XGBoost to Interpret the Opioid Patients' State Based on Cognitive and Physiological Measures

Omid Dehzangi, Arash Shokouhmand, Jad Ramadan, Victor Finomore, Nasser M. Nasrabadi, Ali Rezai
Rockefeller Neuroscience Institute,
West Virginia University, Morgantown, USA

Introduction and Motivation:

- **Dealing with opioid addiction and its long-term consequences is of great importance.**
- **Quitting the opioid requires clinicians to arrange a gradual plan for the patients, which necessitates observing the patients' wellness periodically.**
- **We propose continuous patient monitoring pervasively using wearables and smart-phones.**
- **we aim to predict the wellness of the opioid patients by employing extreme gradient boosting (XGBoost).**

Methodology:

1) **We design and develop a mobile application that is capable of continuous wearable monitoring and gives the ability to patients to actively engage in their state self-reporting and performance oriented cognitive tasks.**
2) **We generate a dataset with ten different patients over the period of three months and conduct pre-processing and annotation on the data.**
3) **We extract a 1st to 4th order statistical hand-crafted features out of the collected data, which helps the task of prediction to be carried out more accurately.**
4) **We configure and train a state-of-the-art machine learning model, i.e., extreme gradient-based boosting, for on average in identifying current patient state of being (e.g., depression) based on their physiological and performance measurements.**
5) **Not only can we predict the well-being state of the patients, but also to a higher degrees of confidence, we recognize those of feature space resulting in such a situation, i.e., interpretability.**
6) **Achievement in this study does not end with the subject level analysis, but a few inter-population analyses are conducted, which investigate possibility of generalization of the methods used for entire population of patients.**

Data Acquisition:

Multimodal recording and feature extraction:

- **12-hour resolution data from cognitive tests, along with heart rate (HR) and heart rate variability (HRV), sampled at each 15 and 180 seconds, respectively. 10 subjects participated in a naturalistic driving experiment with 6 secondary tasks.**

TABLE I
EXTRACTED FEATURES OUT OF THE COGNITIVE TASKS AND PHYSIOLOGICAL SIGNALS.

Cognitive tasks			Physiological measurements
PVT	**Flanker**	**N-Back**	**HR, HRV and sleep**
Score	Score	Score	HR mean
RT median	RT median	RT median	HR median
RT mean	RT mean	RT mean	HR variance
RT variance	RT variance	RT variance	HR skewness
RT skewness	RT skewness	RT skewness	HR kurtosis
RT kurtosis	RT kurtosis	RT kurtosis	HRV mean
Missed rate	Accuracy	RT(2)-RT(1)	HRV median
False rate	-	RT(1)-RT(0)	HRV variance
-	-	False alarm rate	HRV skewness
-	-	Missed detection rate	HRV kurtosis
-	-	Accuracy	Sleep duration
-	-	F1 score	-

Extreme Gradient-based Boosting:

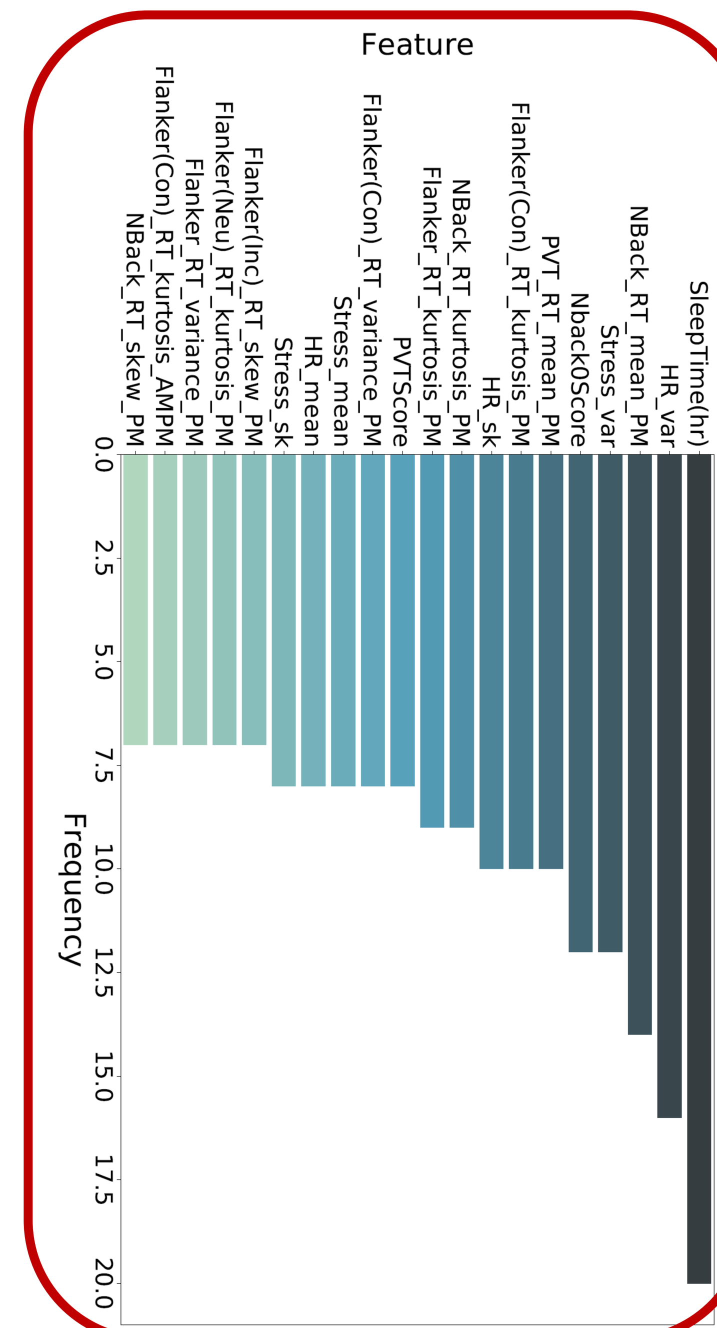
- Popularity of this method stems from its capability of involving the boosting techniques along with preserving the speed of optimization, for which K additive functions are used to predict the output:

$$\hat{y}_i = \phi(\mathbf{x}_i) = \sum_{k=1}^{K} f_k(\mathbf{x}_i),\ f_k \in \mathcal{F}, \tag{3}$$

where, $\mathcal{F} = f(x) = w_{q(x)}$, where q represents the tree structure mapping a sample to the output of tree with T being the number leaves with w equal to the leaves weights. Beside state-of-the-art performances in various tasks, one important feature of the XGBoost is its interpretability. That exactly advocates the reason why we employ XGBoost to deal with our problem.

- **Algorithm 1** Top Features Selection Procedure.

$Q, F.$
for k = 1: Q **do**
  $f_k \leftarrow []$ ▷ f_k : *top features for question k*
  $f_k = Model(X)$ ▷ X : *feature space*
  $append(F, f_k[1:10])$
$F' = set(F)$ ▷ F' : *set of unique members (features)*
$T.F. \leftarrow Sort(F')$ ▷ $T.F.$: *top features*



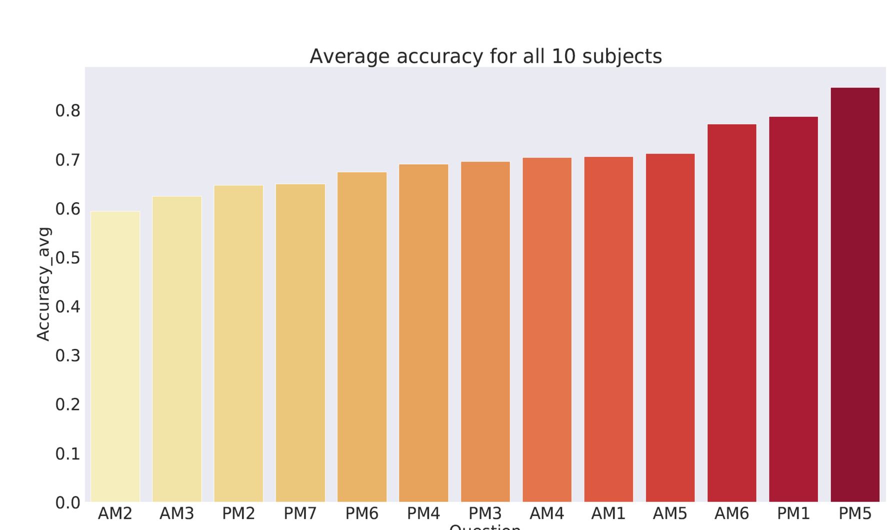
Results:

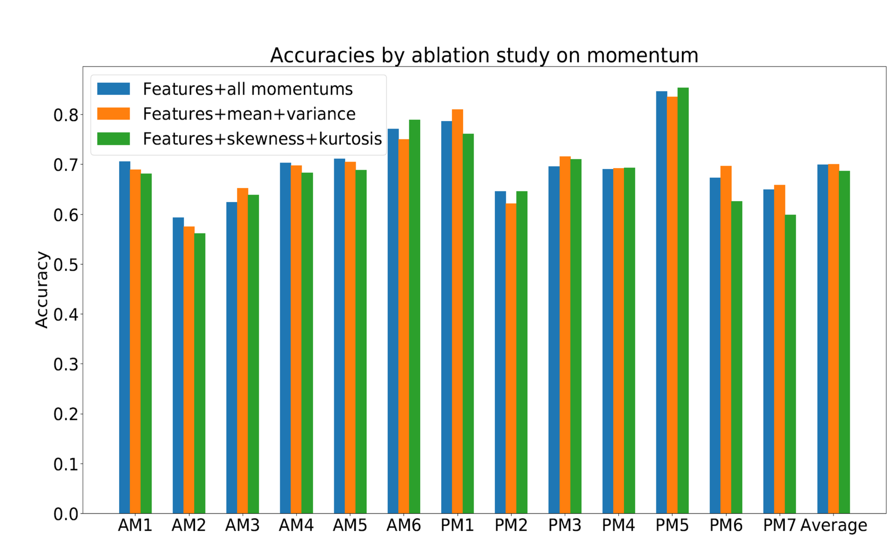
TOP FEATURES CONTRIBUTING TO THE PREDICTION OF THE SUBJECTS' STATES.

Feature rank	Subject Sub. 1	Sub. 2	Sub. 3	Sub. 4	Sub. 5
1	FlankerScore_PM	HR_kurtosis	NBack0_RT_mean_AM	Nback0Score_PM	Flanker(Con)_RT_kurtosis_AM
2	HR_var	Flanker(Inc)_RT_kurtosis_PM	Nback1Score_PM	SleepTime(hr)	Flanker(Inc)_RT_mean_PM
3	PVT_RT_mean_PM	SleepTime(hr)	NBack2_RT_mean_AM	NBack0_RT_skew_PM	HR_var
4	HR_sk	HR_var	PVTScore_PM	FlankerScore_AM	NBack1_RT_mean_AMPM
5	NBack1_RT_variance_PM	Flanker(Neu)_RT_mean_PM	NBack1_RT_variance_PM	Flanker(Neu)_RT_skew_PM	NBack2_RT_median_PM
6	Nback2Score_AM	Nback2Score_PM	NBack_RT_variance_AM	NBack2_RT_mean_PM	Flanker(Inc)_RT_kurtosis_AM
7	SleepTime(hr)	NBack_10_RT_PM	PVTScore_AM	Flanker(Neu)_RT_mean_AM	Flanker(Inc)_RT_skew_AM
8	Stress_var	FlankerScore_PM	NBack0_RT_variance_AM	Stress_mean	NBack1_RT_median_AM
9	NBack_RT_variance_PM	HR_mean	NBack_10_RT_PM	NBack2_RT_skew_PM	Flanker(Inc)_RT_median_AM
10	NBack2_Accuracy_AM	HR_ku	NBack0_RT_skew_AM	Flanker_accuracy_PM	Nback2Score_PM
Feature rank	Subject Sub. 6	Sub. 7	Sub. 8	Sub. 9	Sub. 10
1	SleepTime(hr)	FlankerScore_AM	NBack0_RT_median_AMPM	NBack2_RT_variance_AM	PVT_RT_mean_PM
2	NBack0_RT_mean_AMPM	Nback0Score_PM	Flanker_RT_median_PM	HR_ku	PVT_RT_median_PM
3	NBack0_RT_skew_AM	FlankerScore_PM	Flanker(Neu)_RT_median_PM	Flanker(Neu)_RT_skew_PM	Flanker_RT_kurtosis_PM
4	FlankerScore_PM	NBack1_RT_variance_PM	Flanker(Con)_RT_kurtosis_AM	Flanker_RT_variance_AM	Flanker(Con)_RT_skew_PM
5	Flanker(Con)_accuracy_PM	Flanker(Neu)_RT_variance_PM	Flanker(Con)_accuracy_AM	FlankerScore_AM	Flanker_RT_kurtosis_AM
6	Nback2Score_AM	HR_mean	Flanker(Inc)_accuracy_AMPM	HR_sk	NBack0_RT_skew_PM
7	NBack_21_RT_PM	SleepTime(hr)	NBack0_RT_median_AM	Stress_var	NBack2_RT_median_PM
8	NBack_RT_kurtosis_PM	NBack2_RT_skew_PM	Flanker(Inc)_RT_mean_AMPM	Flanker(Con)_RT_variance_PM	Flanker(Inc)_RT_mean_PM
9	Flanker(Inc)_RT_mean_PM	Stress_ku	NBack_RT_skew_PM	Flanker_RT_kurtosis_AM	SleepTime(hr)
10	PVT_RT_mean_AMPM	Flanker(Con)_RT_variance_PM	Flanker(Inc)_RT_mean_AM	NBack2_RT_skew_PM	NBack2_RT_skew_PM

TOP QUESTIONS WITH RESPECT TO THE ACHIEVED ACCURACY FOR PREDICTION.

Subject (Age- gender)	Top predictable questions (Prediction Accuracy) 1st	2nd	3rd	4th	5th
Sub. 1 (30-M)	PM1(1.0)	PM5(1.0)	AM4(0.9615)	PM2(0.8846)	PM6(0.8846)
Sub. 2 (36-F)	AM4(0.9565)	PM5(0.9565)	AM5(0.7391)	AM6(0.6956)	AM1(0.6521)
Sub. 3 (63-F)	PM5(0.9130)	AM5(0.8695)	PM7(0.8695)	AM1(0.8260)	AM4(0.7826)
Sub. 4 (41-M)	AM3(1.0)	AM5(1.0)	AM6(1.0)	PM1(1.0)	PM5(1.0)
Sub. 5 (55-F)	PM3(0.9565)	PM6(0.9565)	PM4(0.9130)	AM4(0.7826)	PM5(0.7826)
Sub. 6 (52-F)	PM2(1.0)	AM1(0.9375)	AM2(0.9375)	PM1(0.9375)	AM3(0.875)
Sub. 7 (41-M)	AM1(1.0)	AM6(1.0)	AM2(1.0)	AM3(1.0)	PM1(0.9545)
Sub. 8 (30-F)	PM4(0.9444)	AM4(0.8333)	AM6(0.8333)	AM1(0.7777)	PM1(0.7777)
Sub. 9 (39-M)	AM6(0.8421)	PM1(0.7894)	PM2(0.7368)	AM3(0.6315)	AM5(0.6315)
Sub. 10 (48-F)	PM3(1.0)	PM4(1.0)	AM5(0.6666)	AM6(0.6666)	AM1(0.5833)
Average Accuracy	0.9612	0.9342	0.8657	0.8202	0.7923





Conclusion:

- **High correlations were observed between the subjective and objective measures**
- **We were able to rank subjects, wellness questions, and features of multimodal measures using powerful method of extreme gradient boosting**

http://www.wssplab.net/