

California State University, Dominguez Hills

Department of Computer Science

CSC 595

Professor: Dr. Benyamin Ahmadnia
bahmadniayebosari@csudh.edu

Spring 2025

Copyright Notice

These slides are distributed under the Creative Commons License.

[DeepLearning.AI](#) makes these slides available for educational purposes. You may not use or distribute these slides for commercial purposes. You may make copies of these slides and use or distribute them for educational purposes as long as you cite [DeepLearning.AI](#) as the source of the slides.

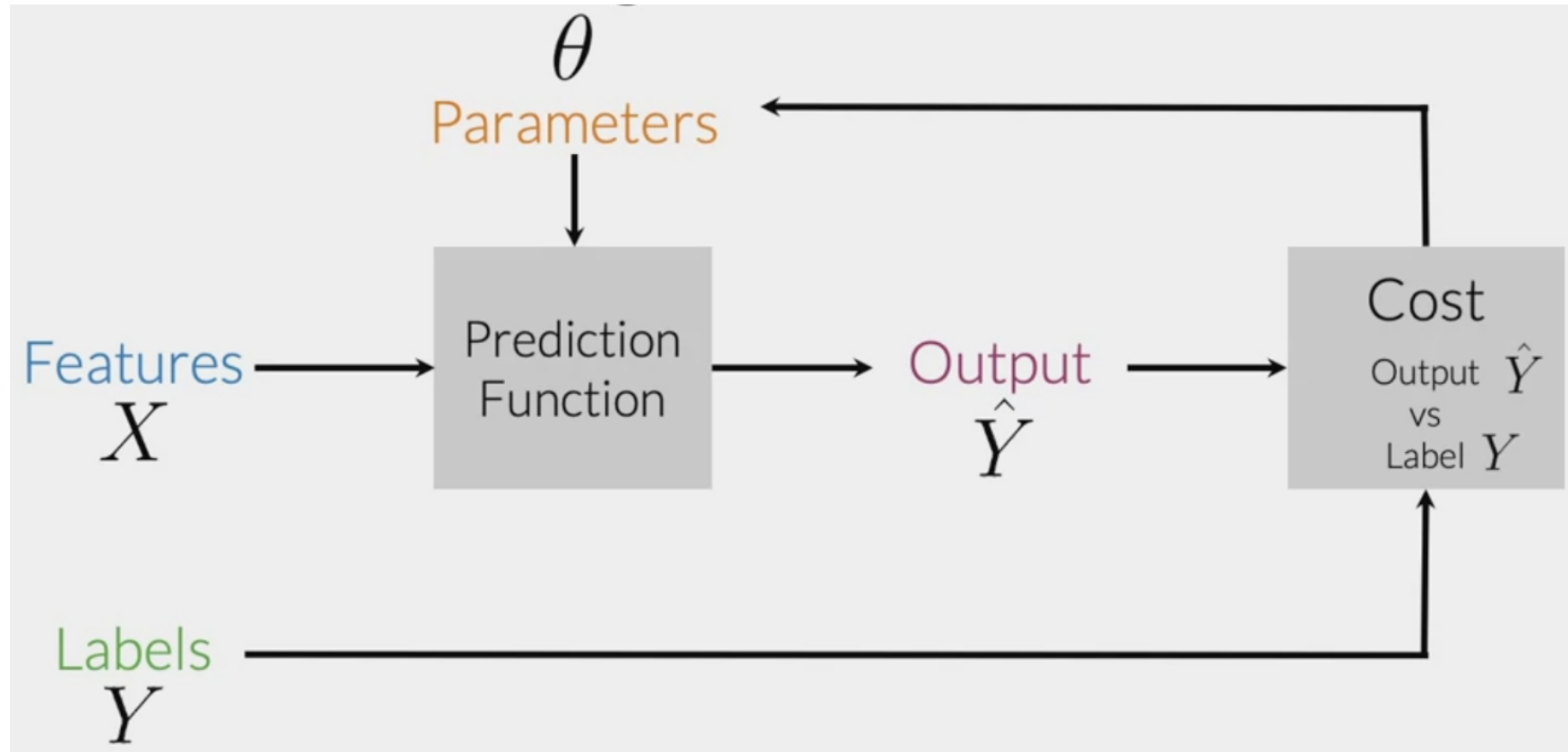
For the rest of the details of the license, see <https://creativecommons.org/licenses/by-sa/2.0/legalcode>

Table of Contents

- Sentiment Analysis with Logistic Regression
- Sentiment Analysis with Naïve Bayes
- Vector Space Models
- Machine Translation and Document Search

Sentiment Analysis with Logistic Regression

Supervised ML (training)



Sentiment Analysis

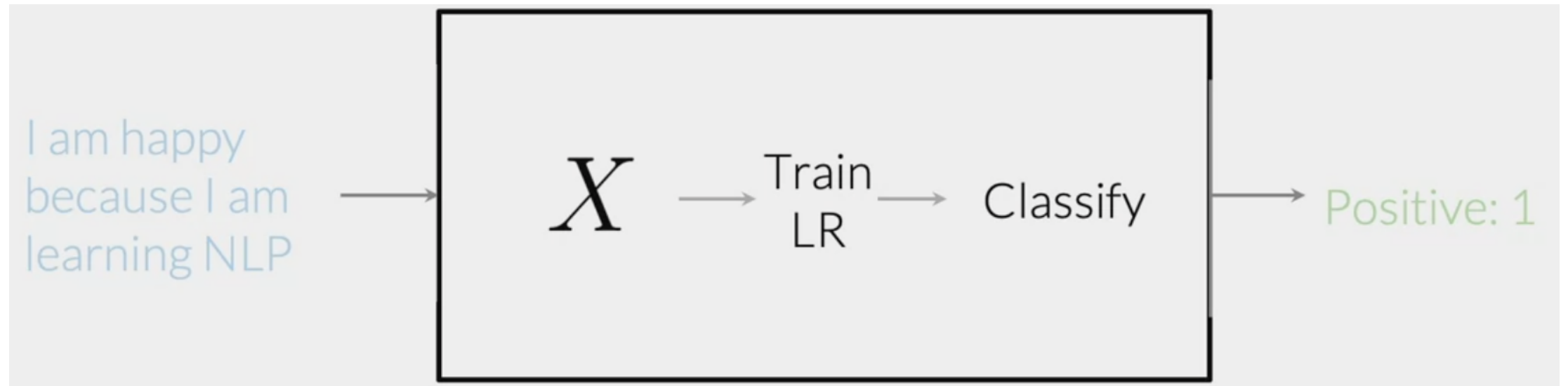
Tweet: I am happy because I am learning NLP

Positive: 1

Negative: 0

Logistic regression

Sentiment Analysis



To Classify a Tweet ...

- Positive or negative
- Extract the features
- Train the model
- Classify the tweet based on the trained model

Vocabulary

Tweets:

[tweet_1, tweet_2, ..., tweet_m]



I am happy because I am learning NLP

...

I hated the movie

$V =$

[I, am, happy, because, learning, NLP, ... hated, the, movie]

Feature Extraction

I am happy because I am learning NLP

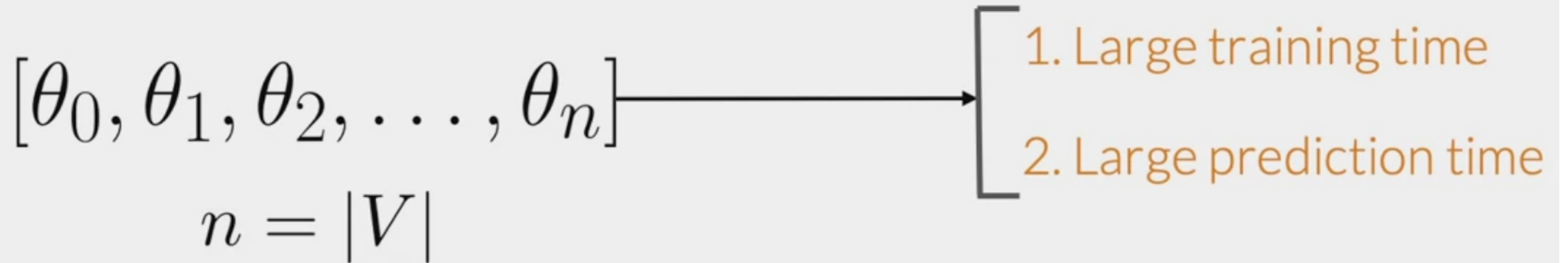
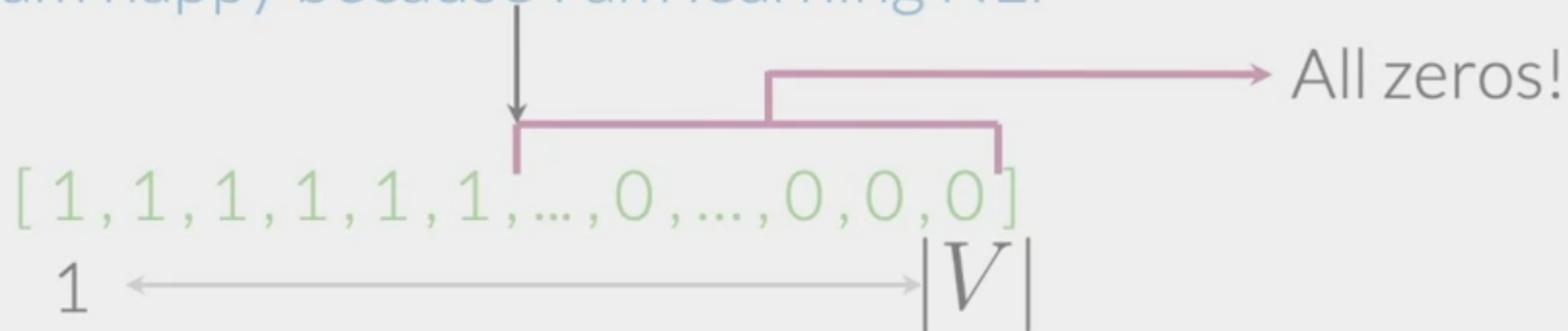
[I, am, happy, because, learning, NLP, ... hated, the, movie]

↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓
[1, 1, 1, 1, 1, 1, ... 0, 0, 0]

A lot of zeros! That's a sparse representation.

Problems with Sparse Representations

I am happy because I am learning NLP



Positive and Negative Counts

Corpus

I am happy because I am learning NLP

I am happy

I am sad, I am not learning NLP

I am sad

Vocabulary

I

am

happy

because

learning

NLP

sad

not

Positive and Negative Counts

Positive tweets

I am happy because I am learning NLP

I am happy

Negative tweets

I am sad, I am not learning NLP

I am sad

Positive Counts

Positive tweets

I am happy because I am learning NLP

I am happy

Vocabulary	PosFreq (1)
------------	-------------

I	
---	--

am	
----	--

happy	2
--------------	---

because	
---------	--

learning	
----------	--

NLP	
-----	--

sad	
-----	--

not	
-----	--

Negative Counts

Vocabulary	NegFreq (0)	Negative tweets
I		
am	3	<u>I</u> am sad, I <u>am</u> not learning NLP
happy		
because		
learning		
NLP		
sad		
not		
		I <u>am</u> sad

Word Frequency in Classes

Vocabulary	PosFreq (1)	NegFreq (0)
I	3	3
am	3	3
happy	2	0
because	1	0
learning	1	1
NLP	1	1
sad	0	2
not	0	1

Word Frequency in Classes

Vocabulary	PosFreq (1)	NegFreq (0)	<i>freqs</i> : dictionary mapping from (word, class) to frequency
I	3	3	
am	3	3	
happy	2	0	
because	1	0	
learning	1	1	
NLP	1	1	
sad	0	2	
not	0	1	

Feature Extraction

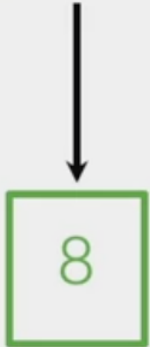
freqs: dictionary mapping from (word, class) to frequency

$$X_m = [1, \sum_w \textit{freqs}(w, 1), \sum_w \textit{freqs}(w, 0)]$$

↓ ↓ ↓ ↓

Features of Bias Sum Pos. Sum Neg.
tweet m Frequencies Frequencies

Feature Extraction

Vocabulary	PosFreq (1)	I am sad, I am not learning NLP
I	<u>3</u>	$X_m = [1, \sum_w \text{freqs}(w, 1), \sum_w \text{freqs}(w, 0)]$ <div></div>
am	<u>3</u>	
happy	2	
because	1	
learning	<u>1</u>	
NLP	<u>1</u>	
sad	<u>0</u>	
not	<u>0</u>	

Feature Extraction

Vocabulary	NegFreq (0)
I	<u>3</u>
am	<u>3</u>
happy	0
because	0
learning	<u>1</u>
NLP	<u>1</u>
sad	<u>2</u>
not	<u>1</u>

I am sad, I am not learning NLP

$$X_m = [1, \sum_w \text{freqs}(w, 1), \sum_w \text{freqs}(w, 0)]$$

↓
11

Feature Extraction

I am sad, I am not learning NLP

$$X_m = [1, \sum_w \textit{freqs}(w, \textcolor{green}{1}), \sum_w \textit{freqs}(w, \textcolor{violet}{0})]$$



$$X_m = [1, \textcolor{green}{8}, \textcolor{violet}{11}]$$

Preprocessing

@YMourri and @AndrewYNg are
tuning a GREAT AI model at
<https://deeplearning.ai!!!>

Stop words

and
is
are
at
has
for
a

Punctuation

,
.
:
!
"
'

Preprocessing

@YMourri and @AndrewYNg are
tuning a GREAT AI model at
<https://deeplearning.ai!!!>

@YMourri @AndrewYNg tuning
GREAT AI model
<https://deeplearning.ai!!!>

Stop words

and

is

are

at

has

for

a

Punctuation

,

.

:

!

“

‘

Preprocessing

@YMourri @AndrewYNg tuning
GREAT AI model
<https://deeplearning.ai!!!>

@YMourri @AndrewYNg tuning
GREAT AI model
<https://deeplearning.ai>

Stop words

and

is

a

at

has

for

of

Punctuation

,

.

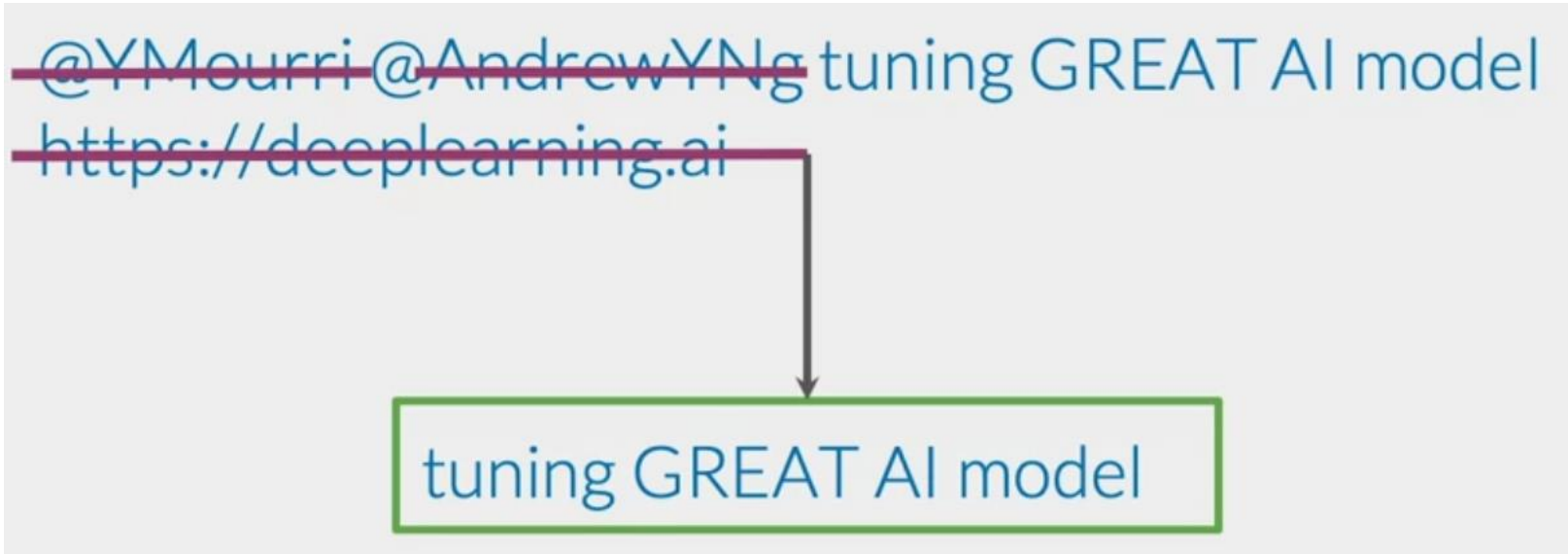
:

!

"

'

Preprocessing



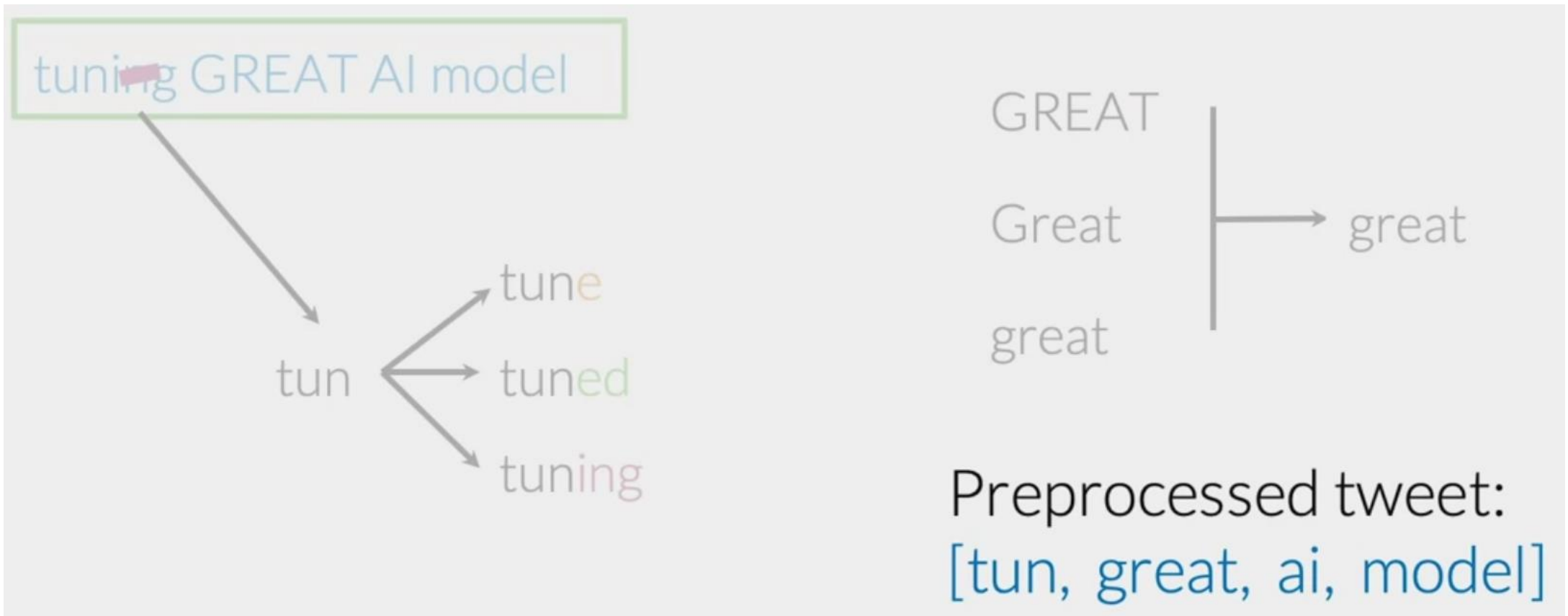
Preprocessing

tuning GREAT AI model

tun →
tune
tuned
tuning

GREAT
Great
great → great

Preprocessing



Putting it all Together

I am Happy Because i am learning NLP @deeplearning

↓ Preprocessing

[happy, learn, nlp]

↓ Feature Extraction

Bias ← [1, 4, 2] → Sum negative frequencies

Sum positive frequencies

Putting it all Together

I am Happy Because i am
learning NLP
@deeplearning

I am sad not learning NLP

...

I am sad :(

[happy, learn, nlp]

[sad, not, learn, nlp]

...

[sad]

[[1, 40, 20],

[1, 20, 50],

...

[1, 5, 35]]

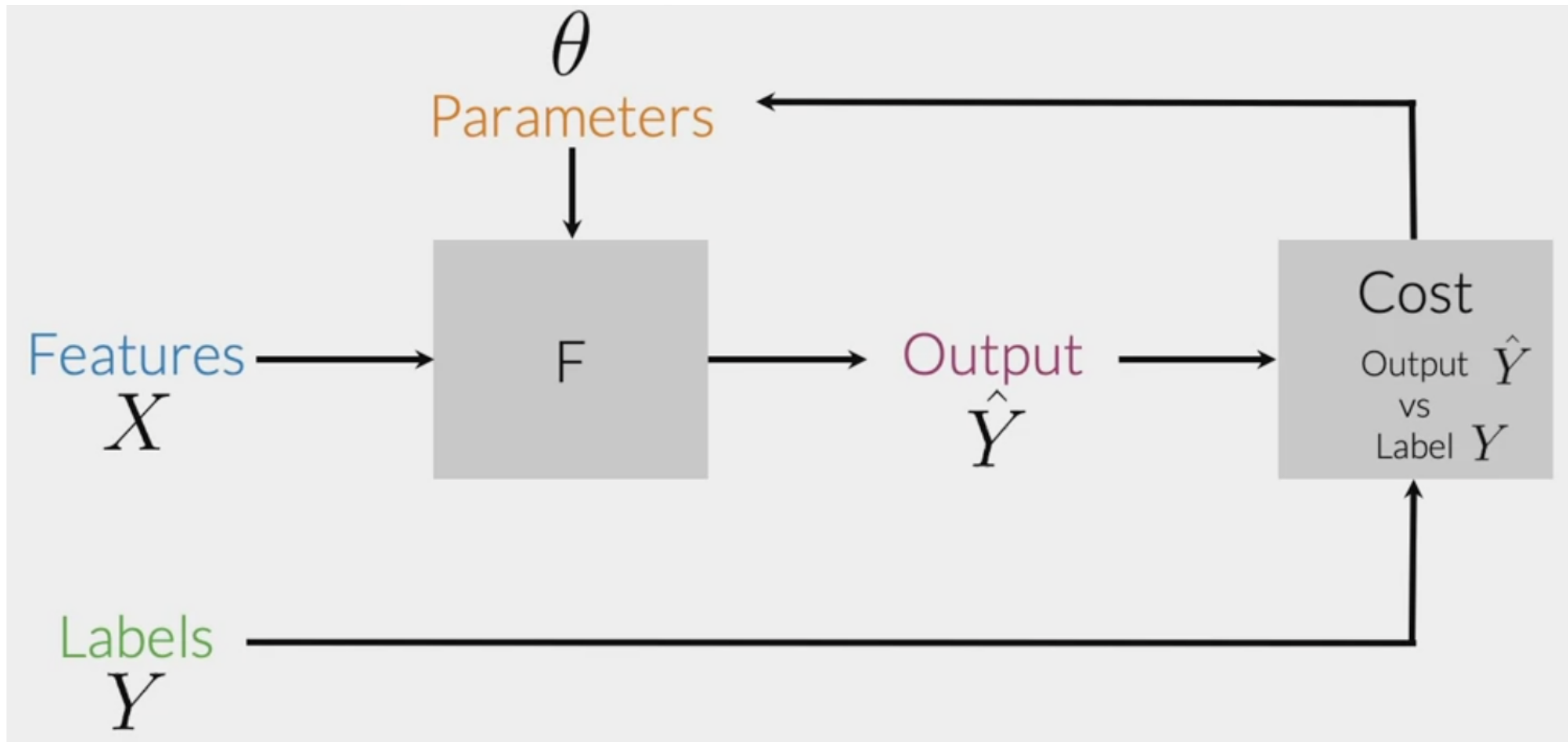
Putting it all Together

$$\begin{bmatrix} 1 & X_1^{(1)} & X_2^{(1)} \\ 1 & X_1^{(2)} & X_2^{(2)} \\ \vdots & \vdots & \vdots \\ 1 & X_1^{(m)} & X_2^{(m)} \end{bmatrix} \longleftrightarrow \begin{matrix} [1, 40, 20] \\ [1, 20, 50], \\ \dots \\ [1, 5, 35] \end{matrix}$$

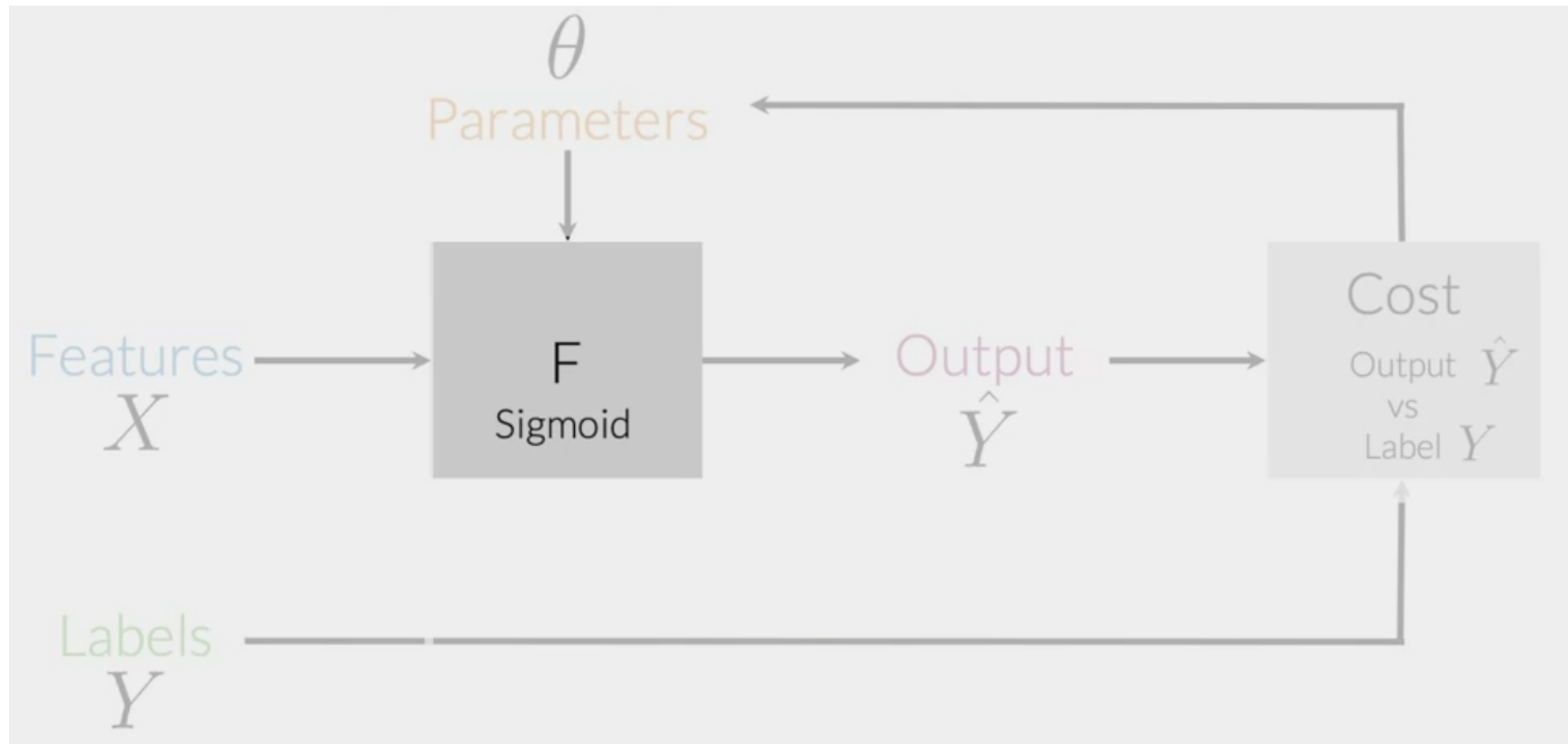
Putting it all Together

```
freqs = build_freqs(tweets, labels) #Build frequencies dictionary
X = np.zeros((m, 3)) #Initialize matrix X
for i in range(m): #For every tweet
    p_tweet = process_tweet(tweets[i]) #Process tweet
    X[i, :] = extract_features(p_tweet, freqs) #Extract Features
```

Overview of Logistic Regression

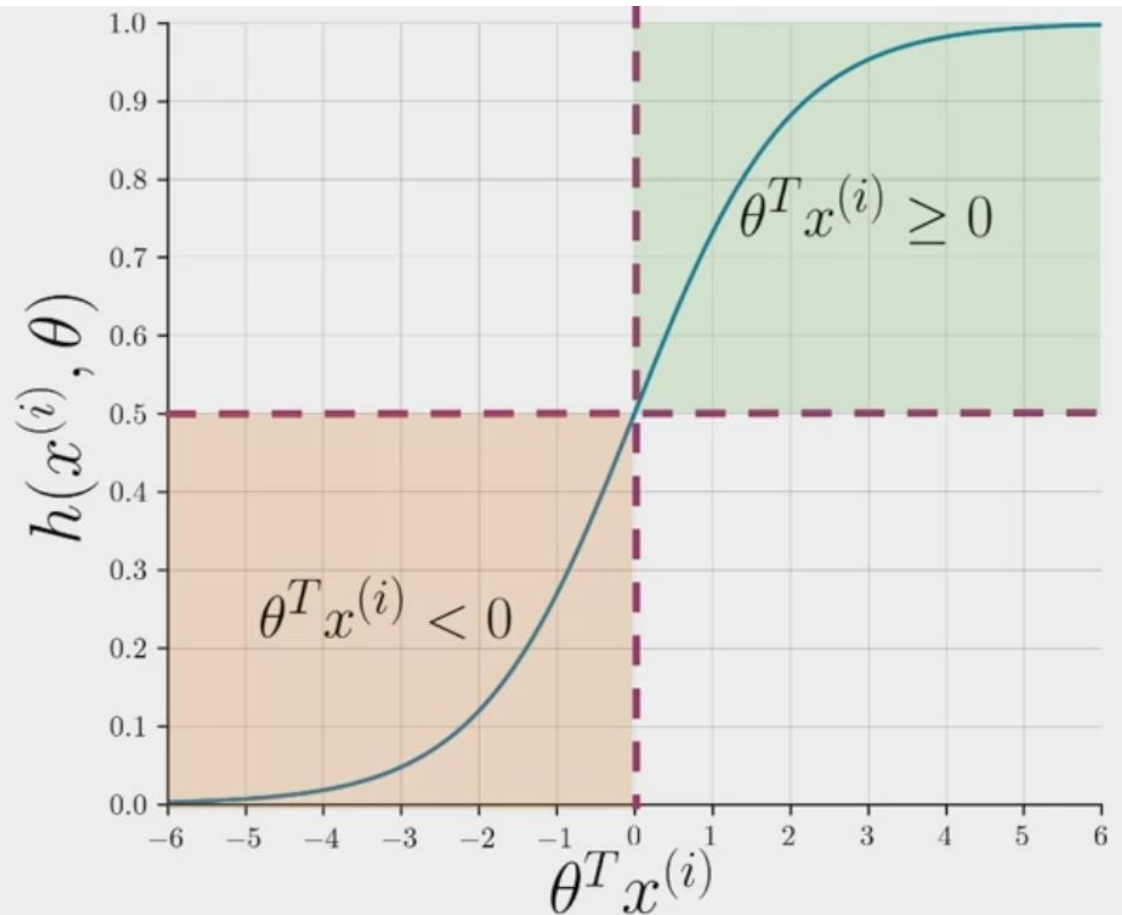


Overview of Logistic Regression



Overview of Logistic Regression

$$h(x^{(i)}, \theta) = \frac{1}{1 + e^{-\theta^T x^{(i)}}}$$

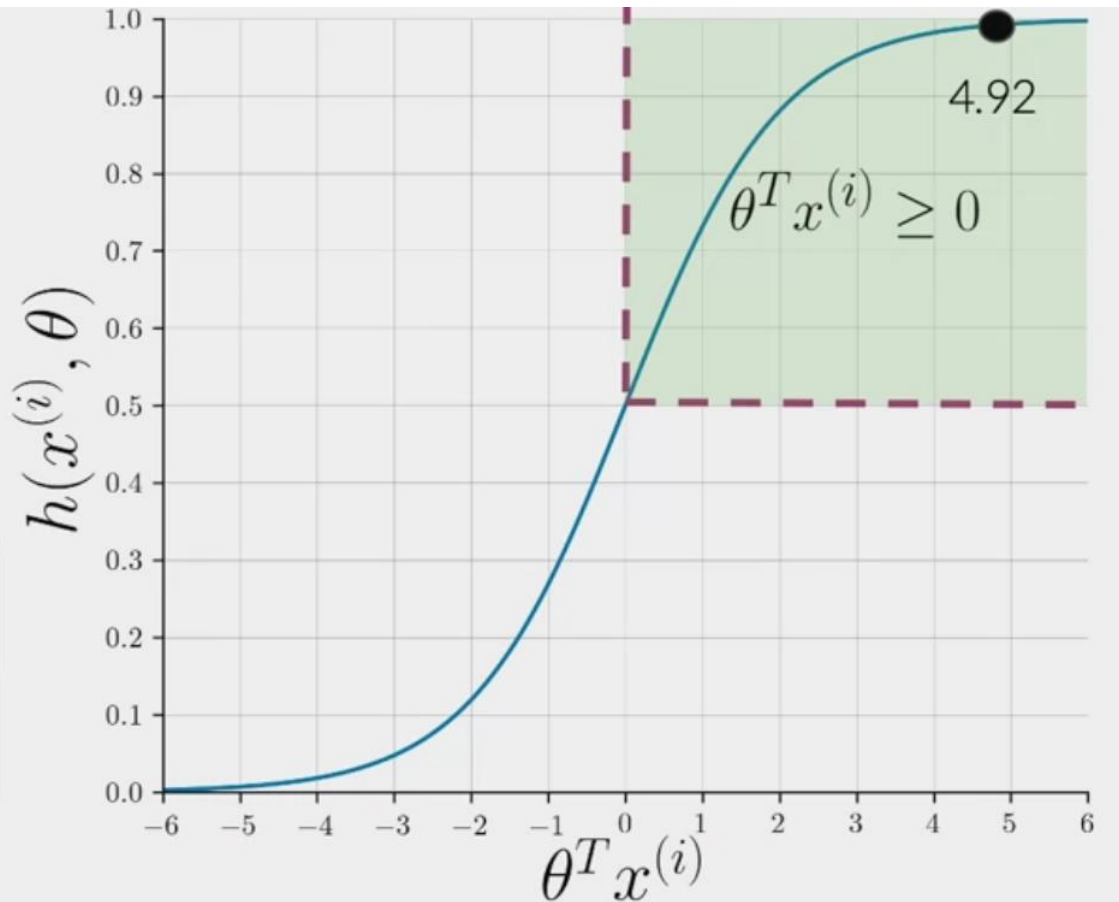


Overview of Logistic Regression

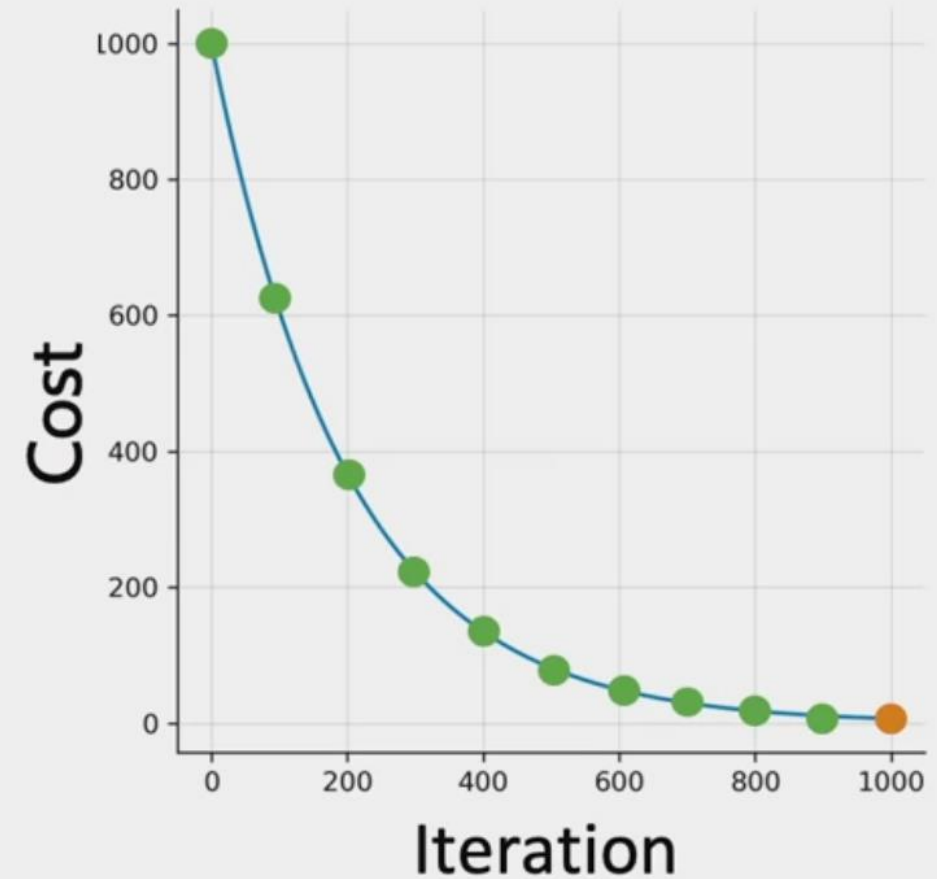
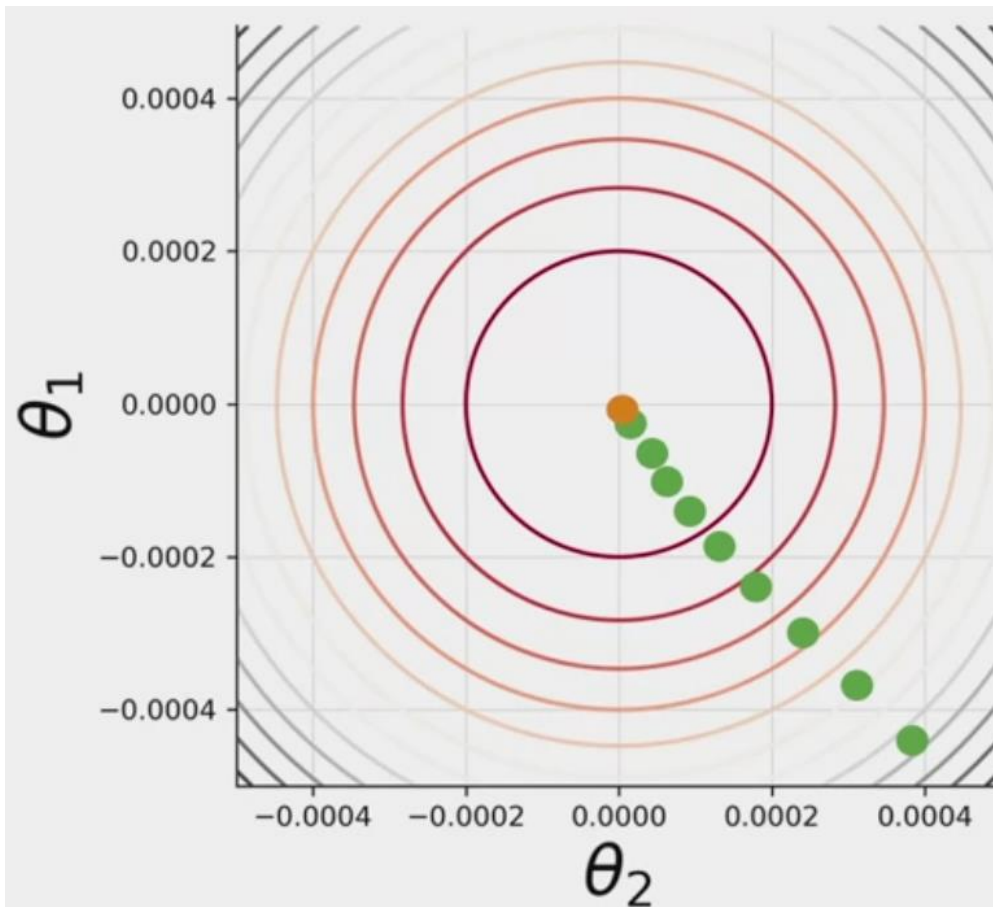
@YMurri and
@AndrewYNg are tuning a
GREAT AI model

[tun, ai, great, model]

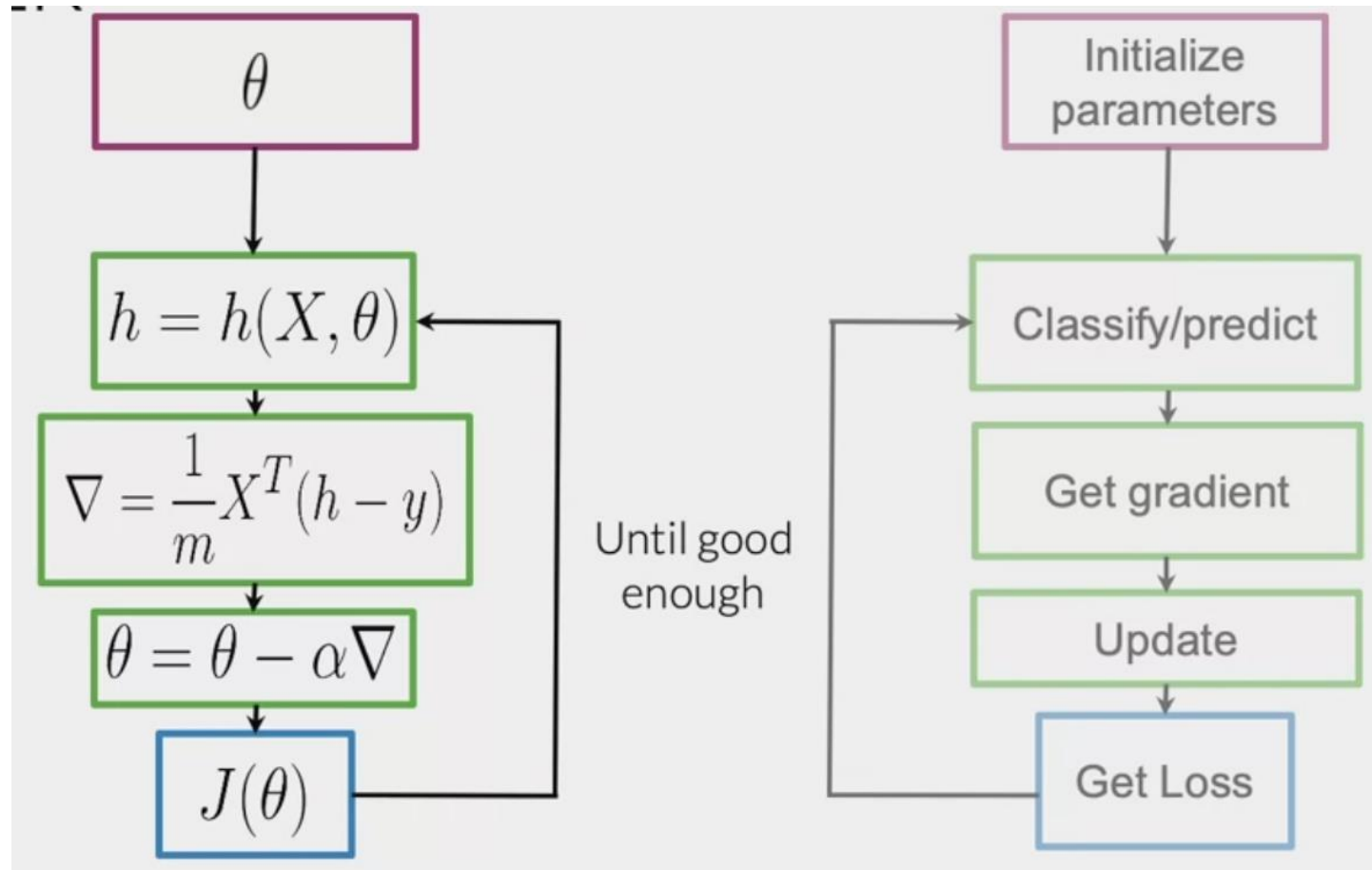
$$x^{(i)} = \begin{bmatrix} 1 \\ 3476 \\ 245 \end{bmatrix} \quad \theta = \begin{bmatrix} 0.00003 \\ 0.00150 \\ -0.00120 \end{bmatrix}$$



Training Logistic Regression



Training Logistic Regression

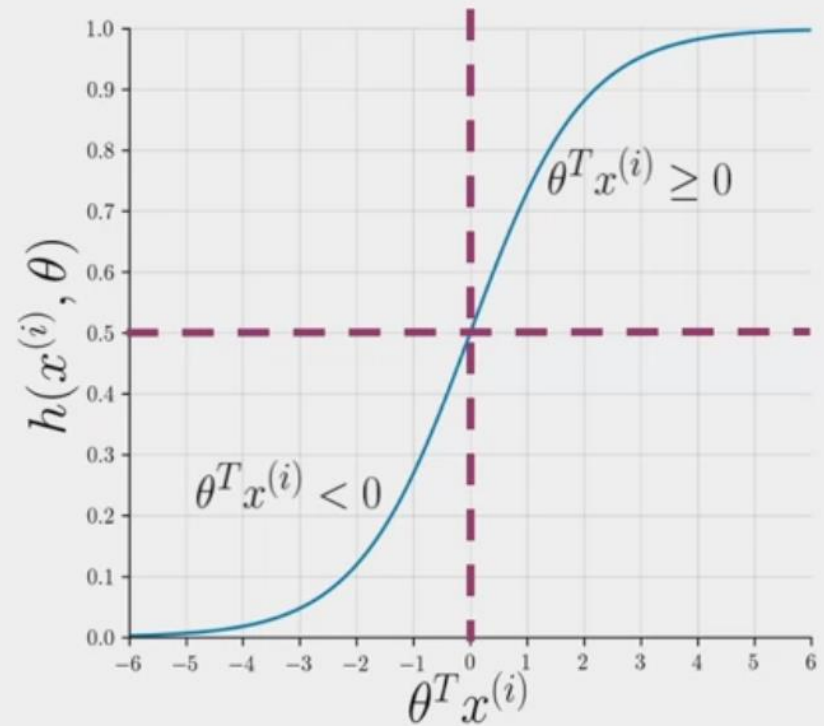


Testing Logistic Regression

- X_{val} Y_{val} θ

$h(X_{val}, \theta)$

$pred = h(X_{val}, \theta) \geq 0.5$



Testing Logistic Regression

- X_{val} Y_{val} θ

$$h(X_{val}, \theta)$$

$$pred = h(X_{val}, \theta) \geq 0.5$$

$$\begin{bmatrix} 0.3 \\ 0.8 \\ 0.5 \\ \vdots \\ h_m \end{bmatrix} \geq 0.5 = \begin{bmatrix} 0.3 \geq 0.5 \\ 0.8 \geq 0.5 \\ 0.5 \geq 0.5 \\ \vdots \\ pred_m \geq 0.5 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ \vdots \\ pred_m \end{bmatrix}$$

Testing Logistic Regression

- X_{val} Y_{val} θ

$h(X_{val}, \theta)$

$pred = h(X_{val}, \theta) \geq 0.5$

$$\sum_{i=1}^m \frac{(pred^{(i)} == y_{val}^{(i)})}{m}$$

Testing Logistic Regression

- X_{val} Y_{val} θ

$h(X_{val}, \theta)$

$pred = h(X_{val}, \theta) \geq 0.5$

$$\sum_{i=1}^m \frac{(pred^{(i)} == y_{val}^{(i)})}{m}$$

$$\begin{bmatrix} \underline{0} \\ 1 \\ 1 \\ \vdots \\ pred_m \end{bmatrix} == \begin{bmatrix} \underline{0} \\ 0 \\ 1 \\ \vdots \\ Y_{val_m} \end{bmatrix}$$

$$\begin{bmatrix} \underline{1} \\ 0 \\ 1 \\ \vdots \\ pred_m == Y_{val_m} \end{bmatrix}$$

Testing Logistic Regression

$$Y_{val} = \begin{bmatrix} 0 \\ 1 \\ \underline{1} \\ 0 \\ 1 \end{bmatrix} \quad pred = \begin{bmatrix} 0 \\ 1 \\ \underline{0} \\ 0 \\ 1 \end{bmatrix}$$

$$(Y_{val} == pred) =$$

$$\begin{bmatrix} 1 \\ 1 \\ \underline{0} \\ 1 \\ 1 \end{bmatrix}$$

$$\text{accuracy} = \frac{4}{5} = 0.8$$