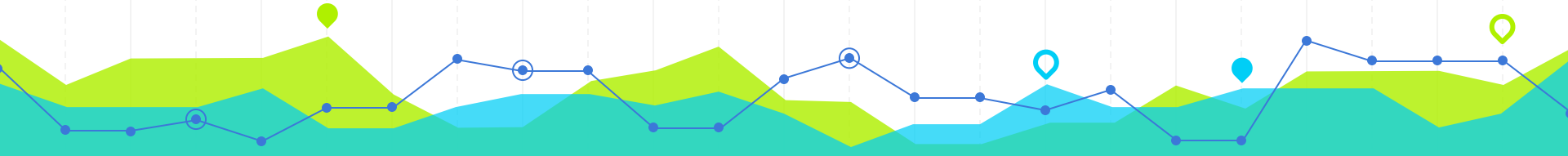
EmotionPush: Emotion and Response Time Prediction towards Human-Like Chatbots

Chieh-Yang Huang[1] and Lun-Wei Ku[1]
[1]Arizona State University, USA, [2] Academia Sinica, Taiwan

Outline

- Introduction
- EmotionPush Dataset
- Experiments
  - Emotion Classification
  - Response Time Prediction
- Conclusion



Introduction 1

Introduction

- A human-like chat abilities can be discussed from two aspects.
  - **Content**: what message to respond.
  - **Behavior**: the current metal states, the time to response.

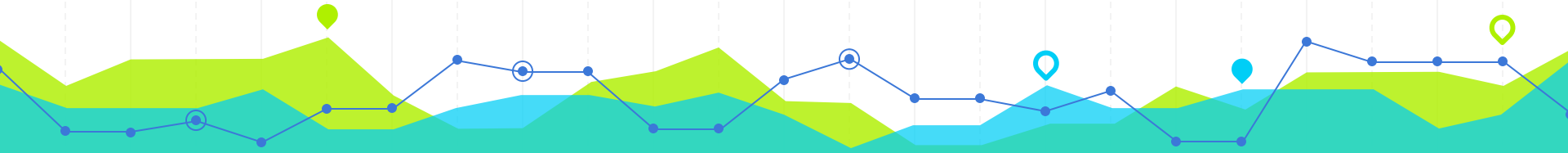
Learn from data!

Introduction

- Publicly-available dataset of private social dialogs is limited due to the privacy issues.
- Two ways to deal with this situation.
  - Building dataset. ➡ **Expensive**
  - Using data from different genres. ➡ **Could yield misleading results.**
    (forum-style web data / task-oriented dialog logs)

Introduction

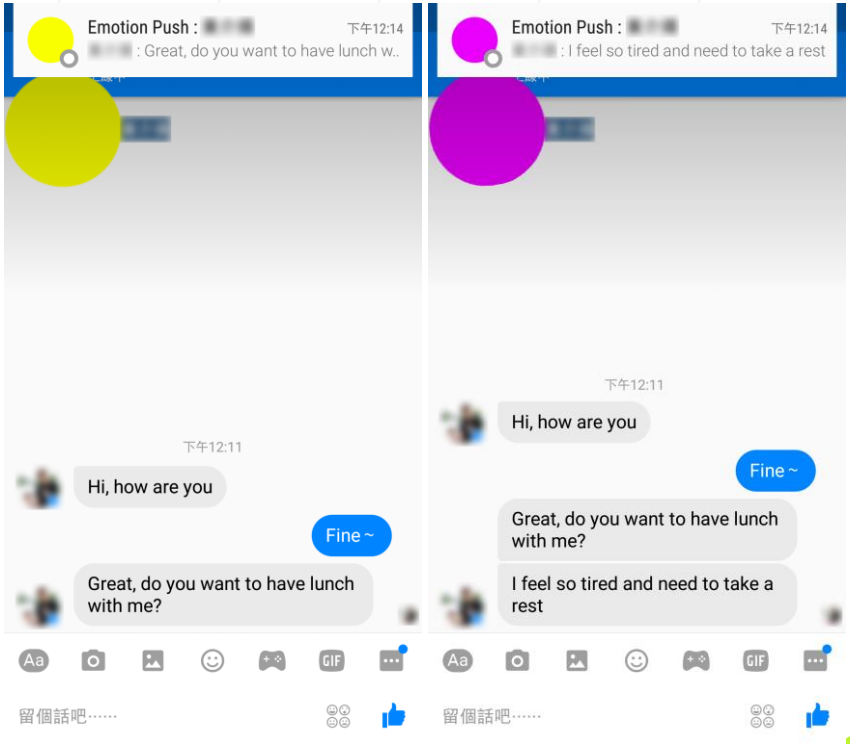
- **EmotionPush** Dataset
  - Collected from Facebook Messenger
  - Private information were Masked
- Two experiments are conducted on this dataset
  - Emotion Classification
  - Response Time Prediction



EmotionPush Dataset 2

EmotionPush Dataset

- **EmotionPush** App
  - An Android App that can sense the emotion of your friend on Facebook Messenger.



EmotionPush Dataset

- **EmotionPush** Backend
  - Feature: $\sum_{w_i \in Sentence} w_i$
  - Model: Liblinear
  - Data: LJ40K
  - Emotions:

EmotionPush Dataset

- **Dataset Description**

Log Type	Attributes
Message Log	logID, currentUserID, threadID, timestamp, senderID, message, emotion
Read Log	logID, currentUserID, threadID, timestamp, readerID

TABLE I
ATTRIBUTES OF THE EMOTIONPUSH LOG.

Property	Number	Property	Number
Message	162,031	Involved User	1,627
Thread	1,477	Word	978,452
Conversation	24,252	Distinct Word	51,170
Turn	101,135	Word/Msg	6.039

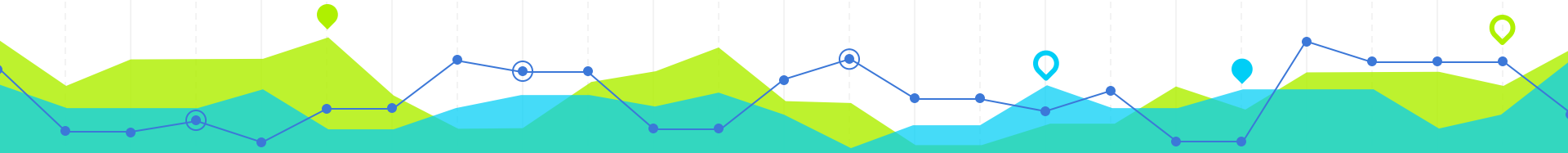
TABLE II
EMOTIONPUSH DATASET STATS.

EmotionPush Dataset

- **Dataset Masking**
  - **Name Entity Recognition.** Person Names, Locations, Organizations are turned into name_1, loc_3, org_7.
  - **Tokenization**.
  - Words appearing less than 5 times are turned into **UNKNOWN**.
  - Turning words into **random indices**.

EmotionPush Dataset

- **Dataset Embedding**
  - Trained by GloVe
  - Two versions:
    - Trained on the EmotionPush Dataset
    - Trained on the EmotionPush Dataset + the Sentiment140 Dataset (200,000 instances)



Experiments 3

Emotion Classification

- **Train**:
  - 70% EmotionPush Dataset
  - LJ40K Dataset
- **Test**:
  - 20% EmotionPush Dataset

Emotion Classification

- **Result**

Model	Joy	Anger	Sadness	Neutral
CNN	.905	.962	.973	.820
LSTM	.906	.965	.964	.816
EmotionPush [1]	.779	.771	.853	.323
EmotionPush [2]	.902	.966	.925	.526

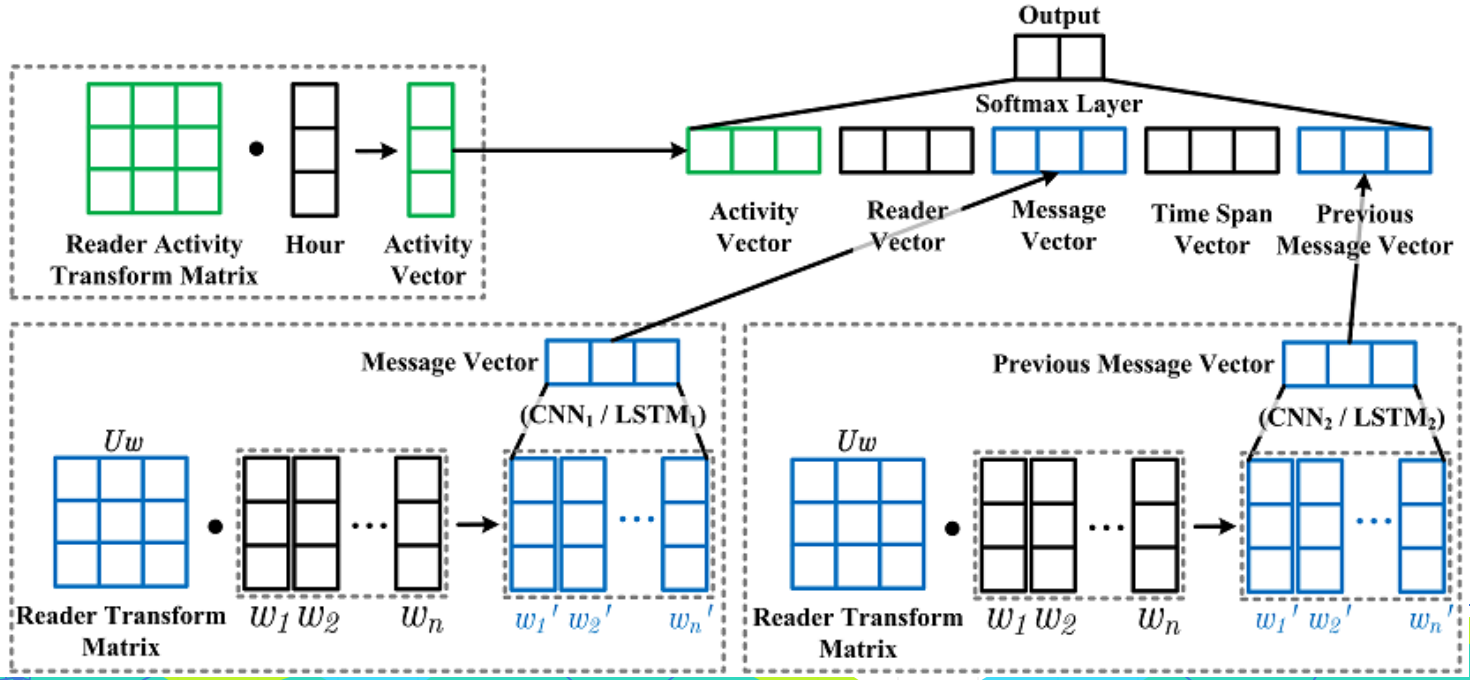
TABLE III
ACCURACY OF THE EMOTION CLASSIFICATION TASK. EMOTIONPUSH [1] AND EMOTIONPUSH [2] ARE TRAINED ON LJ40K AND THE EMOTIONPUSH DATASET BY LIBLINEAR, RESPECTIVELY.

Response Time Prediction

- **Response Time**
  = Time(E) – Time(c)
- Predictions
  - Receive a response / Not
  - 4 time intervals
    [<10s, 10-30s, 30-300s, >300s]

Response Time Prediction

- **Model**



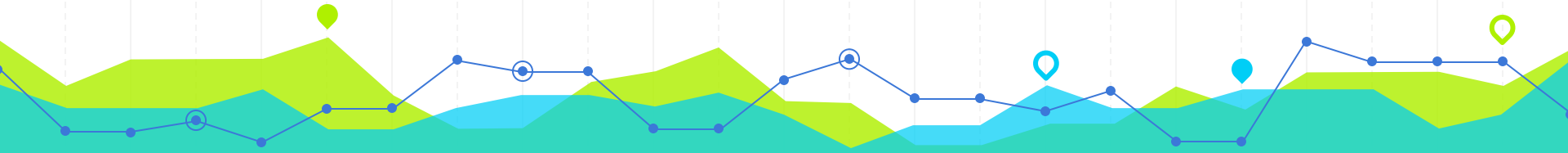
Response Time Prediction

- **Result**

# class	2	4				
Model	**Acc^o**	**Acc^1**	**Acc^2**	**Acc^3**	**Acc^4**	**Acc^o**
CNN	.591	.649	.448	.721	.728	.338
LSTM	.593	.687	.694	.758	.470	.322
CNN+	.878	.681	.614	.692	.919	.500
LSTM+	**.890**	.695	.718	.633	.938	.507

TABLE IV

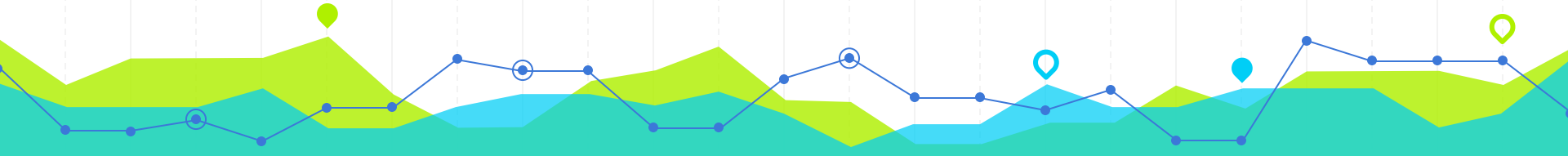
ACCURACY OF RESPONSE TIME PREDICTION. ACC^a STANDS FOR THE OVERALL ACCURACY OF THE RESULT AND ACC^1, ACC^2, ACC^3, ACC^4 STAND FOR THE ACCURACY OF $<$10S, 10-30S, 30-300S, AND $>$300S, RESPECTIVELY.

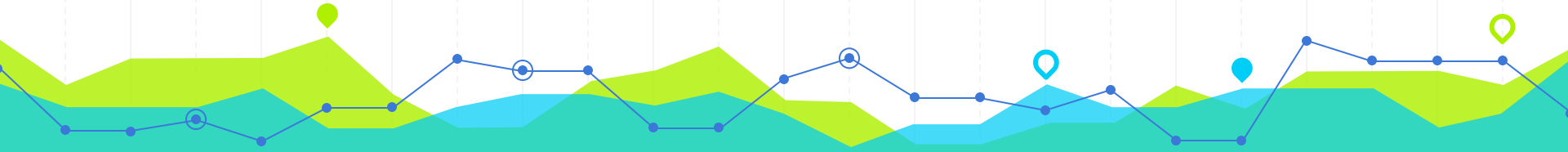


Conclusion 4

Conclusion

- We construct the first **private dataset** and apply **masking mechanism** to protect the private information.
- Two experiments show that the EmotionPush dataset could really help us.
- Please visit our website for the EmotionPush dataset http://academiasinicanlplab.github.io/





THANKS!

Any questions?