Predicting Students' Performance in Online Adaptive Learning System Using Attention Extracted from EEG

Fangli Xu[1*], **Lingfei Wu**[2], **Shimeng Peng**[3], **Richard Jiarui Tong**[1]
[1]Squirrel AI Learning by Yixue Education Inc., New Jersey
[2]IBM T.J. Watson Research Center, Yorktown Heights, New York
[3] Dept. Intelligent Systems, Graduate School of Informatics, Nagoya University, Japan
{lili, richard}@yixue.us, lwu@email.wm.edu, hou@nagao.nagoy-u.ac.jp

Abstract

In this paper, we aim to predict students' performance on questions from four subjects by two-modality variables include students' user records and students' attention sequences extracted from electroencephalography (EEG) that measured while students working on questions generated by an adapted online learning system. Our goal is to use the data to provide appropriate adaptive assistance for students to support their learning engagement and optimize their learning outcomes. We collected brainwave signals of 94 students while they were interacting with an adaptive online learning system. We use RWS (Random Warping Series) to extract students' attention features from raw EEG signals and conducted classification on question answers using different classification algorithms. The results show that we can achieve test accuracy by up to 76.04% using Lib Linear classifier.

1 Introduction

In order to using modern AI and Deep Learning technologies to personalize learning experience for students, the key step is using adaptive learning system to collect suitable data thus people can estimate students' state and respond accordingly [Khedher *et al.*, 2018]. In last decade, rapid research progress has been made in using multi-modality dataset include behaviors such as keyboard clicks, interaction with real time sensors and facial expression captured by cameras to build models [Arroyo *et al.*, 2009]. These models have significant influence on intelligent tutoring system (ITS) research, generating more adaptive systems that could be able to optimize learning outcomes for individuals. However, when considering the typical classroom situation, feasibility of data collection still be an important constraint. Lab-based dataset could create accurate student models, but the limitation to apply to users or students, at least outside a lab situation still could be a concern.

This study was conducted to investigate the potential value of students' attention derived from Electroencephagraphy (EEG) data about students' performance during question solving of four different subjects including Math, English, Physics and Chinese while students interacting with an adaptive online learning system. The application of EEG or other features derived from EEG such as attention in educational research is relatively new. If we could find that EGG data-based features could provide a reliable indication of the students' performance in learning process, such data could be incorporated into a real time system that human could support if the learners are predicted to be in need. Our study result shows that combine EEG based attention features can achieve prediction accuracy on students' question answering performance up to 76.04%.

2 Related Works

2.1 Multimodal Learning Data Analysis

In Education Data Mining (EDM) community, researchers have investigated several approaches to study the performance of learners. One popular approach is to use different sensors that can capture the physical indications of the learner's state. In Kapoor, Burleson and Picard's work, they used variety of sensors including tracking eye movement, mouth movement, skin conductance and chair pressure in a lab setting to measure students' level of frustration and reported that they could predict when a user was about to give up with a 79% accuracy using Support Vector Machine (SVM) and SVM with Gaussian Process Classification (GP) classifier [Kapoor and Burleson, 2007]. Arroyo et al used multiple sensors, including facial expression recognition, skin conductance, mouse pressure, back pressure and students' feedback to estimate their emotional state as they solving math problems in an authentic classroom setting. Their study can achieve accuracy more than 60% in predicting the variance of students' emotional state [Arroyo *et al.*, 2009]. Fincham et al using functional Magnetic Resonance Imaging (fMRI) equipment to test subjects while solving problems in an Intelligent Tutoring System. They applied a cognitive model to predict distribution of problem solution times from measures of problem complexity, and used fMRI to predict students' engagement level while problem solving. They reported that they can predict the mental state of the subject with 83% accuracy [Fincham *et al.*, 2010].

From previous research we could get better result from

*Corresponding Author

fMRI data. However,fMRI data is still costly and not feasible to use in real classroom setting. So recently EEG light weight headset attract more and more education researchers. The related technologies for capturing EEG signals has progressed that users can easily wear an affordable and lightweight headset recording set that transmits data for analysis. Stevens, Galloway, Berka, Johnson and Sprang compared two groups of users namely novices and experts who solve chemistry problems. Their results indicated that the two groups showed distinct patterns of attention in term of problem solve time and accuracy [Stevens *et al.*, 2008]. Mostow, Chang and Nelson used a single-channel EEG recorder with both adults and children as they read text passages with different difficulty levels. The results showed significantly at discriminating adults and children, as well as predicting the difficulty level of the passages [Mostow *et al.*, 2011]. Chaouachi, Jraidi and Frasson used a six channel set to capture EEG activity as users solved cognitive tasks varying in difficulty level, and their experiment verified that their estimates of the users' mental workload were correlated with the difficulty of the tasks [Fincham *et al.*, 2010].From previous researchers, initial results appear to suggest that EEG signals could be a valuable data for assessing a student's level of cognitive effort and performance measurement. Moreover, Xu et al showed that EEG measurements with knowledge tracing can improve estimation of the students' hidden knowledge state [Xu *et al.*, 2014].

2.2 Time Series Data Preprocess

In time series data preprocess, Dynamic Time Warping (DTW) is the most widely adopted technique to measure dissimilarity between time series. We used a method which present a family of align time series. There are lots of study about how to get both effectiveness and efficient global-alignment kernels. In this study, we leverage an alignment-aware positive definite (p.d.) kernels which given by a distribution of Random Warping Series (RWS). The approach reduces the computational complexity of existing DTW-based techniques from quadratic to linear in terms of both the number and the length of time-series [Wu *et al.*, 2018].

3 Experimental Study

3.1 Data Collection

The dataset we gathered is from a truly real-world educational activities versus those from a controlled lab environment. All the data were collected from two offline schools of Squirrel AI learning. All the students are from 7th to 9th grade, doing Math, English, Physics and Chinese questions. All the learning logs are stored in Squirrel AI online learning system data center. Students' attention and EEG signals are collected by a headset called FocusEDU[1] developed by BrainCo as they were solving questions adaptively given by the online learning system. The detailed data collection procedure (shown in figure 1) and the whole dataset could be found via this link.[2]

[1]https://www.brainco.tech/focusedu/

[2]https://tinyurl.com/SAILdata

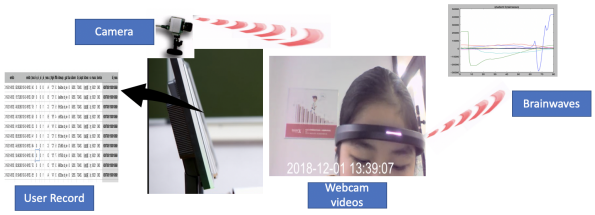

Figure 1: Data Collection Procedure Overview

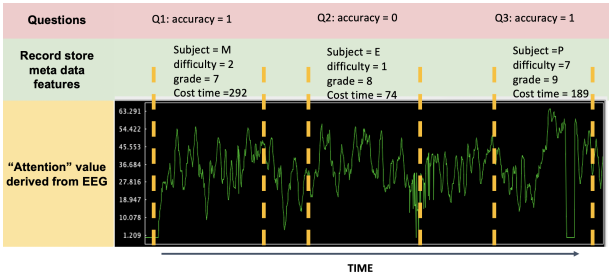

Figure 2: Data segmentation Overview

3.2 Data Preparation

The dataset includes time-synchronized multi-modal data recordings (learning logs, videos, EEG brainwaves) on students as they working on various Squirrel AI Learning products to solve problems that vary in subjects and difficulty levels. In this paper, the dataset was segmented by questions, along with each question, we compiled several features from user records extracted from the learning logs and EEG signals. For each question sample, we have student ID, grade level, difficulty level, subject, answer correctness and cost time for solving the question and students' attention sequence extracted from EEG signals during the time window solving the question as shown in Figure 2.

In total we got 2441 samples from 94 students. For grade level, we have 483 question samples are from grade 7. 1400 question samples are from grade 8, and 558 question samples are from grade 9. The difficulty levels rang from 1 to 9. The overview is shown in Figure 3. We also reported the overview of the dataset in Table 1.

3.3 Attention Feature Set Preprocession

All the attention values are at frequency of 1.0-2.0 Hz which was derived using EEG sensor manufacturer's proprietary algorithm. For each question, we extracted the attention series from the raw EEG signal between the time that the student gets the question and the time to submit the answer of the question. We then use the approach discussed in [Wu *et al.*,

Subject	Math	English	Physics	Chinese
School 1	8	230	-	-
School 2	963	642	450	148
Total	971	872	450	148

Table 1: Dataset Overview

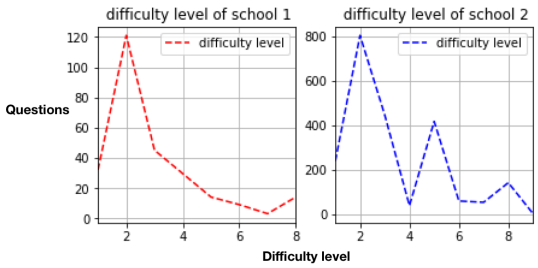

Figure 3: Difficulty level of the dataset

Schools	Train Accuracy(%)	Test Accuracy (%)
School 1	61.45	54.17
School 2	73.74	63.39
Overall	70.4	65.8

Table 2: Train/Test Results Using KNN with User Records

2018] paper to get the best approximation for each varying length attention sequence.

3.4 Students Performance Prediction

In this section, we report the experiments results. We conduct several experiments. At first, we only use students user records features which we complied to predict question answering performance. The features we used include each student grade level, each question difficulty level, subject, cost time of solving each question, answer correctness about the question to predict students' performance on answering question. For classification, we used different classifiers namely SVM, K-Nearest Neighbor, Decision Tree, Random forest and Adaboost and logistic regression. We then used k-fold cross validation where $k = 5$. From table 2, we reported the highest accuracy rates comes from K Nearest Neighbor classifier. School 2 which has more samples can achieve higher training accuracy. The overall train/test accuracy can be achieved to 70.4% / 65.8%.

Then, we only use the students' attention sequence to predict the student answering performance of each question. We leverage the approach discussed in [Wu *et al.*, 2018] paper. The first step, we use the data procession method to preprocess the attention sequence. Then use the script to generate low-rank approximation of latent kernel matrix using random features approach based on DTW like distance for multivariate time series. The third step is using Liblinear to perform grid search with K-fold cross-validation where $k = 10$. After that, we investigate performance changes when varying R using parameters learned from last step. The best results are shown in table 3 where $R = 128$.

From table 3, we can find that both schools and the overall train / test accuracy are higher than the classification performance based on features compiled from students' user records. For overall performance, we can achieve as high as 76.04% based only students' attention sequence as they solving question which means students' attention sequence can help on student question solving performance prediction.

Schools	Train Accuracy(%)	Test Accuracy (%)
School 1	76.04	61.75
School 2	72.99	64.86
Overall	76.04	70.19

Table 3: Train/Test Results Using Liblinear with Attention

We can then further investigate the pattern of the attention sequences and identify the students who may need early assistance to solve question correctly.

4 Discussion and Future Work

These preliminary results suggest that using students' attention sequences extracted from their EEG signal predict question solving outcomes better than based users records extracted from learning logs. These findings confirm that a prediction model can be built based on a combination of features from two or more modalities. We can further suggest that the sensor based multi-modal approach can be used to predict learners' performance during an interaction with a learning system.

In future work, we will process one more modality data in the MIBA system to combine three modalities namely students' user records, attention sequence, facial expression to do more solid performance prediction of learning outcomes and students' engagement of the learning system analysis.

5 Conclusion

In this paper,we processed two-modality data from the MIBA system. Extracted students' attention sequences from EEG. Used RWS to extract feature sets of attention sequence. We first use several features from students' user records including grade, difficulty level, question solving cost time, subject and student ID to predict students' question answering performance with 70.4 / 65.8% train/test accuracy with KNN Classifier. Then we extract students' attention sequence to predict students' question answering performance with 76.04% accuracy. These findings confirm that a prediction model can be built based on a combination of features from two or more modalities. We can further suggest that the sensor based multi-modal approach can be used to predict learners' performance during an interaction with a learning system.

6 Acknowledgement

Special thanks to the two after-school learning centers teachers and students who participated and supported the data collection, and Squirrel AI Learning staff and interns who aided in the data collection, cleaning, and aggregation. The data described and shared here were collected by Squirrel AI Learning, and they also funded this research.

References

[Arroyo *et al.*, 2009] Ivon Arroyo, David G. Cooper, Winslow Burleson, Beverly Park Woolf, Kasia Muldner, and Robert Christopherson. Emotion sensors go to school. In *Frontiers in Artificial Intelligence and Applications*,

volume 200 of *Frontiers in Artificial Intelligence and Applications*, pages 17–24, 2009.

[Fincham *et al.*, 2010] Jon Fincham, John Anderson, Shawn Betts, and Jennifer Ferris. Using neural imaging and cognitive modeling to infer mental states while using an intelligent tutoring system. pages 51–60, 10 2010.

[Kapoor and Burleson, 2007] Ashish Kapoor and Winslow Burleson. Automatic prediction of frustration. *International Journal of Human-Computer Studies*, 65(8):724–736, August 2007.

[Khedher *et al.*, 2018] Asma Ben Khedher, Imène Jraidi, and Claude Frasson. Local sequence alignment for scan path similarity assessment. 2018.

[Mostow *et al.*, 2011] Jack Mostow, Kai-min Chang, and Jessica Nelson. Toward exploiting eeg input in a reading tutor. In Gautam Biswas, Susan Bull, Judy Kay, and Antonija Mitrovic, editors, *Artificial Intelligence in Education*, pages 230–237, Berlin, Heidelberg, 2011. Springer Berlin Heidelberg.

[Stevens *et al.*, 2008] Ronald H. Stevens, Trysha L Galloway, Chris Berka, Robin Johnson, and Marcia Sprang. Assessing student's mental representations of complex problem spaces with eeg technologies. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 52(3):167–171, 2008.

[Wu *et al.*, 2018] Lingfei Wu, Ian En-Hsu Yen, Jinfeng Yi, Fangli Xu, Qi Lei, and Michael Witbrock. Random warping series: A random features method for time-series embedding. In Amos Storkey and Fernando Perez-Cruz, editors, *Proceedings of the Twenty-First International Conference on Artificial Intelligence and Statistics*, volume 84 of *Proceedings of Machine Learning Research*, pages 793–802, Playa Blanca, Lanzarote, Canary Islands, 09–11 Apr 2018. PMLR.

[Xu *et al.*, 2014] Yanbo Xu, Kai min Chang, Yueran Yuan, and Jack Mostow. Eeg helps knowledge tracing ! 2014.