

Ex-Twit: Explainable Twitter Mining on Health Data

Tunazzina Islam

Ph.D. Student

Department of Computer Science

Purdue University, West Lafayette



islam32@purdue.edu



<https://tunazislam.github.io/>



[@Tunaz_Islam](https://twitter.com/Tunaz_Islam)

Social NLP 2019@IJCAI-19, Macao, China

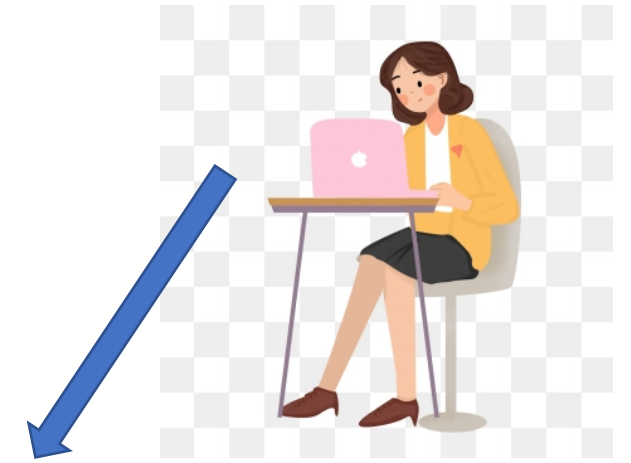
Date: August 12, 2019



Motivation



Balanced diet



@vuthihoangquye1: RT @go1click: Ketogenic Diet The truth:> buff.ly/2NQr4jY
#health #fitness #diet #healthy #fitness #weightloss #exercise #workout #sport
#paleo #yoga #food #nutrition #fat #cbd #keto #wellness #news #ff #inspiration



Exercise



Running



Yoga



Motivation

Tweet 1: *Swimming is great. It's a perfect workout. #fitness #wellness*

Motivation

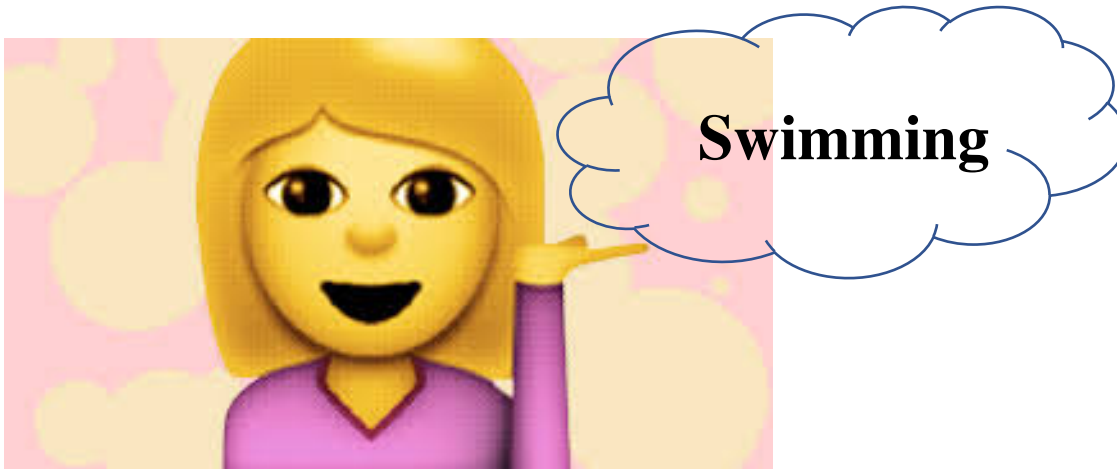
Tweet: *Swimming is great. It's a perfect workout. #fitness #wellness*



What is the topic of this tweet?

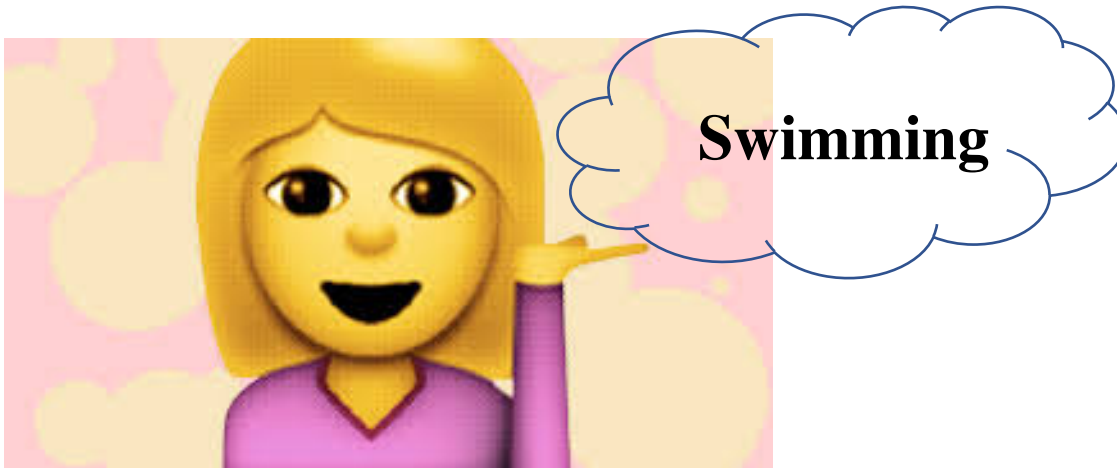
Motivation

Tweet: *Swimming is great. It's a perfect workout. #fitness #wellness*



Motivation

Tweet: *Swimming is great. It's a perfect workout. #fitness #wellness*



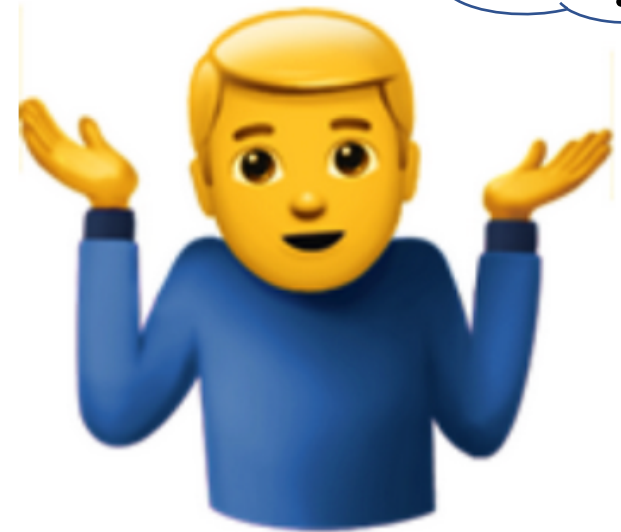
Human prediction

Motivation

Tweet: *Swimming is great. It's a perfect workout. #fitness #wellness*



Swimming

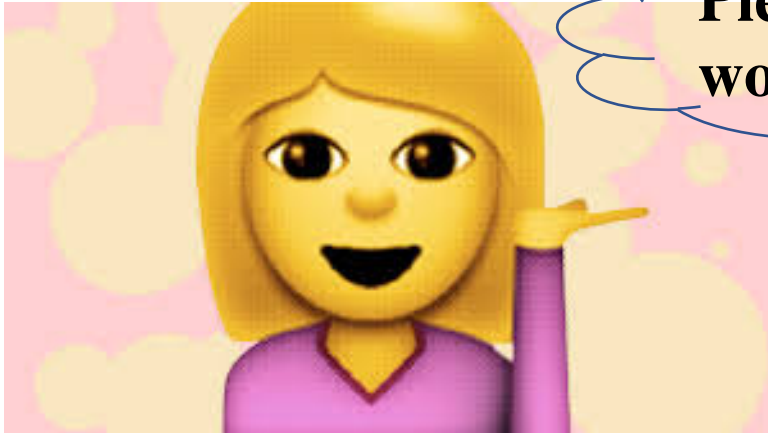


Why?

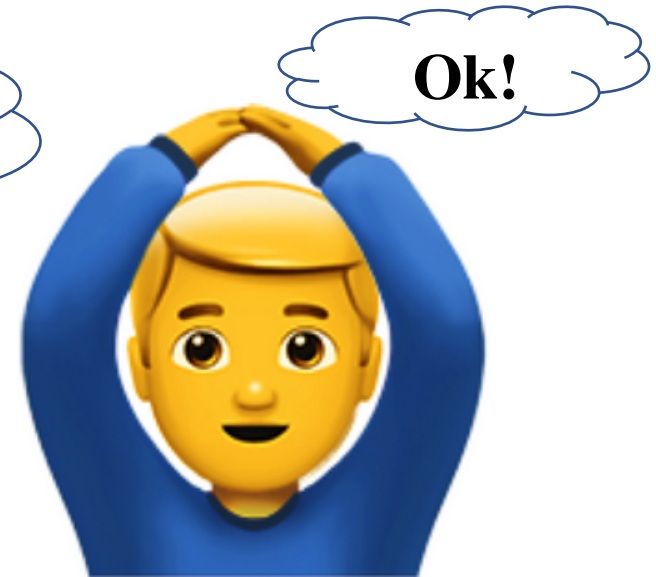
Human prediction

Motivation

Tweet: Swimming is great. It's a perfect workout. #fitness #wellness



**Please look at the first
word of the tweet.**

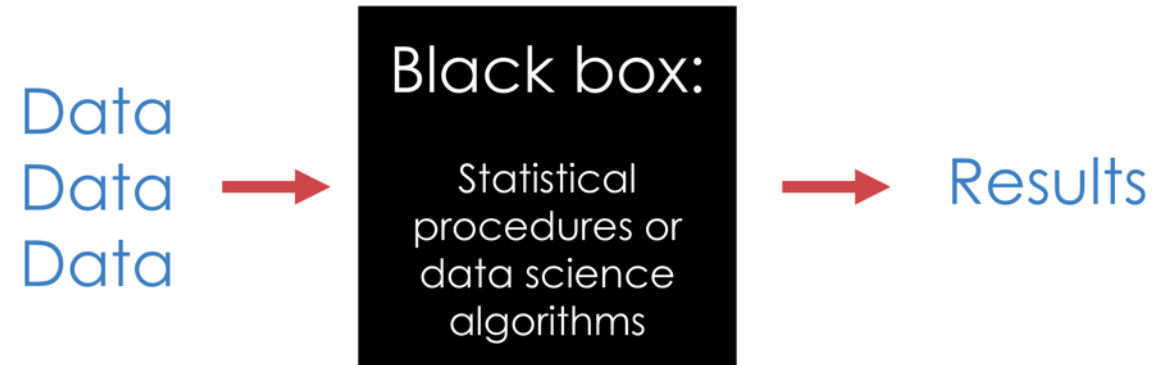


Human Explanation

Motivation

Tweet: *Swimming is great. It's a perfect workout. #fitness #wellness*

What is the prediction by ML algorithm?



Motivation

Tweet: *Swimming is great. It's a perfect workout. #fitness #wellness*

What is the prediction by ML algorithm?

Fitness and wellness

Data
Data
Data



Black box:

Statistical
procedures or
data science
algorithms



Results

Motivation

Tweet: *Swimming is great. It's a perfect workout. #fitness #wellness*

What is the prediction by ML algorithm?

Fitness and wellness

Swimming

Data
Data
Data



Black box:

Statistical
procedures or
data science
algorithms



Results

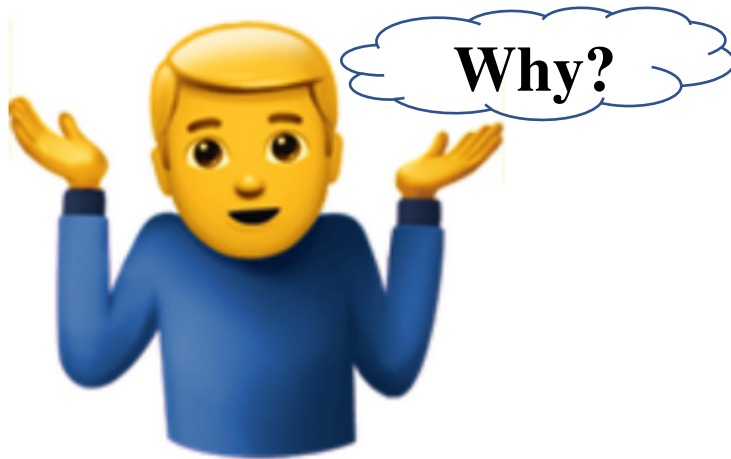
Motivation

Tweet: *Swimming is great. It's a perfect workout. #fitness #wellness*

What is the prediction by ML algorithm?

Fitness and wellness

Swimming



Data
Data
Data



Black box:

Statistical
procedures or
data science
algorithms



Results

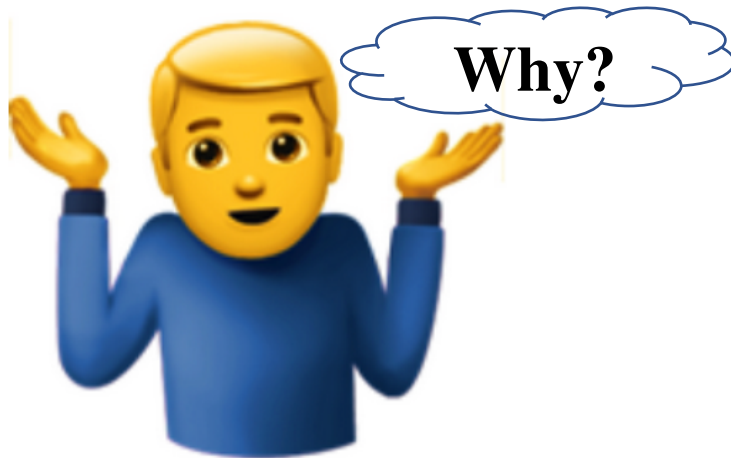
Motivation

Tweet: *Swimming is great. It's a perfect workout. #fitness #wellness*

What is the prediction by ML algorithm?

Fitness and wellness

Swimming



Data
Data
Data

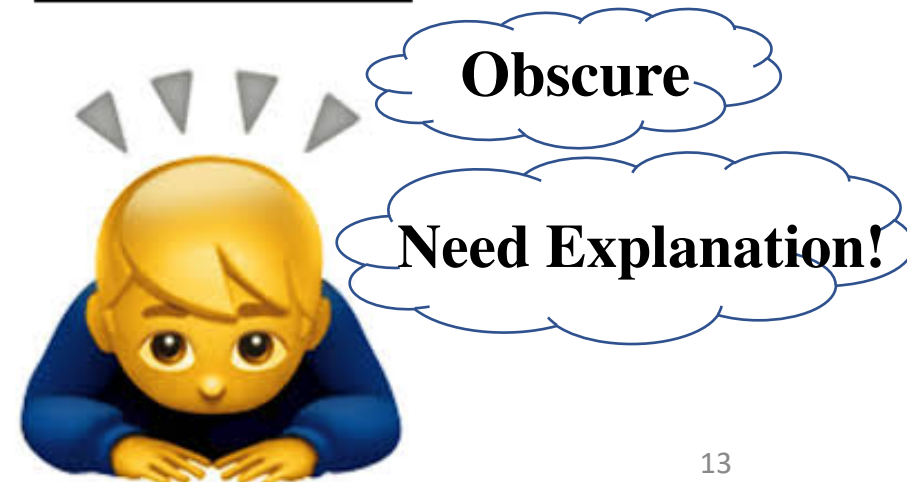


Black box:

Statistical
procedures or
data science
algorithms



Results



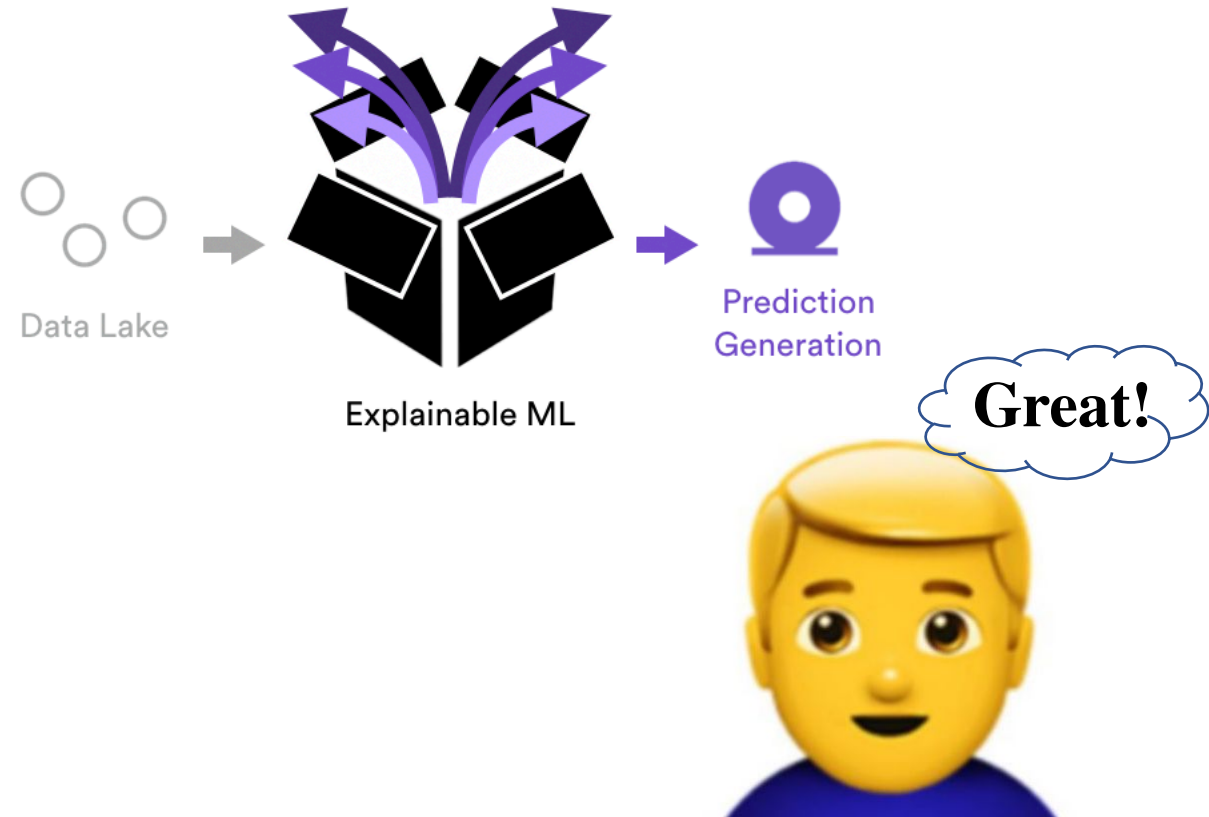
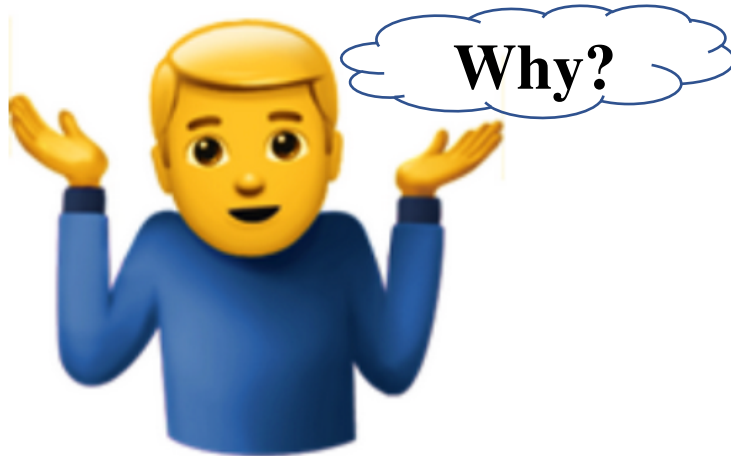
Motivation

Tweet: *Swimming is great. It's a perfect workout. #fitness #wellness*

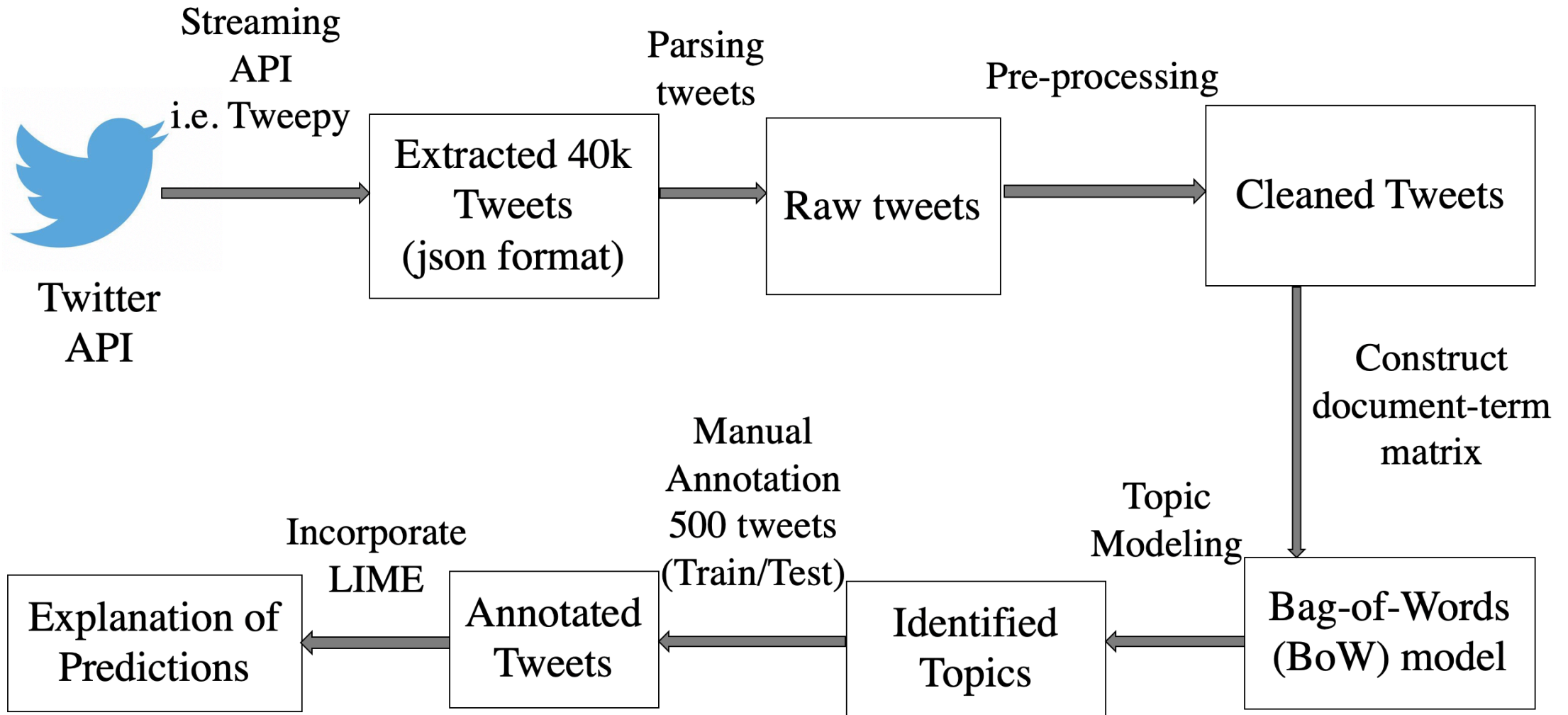
What is the prediction by ML algorithm?

Fitness and wellness

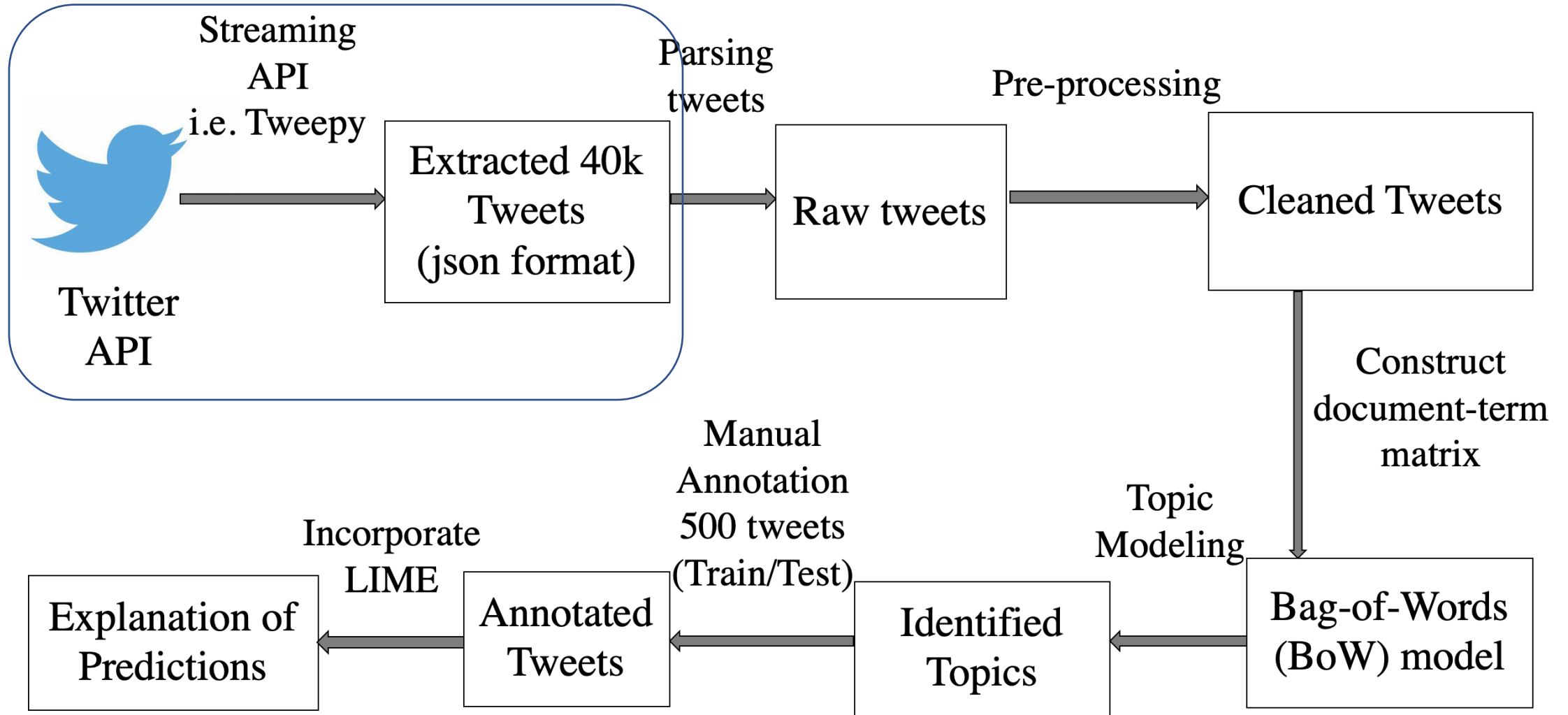
Swimming



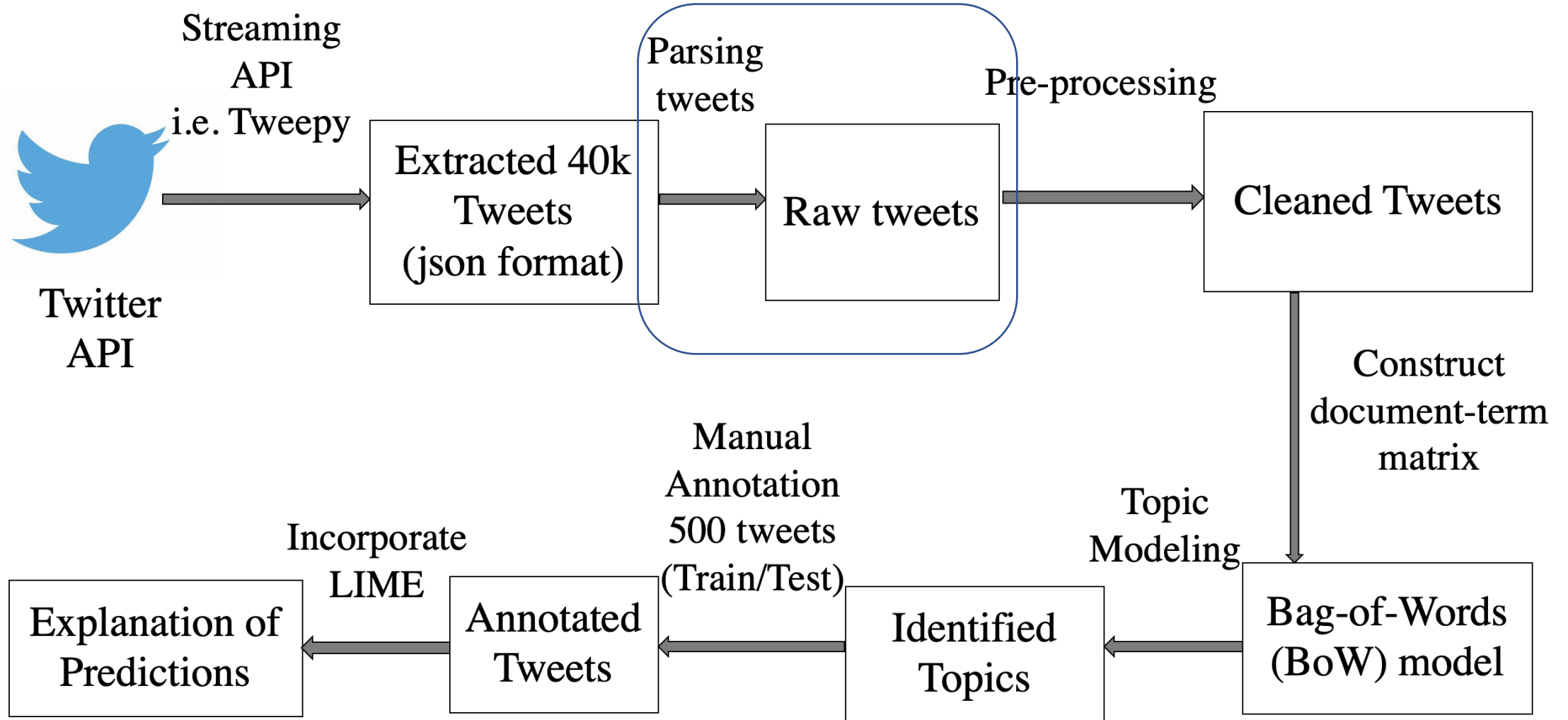
Methodology of Ex-Twit



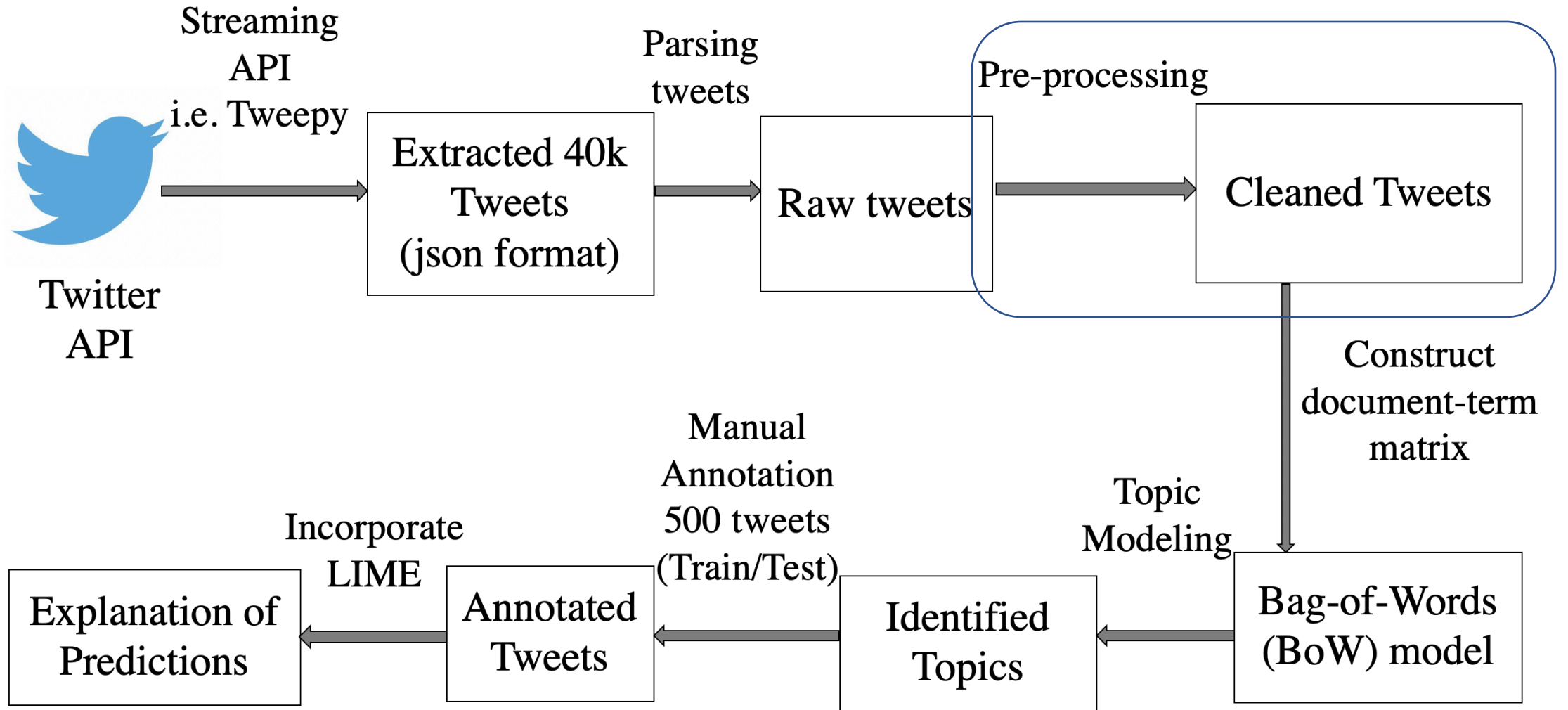
Methodology of Ex-Twit



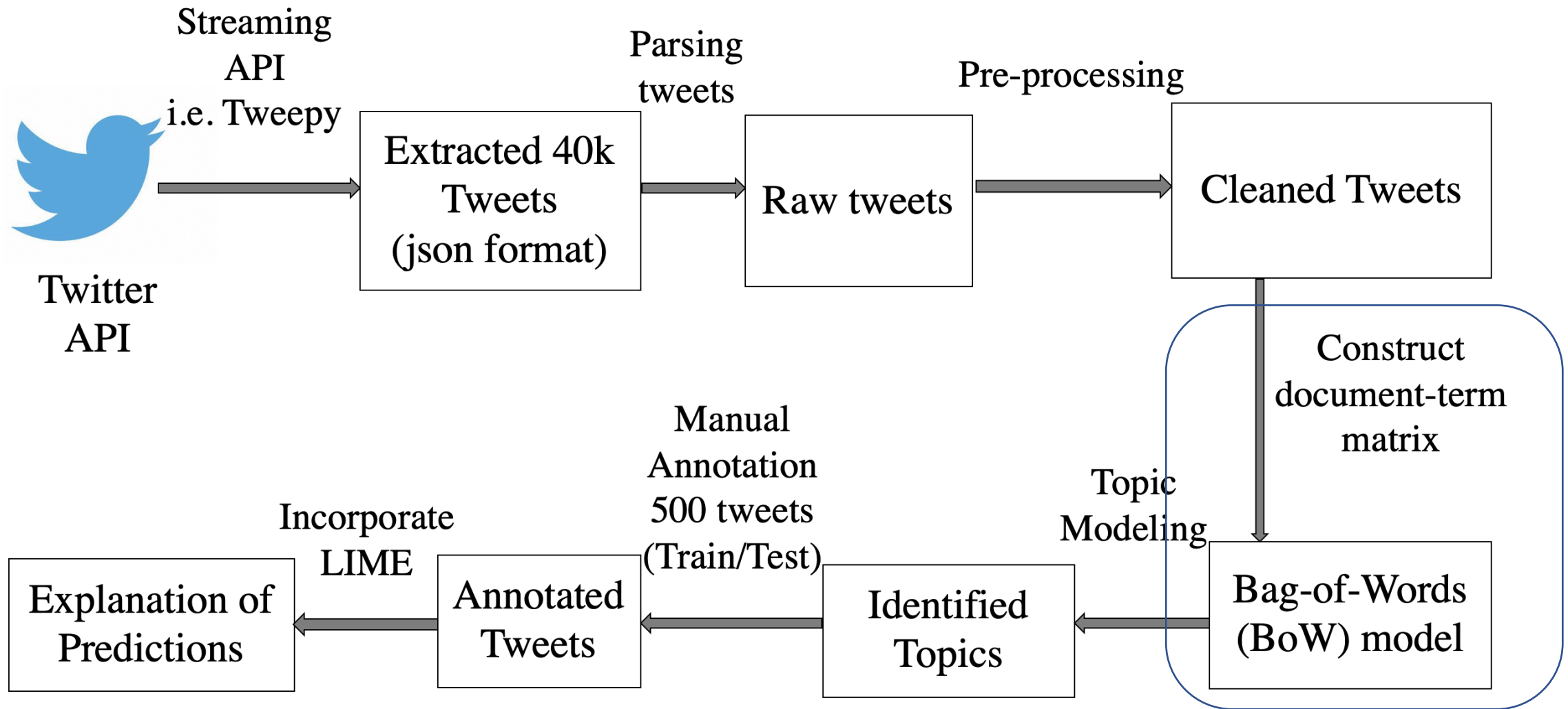
Methodology of Ex-Twit



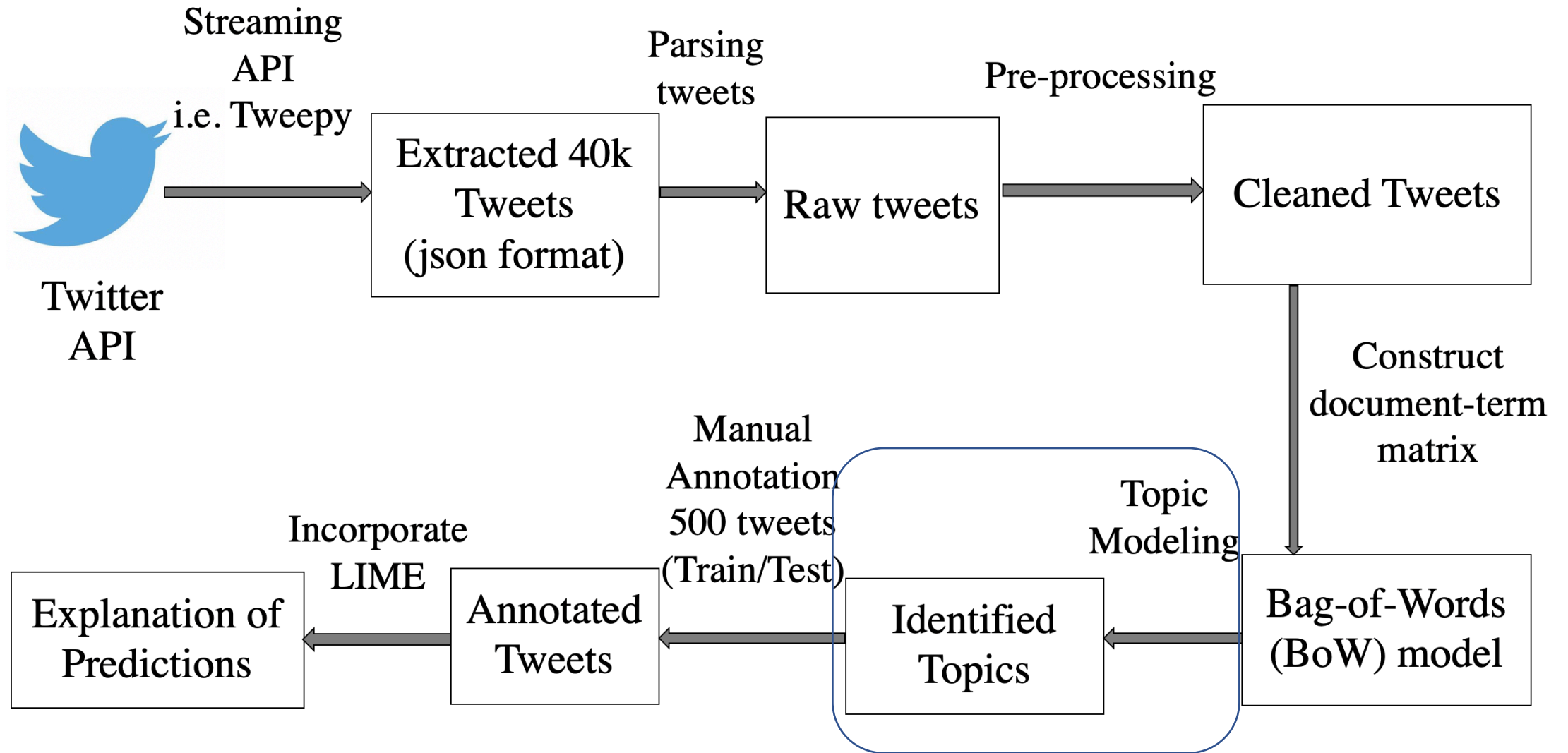
Methodology of Ex-Twit



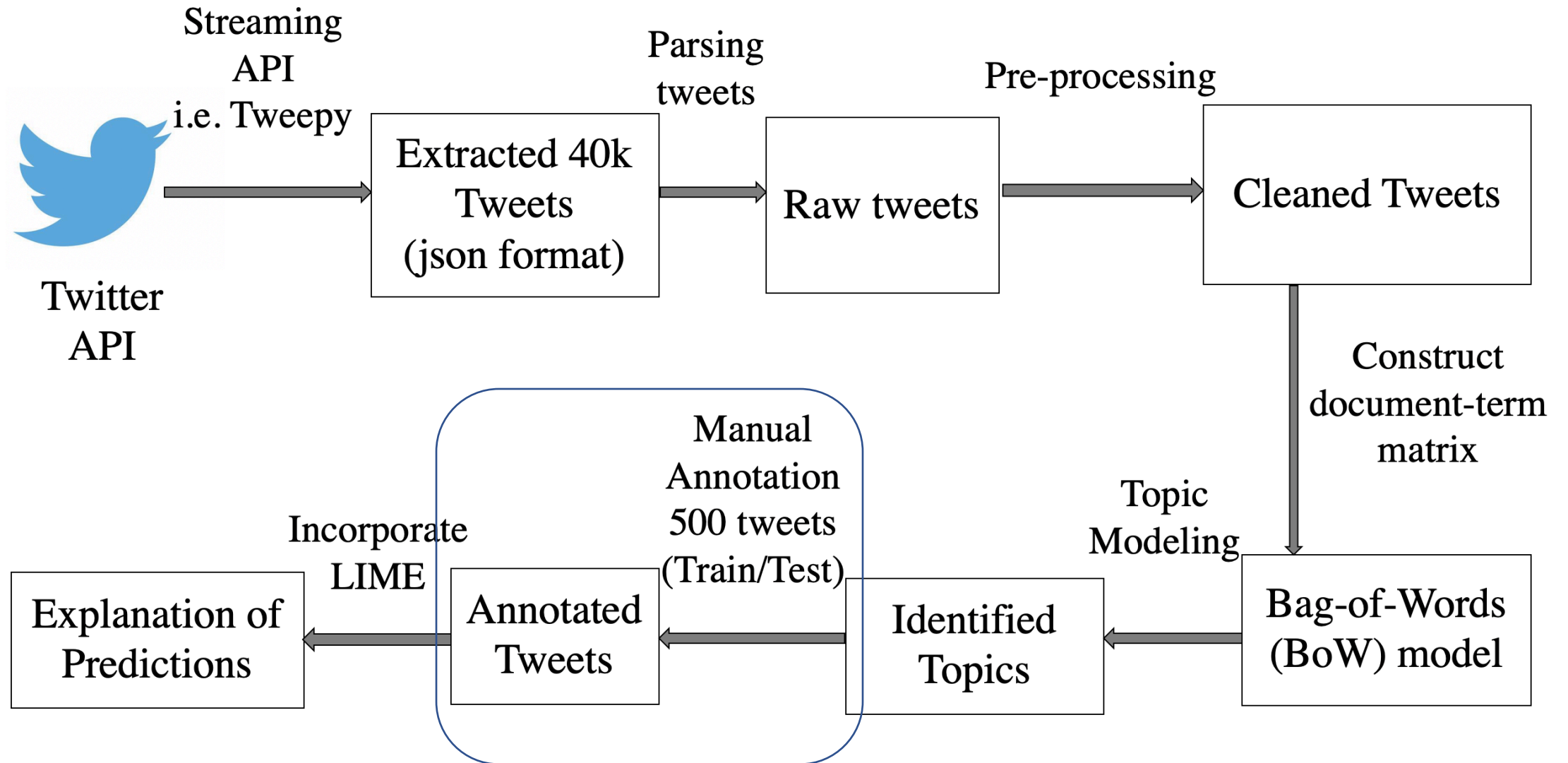
Methodology of Ex-Twit



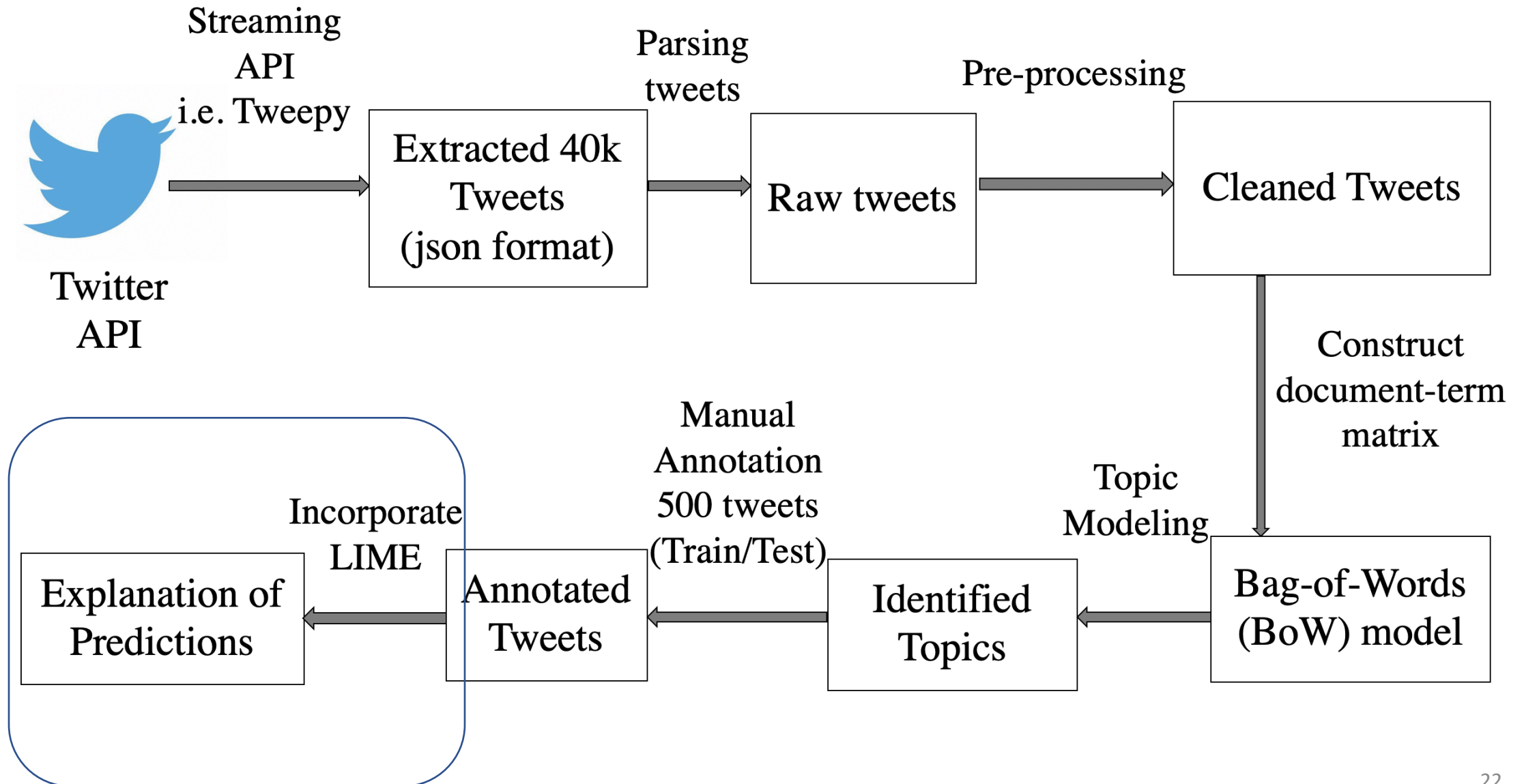
Methodology of Ex-Twit



Methodology of Ex-Twit



Methodology of Ex-Twit



LIME

- Generates distorted versions of the tweets.
- Predicts probabilities for these distorted tweets using the black-box classifier.
- Trains another classifier to predict output of a black-box classifier on the original tweets.
- LIME Evaluation:
 - Accuracy score.
 - KL divergence.

LIME Evaluation- Accuracy score

- Cosine distance between generated sample and original tweets.
- Given two vectors of attributes, A and B, the cosine similarity, $\cos(\theta)$ is represented as following:

$$\begin{aligned}\text{similarity} = \cos(\theta) &= \frac{A \cdot B}{||A|| ||B||} \\ &= \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}}\end{aligned}$$

where A_i and B_i are components of vector A and B respectively.

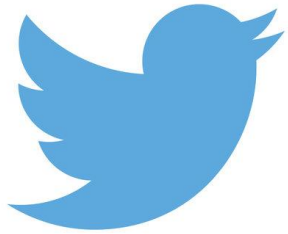
LIME Evaluation- KL divergence

- Measures the difference between two probability distributions.
- For distributions of P and Q of a continuous random variable, the KL divergence is defined as following:

$$D_{KL} (P||Q) = \int_{-\infty}^{\infty} p(x) \log \left(\frac{p(x)}{q(x)} \right) dx$$

- Weighted by distance.
- KL divergence 0.0 means a perfect match.

Topic Modeling

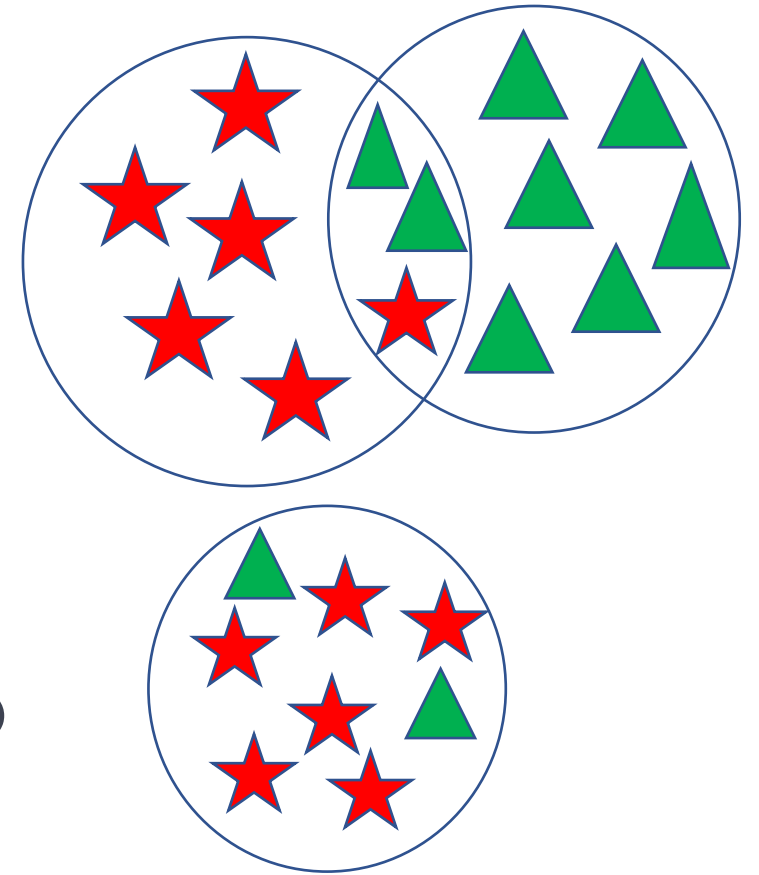


Tweets



Topic Modeling

- Latent Semantic Analysis (LSA)
- Non-negative Matrix Factorization (NMF)
- Latent Dirichlet Allocation (LDA)

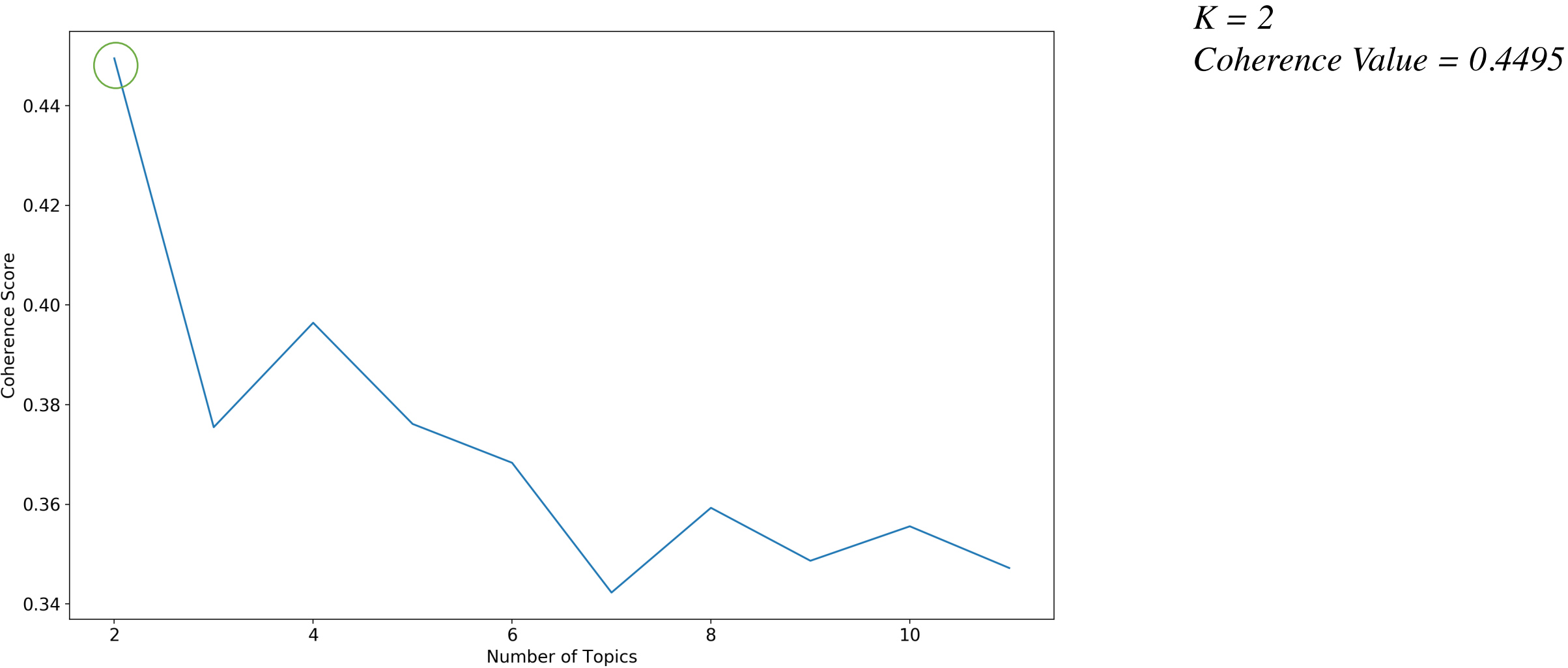


Identified Topics

How to choose optimal Number of Topics?

- Build many LSA, LDA, NMF models with different values of number of topics (k).
- pick k with highest coherence value.

Optimal Number of Topics vs Coherence Score LSA



Topics using LSA

Topic1

yoga
everi
life
job
remember
goe
woman
everyone
cook
therapy

Topic2

diet
vegan
fit
day
new
like
beyonce
amp
eat
workout

Topics using LSA

Topic1

yoga

everi

life

job

remember

goe

woman

everyone

cook

therapy

Topic2

diet

vegan

fit

day

new

like

beyonce

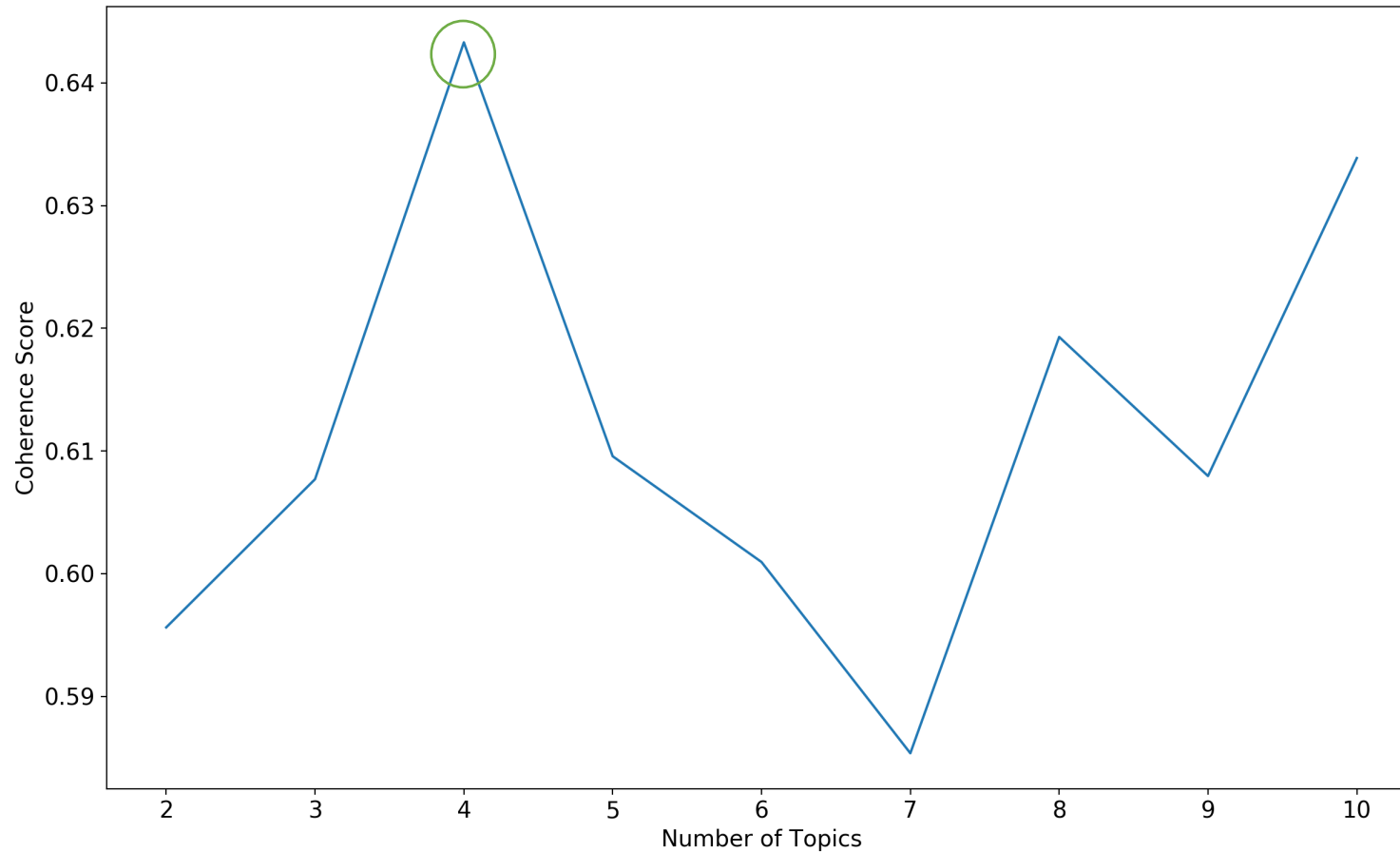
amp

eat

workout

- *unable to capture the meanings of words.*
- *lower accuracy*

Optimal Number of Topics vs Coherence Score NMF



$K = 4$

Coherence Value = 0.6404

Topic coherence measure TC-W2V

Topics using NMF

Topic1

Yoga
job
every_woman
cooks_goe
therapy_remember
life_juggl
everyone_birthday
boyfriend
hot
know

Topic2

diet
beyonce
new
bitch
ciara_prayer
day
eat
go
fat
keto

Topic3

vegan
go
eat
make
food
day
amp
shit
meat
vegetarian

Topic4

fitness
workout
go
good
amp
day
yoga
health
gym
today

Topics using NMF

- *same keywords are repeated in multiple topics.*

Topic1

Yoga
job
every_woman
cooks_goe
therapy_remember
life_juggl
everyone_birthday
boyfriend
hot
know

Topic2

diet
beyonce
new
bitch
ciara_prayer
day
eat
go
fat
keto

Topic3

vegan
go
eat
make
food
day
amp
shit
meat
vegetarian

Topic4

fitness
workout
go
good
amp
day
yoga
health
gym
today

Topics using NMF

- *same keywords are repeated in multiple topics.*

Topic1

Yoga

job

every_woman

cooks_goe

therapy_remember

life_juggl

everyone_birthday

boyfriend

hot

know

Topic2

diet

beyonce

new

bitch

ciara_prayer

day

eat

go

fat

keto

Topic3

vegan

go

eat

make

food

day

amp

shit

meat

vegetarian

Topic4

fitness

workout

go

good

amp

day

yoga

health

gym

today

Topics using NMF

- *same keywords are repeated in multiple topics.*

Topic1

Yoga

job

every_woman

cooks_goe

therapy_remember

life_juggl

everyone_birthday

boyfriend

hot

know

Topic2

diet

beyonce

new

bitch

ciara_prayer

day

eat

go

fat

keto

Topic3

vegan

go

eat

make

food

day

amp

shit

meat

vegetarian

Topic4

fitness

workout

go

good

amp

day

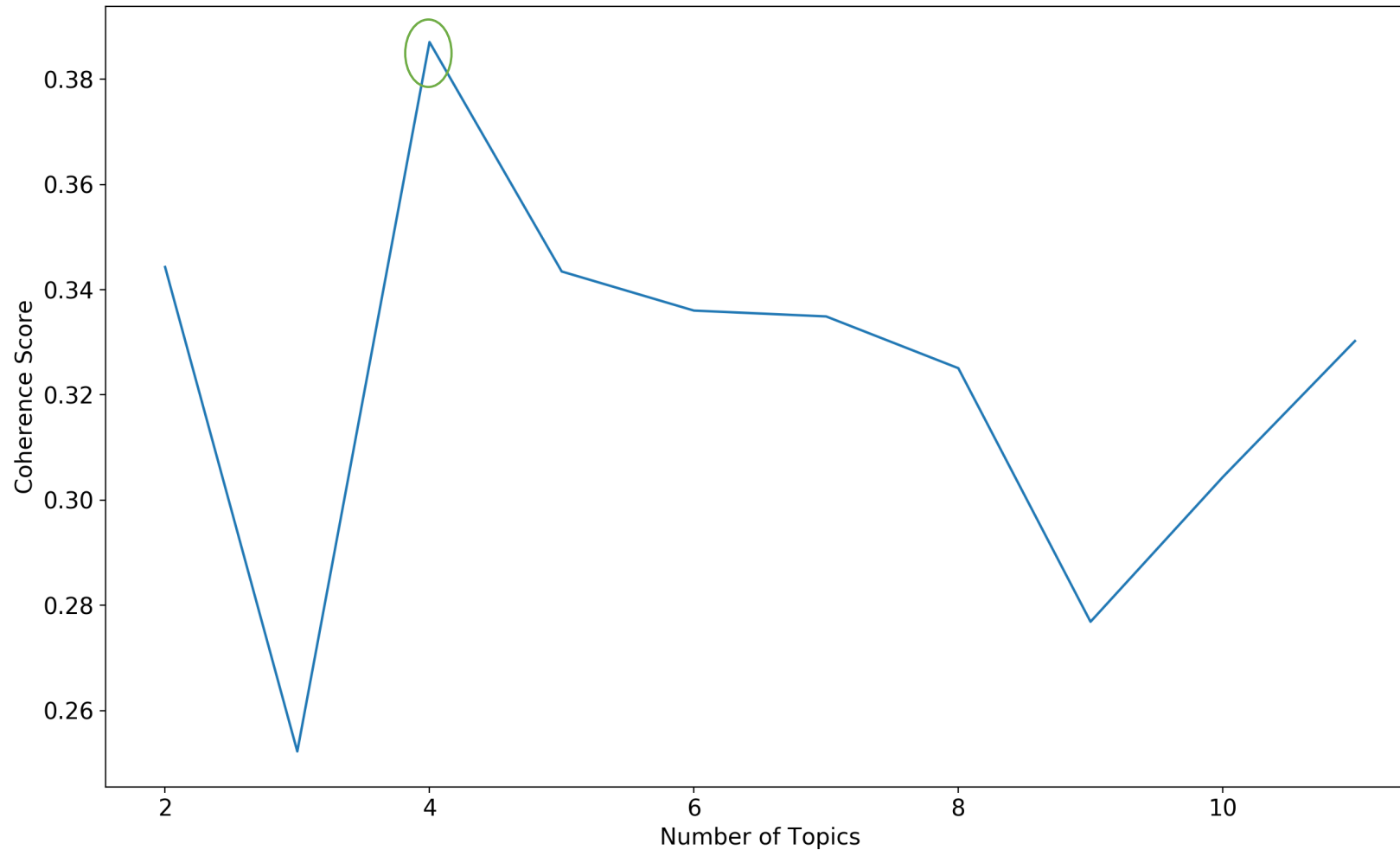
yoga

health

gym

today

Optimal Number of Topics vs Coherence Score LDA



$K = 4$

Coherence Value = 0.3871

Topics using LDA

Topic1

diet
workout
new
go
day
beyonce
get
today
bitch
gym

Topic2

vegan
yoga
job
every_woman
cooks_goe
therapy_remember
life_juggle
everyone_birthday
eat
boyfriend

Topic3

swimming
swim
day
much
support
really
try
always
relationship
pool

Topic4

fitness
amp
wellness
health
time
great
look
hiking
make
love

Topics using LDA

- *coherent topics*

Topic1

diet
workout
new
go
day
beyonce
get
today
bitch
gym

Topic2

vegan
yoga
job
every_woman
cooks_goe
therapy_remember
life_juggle
everyone_birthday
eat
boyfriend

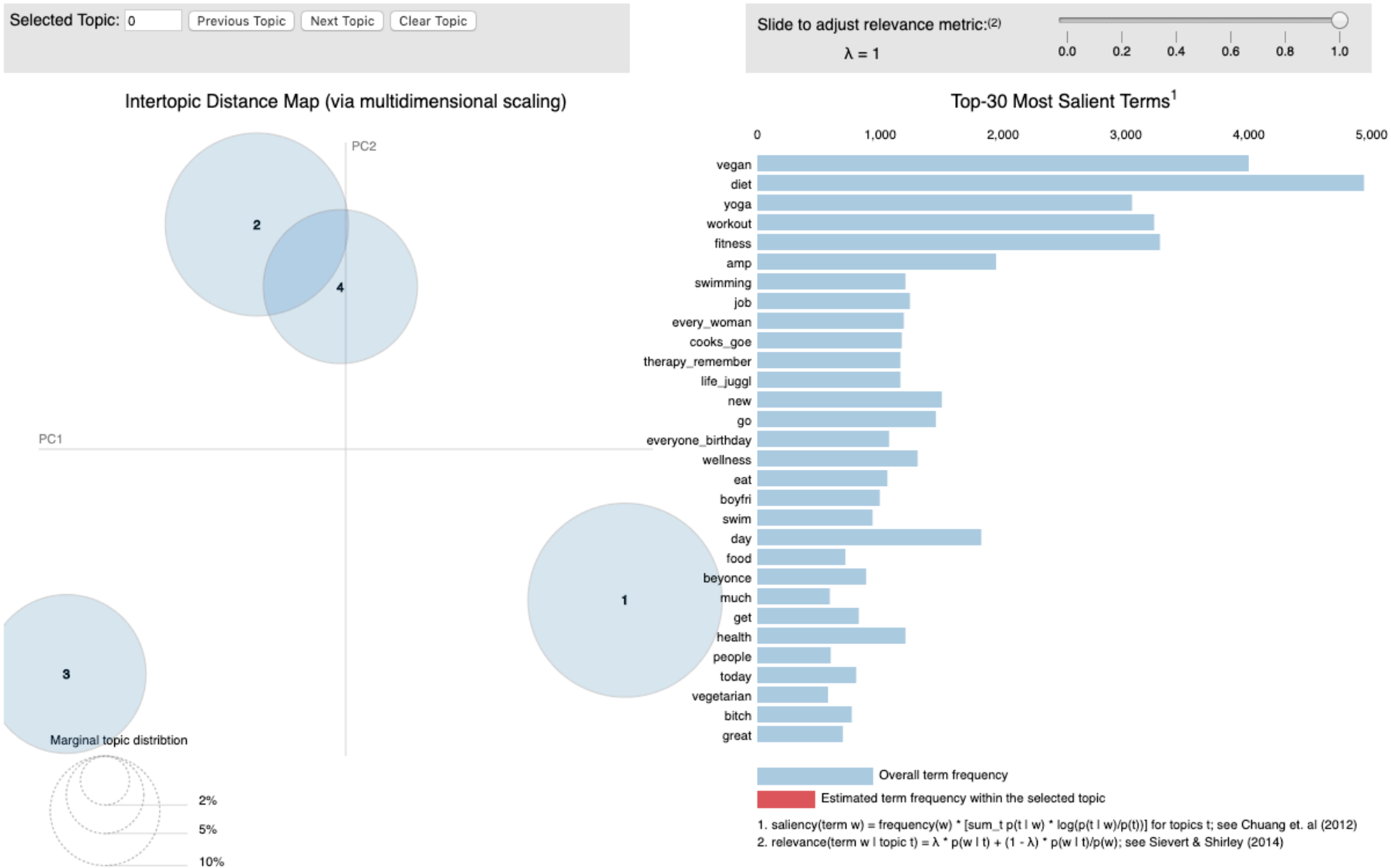
Topic3

swimming
swim
day
much
support
really
try
always
relationship
pool

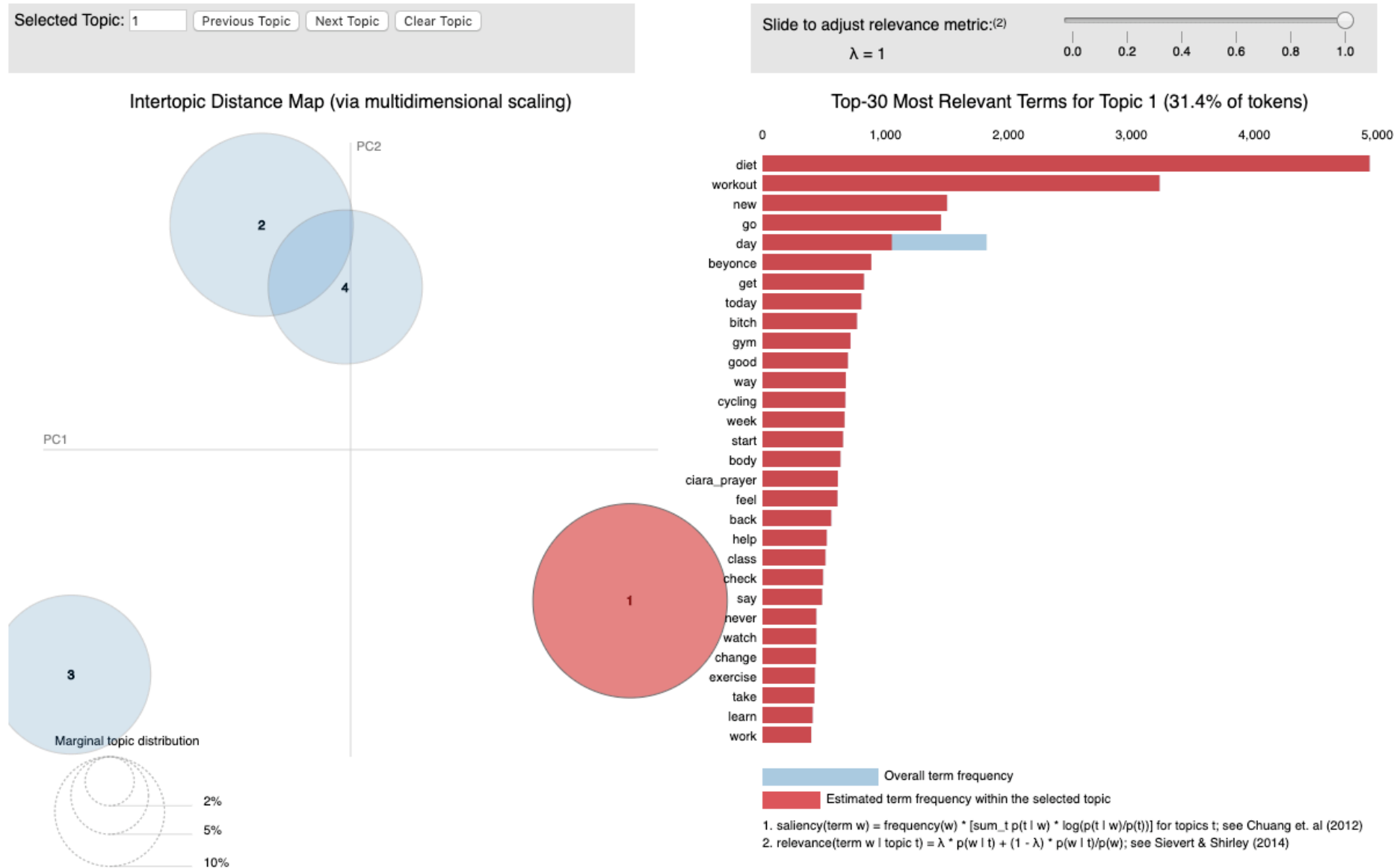
Topic4

fitness
amp
wellness
health
time
great
look
hiking
make
love

Visualization of Topics- pyLDAVIS



Visualization of Topics- pyLDAVIS



Manual Annotation (Train/Test data)

- Intent of tweets
- **500** tweets from train data
- **500 New** tweets for test data
- Calculate accuracy with ground truth

Explanation Observation 1

Tweet 1: *I lost 28kg / 61 lbs in 6 months! I changed my diet and went gym 5/6 times a week. I'm lighter and much happier now.*

Explanation Observation 1

Tweet 1: *I lost 28kg / 61 lbs in 6 months! I changed my diet and went gym 5/6 times a week. I'm lighter and much happier now.*

Annotated topic: **Topic 1**

Predicted topic: **Topic 1**

Explanation Observation 1

Tweet 1: *I lost 28kg / 61 lbs in 6 months! I changed my diet and went gym 5/6 times a week. I'm lighter and much happier now.*

Annotated topic: **Topic 1**

Predicted topic: **Topic 1**

Topic 1
<i>diet</i>
<i>workout</i>
<i>new</i>
<i>go</i>
<i>day</i>
<i>beyonce</i>
<i>get</i>
<i>today</i>
<i>bitch</i>
<i>gym</i>

Explanation Observation 1

Tweet 1: *I lost 28kg / 61 lbs in 6 months! I changed my diet and went gym 5/6 times a week. I'm lighter and much happier now.*

Annotated topic: **Topic 1**

Predicted topic: **Topic 1**

Mean KL Divergence= **0.025**

Accuracy Score= **94%**

Topic 1
<i>diet</i>
<i>workout</i>
<i>new</i>
<i>go</i>
<i>day</i>
<i>beyonce</i>
<i>get</i>
<i>today</i>
<i>bitch</i>
<i>gym</i>

Explanation Observation 1

incomprehensible
topic

Tweet 1: *I lost 28kg / 61 lbs in 6 months! I changed my diet and went gym 5/6 times a week. I'm lighter and much happier now.*

Annotated topic: **Topic 1**

Predicted topic: **Topic 1**

Mean KL Divergence= **0.025**

Accuracy Score= **94%**

Topic 1
diet
workout
new
go
day
beyonce
get
today
bitch
gym

y=0 (probability 0.007, score -4.923) top features

Contribution?	Feature
-0.551	<BIAS>
-4.373	Highlighted in text (sum)

i lost 28kg / 61 lbs in 6 months! i changed my diet and went gym 5/6 times a week. i'm lighter and much happier now.

y=1 (probability 0.893, score 2.409) top features

Contribution?	Feature
+3.039	Highlighted in text (sum)
-0.630	<BIAS>

i lost 28kg / 61 lbs in 6 months! i changed my diet and went gym 5/6 times a week. i'm lighter and much happier now.

y=2 (probability 0.087, score -2.324) top features

Contribution?	Feature
-0.589	<BIAS>
-1.735	Highlighted in text (sum)

i lost 28kg / 61 lbs in 6 months! i changed my diet and went gym 5/6 times a week. i'm lighter and much happier now.

y=3 (probability 0.005, score -5.229) top features

Contribution?	Feature
-0.579	<BIAS>
-4.650	Highlighted in text (sum)

i lost 28kg / 61 lbs in 6 months! i changed my diet and went gym 5/6 times a week. i'm lighter and much happier now.

y=4 (probability 0.008, score -4.847) top features

Contribution?	Feature
-0.243	<BIAS>
-4.605	Highlighted in text (sum)

i lost 28kg / 61 lbs in 6 months! i changed my diet and went gym 5/6 times a week. i'm lighter and much happier now.

Explanation Observation 1

Tweet 1: *I lost 28kg / 61 lbs in 6 months! I changed my diet and went gym 5/6 times a week. I'm lighter and much happier now.*

Annotated topic: **Topic 1**

Predicted topic: **Topic 1**

Mean KL Divergence= **0.025**

Accuracy Score= **94%**

Prediction Probability= **0.89**

Score= **2.4**

Contribution= **+ve**

Topic 1

diet

workout

new

go

day

beyonce

get

today

bitch

gym

y=0 (probability 0.007, score -4.923) top features

Contribution?	Feature
-0.551	<BIAS>
-4.373	Highlighted in text (sum)

i lost 28kg / 61 lbs in 6 months! i changed my diet and went gym 5/6 times a week. i'm lighter and much happier now.

y=1 (probability 0.893, score 2.409) top features

Contribution?	Feature
+3.039	Highlighted in text (sum)
-0.630	<BIAS>

i lost 28kg / 61 lbs in 6 months! i changed my diet and went gym 5/6 times a week. i'm lighter and much happier now.

y=2 (probability 0.087, score -2.324) top features

Contribution?	Feature
-0.589	<BIAS>
-1.735	Highlighted in text (sum)

i lost 28kg / 61 lbs in 6 months! i changed my diet and went gym 5/6 times a week. i'm lighter and much happier now.

y=3 (probability 0.005, score -5.229) top features

Contribution?	Feature
-0.579	<BIAS>
-4.650	Highlighted in text (sum)

i lost 28kg / 61 lbs in 6 months! i changed my diet and went gym 5/6 times a week. i'm lighter and much happier now.

y=4 (probability 0.008, score -4.847) top features

Contribution?	Feature
-0.243	<BIAS>
-4.605	Highlighted in text (sum)

i lost 28kg / 61 lbs in 6 months! i changed my diet and went gym 5/6 times a week. i'm lighter and much happier now.

Explanation Observation 2

Tweet 2: *Swimming is great. It's a perfect workout. #fitness #wellness*

Explanation Observation 2

Tweet 2: *Swimming is great. It's a perfect workout. #fitness #wellness*

Annotated topic: **Topic 3**

Predicted topic: **Topic 4**

Explanation Observation 2

Tweet 2: *Swimming is great. It's a perfect workout. #fitness #wellness*

Annotated topic: **Topic 3**

Predicted topic: **Topic 4**

Topic 3	Topic 4
<i>swimming</i>	<i>fitness</i>
<i>swim</i>	<i>amp</i>
<i>day</i>	<i>wellness</i>
<i>much</i>	<i>health</i>
<i>support</i>	<i>time</i>
<i>really</i>	<i>great</i>
<i>try</i>	<i>look</i>
<i>always</i>	<i>hiking</i>
<i>relationship</i>	<i>make</i>
<i>pool</i>	<i>love</i>

Explanation Observation 2

Tweet 2: *Swimming is great. It's a perfect workout. #fitness #wellness*

Annotated topic: **Topic 3**

Predicted topic: **Topic 4**

Mean KL Divergence= **0.061**

Accuracy Score= **92%**

Topic 3	Topic 4
<i>swimming</i>	<i>fitness</i>
<i>swim</i>	<i>amp</i>
<i>day</i>	<i>wellness</i>
<i>much</i>	<i>health</i>
<i>support</i>	<i>time</i>
<i>really</i>	<i>great</i>
<i>try</i>	<i>look</i>
<i>always</i>	<i>hiking</i>
<i>relationship</i>	<i>make</i>
<i>pool</i>	<i>love</i>

Explanation Observation 2

Tweet 2: *Swimming is great. It's a perfect workout.*
#fitness #wellness

Annotated topic: **Topic 3**

Predicted topic: **Topic 4**

Mean KL Divergence= **0.061**

Accuracy Score= **92%**

Prediction Probability= **0.16**

Score= **-1.75**

Contribution= **-ve**

Topic 3

swimming

swim
day
much
support
really
try
always
relationship
pool

Topic 4

fitness

amp
wellness
health
time
great
look
hiking
make
love

y=0 (probability 0.002, score -6.325) top features

Contribution?	Feature
-0.663	<BIAS>
-5.662	Highlighted in text (sum)

swimming is great. it's a perfect workout. #fitness #wellness

y=1 (probability 0.023, score -3.774) top features

Contribution?	Feature
-0.487	<BIAS>
-3.287	Highlighted in text (sum)

swimming is great. it's a perfect workout. #fitness #wellness

y=2 (probability 0.000, score -8.224) top features

Contribution?	Feature
-0.568	<BIAS>
-7.657	Highlighted in text (sum)

swimming is great. it's a perfect workout. #fitness #wellness

y=3 (probability 0.155, score -1.745) top features

Contribution?	Feature
-0.651	<BIAS>
-1.094	Highlighted in text (sum)

swimming is great. it's a perfect workout. #fitness #wellness

y=4 (probability 0.819, score 1.288) top features

Contribution?	Feature
+1.501	Highlighted in text (sum)
-0.213	<BIAS>

swimming is great. it's a perfect workout. #fitness #wellness

Topic 3

Explanation Observation 2

Tweet 2: *Swimming is great. It's a perfect workout.*
#fitness #wellness

Annotated topic: **Topic 3**

Predicted topic: **Topic 4**

Mean KL Divergence= **0.061**

Accuracy Score= **92%**

Prediction Probability= **0.82**

Score= **1.3**

Contribution= **+ve**

Topic 3

swimming

swim
day
much
support
really
try
always
relationship
pool

Topic 4

fitness

amp
wellness
health
time
great
look
hiking
make
love

y=0 (probability **0.002**, score **-6.325**) top features

Contribution?	Feature
-0.663	<BIAS>
-5.662	Highlighted in text (sum)

swimming is great. it's a perfect workout. #fitness #wellness

y=1 (probability **0.023**, score **-3.774**) top features

Contribution?	Feature
-0.487	<BIAS>
-3.287	Highlighted in text (sum)

swimming is great. it's a perfect workout. #fitness #wellness

y=2 (probability **0.000**, score **-8.224**) top features

Contribution?	Feature
-0.568	<BIAS>
-7.657	Highlighted in text (sum)

swimming is great. it's a perfect workout. #fitness #wellness

y=3 (probability **0.155**, score **-1.745**) top features

Contribution?	Feature
-0.651	<BIAS>
-1.094	Highlighted in text (sum)

swimming is great. it's a perfect workout. #fitness #wellness

y=4 (probability **0.819**, score **1.288**) top features

Contribution?	Feature
+1.501	Highlighted in text (sum)
-0.213	<BIAS>

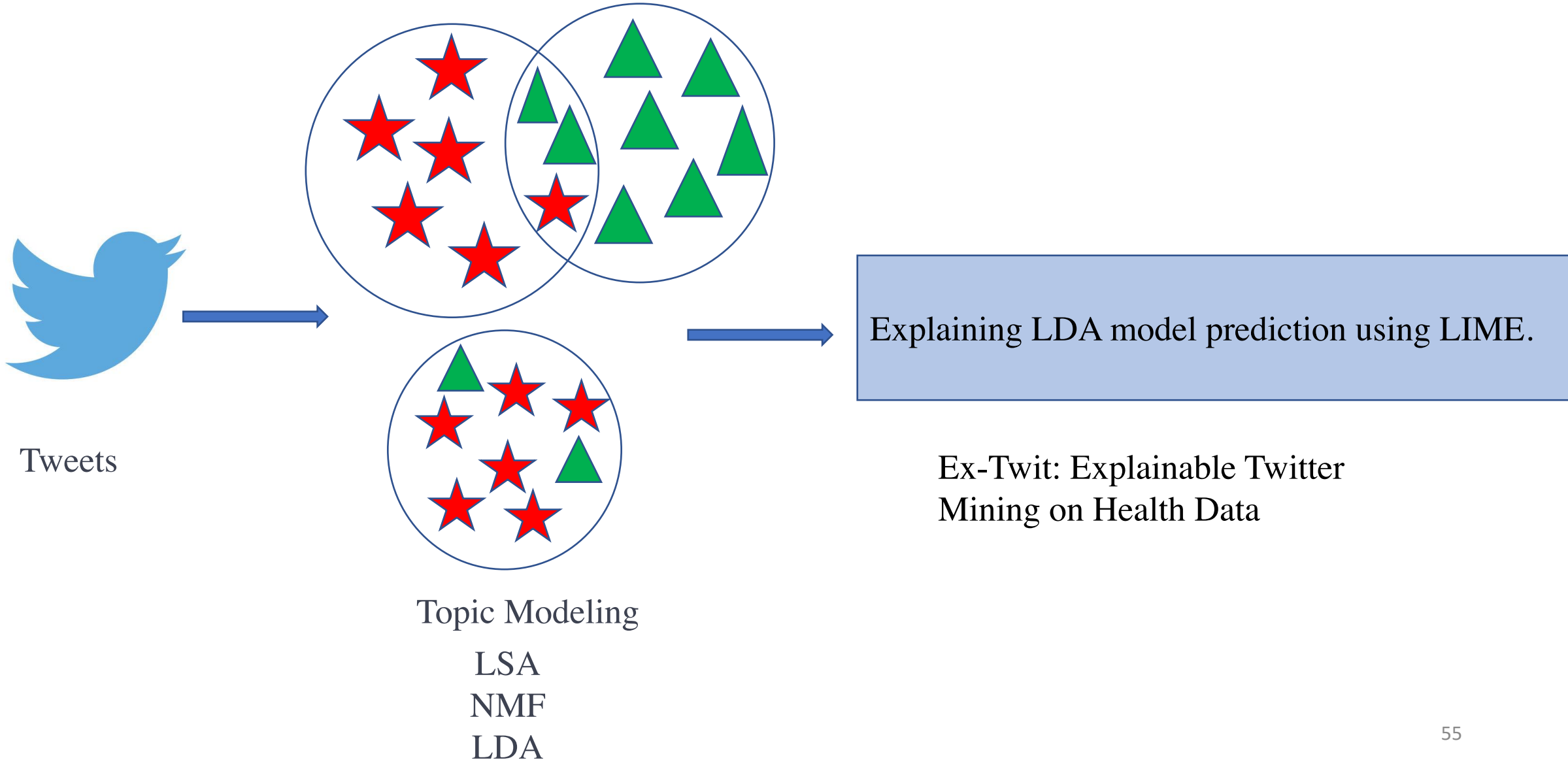
swimming is great. it's a perfect workout. #fitness #wellness

Topic 4

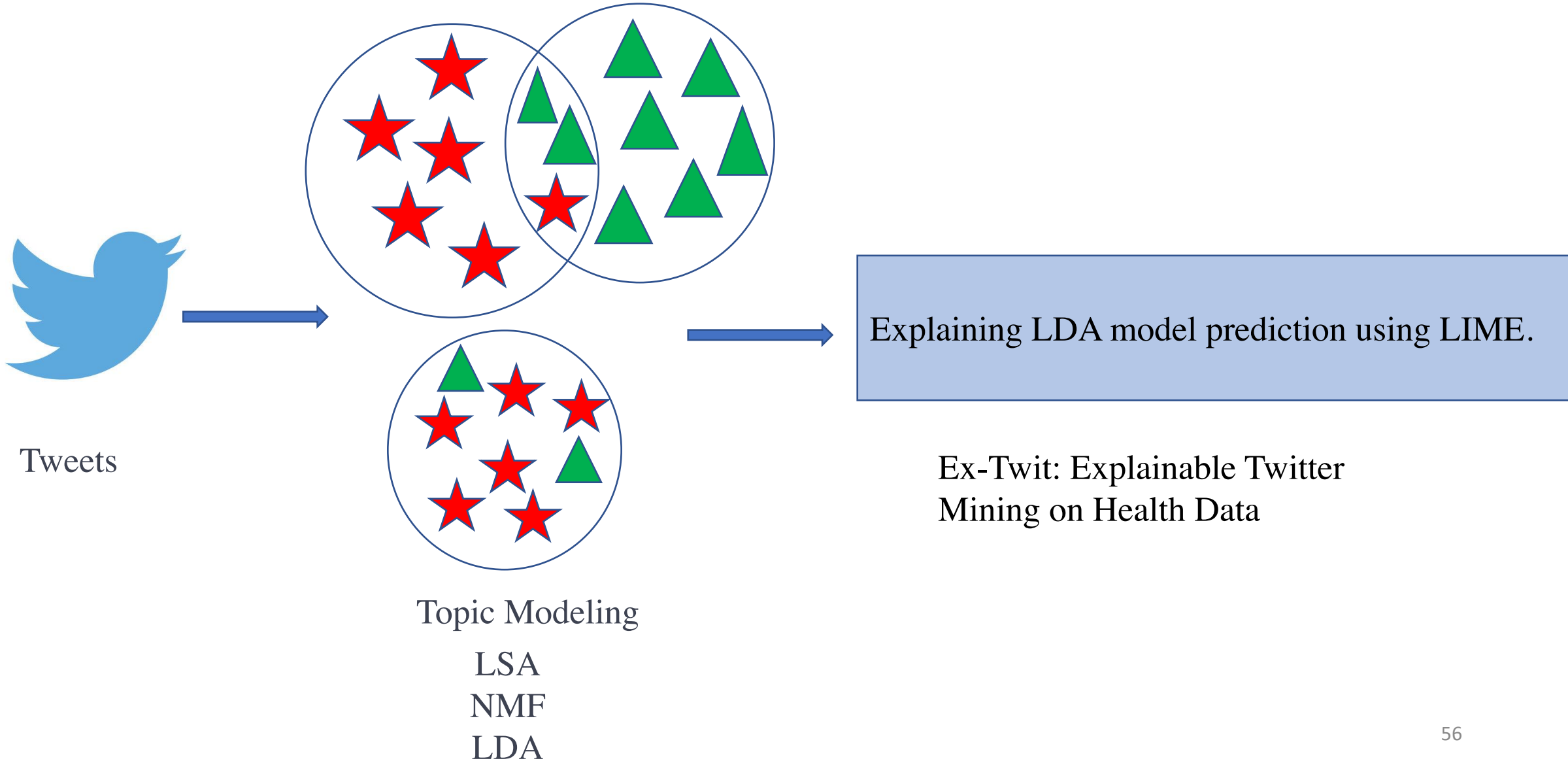
Future Work

- Is there bias in data?
- Observe Scalability.

Summary



QUESTION?



THANK YOU

Paper link: <https://arxiv.org/abs/1906.02132>

Slide: https://tunazislam.github.io/files/SocialNLP_IJCAI_2019_Tunaz.pdf

Tunazzina Islam

Ph.D. Student

Department of Computer Science

Purdue University, West Lafayette



islam32@purdue.edu



<https://tunazislam.github.io/>



@Tunaz_Islam

Backup Slides

Manual Annotation

- Intent of tweets.
- For example:
 - **Tweet 1:** *Learning some traditional yoga with my good friend.*
 - **Tweet 2:** *Why You Should #LiftWeights to Lose #BellyFat #Fitness
#core #abs #diet #gym #bodybuilding #workout #yoga*

Manual Annotation

- Intent of tweets.
- For example:
 - **Tweet 1:** *Learning some traditional yoga with my good friend.*
 - **Tweet 2:** *Why You Should #LiftWeights to Lose #BellyFat #Fitness #core #abs #diet #gym #bodybuilding #workout #yoga*

Yoga activity

Workout, Diet

Manual Annotation

- Intent of tweets.
- For example:
 - **Tweet 1:** *Learning some traditional yoga with my good friend.*
 - **Tweet 2:** *Why You Should #LiftWeights to Lose #BellyFat #Fitness
#core #abs #diet #gym #bodybuilding #workout #yoga*

Topic 2

Topic 1