MACHINE LEARNING CLASSIFICATION OF HEMODYNAMICS TO PREDICT SCIENCE STUDENT LEARNING OUTCOMES IN REAL-TIME DURING VIRTUAL REALITY EAND ONLINE LEARNING SESSIONS

By

Kayleigh Anne Linder

A Junior Honors Project Presented to the

Honors College

East Carolina University

In Partial Fulfillment of the

Requirements for

Graduation with Honors

by

Kayleigh Anne Linder

Greenville

May 2022

Approved by:

Associate Professor Richard L. Lamb

Neurocognition Science Laboratory
College of Education
East Carolina University

Abstract

Students' learning results in science content and practices are expected to be improved through automated interactive learning management systems and linked online video-based learning environments. The goal of this study is to see how hemodynamic response data may be used to build student-level answer predictions using machine learning algorithms in a science classroom while students are using an online learning management system. A charter school in the northeastern United States was used to recruit 40 participants (n=40), 21 females and 19 males. Students viewed a recorded film that included a 20-minute instruction and explanation of the DNA replication process. A female educator on a computer screen presented an overview of the DNA replication process during class. The findings illustrate those hemodynamic responses seen during topic presentations accurately predict student replies to subject-related questions. The results imply that hemodynamic response can be used to gauge degrees of student involvement in video-based tasks, with error rates in the predictive models below 30%. This could lead to the development of unique visual media assessment methodologies, allowing educators to assess whether students can comprehend the material.

Keywords: science student learning; online learning; fNIRS

Introductory STEM courses have one of the highest failure rates for college students compared to other course types. In addition, students who fail these STEM courses are more likely to drop out of college or transfer majors (Alzen et al., 2018). STEM is defined as a curriculum that encompasses science, technology, engineering, and math skill sets and some examples of STEM courses include psychology, chemistry, algebra, and robotics (Gonzalez & Kuenzi, 2012). Difficulties in STEM courses arise from difficulties in making material concrete. This is especially true for math and science lessons. For example, concepts like atomic structures in chemistry are difficult for students to understand without a concrete representation. Due to this, the teaching strategies that educators use are often the most important components of student understanding (Ekwueme et al., 2015). The first experiences that a student has in science and math courses, especially in higher education, affect the likelihood of the student taking more science and math courses later in their college or university experience. Also, the student's performance and motivation are greatly affected when he or she has a difficult experience in a previous STEM course (Dyrberg & Holmegaard, 2018). When students are not able to understand the material they are being taught, the risk of student failure and drop-out becomes greater (Alzen et al., 2018). Redesigning a curriculum in a classroom is a necessary part to manage change and meet student needs before they are at risk of failure. Although it is not always collected in classes, direct student feedback is increasingly valuable. At the College of Business Administration at Valparaiso University instructors implemented a "student-centered" model of teaching, the professors received positive feedback and students showed increased understanding (McCuddy et al., 2008). By implementing a method for educators to directly measure student learning in real-time, the effect is expected to be greater student understanding. Under the theoretical framework of the brain microstate, a higher level of oxygenated blood in the brain correlates to a greater level of student understanding. Therefore, reading students' hemodynamic responses using functional near-infrared spectroscopy (fNIRS) would allow one to gauge the relative difficulty of a lesson for each student and the entire class. This process allows the instructor to adjust their lesson or rework material based on the needs of the students. Another application of this research is to measure the quality of a teacher's curriculum as they are teaching it. If an educator noticed that all of their students' hemodynamic responses showed lower levels of oxygenated hemoglobin, they could deduce that the way they are teaching the curriculum is not effective for student understanding. The ability to collect hemodynamic responses and categorize them via machine learning algorithms would supply instantaneous student feedback that can factor into a teacher's curriculum development. Current approaches to receiving student feedback rely on surveys, assessments, and independent practice assignments to gauge a student's understanding of the material. However, by using machine learning algorithms to categorize students based on a performance level, the teacher could know what aspects of their lesson they need to change while they are teaching it. If students' hemodynamic responses argue for a greater level of student understanding except during one section of the lesson, then it can be argued that the way that material is taught needs to be adjusted to better fit student needs. In addition, these responses allow an educator to immediately know whether a student understands the material without having to plan and grade an assessment first. Collecting this feedback automatically using machine learning technology saves teachers time when adjusting lessons and enhances the understanding of how to tailor a topic to a student's needs (How & Hung, 2019).

Purpose, Research Questions, and Hypothesis

The purpose of this study is to compare the hemodynamic responses of students at different performance levels as they watch science video content. Based upon hemodynamic responses derived from fNIRS data occurring during video-viewing, the author looks to predict student proficiency related to content understanding via machine learning. Research Question 1, what is the range of variability in neurocognitive data as students watch a science-based video? Research Question 2, based on hemodynamic response data collected during the video review, how well does the data fit to the machine learning algorithm, and does it allow classification of student learning outcomes? Hypothesis 1 is that the neurocognitive data compiled during the video review can be grouped into significant quartiles. Hypothesis 2 is that the quartiles of hemodynamic response can be classified by machine learning into multiple significant categories.

Theoretical Framework

Using the underlying framework of the brain microstate, the hemodynamic responses of the brain are measured using the time-dependent oxygenation and deoxygenation in areas of the brain that process various cognitive inputs from the environment (Sikka et al., 2020). Cognitive inputs encompass internal stimuli including memory and external stimuli including educational instruction such as a lecture or video. The activations (hemodynamic responses) within the brain can be connected to its functional state, and the oxygenation of brain tissue reflects the brain's state. Therefore, researchers can use the state of the brain before and after oxygenation to determine when and where the brain is activating in response to a specific task. This method is foundational for using measurement tools such as the fMRI and fNIRS since determining functional states and their sequence in the brain is part of the tools' core function in research. In this study, each microstate of the brain is connected to short-term processing moments that make up a unique cognitive process and task complex. These processes do not appear in random order and form patterns over time, which supplies the ability to predict future processes and outcomes in the brain (Rathore et al., 2017). This framework relies on the basis that the hemodynamic response in localized and global regions of the brain will remain stable over time until a new stimulus is introduced and causes a new response. Hemodynamic states are translated into the functional states of the brain as the researchers pose a study that starts with descriptive actions and moves to studying the mechanics of the actions. The mechanics are used to implement a causal understanding of the states that results in a related prediction. The act of translating hemodynamic responses into information processing states deepens the understanding of the stages at which information is processed rather than focusing on the blood flow in the brain. These patterns of blood flow concerning the activities are used for future states and analysis when needed, such as when a student is answering a content-related question (Haynes & Rees, 2006).

Review of the Literature

Real-Time Assessment

The measurement of hemodynamic responses to evaluate student understanding was the method of real-time assessment used within this study. The importance of real-time assessment in education can be seen through the positive responses and resulting grades of students after engaging in interactive practices in the classroom. One piece of technology with extensive applications is the clicker which is often used in primary, secondary, and higher education classrooms. A clicker is a form of technology that allows instructors to provide live questions to their students and view the responses from the entire class. The use of this divide increases the modes available to assess students in real-time. Studies have shown that students supply positive feedback after using devices such as clickers that help increase their focus and engagement during a lecture (Ruisoto & Juanes, 2019). The effect of a “student-centered” model using various technologies also resulted in increased understanding and increased overall performance from the students in the course (McCuddy et al., 2008). Subjective assessments, or those that measure a student’s understanding of a topic after the lesson is finished, are less beneficial for addressing a student's immediate misconceptions (Trumbull and Lash, 2013). Often assessments take place at least a few days after the corresponding lesson, which argues that a teacher is unable to assess a student's understanding until that point. Real-time assessment is useful in the classroom due to its ability to allow teachers to address any misconceptions as soon as students experience them. This process also prevents students from studying material that they do not have a firm grasp on and allows teachers to reiterate confusing topics as students learn them if needed (Renawi et al., 2021).

DNA Replication

This study used a video on DNA replication to assess student understanding in real-time. The replication of deoxyribonucleic acid (DNA) is the foundation for cell replication in the human body. For a cell to divide and produce a new cell, the DNA of the genome must be replicated so there is a full set of chromosomes in each cell (Brody, 2014). Therefore, DNA replication is an imperative part of life and has the potential to produce life-threatening mutations such as genetic diseases and cancer (Mertz et al., 2017). This process is critical for students to learn as it is the ground on which much biological research is built. Since DNA replication is a fundamental aspect of biological sciences and forms the building blocks for further processes, students must have a strong grasp of the factors that contribute to it. However, genetics and DNA replication can be difficult for students to understand since the concepts occur on an intracellular level that cannot be easily seen (Fossey & Hancock, 2005).

Purpose of Chemistry

Similar to DNA replication, chemistry is a vital aspect of science to grasp because it exists in every part of life on an atomic level. This branch of science is essential to understanding biological processes including digestion, water purification, and energy. Since chemistry is taught on an abstract and minuscule scale, students may struggle to grasp the overarching ideas of the discipline (*Why study chemistry,* 2022). The teacher must be able to transform the materials into a concrete lesson that the students can understand (Taber, 2020). One obstacle to this style of teaching is the lack of ability to gauge how well students are grasping the material as they are learning it.

Machine Learning

Machine learning algorithms were used to derive the predictors for student answers on the science content test. Random forest confusion matrices and penalized logistic regression techniques are two commonly used algorithms in machine learning applications (Bari Antor, 2021). These algorithms make it possible to decipher the structure of current data, construct prediction rules for fresh observations, and create real-time adaptive content. A random forest algorithm is used in high-dimensional neuroimaging to place clinical labels on neuroimaging datasets and feature reduction (Dimitriadis et al., 2018). Clinical labels identify the various samples in categories, while feature reduction reduces the number of samples without sacrificing vital data. Penalized logistic regression techniques, where the dataset is shrunk to limit the number of variables, are commonly used for identifying the possibility of a piece of data fitting into a group using mathematical linear combinations of variables (Eilers et al., 2001). For nonlinear applications, the least-squares--or Levenberg-Marquardt--algorithm is useful for fitting a data set to a mathematical model as well. The Levenberg–Marquardt algorithm varies parameter updates adaptively and divides model node updating between a Gauss-Newton and a gradient descent procedure (Gavin 2019). Using various algorithms, predictors can be generated by inputting the acquired data and training the algorithm to detect patterns in the data such as specific hemodynamic responses associated with understanding content. To ensure accurate predictors are derived after training, K-fold cross-validation can be used to assess the efficacy of the results. In K-fold cross-validation, the data is randomly divided into similar-sized data sets where one section represents the test section and the other represents the training section. This procedure is done multiple times so that each data point gets to be included in both sets at least once to increase the validity of the technique (Jung & Hu, 2015).

Hemodynamic Response

The real-time responses analyzed in this study were the hemodynamic responses relating to the amount of oxygenated hemoglobin in the brain. Blood cells do not possess internal reserves of glucose and oxygen, which are vital ingredients needed to conduct cognitive processes. Similar to a person breathing heavier as they exercise because the body needs more oxygen to fuel the muscles, the brain requires more oxygen during cognitive tasks. As the brain undergoes cognitive processing and engages in more mentally strenuous tasks, more oxygen and glucose are needed to support functionality. This causes glucose and oxygen to be sent to active neurons faster than they would be sent to an inactive neuron. An inactive area of the brain will, therefore, have relatively less oxygen and glucose in the bloodstream. The ratio of oxygenated blood, or oxyhemoglobin, and deoxyhemoglobin can be measured to gain an understanding of what areas of the brain are active during different stimuli (Harris, 2021). These ratios can be interpreted into functional states that are used to make predictions about student understanding in this study.

Functional Near-Infrared Spectroscopy (fNIRS)

Functional magnetic resonance imaging (fMRI) and functional near-infrared spectroscopy (fNIRS) are two techniques used to measure how changes in blood flow relate to brain activity that occurs through hemodynamic response. An fMRI may be used for various applications such as measuring the effects of a stroke, detecting abnormalities in the brain, or developing a treatment plan. An fMRI works by using electromagnetic waves to align hydrogen atoms in the blood which in turn emit a certain amount of energy based on where they are in the brain. This process creates a map that a researcher can use to see brain patterns. During the fMRI procedure, the patient will perform various tasks that cause an increase in hemodynamic response and brain activity that can be measured in the resulting scan (*Magnetic resonance,* 2021). This technique has been widely used in research due to its wide availability and noninvasive methods. However, fMRI systems are limited in their spatial recognition functions and are not portable due to the size of the scanner that is needed. Due to this aspect of the tool, fMRI is unsuited for mobile educational and student-centered research. A technique that is growing in popularity is functional near-infrared spectroscopy (fNIRS). fNIRS works similarly to fMRI by measuring the hemodynamic response in the brain but this technique uses the relative optical absorptivity of oxyhemoglobin and deoxyhemoglobin rather than energy differences through the exposure to electromagnetic waves (Chen et al., 2020). However, fNIRS can only analyze sub-cortical tissue up to a depth of ten millimeters, while fMRI can see the entirety of the sub-cortical tissue. Although this measure(fNIRS) has a limited depth of recording compared to the fMRI, it is highly portable and has a higher tolerance for motion than its counterpart. Due to this, fNIRS have been previously used in studies that require a less constrained scanner environment (Scarapicchia et al., 2017). The capability of using fNIRS instead of fMRI or another technique is that it allows research to be done in a setting like a classroom. Students can also engage in mobile activities as they would in a classroom to increase the validity of the results. This supplies the opportunity for a real-time display of students' reactions to their environmental stimuli such as an educational video and to the lessons they are being taught. The portability, low cost, and wearability of the fNIRS equipment are all benefits related to using this technique. In earlier studies related to student engagement and understanding, the fNIRS has been shown to supply data that distinguishes between high and low levels of engagement as well as the relative difficulty of the task at hand. In a study conducted by Oku and Sato in 2021, a significant correlation between the fNIRS data and their hypotheses was found. Their study was reviewing the connection between hemodynamic response and student engagement and found grounds for a causal relationship. Their research will be expanded to include the predictability of student performance and understanding as well in this study (Oku & Sato, 2021). A future implementation for this research is the use of newly developed fNIRS wearable headsets. This tool would allow the researcher to measure students' hemodynamic responses directly in a classroom setting rather than just a simulation of that environment (Piper et al., 2014).

Method

Hemodynamics data was measured using fNIRS from the prefrontal cortex. The data was used to predict student answers on a subsequent science test. A machine learning algorithm was implemented by applying it to the hemodynamics data and question responses to develop predictors. The authors used random forest and Levenberg-Marquardt algorithms in an Artificial Neural Network (ANN) to evaluate the structures of the acquired data and formulate the prediction rules.

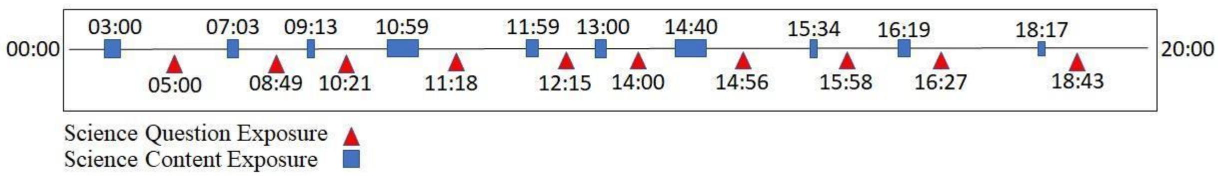
Participants

The study was made up of forty total participants (n=40) that were recruited from a charter school in the northeastern United States. The participants consisted of 21 females and 19 males, and all 40 participants were from a seventh-grade classroom at an urban pre-K through eighth-grade charter school. The participants had a mean age of 11.8 with a standard deviation of 0.6. The students were unfamiliar with the content on DNA replication before the online instruction. The students were also within the range of their grade level by one-half plus or minus for mathematics and reading subject areas. Interviews and a review of the history suggested in the *Compendium of Neuropsychological Tests* were used to conduct additional screenings (Strauss et al., 2006). No participants were eliminated due to the student screening nor during the events of the study.

Movie Condition

A twenty-minute video explaining the process of DNA replication was recorded to present to the participants. A female educator on a computer screen presented an overview of the DNA replication process. Pictorial representations were shown in addition to vocal instructions. Three times in the film, colored graphical representations were flashed on the screen in place of the instructor, who was merely supplying a voice-over. There were no animations in the video that was used. Students were asked questions about the topic delivered at certain points throughout the film (Cronbach's Alpha =.87). The students did not receive any feedback from the video on whether their responses were right. Figure 1 depicts the time during the content at which an associated question was asked.

Figure 1. Graphical representation of content and questions over 20 minutes



Null Condition

A baseline condition was taken before and after the movie presentation during data collection. Participants were encouraged to sit quietly with their eyes closed and relax during the baseline data collection. During this time, hemodynamic readings were taken in the prefrontal cortex to assess rest conditions, which was a vital step in establishing initial reference points for later hemodynamic responses. This activity is frequent in electroencephalography protocols and has been adapted for fNIRS studies (Vernon, Peryer, Louch, & Shaw, 2014). The researchers were able to confirm that the responses detected inside the stimulus were true reactions from the stimulus and not an artifact of the fNIRS by using the second baseline (Kratochwill et al., 2013).

Statistical Analysis

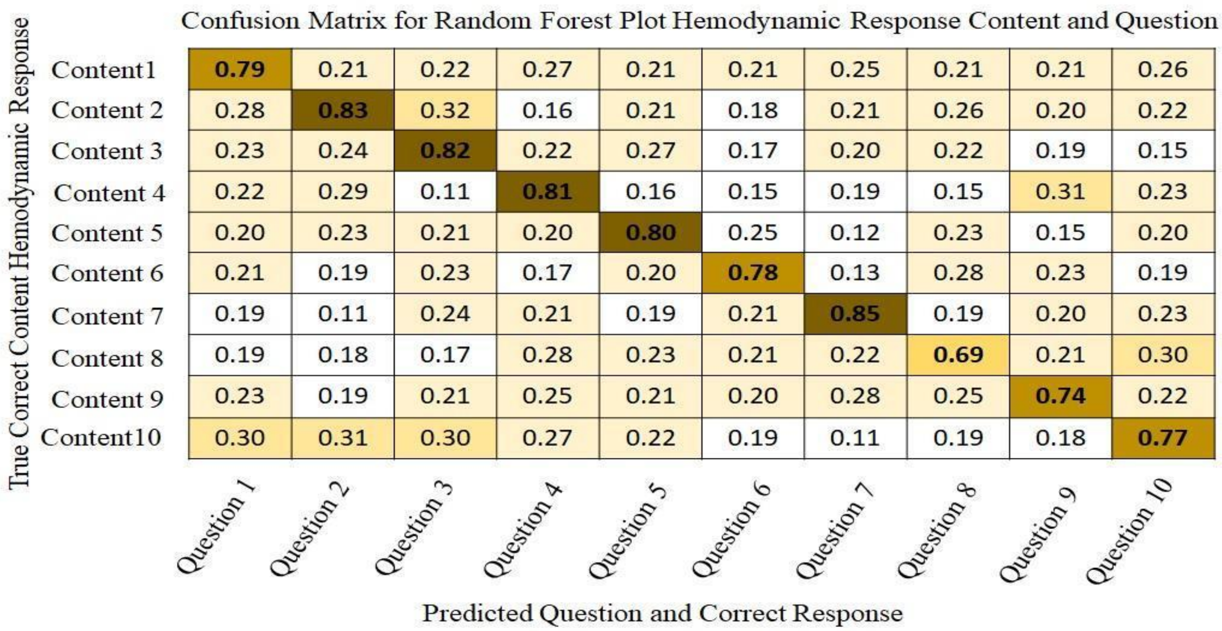
The standardized hemoglobin absorption ratios (*i*=864,000) between oxygenated and deoxygenated hemoglobin were subjected to statistical analysis. The differences between the O_2Hbi and -Hbi standardized hemodynamic responses were statistically assessed using a repeated measures Analysis of Variance (rANOVA) and planned post-hoc comparisons by condition using SPSS version 27. The subjects serve as their own controls in rANOVA, making it especially suitable for investigating A-B-A within-subject designs in studies like this one. The researchers can also use this approach to identify optodes of interest, which are those that show statistically significant variations from baseline. The optodes that differed from the baseline created patterns that were fed into the machine learning system and classified for response prediction on the content test.

The goal of this analysis is to assess the ANN node weightings and find the model with the best accuracy and generalizability. If they forecast test data with the least error, the highest accuracy and generalized fit are given. Multiple measures in the form of Conforming Capability and Generalizing Capability (Conforming = MMSEtr + MMSEtst and Generalizability = MMSEtst MMSEtr) were used to find the architecture that demonstrates the best data and conceptual fit. The model's stability was investigated using SDMSE and variants of the model. These variations take the shape of competing model designs that all meet the same conceptual requirements. The models with the largest standard deviations are eliminated from the analysis. The most generalized architecture is kept after the model with the biggest standard deviation is removed. A model showing the closest training and testing data as seen through the comparison of means is the most generic design. ANNs have been used extensively to stimulate cognitive function and identify learning outcomes, as found by Al-Nafijan et al. (2014) and Xiao et al. (2019). An error backpropagation model and the Levenberg–Marquardt algorithm are employed in this cognitive prediction model. Backpropagation happens when the chain rule for partial derivatives is applied repeatedly. The output nodes are regarded as the probability of correct and incorrect responses concerning the video's query. The ANN that was employed in this investigation was created with R 3.3.1 and the ANN package. A cross-validation strategy is used in the suggested model validation method. The authors partition data into numerous identically sized and randomly selected slices (n1 = 288,000, n2 = 288,000, and n3 = 288,000) using K-fold validation. Both training and validation are done with the segments or slices. In sectors like engineering, this method has proven to be effective (Wijayasekara, Manic, Sabharwall, and Utgikar, 2011).

Results

The findings imply that student hemodynamic responses during topic presentations are indicative of student achievement on the content test. While the students were watching the movie, the classification results of hemodynamic patterns predicted correct and incorrect replies during the content test. A confusion matrix is used to show how these consequences are visualized. To calculate the chance of successfully answering the relevant questions, the confusion matrix utilizes two algorithms in the predicted classes of hemodynamic response during content. The projected model's examples are represented in each row, while the true results of the students' performance are represented in the column. The training and test models had an excellent fit when it came to detecting correct and incorrect answers. For example, hemodynamic response patterns while students viewed Content 1 showed success in predicting the answer for Question 1 78% of the time. Overall, the ANNs model predictions showed a predictive success rate for the content questions between 69% and 85%, as shown by the deep brown blocks. The resulting confusion matrix between the shown content and the content question is depicted in Figure 2.

Figure 2. Confusion matrix illustrating the relationship between content and questions.



Discussion and Implications

Assessment is an ongoing cycle that helps increase student understanding and the improvement of teaching strategies (*The Assessment Process* 2022). “Student-centered” models that promote the use of real-time assessments are proven to increase student engagement and understanding in a virtual and in-person classroom setting (McCuddy et al., 2008). Real-time assessment is also useful for abstract units such as DNA replication, which students have a naturally harder time grasping. By employing real-time assessment, teachers can address misconceptions at once without students internalizing topics they do not fully understand (Renawi et al., 2021). Recent studies, such as Wachtler et al. (2018), have also found that video lectures with quizzes can be used to boost knowledge, heighten engagement, and increase attention. Online adaptive lectures supply the opportunity for real-time assessment and immediate evaluation of teaching techniques so that adjustments can be made to the lesson to supply an understandable experience for students during the class. Although student performance can be measured through the results of in-class exams, the mapping of brain states during the activity has the possibility of supplying immediate feedback that can increase student understanding while the lesson is taking place. Typically, cognitive neuroscience experiments investigate psychological processes by manipulating one of its components' behaviors. This paradigm, on the other hand, is ineffective when trying to generalize the characteristics of new scenarios based on detailed descriptions of the behavior (Varoquaux and Poldrack, 2018). This study can be proposed to assess levels of student understanding during virtual lessons, allowing for the evaluation of whether students can assimilate content using fNIRS signals, with error rates in the models below 30%. The results of this study imply a way for teachers to remove how their students are understanding the material as they teach it. In a classroom where most students are predicted to pass future assessments based on their hemodynamic responses, the teacher could continue with their current strategies and address other students' problems separately. Alternatively, the teacher would also know when none of her students are understanding the material and the teacher could adjust their strategies accordingly. The result is a method for teachers to limit the amount of time needed to plan assessments to supply more time to tailor the lesson to the students' needs. Despite the positive outcomes, the study has some limitations. For example, the model considers the fNIRS signal from a single video lesson. More research is also needed to learn more about the students' behavior and performance during the task. Short-distance detectors were not used in this study but may have a significant impact on future studies in a traditional classroom setting due to the ability to limit environmental feedback (Tachtsidis and Scholkmann, 2016). The algorithms and techniques used in this study could allow for adaptive online learning that can focus on addressing shortcomings in teaching techniques and content when considering students' needs. Although the technology is not currently adapted for the traditional classroom setting, adaptable programs are expected to become prevalent in future classrooms.

References

Alzen, J. L., Langdon, L. S., & Otero, V. K. (2018). A logistic regression investigation of the relationship between the learning assistant model and failure rates in introductory stem courses. *International Journal of STEM Education*, *5*(1). https://doi.org/10.1186/s40594-018-0152-1.

Bari Antor, M., Jamil, A., Mamtaz, M., Monirujjaman Khan, M., Aljahdali, S., Kaur, M., Singh, P., & Masud, M. (2021). A Comparative Analysis of Machine Learning Algorithms to Predict Alzheimer's Disease. *Journal of healthcare engineering*, *2021*, 9917919. https://doi.org/10.1155/2021/9917919

Board of Governors, Missouri State University. (2019). *The Assessment Process*. Assessment of Student Learning - Missouri State. https://www.missouristate.edu/assessment/the-assessment-process.htm.

Brody, L. C. (2014). *DNA replication*. Genome.gov., https://www.genome.gov/genetics-glossary/DNA-Replication.

Chen, W.-L., Wagner, J., Heugel, N., Sugar, J., Lee, Y.-W., Conant, L., Malloy, M., Heffernan, J., Quirk, B., Zinos, A., Beardsley, S. A., Prost, R., & Whelan, H. T. (2020). Functional near-infrared spectroscopy and its clinical application in the field of neuroscience: Advances and future directions. *Frontiers in Neuroscience*, *14*. https://doi.org/10.3389/fnins.2020.00724.

Dimitriadis, S. I., Liparas, D., & ADNI. (2018). How random is the Random Forest? Random Forest algorithm on the service of structural imaging biomarkers for Alzheimer’s disease:

From Alzheimer's disease neuroimaging initiative (ADNI) database. *Neural Regeneration Research*, *13*(6), 962. https://doi.org/10.4103/1673-5374.233433.

Dyrberg, N. R., & Holmegaard, H. T. (2018). Motivational patterns in STEM education: A self-determination perspective on first year courses. *Research in Science & Technological Education*, *37*(1), 90–109. https://doi.org/10.1080/02635143.2017.1421529.

Eilers, P. H., Boer, J. M., van Ommen, G.-J., & van Houwelingen, H. C. (2001). Classification of Microarray Data with Penalized Logistic Regression. *SPIE Proceedings*. https://doi.org/10.1117/12.427987.

Ekwueme, C. O., Ekon, E. E., & Ezenwa-Nebife, D. C. (2015). The impact of hands-on-approach on student academic performance in basic science and mathematics. *Higher Education Studies*, *5*(6), 47. https://doi.org/10.5539/hes.v5n6p47.

Fossey, A., & Hancock, C. (2005). DNA replication and transcription: An innovative teaching strategy. *Biochemistry and Molecular Biology Education*, *33*(6), 411–415. https://doi.org/10.1002/bmb.2005.49403306411.

Gavin, H. P. (2019). The Levenberg-Marquardt algorithm for nonlinear least squares curve-fitting problems. *Department of Civil and Environmental Engineering, Duke University*, 1-19.

Gonzalez, H. B., & Kuenzi, J. J. (2012). *Science, Technology, engineering, and Mathematics (Stem) education: A Primer*. Congressional Research Service, Library of Congress.

Harris, K. A. (2021). Blood-oxygen-level-dependent (BOLD) signal. *Encyclopedia of Autism Spectrum Disorders*, 717–718. https://doi.org/10.1007/978-3-319-91280-6_550.

Haynes, J.-D. (2011). Brain reading: Decoding mental states from brain activity in humans. *Oxford Handbooks Online*. https://doi.org/10.1093/oxfordhb/9780199570706.013.0013.

How, M.-L., & Hung, W. L. (2019). Educing AI-thinking in science, Technology, Engineering, arts, and Mathematics (STEAM) education. *Education Sciences*, *9*(3), 184. https://doi.org/10.3390/educsci9030184.

Jung, Y., & Hu, J. (2015). A k-fold averaging cross-validation procedure. *Journal of Nonparametric Statistics*, *27*(2), 167–179. https://doi.org/10.1080/10485252.2015.1010532.

Kratochwill, T. R., Hitchcock, J. H., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2013). Single-case intervention research design standards. *Remedial and Special Education*, *34*(1), 26-38.

McCuddy, M. K., Pinar, M., & Gingerich, E. (2008). Using student feedback in designing student-focused curricula. *International Journal of Educational Management*, *22*(7), 611–637. https://doi.org/10.1108/09513540810908548.

Mertz, T., Harcy, V., & Roberts, S. (2017). Risks at the DNA replication fork: Effects upon carcinogenesis and tumor heterogeneity. *Genes*, *8*(1), 46. https://doi.org/10.3390/genes8010046.

Oku, A. Y., & Sato, J. R. (2021). Predicting student performance using machine learning in fNIRS Data. *Frontiers in Human Neuroscience*, *15*. https://doi.org/10.3389/fnhum.2021.622224.

Piper, S. K., Krueger, A., Koch, S. P., Mehnert, J., Habermehl, C., Steinbrink, J., Obrig, H., & Schmitz, C. H. (2014). A wearable multi-channel fNIRS system for brain imaging in freely moving subjects. *NeuroImage*, *85*, 64–71. https://doi.org/10.1016/j.neuroimage.2013.06.062.

Radiological Society of North America, Inc. (2021, July 30). *Magnetic resonance, functional (fmri) - brain*. Radiologyinfo. https://www.radiologyinfo.org/en/info/fmribrain.

Rathore, S., Habes, M., Iftikhar, M. A., Shacklett, A., & Davatzikos, C. (2017). A review on neuroimaging-based classification studies and associated feature extraction methods for Alzheimer’s disease and its prodromal stages. *NeuroImage*, *155*, 530–548. https://doi.org/10.1016/j.neuroimage.2017.03.057.

Renawi, A., Alnajjar, F., Parambil, M., Trabelsi, Z., Gochoo, M., Khalid, S., & Mubin, O. (2021). A simplified real-time camera-based attention assessment system for classrooms: Pilot study. *Education and Information Technologies*. https://doi.org/10.1007/s10639-021-10808-5.

Ruisoto, P., & Juanes, J. A. (2019). Fostering student’s engagement and active learning in neuroscience education. *Journal of Medical Systems*, *43*(3). https://doi.org/10.1007/s10916-019-1192-x.

Scarapicchia, V., Brown, C., Mayo, C., & Gawryluk, J. R. (2017). Functional magnetic resonance imaging and functional near-infrared spectroscopy: Insights from combined recording studies. *Frontiers in Human Neuroscience*, *11*. https://doi.org/10.3389/fnhum.2017.00419.

Sikka, I., Chawla, C., Seth, S., Alghadir, A. H., & Khan, M. (2020). Effects of deep cervical flexor training on forward head posture, neck pain, and functional status in adolescents using computer regularly. *BioMed Research International*, *2020*, 1–7. https://doi.org/10.1155/2020/8327565.

Strauss, E., Sherman, E. M., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary*. American Chemical Society.

Taber, K. (2020). *Foundations for Teaching Chemistry: Chemical Knowledge for teaching*. Routledge.

Tachtsidis, I., & Scholkmann, F. (2016). False positives and false negatives in functional near-infrared spectroscopy: issues, challenges, and the way forward. *Neurophotonics*, *3*(3), 031405.

Trumbull, E. & Lash, A.A. (2013). Understanding Formative Assessment Insights from Learning Theory and Measurement Theory.

University of Wisconsin. (2022). *Why study chemistry*. University of Wisconsin. https://www.uwlax.edu/chemistry-and-biochemistry/student-resources/why-study-chemistry/.

Varoquaux, G., Schwartz, Y., Poldrack, R. A., Gauthier, B., Bzdok, D., Poline, J. B., & Thirion, B. (2018). Atlases of cognition with large-scale human brain mapping. *PLoS computational biology*, *14*(11), e1006565.

Vernon, D., Peryer, G., Louch, J., & Shaw, M. (2014). Tracking EEG changes in response to alpha and beta binaural beats. *International Journal of Psychophysiology*, *93*(1), 134-139.

Wachtler, J., Scherz, M., & Ebner, M. (2018, June). Increasing learning efficiency and quality of students' homework by attendance monitoring and polls at interactive learning videos. In *EdMedia+ Innovate Learning* (pp. 1357-1367). Association for the Advancement of Computing in Education (AACE).

Wijayasekara, D., Manic, M., Sabharwall, P., & Utgikar, V. (2011). Optimal artificial neural network architecture selection for performance prediction of compact heat exchanger with the EBaLM-OTR technique. *Nuclear Engineering and Design*, *241*(7), 2549-2557.