2018 Asia University AI Summer Program (Day1-July3) B: Introduction to AI

Frank Hsueh-Ting Chu

Content

- What is Intelligence
- What is Artificial intelligence
- Machine learning
- Deep learning
- AI applications

(1)WHAT IS INTELLIGENCE


Which animals are smart?

10 Most Intelligent Animals In The World

1. Chimpanzees

2. Elephants

Dolphins

Parrots

Sheep

Rats

Dogs

Octopus

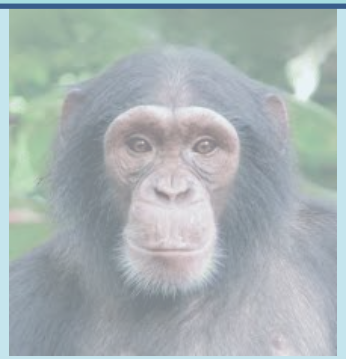
Crows

Pigeons

https://10mosttoday.com/10-most-intelligent-animals-in-the-world/

Animal Intelligence

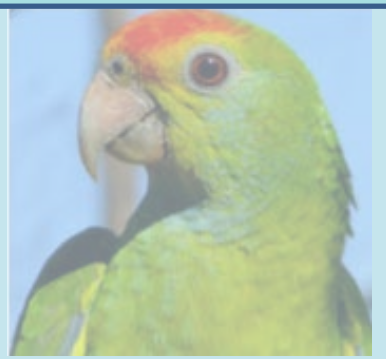
1. Chimpanzees



2. Elephants

Dolphins

Parrots



Sheep

Why we think they are smart?

Rats

Dogs

Octopus

Crows

Pigeons

https://10mosttoday.com/10-most-intelligent-animals-in-the-world/

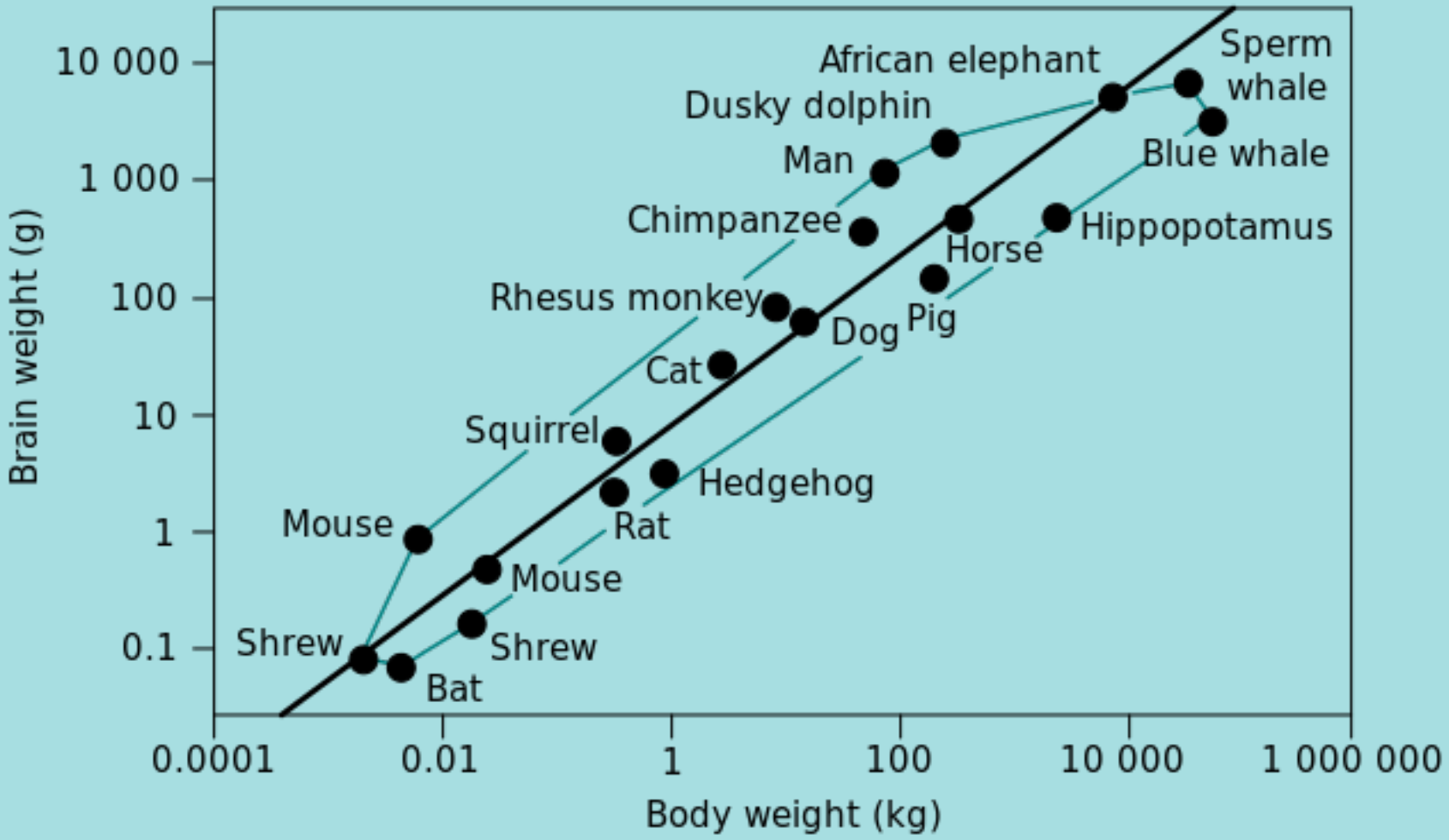
The crow and the pitcher in Aesop's Fables

In a spell of dry weather, when the Birds could find very little to drink, a thirsty Crow found a pitcher with a little water in it. But the pitcher was high and had a narrow neck, and no matter how he tried, the Crow could not reach the water. The poor thing felt as if he must die of thirst.
Then an idea came to him. Picking up some small pebbles, he dropped them into the pitcher one by one. With each pebble the water rose a little higher until at last it was near enough so he could drink.

Metrics of intelligence

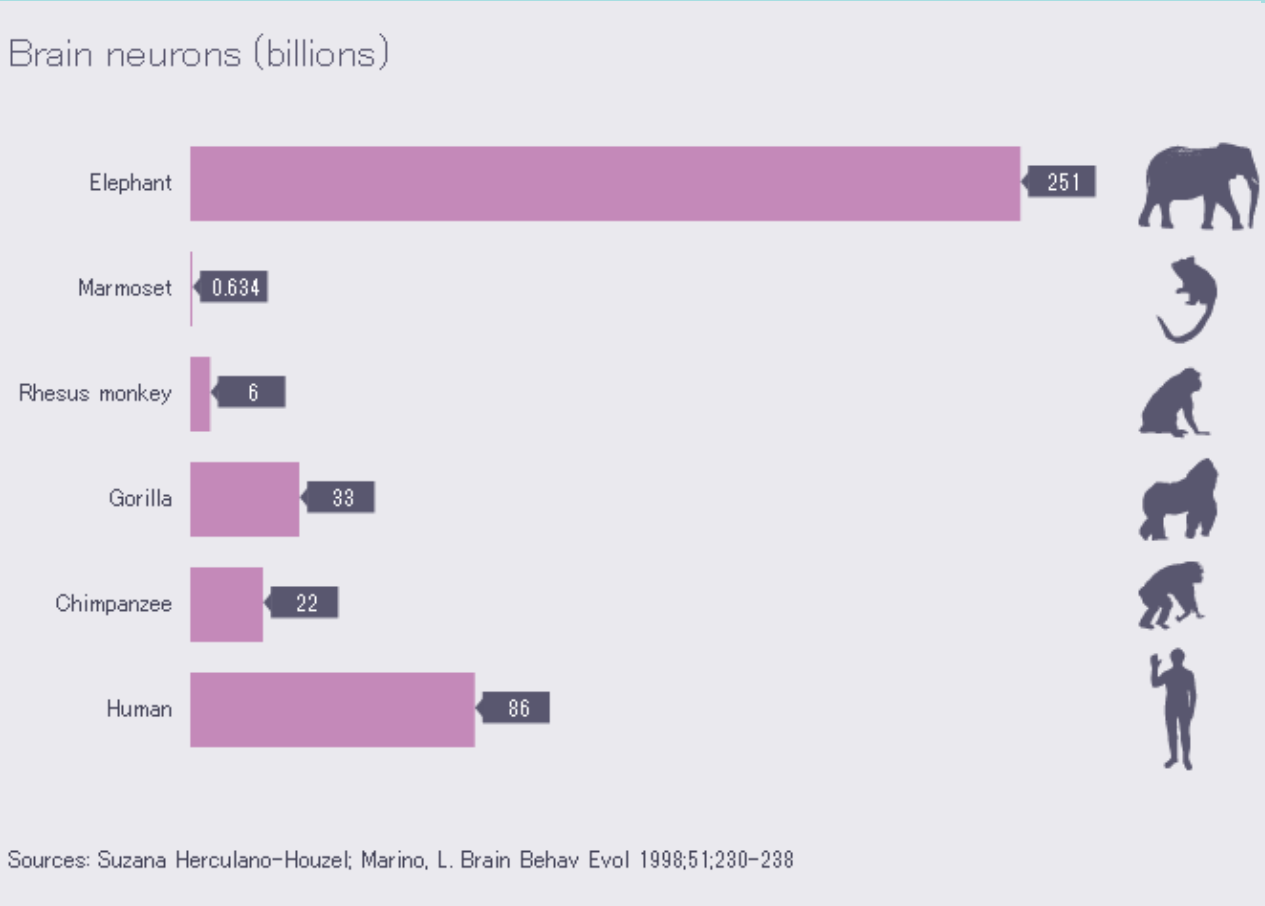
- Brain-to-body mass ratio
- Number of neurons
- Encephalization quotient

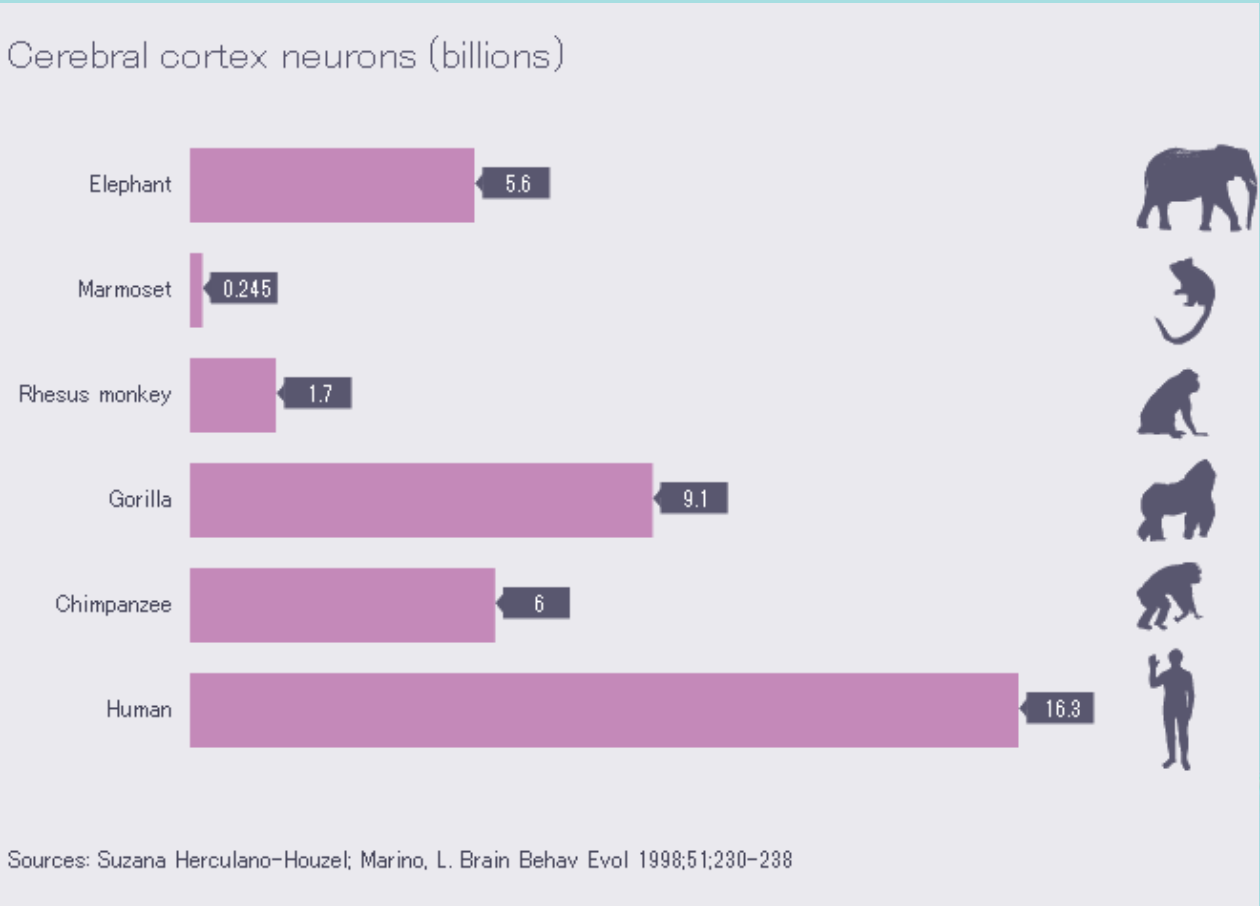
Brain-to-body mass ratio



https://en.wikipedia.org/wiki/Brain-to-body_mass_ratio

Number of neurons

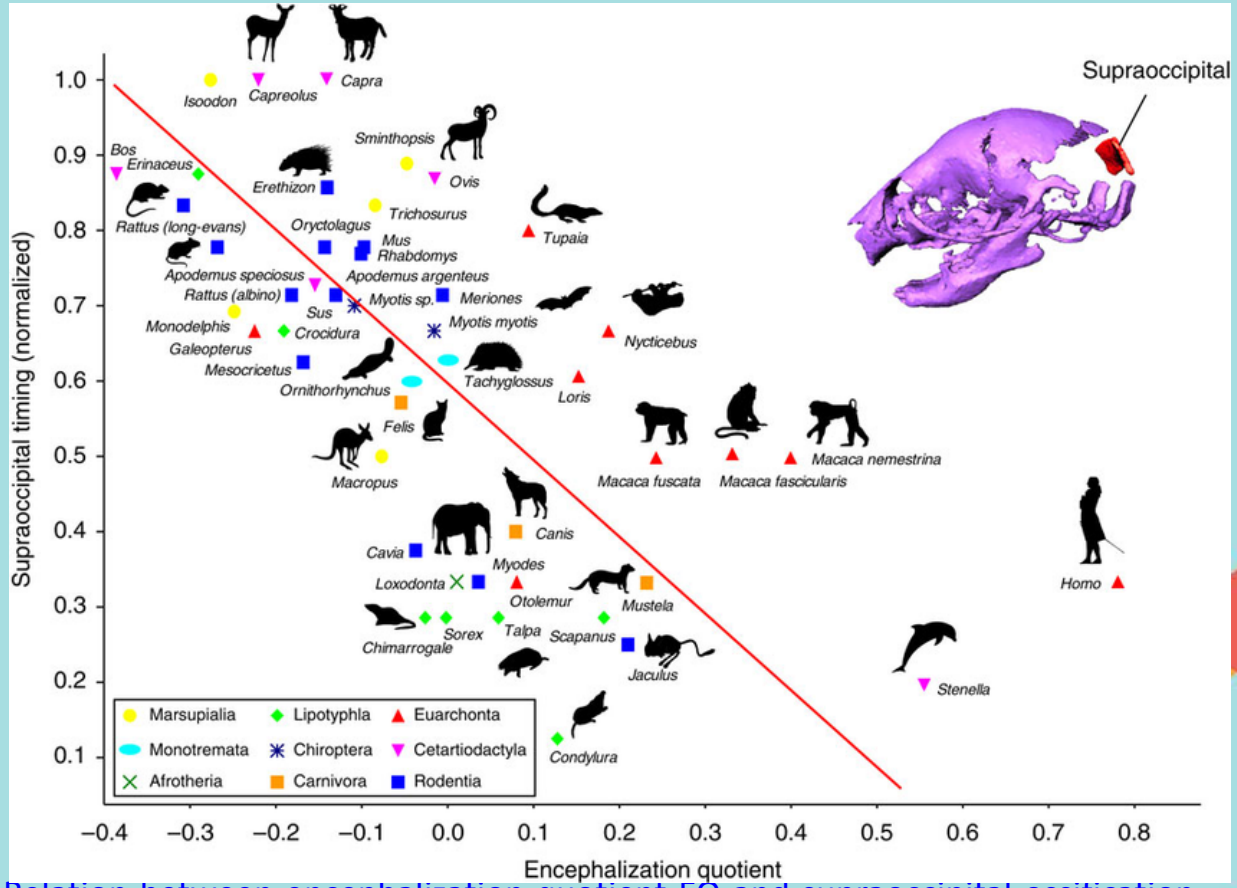




https://en.wikipedia.org/wiki/List_of_animals_by_number_of_neurons

Encephalization quotient

- Encephalization quotient (EQ) is a relative brain size measure that is defined as the ratio between actual brain mass and predicted brain mass for an animal of a given size, which may approximate intelligence level or cognition of the species.



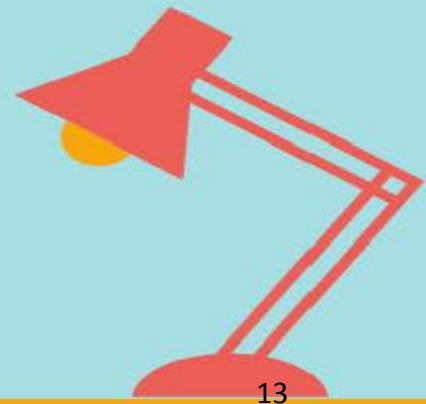
https://en.wikipedia.org/wiki/Encephalization_quotient
https://www.researchgate.net/figure/261407553_fig6_Figure-3-Relation-between-encephalization-quotient-EQ-and-supraoccipital-ossification

(2) INTRODUCTION TO AI

Concept of Artificial intelligence

- Artificial intelligence (AI) is intelligence demonstrated by machines

- Natural intelligence (NI) displayed by humans and other animals.



Intelligence

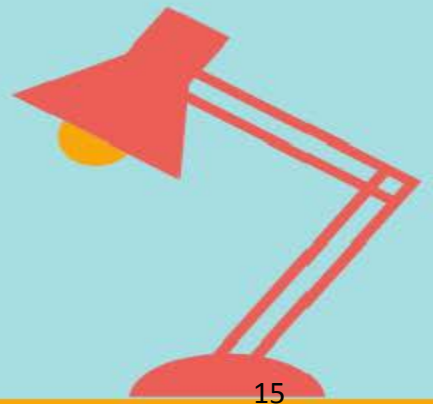
- Intelligence consists of the following skills:
- 1. the ability to reason
- 2. the ability to acquire and apply knowledge
- 3. the ability to manipulate and communicate ideas

http://slideplayer.com/slide/272574/

History of artificial intelligence

https://en.wikipedia.org/wiki/History_of_artificial_intelligence

- The First Boom of AI :
  - 1952–1956
  - Failure of machine translation
- The Second Wave of AI
  - 1980–1987
  - In 1981, Japanese began the Fifth generation computer project.
  - Failure of expert systems
- The Third Wave of AI
  - 2011–present
  - Deep learning and big data



Artificial Intelligence

- Systems that can demonstrate human-like reasoning capability
- AlphaGo Zero: Learning from scratch

Article

Mastering the game of Go without human knowledge

David Silver ✉, Julian Schrittwieser, Karen Simonyan, Ioannis Antonoglou, Aja Huang, Arthur Guez, Thomas Hubert, Lucas Baker, Matthew Lai, Adrian Bolton, Yutian Chen, Timothy Lillicrap, Fan Hui, Laurent Sifre, George van den Driessche, Thore Graepel & Demis Hassabis

Nature **550**, 354–359 (19 October 2017)
doi:10.1038/nature24270
Download Citation

Received: 07 April 2017
Accepted: 13 September 2017
Published online: 18 October 2017

The Prolog programming language

https://en.wikipedia.org/wiki/Prolog

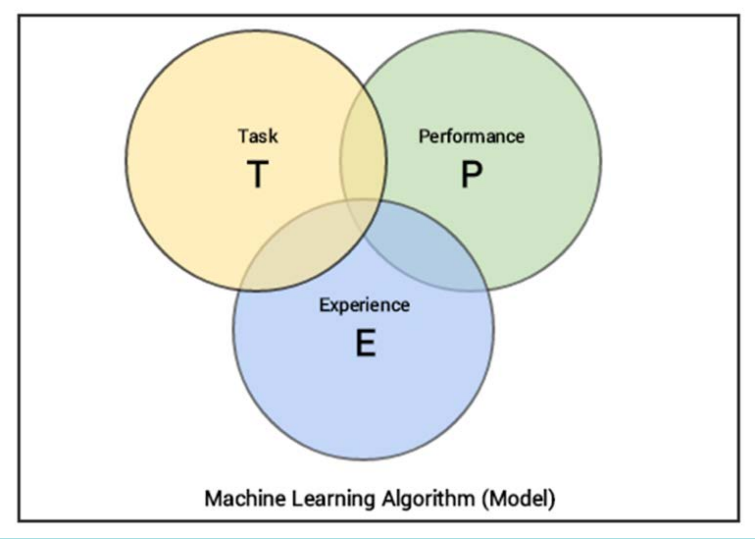
- Rules
  - Head :- Body.
  - animal(X) :- cat(X).
- Facts
  - cat(tom).
- Queries
  - ?- cat(tom).
  - ?- animal(X).

(3) INTRODUCTION TO MACHINE LEARNING

Introduction to Machine Learning

Definition

The art of making machines intelligent without explicit programming



Machine learning is a field which consists of learning algorithms or techniques which

- execute some task T *(Regression/Classification)*
- Improve their performance P *(model performance)*
- with experience E *(Data)*

Introduction contd…

Building Machine learning models is a 3 stage process

- Representation (selection of an algorithm/parameters)
- Evaluation (objective function)
- Optimization (finding the optimal parameters)

Types of Machine Learning techniques

Supervised Machine Learning

Training is done using labelled data

Algorithm learns the mapping function from the input to the output. Y = f(X)

Examples

Regression - used to predict continuous values

Classification - used to predict categorical values

Unsupervised Machine Learning

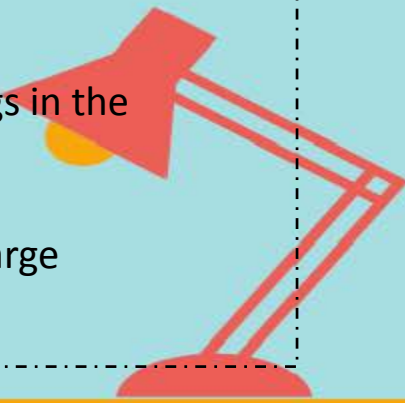
Training is done using unlabelled data

Algorithms are left to their own devises to discover and present the interesting structure in the data.

Examples

Clustering – used to discover the inherent groupings in the data

Association - used to discover rules that describe large portions of the data



Supervised Learning

Supervised M/L – Regression

Definition
Regression is the technique that determines the relationship between one or more independent variables and a dependent variable

Few other naming styles

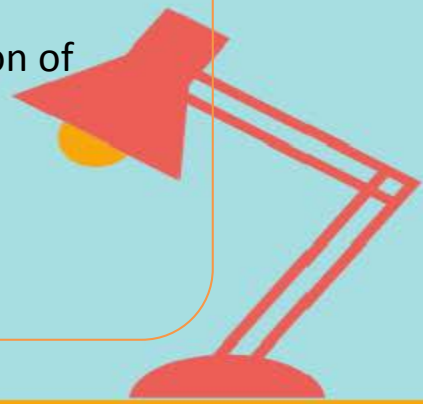
Dependent	: Independent
Target	: Input
Criterion	: Predictor

Linear

- Simple Linear Regression
  Only one independent variable
- Multiple Linear Regression
  Two or more independent variable

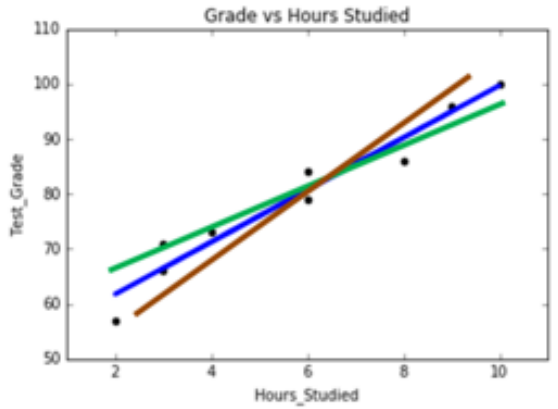
Non Linear Regression

Dependent on non-linear transformation of independent variables

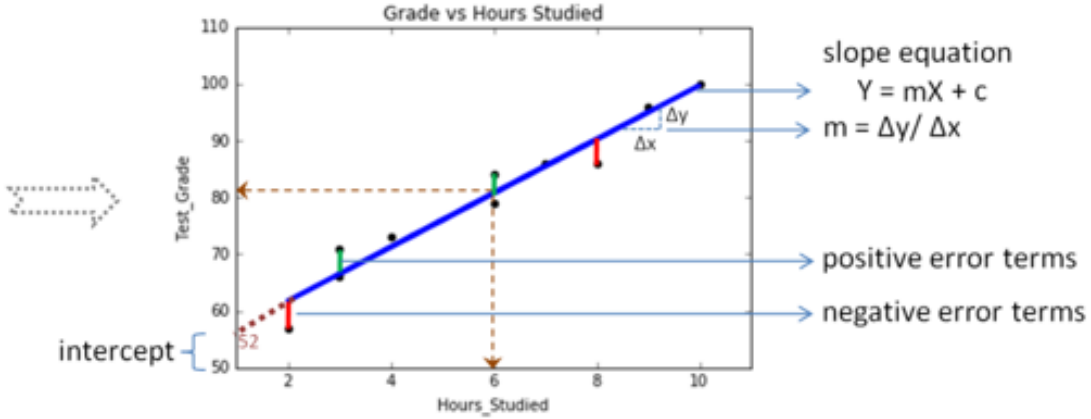


Simple Linear Regression

In Simple Linear Regression, we fit the best line between the dependent variable and the independent variable given as y = mx + c
m = Co-efficient of x (i.e. change in y divided by change in x)
C = intercept (represents the variability in y, unexplained by x)



Different possible slopes



Best slope that gives minimal sum of squared error

Few points to wander-
How is the best line calculated?
- OLS

Why not use just correlation?
- Intercept

Multiple Linear Regression

Real life scenarios will definitely have many more independent variables than just 1
To model an equation that studies the relationship between the dependent variable and multiple independent variables

The representation equation can be extended to

$$y = m0 + m1\ x1 + m2\ x2 + m3\ x3 + . . .+ mn\ xn$$

Where
- m0 is the intercept
- m1 is the coefficient of variable x1
- m2 is the coefficient of variable x2

and so on…..

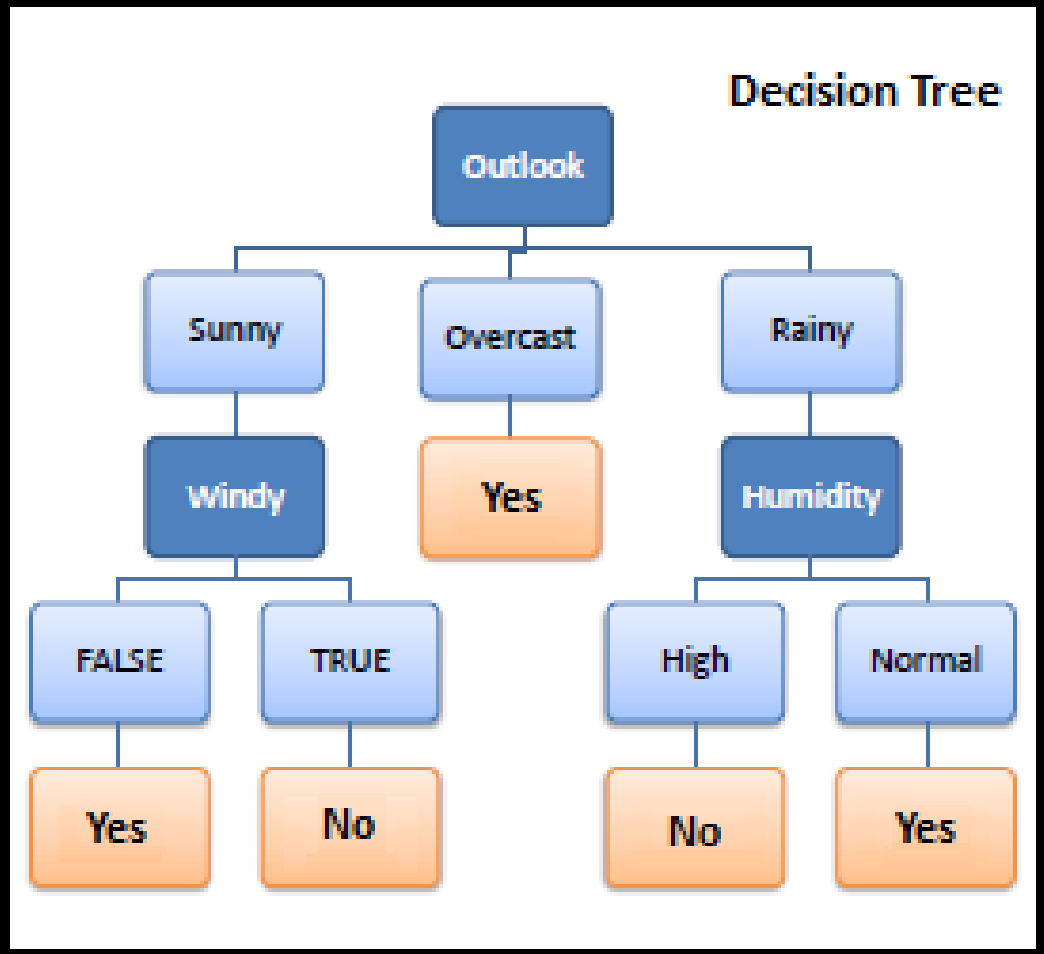
Supervised ML - Decision Trees

Definition

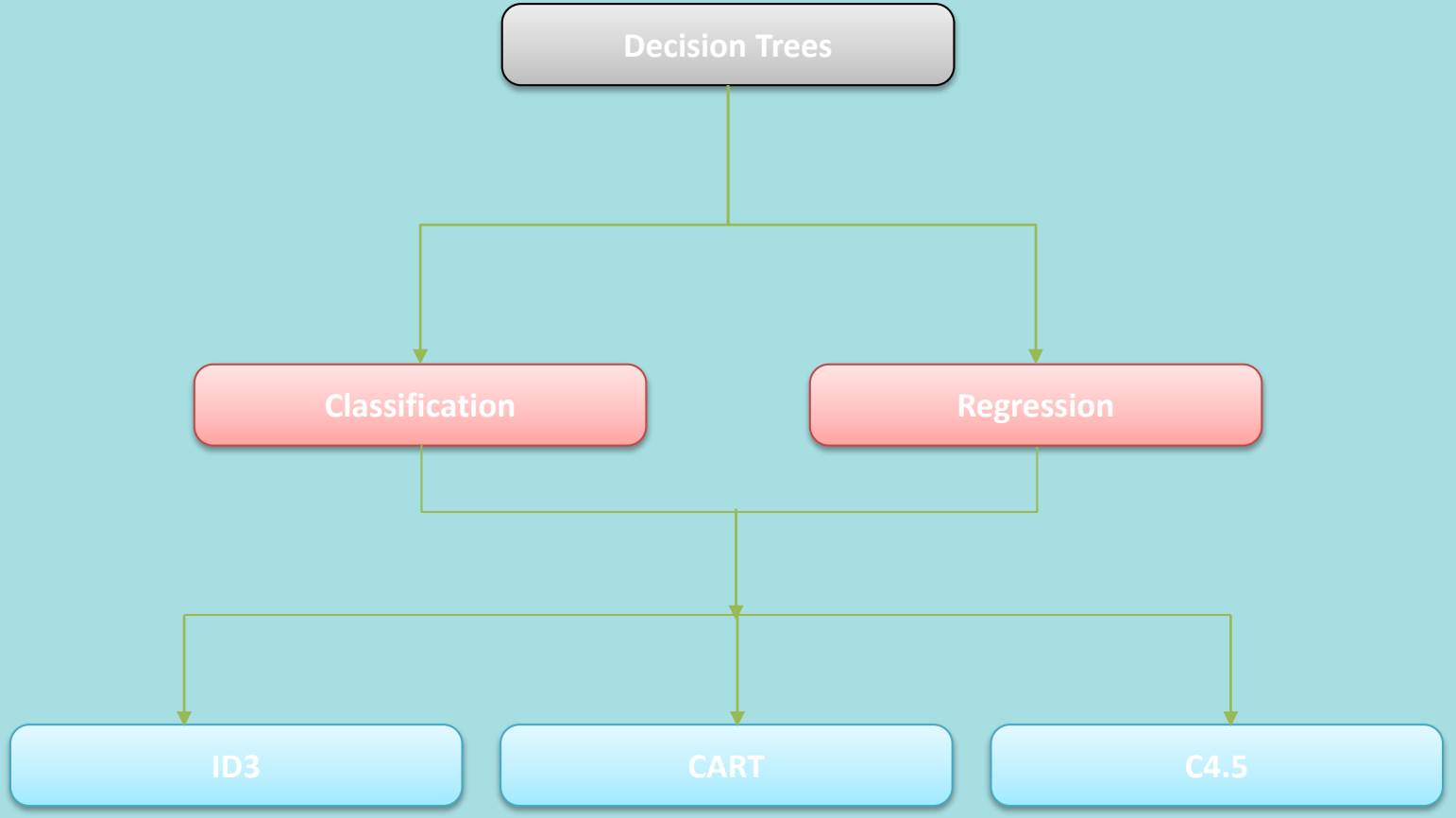
Decision Trees are a class of Supervised Learning Algorithms which can be used for predicting categorical or continuous variables

How does it work?

- It works by breaking data from the root node into smaller and smaller subsets while incrementally building associated decision tree.
- The final result is a tree with a root node, decision nodes and leaf nodes.
- Decision nodes create a rule and leaf nodes deliver a result



Types of Decision Trees



And a few more…

Building Decision Trees

Pseudocode

1) Select Root Node

2) Partition Data into respective groups

3) Create a Decision node

4) Partition Data into respective groups

Repeat until the node size > threshold or Features = empty

Few questions?

- **How** to select the Root Node?
- **How** are the decision nodes ordered/chosen?
- **When** does the branching stop?
- **How** does the tree treat continuous variables?
- **How** different is the process for classification and Regression?

Entropy

What is it?

Entropy (measures the homogeneity of the sample)

-pLogp – qLogq (base 2)

Where p = Probability of event happening
q = probability of event not happening

How is it done?

Consider the problem of
Predicting whether the Employee is a Techie or a Non-Techie?

Techie	Non-techie
46	15

p	0.75
q	0.25

Entropy(Emp) : Entropy(Techie,Non-Techie)
Entropy(46,15) -pLogp – qLogq : -0.75Log(0.75) -0.25Log(0.25)
= 0.8

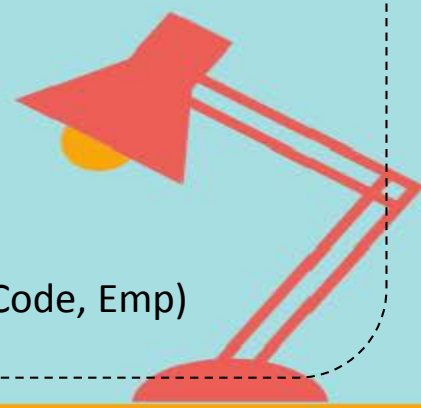
Dress codes	Techie	Non-Techie	Total
Formals	6	8	14
Casuals	31	4	35
Business Casuals	9	3	12
			61

Entropy(Employee,Dress-code)

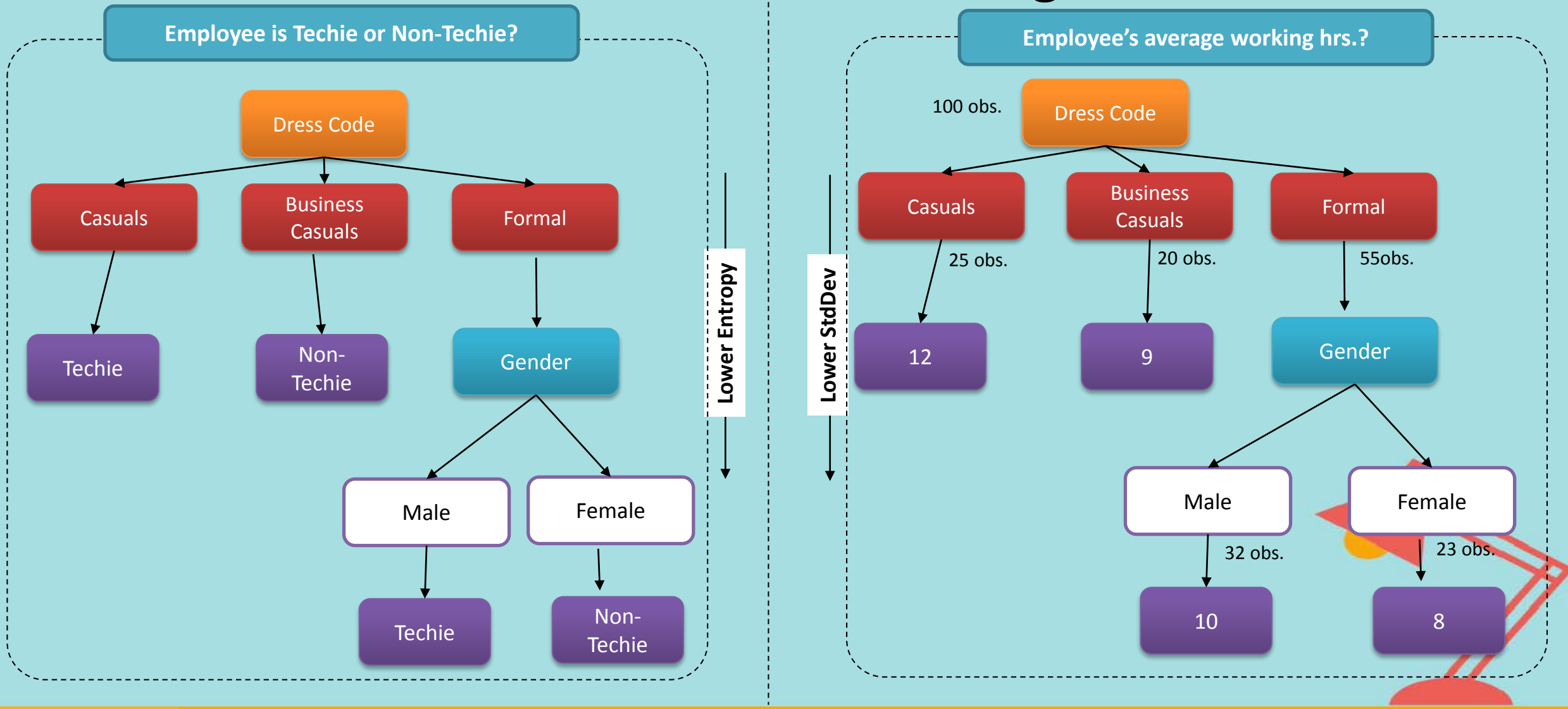
Prob(Formals) * Entropy_formals +
Prob(Casuals) * Entropy_casuals +
Prob(BCasuals) * Entropy_Bcasuals

(14/61) * Entropy(6,8) +
(35/61) * Entropy(31,4) +
(12/61) * Entropy(9,3)
= 0.68

Gain(Emp) =
Entropy(Emp) – Entropy(DressCode, Emp)
= 0.80 – 0.69 = **0.11**



Decision Trees - Classification & Regression



Unsupervised Learning

Clustering

The key objective in clustering is to identify distinct groups/clusters)based on some notion of similarity within a given dataset.
Types of clustering

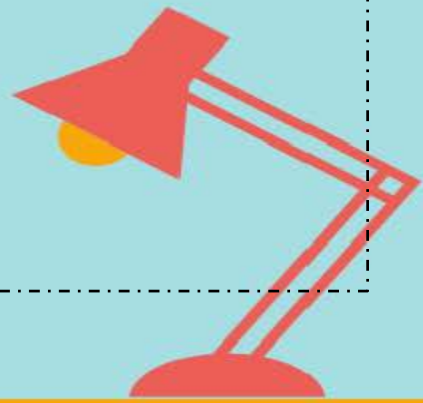
- Agglomerative (hierarchical)
- Divisive (k – means)

K-Means clustering

- Start with random point initialization of the required number of centers. ('K' in K-means stands for the number of clusters)
- Assign each data point to the 'center' closest to it. (distance metric := normal Euclidian distance)
- Recalculate centers by averaging the dimensions of the points belonging to the cluster.
- Repeat with new centers until we reach a point where the assignments become stable.

Hierarchical clustering

- Start with n clusters (n = # of datapoints)
- Combine the 2 closest clusters
- Repeat till only 1 cluster exists



Association Rule Mining

Association rule mining is a procedure to find frequent patterns, correlations, associations, or causal structures from data sets

Algorithm details

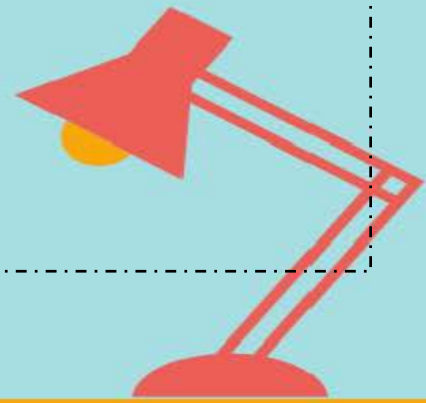
- Item set – The list of all transactions; {milk, bread}, {apples, oranges}, {milk}

- Support - # of times an item appears in a dataset
  i.e. support(milk, bread) = (# transactions with milk and bread)/(Total # of transactions)

- Confidence – Measure of # of times a rule is found to exist in the dataset
  i.e. confidence(milk -> bread) = support(milk and bread)/ support(milk)

Main Applications

- Cross sell/Up-sell
- Market Basket Analysis



(4) INTRODUCTION TO DEEP LEARNING

Brief History - Neural Networks

1958 Perceptron

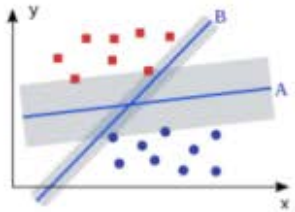
1969
Perceptron criticized

1974 Backpropagation

awkward silence (AI Winter)

Convolution Neural Networks for Handwritten Recognition
1998

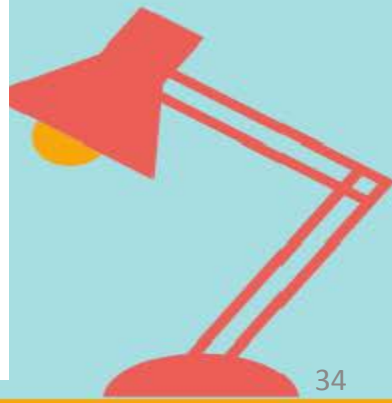
1995
SVM reigns



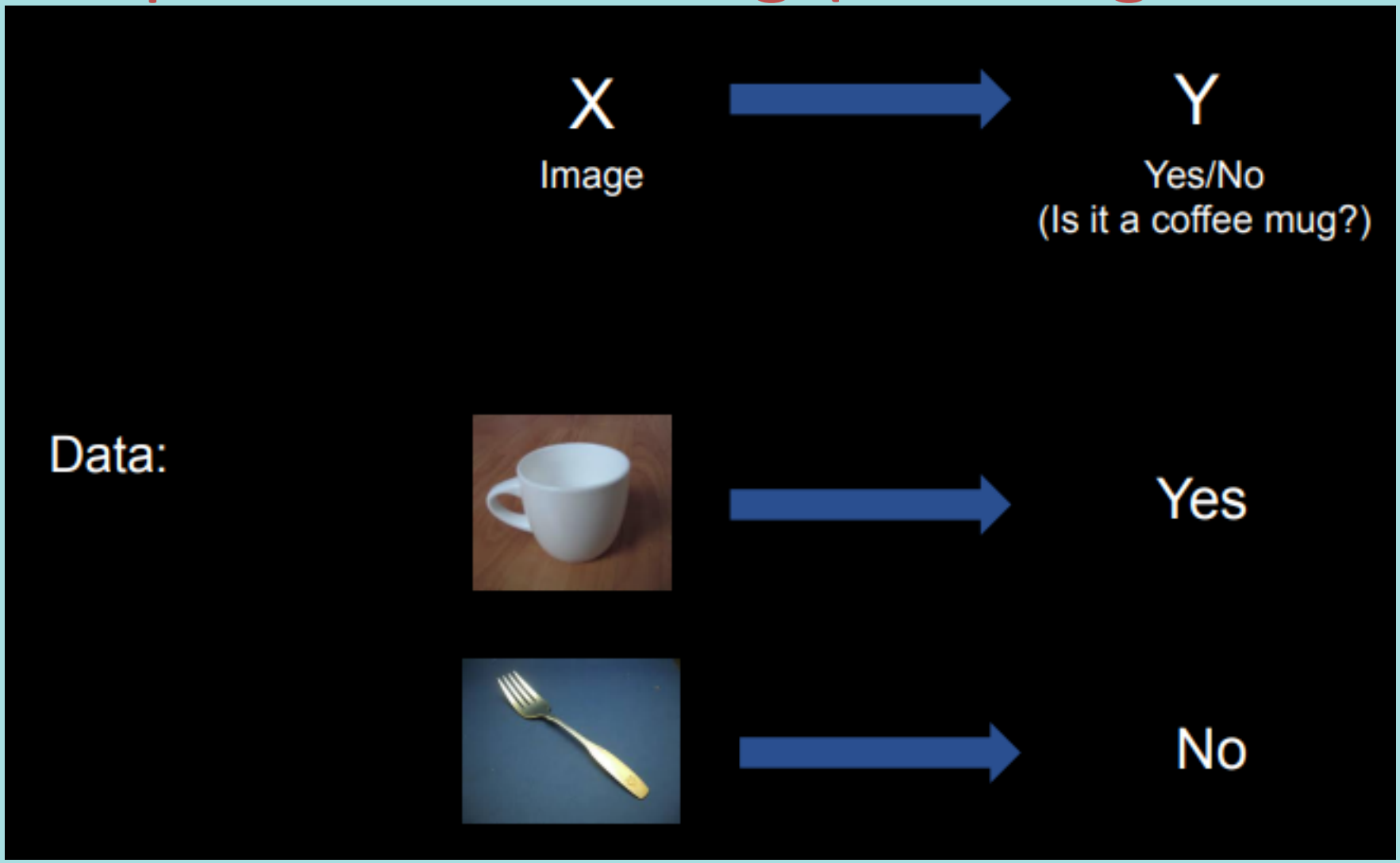
2006
Restricted Boltzmann Machine

Google Brain Project on 16k Cores
2012

2012
AlexNet wins ImageNet
IMAGENET

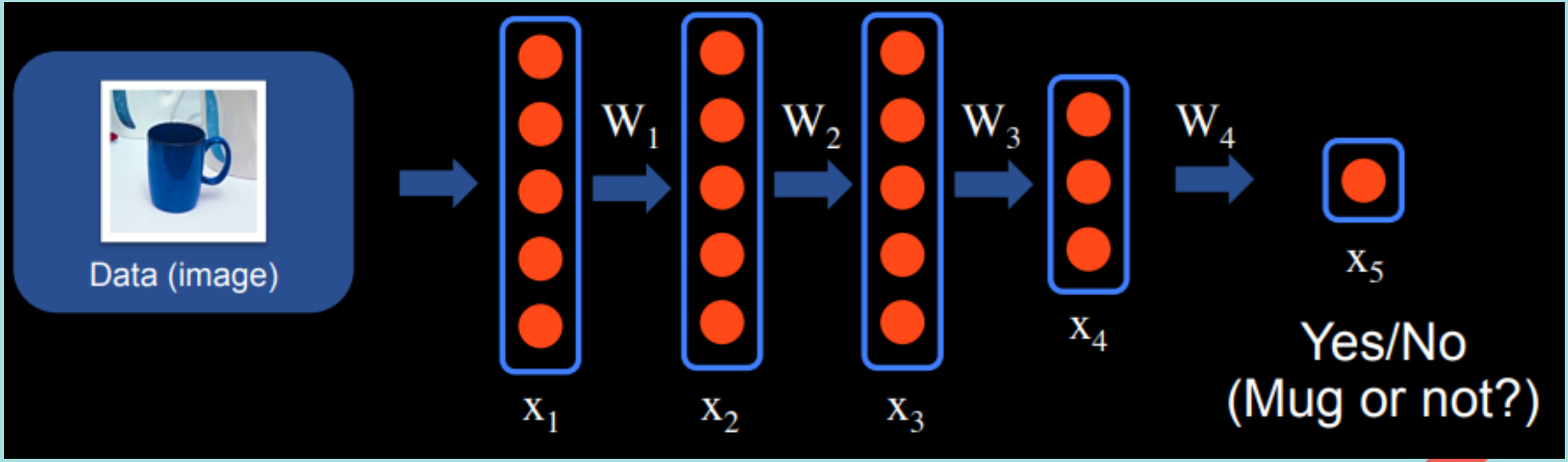


Supervised learning (learning from tagged data)



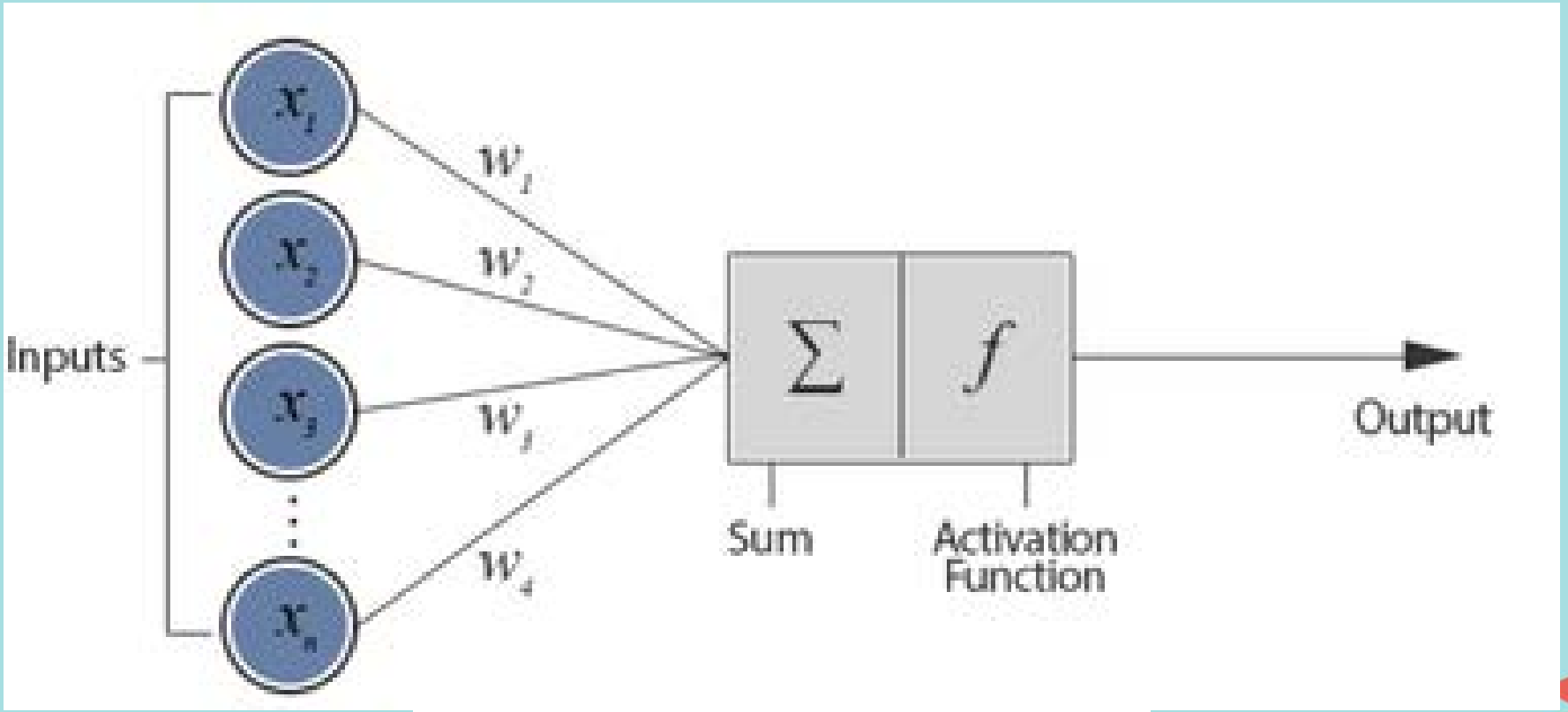
https://www.kth.se/social/files/58d4dca1f276544c5050a5cf/Lecture1.pdf

What is a neural network?



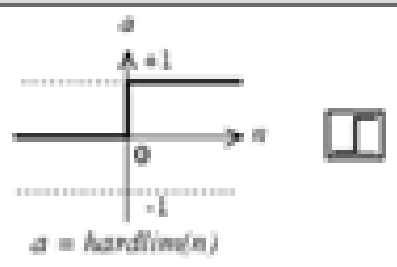
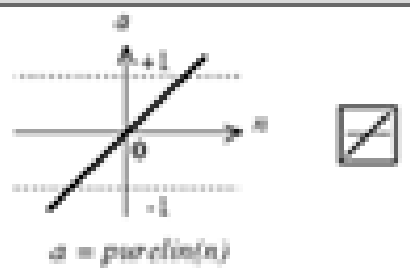
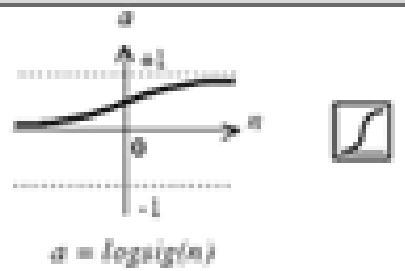
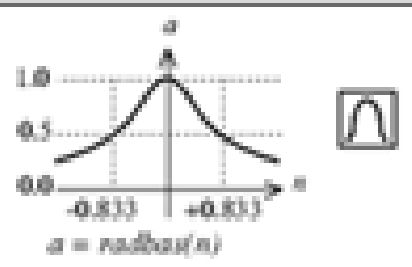
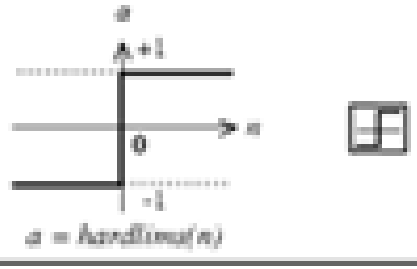
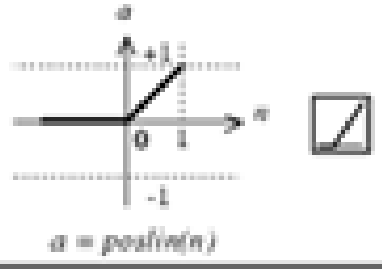
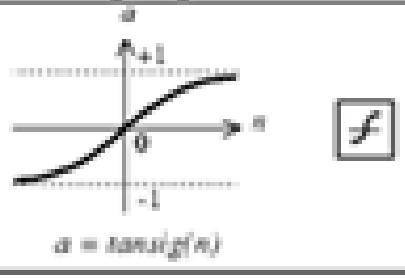
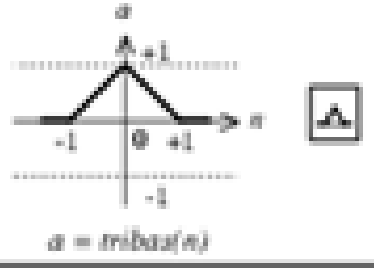
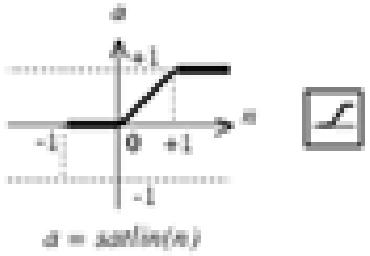
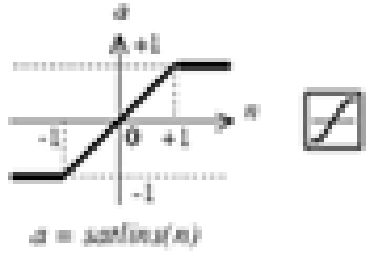
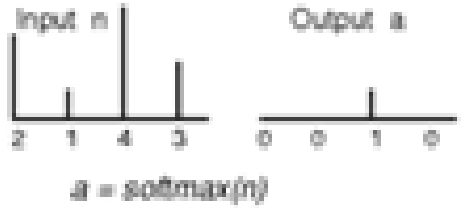
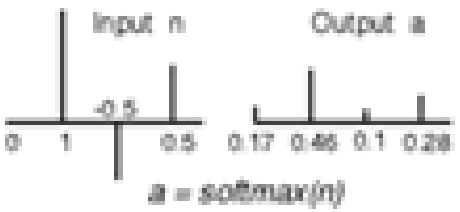
https://www.kth.se/social/files/58d4dca1f276544c5050a5cf/Lecture1.pdf

Perceptron



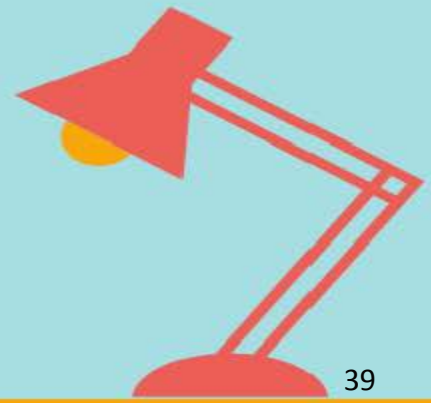
$$\text{Output} = \text{f}\left(\sum_{k=1}^{n} x_k \cdot w_k\right)$$

Activation functions

 a = hardlim(n)	 a = purelin(n)	 a = logsig(n)	 a = radbas(n)
Hard-Limit TF	Linear TF	Log-Sigmoid TF	Radial Basis TF
 a = hardlims(n)	 a = poslin(n)	 a = tansig(n)	 a = tribas(n)
Symmetric Hard-Limit TF	Positive Linear TF	Tan-Sigmoid TF	Triangular Basis TF
 a = satlin(n)	 a = satlins(n)	 a = softmax(n)	 a = softmax(n)
Satlin TF	Satlins TF	Compet TF	Softmax TF

Deep Neural Networks

- Dramatic improvements in accuracy and runtime in multiple domains
  - Image Detection
  - Face Recognition
  - Gesture Recognition
  - Video Search & Analytics
  - Speech Recognition & Translation
  - Recommendation Engines
  - Indexing & Search



Images Classification - Neural Networks

Upload your photo

You can upload your photo or paste any URL to an image

Generated tags

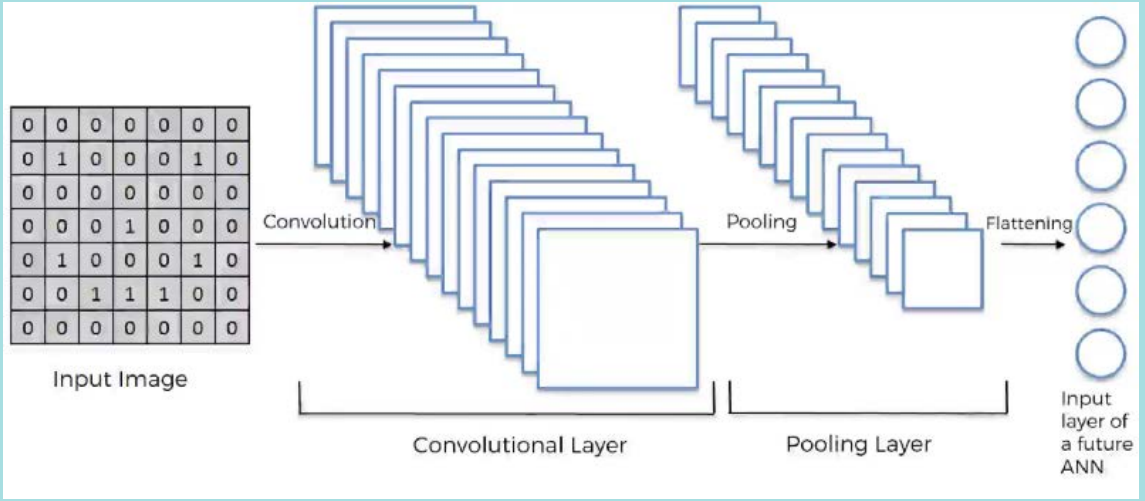
Concepts

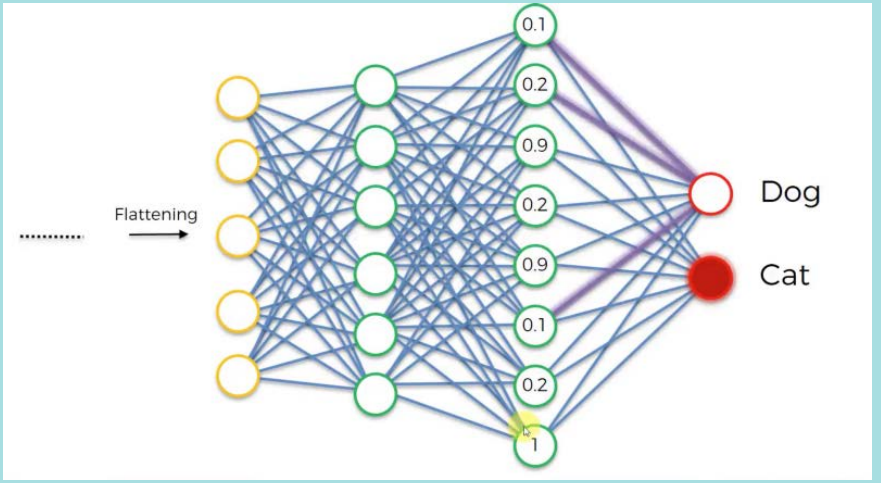
deer	79.76%
bighorn	38.25%
caribou	34.53%
mountain sheep	30.61%
bovid	25.83%
animal	23.11%
wild sheep	22.96%
wilderness	18.55%
goat	17.59%
mammal	15.33%

↪ show me more tags

Convolutional Neural Network

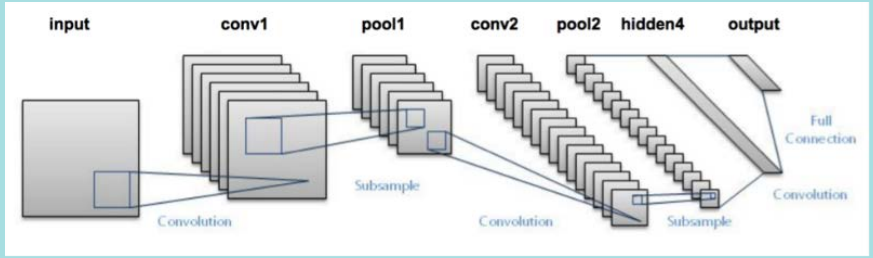
- **Convolution Layer**
  - Feature Detector
- Pooling Layer
  - Max Pooling
- Fully Connected Layer

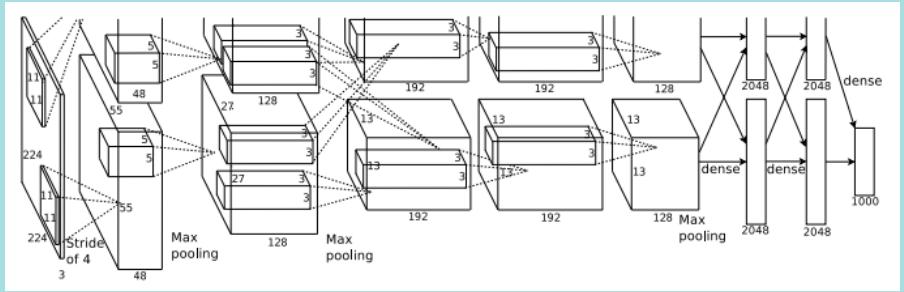


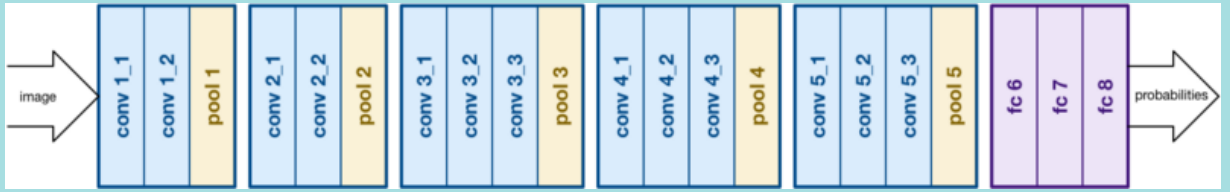


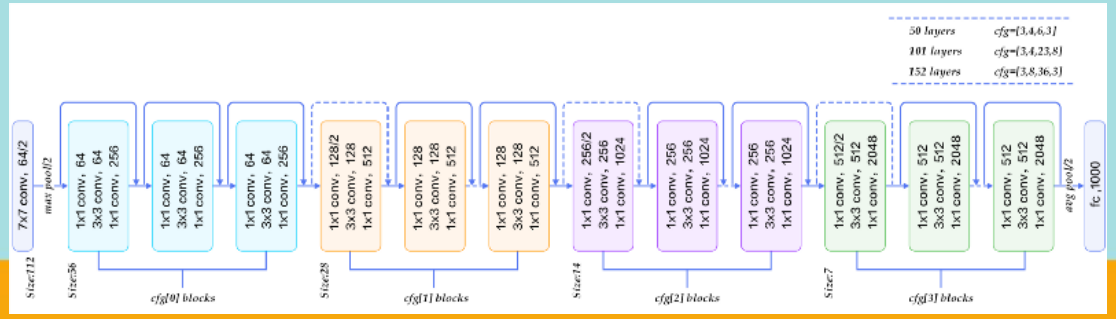
CNN Architectures

- LeNet-5 (1998)
- AlexNet (2012)
- VGG (2014)
- ResNet (2015)



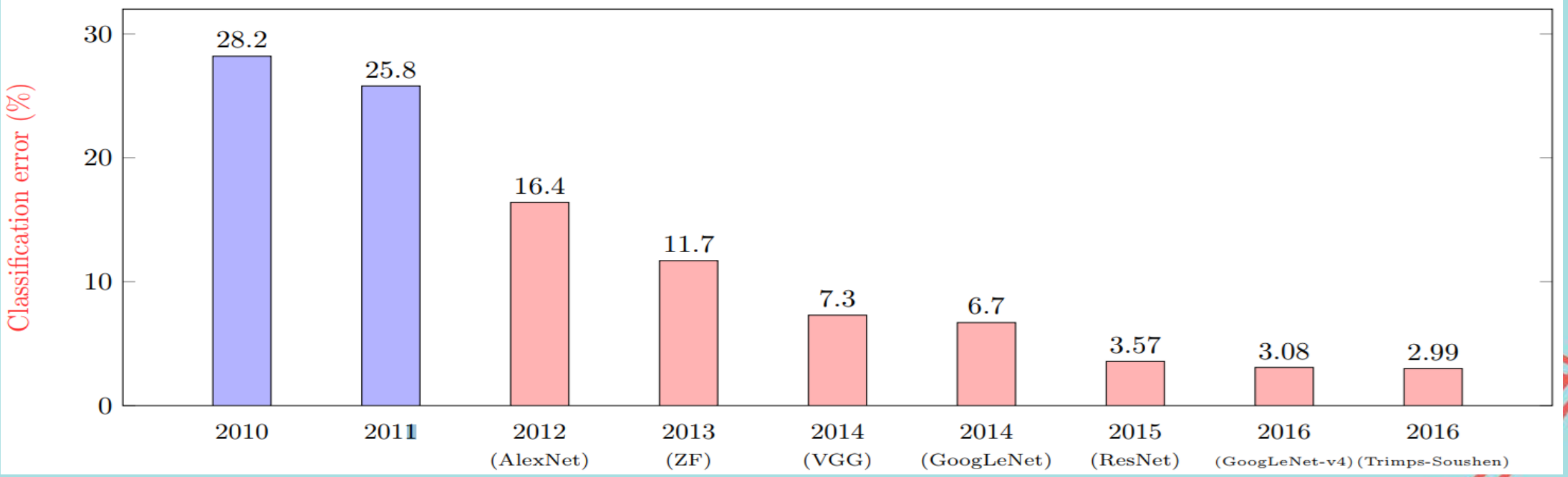




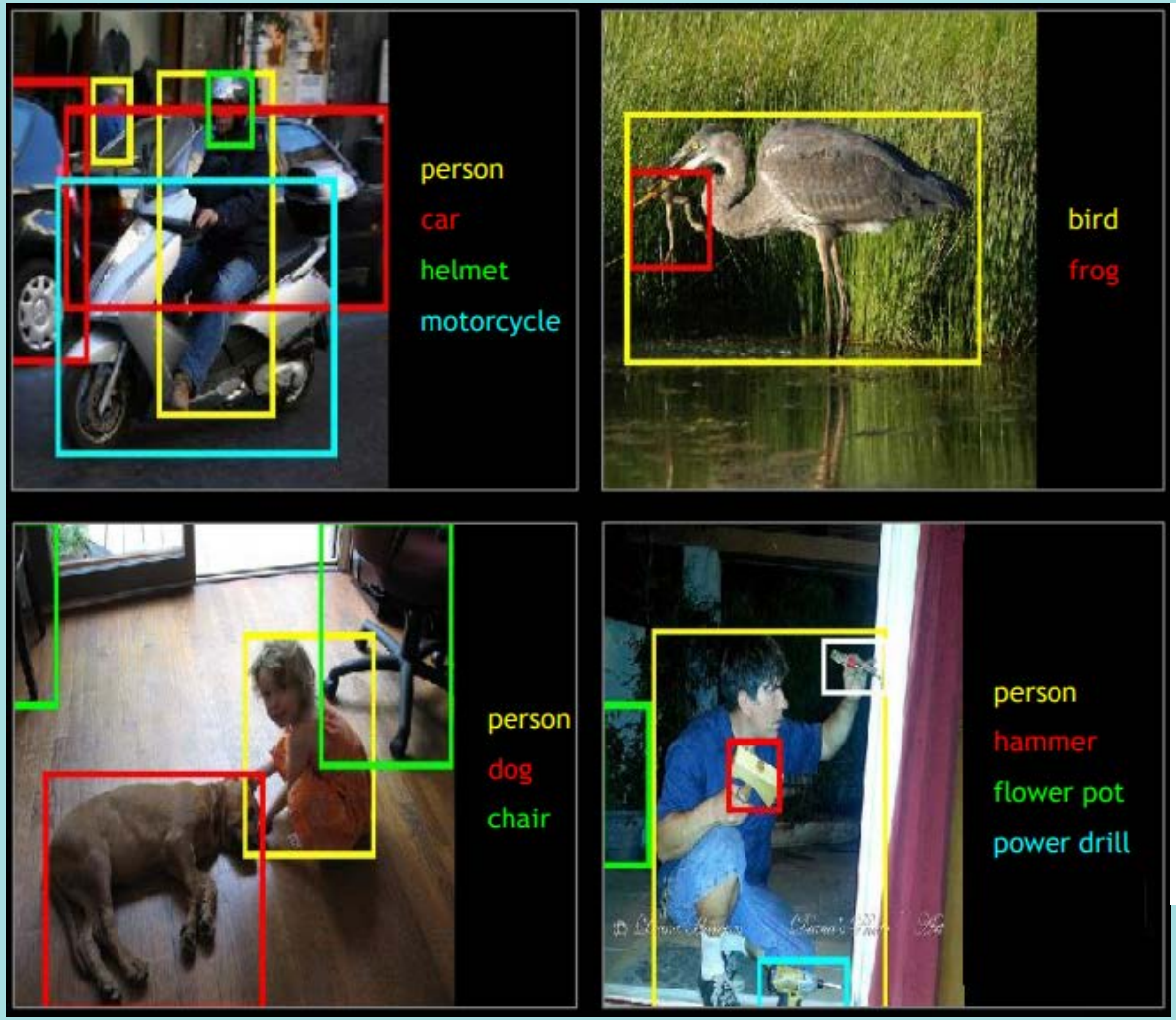


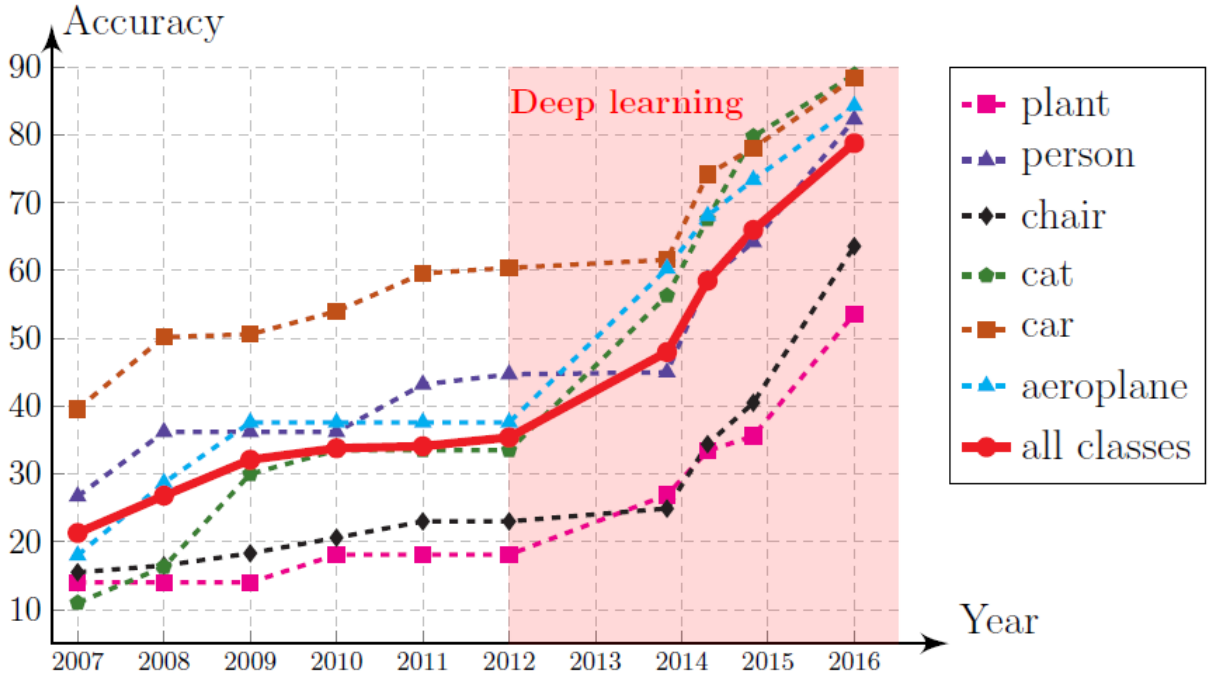
ImageNet Large Scale Visual Recognition Challenge (ILSVRC)

- 1000 object classes, 1.4 million labelled training images



Object Localization - Neural Networks

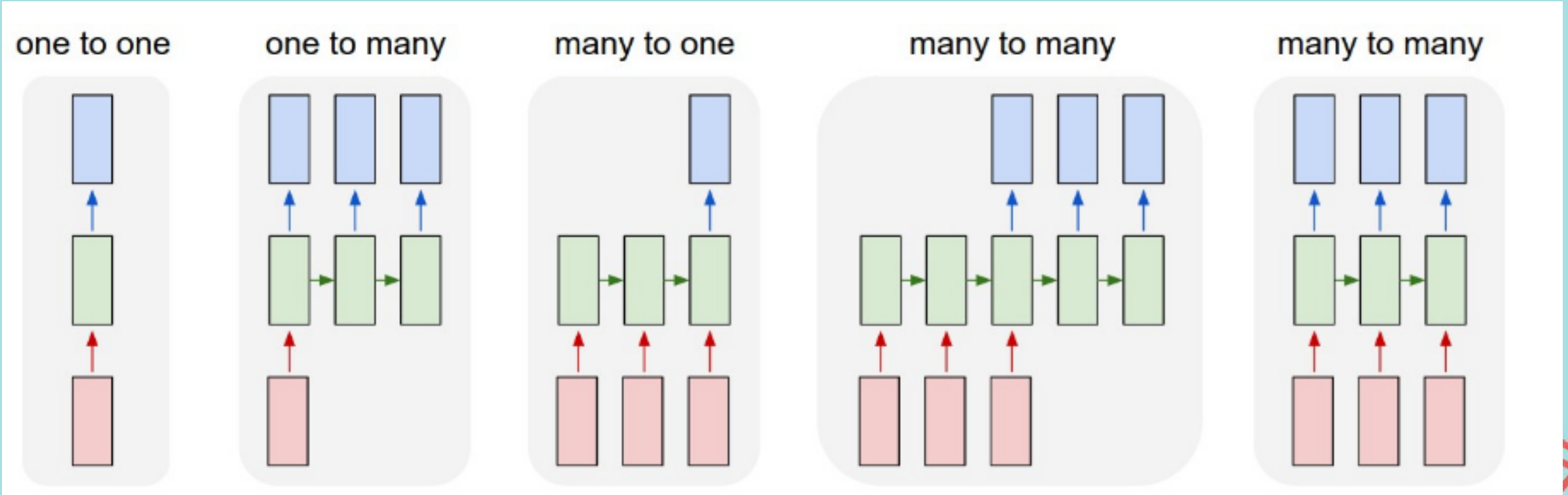




Progress on the Pascal VOC 2007 challenge.

Recurrent Neural Networks

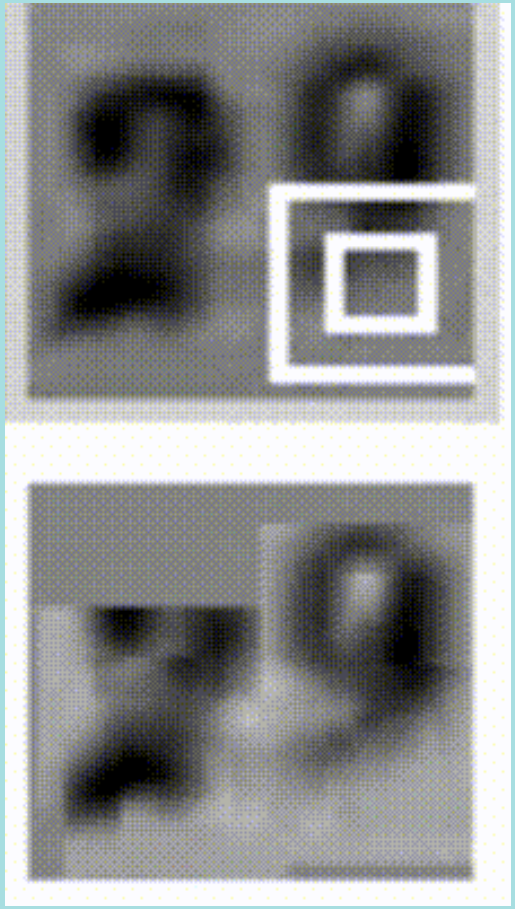
- Sequences - Learn and Generate

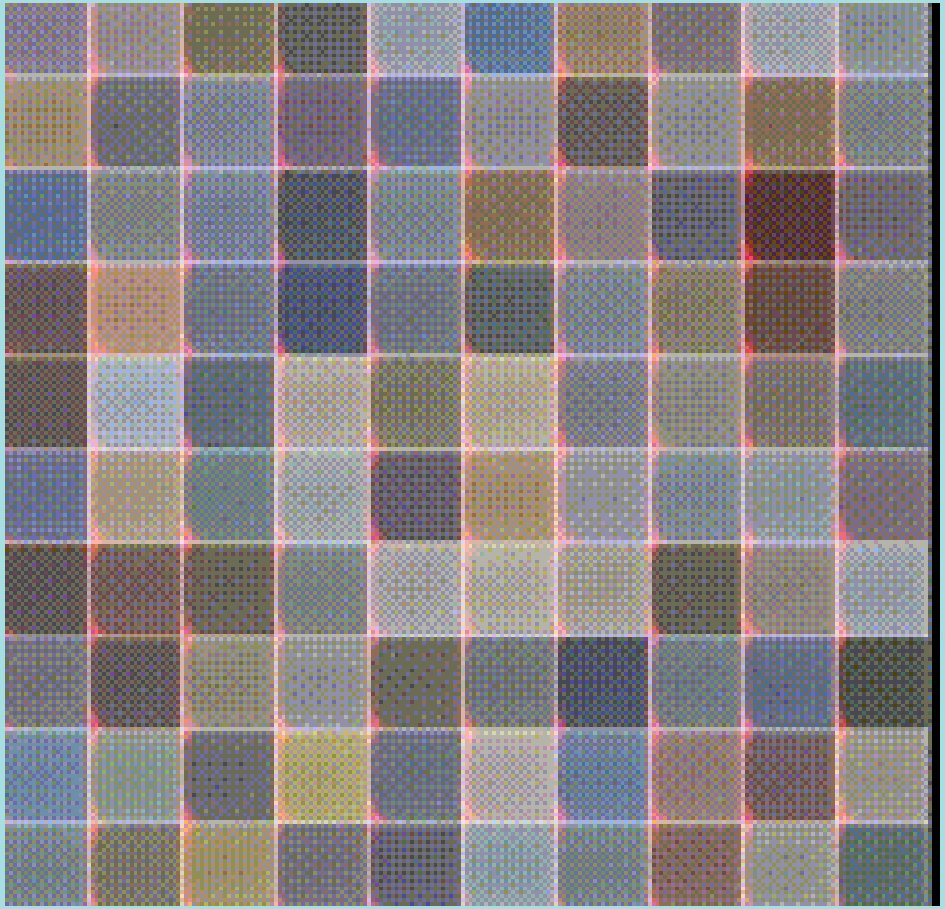


http://karpathy.github.io/2015/05/21/rnn-effectiveness/

Sequential processing in absence of sequences

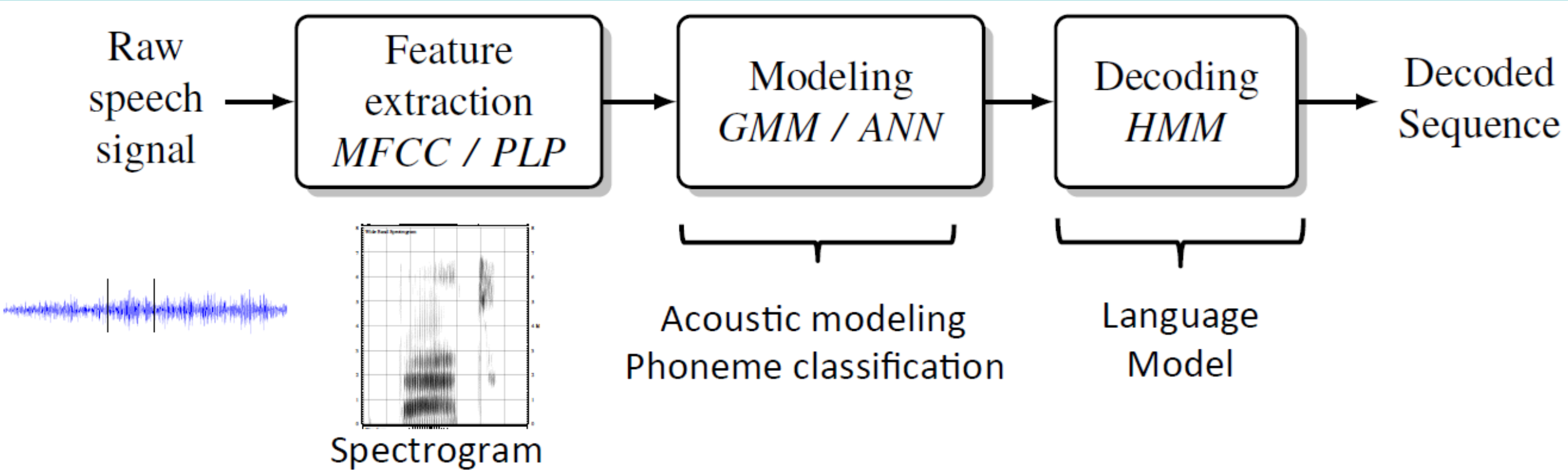
- Left: RNN learns to read house numbers. Right: RNN learns to paint house numbers.





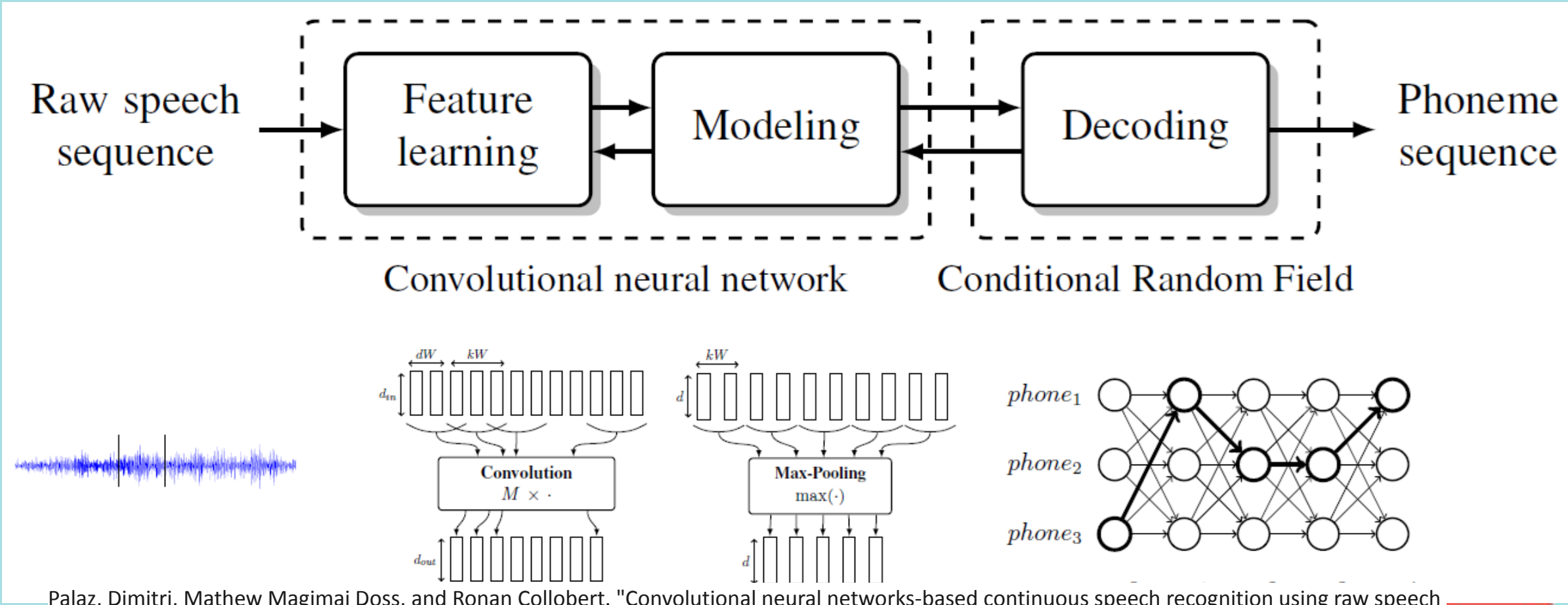
http://karpathy.github.io/2015/05/21/rnn-effectiveness/

Speech recognition



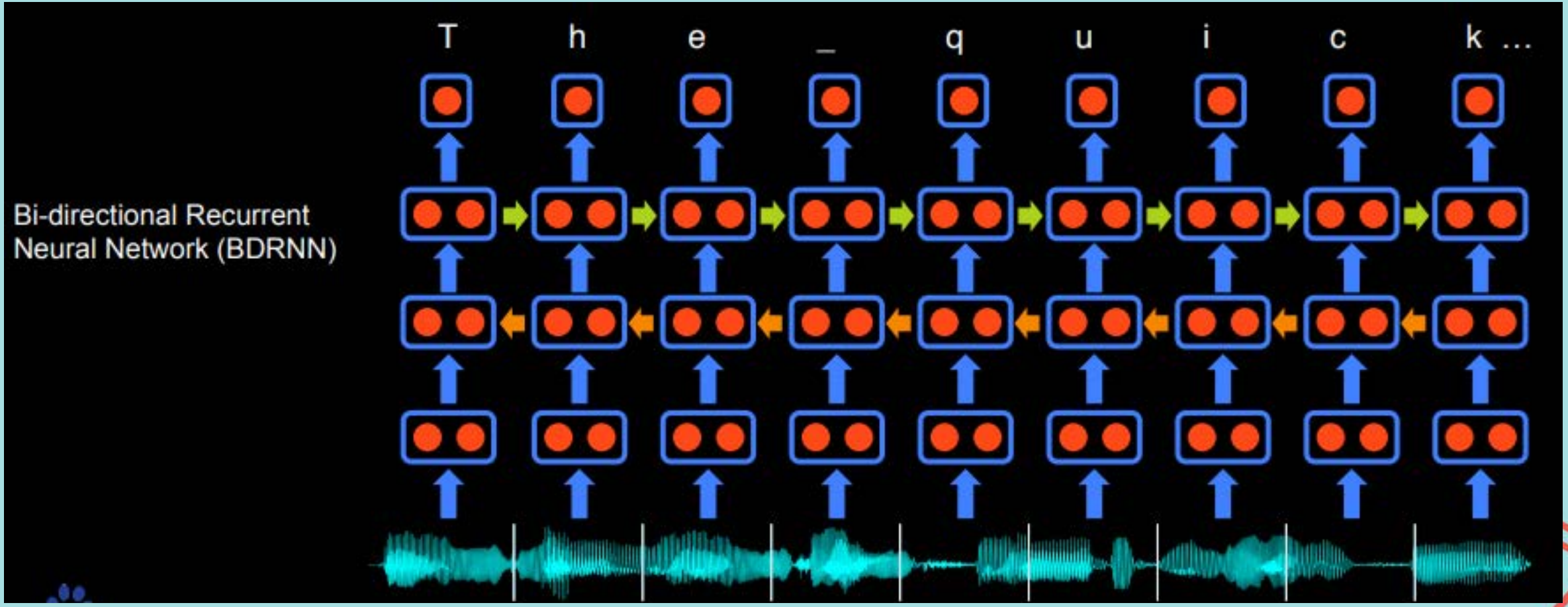
Speech recognition - Neural Networks

- End-to-End Recognition



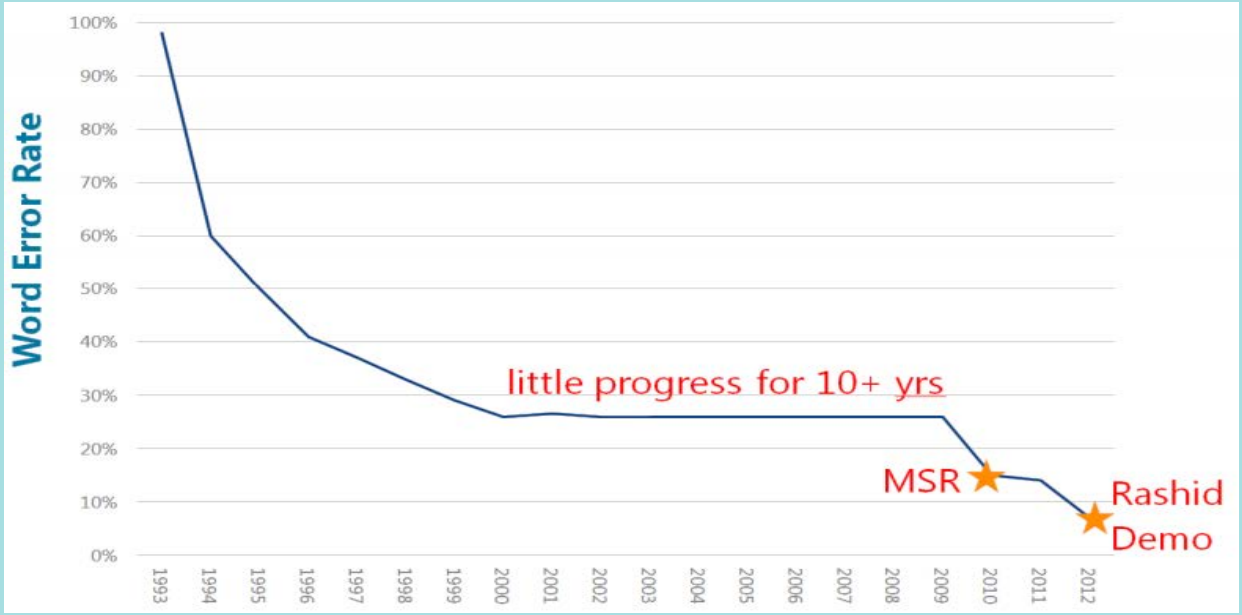
Palaz, Dimitri, Mathew Magimai Doss, and Ronan Collobert. "Convolutional neural networks-based continuous speech recognition using raw speech signal." *Acoustics, Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on*. IEEE, 2015.

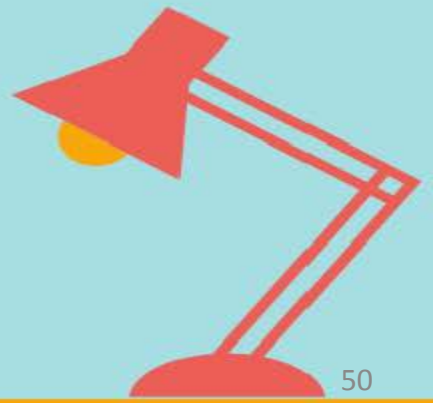
Speech recognition - Recurrent Neural Networks



Progress of speech recognition

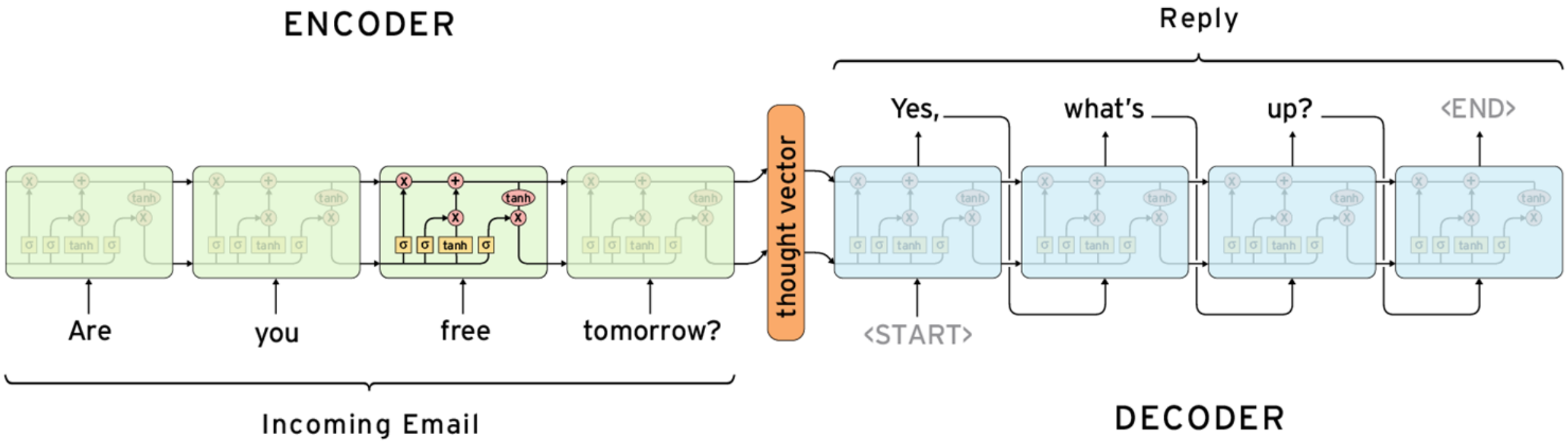
- After no improvement for 10+ years by the research community . . . MSR reduced error from ~27% to < 13% (and under < 7% for Rick Rashid's demo in 2012)!





Dialog Systems - Neural Networks

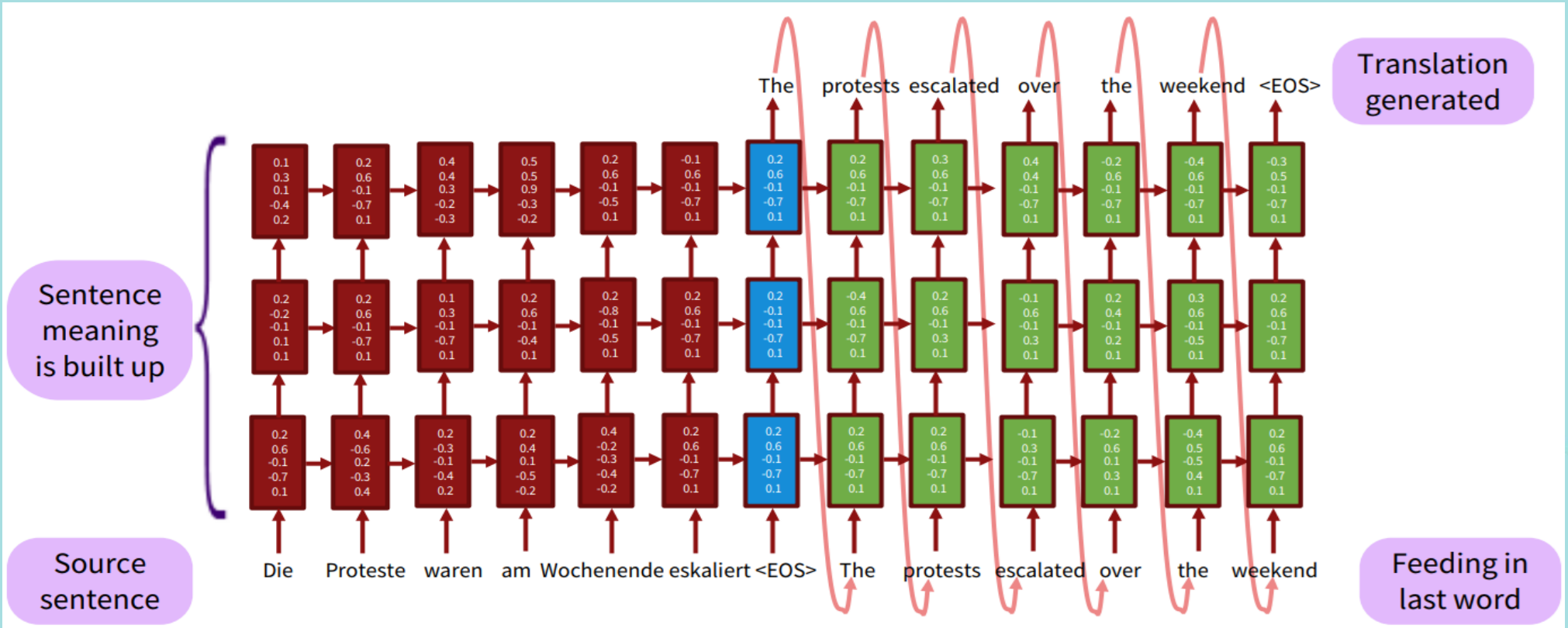
- Sequence-to-Sequence Model



http://www.wildml.com/2016/04/deep-learning-for-chatbots-part-1-introduction/

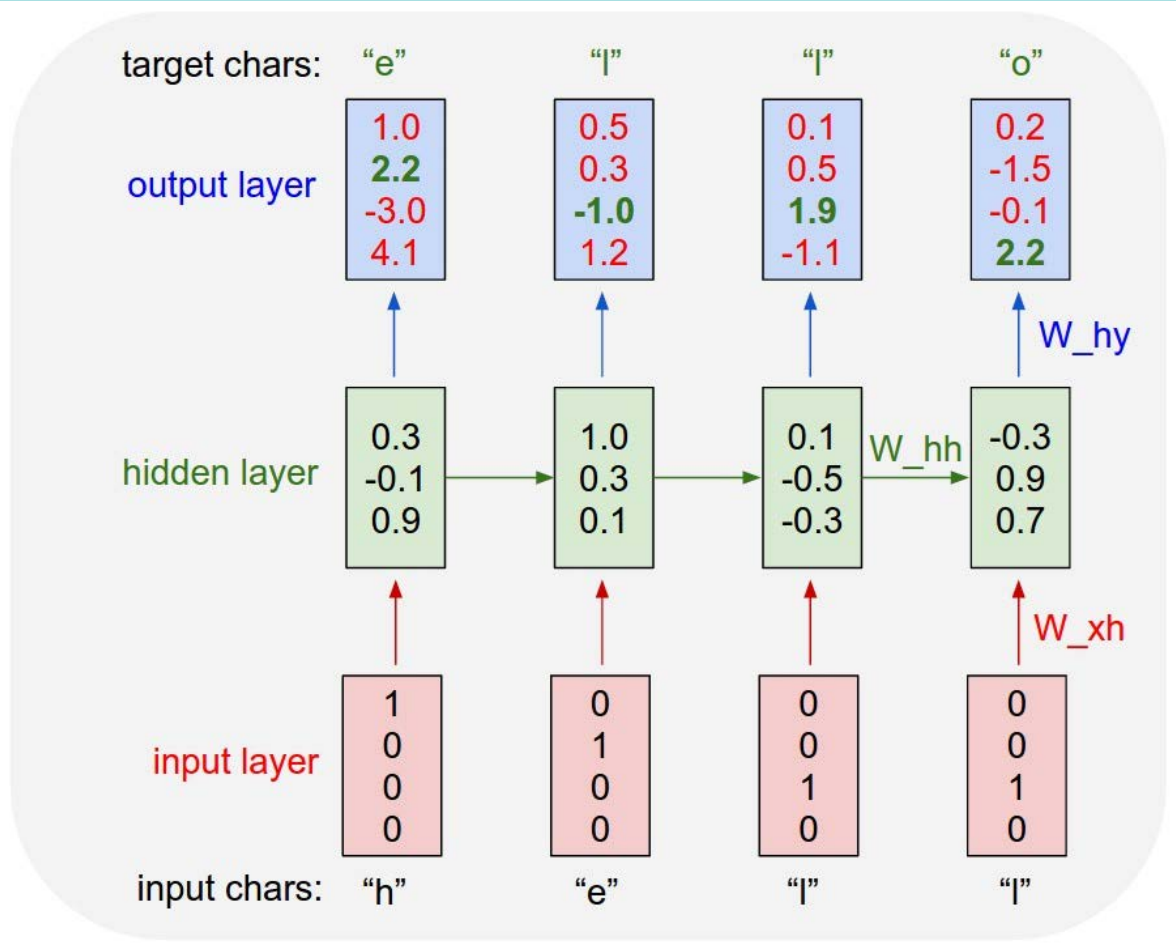
Natural Language Translation - Neural Networks

- Sequence-to-Sequence Model



Sutskever, Ilya, Oriol Vinyals, and Quoc V. Le. "Sequence to sequence learning with neural networks." *Advances in neural information processing systems*. 2014.

Character-Level Language Models

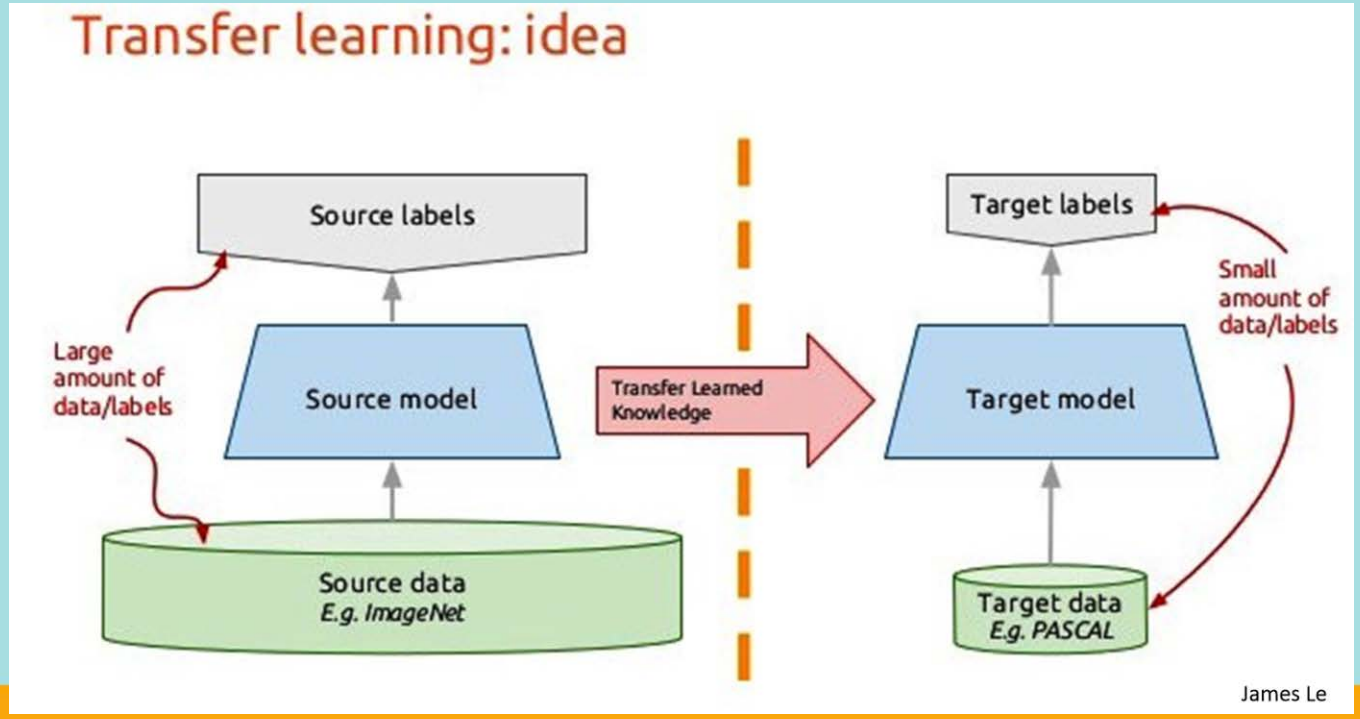


- An example RNN with 4-dimensional input and output layers, and a hidden layer of 3 units (neurons).
- This diagram shows the activations in the forward pass when the RNN is fed the characters "hell" as input.
- The output layer contains confidences the RNN assigns for the next character (vocabulary is "h,e,l,o");
- We want the green numbers to be high and red numbers to be low.

http://karpathy.github.io/2015/05/21/rnn-effectiveness/

Transfer Learning for Deep Learning

- Transfer learning make use of the knowledge gained while solving one problem and applying it to a different but related problem.



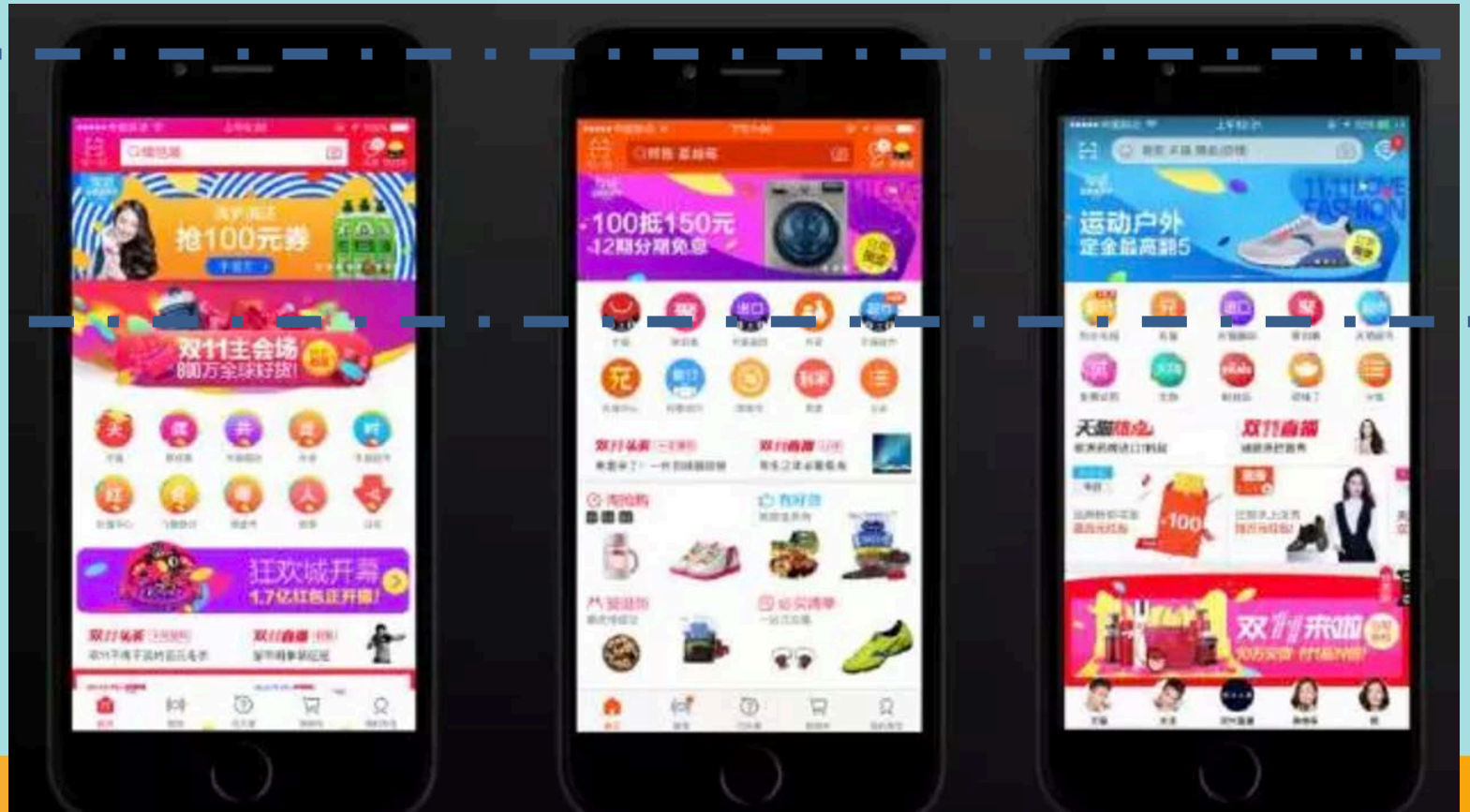
Deep Neural Networks

- Image Detection
- Face Recognition
- Gesture Recognition
- Video Search & Analytics
- Speech Recognition & Translation
- Recommendation Engines
- Indexing & Search

(5) APPLICATION: AI DESIGNER

鲁班 (Alibaba Luban AI banner

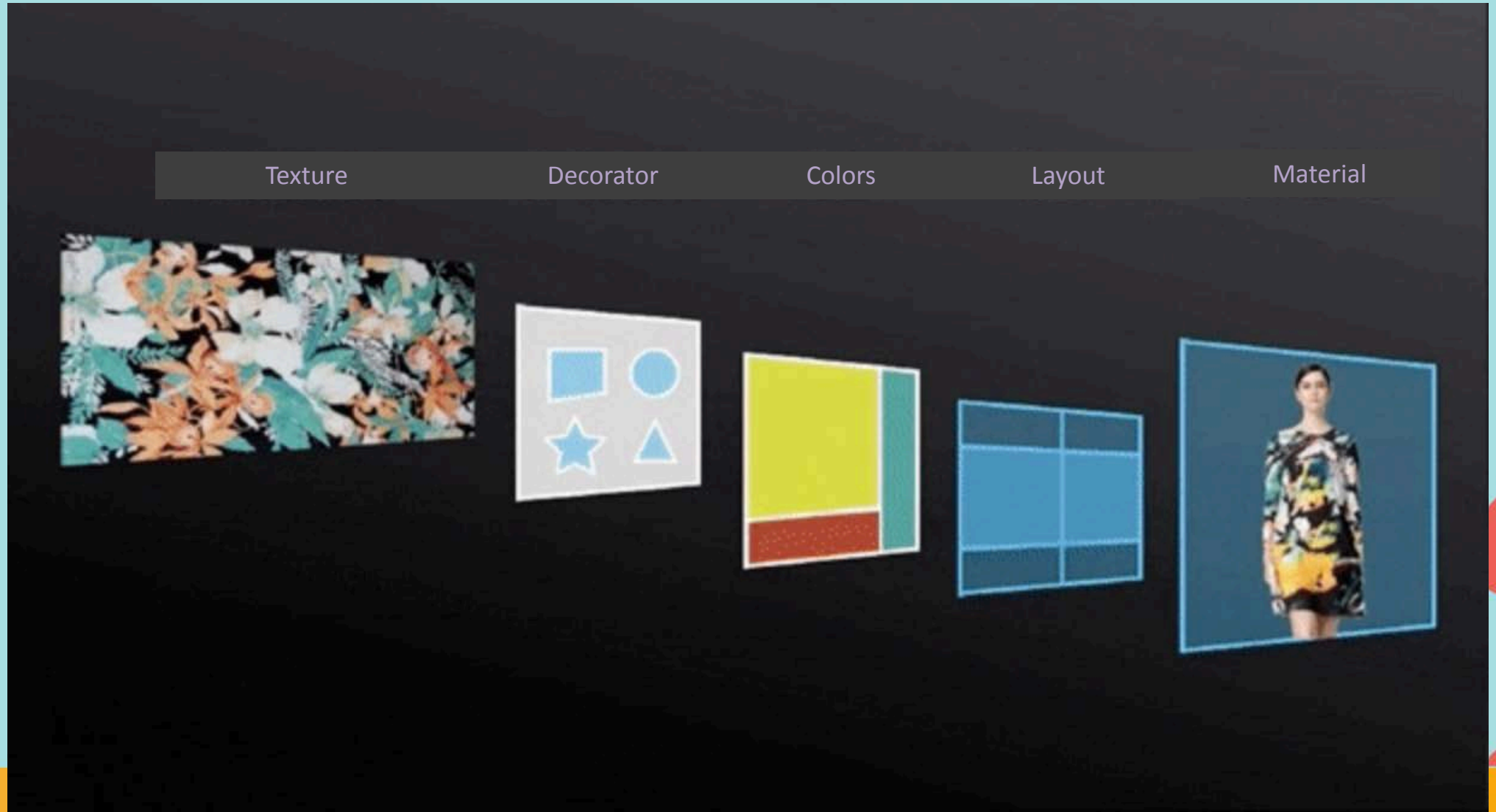
Alibaba AI Design Lab



AI-Made banner

170 Million banners

AI Design for banner



构建元素中心

The dimensions of design

Scoring design

Define the quality of autonomous design

The model of deep learning

Training the AI system

(6) DISCUSSIONS

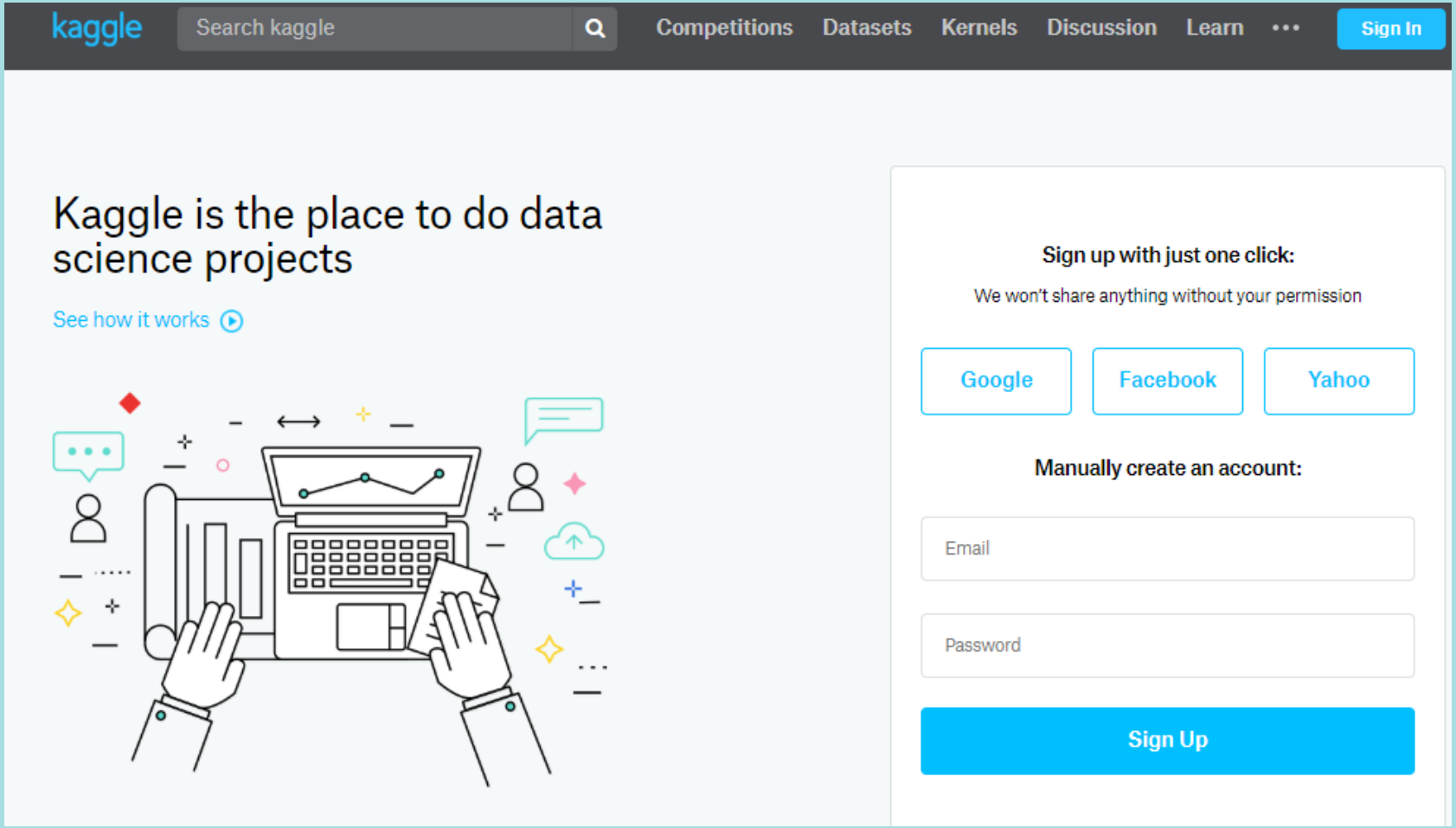
The UAE just appointed a minister for Artificial Intelligence

Saudi Arabia to invest $500 billion in fully automated city spanning three countries

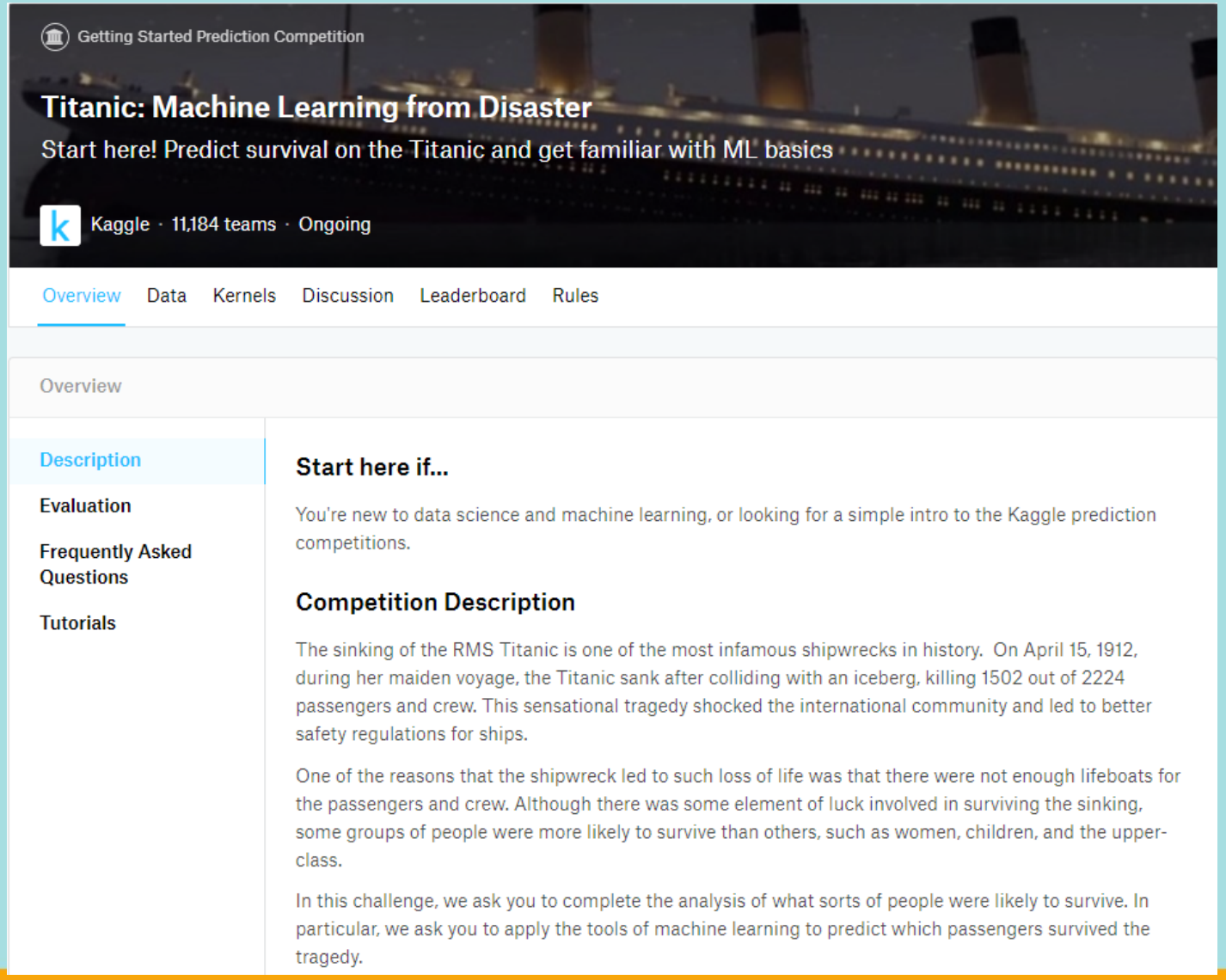
This week, Saudi Arabia granted citizenship to a robot and unveiled plans for a fully autonomous city

Danil Boparai | 28 October 2017 | 3 comments

Kaggle


kaggle
Search kaggle
Competitions
Datasets
Kernels
Discussion
Learn
Sign In
Kaggle is the place to do data science projects
See how it works
Sign up with just one click:
We won't share anything without your permission
Google
Facebook
Yahoo
Manually create an account:
Email
Password
Sign Up

Titanic: Machine Learning from Disaster



Assignment 1B

- Step 1: Open colab and create a new notebook
  - https://colab.research.google.com
- Step 2: Edit the notebook with text cells
  - https://guides.github.com/features/mastering-markdown/
- Step 3:
  - List AI Terms in this topics
  - Search AI challenge from the Internet and introduce one of them
- Step 4:
  - Rename the notebook: Untitled.ipynb->day1b.ipynb
  - Copy the notebook to GitHub
  - Check the notebook on GitHub

Thanks!

Q&A