## how to train your model

#### Jenna Zeigen (she/her) QueensJS 8/5/2020

## senior frontend engineer at Slack

## organizer of BrooklynJS

organizer of
 EmpireJS

### @zeigenvector

#### jenna.is/at-queensjs-2020

## machine learning

creating algorithms
 that improve
automatically through
 experience over
 time

## machine learning

building a mathematical model based on "training data" in order to make predictions



## supervised learning

when the answers are known ahead of time, and the computer tries to find a model to fit the data

i.e. classification

#### unsupervised learning

when the answers aren't known ahead of time, and computer finds patterns

#### i.e. clustering



#### reinforcement learning

when the answers aren't known ahead of time and the algorithm learns by trial and error through "incentives"

popular in teaching computers to play games



	1000	0 1 1-1 11 0 0 1	11011			000 011 000 001	<b>Bud</b>	100 001 101		000	Stand-In-			<b>Budb</b>	1200	0 1 0 1 0 1 0 1 0 0	19.95			0/1 0/1 000 000	1000		0101 0101 0011		111 <b>0</b>		110 011 010		$0^{0}_{0}0^{0}_{1}1$ 110 111	<b>Care</b>			3 <b>3</b> 9 6		
$ \begin{array}{c} 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ \end{array} $	Ĩ	0	0				¢,	1	5			201 1	1	7		24.5	21.00		0_		64	7. 4	1.41		0100	<u>,</u>						0 0 1 1 1 1 0 1 0 1 0		1000	000
										000			0 (0 0 10 1 1 0 0	1			100	0 1	)					010	11011		010	to-tot					010	191	0000
							0			(1 - 1) (1 - 1) (1 - 1) (1 - 1)			.0 40				C TTC				1			1			ţ		01)			101 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	000		
$1^{1}$ $1^{1}$ $1^{1}$ $1^{1}$			422200	) 1 0 0		0						ŋ		ł									1 10 10 11 00 00 1 0 10		6101			0	47		)	$1^{0}1$ $1^{1}1$ $0_{1}0$ $0_{0}1$ $0_{0}1$			
			10	$0^{1}_{1}$	ğ	001 010 010	Por Pd							1			10 10							000	1000			104				00100 0100 111000	100		0111
0 $0$ $1$			0000	1 1 0		010 01 000 000	D THO O							9999			1100			0 0 1 1 0 1 0 1	ļ		0,1951	011	Sapp		2017	012				$     \begin{array}{c}       0 & 1 & 0 \\       1 & 1 & 1 \\       1 & 1 & 1 \\       0 & 0 & 1   \end{array} $	00011		
$\begin{array}{c} 1 & 01 \\ 0 & 101 \\ 0 & 0 \\ 0 & 0 \\ 1 & 1 \\ 1 & 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1$	,00	í ( 1 .	19 1- 0-19	0 0 1 0		000 011 111 10	Series Series				0				1	0	ju O			$     \begin{array}{c}             0 \\             0 \\         $	í g±		81000	011	0101	ĥ.	0 1 0	101	$     \begin{array}{c}       0 \\       1 \\       1 \\       1 \\       0 \\       1 \\       0 \\       1     \end{array} $			1010 1010 1010	0 0		
	Ō		11 00	001	0 0 0	$1^{1}1$ 000 000	1910	0		$     \begin{array}{c}       0 \\       1 \\       1 \\       1 \\       0 \\     $	001	110			0	0 10 10	יק פויט			1 1 0 0 01 0 0	100			101				nh cist				0 <sup>1</sup> 1 101 011 0	000		
		1 1 1 0	1130	011		01 11 01	P				0 1 0	11 10 (01		6044	South Cal	00 01 10	110	1 1 0 1	5	$0^{10}_{11}$ 100	5.42			1	<i>Ş</i> ą	2 			11	1	1			0 1 1 0 1	
			000	0 1 1 0	010	10 1 0 1 1 0	Port of	1 (	011	10 01 01	1 0 1	000		1919	1010	1 1 0	6-4 d d		1	101 001 000	0100			(	C	5				-					1101
0 010 4 0 010 0 0 010 1	)	1000	1 1 1	0000	111	00 01 01	10		801	${}^{1\perp}_{00}_{0}$	0	чо т іі	י 1 s's	'1 the	1' > r	1 na	1 triv	<u>a</u>	1	10	0-10 10 1	0				-	s'							1911	
	)	00101	10	010	100	00 1 1	1	(		b	ec	aus		he	ere	e ai	re a	a lo	ot	of		11	1 0 0	0		5		0 1	P. C			K	X		
1 0 0 <u>1</u> 1 0 10		010	100	1		1	01	(		1	0		ιαι				U	1	0		0	00 11	)	1 11	0	10	1	1	111		ļ	8.	~		
		0	1				1			U	0	1		â			1	0		1 0	1			0	0	1	0	1	0	4	0	1	1	1 1	1

## classifiers

#### put objects into groups based on their characteristics

haha no math yet

## linear classifiers

#### Do this based on a boundary that is a "line"

(or through ~\*linear combination\*~)

ok now math

## linear classifier



#### Who called it Tinder and not Naive Baes Classifier?

11:56 AM - Sep 28, 2016 - Twitter Web Client

395 Retweets and comments 910 Likes





 $\sim$ 

autocomplete ranking is a matter of classification — is it the thing you're looking for or not?

tl;dr, turning logs
into decimals using
supervised learning

## feature extraction

"features" are the attributes of the item that could be influencing the classification

### feature extraction

creates a "feature vector" for each item, a list of all the features and their values p.s. even images can be represented as vectors

## training

feature vectors from
the selected and not
selected items are
 used as data
 to train
 the "model"

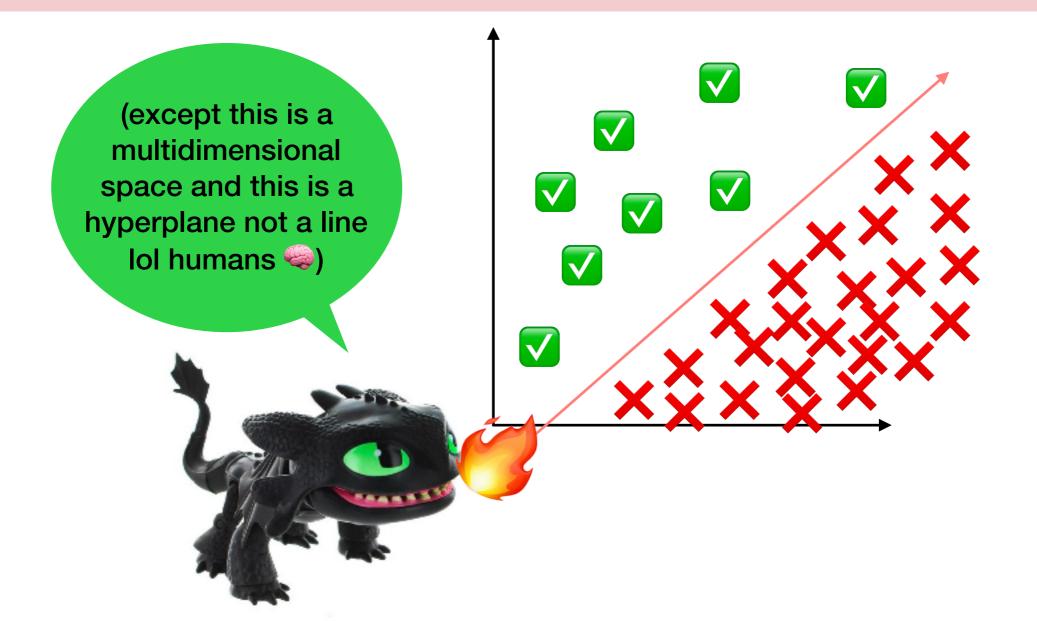
## training

the model will be a vector which has weights for each of the features

## training



## training



## training

2.286

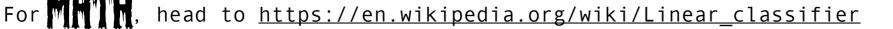
.156

vonktor

minimize costs \$\$\$!!!

## scoring

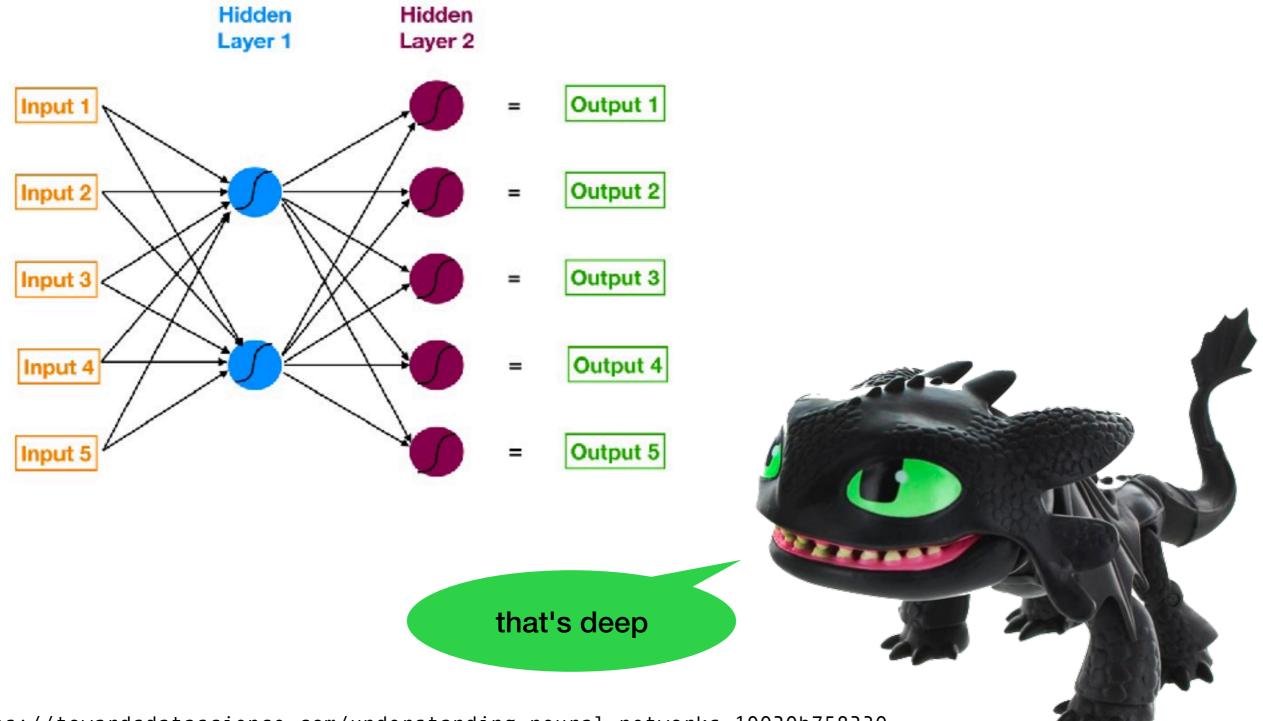
An item's score is the sum of the product of each feature's value and its weight



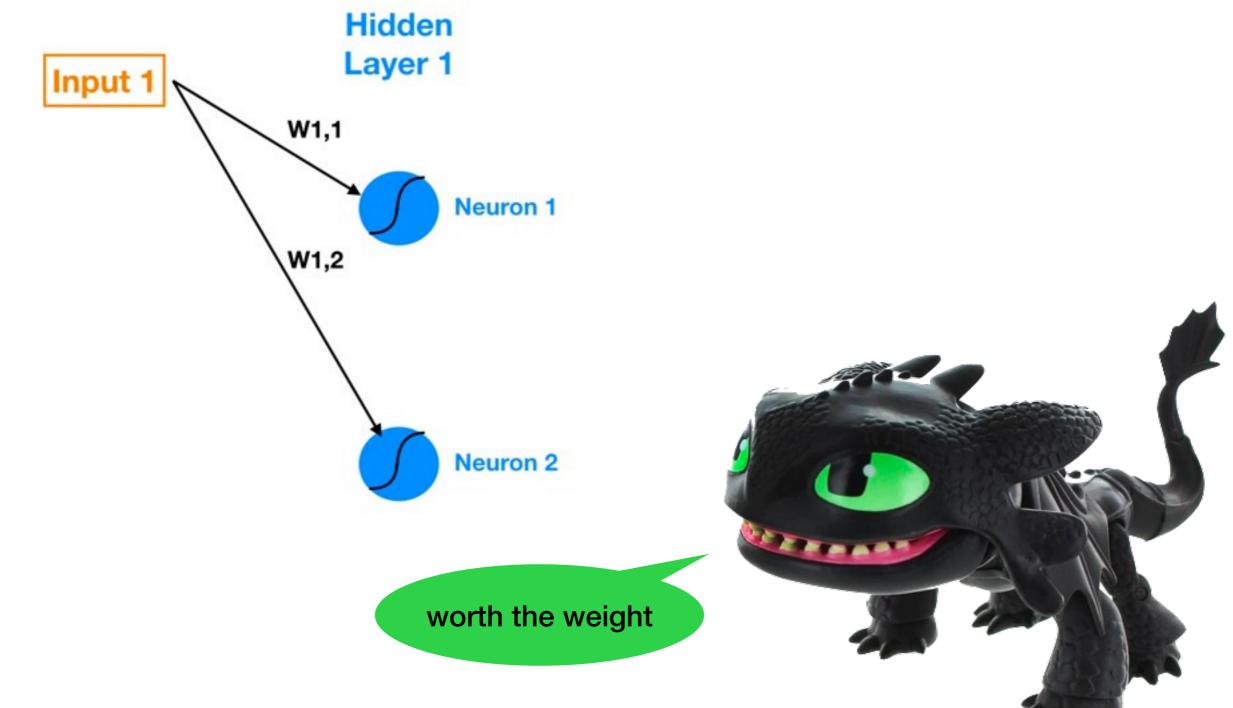
#### "neural networks"

hotness

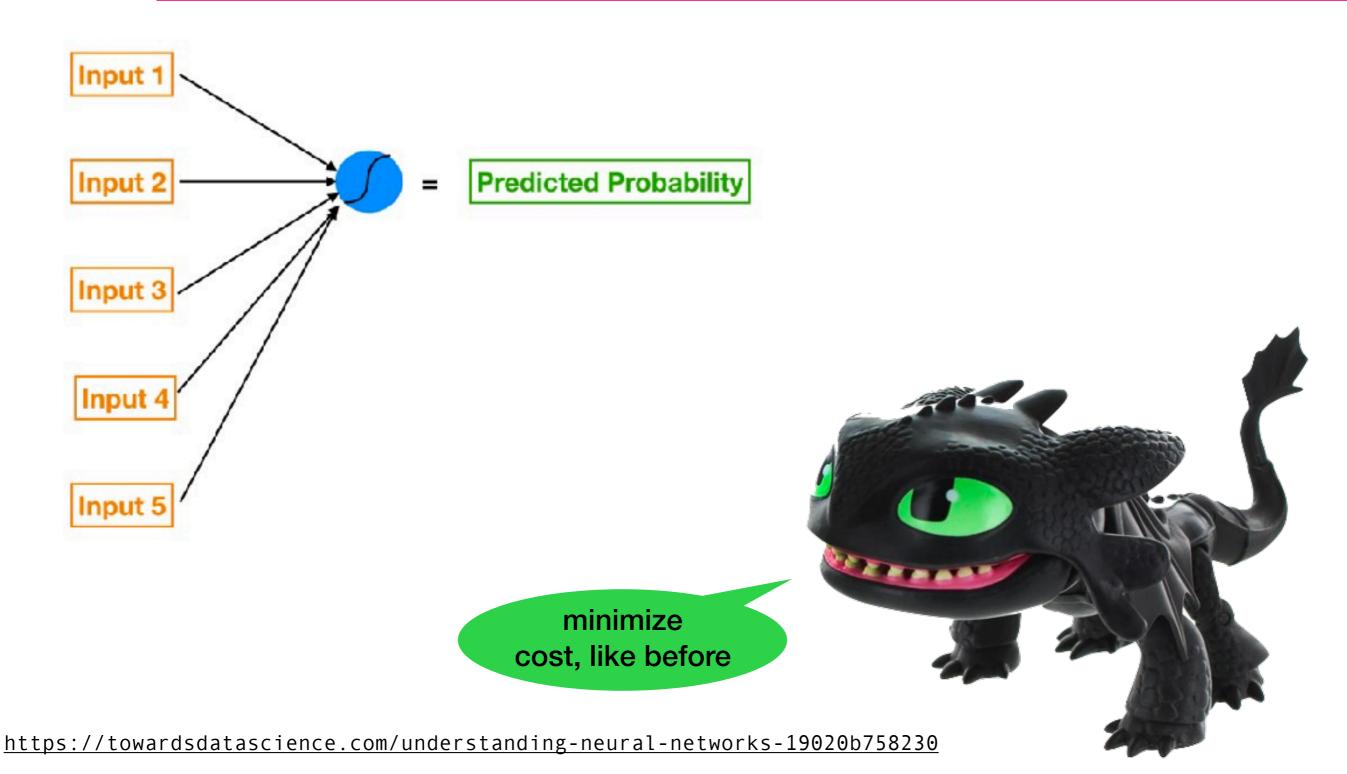
https://towardsdatascience.com/understanding-neural-networks-19020b758230



https://towardsdatascience.com/understanding-neural-networks-19020b758230



https://towardsdatascience.com/understanding-neural-networks-19020b758230



#### how to train your model

## ethically.

spiciness?

#### how to train your model

## algorithmic bias is real.

spiciness?

# @zeigenvector jenna.is/at-queensjs-2020

