Anxiety Disorder Prediction from Virtual Reality Head Movements

Sarah Ciresi, John Hewitt, Cooper Raterink

Overview

This project uses VR-based head movement data to predict presence of an anxiety disorder. Researchers have established a relationship between type of head movements and mental illness [1]. We use votes from an ensemble of machine learners to predict anxiety in patients based on hand-selected expert features, and compare the results to the predictions of a deep convolutional network.

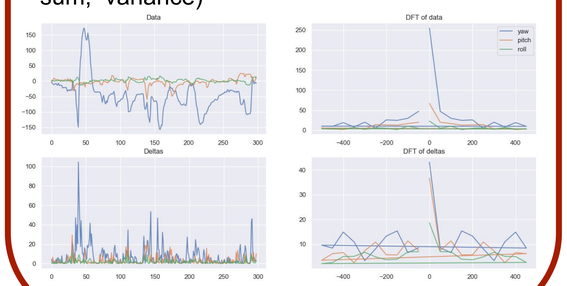
Data

Andrea Goldstein, from Stanford's Williams PANLab, advised us on this project and provided VR head movement data gathered during the ENGAGE study. Participants' head movements were tracked while experiencing virtual reality environments including positive experiences, negative experiences, and calm experiences. For each participant, we have a binary label as to whether they are judged as having high anxiety according to a quantitative metric called GAD7 [3].

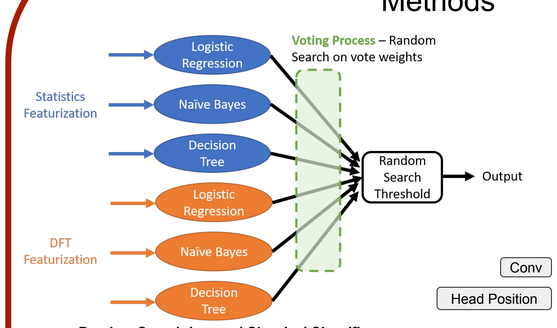
The train/dev data has 118 samples, and the test data had 30.

Features

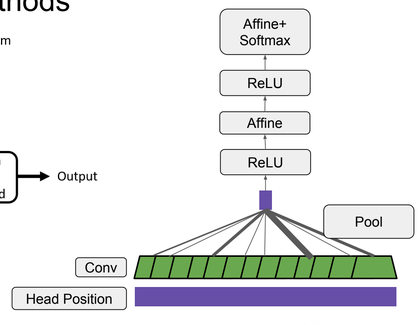
We compute two separate featurizations on both the time-series data and the one-step differences: (1) a 30-point discrete Fourier Transform, and (2) summary statistics (mean, sum, variance)



Methods



Random Search-Learned Classical Classifier Ensemble

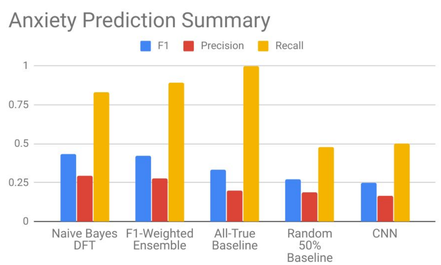


1D Head Movement CNN

As we experimented more and more with classifiers, we realized our results were extremely noisy, and were suffering from high variance. So, we decided to aggregate the outputs of the models we had been using into a ensemble voting system. Instead of naively using uniform weights we do a random search [4] on the voter weights as well as the decision threshold. We conducted this search and all hyperparameter choices using hold-one-out validation on data with disjoint participants from our test participants.

We compared this random search-learned classifier model with a Convolutional Neural Network with five layers: a convolutional layer, an average pooling layer, a ReLU activation layer, and two dense layers. We also used dropout for regularization. The number of output filters, kernel size, and dropout rate were tuned to find the set of hyperparameters that yielded the highest precision, recall, and F1 scores.

Results



CNN Hyperparameter Search, Development Set

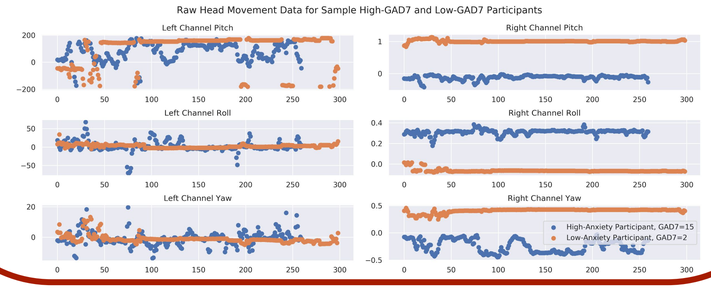
Filter Count	Kernel Size	Dropout Rate	F-score	Precision	Recall
16	5	0.5	22.7	17.9	31.3
10	5	0.3	23.8	17.9	35.7
32	10	0.5	21.4	**42.9**	28.7
16	10	0.3	**38.3**	32.1	**47.4**
32	20	0.3	16.0	11.0	30.1

Anxiety Prediction Model Comparison

Predictor	F1	Precision	Recall
LogReg - summary	23.5	18.2	33.3
D Tree - summary	35.3	25.6	57.3
N Bayes - summary	37.0	23.8	83.3
LogReg - DFT	33.3	25.0	50.0
D Tree - DFT	26.2	19.5	40.3
N Bayes - DFT	**43.5**	**29.4**	83.3
Uniform Weight Voter Ensemble	35.7	27.8	50.0
F1 Ensemble	42.3	27.8	89.0
Precision Ensemble	9.3	12.4	7.7
Recall Ensemble	36.3	22.8	90.3
CNN	25.0	16.6	50.0
Predict all 1s Baseline	33.3	20.0	**100.0**
Predict 50/50 Baseline	26.9	18.8	48.0

Discussion

In this work, we explored a ***head movement, a noisy signal in the medical domain which we confirm to be useful for predicting patient anxiety disorder***. We faced an inherently *small-data* problem, since controlled participation in a VR experience is costly to collect. As such, we focused on featurization and model comparison to determine what features and methods are promising for evaluating anxiety through head movement. ***We conclude that the discrete Fourier Transform tends to be more predictive than summary statistics about the head movement data***, as two of our three models perform better on DFT parameters, and **our *best model, a multinomial Naive Bayes on DFT features, considerably outperforms all baselines***. Qualitatively, we see in the graph below why head movement data can be predictive of anxiety. We report ***largely negative results in our attempts to use random search on model weights to choose a high-quality mixture of experts***, though our F1-optimized model does achieve the best combination of F1 and recall. Our CNN model underperformed classical baselines, likely due to a lack of data, though this was true even for very small CNNs, as shown by our hyperparameter search.



Future Work

For further research, we are interested in "more expert" featurizations on the data, as well as extending the analysis to include the video data. We believe there likely exists a featurization related to movement patterns that would perform well, we just struggled to find it. Also, in an ideal world, we'd like to collect more data so that a deep model would be more effective. We could also incorporate what happens during the virtual reality experience into the analysis. Our chief goal in future work would be increased precision.

References

[1] Sharifa Alghowinem et al. "Head pose and movement analysis as an indicator of depres- sion". In: Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Associa- tion Conference on. IEEE. 2013, pp. 283–288.

[2] Benjamin J Li et al. "A Public Database of Immersive VR Videos with Corresponding Ratings of Arousal, Valence, and Correlations between Head Movements and Self Report Measures". In: Frontiers in psychology 8 (2017), p. 2116.

[3] Spitzer RL, Kroenke K, Williams JBW, Löwe B. A Brief Measure for Assessing Generalized Anxiety DisorderThe GAD-7. *Arch Intern Med.* 2006;166(10):1092–1097. doi:10.1001/archinte.166.10.1092

[4] Bergstra, James, and Yoshua Bengio. "Random search for hyper-parameter optimization." *Journal of Machine Learning Research*13.Feb (2012): 281-305.