



# Predicting Student Performance Using Machine Learning in fNIRS Data

Amanda Yumi Ambriola Oku\* and João Ricardo Sato

Center of Mathematics, Computing and Cognition, Universidade Federal Do ABC, São Bernardo Do Campo, Brazil

Increasing student involvement in classes has always been a challenge for teachers and school managers. In online learning, some interactivity mechanisms like quizzes are increasingly used to engage students during classes and tasks. However, there is a high demand for tools that evaluate the efficiency of these mechanisms. In order to distinguish between high and low levels of engagement in tasks, it is possible to monitor brain activity through functional near-infrared spectroscopy (fNIRS). The main advantages of this technique are portability, low cost, and a comfortable way for students to concentrate and perform their tasks. This setup provides more natural conditions for the experiments if compared to the other acquisition tools. In this study, we investigated levels of task involvement through the identification of correct and wrong answers of typical quizzes used in virtual environments. We collected data from the prefrontal cortex region (PFC) of 18 students while watching a video lecture. This data was modeled with supervised learning algorithms. We used random forests and penalized logistic regression to classify correct answers as a function of oxyhemoglobin and deoxyhemoglobin concentration. These models identify which regions best predict student performance. The random forest and penalized logistic regression (GLMNET with LASSO) obtained, respectively, 0.67 and 0.65 area of the ROC curve. Both models indicate that channels F4-F6 and AF3-AFz are the most relevant for the prediction. The statistical significance of these models was confirmed through cross-validation (leave-one-subject-out) and a permutation test. This methodology can be useful to better understand the teaching and learning processes in a video lecture and also provide improvements in the methodologies used in order to better adapt the presentation content.

**Keywords:** neuroscience, fNIRS, education, prefrontal cortex, machine learning, logistic regression, random forest

## 1. INTRODUCTION

The interactivity in a virtual teaching environment can increase student engagement and, therefore, reinforces learned concepts and provide on-demand learning capacity (Jonassen et al., 1995). Empirical assessments have emerged in recent research, such as studies by Wachtler et al. (2018), which show that video lectures with quizzes can be used to increase knowledge, intensify engagement, and raise attention.

Although it is possible to measure student performance through the results of quizzes in class, a relevant factor to be studied is the involvement of students in the execution of tasks

## OPEN ACCESS

### Edited by:

Guilherme Lepski,  
University of São Paulo, Brazil

### Reviewed by:

Eleonora Gentile,  
University of Bari Aldo Moro, Italy  
Aykut Eken,  
The Institute of Photonic Sciences  
(ICFO), Spain

### \*Correspondence:

Amanda Yumi Ambriola Oku  
amanda.yumi@ufabc.edu.br

### Specialty section:

This article was submitted to  
Cognitive Neuroscience,  
a section of the journal  
Frontiers in Human Neuroscience

**Received:** 28 October 2020

**Accepted:** 08 January 2021

**Published:** 05 February 2021

### Citation:

Oku AYA and Sato JR (2021)  
Predicting Student Performance Using  
Machine Learning in fNIRS Data.  
Front. Hum. Neurosci. 15:622224.  
doi: 10.3389/fnhum.2021.622224

through the mapping of brain states during the task. Usually, cognitive neuroscience experiments study psychological processes through controlled manipulations, reducing the behavior of one of its components. However, this framework is not suitable when one wishes to generalize the characteristics of new situations from full descriptions of the behavior (Varoquaux and Poldrack, 2018). For instance, Barreto et al. (2020) and Noah et al. (2015) indicate the importance that studies involving music and dance be carried out under natural conditions. Similarly, Lamb et al. (2018) performs experiments under naturalistic conditions for the evaluation of science education.

We address this issue by performing an experiment in a more realistic setting. Specifically, we collected brain data with a fNIRS (functional near-infrared spectroscopy) device from students while they were watching a video lecture and answering questions. The fNIRS device was chosen due to its acquisition systems that collect data of hemodynamic states in several brain regions in a naturalistic, comfortable, and safe manner for participants (Noah et al., 2015). Safe levels of light (with wavelengths between 650 and 1,000 nm) were used to infer the variation in the level of oxygenation of brain tissue in a non-invasive way, which penetrates the biological tissue and reaches the cortex, allowing the analysis of oxygenation, hemoglobin ( $\text{HbO}_2$ ), deoxyhemoglobin (HHb) and total hemoglobin (tHb;  $\text{tHb} = \text{HbO}_2 + \text{HHb}$ ) from cerebral blood (Delpy and Cope, 1997). The fNIRS technical limitations include superficial depth cortical evaluation (Ferrari et al., 2004). Specifically, we collected fNIRS data from the Prefrontal cortex (PFC).

The PFC has a central role in cognitive control. It has interconnections with brain areas that process external information (with all the sensory systems and structures of the cortical and subcortical motor system) and with internal information (limbic and midbrain structures involved in affection, memory, and reward). It has access and the means to influence processing in all major forebrain systems and can provide a means of synthesizing the various sources of information related to a given objective (Miller et al., 2002). McGuire and Botvinick (2010) shows there are indications that prefrontal cortex neurons appear to have a crucial ability for cognitive control, transmitting knowledge about a specific goal-directed task. Furthermore, Lamb et al. (2018) shows that fNIRS imaging of the prefrontal cortex can be useful to educators, since this region is responsible for problem solving, memory, and social behavior. However, this study also shows that tasks involving large amounts of unstructured processing, such as video lectures, can be challenging, since they generate less dynamic response within the prefrontal cortex than structured tasks.

In this paper, the fNIRS data from the PFC was used to create predictors for a student's answers. These predictors were obtained by applying machine learning algorithms to the data. In particular, we used random forests and penalized logistic regression algorithms. These algorithms allow one to understand the structure of existing data and generate prediction rules for new observations.

## 2. MATERIALS AND METHODS

### 2.1. Participants

A total of 21 participants were recruited for participation but 3 of them were excluded (one for low signal quality and two for not meeting the health requirements). All 18 participants (10 female, 8 male) were right-handed, had normal vision and hearing, and mean age  $25.6 \pm 4.6$  (range 18–40 years). No subject had a history of neurological or psychiatric disorders. Participants were recruited among undergraduate and graduate students in fields of Science. All participants alleged to have little or no prior knowledge in Astronomy. Signed consent was obtained from all members prior to participation. The Federal University of ABC - Ethics Committee approved the experiment. The experiment was performed in accordance with all local relevant guidelines and regulations. All subjects participated voluntarily and without any financial compensation, as required by federal laws.

### 2.2. Experiment

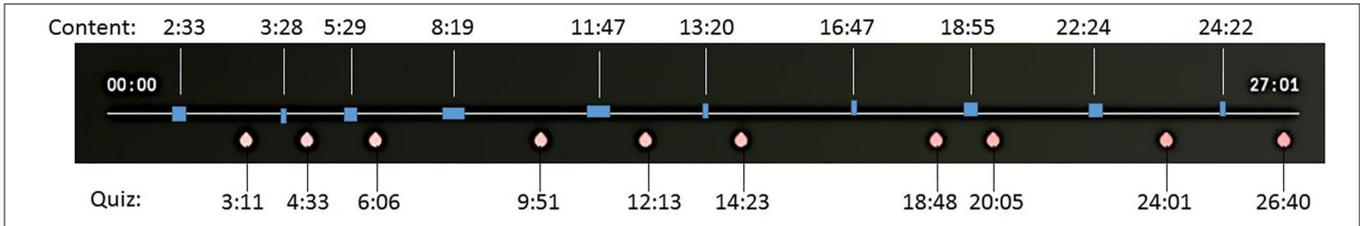
The experiment's tasks consisted of watching the first class in an Astronomy course while answering several multiple choice questions. The class was entitled "Astronomy: A general introduction", and was chosen from a publicly available e-learning course from the Virtual University of São Paulo State (UNIVESP). The video's content usually does not belong to the basic education curriculum and requires reasoning and attention for understanding calculations and order of events. It was chosen since it brings new content to most students and does not require a large amount of previous knowledge.

Before running the main experiment, we tested the hypothesis that answering correctly depended on watching the video. This hypothesis was tested by applying the a quiz with multiple choice questions to a control group with 116 participants who did not watch the video lecture. The probability of a correct answer without watching the video was found based on a binomial test. The test did not reject the hypothesis that, without answering the video, participants answer correctly no better than by chance.

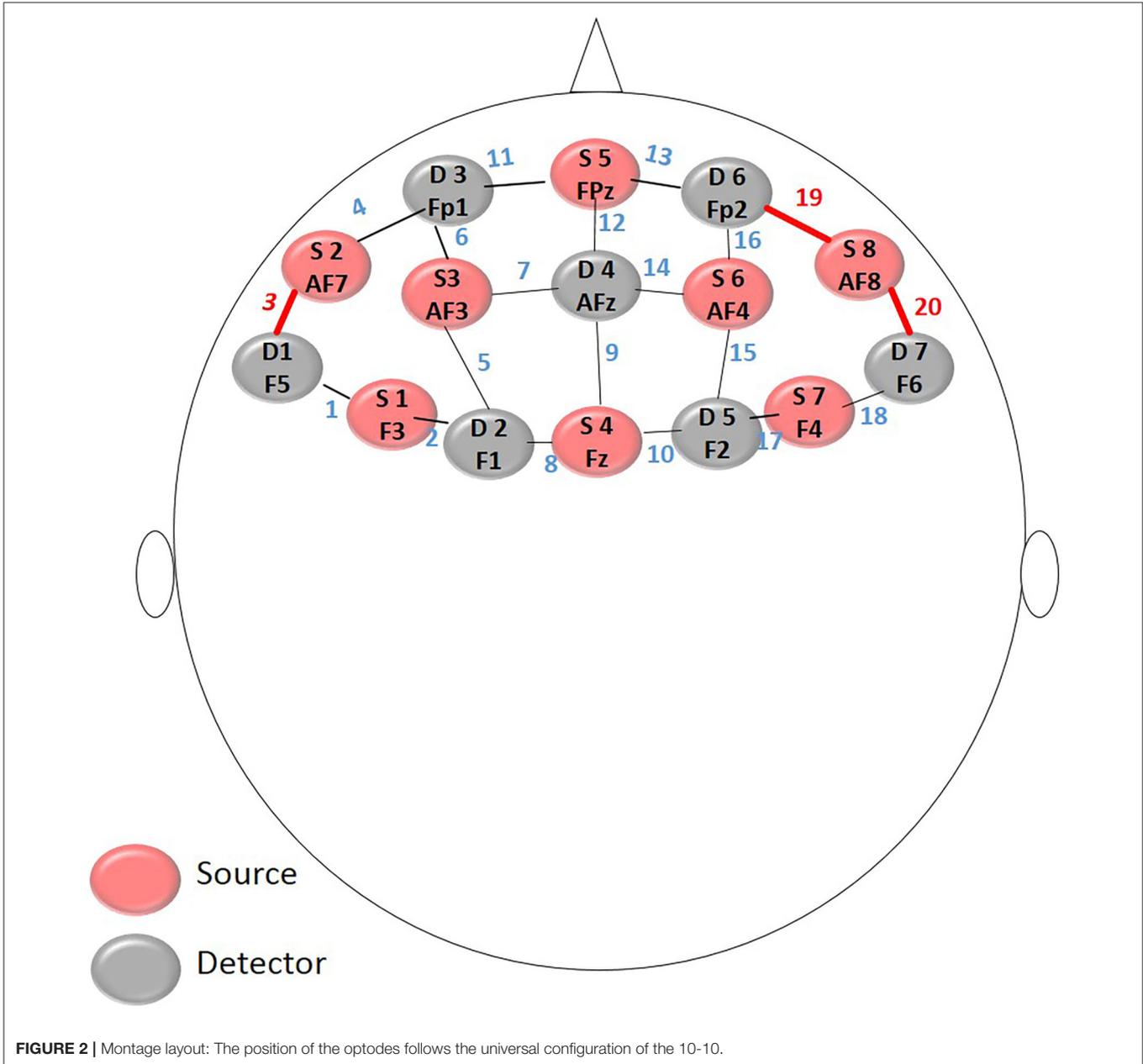
The main experiment was performed using Edpuzzle (<http://edpuzzle.com/>), an American platform for online learning. This platform was validated by Abou Afach et al. (2018) and is used by colleges, open courses, and universities. It was also validated in Brazil by researchers in education, which signaled it could be used successfully by local students (Lombardi and Gitahy, 2017).

We collected data of functional near-infrared spectroscopy (fNIRS) placed over the PFC (responsible for planning complex cognitive behavior, decision making, and moderating social behavior) of 18 undergraduate and graduate students using NIRSport equipment (company NIRx Medical Technologies). In the experiment, subjects were seated in a comfortable chair in a quiet and ventilated room.

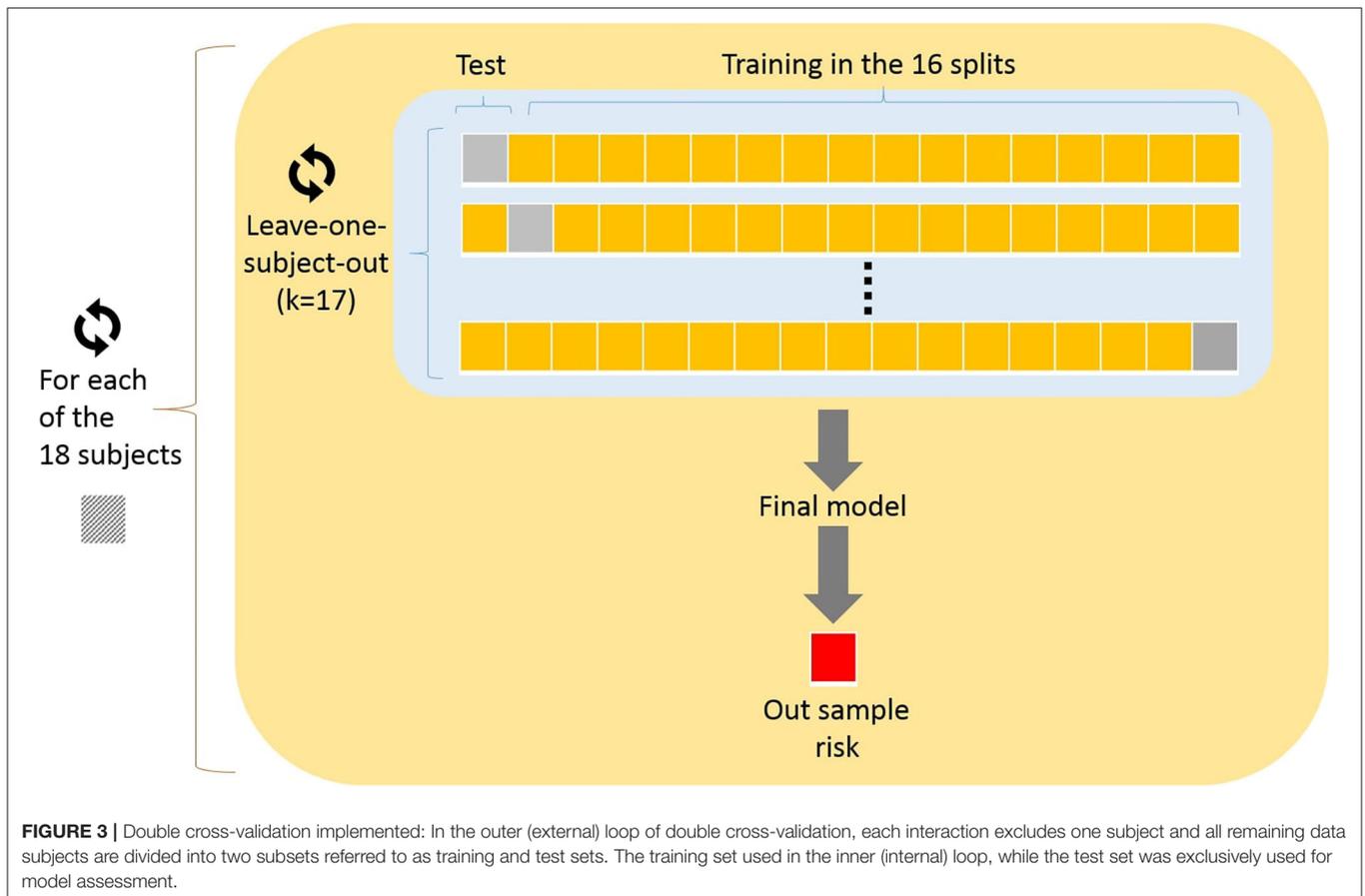
The subjects were asked to relax and to remain still during the experiment. They watched a free recorded lecture (27 min) with 10 multiple-choice exercises (**Figure 1**). As in real classroom situations, there was no indication of the times that they would be asked future questions.



**FIGURE 1** | The questions are based on content exposed at earlier times throughout the video (indicated in blue). The red dots show the exact timing of the questions.



**FIGURE 2** | Montage layout: The position of the optodes follows the universal configuration of the 10-10.



## 2.3. Data Acquisition

The position of the optodes follows the universal configuration of the 10-10 electroencephalogram (EEG) system (Koessler et al., 2009). The 8 emitters and 7 detectors are positioned in the form: Sources in F3, AF7, AF3, Fz, Fpz, AF4, AF8, F4 and the Detectors in F5, F1, Fp1, AFz, F2, Fp2, F6 under an approximate distance of 3 cm between the optodes and resulting in the collection of oxyhemoglobin and deoxyhemoglobin from 20 channels, as **Figure 2**.

The recording of the PFC region was conducted on a multi-channel continuous wave system using NIRSport equipment (company NIRx Medical Technologies). This system consists of 8 illumination sources and 8 detection sensors with two wavelengths of 760–850 nm. The sampling rate of NIRSport is 62.5 Hz, as the device implements time multiplexing, which means that only one LED is turned on at each time, the sampling rate for each data channel is 7.81 Hz. The data were recorded by a computer during the measurements using NIRStar software (NIRx Medizintechnik GmbH, Berlin, Germany)

## 2.4. Data Preprocessing

Raw data from the NIRStar were processed using the NIRSLab-2014 (NIRx Medizintechnik GmbH, Berlin, Germany) via the Matlab 2007b (Mathworks, Natick, MA, USA) (Xu et al., 2014) software using a 0.01–0.2 Hz bandpass filter to reduce

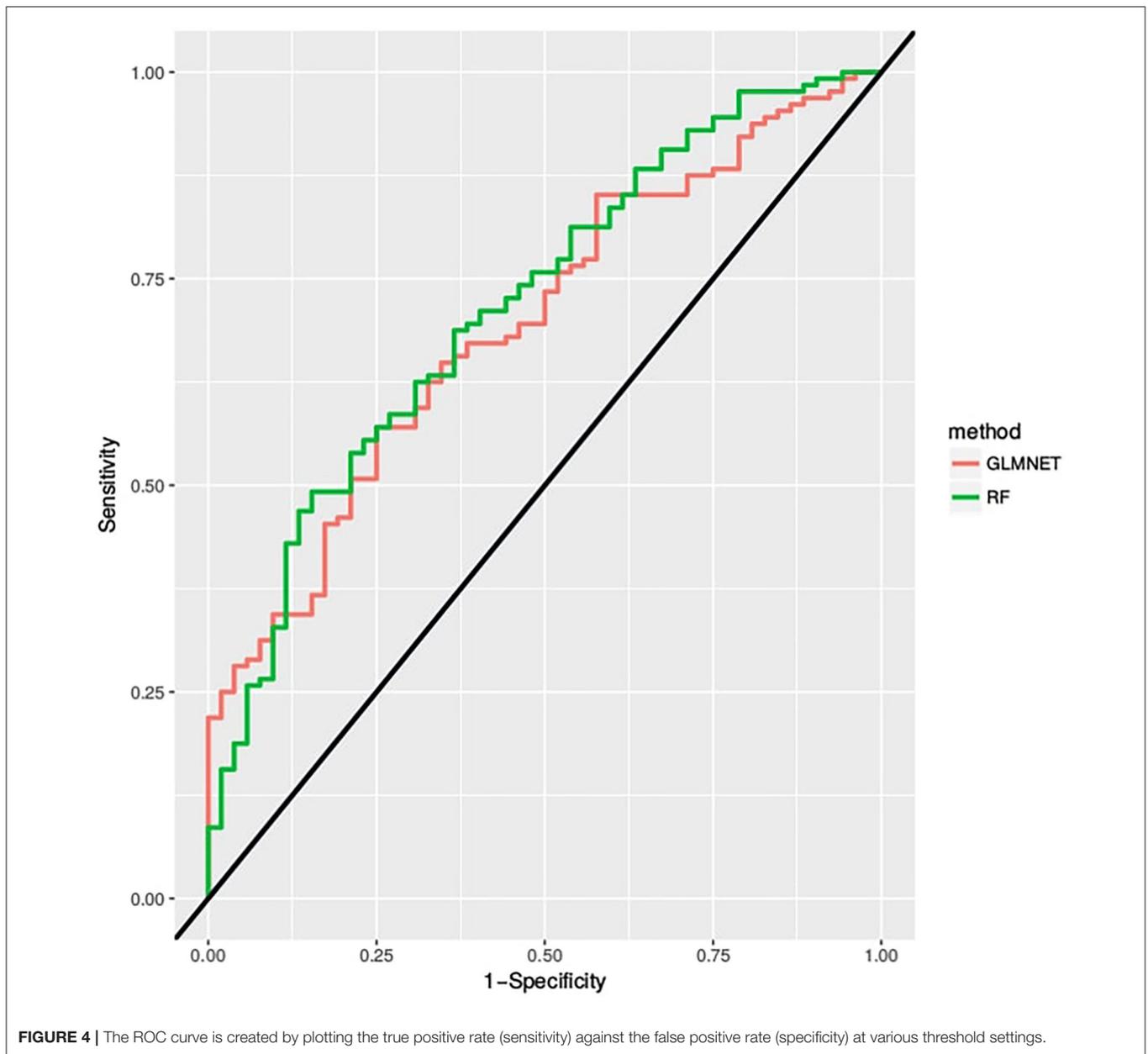
**TABLE 1** | Confusion matrix—random forest.

Predicted \ Actual	Incorrect	Correct
Incorrect	34	44
Correct	18	84

**TABLE 2** | Confusion matrix—GLMNET.

Predicted \ Actual	Incorrect	Correct
Incorrect	32	46
Correct	20	82

physiological signal artifacts at the cutoff frequencies of the global deviations ( $< 0.01$  Hz), systemic interferences such as respiration rate ( $> 0.2$  Hz) and cardiac cycles ( $> 0.5$  Hz). We used the modified Beer-Lambert law (Mesquita and Covolán, 2008), to find the variations in oxygenated hemoglobin (HbO<sub>2</sub>) and deoxygenated hemoglobin (HHb) cited by Delpy and Cope (1997). We removed some motion artifacts manually (spikes) where HbO<sub>2</sub> and HHb increased or decreased in unison based on visual inspection of the record (Lloyd-Fox et al., 2010). Afterward, we used the mean of the entire timeline as a baseline



and differential path length factor (DPF) of 7.25 for the 760 nm and 6 wave, 38 for 850 nm lengths.

After computing the states of oxyhemoglobin ( $\text{HbO}_2$ ) and deoxyhemoglobin (HHb), the signal was averaged and grouped according to 10 exercises and 18 students, totaling 180 observations over 20 channels. The signal's standard deviation was also computed in each of these groups. However, since this feature did not improve the statistical analysis, it was not used in the final model.

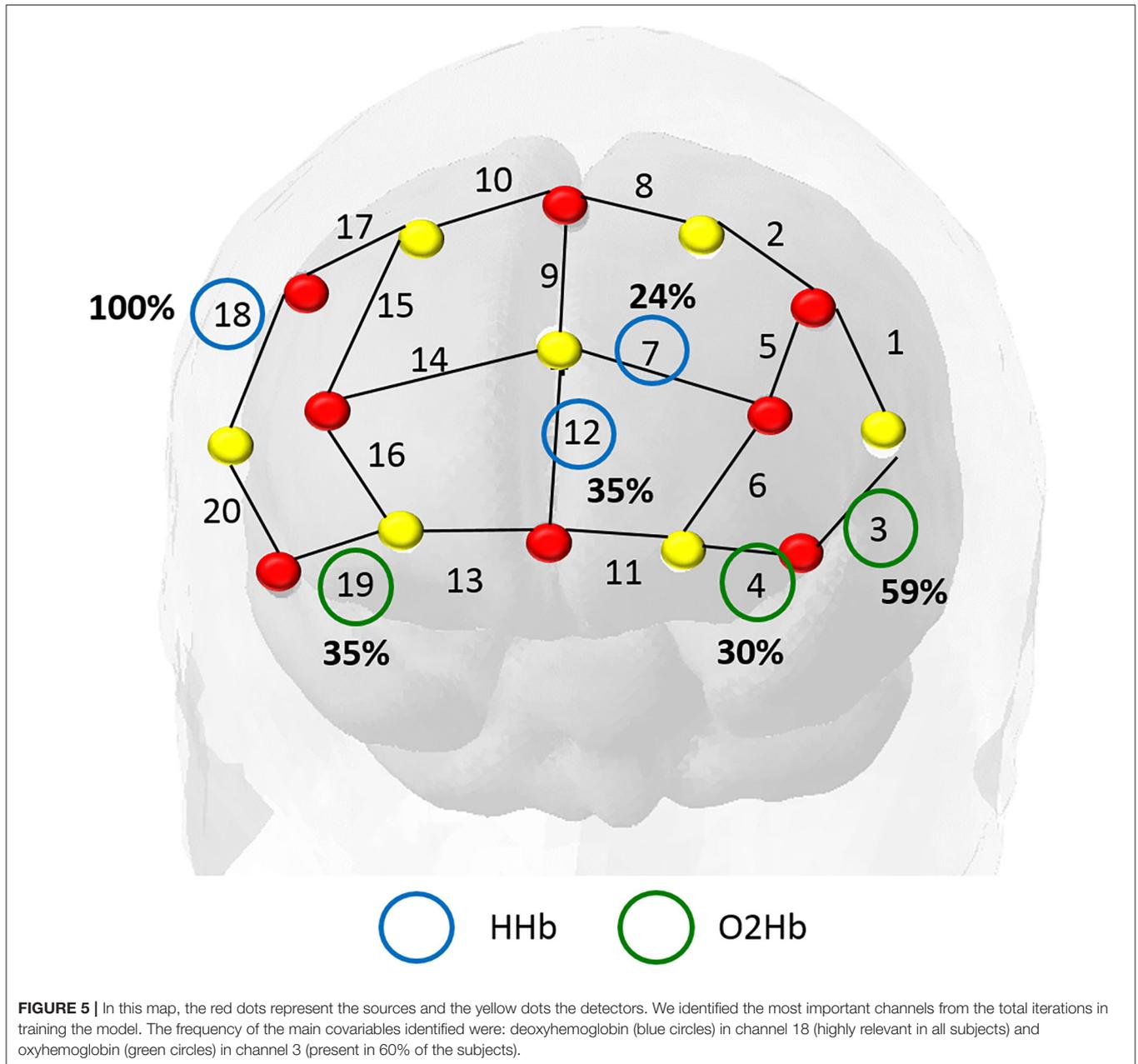
## 2.5. Statistical Analysis

All learning algorithms were implemented in the R language (4.0.3 version). The “magrittr” and “tidyverse” packages were used in building the final database. The packages “randomForest”

and “GLMNET” were used for fitting the Random Forest and Penalized Logistic Regression classifiers. Also, the “ROCR” package was used for performance analysis.

Logistic regression performs binary classification (dichotomous output labels), returning the probability that the object belongs to each class. In this way, the cost function can be the difference between the predicted probability and label 0 or 1. This cost can be estimated by calculating the average loss over all objects in a test set, similarly as done in linear regression.

Simple logistic regression can cause overfitting when dealing with many covariates. To mitigate this problem, we applied LASSO (least absolute shrinkage and selection operator) to our data. This is a regularization method that penalizes large parameter values and usually yields solutions in which the



estimates of several of the parameters are zero (sparse solutions). This method is done by maximizing the log-likelihood added by a penalty factor. More details about LASSO can be found in section A.1 of the **Appendix**.

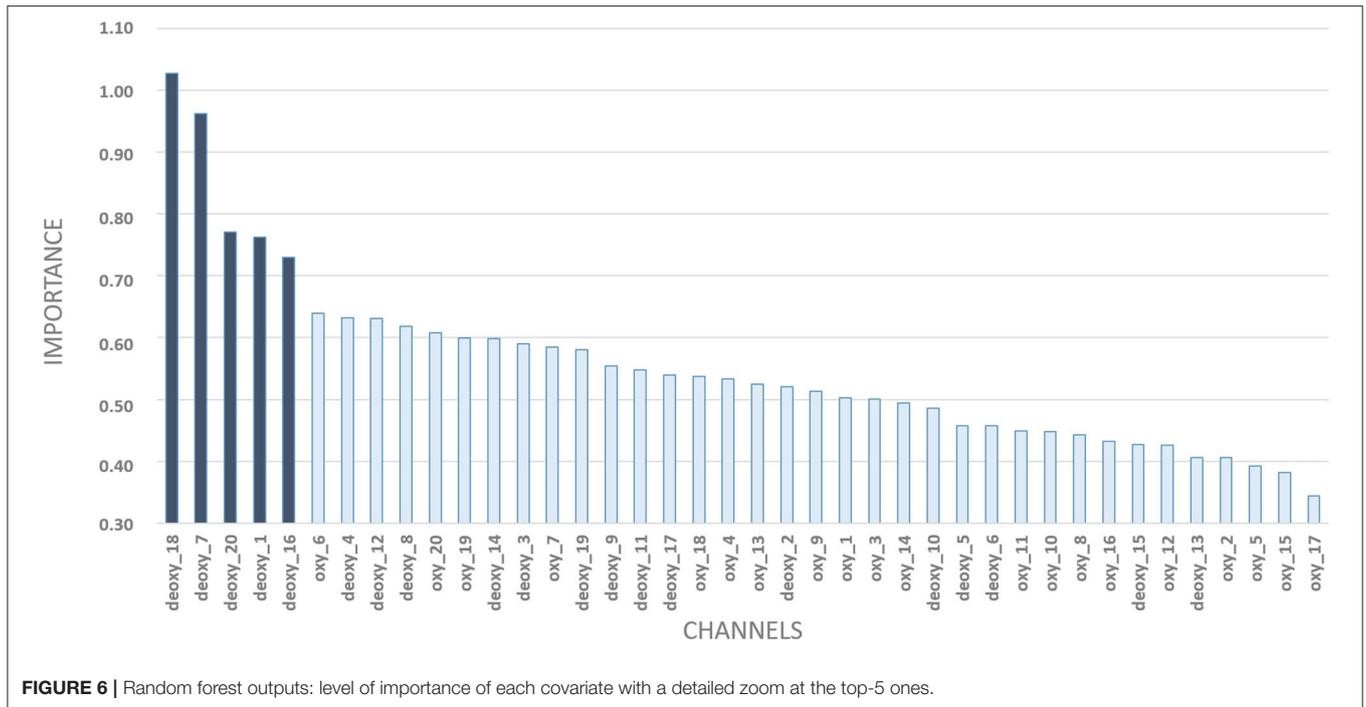
## 2.6. Cross Validation

Both our algorithms (Random Forest and GLMNET with LASSO) involved training 180-response BD (10 video ranges for each of the 18 subjects). Each of these has 40 covariates for prediction [mean ( $\text{HbO}_2$ ) and mean (HHb) for each of the 20 channels obtained in each video snippet].

Using a small database to learn the parameters of a prediction function and testing it on the same data can find a perfect score

but would fail to predict yet-unseen data. This situation is called overfitting and can be overcome by cross-validation.

The performance of Random Forest and LASSO logistic regression was evaluated using different types of cross-validation. The Random Forest was evaluated using simple leave-one-subject-out cross-validation. Also, we assessed the performance of LASSO logistic regression using double cross-validation (leave-one-subject-out) as illustrated in **Figure 3**. The double cross-validation process implemented comprises two nested cross-validation loops which are referred to as internal and external cross-validation loops. In the outer (external) loop of double cross-validation, each interaction excludes one subject and all remaining data subjects are divided into two subsets



**FIGURE 6** | Random forest outputs: level of importance of each covariate with a detailed zoom at the top-5 ones.

referred to as training and test sets. The training set used in the inner (internal) loop of double cross-validation for model building and model selection, while the test set was exclusively used for model assessment.

### 3. RESULTS

The Random Forest and the GLMNET obtained satisfactory results with, respectively, areas of 0.67 and 0.65 under the ROC curve in **Figure 4**. Also, We represented through the confusion matrix of both algorithms (**Tables 1, 2**) the instances of the predicted classes: Each row represents the instances of the predicted model while the column represents the real results of the students' performance. Both models obtained a good fit on identifying actual right answers (correct/correct) and wrong answers (incorrect/incorrect). The GLMNET LASSO had an accuracy of  $0.63 \pm 0.036$ , a sensitivity of  $0.62 \pm 0.067$ , a specificity of  $0.64 \pm 0.042$ , and a Cohen's kappa coefficient of 0.22 (fair on the Kappa scale). The random forest had a slightly better result, with an accuracy of  $0.66 \pm 0.035$ , a sensitivity of  $0.63 \pm 0.066$ , a specificity of  $0.66 \pm 0.042$ , and a Cohen's kappa coefficient of 0.26 (fair on the Kappa scale).

We also showed that the models are in fact better than chance through a permutation test, which evaluates whether the model is uninformative. This test can be easily applied to a wide range of statistical learning methods, including some in which a measure of variability is difficult to obtain and is not automatically produced by the statistical software (Friedman et al., 2001).

We repeated the same procedure of adjusting the models with the shuffled response variables and calculated the AUC (area

under the ROC curve) for each one of the 1,000 iterations. The total number of cases that resulted in a better model than the original was 3 cases for the Random Forest, thus obtaining a  $p$ -value of 0.003 (thus rejecting the null hypothesis) and the total number of cases that resulted in a better model than the original was 1 in GLMNET, thus obtaining a  $p$ -value of 0.001 (also rejecting the null hypothesis).

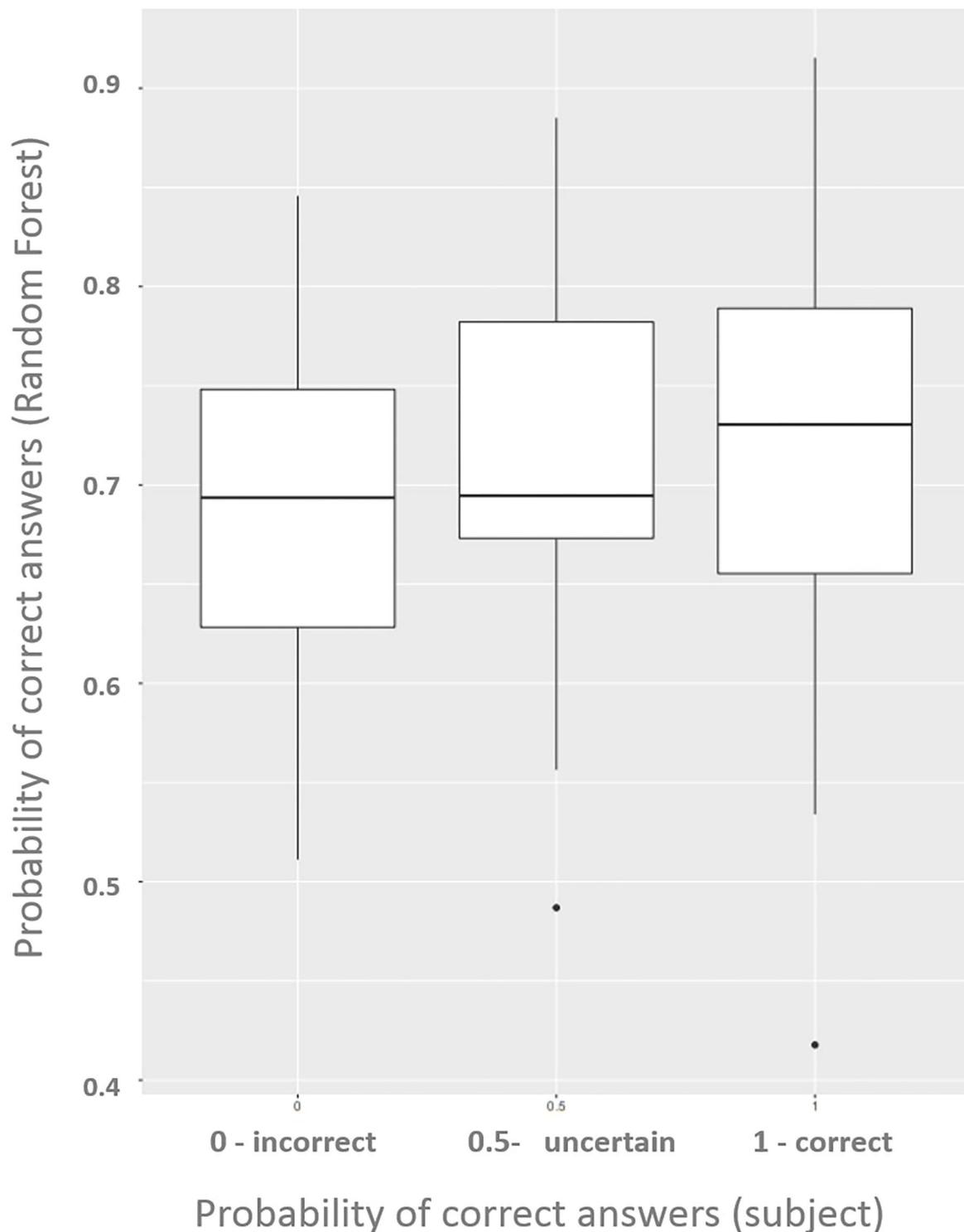
The output of the models identified which channels resulted in better predictors for the exercises.

#### 3.1. Main Predictors – Penalized Logistic Regression

For the GLMNET model, we calculated the frequency of the selected channels in each iteration of the outer loop of the cross-validation, as displayed in **Figure 5**. We verified that the covariates (HHb) in channel 18 (referring to regions F4-F6 in the 10-10 system) and the (HbO<sub>2</sub>) in channel 3 (F5-AF7) had greater weight in the prediction, being used in, respectively, 100 and 59% of the subjects.

The relevant channels according to this model are the areas of channel 4 (AF7- F5), and channel 18, regions F4-F6, both corresponding to middle frontal cortex (Koessler et al., 2009; Balconi and Fronda, 2020). The region belongs to the dorsolateral prefrontal cortex (Bandeira et al., 2019) which is associated with the cognitive process, working memory, cognitive flexibility, planning, inhibition, and abstract reasoning (Zgaljardic et al., 2010).

As for the most important channels for each of the models, it is worth mentioning that the penalty of the channels in the GLMNET with LASSO does not imply that they are not

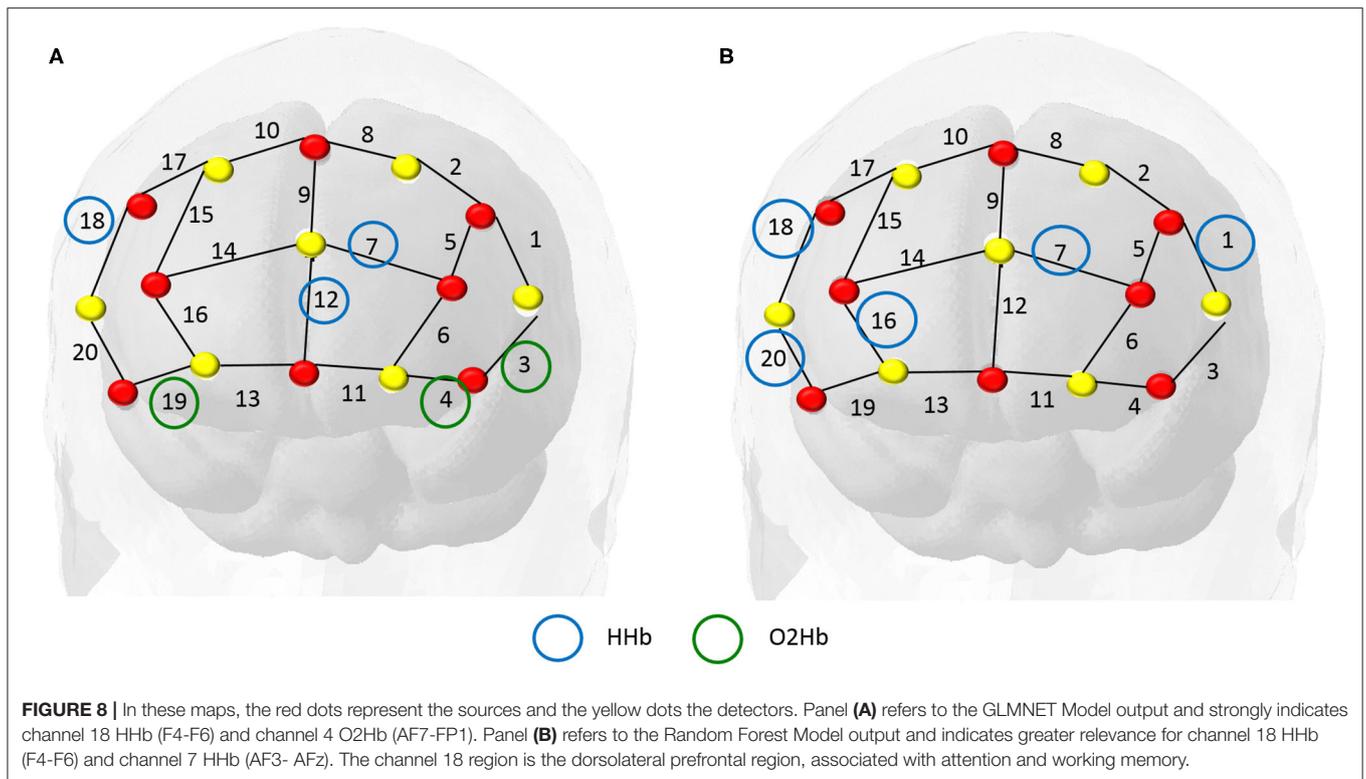


**FIGURE 7** | Boxplots show differences between the groups: 1, certainly right exercise; 0.5, not sure/next idea; 0, probably wrong/random guess.

explanatory for the response variable, but rather, there may be a correlation with another channel that is explanatory and therefore suffered a penalty.

### 3.2. Main Predictors—Random Forest

The Random Forest Model indicated high predictive power from the covariates (HHb) in channel 18 (Figure 6). Besides this



channel, the following were the most relevant: 7 (AF3-AFz), 20 (AF8-F6), 1 (F3 and F5), and 16 (AF4 and Fp2). In addition to working memory, they also show semantic aspects of language.

### 3.3. Hits Expectations vs. Prediction

In addition to identifying which channels are more explanatory for identifying the errors in the questions per individual, it was also possible to evaluate the levels of student involvement in interactive classes. We analyzed which types of questions are more difficult to answer by comparing the error rate with moments when the students declared to have lost concentration.

We compared the results of the random forest prediction with what the volunteers believed they had got right and mentioned at the end of the experiment. The **Figure 7** shows how the model differentiates hits and errors using only signals of oxyhemoglobin and deoxyhemoglobin in each question.

The Random Forest model indicates a slightly higher probability of correct answers among the cases in which the subjects believe they have hit the exercise. Also, it indicates a low probability of correct answers for the cases in which the subjects declared to have felt indecisive or believed to have mistaken the question (in this case, with low differentiation between them).

For the training of the models, it was necessary to identify the hemodynamic signs linked to the questions. We conducted tests to assure that the questions alone were not enough to predict student successes and errors (which would show an error in the design of the experiment).

The analyzed regions of the experiment are only suitable for exercises with the fixation of theoretical content. Mathematical reasoning, calculation, and perception have not been validated.

## 4. DISCUSSION

In this research, we fit a predictive model for a students' correctness of answers in an interactive class based on PFC activity. These models allowed the identification of which regions are most relevant and influence results the most.

Both models (**Figure 8**) indicated that the information from channels F4-F6 (based on the EEG 10-10 system) had the greatest impact on the predictive model (**Figure 4**), suggesting a significant contribution to language understanding and semantic decision tasks.

Our models are consistent with other articles in the literature. For instance, (Liu and Ayaz, 2018) shows that perceived speech can be identified from the listeners' brain signals measured with fNIRS and (Herff et al., 2014) shows that measuring hemodynamic responses in the PFC with fNIRS, they showed the degree of workload a subject was experiencing, instead of only identify if there was an engagement during the tasks. Furthermore, MacDonald et al. (2000) and Dosenbach et al. (2006) use fNIRS data to show that brain activity can distinguish between high and low levels of task engagement. Specifically, they detected differences in the brain activity in the dorsolateral prefrontal cortex (*dorsolateral prefrontal cortex*—DLPFC) while participants alternated between performing and not performing a cognitive task.

With error rates in the models below 30%, our work can be suggested to assess levels of student involvement in tasks to validate new teaching content through videos, allowing us to evaluate whether students can assimilate content from fNIRS signals.

Despite the results obtained, the study has some limitations. For instance, the model considers the NIRS signal related to a single video lesson. Further studies are needed to have more information about students' behavior and performance during the task. Also, in the collection of fNIRS data in this experiment, we did not use short distance detectors, which could assist in the exclusion of extracerebral signals around the sources (Tachtsidis and Scholkmann, 2016).

An unexpected result was the high importance of HHb in both predictive models. Usually fNIRS studies indicate a high influence of HHbO<sub>2</sub> on results, with higher signal-to-noise ratio SNR than HHb. Fishburn et al. (2014) shows the fNIRS sensitivity to detect linear changes in activation and functional connectivity in response to cognitive load, using HHbO<sub>2</sub> and HHb had low correspondence. Also, Fishburn et al. (2014), Leon-Dominguez et al. (2014), and Barreto et al. (2020) show significant results for HHb. The sensitivity and SNR are core parameters during the fNIRS measurement and from the results obtained, further investigation is needed regarding the importance of HHb data in the models and new systematic analysis of SNR.

Since our primary goal was limited to investigating the PFC, we did not acquire signals from other brain regions. Although this assembly of optodes provides favorable conditions for more realistic situations, complementary studies with Functional Magnetic Resonance Imaging (fMRI) could perform to accurately identify other brain regions and also identify a precise location of Brodmann's areas involved during the task.

This study opens perspectives for a better understanding of the PFC during the execution of tasks and experiments in real situations. For further studies, we understand that it is important

to continue assessing the level of sustained attention of students from hemodynamic states through models for classifying the involvement in the task rather than subtasking specific tasks.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Federal University of ABC - Ethics Committee approved all this experiment. The experiment was performed in accordance with all local relevant guidelines and regulations. The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

AO and JS: designed the study, collected, and analyzed the data and revised the manuscript. All authors have read and agreed to the published version of the manuscript.

## FUNDING

The authors are thankful to São Paulo Research Foundation (FAPESP grants number 2018/04654-9; 2018/21934-5; 2019/17907-5).

## ACKNOWLEDGMENTS

The authors are grateful to NIRx for technological support and Virtual University of São Paulo State (UNIVESP), specially Dr. João Evangelista Steiner for the didactic content and the support offered.

## REFERENCES

- Abou Afach, S., Kiwan, E., and Semaan, C. (2018). How to enhance awareness on bullying for special needs students using "edpuzzle" a web 2.0 tool. *Online Submission* 3, 1–7. doi: 10.24331/ijere.372260
- Balconi, M., and Fronza, G. (2020). Morality and management: an oxymoron? fNIRS and neuromanagement perspective explain us why things are not like this. *Cogn. Affect. Behav. Neurosci.* 20, 1–13. doi: 10.3758/s13415-020-00841-1
- Bandeira, J. S., Antunes, L. C., Soldatelli, M. D., Sato, J. R., Fregni, F., and Caumo, W. (2019). Functional spectroscopy mapping of pain processing cortical areas during non-painful peripheral electrical stimulation of the accessory spinal nerve. *Front. Hum. Neurosci.* 13:200. doi: 10.3389/fnhum.2019.00200
- Barreto, C. D. S. F., Morais, G. A. Z., Vanzella, P., and Sato, J. R. (2020). Combining the intersubject correlation analysis and the multivariate distance matrix regression to evaluate associations between fNIRS signals and behavioral data from ecological experiments. *Exp. Brain Res.* 238, 2399–2408. doi: 10.1007/s00221-020-05895-8
- Delpy, D., and Cope, M. (1997). Quantification in tissue near-infrared spectroscopy. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 352, 649–659. doi: 10.1098/rstb.1997.0046
- Dosenbach, N. U., Visscher, K. M., Palmer, E. D., Miezin, F. M., Wenger, K. K., Kang, H. C., et al. (2006). A core system for the implementation of task sets. *Neuron* 50, 799–812. doi: 10.1016/j.neuron.2006.04.031
- Ferrari, M., Mottola, L., and Quaresima, V. (2004). Principles, techniques, and limitations of near infrared spectroscopy. *Can. J. Appl. Physiol.* 29, 463–487. doi: 10.1139/h04-031
- Fishburn, F. A., Norr, M. E., Medvedev, A. V., and Vaidya, C. J. (2014). Sensitivity of fNIRS to cognitive state and load. *Front. Hum. Neurosci.* 8:76. doi: 10.3389/fnhum.2014.00076
- Friedman, J., Hastie, T., and Tibshirani, R. (2001). *The Elements of Statistical Learning*, Vol. 1. New York, NY: Springer. doi: 10.1007/978-0-387-21606-5\_1
- Herff, C., Heger, D., Fortmann, O., Hennrich, J., Putze, F., and Schultz, T. (2014). Mental workload during n-back task-quantified in the prefrontal cortex using fNIRS. *Front. Hum. Neurosci.* 7:935. doi: 10.3389/fnhum.2013.00935
- Jonassen, D., Davidson, M., Collins, M., Campbell, J., and Haag, B. B. (1995). Constructivism and computer-mediated communication in distance education. *Am. J. Distance Educ.* 9, 7–26. doi: 10.1080/08923649509526885
- Koessler, L., Maillard, L., Benhadid, A., Vignal, J. P., Felblinger, J., Vespignani, H., et al. (2009). Automated cortical projection of eeg sensors: anatomical correlation via the international 10-10 system. *Neuroimage* 46, 64–72. doi: 10.1016/j.neuroimage.2009.02.006

- Lamb, R., Antonenko, P., Etopio, E., and Seccia, A. (2018). Comparison of virtual reality and hands on activities in science education via functional near infrared spectroscopy. *Comput. Educ.* 124, 14–26. doi: 10.1016/j.compedu.2018.05.014
- Leon-Dominguez, U., Izzetoglu, M., Leon-Carrion, J., Solis-Marcos, I., Garcia-Torrado, F. J., Forastero-Rodriguez, A., et al. (2014). Molecular concentration of deoxyhb in human prefrontal cortex predicts the emergence and suppression of consciousness. *Neuroimage* 85, 616–625. doi: 10.1016/j.neuroimage.2013.07.023
- Liu, Y., and Ayaz, H. (2018). Speech recognition via fNIRS based brain signals. *Front. Neurosci.* 12:695. doi: 10.3389/fnins.2018.00695
- Lloyd-Fox, S., Blasi, A., and Elwell, C. (2010). Illuminating the developing brain: the past, present and future of functional near infrared spectroscopy. *Neurosci. Biobehav. Rev.* 34, 269–284. doi: 10.1016/j.neubiorev.2009.07.008
- Lombardi, E., and Gitahy, R. R. C. (2017). O uso das tecnologias móveis na educação: uma experiência com o ed puzzle. *Capes/prosup Presidente Prudente-SP* 14, 590–594. doi: 10.5747/ch.2017.v14.nesp.000997
- MacDonald, A. W., Cohen, J. D., Stenger, V. A., and Carter, C. S. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science* 288, 1835–1838. doi: 10.1126/science.288.5472.1835
- McGuire, J. T., and Botvinick, M. M. (2010). Prefrontal cortex, cognitive control, and the registration of decision costs. *Proc. Natl. Acad. Sci. U.S.A.* 107, 7922–7926. doi: 10.1073/pnas.0910662107
- Mesquita, R. C., and Covolan, R. J. (2008). Estudo funcional do cérebro através de NIRS e tomografia Óptica de difusão. *Neurociências e Epilepsia* 1:147.
- Miller, E. K., Freedman, D. J., and Wallis, J. D. (2002). The prefrontal cortex: categories, concepts and cognition. *Philos. Trans. R. Soc. B Biol. Sci.* 357, 1123–1136. doi: 10.1098/rstb.2002.1099
- Noah, J. A., Ono, Y., Nomoto, Y., Shimada, S., Tachibana, A., Zhang, X., et al. (2015). fMRI validation of fNIRS measurements during a naturalistic task. *J. Visual. Exp.* 100:e52116. doi: 10.3791/52116
- Tachtsidis, I., and Scholkmann, F. (2016). False positives and false negatives in functional near-infrared spectroscopy: issues, challenges, and the way forward. *Neurophotonics* 3:031405. doi: 10.1117/1.NPh.3.3.031405
- Varoquaux, G., and Poldrack, R. (2018). Predictive models can overcome reductionism in cognitive neuroimaging. *Curr. Opin. Neurobiol.* 55, 1–6. doi: 10.1016/j.conb.2018.11.002
- Wachtler, J., Scherz, M., and Ebner, M. (2018). “Increasing learning efficiency and quality of students’ homework by attendance monitoring and polls at interactive learning videos,” in *EdMedia+ Innovate Learning*, eds T. Bastiaens, J. Van Braak, M. Brown, L. Cantoni, M. Castro, R. Christensen, G. V. Davidson-Shivers, K. DePryck, M. Ebner, M. Fominykh, C. Fulford, S. Hatzipanagos, G. Knezek, K. Kreijns, G. Marks, E. Sointu, E. K. Sorensen, J. Viteli, J. Voogt, P. Weber, E. Weippl, O. Zawacki-Richter (Amsterdam; Waynesville, NC: Association for the Advancement of Computing in Education), 1357–1367.
- Xu, Y., Graber, H. L., and Barbour, R. L. (2014). “nirsLAB: a computing environment for fNIRS neuroimaging data analysis,” in *Biomedical Optics 2014 OSA Technical Digest* [online] (Miami, FL: Optical Society of America). doi: 10.1364/BIOMED.2014.BM3A.1
- Zgaljardic, D., Mattis, P., and Charness, A. (2010). “Executive dysfunction,” in *Encyclopedia of Movement Disorders*, eds K. Kompolti and L. V. Metman (Oxford: Academic Press), 458–462. doi: 10.1016/B978-0-12-374105-9.00176-3

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Oku and Sato. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

## A. APPENDIX

### A.1. LASSO

Logistic regression is a supervised learning method that is used for binary response variables. Let  $Y_i \in \{0, 1\}$  be a response variable and  $\mathbb{X}_i$  be a vector of covariates. In logistic regression, the logit of  $\mathbb{P}(Y_i = 1|\mathbb{X}_i)$  follows a linear equation, that is,

$$\log\left(\frac{\mathbb{P}_\beta(Y_i = 1|\mathbb{X}_i)}{1 - \mathbb{P}_\beta(Y_i = 1|\mathbb{X}_i)}\right) = \beta^t \mathbb{X}_i, \text{ where } \beta \text{ are unknown coefficients.} \quad (\text{A1})$$

Using Equation (A1), it is possible to compute the log-likelihood of coefficients,  $l(\beta)$ , for the observed sample.

$$l(\beta_0) = \sum_{i=1}^n \log(\mathbb{P}_{\beta_0}(Y_i = y_i|\mathbb{X}_i)) \quad (\text{A2})$$

The value of  $l(\beta_0)$  is a measure of how likely it is that  $\beta = \beta_0$ . Based on this interpretation, a common choice of estimator for

$\beta$  is the one which maximizes  $l(\beta_0)$ , the maximum likelihood estimator. However, this estimator can lead to overfitting when the sample size is small relatively to the number of covariates. In this case, it is common to use regularized maximum likelihood estimators.

LASSO is one alternative for performing regularized logistic regression. In this framework, one estimates  $\beta$  by maximizing

$$s(\beta_0) = l(\beta_0) - \lambda \sum_{i=1}^d |\beta_i|. \quad (\text{A3})$$

Equation A3 leads to a trade-off between how likely is  $\beta$  and how small are its values. This trade-off often avoids overfitting and leads to better estimators. Furthermore, in LASSO one uses a  $l_1$  penalty,  $\sum_{i=1}^d |\beta_i|$ . This penalty often leads to estimates for  $\beta$  that have many zeroes. That is, LASSO estimation often automatically performs feature selection.