Education and Information Technologies
https://doi.org/10.1007/s10639-023-12132-6

Factors influencing the use of information and communication technologies by students for educational purposes

João C. Silva[1] · José Coelho Rodrigues[2] · Vera L. Miguéis[2]

Received: 30 March 2023 / Accepted: 10 August 2023

Abstract

Implementation of information and communication technologies (ICTs) in education is defined as the incorporation of ICTs into teaching and learning activities, both inside and outside the classroom. Despite widely studied, there is still no consensus on how it affects student performance. However, before evaluating this, it is crucial to identify which factors impact students' use of ICT for educational purposes. This understanding can help educational institutions to effectively implement ICT, potentially improving student results. Thus, adapting the conceptual framework proposed by Biagi and Loi (2013) and using the 2018 database of the Program for International Student Assessment (PISA) and a decision tree classification model developed based on CRISP-DM framework, we aim to determine which socio-demographic factors influence students' use of ICT for educational purposes. First, we categorized students according to their use of ICT for educational purposes in two situations: during lessons and outside lessons. Then, we developed a decision tree model to distinguish these categories and find patterns in each group. The model was able to accurately distinguish different levels of ICT adoption and demonstrate that ICT use for entertainment and ICT access at school and at home are among the most influential variables to predict ICT use for educational purposes. Moreover, the model showed that variables related to teaching best practices of Internet utilization at school are not significant predictors of such use. Some results were found to be country-specific, leading to the recommendation that each country adapts the measures to improve ICT use according to its context.

Keywords ICT · Education · Performance · Use of information and communication technologies · PISA · Decision tree model

Extended author information available on the last page of the article

Published online: 07 September 2023

1 Introduction

In a world in constant evolution, the needs of the labor market are also frequently changing (Fareri et al., 2020). Technological innovations lead to the creation of jobs with new requirements. With these innovations, the established tasks evolve, and the skills associated with them change (Fareri et al., 2020). It is then imperative that education is adapted to better prepare each citizen for these new market requirements (Skryabin et al., 2015). Indeed, education is seen as a factor contributing to the growth and development of economies (Hanushek & Woessmann, 2008). It is through education that citizens acquire the necessary skills to enter the labor market and become creators of value for society (Srijamdee & Pholphirul, 2020). As such, education must keep up with the digitalization process of the most diverse sectors of the economy.

The relevance of digitization became even clearer in 2020 with the spread of the COVID-19 pandemic. Educational institutions were some of the most affected by the pandemic. The need to move teaching to a remote setting brought several obstacles and caught many education institutions unprepared, as they were not yet able or accustomed to working with access to information and communication technologies (Gillis & Krull, 2020). Society in general was not prepared for such a sudden change, as were neither teachers nor students (Shim & Lee, 2020).

Aligned with this digitalization phenomenon, many tasks are currently performed with the aid of information and communication technologies (Srijamdee & Pholphirul, 2020). Therefore, it is worth understanding which teaching method best provides the skills required to work with information and communication technologies. The answer seems to lie in education using information and communication technologies, contrary to what has been the norm until now (traditional classrooms) (Fullan & Langworthy, 2013). This means not only learning how to use information and communication technologies well but also using these technologies for learning purposes in general. An example of this trend can be found in the Portuguese education system, where the use of information and communication technologies is seen as a necessity for education purposes (Cnedu, 2019).

To the best of our knowledge, no study has yet determined the factors that promote students' use of information and communication technology (ICT) for educational purposes, even though several studies specify the use of information and communication technologies for educational purposes as a good predictor of student performance (Bielefeldt, 2005; Kubiatko & Vlckova, 2010; Skryabin et al., 2015; Srijamdee & Pholphirul, 2020). Therefore, it is very relevant to find out what factors influence this type of use so that better incentive measures can be designed to promote it. Regarding student performance, several studies have analyzed how it is impacted by information and communication technologies, however, no consensus has yet been reached.

Within this context, our aim in the present study is to determine which factors promote students' use of ICT for educational purposes so that ICT implementation in schools can be carried out with optimal effectiveness, easiness, and speed. Accordingly, we seek to answer two main questions:

How do socio-demographic factors impact students' use of information and communication technologies for educational purposes? To carry out this analysis, we used the 2018 database from the Program for International Student Assessment (PISA).

Which factors impact students' use of information and communication technologies for educational purposes the most?

The paper is organized in five sections. After this introduction to the research topic, we review the literature on the implementation of information and communication technologies in education. In Section 3, we describe the methodology used, based on the construction of a classification machine learning model, using the CRoss-Industry Standard Process for Data Mining framework. Section 4, presents and discusses the results obtained from testing the model with data from PISA 2018. Finally, in Section 6, we outline our conclusions and provide practical recommendations to assist in the adoption and implementation of ICT for educational purposes.

2 Literature review

The literature on the implementation of information and communication technologies in education is vast. However, there is no consensus regarding how the use of information and communication technologies impacts student performance. It is believed that the skills developed through the frequent use of ICT have a positive influence on student performance. The improvements confirmed at various student levels include increased autonomy related to school tasks (Clark & Lee, 2019), higher awareness of news about their curriculum (Chou & Block, 2018), increased creativity and critical thinking (Chou & Block, 2018), better insight into their performance and a higher level of independence and ownership (Clark & Lee, 2019), a more personalized learning experience (Chou & Block, 2018; Harper & Milman, 2016), the development of the skills needed to solve real problems (Chou & Block, 2018); easier interaction with peers and teachers, e.g., to ask questions (Clark & Lee, 2019), and greater enjoyment and motivation to learn (Clark & Lee, 2019; Harper & Milman, 2016).

Most studies show that the relationship between the use of information and communication technologies in education and student performance is mostly positive. For example, the use of ICT in education seems to improve student performance in various subjects such as science, reading, and mathematics (Areepattamannil & Santos, 2019; Bielefeldt, 2005; Ferraro, 2018; Kubiatko & Vlckova, 2010; Skryabin et al., 2015; Srijamdee & Pholphirul, 2020; Wei et al., 2020; Xiao & Hu, 2019a). However, it is important to note that this improvement appears to be higher in students who are average at mathematics than in those who are very good or very bad, i.e., very good students would still be good regardless of access to ICT and very bad students do not become good in a short period of time just because of access to ICT (Wei et al., 2020). The literature further indicates that student performance in mathematics is not related to having access to information and communication technologies at school or at home, but rather to the actual use of these technologies (Bielefeldt, 2005; Srijamdee & Pholphirul, 2020). On the other hand, the use of ICT for recreational purposes such as games and programming seem to negatively influence performance (Kubiatko & Vlckova, 2010; Srijamdee & Pholphirul, 2020). Students' perceptions of ICT, namely regarding perceived autonomy and perceived competence also appear to have a positive influence on their performance in science subjects, with the effect of autonomy being higher (Areepattamannil & Santos, 2019). Finally, Skryabin et al. (2015) found that the level of ICT use in a country is positively related to students' achievement in mathematics, sciences, and reading. The effects of individual usage of ICT proved to be good for 4th-grade students' performance, both at home and at school. For 8th-grade students, the use of ICT at home was positive regardless of whether it was for entertainment or educational purposes, while the use of ICT at school was negative.

On the other hand, some studies found that the impact of ICT use on students' performance is not so positive. Biagi and Loi (2013) found that gaming is the only activity that positively relates the intensity of ICT use with students' achievement. Students' perceived ICT competence does not seem to have a positive influence on student performance in science, which can be explained by students' overconfidence in their ability to master ICT and by cultural factors,[1] although autonomy and interest in ICT appear to have a positive influence (Li et al., 2020; Meng et al., 2019). Petko et al. (2016) argued that ICT use at school and ICT use for entertainment have a negative relationship with student performance, even though ICT use at home is positively associated with student performance. Despite believing that ICT use is good for education, Petko et al. (2016) pointed out that this type of analysis should focus on the quality of ICT use rather than just its intensity and concluded that there is a positive relationship between a positive attitude toward ICT and performance. Lastly, Zhai et al. (2019) concluded that the relationship between ICT-related activities and students' achievements in physics depends on the intensity of ICT use in those activities.

Corroborating the above-mentioned results, the meta-analysis of Odell et al. (2020) concluded that the impact of ICT use on student performance is not consensual, with results existing in all directions.[2] However, these authors mentioned some trends in the results, such as perceived autonomy in ICT being a good influence for students. Competence and interest also appear to be mostly positive for students, although this is not always the case.

Regarding the factors that promote the use of ICT by students, the literature is scarce. The existing studies mostly take into consideration some socio-demographic factors to explain the use of technology in education. However, few other factors are considered, and these socio-demographic factors are only used as controls. Examples of these factors are gender, family structure, nationality, and repetition of school level (Coovadia & Ackermann, 2020; Petko et al., 2016; Skryabin et al., 2015; Xiao & Hu, 2019a, 2019b).

[1] Students from European countries seem to be overconfident about their ability to operate with ICT (they self-assess as very good at ICT). On the other hand, students from Eastern countries, such as China, seem to be more honest and self-evaluate as lacking in terms of ICT skills, which again leads to a negative relationship, since the best students have a higher tendency to self-evaluate.

[2] This conclusion derives from a meta-analysis that considers each variable and each subject in isolation. However, when evaluating the studies from a broader perspective, they point to mostly positive results.

A small number of studies have explored how economic, social, demographic, and technological factors (namely regarding the use of technologies) influence ICT adoption, approaching the topic from different perspectives. The most studied issue relates to the identification of the factors that influence the adoption and use of ICT by teachers in classrooms (Al-Mamary, 2020; Basak, 2014; Juggernath & Govender, 2020; Mirzajani et al., 2016; Prieto et al., 2014; Salinas et al., 2016; Spiteri & Chang Rundgren, 2018; Wang & Han, 2018). Basak (2014) concluded that some of the barriers to ICT use by teachers are a lack of competence, lack of confidence, and lack of time. In terms of the factors enabling ICT use by teachers, Basak (2014) gave the examples of teacher familiarity with ICT, availability of ICT, reliability in ICT, and classes being more enjoyable for students, while Salinas et al. (2016) pointed out ICT training and the perceived contribution of ICT to student learning. Juggernath and Govender (2020) added that most teachers believe that ICT can bring improvements to teaching, that a lack of ICT mastery can make it difficult to implement ICT in teaching, and that most teachers are willing to learn how to work with ICT.

Some other research topics have been less frequently explored. For example, Chen and Hu (2020) discovered a strong relationship between interest in ICT and ICT self-efficacy that was in part mediated by ICT use. Additionally, Xiao and Hu (2019b) found a relationship between socio-economic factors and students' performance in the form of reading test scores that was mediated by ICT use. Finally, Juhaňák et al. (2019) studied the relationship between the age of first contact with ICT and perceived ICT competence and autonomy and concluded that the earlier the first contact with ICT, the better the perceived autonomy and competence. Furthermore, the authors found that the moderating factors[3] were positively related to perceived ICT competence and autonomy (apart from the use of ICT outside of school for schoolwork).

In the present study, we used the most recent data (PISA 2018 database) and approached ICT use very granularly (at the student level) to identify the factors that influence students' use of ICT for educational purposes. We hope to contribute to a consensus regarding the use of ICT in education and introduce a novel approach to studying the factors that promote students' use of ICT. This aspect is just beginning to be addressed in the literature and several studies refer to it not only as an opportunity for future research but also as a contribution to improve future discussions surrounding the impact of ICT use on student performance (Buabeng-Andoh et al., 2018; Eickelmann et al., 2016; Li et al., 2020; Odell et al., 2020).

3 Methodology

We sought to construct a classification machine learning model, more specifically a decision tree model. We chose a decision tree model since it is easy to interpret, handles numerical and categorical variables well, and, unlike traditional statistical methods, is non-parametric and non-linear (Önder & Uyar, 2017).

[3] Student use of ICT outside of school for schoolwork, student use of ICT outside of school for leisure activities, student ICT interest and degree to which ICT is a part of students' daily social life.

We have adopted the CRoss-Industry Standard Process for Data Mining (CRISP-DM) framework to guide the present study from the definition of the problem to the evaluation of the results. This is one of the most popular frameworks to support a data mining project. The CRISP-DM framework was developed by an association of companies involved in data mining and has since been adopted by the research community working in the field.

CRISP-DM is used to provide solutions to problems that are solved using data mining and considers different stages for the data mining process, going through the understanding of the data, the preparation of the data, and the modeling. In the following sections, we describe the work developed within the scope of each CRISP-DM stage.

3.1 Business understanding

It is of the utmost importance to identify the factors that may influence students' use of ICT for educational purposes. This identification may help educational institutions to effectively implement ICT into their teaching and learning practices and to promote better results for their students.

Therefore, in the present study, we adopted and adapted the conceptual framework proposed by Biagi and Loi (2013). The main change to the original framework consisted of not considering "ICT use" but instead "ICT use for educational purposes", as it potentially impacts more directly the student's school performance. Other changes included the separation of Student-level factors into two categories to distinguish between factual variables and interpretative variables, allowing a more comprehensive understanding of the influence of each category on the dependent variables, and the creation of a dedicated group of variables for teachers, as they are potentially the most promising influencer regarding the use of ICT for educational purposes at school. This framework assumes that factors related to school, teachers, family, and students affect students' use of ICT for educational purposes, which, in turn, is expected to influence students' school performance.

In the present study, we focused on whether certain factors have an influence on the use of ICT for educational purposes, considering most of the factors proposed by Biagi and Loi (2013), as seen in Fig. 1. Each group of variables considered included both variables related to ICT and others that were not directly related to ICT.

3.2 Data understanding

The data used in the present study is from the 2018 PISA database[4] of the Organization for Economic Co-operation and Development (OECD). This is an international, standardized educational research study of 15-year-olds. PISA 2018 is the seventh

[4] https://www.oecd.org/pisa/publications/pisa-2018-results.html

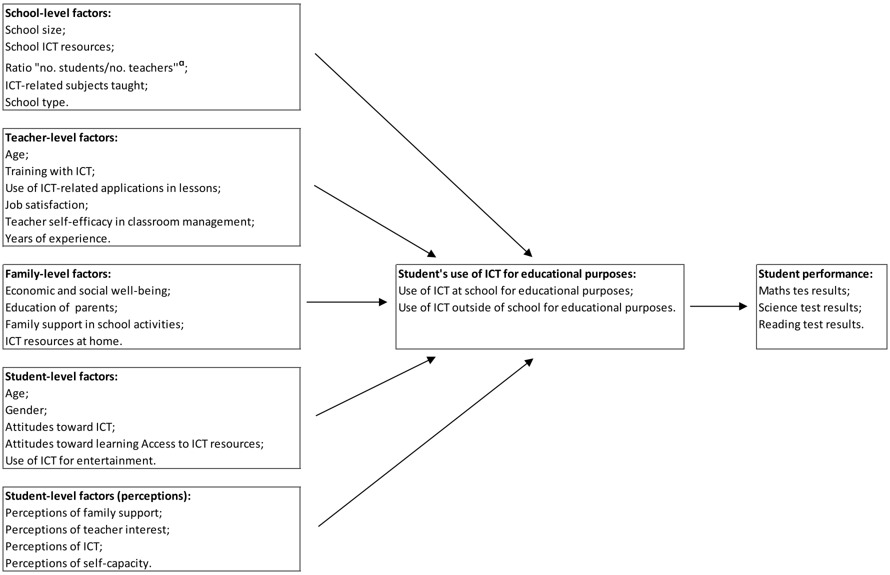

Fig. 1 Conceptual framework systematizing factors that influence the use of ICT for education purposes (adapted from Biagi & Loi, 2013). Note. Own elaboration based on Biagi and Loi (2013). [a] "Ratio 'no. students/no. teachers'" refers to the ratio between the number of students and the number of teachers, to provide an idea of how many students per teacher there are in the respective school

three-year cycle of this OECD research program, which began in 2000. PISA 2018 involved 79 countries. Around 600,000 students participated, representing a population of approximately 32 million 15-year-old students. The sample consists of approximately 50% female students and 50% male students.

PISA databases are constructed based on questionnaires and tests. The variables used for the present study come from the student survey, the teacher questionnaire, the student computer familiarity questionnaire, and the parent questionnaire.

The database used in the present study originally included 612,004 students, 40 variables related to the characteristics of the students, and two variables on the use of ICT by students for educational purposes.

3.3 Data preparation

To accomplish the goal of the present study, and in line with previous studies (Salloum et al., 2017; Jamil et al., 2018), the data preparation step was very important. Data preprocessing allowed the databases to be error-free and ready to use and ultimately enabled us to draw valid conclusions from them.

We selected and downloaded three PISA 2018 databases from https://www.oecd.org/pisa/data/2018database/: SCHQQQQ, which refers to the survey filled out by school officials, TCHQQQ, which refers to the survey filled out by teachers, and STUQQQ, which refers to the survey filled out by students.

After a detailed analysis of the variables in each of the transferred databases, and in line with the framework of Biagi and Loi (2013), we selected the variables with the highest potential of explaining/predicting the evolution of the behavior of the variables that characterize the use of ICT: the use of ICT for educational purposes in the classroom and the use of ICT for educational purposes outside the classroom. To build the databases used in the present study, we used three variables from the SCHQQQ database, 10 variables from the TCHQQQ database, and 29 variables from the STUQQQ database.[5] When preparing the database, we replaced the values of some of the attributes. Values with the following contents were replaced with a "missing values" category: "95—valid skip", "97—not applicable", "98—invalid", "99—no response", "9995—valid skip", "9997—not applicable", "9998—invalid", "9999—no response", "5—valid skip", "7—not applicable", "8—invalid", and "9—no response". Once this procedure was concluded, we performed a descriptive analysis of the database to derive statistics such as the number of missing values, the mean, the standard deviation, the minimum, the maximum, and the mode and thus get a better view of the distribution of variables and their usefulness for the present study.

As a result of this analysis, we removed the variables "How many languages […] do you and your parents speak well enough to converse with others? Your mother" and "How many languages […] do you and your parents speak well enough to converse with others? Your father" due to the excessive lack of data, i.e., missing values. Moreover, we included the variables "School size (sum)" and "School ownership",[6] which were extracted from the SCHQQQ database and added to the database via the ID of the schools, i.e., each student got the values of these variables according to their school.

We had to transform the variables "Teacher's use of specific ICT applications (WLE[7])", "Teacher's satisfaction with the current job environment (WLE)", "Teacher's satisfaction with teaching profession (WLE)", "Teacher's self-efficacy in classroom management (WLE)", and "Teacher's self-efficacy in maintaining positive relationships with students (WLE)" in order to link them to the students. These variables refer to each teacher, and since each student has several teachers, there are several values of these variables for each student. A simple way to solve this problem would be to use the average of the teachers in the school for each student. However, the average can be the same between schools with very different teacher profiles and these are variables that represent differences from an OECD average value, which led us to another transformation to better capture those differences: we decided to create new variables from the old ones. The new variables represent the "% of teachers in the X range of variable Y". We achieved these transformations by discretizing the variables into three intervals where each variable took only three values (interval1, interval2, or interval3, where the 1st represents the worst-case scenario, the 2nd represents the middle-case scenario, and the 3rd represents the best-case scenario). Next, we created binary variables for each interval, with '1'

[5] The tables in Chapter 3.4 ("Modeling and evaluation") allow a detailed analysis of the chosen variables and their origin.

[6] Also subjected to the same data preparation process.

[7] Weighted likelihood estimates.

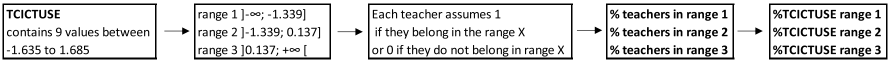

Fig. 2 Example of variable transformations to link it to the students. Note. Source: Own elaboration

meaning a teacher is assigned to interval X and '0' meaning the opposite. Finally, we created variables corresponding to the sum of the teachers in each school to obtain the denominator for the ratios. To finish this transformation, we merged the two databases using the ID of the schools.[8] In the end, we obtained the variables "% of teachers in the X range of variable Y" for the five transformed variables, which resulted in 15 variables. An example is shown in Fig. 2 and explained in the following paragraph.

As an example, the variable "Teacher's use of specific ICT applications" in one school initially contained 9 values, i.e., there were 9 teachers for that school, ranging from -1.635 to 1.685, with a mean of 0.022. After discretization, the variable included the following bins: range 1]-∞; -1.339], range 2]-1.339; 0.137], and range 3]0.137; + ∞ [. Once this division was defined, we created the binary variables for each interval and estimated a variable corresponding to the sum of all "1" values in each of the intervals. Afterward, we aggregated the teachers' information for each school. For this, we estimated the proportion of teachers in each school belonging to each of the previously mentioned intervals. Finally, we merged this database with the one containing the students' variables through the school ID so that all the students in the same school would obtain these values.

Still on the topic of the independent variables, there were other variables we had to modify.[9] In these cases, we only considered the average values of the teachers for each student in the same school, except for the variable "Originally trained teacher (wide definition): standard, in-service, or work-based teacher training (composite)", for which we considered the mode of the teachers' answers for each student, given that this is a categorical variable. We decided what type of transformation to use based on the variance of the values. In case the variance observed was limited, we considered that the mean or the mode of the teachers' values would be an acceptable representation of how these characteristics impact students.

Data preparation also involved eliminating all students for whom there were no values for both dependent variables in the database, which reduced the data by about 50% (from 612,004 to 324,956). Given that the variables from the teacher survey (TCHQQQ) also had a very high number of missing data, we decided to perform two separate analyses. In the first analysis, we used the database excluding the variables from the teachers' survey (TCHQQQ), hereafter called database 1. In the second analysis, we used database 2, which includes all variables. After creating the

[8] The ID of the schools was the only variable in common between those related to students and those related to teachers, i.e., both databases had this variable.

[9] "How old are you?", "Included in teacher education, training, or other qualification: Technology", "Current need for professional development: ICT (information and communication technology) skills for teaching", "How many years of work experience do you have? Year(s) working as a teacher in total", and "Originally trained teacher (wide definition): standard, in-service, or work-based teacher training (composite).".

two databases, we eliminated all the students who had missing values in any of the variables, which made the databases smaller. Database 1 was left with 197,267 students and database 2 was left with 44,899 students.

3.4 Modeling and evaluation

The modeling phase consisted in constructing two decision tree models (Tables 7, 8, 9, and 10 in the Appendix) using the RapidMiner software. The independent and dependent variables included in the models are the ones shown in Tables 1 and 2, respectively. We considered these variables because they were available in the PISA 2018 database, a reliable and reputable database, and because they were able to represent each of the factors that influence students' use of ICT for educational purposes according to Biagi and Loi (2013).

Table 2 presents the variables used to support the creation of the dependent variables. Based on these variables, we assumed that each student could belong to one of four possible clusters: cluster 0 (ICT use during lessons < 0 and ICT use outside of lessons < 0), cluster 1 (ICT use during lessons > 0 and ICT use outside of lessons < 0), cluster 2 (ICT use during lessons < 0 and ICT use outside of lessons > 0), and cluster 3 (ICT use during lessons > 0 and ICT use outside of lessons > 0). This is shown in Fig. 3. Since these variables originally represented the difference from the OECD student average, zero acted as the mean value.

In cluster 0, students' use of ICT during lessons and students' use of ICT outside of lessons were below average. In cluster 1, students' use of ICT during lessons was above average but the use of ICT outside of lessons was below average. In cluster 2, students' use of ICT during lessons was below average and the ICT use outside of lessons was above average. Lastly, in cluster 3, students' use of ICT during lessons and students' use of ICT outside of lessons were above average.

After observing the distribution of students within these clusters, we decided it would be better for the model to predict only two possible extreme situations: above-average use of ICT for educational purposes (cluster 3) (regardless of location, i.e., at school or outside school) and below-average use of ICT for education purposes (cluster 0) (regardless of location), as most of the students fall within one of these two clusters. In database 1, cluster 0 had 61,383 students, cluster 1 had 26,904, cluster 2 had 37,262, and cluster 3 had 71,718. In database 2, cluster 0 had 15,907 students, cluster 1 had 8222, cluster 2 had 7932, and cluster 3 had 12,838. After disregarding all students belonging to the intermediate clusters, database 1 was left with 133,101 students while database 2 was left with 28,745 students.

To obtain reliable classification models, we divided the data into training, testing, and validation groups, which we then used to train the model. This division resulted in 80% training data and 20% test data. Of the 80% training data, 20% was validation data. To determine the optimal parameters, we used a grid search procedure. We used the validation data to determine the optimal decision tree parameters (minimum leaf size, maximum tree depth, distribution criteria, and minimum leaf size for distribution). To evaluate the quality of the models obtained, we used the "precision", "recall", and "accuracy" metrics.

Table 1 Independent variables used

Independent variables		Models	Source
School-level factors			
School ICT resources	ICT available at school	1, 2	STUQQQ
No. of students/no. of teachers ratio	Student/teacher ratio	1, 2	SCHQQQ
ICT-related subjects taught	Taught at school: How to use keywords when using a search engine such as <Google©>, <Yahoo©>, etc.	1, 2	STUQQQ
	Taught at school: How to decide whether to trust information from the Internet	1, 2	STUQQQ
	Taught at school: How to compare different web pages and decide what information is more relevant for your schoolwork	1, 2	STUQQQ
	Taught at school: To understand the consequences of making information publicly available online on <Facebook>, [...]	1, 2	STUQQQ
	Taught at school: How to use the short description below the links in the list of results of a search	1, 2	STUQQQ
	Taught at school: How to detect whether the information is subjective or biased	1, 2	STUQQQ
	Taught at school: How to detect phishing or spam emails	1, 2	STUQQQ
Size of the school	School size (sum)	1, 2	SCHQQQ
Type of school	School ownership	1, 2	SCHQQQ
Teacher-level factors			
Age	How old are you?	2	TCHQQQ
Training with ICT	Included in teacher education, training, or other qualification: Technology	2	TCHQQQ
	Current need for professional development: ICT (information and communication technology) skills for teaching	2	TCHQQQ
Use of ICT-related applications in the classroom	Teacher's use of specific ICT applications (WLE)	2	TCHQQQ
Job satisfaction	Teacher's satisfaction with the current job environment (WLE)	2	TCHQQQ
	Teacher's satisfaction with the teaching profession (WLE)	2	TCHQQQ
Teacher self-efficacy in classroom management	Teacher's self-efficacy in classroom management (WLE)	2	TCHQQQ
	Teacher's self-efficacy in maintaining positive relationships with students (WLE)	2	TCHQQQ
Years of experience	How many years of work experience do you have? Year(s) working as a teacher in total	2	TCHQQQ

Table 1 (Continued)

Independent variables		Models	Source
	Originally trained teacher (wide definition): standard, in-service, or work-based teacher training	2	TCHQQQ
Family-level factors			
Economic and social well-being	Family wealth (WLE)	1, 2	STUQQQ
Parental education	Mother's education (ISCED[a])	1, 2	STUQQQ
	Index of the highest education level of parents (on an internationally agreed scale for years of schooling)	1, 2	STUQQQ
	Father's education (ISCED)	1, 2	STUQQQ
ICT resources at home	ICT available at home	1, 2	STUQQQ
Student-level factors			
Age	Student (standardized) birth year	1, 2	STUQQQ
Gender	Student (standardized) gender	1, 2	STUQQQ
Attitudes toward ICT	Interest in ICT (WLE)	1, 2	STUQQQ
Attitudes toward learning	Resilience (WLE)	1, 2	STUQQQ
	Attitude toward school: learning activities (WLE)	1, 2	STUQQQ
Access to ICT resources	ICT resources (WLE)	1, 2	STUQQQ
Use of ICT for entertainment	ICT use outside of school (leisure) (WLE)	1, 2	STUQQQ
Student ID	Student identifier	1, 2	STUQQQ
Student's home country	Country identifier	1, 2	STUQQQ
Identifier of the student's school	School identifier	1, 2	STUQQQ
Student-level factors (perceptions)			
Perceptions of family support	Parents' emotional support as perceived by the student (WLE)	1, 2	STUQQQ
Perceptions of teacher interest	Perceived teacher's interest (WLE)	1, 2	STUQQQ
Perceptions of self-capacity	Perceived ICT competence (WLE)	1, 2	STUQQQ
	Perceived autonomy related to ICT use (WLE)	1, 2	STUQQQ

Source: Own elaboration based on the PISA 2018 database

[a]International Standard Classification of Education

Table 2 Dependent variables used

Dependent variables		Models	Source
Students' use of ICT for educational purposes			
Students' use of ICT for educational purposes	Subject-related ICT use during lessons (WLE)	1, 2	STUQQQ
	Subject-related ICT use outside of lessons (WLE)	1, 2	STUQQQ

Source: Own elaboration based on the PISA 2018 database

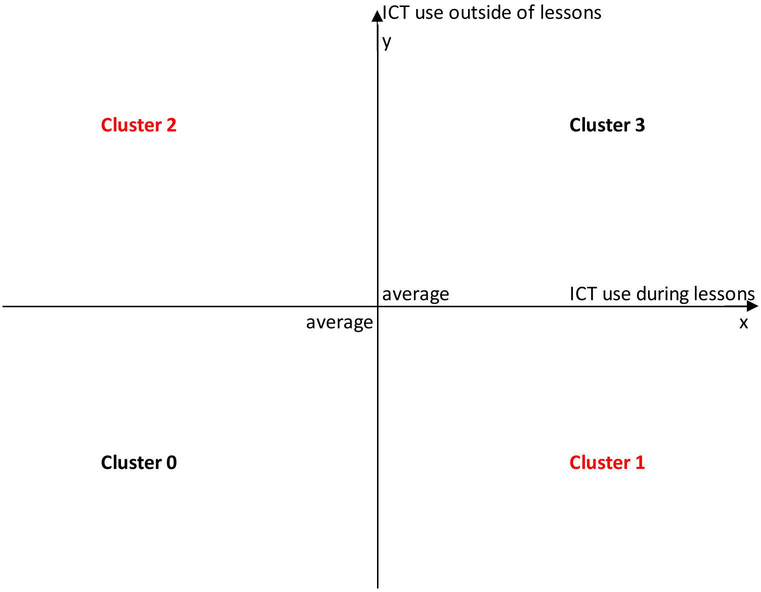

Fig. 3 Student clusters for dependent variables (ICT use during lessons and ICT use outside lessons). Note. Source: Own elaboration based on the PISA 2018 database

4 Results and discussion

In this section, we present and discuss the results obtained. First, we present and discuss the model that does not consider the data from the teacher survey (decision tree 1), and then the model that considers all the data (decision tree 2).

The distribution of the students per cluster for the dataset that considers all the data is presented in Fig. 4. Cluster 0 consists of 61,383 students who had below-average ICT use during lessons and outside of lessons. Cluster 3 consists of 71,718 students who had above-average ICT use during lessons and outside lessons. In database 2 (Fig. 5), cluster 0 consists of 15,907 students who had below-average ICT use during lessons and outside of lessons. Finally, cluster 3 consists of 12,838 students who had above-average ICT use during lessons and outside lessons.

For the construction of decision tree models 1 and 2, in the first parameter (minimum leaf size), we considered a minimum of 2 and a maximum of 100 (considering

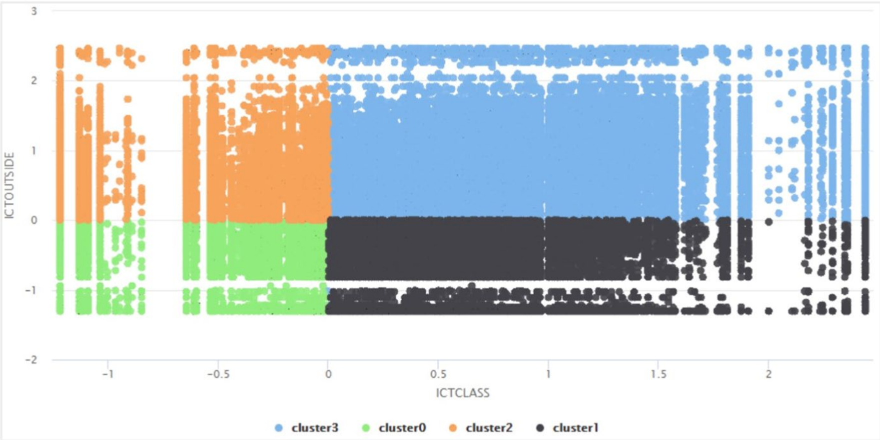

Fig. 4 Division of database 1 (not considering data from teacher survey) into quadrants (according to the clusters in Fig. 3). Note. Source: Own elaboration based on the PISA 2018 database

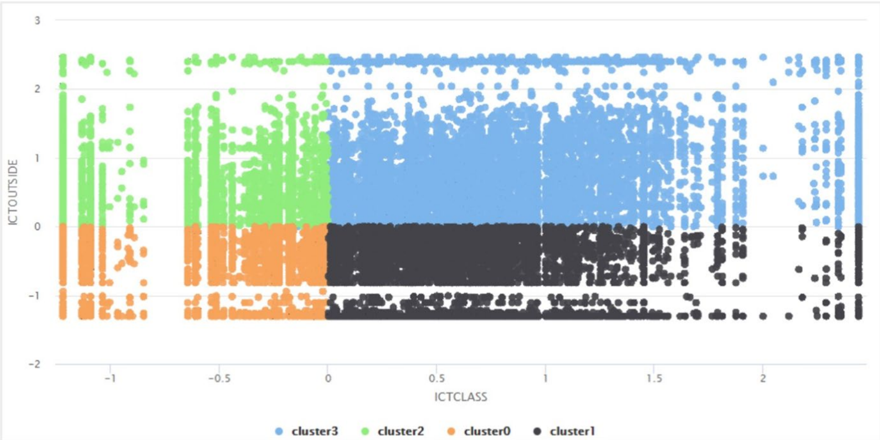

Fig. 5 Division of database 2 (considering all the data) into quadrants (according to the clusters in Fig. 3). Note. Source: Own elaboration based on the PISA 2018 database

20 possible values in this interval[10]), in the second parameter (maximum tree depth), we considered a minimum of 1 and a maximum of 20 (considering 15 possible values in this interval[11]), in the third parameter (distribution criteria), we considered the possibilities of the Gini index, the information gain, and the gain ratio, and finally, in the last parameter (minimum leaf size for distribution), we considered a minimum of 1 and a maximum of 100 (considering 10 possible values in this interval[12]). Once

[10] The considered values were: 2, 7, 12, 17, 22, 27, 31, 36, 41, 46, 51, 56, 61, 66, 71, 76, 80, 85, 90, 95, and 100.

[11] The considered values were: 1, 2, 4, 5, 6, 7, 9, 10, 11, 12, 14, 15, 16, 17, 19, 20.

[12] The considered values were: 1, 11, 21, 31, 41, 51, 60, 70, 80, 90, 100.

we determined the optimal parameters, we used them to train the final decision trees, which we then tested using the test data.

4.1 Summary of results and discussion—decision tree 1

Tables 7 and 8 in Appendix present the classification rules inferred from the decision tree model 1, which disregards data from the teacher survey. Having tuned the parameters, we ended up using the Gini index as the distribution criterion, a minimum leaf size of 22, a maximum decision tree depth of 5, and a minimum leaf size for distribution of 60. This model achieved an accuracy of 68.04% (Fig. 6), which demonstrates it has a good ability to distinguish between cluster0 and cluster3 students.

Regarding the recall metric, the model correctly predicted 72.86% of the true values (cluster = cluster 3) and 62.42% of the false values (cluster = cluster 0). In terms of the precision metric, the model correctly predicted 69.37% of the true values and 66.31% of the false values. Given the high scores achieved in all three metrics, we can assume that this model can distinguish different patterns in ICT use for educational purposes.

Table 3 lists the obtained weight of the attributes in descending order of magnitude.

The weight of an attribute indicates the improvement in the performance of the tree when the attribute in question is selected for a given node.

When analyzing the results, a given level of a variable may be considered high or low depending on the country. Therefore, we analyzed the results per country, given the decision trees obtained. For example, a school size of 1000 students can be considered high in a country where the average is 500 and low in a country where the average is 1200. It is important to have this in mind in order to better understand the discussion of the results. However, we did not follow a per-country analysis for all variables, because, for most of them, the averages were very close.

The decision tree obtained allowed us to conclude that ICT access at school had a positive influence on the use of ICT for education purposes in almost all countries. In fact, access to ICT at school is crucial, as the school is where students are expected to more easily find support on how to use ICT for educational purposes. Access to ICT at home and the existence of ICT resources at home also had a positive influence on the use of ICT for educational purposes. This is entirely reasonable

accuracy: 68.04%

	true false	true true	class precision
pred. false	7663	3893	66.31%
pred. true	4614	10451	69.37%
class recall	62.42%	72.86%	

Fig. 6 Performance of decision tree 1 (not considering data from teacher survey). Note. Source: Own elaboration based on the PISA 2018 database

Table 3 Weight of the attributes (database 1)

Variable	Order of magnitude (%)
ICT available at school	23.6%
ICT use outside of school (leisure)	21.6%
Student/teacher ratio	10.8%
School size (sum)	7.9%
ICT available at home	5.9%
ICT resources	4.1%
Perceived ICT competence	3.6%
Attitude toward school: learning activities	3.6%
Parents' emotional support as perceived by the student	2.7%
Perceived autonomy related to ICT use	2.7%
School ownership	2.5%
Resilience	2.1%
Country identifier	2.0%
Perceived teacher's interest	1.7%
Interest in ICT	1.4%
Family wealth	0.9%
Index of the highest education level of parents (on an internationally agreed scale for years of schooling)	0.6%
Taught at school: How to compare different web pages and decide what information is more relevant for your schoolwork	0.6%
Taught at school: To understand the consequences of making information publicly available online on <Facebook>, […]	0.5%
Student (standardized) gender	0.3%
Taught at school: How to use the short description below the links in the list of results of a search	0.3%
Taught at school: How to detect phishing or spam emails	0.3%
Taught at school: How to detect whether the information is subjective or biased	0.3%
Father's education (ISCED)	0.2%

since the use of ICT for educational purposes outside of school is mainly expected to happen at home. In this sense, both types of access are central to surpassing the OECD average for ICT use for educational purposes in both locations, i.e., in and out of school. These results seem to be in line with the findings of Basak (2014). This author showed that the availability of ICT resources is one of the factors that promotes ICT adoption by teachers and our results suggest that this pattern may also apply to students.

The use of ICT for entertainment had a positive effect on the use of ICT for education purposes in almost all countries. The use of ICT for entertainment contributes to improving ICT skills, which in turn may favor the use of ICT for educational purposes. Nevertheless, it is curious to note that this type of use, which could have a negative influence because of the distraction it can cause, turned out to be generally positive. This conclusion may help to understand why Biagi and Loi (2013) reported that the only activity with a positive relationship between frequency of use and performance improvement was online gaming, i.e., the use of ICT for entertainment led to an increase in the use of ICT for educational purposes. Despite this, according to Srijamdee and Pholphirul (2020), the use of ICT for entertainment should not exceed a certain limit, otherwise, it consumes all the students' free time (Kubiatko & Vlckova, 2010) and compromises their availability to learn.

The interest of teachers (as perceived by students) had a positive influence on the use of ICT for educational purposes in most countries. A more interested teacher should be able to make the lessons more dynamic and try to be up-to-date with the state of the art in their teaching, which includes promoting the use of ICT, especially for educational purposes.

Parental emotional support (as perceived by students) had a negative influence on the use of ICT for educational purposes in general, contrary to what we expected. One explanation for this result may be that students with too much parental support do not feel the need to adopt new ways of studying or get out of their comfort zone, restricting themselves to "normal" studying, i.e., without the use of ICT.

Autonomy in using ICT had a positive influence on the use of ICT for educational purposes in most countries. A student must be able to use ICT autonomously in order to feel ready to use ICT for educational purposes. Likewise, ICT competence had a positive effect on good ICT use for educational purposes. We already expected this result, since a good use of ICT in any kind of activity requires the user to have the necessary skills. These results may help to explain why Areepattamannil and Santos (2019) observed a link between autonomy and competence in using ICT and a greater interest in topics such as sciences. Finally, these results may help to explain why Xiao and Hu (2019a) indicated that these variables can improve literacy.

Financial and social well-being also had a positive effect on the use of ICT for educational purposes. In terms of financial well-being, ICT involves some investment. This is reinforced by the fact that the use of ICT for educational purposes is highest among students enrolled in private schools. As for social well-being, a calm and positive environment should be beneficial for almost all types of student activities, including the use of ICT for educational purposes. This result (as well as those of similar types of influence such as parents' level of education and type of school attended) is consistent with those of Skryabin et al. (2015) and Xiao and Hu (2019a), who showed that the higher the socio-economic status of the student, the better the results. As is the case with other variables, the better performance associated with these variables in the previous studies might be explained by their positive effect on the use of ICT for educational purposes.

The level of education of parents also had a positive effect on ICT use for educational purposes. A high level of education of parents is usually linked with a culture that fosters study, interest in learning, and goal orientation. Thus, these students are more culturally inclined to follow their parents' example and dedicate themselves to studying and achieving good results, which might mean using ICT for educational purposes.

Attitudes toward learning, as well as interest in ICT, proved positive in most of the results. These results are not surprising. Students with a positive and constructive attitude toward the role of school and learning are more willing to engage in study-related activities, namely the use of ICT for educational purposes. As for interest in ICT, students interested in ICT are expected to carry out more ICT-related activities (namely for educational purposes), which may help to explain why Li et al. (2020) found a positive relationship between interest in ICT and performance in science.

Learning about ICT-related topics was positive for the use of ICT for educational purposes. When students master ICT and recognize its dangers and possible uses, they are more apt to use it for educational purposes.

Finally, the results of the no. of students/no. of teachers ratio and school size variables point in both directions. Starting with the ratio, the results show that a lower ratio was beneficial for a better use of ICT for educational purposes, except in the UK, where the opposite was observed (a higher ratio can indicate more autonomy and thus better results, however, this is only a possible explanation, and this result should be explored further), and Hungary and Albania, where an intermediate ratio was more beneficial. A low ratio may be beneficial as it means that teachers can pay more attention to each student and manage the classroom more easily, which are factors that may lead to more dynamic lessons, in which ICT might be used for educational purposes. As for school size, the generality of the results indicates that larger schools were better when it came to ICT use for educational purposes. However, some results point in the opposite direction. Larger schools are typically more modern, located in more developed locations, have good ICT resources, and have more modern teaching methods, which might foster the use of ICT for educational purposes. On the other hand, in locations with more rudimentary schools that have poor ICT resources and worse infrastructure, a lower ratio might be important because it means more ICT resources per student, fostering ICT use for educational purposes. This explanation makes sense in the context of the present study since the countries where smaller school sizes were beneficial for the use of ICT for educational purposes (Bulgaria, Morocco, Switzerland, Turkey, and the Tatarstan region in Russia) have a lower Human Development Index (HDI) than those where the opposite happened (Estonia, France, Hong Kong, Singapore, Slovakia, the Moscow region in Russia, and Uruguay), apart from Switzerland and Uruguay, which did not fit this pattern.[13]

4.2 Summary of results and discussion—decision tree 2

Tables 9 and 10 in Appendix present the classification rules inferred from the decision tree model 2, which considers all variables, namely those collected from the teachers' survey. We based the final tree on the Gini index and considered a minimum leaf size of 27, a maximum decision tree depth of 7, and a minimum distribution size of 4.

[13] In Hungary's and Brunei's results the variable related to school dimension is also considered, but it does not lead to any conclusion.

This model achieved an accuracy of 64.78% (Fig. 7). This decrease in accuracy when compared to decision tree model 1 may be related to the significant reduction in the number of observations when considering all the predictors.

As far as the recall is concerned, the model correctly predicted 62.54% of the true values (cluster = cluster 3) and 66.58% of the false values (cluster = cluster 0). Regarding precision, the model correctly predicted 60.17% of the true values and 68.77% of the false values.

Table 4 lists the obtained weight of the attributes in descending order of magnitude.

The relationships that emerged in this decision tree are similar to those found in decision tree 1, whose model excluded the teacher-related variables. Access to ICT both at school and at home was important for a good ICT use for educational purposes. ICT use for entertainment was once more positive for the use of ICT for educational purposes in most countries. Autonomy in ICT use and resilience were also important for students' use of ICT for educational purposes. Learning about ICT-related topics such as search engines was also a determining factor in dictating the use of ICT for educational purposes.

Teacher-related variables such as satisfaction with the profession, level of ICT use in the classroom, satisfaction with the work environment, effort in maintaining good relationships with students, and need for ICT training also appeared in this decision tree. The majority of these variables had a positive influence on students' ICT use for educational purposes. The only exception was the need for ICT training, which, as expected, had a negative influence, i.e., a greater need for ICT training by teachers was detrimental to students' ICT use for educational purposes. This result is in line with that found by Juggernath and Govender (2020), who showed that teachers' need for ICT training is a barrier to ICT implementation in classrooms.

Unlike in decision tree model 1, the emotional support perceived by students promoted ICT use for educational purposes in most countries. However, it rarely appeared in the decision tree, so it would be interesting to conduct further studies to clarify this relationship.

The no. of students/no. of teachers ratio and the school size were less relevant in this second model. The results show that a larger school size was beneficial for a good ICT use for educational purposes (again, countries confirming this rule have a high HDI, namely Hong Kong and the US). As for the ratio, it appeared only once in this decision tree, being associated with the UK. The UK stood out in decision tree 1 in this regard, and in decision tree 2, the results went in the same direction. Again,

accuracy: 64.78%

	true false	true true	class precision
pred. false	2118	962	68.77%
pred. true	1063	1606	60.17%
class recall	66.58%	62.54%	

Fig. 7 Performance of decision tree 2 (considering all the data). Source: Own elaboration based on the PISA 2018 database

Table 4 Weight of the attributes (database 2)

Variable	Order of magnitude (%)
ICT available at school	15.0%
School size (sum)	11.6%
ICT use outside of school (leisure)	9.9%
Resilience	8.9%
% of teachers in the worst tertile of "teacher's use of specific ICT applications"	6.3%
ICT available at home	4.7%
Perceived autonomy related to ICT use	3.9%
Taught at school: How to compare different web pages and decide what information is more relevant for your schoolwork	3.4%
School ownership	3.4%
Country identifier	3.3%
% of teachers in the middle tertile of "Teacher's satisfaction with the current job environment"	3.0%
Average of "Included in teacher education, training, or other qualifications: Technology" answers	2.9%
% of teachers in the worst tertile of "Teacher's satisfaction with teaching profession"	2.9%
% of teachers in the best tertile of "Teacher's satisfaction with teaching profession"	2.8%
Average of "How old are you?" answers	2.4%
Student/teacher ratio	2.3%
Parent's emotional support as perceived by the student	2.1%
Perceived teacher's interest	2.0%
Taught at school: How to use keywords when using a search engine such as <Google©>, <Yahoo©>, etc	1.8%
Average of "Current need for professional development: ICT (information and communication technology) skills for teaching" answers	1.6%
ICT resources	1.3%
Attitude toward school: learning activities	1.3%
% of teachers in the best tertile of "Teacher's self-efficacy in maintaining positive relationships with students"	1.2%
Mode of "Originally trained teacher (wide definition): standard, in-service, or work-based teacher training"	1.0%
Taught at school: To understand the consequences of making information publicly available online on <Facebook>, […]	0.90%

the need to promote autonomy might be higher in this country for some reason that should be further explored.

Finally, the attributes that influenced ICT use for educational purposes differently from what was expected were the ICT training of the teachers, the teachers' background in education, and the age of the teachers. We expected teachers' ICT training and background in education to be beneficial for the dependent variables and expected teachers' age to have an inverted U-shaped relationship with the dependent variables. However, the results indicate that this was not the case. As for ICT training, one possible explanation is that teachers with no ICT training see ICT as a great benefit and innovation, and even without the necessary skills, they try to implement it into their teaching, making a greater effort than teachers for whom ICT is something more common. However, these results contradict those of Salinas et al. (2016), who found teachers' ICT training to have a positive influence on their adoption of technology. As for teachers' age, older ages were found to be more beneficial. This relationship may be related to younger teachers being less confident in their ability to innovate in teaching and to promote a teaching style that is different from the traditional one, despite their greater ability with ICT. Finally, teachers having a background in education had a negative influence on ICT use for educational purposes. One possible explanation is that teachers coming from other fields have different views about school subjects and more innovative teaching methodologies, which may include promoting ICT use.

5 Conclusions

In the present study, we researched which factors have the greatest impact on students' ICT use for educational purposes with the aim of assisting the implementation and use of ICT in schools. Given that ICT are massively adopted in society it is of crucial importance that schools have identified those factors and consequently adapt their ecosystem to meet the context in which students learn and grow up in today's society. We used two decision tree models to answer the research questions. The obtained decision trees performed well, allowing us to derive valid and valuable conclusions about which factors influence students' use of ICT for educational purposes. We observed strong trends in the results among countries, however, some results turned out to be country-specific, so we recommend that each country analyzes its specificities and adopts the best measures to improve students' use of ICT for educational purposes within its own context.

The database used for one of the decision tree models included variables used to characterize the teachers at the schools included in the study, while the database used for the other decision tree model did not include this data. The results obtained allowed us to conclude that the teacher-related variables were important to correctly predict the use of ICT for educational purposes, since the results of the two models were slightly different, and teacher-related variables were relevant in influencing the dependent variables in the second model. However, it is important to highlight that the model with the teacher-related variables performed worse than the other model.

One possible explanation is the smaller number of observations included in the model with more variables, i.e., the model with the teacher-related variables.

Summarizing the observed trends, we found that the three most important variables for a good use of ICT for educational purposes were ICT use for entertainment, access to ICT at home, and access to ICT at school, which positively impacted ICT use for education purposes in almost all countries. These were followed by school size and the no. of students/no. of teachers ratio. However, the impact of these two variables varied more among countries than those of the previous three variables, so it would be interesting for future studies to focus on these two variables in order to better clarify their impact. There seem to be some trends related to the HDI of the countries that explain these variations, which could be a starting point to further explore this topic. A very interesting observation is that ICT use for entertainment seems to be able to stimulate ICT use for educational purposes as long as it is not excessive. Apart from these variables, there were several other important ones, such as learning ICT-related topics at school, teacher interest, autonomy with ICT, resilience, financial and social well-being, parental education level, attitudes toward school, ICT skills, interest in ICT, the type of school, and emotional support.

As expected, most of the teacher-related variables had a positive impact on promoting the use of ICT for educational purposes. Nonetheless, there were some surprising results, such as higher teacher ages being favorable for the use of ICT for educational purposes. Teachers having training in education or ICT was detrimental to students' use of ICT for educational purposes. Finally, and as expected, teachers' need for ICT training was also detrimental to students' use of ICT for educational purposes.

In the present study, we identified the factors with the highest impact on ICT use for educational purposes, which in turn is a very promising factor influencing student performance in various subjects, as stated in the literature. The present study can be used as a basis for future research seeking to build a model capable of effectively predicting between more than two clusters of ICT use for educational purposes and to stratify the various levels of such use. It would also be interesting for future studies to include more relevant variables in order to improve predictions and to gather more data on teacher-related variables in order to obtain a larger database. Finally, we believe that it would be important to focus on the variables that had nonconsensual or unpredictable results and try to obtain new results, as these could lead to more robust conclusions.

Appendix

Table 5 Exploration of the variables from database 1

Variable	Minimum	Maximum	Average	Notes
Student identifier	800007	98329122	Not applicable	This serves as an identifying variable.
School identifier	Not applicable	Not applicable	Not applicable	This variable enabled the connection between teachers and students.
Country/region identifier	8	983	Not applicable	The country with the most students is Spain (724), with 10,981, and the country with the fewest is Morocco (504), with 471.
ICT resources available at school	0	10	6.476	This variable ranges from 0 to 11 because 11 possible types of ICT resources are considered (fixed computer, laptop, tablet, school computer connected to the internet, wireless internet access, USB stick, storage space for school documents, storage hardware, e-book reader, projector, and interactive whiteboard).
Financial and social well-being of families	–7.327	4.628	–0.303	This variable includes issues such as students having their own room, internet access, television, cars, and a bathroom with a shower. This is a source-transformed variable whose value 0 represents the average OECD student.
Level of education among the student's parents	3	16	13.786	Not applicable.
ICT resources available at home	0	11	8.087	This variable ranges from 0 to 11 because 11 possible types of ICT resources are considered (fixed computer, laptop, tablet, internet connection, video game console, mobile phone without internet access, mobile phone with internet access, portable device for listening to music, printer, storage hardware, and e-book reader).
Year of birth of students	2002	2003	2002.075	Most of the students were born in 2002.
Student interest in ICT	–2.936	2.701	–0.034	This variable encompasses aspects such as "I forget about time when I am using information and communication technologies" and "I get excited when I am discovering new applications and/or digital devices". It is a source-transformed variable whose value 0 represents the average OECD student.

Table 5 (continued)

Variable	Minimum	Maximum	Average	Notes
Student resilience	–3.167	2.767	0.036	This variable encompasses aspects such as feeling proud to accomplish something and being able to manage several things simultaneously.It is a source-transformed variable whose value 0 represents the average OECD student.
Students' attitudes toward school	–2.538	1.084	–0.033	This variable encompasses the aspects "effort at school will help me get a good job", "effort at school will help me get into a good university", and "effort at school is important". It is a source-transformed variable whose value 0 represents the average OECD student.
ICT possessions at home	–3.814	4.007	–0.252	This variable encompasses "having educational software", "having access to the internet", "having a mobile phone with internet access", "having computers", "having tablets", and "having e-book readers". It is a source-transformed variable whose value 0 represents the average OECD student.
Use of ICT for entertainment outside school	–3.594	4.315	0.060	This variable encompasses activities such as online gaming, using email, chatting online, participating in social networks, searching the internet for fun, and downloading movies. It is a source-transformed variable whose value 0 represents the average OECD student.
Parental emotional support as perceived by the student	–2.447	1.035	–0.050	This variable encompasses aspects such as "my parents support my efforts and achievements at school", "my parents support me when I have difficulties at school", and "my parents encourage me to be confident". It is a source-transformed variable whose value 0 represents the average OECD student.
Interest of the teacher as perceived by the student	–2.218	1.825	0.082	This variable reflects the answers of students when asked to rate the teacher's enthusiasm in them, rate the comprehensibility of what was taught, whether the teacher appreciates the subject taught, and whether the teacher showed enjoyment in the activity of teaching. These answers refer to the most recent lessons (at the time of the questionnaire). It is a source-transformed variable whose value 0 represents the average OECD student.

Table 5 (continued)

Variable	Minimum	Maximum	Average	Notes
Students' ICT competence as perceived by themselves	–2.619	2.490	–0.019	This variable encompasses aspects such as comfort using ICT, ease of solving ICT-related problems, and the ability to advise on ICT-related purchases. It is a source-transformed variable whose value 0 represents the average OECD student.
Students' autonomy related to ICT use	–2.514	2.026	–0.032	The variable encompasses aspects such as the ability to install software by oneself, the ability to use digital devices in the intended way, and the autonomy to start solving problems related to digital devices. It is a source-transformed variable whose value 0 represents the average OECD student.
No. of students/no. of teachers ratio of each school	1	100	13.222	Not applicable.
Size of the school in terms of the number of students	2	10,700	853.942	Not applicable.
Learning at school how to use keywords in internet search engines	1	2	Not applicable	This is a binary variable that takes the value "1" if the answer is positive, or the value "2" if the answer is negative. The value "1" has 78,445 responses and the value "2" has 54,656.
Learning at school how to decide what information to trust on the internet	1	2	Not applicable	This is a binary variable that takes the value "1" if the answer is positive, or the value "2" if the answer is negative. The value "1" has 93,712 responses and the value "2" has 39,389.
Learning at school how to compare information from the internet and decide which is the most relevant	1	2	Not applicable	This is a binary variable that takes the value "1" if the answer is positive, or the value "2" if the answer is negative. The value "1" has 84,970 responses and the value "2" has 48,131.
Learning at school about the consequences of making information public on the internet	1	2	Not applicable	This is a binary variable that takes the value "1" if the answer is positive, or the value "2" if the answer is negative. The value "1" has 96,258 responses and the value "2" has 36,843.
Learning at school how to use the short descriptions under links in search results	1	2	Not applicable	This is a binary variable that takes the value "1" if the answer is positive, or the value "2" if the answer is negative. The value "1" has 71,498 responses and the value "2" has 61,603.
Learning at school how to detect whether information on the internet is subjective or biased	1	2	Not applicable	This is a binary variable that takes the value "1" if the answer is positive, or the value "2" if the answer is negative. The value "1" has 76,777 responses and the value "2" has 56,324.

Table 5 (continued)

Variable	Minimum	Maximum	Average	Notes
Learning at school how to detect spam or phishing emails	1	2	Not applicable	This is a binary variable that takes the value "1" if the answer is positive, or the value "2" if the answer is negative. The value "1" has 62,633 responses and the value "2" has 70,468.
Education of the student's guardian	0	6	Not applicable	Value "0" has 2082 responses, value "1" has 5652 responses, value "2" has 12,657 responses, value "3" has 8476 responses, value "4" has 35,402 responses, value "5" has 21,125 responses, and value "6" has 47,707 responses. 0 is the lowest level of education and 6 is the highest.
Education of the student's guardian	0	6	Not applicable	Value "0" has 2159 responses, value "1" has 5824 responses, value "2" has 14,211 responses, value "3" has 10,448 responses, value "4" has 35,579 responses, value "5" has 21,610 responses, and value "6" has 43,270 responses. 0 is the lowest level of education and 6 is the highest.
Gender of students	1	2	Not applicable	The variable takes the value "1" if the student is female and "2" if the student is male. There are 66,359 female students and 66,742 male students.
Type of school	1	3	Not applicable	This variable takes the value of "1" if the school is private, "2" if the school is private but government-dependent, and "3" if it is public. The value "1" has 10,368 responses, the value "2" has 14,305 responses, and the value "3" has 108,428 responses.
Variable	True Values	False Values	Average	Notes
Level of ICT use for educational purposes	71,718	61,383	Not applicable	Binomial. Dependent variable.

Source: Own elaboration based on the PISA 2018 database

Table 6 Exploration of the variables from database 2

Variable	Minimum	Maximum	Average	Notes
Student identifier	7600005	84008625	Not applicable	Same as in database 1.
School identifier	Not applicable	Not applicable	Not applicable	Same as in database 1.
Country/region identifier	76	840	Not applicable	The country with the most students is Spain (724), with 10,760, and the country with the fewest is Morocco (504), with 471.
ICT resources available at school	0	10	6.267	Same as in database 1.
Financial and social well-being of families	–7.327	4.628	–0.237	Same as in database 1.
Level of education among the student's parents	3	16	13.715	Same as in database 1.
ICT resources available at home	0	11	8.146	Same as in database 1.
Year of birth of students	2002	2003	2002.102	Same as in database 1.
Student interest in ICT	–2.936	2.701	0.103	Same as in database 1.
Student resilience	–3.167	2.369	0.062	Same as in database 1.
Students' attitudes toward school	–2.538	1.084	0.104	Same as in database 1.
ICT possessions at home	–3.814	3.612	–0.229	Same as in database 1.
Use of ICT for entertainment outside school	–3.594	4.286	0.035	Same as in database 1.
Parental emotional support as perceived by the student	–2.447	1.035	–0.003	Same as in database 1.
Interest of the teacher as perceived by the student	–2.218	1.825	0.176	Same as in database 1.
Students' ICT competence as perceived by themselves	–2.619	1.964	0.017	Same as in database 1.
Students' autonomy related to ICT use	–2.514	2.026	0.027	Same as in database 1.
No. of students/no. of teachers ratio of each school	1	100	15.221	Same as in database 1.
Size of the school in terms of the number of students	4	10,700	1,059.016	Same as in database 1.
Learning at school how to use keywords in internet search engines	1	2	Not applicable	This is a binary variable that takes the value "1" if the answer is positive, or the value "2" if the answer is negative. The value "1" has 15,502 responses and the value "2" has 13,243.

Table 6 (continued)

Variable	Minimum	Maximum	Average	Notes
Learning at school how to decide what information to trust on the internet	1	2	Not applicable	This is a binary variable that takes the value "1" if the answer is positive, or the value "2" if the answer is negative. The value "1" has 20,184 responses and the value "2" has 8,561.
Learning at school how to compare information from the internet and decide which is the most relevant	1	2	Not applicable	This is a binary variable that takes the value "1" if the answer is positive, or the value "2" if the answer is negative. The value "1" has 18,182 responses and the value "2" has 10,563.
Learning at school about the consequences of making information public on the internet	1	2	Not applicable	This is a binary variable that takes the value "1" if the answer is positive, or the value "2" if the answer is negative. The value "1" has 21,281 responses and the value "2" has 7464.
Learning at school how to use short descriptions under links in search results	1	2	Not applicable	This is a binary variable that takes the value "1" if the answer is positive, or the value "2" if the answer is negative. The value "1" has 14,180 responses and the value "2" has 14,565.
Learning at school how to detect whether information on the internet is subjective or biased	1	2	Not applicable	This is a binary variable that takes the value "1" if the answer is positive, or the value "2" if the answer is negative. The value "1" has 16,425 responses and the value "2" has 12,320.
Learning at school how to detect spam or phishing emails	1	2	Not applicable	This is a binary variable that takes the value "1" if the answer is positive, or the value "2" if the answer is negative. The value "1" has 12,363 responses and the value "2" has 16,382.
Education of the student's female guardian	0	6	Not applicable	Value "0" has 678 responses, value "1" has 1396 responses, value "2" has 3665 responses, value "3" has 2038 responses, value "4" has 5848 responses, value "5" has 4794 responses, and value "6" has 10,326 responses.
Education of the student's male guardian	0	6	Not applicable	Value "0" has 791 responses, value "1" has 1585 responses, value "2" has 4196 responses, value "3" has 2177 responses, value "4" has 5403 responses, value "5" has 4719 responses, and value "6" has 9874 responses.

Table 6 (continued)

Variable	Minimum	Maximum	Average	Notes
Gender of students	1	2	Not applicable	This variable takes the value "1" if the student is female, or the value "2" if the student is male. There are 14,150 female students and 14,595 male students.
Type of school	1	3	Not applicable	Value "1" has 3344 responses, value "2" has 8080 responses, and value "3" has 17,321 responses. It is clear that most students are enrolled in public schools.
Average age of teachers in each school	26.333	60	43.648	Not applicable.
Average ICT training of teachers in each school	0	1	0.168	This variable takes the value "0" if no teacher in the school has had ICT training and "1" if all teachers in the school have had ICT training. Most schools have few teachers with ICT training.
Average ICT training needs of teachers in each school	1	4	2.541	This variable takes on integer values from 1 to 4, where 1 represents that no training is needed and 4 represents that a lot of training is needed.
Average experience of teachers in each school, measured in years	2	32.571	16.543	Not applicable.
Mode of teachers with basic education training in each school	0	1	Not applicable	The variable takes the value "0" if teachers have a different training and "1" if teachers have education training. The value "0" has 354 responses and the value "1" has 28,391 responses. It is evident that most of the teachers have education training.
Percentage of teachers in a given school in the worst tertile of satisfaction with their working environment	0%	36%	3.6%	Not applicable.
Percentage of teachers in a given school in the middle tertile of satisfaction with their working environment	0%	100%	48.5%	Not applicable.
Percentage of teachers in a given school in the best tertile of satisfaction with their working environment	0%	100%	47.9%	Not applicable.
Percentage of teachers in a given school in the worst tertile of job satisfaction	0%	100%	5.1%	Not applicable.

Table 6 (continued)

Variable	Minimum	Maximum	Average	Notes
Percentage of teachers in a given school in the middle tertile of job satisfaction	0%	100%	51.6%	Not applicable.
Percentage of teachers in each school in the best tertile of job satisfaction	0%	100%	43.3%	Not applicable.
Percentage of teachers in a given school in the worst tertile of classroom management as perceived by themselves	0%	33.3%	1.1%	Not applicable.
Percentage of teachers in each school in the middle tertile of classroom management as perceived by themselves	0%	100%	26.8%	Not applicable.
Percentage of teachers in a given school in the best tertile of classroom management as perceived by themselves	0%	100%	72%	Not applicable.
Percentage of teachers in a given school in the worst tertile of maintaining good relationships with students as perceived by the teachers	0%	50%	1.2%	Not applicable.
Percentage of teachers in a given school in the middle tertile of maintaining good relationships with students as perceived by the teachers	0%	100%	38.6%	Not applicable.
Percentage of teachers in a given school in the best tertile of maintaining good relationships with students as perceived by the teachers	0%	100%	60.2%	Not applicable.
Percentage of teachers in a given school in the worst tertile of ICT use in the classroom	0%	100%	7.8%	Not applicable.
Percentage of teachers in a given school in the middle tertile of ICT use in the classroom	0%	100%	88.6%	Not applicable.
Percentage of teachers in a given school in the best tertile of ICT use in the classroom	0%	66.7%	3.6%	Not applicable.
Variable	True Values	False Values	Average	Notes
Level of ICT use for educational purposes	12,838	15,907	Not applicable	Same as in database 1.

Table 7 Decision Tree 1 (cluster 0)

If	CNTRYID=100 (Bulgary); ICTSCH>7,500; ENTUSE>1,327 → cluster 0 (28/51)
If	CNTRYID=100 (Bulgary); ICTSCH>7,500; ENTUSE≤-1,039 → cluster 0 (26/37)
If	CNTRYID=100 (Bulgary); ICTSCH≤7,500; TEACHINT>-0,878; ICTHOME≤5,500 → cluster 0 (56/80)
If	CNTRYID=100 (Bulgary); ICTSCH≤7,500; TEACHINT≤-0,878; SCHSIZE>1064 → cluster 0 (13/23)
If	CNTRYID=152 (Chile); ENTUSE>-0,309; AUTICT≤-0,780; RESILIENCE≤0,373 → cluster 0 (42/68)
If	CNTRYID=152 (Chile); -1,066<ENTUSE≤-0,309; TEACHINT≤-0,319 → cluster 0 (46/69)
If	CNTRYID=152 (Chile); ENTUSE≤-1,066 → cluster 0 (62/78)
If	CNTRYID=158 (Taipei) → cluster 0 (1872/2512)
If	CNTRYID=392 (Japan) → cluster 0 (2336/2505)
If	CNTRYID=188 (Costa Rica); ENTUSE>-0,558; ICTHOME≤7,500; AUTICT≤-0,374 → cluster 0 (190/337)
If	CNTRYID=188 (Costa Rica); ENTUSE≤-0,558 → cluster 0 (354/502)
If	CNTRYID=191 (Croatia); ENTUSE>-0,229; ICTSCH≤9,500 → cluster 0 (685/1162)
If	CNTRYID=191 (Croatia); ENTUSE≤-0,229; ICTRES>-0,329; EMOSUPS>-0,785 → cluster 0 (113/166)
If	CNTRYID=191 (Croatia); ENTUSE≤-0,229; ICTRES≤-0,329 → cluster 0 (323/396)
If	CNTRYID=203 (Czech Republic); ENTUSE>-0,010; ICTSCH≤8,500; WEALTH≤1,244 → cluster 0 (473/761)
If	CNTRYID=203 (Czech Republic); ENTUSE≤-0,010; STRATIO>7,433 → cluster 0 (762/1037).
If	CNTRYID=214 (Dominican Republic); ICTSCH>5,500; STRATIO>13,764; PAREDINT≤13,250 → cluster 0 (37/71)
If	CNTRYID=214 (Dominican Republic); ICTSCH≤5,500; ENTUSE>0,883 → cluster 0 (17/23)
If	CNTRYID=214 (Dominican Republic); ICTSCH≤5,500; ENTUSE≤0,265 → cluster 0 (138/183)
If	CNTRYID=233 (Estonia); -0,358<ENTUSE≤0,030; ICTSCH≤7,500 → cluster 0 (211/411)
If	CNTRYID=233 (Estonia); ENTUSE≤-0,358; ICTSCH>6,500; ICTHOME≤8,500 → cluster 0 (40/66)
If	CNTRYID=233 (Estonia); ENTUSE≤-0,358; ICTSCH≤6,500; SCHSIZE≤979 → cluster 0 (156/212)
If	CNTRYID=246 (Finland); ENTUSE>-0,191; ICTSCH≤8,500; WEALTH≤-0,950 → cluster 0 (19/24)
If	CNTRYID=246 (Finland); -0,627<ENTUSE≤-0,191; ICTHOME≤9,500 → cluster 0 (203/346)
If	CNTRYID=246 (Finland); ENTUSE≤-0,627; STRATIO>9,078 → cluster 0 (147/193)
If	CNTRYID=250 (France); ICTSCH>5,500; ENTUSE>-0,241; SCHSIZE≤1682,500 → cluster 0 (228/455)
If	CNTRYID=250 (France); ICTSCH>5,500; ENTUSE≤-0,241; ICTRES≤0,552 → cluster 0 (123/185)
If	CNTRYID=250 (France); ICTSCH≤5,500; ICTRES>0,414; ATTLNACT≤-0,006 → cluster 0 (19/28)
If	CNTRYID=250 (France); ICTSCH≤5,500; ICTRES≤0,414 → cluster 0 (294/419)
If	CNTRYID=268 (Georgia); ICTSCH>4,500; ICTHOME>9,500; ICTRES>-0,506 → cluster 0 (66/107).
If	CNTRYID=268 (Georgia); ICTSCH>4,500; ICTHOME≤9,500; ATTLNACT>-0,662 → cluster 0 (301/441)
If	CNTRYID=268 (Georgia); ICTSCH≤4,500 → cluster 0 (349/441)
If	CNTRYID=300 (Greece); ICTSCH>8,500; TEACHINT>0,903 → cluster 0 (46/60)
If	CNTRYID=300 (Greece); ICTSCH>8,500; TEACHINT≤0,903; ENTUSE>0,107 → cluster 0 (105/177)
If	CNTRYID=300 (Greece); ICTSCH≤8,500; EMOSUPS>-1,186 → cluster 0 (1186/1468)
If	CNTRYID=300 (Greece); ICTSCH≤8,500; EMOSUPS≤-1,186; COMPICT>-0,265 → cluster 0 (60/83)
If	CNTRYID=344 (Hong Kong); SCHLTYPE = 2; SCHSIZE>1020; STRATIO>11,505 → cluster 0 (37/55)
If	CNTRYID=344 (Hong Kong); SCHLTYPE = 2; SCHSIZE≤1020 → cluster 0 (1016/1329)
If	CNTRYID=344 (Hong Kong); SCHLTYPE = 3 → cluster 0 (131/161)
If	CNTRYID=348 (Hungary); ICTSCH>7,500; STRATIO>12,424 → cluster 0 (40/73)
If	CNTRYID=348 (Hungary); ICTSCH>7,500; STRATIO≤8,120 → cluster 0 (48/79)
If	CNTRYID=348 (Hungary); ICTSCH≤7,500; ENTUSE>-0,155; WEALTH≤0,528 → cluster 0 (382/611)
If	CNTRYID=348 (Hungary); ICTSCH≤7,500; ENTUSE≤-0,155 → cluster 0 (279/357)
If	CNTRYID=352 (Iceland); ENTUSE≤-1,331 → cluster 0 (30/46)

Table 7 (continued)

If	CNTRYID=36 (Australia); SCHLTYPE = 3; ENTUSE≤-0,843; ATTLNACT≤0,774 → cluster 0 (43/72)
If	CNTRYID=380 (Italy); -0,554<ENTUSE≤0,392; ICTSCH≤6,500 → cluster 0 (491/905)
If	CNTRYID=380 (Italy); ENTUSE≤-0,554 → cluster 0 (243/356)
If	CNTRYID=398 (Kazakhstan); ICTSCH>6,500; ENTUSE≤-3,481 → cluster 0 (32/46)
If	CNTRYID=410 (South Korea); ICTSCH>5,500; ST158Q06HA = 1; INTICT≤-2,150 → cluster 0 (24/27)
If	CNTRYID=410 (South Korea); ICTSCH>5,500; ST158Q06HA = 2 → cluster 0 (414/652)
If	CNTRYID=410 (South Korea); ICTSCH≤5,500 → cluster 0 (626/847)
If	CNTRYID=428 (Latvia); ICTSCH≤6,500; -0,151<ENTUSE≤0,036 → cluster 0 (61/100)
If	CNTRYID=428 (Latvia); ENTUSE≤-0,410; ICTSCH>6,500 → cluster 0 (99/175)
If	CNTRYID=428 (Latvia); ENTUSE≤-0,151; ICTSCH≤6,500 → cluster 0 (194/294)
If	CNTRYID=440 (Lithuania); ICTSCH>7,500; COMPICT≤-1,804 → cluster 0 (22/36)
If	CNTRYID=440 (Lithuania); ICTSCH≤7,500; ENTUSE≤-0,562 → cluster 0 (100/159)
If	CNTRYID=442 (Luxembourg); SCHLTYPE = 2; ENTUSE>-0,186; ICTSCH≤6,500 → cluster 0 (20/31)
If	CNTRYID=442 (Luxembourg); SCHLTYPE = 2; ENTUSE≤-0,186 → cluster 0 (42/55)
If	CNTRYID=442 (Luxembourg); SCHLTYPE = 3; ICTSCH>7,500; EMOSUPS>0,011 → cluster 0 (109/177)
If	CNTRYID=442 (Luxembourg); SCHLTYPE = 3; ICTSCH≤7,500; ENTUSE≤1,228 → cluster 0 (558/729)
If	CNTRYID=470 (Malta); 0,821<ICTRES≤1,341; ICTSCH≤5,500 → cluster 0 (33/54)
If	CNTRYID=470 (Malta); ICTRES≤0,821; ICTSCH>4,500; PAREDINT≤10,500 → cluster 0 (86/133)
If	CNTRYID=470 (Malta); ICTRES≤0,821; ICTSCH≤4,500 → cluster 0 (73/102)
If	CNTRYID=484 (Mexico); ICTHOME>6,500; ENTUSE≤-0,675 → cluster 0 (48/77)
If	CNTRYID=484 (Mexico); ICTHOME≤6,500; ENTUSE>-1,413; WEALTH≤-1,211 → cluster 0 (235/381)
If	CNTRYID=484 (Mexico); ICTHOME≤6,500; ENTUSE≤-1,413 → cluster 0 (75/90)
If	CNTRYID=504 (Morocco); ST158Q03HA = 1; EMOSUPS>0,638 → cluster 0 (22/29)
If	CNTRYID=504 (Morocco); ST158Q03HA = 1; EMOSUPS≤0,638; SCHSIZE>918 → cluster 0 (27/47)
If	CNTRYID=504 (Morocco); ST158Q03HA = 2; ICTSCH≤6,500 → cluster 0 (94/124)
If	CNTRYID=591 (Panama); AUTICT>-1,468; STRATIO>8,855; ST004D01T = 1 → cluster 0 (92/159)
If	CNTRYID=591 (Panama); AUTICT≤-1,468 → cluster 0 (23/29)
If	CNTRYID=616 (Poland); 3,500<ICTSCH≤6,500; ENTUSE≤-0,643 → cluster 0 (42/61)
If	CNTRYID=616 (Poland); ICTSCH≤1,500; COMPICT>0,361 → cluster 0 (14/24)
If	CNTRYID=616 (Poland); ICTSCH≤3,500; COMPICT≤0,361 → cluster 0 (151/248)
If	CNTRYID=643 (Russia); ENTUSE≤-0,530; AUTICT>-1,779; ICTRES≤-1,244 → cluster 0 (35/53)
If	CNTRYID=643 (Russia); ENTUSE≤-0,530; AUTICT≤-1,779 → cluster 0 (24/29)
If	CNTRYID=688 (Serbia); ICTSCH>6,500; COMPICT≤-2,602 → cluster 0 (19/23)
If	CNTRYID=688 (Serbia); ICTSCH≤6,500 → cluster 0 (539/819)
If	CNTRYID=702 (Singapore); SCHSIZE>1577,500; ICTSCH≤4,500 → cluster 0 (15/29)
If	CNTRYID=702 (Singapore); SCHSIZE≤1577,500; ICTSCH>8,500; TEACHINT≤-0,391 → cluster 0 (69/110)
If	CNTRYID=702 (Singapore); SCHSIZE≤1577,500; ICTSCH≤8,500; SCHLTYPE = 3 → cluster 0 (1013/1520)
If	CNTRYID=703 (Slovakia); ICTHOME>8,500; ICTSCH>7,500; INTICT≤-1,967 → cluster 0 (15/23)
If	CNTRYID=703 (Slovakia); ICTHOME>8,500; ICTSCH≤7,500; ATTLNACT≤-0,935 → cluster 0 (78/128)
If	CNTRYID=703 (Slovakia); ICTHOME≤8,500; ST158Q05HA = 1; ATTLNACT≤0,465 → cluster 0 (140/257)
If	CNTRYID=703 (Slovakia); ICTHOME≤8,500; ST158Q05HA = 2; SCHSIZE≤755 → cluster 0 (306/462)
If	CNTRYID=705 (Slovenia); ICTSCH>6,500; ENTUSE≤-0,526 → cluster 0 (106/152)
If	CNTRYID=705 (Slovenia); ICTSCH≤6,500 → cluster 0 (699/994)

Table 7 (continued)

If	CNTRYID=724 (Spain); ICTSCH≤5,500 → cluster 0 (1551/2423)
If	CNTRYID=756 (Switzerland); ICTSCH>5,500; ST158Q07HA = 1; SCHSIZE>249 → cluster 0 (110/200)
If	CNTRYID=756 (Switzerland); ICTSCH>5,500; ST158Q07HA = 2 → cluster 0 (322/504)
If	CNTRYID=756 (Switzerland); ICTSCH≤5,500 → cluster 0 (333/414)
If	CNTRYID=764 (Thailand); ICTHOME>6,500; INTICT≤1,232; RESILIENCE>-1,155 → cluster 0 (46/90)
If	CNTRYID=764 (Thailand); ICTHOME≤6,500; ENTUSE≤-1,223 → cluster 0 (53/81)
If	CNTRYID=792 (Turkey); ENTUSE>-0,633; ICTSCH≤3,500; ST158Q04HA = 2 → cluster 0 (117/225)
If	CNTRYID=792 (Turkey); ENTUSE≤-0,633; ICTSCH≤5,500; SCHSIZE>365 → cluster 0 (164/224)
If	CNTRYID=8 (Albania); ICTSCH>4,500; ICTHOME>8,500; AUTICT≤-0,615 → cluster 0 (33/58)
If	CNTRYID=8 (Albania); ICTSCH>4,500; ICTHOME≤8,500; STRATIO≤7,554 → cluster 0 (24/36)
If	CNTRYID=8 (Albania); ICTSCH≤4,500; ENTUSE>-0,379; STRATIO>9,020 → cluster 0 (60/89)
If	CNTRYID=8 (Albania); ICTSCH≤4,500; ENTUSE≤-0,379 → cluster 0 (53/63)
If	CNTRYID=826 (United Kingdom); ICTSCH>6,500; ENTUSE≤-0,456 → cluster 0 (63/119)
If	CNTRYID=826 (United Kingdom); ICTSCH≤6,500; ATTLNACT>0,030; STRATIO≤17,435 → cluster 0 (74/148)
If	CNTRYID=826 (United Kingdom); ICTSCH≤6,500; ATTLNACT≤0,030 → cluster 0 (100/158)
If	CNTRYID=840 (United States); ENTUSE>-1,292; FISCED = 2; ICTSCH≤5,500 → cluster 0 (15/27)
If	CNTRYID=840 (United States); ENTUSE>-1,292; FISCED = 4; ICTSCH≤3,500 → cluster 0 (19/33)
If	CNTRYID=840 (United States); ENTUSE≤-1,292 → cluster 0 (34/58)
If	CNTRYID=858 (Uruguay); ENTUSE>-0,542; AUTICT>1,294 → cluster 0 (34/67)
If	CNTRYID=858 (Uruguay); ENTUSE≤-0,542; SCHSIZE≤1209 → cluster 0 (84/135)
If	CNTRYID=96 (Brunei); ICTSCH>6,500; SCHSIZE>1227; STRATIO>13,995 → cluster 0 (19/33)
If	CNTRYID=96 (Brunei); ICTSCH>6,500; SCHSIZE≤1227; ICTHOME≤10,500 → cluster 0 (273/507)
If	CNTRYID=96 (Brunei); ICTSCH≤6,500; STRATIO>15,044 → cluster 0 (91/107)
If	CNTRYID=96 (Brunei); ICTSCH≤6,500; STRATIO≤15,044; COMPICT≤-0,206 → cluster 0 (251/373)
If	CNTRYID=982 (Moscow Region - Russia); ENTUSE>-0,481; SCHSIZE≤284 → cluster 0 (18/32)
If	CNTRYID=982 (Moscow Region - Russia); ENTUSE≤-0,481; COMPICT≤-1,036 → cluster 0 (25/37)
If	CNTRYID=983 (Tartaristan - Russia); ENTUSE≤-0,498; INTICT>-2,543; SCHSIZE>1080 → cluster 0 (35/55)
If	CNTRYID=983 (Tartaristan - Russia); ENTUSE≤-0,498; INTICT≤-2,543 → cluster 0 (18/22)

Each classification rule ends with the proportion of students in the training set that fit the rule and belong to the cluster assigned

Table 8 Decision Tree 1 (cluster 3)

If	CNTRYID=100 (Bulgary); ICTSCH>7,500; -1.039<ENTUSE≤1,327 → cluster 3 (278/340)
If	CNTRYID=100 (Bulgary); ICTSCH≤7,500; TEACHINT>-0,878; ICTHOME>5,500 → cluster 3 (208/406)
If	CNTRYID=100 (Bulgary); ICTSCH≤7,500; TEACHINT≤-0,878; SCHSIZE≤1064 → cluster 3 (63/82)
If	CNTRYID=152 (Chile); ENTUSE>-0,309; AUTICT>-0,780 → cluster 3 (587/805)
If	CNTRYID=152 (Chile); ENTUSE>-0,309; AUTICT≤-0,780; RESILIENCE>0,373 → cluster 3 (28/45)
If	CNTRYID=152 (Chile); -1,066<ENTUSE≤-0,309; TEACHINT>-0,319 → cluster 3 (112/203)
If	CNTRYID=188 (Costa Rica); ENTUSE>-0,588; ICTHOME>7,500 → cluster 3 (489/682)
If	CNTRYID=188 (Costa Rica); ENTUSE>-0,558; ICTHOME≤7,500; AUTICT>-0,374 → cluster 3 (255/430)
If	CNTRYID=191 (Croatia); ENTUSE>-0,229; ICTSCH>9,500 → cluster 3 (94/137)
If	CNTRYID=191 (Croatia); ENTUSE≤-0,229; ICTRES>-0,329; EMOSUPS≤-0,785 → cluster 3 (22/34)
If	CNTRYID=203 (Czech Republic); ENTUSE>-0,010; ICTSCH>8,500 → cluster 3 (69/106)
If	CNTRYID=203 (Czech Republic); ENTUSE>-0,010; ICTSCH≤8,500; WEALTH>1,244 → cluster 3 (22/32)
If	CNTRYID=203 (Czech Republic); ENTUSE≤-0,010; STRATIO≤7,433 → cluster 3 (22/33)
If	CNTRYID=208 (Denmark) → cluster 3 (1980/1994)
If	CNTRYID=214 (Dominican Republic); ICTSCH>5,500; STRATIO>13,764; PAREDINT>13,250 → cluster 3 (73/111)
If	CNTRYID=214 (Dominican Republic); ICTSCH>5,500; STRATIO≤13,764 → cluster 3 (32/39)
If	CNTRYID=214 (Dominican Republic); ICTSCH≤5,500; 0,265<ENTUSE≤0,883 → cluster 3 (20/37)
If	CNTRYID=233 (Estonia); ENTUSE>-0,358; ICTSCH>7,500 → cluster 3 (361/509)
If	CNTRYID=233 (Estonia); ICTSCH≤7,500; ENTUSE>0,030 → cluster 3 (330/531)
If	CNTRYID=233 (Estonia); ENTUSE≤-0,358; ICTSCH>6,500; ICTHOME>8,500 → cluster 3 (58/96)
If	CNTRYID=233 (Estonia); ENTUSE≤-0,358; ICTSCH≤6,500; SCHSIZE>979 → cluster 3 (13/25)
If	CNTRYID=246 (Finland); ENTUSE>-0,191; ICTSCH>8,500 → cluster 3 (258/332)
If	CNTRYID=246 (Finland); ENTUSE>-0,191; ICTSCH≤8,500; WEALTH>-0,950 → cluster 3 (388/667)
If	CNTRYID=246 (Finland); -0,627<ENTUSE≤-0,191; ICTHOME>9,500 → cluster 3 (63/95)
If	CNTRYID=246 (Finland); ENTUSE≤-0,627; STRATIO≤9,078 → cluster 3 (17/29)
If	CNTRYID=250 (France); ICTSCH>5,500; ENTUSE>-0,241; SCHSIZE>1682,500 → cluster 3 (39/52)
If	CNTRYID=250 (France); ICTSCH>5,500; ENTUSE≤-0,241; ICTRES>0,552 → cluster 3 (18/30)
If	CNTRYID=250 (France); ICTSCH≤5,500; ICTRES>0,414; ATTLNACT>-0,006 → cluster 3 (41/70)
If	CNTRYID=268 (Georgia); ICTSCH>4,500; ICTHOME>9,500; ICTRES≤-0,506 → cluster 3 (78/123)
If	CNTRYID=268 (Georgia); ICTSCH>4,500; ICTHOME≤9,500; ATTLNACT≤-0,662 → cluster 3 (43/83)
If	CNTRYID=300 (Greece); ICTSCH>8,500; TEACHINT≤0,903; ENTUSE≤0,107 → cluster 3 (125/212)
If	CNTRYID=300 (Greece); ICTSCH≤8,500; EMOSUPS≤-1,186; COMPICT≤-0,265 → cluster 3 (51/96)
If	CNTRYID=344 (Hong Kong); SCHLTYPE = 1 → cluster 3 (44/44)
If	CNTRYID=344 (Hong Kong); SCHLTYPE = 2; SCHSIZE>1020; STRATIO≤11,505 → cluster 3 (30/33)
If	CNTRYID=348 (Hungary); ICTSCH>7,500; 8,120<STRATIO≤12,424 → cluster 3 (156/ 238)
If	CNTRYID=348 (Hungary); ICTSCH≤7,500; ENTUSE>-0,155; WEALTH>0,528 → cluster 3 (50/90)
If	CNTRYID=352 (Iceland); ENTUSE>-1,331 → cluster 3 (606/834)
If	CNTRYID=36 (Australia); SCHLTYPE = 1 → cluster 3 (637/658)
If	CNTRYID=36 (Australia); SCHLTYPE = 2 → cluster 3 (833/892)
If	CNTRYID=36 (Australia); SCHLTYPE = 3; ENTUSE>-0,843 → cluster 3 (1398/1714)
If	CNTRYID=36 (Australia); SCHLTYPE = 3; ENTUSE≤-0,843; ATTLNACT>0,774 → cluster 3 (39/56)
If	CNTRYID=380 (Italy); ENTUSE>-0,554; ICTSCH>6,500 → cluster 3 (630/943)
If	CNTRYID=380 (Italy); ENTUSE>0,392; ICTSCH≤6,500 → cluster 3 (276/483)
If	CNTRYID=398 (Kazakhstan); ICTSCH>6,500; ENTUSE>-1,092 → cluster 3 (3504/4219)

Table 8 (continued)

If	CNTRYID=398 (Kazakhstan); ICTSCH>6,500; -3,481<ENTUSE≤-1,092 → cluster 3 (198/292)
If	CNTRYID=398 (Kazakhstan); ICTSCH≤6,500 → cluster 3 (1277/1935)
If	CNTRYID=410 (South Korea); ICTSCH>5,500; ST158Q06HA = 1; INTICT>-2,150 → cluster 3 (419/815)
If	CNTRYID=428 (Latvia); ENTUSE>-0,151; ICTSCH>6,500 → cluster 3 (320/458)
If	CNTRYID=428 (Latvia); ENTUSE>0,036; ICTSCH≤6,500 → cluster 3 (183/298)
If	CNTRYID=428 (Latvia); -0,410<ENTUSE≤-0,151; ICTSCH>6,500 → cluster 3 (83/128)
If	CNTRYID=440 (Lithuania); ICTSCH>7,500; COMPICT>-1,804 → cluster 3 (667/864)
If	CNTRYID=440 (Lithuania); ICTSCH≤7,500; ENTUSE>-0,562 → cluster 3 (586/928)
If	CNTRYID=442 (Luxembourg); SCHLTYPE = 1 → cluster 3 (65/67)
If	CNTRYID=442 (Luxembourg); SCHLTYPE = 2; ENTUSE>-0,186; ICTSCH>6,500 → cluster 3 (40/61)
If	CNTRYID=442 (Luxembourg); SCHLTYPE = 3; ICTSCH>7,500; EMOSUPS≤0,011 → cluster 3 (99/172)
If	CNTRYID=442 (Luxembourg); SCHLTYPE = 3; ICTSCH≤7,500; ENTUSE>1,228 → cluster 3 (25/44)
If	CNTRYID=470 (Malta); ICTRES>0,821; ICTSCH>5,500 → cluster 3 (93/142)
If	CNTRYID=470 (Malta); ICTRES>1,341; ICTSCH≤5,500 → cluster 3 (16/24)
If	CNTRYID=470 (Malta); ICTRES≤0,821; ICTSCH>4,500; PAREDINT>10,500 → cluster 3 (225/443)
If	CNTRYID=484 (Mexico); ICTHOME>6,500; ENTUSE>-0,675 → cluster 3 (405/591)
If	CNTRYID=484 (Mexico); ICTHOME≤6,500; ENTUSE>-1,413; WEALTH>-1,211 → cluster 3 (67/123)
If	CNTRYID=504 (Morocco); ST158Q03HA = 1; EMOSUPS≤0,638; SCHSIZE≤918 → cluster 3 (45/62)
If	CNTRYID=504 (Morocco); ST158Q03HA = 2; ICTSCH>6,500 → cluster 3 (20/36)
If	CNTRYID=591 (Panama); AUTICT>-1,468; STRATIO>8,855; ST004D01T = 2 → cluster 3 (90/157)
If	CNTRYID=591 (Panama); AUTICT>-1,468; STRATIO≤8,855 → cluster 3 (25/34)
If	CNTRYID=616 (Poland); ICTSCH>3,500; ENTUSE>-0,643 → cluster 3 (806/1153)
If	CNTRYID=616 (Poland); ICTSCH>6,500; ENTUSE≤-0,643 → cluster 3 (36/59)
If	CNTRYID=616 (Poland); 1,500<ICTSCH≤3,500; COMPICT>0,361 → cluster 3 (34/46)
If	CNTRYID=643 (Russia); ENTUSE>-0,530 → cluster 3 (1438/1889)
If	CNTRYID=643 (Russia); ENTUSE≤-0,530; AUTICT>-1,779; ICTRES>-1,244 → cluster 3 (160/277)
If	CNTRYID=688 (Serbia); ICTSCH>6,500; COMPICT>-2,602 → cluster 3 (307/519)
If	CNTRYID=702 (Singapore); SCHSIZE>1577,500; ICTSCH>4,500 → cluster 3 (247/287)
If	CNTRYID=702 (Singapore); SCHSIZE≤1577,500; ICTSCH>8,500; TEACHINT>-0,391 → cluster 3 (214/367)
If	CNTRYID=702 (Singapore); SCHSIZE≤1577,500; ICTSCH≤8,500; SCHLTYPE = 1 → cluster 3 (41/57)
If	CNTRYID=703 (Slovakia); ICTHOME>8,500; ICTSCH>7,500; INTICT>-1,967 → cluster 3 (338/451)
If	CNTRYID=703 (Slovakia); ICTHOME>8,500; ICTSCH≤7,500; ATTLNACT>-0,935 → cluster 3 (209/353)
If	CNTRYID=703 (Slovakia); ICTHOME≤8,500; ST158Q05HA = 1; ATTLNACT>0,465 → cluster 3 (53/78)
If	CNTRYID=703 (Slovakia); ICTHOME≤8,500; ST158Q05HA = 2; SCHSIZE>755 → cluster 3 (18/27)
If	CNTRYID=705 (Slovenia); ICTSCH>6,500; ENTUSE>-0,526 → cluster 3 (267/465)
If	CNTRYID=724 (Spain); ICTSCH>5,500; ENTUSE>-0,471 → cluster 3 (2269/3943)
If	CNTRYID=724 (Spain); ICTSCH>5,500; ENTUSE≤-0,471; ICTRES>1,804 → cluster 3 (20/27)
If	CNTRYID=756 (Switzerland); ICTSCH>5,500; ST158Q07HA = 1; SCHSIZE≤249 → cluster 3 (55/82)
If	CNTRYID=76 (Brazil); ICTHOME>7,500; ICTSCH>6,500 → cluster 3 (169/281)
If	CNTRYID=76 (Brazil); ICTHOME>7,500; ICTSCH≤6,500; ICTRES>0,040 → cluster 3 (45/86)
If	CNTRYID=764 (Thailand); ICTHOME>6,500; INTICT>-1,232 → cluster 3 (1242/1541)
If	CNTRYID=764 (Thailand); ICTHOME>6,500; INTICT≤-1,232; RESILIENCE≤-1,155 → cluster 3 (24/28)
If	CNTRYID=764 (Thailand); ICTHOME≤6,500; -1,223<ENTUSE≤-0,177 → cluster 3 (197/368)
If	CNTRYID=792 (Turkey); ENTUSE>-0,633; ICTSCH>3,500 → cluster 3 (1182/1575)

Table 8 (continued)

If	CNTRYID=792 (Turkey); ENTUSE>-0,633; ICTSCH≤3,500; ST158Q04HA = 1 → cluster 3 (114/170)
If	CNTRYID=792 (Turkey); ENTUSE≤-0,633; ICTSCH>5,500 → cluster 3 (147/233)
If	CNTRYID=792 (Turkey); ENTUSE≤-0,633; ICTSCH≤5,500; SCHSIZE≤365 → cluster 3 (32/60)
If	CNTRYID=8 (Albania); ICTSCH>4,500; ICTHOME>8,500; AUTICT>-0,615 → cluster 3 (227/296)
If	CNTRYID=8 (Albania); ICTSCH>4,500; ICTHOME≤8,500; STRATIO>7,554 → cluster 3 (151/255)
If	CNTRYID=8 (Albania); ICTSCH≤4,500; ENTUSE>-0,379; STRATIO≤9,020 → cluster 3 (17/25)
If	CNTRYID=826 (United Kingdom); ICTSCH>6,500; ENTUSE>-0,456 → cluster 3 (537/781)
If	CNTRYID=826 (United Kingdom); ICTSCH≤6,500; ATTLNACT>0,030; STRATIO>17,435 → cluster 3 (30/39)
If	CNTRYID=840 (United States); ENTUSE>-1,292; FISCED = 0 → cluster 3 (16/26)
If	CNTRYID=840 (United States); ENTUSE>-1,292; FISCED = 1 → cluster 3 (24/31)
If	CNTRYID=840 (United States); ENTUSE>-1,292; FISCED = 2; ICTSCH>5,500 → cluster 3 (82/104)
If	CNTRYID=840 (United States); ENTUSE>-1,292; FISCED = 4; ICTSCH>3,500 → cluster 3 (411/527)
If	CNTRYID=840 (United States); ENTUSE>-1,292; FISCED = 5 → cluster 3 (135/165)
If	CNTRYID=840 (United States); ENTUSE>-1,292; FISCED = 6 → cluster 3 (453/522)
If	CNTRYID=858 (Uruguay); ENTUSE>-0,542; AUTICT≤1,294 → cluster 3 (357/515)
If	CNTRYID=858 (Uruguay); ENTUSE≤-0,542; SCHSIZE>1209 → cluster 3 (15/22)
If	CNTRYID=96 (Brunei); ICTSCH>6,500; SCHSIZE>1227; STRATIO≤13,955 → cluster 3 (142/166)
If	CNTRYID=96 (Brunei); ICTSCH>6,500; SCHSIZE≤1227; ICTHOME>10,500 → cluster 3 (92/126)
If	CNTRYID=96 (Brunei); ICTSCH≤6,500; STRATIO≤15,044; COMPICT>-0,206 → cluster 3 (146/287)
If	CNTRYID=982 (Moscow Region - Russia); ENTUSE>-0,481; SCHSIZE>284 → cluster 3 (391/528)
If	CNTRYID=982 (Moscow Region - Russia); ENTUSE≤-0,481; COMPICT>-1,036 → cluster 3 (49/78)
If	CNTRYID=983 (Tartaristan - Russia); ENTUSE>-0,498 → cluster 3 (1217/1588)
If	CNTRYID=983 (Tartaristan - Russia); ENTUSE≤-0,498; INTICT>-2,543; SCHSIZE≤1080 → cluster 3 (169/270)

Each classification rule ends with the proportion of students in the training set that fit the rule and belong to the cluster assigned

Table 9 Decision Tree 2 (cluster 0)

If	CNTRYID=152 (Chile); -0,904<ENTUSE≤-0,292; TEACHINT>-0,319; ICTHOME≤8,500 → cluster 0 (71/139)
If	CNTRYID=152 (Chile); -0,904<ENTUSE≤-0,292; TEACHINT≤-0,319 → cluster 0 (42/60)
If	CNTRYID=152 (Chile); ENTUSE≤-0,904 → cluster 0 (85/110)
If	CNTRYID=158 (Taipei); ICTSCH>5,500; ENTUSE>-0,539; mode(OTT2) = 1 → cluster 0 (671/1024)
If	CNTRYID=158 (Taipei); ICTSCH>5,500; ENTUSE≤-0,539 → cluster 0 (202/247)
If	CNTRYID=158 (Taipei); ICTSCH≤5,500 → cluster 0 (971/1179)
If	CNTRYID=214 (Dominican Republic); ICTSCH>5,500; average (TC002Q01NA)≤41,643 → cluster 0 (53/106)
If	CNTRYID=214 (Dominican Republic); ICTSCH≤5,500; ATTLNACT>0,776 → cluster 0 (99/130)
If	CNTRYID=214 (Dominican Republic); ICTSCH≤5,500; ATTLNACT≤0,776; %SATTEACHrange3≤0,760 → cluster 0 (40/56)
If	CNTRYID=344 (Hong Kong); SCHLTYPE = 2; average(TC185Q05HA)>2,555 → cluster 0 (888/1140)
If	CNTRYID=344 (Hong Kong); SCHLTYPE = 2; average(TC185Q05HA)≤2,555; SCHSIZE≤847,500 → cluster 0 (143/205)
If	CNTRYID=344 (Hong Kong); SCHLTYPE = 3 → cluster 0 (116/150)
If	CNTRYID=410 (South Korea); ICTSCH>5,500; ST158Q03HA = 1; RESILIENCE>0,352; ICTRES≤-0,898 → cluster 0 (28/47)
If	CNTRYID=410 (South Korea); ICTSCH>5,500; ST158Q03HA = 1; RESILIENCE≤0,352 → cluster 0 (313/577)
If	CNTRYID=410 (South Korea); ICTSCH>5,500; ST158Q03HA = 2; ICTHOME>6,500; ST158Q04HA = 2 → cluster 0 (249/381)
If	CNTRYID=410 (South Korea); ICTSCH>5,500; ST158Q03HA = 2; ICTHOME≤6,500 → cluster 0 (122/159)
If	CNTRYID=410 (South Korea); 3,500<ICTSCH≤5,500; AUTICT≤0,856 → cluster 0 (285/404)
If	CNTRYID=410 (South Korea); ICTSCH≤3,500 → cluster 0 (298/375)
If	CNTRYID=504 (Morocco); ST158Q03HA = 2; ICTSCH≤6,500 → cluster 0 (102/135)
If	CNTRYID=724 (Spain); ICTSCH>4,500; ENTUSE>-0,471; SCHLTYPE = 1; %TCICTUSErange1>0,087 → cluster 0 (42/54)
If	CNTRYID=724 (Spain); ICTSCH>4,500; ENTUSE≤-0,471 → cluster 0 (553/889)
If	CNTRYID=724 (Spain); ICTSCH≤4,500 → cluster 0 (935/1344)
If	CNTRYID=76 (Brazil); ICTHOME>7,500; ICTSCH>7,500; %SATTEACHrange1>0,061 → cluster 0 (50/95)
If	CNTRYID=76 (Brazil); ICTHOME>7,500; ICTSCH≤7,500; %SEFFRELrange3>0,679 → cluster 0 (112/168)
If	CNTRYID=76 (Brazil); ICTHOME>7,500; ICTSCH≤7,500; %SEFFRELrange3≤0,679; average(TC018Q04NA)>0,257 → cluster 0 (32/50)
If	CNTRYID=76 (Brazil); ICTHOME≤7,500; ENTUSE>-0,386; ICTSCH>4,500; RESILIENCE≤0,113 → cluster 0 (99/149)
If	CNTRYID=76 (Brazil); ICTHOME≤7,500; ENTUSE>-0,386; ICTSCH≤4,500 → cluster 0 (213/281)
If	CNTRYID=76 (Brazil); ICTHOME≤7,500; ENTUSE≤-0,386 → cluster 0 (168/199)
If	CNTRYID=826 (United Kingdom); STRATIO≤15,235; ST158Q01HA = 1; ENTUSE≤0,419; EMO-SUPS≤-0,651 → cluster 0 (40/64)
If	CNTRYID=826 (United Kingdom); STRATIO≤15,235; ST158Q01HA = 2; AUTICT>-0,262; RESIL-IENCE>-0,426 → cluster 0 (27/39)
If	CNTRYID=826 (United Kingdom); STRATIO≤15,235; ST158Q01HA = 2; AUTICT≤-0,262 → cluster 0 (40/51)
If	CNTRYID=840 (United States); %TCICTUSErange1>0,138; SCHSIZE≤1382; %SATJOBrange2≤0,414 → cluster 0 (24/36)

Each classification rule ends with the proportion of students in the training set that fit the rule and belong to the cluster assigned

Table 10 Decision Tree 2 (cluster 3)

If	CNTRYID=152 (Chile); ENTUSE>-0,292 → cluster 3 (603/883)
If	CNTRYID=152 (Chile); -0,904<ENTUSE≤-0,292; TEACHINT>-0,319; ICTHOME>8,500 → cluster 3 (39/53)
If	CNTRYID=158 (Taipei); ICTSCH>5,500; ENTUSE>-0,539; mode(OTT2) = 0 → cluster 3 (50/78)
If	CNTRYID=214 (Dominican Republic); ICTSCH>5,500; average(TC002Q01NA)>41,643 → cluster 3 (64/87)
If	CNTRYID=214 (Dominican Republic); ICTSCH≤5,500; ATTLNACT≤0,776; %SATTEACHrange3>0,760 → cluster 3 (26/48)
If	CNTRYID=344 (Hong Kong); SCHLTYPE = 1 → cluster 3 (40/40)
If	CNTRYID=344 (Hong Kong); SCHLTYPE = 2; average(TC185Q05HA)≤2,555; SCHSIZE>847,500 → cluster 3 (65/79)
If	CNTRYID=410 (South Korea); ICTSCH>5,500; ST158Q03HA = 1; RESILIENCE>0,352; ICTRES>-0,898 → cluster 3 (149/232)
If	CNTRYID=410 (South Korea); ICTSCH>5,500; ST158Q03HA = 2; ICTHOME>6,500; ST158Q04HA = 1 → cluster 3 (55/106)
If	CNTRYID=410 (South Korea); 3,500<ICTSCH≤5,500; AUTICT>0,856 → cluster 3 (25/48)
If	CNTRYID=504 (Morocco); ST158Q03HA = 1 → cluster 3 (79/138)
If	CNTRYID=504 (Morocco); ST158Q03HA = 2; ICTSCH>6,500 → cluster 3 (26/44)
If	CNTRYID=724 (Spain); ICTSCH>4,500; ENTUSE>-0,471; SCHLTYPE = 1; %TCICTUSErange1≤0,087 → cluster 3 (236/321)
If	CNTRYID=724 (Spain); ICTSCH>4,500 e ENTUSE>-0,471 e SCHLTYPE = 2 → cluster 3 (935/521)
If	CNTRYID=724 (Spain); ICTSCH>4,500; ENTUSE>-0,471; SCHLTYPE = 3 → cluster 3 (1434/2825)
If	CNTRYID=76 (Brazil); ICTHOME>7,500; ICTSCH>7,500; %SATTEACHrange1≤0,061 → cluster 3 (66/90)
If	CNTRYID=76 (Brazil); ICTHOME>7,500; ICTSCH≤7,500; %SEFFRELrange3≤0,679; average(TC018Q04NA)≤0,257 → cluster 3 (37/59)
If	CNTRYID=76 (Brazil); ICTHOME≤7,500; ENTUSE>-0,386; ICTSCH>4,500; RESILIENCE>0,113 → cluster 3 (37/48)
If	CNTRYID=826 (United Kingdom); STRATIO>15,235 → cluster 3 (37/41)
If	CNTRYID=826 (United Kingdom); STRATIO≤15,235; ST158Q01HA = 1; ENTUSE>0,419 → cluster 3 (52/67)
If	CNTRYID=826 (United Kingdom); STRATIO≤15,235; ST158Q01HA = 1; ENTUSE≤0,419; EMOSUPS>-0,651 → cluster 3 (90/145)
If	CNTRYID=826 (United Kingdom); STRATIO≤15,235; ST158Q01HA = 2; AUTICT>-0,262; RESILIENCE≤-0,426 → cluster 3 (25/38)
If	CNTRYID=840 (United States); %TCICTUSErange1>0,138; SCHSIZE>1382 → cluster 3 (95/128)
If	CNTRYID=840 (United States); %TCICTUSErange1>0,138; SCHSIZE≤1382; %SATJOBrange2>0,414 → cluster 3 (40/65)
If	CNTRYID=840 (United States); %TCICTUSErange1≤0,138 → cluster 3 (981/1225)

Each classification rule ends with the proportion of students in the training set that fit the rule and belong to the cluster assigned

Author contribution All authors contributed to the study conception and design. Material preparation and data collection were performed by João C. Silva. Data analysis and discussion were led by João C. Silva with strong participation of Vera L. Miguéis and José Coelho Rodrigues. The first draft of the manuscript was written by João C. Silva and all authors commented on the previous versions and contributed to the following versions of the manuscript. All authors read and approved the final manuscript.

Funding Open access funding provided by FCT|FCCN (b-on). This work has been partially financed by National Funds through the Portuguese funding agency, FCT – Fundação para a Ciência e a Tecnologia, within project UIDB/50014/2020.

Data availability The datasets analysed during the current study are available in the OECD PISA repository (Programme for International Student Assessment), https://www.oecd.org/pisa/data/

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

References

Al-Mamary, Y. H. S. (2020). Examining the factors affecting the use of ICT in teaching in Yemeni schools. *Journal of Public Affairs, 22*(1), e2330. https://doi.org/10.1002/pa.2330

Areepattamannil, S., & Santos, I. M. (2019). Adolescent students' perceived information and communication technology (ICT) competence and autonomy: Examining links to dispositions toward science in 42 countries. *Computers in Human Behavior, 98*, 50–58. https://doi.org/10.1016/j.chb.2019.04.005

Basak, S. K. (2014, 10–12 Nov. 2014). *A Comparison of Barriers and Enhance Factors on the Adoption and Use of ICT into Teaching and Learning for Teachers*. International Conference on Information Society (i-Society 2014).

Biagi, F., & Loi, M. (2013). Measuring ICT Use and Learning Outcomes: Evidence from recent econometric studies. *European Journal of Education, 48*(1), 28–42. https://doi.org/10.1111/ejed.12016

Bielefeldt, T. (2005). Computers and Student Learnings. *Journal of Research on Technology in Education, 37*(4), 339–347. https://doi.org/10.1080/15391523.2005.10782441

Buabeng-Andoh, C., Yaokumah, W., & Tarhini, A. (2018). Investigating students' intentions to use ICT: A comparison of theoretical models. *Education and Information Technologies, 24*(1), 643–660. https://doi.org/10.1007/s10639-018-9796-1

Chen, X., & Hu, J. (2020). ICT-related behavioral factors mediate the relationship between adolescents' ICT interest and their ICT self-efficacy: Evidence from 30 countries. *Computers & Education, 159*, 104004. https://doi.org/10.1016/j.compedu.2020.104004

Chou, C. C., & Block, L. (2018). The Mismatched Expectations of iPad Integration Between Teachers and Students in Secondary Schools. *Journal of Educational Computing Research, 57*(5), 1281–1302. https://doi.org/10.1177/0735633118784720

Clark, S., & Lee, L. (2019). Technology Enhanced Classroom for Low-Income Children's Mathematical Content Learning: A Case Study. *International Journal of Information and Education Technology, 9*(1), 66–69. https://doi.org/10.18178/ijiet.2019.9.1.1175

Cnedu. (2019). Estado da Educação 2018. *Conselho Nacional da Educação*. https://www.cnedu.pt/content/edicoes/estado_da_educacao/Estado_da_Educacao2018_web_26nov2019.pdf

Coovadia, H., & Ackermann, C. (2020). Integrating digital pedagogies into a typical student learning lifecycle and its effect on exam performance. *Accounting Education, 30*(1), 42–62. https://doi.org/10.1080/09639284.2020.1811993

Eickelmann, B., Gerick, J., & Koop, C. (2016). ICT use in mathematics lessons and the mathematics achievement of secondary school students by international comparison: Which role do school level factors play? *Education and Information Technologies, 22*(4), 1527–1551. https://doi.org/10.1007/s10639-016-9498-5

Fareri, S., Fantoni, G., Chiarello, F., Coli, E., & Binda, A. (2020). Estimating Industry 4.0 impact on job profiles and skills using text mining. *Computers in Industry, 118*, 103222. https://doi.org/10.1016/j.compind.2020.103222

Ferraro, S. (2018). Is information and communication technology satisfying educational needs at school? *Computers & Education, 122*, 194–204. https://doi.org/10.1016/j.compedu.2018.04.002
Fullan, M., & Langworthy, M. (2013). Towards a New End: New Pedagogies for Deep Learning. *Seattle: Collaborative Impact*. http://www.michaelfullan.ca/wp-content/uploads/2013/08/New-Pedagogies-for-Deep-Learning-An-Invitation-to-Partner-2013-6-201.pdf
Gillis, A., & Krull, L. M. (2020). COVID-19 Remote Learning Transition in Spring 2020: Class Structures, Student Perceptions, and Inequality in College Courses. *Teaching Sociology, 48*(4), 283–299. https://doi.org/10.1177/0092055x20954263
Hanushek, E. A., & Woessmann, L. (2008). The Role of Cognitive Skills in Economic Development. *Journal of Economic Literature, 46*(3), 607–668. https://doi.org/10.1257/jel.46.3.607
Harper, B., & Milman, N. B. (2016). One-to-One Technology in K–12 Classrooms: A Review of the Literature From 2004 Through 2014. *Journal of Research on Technology in Education, 48*(2), 129–142. https://doi.org/10.1080/15391523.2016.1146564
Jamil, A., Ahsan, M., Farooq, T., Hussain, A., & Ashraf, R. (2018, 6–8 Sept. 2018). *Student Performance Prediction Using Algorithms of Data Mining* 2018 International Conference on Computing, Engineering, and Design (ICCED).
Juggernath, A., & Govender, N. (2020). Natural Sciences Teachers' Beliefs as Barriers for Integrating ICTs in a Technology-rich Context. *African Journal of Research in Mathematics, Science and Technology Education, 24*(1), 105–115. https://doi.org/10.1080/18117295.2020.1736854
Juhaňák, L., Zounek, J., Záleská, K., Bárta, O., & Vlčková, K. (2019). The relationship between the age at first computer use and students' perceived competence and autonomy in ICT usage: A mediation analysis. *Computers & Education, 141*, 103614. https://doi.org/10.1016/j.compedu.2019.103614
Kubiatko, M., & Vlckova, K. (2010). THE RELATIONSHIP BETWEEN ICT USE AND SCIENCE KNOWLEDGE FOR CZECH STUDENTS: A SECONDARY ANALYSIS OF PISA 2006. *International Journal of Science and Mathematics Education, 8*(3), 523–543. https://doi.org/10.1007/s10763-010-9195-6
Li, S., Liu, X., Tripp, J., & Yang, Y. (2020). From ICT availability to student science achievement: Mediation effects of ICT psychological need satisfactions and interest across genders. *Research in Science & Technological Education, 40*(4), 529–548. https://doi.org/10.1080/02635143.2020.1830269
Meng, L., Qiu, C., & Boyd-Wilson, B. (2019). Measurement invariance of the ICT engagement construct and its association with students' performance in China and Germany: Evidence from PISA 2015 data. *British Journal of Educational Technology, 50*(6), 3233–3251. https://doi.org/10.1111/bjet.12729
Mirzajani, H., Mahmud, R., Ayub, A. F. M., & Wong, S. L. (2016). Teachers' acceptance of ICT and its integration in the classroom. *Quality Assurance in Education, 24*(1), 26–40. https://doi.org/10.1108/qae-06-2014-0025
Odell, B., Cutumisu, M., & Gierl, M. (2020). A scoping review of the relationship between students' ICT and performance in mathematics and science in the PISA data. *Social Psychology of Education, 23*(6), 1449–1481. https://doi.org/10.1007/s11218-020-09591-x
Önder, E., & Uyar, Ş. (2017). CHAID Analysis to Determine Socioeconomic Variables that Explain Students' Academic Success. *Universal Journal of Educational Research, 5*(4), 608–619. https://doi.org/10.13189/ujer.2017.050410
Petko, D., Cantieni, A., & Prasse, D. (2016). Perceived Quality of Educational Technology Matters. *Journal of Educational Computing Research, 54*(8), 1070–1091. https://doi.org/10.1177/0735633116649373
Prieto, J. C. S., Migueláñez, S. O., & García-Peñalvo, F. J. (2014). *ICTs integration in education* Proceedings of the Second International Conference on Technological Ecosystems for Enhancing Multiculturality - TEEM '14, Salamanca, Spain. https://doi.org/10.1145/2669711.2669974
Salinas, Á., Nussbaum, M., Herrera, O., Solarte, M., & Aldunate, R. (2016). Factors affecting the adoption of information and communication technologies in teaching. *Education and Information Technologies, 22*(5), 2175–2196. https://doi.org/10.1007/s10639-016-9540-7
Salloum, S. A., Al-Emran, M., & Shaalan, K. (2017). Mining Social Media Text: Extracting Knowledge from Facebook. *International Journal of Computing and Digital Systemss, 6*(2), 73–81. https://doi.org/10.12785/ijcds/060203
Shim, T. E., & Lee, S. Y. (2020). College students' experience of emergency remote teaching due to COVID-19. *Child Youth Serv Rev, 119*, 105578. https://doi.org/10.1016/j.childyouth.2020.105578
Skryabin, M., Zhang, J., Liu, L., & Zhang, D. (2015). How the ICT development level and usage influence student achievement in reading, mathematics, and science. *Computers & Education, 85*, 49–58. https://doi.org/10.1016/j.compedu.2015.02.004

Spiteri, M., & Chang Rundgren, S.-N. (2018). Literature Review on the Factors Affecting Primary Teachers' Use of Digital Technology. *Technology, Knowledge and Learning, 25*(1), 115–128. https://doi.org/10.1007/s10758-018-9376-x

Srijamdee, K., & Pholphirul, P. (2020). Does ICT familiarity always help promote educational outcomes? Empirical evidence from PISA-Thailand. *Education and Information Technologies, 25*(4), 2933–2970. https://doi.org/10.1007/s10639-019-10089-z

Wang, W., & Han, X. (2018, 12–14 Dec. 2018). *Some Factors Influencing Teachers' Intention to use ICT in Teaching in Secondary Vocational Schools in China.* 2018 Seventh International Conference of Educational Innovation through Technology (EITT).

Wei, X., Cheng, I. L., Chen, N.-S., Yang, X., Liu, Y., Dong, Y., Zhai, X., & Kinshuk. (2020). Effect of the flipped classroom on the mathematics performance of middle school students. *Educational Technology Research and Development, 68*(3), 1461–1484. https://doi.org/10.1007/s11423-020-09752-x

Xiao, Y., & Hu, J. (2019a, 19–21 Aug. 2019a). *The influence of ICT attitudes on closing the reading literacy gap of students from different economic, social and cultural backgrounds*. 2019a 14th International Conference on Computer Science & Education (ICCSE).

Xiao, Y., & Hu, J. (2019b). The Moderation Examination of ICT Use on the Association Between Chinese Mainland Students' Socioeconomic Status and Reading Achievement. *International Journal of Emerging Technologies in Learning (iJET), 14*(15). https://doi.org/10.3991/ijet.v14i15.10494

Zhai, X., Zhang, M., Li, M., & Zhang, X. (2019). Understanding the relationship between levels of mobile technology use in high school physics classrooms and the learning outcome. *British Journal of Educational Technology, 50*(2), 750–766. https://doi.org/10.1111/bjet.12700

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

João C. Silva[1] · José Coelho Rodrigues[2] · Vera L. Miguéis[2]

✉ José Coelho Rodrigues
jpcr@fe.up.pt

João C. Silva
jpedroscsilva@gmail.com

Vera L. Miguéis
vera.migueis@fe.up.pt

[1] School of Economics of the University of Porto, Rua Dr. Roberto Frias, S/N, 4200-464 Porto, Portugal

[2] Faculty of Engineering of the University of Porto, INESC TEC, Rua Dr. Roberto Frias, S/N, 4200-465 Porto, Portugal