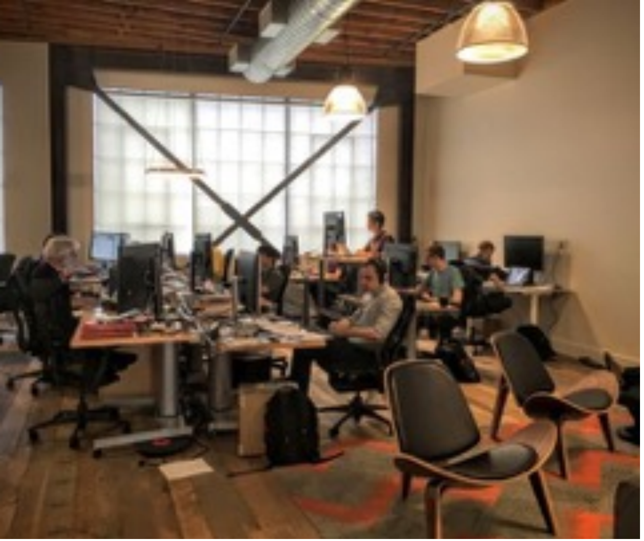
Introductions



OpenAI

CrowdFlower

Google

Bloomberg

TESLA

UBER

WEIGHTS & BIASES

W&B

High Quality Tools for Machine Learning

Dolores Labs

Class Goals

1) Concrete Practical Understanding of Machine Learning/Deep Learning
2) Feel What It's Really Like to Do Machine Learning
3) Put You On Fastest Possible Path to Becoming Proficient in Machine Learning

Agenda

9:00 - 10:00	Introduction, Machine Learning Overview
10:00 - 11:00	Build a text classifier, feature extraction.
11:00 - 12:00	Build a better text classifier, model evaluation.
12:00 - 1:00	Lunch
1:00 - 2:00	Neural Network Overview
2:00 - 3:00	Build a digit classifier, Multi-layer perceptrons
3:00 - 4:00	Build a better digit classifier, Convolutional Neural Networks
4:00 - 5:00	Transfer Learning. Next steps.

Module 1 Begin

Goals:

High level overview of Machine Learning
Look at our data/check software

Questions:

What is machine learning?
What kinds of problems are useful for machine learning?

Introduction to Machine Learning

5 kinds of questions

1. How much / how many?
2. Which category?
3. Which groups?
4. Is it weird?
5. Which action?

How much / how many? (Regression)

Age	Zip Code	Income
58	02138	$95,824
73	94110	$20,708
59	45323	$82,152
66	34134	$25,334

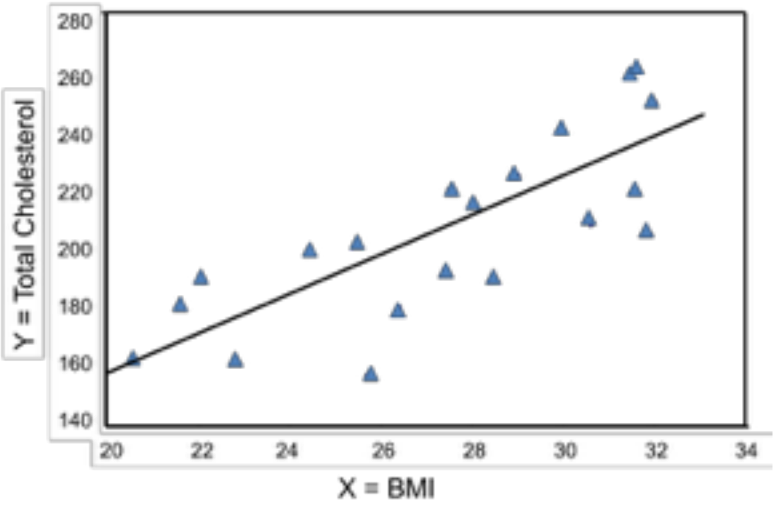
Age	Zip Code	Income
73	01233	
61	34134	
47	92349	
44	81112	

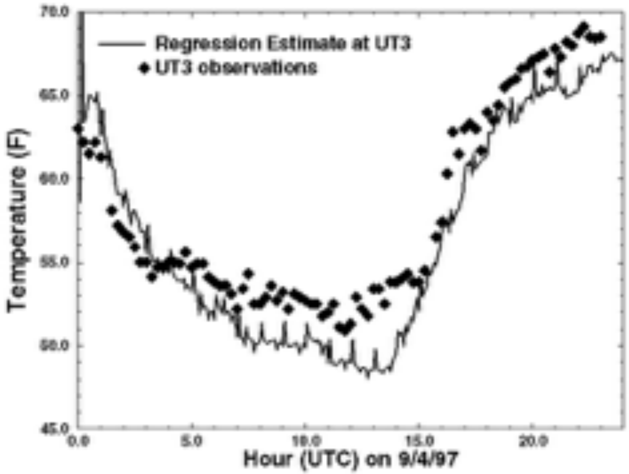
Training Data

Test Data

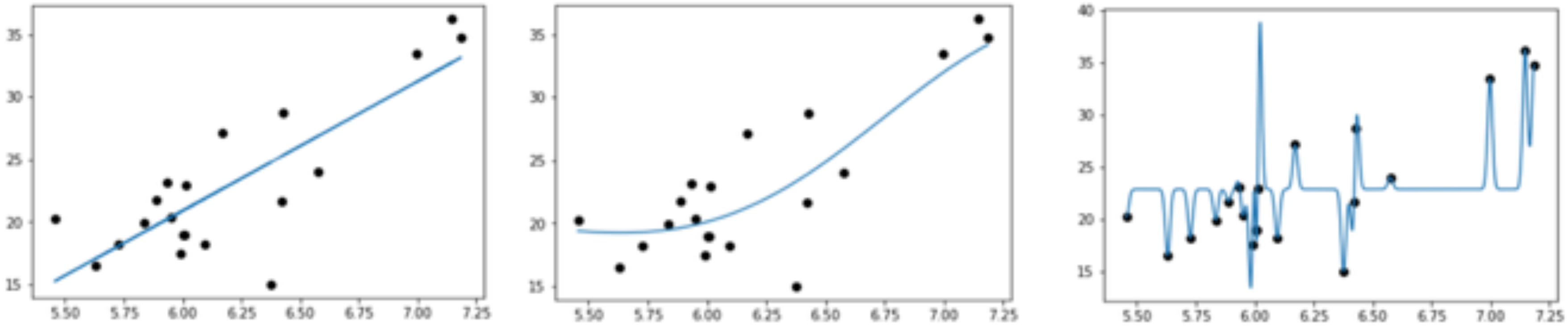
How much / how many? (Regression)

1. What temperature will it be next Tuesday?
2. How promising is this sales lead?
3. How many Twitter followers will I have by the end of the year?





Regression



Which category? (Classification)

Age	Income	Default
58	$95,824	True
73	$20,708	False
59	$82,152	False
66	$25,334	True

Age	Income	Default
73	$53,445	
61	$36,679	
47	$90,422	
44	$79,040	

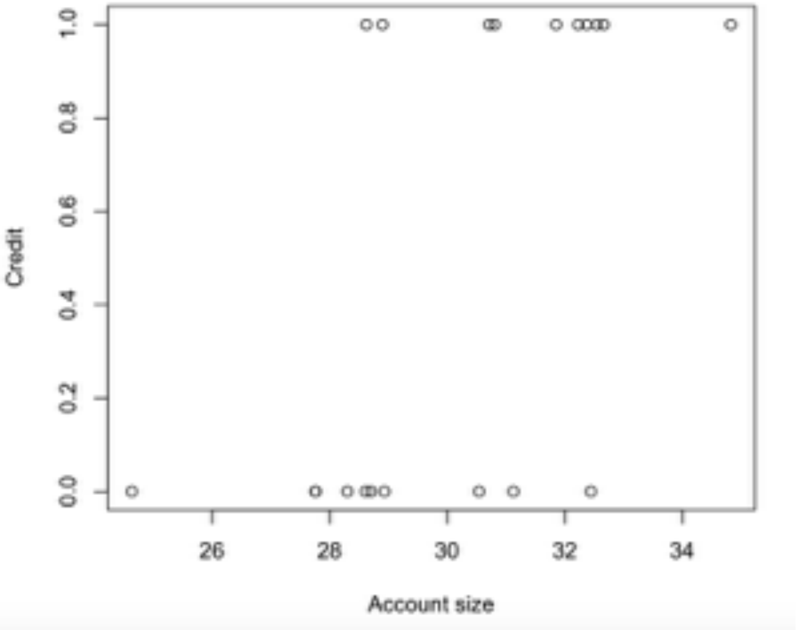
Training Data

Test Data

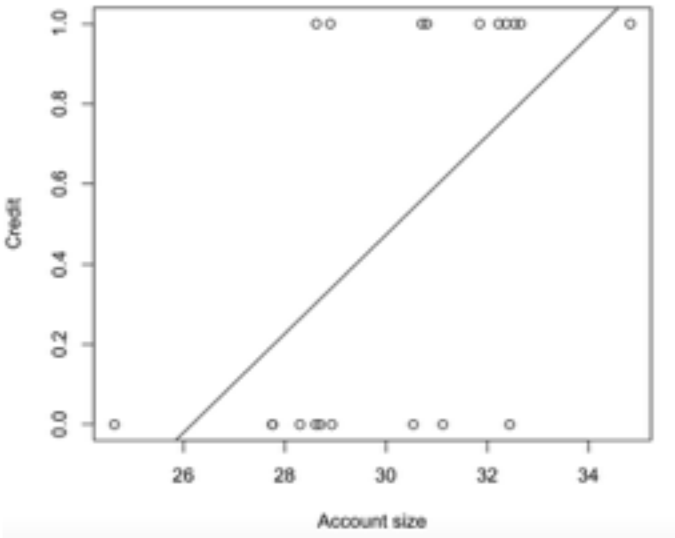
Which category? (Classification)

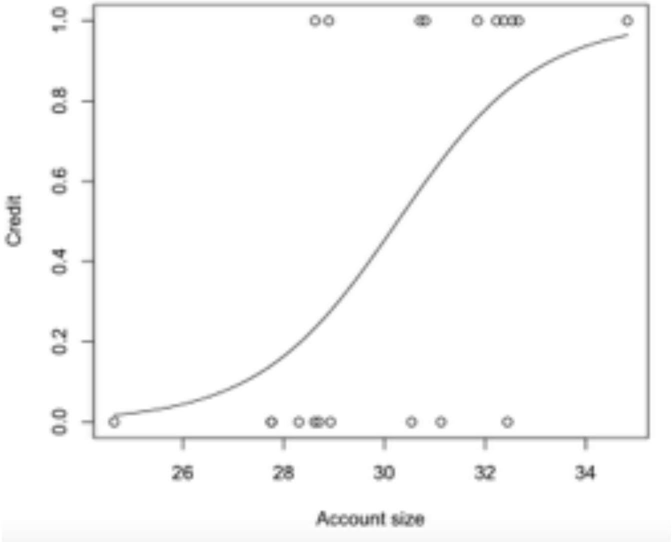
1. Is this a picture of a cat or a dog?
2. Is this email message spam or not?
3. What is the topic of this news article?

Classification as Regression

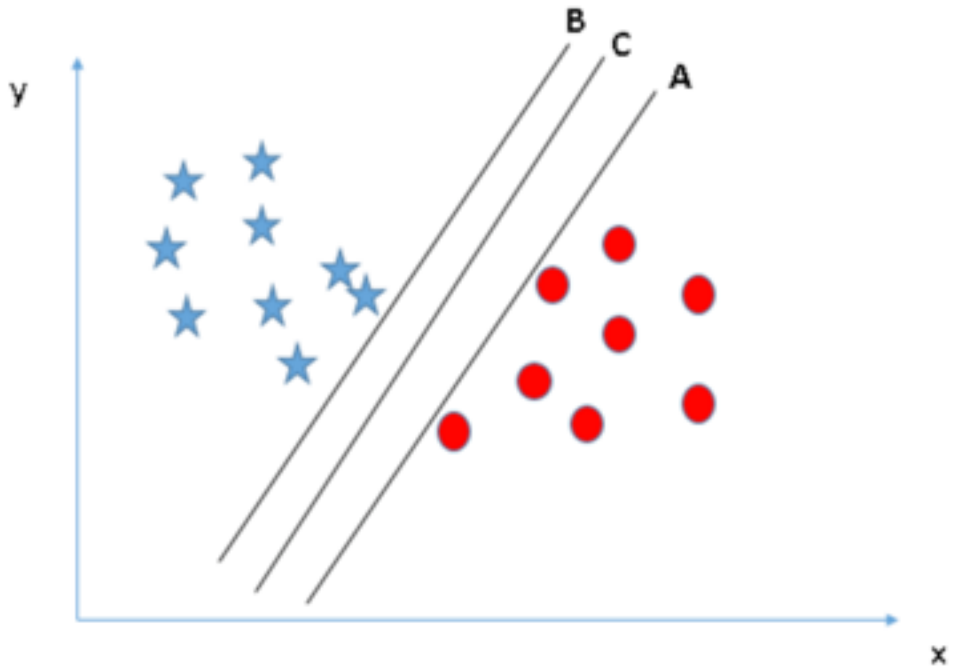


Classification as Regression



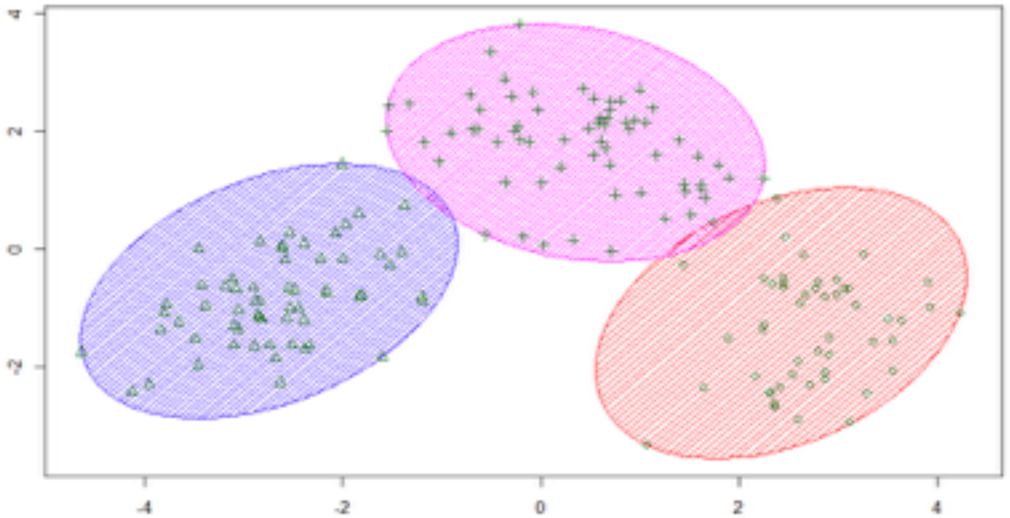


Another way to draw classification



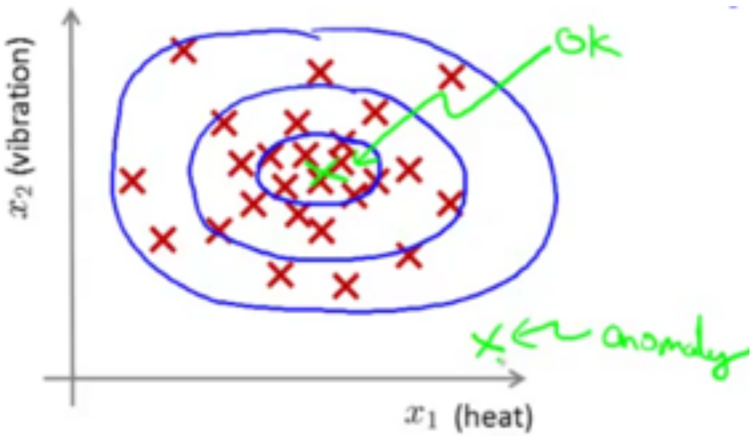
Which groups? (Clustering)

1. Which customers have similar shopping preferences?
2. Which of these images look similar?
3. How can I group these documents together by topic?



Is it weird? (Anomaly detection)

1. Is this sensor reading unusually high?
2. Is someone trying to hack into my system?
3. Does this credit card transaction appear fraudulent?

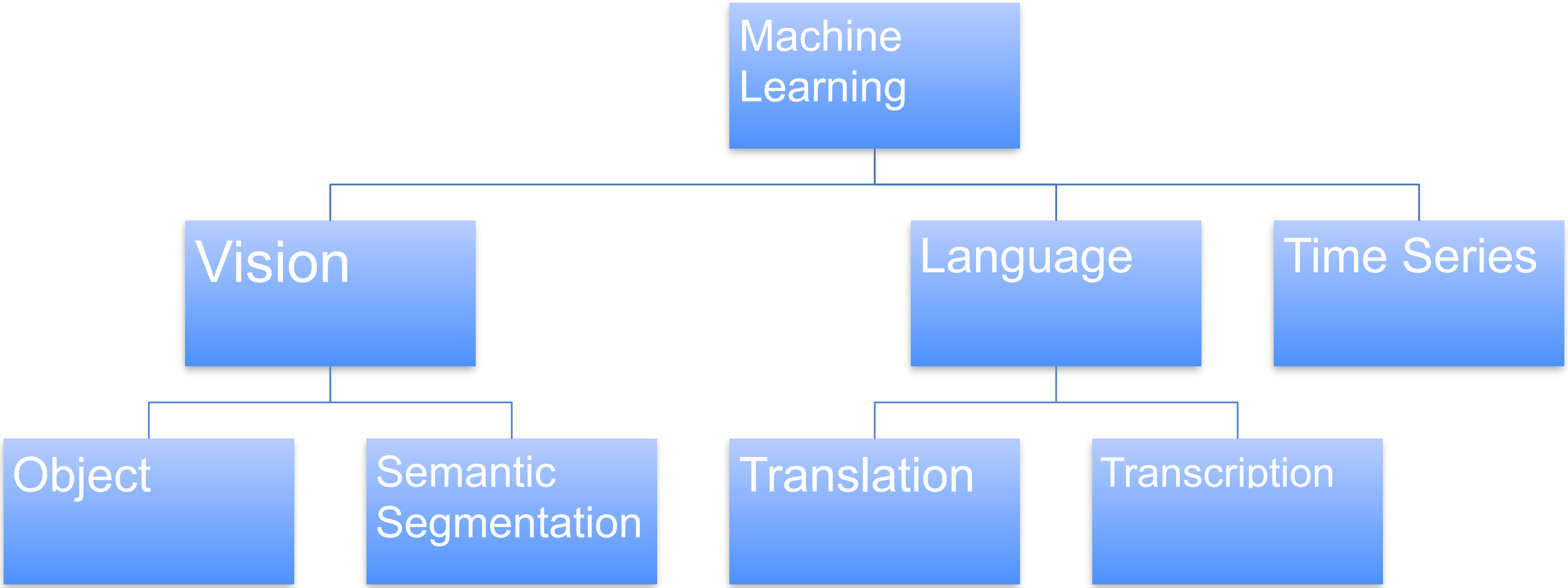


Which action? (Reinforcement Learning)

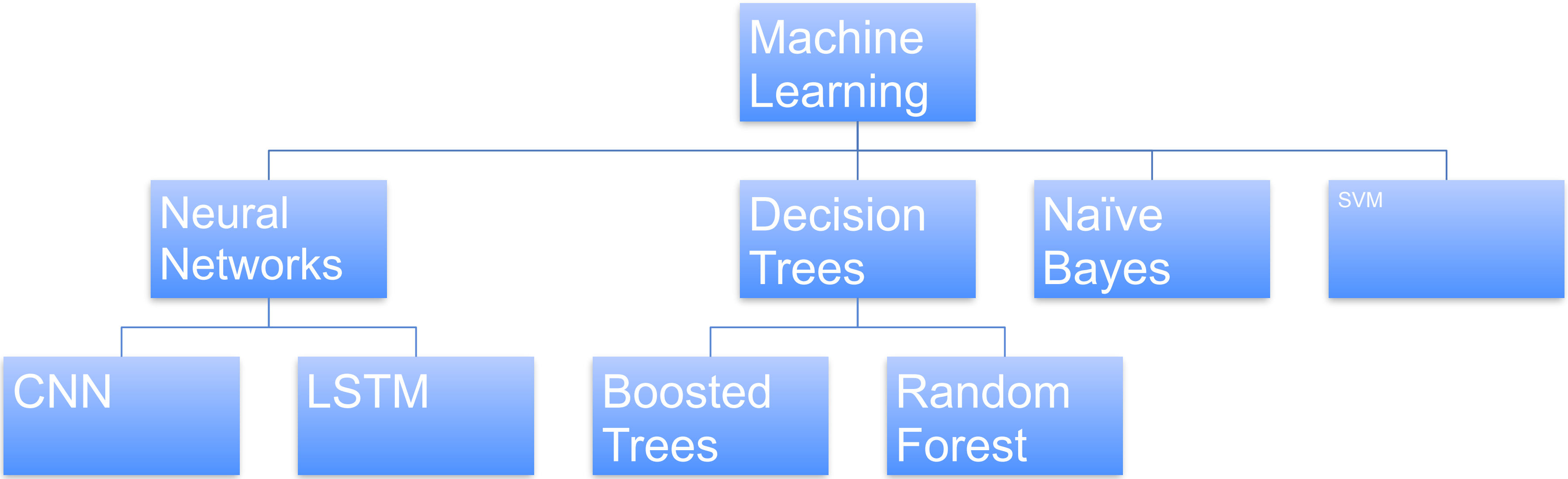
1. Should I speed up or stop at this yellow light?
2. Should I raise or lower the temperature?
3. Do I continue vacuuming or do I return to my charging station?

Supervised vs Unsupervised Learning

Machine Learning by Application



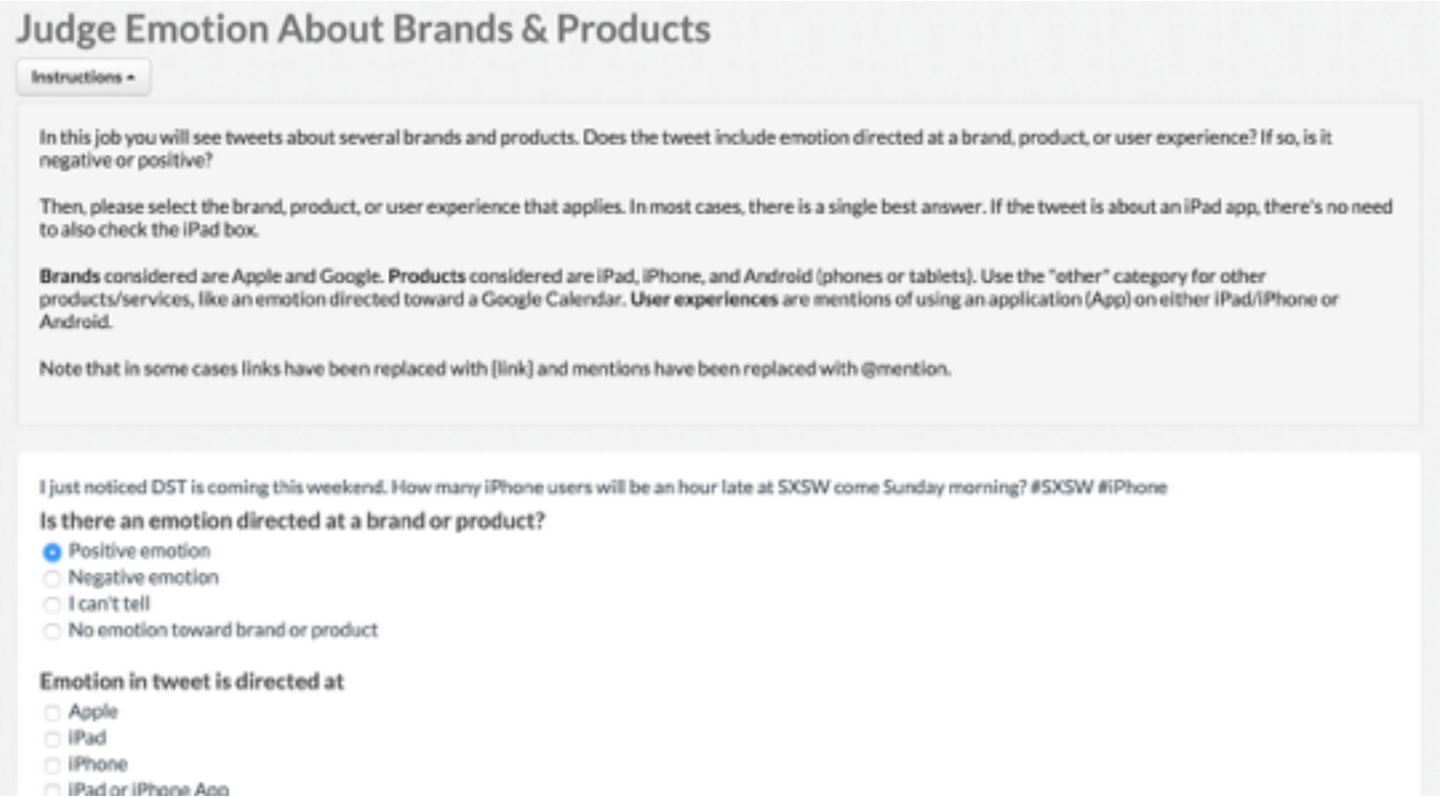
Machine Learning by Technique



Three Steps in all Supervised Machine Learning

Loading and Exploring Data
(The most important topic in ML/DL that no one ever teaches)

Project for Today



Judge Emotion About Brands & Products

Instructions ▴

In this job you will see tweets about several brands and products. Does the tweet include emotion directed at a brand, product, or user experience? If so, is it negative or positive?

Then, please select the brand, product, or user experience that applies. In most cases, there is a single best answer. If the tweet is about an iPad app, there's no need to also check the iPad box.

Brands considered are Apple and Google. **Products** considered are iPad, iPhone, and Android (phones or tablets). Use the "other" category for other products/services, like an emotion directed toward a Google Calendar. **User experiences** are mentions of using an application (App) on either iPad/iPhone or Android.

Note that in some cases links have been replaced with [link] and mentions have been replaced with @mention.

I just noticed DST is coming this weekend. How many iPhone users will be an hour late at SXSW come Sunday morning? #SXSW #iPhone

Is there an emotion directed at a brand or product?

- (•) Positive emotion
- ( ) Negative emotion
- ( ) I can't tell
- ( ) No emotion toward brand or product

Emotion in tweet is directed at

- [ ] Apple
- [ ] iPad
- [ ] iPhone
- [ ] iPad or iPhone App

Look at your data!

A	B	C	D	E	F	G
tweet_text	emotion_in_tweet_is_directed_at	is_there_an_emotion_directed_at_a_brand_or_product				
.@wesley83 I have a 3G iPhone. After 3 hrs tweeting at #RISE_Austin, it was dead! I need tc	iPhone	Negative emotion				
@jessedee Know about @fludapp ? Awesome iPad/iPhone app that you'll likely appreciate f	iPad or iPhone App	Positive emotion				
@swonderlin Can not wait for #iPad 2 also. They should sale them down at #SXSW.	iPad	Positive emotion				
@sxsw I hope this year's festival isn't as crashy as this year's iPhone app. #sxsw	iPad or iPhone App	Negative emotion				
@sxtxstate great stuff on Fri #SXSW: Marissa Mayer (Google), Tim O'Reilly (tech books/conf	Google	Positive emotion				
@teachntech00 New iPad Apps For #SpeechTherapy And Communication Are Showcased At The #SXSW Conference http://ht.ly/49n4M #iear #edchat		No emotion toward brand or product				
		No emotion toward brand or product				
#SXSW is just starting, #CTIA is around the corner and #googleio is only a hop skip and a jum	Android	Positive emotion				
Beautifully smart and simple idea RT @madebymany @thenextweb wrote about our #holler	iPad or iPhone App	Positive emotion				
Counting down the days to #sxsw plus strong Canadian dollar means stock up on Apple gear	Apple	Positive emotion				
Excited to meet the @samsungmobileus at #sxsw so I can show them my Sprint Galaxy S stil	Android	Positive emotion				
Find & Start Impromptu Parties at #SXSW With @HurricaneParty http://bit.ly/gVLrIn I c	Android App	Positive emotion				
Foursquare ups the game, just in time for #SXSW http://j.mp/grN7pK) - Still prefer @Gowall	Android App	Positive emotion				
Gotta love this #SXSW Google Calendar featuring top parties/ show cases to check out. RT @	Other Google product or service	Positive emotion				
Great #sxsw ipad app from @madebymany: http://tinyurl.com/4nqv92l	iPad or iPhone App	Positive emotion				
haha, awesomely rad iPad app by @madebymany http://bit.ly/hTdFim #hollergram #sxsw	iPad or iPhone App	Positive emotion				
Holler Gram for iPad on the iTunes App Store - http://t.co/kfN3f5Q (via @marc_is_ken) #sxsw		No emotion toward brand or product				
I just noticed DST is coming this weekend. How many iPhone users will be an hour late at SX	iPhone	Negative emotion				
Just added my #SXSW flights to @planely. Matching people on planes/airports. Also downlo	iPad or iPhone App	Positive emotion				
Must have #SXSW app! RT @malbonster: Lovely review from Forbes for our SXSW iPad app	iPad or iPhone App	Positive emotion				
Need to buy an iPad2 while I'm in Austin at #sxsw. Not sure if I'll need to Q up at an Austin A	iPad	Positive emotion				
Oh. My. God. The #SXSW app for iPad is pure, unadulterated awesome. It's easier to browse	iPad or iPhone App	Positive emotion				
Okay, this is really it: yay new @Foursquare for #Android app!!!!11 kthxbai. #sxsw	Android App	Positive emotion				
Photo: Just installed the #SXSW iPhone app, which is really nice! http://tumblr.com/x6t1pi6	iPad or iPhone App	Positive emotion				
Really enjoying the changes in Gowalla 3.0 for Android! Looking forward to seeing what else	Android App	Positive emotion				
RT @LaurieShook: I'm looking forward to the #SMCDallas pre #SXSW party Wed., and hopin	iPad	Positive emotion				
RT haha, awesomely rad iPad app by @madebymany http://bit.ly/hTdFim #hollergram #sxsw	iPad or iPhone App	Positive emotion				
someone started an #austin @PartnerHub group in google groups, pre-#sxsw. great idea	Other Google product or service	Positive emotion				
The new #4sq3 looks like it is going to rock. Update for iPhone and Android should push ton	iPad or iPhone App	Positive emotion				
They were right, the @gowalla 3 app on #android is sweeeeet! Nice job by the team there. #	Android App	Positive emotion				
Very smart from @madebymany #hollergram iPad app for #sxsw! http://t.co/A3xvWc6 (ma	iPad or iPhone App	Positive emotion				
You must have this app for your iPad if you are going to #SXSW http://itunes.apple.com/us/	iPad or iPhone App	Positive emotion				
Attn: All #SXSW frineds, @mention Register for #GDGTLive and see Cobra iRadar for Android. {link}		No emotion toward brand or product				
Anyone at #sxsw want to sell their old iPad?		No emotion toward brand or product				
Anyone at #SXSW who bought the new iPad want to sell their older iPad to me?		No emotion toward brand or product				
At #sxsw. Oooh. RT @mention Google to Launch Major New Social Network Called Circles. Possibly Today {link}		No emotion toward brand or product				

load-data.py

Questions:
How many rows in our dataset?
What is “text” and what is “target”?

How do we turn the text into numbers?

	A	aardvark	…	hate	I	iphone	love	my	…	Zyzzyva
I love my iphone	0	0	…	0	1	1	1	1	…	0
I hate my iphone	0	0	…	1	1	1	0	1	…	0

http://www.python-course.eu/text_classification_python.php

Module 1 End

Goals:
Look at our data/check software
High level overview of Machine Learning

Questions:
What is machine learning?
What kinds of problems are useful for machine learning?

Common Pitfalls:
Hard to research machine learning problems online

Module 2 Begin

Goals:
Build a Machine Learning text classifier

Questions:
What is feature extraction?
Why is feature extraction so important?

feature-extraction-1.py

Questions:
What happened?

feature-extraction-2.py

Questions:
How many columns did we create?
Are the words in alphabetical order?
What happened to capitalization?

feature-extraction-3.py

Questions:
What is the difference between count_vect.fit and count_vect.transform?
What happens if we try to do transform on a word we haven’t seen before?

Takeaways from Feature Selection:

What is feature selection?
Why is it so important?
Do we think the classifier is good?

Module 3 Begin

Goals:

Improve our Machine Learning Classifier

Questions:

How do we evaluate models?

Why is it so important to evaluate models?

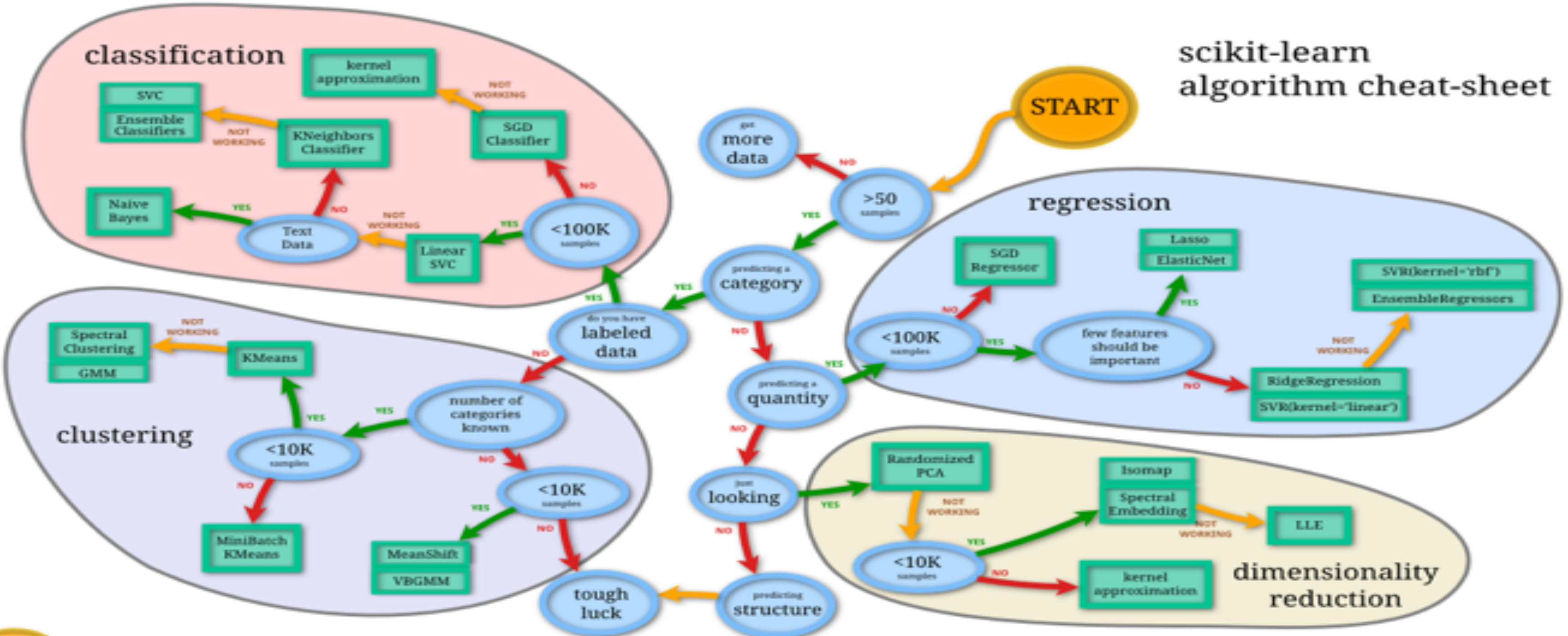
What are the common machine learning algorithms?

Lots of important choices already!

- ```
  class
  sklearn.feature_extraction.text.CountVectorizer(input=u'content',
  encoding=u'utf-8', decode_error=u'strict', strip_accents=None,
  lowercase=True, preprocessor=None, tokenizer=None, stop_words=None,
  token_pattern=u'(?u)\b\w\w+\b', ngram_range=(1, 1), analyzer=u'word',
  max_df=1.0, min_df=1, max_features=None, vocabulary=None,
  binary=False, dtype=<type 'numpy.int64'>)
  ```

- Should we remove really rare words?
- Should we remove really common words?
- Should we remove "stop words"?
- Should we lower case all the words?
- What is a word?
  - For those at #SXSW: Apple sets up 5,000-square-foot temporary store at SXSW to sell new iPads, test potential traffic
  - If ur not at the #google #aclu 80's party....u should be! #sxsw
  - My iPhone battery at 100%. #winning at #SXSW
  ```

Choose an algorithm



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

classifier.py

Questions:

Can we find some examples that confuse the classifier?

What kinds of input does np.predict take?

Takeaways from Feature Selection:

What is feature selection?
Why is it so important?

Evaluating Models

Lots of scary choices already!

- MultinomialNB vs GaussianNB vs BernoulliNB

Attributes:	**class_prior_** : array, shape (n_classes,) probability of each class. **class_count_** : array, shape (n_classes,) number of training samples observed in each class. **theta_** : array, shape (n_classes, n_features) mean of each feature per class **sigma_** : array, shape (n_classes, n_features) variance of each feature per class

- Do I want to weight one class more than the other?
- Do I want to monkey around with the algorithm?

test-algorithm-1.py

Questions:

Do we believe this probability?

test-algorithm-2.py

Questions:

Why are we doing this?

What would happen if we increased the proportion of training data?

What would happen if we decreased the proportion of training data?

Why wasn't the last classifier perfect?

test-algorithm-3.py

Questions:

What patterns do you see in the confusion matrix?
Are all the labels equally represented?
Is the classifier equally accurate on all the labels?
Do you see a relationship between label representation and classifier predictions?

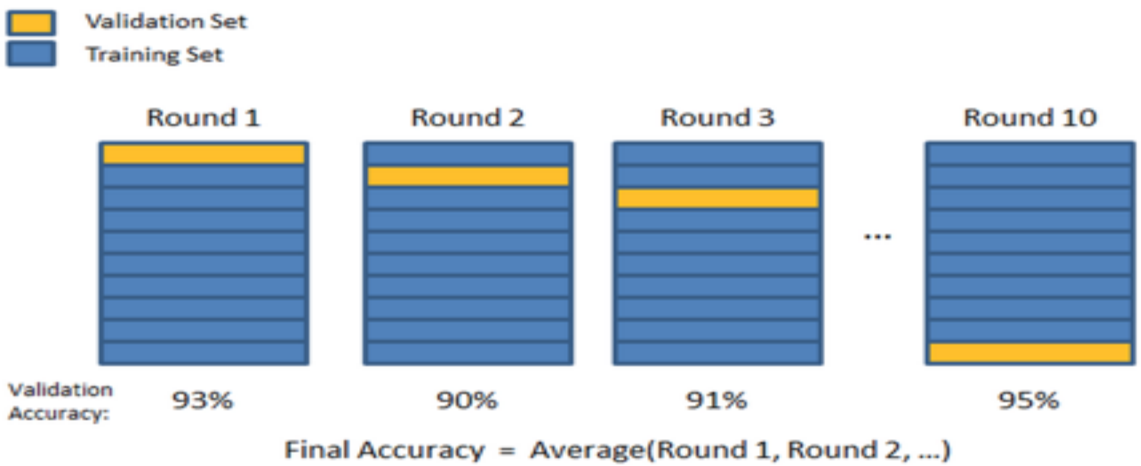
test-algorithm-dummy.py

Questions:

What is the dummy classifier’s strategy?

Do we now think our model is good?

Cross Validation



https://chrisjmccormick.wordpress.com/2013/07/31/k-fold-cross-validation-with-matlab-code/

test-algorithm-cross-validation.py

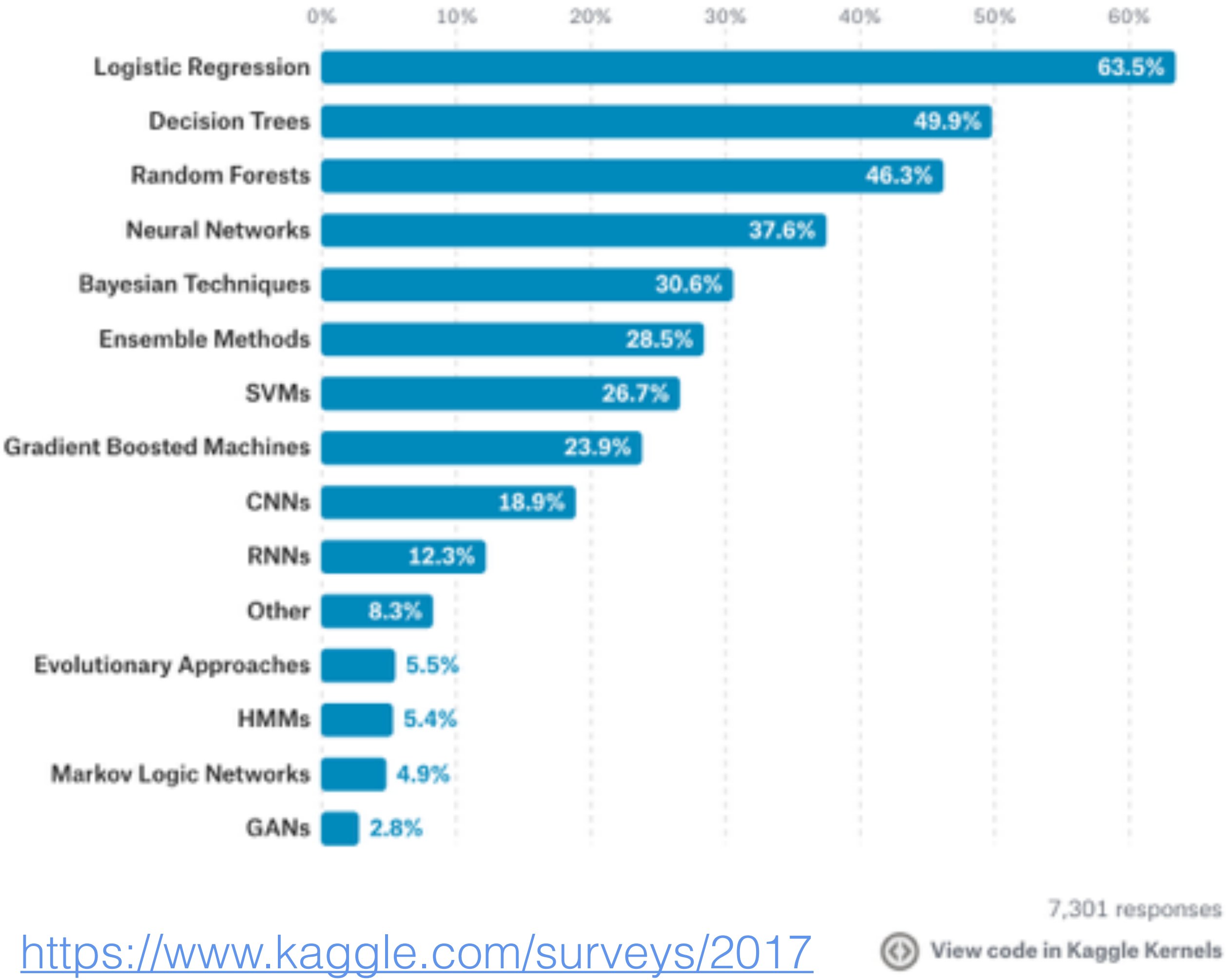
Questions:
Is our model better than a baseline?

Takeaways from Model Evaluation:

Why shouldn’t you test on training data?
Why are dummy classifiers important?
What is Cross Validation?
How should I pick a good algorithm?

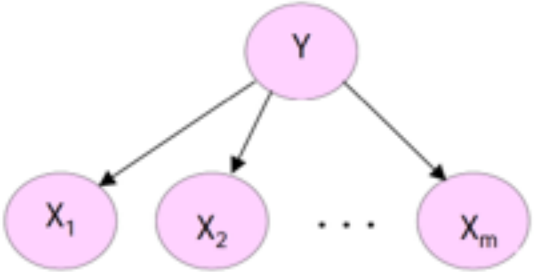
Model Building/Iterating/Deploying

Tour of Useful Non-Deep Learning Algorithms



https://www.kaggle.com/surveys/2017

What is Naïve Bayes?

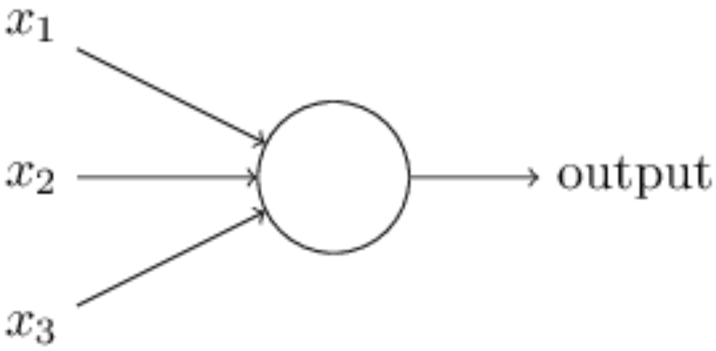


1. Estimate P(Y=v) as fraction of records with Y=v
2. Estimate P(X_i=u | Y=v) as fraction of "Y=v" records that also have X=u.
3. To predict the Y value given observations of all the X_i values, compute

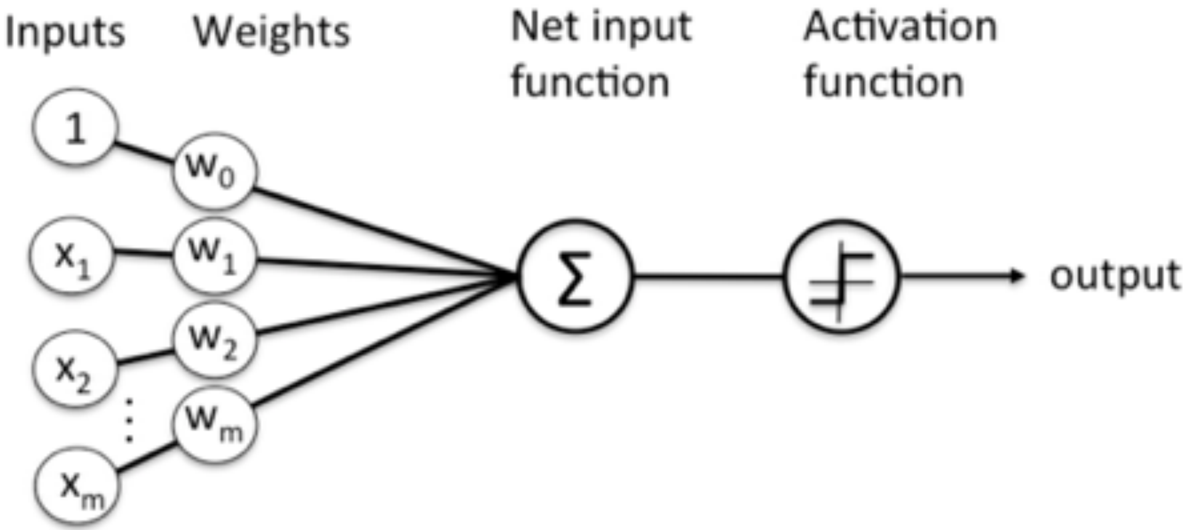
$$Y^{\text{predict}} = \operatorname{argmax} P(Y = v \mid X_1 = u_1 \cdots X_m = u_m)$$

- http://norvig.com/spell-correct.html

Perceptron

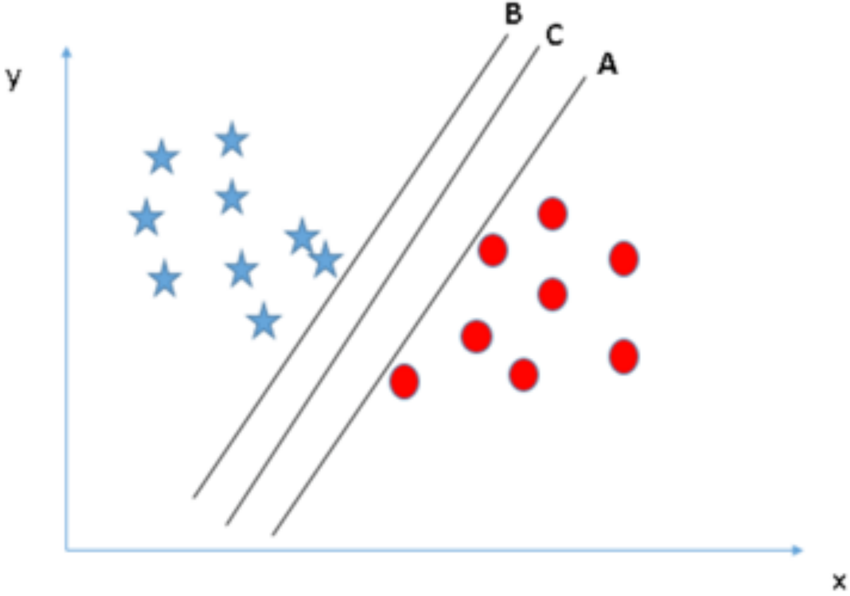


Perceptron



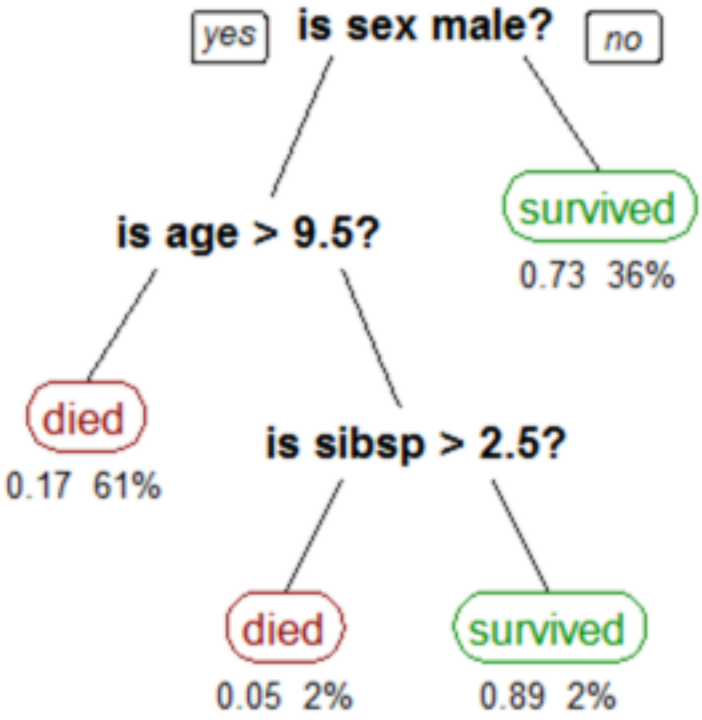
Schematic of Rosenblatt's perceptron.

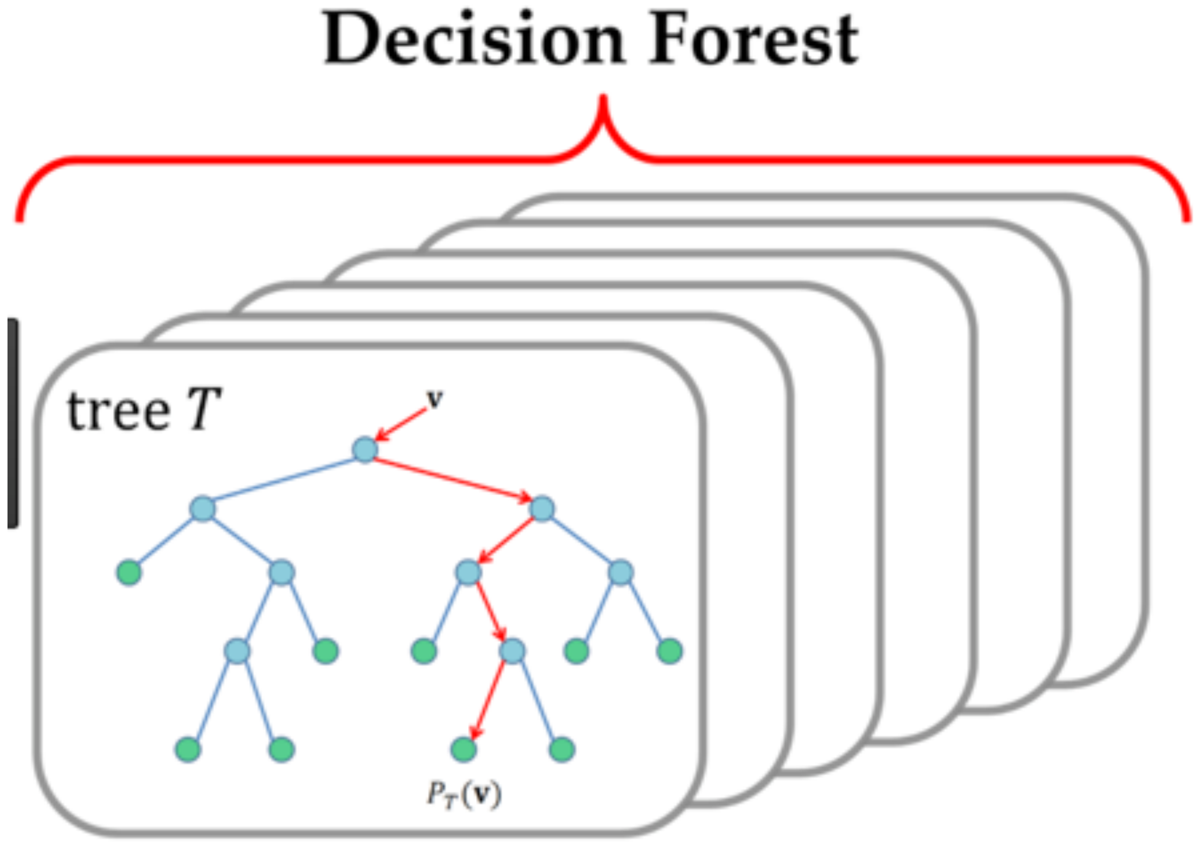
Other Algorithms: SVM



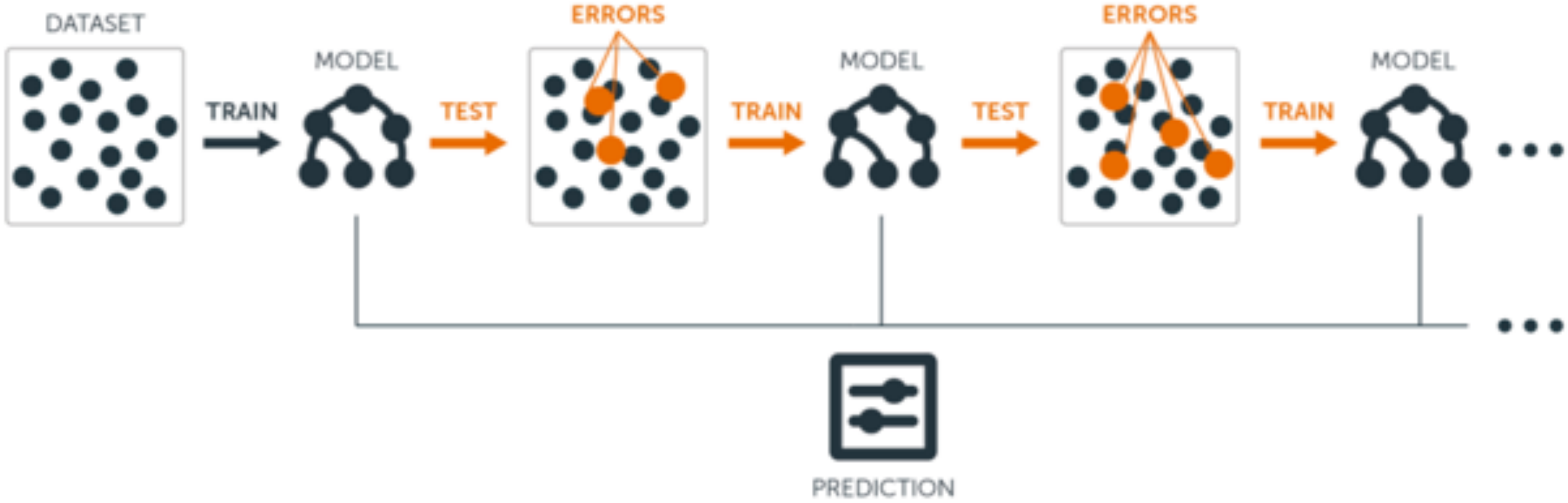
Cares more about correct classification than understanding probabilities.

Decision Trees




Decision Forest
tree T
$\mathbf{v}$
$P_T(\mathbf{v})$

Boosted Trees



XGBoost

dmlc

XGBoost eXtreme Gradient Boosting

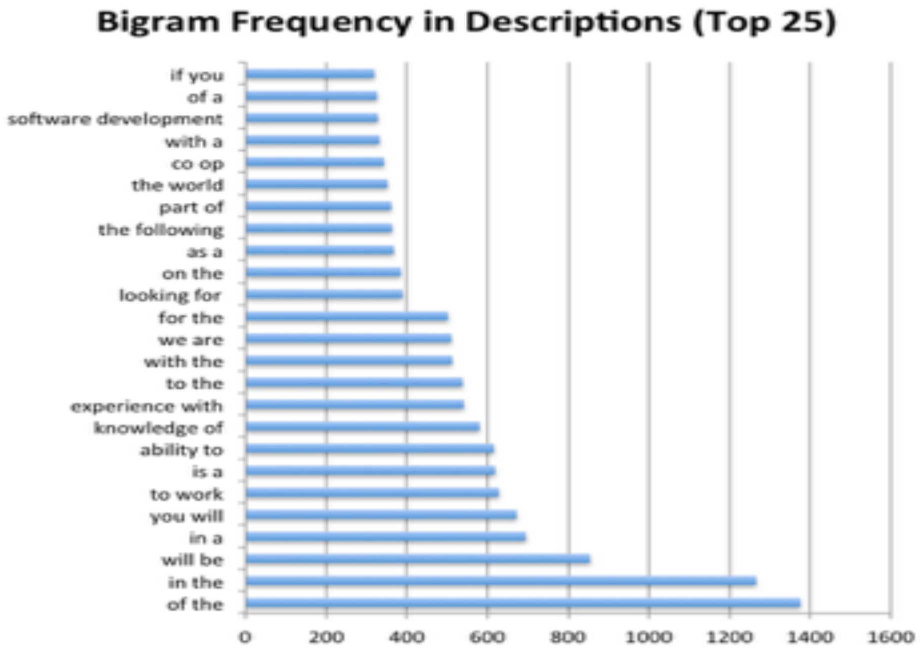
test-algorithm-cross-validate-svm.py
test-algorithm-cross-validate-rf.py

Questions:
Any observations?

pipeline.py

Questions:
Why might this be useful?

N-Grams


Bigram Frequency in Descriptions (Top 25)
if you
of a
software development
with a
co op
the world
part of
the following
as a
on the
looking for
for the
we are
with the
to the
experience with
knowledge of
ability to
is a
to work
you will
in a
will be
in the
of the
0
200
400
600
800
1000
1200
1400
1600

pipeline-bigrams.py

Questions:
Did this improve things?

pipeline-bigrams-cross-validate.py

Questions:
Did this improve things?

grid-search.py

Questions:
Can I find a way to improve things?

Module 3 End

Goals:
Improve our Machine Learning Classifier

Questions:
How do we evaluate models?
Why is it so important to evaluate models?
What are the common machine learning algorithms?

Common Pitfalls:
Testing on training data
Premature optimization

Deep Learning (and Vision)

Module 4 Begin

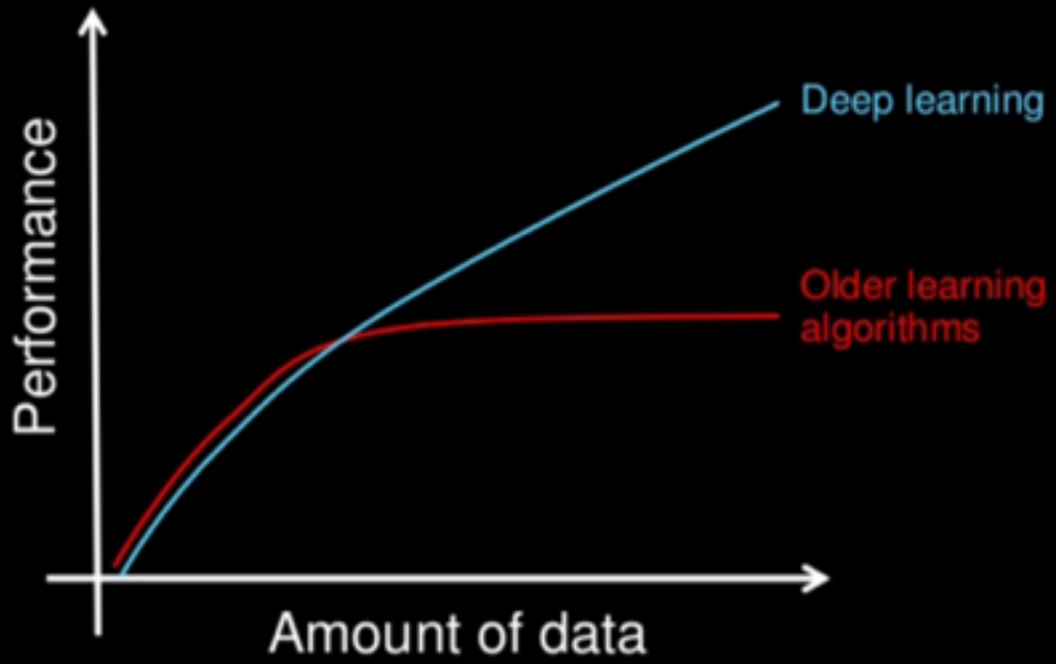
Goals:
Overview of Neural Networks

Questions:
What is a perceptron?
How do we frame a vision problem as a deep learning problem?

Image Classification Errors

What is the difference between Deep Learning, Machine Learning and Artificial Intelligence?

Why deep learning



How do data science techniques scale with amount of data?


Keras
TensorFlow
CudNN
CUDA
GPU

Set up a box

		Desktop CPU				Server CPU						GPU		
		1	2	4	8	1	2	4	8	16	32	G.980	G.1080	T.K80
FCN-5	Caffe	0.919	0.495	**0.480**	-	0.769	0.446	0.354	0.269	**0.287**	0.688	0.020	**0.017**	0.028
	CNTK	2.351	1.239	0.961	**0.810**	2.311	1.229	0.827	0.546	**0.530**	0.549	0.043	**0.033**	0.052
	TF	7.205	4.904	2.626	**1.933**	7.449	5.203	2.803	1.574	0.857	**0.594**	0.071	**0.063**	0.098
	Torch	1.227	0.655	**0.661**	-	1.030	0.740	0.535	0.440	**0.425**	0.892	0.044	**0.039**	0.056
FCN-8	Caffe	1.035	0.857	**0.572**	-	0.888	0.613	0.391	0.319	**0.316**	0.810	0.023	**0.019**	0.033
	CNTK	2.641	1.402	1.393	**0.919**	2.514	1.391	0.884	0.633	**0.579**	0.653	0.048	**0.037**	0.059
	TF	7.166	4.863	2.629	**1.955**	7.759	5.198	2.896	1.577	0.891	**0.619**	0.074	**0.065**	0.106
	Torch	1.316	0.706	**0.448**	0.881	1.106	0.774	0.559	0.475	**0.443**	0.975	0.046	**0.046**	0.057
AlexNet	Caffe	2.507	1.492	**1.005**	1.460	1.917	1.281	0.975	0.996	**1.035**	1.239	0.042	**0.038**	0.089
	CNTK	6.661	3.556	**2.123**	4.232	6.716	3.966	2.618	1.987	**1.446**	1.578	0.052	**0.043**	0.089
	TF	3.192	2.219	1.346	**1.134**	3.720	2.671	1.416	0.812	**0.516**	0.627	0.037	**0.012**	0.064
	Torch	4.689	2.473	**2.090**	4.012	3.293	1.883	1.156	1.145	**1.083**	1.182	0.036	**0.034**	0.073
ResNet	Caffe	7.810	5.312	**4.056**	5.876	6.150	5.390	4.314	**4.124**	4.500	5.034	-	**0.208**	0.353
	CNTK	-	-	-	-	-	-	-	-	-	-	0.289	**0.261**	0.468
	TF	21.63	12.19	7.655	**6.340**	20.49	14.340	7.703	4.600	**2.890**	3.937	0.226	**0.085**	0.392
	Torch	12.10	7.147	-	-	10.16	6.928	4.856	3.757	**3.524**	4.165	0.216	**0.181**	0.412
LSTM-32	CNTK	0.579	0.391	**0.306**	1.153	0.591	0.418	0.353	0.338	**0.342**	0.442	0.433	**0.366**	0.602
	TF	9.305	3.432	**2.020**	1.722	6.453	3.782	2.167	1.228	**0.769**	0.706	0.086	**0.083**	0.122
	Torch	4.872	2.680	**2.366**	3.645	4.704	2.971	2.067	1.706	**1.763**	2.900	0.124	**0.098**	0.204
LSTM-64	CNTK	1.026	0.690	**0.535**	1.860	1.043	0.756	0.622	0.585	**0.648**	0.790	0.779	**0.649**	1.052
	TF	11.69	7.292	3.515	**3.476**	12.76	7.823	4.402	2.524	1.590	**1.469**	0.178	**0.173**	0.233
	Torch	9.622	5.323	**4.980**	6.975	9.364	5.613	4.054	**3.252**	3.357	5.815	0.247	**0.194**	0.406

https://www.nextplatform.com/2016/09/01/cpu-gpu-put-deep-learning-framework-test/

Build your own box

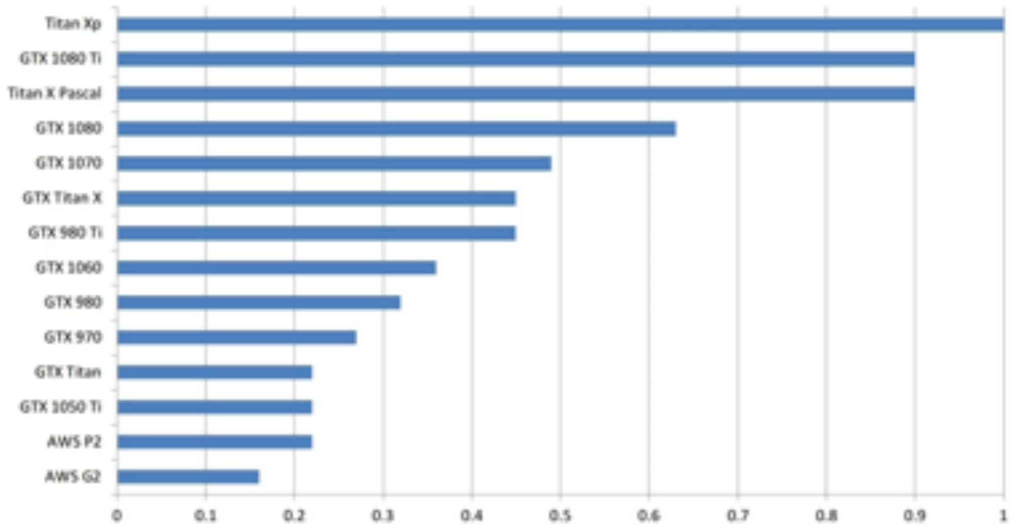
Build a super fast deep learning machine for under $1,000

The adventures in deep learning and cheap hardware continue!

By Lukas Biewald. February 1, 2017

https://www.oreilly.com/learning/build-a-super-fast-deep-learning-machine-for-under-1000

Performance Benchmarks



Rough performance comparisons between GPUs. This comparison is only valid for large workloads.

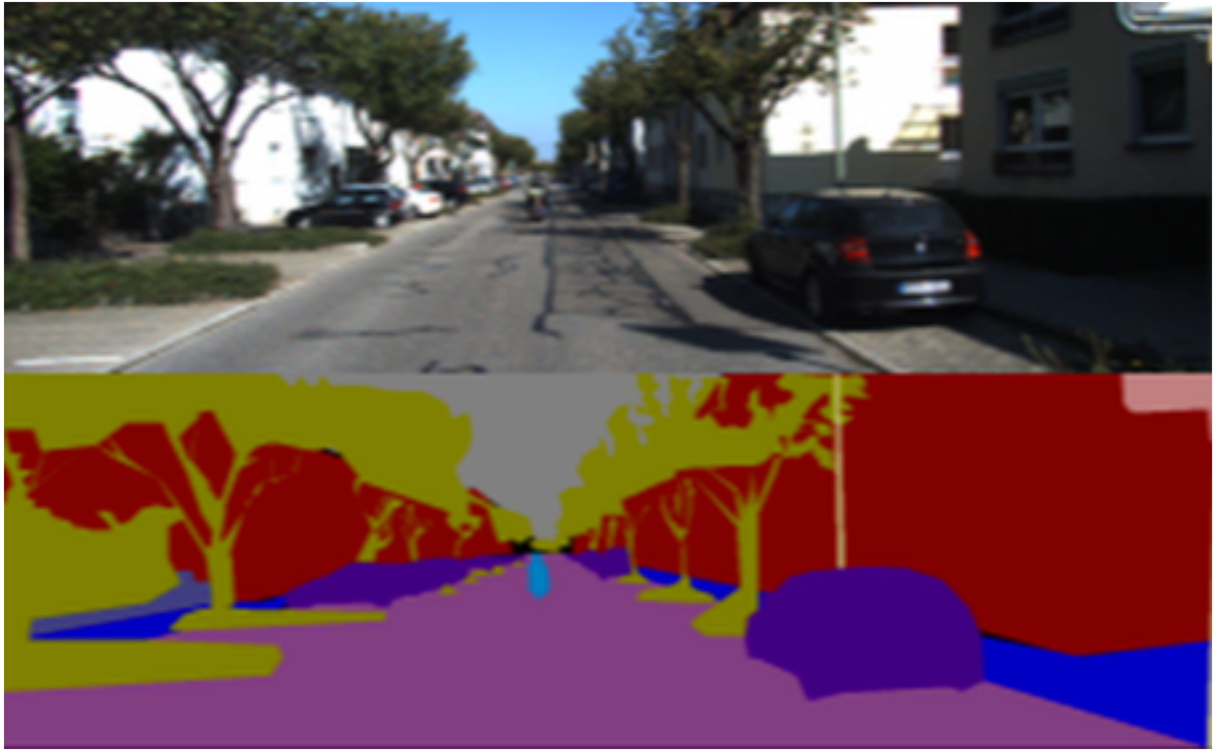
http://timdettmers.com/2017/04/09/which-gpu-for-deep-learning/

Framing Problems: Object Recognition

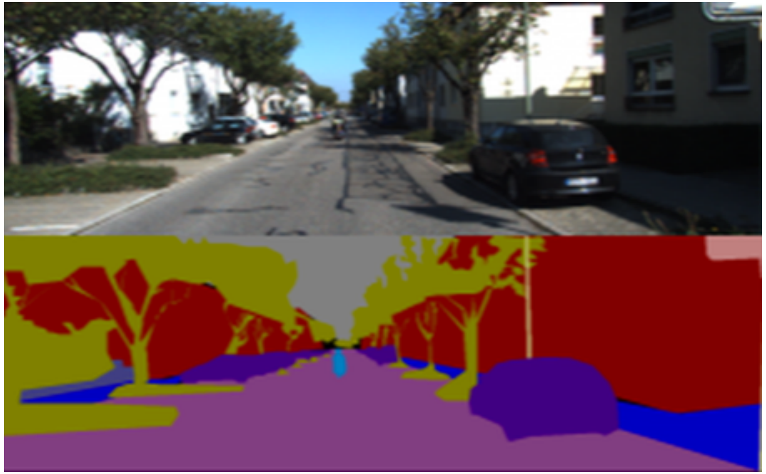
Framing Problems: Object Recognition

(0,0)	(0,1)	(0,2)	....	(1,0)	(1,1)	(1,2)	....	(2,0)	(2,1)	(2,2)	Label
23	15	3		56	23	12		56	23	12	Cat

Framing Problems: Vision Applications

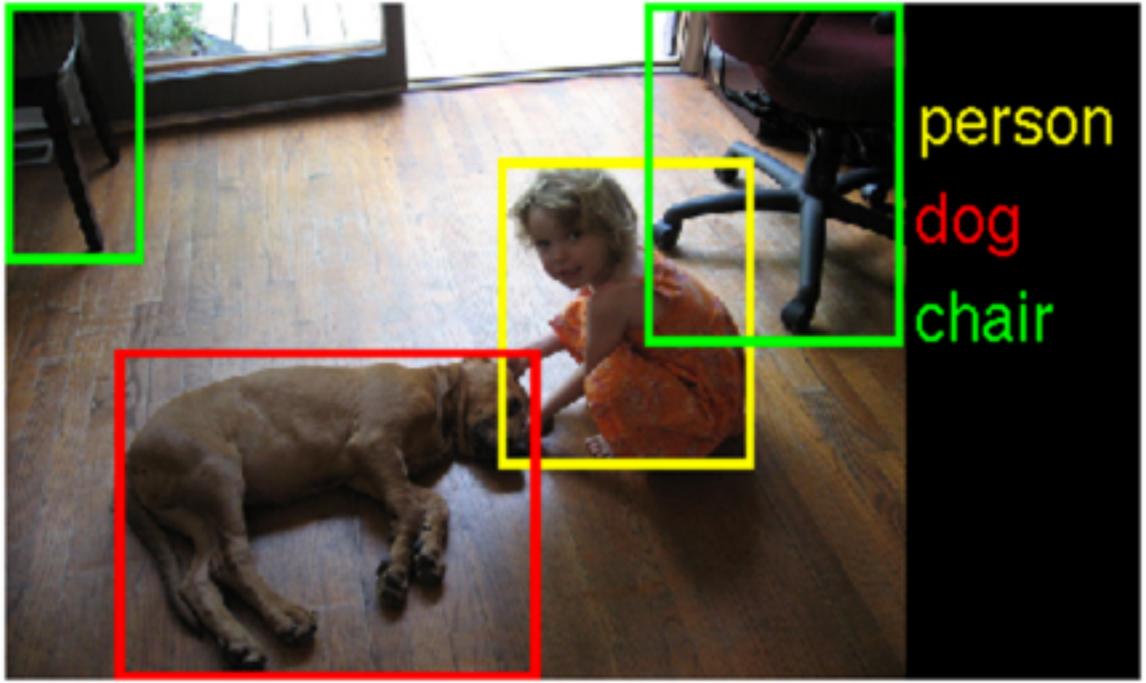


Semantic Segmentation

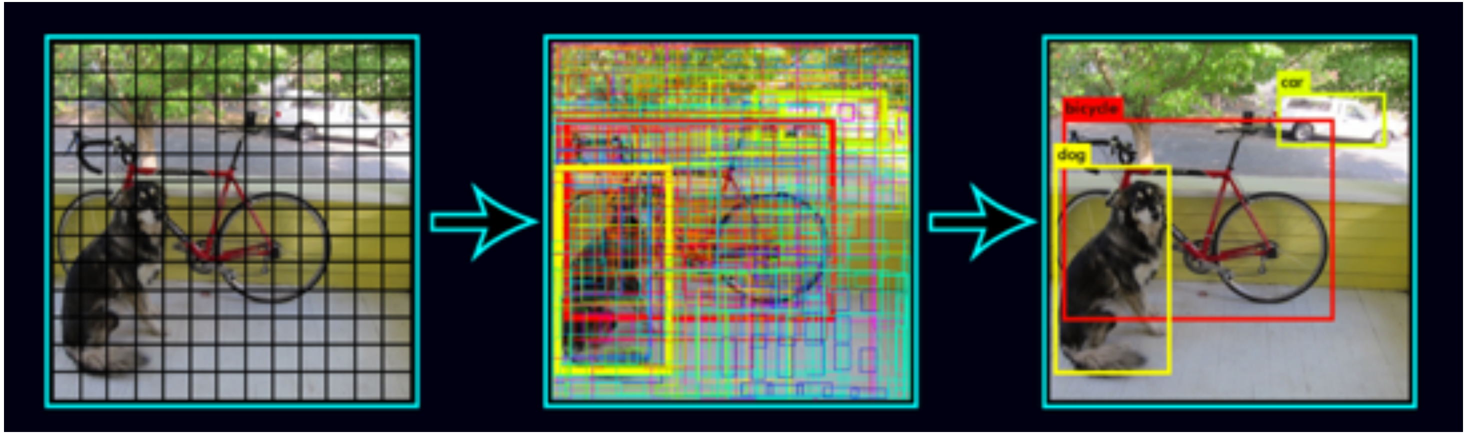


(0,0)	(0,1)	(0,2)	....	(1,0)	(1,1)	(1,2)	....	Label (0,0)	Label (0,1)	Label (1,0)	Label (1,1)
23	15	3		56	23	12		Tree	Tree	Car	Car

Bounding Box

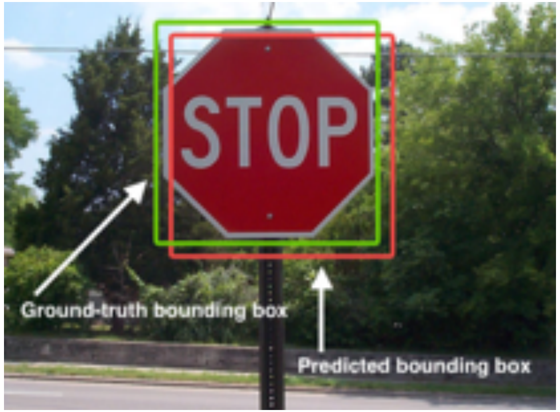


Bounding Box as Classification



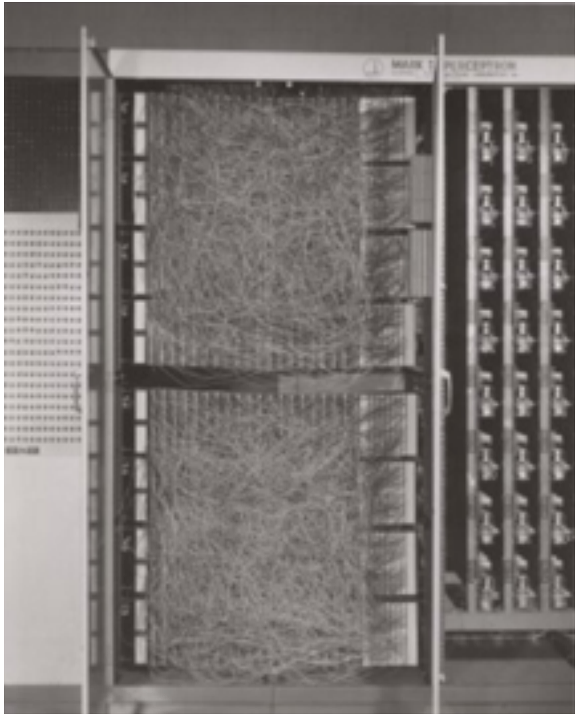
(0,0)	(0,1)	(0,2)		(1,0)	(1,1)	(1,2)	Box Upper Left X	Box Upper Left Y	Good Box	Object
23	15	3	····	123	89	56	15	90	True	Bicycle
23	15	3		123	89	56	23	23	False	Bicycle
23	15	3		123	89	56	56	23	False	Dog

Bounding Box as Regression



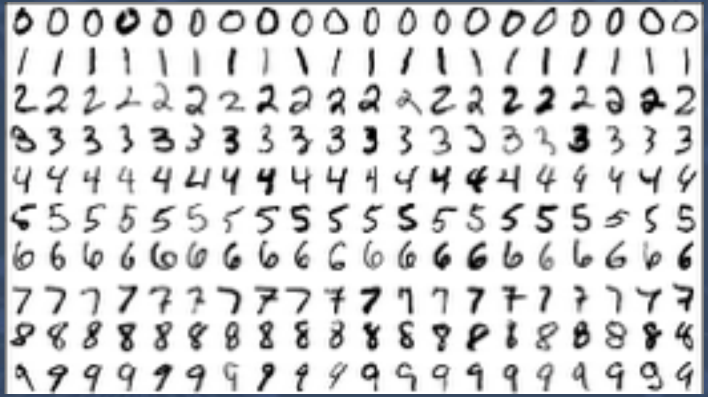
(0,0)	(0,1)	(0,2)		(1,0)	(1,1)	(1,2)	Box Upper Left X	Box Upper Left Y
23	15	3	….	123	89	56		
123	143	23		78	54	1		
12	17	6		90	9	30		

The First Perceptron

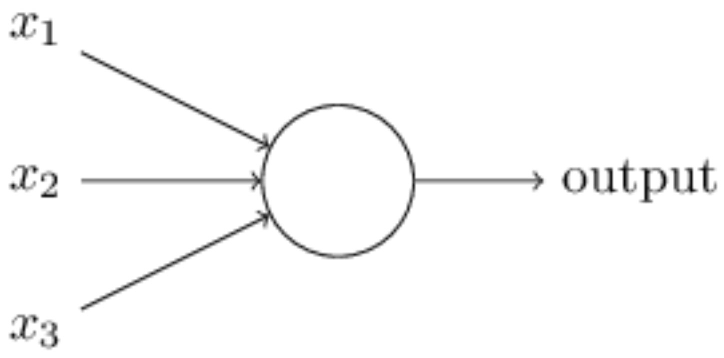


keras-digits.py

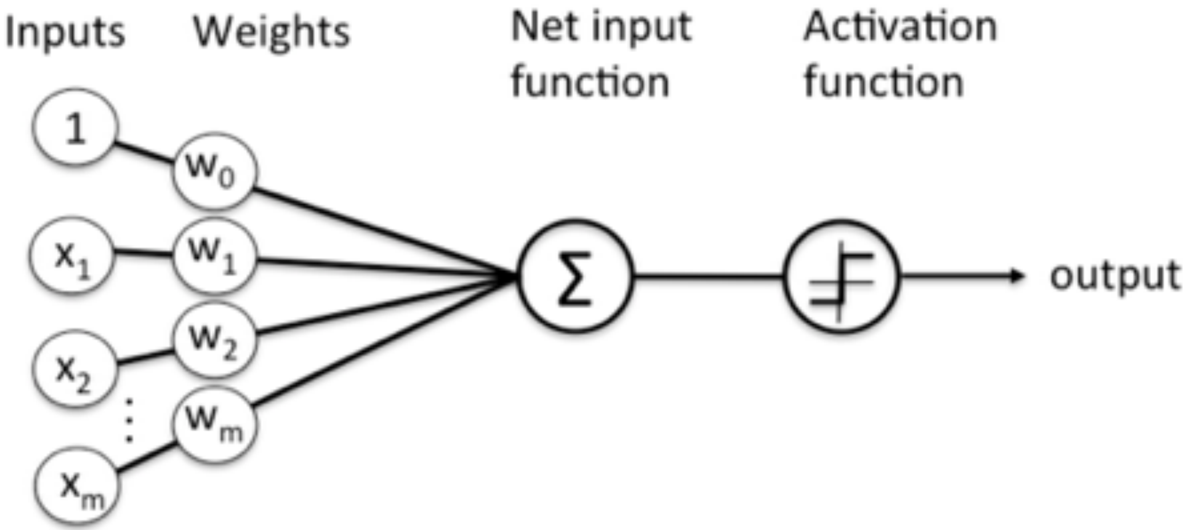
Questions:
What is the data type of X_train?



Perceptron

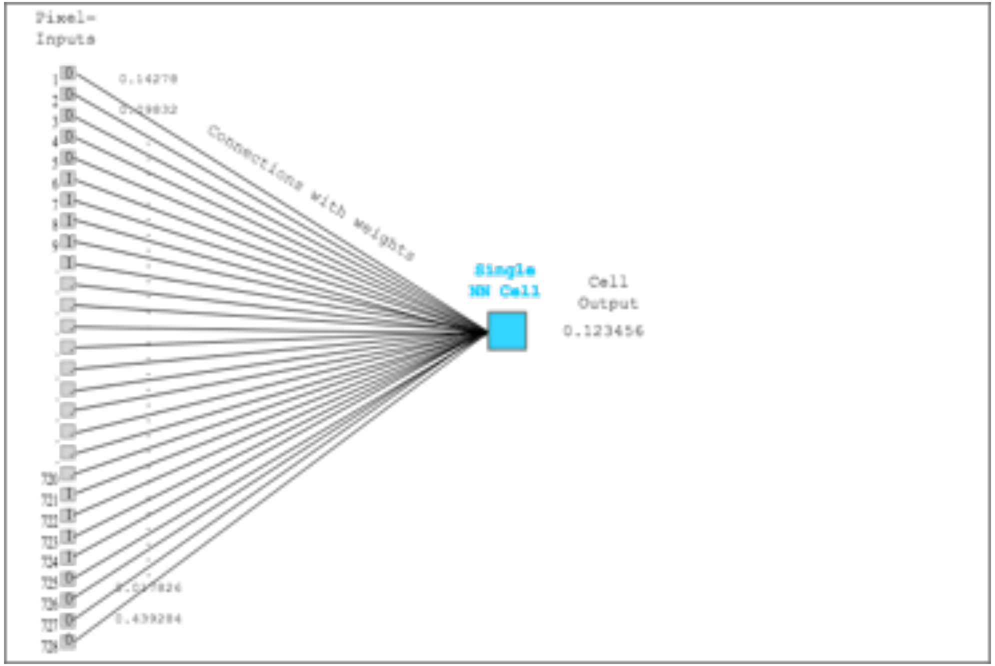


Perceptron



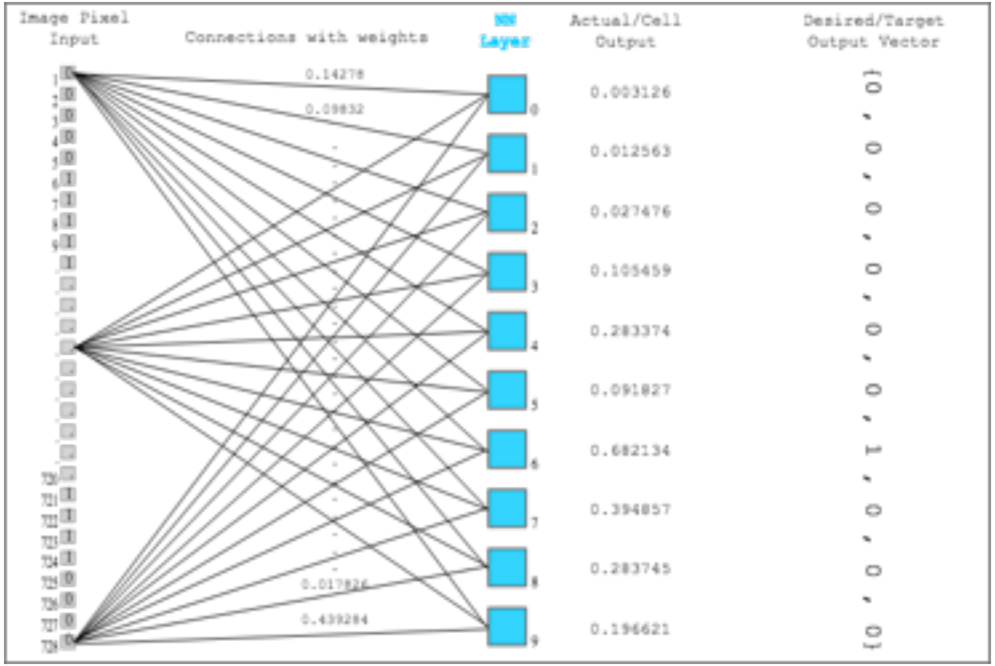
Schematic of Rosenblatt's perceptron.

Perceptron on MNist



https://mmlind.github.io/Simple_1-Layer_Neural_Network_for_MNIST_Handwriting_Recognition/

Classifiers



One hot encoding

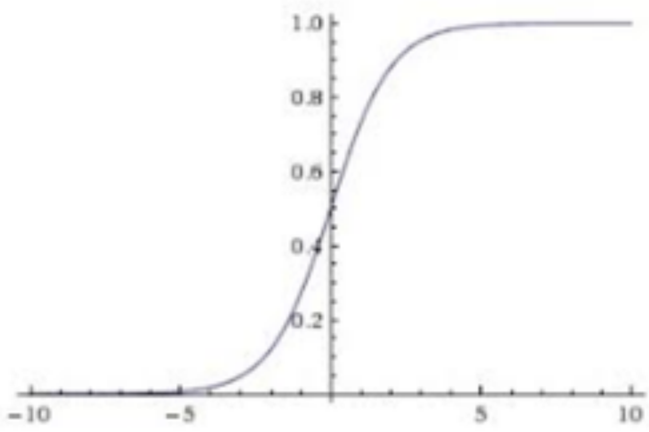
Label	0	1	2	3	4	5	6	7	8	9
0	1	0	0	0	0	0	0	0	0	0
4	0	0	0	0	1	0	0	0	0	0
4	0	0	0	0	1	0	0	0	0	0
3	0	0	0	1	0	0	0	0	0	0
0	1	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	1

wandb run perceptron-linear.py

Questions:

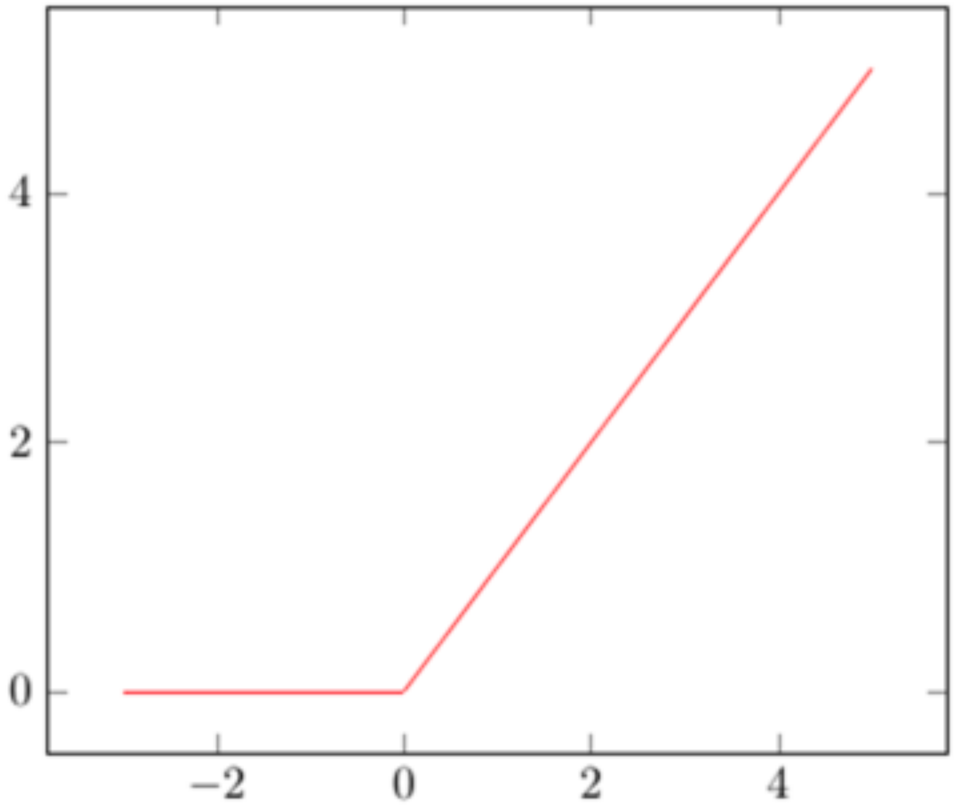
Is our model better than a baseline?

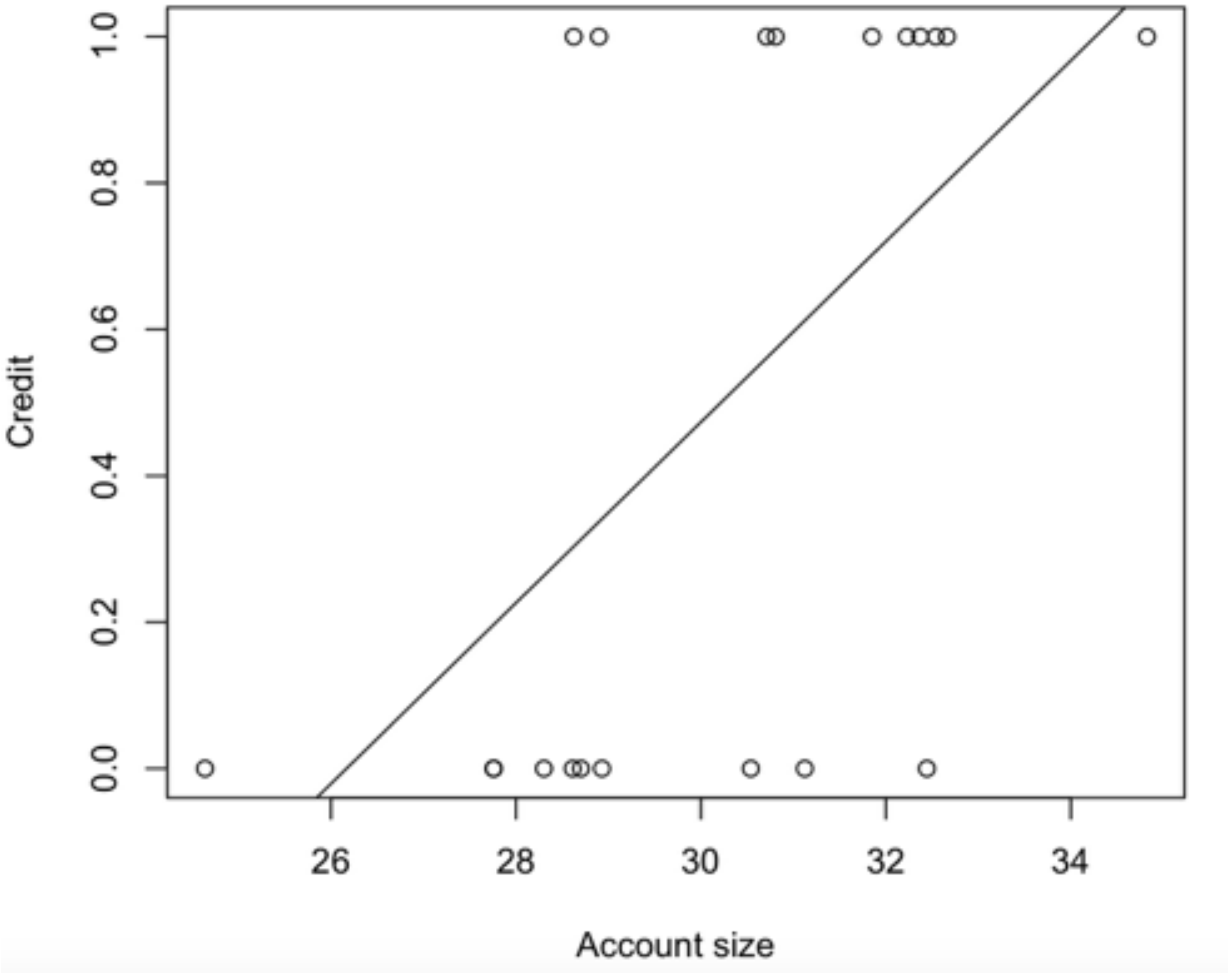
Activation Functions: Sigmoid



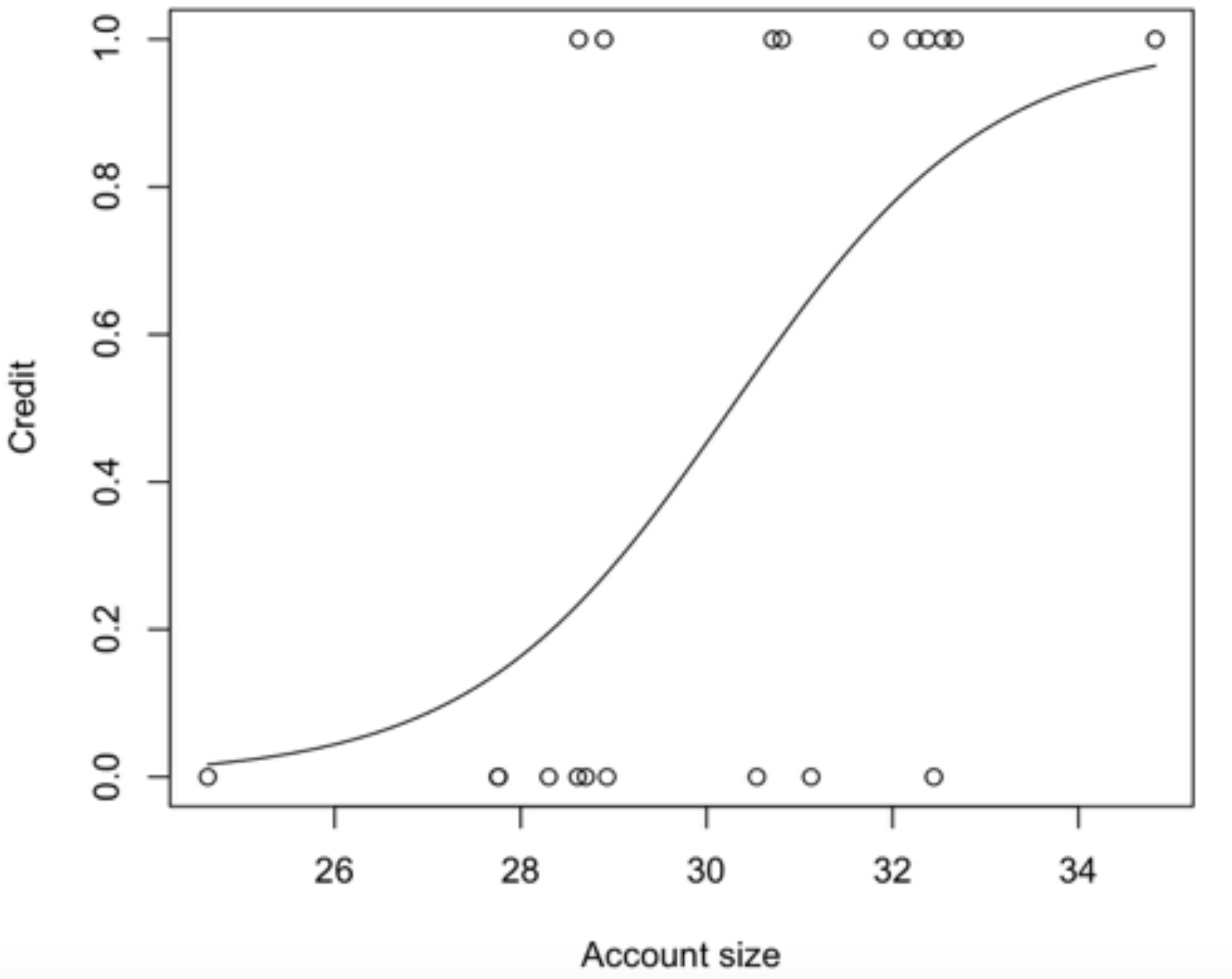
(special case of softmax and logistic)

Activation Functions: ReLU





Credit
1.0
0.8
0.6
0.4
0.2
0.0
26
28
30
32
34
Account size



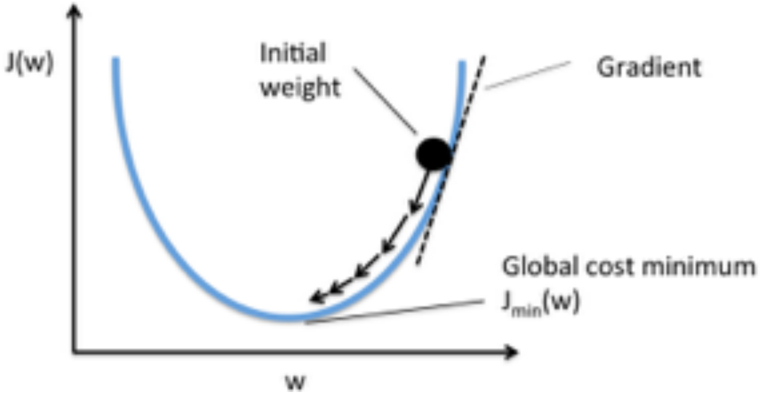
Credit
0.0
0.2
0.4
0.6
0.8
1.0
26
28
30
32
34
Account size

wandb run perceptron-logistic.py

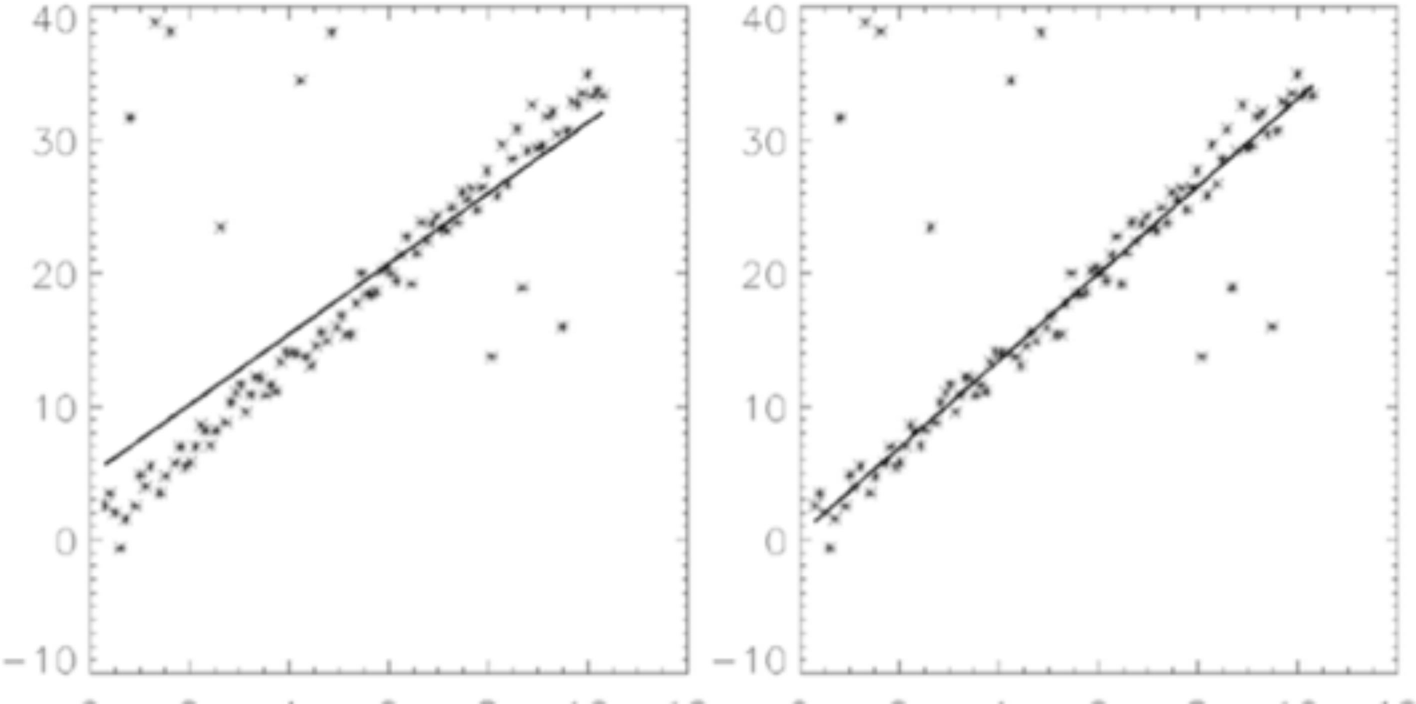
Questions:

Why is this so much better?

Gradient Descent



Squared Error vs Absolute Error



wandb run perceptron.py

Questions:

What does Flatten do?

What does activation="softmax" do?

What does optimizer='adam' do?

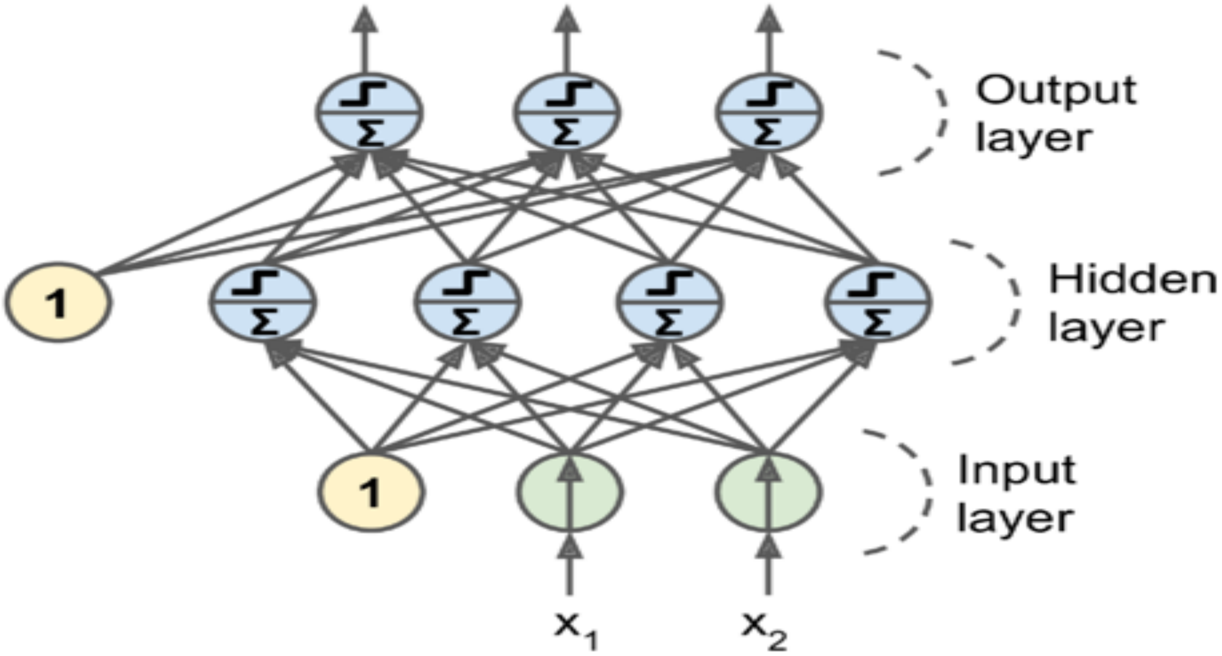
What does epochs=100 do?

wandb run perceptron-normalize.py

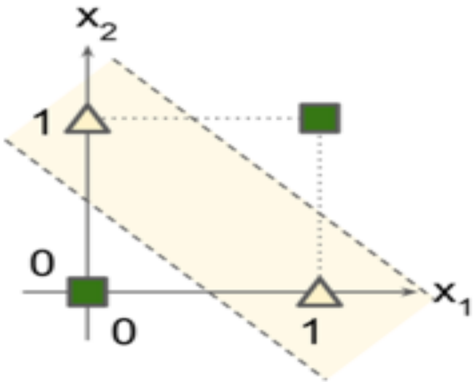
Questions:
What did I change?
Any issues with how we're computing accuracy?

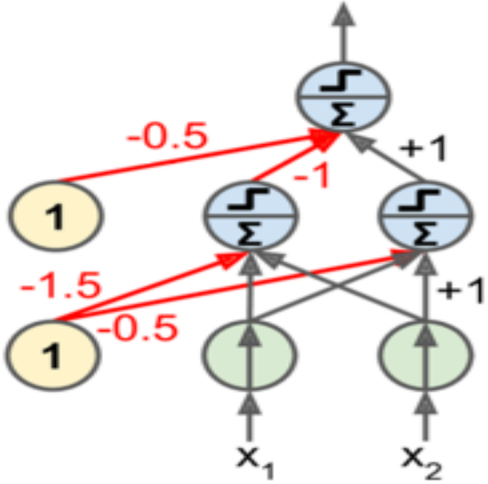
Multi-Layer Perceptrons and CNNs

Two Layers of Perceptrons



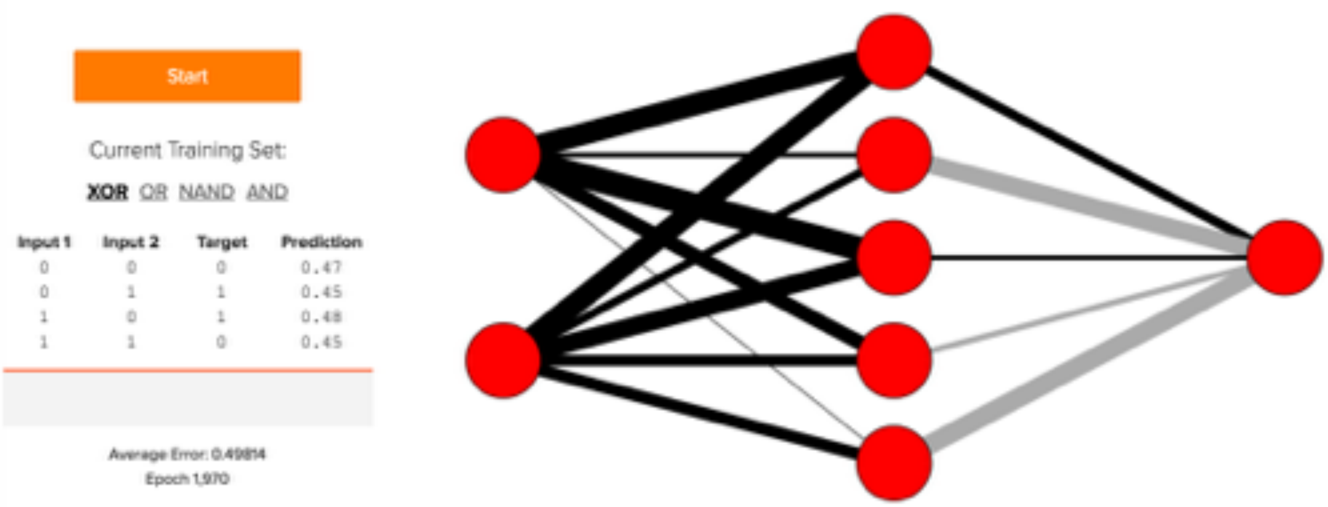
Multi-Payer Perceptrons Solving the “XOR” Problem





Backpropagation (1985)

http://www.emergentmind.com/neural-network



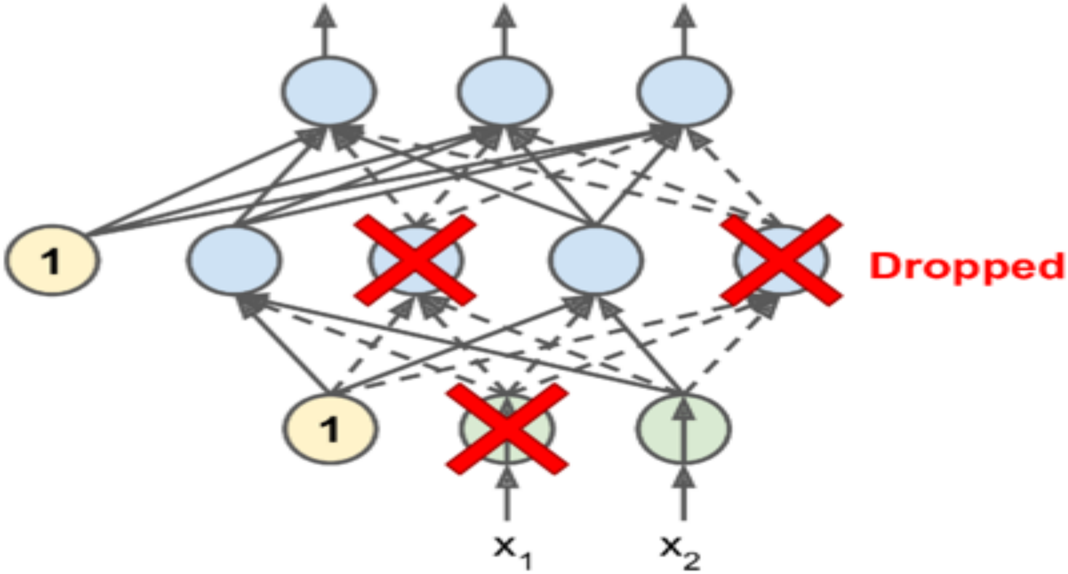
Neural Network Visualization

- http://playground.tensorflow.org/

cd keras-mlp
wandb run python mlp.py

Questions:
How did this change run-time and accuracy?
How could we make a deeper model?
Why do we use activation="relu" in one layer?

Dropout



wandb run python dropout.py

Questions:

How does this change training accuracy and validation accuracy?

When would we consider using dropout?

Module 5 End

Goals:
Improved perceptrons and multilayer perceptrons

Questions:
What is gradient descent?
What is an activation function?
What is a loss function?
What is a multilayer perceptron?
Why?

Common Pitfalls:
Premature optimization
Lots of hyper parameters
Can learn test data through hyper parameter search

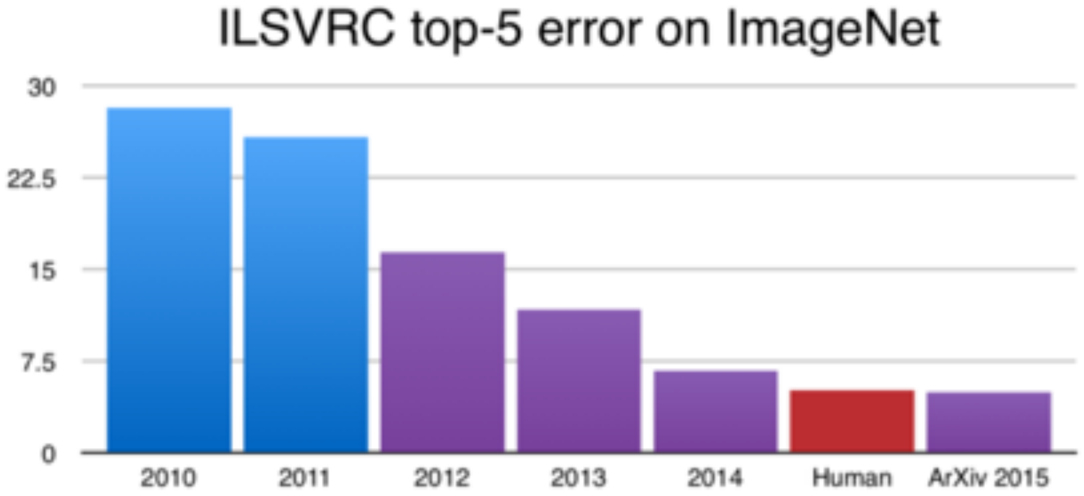
Module 6 Begin

Goals:
Convolutional Neural Networks

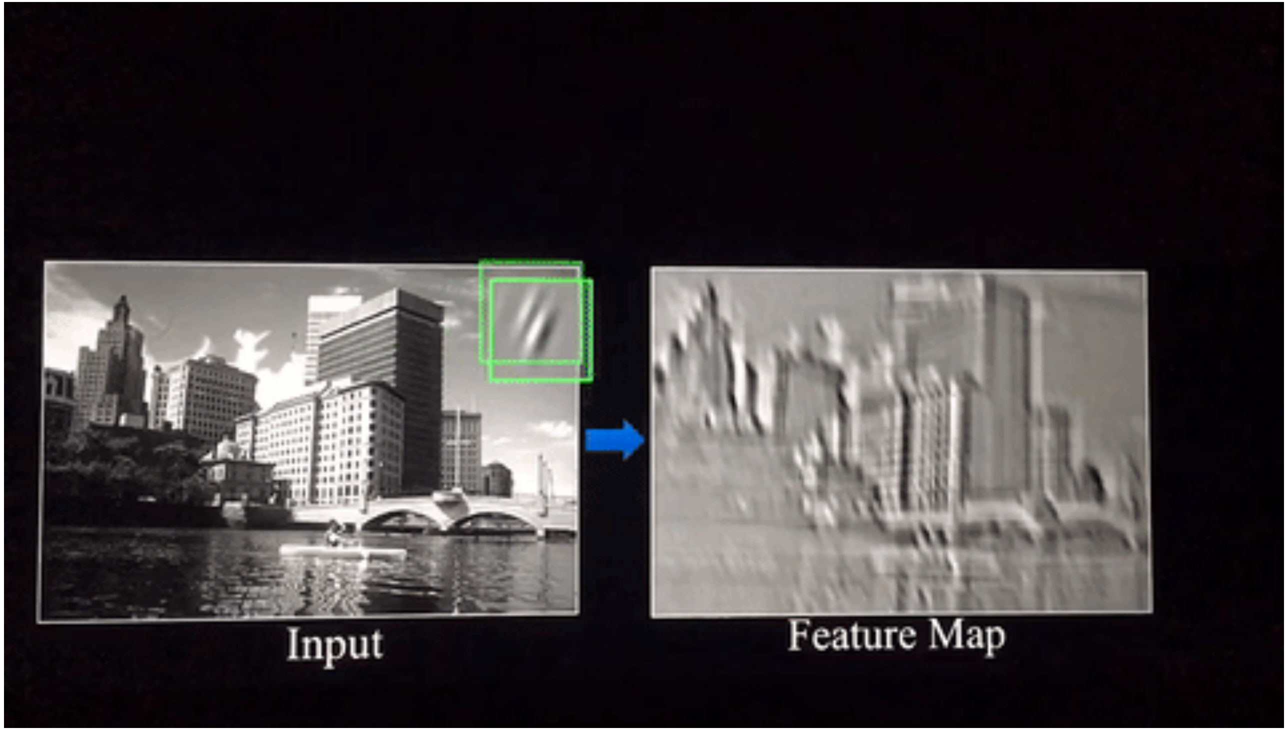
Questions:
What is a convolution?
What is max pooling?
Why?

Convolutional Neural Networks

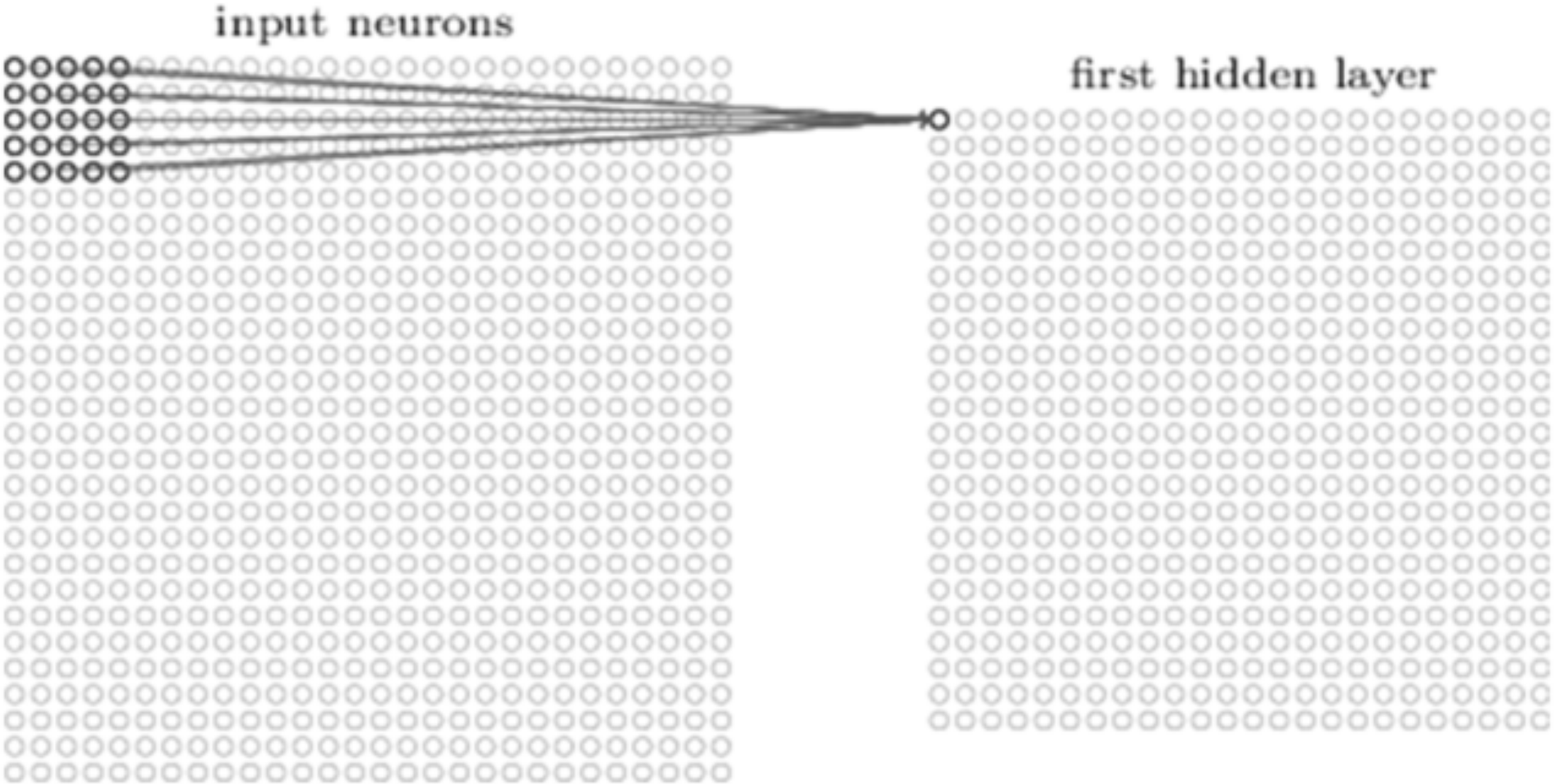
- Explanationshttps://ujjwalkarn.me/2016/08/11/intuitive-explanation-convnets/
- https://adeshpande3.github.io/adeshpande3.github.io/A-Beginner's-Guide-To-Understanding-Convolutional-Neural-Networks/



Convolution

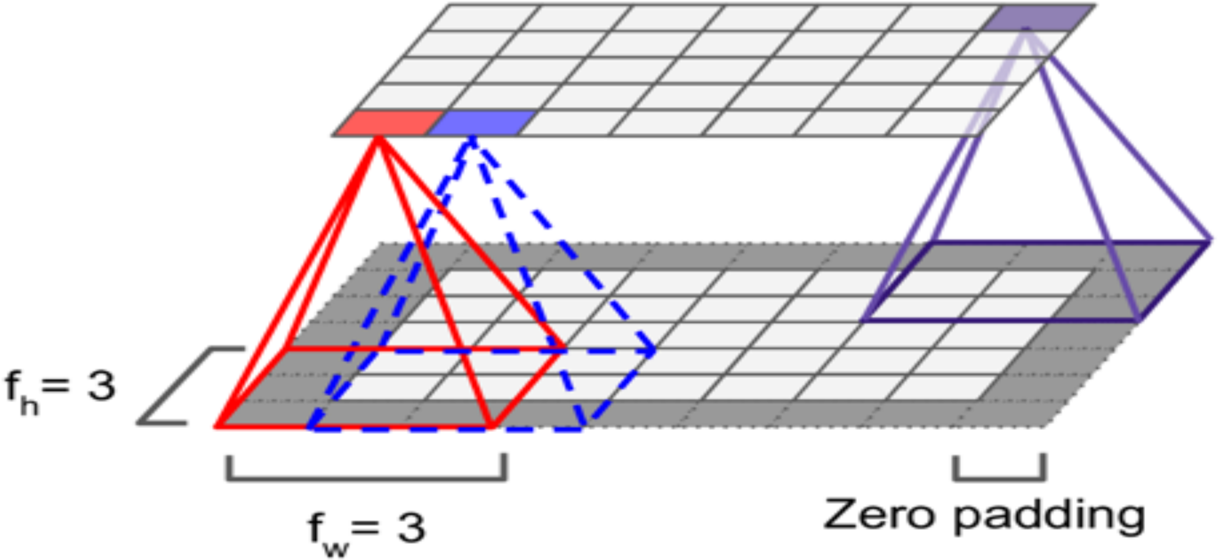


https://ujwlkarn.files.wordpress.com/2016/08/giphy.gif?w=748



Visualization of 5 x 5 filter convolving around an input volume and producing an activation map

Convolutions



convolution-demo.py

Questions:

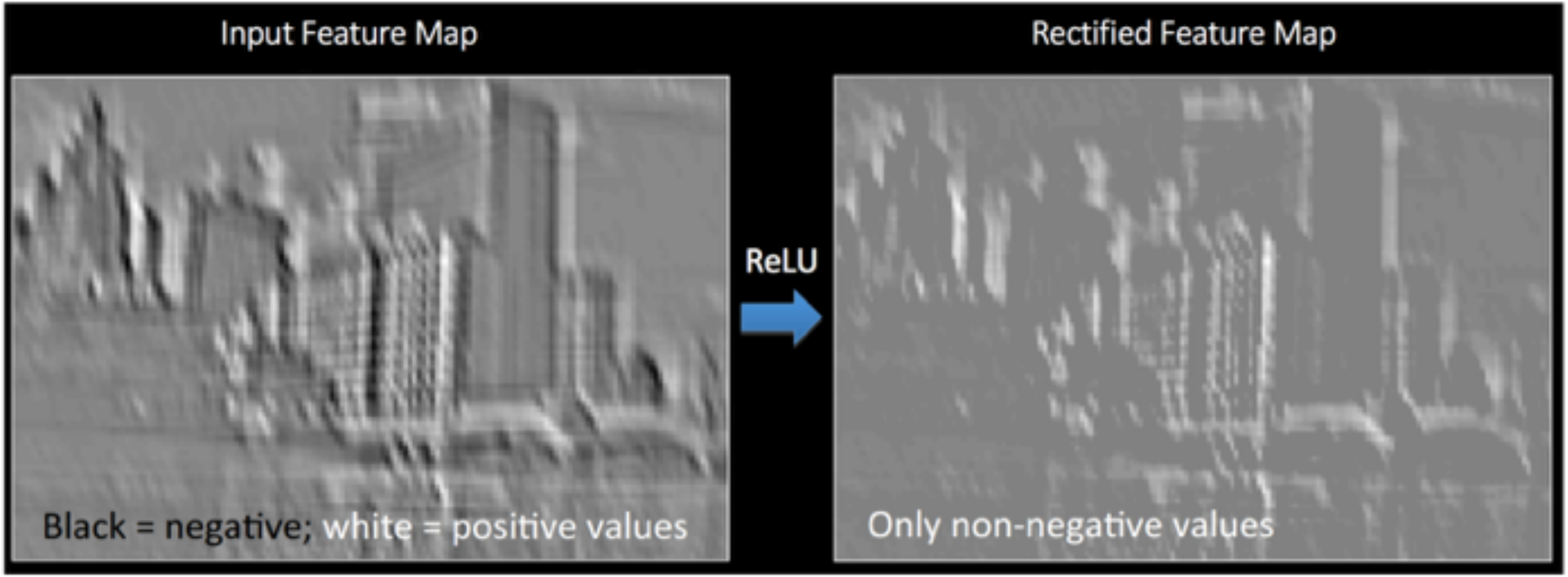
What happens if all the numbers are zeros?

What happens if the numbers are large?

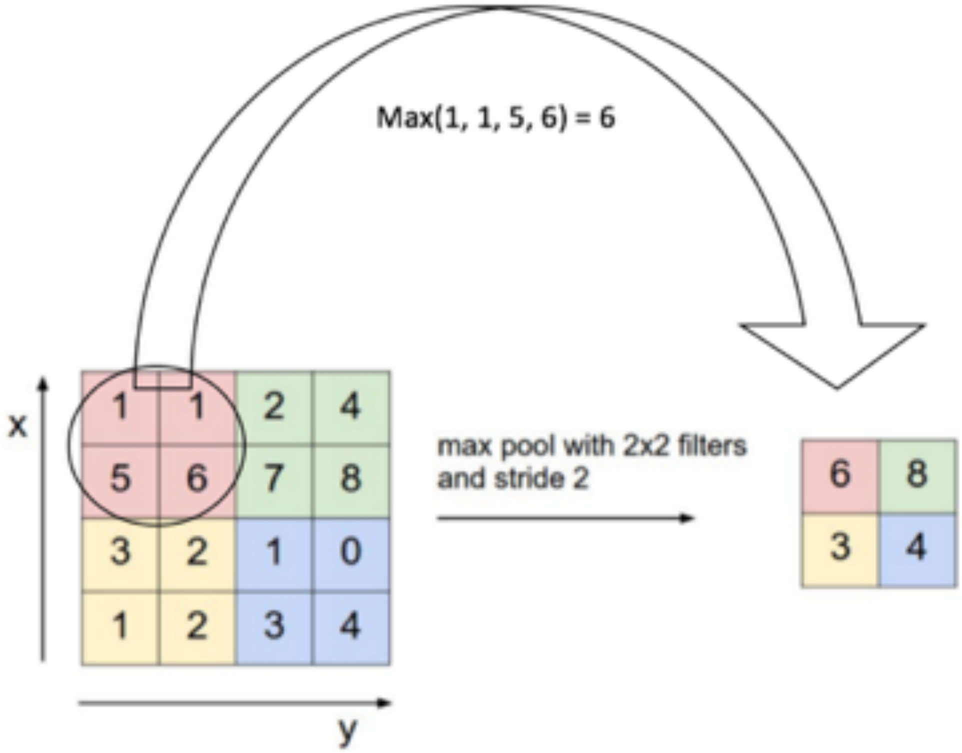
What convolution would leave the image unaffected?

(For the math nerds) What is the derivative of the coefficients with respect to the loss function?

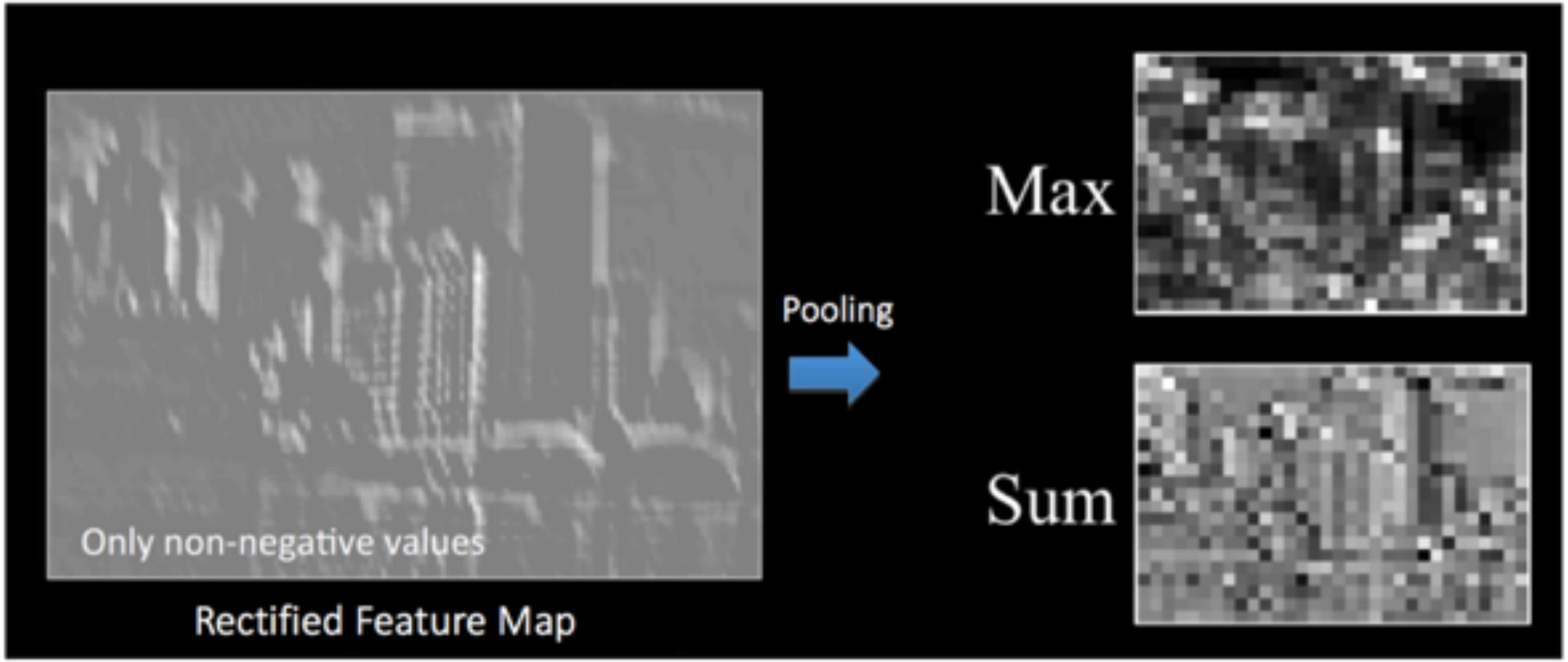
Activation Function (ReLU)



Pooling


Max(1, 1, 5, 6) = 6
x
1
1
2
4
5
6
7
8
3
2
1
0
1
2
3
4
y
max pool with 2x2 filters
and stride 2
6
8
3
4

Pooling

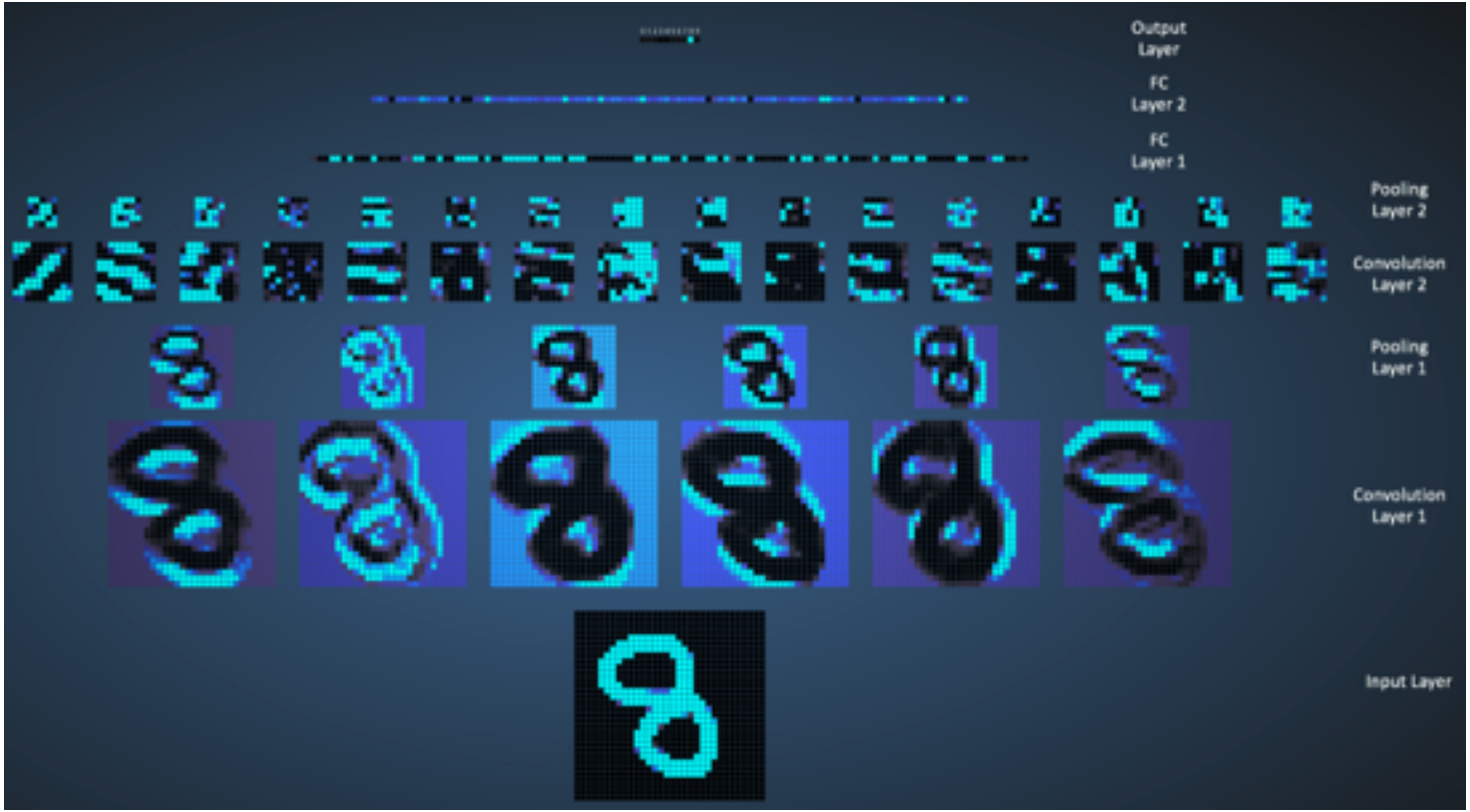


Max
Pooling
Sum
Only non-negative values
Rectified Feature Map

maxpool-demo.py

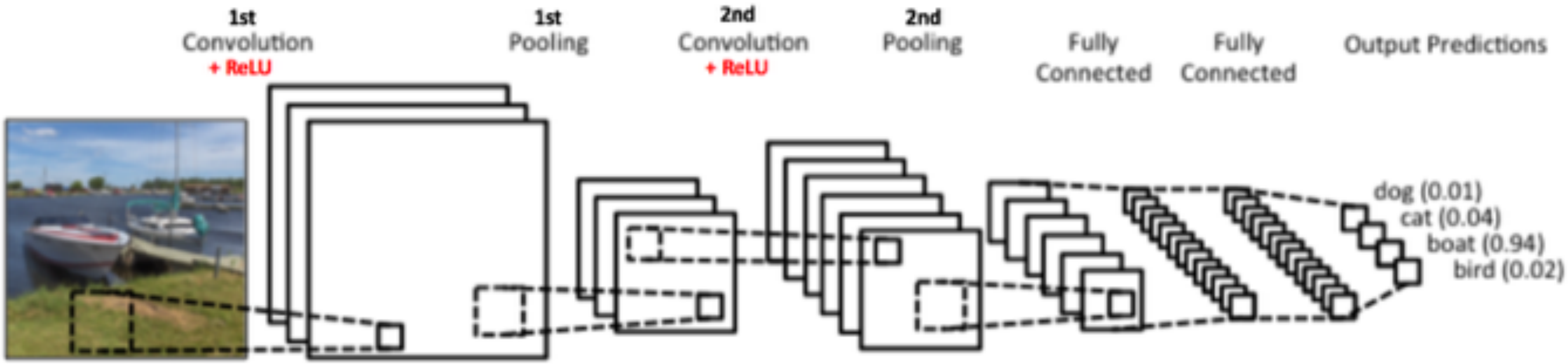
Questions:
What is happening here?

CNN



Output Layer
FC Layer 2
FC Layer 1
Pooling Layer 2
Convolution Layer 2
Pooling Layer 1
Convolution Layer 1
Input Layer

Architecture



cd keras-cnn

Run:
wandb run python cnn-1.py

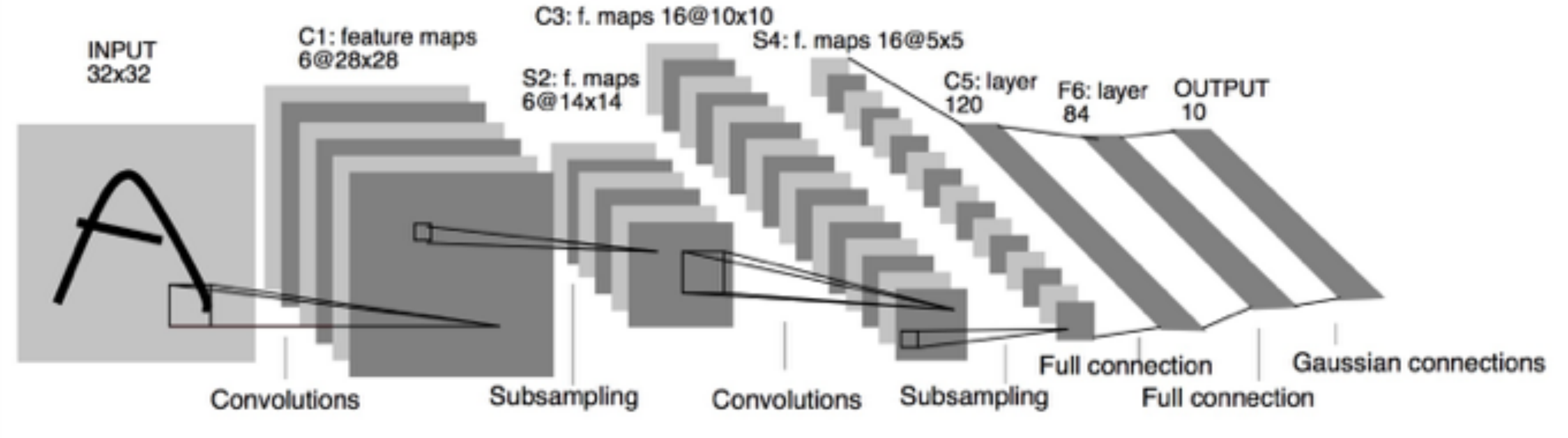
Questions:
Why am I reshaping the input data?
How could I add more convolutional and pooling layers?

Transfer Learning

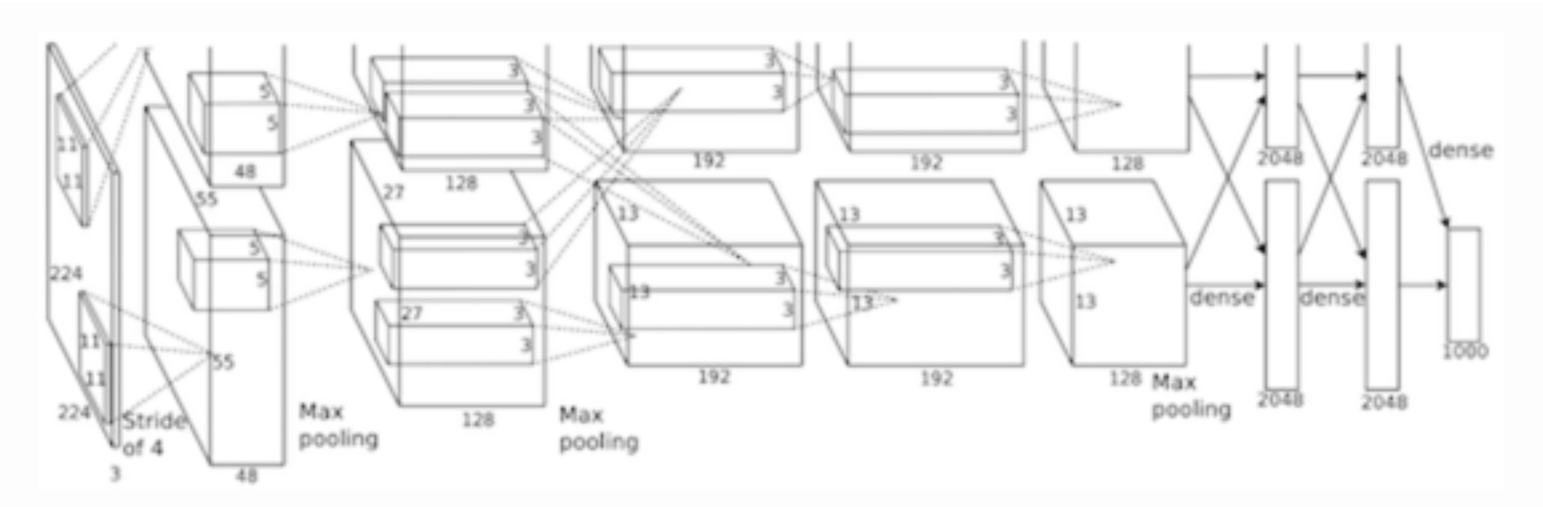
Image Models – Quick History

- https://culurciello.github.io/tech/2016/06/04/nets.html

1994 LeNet

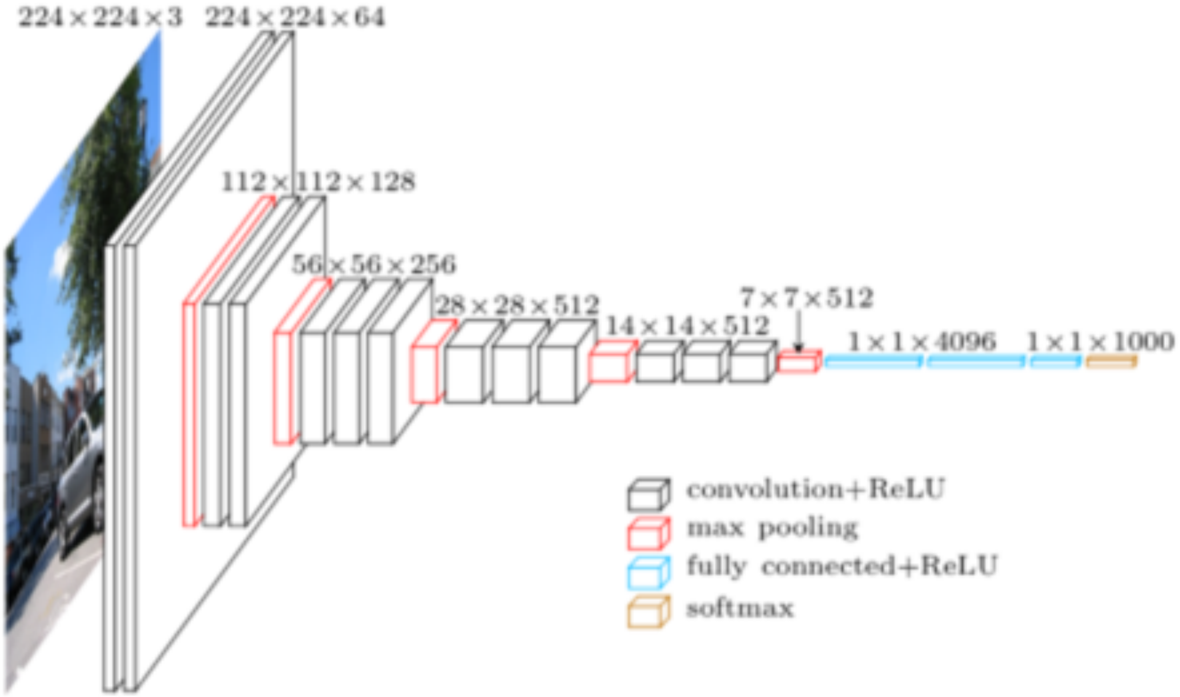

C3: f. maps 16@10x10
C1: feature maps
6@28x28
INPUT
32x32
S4: f. maps 16@5x5
S2: f. maps
6@14x14
C5: layer
120
F6: layer
84
OUTPUT
10
Convolutions
Subsampling
Convolutions
Subsampling
Full connection
Full connection
Gaussian connections

AlexNet



224
11
11
224
3
Stride of 4
55
55
5
5
48
48
Max pooling
27
27
128
128
3
Max pooling
13
13
192
192
13
13
192
192
13
13
128
128
Max pooling
dense
dense
dense
2048
2048
2048
2048
1000

VGG architecture (2014)

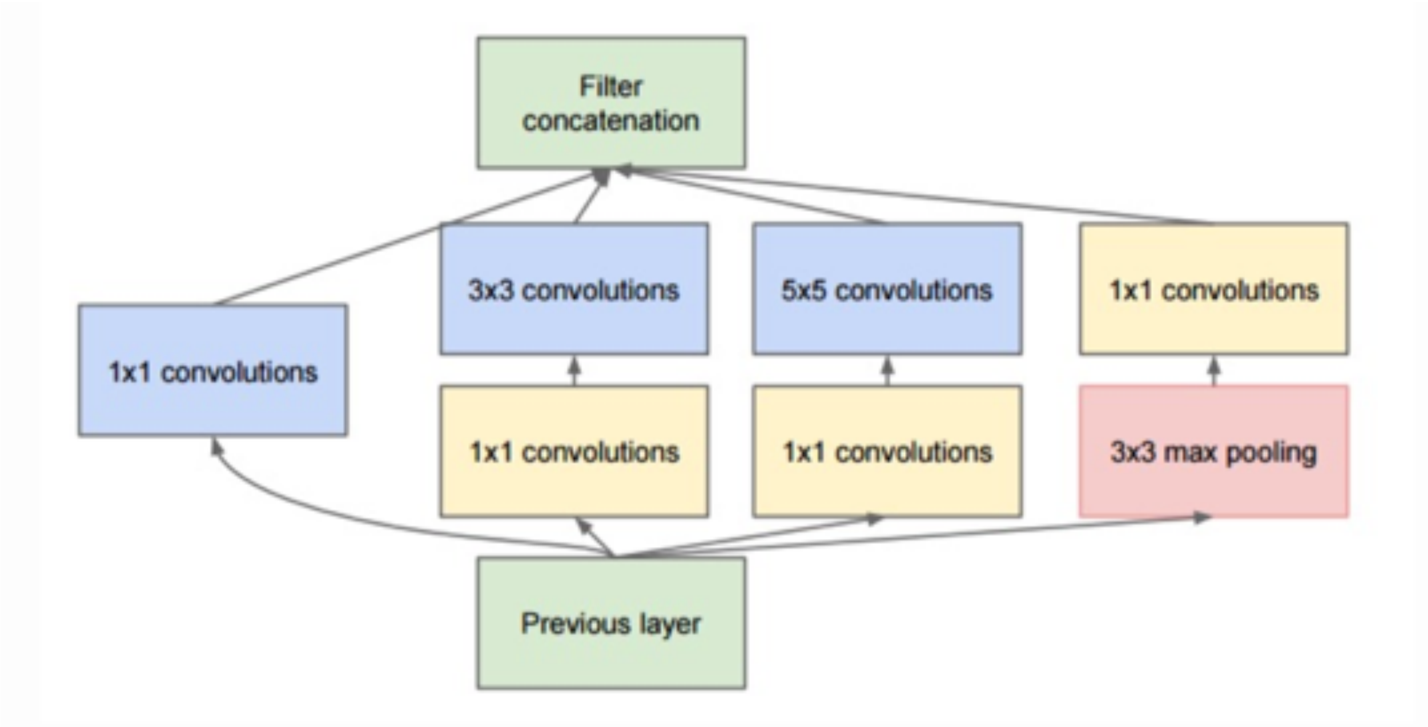


http://blog.christianperone.com/2016/01/convolutional-hypercolumns-in-python/

keras-vgg-inspect.py

Questions:
How many parameters does this model have?
Why doesn't it overfit the training data?

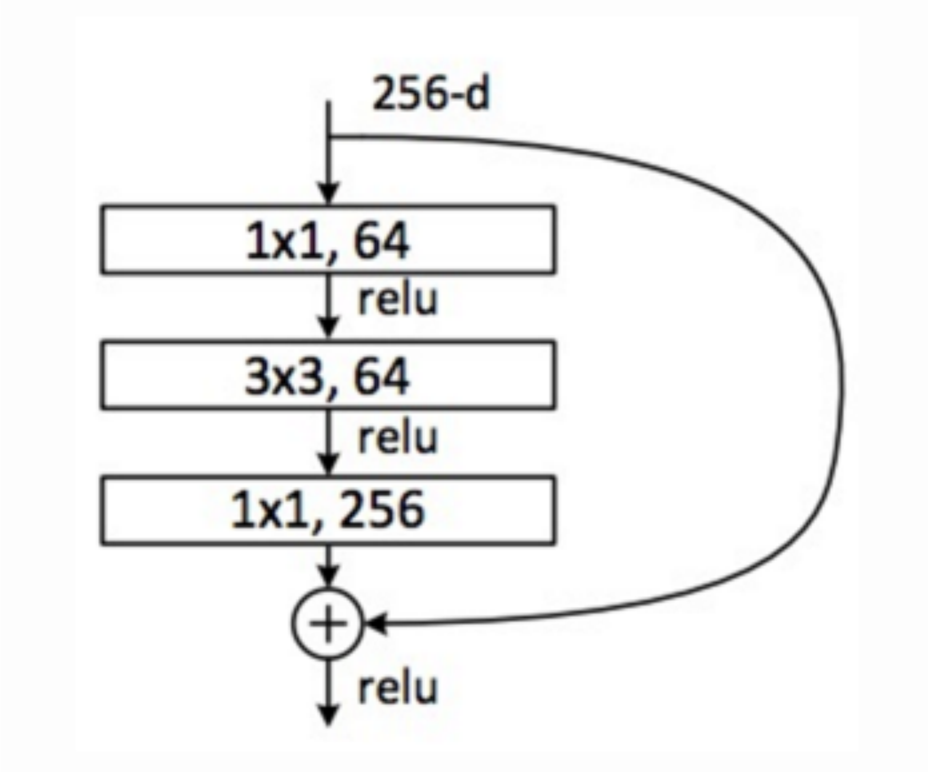
Google Inception (2014,15,16,17)



keras-inception.py

Questions:
Can you trick this model?

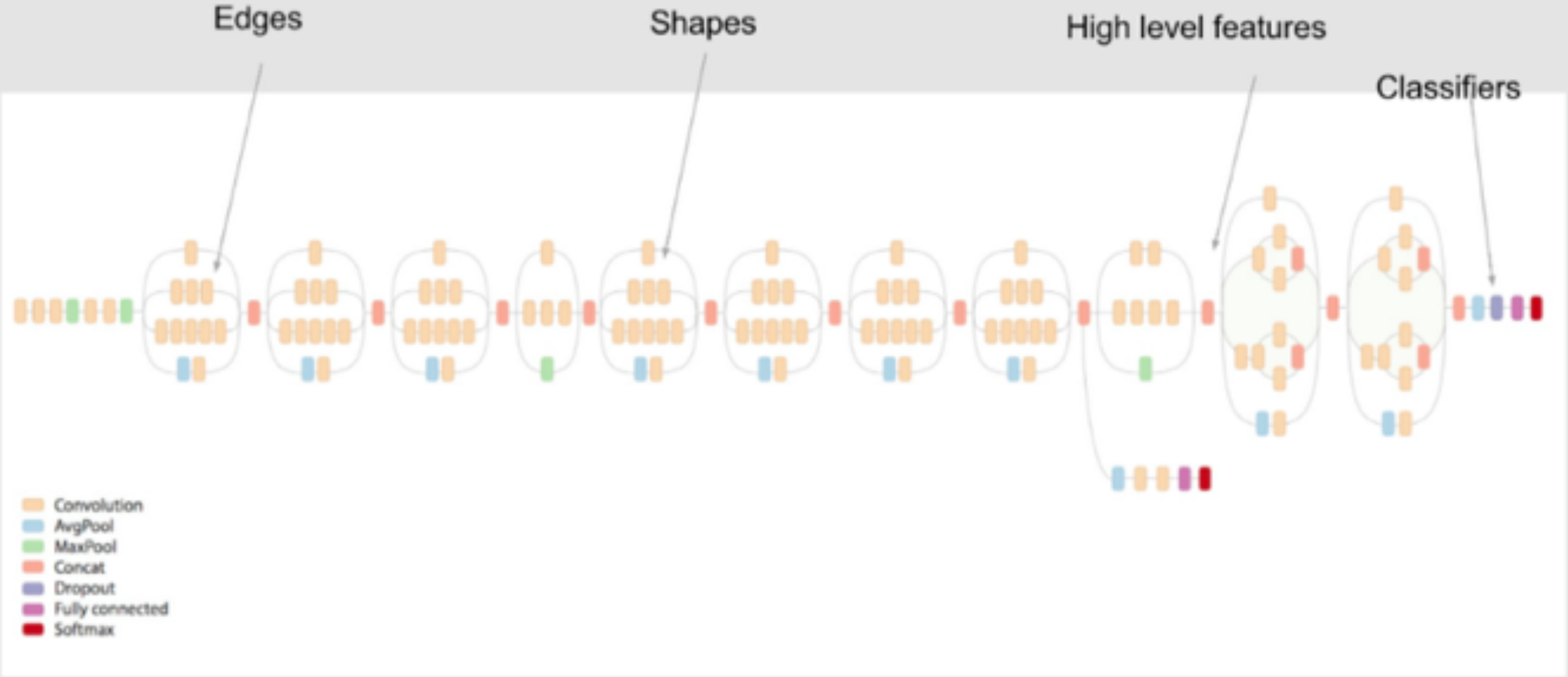
Resnet



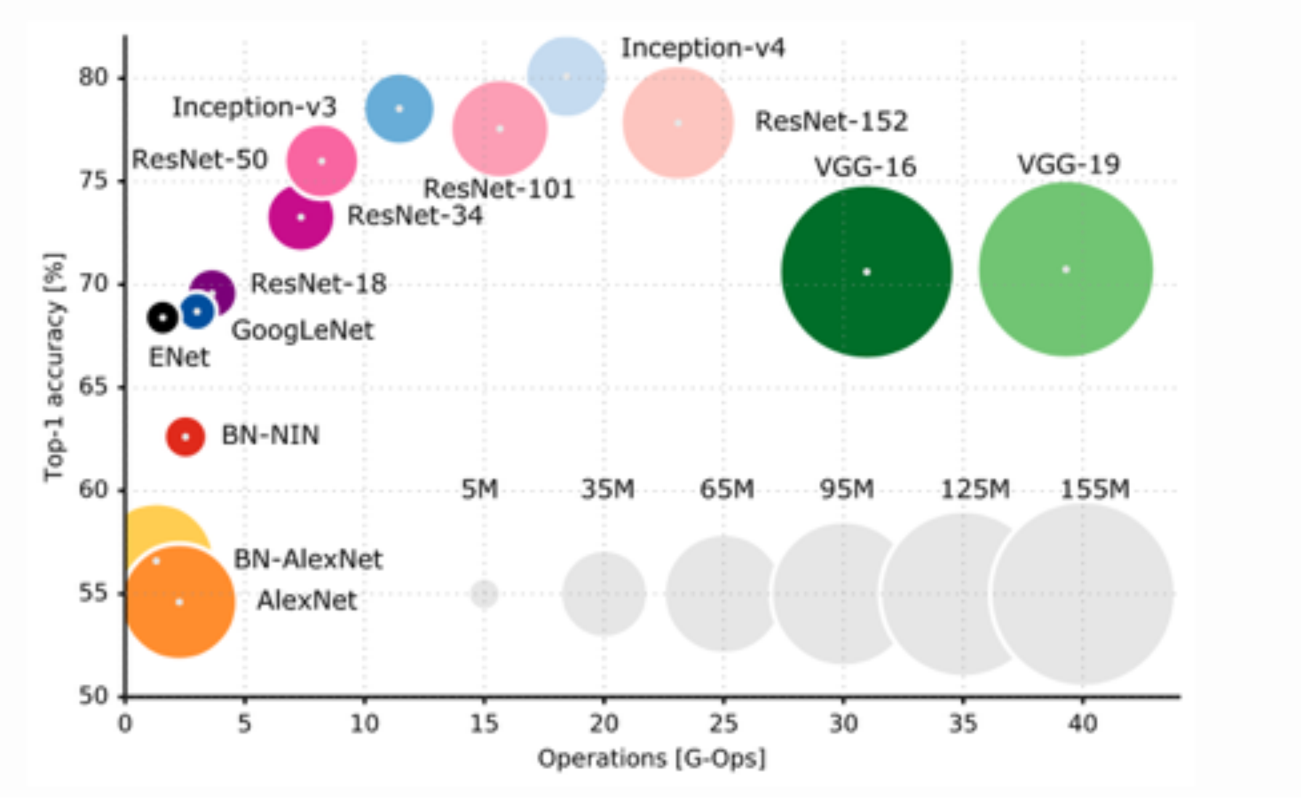
keras-resnet.py

Questions:
How is this model different in performance to inception?

What does the layers learn?



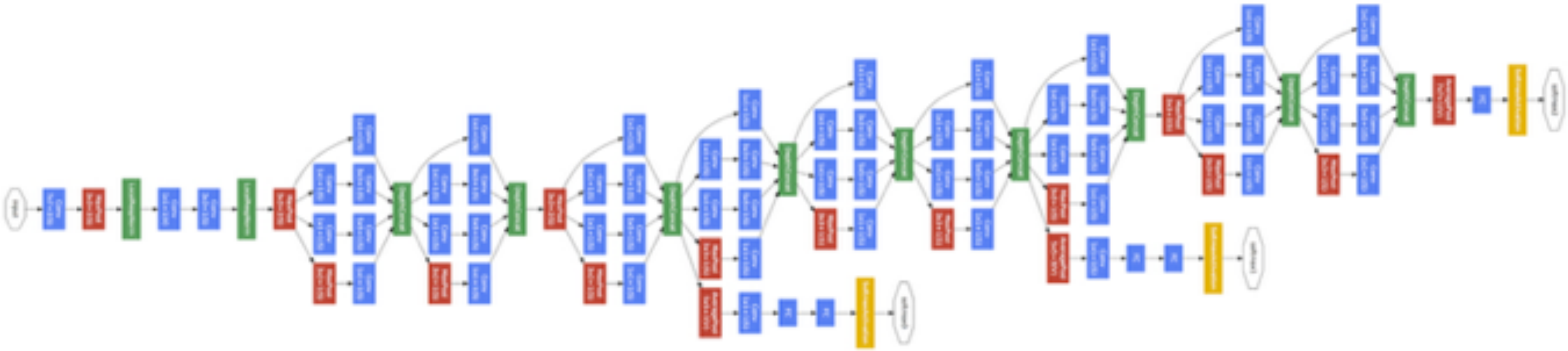
Overview of Neural Networks



Transfer Learning

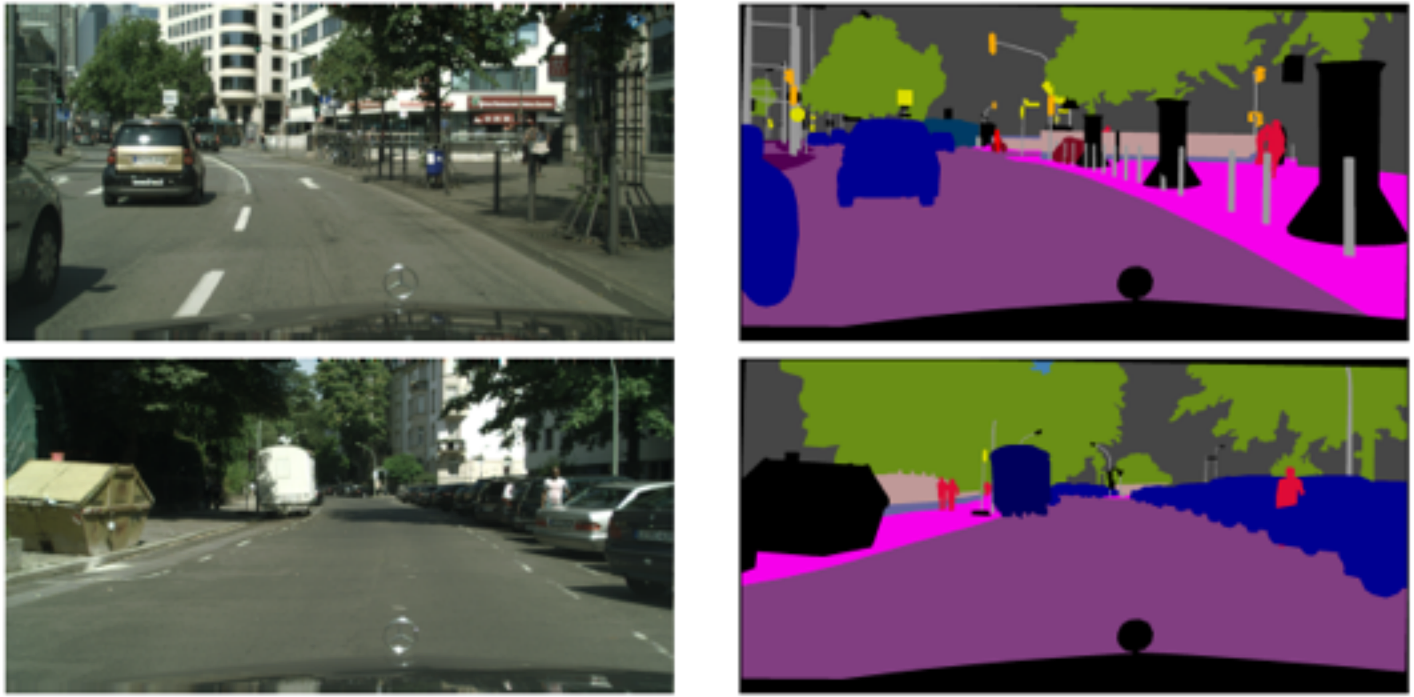
Code https://github.com/fchollet/deep-learning-models

Overview http://sebastianruder.com/transfer-learning/

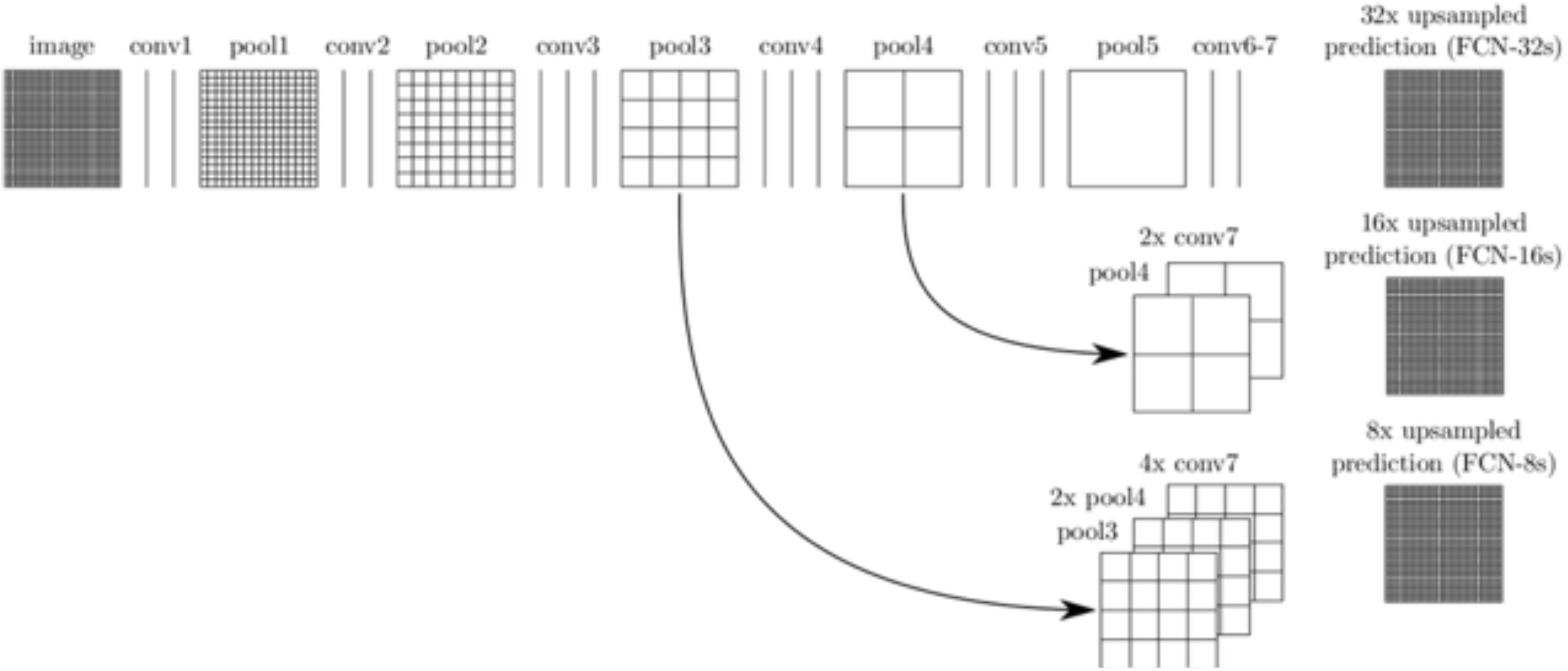


Architecture of Google's inception network

Using Models for Other Applications



FCN Architecture




LG

Module 6 End

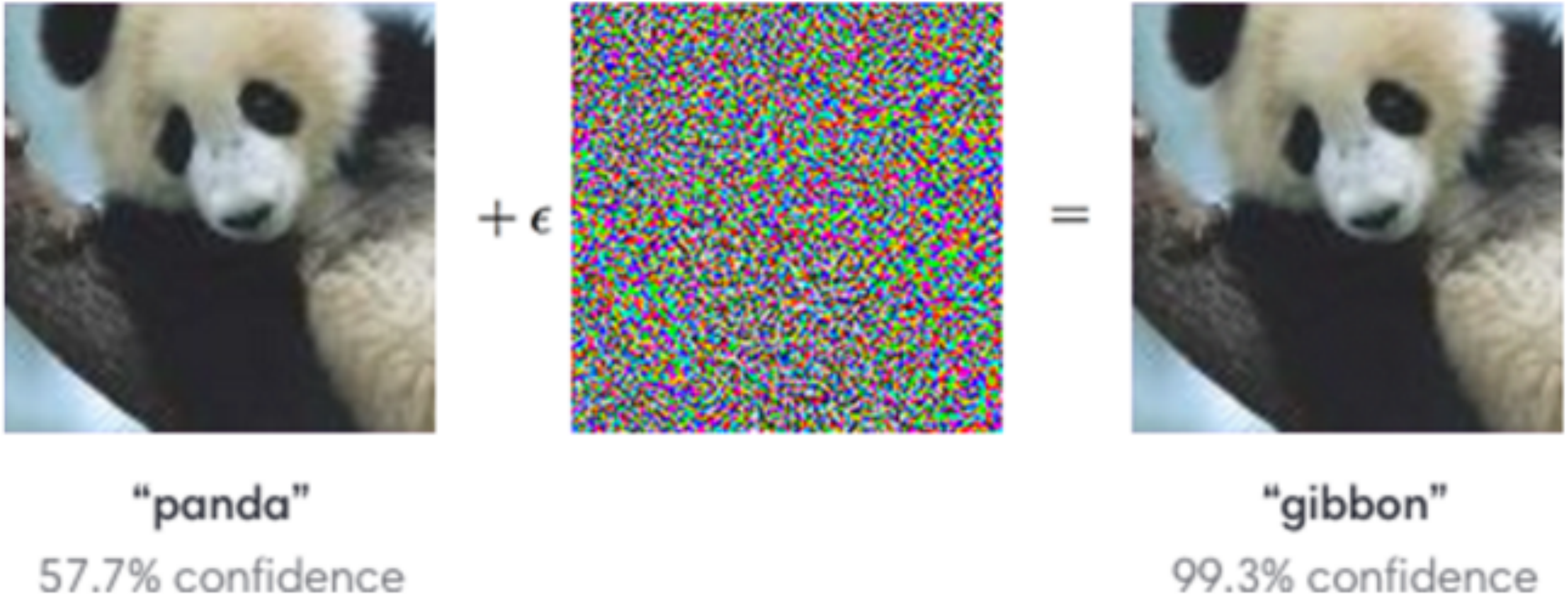
Goals:
Convolutional Neural Networks

Questions:
What is a convolution?
What is max pooling?
Why?

Common Pitfalls:
Training can be slow
Many details to get right
Accidentally cause the singularity?

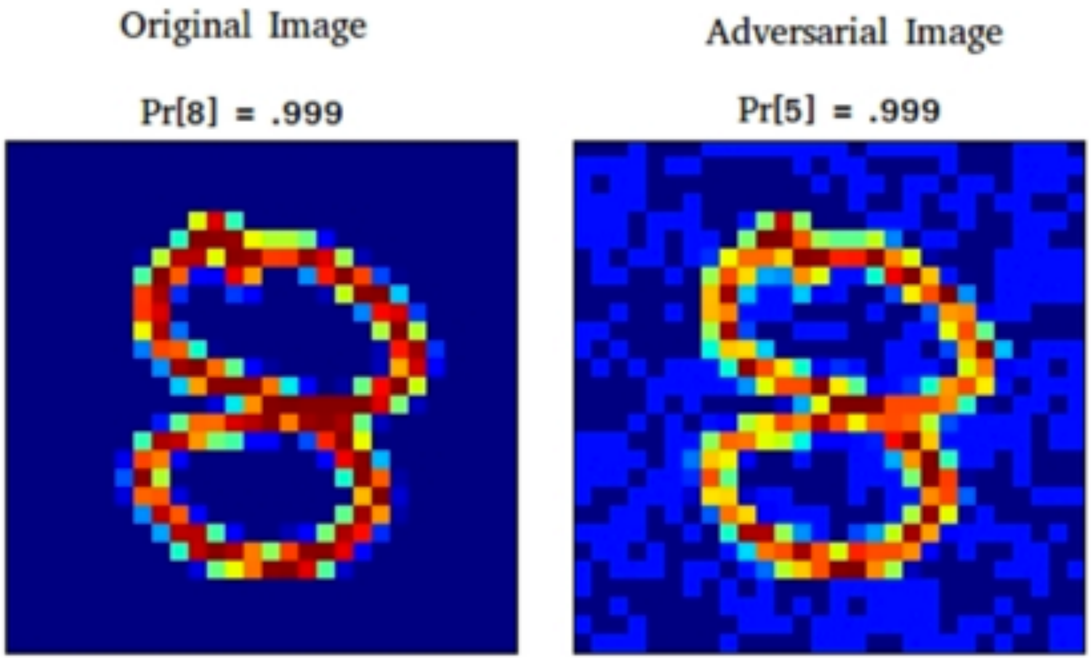
Where do we go from here?

Adversarial Examples and Adversarial Networks

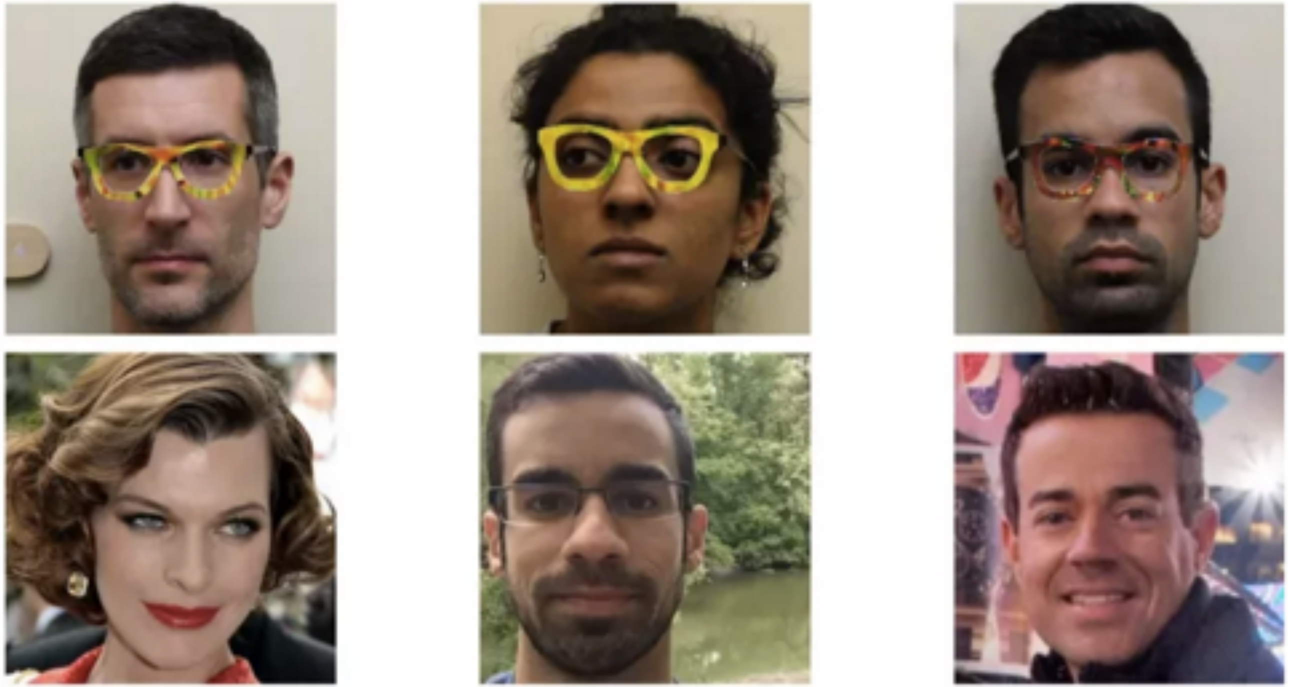


+ ϵ
=
"panda"
57.7% confidence
"gibbon"
99.3% confidence

https://github.com/jmgilmer/AdversarialMNIST



Glasses Fooling Face Recognition

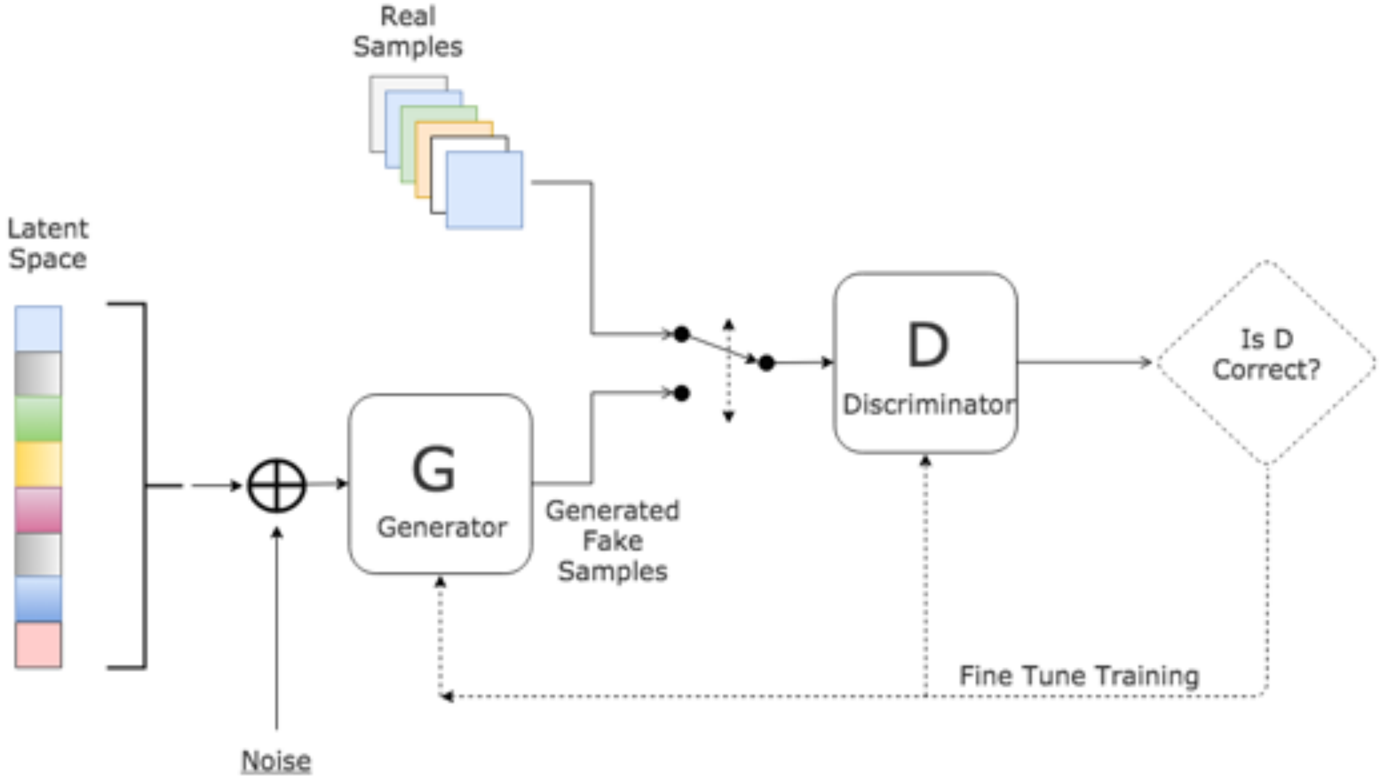


Researchers wearing simulated pairs of fooling glasses, and the people the facial recognition system thought they were.
| Image by Mahmood Sharif, Sruti Bhagavatula, Lujo Bauer, and Michael K. Reiter

You Can Trick Self-Driving Cars by Defacing Street Signs

By Catalin Cimpanu August 7, 2017 07:31 AM 6

Generative Adversarial Network



http://bit.ly/QCAIfeedback1

Learn More

- Books
  - Deep Learning Book (http://www.deeplearningbook.org/)
  - Artificial Intelligence: A Modern Approach
- Classes
  - Stanford CS229, CS231n
  - Fast.AI
  - Udacity/Coursera
  - My classes (lukasbiewald.com, doloreslabs.com)
- Hands-on
  - kaggle.com
- Facebook Group
  - Dolores Labs

Survey:
http://bit.ly/QCAIfeedback1