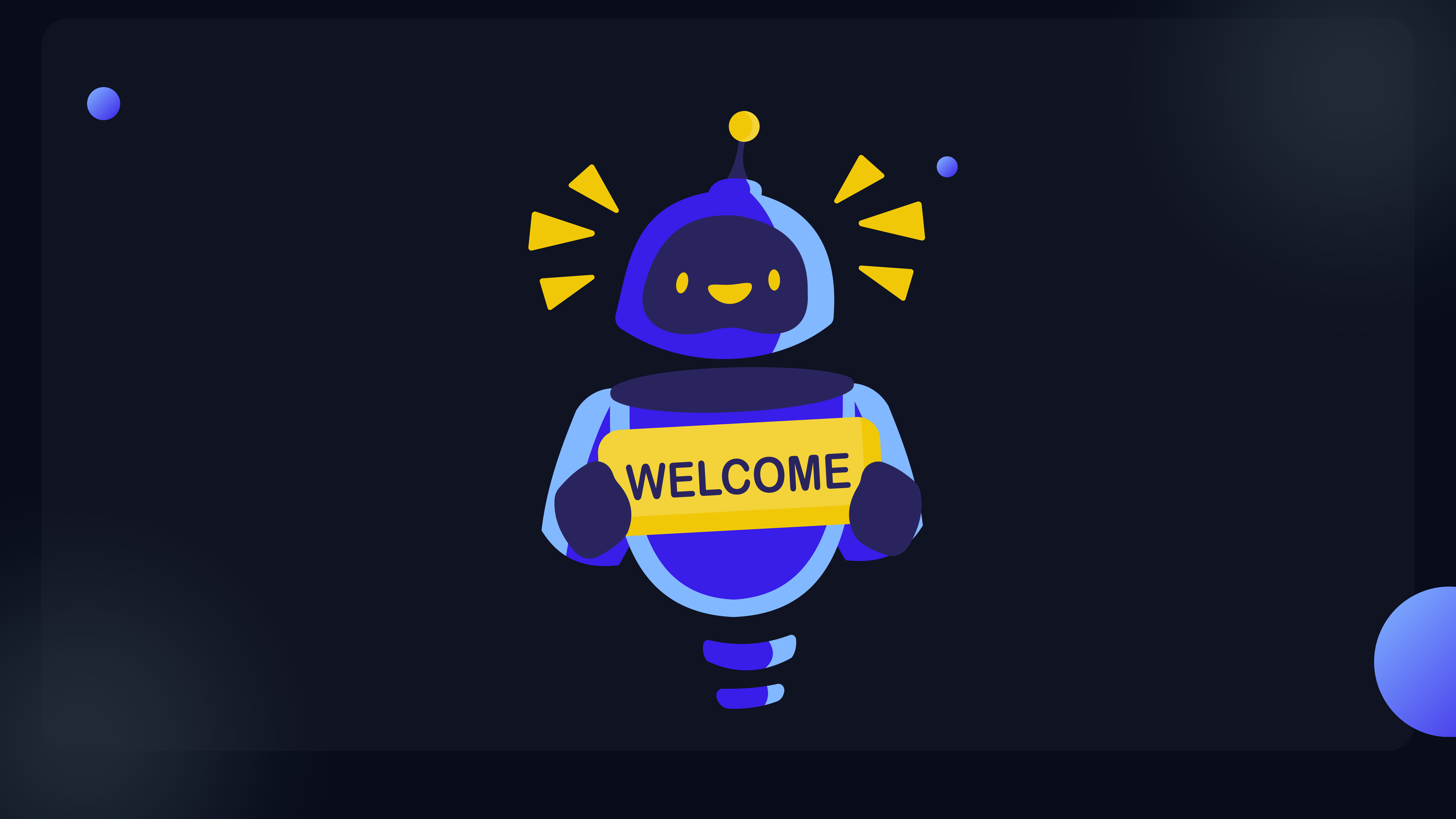
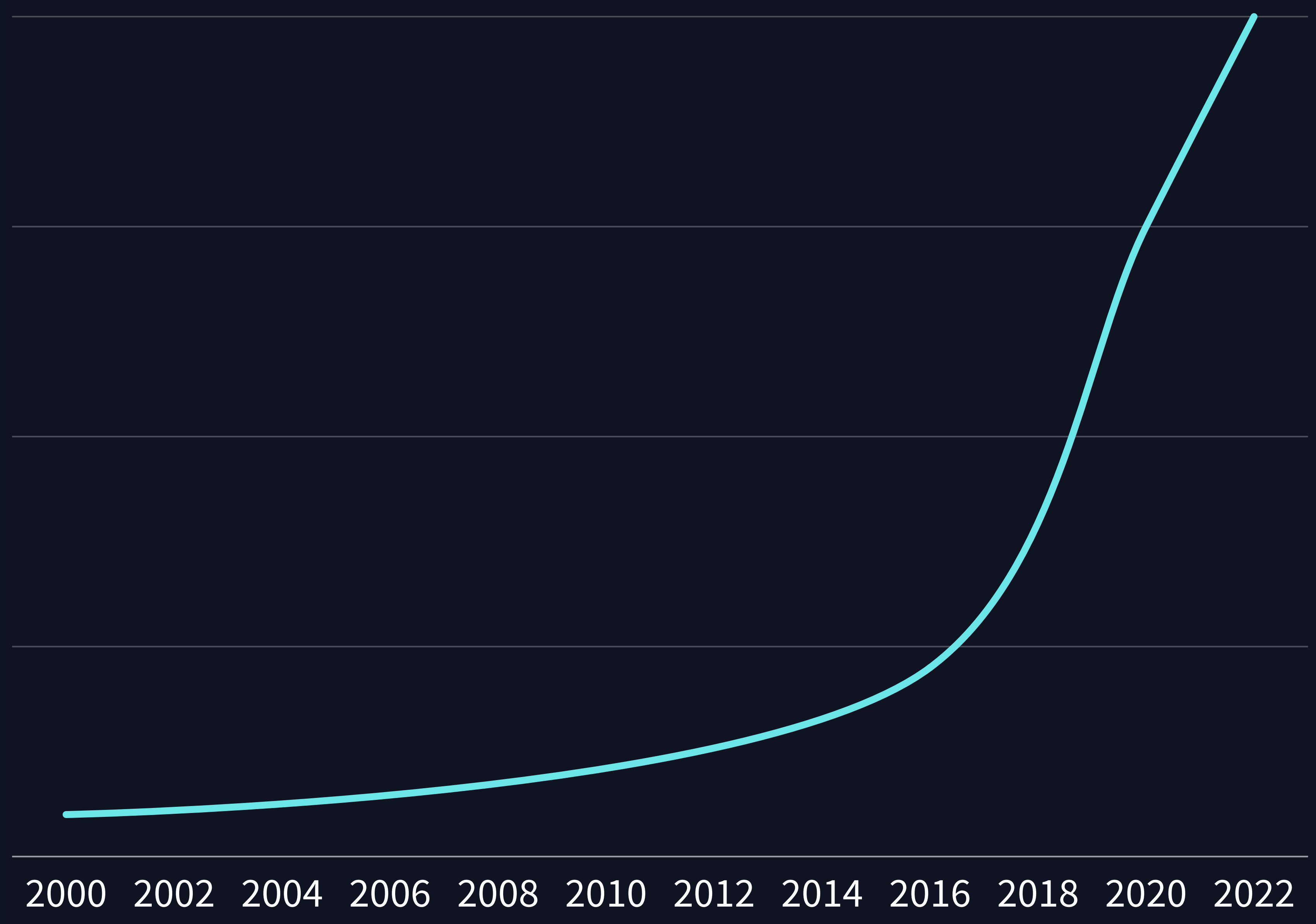
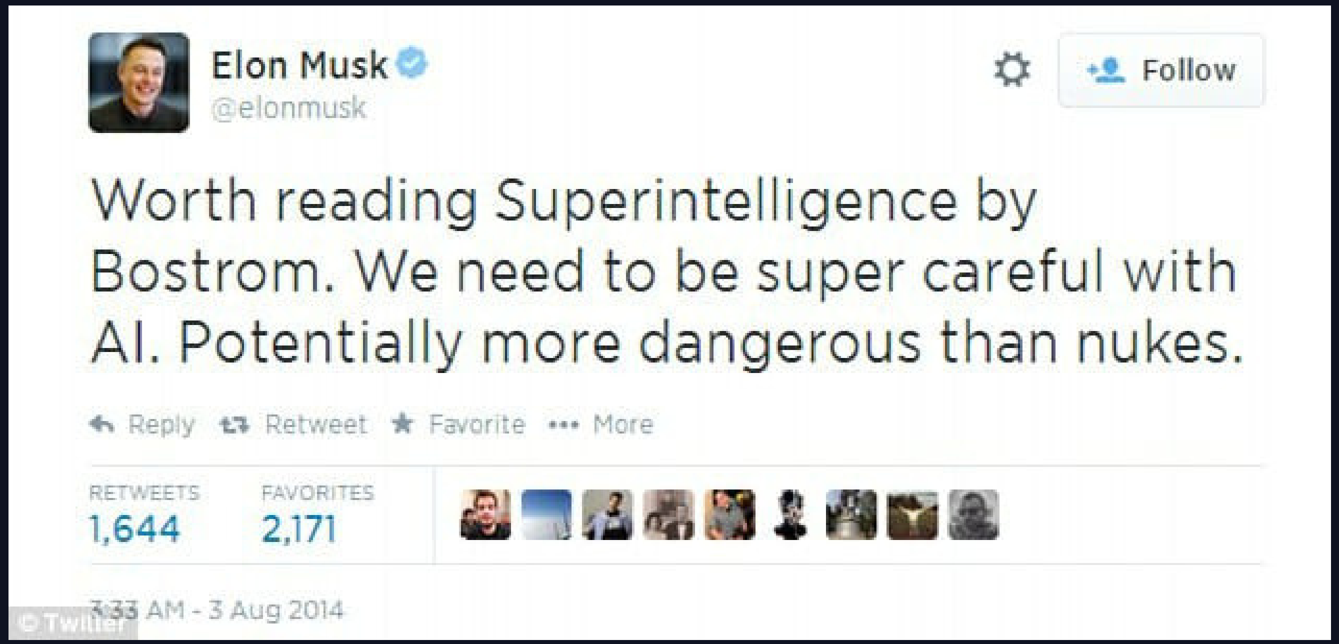

WELCOME

WHAT IS MACHINE LEARNING

WHY ML IS GETTING DEMAND NOW?



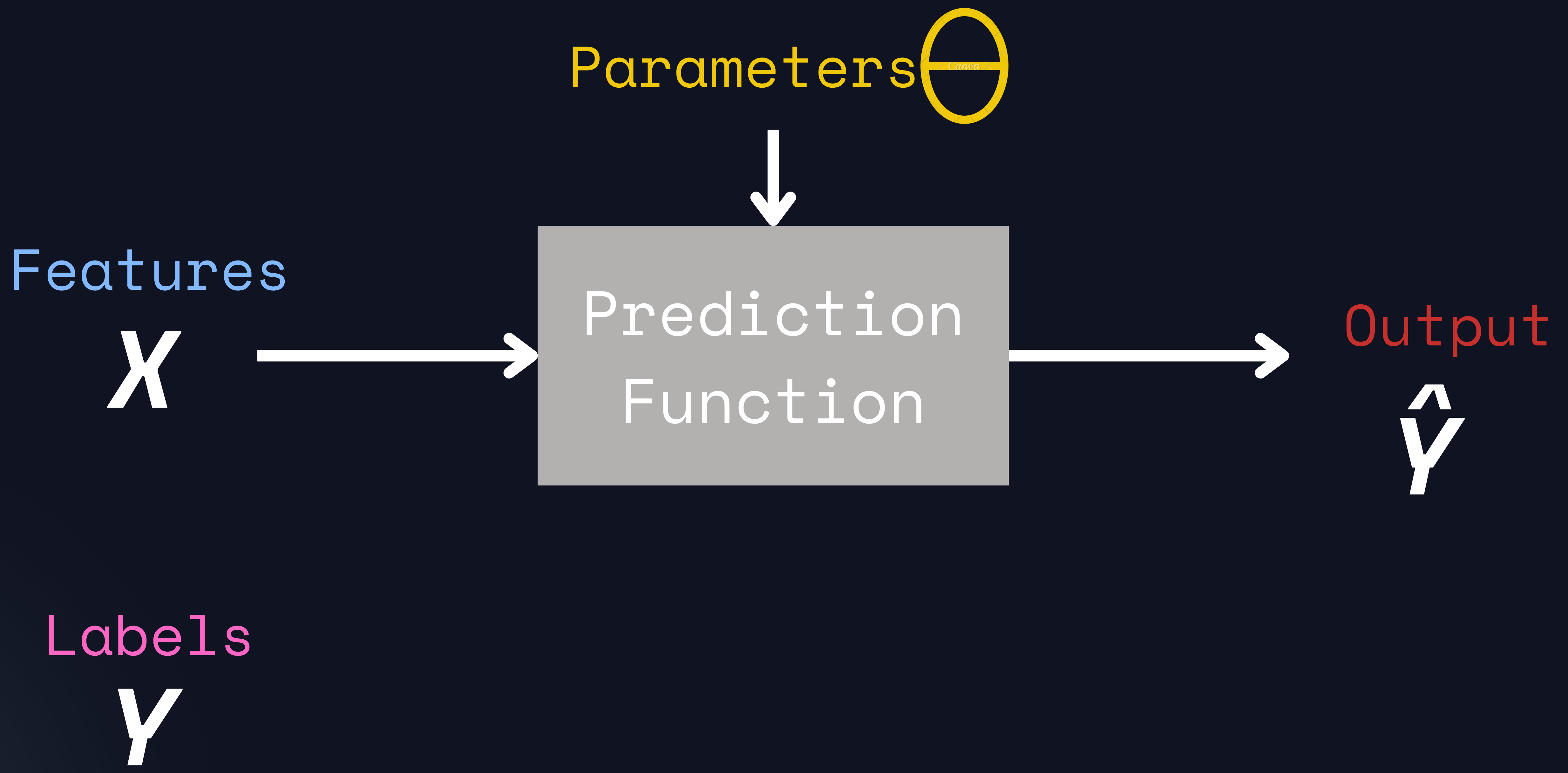
NATURAL LANGUGE PROCESSING


Elon Musk
@elonmusk
Follow
Worth reading Superintelligence by Bostrom. We need to be super careful with AI. Potentially more dangerous than nukes.
Reply Retweet Favorite More
RETWEETS
1,644
FAVORITES
2,171
7:33 AM - 3 Aug 2014
© Twitter


Elon Musk
@elonmusk · Follow
He's fired
9:02 PM · Nov 14, 2022
40.3K
Reply
Share
Read 7.1K replies

SUPERVISED ML &
SENTIMENT ANALYSIS

SUPERVISED ML (TRAINING)



SENTIMENT ANALYSIS

Tweet: I am happy I am learning NLP

Positive

< 0 <

Negative

SENTIMENT ANALYSIS

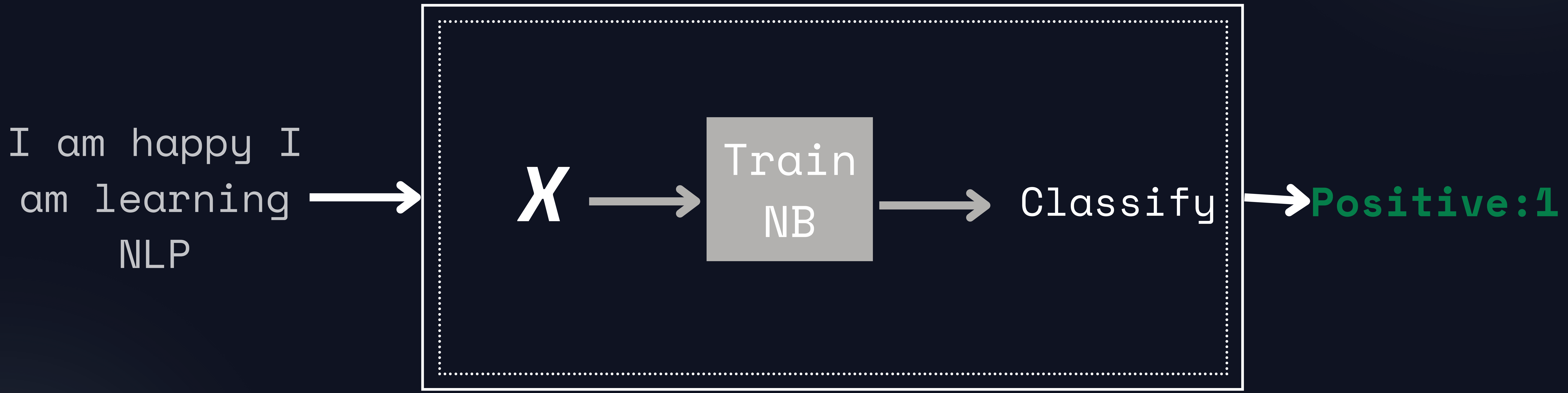
Positive:1

Tweet: I am happy I am learning NLP

Negative:0

Naive Bayes

SENTIMENT ANALYSIS



VOCABULARY & FEATURE EXTRACTION

VOCABULARY

Tweets:

[tweet_1,tweet2, ..., tweet_m]

```
I am happy I am learning NLP
...
...
...
I hate the movie
```

$V=$

[I,am, happy, learning, NLP,..., hate, the, movie]

FEATURE EXTRACTION

I am happy I am learning NLP

FEATURE EXTRACTION

I am happy I am learning NLP

[I, am, happy, learning, NLP,..., hate, the, movie]

[1, 1

FEATURE EXTRACTION

I am happy I am learning NLP

[I, am, happy, learning, NLP,..., hate, the, movie]

[1, 1 1

FEATURE EXTRACTION

I am happy I am learning NLP

[I, am, happy, learning, NLP,..., hate, the, movie]

[1, 1 1 1,

FEATURE EXTRACTION

I am happy I am learning NLP

[I, am, happy, learning, NLP,..., hate, the, movie]

[1, 1, 1, 1, 1,

FEATURE EXTRACTION

I am happy I am learning NLP

[I, am, happy, learning, NLP,..., hate, the, movie]

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

FEATURE EXTRACTION

I am happy I am learning NLP

[I, am, happy, learning, NLP,..., hate, the, movie]

$$\begin{array}{cccccccc} I, & am, & happy, & learning, & NLP, \ldots, & hate, & the, & movie \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ 1, & 1, & 1, & 1, & 1, \ldots, & 0, & 0, & 0 \end{array}$$

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

FEATURE EXTRACTION

I am happy I am learning NLP

[I, am, happy, learning, NLP,..., hate, the, movie]

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

VOCABULARY

I am happy I am learning NLP

VOCABULARY

I am happy I am learning NLP

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

VOCABULARY

I am happy I am learning NLP

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

$1 \longleftrightarrow |V|$

VOCABULARY

I am happy I am learning NLP

All Zeros

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

1 $\longleftrightarrow$ $|V|$

VOCABULARY

I am happy I am learning NLP

All Zeros

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

$1 \longleftrightarrow |V|$

$$[\theta_1, \theta_2, \theta_3, \ldots, \theta_n]$$

VOCABULARY

I am happy I am learning NLP

All Zeros

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

$1 \longleftrightarrow |V|$

$$[\theta_1, \theta_2, \theta_3, \ldots, \theta_n]$$

$$n = |V|$$

VOCABULARY

I am happy I am learning NLP

All Zeros

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

$1 \longleftrightarrow |V|$

$$[\theta_1, \theta_2, \theta_3, \ldots, \theta_n]$$

1) Large Traning Time
2) Large Prediction Time

$$n = |V|$$

POSITIVE & NEGATIVE FREQUENCE

POSITIVE & NEGATIVE COUNT

Corpos:

I am happy because I am learning NLP

I am happy

I am sad I am not Learning NLP

I am sad

POSITIVE & NEGATIVE COUNT

Corpos:

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 & NEGATIVE COUNT

Positive:

```
I am happy becaues I am

learning NLP

I am happy
```

POSITIVE & NEGATIVE COUNT

Positive:

I am happy becaues I am learning NLP

I am happy

Negative:

I am sad I am not Learning NLP

I am sad

POSITIVE & NEGATIVE COUNT

Positive:

I am happy becaues I am

learning NLP

I am happy

VOCABULARY
I
am
happy
because
learning
NLP
sad
not

POSITIVE & NEGATIVE COUNT

Positive:

I am happy becaues I am learning NLP

I am happy

VOCABULARY	PosFreq(1)
I	
am	
happy	
because	
learning	
NLP	
sad	
not	

POSITIVE & NEGATIVE COUNT

Positive:

I am happy becaues I am learning NLP

I am happy

VOCABULARY	PosFreq(1)
I	
am	
happy	2
because	
learning	
NLP	
sad	
not	

POSITIVE & NEGATIVE COUNT

Positive:

I am happy becaues I am

learning NLP

I am happy

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

POSITIVE & NEGATIVE COUNT

Negative:

I am sad I am not Learning NLP

I am sad

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

POSITIVE & NEGATIVE COUNT

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

FEATURE EXTRACTION WITH FREQUENCIES

WORD FREQUENCE IN CLASSES

WORD FREQUENCE 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 FREQUENCE 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

Freq: Dictionary mapping from (word, class) to frequency

FEATURE EXTRACTION

FEATURE EXTRACTION

Freq: Dictionary mapping from (word, class) to frequency

FEATURE EXTRACTION

Freq: Dictionary mapping from (word, class) to frequency

$$X_m = [\quad]$$

Features of tweets m

FEATURE EXTRACTION

Freq: Dictionary mapping from (word, class) to frequency

$$X_m = [\ 1, \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad]$$

Features of tweets m

Bias

FEATURE EXTRACTION

Freq: Dictionary mapping from (word, class) to frequency

$$X_m = [\ 1,\ \sum freqs(w,1),\ \ ]$$

Features of tweets m

Bias

Sum Pos Frequence

FEATURE EXTRACTION

Freq: Dictionary mapping from (word, class) to frequency

$$X_m = [\ 1, \sum freqs(w,1), \sum freqs(w,0)\ ]$$

Features of tweets m

Bias

Sum Pos Frequencies

Sum Neg Frequencies

FEATURE EXTRACTION

I am sad, I am not Learning NLP

FEATURE EXTRACTION

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

I am sad, I am not Learning NLP

FEATURE EXTRACTION

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

I am sad, I am not Learning NLP

FEATURE EXTRACTION

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

I am sad, I am not Learning NLP

$$X_m = [\ 1, \sum freqs(w,1), \sum freqs(w,0)\ ]$$

FEATURE EXTRACTION

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

I am sad, I am not Learning NLP

$$X_m = [\ 1, \sum freqs(w,1), \sum freqs(w,0)\ ]$$

FEATURE EXTRACTION

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

I am sad, I am not Learning NLP

$$X_m = [\ 1, \sum freqs(w,1), \sum freqs(w,0)\ ]$$

8

FEATURE EXTRACTION

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

I am sad, I am not Learning NLP

$$X_m = [\ 1, \sum freqs(w,1), \sum freqs(w,0)\ ]$$

FEATURE EXTRACTION

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

I am sad, I am not Learning NLP

$$X_m = [\ 1, \sum freqs(w,1), \underbrace{\sum freqs(w,0)}_{11}\ ]$$

FEATURE EXTRACTION

I am sad, I am not Learning NLP

$$X_m = [\ 1, \sum freqs(w,1), \sum freqs(w,0)\ ]$$

$$\downarrow$$

$$X_m = [1, 8, 11]$$

PREPROCESSING

PREPROCESSING: STOP WORD & PUNCTUATION

PREPROCESSING: STOP WORD & PUNCTUATION

@PSGiTech is conducting YUKTA

to register https://yuktaha.psgitech.ac.in/

PREPROCESSING: STOP WORD & PUNCTUATION

@PSGiTech is conducting YUKTA!!!

to register

https://yuktaha.psgitech.ac.in/

STOPWORD
and
is
a
at
has
for
of

PREPROCESSING: STOP WORD & PUNCTUATION

@PSGiTech ~~is~~ conducting YUKTA!!!

~~to~~ register

https://yuktaha.psgitech.ac.in/

STOPWORD
and
is
a
at
has
for
of

PREPROCESSING: STOP WORD & PUNCTUATION

@PSGiTech ~~is~~ conducting YUKTA!!!

~~to~~ register

https://yuktaha.psgitech.ac.in/

@PSGiTech conducting YUKTA!!!

register

https://yuktaha.psgitech.ac.in/

STOPWORD
and
is
a
at
has
for
of

PREPROCESSING: STOP WORD & PUNCTUATION

@PSGiTech conducting YUKTA!!!

register

https://yuktaha.psgitech.ac.in/

PUNCTUATION
,
.
:
!
"
'

PREPROCESSING: STOP WORD & PUNCTUATION

@PSGiTech conducting YUKTA!~~!~~!

register

https://yuktaha.psgitech.ac.in/

PUNCTUATION

,

.

:

!

"

'

PREPROCESSING: STOP WORD & PUNCTUATION

@PSGiTech conducting YUKTA!~~!~~!

register

https://yuktaha.psgitech.ac.in/

@PSGiTech conducting YUKTA

register

https://yuktaha.psgitech.ac.in/

PUNCTUATION

’
.
:
!
"
,

PREPROCESSING: STOP WORD & PUNCTUATION

@PSGiTech conducting YUKTA register

https://yuktaha.psgitech.ac.in/

PREPROCESSING: HANDLES AND URLS

~~@PSGiTech~~ conducting YUKTA register

~~https://yuktaha.psgitech.ac.in/~~

↓

conducting YUKTA register

PREPROCESSING: STEMMING & LOWERCASE

conducting YUKTA register

PREPROCESSING: STEMMING & LOWERCASE

conduct~~ing~~ YUKTA register

conduct

PREPROCESSING: STEMMING & LOWERCASE

conduct~~ing~~ YUKTA register

conduct

conductivity

PREPROCESSING: STEMMING & LOWERCASE

conduct~~ing~~ YUKTA register

conduct

conductivity

conducted

PREPROCESSING: STEMMING & LOWERCASE

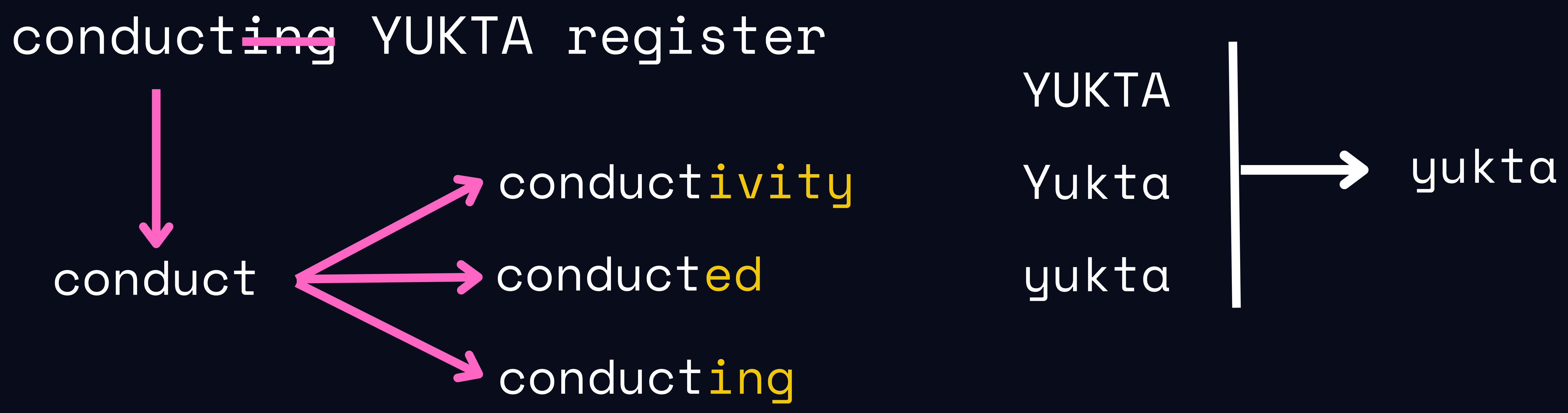
conduct~~ing~~ YUKTA register

conduct → conductivity

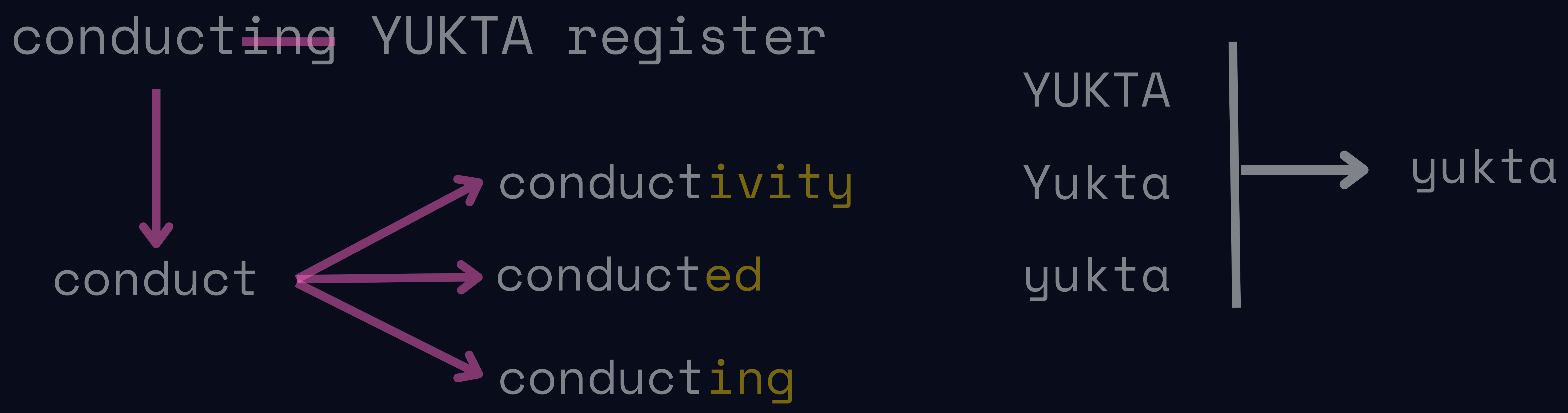
conduct → conducted

conduct → conducting

PREPROCESSING: STEMMING & LOWERCASE



PREPROCESSING: STEMMING & LOWERCASE



PROCESSED TWEET:

conduct yukta register

NAIVEE BAYES FOR SENTIMENT ANALAYSIS

NAIVEE BAYES FOR SENTIMENT ANALAYSIS

Positive tweets

I am happy because I am learning NLP

I am happy, not sad

Negative tweets

I am sad I am not Learning NLP

I am sad, not happy

NAIVEE BAYES FOR SENTIMENT ANALAYSIS

Positive tweets

I am happy because I am learning NLP

I am happy, not sad

Negative tweets

I am sad I am not Learning NLP

I am sad, not happy

WORD	Pos	Neg
I	3	3
am	3	3
happy	2	1
because	1	0
learning	1	1
NLP	1	1
sad	1	2
not	1	2

NAIVEE BAYES FOR SENTIMENT ANALAYSIS

Positive tweets

I am happy because I am learning NLP

I am happy, not sad

Negative tweets

I am sad I am not Learning NLP

I am sad, not happy

WORD	Pos	Neg
I	3	3
am	3	3
happy	2	1
because	1	0
learning	1	1
NLP	1	1
sad	1	2
not	1	2
N_{class}	13	13

NAIVEE BAYES FOR SENTIMENT ANALAYSIS

$P(w_i | class)$

WORD	Pos	Neg
I	3	3
am	3	3
happy	2	1
because	1	0
learning	1	1
NLP	1	1
sad	1	2
not	1	2
N_{class}	13	13

WORD	Pos	Neg

NAIVEE BAYES FOR SENTIMENT ANALAYSIS

$P(w_i \mid class)$

WORD	Pos	Neg
I	3	3
am	3	3
happy	2	1
because	1	0
learning	1	1
NLP	1	1
sad	1	2
not	1	2
N_{class}	13	13

WORD	Pos	Neg
I		

NAIVEE BAYES FOR SENTIMENT ANALAYSIS

$P(w_i | class)$

WORD	Pos	Neg
I	3	3
am	3	3
happy	2	1
because	1	0
learning	1	1
NLP	1	1
sad	1	2
not	1	2
N_{class}	13	13

$$P(I|pos) = \frac{3}{13}$$

WORD	Pos	Neg
I	0.24	

NAIVEE BAYES FOR SENTIMENT ANALAYSIS

$P(w_i \mid class)$

WORD	Pos	Neg
I	3	3
am	3	3
happy	2	1
because	1	0
learning	1	1
NLP	1	1
sad	1	2
not	1	2
N_{class}	13	13

$P(I|pos) = \frac{3}{13}$

WORD	Pos	Neg
I	0.24	0.24

NAIVEE BAYES FOR SENTIMENT ANALAYSIS

$P(w_i \mid class)$

WORD	Pos	Neg
I	3	3
am	3	3
happy	2	1
because	1	0
learning	1	1
NLP	1	1
sad	1	2
not	1	2
N_{class}	13	13

WORD	Pos	Neg
I	0.24	0.24
am	0.24	0.24
happy	0.15	0.08
because	0.08	0
learning	0.08	0.08
NLP	0.08	0.08
sad	0.08	0.15
not	0.08	0.15

NAIVEE BAYES FOR SENTIMENT ANALAYSIS

$P(w_i \mid class)$

WORD	Pos	Neg
I	3	3
am	3	3
happy	2	1
because	1	0
learning	1	1
NLP	1	1
sad	1	2
not	1	2
N_{class}	13	13

WORD	Pos	Neg
I	0.24	0.24
am	0.24	0.24
happy	0.15	0.08
because	0.08	0
learning	0.08	0.08
NLP	0.08	0.08
sad	0.08	0.15
not	0.08	0.15
Sum	1	1

NAIVEE BAYES FOR SENTIMENT ANALAYSIS

$P(w_i | class)$

WORD	Pos	Neg
I	0.24	0.24
am	0.24	0.24
happy	0.15	0.08
because	0.08	0
learning	0.08	0.08
NLP	0.08	0.08
sad	0.08	0.15
not	0.08	0.15

NAIVEE BAYES FOR SENTIMENT ANALAYSIS

$P(w_i \mid class)$

WORD	Pos	Neg
I	0.24	0.24
am	0.24	0.24
happy	0.15	0.08
because	0.08	0
learning	0.08	0.08
NLP	0.08	0.08
sad	0.08	0.15
not	0.08	0.15

NAIVEE BAYES FOR SENTIMENT ANALAYSIS

$P(w_i \mid class)$

WORD	Pos	Neg
I	0.24	0.24
am	0.24	0.24
happy	0.15	0.08
because	0.08	0
learning	0.08	0.08
NLP	0.08	0.08
sad	0.08	0.15
not	0.08	0.15

NAIVEE BAYES FOR SENTIMENT ANALAYSIS

$P(w_i \mid class)$

WORD	Pos	Neg
I	0.24	0.24
am	0.24	0.24
happy	0.15	0.08
because	0.08	0
learning	0.08	0.08
NLP	0.08	0.08
sad	0.08	0.15
not	0.08	0.15

NAIVEE BAYES

Tweet: I am happy today; I am learning

NAIVEE BAYES

Tweet: I am happy today; I am learning

WORD	Pos	Neg
I	0.20	0.20
am	0.20	0.20
happy	0.14	0.10
because	0.10	0.05
learning	0.10	0.10
NLP	0.10	0.10
sad	0.10	0.14
not	0.10	0.14

NAIVEE BAYES

Tweet: I am happy today; I am learning

$$\prod_{i=1}^{m} \frac{P(w_i|Pos)}{P(w_i|Neg)}$$

WORD	Pos	Neg
I	0.20	0.20
am	0.20	0.20
happy	0.14	0.10
because	0.10	0.05
learning	0.10	0.10
NLP	0.10	0.10
sad	0.10	0.14
not	0.10	0.14

NAIVEE BAYES

Tweet: I am happy today; I am learning

$$\prod_{i=1}^{m} \frac{P(w_i|Pos)}{P(w_i|Neg)}$$

$$\frac{0.20}{0.20}$$

WORD	Pos	Neg
I	0.20	0.20
am	0.20	0.20
happy	0.14	0.10
because	0.10	0.05
learning	0.10	0.10
NLP	0.10	0.10
sad	0.10	0.14
not	0.10	0.14

NAIVEE BAYES

Tweet: I am happy today; I am learning

$$\prod_{i=1}^{m} \frac{P(w_i|Pos)}{P(w_i|Neg)}$$

$$\frac{0.20}{0.20} * \frac{0.20}{0.20}$$

WORD	Pos	Neg
I	0.20	0.20
am	0.20	0.20
happy	0.14	0.10
because	0.10	0.05
learning	0.10	0.10
NLP	0.10	0.10
sad	0.10	0.14
not	0.10	0.14

NAIVEE BAYES

Tweet: I am happy today; I am learning

$$\prod_{i=1}^{m} \frac{P(w_i|Pos)}{P(w_i|Neg)}$$

$$\frac{0.20}{0.20} * \frac{0.20}{0.20} * \frac{0.14}{0.10}$$

WORD	Pos	Neg
I	0.20	0.20
am	0.20	0.20
happy	0.14	0.10
because	0.10	0.05
learning	0.10	0.10
NLP	0.10	0.10
sad	0.10	0.14
not	0.10	0.14

NAIVEE BAYES

Tweet: I am happy today; I am learning

$$\prod_{i=1}^{m} \frac{P(w_i|Pos)}{P(w_i|Neg)}$$

$$\frac{0.20}{0.20} * \frac{0.20}{0.20} * \frac{0.14}{0.10}$$

WORD	Pos	Neg
I	0.20	0.20
am	0.20	0.20
happy	0.14	0.10
because	0.10	0.05
learning	0.10	0.10
NLP	0.10	0.10
sad	0.10	0.14
not	0.10	0.14

NAIVEE BAYES

Tweet: I am happy today; I am learning

$$\prod_{i=1}^{m} \frac{P(w_i|Pos)}{P(w_i|Neg)}$$

$$\frac{0.20}{0.20} * \frac{0.20}{0.20} * \frac{0.14}{0.10} * \frac{0.20}{0.20}$$

WORD	Pos	Neg
I	0.20	0.20
am	0.20	0.20
happy	0.14	0.10
because	0.10	0.05
learning	0.10	0.10
NLP	0.10	0.10
sad	0.10	0.14
not	0.10	0.14

NAIVEE BAYES

Tweet: I am happy today; I am learning

$$\prod_{i=1}^{m} \frac{P(w_i|Pos)}{P(w_i|Neg)}$$

$$\frac{0.20}{0.20} * \frac{0.20}{0.20} * \frac{0.14}{0.10} * \frac{0.20}{0.20} * \frac{0.20}{0.20}$$

WORD	Pos	Neg
I	0.20	0.20
am	0.20	0.20
happy	0.14	0.10
because	0.10	0.05
learning	0.10	0.10
NLP	0.10	0.10
sad	0.10	0.14
not	0.10	0.14

NAIVEE BAYES

Tweet: I am happy today; I am learning

$$\prod_{i=1}^{m} \frac{P(w_i|Pos)}{P(w_i|Neg)}$$

$$\frac{0.20}{0.20} * \frac{0.20}{0.20} * \frac{0.14}{0.10} * \frac{0.20}{0.20} * \frac{0.20}{0.20} * \frac{0.10}{0.10}$$

WORD	Pos	Neg
I	0.20	0.20
am	0.20	0.20
happy	0.14	0.10
because	0.10	0.05
learning	0.10	0.10
NLP	0.10	0.10
sad	0.10	0.14
not	0.10	0.14

NAIVEE BAYES

Tweet: I am happy today; I am learning

$$\prod_{i=1}^{m} \frac{P(w_i|Pos)}{P(w_i|Neg)}$$

WORD	Pos	Neg
I	0.20	0.20
am	0.20	0.20
happy	0.14	0.10
because	0.10	0.05
learning	0.10	0.10
NLP	0.10	0.10
sad	0.10	0.14
not	0.10	0.14

$$\cancel{\frac{0.20}{0.20}} * \cancel{\frac{0.20}{0.20}} * \frac{0.14}{0.10} * \cancel{\frac{0.20}{0.20}} * \cancel{\frac{0.20}{0.20}} * \cancel{\frac{0.10}{0.10}}$$

NAIVEE BAYES

Tweet: I am happy today; I am learning

$$\prod_{i=1}^{m} \frac{P(w_i|Pos)}{P(w_i|Neg)}$$

$$\cancel{\frac{0.20}{0.20}} * \cancel{\frac{0.20}{0.20}} * \frac{0.14}{0.10} * \cancel{\frac{0.20}{0.20}} * \cancel{\frac{0.20}{0.20}} * \cancel{\frac{0.10}{0.10}}$$

WORD	Pos	Neg
I	0.20	0.20
am	0.20	0.20
happy	0.14	0.10
because	0.10	0.05
learning	0.10	0.10
NLP	0.10	0.10
sad	0.10	0.14
not	0.10	0.14

NAIVEE BAYES

Tweet: I am happy today; I am learning

$$\prod_{i=1}^{m} \frac{P(w_i|Pos)}{P(w_i|Neg)} = \frac{0.14}{0.10} = 1.4$$

$$\cancel{\frac{0.20}{0.20}} * \cancel{\frac{0.20}{0.20}} * \frac{0.14}{0.10} * \cancel{\frac{0.20}{0.20}} * \cancel{\frac{0.20}{0.20}} * \cancel{\frac{0.10}{0.10}}$$

WORD	Pos	Neg
I	0.20	0.20
am	0.20	0.20
happy	0.14	0.10
because	0.10	0.05
learning	0.10	0.10
NLP	0.10	0.10
sad	0.10	0.14
not	0.10	0.14

NAIVEE BAYES

Tweet: I am happy today; I am learning

$$\prod_{i=1}^{m} \frac{P(w_i|Pos)}{P(w_i|Neg)} = \frac{0.14}{0.10} = 1.4 > 1$$

WORD	Pos	Neg
I	0.20	0.20
am	0.20	0.20
happy	0.14	0.10
because	0.10	0.05
learning	0.10	0.10
NLP	0.10	0.10
sad	0.10	0.14
not	0.10	0.14

$$\cancel{\frac{0.20}{0.20}} * \cancel{\frac{0.20}{0.20}} * \frac{0.14}{0.10} * \cancel{\frac{0.20}{0.20}} * \cancel{\frac{0.20}{0.20}} * \cancel{\frac{0.10}{0.10}}$$

LAPLACIAN SMOOTHING

LAPLACIAN SMOOTHING

$$P(w_i \mid class) = \frac{freq(w_i, class)}{N_{class}}$$

class - {Positive,Negative}

N_{class} = frequency of all words in class

V_{class} = number of unique words in class

LAPLACIAN SMOOTHING

$$P(w_i \mid class) = \frac{freq(w_i, class)}{N_{class}}$$

class - {Positive,Negative}

$$P(w_i \mid class) = \frac{freq(w_i, class) + 1}{N_{class} + V_{class}}$$

N_{class} = frequency of all words in class

V_{class} = number of unique words in class

LAPLACIAN SMOOTHING

$$P(w_i \mid class) = \frac{freq(w_i, class)}{N_{class}}$$

class - {Positive,Negative}

$$P(w_i \mid class) = \frac{freq(w_i, class) + 1}{N_{class} + V_{class}}$$

N_{class} = frequency of all words in class

V_{class} = number of unique words in class

LAPLACIAN SMOOTHING

$$P(w_i \mid class) = \frac{freq(w_i, class)}{N_{class}}$$

class - {Positive,Negative}

$$P(w_i \mid class) = \frac{freq(w_i, class) + 1}{N_{class} + V_{class}}$$

N_{class} = frequency of all words in class

V_{class} = number of unique words in class

LAPLACIAN SMOOTHING

Introducing $P(w_i|class)$ with smoothing

WORD	Pos	Neg
I	3	3
am	3	3
happy	2	1
because	1	0
learning	1	1
NLP	1	1
sad	1	2
not	1	2
N_{class}	13	13

WORD	Pos	Neg

LAPLACIAN SMOOTHING

Introducing $P(w_i|class)$ with smoothing

WORD	Pos	Neg
I	3	3
am	3	3
happy	2	1
because	1	0
learning	1	1
NLP	1	1
sad	1	2
not	1	2
N_{class}	13	13

$V = 8$

WORD	Pos	Neg
I		

LAPLACIAN SMOOTHING

Introducing $P(w_i|class)$ with smoothing

WORD	Pos	Neg
I	3	3
am	3	3
happy	2	1
because	1	0
learning	1	1
NLP	1	1
sad	1	2
not	1	2
N_{class}	13	13

$$P(I|pos) = \frac{3 + 1}{13 + 8}$$

WORD	Pos	Neg
I	0.20	

$V = 8$

LAPLACIAN SMOOTHING

Introducing $P(w_i|class)$ with smoothing

WORD	Pos	Neg
I	3	3
am	3	3
happy	2	1
because	1	0
learning	1	1
NLP	1	1
sad	1	2
not	1	2
N_{class}	13	13

$$P(I|pos) = \frac{3+1}{13+8}$$

$V = 8$

WORD	Pos	Neg
I	0.20	0.20

LAPLACIAN SMOOTHING

Introducing $P(w_i|class)$ with smoothing

WORD	Pos	Neg
I	3	3
am	3	3
happy	2	1
because	1	0
learning	1	1
NLP	1	1
sad	1	2
not	1	2
N_{class}	13	13

$V = 8$

WORD	Pos	Neg
I	0.20	0.20
am	0.20	0.20
happy	0.14	0.10
because	0.10	0.05
learning	0.10	0.10
NLP	0.10	0.10
sad	0.10	0.14
not	0.10	0.14

LAPLACIAN SMOOTHING

Introducing $P(w_i|class)$ with smoothing

WORD	Pos	Neg
I	3	3
am	3	3
happy	2	1
because	1	0
learning	1	1
NLP	1	1
sad	1	2
not	1	2
N_{class}	13	13

$V = 8$

WORD	Pos	Neg
I	0.20	0.20
am	0.20	0.20
happy	0.14	0.10
because	0.10	0.05
learning	0.10	0.10
NLP	0.10	0.10
sad	0.10	0.14
not	0.10	0.14
Sum	1	1

LAPLACIAN SMOOTHING

Introducing $P(w_i|class)$ with smoothing

WORD	Pos	Neg
I	3	3
am	3	3
happy	2	1
because	1	0
learning	1	1
NLP	1	1
sad	1	2
not	1	2
N_{class}	13	13

$V = 8$

WORD	Pos	Neg
I	0.20	0.20
am	0.20	0.20
happy	0.14	0.10
because	0.10	0.05
learning	0.10	0.10
NLP	0.10	0.10
sad	0.10	0.14
not	0.10	0.14
Sum	1	1

LOG LIKELIHOOD

Ratio of Probabilities

WORD	Pos	Neg
I	0.20	0.20
am	0.20	0.20
happy	0.14	0.10
because	0.10	0.05
learning	0.10	0.10
NLP	0.10	0.14
sad	0.10	0.14
not	0.10	0.14

LOG LIKELIHOOD

Ratio of Probabilities

WORD	Pos	Neg
I	0.20	0.20
am	0.20	0.20
happy	0.14	0.10
because	0.10	0.05
learning	0.10	0.10
NLP	0.10	0.14
sad	0.10	0.14
not	0.10	0.14

$$\text{ratio}(w_i) = \frac{P(w_i \mid \text{Pos})}{P(w_i \mid \text{Neg})}$$

LOG LIKELIHOOD

Ratio of Probabilities

WORD	Pos	Neg	ratio
I	0.20	0.20	
am	0.20	0.20	
happy	0.14	0.10	
because	0.10	0.05	
learning	0.10	0.10	
NLP	0.10	0.14	
sad	0.10	0.14	
not	0.10	0.14	

$$\text{ratio}(w_i) = \frac{P(w_i \mid \text{Pos})}{P(w_i \mid \text{Neg})}$$

LOG LIKELIHOOD

Ratio of Probabilities

WORD	Pos	Neg	ratio
I	0.20	0.20	1
am	0.20	0.20	
happy	0.14	0.10	
because	0.10	0.05	
learning	0.10	0.10	
NLP	0.10	0.14	
sad	0.10	0.14	
not	0.10	0.14	

$$\text{ratio}(w_i) = \frac{P(w_i \mid \text{Pos})}{P(w_i \mid \text{Neg})}$$

LOG LIKELIHOOD

Ratio of Probabilities

WORD	Pos	Neg	ratio
I	0.20	0.20	1
am	0.20	0.20	1
happy	0.14	0.10	
because	0.10	0.05	
learning	0.10	0.10	
NLP	0.10	0.14	
sad	0.10	0.14	
not	0.10	0.14	

$$\text{ratio}(w_i) = \frac{P(w_i \mid \text{Pos})}{P(w_i \mid \text{Neg})}$$

LOG LIKELIHOOD

Ratio of Probabilities

WORD	Pos	Neg	ratio
I	0.20	0.20	1
am	0.20	0.20	1
happy	0.14	0.10	1.4
because	0.10	0.05	
learning	0.10	0.10	
NLP	0.10	0.14	
sad	0.10	0.14	
not	0.10	0.14	

$$\text{ratio}(w_i) = \frac{P(w_i \mid \text{Pos})}{P(w_i \mid \text{Neg})}$$

LOG LIKELIHOOD

Ratio of Probabilities

WORD	Pos	Neg	ratio
I	0.20	0.20	1
am	0.20	0.20	1
happy	0.14	0.10	1.4
because	0.10	0.05	2
learning	0.10	0.10	
NLP	0.10	0.14	
sad	0.10	0.14	
not	0.10	0.14	

$$\text{ratio}(w_i) = \frac{P(w_i \mid \text{Pos})}{P(w_i \mid \text{Neg})}$$

LOG LIKELIHOOD

Ratio of Probabilities

WORD	Pos	Neg	ratio
I	0.20	0.20	1
am	0.20	0.20	1
happy	0.14	0.10	1.4
because	0.10	0.05	2
learning	0.10	0.10	1
NLP	0.10	0.14	
sad	0.10	0.14	
not	0.10	0.14	

$$\text{ratio}(w_i) = \frac{P(w_i|\text{Pos})}{P(w_i|\text{Neg})}$$

LOG LIKELIHOOD

Ratio of Probabilities

WORD	Pos	Neg	ratio
I	0.20	0.20	1
am	0.20	0.20	1
happy	0.14	0.10	1.4
because	0.10	0.05	2
learning	0.10	0.10	1
NLP	0.10	0.14	0.6
sad	0.10	0.14	
not	0.10	0.14	

$$\text{ratio}(w_i) = \frac{P(w_i \mid \text{Pos})}{P(w_i \mid \text{Neg})}$$

LOG LIKELIHOOD

Ratio of Probabilities

Neutral 1

WORD	Pos	Neg	ratio
I	0.20	0.20	1
am	0.20	0.20	1
happy	0.14	0.10	1.4
because	0.10	0.05	2
learning	0.10	0.10	1
NLP	0.10	0.14	0.6
sad	0.10	0.14	0.6
not	0.10	0.14	0.6

$$\mathrm{ratio}(w_i) = \frac{P(w_i \mid Pos)}{P(w_i \mid Neg)}$$

LOG LIKELIHOOD

Ratio of Probabilities

Pos ∞

Neutral 1

WORD	Pos	Neg	ratio
I	0.20	0.20	1
am	0.20	0.20	1
happy	0.14	0.10	1.4
because	0.10	0.05	2
learning	0.10	0.10	1
NLP	0.10	0.14	0.6
sad	0.10	0.14	0.6
not	0.10	0.14	0.6

$$\text{ratio}(w_i) = \frac{P(w_i \mid \text{Pos})}{P(w_i \mid \text{Neg})}$$

LOG LIKELIHOOD

Ratio of Probabilities

Pos ∞

Neutral 1

Neg 0

WORD	Pos	Neg	ratio
I	0.20	0.20	1
am	0.20	0.20	1
happy	0.14	0.10	1.4
because	0.10	0.05	2
learning	0.10	0.10	1
NLP	0.10	0.14	0.6
sad	0.10	0.14	0.6
not	0.10	0.14	0.6

$$\text{ratio}(w_i) = \frac{P(w_i \mid \text{Pos})}{P(w_i \mid \text{Neg})}$$

LOG LIKELIHOOD

Naive Bayes' inference

LOG LIKELIHOOD

Naive Bayes' inference

$$\prod_{i=1}^{m} \frac{P(w_i|Pos)}{P(w_i|Neg)} > 1$$

LOG LIKELIHOOD

Naive Bayes' inference

$$\frac{P(Pos)}{P(Neg)} \prod_{i=1}^{m} \frac{P(w_i|Pos)}{P(w_i|Neg)} > 1$$

LOG LIKELIHOOD

- Products bring the risk of underflow
- log(a * b) = log(a) + log(b)

LOG LIKELIHOOD

- Products bring the risk of underflow
- log(a * b) = log(a) + log(b)

$$\frac{P(Pos)}{P(Neg)} \prod_{i=1}^{m} \frac{P(w_i|Pos)}{P(w_i|Neg)}$$

LOG LIKELIHOOD

$$\frac{P(Pos)}{P(Neg)} \prod_{i=1}^{m} \frac{P(w_i|Pos)}{P(w_i|Neg)}$$

- Products bring the risk of underflow
- $\log(a * b) = \log(a) + \log(b)$
- $\log\left(\frac{P(Pos)}{P(Neg)} \prod_{i=1}^{m} \frac{P(w_i|Pos)}{P(w_i|Neg)}\right)$

LOG LIKELIHOOD

- Products bring the risk of underflow $\frac{P(Pos)}{P(Neg)} \prod_{i=1}^{m} \frac{P(w_i|Pos)}{P(w_i|Neg)}$
- log(a * b) = log(a) + log(b)
- $\log\left(\frac{P(Pos)}{P(Neg)} \prod_{i=1}^{m} \frac{P(w_i|Pos)}{P(w_i|Neg)}\right) \Rightarrow \log \frac{P(Pos)}{P(Neg)} + \sum_{i=1}^{n} \log \frac{P(w_i|Pos)}{P(w_i|Neg)}$

LOG LIKELIHOOD

- Products bring the risk of underflow

$$\frac{P(Pos)}{P(Neg)} \prod_{i=1}^{m} \frac{P(w_i|Pos)}{P(w_i|Neg)}$$

- $\log(a * b) = \log(a) + \log(b)$

- $$\log\left(\frac{P(Pos)}{P(Neg)} \prod_{i=1}^{m} \frac{P(w_i|Pos)}{P(w_i|Neg)}\right) \Rightarrow \log \frac{P(Pos)}{P(Neg)} + \sum_{i=1}^{n} \log \frac{P(w_i|Pos)}{P(w_i|Neg)}$$

log prior + log likelihood

LOG LIKELIHOOD

Calculating Lambda

tweet: I am happy because I am learning.

WORD	Pos	Neg
I	0.05	0.05
am	0.04	0.04
happy	0.09	0.01
because	0.01	0.01
learning	0.03	0.01
NLP	0.02	0.02
sad	0.01	0.09
not	0.02	0.03

LOG LIKELIHOOD

Calculating Lambda

tweet: I am happy because I am learning.

$$\lambda(w) = \log \frac{P(w|Pos)}{P(w|Neg)}$$

WORD	Pos	Neg
I	0.05	0.05
am	0.04	0.04
happy	0.09	0.01
because	0.01	0.01
learning	0.03	0.01
NLP	0.02	0.02
sad	0.01	0.09
not	0.02	0.03

LOG LIKELIHOOD

Calculating Lambda

tweet: I am happy because I am learning.

$$\lambda(w) = \log \frac{P(w|Pos)}{P(w|Neg)}$$

$$\lambda(I) = \log \frac{0.05}{0.05} = 0$$

WORD	Pos	Neg	λ
I	0.05	0.05	0
am	0.04	0.04	
happy	0.09	0.01	
because	0.01	0.01	
learning	0.03	0.01	
NLP	0.02	0.02	
sad	0.01	0.09	
not	0.02	0.03	

LOG LIKELIHOOD

Calculating Lambda

tweet: I am happy because I am learning.

$$\lambda(w) = \log \frac{P(w|Pos)}{P(w|Neg)}$$

$$\lambda\,(am) = \log \frac{0.04}{0.04} = 0$$

WORD	Pos	Neg	λ
I	0.05	0.05	0
am	0.04	0.04	0
happy	0.09	0.01	
because	0.01	0.01	
learning	0.03	0.01	
NLP	0.02	0.02	
sad	0.01	0.09	
not	0.02	0.03	

LOG LIKELIHOOD

Calculating Lambda

tweet: I am happy because I am learning.

$$\lambda(w) = \log \frac{P(w|Pos)}{P(w|Neg)}$$

$$\lambda(happy) = \log \frac{0.09}{0.01} = 2.2$$

WORD	Pos	Neg	λ
I	0.05	0.05	0
am	0.04	0.04	0
happy	0.09	0.01	2.2
because	0.01	0.01	
learning	0.03	0.01	
NLP	0.02	0.02	
sad	0.01	0.09	
not	0.02	0.03	

LOG LIKELIHOOD

Calculating Lambda

tweet: I am happy because I am learning.

$$\lambda(w) = \log \frac{P(w|Pos)}{P(w|Neg)}$$

$$\lambda(happy) = \log \frac{0.09}{0.01} = 2.2$$

WORD	Pos	Neg	λ
I	0.05	0.05	0
am	0.04	0.04	0
happy	0.09	0.01	2.2
because	0.01	0.01	
learning	0.03	0.01	
NLP	0.02	0.02	
sad	0.01	0.09	
not	0.02	0.03	

LOG LIKELIHOOD

Calculating Lambda

tweet: I am happy because I am learning.

$$\lambda(w) = \log \frac{P(w|Pos)}{P(w|Neg)}$$

$$\lambda(\text{because}) = \log \frac{0.01}{0.01} = 0$$

WORD	Pos	Neg	λ
I	0.05	0.05	0
am	0.04	0.04	0
happy	0.09	0.01	2.2
because	0.01	0.01	0
learning	0.03	0.01	
NLP	0.02	0.02	
sad	0.01	0.09	
not	0.02	0.03	

LOG LIKELIHOOD

Calculating Lambda

tweet: I am happy because I am learning.

$$\lambda(w) = \log \frac{P(w|Pos)}{P(w|Neg)}$$

$$\lambda(\text{learning}) = \log \frac{0.03}{0.01} = 1.1$$

WORD	Pos	Neg	λ
I	0.05	0.05	0
am	0.04	0.04	0
happy	0.09	0.01	2.2
because	0.01	0.01	0
learning	0.03	0.01	1.1
NLP	0.02	0.02	
sad	0.01	0.09	
not	0.02	0.03	

LOG LIKELIHOOD

Calculating Lambda

tweet: I am happy because I am learning.

$$\lambda(w) = \log \frac{P(w|Pos)}{P(w|Neg)}$$

$$\lambda(NLP) = \log \frac{0.02}{0.02} = 0$$

WORD	Pos	Neg	λ
I	0.05	0.05	0
am	0.04	0.04	0
happy	0.09	0.01	2.2
because	0.01	0.01	0
learning	0.03	0.01	1.1
NLP	0.02	0.02	0
sad	0.01	0.09	
not	0.02	0.03	0

LOG LIKELIHOOD

Calculating Lambda

tweet: I am happy because I am learning.

$$\lambda(w) = \log \frac{P(w|Pos)}{P(w|Neg)}$$

$$\lambda(sad) = \log \frac{0.01}{0.08} = -2.2$$

WORD	Pos	Neg	λ
I	0.05	0.05	0
am	0.04	0.04	0
happy	0.09	0.01	2.2
because	0.01	0.01	0
learning	0.03	0.01	1.1
NLP	0.02	0.02	0
sad	0.01	0.09	-2.2
not	0.02	0.03	0- 0

LOG LIKELIHOOD

Calculating Lambda

tweet: I am happy because I am learning.

$$\lambda(w) = \log \frac{P(w|Pos)}{P(w|Neg)}$$

$$\lambda(\text{not}) = \log \frac{0.02}{0.03} = -0.4$$

WORD	Pos	Neg	λ
I	0.05	0.05	0
am	0.04	0.04	0
happy	0.09	0.01	2.2
because	0.01	0.01	0
learning	0.03	0.01	1.1
NLP	0.02	0.02	0
sad	0.01	0.09	-2.2
not	0.02	0.03	-0.4

LOG LIKELIHOOD

$$\text{Word sentiment} \begin{cases} \text{ratio}(w) = \dfrac{P(w|Pos)}{P(w|Neg)} \\[2em] \lambda(w) = \log \dfrac{P(w|Pos)}{P(w|Neg)} \end{cases}$$

LOG LIKELIHOOD

tweet: I am happy because I am learning.

$$\sum_{i=1}^{n} \log \frac{P(w_i|Pos)}{P(w_i|Neg)} = \sum_{i=1}^{n} \lambda(w_i)$$

WORD	Pos	Neg	λ
I	0.05	0.05	0
am	0.04	0.04	0
happy	0.09	0.01	2.2
because	0.01	0.01	0
learning	0.03	0.01	1.1
NLP	0.02	0.02	0
sad	0.01	0.09	-2.2
not	0.02	0.03	-0.4

total log likelihood =

LOG LIKELIHOOD

tweet: I am happy because I am learning.

$$\sum_{i=1}^{n} \log \frac{P(w_i|Pos)}{P(w_i|Neg)} = \sum_{i=1}^{n} \lambda(w_i)$$

total log likelihood = 0

WORD	Pos	Neg	λ
I	0.05	0.05	0
am	0.04	0.04	0
happy	0.09	0.01	2.2
because	0.01	0.01	0
learning	0.03	0.01	1.1
NLP	0.02	0.02	0
sad	0.01	0.09	-2.2
not	0.02	0.03	-0.4

LOG LIKELIHOOD

tweet: I am happy because I am learning.

$$\sum_{i=1}^{n} \log \frac{P(w_i|Pos)}{P(w_i|Neg)} = \sum_{i=1}^{n} \lambda(w_i)$$

total log likelihood = 0 + 0

WORD	Pos	Neg	λ
I	0.05	0.05	0
am	0.04	0.04	0
happy	0.09	0.01	2.2
because	0.01	0.01	0
learning	0.03	0.01	1.1
NLP	0.02	0.02	0
sad	0.01	0.09	-2.2
not	0.02	0.03	-0.4

LOG LIKELIHOOD

tweet: I am happy because I am learning.

$$\sum_{i=1}^{n} \log \frac{P(w_i|Pos)}{P(w_i|Neg)} = \sum_{i=1}^{n} \lambda(w_i)$$

total log likelihood = 0 + 0 + 2.2

WORD	Pos	Neg	λ
I	0.05	0.05	0
am	0.04	0.04	0
happy	0.09	0.01	2.2
because	0.01	0.01	0
learning	0.03	0.01	1.1
NLP	0.02	0.02	0
sad	0.01	0.09	-2.2
not	0.02	0.03	-0.4

LOG LIKELIHOOD

tweet: I am happy because I am learning.

$$\sum_{i=1}^{n} \log \frac{P(w_i|Pos)}{P(w_i|Neg)} = \sum_{i=1}^{n} \lambda(w_i)$$

total log likelihood = 0 + 0 + 2.2 + 0 + 0 + 0

WORD	Pos	Neg	λ
I	0.05	0.05	0
am	0.04	0.04	0
happy	0.09	0.01	2.2
because	0.01	0.01	0
learning	0.03	0.01	1.1
NLP	0.02	0.02	0
sad	0.01	0.09	-2.2
not	0.02	0.03	-0.4

LOG LIKELIHOOD

tweet: I am happy because I am learning.

$$\sum_{i=1}^{n} \log \frac{P(w_i|Pos)}{P(w_i|Neg)} = \sum_{i=1}^{n} \lambda(w_i)$$

WORD	Pos	Neg	λ
I	0.05	0.05	0
am	0.04	0.04	0
happy	0.09	0.01	2.2
because	0.01	0.01	0
learning	0.03	0.01	1.1
NLP	0.02	0.02	0
sad	0.01	0.09	-2.2
not	0.02	0.03	-0.4

total log likelihood = 0 + 0 + 2.2 + 0 + 0 + 0 + 1.1

LOG LIKELIHOOD

tweet: I am happy because I am learning.

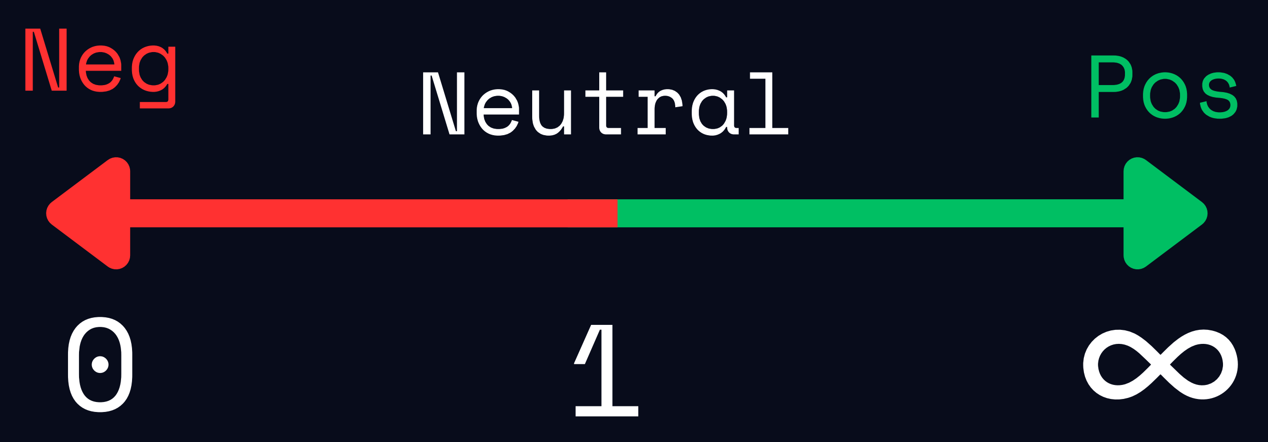
$$\sum_{i=1}^{n} \log \frac{P(w_i|Pos)}{P(w_i|Neg)} = \sum_{i=1}^{n} \lambda(w_i)$$

WORD	Pos	Neg	λ
I	0.05	0.05	0
am	0.04	0.04	0
happy	0.09	0.01	2.2
because	0.01	0.01	0
learning	0.03	0.01	1.1
NLP	0.02	0.02	0
sad	0.01	0.09	-2.2
not	0.02	0.03	-0.4

total log likelihood = 0 + 0 + 2.2 + 0 + 0 + 0 + 1.1
= 3.3

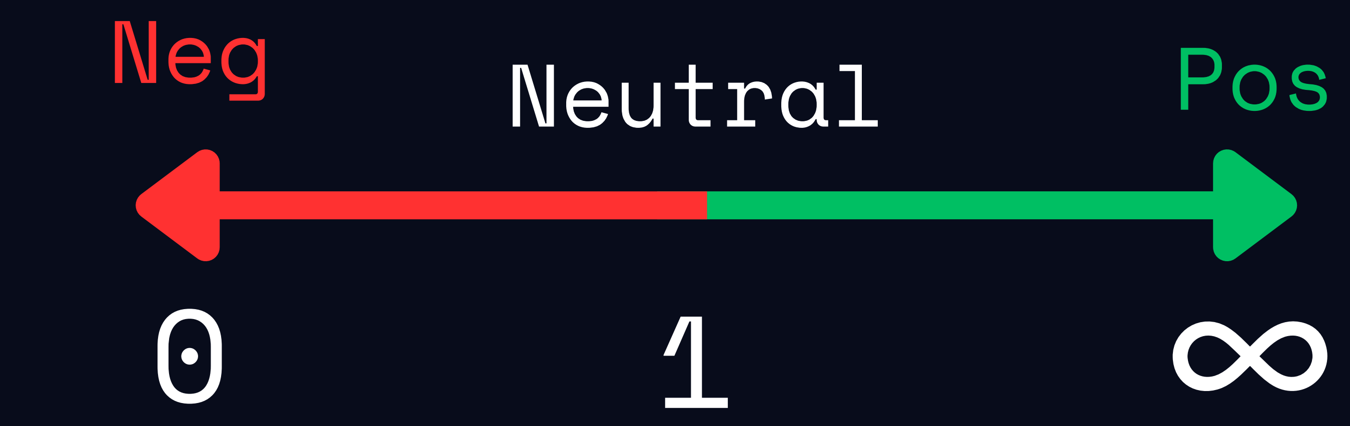
LOG LIKELIHOOD

$$\prod_{i=1}^{m} \frac{P(w_i|Pos)}{P(w_i|Neg)} > 1$$

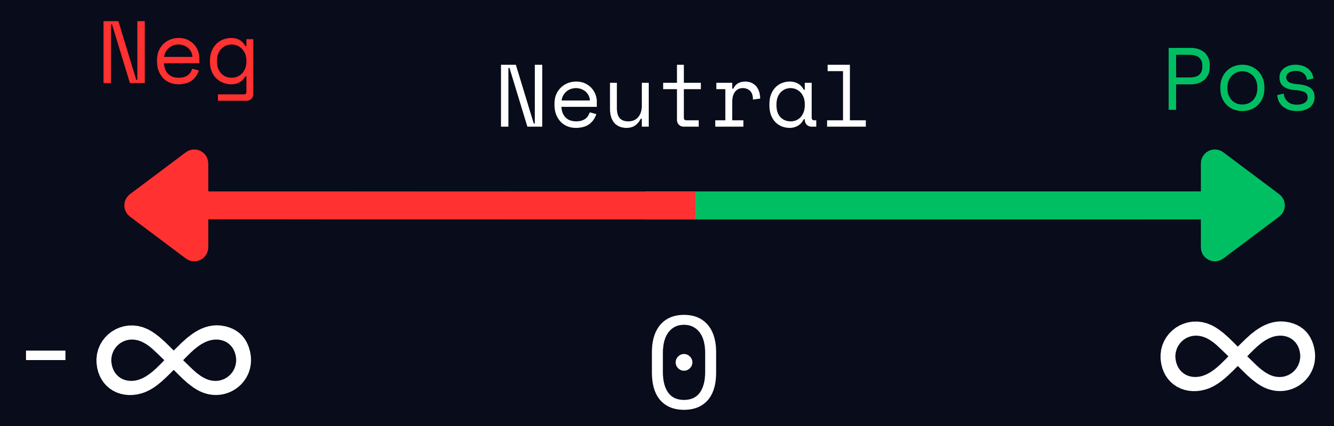


LOG LIKELIHOOD

$$\prod_{i=1}^{m} \frac{P(w_i|Pos)}{P(w_i|Neg)} > 1$$

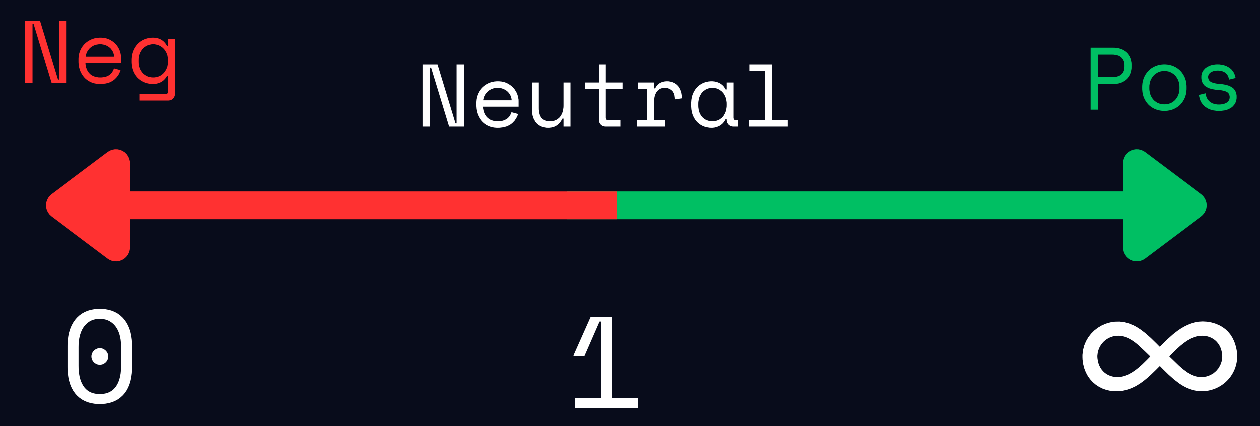


$$\sum_{i=1}^{n} \log \frac{P(w_i|Pos)}{P(w_i|Neg)} > 0$$



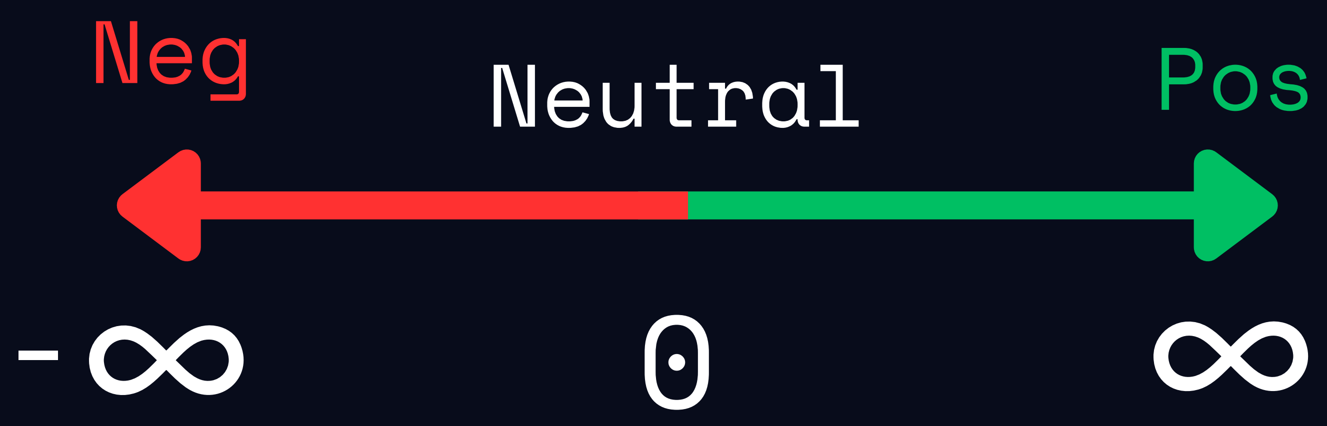
LOG **LIKELIHOOD**

$$\prod_{i=1}^{m} \frac{P(w_i|Pos)}{P(w_i|Neg)} > 1$$



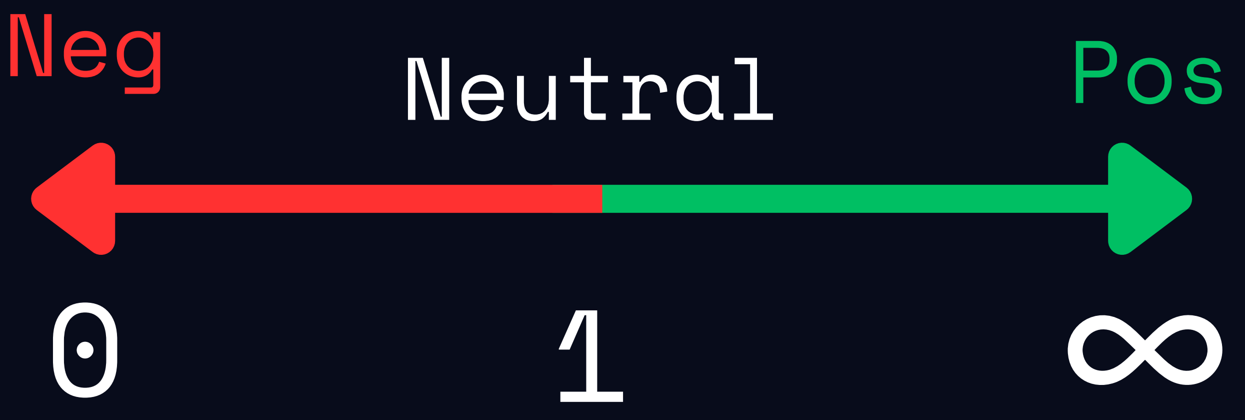
$$\sum_{i=1}^{n} \log \frac{P(w_i|Pos)}{P(w_i|Neg)} > 0$$

$3.3 > 0$



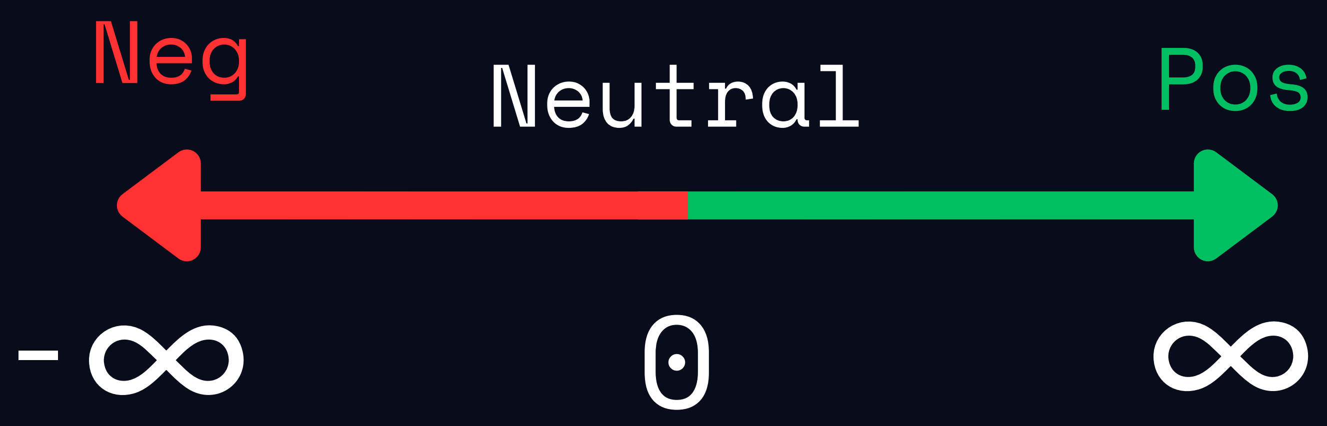
LOG LIKELIHOOD

$$\prod_{i=1}^{m} \frac{P(w_i|Pos)}{P(w_i|Neg)} > 1$$



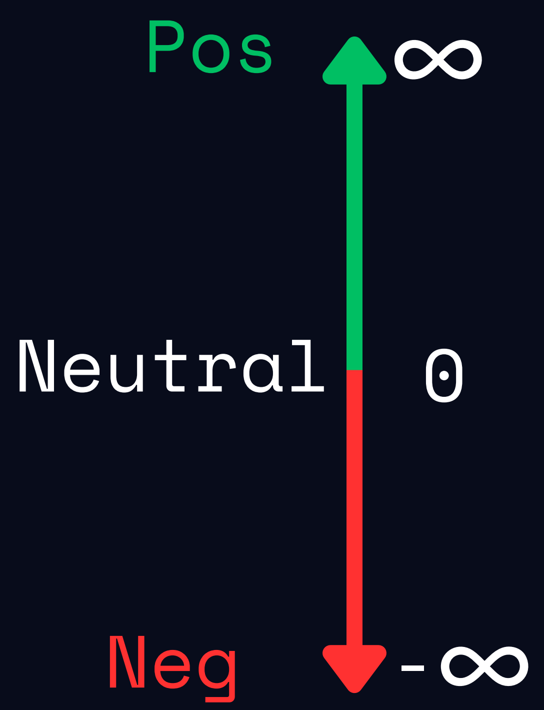
$$3.3 > 0$$

$$\sum_{i=1}^{n} \log \frac{P(w_i|Pos)}{P(w_i|Neg)} > 0$$



LOG LIKELIHOOD

$$\log \prod_{i=1}^{m} \text{ratio}(w_i) = \sum_{i=1}^{n} \lambda(w_i) > 0$$



TRAINING NAIVE BAYES

Step 0: Collect and annotate corpus

Positive tweets

I am happy because I am learning NLP

I am happy, not sad

Negative tweets

I am sad I am not Learning NLP

I am sad, not happy

TRAINING NAIVE BAYES

Step 0: Collect and annotate corpus

Positive tweets

I am happy because I am learning NLP

I am happy, not sad

Negative tweets

I am sad I am not Learning NLP

I am sad, not happy

Step 1: Preprocess

TRAINING NAIVE BAYES

Step 0: Collect and annotate corpus

Positive tweets

I am happy because I am learning NLP

I am happy, not sad

Negative tweets

I am sad I am not Learning NLP

I am sad, not happy

Step 1: Preprocess

- Lowercase
- Remove punctuation, URLs, names
- Remove stop words
- Stemming
- Tokenize sentences

TRAINING NAIVE BAYES

Step 0: Collect and annotate corpus

- Lowercase
- Remove punctuation, URLs, names
- Remove stop words
- Stemming
- Tokenize sentences

Positive tweets

I am happy because I am learning NLP

I am happy, not sad

Negative tweets

I am sad I am not Learning NLP

I am sad, not happy

Step 1: Preprocess

Positive tweets

[happi, because, learn, NLP]

[happi, not, sad]

Negative tweets

[sad, not, learn, NLP]

[sad, not, happi]

TRAINING NAIVE BAYES

Positive tweets

[happi, because, learn, NLP]

[happi, not, sad]

Negative tweets

[sad, not, learn, NLP]

[sad, not, happi]

TRAINING NAIVE BAYES

freq(w, class)

Positive tweets

[happi, because, learn, NLP]

[happi, not, sad]

Negative tweets

[sad, not, learn, NLP]

[sad, not, happi]

TRAINING NAIVE BAYES

Positive tweets

[happi, because, learn, NLP]

[happi, not, sad]

Negative tweets

[sad, not, learn, NLP]

[sad, not, happi]

Step 2: Word coun

freq(w, class)

WORD	Pos	Neg
happi	2	1
beacuse	1	0
learn	1	1
NLP	1	1
sad	1	2
not	1	2
N_{class}	7	7

TRAINING NAIVE BAYES

freq(w, class)

WORD	Pos	Neg
happi	2	1
beacuse	1	0
learn	1	1
NLP	1	1
sad	1	2
not	1	2
N_{class}	7	7

TRAINING NAIVE BAYES

freq(w, class)

WORD	Pos	Neg
happi	2	1
beacuse	1	0
learn	1	1
NLP	1	1
sad	1	2
not	1	2
N_{class}	7	7

Step 3:
$P(w|class)$

$$V_{class} = 6$$

$$\frac{freq(w_i, class) + 1}{N_{class} + V_{class}}$$

TRAINING NAIVE BAYES

freq(w, class)

WORD	Pos	Neg
happi	2	1
beacuse	1	0
learn	1	1
NLP	1	1
sad	1	2
not	1	2
N_{class}	7	7

Step 3:
$P(w|class)$

$V_{class} = 6$

$$\frac{freq(w_i, class) + 1}{N_{class} + V_{class}}$$

WORD	Pos	Neg
happi	0.23	0.15
beacuse	0.15	0.07
learn	0.08	0.08
NLP	0.08	0.08
sad	0.08	0.17
not	0.08	0.17

TRAINING NAIVE BAYES

$freq(w, class)$

WORD	Pos	Neg
happi	2	1
beacuse	1	0
learn	1	1
NLP	1	1
sad	1	2
not	1	2
N_{class}	7	7

Step 3: $P(w|class)$

$V_{class} = 6$

$$\frac{freq(w_i, class) + 1}{N_{class} + V_{class}}$$

$$\lambda(w) = \log \frac{P(w|Pos)}{P(w|Neg)}$$

WORD	Pos	Neg
happi	0.23	0.15
beacuse	0.15	0.07
learn	0.08	0.08
NLP	0.08	0.08
sad	0.08	0.17
not	0.08	0.17

TRAINING NAIVE BAYES

freq(w, class)

WORD	Pos	Neg
happi	2	1
beacuse	1	0
learn	1	1
NLP	1	1
sad	1	2
not	1	2
N_{class}	7	7

Step 3:
P(w|class)

$V_{class} = 6$

$$\frac{freq(w_i, class) + 1}{N_{class} + V_{class}}$$

$$\lambda(w) = \log \frac{P(w|Pos)}{P(w|Neg)}$$

Step 4:
Get lambda

WORD	Pos	Neg
happi	0.23	0.15
beacuse	0.15	0.07
learn	0.08	0.08
NLP	0.08	0.08
sad	0.08	0.17
not	0.08	0.17

TRAINING NAIVE BAYES

freq(w, class)

WORD	Pos	Neg
happi	2	1
beacuse	1	0
learn	1	1
NLP	1	1
sad	1	2
not	1	2
N_{class}	7	7

Step 3:
P(w|class)

$$V_{class} = 6$$

$$\frac{freq(w_i, class) + 1}{N_{class} + V_{class}}$$

$$\lambda(w) = \log \frac{P(w|Pos)}{P(w|Neg)}$$

Step 4:
Get lambda

WORD	Pos	Neg	λ
happi	0.23	0.15	0.43
beacuse	0.15	0.07	0.6
learn	0.08	0.08	0
NLP	0.08	0.08	0
sad	0.08	0.17	-0.75
not	0.08	0.17	-0.75

TRAINING NAIVE BAYES

Step 5:
Get the
log prior

TRAINING NAIVE BAYES

D_{pos} = Number of positive tweets

D_{neg} = Number of negative tweets

Step 5:
Get the
log prior

TRAINING NAIVE BAYES

D_{pos} = Number of positive tweets

D_{neg} = Number of negative tweets

Step 5:
Get the
log prior

$$logprior = \log \frac{D_{pos}}{D_{neg}}$$

TRAINING NAIVE BAYES.

D_{pos} = Number of positive tweets

D_{neg} = Number of negative tweets

Step 5:
Get the
log prior

$$\text{logprior} = \log \frac{D_{pos}}{D_{neg}}$$

If dataset is balanced, $D_{pos} = D_{neg}$ and logprior = 0

TRAINING NAIVE BAYES

0. Get or annotate a dataset with positive and negative tweets
1. Preprocess the tweets: process_tweet(tweet)-->[W1, W2, W3,..]
2. Compute freq(w,class)
3. Get P(w | pos), P(w | neg)
4. Get λ(w)
5. Compute logprior = log(P(pos) / P(neg))

TESTING NAIVE BAYES

Outline

- Predicting using a Naive Bayes Model

TESTING NAIVE BAYES

Predicting using a Naive Bayes Model

- Log-liklihood dictionary $\lambda(w) = \log \frac{P(w|Pos)}{P(w|Neg)}$

WORD	λ
happi	0.43
beacuse	0.6
learn	0
NLP	0
sad	-0.75
not	-0.75

TESTING NAIVE BAYES

Predicting using a Naive Bayes Model

- Log-liklihood dictionary $\lambda(w) = \log \frac{P(w|Pos)}{P(w|Neg)}$
- logprior $= \log \frac{D_{pos}}{D_{neg}} = 0$

WORD	λ
happi	0.43
beacuse	0.6
learn	0
NLP	0
sad	-0.75
not	-0.75

TESTING NAIVE BAYES

Predicting using a Naive Bayes Model

- Log-liklihood dictionary $\lambda(w) = \log \frac{P(w|Pos)}{P(w|Neg)}$
- logprior = $\log \frac{D_{pos}}{D_{neg}} = 0$
- Tweet: **I am happy because I am learning NLP**

WORD	λ
happi	0.43
beacuse	0.6
learn	0
NLP	0
sad	-0.75
not	-0.75

TESTING NAIVE BAYES

Predicting using a Naive Bayes Model

- Log-liklihood dictionary $\lambda(w) = \log \frac{P(w|Pos)}{P(w|Neg)}$
- logprior = $\log \frac{D_{pos}}{D_{neg}} = 0$
- Tweet: **[happi, because, learn, NLP]**

WORD	λ
happi	0.43
because	0.6
learn	0
NLP	0
sad	-0.75
not	-0.75

TESTING NAIVE BAYES

Outline

- Log-liklihood dictionary $\lambda(w) = \log \frac{P(w|Pos)}{P(w|Neg)}$
- logprior $= \log \frac{D_{pos}}{D_{neg}} = 0$
- Tweet: **[happi, because, learn, NLP]**

score = 0.43 + 0.6 + 0 + 0 + logprior = 1.03

WORD	λ
happi	0.43
because	0.6
learn	0
NLP	0
sad	-0.75
not	-0.75

TESTING NAIVE BAYES

Predicting using a Naive Bayes Model

- Log-liklihood dictionary $\lambda(w) = \log \frac{P(w|Pos)}{P(w|Neg)}$
- $\text{logprior} = \log \frac{D_{pos}}{D_{neg}} = 0$
- Tweet: **[happi, because, learn, NLP]**

$$\textit{score} = 0.43 + 0.6 + 0 + 0 + \text{logprior} = 1.03$$

$$\textit{pred} = \textit{score} > 0$$

WORD	λ
happi	0.43
because	0.6
learn	0
NLP	0
sad	-0.75
not	-0.75

TESTING NAIVE BAYES

Predicting using a Naive Bayes Model

- Log-liklihood dictionary $\lambda(w) = \log \frac{P(w|Pos)}{P(w|Neg)}$
- $\text{logprior} = \log \frac{D_{pos}}{D_{neg}} = 0$
- Tweet: **[happi, because, learn, NLP]**

WORD	λ
happi	0.43
because	0.6
learn	0
NLP	0
sad	-0.75
not	-0.75

$$score = 0.43 + 0.6 + 0 + 0 + \text{logprior} = 1.03$$

$$pred = score > 0$$

Karthik @iamnkarthik · 08 Feb

Even though there are so many workshops in this world, why did I participate in this workshop?

1 2 5 420