Proceedings of the 2023 International Conference on Software Engineering and Machine Learning
DOI: 10.54254/2755-2721/27/20230190

Study on the human-brain confusion levels and corresponding EEG levels

Yuhan Luo[1, 5, †], Zijie Zhou[2, †], Hao Yang[3, †], Bin He[4, †]

[1]Communication School, Xidian University, Xi'an City, Shannxi Province, 710119, China, 410512556@qq.com
[2]The Nanyang Technological University, Singapore, 639798, Singapore, 2069351871@qq.com
[3]Huaer Zizhu Academy, Shanghai, 201102, China, hhappier@163.com
[4]Basis International School Guangzhou, Guangzhou City, Guangdong Province, 510700, China, hebinbenny@hotmail.com

[5]410512556@qq.com

[†]These authors contributed equally and should be considered as co-first author.

Abstract. Confusion is a state in which an individual is not clear about the situation at hand, does not understand the logic of the matter, or cannot arrive at a reasoned outcome. When people are confused, they often make poor decisions or even fail to make decisions. In this paper, we tested and validated our two models using the collected EEG signals of ten students while watching online courses of different difficulty levels. Both the LSTM model and the DNN model obtained an accuracy rate close to 70% in the validation part, but they showed different characteristics. In the end, we compare and summarize the results of the two models and try to generalize them to other domains in the future.

Keywords: EGG, LSTM, DNN.

1. Introduction

Confusion is a term that refers to a decline in cognitive ability, that is, our power of thinking, learning ability, and comprehension. Confusion is a common phenomenon in people's daily life. Decline in cognitive ability is often associated with dementia, while there is also a large portion of confusion that results from receiving information that is beyond the scope of pre-existing cognition. However, once the confusion is solved successfully, deep learning and understanding will occur [1]. Especially when online teaching is becoming more and more popular and urgent, this situation is even more prominent.

Under the impact of the pandemic, mass open online courses (MOOC) has become a popular education method. However, network teaching also has many limitations, the biggest limitation is the lack of interaction between teachers and students. Since there is no communication between teachers and students, it is easy to evaluate the effectiveness of online teaching. Actually, confusion occurs commonly in classes, where confusion refers to a lack of understanding or a state of disorder. Students are easy to miss some key content because they are distracted by something or they feel shy to ask questions during the class. The teacher would know whether a student is confused on the topic through students' verbal queries and body language in conventional classes, but the current online course platform has clearly lost this convenience [3]. The pattern of teaching that leaving the face-to-face interactions means that teachers have little or no chance to know students' cognitive and affective states. For the sake of solving this problem, researchers have tried various ways to detect confusion, and recently they started to move their eyes on EEG signals-detecting equipment.

Traditional EEG signal analysis techniques usually go through three steps: removing artifacts, extracting features, and selecting features. The basic idea of the algorithm is: first, through the waveform segmentation of the EEG signal, and then extract the waveform according to the waveform, and identify the waveform. Before the training results are used in this method, a large number of noises and illusions are eliminated. ICA will split this message into separate components. On this basis, the frequency domain unmixing of EEG is carried out by using Fourier transform technology.

The inspiration for our study comes from the paper "Using EEG to Improve Massive Open Online Courses Feedback Interaction [3]". We work on a pioneer study that records EEG signals from 10 college students when watching 20 different online courses. The courses are divided into 10 easy ones (e.g. an introduction to algebra) and 10 complicated ones (e.g. quantum mechanics). 3 student observers would estimate the confusion level of the students rated on a scale of 1-7 through the body language and facial expressions of the students. And students also rate their confusion levels of themselves, which are treated as the actual confusion levels. A device named "MindSet" will also be worn on the forehead and ears of each student to record the EEG signal.

The researchers trained Gaussian Bayes classifiers to estimate the probability that a student is confused based on the EEG signals produced. The classifiers' ratings are then compared to the observers' ratings, to see whether the classifiers can estimate the confusion levels via EEG signals as well as human observers via body language and facial expressions.

The result of the experiment is that the classifiers have comparable performance to human observers when estimating the confusion levels of students.

2. Literature review

To enrich our background knowledge and to obtain some references, we have collected and organized some relevant literature.

The phenomenon of confusion in EEG data is identified by using a two-dimensional LSTM cyclic neural network [4]. Its BI-LSTM pattern has a good recognition effect, and it can correctly judge the confusion or confusion of the subjects, and the correct degree can reach 73.3%.

Based on tests of Raven's Standard Progressive Matrices, Hussain et al. [5] made an experiment design to induce confusion which recorded and used EEG data of 16 participants in 2018. The results of this preliminary study are very promising. When users infer rules in the test, the results show an accuracy rate of 71.36%, which can classify users' confused states and unconfused states.

In the same year, Yun et al. [1] deliberated on the ability to use TEL system and MATLAB programming language to predict the difficulty of students in subsequent courses. Two approaches (five-fold cross-validation and random division of the data into portions) were applied to assess the performance of their models. The results show that the ANN and SVM models can achieve high accuracy (75%) in predicting the difficulties that students will encounter in the next digital design course.

In the same year, Yin et al. discussed the method of predicting the difficulty of students in the future teaching process by using TEL system and MATLAB programming program [1]. We will evaluate the model in two ways (five-fold interaction and grouping). It is found that the prediction accuracy of ANN and SVM models for the problems encountered in the next digital design class can reach 75%.

The accuracy of SVM model was raised to 80% via the Alpha-investing method.

In 2019, Zhou et al. investigated that in the play-based teaching, the cognitive and emotional aspects other than input are studied [6]. The goal was to detect confusion state via EEG signals. The team first collected EEG signals from the subject while he performed the Raven's test (a standardized cognitive test), which were then used to train the classifier based on end-to-end learning with convolutional neural networks (ConvNets). The team then tested the classifier via signals collected from the subjects as he played a Sokuban Game. This technique lays a foundation for building a classifier with fewer samples. Experiments show that the designed classifier has a high recognition accuracy of 91.04%.

In one of the most recent publications we read, there's a significant progress. Dakoure et al. [7] built an effective model with EEG signals, which can not only predict confusion, but also break it down into three levels (low, medium, high confusion). By recording and analyzing the EEG signals of ten participants when solving five series of different cognitive exercised, they extracted the power spectra from these cleaned EEG signals and used them as input to train the models. Their best model for classifying three levels of confusion reached 68.0% accuracy. As for predicting the confused/unconfused state, they reached an accuracy of 78.6%. The model developed in this literature achieved a breakthrough.

3. Methodology

3.1. Equipment

Mindset

MindSet is an audio headset" developed by NeuroSky based on its latest Brain Computer Interface (BCI) technology. It allows the user to control the characters and scenes in the computer with the mind through real-time acquisition and analysis of the wearer's brainwave signals.

Bioelectric phenomena are one of the fundamental characteristics of life activity. The bioelectric signals generated in the brain during human thinking activities are brain waves, and these brain wave signals can be measured and studied by sensors placed in the scalp.

The Mindset equipped with an EEG sensor uses a dry electrode sensor to collect the EEG generated by the brain and transmit the collected information to ThinkGear (ThinkGear) [8]. Compared with the traditional multi-channel electrode network, this new type of electrode network does not need to use instruments such as colloid or normal saline to collect data, and does not need to have a certain technical basis. A previous survey [9] showed that the "mind" can distinguish between two very similar minds ("neutral" and "concentrated") with an accuracy of 86%. In the intelligent education system, the way of thinking has been used to detect dyslexia [10] and human emotional response [11].

ThinkGear chip

The ThinkGear chip filters out the noise mixed in the signal and the interference generated by motion, amplifies the useful signals, and then interprets the eSense parameters (concentration, relaxation) describing the user's current mental state through the NeuroSky eSense algorithm. Then the eSense parameters (concentration, relaxation), describing the user's current mental state, are interpreted by the NeuroSky eSense algorithm, and finally the brainwave-based human-computer interaction is achieved by outputting these quantified parameters to computers, cell phones and other smart devices, which is commonly referred to as mind control.

3.2. Dataset

First of all, it collects the EEG of college students when they watch the video of Moo class. We have selected some online teaching videos that are less difficult for college graduates, such as an introduction to algebra and geometry. At the same time, we also provide films that can confuse ordinary college students for students who don't know much about topics such as quantum physics and stem cells. We made 20 films and then divided them into two categories. The length of each video is about 2 minutes. We intercepted two minutes of one of the themes to make the film look a little confusing.

We gather records from 10 college students. One student was eliminated because of missing data due to technical difficulty. Each student must watch ten videos during the experiment, including five videos randomly selected from each of the two categories. The researchers disrupted the order of the videos so that students could not predict the difficulty of the videos.

Before each test, students were asked to relax for 30 seconds. Then, the researchers showed the students a video clip and instructed him/her to learn from the video as much as possible. At the end of each session, students must grade their confusion on a scale of 1-7. The higher the value, the greater the

DOI: 10.54254/2755-2721/27/20230190

confusion. In addition, students' body language was under observation of three student observers. Each observer rated the students' confusion level from 1-7 in each session, where the conventional scale of 1-7 was kept. Four observers were asked to observe 1-8 students, so that the observer would not study on only one student.

The students used a wireless single-channel thinking system to measure their prefrontal cortex activity. Minsett measures the voltage between two electrodes connected to the ear on the forehead, one on the ground and the other as a reference. Minsett placed three electrodes on his forehead and ears (one on the ground and the other as a reference) and measured the potential between the two. More specifically, as defined in the 10-20 system [12], the forehead is FP1 (from the left eyebrow to the hairline). NeuroSky API used to collect the following signals:

a. Primary EEG signals sampled at 512 Hz

b. Signal quality indicator reported at a rate of 1 Hz

c. Minsett's patented "attention" and "meditation" signals, which are used to measure users' concentration and calmness, are reported at a frequency of 1 Hz.

d. The reported energy spectrum of 8 Hz is combined into a standard name band: Delta at (1-3Hz), Theta at (4-7 Hz), Alpha at (8-11 Hz), Beta at (12-29 Hz), and Gamma at (30-100 Hz).

In selecting features of EEG to evaluate participants' level of focus, two categories were specifically aimed to be found in classification, which were EEG representative to attention (proprietary measure of mental concentration), and EEG representative to meditation (proprietary measure of calmness). Raw EEG data was collected and selected into Alpha, Beta, Delta and Theta signal. The data are divided into training values and trial values, which is convenient for the study of these materials.

4. Experimental result and discussion

4.1. Models

This project intends to use LSTM (Logistic Strategy Memory, LSTM) and DNN (Deep Net Net, DNN) methods to simulate the confusion in various situations faced by students in online classes.

LSTM

LSTM is a developed version of machine learning model derived from RNN (recurrent neural network). Choosing LSTM for our study has two obvious advantages. As it solved the problem of vanishing gradient, maintaining the depth of machine learning throughout the training and testing, LSTM is advanced in learning a long-lasting characteristic which should combine many features in the study; As the time is a crucial factor for one's change in attention and meditation during the class, LSTM could be suitable in predicting the EEG data in the study.

DNN

DNN (Deep Neural Networks) is the basis of many deep learning models, from which many models branched out. Instead of having time taken as an important factor of its "depth", or having its gradient change with strong correlation with time passing, DNN doesn't specialize in modeling variation in time series. It has its own advantage as a typical model in dealing with data and classification problems, and thus the outcome of it is of value for reference.

4.2. Result

In this section, we compare our results when using different algorithms to analyze the EEG signals. Specifically, these are results obtained when running signal analysis with different types of neural network models. We have two diagrams that will likely show the difference between the different neural networks, in this case, DNN and LSTM.

DOI: 10.54254/2755-2721/27/20230190

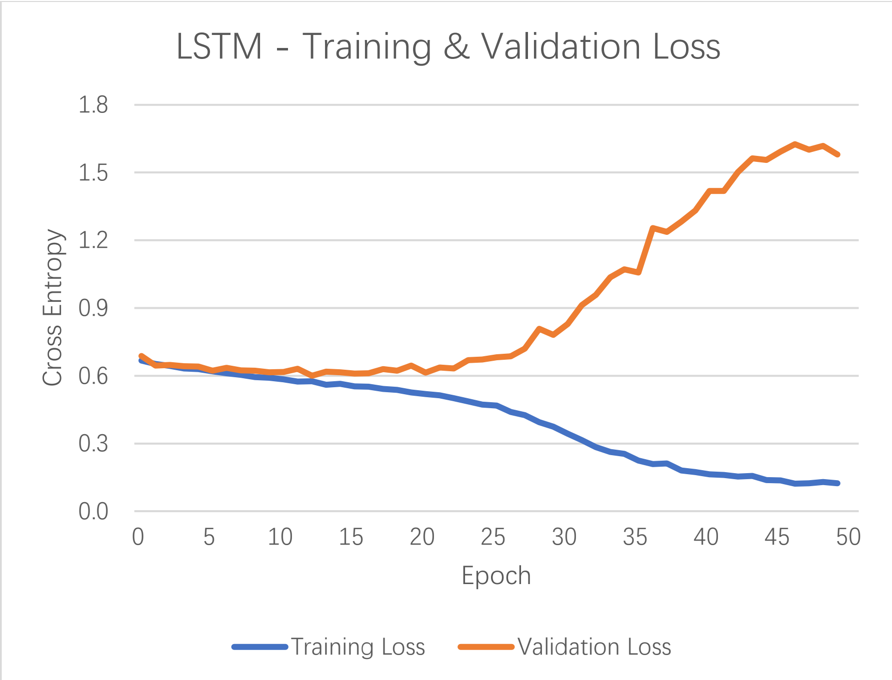

Figure 1. Training and validation loss of LSTM model.

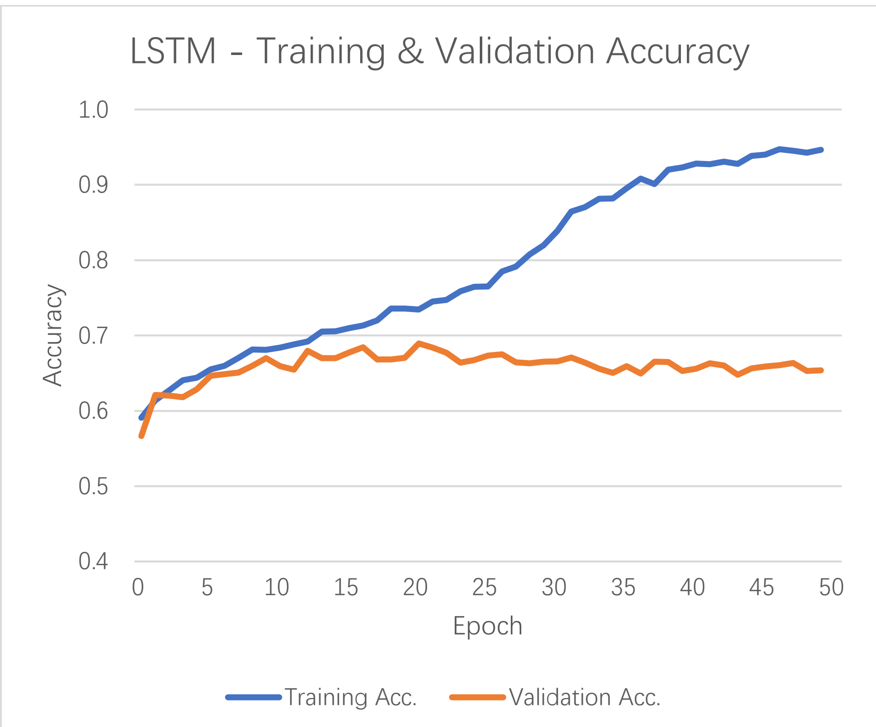

Figure 2. Training and validation accuracy of LSTM model.

In the LSTM accuracy graph, the training accuracy quickly exceeded 90% on the 36th epoch, while the validation set accuracy is significantly lower, settling in at around a consistent 65% starting from the 6th epoch. For its loss, the entropy of the validation loss rose as high as 1.6 on the 47th epoch to the 49th epoch, but its training entropy is kept quite low, below 0.13 on the last three epochs.

DOI: 10.54254/2755-2721/27/20230190

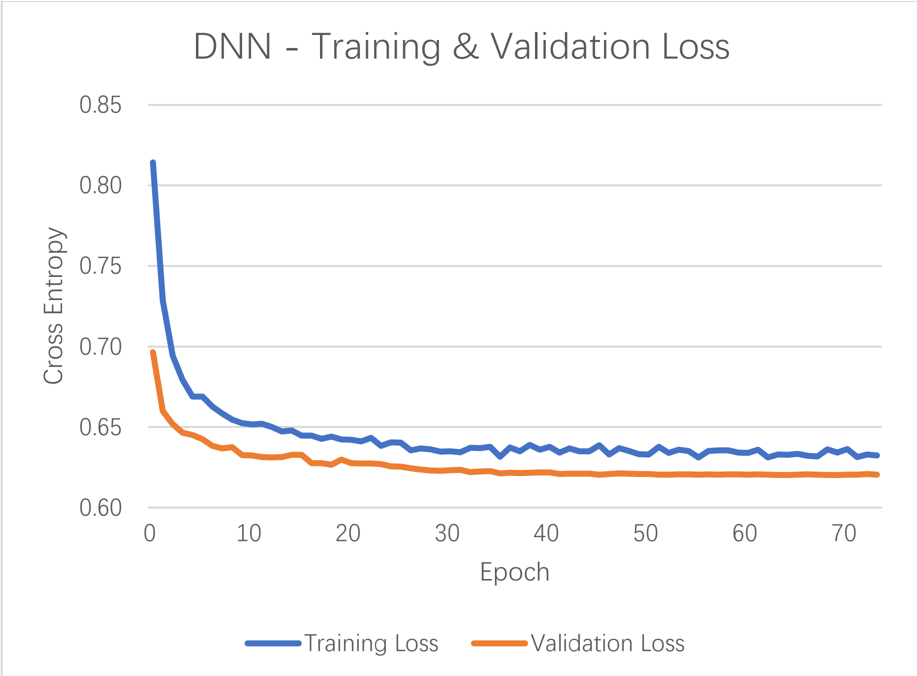

Figure 3. Training and validation loss of DNN model.

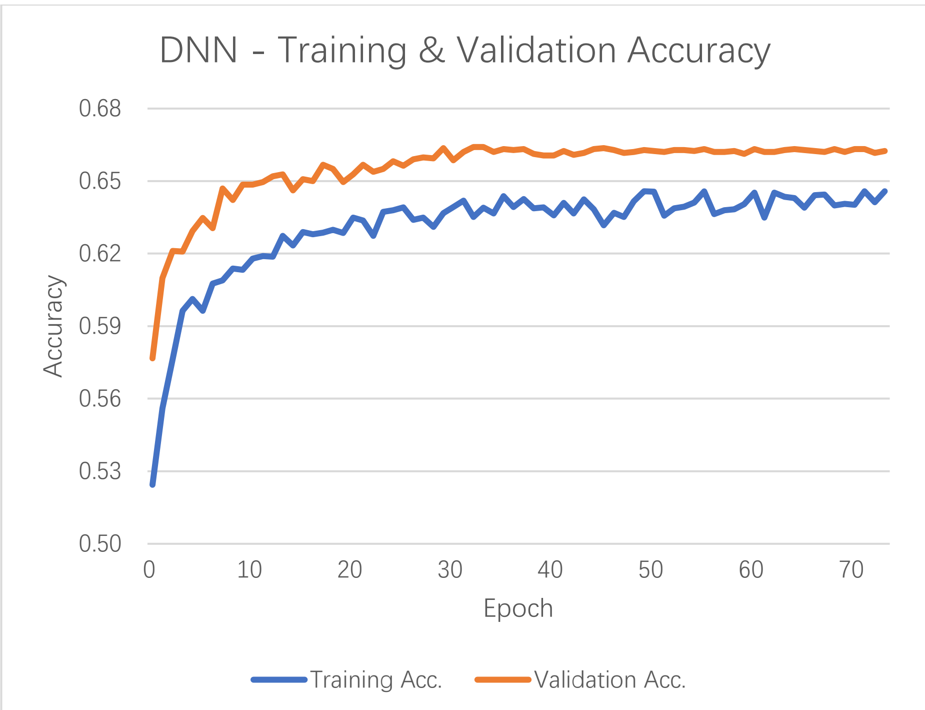

Figure 4. Training and validation accuracy of DNN model.

For DNN, its training and validation accuracy difference are much smaller, and both reached about 0.65 eventually, but in contrast LSTM's training accuracy is a lot higher while its validation accuracy is not that far off from the ones from the DNN. Additionally, similar to the trend of accuracies, the training and validation losses also follow similar trends: both approach 0.63 eventually when following a similar path to that end result, and running both datasets yield lower entropy than the LSTM when running the validation set, but both sets in DNN failed to keep up with the high accuracy of the training entropy of LSTM.

4.3. Discussion

In summary, the LSTM reaches higher peak accuracy when dealing with its training dataset, exceeding DNN by having both a lower entropy and a higher accuracy. When dealing with the validation dataset, LSTM's accuracy is similar to DNN's, while its entropy is way higher than DNN's. However, there is one interesting thing for DNN: its curves for the training or validation datasets follow a trend. For example, they both followed a similar path that yielded them the mediocre accuracy, and their entropies both decreased to a similar value. Therefore, for practical uses, DNN might be the preferred choice, as its higher consistency with the validation dataset suggests its tolerance for EEG signals that the model have not seen before, which happens all the time in real situations. Its more consistent results between the training and validation dataset also make it advantageous compared to LSTM, because it will be easier to predict the results that can be achieved with the real situations' dataset to training datasets.

5. Conclusion

It is of great significance to detect the level of confusion of students. The teacher can use the results of the student confusion level detection to keep track of the students' listening status and understanding during the online courses, and adjust their lecture schedule based on the feedback or further explain the important points that students feel confused about. Based on our dataset, we propose two models with different characteristics: the LSTM model shows amazing results in training, while the DNN model may be more practical and generalized. In the course of subsequent research, we will further refine both models, such as improving the prediction accuracy of the models and extending the models to areas beyond just student confusion detection. For different domains, it is likely that different models will need to be designed and applied to suit their specific situations to obtain better detection results.

Acknowledgement

All the authors contributed equally to this work and should be considered as co-first author.

References

[1] Yun, Zhou, Tao, Xu, Shiqian, & Li, et al. (2018). Confusion State Induction and EEG-based Detection in Learning. Conference proceedings:.Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference.
[2] S D' Mello, Lehman, B., Pekrun, R., & Graesser, A. (2014). Confusion can be beneficial for learning. Learning & Instruction, 29, 153-170.
[3] Wang, H., Li, Y., Hu, X., Yang, Y., Meng, Z., & Chang, K. M. (2013, June). Using EEG to Improve Massive Open Online Courses Feedback Interaction. In AIED Workshops.
[4] Ni, Z., Yuksel, A. C., Ni, X., Mandel, M. I., & Lei, X (2017). Confused or not confused?: disentangling brain activity from eeg data using bidirectional lstm recurrent neural networks. ACM.
[5] Hussain, M., Zhu, W., Zhang, W., Abidi, S., & Ali, S. (2018). Using machine learning to predict student difficulties from learning session data. Artificial Intelligence Review.
[6] Zhou, Y., Xu, T., Li, S., & Shi, R. (2019). Beyond engagement: an eeg-based methodology for assessing user's confusion in an educational game. Universal Access in the Information Society, 18(1), 1-13.
[7] Dakoure, C., Benlamine, M. S., & Frasson, C. (2021). Confusion detection using cognitive ability tests.
[8] NeuroSky, Brain wave signal (EEG), 2009, Neurosky, Inc.
[9] NeuroSky, NeuroSky's eSense™ meters and dtection of mntal sate, 2009, Neurosky, Inc.

DOI: 10.54254/2755-2721/27/20230190

[10] Mostow, J., Chang, K., & Nelson, J. (2013). Toward Exploiting EEG Input in a Reading Tutor. IOS Press (Vol.22, pp.19-38). IOS Press.
[11] Crowley, K., Sliney, A., Pitt, I., & Murphy, D (2010). Evaluating a Brain-Computer Interface to Categorise Human Emotional Response. IEEE International Conference on Advanced Learning Technologies. IEEE.
[12] Jasper, H.. (1958). The ten-twenty electrode system of the international federation. Electroencephalography and Clinical Neurophysiology, 10.