Automatic Text-Based Estimation of Depression Symptoms

Kirill Milintsevich, Kairit Sirts, Gaël Dias
kirill.milintsevich@unicaen.fr, kairit.sirts@ut.ee, gael.dias@unicaen.fr

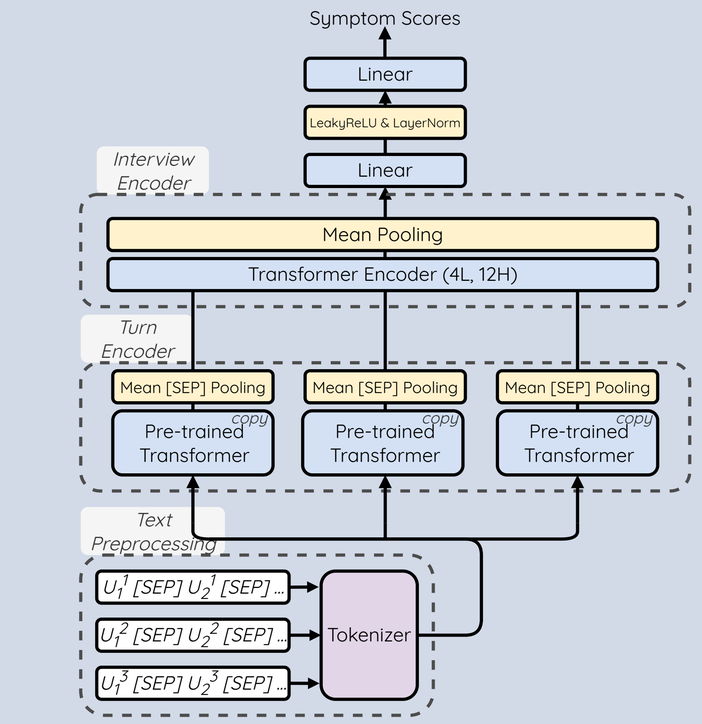
Introduction

- Major Depressive Disorder (MDD) affects one's linguistic footprint, which is reflected by subtle changes in speech production.
- Most of the MDD classifiers disregard symptoms and focus only on binary classification.
- DAIC-WOZ patient-psychiatrist interview dataset is used for training.
- Several emotion and sentiment lexicons are used to introduce external knowledge to the neural models.
- **Symptom network analysis** (SNA) approach views mental health disorders as a causal system of interacting symptoms.

Experimental Setup

- DAIC-WOZ dataset contains 189 clinical interviews, each accompanied with a PHQ-8 assessment - eight questions about symptoms.
- The eight symptoms are: S1 (lack of interest), S2 (feeling down), S3 (sleeping disorder), S4 (lack of energy), S5 (eating disorder), S6 (low self-esteem), S7 (concentration problem), S8 (hyper/lower activity).
- Three lexicons are used to mark important words directly in the input with `@` token from both sides: SDD (depression), AFINN (sentiment), and NRC (sentiment and emotion).
- Two pre-trained models are used: BERT-Base (BERT) and MentalBERT-Base (mBERT).

Architecture



Results

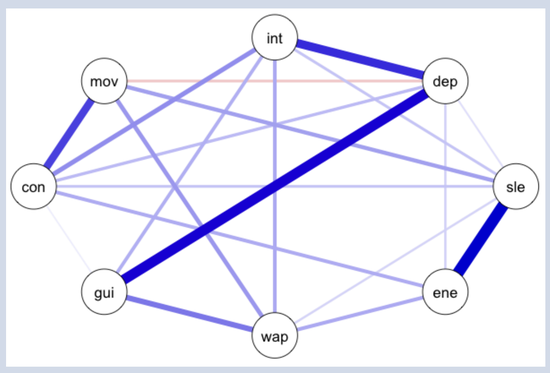


Figure 1: Correlation graph of symptoms computed on the training set of the DAIC-WOZ data set. Thicker edges show a stronger correlation. Blue edges show a positive correlation, and red edges show a negative correlation.

Model	S1	S2	S3	S4	S5	S6	S7	S8	PHQ-8
BERT	$0.56_{\pm.05}$	$\mathbf{0.63_{\pm.02}}$	$0.77_{\pm.05}$	$0.87_{\pm.04}$	$\mathbf{0.81_{\pm.03}}$	$0.78_{\pm.06}$	$0.74_{\pm.01}$	$0.34_{\pm.01}$	$4.38_{\pm.21}$
+SDD	$0.70_{\pm.02}$	$0.88_{\pm.05}$	$0.94_{\pm.05}$	$0.94_{\pm.04}$	$1.00_{\pm.07}$	$0.97_{\pm.04}$	$0.87_{\pm.02}$	$0.34_{\pm.00}$	$5.60_{\pm.18}$
+AFINN	$\mathbf{0.50_{\pm.03}}$	$0.70_{\pm.03}$	$0.79_{\pm.03}$	$0.81_{\pm.04}$	$0.85_{\pm.03}$	$0.72_{\pm.02}$	$0.77_{\pm.02}$	$0.34_{\pm.00}$	$4.56_{\pm.22}$
+NRC	$\mathbf{0.50_{\pm.03}}$	$0.66_{\pm.05}$	$\mathbf{0.73_{\pm.05}}$	$0.77_{\pm.03}$	$\mathbf{0.81_{\pm.05}}$	$0.71_{\pm.07}$	$\mathbf{0.73_{\pm.05}}$	$0.34_{\pm.00}$	$\mathbf{4.31_{\pm.18}}$
+ALL	$\mathbf{0.50_{\pm.04}}$	$0.69_{\pm.03}$	$0.81_{\pm.12}$	$\mathbf{0.74_{\pm.06}}$	$\mathbf{0.81_{\pm.07}}$	$\mathbf{0.69_{\pm.05}}$	$0.74_{\pm.03}$	$0.34_{\pm.00}$	$4.56_{\pm.42}$
mBERT	$0.59_{\pm.02}$	$0.64_{\pm.06}$	$0.91_{\pm.05}$	$0.92_{\pm.04}$	$0.89_{\pm.04}$	$0.71_{\pm.02}$	$0.71_{\pm.04}$	$0.35_{\pm.01}$	$4.71_{\pm.23}$
+SDD	$0.69_{\pm.07}$	$0.72_{\pm.08}$	$0.89_{\pm.07}$	$0.92_{\pm.02}$	$0.93_{\pm.07}$	$0.85_{\pm.07}$	$0.78_{\pm.06}$	$0.34_{\pm.00}$	$5.07_{\pm.38}$
+AFINN	$0.48_{\pm.04}$	$0.62_{\pm.02}$	$0.71_{\pm.05}$	$0.78_{\pm.04}$	$0.79_{\pm.03}$	$0.70_{\pm.03}$	$0.74_{\pm.03}$	$0.34_{\pm.00}$	$4.27_{\pm.22}$
+NRC	$0.60_{\pm.05}$	$0.68_{\pm.03}$	$0.71_{\pm.05}$	$0.78_{\pm.04}$	$0.80_{\pm.08}$	$0.74_{\pm.02}$	$0.71_{\pm.05}$	$0.34_{\pm.00}$	$4.35_{\pm.26}$
+ALL	$\mathbf{0.44_{\pm.06}}$	$\mathbf{0.55_{\pm.04}}$	$\mathbf{0.63_{\pm.06}}$	$\mathbf{0.72_{\pm.07}}$	$\mathbf{0.69_{\pm.03}}$	$\mathbf{0.67_{\pm.04}}$	$\mathbf{0.67_{\pm.03}}$	$0.34_{\pm.00}$	$\mathbf{3.59_{\pm.31}}$
SOTA	$0.53_{\pm.05}$	$\mathbf{0.55_{\pm.03}}$	$0.75_{\pm.07}$	$\mathbf{0.64_{\pm.03}}$	$0.81_{\pm.05}$	$\mathbf{0.62_{\pm.02}}$	$0.83_{\pm.04}$	$0.44_{\pm.02}$	$3.78_{\pm.13}$

Table 1: Results for the DAIC-WOZ test set. The mean MAE and standard deviation are reported for five runs. `S[1-8]` correspond to the symptoms from the PHQ-8 questionnaire. The best MAE for each symptom is **in bold**. SOTA means current state-of-the-art results in the literature [1].

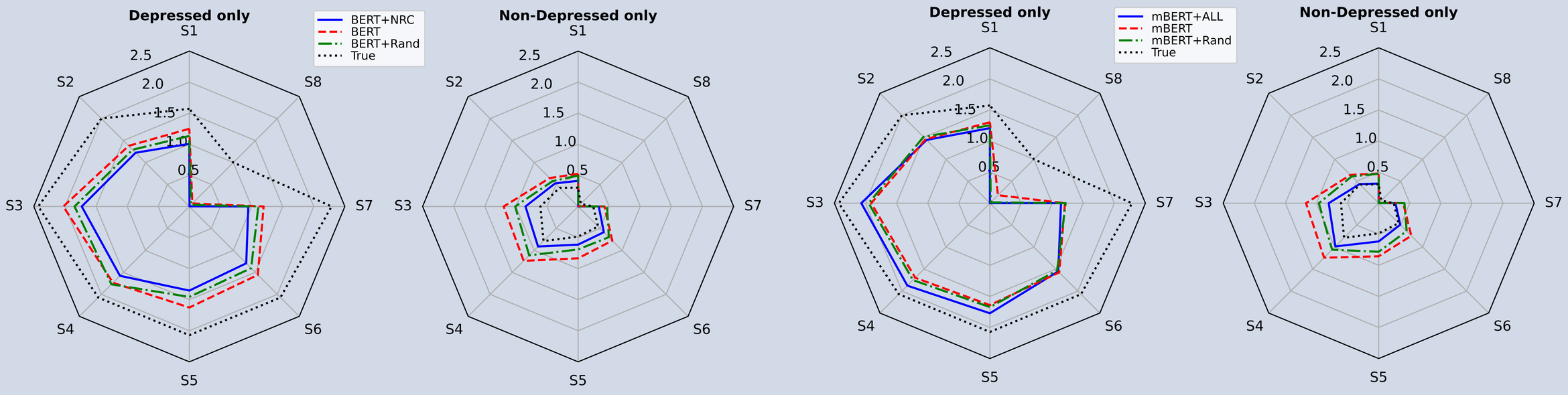


Figure 2: Average predicted values for depressed and non-depressed patients of the DAIC-WOZ test set. +RAND specifies a model with random input marking.

Conclusion

The set of endeavoured experiments shows that introducing sentimental, emotional and/or domain-specific lexicons can correlate with overall performance improvement if adapted to the targeted task. Moreover, the model produces a symptom correlation graph that is structurally identical to the real one. This novel method provides more in-depth information about the depressive condition by focusing on the individual symptoms rather than a general binary diagnosis.

Acknowledgements

This research is part of a joint PhD thesis in cooperation with the University of Tartu, University of Caen, CNRS, ENSICAEN, CHU de Caen.

References

[1] Kirill Milintsevich, Kairit Sirts, and Gaël Dias. Towards automatic text-based estimation of depression through symptom prediction. *Brain Informatics*, 10(1):1–14, 2023.

[2] Stephanie Rodgers et al. Symptom-based subtypes of depression and their psychosocial correlates: a person-centered approach focusing on the influence of sex. *Journal of Affective Disorders*, 156:92–103, 2014.

[3] Claudia van Borkulo et al. Association of symptom network structure with the course of depression. *JAMA psychiatry*, 72(12):1219–1226, 2015.