acoueTd [(aF)-373(measures)-372(of)-373(v)20(oice)-372(from)-373(a)-373(sample)-372(of)-373(patients)-373(and)]TJ 0 -10.959[(acouco

entire experimental process is shown in Figure 2.

Figure 2: Experimental process

Results

We first ran the experiment across the dataset across all 526 features, without feature selection. As we would expect [Cawley and Talbot, 2014], the results are terrible as the number of features is more than the number of data points, potentially causing overfitting. The results of the experiment is illustrated in figure 3. We would expect the GP ARD would be able to theoritcally extract relewe features and m(pro)15(v)15(a)-250(pediection.)291(Hho)25(le)25(w)15(e)440,ara54 645.791 cmqlly
ARD GP
results all the features are shown in figure

GP's advantages, we conclude the GP Dot Product trained using MFCC feature set is a good model for depression prediction.

Further Work

For this experiment, we only used machine learning algorithms with their default parameters. We hope to expand our experiment to perform automatic hyper-parameter optimization across all the machine learning algorithms to fine-tune each model's performance. In particular, we can try Hyperopt-sklearn [Komer, Bergstra, and Eliasmith, 2014] or GP based hyper-parameter tuner. We opine that with hyper-parameter tuning, we can predict PHQ-8 scores better and can havetera bettercomparison ofthe differentlearning algorithms.

Contributions

Antoine Charles Vincent Garcia: Scripting the program, setting up machine learning libraries and running tests.