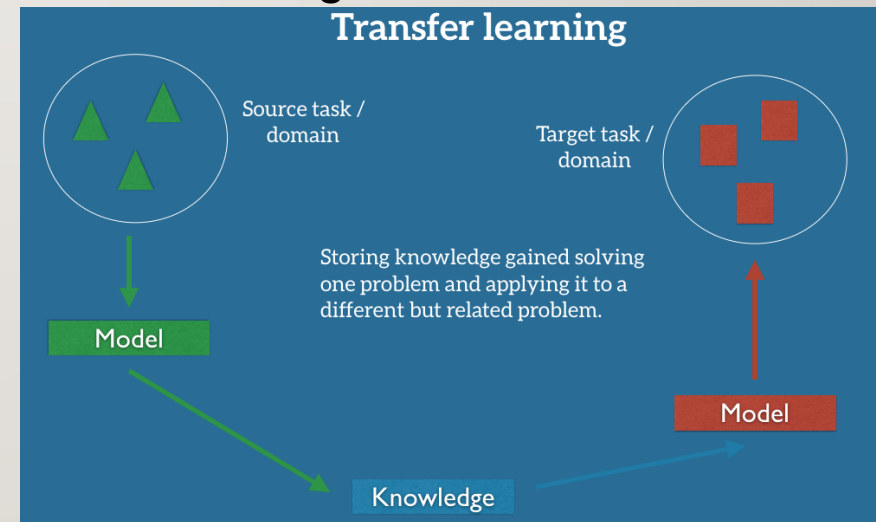
YUP! YOU POUR THE DATA INTO THIS BIG PILE OF LINEAR ALGEBRA, THEN COLLECT THE ANSWERS ON THE OTHER SIDE.

WHAT IF THE ANSWERS ARE WRONG?

JUST STIR THE PILE UNTIL THEY START LOOKING RIGHT.

https://xkcd.com/1838/

DEEP LEARNING - EXAMPLES

Edges   Strokes   Characters   Meaning   High level meaning

Create an image from a high level concept

Computer generated image

"brown horse in a field of green grass in front of a dark wood"

Quick Demo - layers_example .xlsx

Why it works?

Infinitely flexible function   All-purpose parameter fitting   Fast and scalable

Transfer Learning:

Transfer learning

Source task / domain

Target task / domain

Storing knowledge gained solving one problem and applying it to a different but related problem.

Model

Model

Knowledge

From deep learning lecture by Jeremy Howard - https://github.com/fastai/courses/tree/master/deeplearning1/excel

DEEP LEARNING FRAMEWORKS

- **TensorFlow from Google:** is an open source software library released in 2015 by Google to make it easier for developers to design, build, and train deep learning models. ... At a high level, **TensorFlow** is a Python library that allows users to express arbitrary computation as a graph of data flows.

- **Pytorch:** A Python API for Torch, known as Pytorch, was open-sourced by **Facebook** in January 2017. PyTorch offers dynamic computation graphs, which let you process variable-length inputs and outputs, which is useful when working with RNNs

- **Caffe2:** is the second deep-learning framework to be backed by **Facebook** after Torch/PyTorch. The main difference seems to be the claim that Caffe2 is more scalable and light-weight. Like **Caffe** and PyTorch, Caffe2 offers a Python API running on a C++ engine.

- **CNTK:** is **Microsoft's** open-source deep-learning framework. The acronym stands for "Computational Network Toolkit." The library includes feed-forward DNNs, convolutional nets and recurrent networks. CNTK offers a Python API over C++ code.

- **Paddle:** is a deep-learning framework created and supported by **Baidu**. Its name stands for PArallel Distributed Deep Learning. It offers a Python API.

- **Keras:** s a deep-learning library that sits atop Theano and TensorFlow, providing an intuitive API inspired by Torch. Perhaps the best Python API in existence.

WHERE TO LEARN?

- Follow pioneers, popular developers – LinkedIn, Twitter, Medium

- Online MOOCs - Udacity/edX/Coursera/ Udemy

    - https://medium.freecodecamp.com/every-single-machine-learning-course-on-the-internet-ranked-by-your-reviews-3c4a7b8026c0

    - https://www.coursera.org/learn/machine-learning -- Andrew Ng

- https://www.kaggle.com/ - for datasets, competitions

- Deep Learning - http://course.fast.ai/

THANK YOU