# Digital Game-Based Learning Implications for Education

Anthony L Wilson

College of Science and Technology, Bellevue University

DSC 680: Applied Data Science

Professor Catie Williams

4/29/2023

# Digital Game-Based Learning Implications for Education

## Background and History

Digital Game-based learning (DGBL) is integral to our educational system. Game-based learning allows the user to learn through trial and error. Users get to experience consequences immediately after each decision. ("What is Game-Based Learning," 2021) If the user answers incorrectly, they can navigate different options and learn by repetition. Users are immediately rewarded when making a correct decision and get to move on to the next level. There are benefits for educators where they can quickly see how fast students are learning and advancing in the lessons being taught. Traditional learning with textbooks and paper assignments offers a passive approach to learning, whereas game-based learning is active ("What is Game-Based Learning," 2021). With advancements in data science and machine learning models, there is an opportunity to help identify learning patterns for users and progress in our education system.

To better understand DGBL, we have used the Kaggle Predict Student Performance from the Game Play data set. (*Predict student performance from gameplay*) This data was published on Kaggle by the Learning Agency Lab and came from the game Jo Wilder and the Capitol Case. The whole point of this model is to predict the user's answer to 18 questions. The game is trace-based, meaning that we can see everywhere the user clicks on the screen. Based on where they click can help identify if they will be able to answer the question correctly.

## Business Problem

The problem to be explored is that DGBL is a legitimate method for educators. Middle school guidance counselor Rachelle Vallon suggests many ways DGBL can benefit students. She states, "There are so many amazing ways to use games, game-like experiences, and the design process to engage students, not just in academic things, but in things that have to do with conflict or processing feelings"

(McMahon, 2022). Previous studies have shown that DGBL compared to traditional lectures, can improve scores between 7% and 40% (Eck, 2015). This study will utilize new data to help determine if educators can match students' intellect and ultimately meet the needs of student's individual needs.

**Data Explanation**

Before building and running the model, the data needed to be prepped. First, the data was examined for anomalies and missing data; missing values were filled in. It was essential to ensure the sessions followed a sequential order. Next, the data was checked to see if the users' sessions leveled consistently. Another issue was identifying when the question was asked in the game. The creators of the competition removed some of the essential information specified when a question was asked. So we aggregated the data. A data dictionary is included in Appendix A for reference.

**Methods**

We used the Extreme Gradient Boosting, XGBoost, model, a distributed gradient-boosted decision tree method. XGBoost does well with classification problems. The creators removed quite a bit of data; this limited the ability to look at the data line by line. The data was aggregated by session ID and level group. Levels went from zero to twenty-two, and level groups consisted of levels zero to four, five to twelve, and thirteen to twenty-two. The mean and standard deviation were calculated to help identify where the average session would click or the average elapsed time between each interaction in each level group by session. For each categorical variable the unique values were summed up. Some variables were removed due to their low correlation with the data set. Using this XGBoost, we created ten k-fold cross-validations.

**Analysis**

Based on the XGBoost, the best threshold to convert probabilities for the user's answer was 0.63. Based on the training data set, the model's F1 score was .677, as shown in Figure 1. Based on the training data set, the model is 68% accurate in predicting users' answers within the sample.

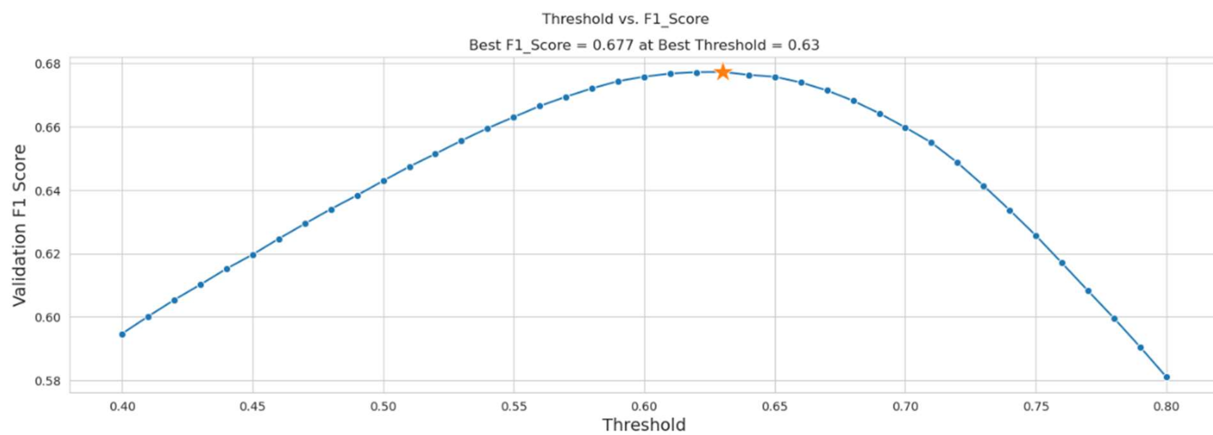


*Figure 1: Threshold vs. F1_Score* (Wilson, 2023).

Figure 2 shows that the model is anywhere from just above 40% accurate for each question to about 60% for this sample. Question thirteen is the only one with an F1 score below 0.45, and eight questions have an F1 score above 0.55.

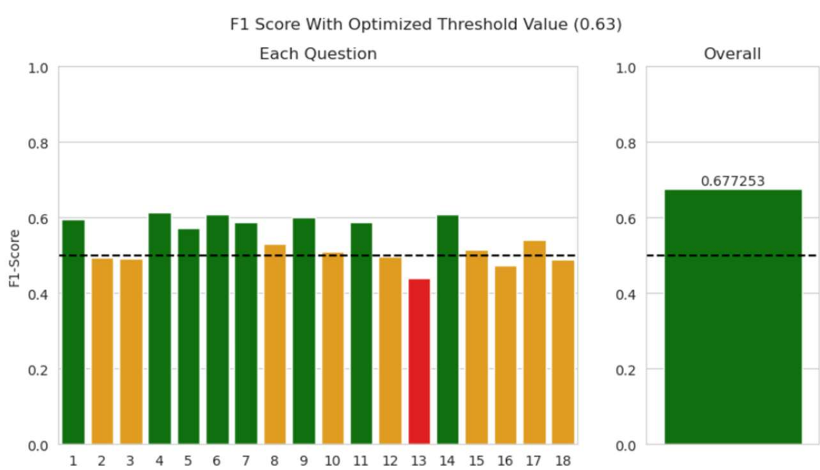


*Figure 2: F1 Score with Optimized Threshold Value (0.63)* (Wilson, 2023).

When running the test data to see how accurate the model is, the F1 score was only 61% correct at 0.63 threshold conversion. Indicating that the model would still need some more fine-tuning, but it is still more accurate than guessing.

As mentioned, our process involved filling in data for missing values. After the data was aggregated, it would be populated with a -1 because the model had to have a value for the decision trees to work appropriately. The graph in Figure 2 depicts the percentage of missing values for each column in the data set. Because some of those values were removed, it may have affected the model's performance.

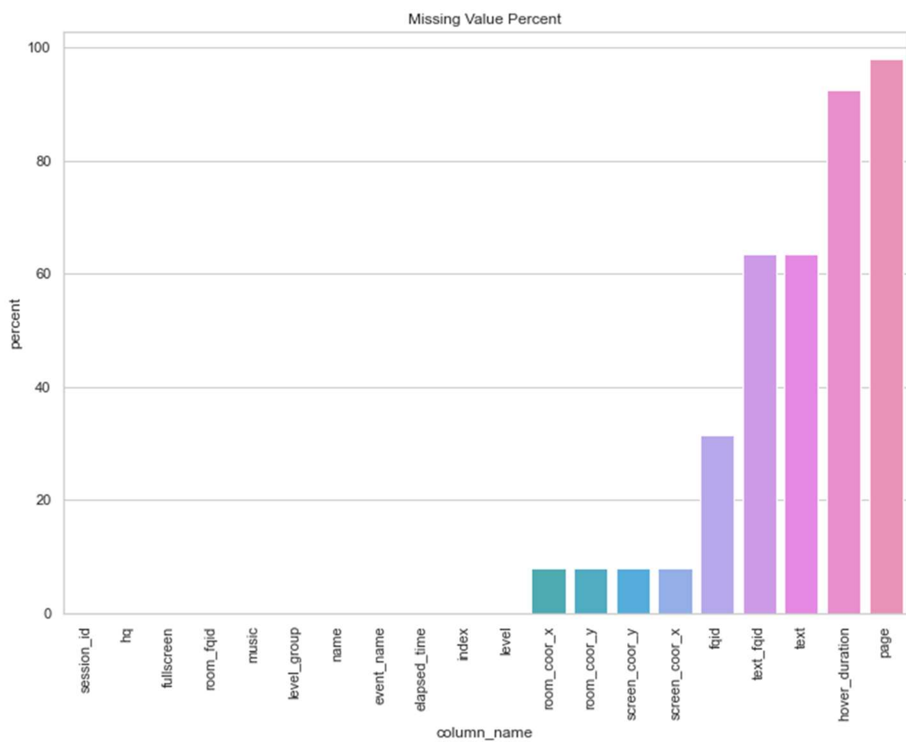


*Figure 3: Missing Value Percent* (Wilson, 2023).

**Assumptions**

Certain assumptions have been made about the data set that guided this study. It is assumed that the data was gathered ethically without bias. Although the game is widely available online, it is also believed that the data was only gathered from students between the $4^{th}$ and $6^{th}$ grades as specified. It is also assumed that the coordinates for the correct answers are all inserted in the same spot and order.

**Limitations**

Specific limitations of this study affect the reliability and validity of the study. First, the data is not necessarily complete and accurate because certain values were removed. Another potential limitation is that little information about the end user could negatively or positively affect the model's outcome. Third, once the user reaches the end of their session, they can level down. When they level up again, it can cause the predictions to be less accurate. Inaccurate leveling up may be a bug in the game's creation and influence the model. Another limitation surrounds the elapsed time. Elapsed time would reset so it would seem like the user would start over or go back in time, which may also influence the model.

**Challenges**

Some challenges inherent in this model due to assumptions and limitations are that it is difficult to do it line by line because the data has been removed, so averages had to be used. Additionally, it is hard to determine the precise amount of time the student used to decide, which influences the model's integrity.

**Future Uses/Additional Applications**

The specific intent of this model applies to education; however, other practices and areas in this data could be of use. DGBL methods may be applied to aptitude tests, certifications, training in the business world, job interviews, and even promotions within the workplace. The ability to predict can help identify knowledge gaps and be used in various settings.

**Recommendations**

Based on the model, DGBL methods may be helpful in educational settings. The model predicted with 68% accuracy in the training data and 61% in the test data what the user would answer. The accuracy suggests that DGBL may successfully further the students' learning opportunities within the sample. However, further collection and refinement of the data and model are needed to propose this as an alternative solution more confidently.

**Implementation Plan**

If this method were to be used in educational settings, we would start with a small group. Next, the plan would need to be refined based on initial results. Third, more data would be gathered to start adding in other groups. Again, this must be refined until sufficient accuracy justifies widespread implementation.

**Ethical Assessment**

The ethical implications must be considered to retain this study's validity and reliability. Nothing unethical seems to be apparent in this method or model. There is no demographic information specifically provided within the data. Missing data has been noted as a limitation, so if we were to add a demographic component to the data, there could be some ethical issues with the model.

**Conclusion**

In education, there is an increasing shortage of educational resources to fill the needs of students. DGBL is a potential solution to help bridge that gap. Even with these limitations, educators can use programs to help engage students in interactive learning. To determine if this is a viable solution, we developed a model utilizing the XGBoost method to analyze the data surrounding DGBL. Amidst limitations and challenges inherent in the data, we found that the model was 61% accurate based on the sample. There is still more development and refining of the model. It is recommended that DGBL still be researched and considered.

**References**

Eck, R. V. (2015, October 12). *Digital game-based learning: Still restless, after all these years*. EDUCAUSE. Retrieved April 25, 2023, from https://er.educause.edu/articles/2015/10/digital-game-based-learning-still-restless-after-all-these-years

McMahon, W. (2022, November 7). *How game-based learning develops real-world skills - edsurge news*. EdSurge. Retrieved April 25, 2023, from https://www.edsurge.com/news/2022-11-07-how-game-based-learning-develops-real-world-skills

*Predict student performance from game play*. Kaggle. (n.d.). Retrieved April 1, 2023, from https://www.kaggle.com/competitions/predict-student-performance-from-game-play/overview

*Research guides: Digital Pedagogy - A Guide for librarians, faculty, and students: Game based learning - why do it: Benefits, challenges*. Game Based Learning - Why Do it: Benefits, Challenges - Digital Pedagogy - A Guide for Librarians, Faculty, and Students - Research guides at University of Toronto. (n.d.). Retrieved April 23, 2023, from https://guides.library.utoronto.ca/c.php?g=448614&p=3505475

*What is game-based learning?* GCU. (2021, October 11). Retrieved April 23, 2023, from https://www.gcu.edu/blog/teaching-school-administration/what-game-based-learning

Wilson, A. (2023, April 29). *DGBL*. GitHub. Retrieved April 29, 2023, from https://github.com/wilsanth/DSC_680/tree/main/DGBL

**Appendix A: Data Dictionary**

| Column | Definition |
|---|---|
| session_id | the ID of the session the event took place in |
| index | the index of the event for the session |
| elapsed_time | how much time has passed (in milliseconds) between the start of the session and when the event was recorded |
| event_name | the name of the event type |
| name | the event name (e.g. identifies whether a notebook click is opening or closing the notebook) |
| level | what level of the game the event occurred in (0 to 22) |
| page | the page number of the event (only for notebook-related events) |
| room_coor_x | the coordinates of the click in reference to the in-game room (only for click events) |
| room_coor_y | the coordinates of the click in reference to the in-game room (only for click events) |
| screen_coor_x | the coordinates of the click in reference to the player's screen (only for click events) |
| screen_coor_y | the coordinates of the click in reference to the player's screen (only for click events) |
| hover_duration | how long (in milliseconds) the hover happened for (only for hover events) |
| text | the text the player sees during this event |
| fqid | the fully qualified ID of the event |
| room_fqid | the fully qualified ID of the room the event took place in |
| text_fqid | the fully qualified ID of the |
| fullscreen | whether the player is in Fullscreen mode |
| hq | whether the game is in high-quality |
| music | whether the game music is on or off |
| level_group | which group of levels - and group of questions - this row belongs to (0-4, 5-12, 13-22) |

(*Predict student performance from gameplay*)

**Appendix B: Questions**

1. ***Did you consider running any other models?***

    Yes, this model is known for classification problems. I want to take more time and look at other potential models to see if something is better. For example, I would like to look into something LightGBM potentially.

2. ***The data was aggregated. Is there a way to run the model row by row?***

    Yes, but due to resources, it was a better fit to aggregate the data for this project. This would be something that could be addressed in future enhancements of the model/modeling.

3. ***Is there more data available that is similar to the project to gain a better understanding of DGBL?***

    There are not a lot of DGBL datasets that are publicly available. Most likely private companies have this data available for their teams to analyze and use.

4. ***What demographics would be good for this study?***

    At a minimum, the last completed grade level and age. Demographics like grade and age would potentially help in understanding skills in analyzing the questions. It would be good to know the approximate location, rural vs. urban, and parents highest completed degree and career.

5. ***More refinement would need to be done in this model; what would that entail?***

    Diving a bit deeper into the coordinates and pages would be where I would start. In the game, you have to use the journal, represented by the page's column, to answer questions to move on to the next level, with some of the text columns missing that identify the boss and questions being asked. I would assume a specific order must be followed to move to the next round. This is where a row-by-row model would be effective.

6. ***Would the current model include the ability to determine more difficult questions for the user to challenge them?***

   No, this model is not running while the game is live. We didn't have access to all of the live data. As we continue to develop the model and its accuracy, we could use it to help build something offering more advanced features, such as increasing the difficulty level. This would also require us to be able to work with the game designers and developers.

7. ***Are there any advantages for those who have disabilities using DGBL?***

   Future studies may address this, and there could be studies that have already identified the benefits for those with disabilities. This study doesn't tackle this question due to a lack of information on the users.

8. ***Why are there different thresholds tested when converting probabilities?***

   Naturally, we want to set the rounding up to one to be anything greater than or equal to 0.5, but that causes an issue when there is an imbalance. This will cause poor performance in the model. We use a grid search or a ROC (Receiving Operating Characteristics curve) to help identify the best threshold for converting classification variables***.***

9. ***Why did the creators remove important data when predicting the user's answer?***

   I am not sure why they chose to remove the data.

10. ***How much DGBL is being implemented in schools now?***

    I am unsure about the amount, but it is becoming more widely available.