Journal of Positive School Psychology http://journalppw.com
2022, Vol. 6, No. 7, 3588-3597

From Face-To-Face To Virtual Education. An Evolutionary Analysis Of The Academic Performance Of First-Level Students Of The Universidad Politécnica Estatal Del Carchi Through Neural Networks

Marco Rubén Burbano Pulles[1] , Eduardo Javier Pozo Burgos[2] , Gabriela Elizabeth Revelo Salgado[3] , Jack Iván Vidal Chica[4] , Erick Patricio Herrera Granda[5]

[1]*https://orcid.org/0000-0002-2036-7032 Universidad Politécnica Estatal del Carchi, marco.burbano@upec.edu.ec*
[2]*https://orcid.org/0000-0002-3240-1648 , Universidad Politécnica Estatal del Carchi, eduardo.pozo@upec.edu.ec*
[3]*https://orcid.org/0000-0001-5238-223X, Universidad Politécnica Estatal del Carchi, gabriela.revelo@upec.edu.ec*
[4]*https://orcid.org/0000-0003-2060-0946, Universidad Politécnica Estatal del Carchi , vicerrectorado@tecnologicosucre.edu.ec*
[5]*https://orcid.org/0000-0002-3490-7917 , Universidad Politécnica Estatal del Carchi , erick.herrera@upec.edu.ec*

ABSTRACT

Prior to the confinement caused by covid19, the face-to-face modality covered a series of aspects that were related to the academic performance of students and that, because of the migration to virtual education, gave way to other variables that now constitute the new causes of their success or failure. This research develops a longitudinal analysis of the evolution of the academic performance of students in the first levels of the Universidad Politécnica Estatal del Carchi, through a comparative statistical methodology based on neural networks that takes as moments of study two academic periods before and two after the pandemic. The research yielded a predictive model that offers a 78% of correct classification in relation to the main variables that can influence academic performance, being results that coincide to a great extent with the research taken as references.

KEYWORDS: Higher education, mathematical models, academic performance, learning environment, Pandemics.

1. INTRODUCTION

The new order established worldwide by the pandemic caused by covid 19 has generated a series of challenges for the different spheres of the daily actions of humanity. With the spread of the virus around the world, governments imposed quarantines and bans of various kinds [1, 6, 17]; in this context, education does not escape from these vicissitudes and has had to adapt to the new needs due to the new measures that seek to preserve the integrity of the actors in the teaching-learning process. The above has led education to apparent virtualization within a learning modality. Despite claims that e-learning can improve the quality of education, Arkorful, Valentina and Abaidoo [3] argue that making learning materials available online improves learning only for specific forms of collective.

The development of multimedia and information technologies, as well as the use internet as a new technique of teaching, has made radical changes in the traditional process of teaching. Development in information technology has generated more choices for today's education. E-learning, has come to be more and more important in institutions of higher education. The introduction and expansion of a range of e-Learning tools has been initiating several changes in higher education institutions, particularly when it comes to their educational delivery and support processes [12,20].

However, the lack of preparation and experience to work in virtual environments, which was implemented almost immediately in the face of recent events, implied certain risks that were not taken into account, such as socio-educational variables and the digital divide. Since virtuality implies the use of technological resources and connectivity in a mandatory manner [19], now becoming the new challenges that, in one way or another, affect students in academic performance.

One of the indicators of the level of student learning is academic performance, and it is associated with the effectiveness in achieving the learning outcomes contemplated within the curricular objectives of the different subjects. However, a diversity of causes originate from low academic performance, generating different types of research associated with the variability of the educational contexts in which they are carried out [15]. nowadays, these contexts are directly affected by the new reality facing the world and in which virtuality marks the roadmap of higher education.

In educational research, the implementation of classification techniques is beneficial for predicting academic performance concerning the student's profile; in this sense, procedures have been implemented with various classification algorithms. One of the techniques that are becoming popular is Neural Networks. However, it is not a new concept since Warren McCulloh, and Walter Pitts developed the first artificial neuron in 1943, almost eighty years ago [5].

Mushtaq and Singh [16] conducted a study in India with students between 15 and 22 years old to generate a model to predict college performance, using five algorithms: Naïve Bayes, Logistic Regression, Multilayer Perceptron, J48, and Random Tree. The model with the highest accuracy was the neural network model with an F-value of 94.5%.

In a study with students of professional technological education, with samples of 922 students from Turkey and 1050 students from Malaysia, students were administered a 34-item demographic questionnaire. As a result, a prediction model of academic performance was obtained using Neural Networks with a sensitivity of 98.0% in the Turkish students and 95.7% in the Malaysian sample [21]. This study suggests that, in the subject of educational research, the particularities of a region or country generate research proposals that must be precisely analyzed according to the geographical location of the sample.

Educational data mining can also be applied to analyze specific subjects. For example, as proposed by Dabhade [8], who developed a model based on Neural Networks to analyze grades obtained in mathematics, The resulting data set had 112 variables. After detailed analysis, 92 variables have been chosen for study. Using factor analysis, the dimensionality has been reduced to 2. The data sets have been divided into training data set and test data set. The model has been trained. It is observed that the linear model provides the best fit with an accuracy of 83.44%. However, it does not work satisfactorily to identify the passing status when identifying the repetition of the course because an additional parameter is introduced, which is the additional or supplementary exam.

2. DATA

The data correspond to the grades of the first-semester students of the nine careers of the Universidad Politécnica Estatal del Carchi. The same was obtained from the academic records in the academic periods of (1) April 2019 - August 2019, (2) October 2019 - February 2020, (3) June 2020 - September 2020, and (4) November 2020 - March 2021, of which, the first two were developed in classroom modality and the last two in virtual modality.

Table 1 summarizes the primary measures of central tendency and the final average, which corresponds to the weighted average of the grades obtained by the students in the different subjects of each course.

The grades and the final average are graded out of 10.

Table 1. Descriptive summary of data

Statistic	Semester			
	1	**2**	**3**	**4**
Mean	7,35	7,58	8,04	8,22
Median	7,40	7,73	8,00	8,44
Standard deviation	,7684	,8156	,7919	,9137
Range	4,90	5,19	5,08	5,33
Minimum	4,67	4,11	4,55	4,00
Maximum	9,57	9,30	9,63	9,33

3. METHODOLOGY

3.1 Measured Variables

Based on the fact that academic performance is a multifactorial or multicausal phenomenon, different variables can be found that attempt to explain it, whether internal or external factors are associated with the student. These variables can be grouped into two groups: those related to the non-academic characteristics of the students, such as gender, economic level of the family group, socioeconomic and sociocultural environment, educational level of the father and mother, etc.; and those related to their academic characteristics, such as their grades in the admission tests or their school performance [14].

The admission processes that involve applying specific tests in higher education institutions have as primary purposes to assign scores to applicants that provide estimated information of those students with greater possibilities of successfully facing their professional training. Likewise, they constitute a diagnosis of the academic aptitudes of the new students [2].

Attendance to regular classes in higher education is conceived as the recording of the daily attendance of students to their different classes. Considered an obligation for the face-to-face modality, which implies that the student is punctually and participating in the subjects for which he/she was enrolled during that academic period. Teachers are directly responsible for recording this parameter and carrying out the respective follow-up [10].

The present research used data analysis based on artificial neural networks (ANNs), which are algorithms that can solve complex problems by mimicking the human brain's functionality with certain limitations. The main application of neural networks is data classification, e.g., a value of 1 if the network output is more significant than zero or 0 if the output is less than or equal to zero.

$$\text{Clasification} = \begin{cases} 1 \text{ if } f(X) > 0 \\ 0 \text{ if } f(X) \leq 0 \end{cases} \quad (1)$$

The input variables of a neural network (x_i) are multiplied by a weight (w_i), the sum of all these values being the nerve impulse of the neuron (f(X)); this impulse is adjusted by a bias value (b), in order to obtain a more accurate classification in the output.

$$z = f(X) = b + \sum_{i=0}^{n} x_i w_i \quad (2)$$

ANNs are organized by levels or layers, the best known being the multilayer perceptron (PM); these networks have two defined functions: activation and output, which can be of various types, highlighting the use of the sigmoid function.

$$\sigma(z) = \frac{1}{1+e^{-z}} \quad (3)$$

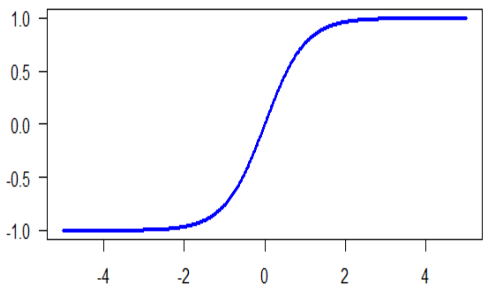

Figure 1 Sigmoid activation function

In a study conducted in 2011 on different activation functions [4], it was concluded that for most applications, the use of the Hyperbolic Tangent activation function allows obtaining better results, even using it as activation or in the output. However, most of the time, the data analysis processes depend on the type of information being handled; that is why in the present research, the Hyperbolic Tangent was used as the activation variable, and Softmax as the output function, since the best prediction results were obtained with this combination.

$$\tanh(x) = \frac{(1-e^{-x})}{(1+e^{-x})} \quad (4)$$

$$\sigma(X_i) = \frac{e^{x_i}}{\sum_{j=0}^{k} e^{x_j}} \quad (5)$$

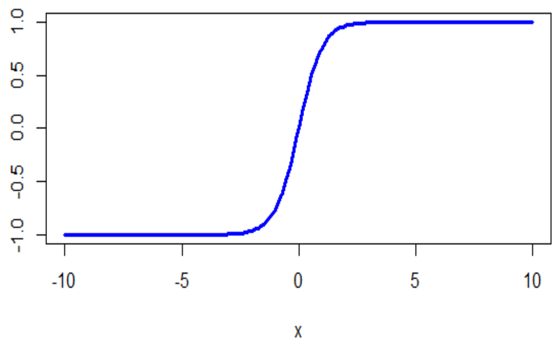

Figure 2 Hyperbolic Tangent Tangent Activation Function

The PM used in this analysis has a hidden layer and five input variables that feed the network to predict students' academic performance.

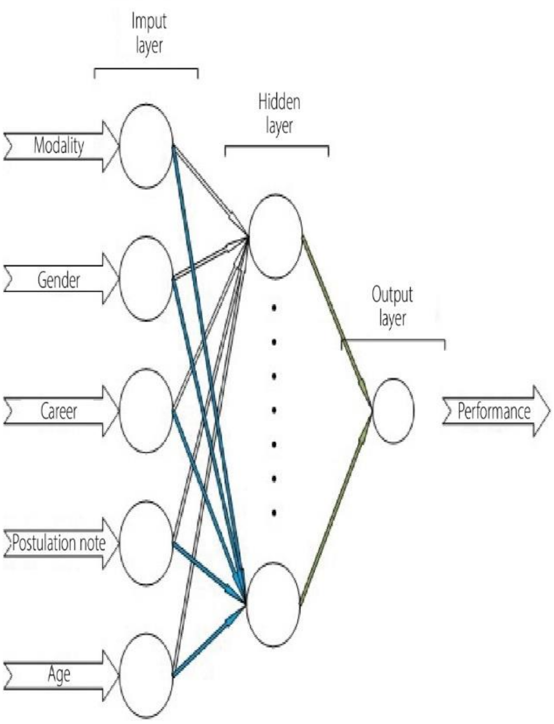

Figure 3 Multilayer perceptron used

Academic performance was measured by calculating the weighted average of the final average, based on the number of academic hours in each course.

The independent variables (factors) and covariates included in the model and the categories assumed are:

Table 2. Variable categorization

Variable	Type of variable	Categories assigned
Modality	Qualitative	1 = On-site 2 = Virtual
Gender	Qualitative	1 = Female 2 = Male
Career	Qualitative	For this specific case, the careers were coded as follows: 1 = Bachelor's degrees 2 = Engineering 3 = Health Sciences
Workday	Qualitative	1 = Morning 2 = Evening
Age	Quantitative	Age in years
University entrance grade	Quantitative	The score obtained by the student in the university entrance exam graded out of 1000

The academic performance variable was classified as high in cases where the weighted mean is greater than or equal to 7 and low to those with a weighted mean less than 7.

Several tests were performed to choose the neural network, variables were eliminated, and different types of training were tested. Below are several results obtained, which were similar in the overall percentage correct, so the network with the lowest cross-entropy error was chosen.

Table 3. Results comparison

	Cross entropy error	Overall Percent Correct	
		Training	Test
Perceptron Multilayer 1	185,145	85,90%	87,10%
Perceptron Multilayer 2	159,45	88,80%	87,90%
Perceptron Multilayer 3	130,309	86,70%	89,30%
Perceptron Multilayer 4	119,08	86,90%	90,10%

The Perceptron Multilayer model, PM4, is summarized as follows:

Table 4. Application of the Perceptron model

Network information			
Input layer	Factors	1	CAREER
		2	GENDER
		3	MODALITY
	Covariants	1	AGE
		2	APPLICATION SCORE
	Number of units		16
	Method of scaling covariates		Typified
Hidden layers	Number of hidden layers		1
	Number of units of hidden layer 1		8
	Activation function		Hyperbolic Tangent
Output layer	Dependent variables		PERFORMANCE
	Number of units		2
	Activation function		Softmax
	Error function		Cross-Entropy

4. Results

1116 (72.9%) individuals were assigned to the training set and 414 (27.1%) to the test set.

For the PM model trained with the data set corresponding to all first-semester students, a total correct classification rate of 90.1% was obtained (Table 5).

Table 5. Classification

Sample		Forecast Low	Forecast High	Correct percentage
Training	Low	50	110	31,3%
	High	36	920	96,2%
	Overall percentage	7,7%	92,3%	86,9%
Test	Low	17	33	34,0%
	High	8	356	97,8%
	Overall percentage	6,0%	94,0%	90,1%

The percentage of incorrect predictions is similar in the training and test set, thus suggesting that there was no overtraining of the network and, therefore, the network's training was good.

The area under the COR curve (Figure 4) allows us to conclude that the model discriminates high achievers from low achievers with a probability of correct classification of 0.828.

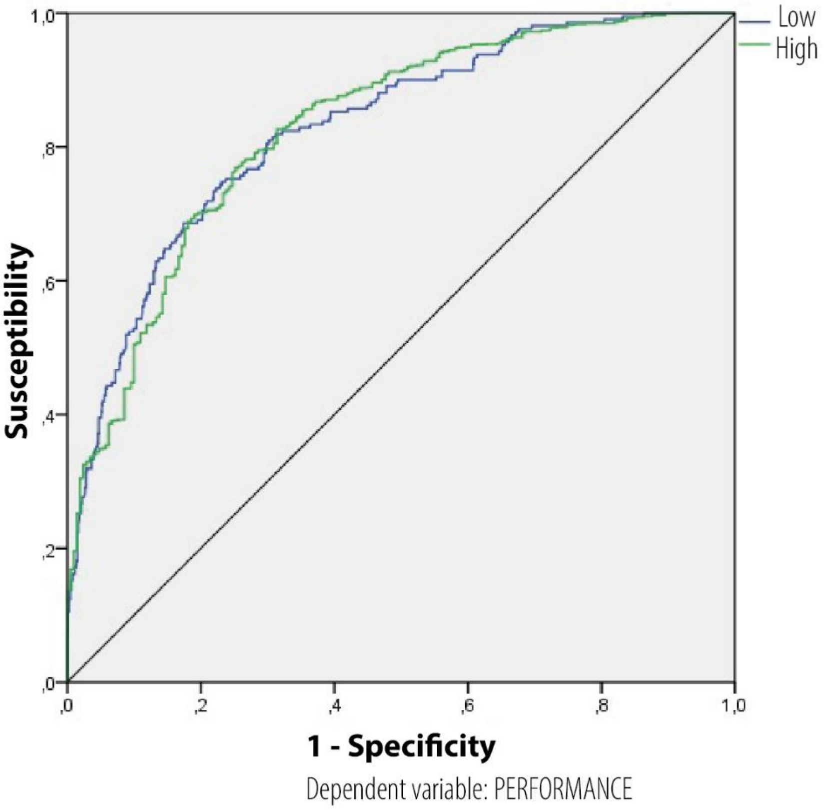

Figure 4- COR curve of the Neural Network.

Finally, the normalized importance of each variable used in the network is analyzed, that is, the change in the predictive value for different independent variable values. For example, the most important predictive variables for the network were Note of application and Career, followed by age. On the other hand, the study mode and gendre were the variables with the least predictive power.

5. Discussion

The analysis of the history of students' academic performance in the first levels of UPEC shows changes in the variables involved for the periods before and after March 2020. Previously, factors such as the working day had an impact on the results of the grades achieved by them, coinciding with research such as that conducted by Roeser, Schlarb, and Kübler [18], in which it is established that adolescents experience a discordance between their biological time, which is oriented towards the eve of school, and the school start time. Given this social mismatch, the evening schedule is negatively correlated with academic performance. Currently, virtuality offers new challenges, according to Karattuthodi et al. [13], who states that there are new challenges for the classical university, which were identified due to the sudden transition to virtual education, which generated tensions that evidenced various types of deficits, such as equity both in access and in the development of academic activities.

Several research studies have examined whether class attendance has a positive impact on student outcomes as measured by grades [10]. They examined this relationship using new data from students who have taken a class called Consumer Decision Making. The results suggest that class attendance has a positive and significant effect on standardized test scores. Thus, those who increase their attendance by 10 percentage points have an improvement of 0.08 standard deviations, or 1.14 points in actual test scores

Dayioğlu and Türüt-Aşik [9] establish the positive and significant effect between the grades of the entrance exams and those related to teaching at the intermediate level on the academic performance. Hanham, Lee, and Teo [11] found that women had a higher rating of perceived ease of use and favorable conditions for an evaluation based on the use of digital media, which could be an important factor to consider when migrating from the traditional to the virtual modality.

In general, it is possible to agree with Cordero-Ferrera et al [7], who point out that it is not possible to generalize a single methodology for measuring efficiency in educational processes, nor is there a standardized indicator for its assessment. The multidimensional nature of academic performance generates a dependence on multiple aspects such as those addressed in this research; likewise, different techniques and methodologies have been integrated seeking to generate a predictive model according to the investigated context.

6. Conclusions

- The neural network established offers 90.1% of correct classification concerning identifying the variables that directly influence the academic performance of the experimental units treated in this research.
- The variables that contributed from the statistical point of view to the established model were the mode of education, gendre, type of career, age, and university entrance grade.
- The most important predictor variables for the network were grade and career, followed by age. In contrast, the study mode and gendre were the variables with the least predictive power.
- A good selection of the activation functions in the hidden and output layers allows obtaining better results in the classification process. Unfortunately, due to the variable nature of the data, especially in the educational field, it is not possible to define a rule that designates the use of a specific function; it is the researcher's job to make the different modifications in order to approach an optimal result in the prediction.
- In the beginning, the change of study modality generated, both in students and

teachers, a resistance to the use of technology and made learning explosive, storing the information in short-term memory, causing emotional stress and even lack of concern and overconfidence due to the conditions presented worldwide.

References

1. Abumalloh, Rabab Ali, Shahla Asadi, Mehrbakhsh Nilashi, Behrouz Minaei-Bidgoli, Fatima Khan Nayer, Sarminah Samad, Saidatulakmal Mohd, and Othman Ibrahim. 2021. “The Impact of Coronavirus Pandemic (COVID-19) on Education: The Role of Virtual and Remote Laboratories in Education.” Technology in Society 67(September):101728. doi: 10.1016/j.techsoc.2021.101728.
2. Althewini, Abdulaziz, and Abdulmohsen Alkushi. 2020. “Predictive Validity of Saudi Admission Criteria for Freshmen Students’ English Performance: Experience of King Saud Bin Abdulaziz University for Health Sciences.” Journal of Language Teaching and Research 11 (1): 108–14. https://doi.org/10.17507/jltr.1101.12.
3. Arkorful, Valentina, and Nelly Abaidoo. 2014. “The Role of E-Learning, Advantages and Disadvantages of Its Adoption in Higher Education.” International Journal of Education and Research 12 (1): 29–42. www.ijern.com.
4. Bekir, Karlik, and Vehbi Ahmet. 2011. “Performance Analysis of Various Activation Functions in Generalized MLP Architectures of Neural Networks.” International Journal of Artificial Intelligence And Expert Systems (IJAE) 1(4):111–22. doi: 10.1088/1742-6596/1237/2/022030.
5. Benko, Attila, and Cecília Sik Lányi. 2009. “History of Artificial Intelligence.” Chapter. Encyclopedia of Information Science and Technology, Second Edition. IGI Global. 2009. https://doi.org/10.4018/978-1-60566-026-4.ch276.
6. Chen, Qiongni, Mining Liang, Yamin Li, Jincai Guo, Dongxue Fei, Ling Wang, Li He, Caihua Sheng, Yiwen Cai, Xiaojuan Li, Jianjian Wang, and Zhanzhou Zhang. 2020. “Mental Health Care for Medical Staff in China during the COVID-19 Outbreak.” The Lancet Psychiatry 7(4):e15–16. doi: 10.1016/S2215-0366(20)30078-X.
7. Cordero-Ferrera, Jose Manuel, Francisco Pedraja-Chaparro, and Javier Salinas-Jiménez. 2008. “Measuring Efficiency in Education: An Analysis of Different Approaches for Incorporating Non-Discretionary Inputs.” Applied Economics 40 (10): 1323–39. https://doi.org/10.1080/00036840600771346.
8. Dabhade, Pranav, Ravina Agarwal, K. P. Alameen, A. T. Fathima, R. Sridharan, and G. Gopakumar. 2021. “Educational Data Mining for Predicting Students’ Academic Performance Using Machine Learning Algorithms.” Materials Today: Proceedings 47: 5260–67. https://doi.org/10.1016/j.matpr.2021.05.646.
9. Dayioğlu, Meltem, and Serap Türüt-Aşik. 2007. “Gender Differences in Academic Performance in a Large Public University in Turkey.” Higher Education 53 (2): 255–77. https://doi.org/10.1007/s10734-005-2464-6.
10. Dey, Ishita. 2018. “Class Attendance and Academic Performance: A Subgroup Analysis.” International Review of Economics Education 28 (March): 29–40. https://doi.org/10.1016/j.iree.2018.03.003.
11. Hanham, José, Chwee Beng Lee, and Timothy Teo. 2021. “The Influence of Technology Acceptance, Academic Self-Efficacy, and Gender on Academic Achievement through Online Tutoring.” Computers and Education 172(July 2020).

doi: 10.1016/j.compedu.2021.104252.
12. Innocent, Willy A. 2020. “Applicability of E-Learning in Higher Learning Institutions in Tanzania.” International Journal of Education and Development Using Information and Communication Technology (IJEDICT) 16 (2): 242–49.
13. Karattuthodi, Mohammed Salim, Sarath Chandran C, Shabeer Ali Thorakkattil, Dilip Chandrasekhar, Naseef Punnoth Poonkuzhi, Mamdouh Mohammed Ahmed Ageeli, and Hafees Madathil. 2022. “Pharmacy Student’s Challenges in Virtual Learning System during the Second COVID 19 Wave in Southern India.” Social Sciences & Humanities Open 5(1):100241. doi: 10.1016/j.ssaho.2021.100241.
14. Lau, E. T., L. Sun, and Q. Yang. 2019. “Modelling, Prediction and Classification of Student Academic Performance Using Artificial Neural Networks.” SN Applied Sciences 1 (9): 1–10. https://doi.org/10.1007/s42452-019-0884-7.
15. Li, Ying, Tongtong Hu, Tingshuai Ge, and Emma Auden. 2019. “The Relationship between Home-Based Parental Involvement, Parental Educational Expectation and Academic Performance of Middle School Students in Mainland China: A Mediation Analysis of Cognitive Ability.” International Journal of Educational Research 97(January):139–53. doi: 10.1016/j.ijer.2019.08.003.
16. Mushtaq, Sheena, and Shailendra Narayan Singh. 2021. “Implementation of Classification Techniques in Education Sector for Prediction and Profiling of Students.” Emerging Technologies in Data Mining and Information Security 435–44. doi: 10.1007/978-981-33-4367-2_42.
17. Parmet, Wendy, and Michael Sinha. 2020. “Covid-19 — The Law and Limits of Quarantine.” New England Journal of Medicine 6(1):510–12.
18. Roeser, Karolin, Angelika Anita Schlarb, and Andrea Kübler. 2013. “The Chronotype-Academic Performance Model (CAM): Daytime Sleepiness and Learning Motivation Link Chronotype and School Performance in Adolescents.” Personality and Individual Differences 54(7):836–40. doi: 10.1016/j.paid.2012.12.021.
19. Simonson, Michael, Charles Schlosser, and Dan Hanson. 1999. “Theory and Distance Education: A New Discussion.” American Journal of Distance Education 13 (1): 60–75. https://doi.org/10.1080/08923649909527014.
20. Warburton, Steven. 2009. “Second Life in Higher Education: Assessing the Potential for and the Barriers to Deploying Virtual Worlds in Learning and Teaching.” British Journal of Educational Technology 40 (3): 414–26. https://doi.org/10.1111/j.1467-8535.2009.00952.x.
21. Yağci, Ali, and Mustafa Çevik. 2019. “Prediction of Academic Achievements of Vocational and Technical High School (VTS) Students in Science Courses through Artificial Neural Networks (Comparison of Turkey and Malaysia).” Education and Information Technologies 24(5):2741–61. doi: 10.1007/s10639-019-09885-4.