Do You Do Yoga? Understanding Twitter Users' Types and Motivations using Social and Textual Information

Tunazzina Islam, Dan Goldwasser

Department of Computer Science

Purdue University, West Lafayette, IN

IEEE ICSC 2021

Date: January 27-29, 2021

PURDUE UNIVERSITY®

Motivation

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. Good for spine. #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: RT: Vegan Drive Through experience. we tried getting one of everything.

follow

GYM

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. Join us for free class.
T2: Join us for Saturday morning yoga class. #weekendyoga

Motivation

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. Good for spine. #yoga #scorpionpose.
T2: Busy day at work. Feeling energized after dancing **@lafitness**.

Description

Location

Tweets

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

Denver, CO

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

follow

GYM

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. Join us for free class.
T2: Join us for Saturday morning yoga class. #weekendyoga

Motivation

follow

Description

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

Location

Los Angeles, CA

Tweets

T1: Scorpion pose or Vrischikasana with a wall. Good for spine. **#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: RT: Vegan Drive Through experience. we tried getting one of everything.

follow

GYM

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. Join us for free class.
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. Good for spine. **#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: RT: Vegan Drive Through experience. we tried getting one of everything.

GYM

follow

Promotional

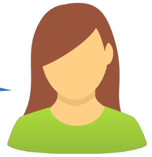
lafitness
A welcoming dance & fitness facility for movers ages 13+. Streaming classes daily on YouTube, support us by joining our Patreon Community!

Los Angeles, CA

T1: Gentle **#Yoga** & Meditation w/ Michael. BreakOut Studios Online Classes. Join us for free class.
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

Health

T1: Scorpion pose or Vrischikasana with a wall. **Good for spine**. #yoga #scorpionpose.

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

Motivation

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

Other (Business)

T1: Gentle **#Yoga** & Meditation w/ Michael. BreakOut Studios Online Classes. **Join us for free class**.

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

Health

T1: Scorpion pose or Vrischikasana with a wall. **Good for spine**. **#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: RT: Vegan Drive Through experience. we tried getting one of everything.

Other

GYM

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

Promotional

Other (Business)

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

Methodology

- We use following sources of information to build our model:

    1) Tweet text;
    2) User network;
    3) Metadata including user location and description.

- Our model employs those sources and then jointly builds a neural network model to generate a dense vector representation for each field and finally concatenates these representations.

Downstream Tasks

We demonstrate our model on two downstream tasks:

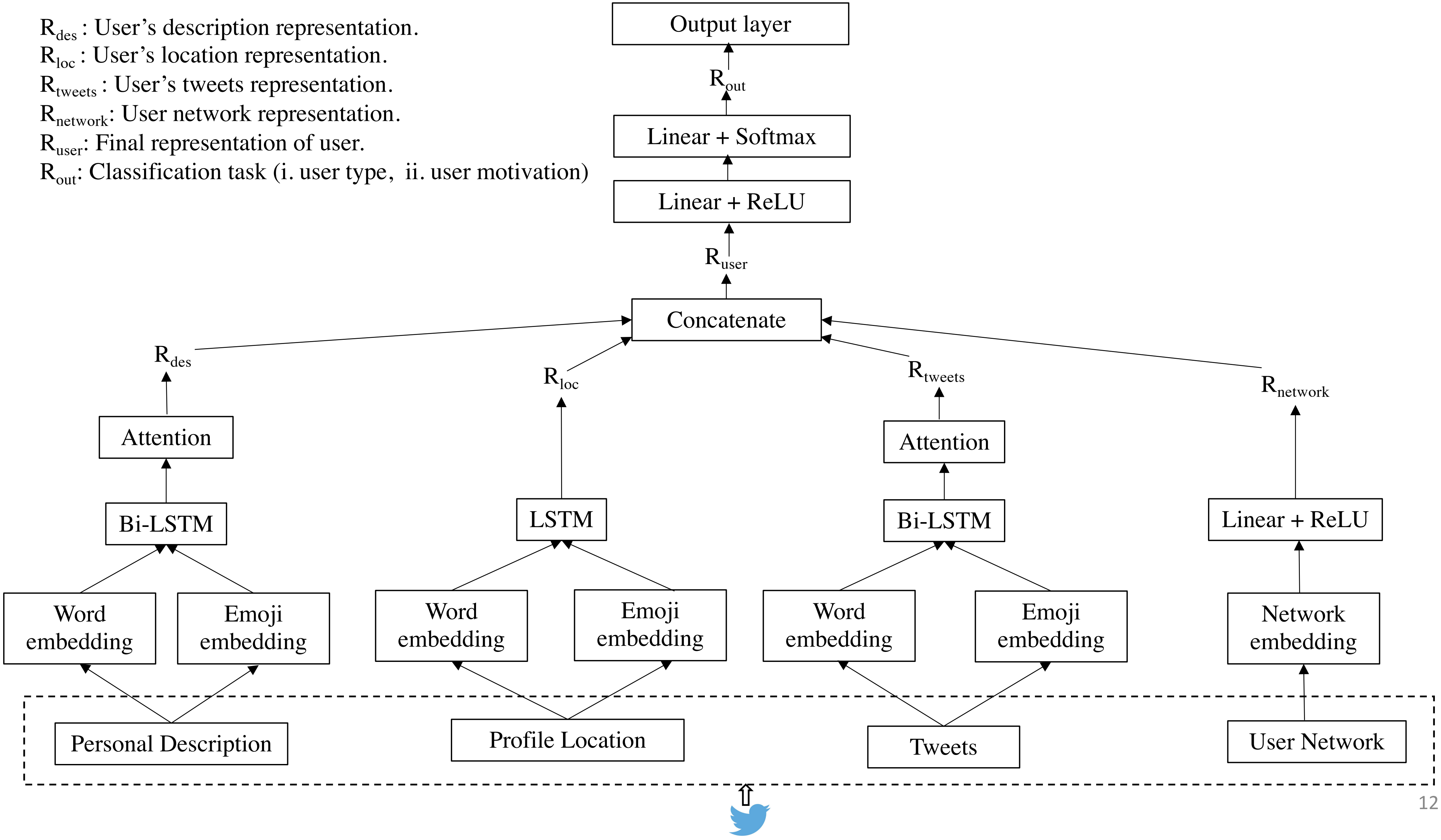
1) Finding user type
    1) Practitioner
    2) Promotional
    3) Other
2) Finding user motivation
    1) Health
    2) Spiritual
    3) Other

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 user type and motivation.

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)



Dataset

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

  - ~**1300** users have at least a yoga-related tweet in their timelines.

  - ~ **3 million** of timeline tweets.

Baseline Models

- User type and motivation 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).

Results and Analysis

Evaluation Metrics:

- Accuracy
- Macro-avg F1 score

TABLE II: Performance comparisons on test data

	user type		**user motivation**	
Model	**Accuracy**	**Macro F1 score**	**Accuracy**	**Macro F1 score**
Description	0.725	0.693	0.782	0.575
Location	0.676	0.563	0.695	0.470
Tweets	0.721	0.687	0.744	0.504
Network	0.752	0.557	0.790	0.541
Description_BERT	0.718	0.681	0.771	0.528
Location_BERT	0.679	0.606	0.695	0.476
Tweets_BERT	0.760	0.669	0.805	0.551
Des + Loc	0.733	0.693	0.775	0.599
Des + Loc + Twt	0.760	0.725	0.794	0.580
Des + Loc + Net	0.775	0.723	0.828	**0.647**
YUN	**0.790**	**0.742**	**0.844**	0.619

Results and Analysis

YUN (Des + Loc + Twt + Net) outperforms the baselines.

- Accuracy (user type): 79.0%
- Macro-avg F1 score (user type): 74.2%
- Accuracy (user motivation): 84.4%

TABLE II: Performance comparisons on test data

	user type		**user motivation**	
Model	**Accuracy**	**Macro F1 score**	**Accuracy**	**Macro F1 score**
Description	0.725	0.693	0.782	0.575
Location	0.676	0.563	0.695	0.470
Tweets	0.721	0.687	0.744	0.504
Network	0.752	0.557	0.790	0.541
Description_BERT	0.718	0.681	0.771	0.528
Location_BERT	0.679	0.606	0.695	0.476
Tweets_BERT	0.760	0.669	0.805	0.551
Des + Loc	0.733	0.693	0.775	0.599
Des + Loc + Twt	0.760	0.725	0.794	0.580
Des + Loc + Net	0.775	0.723	0.828	**0.647**
YUN	**0.790**	**0.742**	**0.844**	0.619

Results and Analysis

Ablation study by training individual neural network model for each field i.e., description, location, tweets and network.

TABLE II: Performance comparisons on test data

	user type		user motivation	
Model	**Accuracy**	**Macro F1 score**	**Accuracy**	**Macro F1 score**
Description	0.725	0.693	0.782	0.575
Location	0.676	0.563	0.695	0.470
Tweets	0.721	0.687	0.744	0.504
Network	0.752	0.557	0.790	0.541
Description_BERT	0.718	0.681	0.771	0.528
Location_BERT	0.679	0.606	0.695	0.476
Tweets_BERT	0.760	0.669	0.805	0.551
Des + Loc	0.733	0.693	0.775	0.599
Des + Loc + Twt	0.760	0.725	0.794	0.580
Des + Loc + Net	0.775	0.723	0.828	**0.647**
YUN	**0.790**	**0.742**	**0.844**	0.619

Results and Analysis

Ablation study by training individual neural network model for each field i.e., description, location, tweets and network.

- profile description is the most informative.

TABLE II: Performance comparisons on test data

	user type		user motivation	
Model	**Accuracy**	**Macro F1 score**	**Accuracy**	**Macro F1 score**
Description	0.725	0.693	0.782	0.575
Location	0.676	0.563	0.695	0.470
Tweets	0.721	0.687	0.744	0.504
Network	0.752	0.557	0.790	0.541
Description_BERT	0.718	0.681	0.771	0.528
Location_BERT	0.679	0.606	0.695	0.476
Tweets_BERT	0.760	0.669	0.805	0.551
Des + Loc	0.733	0.693	0.775	0.599
Des + Loc + Twt	0.760	0.725	0.794	0.580
Des + Loc + Net	0.775	0.723	0.828	**0.647**
YUN	**0.790**	**0.742**	**0.844**	0.619

Results and Analysis

Ablation study by training individual neural network model for each field i.e., description, location, tweets and network.

- profile description is the most informative.
- User location has the lower performance.

TABLE II: Performance comparisons on test data

	user type		user motivation	
Model	**Accuracy**	**Macro F1 score**	**Accuracy**	**Macro F1 score**
Description	0.725	0.693	0.782	0.575
Location	0.676	0.563	0.695	0.470
Tweets	0.721	0.687	0.744	0.504
Network	0.752	0.557	0.790	0.541
Description_BERT	0.718	0.681	0.771	0.528
Location_BERT	0.679	0.606	0.695	0.476
Tweets_BERT	0.760	0.669	0.805	0.551
Des + Loc	0.733	0.693	0.775	0.599
Des + Loc + Twt	0.760	0.725	0.794	0.580
Des + Loc + Net	0.775	0.723	0.828	**0.647**
YUN	**0.790**	**0.742**	**0.844**	0.619

Results and Analysis

Ablation study by training individual neural network model for each field i.e., description, location, tweets and network.

- profile description is the most informative.
- User location has the lower performance.
- Excluding user network information (Des + Loc + Twt model) declines the performance of the final model.

TABLE II: Performance comparisons on test data

	user type		**user motivation**	
Model	**Accuracy**	**Macro F1 score**	**Accuracy**	**Macro F1 score**
Description	0.725	0.693	0.782	0.575
Location	0.676	0.563	0.695	0.470
Tweets	0.721	0.687	0.744	0.504
Network	0.752	0.557	0.790	0.541
Description_BERT	0.718	0.681	0.771	0.528
Location_BERT	0.679	0.606	0.695	0.476
Tweets_BERT	0.760	0.669	0.805	0.551
Des + Loc	0.733	0.693	0.775	0.599
Des + Loc + Twt	0.760	0.725	0.794	0.580
Des + Loc + Net	0.775	0.723	0.828	**0.647**
YUN	**0.790**	**0.742**	**0.844**	0.619

Error Analysis

- Misclassifications
  - Did not use contextualized word embedding for text.
  - Misleading or absence of profile description.
    - Users do not have their profile description in Twitter.
    - Users who provide profile description related to “yoga” but usually retweet yoga-related quotes.
  - Absence of user location information.

Discussion and Future Work

- Misclassifications in user type and motivation detection.
- Expensive data annotation.
- Develop a contextualized model to predict user type and motivation using minimal supervision.
- Use the model to understand users’ type and motivation for different lifestyle choices i.e. “keto diet”, “veganism”.

THANK YOU ☺

Slide: https://tunazislam.github.io/files/IEEE_ICSC_YUN.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