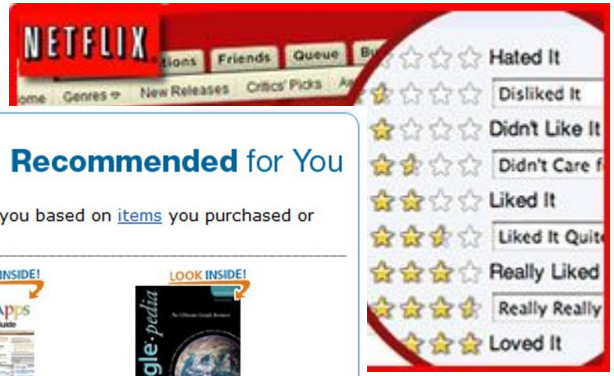
What is Machine Learning, Really?

Scott Kirkland
AppDev SIG 12/12/17

Artificial Intelligence


36
32
3
Stinger Minigun
78
98


www.zdsoft.com
100



amazon.com

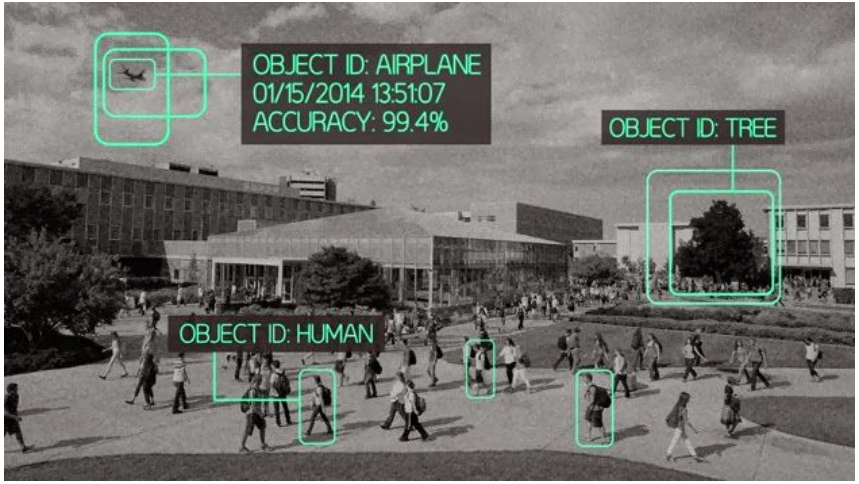
Recommended for You

Amazon.com has new recommendations for you based on items you purchased or told us you own.

Google Apps Deciphered: Compute in the Cloud to Streamline Your Desktop

Google Apps Administrator Guide: A Private-Label Web Workspace

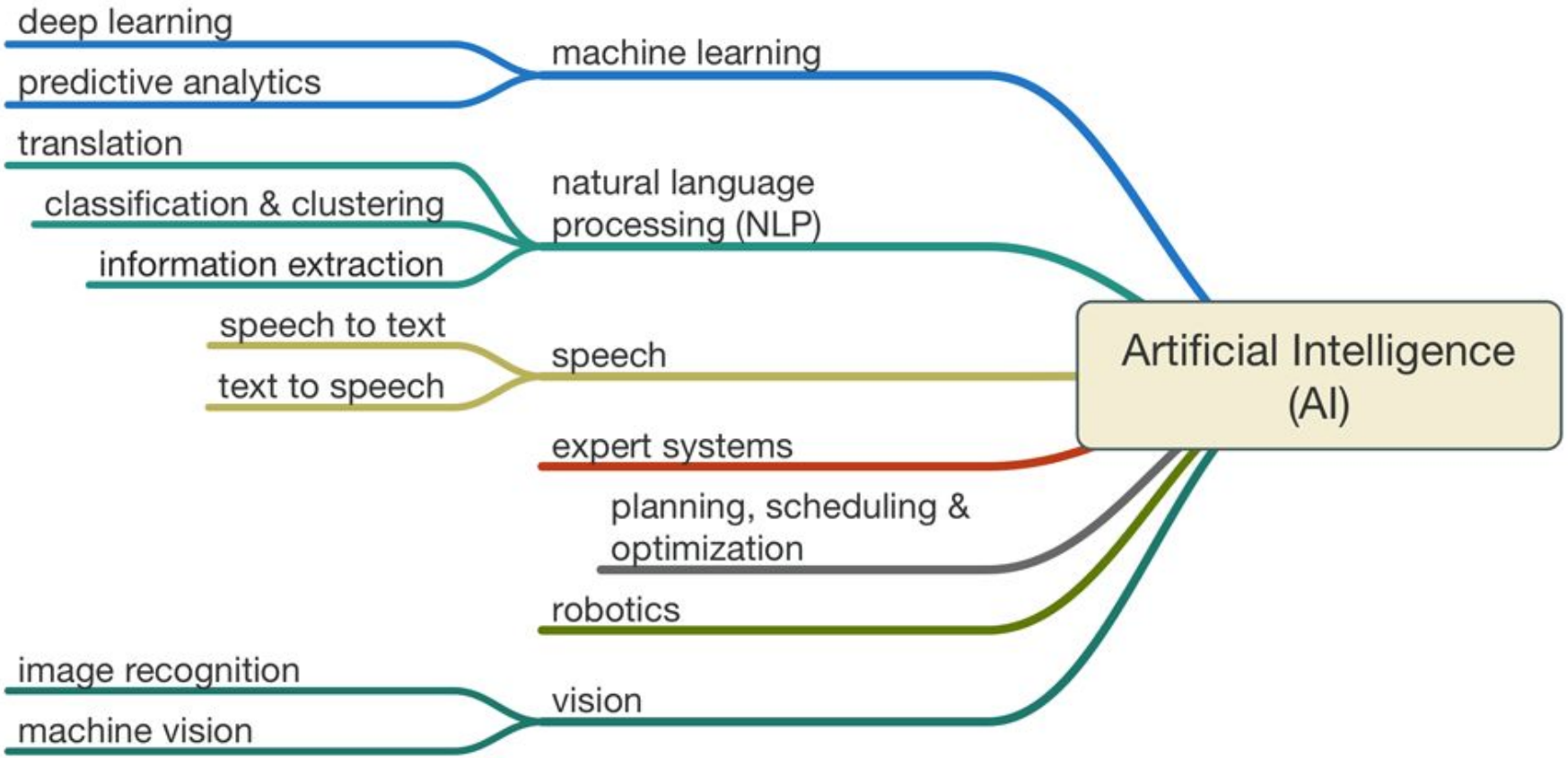
Googlepedia: The Ultimate Google Resource (3rd Edition)



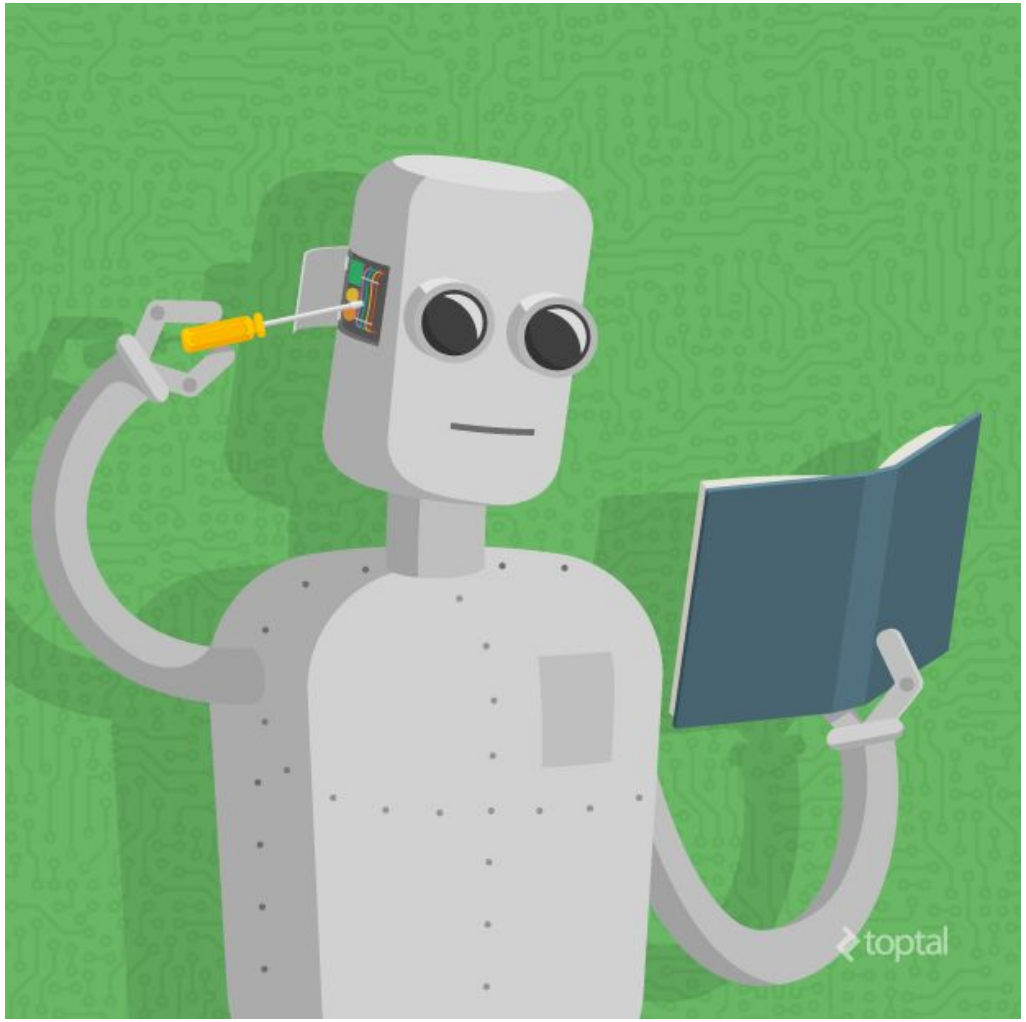
amazon alexa

works with the Google Assistant

Artificial Intelligence is the ability of a digital computer or robot to perform tasks commonly associated with intelligent beings

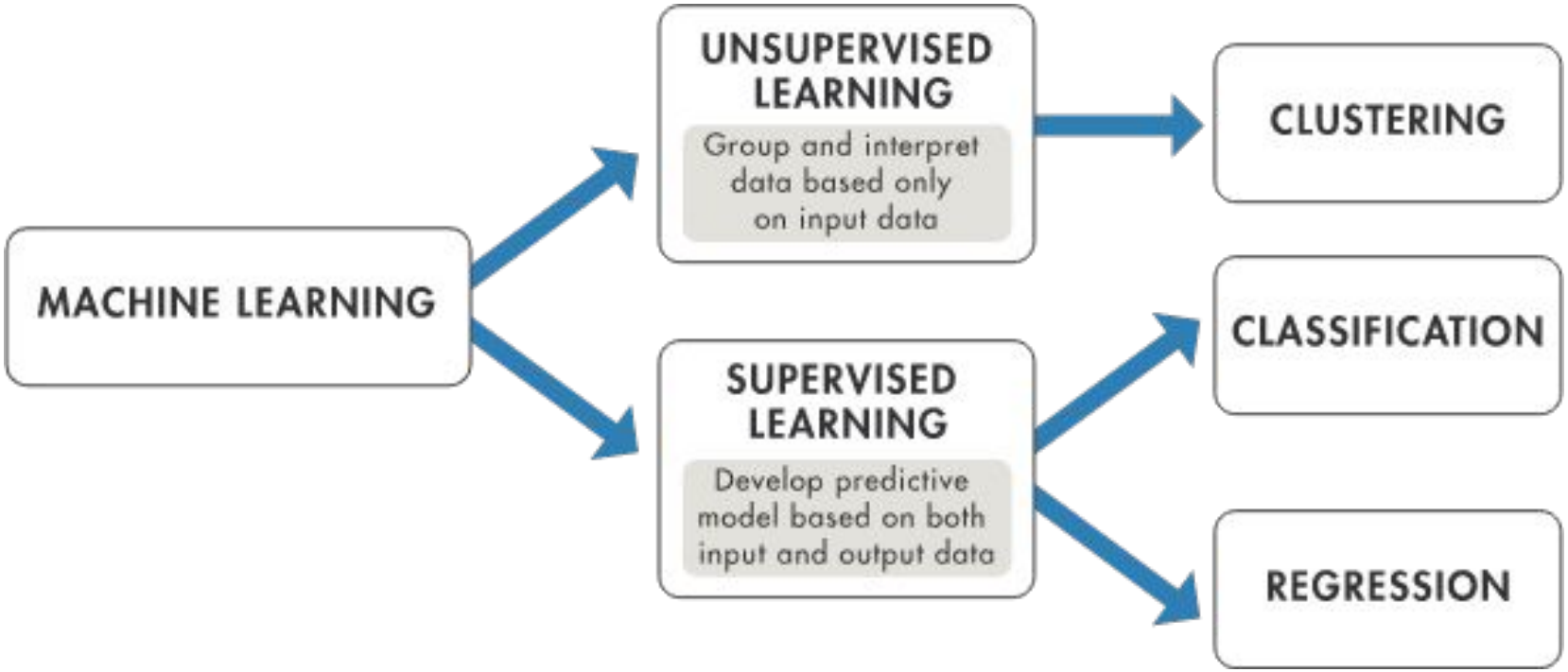

Artificial Intelligence (AI)
machine learning
deep learning
predictive analytics
natural language processing (NLP)
translation
classification & clustering
information extraction
speech
speech to text
text to speech
expert systems
planning, scheduling & optimization
robotics
vision
image recognition
machine vision

What is Machine Learning?


toptal


MACHINE
LEARNING
acer
acer

Machine learning is a field of computer science that gives computers the ability to learn without being explicitly programmed


MACHINE LEARNING
UNSUPERVISED LEARNING
Group and interpret data based only on input data
SUPERVISED LEARNING
Develop predictive model based on both input and output data
CLUSTERING
CLASSIFICATION
REGRESSION

Example: Student Performance

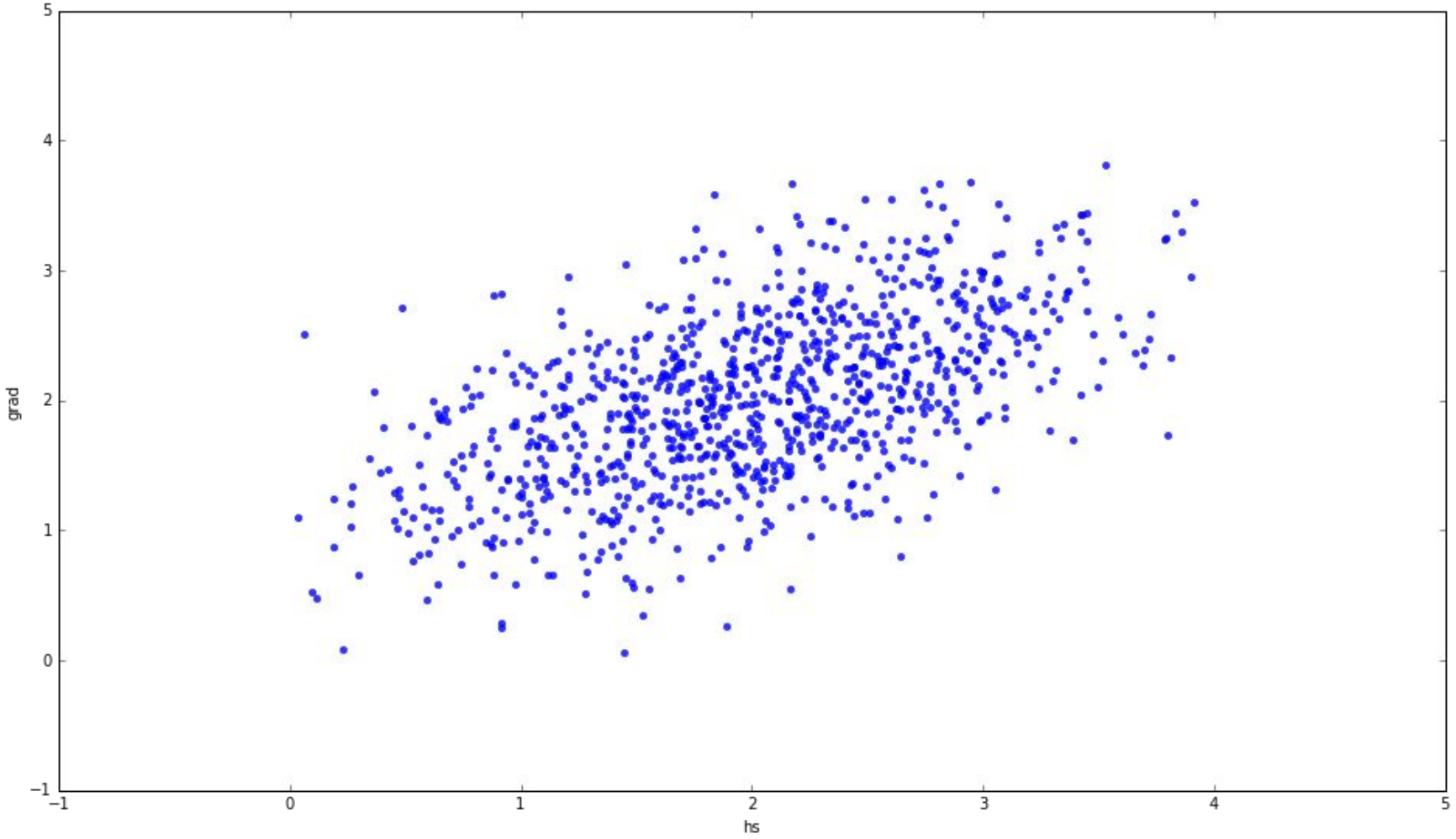
- You have the high school + college GPA of every student ever at UCD
- You want to know, given a new applicant, what their future GPA at UCD is likely to be

HIGHSCHOOL GPA	GRADUATED GPA
3.3	3.5
2.3	3.0
3.9	3.5
2.6	2.2
2.0	2.5
3.4	3.8
2.2	2.5
3.5	3.7
2.9	3.1

X

Y

HIGHSCHOOL GPA	GRADUATED GPA
3.3	3.5
2.3	3.0
3.9	3.5
2.6	2.2
2.0	2.5
3.4	3.8
2.2	2.5
3.5	3.7
2.9	3.1


grad
hs
5
4
3
2
1
0
−1
−1
0
1
2
3
4
5

$$h(x) = \hat{y}$$ 

This is what we want

A magic function

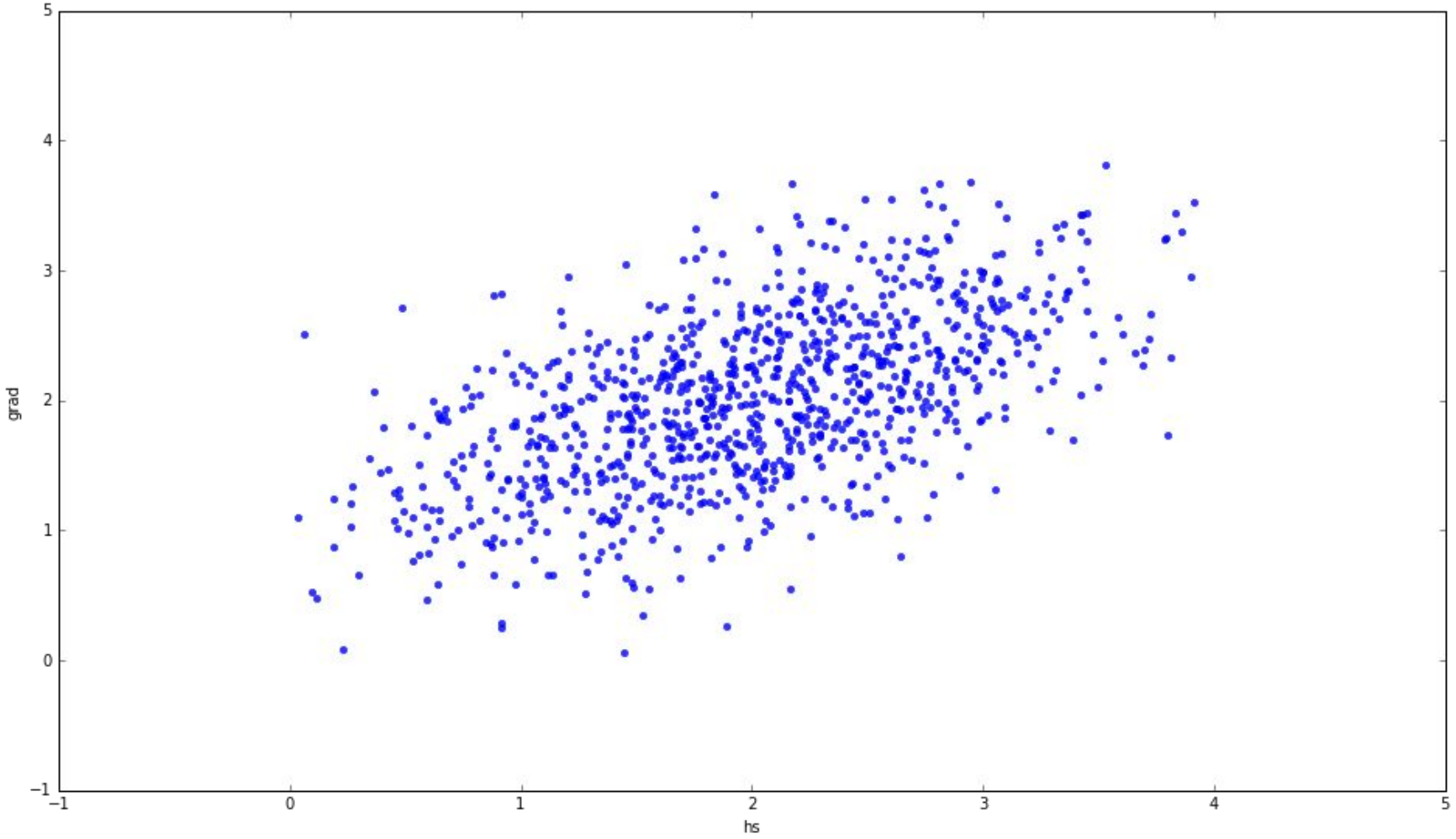
The goal of ML is never to make "perfect" guesses, because ML deals in domains where there is no such thing. The goal is to make guesses that are good enough to be useful.

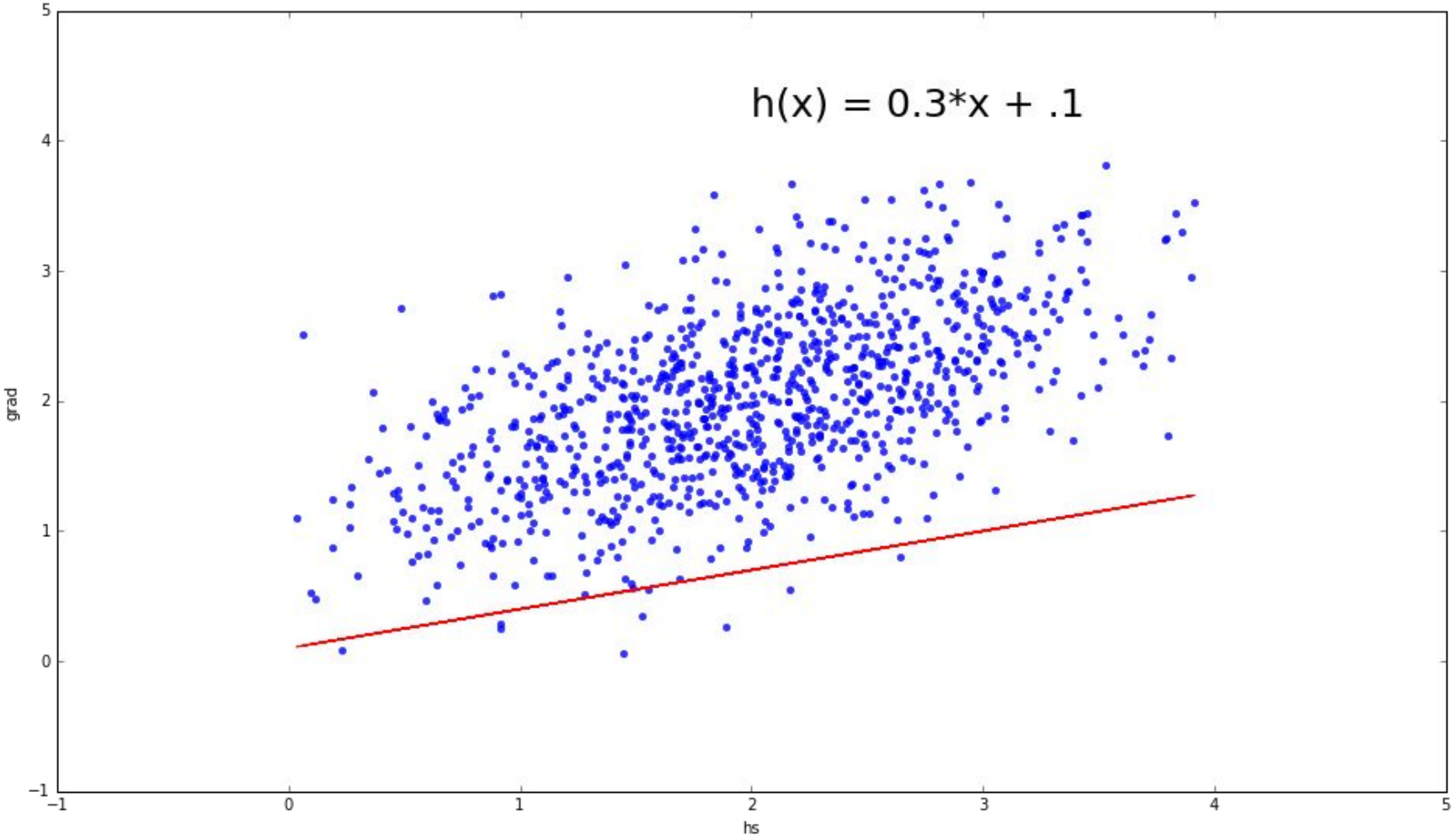
Technique #1

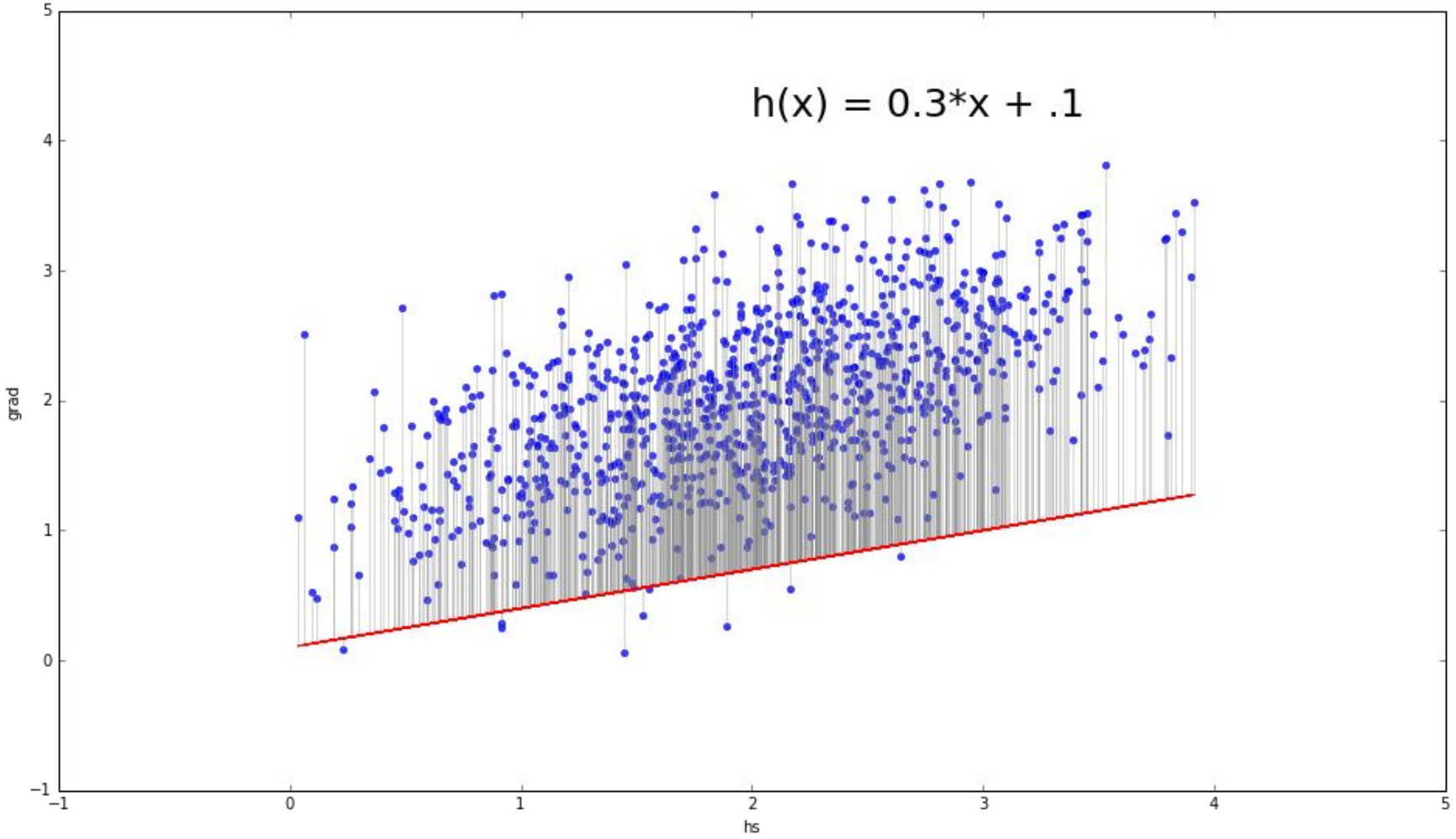
Linear Regression

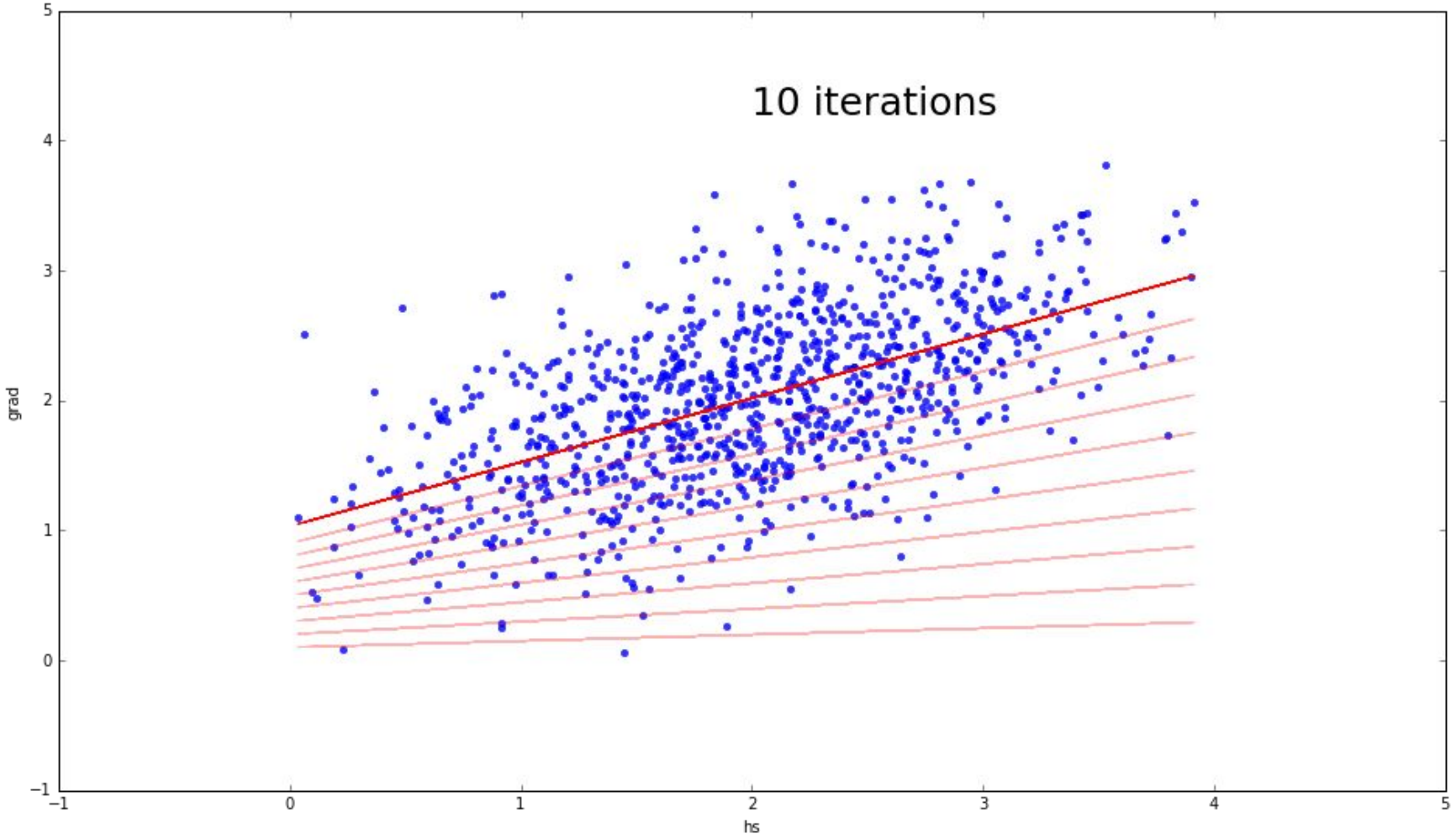
$$h(x) = W^{*}x + b$$

Linear Regression


grad
hs
5
4
3
2
1
0
−1
−1
0
1
2
3
4
5


h(x) = 0.3*x + .1
grad
hs
5
4
3
2
1
0
−1
−1
0
1
2
3
4
5


h(x) = 0.3*x + .1
grad
hs
5
4
3
2
1
0
−1
−1
0
1
2
3
4
5


10 iterations
grad
hs
5
4
3
2
1
0
−1
−1
0
1
2
3
4
5

How do we iterate?

First, we need to figure out what it means for our model to be “better”

$$J(W,b) = \Sigma(\hat{Y} - Y)^2$$

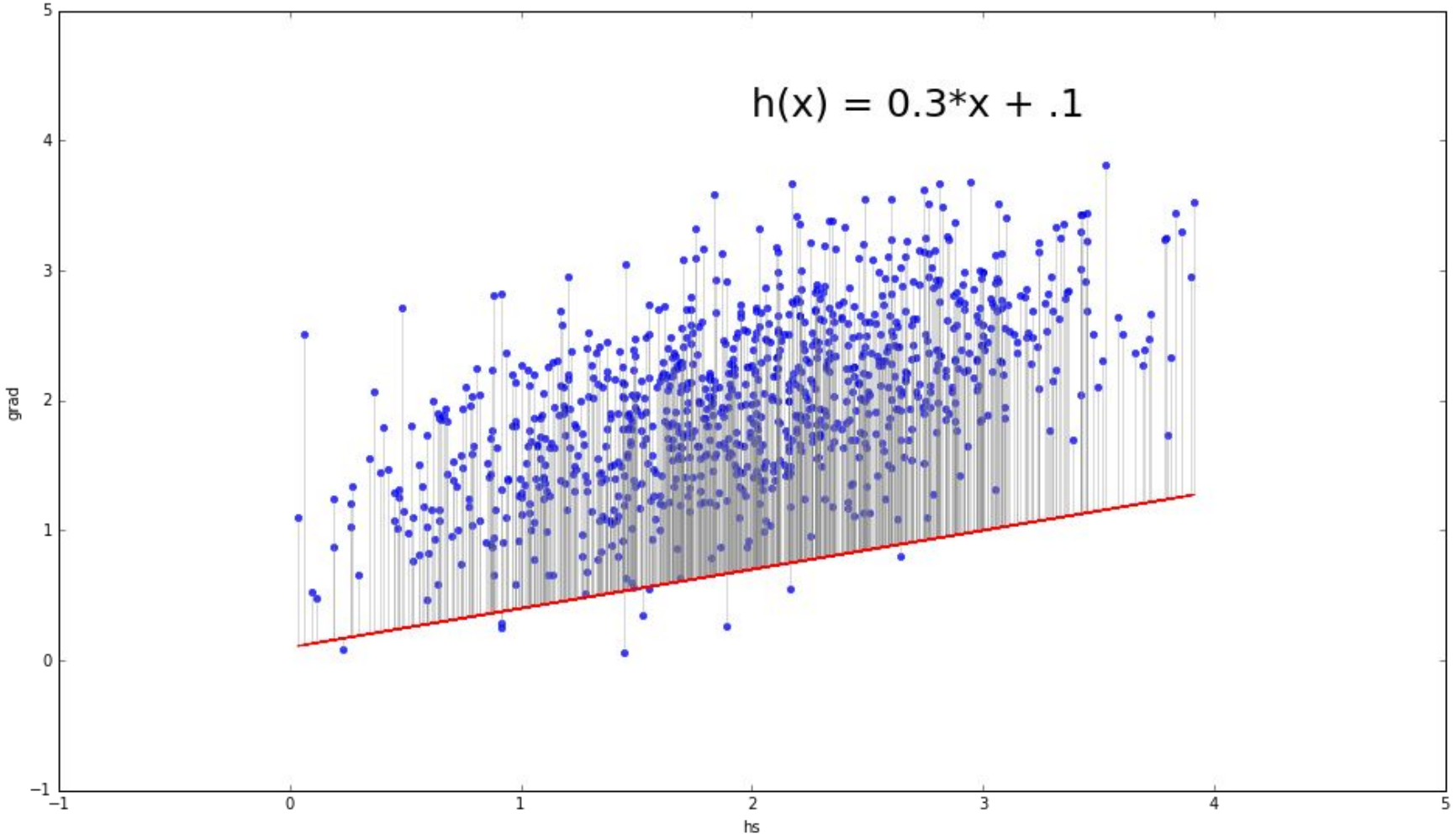
Cost Function

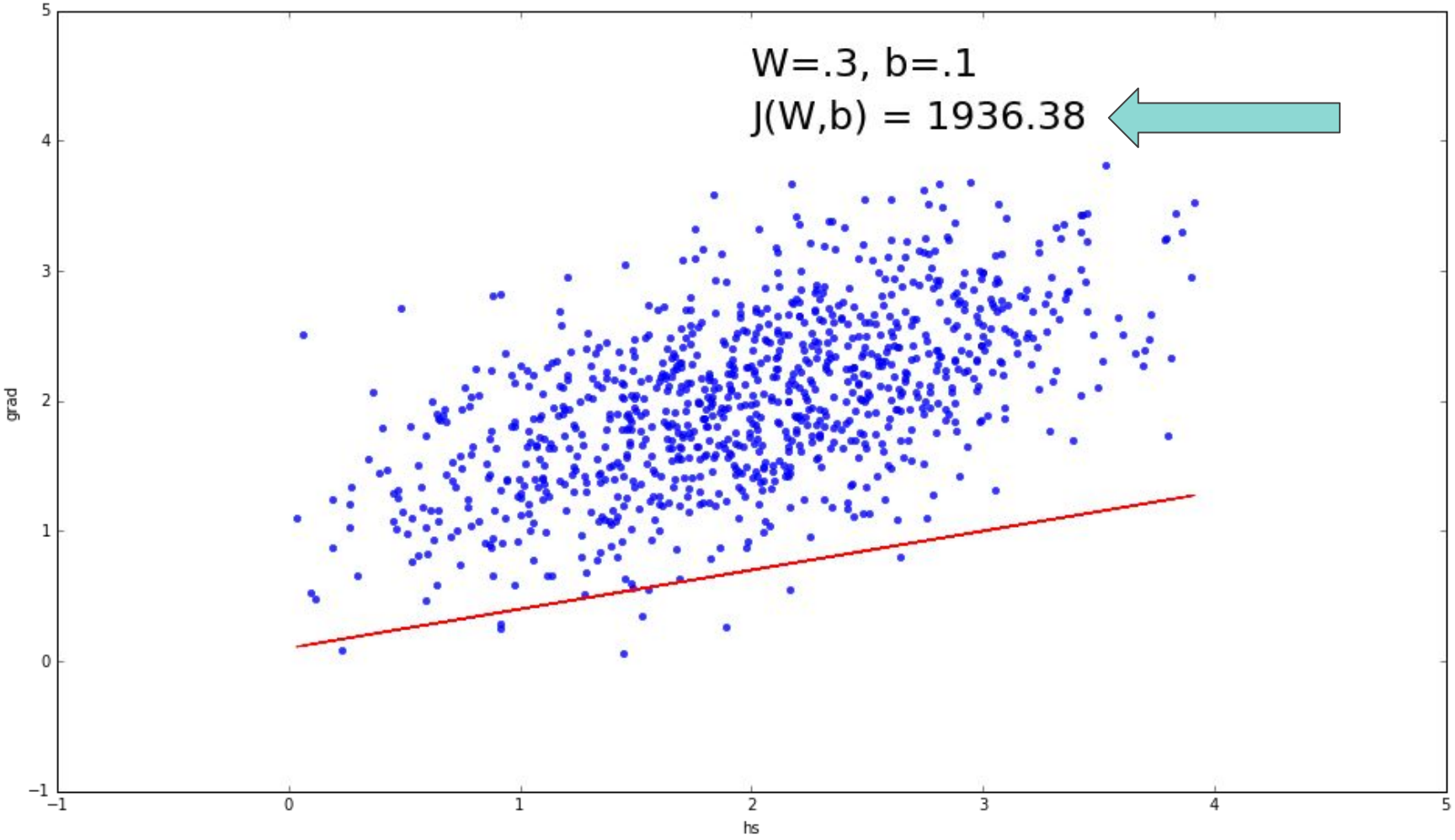
Aka: How wrong are we?

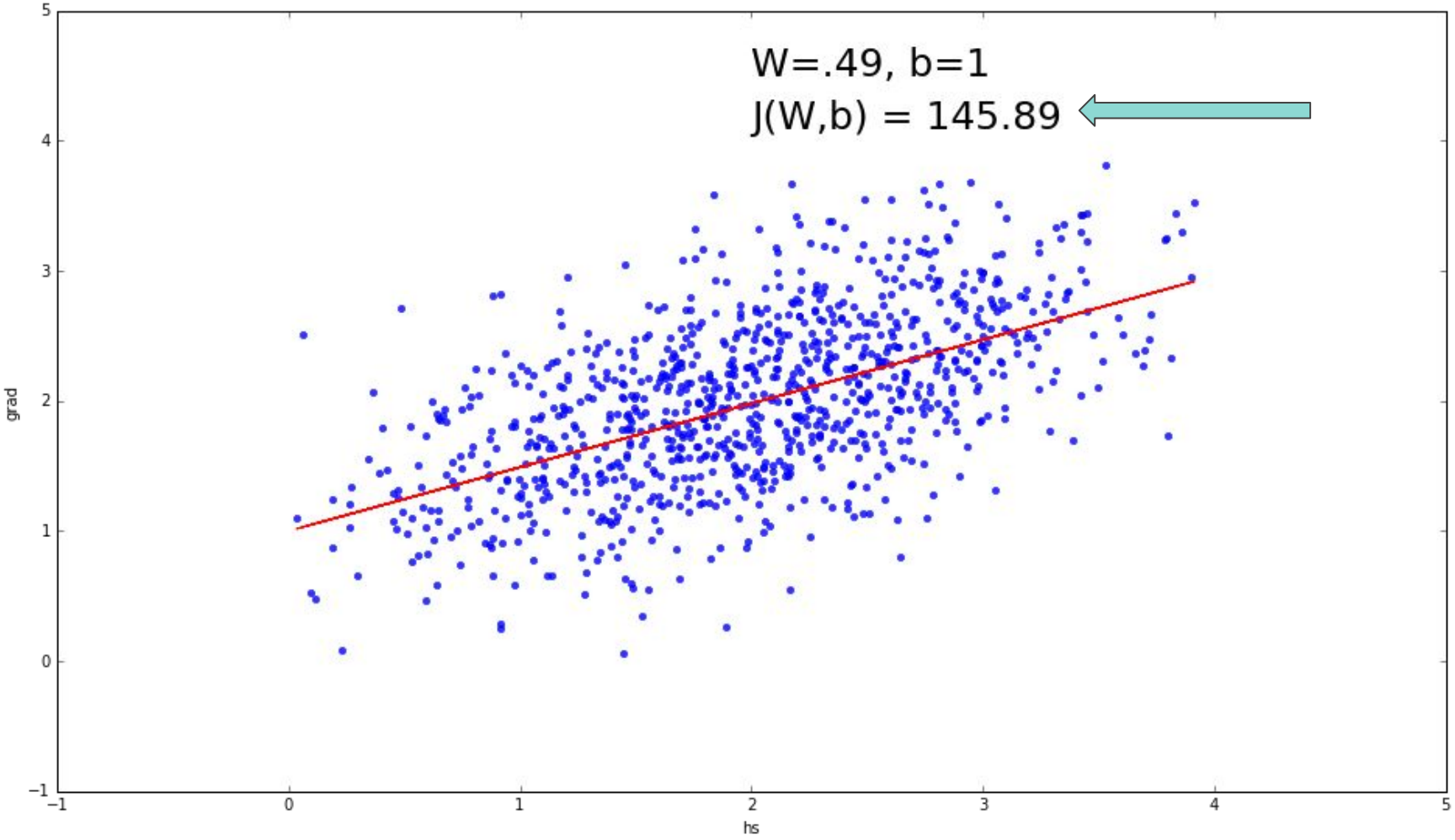
Cost Function = $J(W,b) = \Sigma(\hat{Y}-Y)^2$

- Basically we add up how different our guess was from the real value for every row of data

- If our line isn't very good, we get a big value


h(x) = 0.3*x + .1
grad
hs
5
4
3
2
1
0
−1
−1
0
1
2
3
4
5

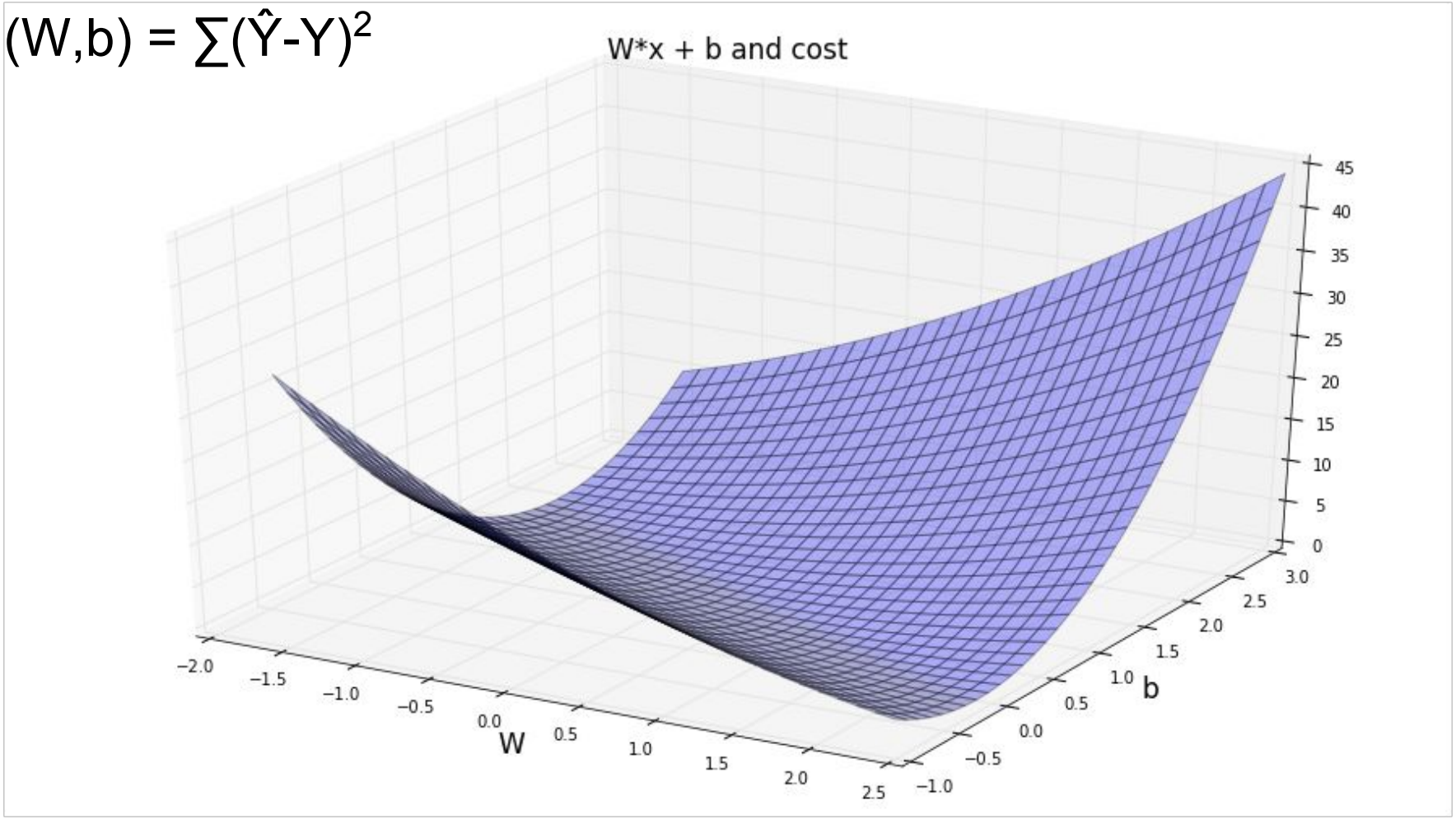

W=.3, b=.1
J(W,b) = 1936.38
grad
hs
5
4
3
2
1
0
−1
−1
0
1
2
3
4
5


W=.49, b=1
J(W,b) = 145.89
grad
hs
5
4
3
2
1
0
−1
−1
0
1
2
3
4
5

How do we go from a bad model to a better model?

Gradient Descent

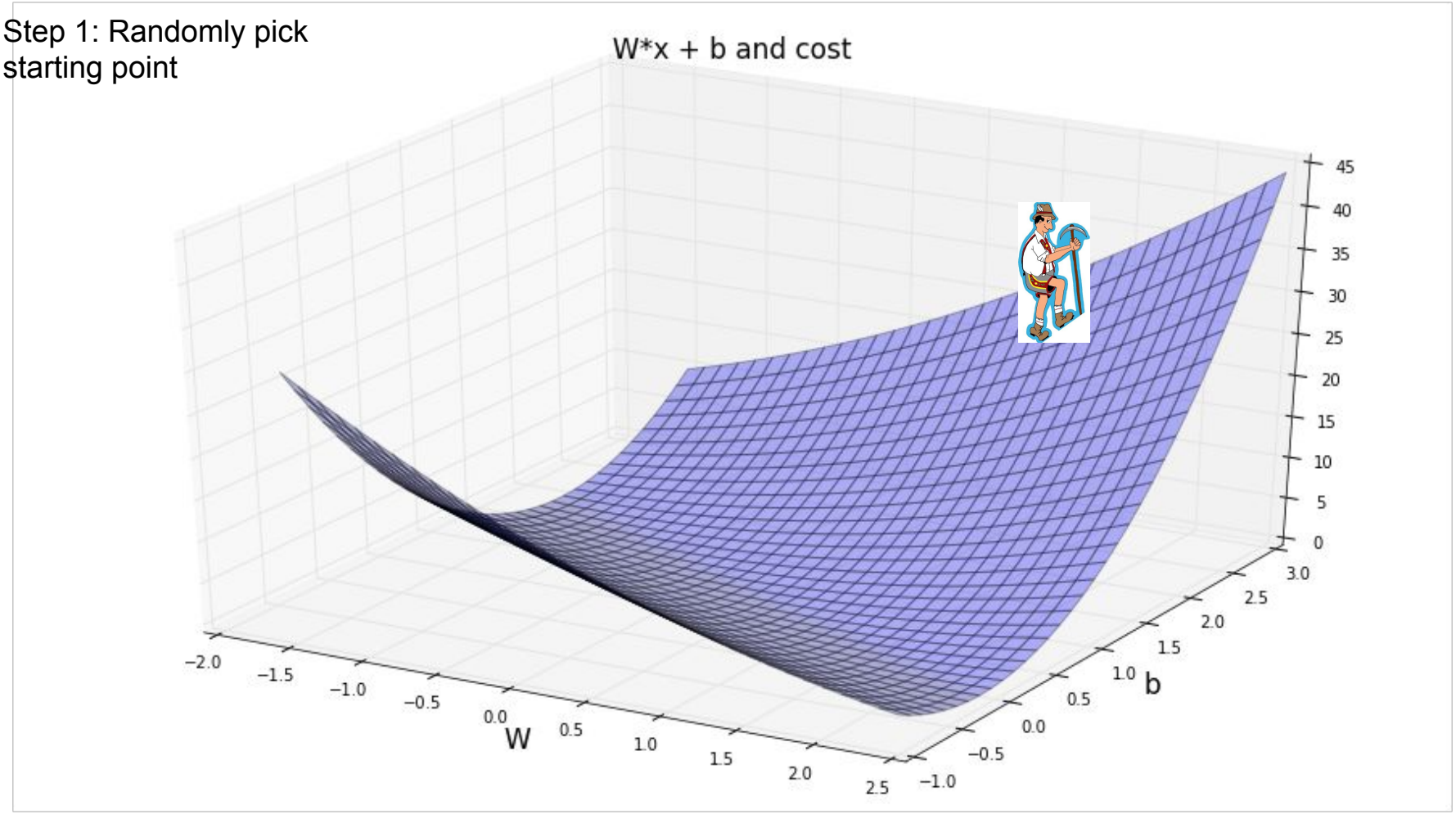
$$J(W,b) = \sum(\hat{Y}-Y)^2$$



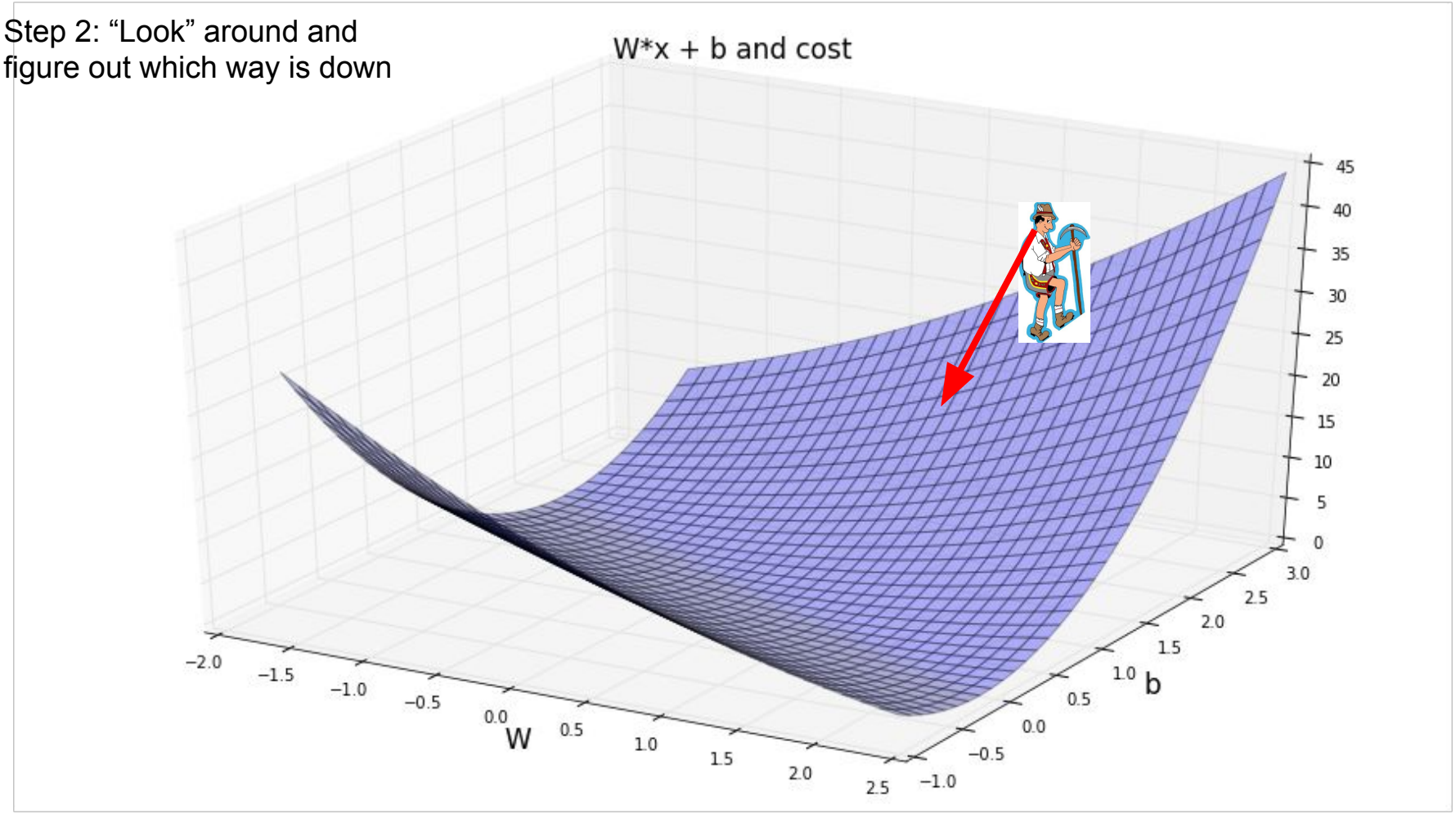
So let’s start by making a guess…

Start by picking random numbers

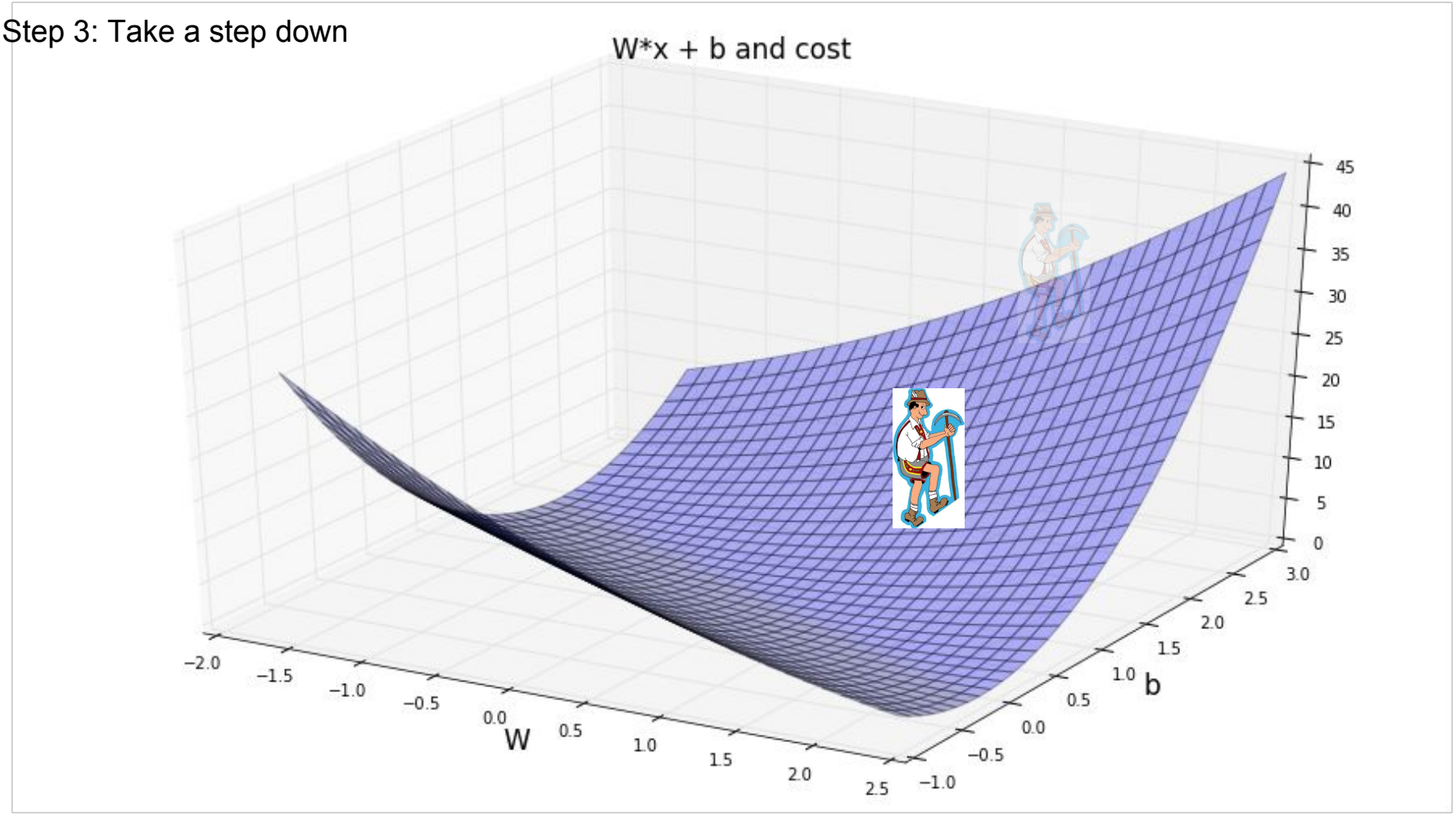
Step 1: Randomly pick starting point



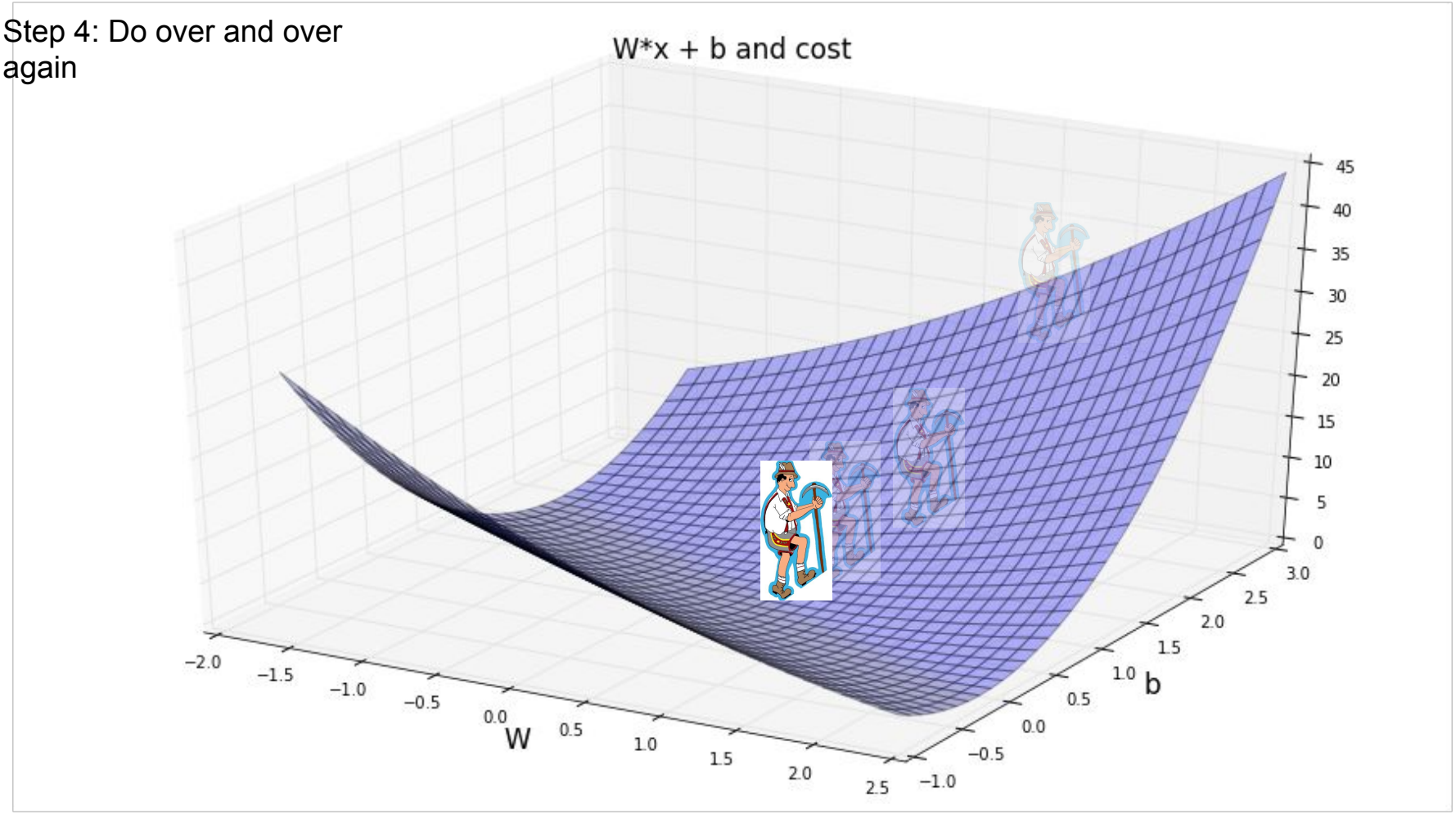
Step 2: “Look” around and figure out which way is down



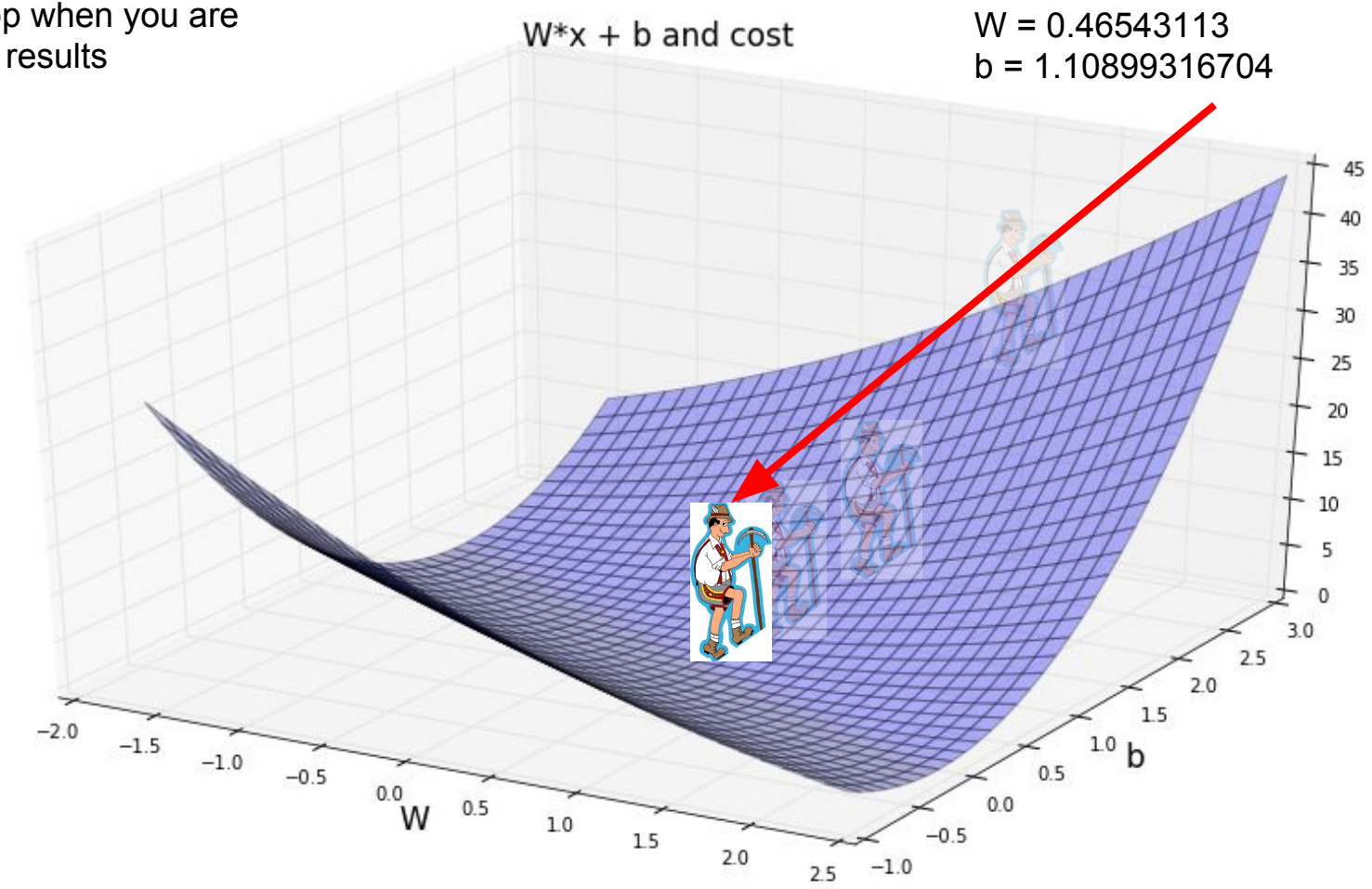
Step 3: Take a step down



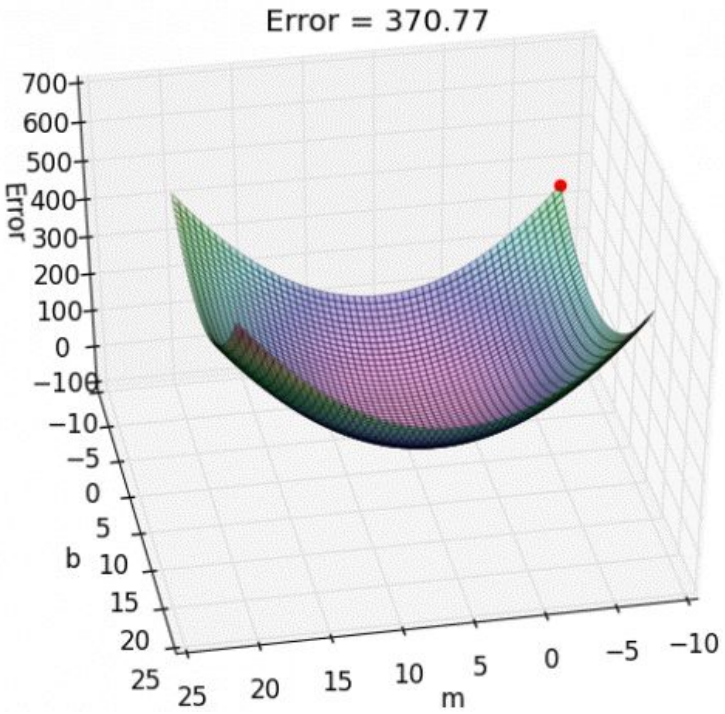
Step 4: Do over and over again

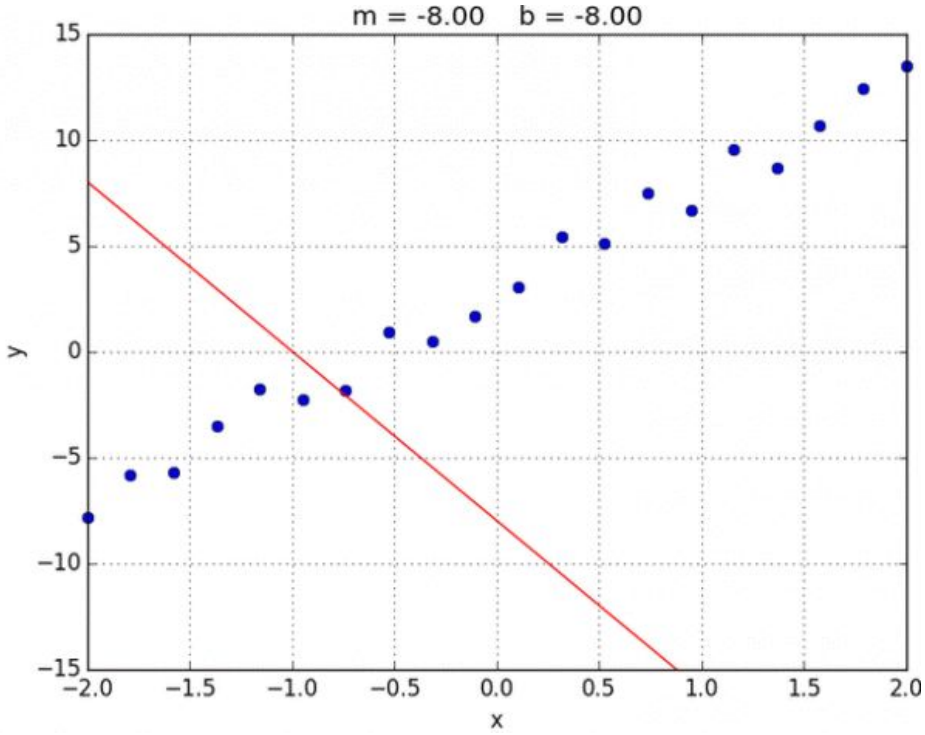


Step 5: Stop when you are happy with results



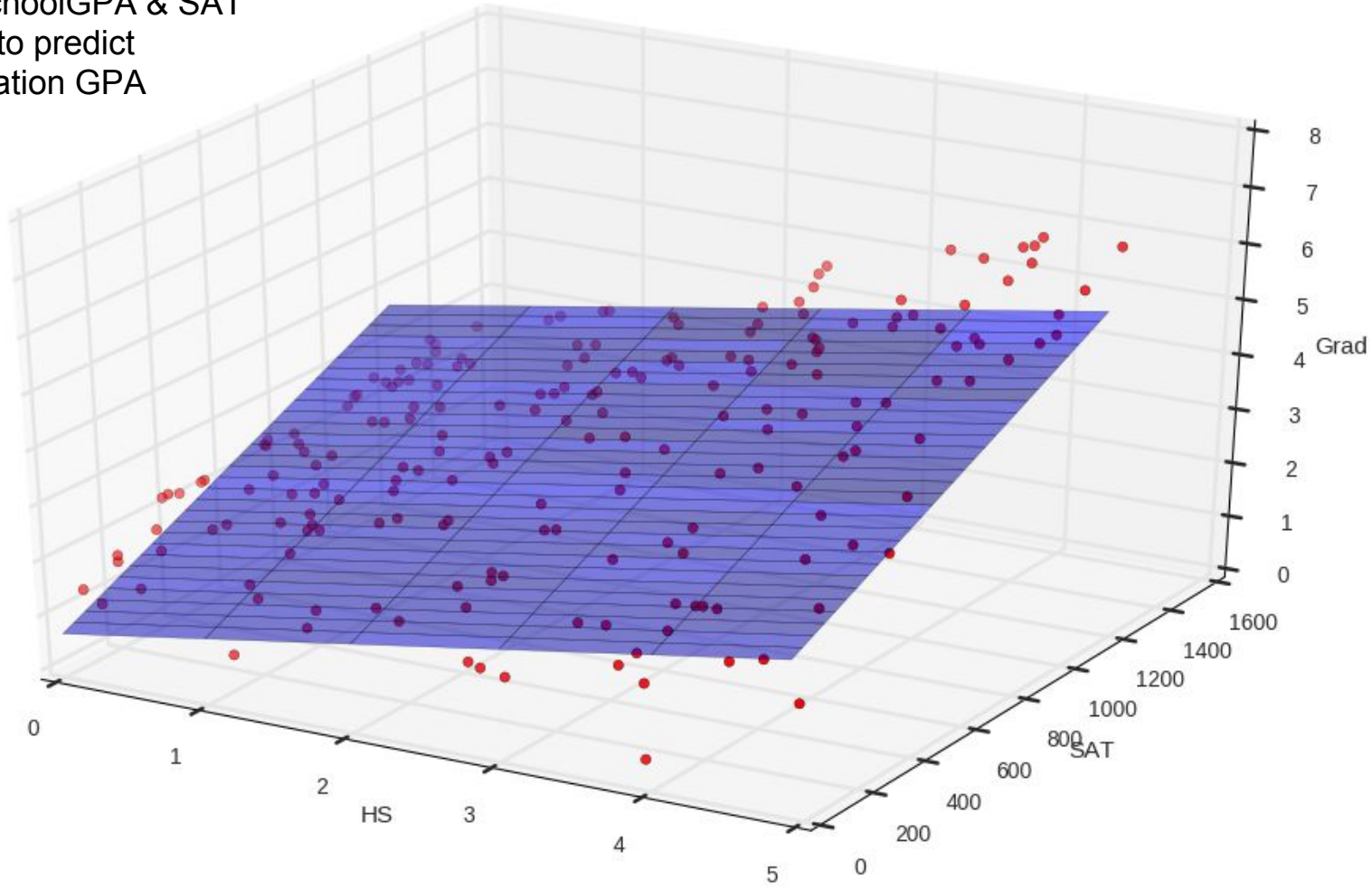
Now our machine has “learned” the solution which minimizes our cost


Error = 370.77
Error
700
600
500
400
300
200
100
0
−100
b
−10
−5
0
5
10
15
20
25
m
25
20
15
10
5
0
−5
−10


m = -8.00 b = -8.00
y
15
10
5
0
−5
−10
−15
x
−2.0
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0

What if we have more features?

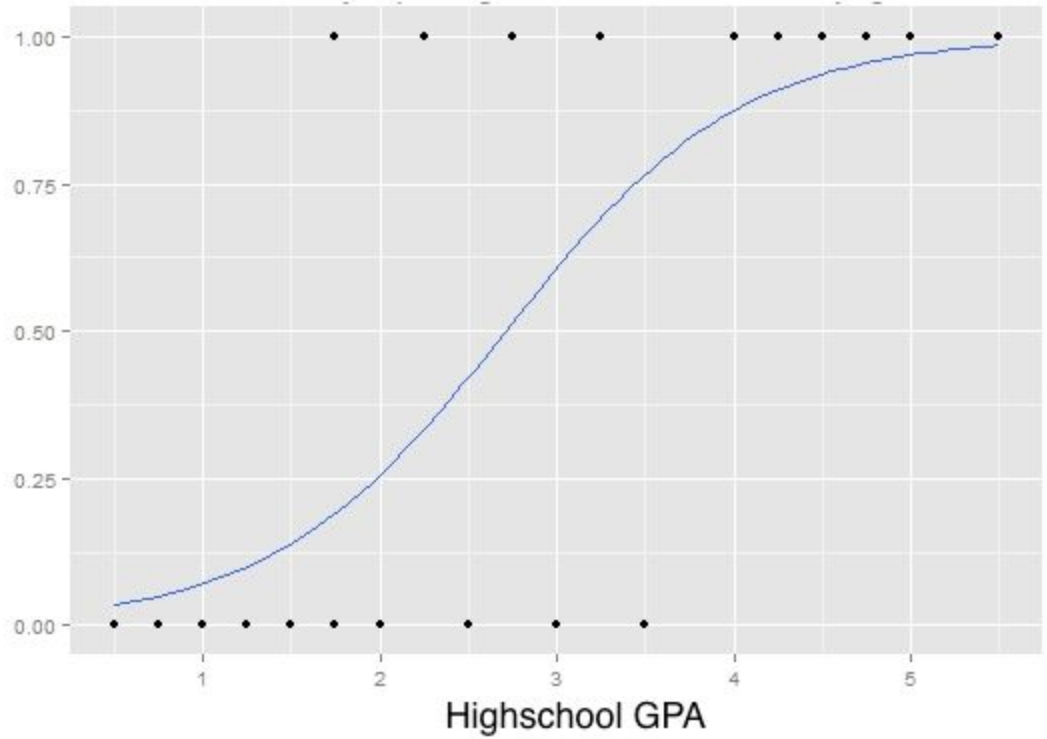
HighschoolGPA & SAT
Score to predict
Graduation GPA



What if we aren’t trying to predict a number?

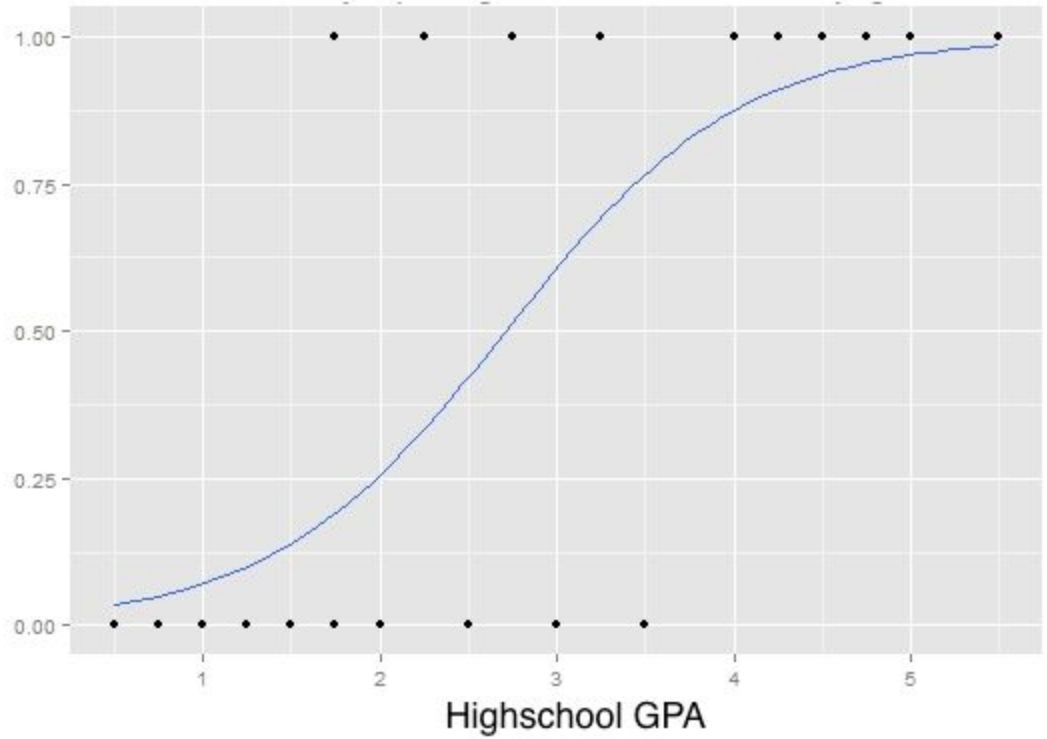
HIGHSCHOOL GPA	GRADUATED Status
3.3	DROPPED
2.3	GRADUATED
3.9	GRADUATED
2.6	GRADUATED
2.0	DROPPED
3.4	GRADUATED
2.2	DROPPED
3.5	GRADUATED
2.9	GRADUATED

HIGHSCHOOL GPA	GRADUATED Status
3.3	0
2.3	1
3.9	1
2.6	1
2.0	0
3.4	1
2.2	0
3.5	1
2.9	1


1.00
0.75
0.50
0.25
0.00
1
2
3
4
5
Highschool GPA

Logistic Regression: $\sigma(t) = \frac{e^t}{e^t + 1} = \frac{1}{1 + e^{-t}}$

Where t = W*x + b



Two Machine Learning problem types

Regression

- Output is a number
- Use linear regression

Classification

- Output is a category (or categories)
- Use logistic regression

What if your relationship is complex?

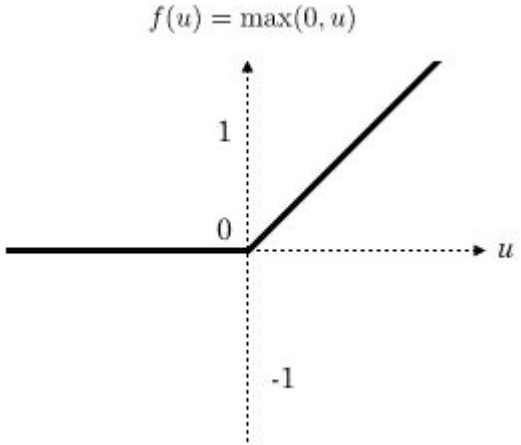
What if your relationship is complex?

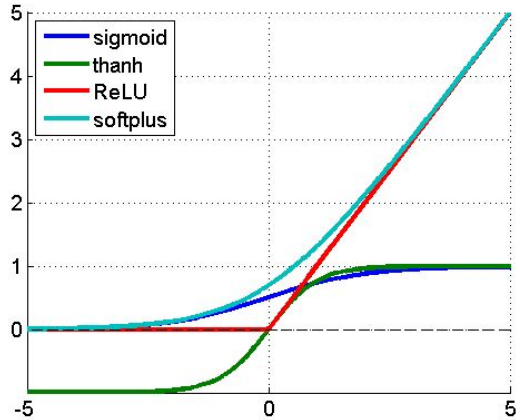
- Add more data features and combine them together

HighSchoolGPA * TransferUnits?

What if your relationship is complex?

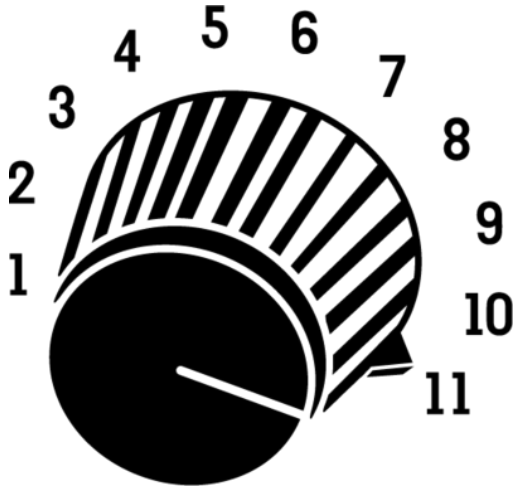
- Make your relationship function more complex.

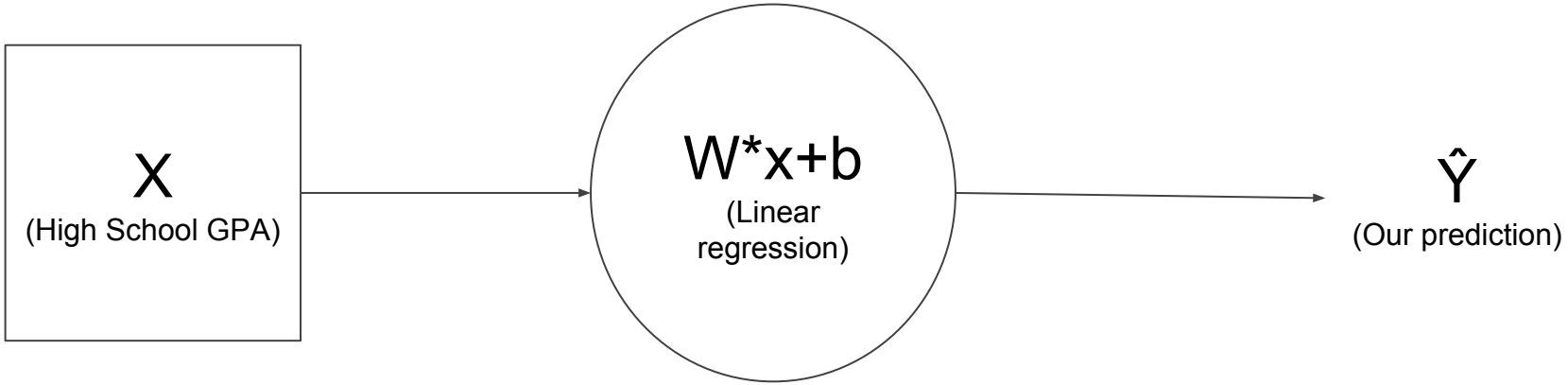


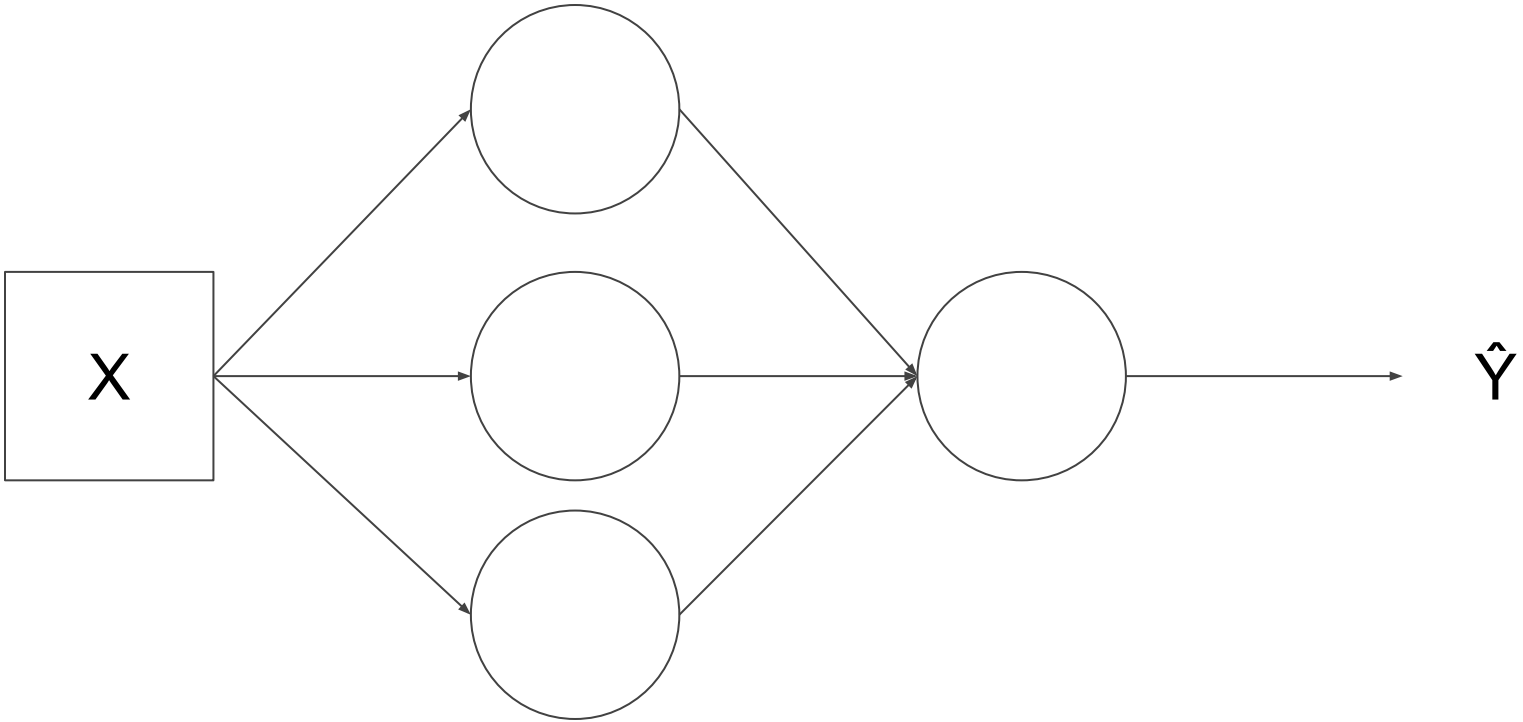


What if your relationship is complex?

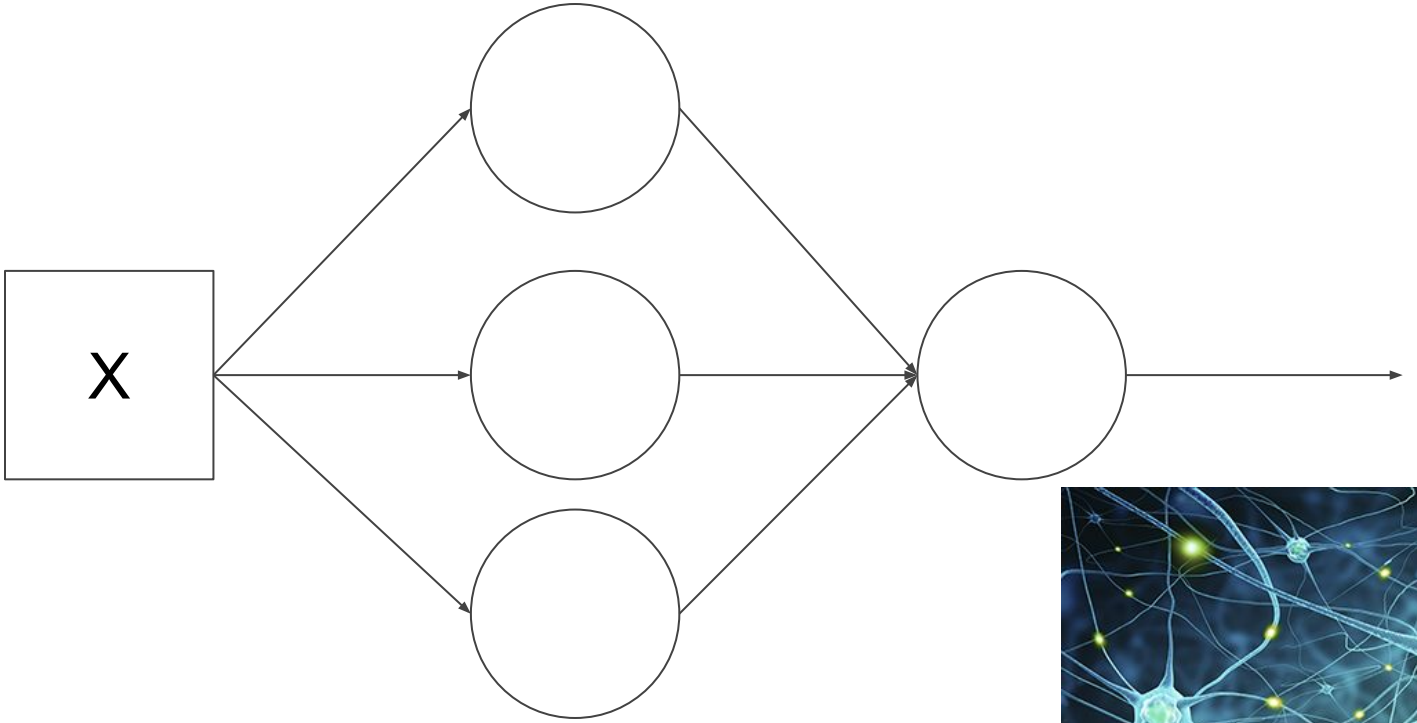
- Do the same thing, but a whole bunch of times!

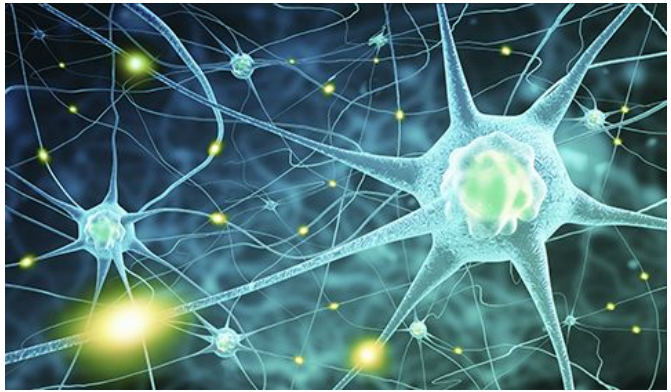


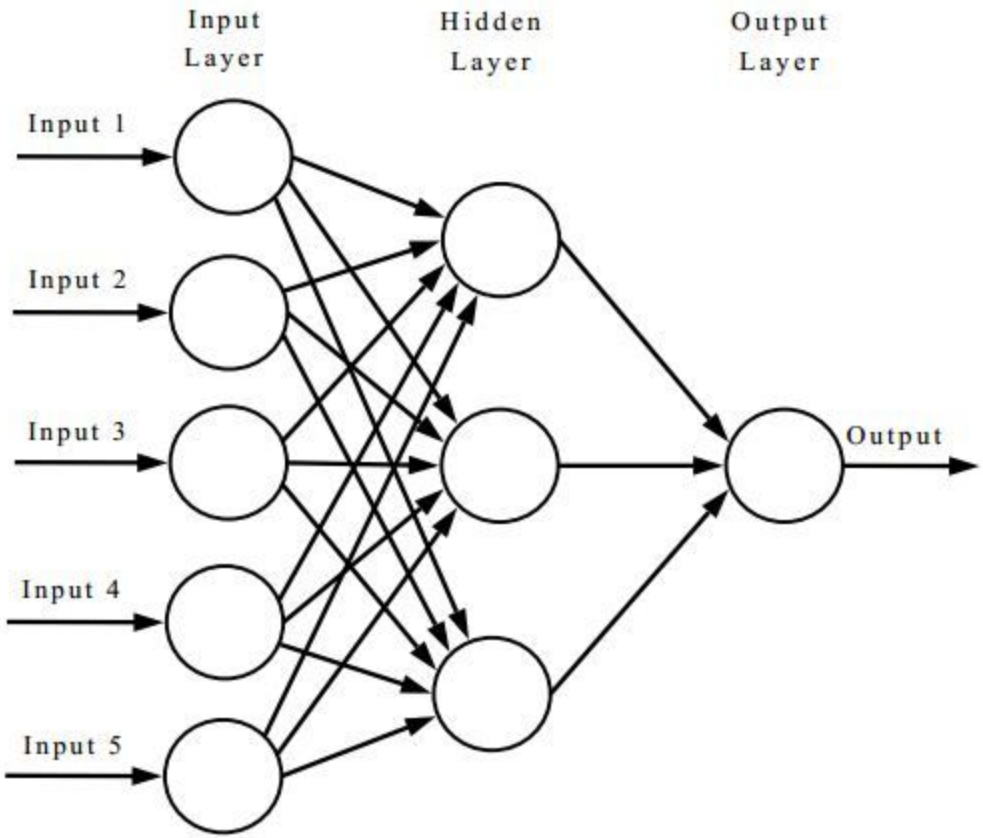

X
(High School GPA)
W*x+b
(Linear regression)
Ŷ
(Our prediction)

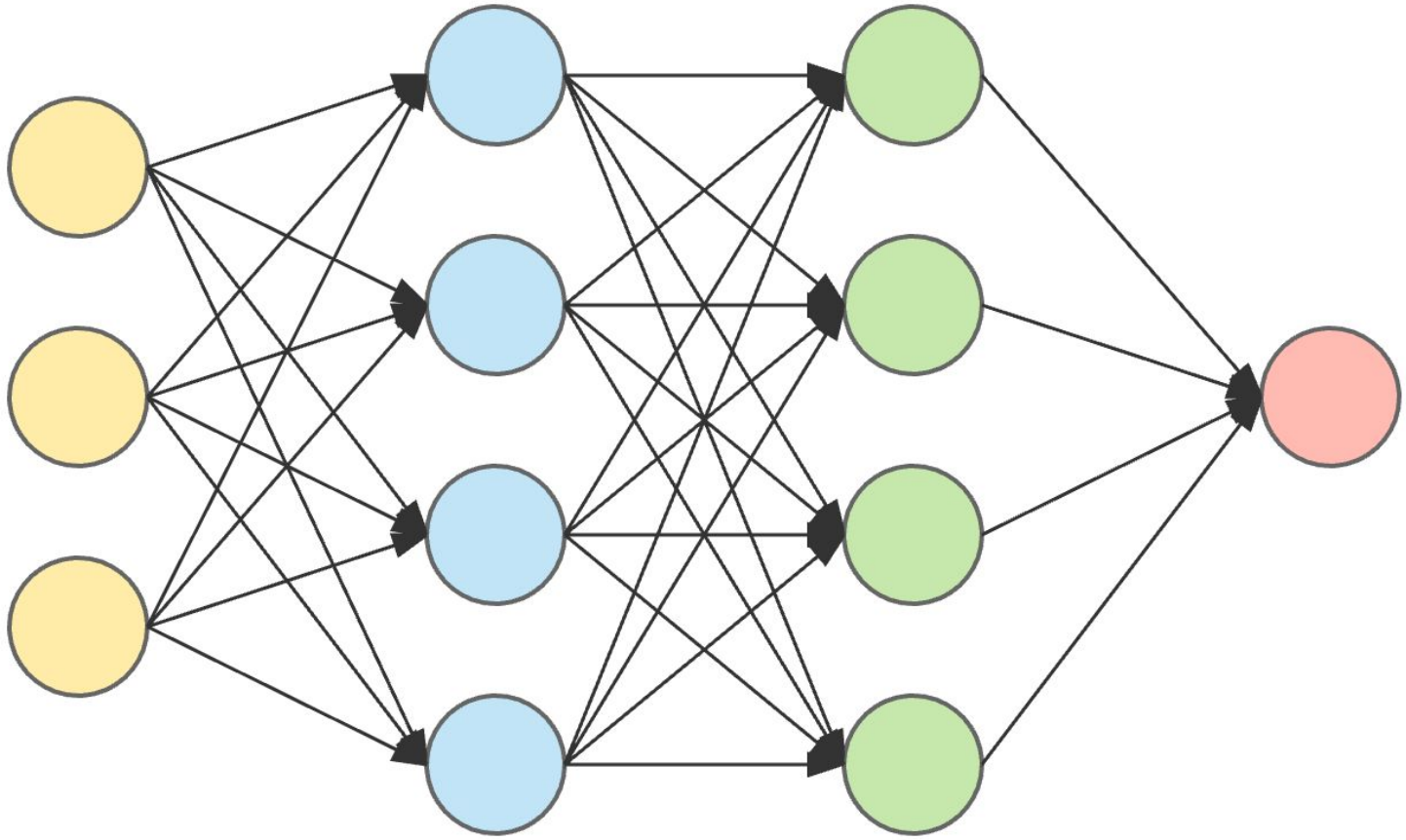

X
Ŷ

Neural Network






Input Layer
Hidden Layer
Output Layer
Input 1
Input 2
Input 3
Input 4
Input 5
Output


input layer
hidden layer 1
hidden layer 2
output layer

DEMO: Neural Networks

What about different data types, like images?

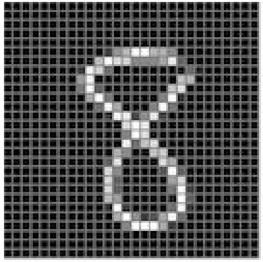







120	230	...	54
80	42	...	90
...	...	...	200
20	121	128	250

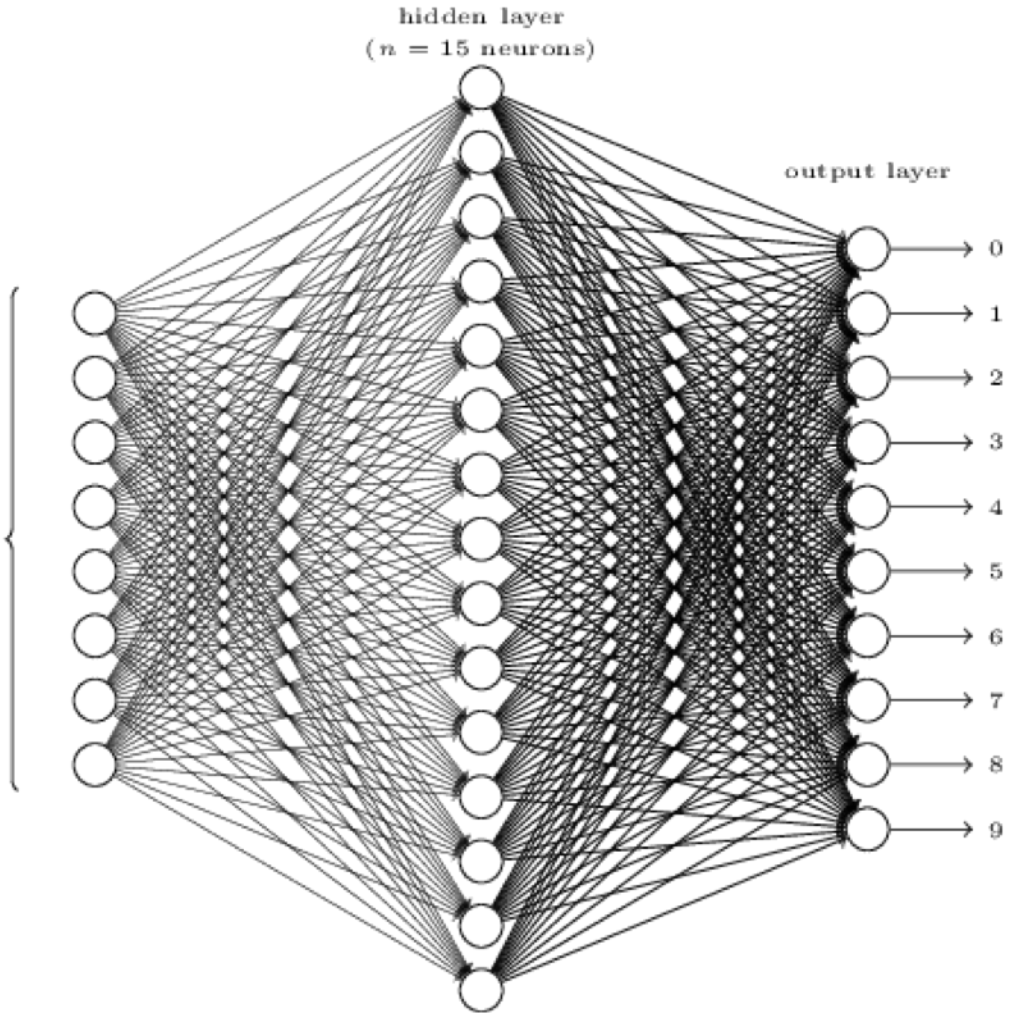
120
230
...
...
...
250



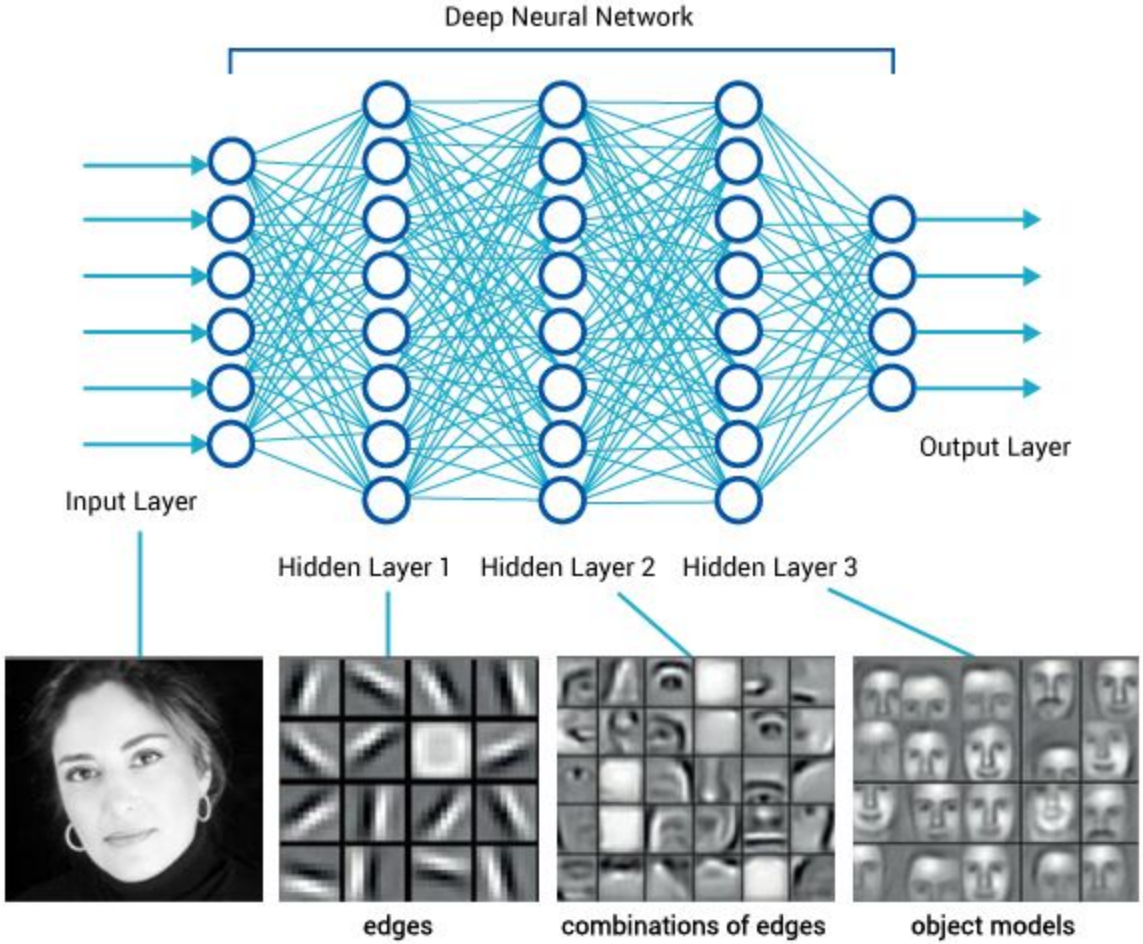
28 x 28
784 pixels

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	4	62	146	182	254	254	181	176	139	15	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	34	186	253	217	208	136	136	136	166	232	99	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	61	242	208	111	3	0	0	0	0	0	18	32	107	43	0	0	0	0	0	0
0	0	0	0	0	0	0	0	156	242	23	0	0	0	0	0	0	0	13	191	181	6	0	0	0	0	0	0
0	0	0	0	0	0	0	0	121	255	98	3	0	0	0	0	0	8	194	225	12	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	169	253	120	3	0	0	0	128	247	51	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	3	111	244	169	19	0	14	131	249	117	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	59	241	235	72	142	229	66	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	25	218	254	231	36	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	133	253	221	33	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	19	237	111	196	217	19	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	174	138	0	23	193	204	18	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	96	224	0	0	0	25	218	169	3	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	215	138	0	0	0	0	86	253	99	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	215	97	0	0	0	0	3	162	214	11	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	215	97	0	0	0	0	0	118	253	68	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	185	157	0	0	0	0	0	40	254	98	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	50	244	61	0	0	0	0	112	244	58	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	174	251	142	59	83	167	244	111	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	6	133	253	253	253	169	61	3	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

pixel 1
pixel 2
pixel 3
pixel 4
pixel 5
pixel 6
pixel 7
pixel 8
pixel 9
pixel 10
pixel 11
pixel 12
pixel 13
pixel 14
pixel 15
pixel 16
pixel 17
pixel 18
pixel 19
pixel 784

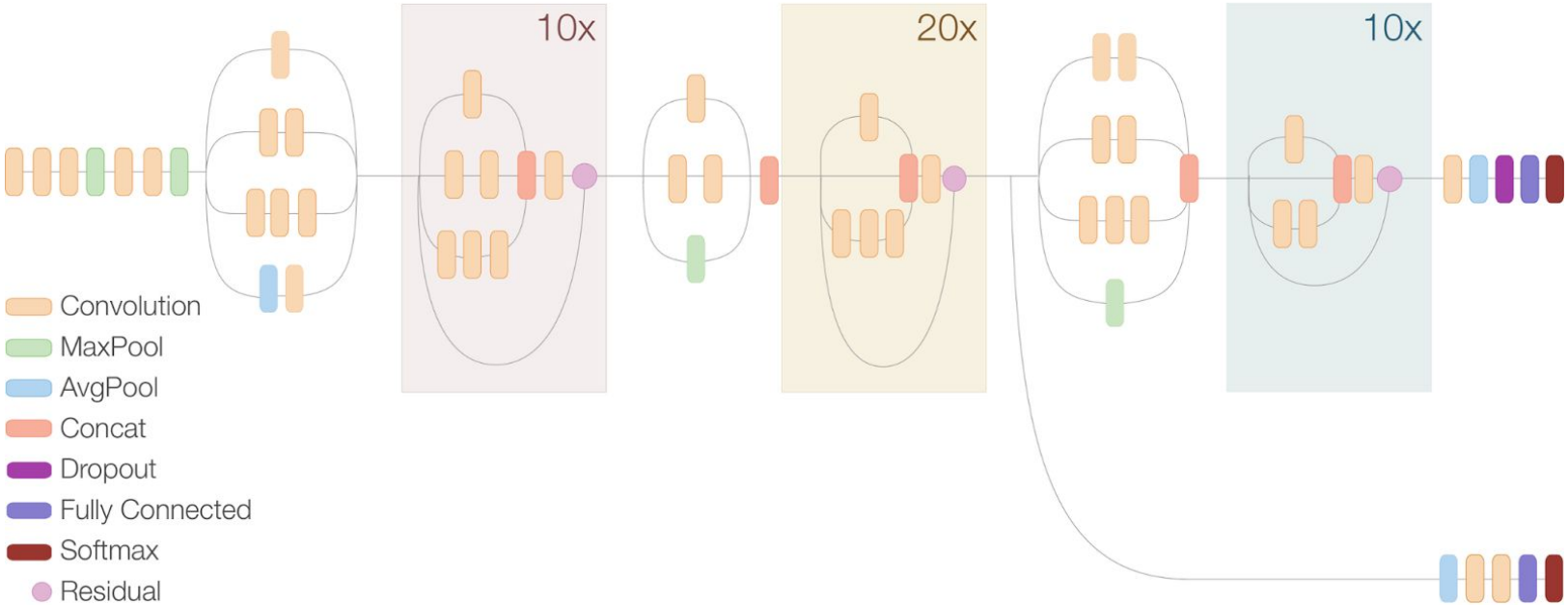

hidden layer
(n = 15 neurons)
output layer
input layer
(784 neurons)
0
1
2
3
4
5
6
7
8
9

DEMO:
Image Recognition


Deep Neural Network
Input Layer
Hidden Layer 1
Hidden Layer 2
Hidden Layer 3
Output Layer
edges
combinations of edges
object models

Inception Resnet V2 Network

Compressed View



Machine Learning is powerful, but it’s just data+math (and fast computers)!

Scott Kirkland

AppDev: 12/12/2017