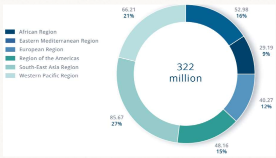
EmoKey: An Emotion-aware Smartphone Keyboard for Mental Health Monitoring

Surjya Ghosh*, Sumit Sahu$, Niloy Ganguly*, Bivas Mitra* & Pradipta De#

*Department of Computer Science & Engineering, Indian Institute of Technology, Kharagpur, India
$Intellicus Technologies Pvt Ltd., India
#Department of Computer Sciences, Georgia Southern University, USA

Introduction

- **Motivation**
  - Approximately320 million people suffering from depressive disorder symptoms across the globe [1]
  - Early diagnosis and counselling can help in great extent [2]
  - Manifestations of depressive symptoms are difficult to track



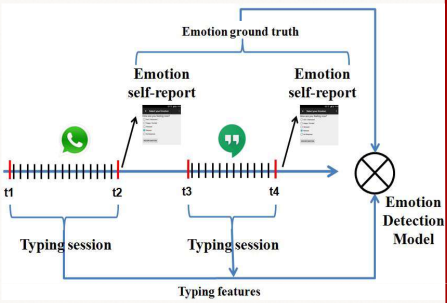
Cases of depressive disorder (millions), by WHO region [1]

- **Background**
  - Sensor-rich ubiquitous smartphones can unobtrusively track interaction pattern
  - Significant portion of interaction is based on text input (WhatsApp, FB Messenger)
  - Typing activity on smartphone carries emotion signature [3]
  - However, such emotion detection models mostly rely on cloud-based services, thereby suffer
    - → Privacy concern
    - → Network delay
- **Problem Statement**
  - Can different emotions be tracked based on typing interaction on the smartphone itself with monitoring capability ?

EmoKey Design

- **Typing-based Emotion Detection Scenario**

- Identify typing session
  - time spent on a single app uninterruptedly
  - extract typing features
- Collect emotion self-reports
  - four emotions based on Circumplex model [4]
  - *happy, sad, stressed, relaxed*
- Construct model to detect emotion
  - combining typing features and self-reports

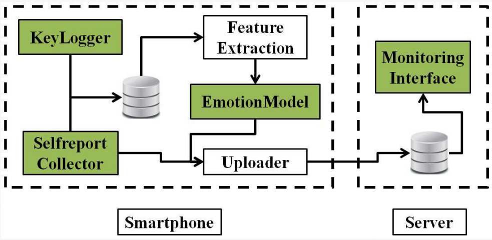


Schematic of typing-based emotion detection

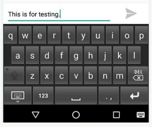
- **Emotion-monitoring interface**

- Store emotion self-reports
  - Self-reports during training
  - Predicted emotions during deployment

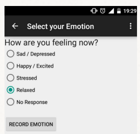
EmoKey Implementation



EmoKey architecture



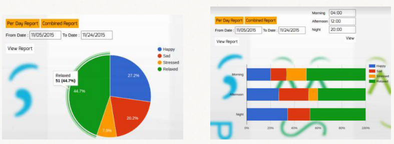
EmoKey keyboard



Emotion self-reporting UI

- Android based application
- On-device, personalized Random Forest based model for emotion detection
- Features

Category	Feature Name
Keystroke Features	Mean Session ITD (MSI)
	Refined Mean Session ITD (RMSI)
	Number of special characters
	Number of backspaces (or delete)
	Session duration
	Session text length
Auxiliary Features	Last ESM Response



Emotion monitoring interface

Field Study and Dataset

- Installed *EmoKey* app in the smartphone of the volunteers for collecting typing details and emotion self-reports
- 22 students (20 male, 2 female, aged between 24-33 years)
- 3-week in-the-wild study

Total typing events	529698
Total typing sessions	2705
Total typing duration	135 Hr.
Per user typing sessions (mean, SD, minimum)	123, 105, 40
Median session duration	98 sec.
Median session length	114

- Total typing sessions: 2705
- Eliminated *No Response* sessions: 2.5%

Final dataset

- Distribution of emotion samples is found to be skewed as users often reported *relaxed* or *stressed* emotion
- *Happy*: 18%, *Sad*: 9%, *Stressed*: 21%, *Relaxed*: 52%

Experiment Setup

- Performed 10-fold cross-validation and measure emotion classification performance
  a) AUCROC *(auc_{wt})* b) F-score

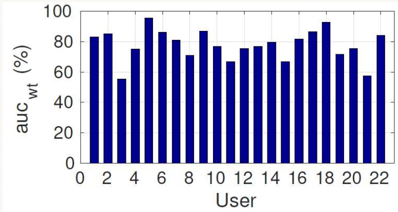
$$auc_{wt} = \sum_{\forall i \in \{happy, sad, stressed, relaxed\}} f_i * auc_i$$

- f_i → fraction of samples for emotion *i*
- auc_i → AUCROC for emotion *i*

- Measured following in terms on training overhead on the device
  a) Training latency b) Battery consumption

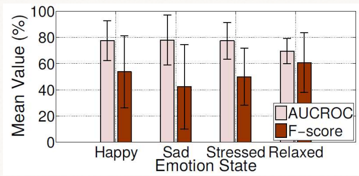
Evaluation: Emotion Classification

- Mean auc_{wt}: 78% (std dev 10%) [Min: 56%, Max: 95%]
- 45% users having AUCROC > 80%
- For all but 2 users, the AUCROC is greater than 70%



User-wise Emotion Detection AUCROC

- All emotions except *relaxed* are identified with an AUCROC greater than 75%
- *Relaxed* emotion is identified with highest F-score (61%)



Emotion-wise AUCROC, F-score

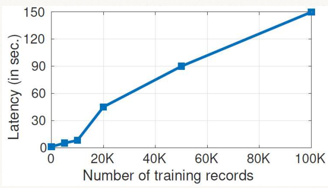
- Prediction result (80 - 20% split) after deployment for one representative user
- Average accuracy: 75%

Model validation for sample user on the monitoring interface

Evaluation: Resource Overhead

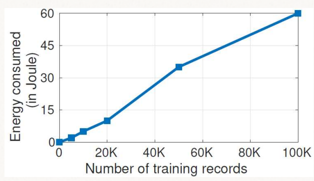
- **Setup**
  - OnePlus X (2.3 GHz quad-core Qualcomm Snapdragon 801 3GB RAM)
  - Synthetically added training records
- **Training Latency**

Latency < 10 secs for 10K records, however increases with high training volume



- **Battery Consumption**

Energy consumption < 5 joules for 10K records, increases significantly with high training volume (> 50K records)



Conclusion

- Design and develop an emotion-aware smartphone keyboard, which detects four emotions*(happy, sad, stressed, relaxed)* based on text input interactions deploying an on-device prediction model
- It returns an average accuracy of 78%, (std. dev. 10%)
- Additionally, provides an interface for mental health monitoring
- Reveals scope for devising efficient on-device models for long-term mental health monitoring

References

[1] World Health Organization et al. Depression and other common mental disorders: global health estimates. 2017.
[2] Lu, Hong, et al. Stresssense: Detecting stress in unconstrained acoustic environments using smartphones. In Proceedings of ACM UbiComp, 2012
[3] Ghosh, S., Ganguly, N., Mitra, B., & De, P. 2017. TapSense: combining self-report patterns and typing characteristics for smartphone based emotion detection. In ACM MobileHCI.
[4] James A Russell. 1980. A circumplex model of affect. Journal of Personality and Social Psychology 39, 6(1980), 1161–1178.

$The work done by the author during his tenure at IIT KGP.

Email: surjya.ghosh@iitkgp.ac.in

Acknowledgement: We thank the "COMSNETS 2019" organizers for the student travel grant.