Beatriz-Afonso, A., Cruz-Jesus, F., Nunes, C., Castelli, M., Oliveira, T. & Canto e Castro, L., (2025). Drivers of academic achievement in high school: Assessing the impact of COVID-19 using machine learning techniques. *Educational Technology & Society*, *28*(2), 148-168. https://doi.org/10.30191/ETS.202504_28(2).RP09

Drivers of academic achievement in high school: Assessing the impact of COVID-19 using machine learning techniques

Ana Beatriz-Afonso[1*], Frederico Cruz-Jesus[1], Catarina Nunes[1], Mauro Castelli[1], Tiago Oliveira[1] and Luísa Canto e Castro[2]

[1]NOVA Information Management School (NOVA IMS), Universidade NOVA de Lisboa, Campus de Campolide, 1070–312 Lisboa, Portugal // [2]Faculdade de Ciências, Centro de Estatística e Aplicações e Fundação Francisco Manuel dos Santos, Universidade de Lisboa, Lisbon, Portugal // aafonso@novaims.unl.pt // fjesus@novaims.unl.pt // cnunes@novaims.unl.pt // mcastelli@novaims.unl.pt // toliveira@novaims.unl.pt// lloura@ffms.pt

[*]Corresponding author

(Submitted November 21, 2023; Revised October 29, 2024; Accepted December 17, 2024)

ABSTRACT: Education is crucial for individual and societal growth. However, it was significantly impacted by the COVID-19 pandemic, with long-lasting effects. Estimates suggest that students' learning decreased by up to 50% compared to a typical year, though the full impact remains unclear. This paper evaluates primary AA drivers to guide efforts addressing pandemic-related educational inequities. Using government data from virtually all public high school students in a European country, we applied advanced data science methods—Multiple Linear Regression, Decision Trees, Neural Networks, Support Vector Machines, Random Forest, and Extreme Gradient Boosting—to analyze AA determinants before and during the pandemic (2019 and 2020, respectively). Our data includes the most well-known potential AA drivers across four dimensions: students, parents, schools, and teachers. Our substantive findings highlight that student age and legal guardian education were key AA drivers, while Internet access and gender gained importance during the pandemic. Additional drivers, including school size, family nationality, and socioeconomic factors (such as the rate of students receiving school support), also emerged as relevant, particularly under pandemic conditions. This study quantitatively assesses these AA determinants across two distinct academic years, providing nuanced insights into the impact of COVID-19 on education. These results offer valuable guidance for policymakers to implement interventions addressing evolving needs and disparities exacerbated by remote learning. This study contributes to AA literature by utilizing extensive data and machine learning models to reveal enduring and emerging factors affecting educational outcomes during challenging times.

Keywords: Education, Academic achievement, Data science, COVID-19

1. Introduction

Education is the first pillar of social rights in the European Union (European Commission, 2017), deeply intertwined with all aspects of life and well-being. It aligns with the United Nations' target of ensuring all youth complete primary and secondary education with relevant skills by 2030 (United Nations, 2015). Education not only enhances job prospects and income but also provides access to culture and knowledge, equipping individuals to navigate life's complexities. Education is crucial for addressing economic and social challenges. Education fosters competitive and adaptable societies by supplying knowledge and skilled individuals (OECD, 2019).

Historically, pandemics like the Spanish Flu of 1918 have significantly impacted education, leading to innovations like distance learning (Spielman & Sunavala-Dossabhoy, 2021). While diseases like HIV and Ebola had severe effects on society, COVID-19 proved to be on a different scale entirely. COVID-19 affected virtually every country worldwide, resulting in widespread lockdowns, travel restrictions, and drastic changes in daily life for billions of people. Its global reach led to profound social impacts, disrupting various aspects of daily life, including work, education, social interactions, and mental health. Lockdowns and social distancing measures led to isolation, economic hardships, and challenges in accessing healthcare services (Sampath et al., 2021). It is estimated that millions of people may have been pushed below the poverty threshold during this period (Alkire et al., 2021). This disruption has highlighted the challenges students face without adequate Internet access or home learning environments (García & Weiss, 2020). COVID-19 represents a unique opportunity to analyze academic achievement (AA) under extreme conditions and compare it with pre-pandemic data.

The AA investigation used self-reported data and classical statistical analyses (Hattie & Clarke, 2018). However, modern studies have explored macro-level institutional data collected by schools (Cruz-Jesus et al., 2020) by applying machine learning (ML) methods to extract insights from the data. Institutional data includes decades of historical data vital for educational research, such as socio-demographics, educational background, and academic records. It is not typically used due to the sensitivity of the information, as accessing and analyzing this type of data poses availability and privacy issues (Fischer et al., 2020). However, there are some solid arguments, such as a higher predictability power (Costa-Mendes et al., 2021), which can complement traditional research methods and profoundly contribute to progress in identifying the main AA drivers. This research uses ML models to create prototypes of fictional students, analyzing how variations in independent variables affect academic performance across two academic years—2019 and 2020. To our knowledge, no other studies can analyze this problem using virtually every student in a public education system. Hence, in this paper, we intend to respond to the following research questions:

- What were the most important drivers of AA in Portuguese high schools before and during the COVID-19 pandemic?
- What was each one of the main drivers in Portugal's mathematics and native language national exam grades?

The first question centers on understanding the shifts in the AA drivers in Portuguese high schools brought by the COVID-19 pandemic. The second question builds upon this exploration and delves deeper into the specific impact of these drivers on performance in national exams, particularly in mathematics and the mother tongue. By examining the relationship between these main drivers and performance in these core subjects, our research seeks to uncover nuanced insights into how the pandemic has influenced academic outcomes.

2. Theoretical background

AA is conceptualized as the outcomes and abilities of students among different subjects of study in school that enable students to be successful (Genesee et al., 2006). AA is an essential factor for shaping the course of life and determining many opportunities, such as access to higher education, employment opportunities, or self-perceptions (Salmela-aro & Tynkkynen, 2012).

2.1. Drivers of AA

Over the past 50 years, researchers have sought to identify what drives a student to success. Intending to reduce disparities in education, Coleman (1968) studied which factors cause variations in students' success. This author offered a starting point for many other studies that have attempted to identify the underlying factors in educational attainment. In the Appendix A, a summary Table of the literature is found.

2.1.1. Students' characteristics

Previous research shows that the power of cognitive skills and past academic behavior is notorious compared with the remaining factors. Grades such as mathematics (Aaronson et al., 2007; Opdenakker & Van Damme, 2001), mother tongue (Workman, 2021), foreign language (Ömeroğulları et al., 2020) or science (Pokropek et al., 2015), among others, are the strongest predictors of AA. Studies demonstrate that students tend to be homogenous with their grades during their academic life (Asif et al., 2017). Therefore, a 'snowball effect' is created for students who suffer retention since they become more likely to be retained again (Costa-Mendes et al., 2020; Cruz-Jesus et al., 2020).

The Internet has been emerging globally for decades, impacting all human lives and businesses. Despite the positive impact in many areas, the effect on education is still especially controversial. Costa-Mendes et al. (2020) recently found that the impact of the possession of a computer and access to the Internet on a student's performance depends on the subject being studied.

2.1.2. Household characteristics

Despite the Miguéis et al. (2018) study stating that socioeconomic variables are weak estimators compared with students' previous results, socioeconomic factors have played a portentous role in predicting AA in many previous studies. Socioeconomic status (SES) is most frequently measured, among others, by parental education, parental occupation, and income (Long & Renbarger, 2023). Legal guardians' (typically parents') educational level is a common and powerful proxy for students' socioeconomic background. The parent's level of education is a major contributor to their children's success in school (Steinmayr et al., 2010).

2.1.3. School's characteristics

The characteristics intrinsic to the schools that affect students' AA and their magnitude have been studied for many years. As with other environmental aspects, the general poverty level in a school is related to AA. In addition, the size of the school, which has been a variable of study among academicians, has shown contradictory results. While Archibald (2006) stated that the impact is negative, research by Costa-Mendes et al. (2020) supported positive effects. However, the authors believed that this might reflect the socio-demographic characteristics of the location where the school was inserted.

2.1.4. Teachers' characteristics

The quality of teachers is another basis for the educational achievement of students (Aaronson et al., 2007; Archibald, 2006; Coleman, 1968). The teacher-student relationship impacts students' academic performance, learning attitudes, school adjustment, and well-being (Heinla & Kuurme, 2022). Teachers unveil a fundamental position in creating a positive classroom environment and avoid social marginalization (Andersen, 2023).

2.2. COVID-19 pandemic and the impact on school life

The COVID-19 pandemic, provoked by the SARS-CoV-2 virus, was declared by the World Health Organization in early 2020 and is still ongoing. Most countries have imposed restrictive measures to contain the spread of the virus. At the height of the pandemic, one of these preventive measures was closing schools and turning all classes to an online format. In Portugal, students from grades one to nine had televised learning support aired on public access television about the curricular topics for each grade, complementary to their small classes. However, high school students were limited to online classes during most of the third term of the 2019/20 school year. One of the immediate consequences of remote instruction is related to the digital divide. Unequal access to technology and the Internet has created significant disparities in students' ability to engage in remote learning, develop essential digital skills, and access educational resources (Golden et al., 2023). Inequities in access to resources and support have worsened the impact, especially for disadvantaged groups like female students with poor AA, isolated children, low-income families, and Indigenous and disabled communities (Tang, 2023). Students with poor or no Internet access at home, who do not own a computer or mobile device, were disadvantaged compared to classmates who could attend online classes in good conditions and pursue their learning goals. In mid-May 2020, some high schools reopened, allowing students to participate in in-person classes to better prepare them for the 11^{th} and 12^{th}-grade final national exams that have an actual weight for students' grade point average and access to university. However, presence in class was not compulsory as the lockdown was ongoing. A systematic review by Cortés-Albornoz et al. (2023) shows how COVID-19 negatively affects students' AA. The review, analyzing 24 studies from different parts of the world, found that many students struggled academically during the pandemic. The online learning ineffectiveness at the beginning of the pandemic was evident, with female students being more affected than males by the change from physical to remote learning (Hong et al., 2021). Estimates are that the learning gains from this school year can be impaired by 30 to 50% compared to a typical setting (Kuhfeld et al., 2020).

As the learning setting shifted from school to home, successful learning during this period and subsequent AA was contingent on the conditions (both physical and emotional) provided by the home environment. Once again, low-income students are more vulnerable to the negative impact of COVID-19 (Santibañez & Guarino, 2021). Apart from having access to a computer with Internet access, there were other concerns regarding the housing conditions or the economic availability of families to ensure adequate nourishment, as socioeconomically challenged students usually have more than one daily meal at school (Esteves et al., 2021). Also, not all parents can help with school tasks and keep their children engaged in remote learning.

The sudden and unprecedented shift from in-person classes to remote instruction was highly challenging for teachers, who generally felt unprepared for the transition (Steinmayr et al., 2021) due to technological impairments and a lack of adequate class planning. These new challenges and demands led to extraordinary stress in teaching, communicating with parents, and providing support (Pressley, 2021). They reported working longer hours and experiencing difficulties despite their disposition to support their students. In general, students experienced less assistance and feedback from their teachers, resulting in a decrease in the work efforts of struggling students and lower self-efficacy (Mælan et al., 2021), indicating that only allocating chores is not enough to enrich the student's learning process (Steinmayr et al., 2021). Students who reported having more support from their teachers during this period had higher levels of academic motivation (Camacho et al., 2021), an essential predictor of AA.

Recent studies have shown that the COVID-19 pandemic had significant psychological impacts on high school students. With the closure of schools and remote learning, students were deprived of the usual social interactions and routines in school and the sports and extracurricular activities they were used to, leading to a decrease in emotional well-being, which was contaminated with feelings of isolation and loneliness (Jack & Oster, 2023). Also, according to Camacho-Zuñiga et al. (2021), students reported feelings of anxiety, depression, tiredness, stress, and overwhelm. Additionally, research by Walters et al. (2022) has highlighted the adverse effects of online learning on students' overall experience and mental well-being. Students, especially those with specific learning difficulties, reported lower concentration, engagement, perceived learning, and self-worth scores than traditional classroom education. The lack of extracurricular activities or sports also affected the student's emotional well-being.

ICT use increased drastically during the lockdown, as it was the preferred tool for remote working, learning, and indoor leisure activities, reducing social isolation and maintaining contact with peers. Usage of social networking sites has shown positive influences on AA, although excessive use has been correlated with poorer sleep habits, absenteeism, and lower AA (Astatke et al., 2021). Adolescents reported suffering from sleep disorders to manage academic demands during COVID-19, which impacted their mental health, as those with greater severity of sleep disturbances had more elevated levels of anxiety. Youths who spend time on screens at bedtime tend to have greater severity of sleep-related disorders (Zhang et al., 2021).

Considering the most relevant factors identified in the literature as determinants of AA, one can hypothesize that, in extreme situations like the COVID-19 pandemic, some became more pressing, while others may have lost significance, and even new ones may have emerged.

2.3. Machine learning applications on AA

In the past few years, research has broadened its techniques and methods for a better understanding of the drivers of AA, extending classical statistical methods to the application of artificial intelligence. Numerous studies have underscored the significance and efficacy of ML methods in predicting AA, yet determining the optimal approach remains a subject of debate, characterized by diverse viewpoints. Şen et al. (2012) delved into this question by investigating various methodologies to predict student success or failure in Turkey's 7th and 8th grades. This classification challenge was addressed through a comparative analysis of decision trees, support vector machines, neural networks, and logistic regression. In this problem, decision trees proved to be the most accurate technique, emphasizing the interpretability and facility to code on a decision system of the decision trees. Another advantage of decision trees is the transparency this method provides in understanding the underlying conditions that lead to the outcome and supporting the creation of more precise analysis (Abad & López, 2017). Another method that has proved its share is neural networks, which are extremely efficient in predicting AA (Musso et al., 2020). On the other side, several studies endorsed that when traditional machine learning methods are compared with ensemble ones, the ensemble techniques are more potent and bring the advantage of providing robust prediction systems (Costa-Mendes et al., 2020; Cruz-Jesus et al., 2020; Delen, 2010; Miguéis et al., 2018). Using 353 primary students, Sun et al. (2024) conducted a study utilizing individual and ensemble methods to forecast AA. Their findings revealed that the accuracy of ensemble models, specifically eXtreme Gradient Boosting, Random Forest, and Stacking models, exceeded that of individual learners. This superiority was consistently validated across test samples, highlighting the efficacy of ensemble techniques in enhancing predictive performance beyond that achievable by standalone algorithms. Advanced ML methods are often considered superior predictors due to their predictive power. However, this advantage comes at the expense of increased complexity, which can severely limit interpretability, unlike simpler methods such as regressions or decision trees. Table 7 in Appendix A presents a succinct overview of the methodologies employed in previous literature.

3. Methodology

This chapter explains the methodology used in this study. It is divided into two subsections: machine learning methods, where an overview of the methods used is given, and feature selection, i.e., choosing the most important features to model.

3.1. Machine learning methods

3.1.1. Multiple linear regression

Multiple Linear Regression (MLR) is a linear regression that uses various independent variables to predict the value of the dependent one, which is usually continuous. Each independent variable has a slope, and the ordinary least squares method (OLS) estimates the parameters. MLR assumes a linear relationship, multivariate normality, no multicollinearity, no auto-correlation, and homoscedasticity (James et al., 2013).

3.1.2. Decision tree regressor

Decision trees (DT) are a tree-like approach. The process is done interactively; it starts at the root node, which contains the entire dataset, and proceeds with sequential divisions into smaller datasets, leading to different decision nodes (Asif et al., 2017). The process of the DT Regressor starts by choosing the attribute with the highest goodness of fit, which can take other measures and is applied until a pre-defined criterion is met. A common splitting criterion is the mean squared residual (Zhang & Ma, 2012).

3.1.3. Artificial neural networks

Artificial neural networks (NN) are biologically inspired methods mimicking the structure of a human brain (Jain & Mao, 1996).

Multilayer perceptron (MLP) is an architecture of NNs with an input layer, multiple hidden layers, and an output layer. Neurons form these layers, and each neuron in the input layer connects to every neuron in the first hidden layer, which in turn connects to every neuron of the following hidden or output layer. All relationships have a weight associated with them, and the set of weights represents the model's parameters. The learning process is performed by applying a learning rule which redefines the weights as needed.

First, a random weight is assigned to each channel, a bias for the neurons of the hidden layers, and all the nodes, except for the input node, receive an activation function. The algorithm starts with the feedforward process, where the neurons in a given layer receive input from the outputs of the neurons of the previous layer. The feedforward process is then performed by sending the data from the input to the output layer, neuron-by-neuron.
The second step is backpropagation, which follows the opposite direction, updating weights using a delta rule to minimize the cost function. The weights are updated using the feedback regarding the error in the feedforward process. Backpropagation aims to minimize the cost function concerning the NN's weights. This process can be repeated until the error is below a threshold or the maximum number of iterations is reached (Vanneschi & Castelli, 2018).

3.1.4. Support vector regressor

Support Vector Machines (SVM) belong to the supervised learning group of the ML methods (Hearst et al., 1998). SVM is an algorithm that maximizes data collection (Noble, 2006). SVM can be applied to continuous targets and is called a Support Vector Regressor (SVR). SVR uses a kernel function to map multidimensional data into a high-dimensional feature space. The goal is to find the best-fit line, a hyperplane that collects the maximum number of data points. The solution is the support vectors, a subset of the training dataset selected to determine the fitted surface (Smola & Scholkopf, 2004).

3.1.5. Random forests

Random forests (RF) are an ML ensemble method used for classification or regression (Breiman, 2001). An RF independently selects random samples of the original dataset and creates decision trees using only a randomly selected set of variables in each node. It then selects the one that produces the maximum information gain—repeating this process until a stopping criterion is met. In the regression context, the predictions are provided by the unweighted average.

3.1.6. Extreme gradient boosting

Gradient boosting is an ensemble ML technique that successively creates weak learns that are improved by the residuals of the previous one. The goal is to minimize the loss function, which varies according to the task. The squared loss is used for regression, while the exponential error is applied for a classification problem (Friedman, 2001). The starting point of XGBoost is with a low variance and high bias model but using a penalization of the individual learners. The model can reduce the bias progressively. The penalization diminishes the leaf weights more heavily with less data evidence, continuously changing and reshaping the parameters, reducing the variance in each tree. It is a highly adaptive model since it considers the trade-off between variance and bias in every stage (Nielsen, 2016).

3.2. Feature selection

Feature selection leads to variance reduction and simpler models, meaning less computational power, and it avoids the curse of dimensionality. A voting system guarantees that the most important variables are used and eliminates redundant or least relevant features. First, Recursive Feature Elimination (RFE) selects the optimal number of features in each dataset (N). Subsequently, RFE, Lasso, and Ridge Regressions were used to determine the N most essential features to predict each target. Afterward, the features were ordered from the most voted to be selected for the model to the least voted. Then, the top N features were selected as part of the model. Combining these three methods allowed the reduction from over 50 variables to around 15 in each dataset.

3.2.1. Recursive feature elimination

RFE ranks the variables based on their performance on a model defined where the features with the least importance are excluded. RFE operates by initially training the model with all the features to obtain the significance of each one of them. In the next step, the least important features are removed from the current set of features, and the process is repeated on a smaller and smaller group of features until the optimal set is found (Guyon et al., 2002).

3.2.2. Ridge regression

Ridge regression, a derivative of linear regression, is a technique for estimating the coefficients of multiple regression models. The method penalizes the coefficients of the features, allowing for the minimizing of errors. The coefficients are shrunk toward zero (and each other), reducing complexity and multicollinearity (Hastie et al., 2009). The features with higher coefficients are chosen to be part of the model.

3.2.3. Lasso regression

Also deriving from liner regression, Lasso considers a constraint on the coefficients' value, which allows shrinkage usage and minimization of the prediction error. The method starts with the creation of an upper bound. A constraint is then applied in the sum of absolute values of the model parameters. Later, the shrinking process penalizes the regression coefficients, shrinking specific values to zero. The larger the penalty, the further the estimates are shrunk toward zero. After the shrinking process, the variables that remain with a non-zero coefficient are selected to be part of the model (Fonti, 2017).

3.3. Data

This study used an anonymized dataset provided by the Directorate-General of Statistics for Education and Science (DGEEC) of the Portuguese Ministry of Education, containing information about virtually all public high school students who have taken national upper secondary exams. The dataset includes variables of four categories: students, legal guardians, teachers, and schools, the more relevant categories in AA research. Within each category, there is a wide variety of potential AA drivers. The AA measure selected was the 2020 grade (in the middle of the first COVID-19 wave in Europe, in the academic year of 2019/2020) upper secondary national exams of Mathematics and mother tongue (Portuguese). The national exam grades were chosen to provide an objective measure of AA, comparable nationwide, avoiding the potential bias present in internal grades given by teachers. Therefore, the instrument that captured the target variable in our data was precisely the same for every student included. The variables chosen by the feature selection system are validated by scholarly evidence, as illustrated in Table 1.

Table 1. Literature support of the variables chosen by the feature selection process

Variable	Support
Failed years	(Coleman, 1968; Cruz-Jesus et al., 2020; Pov et al., 2022; Vandelannote & Demanet, 2020)
Gender	(Coleman, 1968; Costa-Mendes et al., 2020; Cruz-Jesus et al., 2020; Fischbein, 1990; King, 2016; Mensah & Kiernan, 2010; Musso et al., 2020; Sun et al., 2024)
LG Educ	(Coleman, 1968; Mensah & Kiernan, 2010; Miguéis et al., 2018; Opdenakker & Van Damme, 2001; Pokropek et al., 2015; Steinmayr et al., 2010; Tesfagiorgis et al., 2020)
Type of School	(Marks et al., 2006; Nunes et al., 2022)
Internet at Home	(Costa-Mendes et al., 2021; Kubey et al., 2001; Lee, 2022; Trakunphutthirak & Lee, 2022)
Who is the LG	(Aaronson et al., 2007; Nunes et al., 2022)
Teacher Quality	(Aaronson et al., 2007; Rivkin et al., 2005)
State support	(Aaronson et al., 2007; Archibald, 2006; Costa-Mendes et al., 2020; Cruz-Jesus et al., 2020; Şen et al., 2012)

This study included two target variables: the mother tongue and the mathematics national exam grades. The data was collected and pre-processed using Microsoft SQL Server Management Studio®, whereas the analytical modeling was performed using Python and SAS®. The dataset was split into training and test sets, comprising 75% and 25%, respectively.

3.4. Methodology application

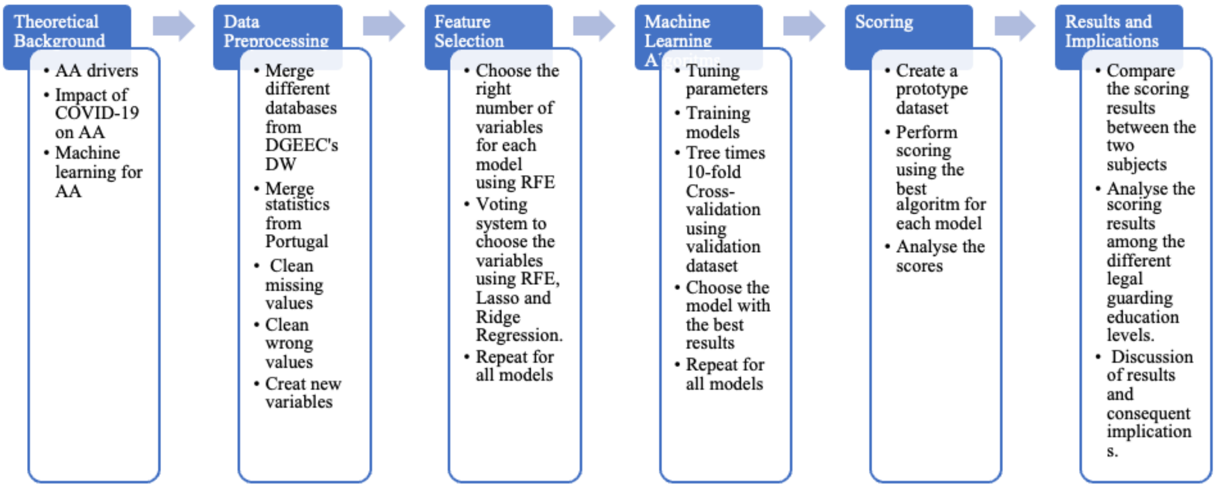
A cross-validation grid search provided by Scikit-learn was used to tune the parameters in all the Machine Learning algorithms (DT, NN, RF, SVR, and XGB). The tuned parameters were the criterion, maximum number of features to use, maximum depth, minimum samples split, subsample, number of estimators, and the penalizations for the tree type models; hidden layers sizes, activation function, solver, and learning rate for the NN; regularization, gamma, and kernel for SVM. Subsequently, the algorithms were trained using the hyperparameters chosen by the process selected. Lastly, ten-fold cross-validation was used to evaluate the models objectively. Cross-validation is a vital tool for understanding the power of model generalization and against overfitting. It functions by splitting the data into K parts of equal sizes (in this case, ten samples), and in a loop, the function learns using K-1, while the remainder is used for testing. Three repetitions of this process were done, thus resulting in the evaluation of 30 models. The outcome is the mean R-squared of all the iterations for each model. The R-squared measure is employed since it represents how much the model explains the dependent variable variation. R-squared is commonly used (e.g., Mendonça et al., 2022; Sultan et al., 2011; Vanderlinde & Van Braak, 2011). It has several advantages: it is more informative than the symmetric mean absolute percentage error, for example, and does not have the interpretation limitations of other measures, such as the mean square error (Chicco et al., 2021).

To evaluate feature significance, we assessed the top ten features of the best model for each target using Scikit-learn's feature importance. Features significantly impacting performance were identified as most valuable. Features with the greatest impact on performance are deemed most valuable. We then created "prototype" datasets for each target variable to understand the importance of the variable using RFs. Each prototype represents a fictitious student, with all variables set to their mean except for one, which varies by one standard deviation or binary value. The baseline dataset comprises average values for all variables, allowing us to quantify the impact of variable changes on AA while keeping other variables constant. The baseline is a row containing the average values for all variables. This approach allows us to quantify how changes in a variable while keeping others constant, impact AA, yielding measurable and understandable effects.

One of the most important questions this work aims to answer is how AA drivers were impacted by the pandemic and the subsequent measures implemented to cope with it (e.g., social distancing and remote classes). In doing so, we aim to replicate the same methodology for 2019. However, caution should be taken before analyzing the results between the two periods: With social distancing and remote classes, how high school students were evaluated underwent meaningful changes between 2019 and 2020. Besides the noticeable changes in the teaching methods, the exams' structures in the two years and the exams' enrolment criteria were also different. Regarding structures, in 2020 (during the pandemic), the exams comprised alternative questions, which students could answer as they saw fit. This aspect simplified the exams as students could answer the questions and felt better prepared. Perhaps more importantly, until 2020 (during the pandemic), the national exams were compulsory for every high school student who wanted to complete their secondary program. However, during the pandemic, the national exams were optional, as only students who wanted to enroll in a tertiary degree that demanded that same exam needed to take it.

For this reason, the number of exams is lower in 2020 than in 2019. However, this new reality has a slighter impact on the mathematics exam. This inconvenience is diminished because the students still need to perform the exam to enter university, as it is required for most degrees in STEM areas. While these changes still make it possible to compare the relative importance of each driver before and during the pandemic with some caution, it needs to be acknowledged that comparisons in absolute differences are not adequate, as there is a sample selection problem between the two years, at least in the Portuguese exam. A graphical representation of the methodology applied can be found in Figure 1.

Figure 1. Schema of the methodology used



4. Results

The methodology mentioned in the previous chapter was applied. The *R*-squared results after the cross-validation for the two datasets (2019 and 2020) are summarized in Table 2.

Table 2. R-squared of the models after ten-fold cross-validation

Model	Mathematics exam grade (2020)	Mother tongue exam grade (2020)
MLR	0.14	0.11
DT	0.14	0.11
NN	0.10	0.10
RF	0.23	0.19
SVM	0.13	0.02
XGB	0.19	0.15

Table 3 shows ensemble ML methods, which combine multiple naive algorithms, outperform the remaining ML methods and the classical statistical approach. The champion model was the unanimous RF, providing the best R-squared average on all targets. However, it is essential to highlight that since the academic year of 2019/2020 had different rules due to COVID-19, the number of observations in this dataset was considerably lower.

4.1. Using prototypes to quantify AA drivers' impact

The goal is to understand the impact of the most important features on the exam grades. The variables under evaluation are the ten variables with the highest feature importance of the model that revealed the highest R-squared, i.e., the RF (please see Table 1). By analyzing the mathematics and mother tongue exam grades separately, students' age, gender (female), and legal guardians' education were the drivers that most stood out. The results of the approach mentioned in Chapter 3 are summarized in Table 3. In Tables 3 and 4, the beta values indicate the impact of each variable relative to a student, with average values across all variables, as outlined in Section 3.4 (Methodology Application). For example, in the 2020 mother tongue exam (please see Table 3), increasing age by one standard deviation (0.7) leads to a decrease of 2.5 points in the exam grade. For gender, we created two prototypes: Stu_Fem = 0 (male) and Stu_Fem = 1 (female), keeping every other variable constant with average values. In this case, in the 2020 mother tongue exam, a female student is predicted to score 2.3 points higher than a male student, on average, keeping everything else constant (*ceteris paribus*).

Table 3. Predicted impacts in mother tongue and mathematics exam grades 2020 (during the pandemic)

Target	Mother tongue exam grade 2020 (during the pandemic)			Mathematics exam grade 2020 (during the pandemic)		
Rank	Variable	Unit	$\hat{\beta}$	Variable	Unit	$\hat{\beta}$
1ˢ	Stu_Age	0.7	-2.5	Stu_Age	0.5	-3.3
2	Stu_Fem	N=0/Y=1	2.3	LG_Educ	3.6	1.3
3	Sch_Size	79.4	1.1	Stu_Net	N=0/Y=1	0.9
4	LG_Educ	3.7	0.7	Sch_FailR	0.1	-0.6
5	Stu_Net	N=0/Y=1	0.7	Stu_Fem	N=0/Y=1	0.4
6	Eff_Growth	0.6	0.4	Sch_Size	80.0	0.3
7	Stu_SocSup	N=0/Y=1	-0.3	Sch_MScPhD	0.1	0.2
8	Sch_Sup	0.3	-0.2	LG_Nat_PT	N=0/Y=1	0.1
9	Sch_FailR	0.1	0.1	Sch_SocSup	0.1	0.1
10	Sch_SocSup	0.1	0	Stu_SocSup	N=0/Y=1	0

Note. effective growth of the school municipality (Eff_Growth), years of education of the legal guardian (LG_Educ), father being the legal guardian (LG_father), the student himself being the legal guardian (LG_own), the mother being the legal guardian (LG_mother), legal guardian being of a nationality other than Portuguese (LG_Nat_PT), the teacher holding an MSc or a Ph.D. (Prof_MscPhD), the school offering elementary and high school (Sch_Elem&High), rate of students who have failed the year in that school (Sch_FailR), the rate of teachers with an MSc or a Ph.D. in the school (Sch_MscPhD), school size (Sch_Size), rate of students with the highest level of social support in the school (Sch_SocSup), students with any level of social support (Sch_Sup), student's age (Stu_Age), student being a female (Stu_Fem), the student being of a nationality other than Portuguese (Stu_Nac_PT), the student having Internet access (Stu_Net), the student received the highest level of social support (Stu_SocSup).

4.1.1. Comparison before and during the COVID-19 pandemic

Table 4 presents the results obtained by applying the same method to 2019, coincidently also using RFs as the champion model (please see Appendix B). Before COVID-19, the student's age, the legal guardian's education, and gender are the top drivers in predicting the grades on both exams. Additionally, the negative impact of different cycles of studies in the same school is emphasized in both subjects. During COVID-19, this variable was no longer a top 10 feature for predicting AA. Now, performing a relative comparison concerning 2020 (Table 3) and 2019 (Table 4), school size reached the top in 2020, evidencing that students in larger schools perform better. Regarding the mother tongue exam, in particular, the rate of students who failed the year in that school loses positions in the ranking of most important drivers for 2020, while the effective growth of the municipality where the school is placed becomes an important driver. In the case of the mathematics exam in 2020, the nationality of the legal guardian becomes, although slightly, a significant factor, with a positive effect of having national legal guardians.

Table 4. Predicted impacts on mother tongue and mathematics exam grades 2019 (pre-pandemic)

Target	Mother tongue exam grade 2019			Mathematics exam grade 2019		
Rank	Variable	Unit	^β	Variable	Unit	^β
1	Stu_Age	0.6	-1.5	Stu_Age	0.5	-2.6
2	Stu_Fem	N=0/Y=1	1.1	LG_Educ	3.3	1.9
3	LG_Educ	3.3	1	Sch_FailR	0.1	-0.7
4	Sch_FailR	0.1	-0.8	Sch_Elem&High	N=0/Y=1	-0.4
5	Sch_Elem&High	N=0/Y=1	-0.3	LG_mother	N=0/Y=1	-0.4
6	Stu_Net	N=0/Y=1	0.2	Stu_Net	N=0/Y=1	0.3
7	LG_mother	N=0/Y=1	0.1	Prof_MscPhD	N=0/Y=1	-0.1
8	Sch_MsCPhD	0.2	0	Stu_Fem	N=0/Y=1	0.1
9	Sch_SocSup	0.1	0	Sch_SocSup	0.1	-0.1
10	Stu_SocSup	N=0/Y=1	0	LG_own	N=0/Y=1	-0.1

Figure 2. Absolute impact of each variable for the 2020 Mother Tongue and Mathematics national exam

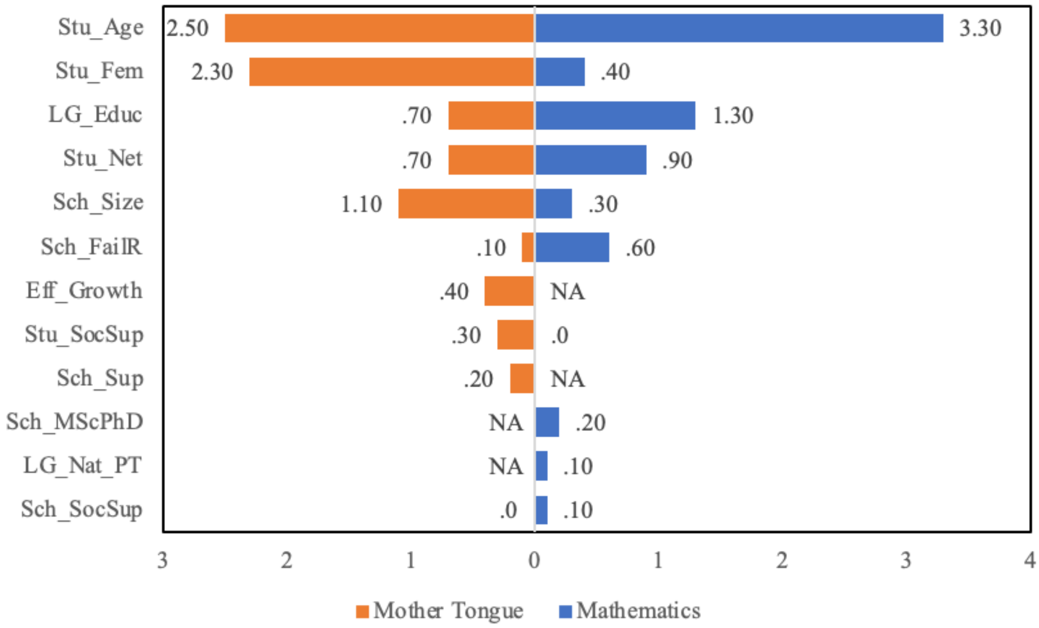
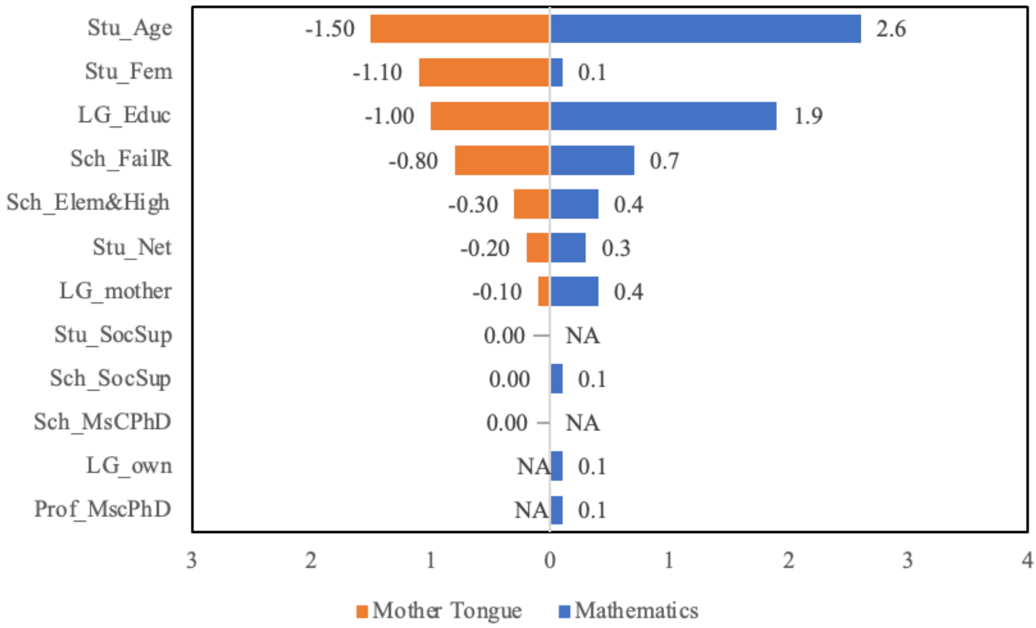


Figure 3. Absolute impact of each variable for the 2019 Mother Tongue and Mathematics national exam



Given the importance of legal guardians' education (see Figures 2 and 3), we decided to test the difference in the different cycles of studies. Table 5 provides a quantifiable comparison of the effects of having legal guardians with nine years (middle school), 12 years (high school), and 16 years (tertiary level) of education. Notably, the mean effect centers around high school education, ranging from 10.8 to 11.8, which explains why having 12 years of education shows only a slight impact. Most importantly, it is worth pointing out that the significant impact on AA is seen around 16 years of education, i.e., at the tertiary level: a student whose guardian has a tertiary degree scores an average of 0.6 points higher in the mother tongue national exam and 0.9 points higher in the mathematics one, in 2020. Note that in 2019, the legal guardian effect was 1.3 and 1.9, respectively, which is noticeably higher than in the pandemic year. As mentioned earlier, a possible reason is due to adjustments in the national exams' enrollment criteria.

Table 5. Predicted impacts on different years of legal guardians' education

	2019		2020	
Target	Mathematics	Mother tongue	Mathematics	Mother tongue
Legal Guardian Education	^β	^β	^β	^β
LG_Educ = 9 years (basic education)	0	-0.1	-0.9	-0.1
LG_Educ = 12 years (secondary education)	-0.1	0.1	0	-0.2
LG_Educ = 16 years (≈ tertiary education)	1.9	1.3	0.9	0.6
Mean LG_Educ	11.4	11	11.8	10.8

5. Discussion, implications, and limitations

5.1. Discussion of findings

This study allowed us to analyze and classify the impact of several AA drivers on Portuguese public high school students in two consecutive years: 2019, a typical school year, and 2020, after the COVID-19 pandemic emerged. Results show an arranged ranking of drivers by order of importance, an explanation of final exam grades, and a quantitative estimation of each driver's impact (signal and effect size). This quantitative estimate is crucial for comparing exam scores in two different years and thus inferring that AA drivers become more relevant, considering a school year so challenging and exceptional as 2020, and students are more at risk. The results indicate that ML methods, especially ensemble ML techniques, exhibit superior predictive power, aligning with findings from previous research (Miguéis et al., 2018; Nunes et al., 2022; Yağcı, 2022). This enhanced predictive capability means these methods are more adept at uncovering insights that reflect real-world conditions, making them more reliable for guiding educational interventions.

Furthermore, these models excel at handling complex data (Luan & Tsai, 2021) without requiring the often impractical assumptions typical of traditional methods. However, a notable challenge with these sophisticated techniques is the black-box effect. Nevertheless, our work brought an innovative approach to this research stream. We used prototypes to quantify each variable's impact on AA, which was almost exclusive to classical methods until recently. It allowed us to use a larger data sample than traditional AA research, relying on more sophisticated and proven superior data science methods (Costa-Mendes et al., 2020; Cruz-Jesus et al., 2020).

According to our results, the drivers that most impact AA are the student's age, legal guardians' education level, Internet at home, gender, the rate of failing students in the school, family nationality (of the legal guardian and the student), school size, and the rate of students who receive a school allowance in the school. Nevertheless, nationality and school size only appeared as top drivers after COVID-19 started. Student's age is a proxy for a much-studied driver in AA research, i.e., previous retention. Our analysis highlighted this variable as one of the most negatively impactful on AA. Retention policies have been among the widely discussed topics in education, and our results help shed some light on this debate. While impending retention may motivate some students to achieve, there is evidence that retention does not prevent future retentions either, and most retained students suffer adverse academic and psychosocial effects (Jimerson et al., 2002). Recent research on retention drivers conducted in Portugal indicates that teachers are less likely to retain students who have been previously held back. The findings also reveal that females are retained less frequently than males, while being an immigrant from another Portuguese-speaking country and having a mother with a lower education level increase the likelihood of retention (Nunes et al., 2018). The negative effect of retention is not exclusively noticeable at a student level, as we found evidence that it also impacts the school level. In schools with higher rates of failing students, exam results worsen.

Also related to age and likely previous retention is the impact of students' legal guardian status, meaning they are 18 or over. This factor negatively affected exam scores during the pandemic, potentially due to reduced parental supervision or previous academic failures. Students over 18 taking final exams are often high school graduates seeking university admission, making them likely to have legal guardians involved in their education.

Another of the main drivers of AA highlighted in our results is the education level of legal guardians (typically parents). Students whose parents have tertiary education outperform their peers. This positive effect of higher parental education on upper-secondary AA is especially significant in 2019 grades, probably due to the students' flexibility during the pandemic, where the students who did the exam were those who needed it to enter university. Our results also prove that the difference between having a legal guardian in middle school and high school is negligible.

Internet access became crucial during the pandemic. Although previous research reports conflicting findings concerning the influence of Internet use on AA, with both positive (Bowers & Berland, 2013) and adverse effects (Rozgonjuk et al., 2021), the necessity to transition from face-to-face classes to online teaching made Internet access an essential tool for online classes. Higher family income generally provided better Internet access for remote learning (Benalcázar et al., 2022). Students' confidence and skill in using the Internet significantly impacted learning effectiveness (Hong et al., 2022). Thus, it seems that the COVID-19 pandemic reinforced the disadvantage of belonging to low SES and rural families. Lin et al. (2023) proposed a novel method to bridge the digital divide between rural and urban students by implementing a dual scaffolding-embedded mobile augmented reality (AR) learning approach to fight this disparity. The approach provided tailored support for teachers and students, benefiting rural students and emphasizing the need for customized strategies.

Immigrant families, including international students, predicted lower exam results, with notable struggles in mathematics during the pandemic. These students seemingly benefit more from the presence of face-to-face classes and teacher support, as the evident decrease in teacher-student communication during the lockdown had adverse effects. In addition, non-national parents are likely less equipped (due to cultural differences, perhaps) to assist children in remote learning, leading to more difficulties and lower exam scores.

According to our results, school-related drivers are also the most relevant to explaining AA. AA is lower in schools with higher rates of economically disadvantaged students who receive the highest school allowance. In Portugal, where public school enrolment is determined by home address (students typically attend the school closest to where they live), this may result in a downside for students from underprivileged neighborhoods, leading to a snowball effect that may be difficult to escape. In light of our results, it would probably be adequate to implement specific learning programs in such neighborhoods to avoid having schools with primarily underprivileged students. As our results show, it will be more challenging to thrive academically and break the disadvantage cycle. Recent research shows that when adolescents believe their socioeconomic conditions can be changed, they achieve more, highlighting the importance of focusing on a growth mindset, hopefulness, and school engagement to enhance AA (Zhao et al., 2021).

Further conclusions of this study show other relevant drivers of AA, namely gender, as females outperform males in exam grades. Even though this gap was already present, the persistence in 2020 exams may arise from the better adaptability of females towards learning at home. Additionally, our results show that in 2019, in schools where middle and high school students are together, 12^{th}-grade exam scores tend to be lower. This aspect shows that being in a school aggregating different levels of studies negatively affects the students' performance. However, during 2020, the effect of the different cycles vanished from the most important, while school size increased, confirming that students from larger schools had slightly better results. Costa-Mendes et al. (2020) exposed that school size may reflect the socioeconomic individualities of the school's location. In accordance, it may seem that during the pandemic, the differences among the type of locations where schools are placed (e.g., rural vs. urban areas) were emphasized, which is the strength of the ascension of effective growth of the municipality where the school is established to the most important variables (2020 mother tongue exam).

Several frameworks for remote learning can be applied to address the challenges that have been proven to arise with the COVID-19 pandemic. Before COVID-19, Aparicio et al. (2016) had already recognized the importance of a structured e-learning approach. The authors proposed a framework that focuses on engaging stakeholders to support learning. Considering the challenges posed by the COVID-19 pandemic, Huang et al. (2022) underscored the importance of adapting STEM education to ensure continued learning. This adaptation is crucial to address the disruptions caused by the pandemic and advocate for innovative solutions that foster educational resilience. The authors showed that through a transdisciplinary approach, where STEM and social service components are integrated, educators guarantee a holistic learning experience for students. Also, video facilitation plays a central role, enabling remote engagement and deepening understanding of complex concepts. Another important factor mentioned is the pandemic's impact on learning gains and how this can lead to low levels of AA. Liu et al. (2024) proposed using Conversational Agents (CAs) to offer personalized feedback, guidance, and support through combined knowledge and emotional scaffolding. This approach enhances learning outcomes by addressing knowledge gaps and providing emotional support like empathy. CAs that act as tutors and companions use human-like language and give timely feedback for cognitive and emotional needs to engage learners and maintain positive emotional states effectively. The study demonstrated that positive emotions strongly predict learning success, emphasizing the need to address both cognitive and emotional aspects of online learning.

5.2. Limitations and future work

Like any other study, we need to acknowledge some limitations. First, as mentioned earlier, we cannot infer a direct comparison between the 2019 and 2020 exams as students' enrolment criteria differ between 2019 and 2020. Future studies should directly compare the magnitudes of the differences between the drivers of AA before and during COVID-19, as this will allow the creation of more accurate policies and measures to fight the consequences of the pandemic. Another limitation is that our study only uses secondary data, which may lead us to exclude some potentially important drivers (e.g., personality traits or sleeping habits). This limitation could stimulate future research that combines both primary and secondary data. However, it should be noted that using primary data would undoubtedly affect the sample size, as it is not practically possible to collect the kind of data that would be interesting to collect (e.g., sleeping habits, personality traits, parents' and teachers' involvement) from tens of thousands of students. Therefore, although a limitation, the fact that we use secondary data only also poses some advantages. Finally, as a third limitation, our data is cross-sectional. Therefore, it would be noteworthy that future researchers use longitudinal data to understand how different drivers change from cycle to cycle of studies in the same student. Doing so would allow pointing out different needs at the different stages of a student's life.

6. Conclusions

This study allowed us to understand the primary drivers of AA in high school while shedding light on the impacts that the COVID-19 pandemic had on school life in Portugal in early 2020. We found that non-Portuguese students were particularly affected by school closures and remote teaching. Students who had previously failed for at least a year also struggled in this period. Older students who are their own legal guardians also performed worse than their peers, meaning parental involvement may have been beneficial; highly educated parents helped students overcome the pandemic challenges, yielding higher exam scores this school year.

The results of our study emphasize the significance of encouraging success in school and offer valuable insights into increasing mandatory graduation rates and the speed of university enrolments and graduates. Investigating the data from virtually all Portuguese public high schools led us to identify new AA features, such as parents' university degrees and schools with highly educated teachers. Legal guardians with university education and teachers with higher degrees (M.Sc. or Ph.D.) are positive AA drivers, leading to a contagion effect. Higher education levels lead to higher personal development and better quality of life. It provides parents with the tools to help their children succeed at school. Also, the importance of the Internet at home during the pandemic is undeniable, which can also be seen as a proxy for a student's socioeconomic level and how the conditions at home with the lockdowns became even more relevant. The reinforcement of the role of gender in the subject of mother language is also pertinent. Hence, these insights provide an undoubted path for education decision-makers.

Acknowledgment

This work was supported by national funds through FCT (Fundação para a Ciência e Tecnologia), under the project - UIDB/04152/2020 - Centro de Investigação em Gestão de Informação (MagIC)/NOVA IMS) (https://doi.org/10.54499/UIDB/04152/2020).

References

Aaronson, D., Barrow, L., & Sander, W. (2007). Teachers and student achievement in the Chicago public high schools. *Journal of Labor Economics*, *25*(1), 95–135. https://doi.org/https://doi.org/10.1086/508733

Abad, F. M., & López, A. A. C. C. (2017). Data-mining techniques in detecting factors linked to academic achievement. *School Effectiveness and School Improvement*, *28*(1), 39–55. https://doi.org/10.1080/09243453.2016.1235591

Alkire, S., Nogales, R., Quinn, N. N., & Suppa, N. (2021). Global multidimensional poverty and COVID-19: A decade of progress at risk? *Social Science and Medicine*, *291*(October), 114457. https://doi.org/10.1016/j.socscimed.2021.114457

Andersen, M. B. (2023). Measuring social marginalisation: using SEM to uncover essential social factors underlying social marginalisation in public schools. *School Effectiveness and School Improvement*, *34*(1), 90–108. https://doi.org/10.1080/09243453.2022.2113547

Aparicio, M., Bacao, F., & Oliveira, T. (2016). An e-Learning theoretical framework. *Educational Technology & Society*, *19*(1), 292–307.

Archibald, S. (2006). Narrowing in on educational resources that do affect student achievement. *Peabody Journal of Education*, *81*(4), 23–42. https://doi.org/10.1207/s15327930pje8104_2

Asif, R., Merceron, A., Ali, S. A., & Haider, N. G. (2017). Analyzing undergraduate students' performance using educational data mining. *Computers and Education*, *113*, 177–194. https://doi.org/10.1016/j.compedu.2017.05.007

Astatke, M., Weng, C., & Chen, S. (2021). A literature review of the effects of social networking sites on secondary school students' academic achievement. *Interactive Learning Environments, 31*(4), 2153–2169. https://doi.org/10.1080/10494820.2021.1875002

Benalcázar, M. E., Barona, L., Valdivieso, Á. L., Vimos, V. H., Velastegui, D., & Santacruz, C. J. (2022). Educational impact on Ecuadorian university students due to the COVID-19 context. *Education Sciences*, *12*(1), 17. https://doi.org/10.3390/EDUCSCI12010017

Bowers, A. J., & Berland, M. (2013). Does recreational computer use affect high school achievement? *Educational Technology Research and Development*, *61*(1), 51–69. https://doi.org/10.1007/s11423-012-9274-1

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*, 5–32.

Camacho, A., Correia, N., Zaccoletti, S., & Daniel, J. R. (2021). Anxiety and social support as predictors of student academic motivation during the COVID-19. *Frontiers in Psychology*, *12*, 644338. https://doi.org/10.3389/fpsyg.2021.644338

Camacho-Zuñiga, C., Pego, L., Escamilla, J., & Hosseini, S. (2021). The impact of the COVID-19 pandemic on students' feelings at high school, undergraduate, and postgraduate levels. *Heliyon*, *7*(3). https://doi.org/10.1016/j.heliyon.2021.e06465

Chicco, D., Warrens, M. J., & Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science*, *7*, 1–24. https://doi.org/10.7717/PEERJ-CS.623

Coleman, J. S. (1968). Equality of educational opportunity. *Equity and Excellence in Education*, *6*(5), 19–28. https://doi.org/10.1080/0020486680060504

Cortés-Albornoz, M. C., Ramírez-Guerrero, S., García-Guáqueta, D. P., Vélez-Van-Meerbeke, A., & Talero-Gutiérrez, C. (2023). Effects of remote learning during COVID-19 lockdown on children's learning abilities and school performance: A systematic review. *International Journal of Educational Development, 101*, 102835. https://doi.org/10.1016/j.ijedudev.2023.102835

Costa-Mendes, R., Cruz-Jesus, F., Oliveira, T., & Castelli, M. (2021). Machine learning bias in predicting high school grades: A knowledge perspective. *Emerging Science Journal*, *5*(5), 576–597. https://doi.org/10.28991/esj-2021-01298

Costa-Mendes, R., Oliveira, T., Castelli, M., & Cruz-Jesus, F. (2020). A machine learning approximation of the 2015 Portuguese high school student grades: A hybrid approach. *Education and Information Technologies*, *26*(2), 1527–1547. https://doi.org/10.1007/s10639-020-10316-y

Cruz-Jesus, F., Castelli, M., Oliveira, T., Mendes, R., Nunes, C., Sa-Velho, M., & Rosa-Louro, A. (2020). Using artificial intelligence methods to assess academic achievement in public high schools of a European Union country. *Heliyon*, *6*(6), e04081. https://doi.org/10.1016/j.heliyon.2020.e04081

Delen, D. (2010). A comparative analysis of machine learning techniques for student retention management. *Decision Support Systems*, *49*(4), 498–506. https://doi.org/10.1016/j.dss.2010.06.003

Esteves, M., Freitas, P., Herdade, M., Carvalho, B. P., & Peralta, S. (2021). Crianças em Portugal e ensino a distância: um retrato [Children in Portugal and distance learning: A portrait]. *Zenodo*.

European Commission. (2017). *European Pillar of Social Rights*. https://doi.org/10.2792/95934

Fischbein, S. (1990). Biosocial influences on sex differences for ability and achievement test results as well as marks at school. *Intelligence*, *14*(1), 127–139. https://doi.org/10.1016/0160-2896(90)90018-O

Fischer, C., Pardos, Z. A., Baker, R. S., Williams, J. J., Smyth, P., Yu, R., Slater, S., Baker, R., & Warschauer, M. (2020). Mining big data in education: Affordances and challenges. *Review of Research in Education*, *44*(1), 130–160.

Fonti, V. (2017). *Feature Selection using LASSO*.

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, *29*(5), 1189–1232.

García, E., & Weiss, E. (2020). COVID-19 and Student performance, equity, and U.S. education policy. *Economic Policy Institute*, 1–60.

Genesee, F., Lindholm-Leary, K., Saunders, W., & Christian, D. (2006). *Educating English Language Learners*. Cambridge University Press.

Golden, A. R., Srisarajivakul, E. N., Hasselle, A. J., Pfund, R. A., & Knox, J. (2023). What was a gap is now a chasm: Remote schooling, the digital divide, and educational inequities resulting from the COVID-19 pandemic. *Current Opinion in Psychology, 52*, 101632. https://doi.org/10.1016/j.copsyc.2023.101632

Guyon, I., Weston, J., Barnhill, S., & Vapnik, V. (2002). Gene selection for cancer classification using support vector machines. *Machine Learning*, *46*(1–3), 389–422. https://doi.org/10.1023/A:1012487302797

Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction* (Vol. 2, pp. 1-758). Springer New York. https://doi.org/10.1007/978-0-387-84858-7

Hattie, J., & Clarke, S. (2018). *Visible Learning: Feedback*. Routledge. https://doi.org/10.4324/9780429485480

Hearst, Marti. A., Scholkopf, Bernhard., Dumais, Susan., Osuna, Edgar., & Platt, J. (1998). Support vector machines. *IEEE Intelligent Systems and Their Applications*, *13*(4), 18–28.

Heinla, E., & Kuurme, T. (2022). The impact of school culture, school climate, and teachers' job satisfaction on the teacher-student relationship: A case study in four Estonian schools. *Research Papers in Education*, *39*(3), 439–465. https://doi.org/10.1080/02671522.2022.2150883

Hong, J. C., Liu, Y., Liu, Y., & Zhao, L. (2021). High school students' online learning ineffectiveness in experimental courses during the COVID-19 Pandemic. *Frontiers in Psychology*, *12*, 738695. https://doi.org/10.3389/fpsyg.2021.738695

Jack, R., & Oster, E. (2023). COVID-19, school closures, and outcomes. *Journal of Economic Perspectives*, *37*(4), 51–70. https://doi.org/10.1257/jep.37.4.51

Jain, A. K., & Mao, J. (1996). Artificial neural networks - A tutorial. *Computer*, *29*(3), 31–44.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An introduction to statistical learning. In *Springer* (Vol. 103, Issue 10). Springer New York. https://doi.org/10.1007/978-1-4614-7138-7

Jimerson, S. R., Anderson, G. E., & Whipple, A. D. (2002). Winning the battle and losing the war: Examining the relation between grade retention and dropping out of high school. *Psychology in the Schools*, *39*(4), 441–457. https://doi.org/10.1002/pits.10046

King, R. B. (2016). Gender differences in motivation, engagement and achievement are related to students' perceptions of peer—but not of parent or teacher—attitudes toward school. *Learning and Individual Differences*, *52*, 60–71. https://doi.org/10.1016/j.lindif.2016.10.006

Kubey, R. W., Lavin, M. J., & Barrows, J. R. (2001). Internet use and collegiate academic performance decrements: Early findings. *Journal of Communication*, *51*(2), 366–382. https://doi.org/10.1111/j.1460-2466.2001.tb02885.x

Kuhfeld, M., Soland, J., Tarasawa, B., Johnson, A., Ruzek, E., & Liu, J. (2020). Projecting the potential impact of COVID-19 school closures on academic achievement. *Educational Researcher*, *49*(8), 549–565. https://doi.org/10.3102/0013189X20965918

Lee, H. (2022). What drives the performance of Chinese urban and rural secondary schools: A machine learning approach using PISA 2018. *Cities*, *123*, 103609. https://doi.org/10.1016/j.cities.2022.103609

Lin, X.-F., Tiang, T., Luo, G., Huang, X., Li, W., Zou, T., Wang, Z., & Hu, Q. (2023). Mitigating the urban-rural digital divide: A dual scaffolding-embedded mobile augmented reality learning approach in the post COVID-19 pandemic. *Educational Technology & Society*, *26*(4), 108–122. https://doi.org/10.30191/ETS.202310_26(4).0008

Liu, Z., Duan, H., Liu, S., Mu, R., Liu, S., & Yang, Z. (2024). Improving knowledge gain and emotional experience in online learning with knowledge and emotional scaffolding-based conversational agent. *Educational Technology & Society*, *27*(2), 197–219. https://doi.org/10.30191/ETS.202404_27(2).RP08

Long, K., & Renbarger, R. (2023). Persistence of poverty: How measures of socioeconomic status have changed over time. *Educational Researcher*, *52*(3), 144–154. https://doi.org/10.3102/0013189X221141409

Luan, H., & Tsai, C.-C. (2021). A review of using machine learning approaches for precision education. *Educational Technology & Society*, *24*(1), 250–266.

Mælan, E. N., Gustavsen, A. M., Stranger-Johannessen, E., & Nordahl, T. (2021). Norwegian students' experiences of homeschooling during the COVID-19 pandemic. *European Journal of Special Needs Education*, *36*(1), 5–19. https://doi.org/10.1080/08856257.2021.1872843

Marks, G. N., Cresswell, J., & Ainley, J. (2006). Explaining socioeconomic inequalities in student achievement: The role of home and school factors. *Educational Research and Evaluation*, *12*(2), 105–128. https://doi.org/10.1080/13803610600587040

Mendonça, Y. V. S., Naranjo, P. G. V., & Pinto, D. C. (2022). The role of technology in the learning process: A decision tree-based model using machine learning. *Emerging Science Journal*, *6*, 280–295. https://doi.org/10.28991/ESJ-2022-SIED-020

Mensah, F. K., & Kiernan, K. E. (2010). Gender differences in educational attainment: Influences of the family environment. *British Educational Research Journal*, *36*(2), 239–260. https://doi.org/10.1080/01411920902802198

Miguéis, V. L., Freitas, A., Garcia, P. J. V., & Silva, A. (2018). Early segmentation of students according to their academic performance: A predictive modelling approach. *Decision Support Systems*, *115*, 36–51. https://doi.org/10.1016/j.dss.2018.09.001

Musso, M. F., Hernández, C. F. R., & Cascallar, E. C. (2020). Predicting key educational outcomes in academic trajectories: a machine-learning approach. *Higher Education*, *80*(5), 875–894. https://doi.org/10.1007/s10734-020-00520-7

Nielsen, D. (2016). *Tree Boosting With XGBoost - Why Does XGBoost Win "Every" Machine Learning Competition?*

Noble, W. S. (2006). What is a support vector machine? *Nature Biotechnology*, *24*(12), 1565–1567. https://doi.org/10.1038/nbt1206-1565

Nunes, C., Beatriz-Afonso, A., Cruz-Jesus, F., Oliveira, T., & Castelli, M. (2022). Mathematics and mother tongue academic achievement: A Machine learning approach. *Emerging Science Journal*, *6*, 137–149. https://doi.org/10.28991/ESJ-2022-SIED-010

Nunes, L. C., Balcão Reis, A., & Seabra, C. (2018). Is retention beneficial to low-achieving students? Evidence from Portugal. *Applied Economics*, *50*(40), 4306–4317. https://doi.org/10.1080/00036846.2018.1444261

OECD. (2019). OECD Employment outlook 2019: The future of work. In *The Future of Work*. OECD Publishing. https://doi.org/10.1787/9ee00155-en

Ömeroğulları, M., Guill, K., & Köller, O. (2020). Effectiveness of private tutoring during secondary schooling in Germany: Do the duration of private tutoring and tutor qualification affect school achievement? *Learning and Instruction*, *66*, 101306. https://doi.org/10.1016/j.learninstruc.2020.101306

Opdenakker, M.-C., & Van Damme, J. (2001). Relationship between school composition and characteristics of school process and their effect on mathematics achievement. *British Educational Research Journal*, *27*(4), 407–432.

Pokropek, A., Borgonovi, F., & Jakubowski, M. (2015). Socio-economic disparities in academic achievement: A comparative analysis of mechanisms and pathways. *Learning and Individual Differences*, *42*, 10–18. https://doi.org/10.1016/j.lindif.2015.07.011

Pov, S., Kawai, N., & Murakami, R. (2022). Identifying causes of lower secondary school dropout in Cambodia: a two-level hierarchical linear model. *International Journal of Inclusive Education*, *26*(8), 834–847. https://doi.org/10.1080/13603116.2020.1735542

Pressley, T. (2021). Factors contributing to teacher burnout during COVID-19. *Educational Researcher, 50*(5), 325–327. https://doi.org/10.3102/0013189X211004138

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. E*conometrica*, *73*(2), 417–458.

Rozgonjuk, D., Täht, K., & Vassil, K. (2021). Internet use at and outside of school in relation to low- and high-stakes mathematics test scores across 3 years. *International Journal of STEM Educatio*n, 8, 27. https://doi.org/10.1186/s40594-021-00287-y

Salmela-aro, K., & Tynkkynen, L. (2012). Gendered pathways in school burnout among adolescents. *Journal of Adolescence*, *35*(4), 929–939. https://doi.org/10.1016/j.adolescence.2012.01.001

Sampath, S., Khedr, A., Qamar, S., Tekin, A., Singh, R., Green, R., & Kashyap, R. (2021). Pandemics throughout the history. *Cureus*, *13*(9), e18136. https://doi.org/10.7759/cureus.18136

Santibañez, L., & Guarino, C. M. (2021). The effects of absenteeism on academic and social-emotional outcomes: Lessons for COVID-19. *Educational Researcher*, *50*(6), 392–400. https://doi.org/10.3102/0013189X21994488

Şen, B., Uçar, E., & Delen, D. (2012). Predicting and analyzing secondary education placement-test scores: A data mining approach. *Expert Systems with Applications*, *39*(10), 9468–9476. https://doi.org/10.1016/j.eswa.2012.02.112

Smola, A. J., & Scholkopf, Bernhard. (2004). A tutorial on support vector regression. *Statistics and Computing*, *14*, 199–222.

Spielman, A. I., & Sunavala-Dossabhoy, G. (2021). Pandemics and education: A historical review. *Journal of Dental Education*, *85*(6), 741–746. https://doi.org/10.1002/jdd.12615

Steinmayr, R., C. Dinger, F., & Spinath, B. (2010). Parents' education and children's achievement: The role of personality. *European Journal of Personality*, 1–16. https://doi.org/10.1002/per.755

Steinmayr, R., Lazarides, R., Weidinger, A. F., & Christiansen, H. (2021). Teaching and learning during the first COVID-19 school lockdown: Realization and associations with parent-perceived students' academic outcomes: A study and preliminary overview. *Zeitschrift Fur Padagogische Psychologie*, *35*(2–3), 85–106. https://doi.org/10.1024/1010-0652/a000306

Sultan, W. H., Woods, P. C., & Koo, A.-C. (2011). A constructivist approach for digital learning: Malaysian schools case study. *Educational Technology & Society*, *14*(4), 149–163.

Sun, Z., Yuan, Y., Xiong, X., Meng, S., Shi, Y., & Chen, A. (2024). Predicting academic achievement from the collaborative influences of executive function, physical fitness, and demographic factors among primary school students in China: ensemble learning methods. *BMC Public Health*, *24*(1). https://doi.org/10.1186/s12889-024-17769-7

Tang, K. H. D. (2023). Impacts of COVID-19 on primary, secondary and tertiary education: a comprehensive review and recommendations for educational practices. *Educational Research for Policy and Practice*, *22*(1), 23–61. https://doi.org/10.1007/s10671-022-09319-y

Tesfagiorgis, M., Tsegai, S., Mengesha, T., Craft, J., & Tessema, M. (2020). The correlation between parental socioeconomic status (SES) and children's academic achievement: The case of Eritrea. *Children and Youth Services Review*, *116*, 105242. https://doi.org/10.1016/j.childyouth.2020.105242

Trakunphutthirak, R., & Lee, V. C. S. (2022). Application of educational data mining approach for student academic performance prediction using progressive temporal data. *Journal of Educational Computing Research*, *60*(3), 742–776. https://doi.org/10.1177/07356331211048777

United Nations. (2015). *Transforming our world: the 2030 Agenda for Sustainable Development*. https://sdgs.un.org/2030agenda

Vandelannote, I., & Demanet, J. (2020). "What's high school got to do with it?" secondary school composition, school-wide social capital and higher education enrollment. *Research in Higher Education*, *62*(5), 680–708. https://doi.org/10.1007/s11162-020-09617-5

Vanderlinde, R., & Van Braak, J. (2011). A new ICT curriculum for primary education in Flanders: Defining and predicting teachers' perceptions of innovation attributes. *Educational Technology & Society*, *14*(2), 124–135.

Vanneschi, L., & Castelli, M. (2018). Multilayer perceptrons. *Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics*, *1–3*, 612–620. https://doi.org/10.1016/B978-0-12-809633-8.20339-7

Walters, T., Simkiss, N. J., Snowden, R. J., & Gray, N. S. (2022). Secondary school students' perception of the online teaching experience during COVID-19: The impact on mental wellbeing and specific learning difficulties. *British Journal of Educational Psychology*, *92*(3), 843–860. https://doi.org/10.1111/bjep.12475

Workman, J. (2023). Income inequality and student achievement: trends among US States (1992–2019). *Educational Review*, *75*(5), 871–893. https://doi.org/10.1080/00131911.2021.1974349

Yağcı, M. (2022). Educational data mining: prediction of students' academic performance using machine learning algorithms. *Smart Learning Environments*, *9*(1), 11. https://doi.org/10.1186/s40561-022-00192-z

Zhang, C., & Ma, Y. (2012). *Ensemble Machine Learning: Methods and applications*. https://doi.org/10.1007/978-1-4419-9326-7

Zhang, X., Dimitriou, D., & Halstead, E. J. (2021). Sleep, anxiety, and academic performance: A study of adolescents from public high schools in China. *Frontiers in Psychology*, *12*(July), 1–7. https://doi.org/10.3389/fpsyg.2021.678839

Zhao, S., Chen, L., King, R. B., Nie, Y., Du, H., & Chi, P. (2021). When adolescents believe that SES can be changed, they achieve more: The role of growth mindset of SES. *Personality and Individual Differences*, *183*, 111104. https://doi.org/10.1016/j.paid.2021.111104

Appendix A. Literature review of the AA drivers and methods used

References	Data	Covid effects?	Methods	St	Pa	Sc	Te
Hanushek and Kimko (2000)	Cognitive skills for 39 countries and 31 countries of economic performance	No. Pre-pandemic	Regressions models	x		x	
Hoxby (2000)	649 elementary schools with data from 1992-1993 to 1997-1998 and 146 elementary districts with data from 1986-1987 to 1997-1998, Connecticut, USA	No. Pre-pandemic	Regressions models	x		x	
Fan and Chen (2001)	Meta-analysis including 25 different studies	No. Pre-pandemic	General linear model	x	x		
Barnett et al. (2002)	152 high schools between 1994-1995 and 1995-1996 academic years in Northern Ireland	No. Pre-pandemic	Linear Programming techniques			x	
Parker et al. (2004)	667 students at a High School in Alabama (different grades)	No. Pre-pandemic	ANOVA tests	x			
Rockoff (2004)	10,000 students and 300 teachers from elementary schools in New Jersey. In District A between 1989-1990 to 2000-2001 and District B between 1989-1990 to 1999-2000 academic years	No. Pre-pandemic	Regressions models				x
Driessen et al. (2005)	Elementary school from the Netherlands, with over 500 schools and 12,000 students, in the academic year of 1994-1995	No. Pre-pandemic	Frequency, Variance, and Structural models	x	x	x	
Rivkin et al. (2005)	Public school students from Texas. Data for three cohorts between the 1993-1995 academic year	No. Pre-pandemic	Regression models			x	x
Archibald (2006)	Elementary schools from Nevada, USA, with more than 60,000 students, between the 2002-2003 academic year	No. Pre-pandemic	Hierarchical linear models	x		x	x
Jackson et al. (2006)	140 children (average age of 13.8) from the USA between December 2000 and June 2002	No. Pre-pandemic	Internet recorded	x			
Lee and Bowen (2006)	415 students, from 3rd until 5th grade, in the southeastern USA in 2004	No. Pre-pandemic	Hierarchical linear models	x	x		
Marks et al. (2006)	6,000 schools across 32 countries – PISA 2000	No. Pre-pandemic	Item Response Theory Regressions models	x	x	x	
Aaronson et al. (2007)	Panel data of Chicago Public Schools, from 1996/1997 to 1998/1999, 8th and 9th grade (n=53000, t=3)	No. Pre-pandemic	Ordinary Least Squares				x
Codjoe (2007)	Black students in Edmonton, Canada	No. Pre-pandemic	Interviews	x			
Croninger et al. (2007)	Longitudinal Study From Early Childhood, Kindergarten Class of 1998–1999	No. Pre-pandemic	Hierarchical linear models	x			x
Lee (2007)	80 high schools and 52 middle schools with students from grades 7 to 12, in 1994, in the USA	No. Pre-pandemic	Hierarchical linear models Classic linear regression model	x	x	x	
Jeynes (2007)	Meta-analysis with 52 studies ranging from 1972 to 2000	No. Pre-pandemic	Regression models		x		
Lei and Zhao (2007)	237 middle school students from Ohio, USA, between 2003–2004 academic year	No. Pre-pandemic	Hierarchical linear models ANOVA tests	x			
Steinmayr and Spinath (2008)	342 students, 11th and 12th graders, in Germany	No. Pre-pandemic	Regression models	x			
Caro et al. (2009)	6290 students from Canada's National Longitudinal Study	No. Pre-pandemic	Hierarchical linear models	x			

	during the 1994-2001 academic years		Panel data models			
Mensah and Kiernan (2010)	Millennium Cohort Study, with children in the first year of school in England, between 2005-2006 academic year	No. Pre-pandemic	Tobit regression models Univariate and Multivariate analyses	x	x	
Hartas (2011)	Longitudinal sample from Millennium Cohort Study, from England, for children between 3 and 5 years	No. Pre-pandemic	Univariate analyses of variance Chi-square tests		x	
Patterson and Pahlke (2011)	211 students in a public middle school in the USA from the academic years 2007 to 2011	No. Pre-pandemic	Regression models	x	x	
Casillas et al. (2012)	4,660 middle-school students,7th and 8th grade, at 24 middle schools from 13 districts in the USA	No. Pre-pandemic	Multilinear regression Hierarchical linear model	x		
Hanushek and Woessmann (2012)	64 different countries from 1964 to 2003	No. Pre-pandemic	Regression models	x		x
Şen et al. (2012)	5000 students from 8th grade in Turkey	No. Pre-pandemic	Regression models Trees Artificial Neural Networks Support Vector Machines	x	x	x
Brunner et al. (2013)	275,369 15th years old students from 41 nations – PISA 2003	No. Pre-pandemic	Multiple group factor analytic models Full maximum likelihood method "MLR"	x		
Wally-Dima and Mbekomize (2013)	660 bachelor students at the University of Botswana during the 2011-2012 academic year	No. Pre-pandemic	Descriptive statistics T-tests	x		
Bosworth (2014)	4th and 5th grade students from a public school in North Carolina, USA, during the 2000-2001 academic year	No. Pre-pandemic	Regression models	x		x
Jayakar and Liu (2014)	144 students from a tertiary institute in Singapore	No. Pre-pandemic	Correlation research design	x		
Krassel and Heinesen (2014)	9th and 10th-grade high school students in Denmark between 2003-2006 academic years	No. Pre-pandemic	Regression discontinuity design Control for school fixed effects Ordinary Least Squares	x	x	x
Vigdor et al. (2014)	Students 5th to 8th grade in public schools, between 2002-2005 academic years, North Carolina	No. Pre-pandemic	Probit regression Regression models	x		
Lee and Mallik (2015)	Students from the University of Western Sydney, from 2007 to 2012	No. Pre-pandemic	Ordinary Least Squares	x		
Hodis et al. (2015)	782 high school students in New Zeland	No. Pre-pandemic	Hierarchical linear models	x		
Pokropek et al. (2015)	29,0361 students from secondary schools in 33 OCDE countries, PISA data	No. Pre-pandemic	Structural Equations Model	x	x	
King (2016)	848 Filipino secondary students from two public schools in Manila	No. Pre-pandemic	Univariate analysis	x		
Abad and López (2017)	18,935 high school students from 99 schools in Baja California state, Mexico.	No. Pre-pandemic	Trees Clustering	x		x
Asif et al. (2017)	4-year bachelor's degree in Information Technology at a public engineering university in Pakistan, with 210 students	No. Pre-pandemic	Trees Hierarchical clustering	x		

	enrolled during the 2007-08 and 2008-09 academic years						
Miguéis et al. (2018)	2,459 students at a European engineering and technology school across five academic years (2003 to 2007) were tracked until 2015.	No. Pre-pandemic	Trees Support Vector Machines Naïve Bays	x			
Costa-Mendes et al. (2020)	Population of Public High School Students from Portugal	No. Pre-pandemic	Regression models Trees Artificial Neural Networks Support Vector Machines xGBoost	x		x	
Cruz-Jesus et al. (2020)	Population of Public High School Students from Portugal	No. Pre-pandemic	Regression models K – Nearest Neighbors Trees Artificial Neural Networks Support Vector Machines	x			
Musso et al. (2020)	655 students at a private university in Buenos Aires	No. Pre-pandemic	Artificial Neural Networks	x	x		
Tesfagiorgis et al. (2020)	Students of the grade 8 from two cities in Eritrea. Survey with 397 students and interviews with 32 of these students, 27 parents, 8 principals, 8 teachers and 4 educational officers.	No. Pre-pandemic	Regression models	x	x		
Vandelannote and Demanet (2020)	1,132 Flemish students from 30 schools	No. Pre-pandemic	Regression models	x		x	x
Zaccoletti et al. (2020)	567 parents from various regions in Italy (n = 173, 89% mothers) and Portugal (n = 394, 93% mothers)	Partially. Post-pandemic only.	Multi-group latent change score model	x			
Camacho et al. (2021)	394 Portuguese parents of students in grades 1 through 9	Partially. Post-pandemic only.	Multi-group latent change score model	x			x
Clark et al. (2021)	2,025 9th graders	Partially. Post-pandemic only.	Difference-in-differences estimation	x			
Hampton et al. (2021)	3,258 students aged 13 and older in grades eight to eleven from Michigan, USA	Partially. Post-pandemic only.	Regression models	x			
J. Lee et al. (2021)	268 middle school students in South Korea	Partially. Post-pandemic only.	Regression models	x			
K. Liu et al. (2021)	1,550 students and their parents from 8 middle schools located in eastern China	Partially. Post-pandemic only.	Regression models	x	x		x
Mælan et al. (2021)	1,755 students in grades 8 to 10 from 93 schools in Inland Norway	Yes. Pre- vs Post-pandemic	ANOVA tests	x			x
Martin et al. (2021)	1,548 Australian high school students in nine schools	Partially. Post-pandemic only.	Confirmatory Factor Analysis Structural Equation Modeling	x	x		
Santibañez and Guarino (2021)	600,000 students from California, with 4 years of data from 2014/2015 through 2017/2018	Partially. Post-pandemic only.	Fixed effect model	x		x	
S. J. Lee et al.	405 parents with at least one	Partially.	Regression models	x	x		

(2021)	child 0–12 years of age in the USA	Post-pandemic only.					
Spitzer and Musslick (2021)	10,693 K-12 German students (1,373 classes)	Yes. Pre- vs Post-pandemic	Linear mixed model	x			
Steinmayr et al. (2021)	2,647 parents from Germany	Partially. Post-pandemic only.	Structural Equation Modeling	x			x
X. Zhang et al. (2021)	99 adolescents aged 15–17 years old from two public schools in Baishan City, Jilin Province, China	Partially. Post-pandemic only.	Spearman's Rho correlations Regression models Parametric and non-parametric tests	x			
Albrecht et al. (2022)	8,972 adolescents from Swiss high schools	Yes. Pre- vs Post-pandemic	Regression models	x			
Barbosa-Camacho et al. (2022)	610 university students from Mexico	Partially. Post-pandemic only.	ANOVA tests	x	x		
Dean et al. (2023)	73,371 students and 772 high schools located in Australia in 2017	No. Pre-pandemic	Hierarchical linear models	x	x	x	
Hillier et al. (2022)	38,000 students aged 15 to age 25	No. Pre-pandemic	Linear Regression	x	x		
Kurt et al. (2022)	20 students and 22 teachers from public high schools in Turkey	Partially. Post-pandemic only.	Interviews	x			x
Lee (2022)	15-year-old students from 361 schools in China in 2018	No. Pre-pandemic	Linear Regression xGBoost	x		x	
Nunes et al. (2023)	220 Portuguese students	Partially. Post-pandemic only.	Structural Equations Model	x	x		x
Tang (2023)	42 studies	Yes. Pre- vs Post-pandemic	Comprehensive review	x	x	x	x

Notes. St – students' related variables; Pa – parents' related variables; Sc – School's related variables; and Te – Teachers' related variables.

Appendix B. R-squared of the models after 10-fold Cross-Validation for the 2019 dataset

	Mathematics exam grade 2019	Mother tongue exam grade 2019
Model		
MLR	0.15	0.14
DT	0.14	0.12
NN	0.12	0.12
RF	0.22	0.20
SVM	0.08	0.11
XGB	0.22	0.20