Does Emotion Influence the Use of Auto-suggest during Smartphone Typing?

Surjya Ghosh, Kaustubh Hiware, Niloy Ganguly, Bivas Mitra
Indian Institute of Technology Kharagpur

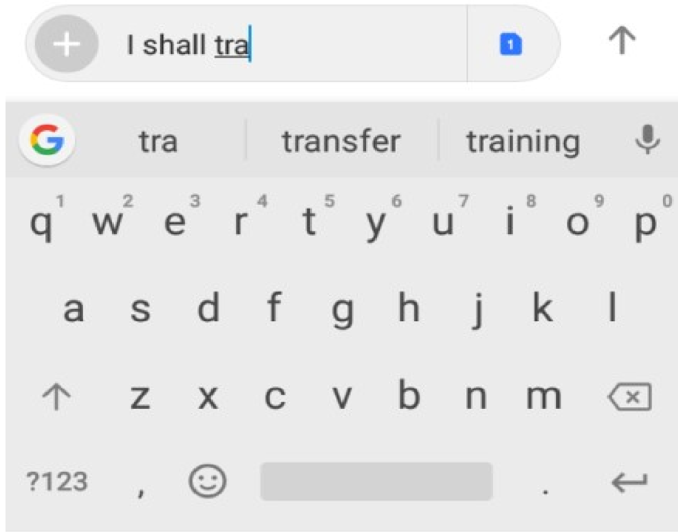
Pradipta De
Georgia Southern University

IUI 2019
Where HCI meets A.I.
Los Angeles, USA
March 16 - 20, 2019

Background

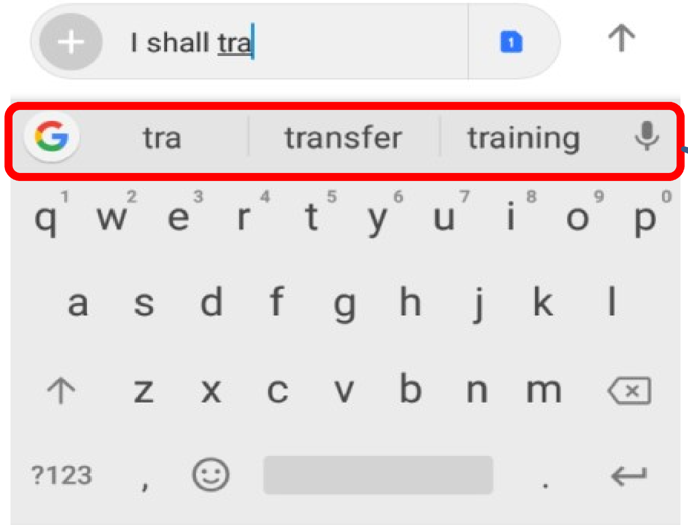
- Text-entry in small touch-based devices
  - relatively inconvenient due to limited space

- To facilitate, different techniques like ***auto-complete, auto-suggest*** are used



Background

- However,
  - reserves fixed space in keyboard layout



- Occupy space in the layout

  - increase cognitive load
    - to read, parse and select correct word [CHI 2016]

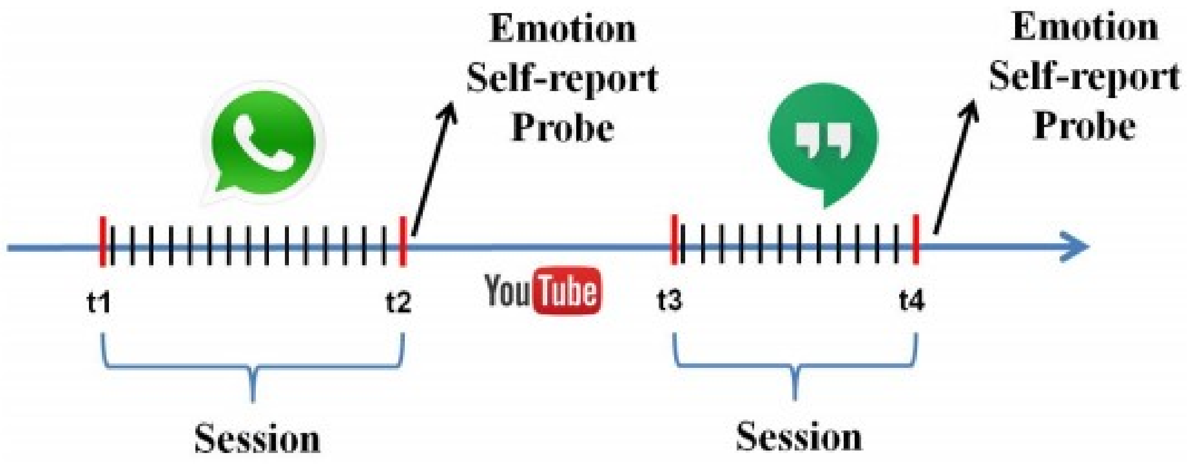
Opportunity to improve the layout by making these techniques adaptive

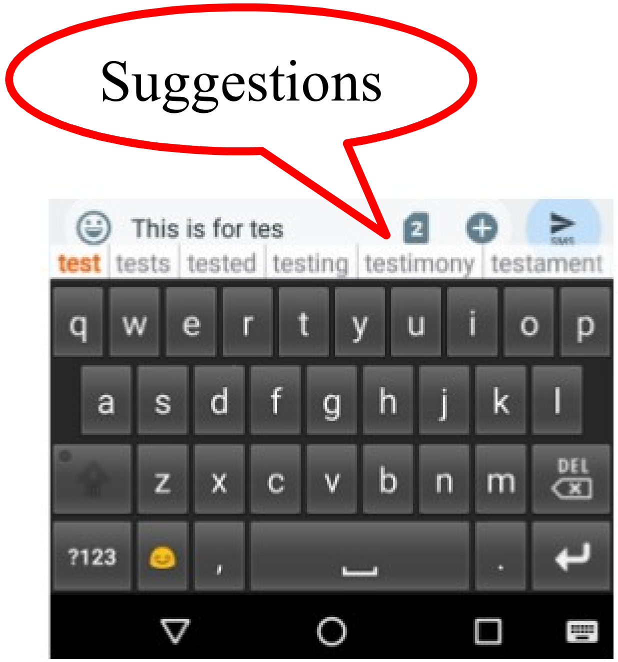
Objective

- Multiple factors can guide auto-suggest usage
  - typing volume
  - application type
  - user emotion
    - as it influences typing behaviour [MobileHCI 2017, ACII 2017]

- Can auto-suggestion usage be determined based on human emotion?

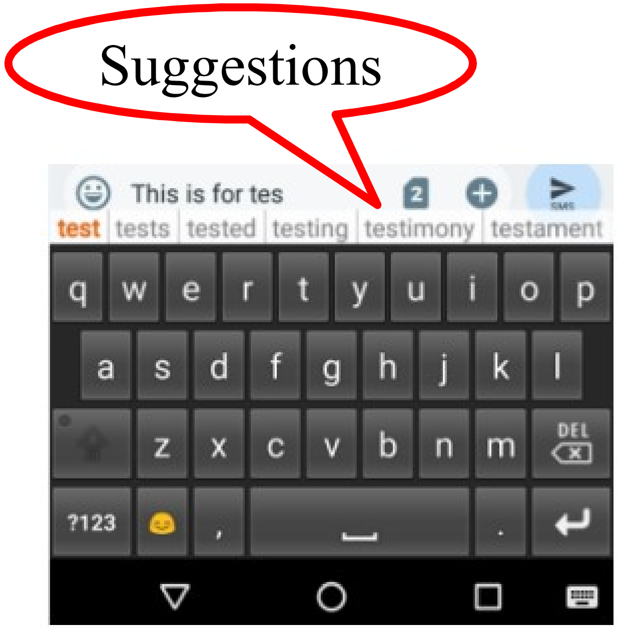
Auto-suggest Usage Scenario



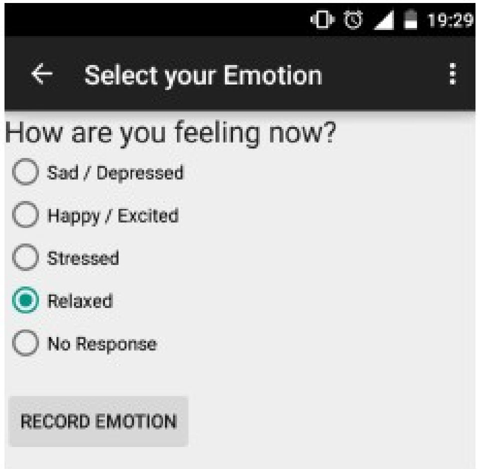


- Auto-suggest usage
  - tracing user typing and providing suggestions
  - labelling auto-suggest usage
    - **accepted** → if at least one suggestion is used in a session
    - **skipped** → if no suggestion is used in a session
  - tracking emotion during typing session

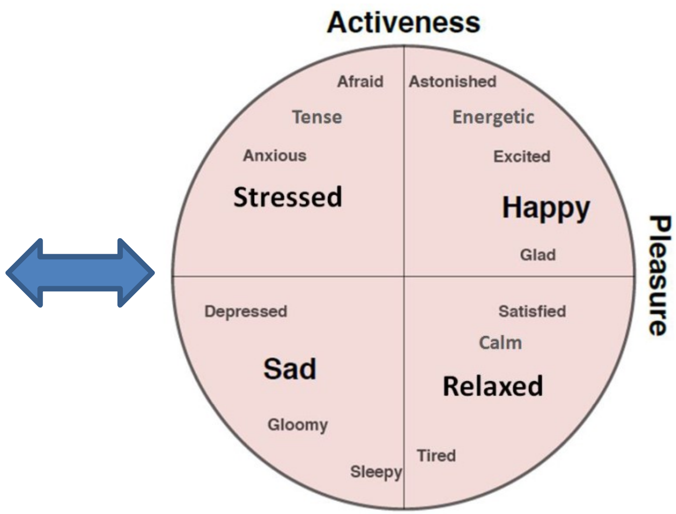
Experiment Apparatus



(a) App keyboard



(b) Emotion collection UI



(c) Circumplex model

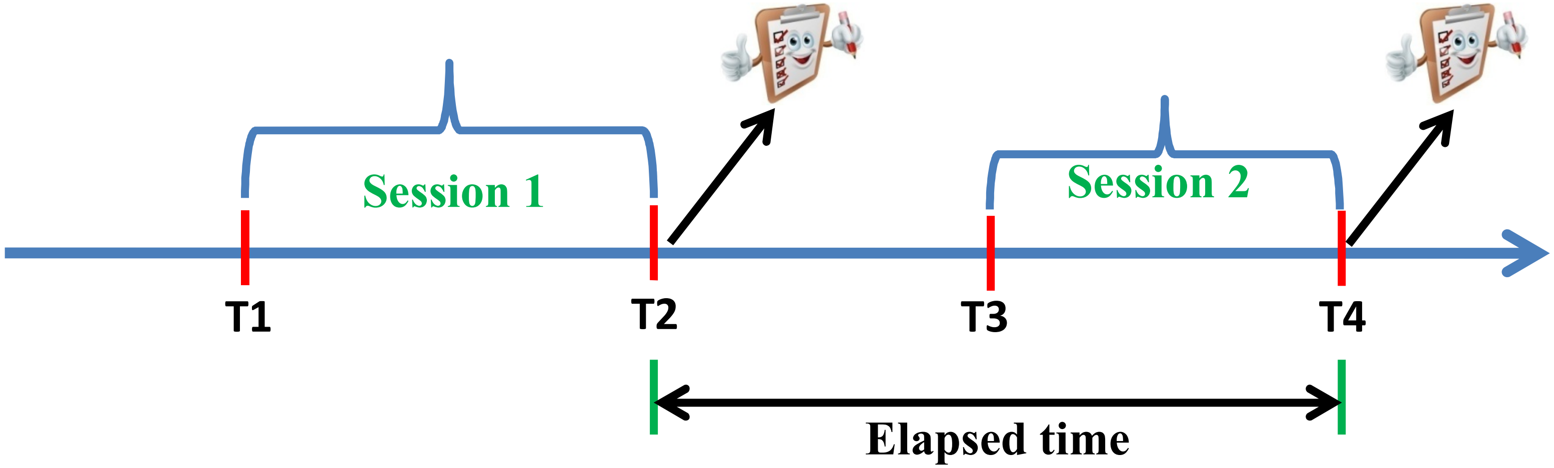
- Custom keyboard
  - tracing user typing and providing suggestions (based on English dictionary)
- Emotion self-report collection UI
  - collects emotions at end of a session
  - four emotions - *happy, sad, stressed, relaxed* (based on Circumplex model)

Auto-suggest Usage Prediction Model

- Auto-suggest usage prediction model
  - Personalized
    - Random Forest
  - Two classes
    - accepted
    - skipped
  - Features
    - Emotion-related features

Feature name	Feature description
Emotion_{curr}	Emotion associated with current session
Emotion_{prev}	Emotion associated with previous session
$\text{Time}_{elapsed}$	Elapsed time between previous and current session emotion recording timestamp

Table 2: Features used for auto-suggest usage prediction

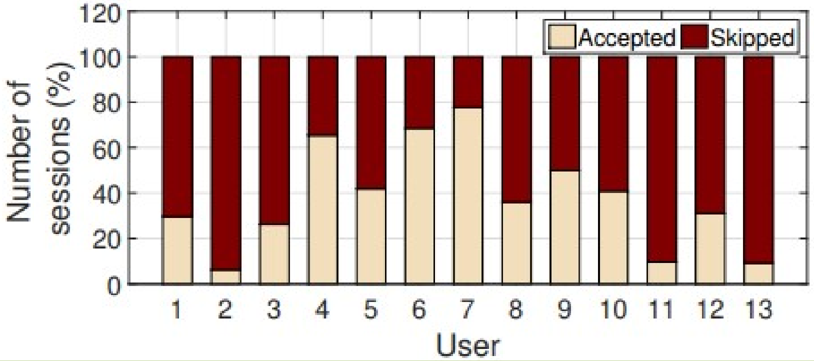


User Study

Study parameters	**Value**
Study duration	3-weeks (in-the-wild)
Total participants	20 (15 M, 5 F) → university students
Age range	20 – 35 years
Installed the app in the participants mobile phones	
Excluded participants	7 (as recorded less than 20 suggestions)
Final participants	**13 (10 M, 3 F)**

Dataset

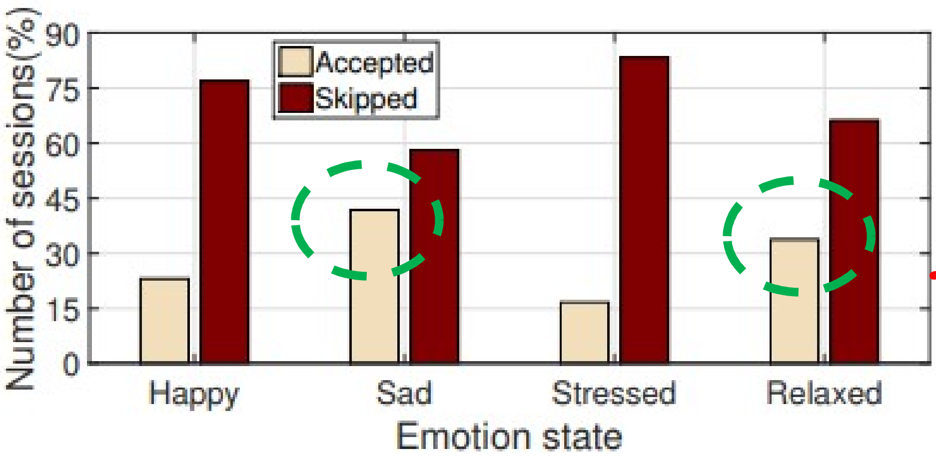
Parameter	Value
Total typing sessions	3,284
No Response sessions	330 (~ 10%)
Effective sessions	2954
Avg. session per user	227 (std dev. 151.7)
Auto-suggest **accepted**	**841 (~28%)**
Auto-suggest **skipped**	**2,113 (~ 72%)**



User-wise auto-suggest usage reveals that most of the users have more skipped session

Role of Emotion on Auto-suggest Usage

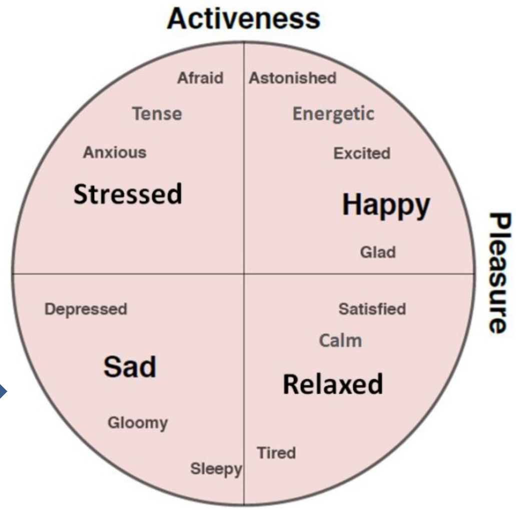
Does auto-suggest usage vary across emotions?



Auto-suggest usage across different emotions for all users

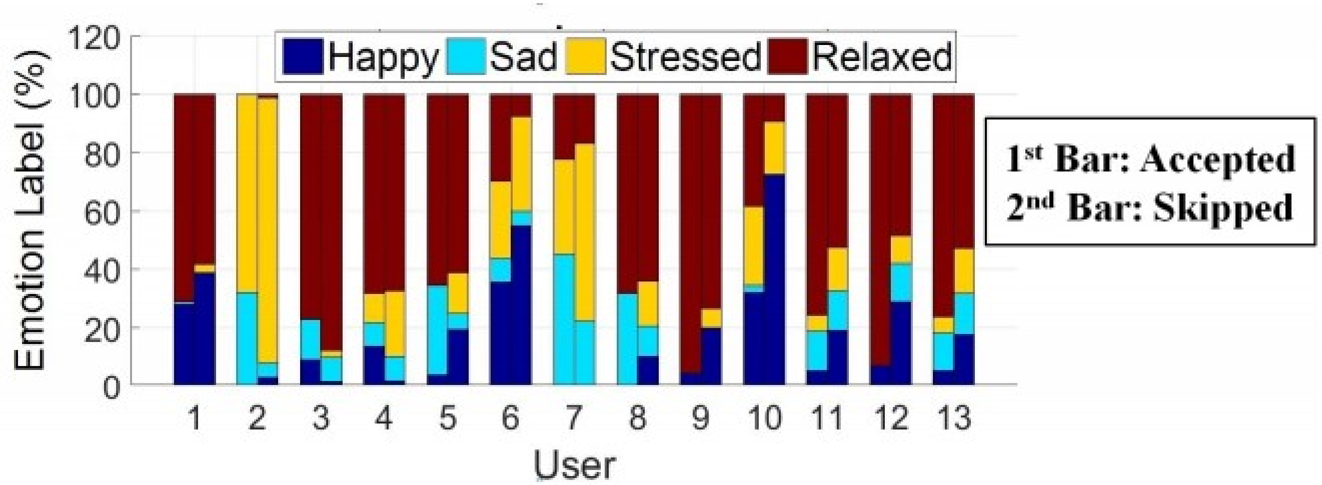
Users more likely to use auto-suggest when **sad or relaxed**

- Relaxed, sad → **Low activity level**
  - May influence more auto-suggest usage



Role of Emotion on Auto-suggest Usage

Does auto-suggest usage vary across individual user emotions?



Auto-suggest usage comparison across different emotions for individual user

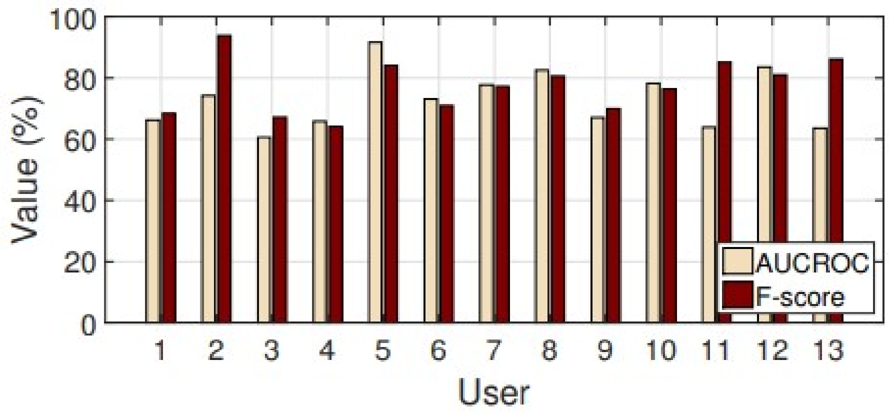
- Majority of the users use auto-suggest in relaxed
- Distribution of emotions in accepted and skipped is statistically significant *(p <0.05)* using chi-square test

Evaluation

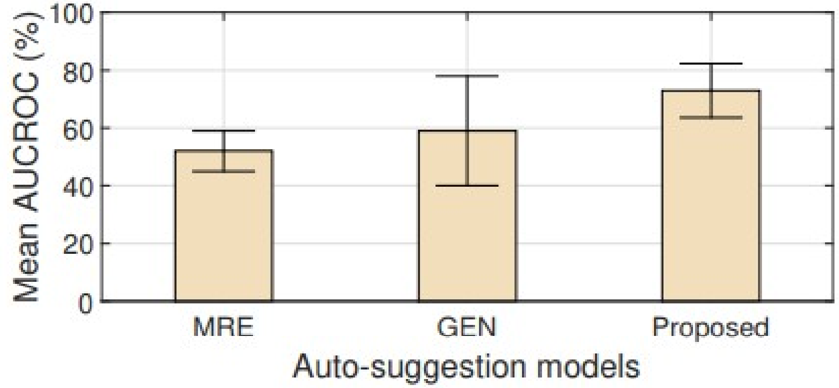
- Experiment setup
  - 10-fold cross validation
  - Metrics: AUCROC, F-score
- Baselines
  - Most-represented Emotion (MRE)
    - Auto-suggest usage → highest in one emotion
    - Personalized model, which determines auto-suggest usage in these emotions
  - Generalized (GEN)
    - Aggregating data from all other users
    - Leave-one-out-cross-validation

Evaluation

- How accurate is the auto-suggest usage prediction model?



(a) Classification performance



(b) Comparison with baselines

Avg. accuracy (AUCROC) – 73% (std dev 9%)
Outperforms the baselines

Evaluation

- How effective are the features to predict auto-suggest usage?

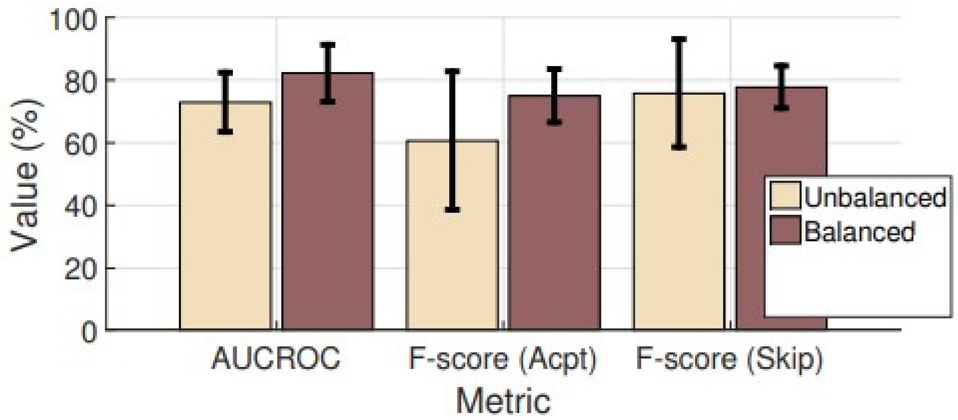
Feature	Rank	Avg. IG
$Emotion_{curr}$	1	0.1194
$Emotion_{prev}$	2	0.1098
$Time_{elapsed}$	3	0.0794

Ranking different features based on InformationGain

Current session emotion is the most discriminating one

Evaluation

- How to improve auto-suggest usage prediction performance?
  - Balancing the dataset (accepted and skipped class)



Classification performance after balancing the dataset

✔✔ Classification performance improves after balancing dataset

– AUCROC – 82%
– F-score (Accepted) – 75%
– F-score (Skipped) – 78%

Conclusion

- Auto-suggest usage is related with human emotion
    - more likely to use in *sad* or *relaxed* state
- Based on emotion, auto-suggest usage can be detected with an accuracy of 82%

Thank You!!

Acknowledgement

- Microsoft Research India
- LRN Foundation India
- IUI student travel grant

http://cse.iitkgp.ac.in/~surjya.ghosh/

Complex Network Research Group (CNeRG)

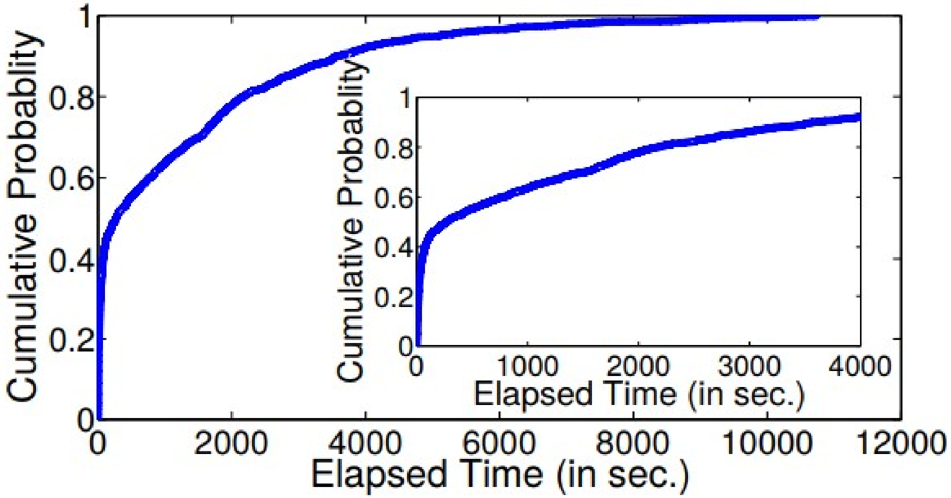
@iitkgpcnerg

@cnerg

http://www.cnergres.iitkgp.ac.in/

Dataset

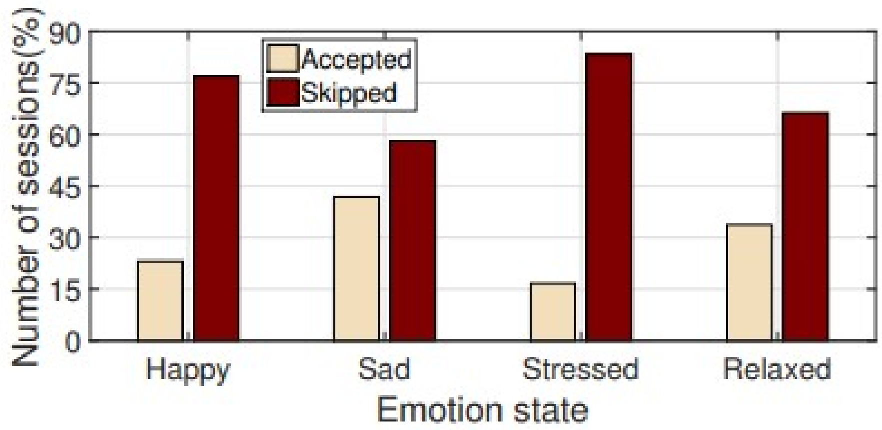
- How close to the typing session, the emotion is collected?



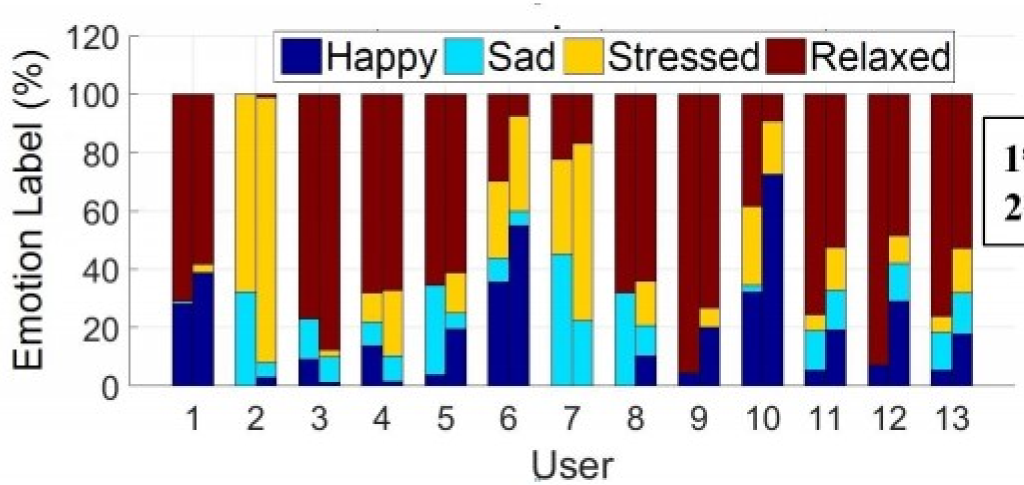
Distribution of elapsed time between typing and emotion recording for all sessions across all users.

- Median elapsed time is less than 5 minutes; 75th and 90th percentile elapsed time is less than 30 minutes and 1 hour respectively.

Dataset



Users more like to use auto-suggest when **sad or relaxed**



1st Bar: Accepted
2nd Bar: Skipped

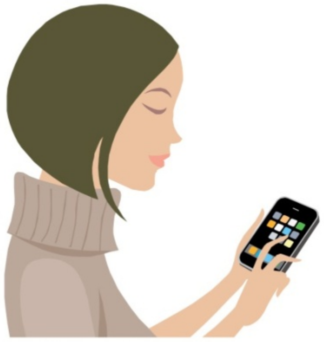
- Majority of the users use auto-suggest in relaxed
- Distribution of emotions in accepted and skipped is statistically significant using chi-square test

Outline

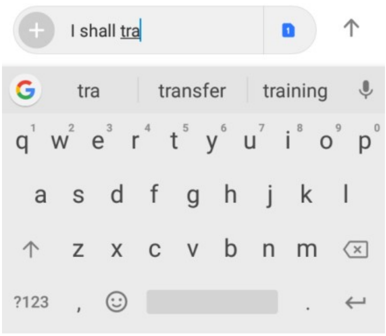
- Auto-suggest Usage Scenario
- Methodology
  - Experiment Apparatus
  - Model Construction
- User Study
  - Field Study
  - Dataset
- Evaluation
- Take-home Points

Background

- Text-entry in small touch-based devices
  - relatively inconvenient due to limited space



- To facilitate, different techniques like ***auto-complete, auto-suggest*** are used
  - However,
    - need additional space
    - increase cognitive load to read, parse and select correct word [CHI 2016]



- Making auto-suggestions adaptive, when the users are more likely to use can overcome these

Thank You!!

Acknowledgement

- Microsoft Research India
- LRN Foundation India
- IUI student travel grant

http://cse.iitkgp.ac.in/~surjya.ghosh/

Complex Network Research Group (CNeRG)

@iitkgpcnerg

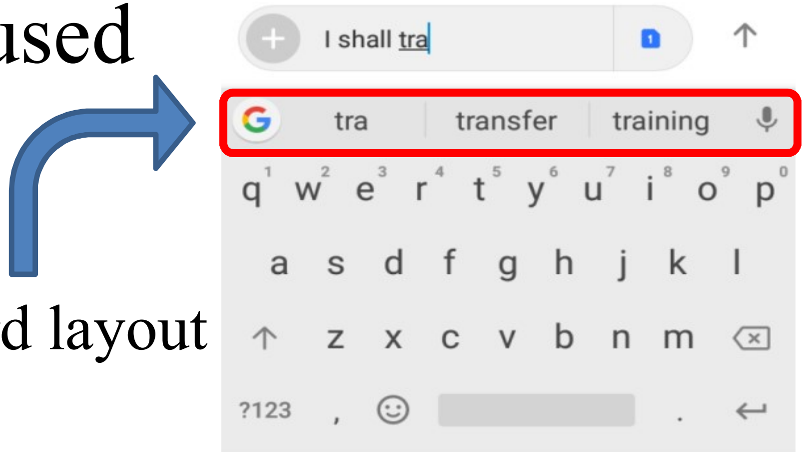
@cnerg

http://www.cnergres.iitkgp.ac.in/

Background

- Text-entry in small touch-based devices
  - relatively inconvenient due to limited space

- To facilitate, different techniques like ***auto-complete, auto-suggest*** are used



  - However,
    - reserves fixed space in keyboard layout
    - increase cognitive load
      - to read, parse and select correct word [CHI 2016]

Opportunity to improve the layout by making these techniques adaptive