**Predicting student's performance in a Serious Game with Machine Learning using Convolution Neural Network**

Dongwook An

The University of Texas at Austin

**Introduction**

**Predictive Learning Analytics**

The recent availability of large volumes of log data generated from Learning Management Systems (LMS) or Serious Games (SG) has led to extensive research on Learning Analytics (LA). Learning analytics referes to “the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs” (Long & Siemens, 2011, p. 34). Specifically, the development of data mining as well as machine learning techniques has enabled researchers to develop more sophisticated model to predict student performance in LMS or SG. Predictive Learning Analytics (PLA) refers to the analysis and modeling of log data to accurately predict future outcomes, for example, whether a specific student will drop out or not (ECARANALYTICS, 2015). Many researchers believe that predictive modeling can help improve students’ performance (Siemens, 2012)

PLA is not a new concept in education, and many researchers have already been developing predictive models and applying them to solve their own problems. San Pedro, Baker, and Heffernan (2017) created a model to predict engagement level using user logs of a math tutoring system with machine learning as well as data mining techniques, while they were trying to investigate a relationship between engagement in the tutoring system and college enrollment. Mills et al. (2014) collected reading time information that how long readers spent on each page in a tutoring system and built a predictor of behavioral disengagement. Muratov, Lewis, Fourches, Tropsha, & Cox (2017) developed predictive models for student admissions using Random Forest machine learning technique showing that their model outperformed over some linear regression models for resident candidates while non-resident model were very similar. Seidel and Kutieleh (2017) built a decision tree model to predict more than 3500 first year students’ risk of attrition in a university using five years of students’ behavior data in their Learning Management System (LMS) as well as students’ demographic data. They also implemented intervention strategy for students with high predicted attrition probabilities before peak attrition time point, which was one of the main goals of PLA, to improve students’ retention and performance.

**Learning Analytics in Serious Games**

A variety of LA techniques has been applied to Serious Games Analytics such as cluster analysis, Decision Tree, Naive Bayes, Support Vector Machine, Logistic Regression, and Sequential Pattern Mining. Kerr (2015) identified students’ learning strategy through cluster analysis and then modified the educational game not only to reduce players’ irrelevant behaviors but also to improve positive perceptions of the game compared with the original setting. Paquette et al. (2016) built a predictive model of affect detection with game log data combined with ground-truth affect data via field observation in a SG. Kang, Liu, & Qu (2017) did Sequential Pattern Mining to explore students’ sequential behavior patterns in a SG and discovered that the patterns are various on different performance groups as well as different stages. Hicks et al. (2016) used Survival Analysis technique and Interaction Network Analysis for analyzing drop-out rates in a SG to find out potential areas of improvement. These efforts have been facilitating understanding students' engagement, behavioral patterns, learning process and performance levels in SG environments.

**Prediction problem in Complex Serious Games**

A serious game can be complicated enough to have diverse paths when solving a problem in it. This can make it difficult to build predictive models even though researchers make use of several machine learning techniques. One reason of it is that they could fail to construct a labeled dataset to make a predictive model or they could fail to predict performance significantly due to the game’s higher level of complexity although they could build dataset.

Numerous prior research in PLA was able to have labeled data with predictors in the form of multiple row to solve simple prediction problem in its context. For example, Huang and Fang (2013) used cumulative GPA, grades in four pre-requisite courses, and three midterm exam grades as features (predictors). Its target label (predicted value) was students’ performance at the end of the semester. And each row of the dataset represents each student’s information, as is usual in a traditional statistical dataset such as linear regression. For the dataset, many researchers have utilized frequency or sum of time log data for predictors as well. Mills et al. (2014) collected a dataset with students’ the location of quit (page number) as well as reading time information that imply how long readers spent on each page in a tutoring system. Each row of the dataset was comprised of each reader’s predictors with a target label: quitted or not quitted.

But in a complex research context or a sophisticated SG environment, dataset should have multiple rows per a user or multiple target labels per a user since the log data of a user should not be aggregated in a row. In that case, researchers can’t use popular machine learning techniques such as Decision Tree, Logistic regression, Neural Network, NaiveBayes, and Support Vector Machine since those methods need a dataset that should have each person each row as well as single target label (e.g. quitted or not quitted). For example, Min et al. (2016) investigated multiclass classification (multiple target label) problem in a serious game using long short-term memory networks (LSTMs) which is a variant of Recurrent Neural Network (RNN), a popular deep learning method for sequence observation. They found that LSTM models yielded the best performing model in classifying player goal with game logs of players.

In this paper, researchers explained a research context of multimodal (multi-label) learning analytics with multimodal data that need to have multiple row per a user as well as multiple target label in the dataset for PLA in a SG environment, and presented a predictive model with Machine Learning technique using Convolutional Neural Network. The following research questions guided this research:

1. How predictive model can be built to predict performance of users whose trajectories are vary in a complex serious game?

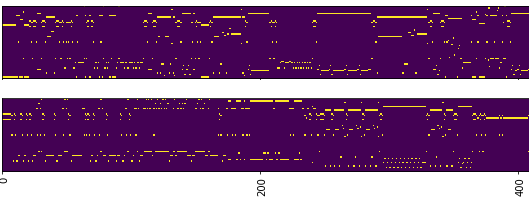
And this, in turn, leads to sub-questions such as:

1. What can be the predictors for the predictive model in a complex serious game to explain students’ performance?

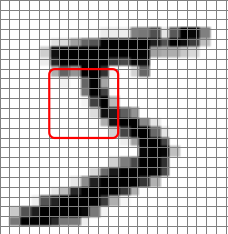
2. In case target labels are multiple, what kind of modeling technique can be used to build predictive model?

**Convolutional neural networks (CNN)**

The Convolutional Neural Network (CNN) is a class of deep learning, which was designed to process data that come in the form of multiple arrays, and has been widely used for challenging many machine-learning tasks such as text classification, image classification (Lecun, Bengio, & Hinton, 2015). The CNN captures spatial topology much better than standard neural networks since it detects distinctive features wherever they appear on the input (LeCun et al., 1998).

SG log data has spatial-temporal nature (Loh & Sheng, 2015). Figure 1 shows the spatial-temporal nature of multimodal log data in a SG environment. In the graph, two game players were randomly selected and visualized according to time (x-axis) and distinct action in different game scenes (y-axis). The action and navigation pattern of two users were quite different, but researchers in this research hypothesized that similar patterns between players could emerge at a different time period since there were frequent sequential patterns across players’ problem-solving stages in a SG environment, according to Kang, Liu, & Qu (2017)’s research. The CNN has been applied to analyze image classification such as hand-written digit image classification in Figure 2, since image pixels have a strong 2-dimension local structure in which the nearby pixels are highly correlated each other spatially or temporally (LeCun et al., 1998).

*Figure 1.* Spatial-temporal nature of interaction log data

  
*Figure 2.* CNN Image classification [Hand-written digital image] (2014, August 14). Retrieved July 13, 2018, from http://deepdish.io/public/images/PCANet\_mnist5\_patch.png

The visualized log graphs in Figure 1 also showed high correlation between nearby pixels in each graph since there were sequential action patterns in it for players to solve a problem in the SG environment. In this research, researchers worked directly with raw pixel image data like Figure 1, and the use of CNN allowed to capture the spatial-temporal patterns of log data to eventually build a predictive model.

**Method**

**Research Context**

This study focused on problem-solving activities of players in a 3D immersive SG environment called *Alien Rescue* (http://alienrescue.edb.utexas.edu), which is designed for

middle school science learning. *Alien Rescue* incorporated problem-based learning pedagogy, so that students could use various approaches as well as hypothesis testing to solve problems in it.

In *Alien Rescue*, students are asked to find new habitats for six Aliens in our Solar system who were displaced from a distant galaxy. Students navigate in a spaceship, a 3d game environment, and gather information of planets such as atmosphere, gravity, water and so forth. Then they need to compare what an alien species needs with what a planet’s data for becoming a new habitat of six aliens. Finally, students select a planet for each alien, and submit it in the game. Answers are multiple, for example, for the first alien, its new habitat can be Callisto or Charon, and for the second alien it’s Deimos, Ganymede or Callisto.

**Participants**

Participants were 535 six graders from middle schools in the Southwestern area of the United States. Initially there were 1,855 students but 1320 students were removed from dataset since they didn’t submit any solution for aliens’ new habitat. Those schools used *Alien Rescue* as part of their science curriculum in 2017, and the period of use varied from one week to over one month given their school schedule or teacher’s own preference.

**Data Sources & Preprocessing**

Gameplay data contains each player’s action type, action item, screen name on which the action is taking, and timestamp for each action (see Table 1). For example, if a student click a solar panel in Probe Design Tool screen to design a probe, then log data was {"action" : "Select: probeSource", "note" : "Solar Panels", "timestamp" : "Friday, March 10, 2017 3:06 PM", "tool" : "Probe Design", "userId" : "1QC1LbaCp3a0yS4QSpkJ7SLxQC82"} as JSON file format. The total log count was 1,517,196 from 1,855 users, but only 365,583 count log data was used for the predictive model since only 535 students submitted at least one solution. Some students didn’t submit an answer for some reason, such as absence or school schedule, individual computer malfunction or teacher’s preference using Google doc when students submit answers instead of putting the answers in the game.

Target labels of dataset was students’ submitted solution for six aliens’ new habitat. For example, a students submitted six planets as answers for each aliens and got full score if the criteria met, or a student submitted three planets for three aliens and would get one score if only one criteria met for the answers. Since answer planets are different for each aliens and students needed to send probes to different planets as hypothesis testing to get information that fits with the aliens, the labels were coded separately for each alien. For example, if a student submitted correct answer only for the first alien, then labels became 1 by 6 array, (1, 0, 0, 0, 0, 0), and if a student submitted correct answer for the first alien and the third alien, then then labels were (1, 0, 1, 0, 0, 0) for the student (see Table 2). Finally, label dataset for the predictive model was built as a form of 3-dimension array (535, 1, 6) for the 535 students.

Table 1

*An example of gameplay data*

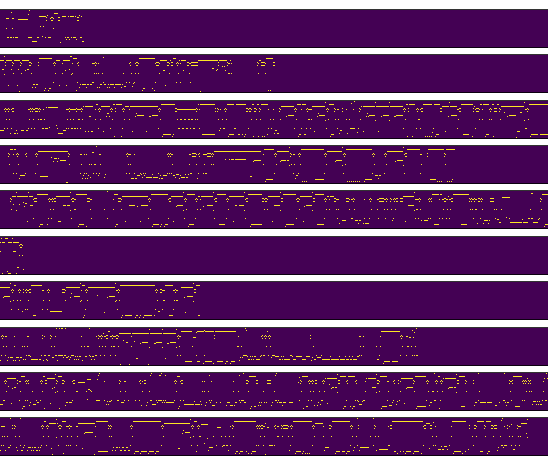
|  |  |
| --- | --- |
| Log | Description |
| "-Keu1n02ZvJ2MMeDLLga" : {  "dataLog" : {  "action" : "Select: probeSource",  "note" : "Solar Panels",  "timestamp" : "Friday, March 10, 2017 3:06 PM",  "tool" : "Probe Design",  "userId" : "1QC1LbaCp3a0yS4QSpkJ7SLxQC82"  }  } | Log id  Action type  Action item  Timestamp  Action screen  User id |

Table 2

*An example of label of dataset for each student’s submitted answer (Multi label classification)*

|  |  |  |
| --- | --- | --- |
| User id | Label (Target variable) | Total Score |
| 1QC1LbaCp3a0yS4QSpkJ7SLxQC82 | (1, 0, 0, 0, 0, 0) | 1 |
| oOjXUOiIAUPMHaWd7nVdVjWeX3t2 | (1, 1, 0, 0, 0, 0) | 2 |
| tlvqCSOAEuax3CFIXaTSyLFv8Tq2 | (0, 1, 0, 0, 1, 0) | 2 |
| mEKCOucQO0OXIEHPSc54KqcKZX33 | (0, 0, 0, 0, 0, 1) | 1 |
| OO0q44lAVhgMlwgvWqoHeDwf9Sq2 | (1, 0, 1, 1, 1, 1) | 5 |
| oOjXUOiIAUPMHaWd7nVdVjWeX3t2 | (1, 1, 1, 1, 1, 1) | 6 |

\* Total score was not used as features or labels

Predictor dataset was constructed as a form of 3-dimenstion array (535, 1180, 56). 2-D array data (1180, 56) was made for each of the 535 students from log data. First, Action field with tool field in Table 1 were dummy-coded and used as features. For example, the name of 56 dummy-coded features were: “action\_Click Back Button”, “action\_Click Destination Planet”, “action\_Click Forward Button”, “Menu,action\_Click Nav Menu”, “action\_Creat Note in: PLANET”, “action\_Lauch Probe”, “tool\_Concepts DB”, “tool\_MissionDB”, and etc. Second, all actions log of a student constructed to 2-D array data (1180, 56). For example, the first row of 1180 x 56 dataset indicates the player’s first action on a game scene, and the second row shows second action on a game scene. As 535 students have their own 2-D array data, it aggregated into 3-D array data (535, 1180, 56). Figure 3 shows the visualization of 10 students’ 2-D array data. 

*Figure 3.* Example of 10 students’ 2D array log data

Then the dataset was shuffled, split into a training set of 373 users (80%) and a test set of 94 users (20%). The training dataset was used for training the CNN, while the test dataset was used to evaluate the trained network.

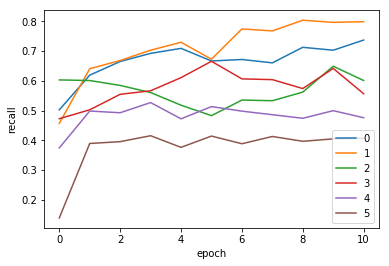
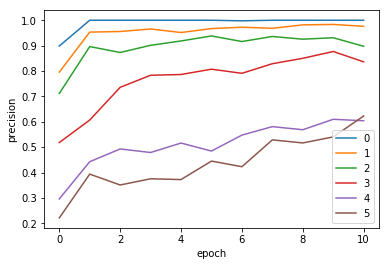
**Data Analysis**

The model is implemented and run using Keras (Chollet, 2017) with Tensorflow backend. The effectiveness of the model was evaluated by processing the test dataset. At the first time, the best test result of trained model reaches an average accurate rate of 9.235%, which means that the model can’t properly detect the student’s navigation patterns. Then researchers up-sampled the training dataset to handle imbalanced classes in the label dataset. For example, only 28 students out of 535 users had correct answer for the first alien, which means that the first column of 507 rows of the label dataset was 0. In that case, the model could predict the first label to all 0 then it could get 507/535 = 94.7% accuracy. So the researchers used up-sampling technique for every column in the label dataset to fix falsely predicting dominant class (zero) and to make balanced dataset.

**Results**

Researchers trained the network for 10 epochs. As it trains, the accuracy increased. After the first pass (epoch), it got 92.73% of training accuracy. After just 10 passes, it was achieving 99.95% multi-label classification accuracy on the training set. The best performance of .89 classification precision was achieved on the first label (the first alien). The worst performance of .59 classification precision was achieved on the fifth label (the fifth alien). Figure 4 shows the results of the CNN model precision and recall of each label. The true positive ratio, which also called precision, was less than 60% for the fifth and fourth aliens. The recall, which is defined as the number of true positives divided by the number of true positives plus the number of false negatives, was below 50% for the fifth and fourth aliens.

After developing the CNN model on the training data, researchers validated its performance on the holdout test data (20% of the whole dataset). The CNN model achieved 57.89% multi-label classification accuracy on the testing set, which is below the pure chance level of 67% prediction accuracy when predict it all zeros.



*Figure 4.* Model precision and recall of each label

**Findings and Discussion**

While the training test accuracy was achieved up to 99.95%, the test set accuracy was below pure chance accuracy. The precision and recall graph shows that the fourth and fifth label had least classification precision and recall of 50%. The second and the third label had less classification recall, which was below 60%, as shown in Figure 4.

The suggested model could not detect students’ navigation patterns properly. Labels mostly misclassified were label 5 and label 6 in the training dataset, which is likely caused by the different frequencies of in the original dataset. The label 1 and label 2 were reliably detected in the training dataset, but their accuracy were low in the test dataset, which is likely cause by the up-sampling on the training dataset. The model’s accuracy was below chance accuracy since it overfitted to the training dataset.

Although the CNN is the most powerful algorithm to capture patterns in many image classification problem, it failed to detect patterns in the game log. Future research applying another way of data pre-processing or another deep learning algorithm like Recurrent Neural Network (RNN) have possibility to attain meaningful results. But, this study demonstrated how deep learning algorithm can be incorporated into Learning Analytics in order to predict students’ performance in a Serious Game environment.

categories were mostly misclassifed as belonging to two former large categories, which is likely caused by the diﬀerent frequencies of coding categories in the original dataset

have achieved great success and applied to various fields such as image and speech recognition.

The main advantage of these methods lies in the strong capability of processing unstructured data, which also makes it possible to detect learning styles dynamically via the behavior sequence of learners.

By contrast to high cost of determining student learning style in face to face classroom, the suggested model can detect learning style dynamically in an acceptable cost by taking the advantage of LMSs and intelligent analysis method, makes that is possible make interventions by individual learning style.

It is found by further investigations on the logs that the change is usually caused by forum related behaviors, which implies that it is possible to enhance engagement of this kind of students by sending them messages about the latest discussions on course forums.

A simple pedagogical model was designed, in which proper actions would be taken while the event for learning style change was detected, as listed in in Table 5. The investigation was limited to providing only different course contents at this time, therefore content-related events “Verbal to Visual” and “Visual to Verbal” were ignored.

Table 5. Action list of learning style change event detected.

The users can significantly benefit from the individualized pedagogical model in accord with their learning styles, especially for the study of complex and cross-disciplinary subjects. Another valuable finding is the remarkable improvement on the interactions among the users in the forums, this could positively help users learn from each other rather than only from courses or instructors which is very important for sustainability education. Of course, the more general conclusion still depend on the further investigation with more rigorous control on the experiments.

After 50 or so passes, it capped out around 95.5% accuracy and additional training didn’t help, so I stopped it there.

The confusion matrix for the test data is shown in Table 7. We see that error rate for the Goal category is the lowest, followed by the

moderate error rate for the Observation category. In contrast, we see that the Other and Motive categories were mostly misclassifed as belonging to two former large categories, which is likely caused by the diﬀerent frequencies of coding categories in the original dataset (Table 3).

Finally, to examine the value of the SMOTE preprocessing, we examined the confusion matrix of the random forest model developed using the original training and test datasets. The optimal

mtry value was 500 by which the classifer obtained .73 (SD = .02) classifcation accuracy and Cohen’s κ of .48 (SD = .04). Further validation of the classifer performance on the holdout test data showed .74 classifcation accuracy (95% CI[.72, .77]) and Cohen’s κ of 0.50 which was slightly lower than the classifer performance obtained after the SMOTE pre-processing

The following

three fundamental problems have to be addressed for this design

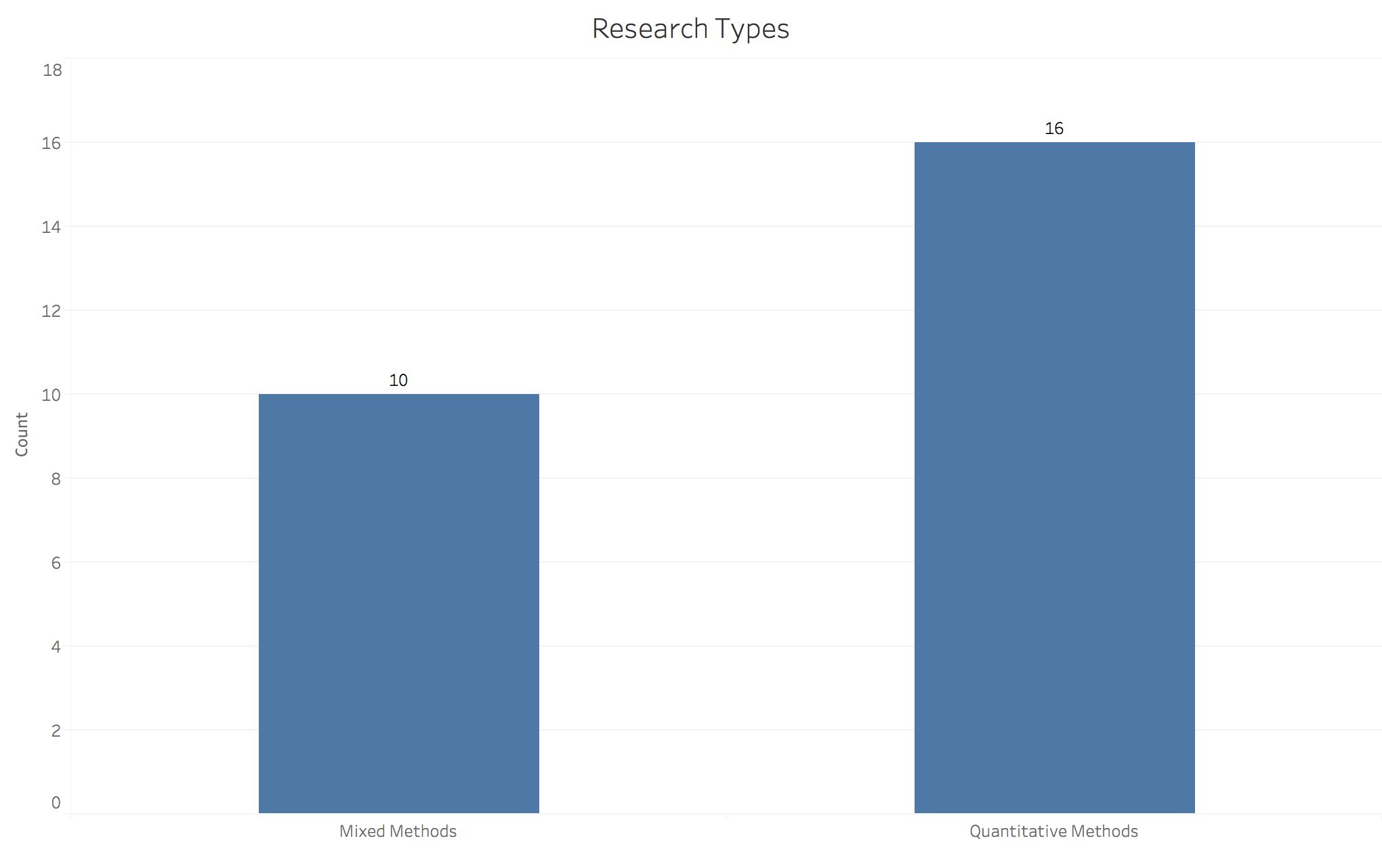
This study demonstrated how data mining can be incorporated into program evaluation in order to

generate in-depth information for decision making. In addition, it explored potential EDM applications at the K-

12 level that have already been broadly adopted in higher education institutions.

*Figure 1.* Frequency of technologies involved.

**Types of research designs.** In these 26 articles, the findings showed that quantitative research design was the most frequently used (*n* = 16, see Figure 2). It also appeared mixed methods research design was also getting popular for learning analytics research (*n* = 10). No qualitative research design was used in these articles, which was expected.



*Figure 2.* Frequency of research types.

Figure 3 also showed the majority of learning analytics research was conducted with the participants from higher education (*n* = 18, 69%). Other studies included participants from middle school (*n* = 5, 19%), high school (*n* = 2, 8%), and adult education (*n* = 1, 4%). It is interesting to find that there were no studies at the elementary school level. The fact the majority of the studies occurred at higher education level could mean the convenience of conducting such research with higher education students as they probably were more readily accessible to researchers.

*Figure 3.* Frequency of subject matters in different educational levels.

**Predicting learning performances.** Research has shown learning analytics mainly served two roles in predicting students’ learning performances and outcomes. The first role was to help instructors identify performance indicators or predictors across various platforms including MOOCs (de Barba et al., 2016; Pursel et al., 2016), learning forums (Goggins & Xing, 2016), serious games (Gauthier et al., 2015), and learning management systems such as Moodle (Strang, 2017). The second role was to create certain prediction models in order to help educators identify at-risk students (Marbouti et al., 2016; Ornelas & Ordonez, 2017), classify novice and expert learners (Loh, Sheng, & Li, 2015), and predict students’ future college attendance (San Pedro et al., 2017).

In the study by de Barba et al. (2016), learner’s participation data were measured by video hits and quiz attempts in a MOOC course. These two factors were identified as predictors of students’ final performance and their levels of motivation. Students’ participation as a vital predictor of learning performance was also revealed in a MOOC (Pursel et al., 2016). In this study, video views, which was considered as an important indicator of learner participation level, was shown to be a strong predictor of course completion rate, so did forum engagement. Moreover, aiming to gain more insights about students’ participation in online courses, Goggins and Xing (2016) focused on discussion forums, and proposed that instead of only relying on a number of posts, the time dimension of students’ behaviors such as the time spent in reading posts should also be considered. In addition, learning analytics strategies also helped researchers identify other predictors. For example, course login counts and quiz score were shown to be important learning indicators in a Moodle course (Strang, 2017) and tasks completion rate was also found as a significant predictor of post-test in a study about games (Gauthier et al., 2015).

Learning analytics also used to construct prediction models. Ornelas and Ordonez (2017) built a model using students’ log data and their performance scores from eleven courses based on Naive Bayes algorithm to classify at-risk and successful students. Their model showed accuracy over 90%. Another study by Marbouti et al. (2016) aimed to identify at-risk students using analytics. Based on students’ performance data such as in-class quizzes and team participation, the researchers developed a model consisting of Naive Bayes, Support Vector Machine and K-Nearest Neighbors, which was shown to achieve a high accuracy. In addition, learning analytics were also used in building models to predict college enrollment. In a study by San Pedro et al. (2017), the researchers used linear regression based on student data from the online system and assessed their learning experience in terms of knowledge, academic emotions, and engaged and disengaged behaviors. Their results showed that high school course selection and college attendance were related to students’ performance and engagement at the middle school level. Furthermore, learning analytics were used in studying serious games to predict expert–novice performances (Loh et al., 2015). In this study, the researchers coded players’ action sequences by tracking the path they had traversed in the game to differentiate and compare expert and novice players. The results showed the efficiency of incorporating tile-based action sequence coding approach in a serious games analytic model.

The findings of these studies in using analytics to identify learning performance predictors and creating prediction model construction can help instructors and researchers across various subjects and platforms to better design learning environments so as to enhance students’ learning.

**Detecting behaviors and learning patterns.** Learning analytics were also used by researchers and educators to detect learning patterns across various educational platforms including serious games (Cagiltay et al., 2015; Cheng et al., 2015; Cheng, Rosenheck, Lin, & Klopfer, 2017; Gauthier et al., 2015; Kerr, 2015), simulations (Angeli et al., 2017; Chang et al., 2017), learning management systems (Giannakos et al., 2016; Yang, Li, & Xing, 2018), collaboration platforms (Zhang, Meng, Ordóñez de Pablos, & Sun, 2017) and blended learning environments (Van Laer & Elen, 2016). The findings of these studies were discussed in the following sections.

***Learning patterns in serious games*.** A study from Cheng et al. (2015) examined the effects of a serious game for 7th-grade biology and the interplay of student concept learning, gaming performance, and in-game behaviors. Analysis showed that students’ in-game behaviors, such as frequencies and duration of viewing specific content in the game, were significantly correlated with their game performance which subsequently influenced their learning outcomes. Moreover, cluster analysis revealed three groups of learners featuring low learning outcomes/low gaming performance, high learning outcomes, and high gaming performance. The study by Cagiltay et al. (2015) also used similar descriptors such as duration of viewing concept explanations, total response time to questions and accuracy (game performance) to understand students’ in-game behavioral patterns in relation to their motivation and post-test scores. The results of this study showed that learners’ motivation and post-test scores in a game environment with competition settings were significantly higher than those who played the non-competition game.

Cheng et al. (2017) collected log data from high school students as they played a Massively Multiplayer Online Game (MMOG) for science learning. Data from a specific genetics quest in the game were analyzed by using data mining techniques to examine the relationship between tool use and quest completion status, how the use of certain tools may influence content-related game choices, and the multiple pathways available to players in the game. The study identified that in this particular quest, learners’ use of two tools, the “trait examiner” and “trait decoder,” was more likely to lead to success.

Gauthier et al. (2015) investigated the educational impact of a serious game for studying human vascular anatomy versus a similar non-game study aid and how it related to the participants’ demographic traits and voluntary use over a 35-day period. Their analyses suggested that game mechanics encouraged more specific problem-solving strategies than the non-game study aid did, leading to greater predictability of learning outcomes. However, there was no significant difference in frequencies of tool-use between the experimental and control groups. It also appeared that students' studying habits had the greatest influence on the level of engagement indicated by frequencies of tool-use.

Kerr and Chung’s study (2012) on a 6th grade math game found that (a) it was common for students to pass certain levels using incorrect mathematical strategies and (b) throughout the game, a large number of students used order-based strategies to solve problems rather than strategies based on mathematics. Based on this study, Kerr’s (2015) modified this game and randomly assigned students to play either the original version or the revised version. The study examining the effect on student performance used two in-game measures of performance (the number of attempts per level and the percentage of first attempts that were solutions) and two paper-based measures of performance (immediate post-test score and delayed post-test score). The results showed that students who played the revised version significantly reduced the use of incorrect mathematical strategies and order-based strategies than students who played the original version.

***Learning patterns in simulations.*** Chang et al. (2017) studied students’ different behavioral patterns while solving physics problems in both individual-based and collaborative simulations. Lag sequential analysis revealed that students’ learning patterns in these two simulations were different significantly. Students in collaborative simulations, although they presented a higher level of collaborative activity, did not transform discussions into suitable problem-solving activities, while students in the individual-based simulation completed individual learning by exploring independently before heading towards group reflection. Angeli et al. (2017) conducted a similar study, but instead of dividing learners into two simulations, the researchers separated learners based on their cognitive styles (i.e. field dependent and field independent) by using educational data mining techniques. The results showed that learners with different cognitive styles exerted different patterns of interactions in the same simulation during problem-solving process. These two studies illustrated how learning analytics were used in a simulation to detect students’ interaction patterns.

***Learning patterns in learning management systems.*** Giannakos et al. (2016) analyzed learner’s usage patterns such as watching period, platforms used and video duration in an online video-assisted software engineering course and found out that these patterns were related to students’ attitudes while learning and interacting on this platform. Thus, this research provided a new direction of analyzing learner attitudes via behaviors patterns. Moreover, behavior patterns discovered by learning analytics were also used to identify students’ levels of knowledge construction and engagement. Yang et al. (2018) conducted a study on students’ knowledge construction in translation activities using log data. The analysis from the lag sequential analysis showed that knowledge co-construction behaviors occurred more continually and frequently in the higher-engagement group.

***Learning patterns in collaboration tools/blended learning.*** Zhang et al. (2017) examined students’ collaboration patterns on a group collaborative platform—Slack. The results using partial least squares revealed that students’ interaction patterns involving social influence, mutual trust, as well as social and academic reward valence had positive influences on students' teamwork engagement. Van Laer and Elen (2016) analyzed log data from a Moodle course using K-means cluster analysis, one-way ANOVA, and MANOVA, and three clusters were found. The cluster with the highest standardized scores showed that the frequency of a learners’ sequential pattern such as from “course module viewed” to “discussion made” was much higher than the other two groups, while the cluster with the lowest scores had higher frequency of the sequential pattern from “discussion made” to “test made.” These findings based on learning analytics revealed important self-regulatory behaviors patterns in a blended learning environment.

In summary, the studies in this category mainly focused on two aspects of learning when analyzing the log data: learners’ engagement and strategies. To measure learners’ engagement, the researchers often calculated the duration and frequency of performing specific tasks or accessing specific tools (Cagiltay et al., 2015; Cheng et al., 2015; Gauthier et al., 2015; Sipiyaruk et al., 2017; Van Laer & Elen, 2016). In contrast, to investigate learners’ strategies, the researchers mainly examined learners’ choices of using certain tools/performing certain actions, or the sequence of performing different tasks/actions (Angeli et al., 2017; Chang, et al., 2017; Cheng et al., 2017; Gauthier et al., 2015; Kerr, 2015; Yang et al., 2018). In some cases, both engagement and strategies were analyzed in order to present a more comprehensive view of learners’ behaviors in certain platforms (Chang et al., 2017; Cheng et al., 2017).

It is also noteworthy that, instead of discussing learning patterns in isolation, many studies attempted to understand the causes and implications of diverse learning patterns. In this regard, researchers often examined the correlation between learning patterns and other elements, such as different system settings that may trigger varied behaviors (Cagiltay et al., 2015; Chang et al., 2017; Gauthier et al., 2015), learning outcomes associated with certain learning patterns (Cagiltay et al., 2015; Cheng et al., 2015; Cheng et al., 2017; Kerr, 2015; Snow, Allen, Jacovina, & McNamara, 2015), learners’ motivations (Cagiltay et al., 2015; Giannakos et al., 2016) and cognitive/learning styles (Angeli et al., 2017; Van Horne et al., 2018).

**Facilitating learning experiences.** Previous research found learning analytics had played a crucial role in facilitating learning by adapting to individual needs (Martin & Wintmer, 2015; Van Horne et al., 2018; Van Leeuwen et al., 2015) and by optimizing learning platforms (Angeli et al., 2017; de Kock & Harskamp, 2016; Giannakos et al., 2016; Kerr, 2015). Learning analytics offer new opportunities to adapt learning, such as using analytics to create dashboards which can not only improve students’ engagement levels and performances (Horn et al., 2017; Martin & Wintmer, 2015), but also facilitate teachers to monitor student learning progress (Van Leeuwen et al., 2015). In a study by Van Horne et al. (2018), researchers created a dashboard that can provide students real-time learning progress, assessments scores, and learning suggestions, using the data generated from students’ log activities as well as their learning profiles such as their high school GPA. The result revealed that users with higher frequencies of accessing the dashboard obtained significantly higher grades than those who had lower frequencies. Van Leeuwen et al. (2015) used a dashboard to assist teachers to diagnose students’ learning obstacles thus to intervene accordingly during a computer-supported collaborative writing activity. Their result showed that using the dashboard encouraged teachers to intervene more often in general and especially with targeted groups who were undergoing difficulties with studies.

Learning analytics techniques also enabled educators to optimize educational platforms to further improve students’ learning performances. Kerr (2015) modified their educational game based on data mining results so that to reduce construction of irrelevant behaviors in the game. This modification also led to more positive perceptions of the game compared with the original setting. In another research conducted by de Kock and Harskamp (2016), the researchers applied learning analytics techniques to compare the effects of two types of hints on student performance during mathematical word problem-solving activities in a learning management system. Researchers used pre- and post-test to examine the differences in problem-solving skills and collected log files. They found that procedural-content hints could better facilitate students’ problem-solving skills than procedural only hints, which led to an enhanced approach for this learning platform by incorporating more procedural-content hints. Moreover, Giannakos et al. (2016) analyzed students’ learning patterns and discovered that the patterns were related to students’ attitudes towards the lectures. The finding offered designers useful insights for improving video-based lectures to fulfill their teaching objectives. These findings demonstrated the potentials of using analytics to improve and enhance efficient learning environments and cultivate more positive learning behaviors and attitudes.

Although we discussed the recent research on learning analytics’ role to support students’ learning in three categories, these categories are not isolated from each other. In fact, they are interrelated. The capability of a learning analytic method to predict students’ learning performance was also determined by how well this method can detect students’ behavior patterns (de Barba et al., 2016; Strang, 2017) such as access frequencies in an LMS (de Kock & Harskamp, 2016; Giannakos et al., 2016). Instructors could gather and interpret the data to give feedback accordingly (Van Leeuwen et al., 2015). Moreover, in order to better optimize a learning platform such as a serious game, researchers would need to analysis how learners’ behavioral patterns had changed (Cagiltay et al., 2015; Cheng et al., 2015) and whether learning performance had improved (Cheng et al., 2017; Gauthier et al., 2015) thus to evaluate a modification (Kerr, 2015). The research has shown the efficacy of using analytics as performance predictors and learning behavior detectors to facilitate teaching and learning and improve the designs of learning environments.

**Summary and Conclusion**

This review of literature presented a systematic review of how learning analytics have been used in different educational settings from 2015 to present and identified factors that were related to learning in these studies. A total of 26 peer-reviewed studies were included in this review. The discussion focused on technologies involved, analytic techniques and data used, types of research designs, subject matter areas, and educational levels. Importantly, the use of analytics in connection to learning was analyzed. The findings showed that the most common categories as emerged in these studies included predicting learning performances, detecting behaviors and learning patterns, and facilitating learning experiences. Specifically, learning analytics were used to build predictive models in order to predict student performance so as to help instructors provide more effective interventions to assist students’ learning. Studies have shown by identifying the learning patterns, learning analytics provided useful insights into ways to motivate and engage learners and stimulate a deeper understanding of knowledge. Many studies also utilized more than two data sources as well as various analytic techniques in one study to address research questions. Survey data and user logs were the two most frequently used data types and analytic techniques such as t-test, ANOVA, and cluster analysis were more frequently used to discover user behavior patterns. We hope the findings of this literature review on most recent literature in learning analytics for educational purposes will provide useful information and insights to researchers interested in the topic and inform educational practices.

As the problem is complex enough for students to spend at least a week of science class or even a full month, former research on the game have found that

Although the above two approaches can detect and predict learning styles to some extent with

certain accuracy, there are obvious limitations when apply. For example, learning behavior could

change in different courses due to the knowledge backgrounds and learning habits of learners; learning style could even change in different parts of the same course, for the difficulty level or knowledge type changes. In a machine-based learning environment, however, the current system is not capable to detect such changes in learning styles of learners and hence might make the improper recommendations. Therefore, dynamic detecting is an essential requirement for modern massive online learning facilities.

Learning style detection module will continuously track the user behavior records and take time

frame into consideration as well. Each log(channel), is mapped into a time(by day)-content(by chapter) based 2-dimensions array, the nodes of each feature events are associated with a timestamp and normalized to the same level.

as an example, logs from “Data Structure” (Part 1, 2017), one of most popular MOOC courses

on xuetangX.com with more than 23,000 students enrolled is examined. The course includes six

chapters and lasts 73 days. Two learners are randomly selected and the logs for their effective learning points density are visualized as Figure

Figure 4. Log for effective learning points density of two users.

Based on the aforementioned FSLSM summarized in Table 1 we can infer the difference in learning styles of the two users: (a) User #4441580 learning one by one chapter in sequence (Sequential progress) and got a higher point in learning point density than average (Think more detail, Detail Oriented), evidently expressed the tendency of Reflective and Sequential type; (b) User #6565569 learning more than one chapter in one day four times (Try something out, Overall picture, Non-sequential progress), thus tend to behave in a more Active and Global manner.

**References**

(\* Indicates articles included in this literature review)

\*Angeli, C., Howard, S. K., Ma, J., Yang, J., & Kirschner, P. A. (2017). Data mining in educational technology classroom research: Can it make a contribution? *Computers and Education*, *113*, 226–242. <https://doi.org/10.1016/j.compedu.2017.05.021>

\*Cagiltay, N. E., Ozcelik, E., & Ozcelik, N. S. (2015). The effect of competition on learning in games. *Computers and Education*, *87*, 35–41. <https://doi.org/10.1016/j.compedu.2015.04.001>

\*Chang, C. J., Chang, M. H., Liu, C. C., Chiu, B. C., Fan Chiang, S. H., Wen, C. T., … Chai, C. S. (2017). An analysis of collaborative problem-solving activities mediated by individual-based and collaborative computer simulations. *Journal of Computer Assisted Learning*, *33*(6), 649–662. <https://doi.org/10.1111/jcal.12208>

\*Cheng, M. T., Lin, Y. W., & She, H. C. (2015). Learning through playing Virtual Age: Exploring the interactions among student concept learning, gaming performance, in-game behaviors, and the use of in-game characters. *Computers and Education*, *86*, 18–29. <https://doi.org/10.1016/j.compedu.2015.03.007>

\*Cheng, M. T., Rosenheck, L., Lin, C. Y., & Klopfer, E. (2017). Analyzing gameplay data to inform feedback loops in The Radix Endeavor. *Computers and Education*, *111*, 60–73. <https://doi.org/10.1016/j.compedu.2017.03.015>

\*de Barba, P. G., Kennedy, G. E., & Ainley, M. D. (2016). The role of students’ motivation and participation in predicting performance in a MOOC. *Journal of Computer Assisted Learning*, *32*(3), 218–231. <https://doi.org/10.1111/jcal.12130>

\*de Kock, W. D., & Harskamp, E. G. (2016). Procedural versus content-related hints for word problem solving: an exploratory study. *Journal of Computer Assisted Learning*, *32*(5), 481–493. <https://doi.org/10.1111/jcal.12148>

\*Gauthier, A., Corrin, M., & Jenkinson, J. (2015). Exploring the influence of game design on learning and voluntary use in an online vascular anatomy study aid. *Computers and Education*, *87*, 24–34. <https://doi.org/10.1016/j.compedu.2015.03.017>

\*Giannakos, M. N., Jaccheri, L., & Krogstie, J. (2016). Exploring the relationship between video lecture usage patterns and students’ attitudes. *British Journal of Educational Technology*, *47*(6), 1259–1275. <https://doi.org/10.1111/bjet.12313>

\*Goggins, S., & Xing, W. (2016). Building models explaining student participation behavior in asynchronous online discussion. *Computers and Education*, *94*, 241–251. <https://doi.org/10.1016/j.compedu.2015.11.002>

\*Kerr, D. (2015). Using Data Mining Results to Improve Educational Video Game Design. *Journal of Educational Data Mining*, *7*(3), 1–17.

Liu, M., Kang, J., Liu, S., Zou, W., & Hodson, J. (2017). Learning analytics as an assessment tool in serious games: A review of literature. In *Serious Games and Edutainment Applications: Volume II* (pp. 537–563). Springer International Publishing. https://doi.org/10.1007/978-3-319-51645-5\_24

\*Liu, M., Lee, J., Kang, J., & Liu, S. (2016). What We Can Learn from the Data: A Multiple-Case Study Examining Behavior Patterns by Students with Different Characteristics in Using a Serious Game. *Technology, Knowledge and Learning*, *21*(1), 33–57. <https://doi.org/10.1007/s10758-015-9263-7>

\*Loh, C. S., Sheng, Y., & Li, I. H. (2015). Predicting expert-novice performance as serious games analytics with objective-oriented and navigational action sequences. *Computers in Human Behavior*, *49*, 147–155. <https://doi.org/10.1016/j.chb.2015.02.053>

Siemens, G., & Long, P. (2011). Penetrating the Fog: Analytics in Learning and Education. EDUCAUSE Review, *46*, 30–32. https://doi.org/10.1145/2330601.2330605

\*Marbouti, F., Diefes-Dux, H. A., & Madhavan, K. (2016). Models for early prediction of at-risk students in a course using standards-based grading. *Computers and Education*, *103*, 1–15. <https://doi.org/10.1016/j.compedu.2016.09.005>

\*Martin, F., & Whitmer, J. C. (2016). Applying Learning Analytics to Investigate Timed Release in Online Learning. *Technology, Knowledge and Learning*, *21*(1), 59–74. <https://doi.org/10.1007/s10758-015-9261-9>

\*Ornelas, F., & Ordonez, C. (2017). Predicting Student Success: A Naïve Bayesian Application to Community College Data. *Technology, Knowledge and Learning*, *22*(3), 299–315. <https://doi.org/10.1007/s10758-017-9334-z>

\*Pursel, B. K., Zhang, L., Jablokow, K. W., Choi, G. W., & Velegol, D. (2016). Understanding MOOC students: Motivations and behaviours indicative of MOOC completion. *Journal of Computer Assisted Learning*, *32*(3), 202–217. <https://doi.org/10.1111/jcal.12131>

\*San Pedro, M. O. Z., Baker, R. S., & Heffernan, N. T. (2017). An Integrated Look at Middle School Engagement and Learning in Digital Environments as Precursors to College Attendance. *Technology, Knowledge and Learning*, *22*(3), 243–270. <https://doi.org/10.1007/s10758-017-9318-z>

Siemens, G. (May 2012). Learning Analytics: Envisioning a Research Discipline and a Domain of Practice. *2nd International Conference on Learning Analytics & Knowledge*, 4–8. https://doi.org/10.1145/2330601.2330605

\*Sipiyaruk, K., Gallagher, J. E., Hatzipanagos, S., & Reynolds, P. A. (2017). Acquiring Critical Thinking and Decision-Making Skills: An Evaluation of a Serious Game Used by Undergraduate Dental Students in Dental Public Health. *Technology, Knowledge and Learning*, *22*(2), 209–218. <https://doi.org/10.1007/s10758-016-9296-6>

\*Snow, E. L., Allen, L. K., Jacovina, M. E., & McNamara, D. S. (2015). Does agency matter?: Exploring the impact of controlled behaviors within a game-based environment. *Computers and Education*, *82*, 378–392. <https://doi.org/10.1016/j.compedu.2014.12.011>

\*Strang, K. D. (2017). Beyond engagement analytics: which online mixed-data factors predict student learning outcomes? *Education and Information Technologies*, *22*(3), 917–937. <https://doi.org/10.1007/s10639-016-9464-2>

\*Van Horne, S., Curran, M., Smith, A., VanBuren, J., Zahrieh, D., Larsen, R., & Miller, R. (2018). Facilitating Student Success in Introductory Chemistry with Feedback in an Online Platform. *Technology, Knowledge and Learning, 23*(1), 21-40.

\*Van Laer, S., & Elen, J. (2018). Adults’ Self-Regulatory Behaviour Profiles in Blended Learning Environments and Their Implications for Design. *Technology, Knowledge and Learning*, 1–31. <https://doi.org/10.1007/s10758-017-9351-y>

\*Van Leeuwen, A., Janssen, J., Erkens, G., & Brekelmans, M. (2015). Teacher regulation of cognitive activities during student collaboration: Effects of learning analytics. *Computers and Education*, *90*, 80–94. <https://doi.org/10.1016/j.compedu.2015.09.006>

\*Yang, X., Li, J., & Xing, B. (2018). Behavioral patterns of knowledge construction in online cooperative translation activities. *Internet and Higher Education*, *36*, 13–21. <https://doi.org/10.1016/j.iheduc.2017.08.003>

\*Zhang, X., Meng, Y., Ordóñez de Pablos, P., & Sun, Y. (2017). Learning analytics in collaborative learning supported by Slack: From the perspective of engagement. *Computers in Human Behavior*. https://doi.org/10.1016/j.chb.2017.08.012

Chollet, F. (2017). *Deep learning with python*. Manning Publications Co..

ECAR-Analytics Working Group. (2015). The predictive learning analytics revolution: Leveraging learning data for student success. *ECAR working group*.

Hicks, D., Eagle, M., Rowe, E., Asbell-Clarke, J., Edwards, T., & Barnes, T. (2016, April). Using game analytics to evaluate puzzle design and level progression in a serious game. *In Proceedings of the Sixth International Conference on Learning Analytics & Knowledge* (pp. 440-448). ACM.

Huang, S., & Fang, N. (2013). Predicting student academic performance in an engineering dynamics course: A comparison of four types of predictive mathematical models. *Computers & Education*, 61, 133-145.

Kang, J., Liu, M., & Qu, W. (2017). Using gameplay data to examine learning behavior patterns in a serious game. *Computers in Human Behavior*, 72, 757-770.

Kizilcec, R. F., Pérez-Sanagustín, M., & Maldonado, J. J. (2017). Self-regulated learning strategies predict learner behavior and goal attainment in Massive Open Online Courses. *Computers and Education*, 104, 18–33. <https://doi.org/10.1016/j.compedu.2016.10.001>

Kerr, D. (2015). Using Data Mining Results to Improve Educational Video Game Design. *Journal of Educational Data Mining*, *7*(3), 1–17.

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE, 86*(11), 2278-2324.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, 521(7553), 436.

Loh, C. S., & Sheng, Y. (2015). Measuring expert performance for serious games analytics: From data to insights. In *Serious Games Analytics* (pp. 101-134). Springer, Cham.

Mills, C., Bosch, N., Graesser, A., & D’Mello, S. (2014, June). To quit or not to quit: predicting future behavioral disengagement from reading patterns. In *International Conference on Intelligent Tutoring Systems* (pp. 19–28). Springer, Cham.

Min, W., Wiggins, J., Pezzullo, L., Vail, A., Boyer, K. E., Mott, B., Frankosky, M., Wiebe, E., & Lester, J. (2016). Predicting dialogue acts for intelligent virtual agents with multimodal student interaction data. In T. Barnes, M. Chi, & M. Feng (Eds.), *Proceedings of the 9th international conference on educational data mining.* Raleigh, NC: International Educational Data Mining Society.

Muratov, E., Lewis, M., Fourches, D., Tropsha, A., & Cox, W. C. (2017). Computer-assisted decision support for student admissions based on their predicted academic performance. *American journal of pharmaceutical education*, 81(3), 46.

Paquette, L., Rowe, J., Baker, R., Mott, B., Lester, J., DeFalco, J., ... & Georgoulas, V. (2016). Sensor-Free or Sensor-Full: A Comparison of Data Modalities in Multi-Channel Affect Detection. *International Educational Data Mining Society.*

San Pedro, M. O. Z., Baker, R. S., & Heffernan, N. T. (2017). An Integrated Look at Middle School Engagement and Learning in Digital Environments as Precursors to College Attendance. *Technology, Knowledge and Learning*, *22*(3), 243–270. <https://doi.org/10.1007/s10758-017-9318-z>

Seidel, E., & Kutieleh, S. (2017). Using predictive analytics to target and improve first year student attrition. *Australian Journal of Education*, 61(2), 200–218.

Siemens, G., & Long, P. (2011). Penetrating the Fog: Analytics in Learning and Education. *EDUCAUSE Review*, *46*, 30–32. <https://doi.org/10.1145/2330601.2330605>

Siemens, G. (May 2012). Learning Analytics: Envisioning a Research Discipline and a Domain of Practice. *2nd International Conference on Learning Analytics & Knowledge*, 4–8. https://doi.org/10.1145/2330601.2330605

Stoehr, M. (2014, August 14). *Deep PCA Nets* [Hand-written digital image]. Retrieved July 13, 2018, from http://deepdish.io/public/images/PCANet\_mnist5\_patch.png