Does Yoga Make You Happy? Analyzing Twitter User Happiness using Textual and Temporal Information

Tunazzina Islam, Dan Goldwasser

Department of Computer Science

Purdue University, West Lafayette, IN

IEEE BigData 2020

Date: December 10-13, 2020

PURDUE UNIVERSITY

Motivation

follow

Emily
@UCLA alumni, traveler, yoga and nature lover. Tend to think in 3+ different languages.

Los Angeles, CA

Smith
@Geologist, Believer in the Love religion, spontaneous fun maker.

Denver, CO

T1: Scorpion pose or Vrischikasana with a wall. Love this feeling. #yoga #scorpionpose.
T2: Busy day at work. Feeling energized after dancing @lafitness.

T1: Great job @Emily.
T2: Vegan Drive Through experience. we tried getting one of everything.

follow

lafitness
A welcoming dance & fitness facility for movers ages 13+. Streaming classes daily on YouTube, support us by joining our Patreon Community!

Los Angeles, CA

T1: Gentle #Yoga & Meditation w/ Michael. BreakOut Studios Online Classes.
T2: Join us for Saturday morning yoga class. #weekendyoga

Motivation

follow

Description

Location

Tweets

Emily
@UCLA alumni, traveler, yoga and nature lover. Tend to think in 3+ different languages.

Los Angeles, CA

T1: Scorpion pose or Vrischikasana with a wall. Love this feeling. #yoga #scorpionpose.
T2: Busy day at work. Feeling energized after dancing **@lafitness**.

Smith
@Geologist, Believer in the Love religion, spontaneous fun maker.

Denver, CO

T1: Great job **@Emily**
T2: Vegan Drive Through experience. we tried getting one of everything.

follow

lafitness
A welcoming dance & fitness facility for movers ages 13+. Streaming classes daily on YouTube, support us by joining our Patreon Community!

Los Angeles, CA

T1: Gentle #Yoga & Meditation w/ Michael. BreakOut Studios Online Classes.
T2: Join us for Saturday morning yoga class. #weekendyoga

Motivation

follow

Description

Location

Tweets

Emily
@UCLA alumni, traveler, yoga and nature lover. Tend to think in 3+ different languages.

Los Angeles, CA

T1: Scorpion pose or Vrischikasana with a wall. Love this feeling. **#yoga** #scorpionpose.
T2: Busy day at work. Feeling energized after dancing **@lafitness**.

Smith
@Geologist, Believer in the Love religion, spontaneous fun maker.

Denver, CO

T1: Great job **@Emily**
T2: Vegan Drive Through experience. we tried getting one of everything.

follow

lafitness
A welcoming dance & fitness facility for movers ages 13+. Streaming classes daily on YouTube, support us by joining our Patreon Community!

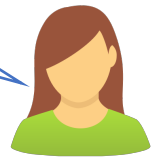
Los Angeles, CA

T1: Gentle **#Yoga** & Meditation w/ Michael. BreakOut Studios Online Classes.
T2: Join us for Saturday morning yoga class. #weekendyoga

Motivation

Practitioner

Other



follow

Emily
@UCLA alumni, traveler, yoga and nature lover. Tend to think in 3+ different languages.

Los Angeles, CA

T1: Scorpion pose or Vrischikasana with a wall. Love this feeling. **#yoga** #scorpionpose.
T2: Busy day at work. Feeling energized after dancing **@lafitness**.

Smith
@Geologist, Believer in the Love religion, spontaneous fun maker.

Denver, CO

T1: Great job **@Emily**
T2: Vegan Drive Through experience. we tried getting one of everything.

follow

Promotional

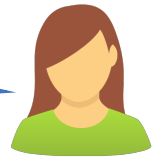
lafitness
A welcoming dance & fitness facility for movers ages 13+. Streaming classes daily on YouTube, support us by joining our Patreon Community!

Los Angeles, CA

T1: Gentle **#Yoga** & Meditation w/ Michael. BreakOut Studios Online Classes.
T2: Join us for Saturday morning yoga class. #weekendyoga

Motivation

Practitioner



Emily
@UCLA alumni, traveler, yoga and nature lover. Tend to think in 3+ different languages.

Los Angeles, CA

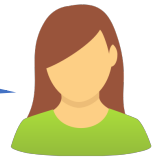
Positive emotion

T1: Scorpion pose or Vrischikasana with a wall. **Love this feeling**. #yoga #scorpionpose.

T2: Busy day at work. Feeling energized after dancing @lafitness.

Motivation

Practitioner



Emily
@UCLA alumni, traveler, yoga and nature lover. Tend to think in 3+ different languages.

Los Angeles, CA

Positive emotion

T1: Scorpion pose or Vrischikasana with a wall. **Love this feeling**. #yoga #scorpionpose.

T2: Busy day at work. Feeling energized after dancing @lafitness.

Find Causal relationship between "*Practicing Yoga*" and "*Being Happy*"

Methodology

To find out causal features from the text, we measure two variables:

- **Yoga activity level** - content analysis.
- **Happiness level** - emotional state.

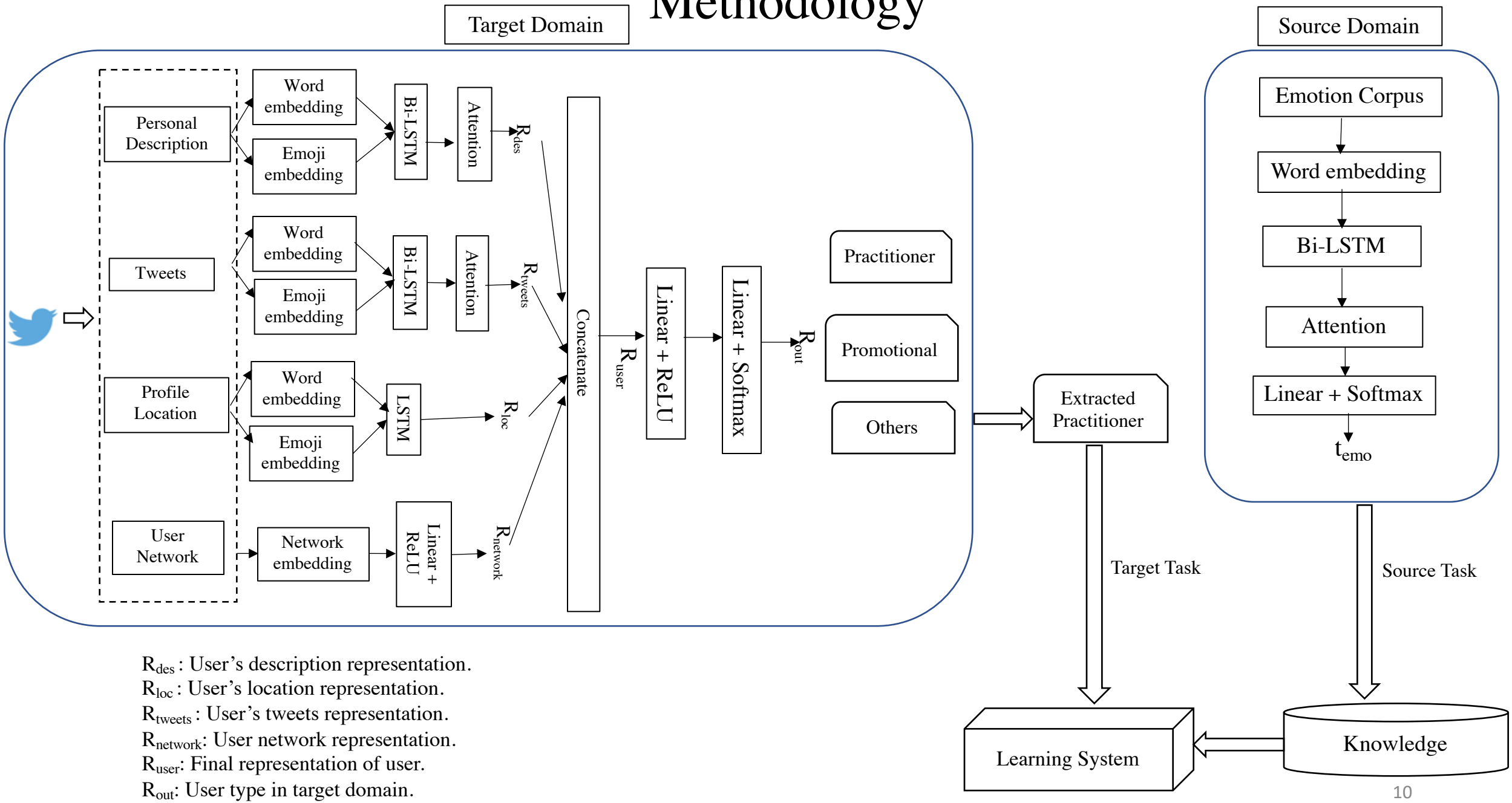
Model

- **Yoga User Network (YUN)** - a joint embedding model based on the fusion of neural networks with attention mechanism leveraging users' social and textual information to understand users' yoga activity.

- **Emotion Detection Model (BiLSTMAttEmo)** - an attention-based neural network model trained on emotion corpus (source domain).

- **Transfer learning** - transfers knowledge from source domain (emotion corpus) to measure the emotional state of yoga users (target domain).

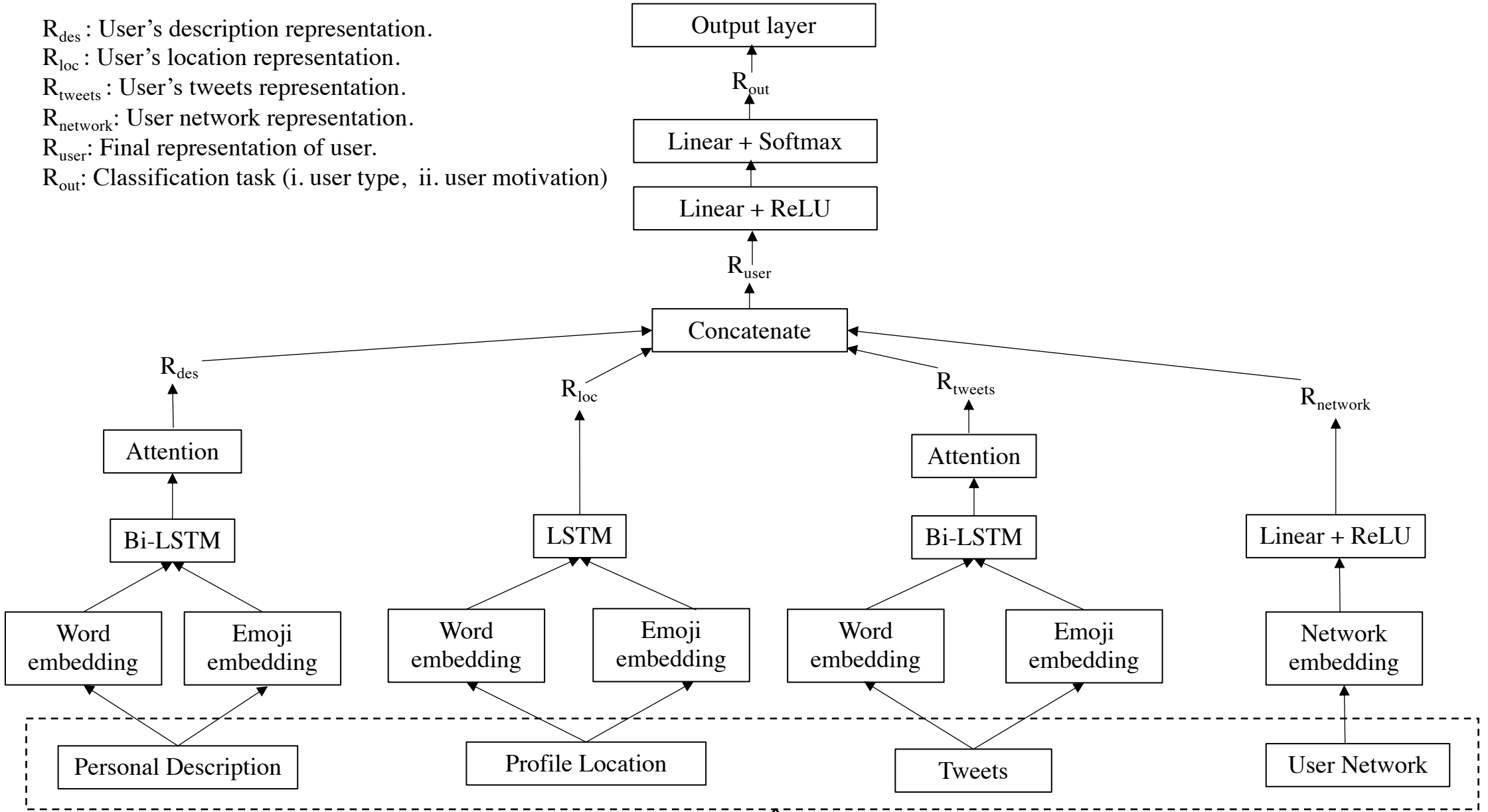
Methodology



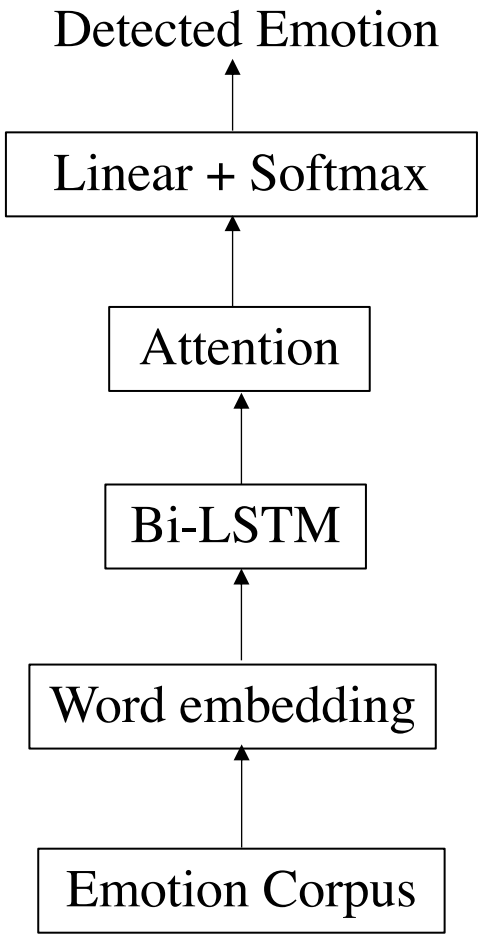
R_{des} : User's description representation.
R_{loc} : User's location representation.
R_{tweets} : User's tweets representation.
$R_{network}$: User network representation.
R_{user}: Final representation of user.
R_{out}: User type in target domain.
t_{emo}: Emotion of tweets in source domain.

YUN Model

R_{des} : User's description representation.
R_{loc} : User's location representation.
R_{tweets} : User's tweets representation.
$R_{network}$: User network representation.
R_{user}: Final representation of user.
R_{out}: Classification task (i. user type, ii. user motivation)



BiLSTMAttEmo Model



Dataset

- YUN model
  - ~ **0.4 million** yoga-related tweets from Twitter using Twitter streaming API (May to November of 2019) containing specific keywords.

  - **~15k** users have at least a yoga-related tweet in their timelines.
  - ~ **35 millions** of timeline tweets.

- Emotion Dataset
  - **~0.4 million** tweets annotated with 6 emotions *{joy, love, sadness, anger, fear, surprise}*

Baseline Models

- Emotion detection baseline – GRUEmo.
  - BiLSTMAttEmo performs better on target task.
- User type detection baseline – 10 baselines
  1. Description only;
  2. Location only;
  3. Tweets only;
  4. Network only;
  5. BERT finetuned with Description (Description_BERT);
  6. BERT fine-tuned with Location (Location_BERT);
  7. BERT fine-tuned with Tweets (Tweets_BERT);
  8. joint embedding on description and location (Des + Loc);
  9. joint embedding on description, location, and tweets (Des + Loc + Twt);
  10. joint embedding on description, location, and network (Des + Loc + Net).

  - YUN (Des + Loc + Twt + Net) outperforms those baselines.
    - Macro-avg F1 score: 74.2%

Causal Features

1. Yoga activity level
   - focuses on practitioner's first-hand experience.

   case (i): Explicitly having 1st Person Singular and 1st Person Plural Number.

   i.e. ***I loved yesterday's yoga session #motivational #calm.***

   - Rule-based approach

   case (ii): Implicit first-hand experience.

   i.e. ***feeling peaceful after doing morning yoga.***

   - Parts of speech tagging approach
     - filter out 2nd and 3rd Person Singular and Plural Number.
     - VBZ, NNP, NNS, NNPS, PRP, PRP$.

2. Yoga practitioner's happiness level
   - focuses on practitioner's two type of emotion *{joy, love}*

Granger Causality

- Granger causality has following two assumptions:
  (i) a cause occurs before its effect.
  (ii) knowledge of a cause can be used to predict its effect.

- A time series *X* (source) is said to Granger-cause a time series *Y* (target) if past values x_{t-i} are significant indicators in predicting y_t.

- In our work, we have source time series - Yoga activity level, *a* (cause) and target time series - Happiness level, *p* (effect). We calculate Granger causality as follows:

$$GC(p_t|p_{<t}, a_{<t}) = \sum_{i=1}^{m} \alpha_i p_{t-i} + \sum_{j=1}^{n} \beta_j a_{t-j} \ldots\ldots\ldots\ldots (1)$$

m *and* n *are the size of lags in the past observation,* α *and* β *are learnable parameters.*

Granger Causality

- Sort yoga activity level and happiness level based on DateTime.
- Apply Granger causality from (eqn.(1)) with different lags = 1, 2, 3, 4, 5.
- Null hypothesis "yoga activity does not Granger-cause happiness".
- Reject the null hypothesis if **p-value ≤ 0.05**, otherwise keep the null hypothesis.

Results and Analysis

- All Practitioners (~ 8k)
  1. **y + h** - considers all tweets containing 'yoga' and it does not include practitioner's first-hand experience.

  2. **Y + 1st + h** - focuses on yoga activity based on practitioner's first-hand experience.

TABLE 1: Granger causality result for lag = 5

Users	Feature	rn	kn	nc
All	y + h	1663	7120	2271
	y + 1^{st} + h	1447	4524	5083
Top 10%	y + h	700	399	6
	y + 1^{st} + h	546	551	8

rn: Number of users for whom we reject null hypothesis.
kn: Number of users for whom we keep null hypothesis.
nc: Number of users whose Granger Causality is not calculated.
y + h: Yoga and happiness.
y + 1^{st} + h: Yoga with 1^{st} hand experience and happiness.

Results and Analysis

- All Practitioners (~ 8k)
  1. **y + h** - considers all tweets containing 'yoga' and it does not include practitioner's first-hand experience.

  2. **Y + 1st + h** - focuses on yoga activity based on practitioner's first-hand experience.

TABLE 1: Granger causality result for lag = 5

Users	Feature	rn	kn	nc
All	y + h	1663	7120	2271
	y + 1^{st} + h	1447	4524	5083
Top 10%	y + h	700	399	6
	y + 1^{st} + h	546	551	8

rn: Number of users for whom we reject null hypothesis.
kn: Number of users for whom we keep null hypothesis.
nc: Number of users whose Granger Causality is not calculated.
y + h: Yoga and happiness.
y + 1^{st} + h: Yoga with 1^{st} hand experience and happiness.

Results and Analysis

- Top 10% (~ 1.1k) – tweeted most about "Yoga".
  1. **y + h** - considers all tweets containing 'yoga' and it does not include practitioner's first-hand experience.

  2. **Y + 1st + h** - focuses on yoga activity based on practitioner's first-hand experience.

TABLE 1: Granger causality result for lag = 5

Users	Feature	rn	kn	nc
All	y + h	1663	7120	2271
	y + 1^{st} + h	1447	4524	5083
Top 10%	y + h	700	399	6
	y + 1^{st} + h	546	551	8

rn: Number of users for whom we reject null hypothesis.
kn: Number of users for whom we keep null hypothesis.
nc: Number of users whose Granger Causality is not calculated.
y + h: Yoga and happiness.
y + 1^{st} + h: Yoga with 1^{st} hand experience and happiness.

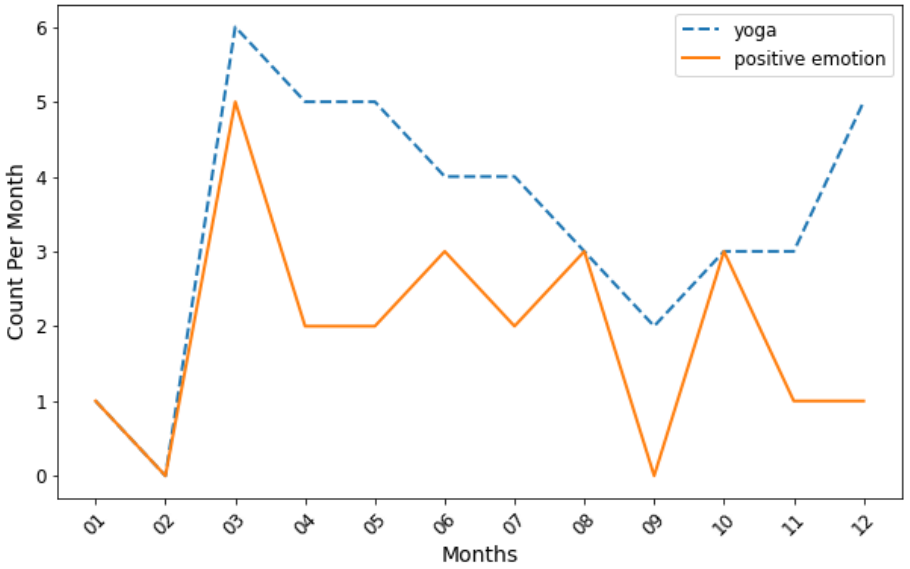
Results and Analysis

- Null hypothesis "Tweeting more does not Granger-cause positive emotion".

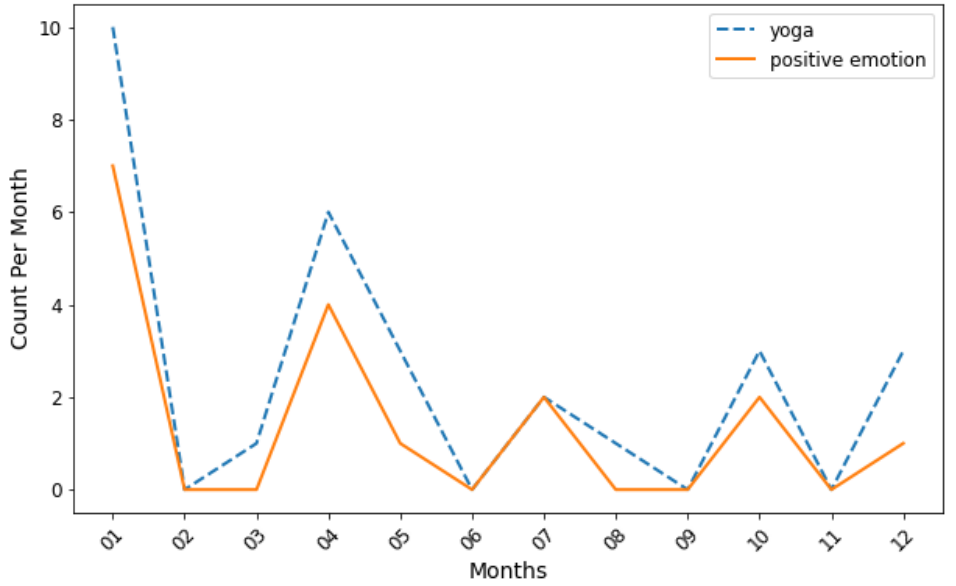
- Two causal features
  (i) number of tweets (cause)
  (ii) number of positive emotions (effect).

- **p-value > 0.05**, keep the null hypothesis which means the overall activity of the practitioners has no effect on happiness.

Results and Analysis



(a) Year 2018



(b) Year 2019

- Causal features: Yoga activity (blue dotted line) and Positive emotion (orange solid line).

- The causal feature temporally causes rise of positive emotion of the user in March 2018 and April 2019.

Discussion and Future Work

- Misclassifications in user type detection.
- Noisy transfer learning approach.
- Expensive data annotation.
- Develop a contextualized model to predict user type using minimal supervision.
- Similar for emotion detection.

THANK YOU ☺

Slide: https://tunazislam.github.io/files/IEEEBigData2020_causal_yoga_happiness.pdf

Questions?

Tunazzina Islam

Department of Computer Science,
Purdue University, West Lafayette, IN.
Email: **islam32@purdue.edu**

https://tunazislam.github.io/

@Tunaz_Islam

PURDUE UNIVERSITY

Backup Slides

Manual Annotation

- Intent of tweets.
- User description
- For example:
  - **Tweet 1:** *Learning some traditional yoga with my good friend*.
  - **Tweet 2:** *Our mission at 532Yoga is pretty simple; great teachers, great classes and superbly happy students #yoga*

Manual Annotation

- Intent of tweets.
- User description
- For example:
  - **Tweet 1:** *Learning some traditional yoga with my good friend.*
  - **Tweet 2:** *Our mission at 532Yoga is pretty simple; great teachers, great classes and superbly happy students #yoga*

Practitioner

Promotional