Representation of the Student's Controllable Performance Features
Based on PS2CLH Model

Arlindo Almada, London Metropolitan University, United Kingdom
Qicheng Yu, London Metropolitan University, United Kingdom
Preeti Patel, London Metropolitan University, United Kingdom

The Barcelona Conference on Education 2022
Official Conference Proceedings

Abstract
Nowadays, the number of studies measuring and representing students' learning and performance has increased. However, there remains a lack of research that represents and measures factors or features within students' control that impact their performances. For university managers, subject tutors and academic mentors, it is essential to represent, measure, analyse and monitor student performance alongside controllable factors affecting their academic achievement to enhance the student experience. This research evaluates the connection among students' performance and their lifestyles, particularly the controllable factors. Controllable factors incorporated in our PS2CLH model are the perspectives of Psychology, Self-responsibility, Sociology, Communication, Learning and Health & wellbeing. This paper proposes a controllable performance features representation in three-dimensional space based on the PS2CLH model. A cluster presentation of the features allows for targeted interventions for students who need additional support. It also indicates clearly where each student stands by using a student web profile and the necessary direction each student needs to take to get to the desired cluster. Initial data presents a clear pattern of creating a diagonal of seven clusters or students' stages from the bottom (0, 0, 0) to the top (100, 100, 100) and leading to the use of filters or queries to represent better features such as sleep-problem, stress, practice exercises and time management. Preliminary results highlight patterns of best-performing students with specific factors/features located in the highest clusters on the rank. This insight facilitates data-driven decisions leading to effective student interventions.

Keywords: Representation, Students' Controllable Features, Performance

Introduction & background

A clear picture representing the extent of a specific issue gives us clues to find the best interventions to solve it. Accordingly, in our research dealing with student performance, remains a lack of research that represents or gives a big picture and measures factors under students' control which impact their performance.

However, there are works in education that measure teaching and learning performance, such as "*Measuring teaching and learning performance in higher education*" (Muda, H., et al., 2017) and "*Effect of learning management system on Student's performance in educational measurement and evaluation*" (Oguguo, B.C.E., et al., 2021). This study determined the effect of a Learning Management System (LMS) on students' performance in educational measurement and evaluation courses. Furthermore, in 2018, Hattie's Visible Learning research synthesised findings from 1500 meta-analyses of 90,000 studies (Hattie, J., 2018). Then, Rossi and Montgomery's model focuses mainly on societal student's context, which points to two distinct scenarios. Firstly, the community environment and home quality, secondly the quality of the school, such as the classroom conditions, curriculum and student's incentives (Akama, E., 2017). A research group led by Dunlosky from Kent State University in 2013 presented ten years of literature indicating the possible enhancement of student accomplishment in different conditions (Ericsson, A. & Pool, R., 2016). Lastly, the "Chemer, Hu, and Garcia's model" is a longitudinal study developed by Martin M. Chemers, Li-tze Hu, and Ben F. Garcia at the University of California. They inspected the effects of optimism and academic self-efficacy on students' achievement, commitment to continuing in school, health and stress (Chemers, M., et al., 2001).

The representation of the student's controllable performance features was based on the PS2CLH model (Almada, A., et al., 2019). The PS2CLH model was inspired by the child development and early learning field (Landry, S. H., 2014). This field develops children's critical skills through interactive play in a safe and engaging environment. However, what are students' controllable factors that affect their results: It may be seen as students' lifestyle, habits, and daily life issues/problems/concerns, which are under students' control and influences their academic performances. In the same way, the PS2CLH model was developed to bring university students the necessary awareness of the controllable issues that affect their performance in ways they can act upon. In addition, looking to help university managers build a clear landscape of students' factors, we propose a new representation of the students' controllable performance features.

In our case, the controllable performance features are extracted from the PS2CLH model, which contemplates the perspectives of Psychology, Self-responsibility, Sociology, Communication, Learning and Health & wellbeing. For each perspective, we have a range of features. For instance, from Psychology's perspective, features that affect students' performance include stress, anxiety, fear and loneliness. Those features mentioned before were factors used in the author's previous paper (Almada, A., et al., 2019) applied to predict students' performance and cluster into groups of students with similar factors.

Referencing Bhargavi and Gowda, *Clustering aims to group data into coherent groups based on the nearness of samples in multiple feature space where the coherency enriches the uniqueness of the clusters with respect to others* (Bhargavi, M.S., & Gowda, S.D, 2018). In this study, we applied the K-means cluster algorithm to group students with a similar number of factors. Thus to Xin Jin and Jiawei Han, the procedure of k-means clustering is the following: *Given an initial but not optimal clustering, relocate each point to its new nearest center, update the clustering centers by calculating the mean of the member points, and repeat the relocating-and-updating process until convergence criteria (such as predefined number of iterations, difference on the value of the distortion function) are satisfied* (Jin, X., & Han, J., 2011). Using clusters of the students' controllable performance features aims to give university decision-makers more data to take better decisions concerning their students.

Looking at data-driven decision-making, according to a survey of more than 1,000 senior executives conducted by PwC, highly data-driven organizations are three times more likely to report significant decision-making improvements compared to those relying less on data. Data-driven decision-making (sometimes abbreviated as DDDM) uses data to inform your decision-making process and validate a course of action before committing to it (Stobierski, T., 2021)(Barbu, S.J., et al., 2022)(Namvar, M., & Intezari, A. 2021).

Representation of the Student's Controllable Performance Features

In recent years, there has been a growth in the number of studies measuring and representing students' learning and performance. However, there is a lack of research on defining, measuring and monitoring controllable factors affecting students' performance. From an assistant or mentor's point of view, measuring, visually representing, and keeping track of the student's performance alongside factors that affect their academic achievement is essential. The problem identified by this research is a lack of research that represents and measures factors within students' control that impact their academic success. For university managers, subject tutors and academic mentors, it is essential to measure, visually represent, analyse and monitor student performance alongside factors affecting their academic achievement to enhance the student experience.

In this study, we collected data from students' controllable factors which affect their performance using a web multi-choice self-evaluate questionnaire. The population sample target was around 500 students from different courses and ages between 20 to 25 years old. With considerable knowledge about Universidade Católica de Angola students, we prepared the place to collect data. After the collection process, the clean data process was done, and we had around 432 students. This study aims to represent the students' controllable features in a student's three-dimensional space features representation using the PS2CLH model.

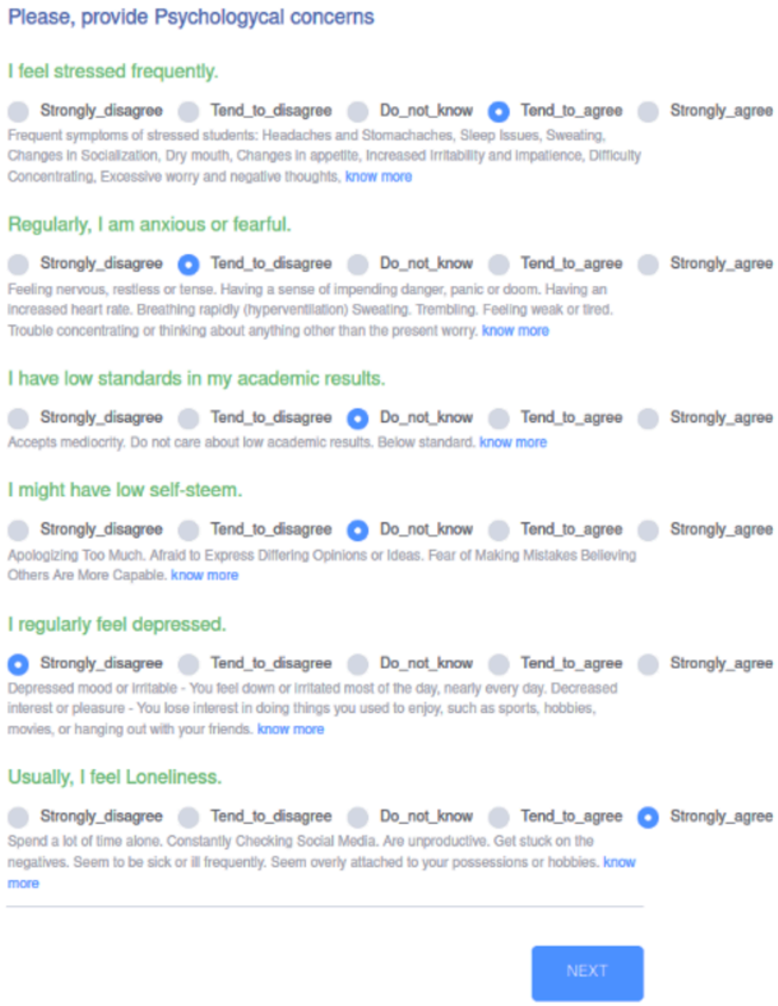
It is essential to point out that we use 66 features from the PS2CLH model in this experiment, which means that the standard representation of the features should be 66 dimensions. However, two essential concepts in data science may confuse us with what we will present. First, feature extraction is a method of reducing the dimensionality by which an initial set of raw data is reduced to more manageable groups for processing (Subasi. A., 2019) (Umamaheswari C., et al., 2018) (Guyon, I., & Elisseeff, A., 2006). The other concept is feature selection. Feature selection is the process of selecting specific variables to increase efficiency in choosing the most relevant features to apply in model construction (Ramesh, A., et al., 2022)(De Silva, A.M., 2015)(Das, T., et al., 2021). At this point, we introduce a new way to represent all these controllable features' performance, in which we use all the features or factors to build the coordinates. Please look at our previous paper to learn about those 66 factors or features (Almada, A., et al., 2019).

Students' controllable features performance 3D representation

It would be a natural question at this point to choose the type of visualization of student clusters. Why visually represent students? And why 3D, not 2D, 4D and 5D? Answering the first question, we can argue that our goal is to help students throughout their studies, and we need to know how much impact the student assistance has been. Therefore, we measure the student's evolution at the level of clusters. It is necessary to have an initial reference of the student's status and thus create a history of their trajectory in their academic life.

The second question concerns the most efficient way to represent the students in the cluster, so we think the 3D representation is ideal for our data. If we meant 2D, we would lose information given the number of areas we have. That is, we would have to group 3 areas to coordinate. In addition, considering the number of students at the university, there would be an overlap in the 2D representation, which is ineffective. Finally, the representation in 4D and 5D would be too complex to read the data. Therefore, we conclude that 3D representation is ideal for our data and the PS2CLH model.

Figure 1 below presents a part of the questionnaire used to collect the students' data. The question types are multiple-choice and a 5-point Likert scale response; we used the radio button for student responses. The questionnaire was constructed in ways that each question has a value depending on the student's answer.


Please, provide Psychologycal concerns

I feel stressed frequently.

Strongly_disagree Tend_to_disagree Do_not_know Tend_to_agree Strongly_agree

Frequent symptoms of stressed students: Headaches and Stomachaches, Sleep Issues, Sweating, Changes in Socialization, Dry mouth, Changes in appetite, Increased Irritability and Impatience, Difficulty Concentrating, Excessive worry and negative thoughts. know more

Regularly, I am anxious or fearful.

Strongly_disagree Tend_to_disagree Do_not_know Tend_to_agree Strongly_agree

Feeling nervous, restless or tense. Having a sense of impending danger, panic or doom. Having an increased heart rate. Breathing rapidly (hyperventilation) Sweating. Trembling. Feeling weak or tired. Trouble concentrating or thinking about anything other than the present worry. know more

I have low standards in my academic results.

Strongly_disagree Tend_to_disagree Do_not_know Tend_to_agree Strongly_agree

Accepts mediocrity. Do not care about low academic results. Below standard. know more

I might have low self-steem.

Strongly_disagree Tend_to_disagree Do_not_know Tend_to_agree Strongly_agree

Apologizing Too Much. Afraid to Express Differing Opinions or Ideas. Fear of Making Mistakes Believing Others Are More Capable. know more

I regularly feel depressed.

Strongly_disagree Tend_to_disagree Do_not_know Tend_to_agree Strongly_agree

Depressed mood or irritable - You feel down or irritated most of the day, nearly every day. Decreased interest or pleasure - You lose interest in doing things you used to enjoy, such as sports, hobbies, movies, or hanging out with your friends. know more

Usually, I feel Loneliness.

Strongly_disagree Tend_to_disagree Do_not_know Tend_to_agree Strongly_agree

Spend a lot of time alone. Constantly Checking Social Media. Are unproductive. Get stuck on the negatives. Seem to be sick or ill frequently. Seem overly attached to your possessions or hobbies. know more

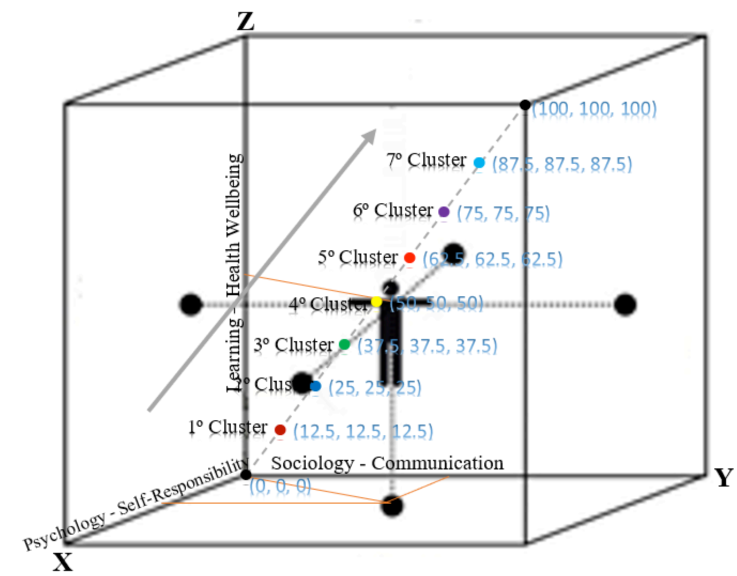
NEXT

Figure 1: A part of the questionnaire used to collect the students' data.

This research proposes a visual representation and measure of a student-controllable learning factor that affects their performance, based on the academic model that combines psychology, Self-responsibility, Sociology, Communication, Learning and Health & wellbeing (PS2CLH). We associate psychology & self-responsibility (coordinate/axes X), social & communication (axes Y), learning and health & wellbeing (axes Z). It results in student representation of a point in three-dimensional space 3D. Consequently, it will be possible to represent students into different clusters, effectively monitor their issues, and understand the PS2CLH's model patterns, leading to a better understanding of academic performance.

	Strongly Disagree	Tend to Disagree	Do not know	Tend to Agree	Strongly Agree
	○	○	○	○	○
Value:	5	4	3	2	1

Apply rationalisation for the pair (Psychology & Self-responsibility) (Sociology & Communication) (Learning & Health and Wellbeing) Then calculate the students' clustering accordingly with their coordinate representation.



Distance from students' point to the cluster centroid.
$$\mathbf{D = \sqrt{(x_2-x_1)^2 + (y_2-y_1)^2 + (z_2-z_1)^2}}$$

Mid-Point.
$$\mathbf{M = ((x_2+x_1)/2, (y_2+y_1)/2, (z_2+z_1)/2)}$$

Figure 2: PS2CLH Visual 3D representation.

The figure above introduces the students' 3D representation. When Students fill in the questionnaire, there is a value for each answer. Their Sum makes the coordinate, and when they finish filling in, it automatically calculates their clustering. In the questionnaire, each question has a weight. According to the answer, this weight is attached to that question in ways that each area weight will be the sum weight of questions, having six areas. In Psychology, Self-Responsibility, Sociology, Communication, Learning and Health-wellbeing, we put a pair of two areas representing a coordinate: PS coordinate X, SC coordinate Y and LH coordinate Z.

To monitor students' evolution or growth, we need to represent and observe the initial state and then monitor their evolution through clusters. Expressing in a 3D students' factors that affect their performance allows the system to know the distance among each student, leading us to build clusters of students—clustering students into groups according to the student controllable learner model and students' questionnaire.

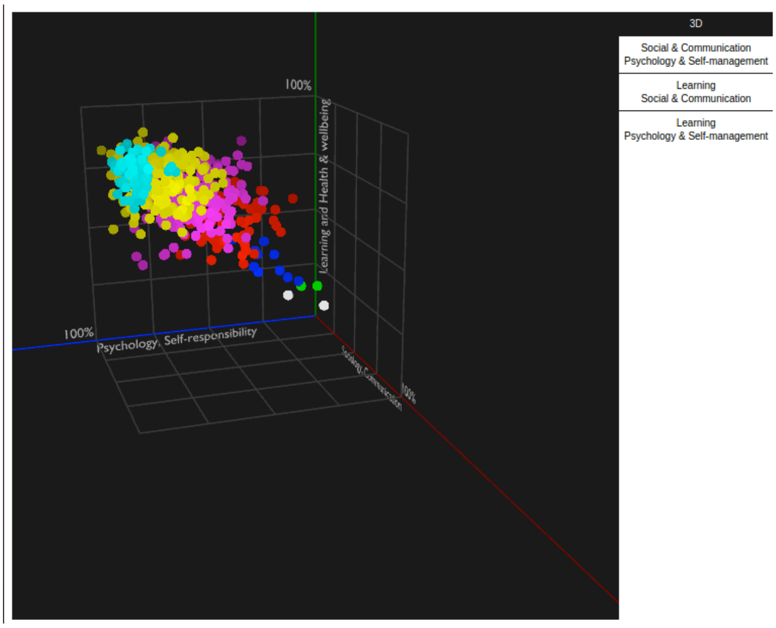


Figure 3: Students represented into PS2CLH Visual 3D representation.

In figure 3, there are clusters, and different colour represents one cluster; each point represents one student. Meaning students using the PS2CLH model in a 3D visualisation based on these students' controllable factors has the following implications. With the representation of students, it will be possible to visualise the students' clusters, thus showing the groups needing more help. This view will also show the pattern and set of the best students, which can guide university decision-makers to act proactively. In addition, the visual image of the student's academic factors that affect students' performance allows lecturers to have a visual idea of their student's academic controllable factors, which affects their performance.

Conclusion & Future works

Initial data presents a clear pattern of creating a diagonal of seven clusters or students' stages from the bottom (0, 0, 0) to the top (100, 100, 100) and leading to the use of filters or queries to represent better features such as sleep-problem, stress, practice exercises and time management. In addition, preliminary results highlight patterns of best-performing students with specific factors/features located in the highest clusters on the rank. This insight facilitates data-driven decisions, creating an intelligent tool for university managers and giving a clear picture of the students' controllable factors, allowing decision-makers to take proactive action leading to effective student interventions.

The proposed future work is: With a clear representation of different clusters, the proactive chatbot (Almada, A., et al., 2022) can have other behaviours related to a particular group of students. This approach gives the flexibility to pay attention to students who need the most and clearly understand where each student stands and the direction. Students must take the necessary steps and phases to reach the desired cluster. Results show that the best students are in clusters six and seven, with fewer problems or factors that affect their performances.

Therefore, the goal is to work with the individual student during their academic year to tackle their problems leading them to move to clusters six and seven.

References

Akama, E., 2017. Student's Performance Prediction Using FP-Tree Data Mining Techniques. International Journal of Science and Research (IJSR).

Almada, A., Yu, Q. & Patel, P., 2019. PS2CLH: A Learning Factor Model for Enhancing Students' Ability to Control Their Achievement. Tokyo, ACE2019.

Almada, A., Yu, Q., Patel, P., 2023. Proactive Chatbot Framework Based on the PS2CLH Model: An AI-Deep Learning Chatbot Assistant for Students. In: Arai, K. (eds) Intelligent Systems and Applications. IntelliSys 2022. Lecture Notes in Networks and Systems, vol 542. Springer, Cham. https://doi.org/10.1007/978-3-031-16072-1_54

Barbu, S.J., McDonald, K., Brazil-Cruz, L., Sullivan, L., Bisson, L.F. (2022). Data-Driven Decision-Making. In: Bisson, L.F., Grindstaff, L., Brazil-Cruz, L., Barbu, S.J. (eds) Uprooting Bias in the Academy. Springer, Cham. https://doi.org/10.1007/978-3-030-85668-7_3

Bhargavi, M.S., Gowda, S.D., 2018. Cluster Representation and Discrimination Based on Regression Line. In: Nagabhushan, T., Aradhya, V.N.M., Jagadeesh, P., Shukla, S., M.L., C. (eds) Cognitive Computing and Information Processing. CCIP 2017. Communications in Computer and Information Science, vol 801. Springer, Singapore. https://doi.org/10.1007/978-981-10-9059-2_20

Chemers, M., Hu, L.-t. & Garcia, B., 2001. Academic Self-Efficacy and First-Year College Student Performance and Adjustment. Journal of Educational Psychology, Vol. 93(No. 1), pp. 55-64.

Das, T., Paitnaik, S., Mishra, S.P. 2021. Feature Selection on Linked Data: A Review. In: Gao, XZ., Kumar, R., Srivastava, S., Soni, B.P. (eds) Applications of Artificial Intelligence in Engineering. Algorithms for Intelligent Systems. Springer, Singapore. https://doi.org/10.1007/978-981-33-4604-8_47

De Silva, A.M., Leong, P.H.W. 2015. Feature Selection. In: Grammar-Based Feature Generation for Time-Series Prediction. SpringerBriefs in Applied Sciences and Technology. Springer, Singapore. https://doi.org/10.1007/978-981-287-411-5_2

Ericsson, A. & Pool, R., 2016. Peak: Secrets from the New Science of Expertise. ISBN 978-0-5-45623-5 ed. New York: New York 10016.

Guyon, I., Elisseeff, A. 2006. An Introduction to Feature Extraction. In: Guyon, I., Nikravesh, M., Gunn, S., Zadeh, L.A. (eds) Feature Extraction. Studies in Fuzziness and Soft Computing, vol 207. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-35488-8_1

Hattie, J., 2018. "Visible Learning plus," 250+ Influences on Student Achievement. [Online] Available at: https://www.visiblelearningplus .com/content/research-john-hattie, 250_influences_10.1.2018.pdf [Accessed 7 December 2018].

Jin, X., Han, J., 2011. K-Means Clustering. In: Sammut, C., Webb, G.I. (eds) Encyclopedia of Machine Learning. Springer, Boston, MA. https://doi.org/10.1007/978-0-387-30164-8_425

Landry, S. H., 2014. The role of parents in early childhood learning, Texas: Children's Learning Institute, University of Texas Health Science Center.

Muda, H., Ali, M.H., and Jusoh, M., 2017. Measuring Teaching andLearning Performance in Higher Education. International Journal of Education, Psychologyand Counselling, 2 (6), 57-70.

Namvar, M., Intezari, A. (2021). Wise Data-Driven Decision-Making. In: Dennehy, D., Griva, A., Pouloudi, N., Dwivedi, Y.K., Pappas, I., Mäntymäki, M. (eds) Responsible AI and Analytics for an Ethical and Inclusive Digitized Society. I3E 2021. Lecture Notes in Computer Science, vol 12896. Springer, Cham. https://doi.org/10.1007/978-3-030-85447 8_10

Oguguo, B.C.E., Nannim, F.A., Agah, J.J. et al., 2021. Effect of learning management system on Student's performance in educational measurement and evaluation. Educ Inf Technol 26, 1471–1483. https://doi.org/10.1007/s10639-020-10318-w

Ramesh, A., Reddy, K.P., Sreenivas, M., Upendar, P. 2022. Feature Selection Technique-Based Approach for Suggestion Mining. In: Bhateja, V., Tang, J., Satapathy, S.C., Peer, P., Das, R. (eds) Evolution in Computational Intelligence. Smart Innovation, Systems and Technologies, vol 267. Springer, Singapore. https://doi.org/10.1007/978-981-16-6616-2_53

Stobierski, T., 2021. The advantages of data-driven decision-making. Harvard Business School Online. https://online.hbs.edu/blog/post/data-driven-decision-making. Assessed 04, September of 202.

Subasi. A., 2019. Feature Extraction and Dimension Reduction. Practical Guide for Biomedical Signals Analysis Using Machine Learning Techniques. ScienceDirect. https://doi.org/10.1016/B978-0-12-817444-9.00004-0 pp 193-275

Umamaheswari C., Dr. Bhavani R. and Dr. Thirunadana Sikamani K. 2018. Texture and Color Feature Extraction from Ceramic Tiles for Various Flaws Detection Classification International Journal on Future Revolution in Computer Science & Communication Engineering 4 169-179.

Contact email: ARA0223@my.londonmet.ac.uk