Term Project

정신 건강 분야에서의 음성 인식 기술 동향 조사 연구

Literature Review on the Trend of Speech Recognition Technology in Mental Health

Dept. of Computer Science & Engineering
Artificial Intelligence and Data Mining Lab(AIDML)
202122029 Meeyun Kim

2021. 11. 25.

Contents

1. Introduction

Diagnosis & Treatment in Mental Health

Diagnosis

정신 질환을 **진단**하여 정확한 병명을 환자에게 부여하는 것.

Treatment

정신 질환의 **종류, 중증도** 등을 **파악**하여 환자에게 가장 적합한 치료를 제공하는 것.

Challenge

- 전문성을 요구하는 **진단 Manual**(HAMD, MADRS, DSM-5 등)을 사용함.

- Therapist들의 **주관**에 따라 병명, 중증도 진단에 차이가 생길 수 있음.

2. Literature Review

Diagnosis Support with Speech Recognition (1)

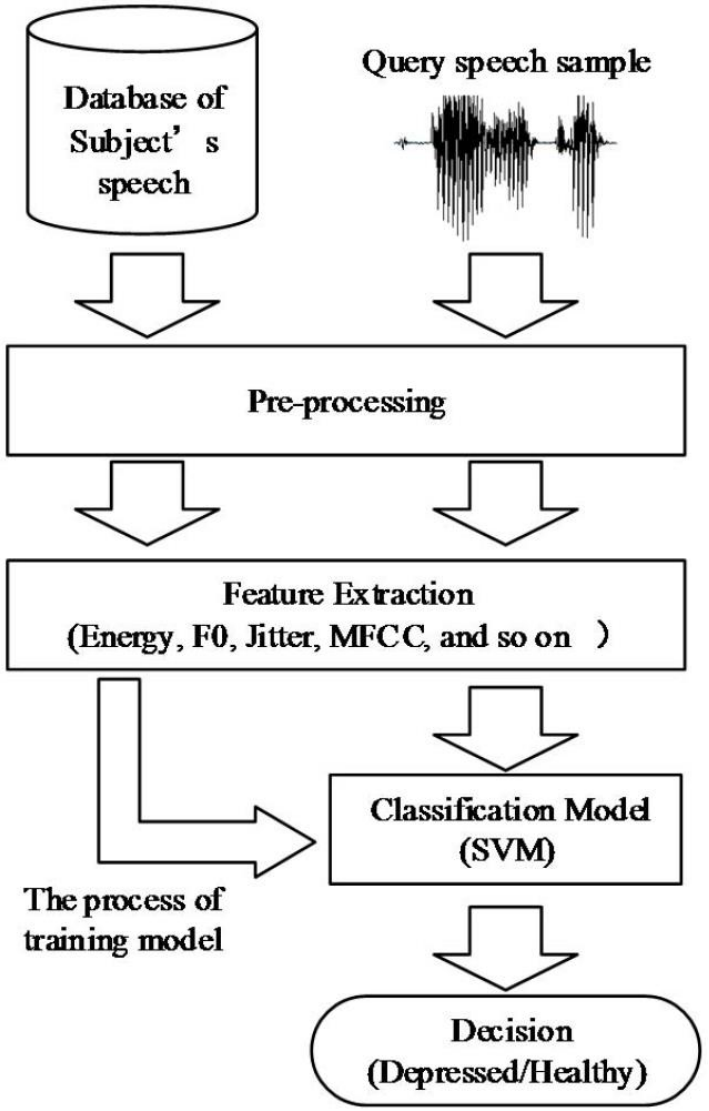
A. Data Collection and Preprocessing

Parameter	Healthy Subjects *HP*		Depressive Subjects *DP*	
	Male	*Female*	*Male*	*Female*
Number	19	18	19	18
Age (years)	Avg[a]:41.8 Dev[b]:10.2	Avg:31.9 Dev:12.9	Avg:34.7 Dev:9.2	Avg:36.7 Dev:12.8
BDI Score	Avg:4.5 Dev:4.0	Avg:5.2 Dev:6.2	Avg:27.1 Dev:9.3	Avg:30.1 Dev:9.4

- **Reading, Interview, Picture Description**
- **Frame size** : 25ms
- **Shift size** : 10ms (using Hamming window)

B. Feature Extraction

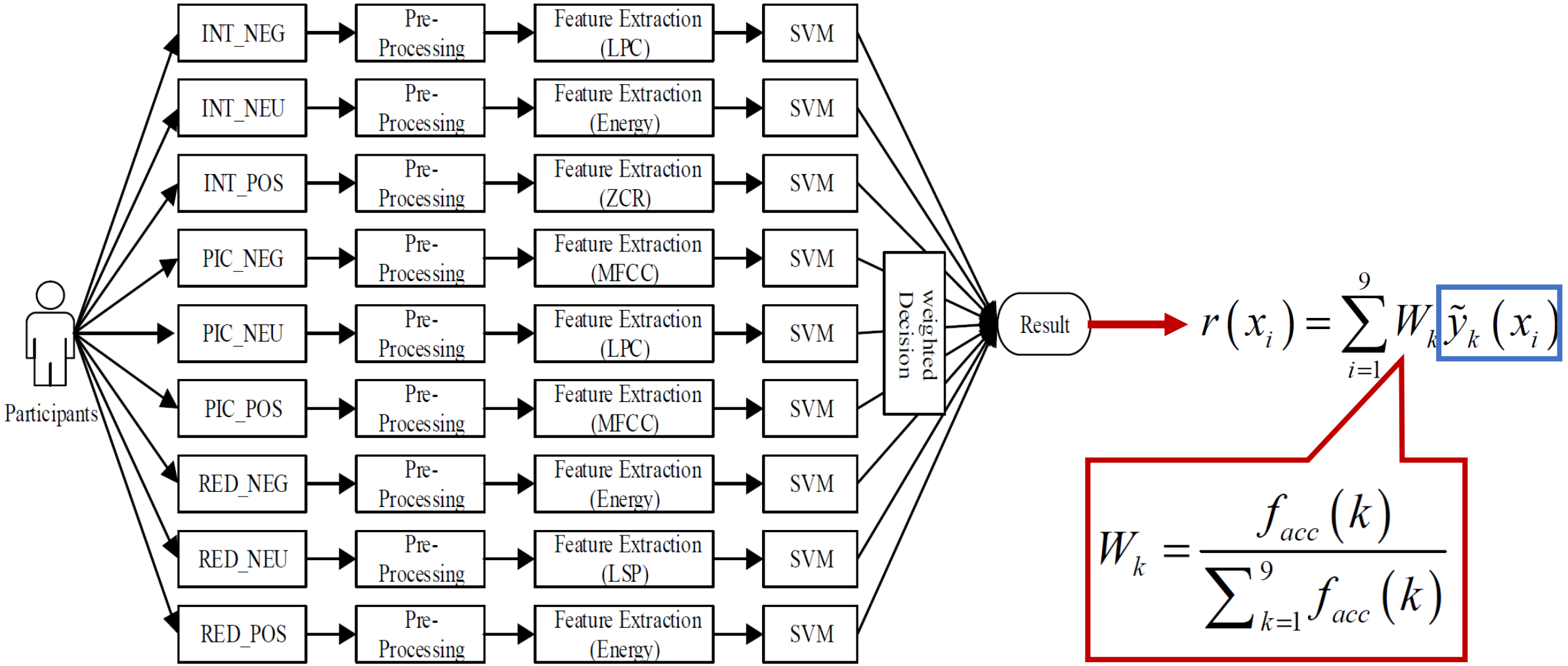
- Only **acoustic features** were extracted.



[IEEE 2017] Detecting Depression in Speech: Comparison and Combination between Different Speech Types

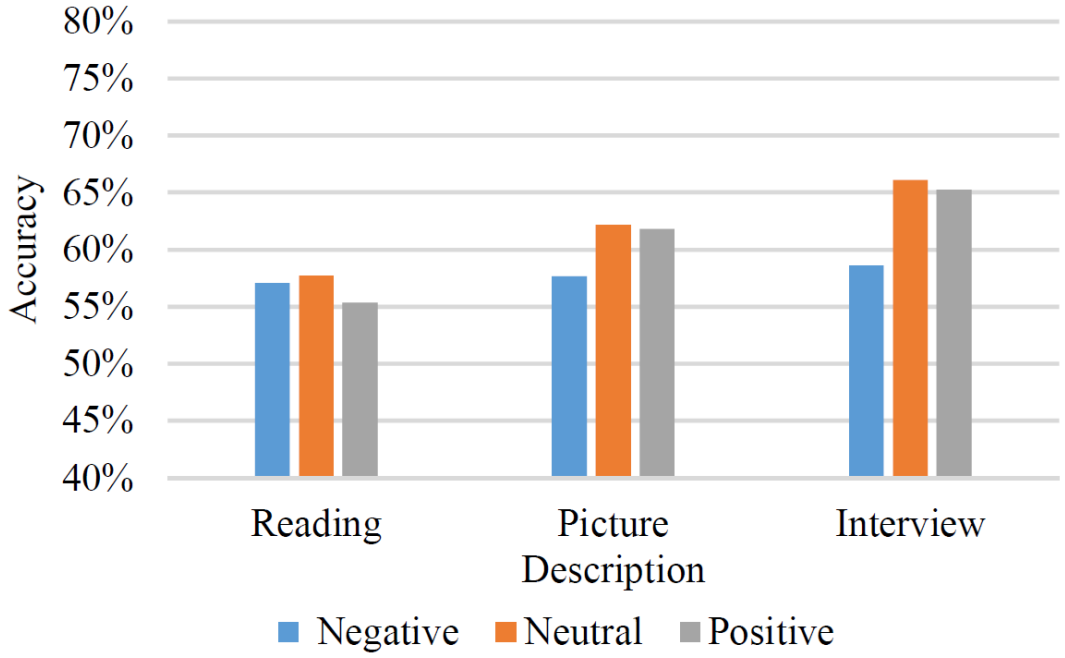
Diagnosis Support with Speech Recognition (1) (Cont'd)

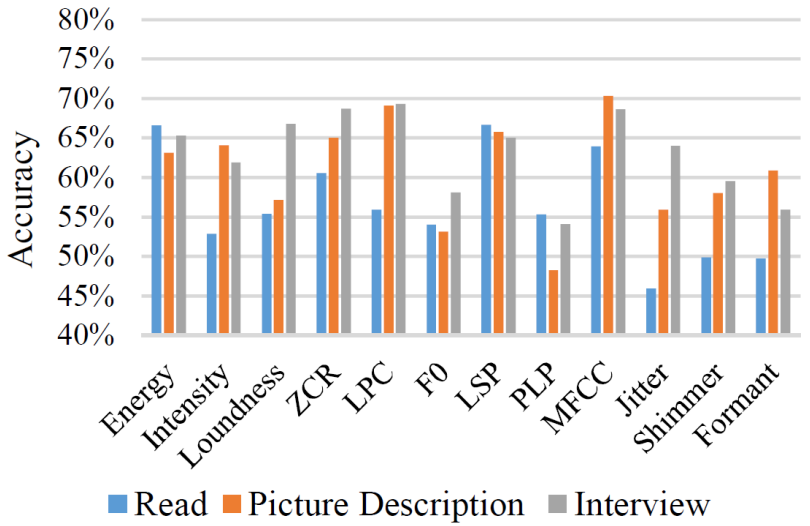
C. New Multiple Classifier System Using Different voice data

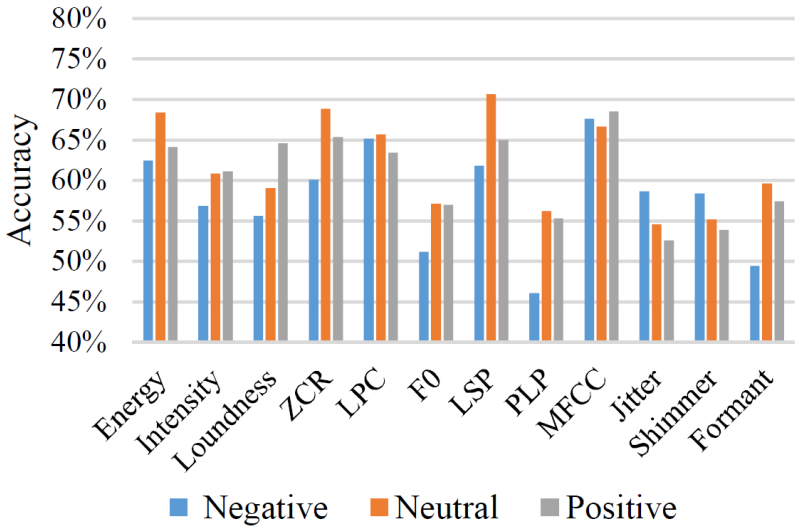


[IEEE 2017] Detecting Depression in Speech: Comparison and Combination between Different Speech Types

Diagnosis Support with Speech Recognition (1) (Cont'd)







Final prediction accuracy : 78.02%

[IEEE 2017] Detecting Depression in Speech: Comparison and Combination between Different Speech Types

Diagnosis Support with Speech Recognition (2)

Table 3. Performance on clinically-relevant utterances by patients.

PHQ	Keywords[a]	Number of positives	True positives	False negatives	False positives	Sensitivity	Positive predictive value
1	Interest, interested, interesting, interests, pleasure	169	127	42	38	75%	77%
2	Depressed, depressing, feeling down, hopeless, miserable	74	63	11	12	85%	84%
3	Asleep, drowsy, sleepiness, sleeping, sleepy	114	85	29	19	75%	82%
4	Energy, tired	143	115	28	22	80%	84%
5	Overeat, overeating	5	3	2	0	60%	100%
6	Bad, badly, poorly	405	336	69	56	83%	86%
7	Mindfulness	11	9	2	0	82%	100%
8	Fidget, fidgety, restless, slow, slowing, slowly	39	28	11	13	72%	68%
9	Dead, death, depression, died, suicide	103	86	17	18	83%	83%
	Weighted average	1063	852	211	178	80%	83%

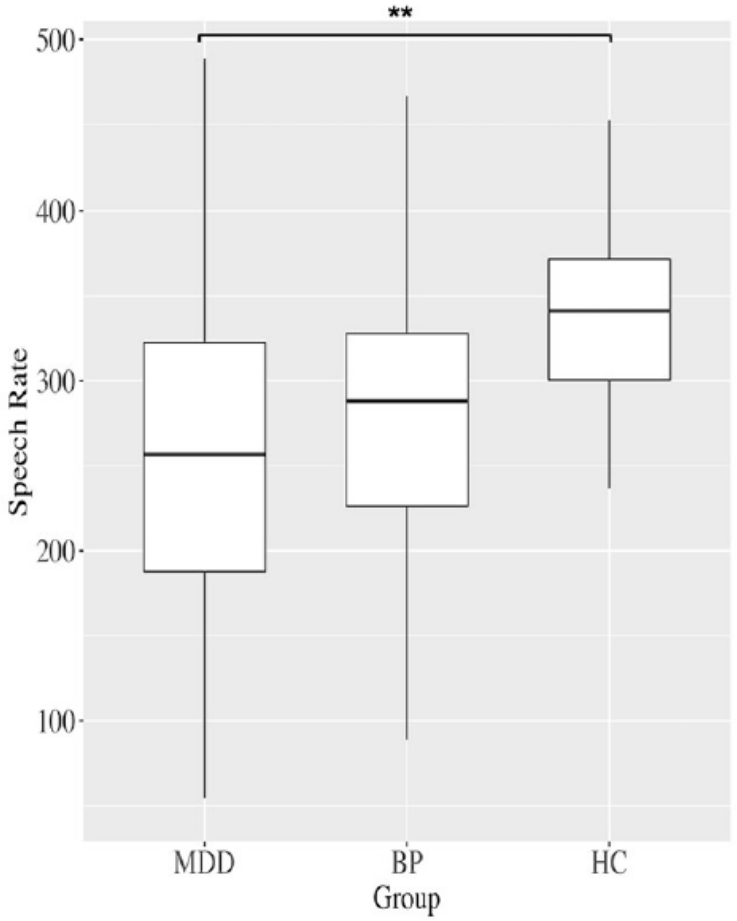
[a]For each question of the Patient Health Questionnaire (PHQ-9), relevant keywords were identified by querying the Unified Medical Language System using each PHQ question to generate search terms. Each table row denotes a different question from the PHQ-9. Number of occurrences refer to how often the keywords appear in our transcribed therapy sessions. True positives refer to a correct transcription by the automatic speech recognition system. False negatives and false positives denote incorrect transcriptions. Sample size is denoted by the number of positives.

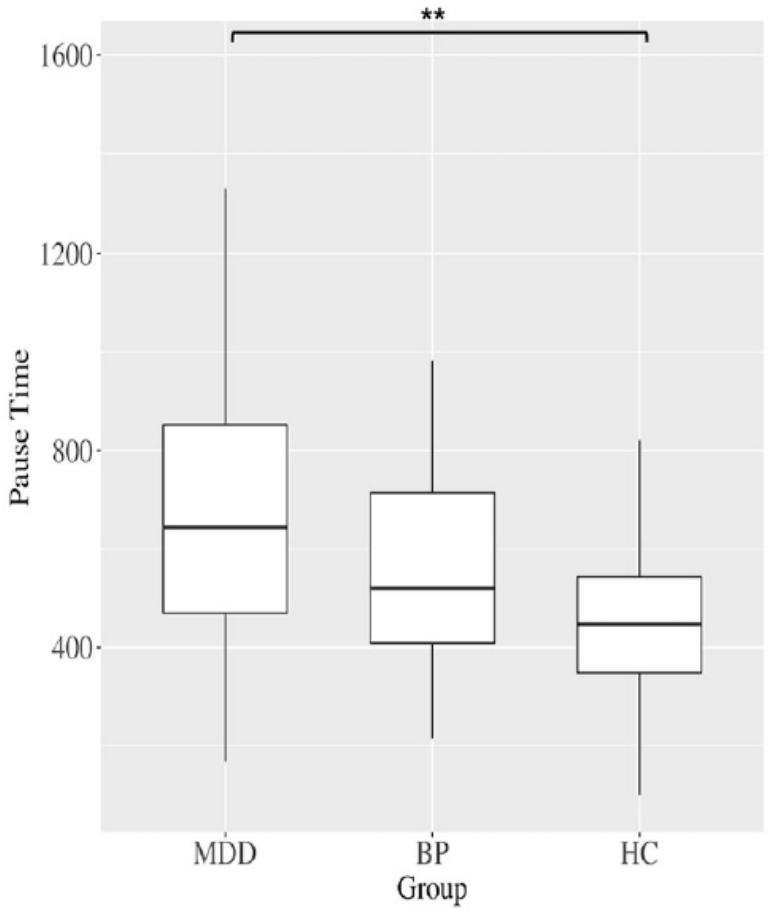
https://www.mdcalc.com/phq-9-patient-health-questionnaire-9

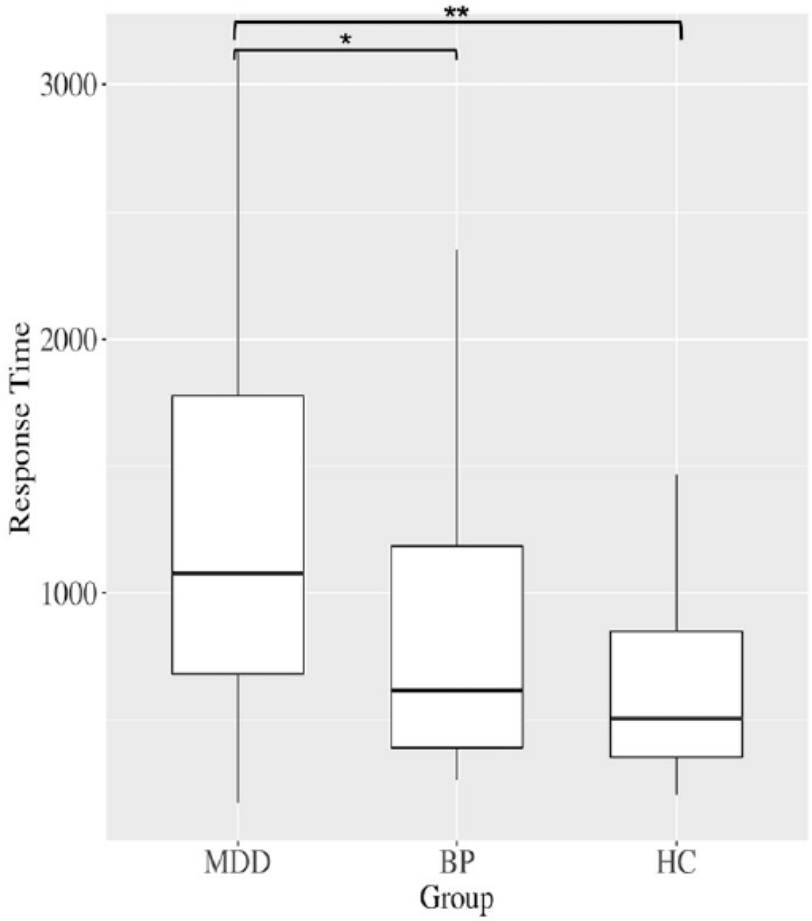
[NPJ Digit. Med. 2020] Assessing the accuracy of automatic speech recognition for psychotherapy

Treatment Support with Speech Recognition

- 우울감과 **speech rate**는 **반비례**, **pause time, response time**은 **비례**한다는 연구.







MDD = major depressive disorder, **BP** = bipolar disorder, **HC** = healthy controls

[PLOS ONE 2020] Using speech recognition technology to investigate the association between timing-related speech features and depression severity

Psychiatrist Experience with Speech Recognition

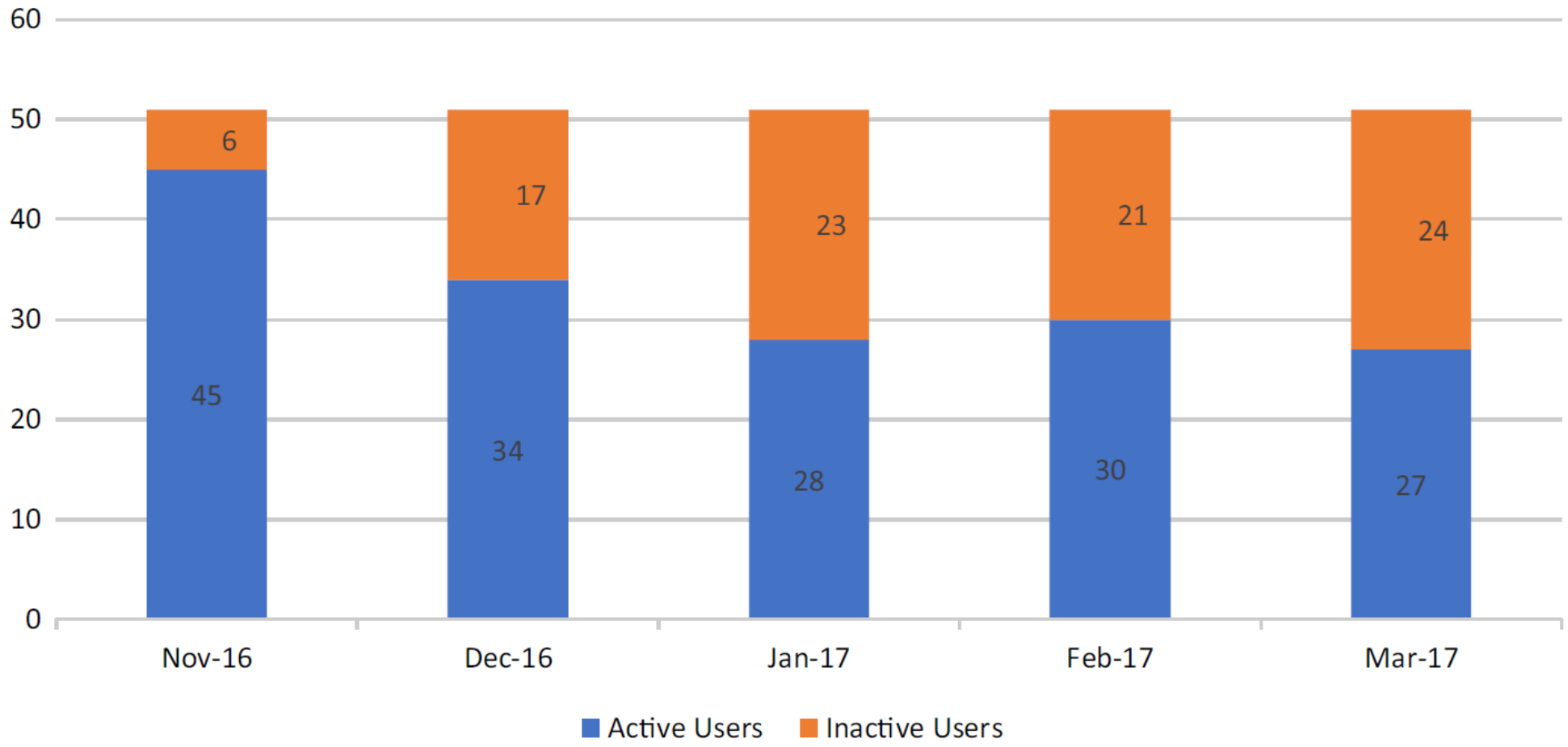


Fig. 1 Active and inactive physician users. The number of physician users is on the y-axis. Month is on the x-axis

[BMC Res Notes 2018] Physician experience with speech recognition software in psychiatry: Usage and perspective

3. Future work

A. Online 진료, 화상 및 전화 진료 등으로의 확장

- 다른 생체 지표를 확인할 수 없는 비대면 음성 혹은 화상 진료에서의 유용성 기대.

B. 청소년 상담, 자살 예방 상담 전화 등으로의 확장

- Mental health 비전문가의 상담 과정을 보조.

C. 다양한 Mental health disorder Diagnosis, Treatment에 응용

- Depression diagnosis 외의 분야로 확장될 것을 기대.

References

- Long H, Guo Z, Wu X, et al. Detecting depression in speech: comparison and combination between different speech types. IEEE., 1052–1058 (2017).

- M. Yamamoto et al. Using Speech Recognition Technology to Investigate the Association between Timing-related Speech Features and Depression Severity, Plos one, vol. 15, no. 9, pp. e0238726 (2020).

- Miner, A. S. et al. Assessing the accuracy of automatic speech recognition for psychotherapy, NPJ Digit. Med. 3, pp. 82 (2020).

- Fernandes J. et al. Physician experience with speech recognition software in psychiatry: Usage and perspective, BMC Res Notes[Internet], 11(1), pp. 15 (2018).