



# Detecting depression: how to have a happier campus

Diogo Braganca – braganca@stanford.edu  
Department of Physics, Stanford University

## Motivation and summary

- Depression is becoming more and more a serious problem in public health;
- Rate of depression among graduate students is much higher than the average population;
- Being able to detect it early could be of major importance for the well being at universities;
- We built and compared 4 models to detect depression using the DAIC-WOZ dataset;
- We found that it is easier to detect depression with the transcripts than with the audio clips.

## Data and Features

- The DAIC-WOZ dataset contains audio and video information about 189 interviews;
- It contains audio features previously extracted;
- Labels are given by the result of the PHQ-8 test that assesses the level of depression;
- Divided in levels: none, mild, moderate, moderately severe, severe;
- We only used the transcripts and applied a short-time Fourier transform to the audio;
- For the transcript, we used Word2Vec encoding with GloVe.

## Models


- **Transcripts**
  - Model 1: LSTM + (2 x dense)
  - Model 2: (2 x LSTM) + dense
- **Audio**
  - Model 3: 1D Conv + (2 x GRU)
  - Model 4: (2 x 2D Conv) + dense


## Discussion

- Model 1 better than Model 2
- Intriguing that Dev and Test acc. are higher than Train.
- Model 1 can predict any sentence!
- Audio models were not learning, they predicted the same outcome.
- Model 3 stuck on "Mod. severe" while Model 4 stuck on "Mild"
- Next steps: increase training set for audio models and deepen architecture, gather more "real" data for transcript models

## Results

Model 1	None	Mild	Moderate	Mod. Severe	Severe
None	3533	19	21	8	12
Mild	8	2881	23	7	15
Moderate	8	9	3084	5	18
Mod. Severe	12	11	34	2770	16
Severe	7	6	5	5	2398

I am a graduate student  none

I am getting married  severe

I detest my horrible job  moderate

Model 4	None	Mild	Moderate	Mod. Severe	Severe
None	0	40	0	0	1
Mild	0	111	0	0	0
Moderate	0	73	0	0	0
Mod. Severe	0	38	0	0	0
Severe	0	35	0	0	0

	Train acc.	Dev acc.	Test acc.
Model 1	0.92	0.98	0.98
Model 2	0.78	0.91	0.91
Model 3	0.50	0.14	0.15
Model 4	0.37	0.37	0.37

## References

- [1] X. Ma *et al.*, DepAudioNet: An Efficient Deep Model for Audio based Depression Classification, Proceedings of the 6th International Workshop on Audio/Visual Emotion Challenge, 35 (2016).
- [2] J. Pennington *et al.*, GloVe: Global Vectors for Word Representation (2014).
- [3] M. Valstar *et al.*, AVEC 2016 – Depression, Mood, and Emotion Recognition Workshop and Challenge, CoRR (2016).