Eye State Prediction Using EEG

Jeff Ho
Columbia University
wh2318@columbia.edu

Abstract

In this paper, we try various methods and algorithms to enhance the prediction accuracy of eye state by analyzing brainwave data collected with EEG headset. We find that the data should be realigned with the consideration of the gap between the brain activity and the eye action.

After realigning data, the improvement is universal over classification algorithms. The best-performing classifier is Random Forest, which produced a classification error rate of only 2.4%, which is around 7% lower than the previous research.

I. INTRODUCTION

Nowadays, many developing applications requiring human-machine interaction. Training machine to precisely capture human actions and thoughts are the key technology. Among all kinds of human actions, brain activity might be the one that interests researchers most. Electroencephalography (EEG) is the most common way to measure brain waves and provide evidence of how the brain functions over time. Brain stimuli have been used as input mode for computer games [1], to track emotions [2], for handicapped persons to control devices [3], or for military scenarios [4].

Several papers have investigated eye states and reached some conclusions. [5] came to the conclusion that the "greatest difference between two states was that the power in the eye closed state was much higher than that in the eye open state." but did not attempt to use power as a feature for predicting the eye state. [6] investigated how to track eye blinking based on EEG input but only used artificial neural network. [7] adopted various algorithms on brainwave data and concluded that the instance-based learner KStar [8] could attain an error rate lower than 3% and that eye state prediction has the potential to be used as accurate binary input channel. [9] concluded that time-series classification can improve the accuracy of the classification results and the relation among time-series data is crucial to the data analysis. However, the result showed a slim improvement from 30% to 27.5%.

Our research is based on the result and the dataset of Rösler and Suendermann's paper [7]. We realign the dataset based on the observation that gaps exist between brain activity peaks and change points of eye state. Then, we try some classic classification algorithms to verify our thought. The result is outstanding and the improvement is universal. The best performer is Random Forest. The accuracy rate can be as high as 97.42%, which is better than any previous result.

The rest of the paper is as follows: Section II gives more information about the data and the softwares we use to perform machine learning algorithms. Section III describes the methods we use to find the best alignment. Section IV tries some classic machine learning algorithms to prove our thinking and measure improvements. Finally, section V draws conclusions and outlines future perspectives and applications.

II. DATA AND SOFTWARES

A. Data

The eye state was detected via a camera during the EEG measurement and manually aligned later by analyzing the video frames. The dataset is collected and provided by Rösler and Suendermann from Baden-Wuerttemberg Cooperative State University (DHBW), Stuttgart, Germany. Their paper [7] also was conducted with the same dataset. Their eye state corpus is now a benchmark problem saved by Machine Learning Repository, University of California, Irvine [10].

The duration of the measurement was 117 seconds. There are 14980 patterns and 14 features in the original dataset, where the 14 features are the data obtained by 14 sensors. We remove 4 instances with obvious transmission errors. They are no. 899, 10387, 11510 and 13180, so we have 14976 instances as our data. The label data is binary, [0,1]=[eye open, eye close]. 55% instances correspond to the eye open and 45% instances correspond to the eye closed state.

B. Softwares

We use R to create plots in this paper. LIBSVM [11] is our main software for linear SVM classifiers in section III and for linear and kernel SVM classifiers in section IIII. For AdaBoost and Random Forest algorithms in section IIII, we use Scikit-learn [12] to perform computations. All codes are available at [13].

III. SVM CLASSIFIER FOR ALIGNMENT

Brainwave plots show a very similar pattern for all features. We can see from the brainwave data collected from sensor 1 (Figure 1). When the eye state changes, the brainwaves tend to surge up. However, we observe gaps between surges of brain activity the eye state transition point. The enlarged plot between time 6500 and 6700 (Figure 2) shows the gaps exist in many features at different times and the scale is around 50 instances or 390 ms. It is a reasonable time to passing signal from brain to eye and is very lose to the previous neuroscience research [14][15].

To prove our thought, We manipulate data to form different alignments, from 0 to 20. Alignment 0 is the original setting. We manipulate the data backward or forward by 10 each time to create new alignments, each has 140 features. Then we train linear SVM on this data to see what the prediction accuracy. We split data 80% and 20% data as training and testing sets. The 20% testing data is split to two sets for validation on parameter C. Each alignment was trained with many values of C. The particular value which gave optimal performance on set A was used for testing the model on set B, and vice versa.

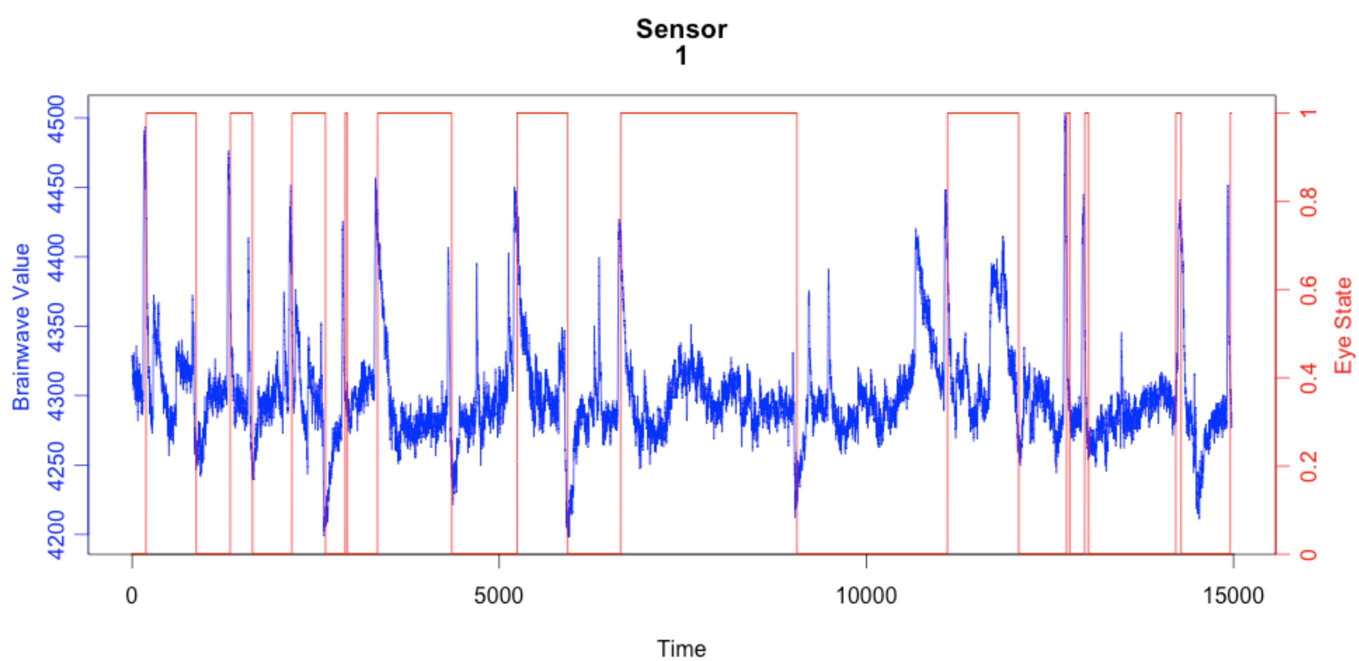

Figure 1. Brainwave value of sensor 1 & eye state

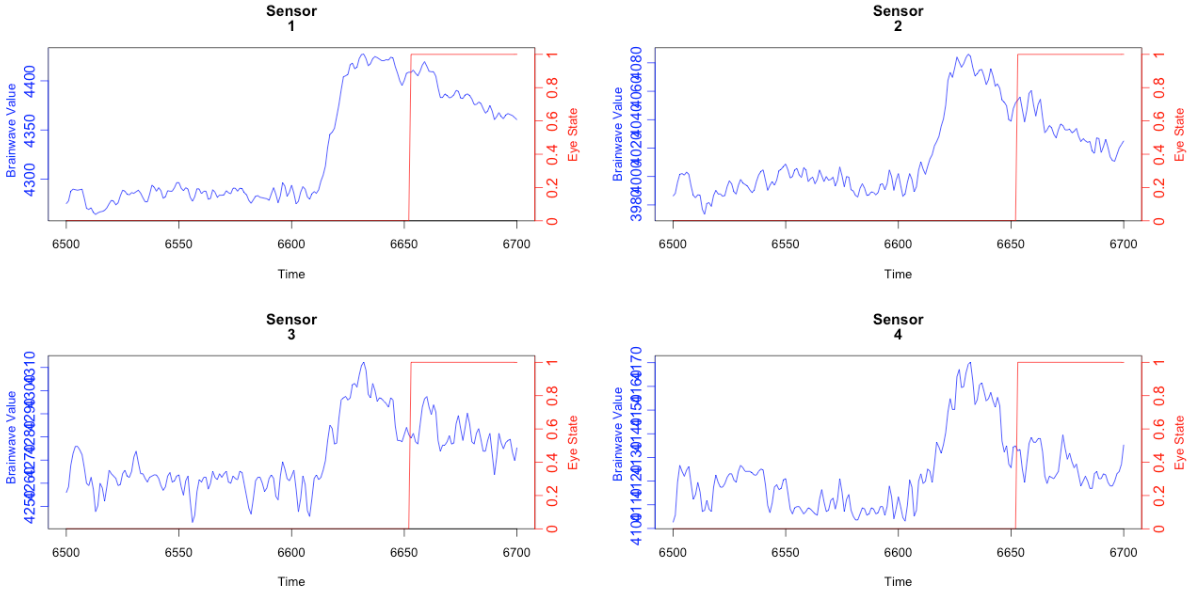

Figure 2. Enlarged plots

-100 -90 -80 -70 -60 -50 -40 -30 -20 -10 -1 1 10 20 30 40 50 60 70 80 90 100

Alignment	10	9	8	7	6	5	4	3	2	1	0	11	12	13	14	15	16	17	18	19	20
# Features	140	140	140	140	140	140	140	140	140	140	14	140	140	140	140	140	140	140	140	140	140
Mean	72.03	72.51	72.07	72.69	73.95	73.86	73.65	72.67	71.52	70.26	65.39	67.91	66.43	64.92	65.23	64.74	64.40	64.67	64.99	65.00	64.31
SD	0.56	0.73	0.57	0.70	1.17	0.92	0.77	0.78	0.76	0.59	0.97	0.81	0.62	0.90	1.01	1.01	1.12	0.70	0.58	0.88	0.74

Table 1. Linear SVM on realignments

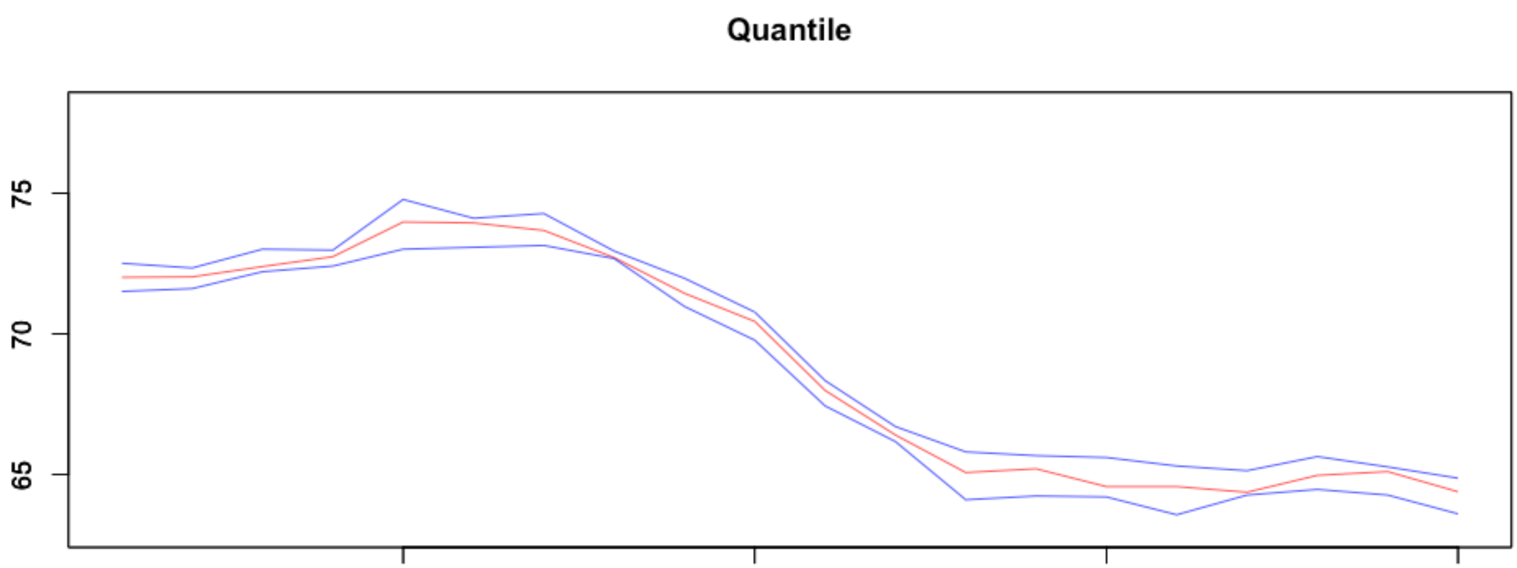

Figure 3. Linear SVM on realignments

The result (Table 1 & Figure 3) shows that the alignments 5 and 6 are the best performers and that realigning data 50 instances or 390 ms backward significantly improves the prediction accuracy. The result is very close to our assumption and expectation. It also provides a different scope for machine learning researcher to improve prediction based on the nature of brainwave data.

IIII. MACHINE LEARNING ALGORITHMS

In this part of the analysis, we adopt different kinds of SVM (linear, polynomial kernel with degree 2 & 3, and RBF kernel) and AdaBoost and Random Forest to validate the the generality of the effect of realignment. We combine alignment 5 & 6 to form alignment 21 with 280 features and use same cross validation method for C and gamma of SVMs. For AdaBoost and Random Forest, we adopt the same training strategy and 500 estimators to train data.

The result (Table 2 & Figure 4) shows performances of all algorithms. Compared with the previous paper [7], the accuracy of SVM improves from 56% to 74.36%, the accuracy of AdaBoost improves from 68% to 84.38%, and the accuracy of Random Forest improves from 91% to 97.74%. The result is very encouraging and shows strong improvements from the previous result. The effect of realignment data is significant.

IV. CONCLUSIONS AND FUTURE WORK

This paper demonstrates the overall improvements with realignment of brainwave data and the possibility to predict eye state with an accuracy of more than 97%.

Under the assumption that the data is properly aligned in terms of timeframe. The realignment process might be beneficial to make the data more predictable and enhance the accuracy rate. Also, the process could be generalizable and applied to EEG data and studies in the future.

Algorithm	Mean	SD
SVM (Linear)	74.36	0.95
SVM (Polynomial, d=2)	70.32	0.65
SVM (Polynomial, d=3)	69.08	0.66
SVM (RBF)	55.67	0.64
AdaBoost (Tree)	84.38	0.73
Random Forest	97.38	0.33

Table 2. Linear, polynomial (d=2), polynomial (d=3), RBF SVM, Adaboost, and Random Forest.

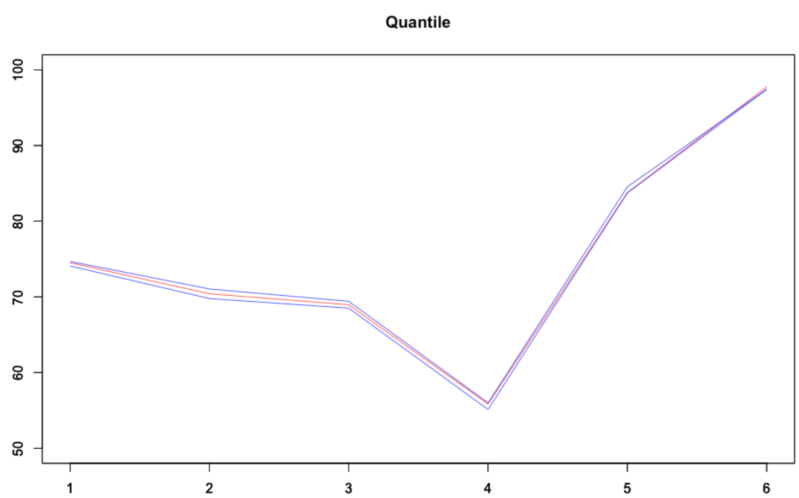

Figure 4. Linear, polynomial (d=2), polynomial (d=3), RBF SVM, Adaboost, and Random Forest.

REFERENCES

[1] P. Pour, T. Gulrez, O. AlZoubi, G. Gargiulo, and R. Calvo, “Brain- Computer Interface: Next Generation Thought Controlled Distributed Video Game Development Platform,” in *Proc. of the CIG*, Perth, Australia, 2008.

[2] T. Pham and D. Tran, “Emotion recognition using the emotiv epoc device,” *Lecture Notes in Computer Science*, vol. 7667, 2012.

[3] O. Ossmy, O. Tam, R. Puzis, L. Rokach, O. Inbar, and Y. Elovici, “MindDesktop - Computer Accessibility for Severely Handicapped,” in *Proc. of the ICEIS*, Beijing, China, 2011.

[4] J. van Erp, S. Reschke, M. Grootjen, and A.-M. Brouwer, “Brain Performance Enhancement for Military Operators,” in *Proc. of the HFM*, Sofia, Bulgaria, 2009.

[5] L. Li, L. Xiao, and L. Chen, “Differences of EEG between Eyes-Open and Eyes-Closed States Based on Autoregressive Method,” *Journal Of Electronic Science And Technology Of China,*, vol. 7, no. 2, 2009.

[6] B. Chambayil, R. Singla, and R. Jha, “EEG Eye Blink Classification Using Neural Network,” in *Proc. of the World Congress on Engineering*, London, UK, 2010.

[7] Oliver Ro¨sler and David Suendermann, “A First Step towards Eye State Prediction Using EEG”, Baden-Wuerttemberg Cooperative State University (DHBW), Stuttgart, Germany

[8] J. Cleary and L. Trigg, “K*: An Instance-based Learner Using an Entropic Distance Measure,” in *Proc. of the ICML*, Lake Tahoe, USA, 1995.

[9] I. Witten, E. Frank, and M. Hall, Data Mining: Practical Machine Learning Tools and Techniques. Los Altos, USA: Morgan Kaufmann, 2011.

[10] EEG Eye State Data Set https://archive.ics.uci.edu/ml/datasets/EEG+Eye+State

[11] Chih-Chung Chang and Chih-Jen Lin, LIBSVM : a library for support vector machines. ACM Transactions on Intelligent Systems and Technology, 2:27:1--27:27, 2011. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm

[12] Scikit-learn: Machine Learning in Python, Pedregosa et al., JMLR 12, pp. 2825-2830, 2011. http://scikit-learn.org/dev/about.html

[13] Source codes: https://github.com/Jeff88Ho/Eye-State-Prediction-Using-EEG

[14] Dodge F. A. Jr. andRahamimoff R., Co-operative action of calcium ions in transmitter release at the neuromuscular junction, 1 NOV 1967, http://onlinelibrary.wiley.com/doi/10.1113/jphysiol.1967.sp008367/abstract

[15] M R Franz, C D Swerdlow, L B Liem, and J Schaefer, Cycle length dependence of human action potential duration in vivo. Effects of single extrastimuli, sudden sustained rate acceleration and deceleration, and different steady-state frequencies., http://www.ncbi.nlm.nih.gov/pmc/articles/PMC303610/