Week 8

Halle Berry Neuron

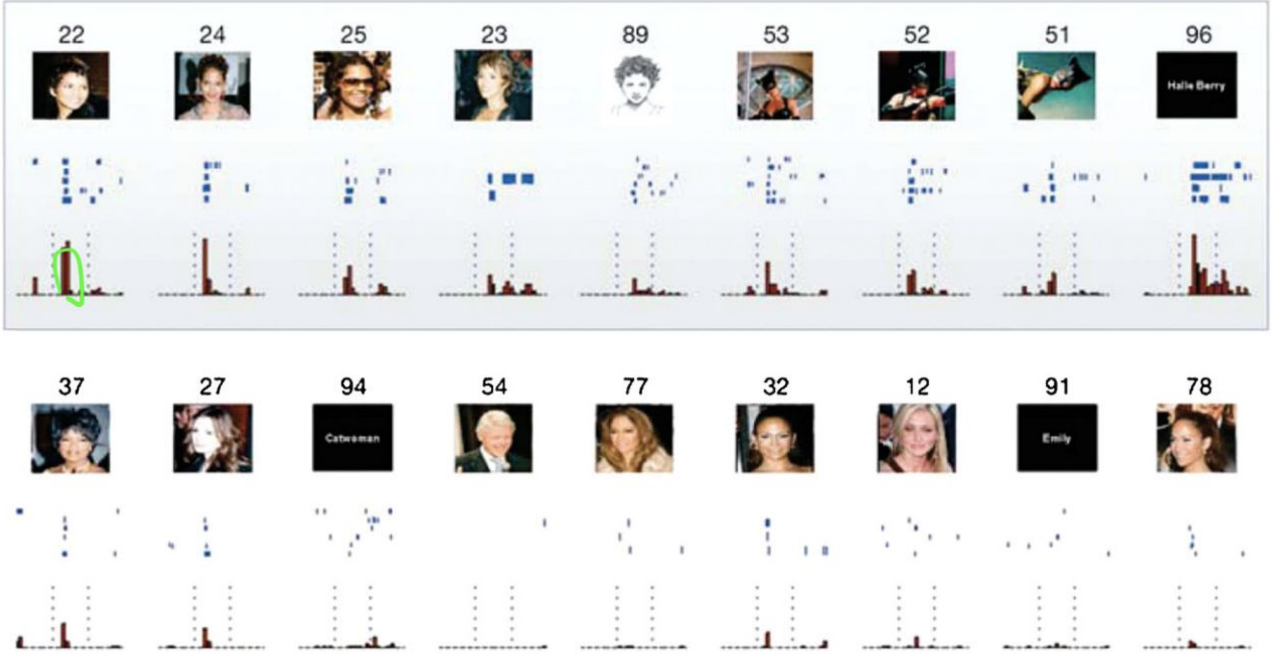


Fig. 8. A single unit in the human right anterior hippocampus that responds to different pictures of the actress Halle Berry including in costume and to the letter string of her name but not to other facial images or letter strings (Quiroga et al., 2005).

Agenda

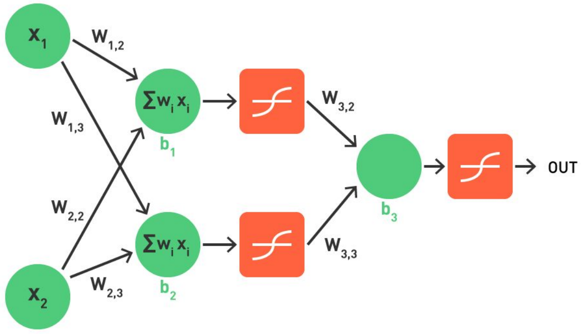
1. Check-in
2. NN recap
3. Convolutional Networks
4. Finish deep learning notebook
5. Brief SVM discussion
6. Supervised learning comparison

Neural Network Recap

1. What happens in forward propagation?
2. What happens in backprop?
3. What are the benefits of SGD and Mini-batches?
4. Why do GPUs speed up computation?
5. How do we handle regression problem?

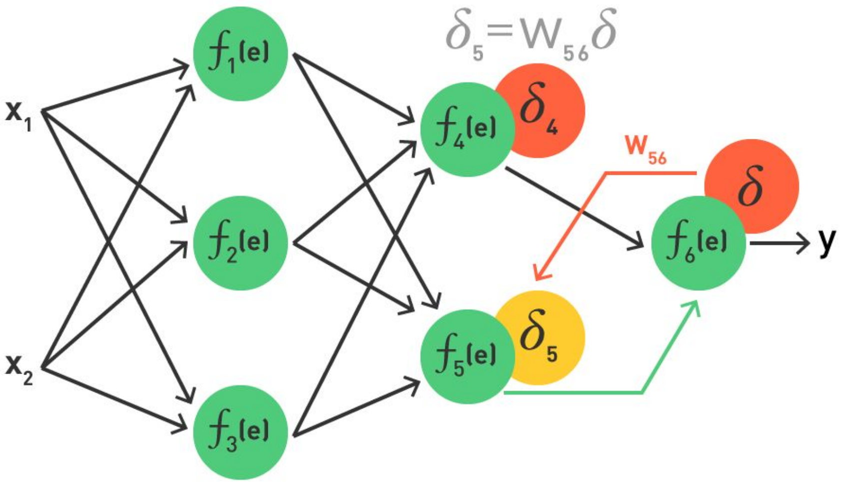
Intuition: Forward Propagation

- Given a training example (X_1, X_2) and output Y_i
- Propagate inputs/activations forward, applying sigmoid function on dot products



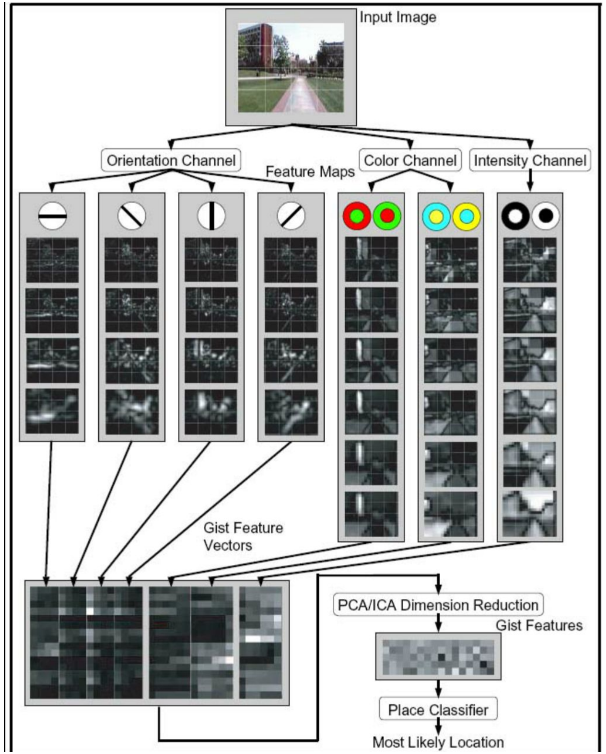
Intuition: Backward Propagation (cont.)

Propagate costs backward to earlier nodes:



datascience@berkeley

Computer Vision



Conferences

- CV is discussed at most ML and AI conferences
- CVPR is main CV conference

Datasets

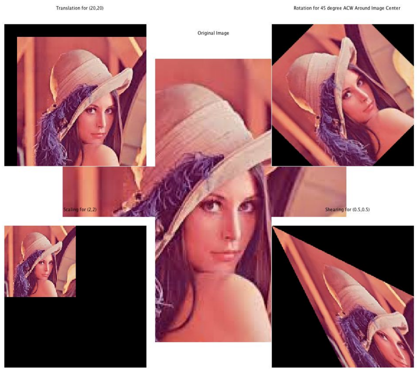
- MNIST -- 60,000 images
- SVHN (http://ufldl.stanford.edu/housenumbers/) -- 600,000 images
- ImageNet (http://image-net.org/about-stats) -- 14M images, 1TB, mapped to WordNet, includes features and hand labels

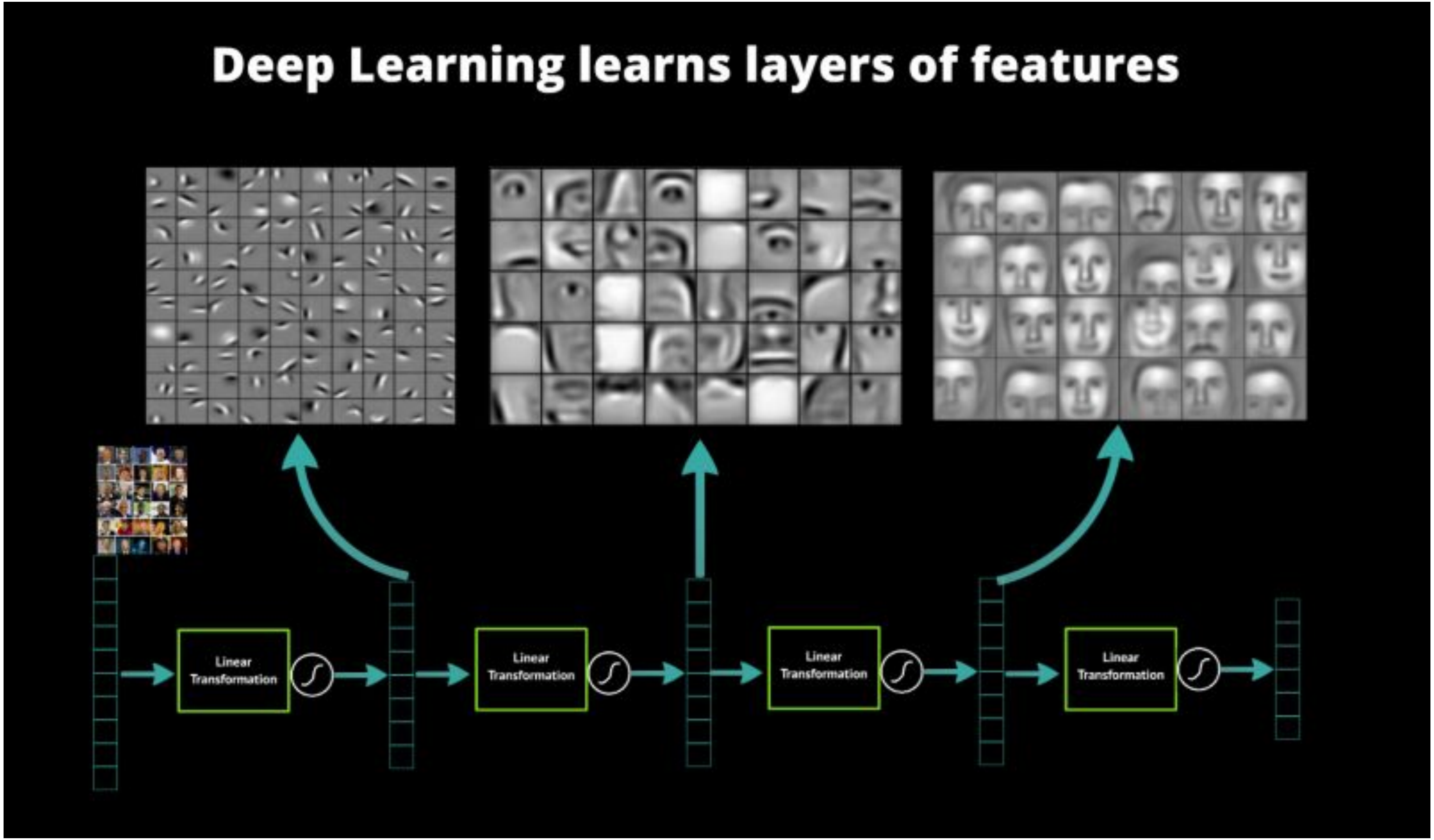
Feature Engineering

- A major focus of field
- Gradient based features popular
- SIFT: 1999, patented by BC.
- Also SURF, GIST, HOG

Engineering Examples

- Common in CV
- Do things that maintain label:
- Rotate, translate, skew, scale, etc




Deep Learning learns layers of features
Linear Transformation
Linear Transformation
Linear Transformation
Linear Transformation

Toothbrush

Small brush; has long handle; used to clean teeth

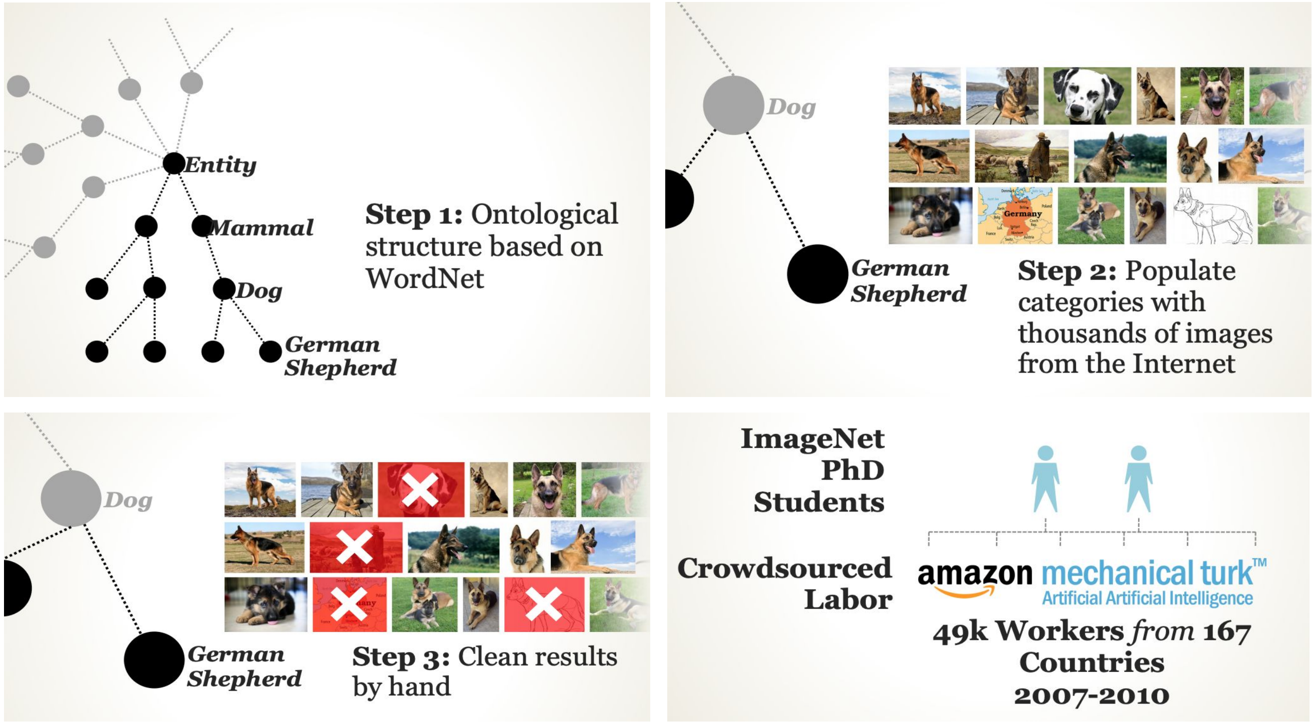
1974 pictures | 62.34% Popularity Percentile | Wordnet IDs

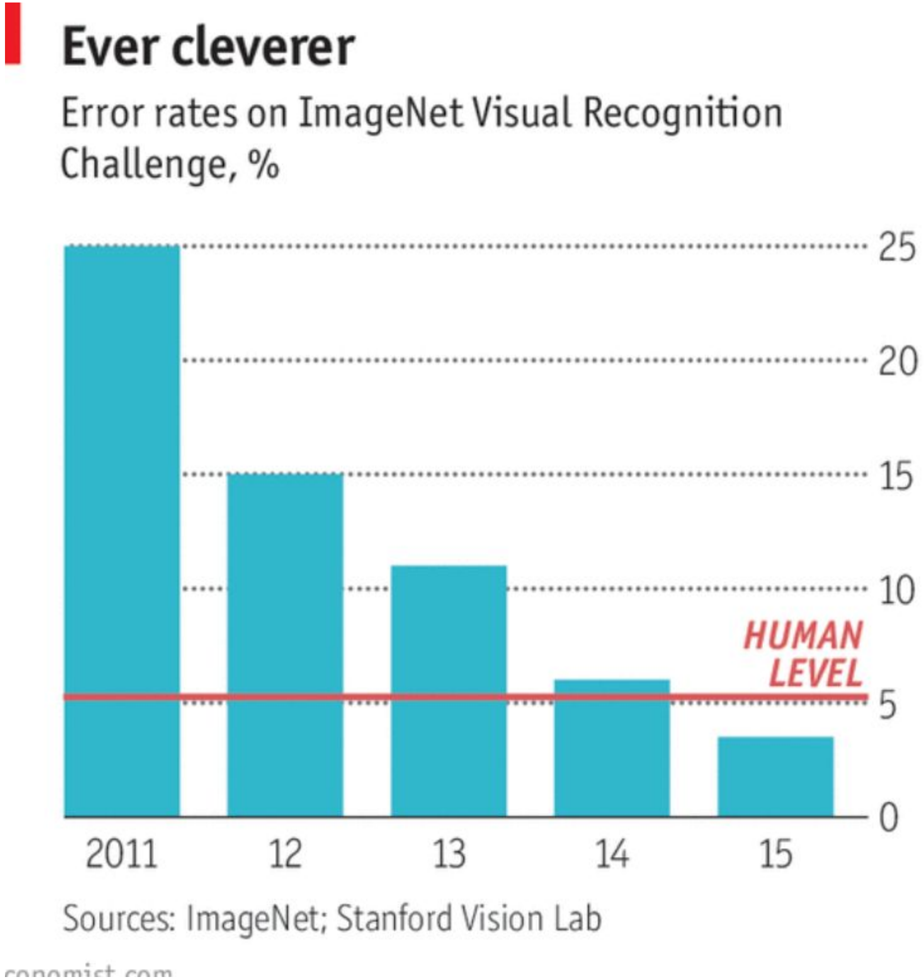
- spear, gig, fizgig, fishgig,
- swatter, flyswatter, flysw
- writing implement (18)
- beater (2)
- fire iron (3)
- needle (9)
- iron, branding iron (0)
- stick (41)
- bar (78)
- sports implement (11)
- container (744)
- hardware, ironware (0)
- equipment (479)
- ceramic (6)
- means (0)
- toiletry, toilet articles (57)
  - cream, ointment, emollier
  - hairdressing, hair tonic, ha
  - bath salts (0)
  - bath oil (0)
  - powder (7)
  - toothbrush (1)
    - electric toothbrush (0)
  - mousse, hair mousse, hair
  - perfume, essence (6)
  - cosmetic (17)
  - antiperspirant (0)
  - lotion (5)
  - hair spray (0)
  - shaving cream, shaving sc
  - shaving foam (0)
  - deodorant, deodourant (C
- conveyance, transport (566)

Treemap Visualization | Images of the Synset | Downloads

*Images of children synsets are not included. All images shown are thumbnails. Images may be subject to copyright.

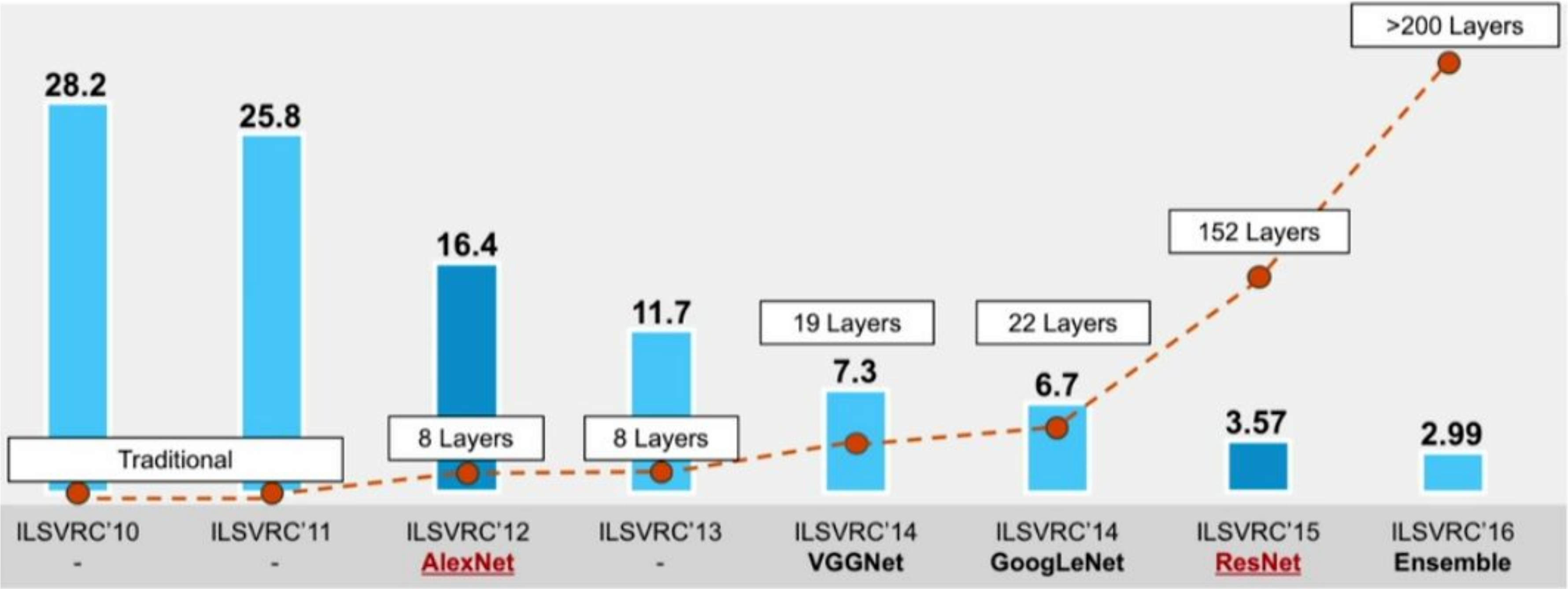
Prev 1 2 3 4 5 6 7 8 9 10 ... 84 85 Next


Entity
Mammal
Dog
German Shepherd
Step 1: Ontological structure based on WordNet
Dog
German Shepherd
Step 2: Populate categories with thousands of images from the Internet
Dog
German Shepherd
Step 3: Clean results by hand
ImageNet PhD Students
Crowdsourced Labor
amazon mechanical turk™
Artificial Artificial Intelligence
49k Workers from 167 Countries
2007-2010



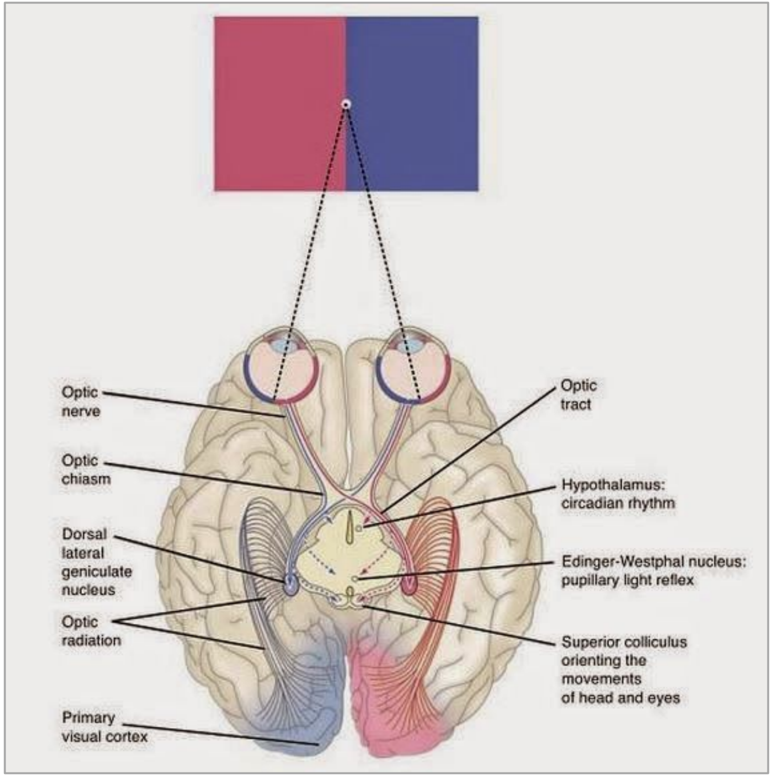
Deep (Feature/Representation) Learning

- Move away from feature engineering (still some and some Architechtural design)
- Today learned features generally outperform
- Learn similar gradient based features at early layers


>200 Layers
28.2
25.8
16.4
152 Layers
11.7
19 Layers
22 Layers
7.3
6.7
3.57
2.99
8 Layers
8 Layers
Traditional
ILSVRC'10
-
ILSVRC'11
-
ILSVRC'12
AlexNet
ILSVRC'13
-
ILSVRC'14
VGGNet
ILSVRC'14
GoogLeNet
ILSVRC'15
ResNet
ILSVRC'16
Ensemble

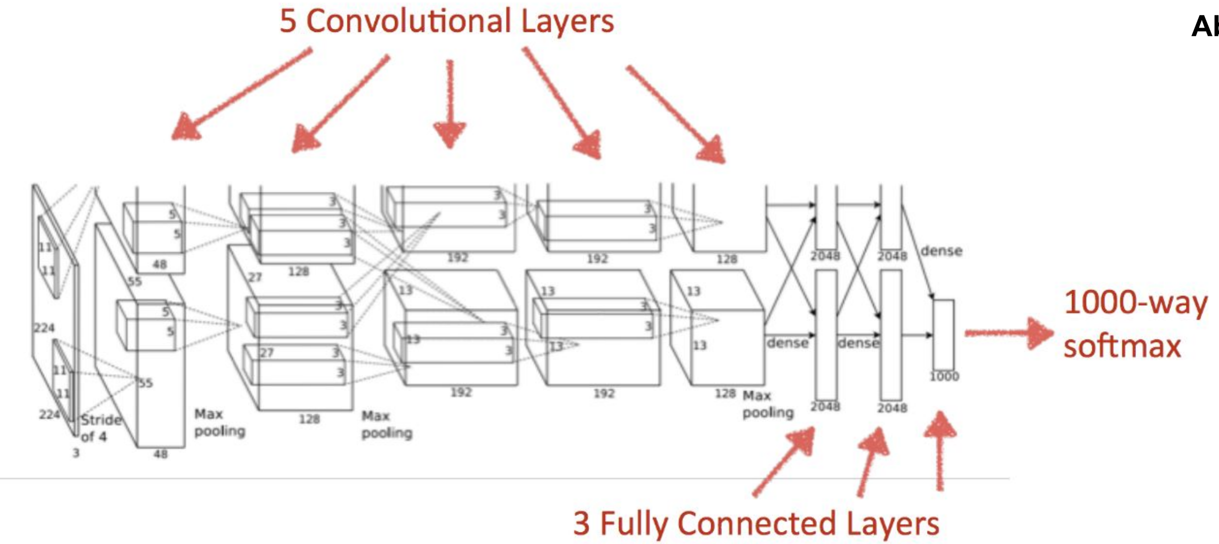
Convolution Nets

Learning vs Evolution vs Systems



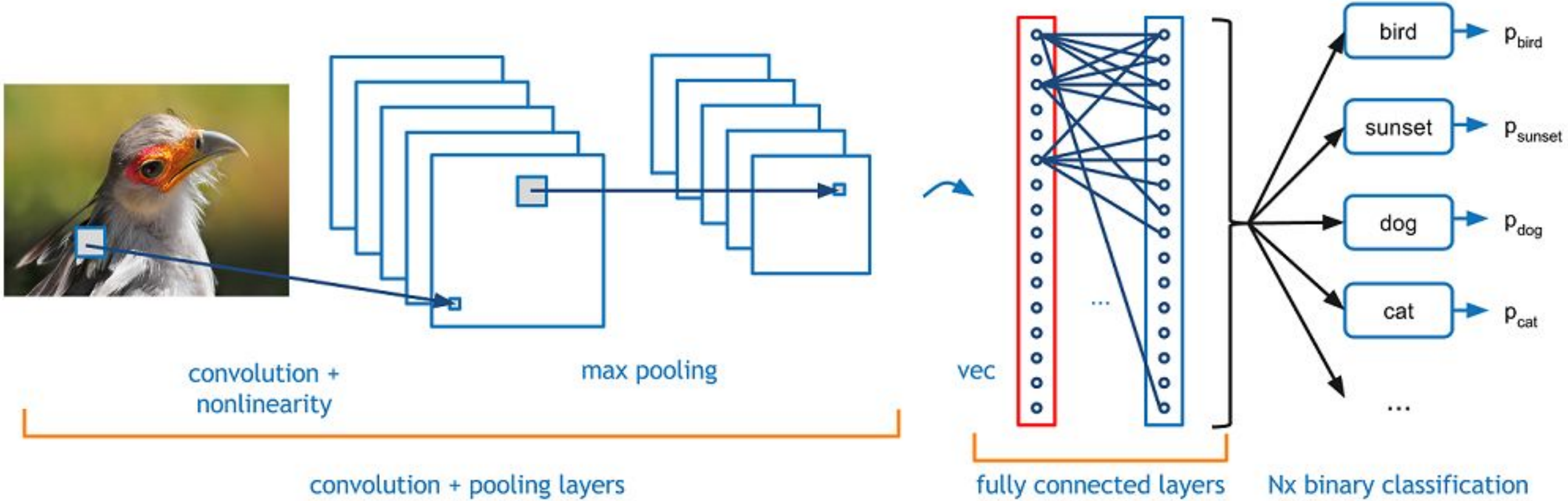
Convolution Networks

The convolutional and pooling layers in ConvNets are directly inspired by the classic notions of simple cells and complex cells in visual neuroscience[43], and the overall architecture is reminiscent of the LGN–V1–V2–V4–IT hierarchy in the visual cortex ventral pathway[44]. When ConvNet models and monkeys are shown the same picture, the activations of high-level units in the ConvNet explains half of the variance of random sets of 160 neurons in the monkey's inferotemporal cortex[45]. ConvNets have their roots in the neocognitron[46],

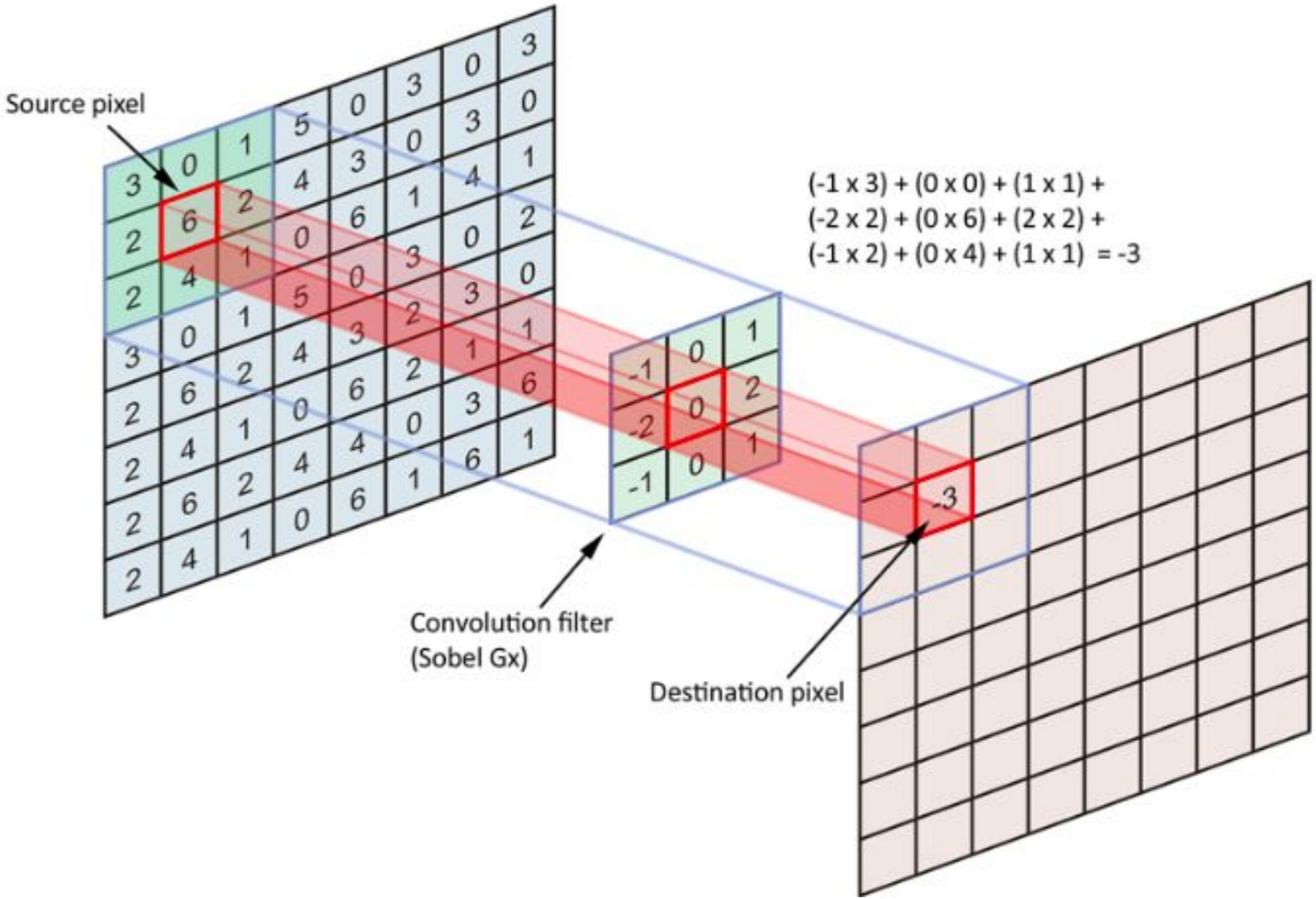


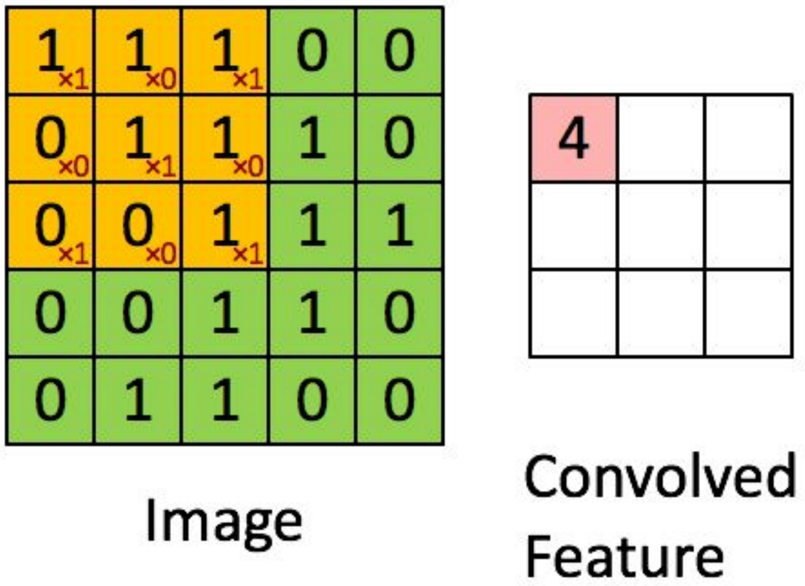
About

- Yann LeCun. LeNet: http://yann.lecun.com/exdb/lenet/ (1989-1998)
- Inspired by Visual Cortex in cats (receptive fields)
- Designed with image recognition in mind--input and layers often shown as 2D or 3D which may look odd coming from 1D.
- Composition of layers. Rightmost feature layers are most similar to output in representation
- Feature learning layers are of different types: (1) convolution and (2) pooling
- AlexNet 2012 (with Hinton)
- (https://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf)



convolution +
nonlinearity
max pooling
convolution + pooling layers
vec
...
fully connected layers
bird
sunset
dog
cat
...
p_{bird}
p_{sunset}
p_{dog}
p_{cat}
Nx binary classification


Source pixel
(-1 x 3) + (0 x 0) + (1 x 1) +
(-2 x 2) + (0 x 6) + (2 x 2) +
(-1 x 2) + (0 x 4) + (1 x 1) = -3
Convolution filter
(Sobel Gx)
Destination pixel


1 ×1
1 ×0
1 ×1
0
0
0 ×0
1 ×1
1 ×0
1
0
0 ×1
0 ×0
1 ×1
1
1
0
0
1
1
0
0
1
1
0
0
4
Image
Convolved
Feature

Visualization of the filter on the image

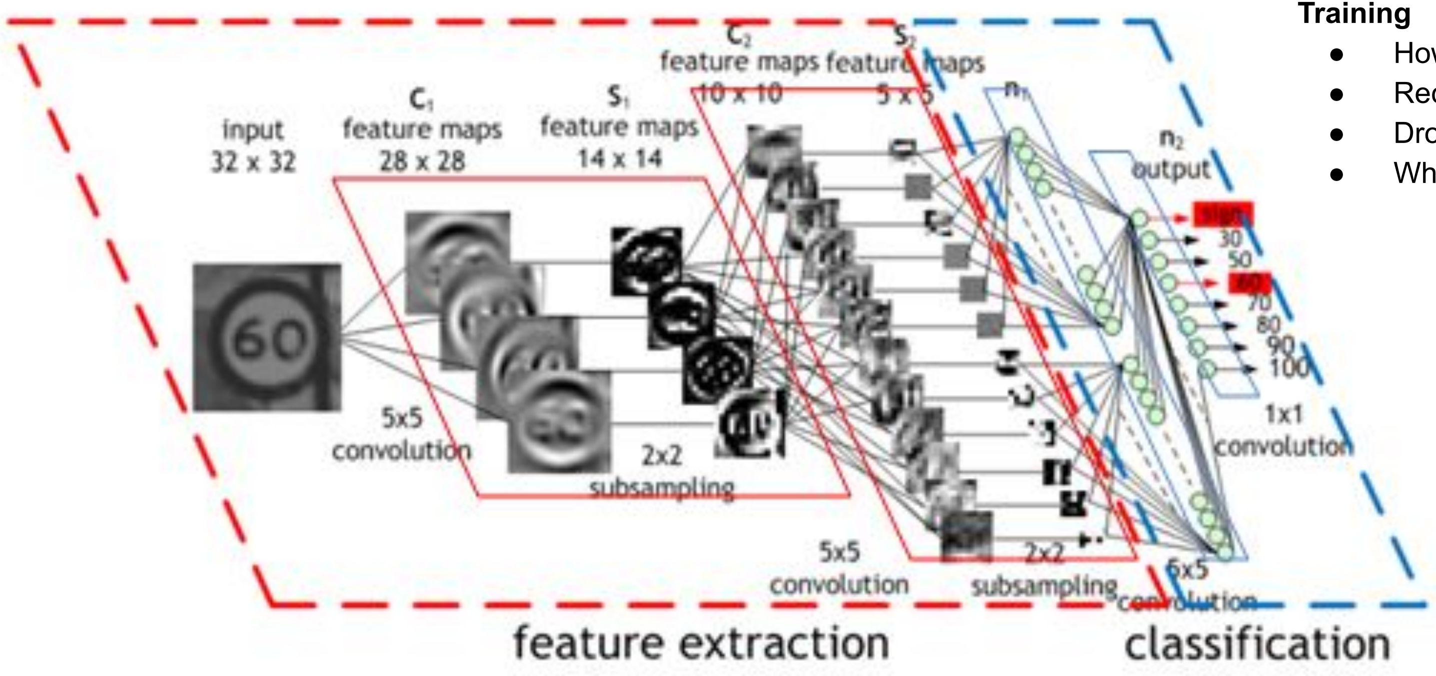
0	0	0	0	0	0	0
0	40	0	0	0	0	0
40	0	40	0	0	0	0
40	20	0	0	0	0	0
0	50	0	0	0	0	0
0	0	50	0	0	0	0
25	25	0	50	0	0	0

Pixel representation of receptive field

*

0	0	0	0	0	30	0
0	0	0	0	30	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	0	0	0	0

Pixel representation of filter

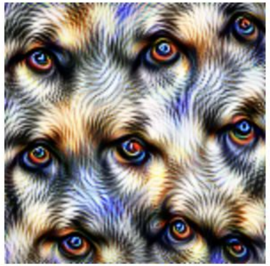


Training

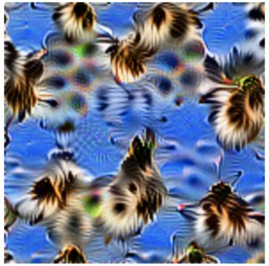
- How many parameters to learn?
- Rectifier activation
- Dropouts
- Why do you think this works well?

Layer 4a


Bookshelves


Dog eyes

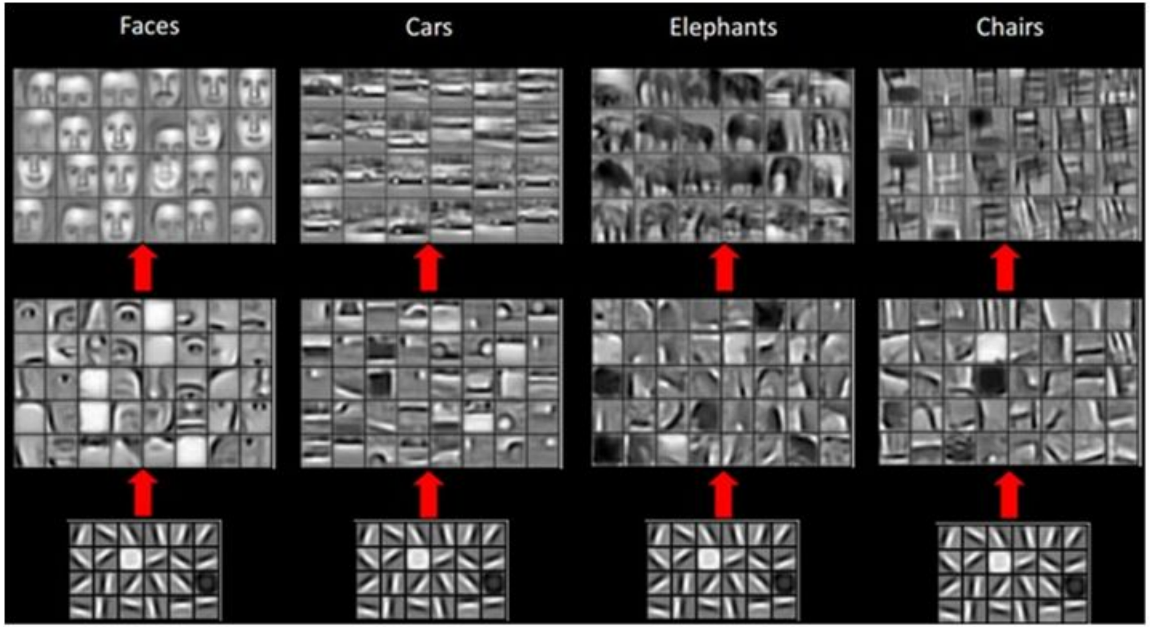

Text, rivets


Birds

In this layer, which follows a pooling step, we see a signficant increase in complexity. We begin to see more complex patterns, and even parts of objects.

Understanding Deep Networks

- Feature visualization: https://distill.pub/2017/feature-visualization/
- Fear and Loathing in LV: https://www.youtube.com/watch?v=oyxSerkkP4o

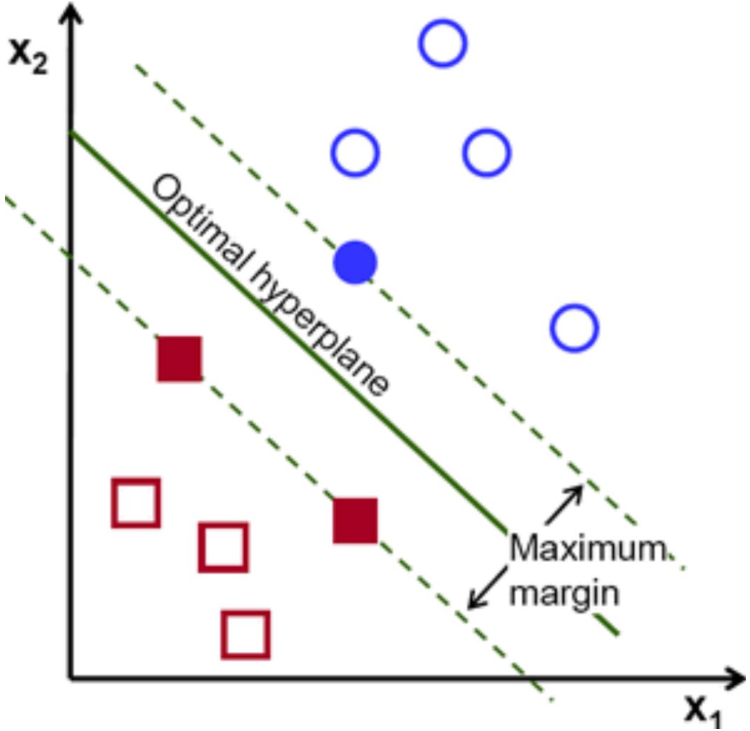


Transfer Learning

- Train on one task, and use trained network or part of trained network when training for a different task
- Model Zoo (e.g. http://caffe.berkeleyvision.org/model_zoo.html)
- https://www.kaggle.com/c/state-farm-distracted-driver-detection/forums/t/20141/official-pre-trained-models-and-external-data-thread/116805

Super Brief Look at SVMs

SVM



Implementations

- LIBSVM/Liblinear -- National Taiwan; used in SK_Learm, e1071, Matlab
- SVMLight/SVMPerf -- Cornell

People

- Vapnik (AT&T, FB)
- Yann LeCun (AT&T, FB)
- Yoshua Bengio (AT&T, Montreal)
- Leon Bottou (AT&T, Google)
- Christopher Bishop (Edinburg, MS)
- Chris Burgess (AT&T, MS)
- Patrick Haffner (AT&T)

Review

- http://www.tristanfletcher.co.uk/SVM%20Explained.pdf
- http://svmlight.joachims.org/

Supervised Learning Comparison

Supervised learning

Training Complexity
KNN: none
NB: 1 epoch
DT: (max depth) epoch
LR: until convergence
NN: until convergence

Prediction Complexity
KNN: 1 epoch
NB: constant
DT: constant
LR: constant
NN: layers * avg nodes

Representation
KNN: store train data
NB: counts
DT: tree
LR: parms
NN: parms + architechture

Pros/Cons & When to Consider

KNN:

- Pros: Simple / explainable, Works with small data, online
- Cons: Storage, slow, No feature weighting

NB:

- Pros: Online, fast train/predict, scales to big col/rows
- Cons: strong ind assumption

DT:

- Pros: Feature selection, explainable, accurate as ensembles, fast train/predict
- Cons: Store trees, single tree can overfit

LR:

- Pros: Fast, explainable, widely useful, scales well
- Cons: linear, can take a lot of data

NN:

- Pros: accurate, nonlinear, unstructured/high dim data
- Cons: Expensive, black box, gpus, lots of data

Final Thoughts?