



Research Article

Predicting Elementary School GPA Using Machine Learning Approaches

Ivana Isailović¹  , Isidora Gatarić¹ , and Ilija Milovanović¹ 

¹*Department of Psychology, Faculty of Philosophy, University of Novi Sad, Serbia*

ABSTRACT

Predicting academic success, quantified as Grade Point Average (GPA), is one of the key research focuses in educational psychology, with ongoing debate regarding the most influential predictors. Previous studies suggest that cognitive readiness for school (e.g., attention, working memory, etc.) is among the strongest predictors of later academic achievement. In addition, social factors such as parents' education and gender also show consistent, though somewhat weaker, associations with GPA. However, many earlier studies have relied on traditional statistical models, such as linear regression, which assume linearity and often overlook complex, nonlinear interactions among variables. This limits their ability to uncover the true structure of influential predictors. In contrast, advanced machine learning (ML) methods, such as decision trees, random forest, and gradient boosting, can model such complexity, offering greater accuracy and deeper insight into predictor importance. This study applied these ML algorithms, along with linear regression, to predict GPA in 4th and 7th grade among 218 elementary school students, using measures of cognitive readiness and socio-demographic variables as predictors. Results indicated that linear regression and random forest yielded the most accurate predictions. The strongest predictors of GPA in both grades were measures of cognitive readiness (*Coding, Visual Memory, General Knowledge, Block Assembly*), while other predictors had minimal or no effect. These findings underscore the value of ML models in improving early identification of at-risk students and informing targeted academic support, while also illustrating the contexts in which traditional methods may still perform comparably.

Keywords: elementary school GPA, decision trees, random forest, gradient boosting, cognitive readiness for school

UDK: 37.015.3

DOI: [10.19090/pp.v19i1.2625](https://doi.org/10.19090/pp.v19i1.2625)

Received: 09.04.2025.

Revised: 14.06.2025.

Accepted: 14.07.2025.



Copyright © 2026 The Author(s).

This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

✉ Corresponding author's email: ivana.strizovic@gmail.com

Introduction

Grade Point Average (GPA) is widely recognized as a key measure of academic achievement at all levels of education. It is frequently used to evaluate a student's readiness for future educational and career opportunities (Imose & Barber, 2015). Studies have demonstrated that students with lower GPAs are more likely to face challenges in high school and beyond, including lower graduation rates and limited post-secondary education opportunities (e.g., Tyler & Lofstrom, 2009). Lower academic performance during this period is also associated with an increased risk of dropping out of formal education, which can have lasting adverse effects on an individual's career trajectory and socioeconomic status (Entwisle et al., 2005; Lawson et al., 2020). Given its critical role in long-term academic success, predicting GPA in elementary school provides valuable insights into early academic performance and potential future career outcomes (Entwisle et al., 2005; Lawson et al., 2020). Understanding GPA predictions involves examining the complex interplay of individual and contextual factors, such as cognitive abilities, socioeconomic background, and parental involvement, which collectively shape academic performance (Eccles & Roeser, 2011).

Cognitive determinants of GPA

A growing body of research highlights the critical role of cognitive readiness for elementary school, defined as a child's ability to engage in and benefit from classroom learning (Blair, 2002), in shaping long-term academic success. Key components such as verbal and visuospatial working memory (Bull et al., 2008; Soltanlou et al., 2019) and numerical representations (Stock et al., 2009) are consistently identified as robust predictors of achievement (Duncan et al., 2006; Jovanović et al., 2010; McClelland et al., 2017). Among them, early math and logical abilities are the strongest determinants of later performance, followed by reading skills and attentional control (Duncan et al., 2006). Soltanlou et al. (2019) found that children with superior working memory capacities achieve higher GPAs across subjects, highlighting the role of these faculties in learning, problem-solving, and retention. Although much of the research originates from Western cultures, similar findings were reported in Serbia (e.g., Jovanović et al., 2010). In Serbia, the most commonly used tools for assessing school readiness are the *New*

Belgrade Revision of the Binet-Simon Scale (Ivić et al., 1976) and *the First Grader Examination Test* (Ivić et al., 1995). More recently, *the School Maturity Test* (Novović et al., 2007) has gained popularity. It assesses memory, attention, visuomotor coordination, practical knowledge, and social adjustment. Its advantage lies in minimizing cultural, social, and linguistic bias, making it a reliable and valid cognitive measure (Biro et al., 2006; Jovanović et al., 2010).

Socio-economics, parental actors, and GPA prediction

Beyond cognitive abilities, socioeconomic status (SES) is a major determinant of academic performance. Research consistently demonstrates that students from higher-income families tend to have greater access to high-quality educational resources, including well-funded schools, private tutoring, and enrichment programs that foster cognitive and academic growth (Sirin, 2005). These advantages create cumulative effects that reinforce achievement gaps (Reardon, 2018). A key SES component is parental education, a strong predictor of student success. Educated parents often provide cognitive stimulation, academic support, and emphasize the value of education (Davis-Kean, 2005; Li et al., 2024; Tamayo Martinez et al., 2022). They are also better equipped to navigate the education system and secure opportunities such as advanced coursework (Lareau, 2018). Furthermore, students whose parents set high expectations, monitor homework, and engage with schools typically earn higher GPAs and test scores (Fan & Chen, 2001; Li et al., 2024; Naite, 2021). Moreover, previous research shows that the supportive, autonomy-granting parenting positively affects children's academic performances (Bucci Liddy et al., 2021; Jeynes, 2007).

Machine learning algorithms and GPA prediction

While prior research has identified GPA determinants (Imose & Barber, 2015; Li et al., 2024; Soltanlou et al., 2019), most relied on traditional statistical methods, such as regression analysis, which often miss the complexity and non-linearity of influencing factors. Recent advancements in machine learning (ML) have addressed these limitations, enabling researchers to use more sophisticated predictive models that account for

complex interactions between multiple predictors (Burke & Sass, 2013). ML approaches, particularly decision trees (DT), random forest (RF), and gradient boosting (GB), have demonstrated superior accuracy in GPA prediction compared to conventional statistical models (Zhang et al., 2020). These ML techniques provide distinct advantages over traditional regression models by uncovering hidden patterns in student performance data, facilitating early interventions, and improving educational decision-making (Baker & Inventado, 2014). By leveraging these models for GPA prediction, researchers can generate more precise predictions and develop targeted support strategies for at-risk students. Advanced techniques, such as deep learning and support vector machines, are often mentioned in GPA prediction. While deep learning models achieve high accuracy, they require large datasets and lack transparency, making them less suitable for educational policy applications where explainability is crucial (Lipton, 2016). Similarly, support vector machines, though effective for classification, are computationally demanding and less interpretable, especially in large, multi-dimensional datasets (Vapnik, 1995). Given these limitations, DT, RF, and GB are more suitable for GPA prediction in real-world educational datasets of limited size.

Accurate GPA prediction using ML has significant implications for policy, instruction, and support systems. Early identification of at-risk students enables targeted interventions, resource allocation, and personalized instruction (D'Mello & Graesser, 2012). International studies confirm ML's effectiveness in predicting achievement, using demographic, cognitive, and behavioral data. Cognitive abilities emerge as the most reliable predictors, followed by SES (Chen & Ding, 2023; Ojajuni et al., 2021; Rajendran et al., 2022). In Serbia, research on the application of ML in education is still in its early stages, with a handful of studies focusing on high school and college students (Gerasimović, 2012; Stanković, 2021). Gerasimović (2012) demonstrated that socioeconomic status and prior academic achievement can be successfully used to predict students' college choices. Stanković (2021) further explored ML applications in education, showing that student engagement metrics and assessment patterns can effectively predict academic success in higher education.

The present study

The review of the available literature suggests there is a notable research gap in exploring how early predictors (e.g., cognitive abilities, SES, etc.) contribute to GPA outcomes in elementary school. The present study aims to address this gap by evaluating the predictive performance of linear regression, DT, RF, and GB in predicting 4th and 7th – grade GPA, based on cognitive, demographic, and socioeconomic variables.

The decision to focus on these grades is grounded in key developmental and educational transitions that occur at these stages. Fourth grade marks a shift from basic skill acquisition to their application across subjects, requiring more cognitive flexibility and independence (Mullis et al., 2012). In Serbia, the fourth grade is also important as it marks the end of the first educational cycle, after which students transition from being taught by a single teacher to subject-specific teachers. Seventh grade coincides with early adolescence and greater academic complexity (Eccles & Roeser, 2011). In the Serbian educational system, a broader curriculum, including physics, chemistry, and a second foreign language, is introduced in the seventh grade, demanding cross-disciplinary integration and increasing cognitive load.

While linear regression is a foundational and widely used predictive model in educational research, it imposes certain statistical assumptions, such as linearity, homoscedasticity, and lack of multicollinearity, that are often violated in real-world educational data (Kutner et al., 2005). More flexible ML algorithms, such as RF, DT, and GB, are designed to accommodate nonlinearity, interaction effects, and complex, high-dimensional feature spaces without relying on these assumptions (Fernández-Delgado et al., 2014; Zhang et al., 2020). These models can better capture the multifaceted and nonlinear nature of educational performance, offering improved generalization, robustness to overfitting, and enhanced interpretability through tools like variable importance measures.

Importantly, this study does not frame linear regression and modern ensemble-based methods as belonging to entirely separate analytical paradigms. Rather, all models applied in this study fall within the broader domain of supervised ML. However, their methodological capacities differ significantly. Linear regression represents a simpler, parametric approach,

while ensemble models offer more adaptive, data-driven learning mechanisms capable of modeling complex relationships. Emphasizing these distinctions is essential for understanding the potential added value that newer ML approaches bring to educational prediction tasks (Smith & Lee, 2022).

Building on the findings of previous studies (e.g., Zhang et al., 2020), our study aims to compare predictive performance across models and identify the variables that most strongly contribute to GPA, thereby providing both theoretical and practical contributions. It is expected that ensemble methods, particularly RF, will yield superior predictive performance due to their ability to manage noisy, high-dimensional data and uncover non-obvious patterns. Additionally, cognitive readiness and SES-related variables, such as parental education and academic expectations, are hypothesized to emerge as key predictors. Through this comparative approach, the study aims to offer a nuanced perspective on the utility of different supervised ML models in educational research and to provide a foundation for their future integration into data-informed educational decision-making.

Method

Sample and procedure

This study involved 218 primary school students from Stara Pazova (112 male, 106 female), all native Serbian speakers, aged 13–14 years ($M = 13.39$, $SD = 0.49$). Data collection spanned three phases over approximately seven years. In the first phase, a certified school psychologist or pedagogue conducted cognitive and developmental assessments during the students' final preschool year, alongside a parental sociodemographic questionnaire. Academic performance data for 4th and 7th grades were obtained from school records. The 7th grade data were collected prospectively, while 4th grade GPA data were retrieved retrospectively.

In Serbia, preschool cognitive testing is required for school enrollment. Parents provided formal applications and signed informed consent forms permitting scientific use of their children's data. The study adhered to strict ethical standards, ensuring anonymity and confidentiality.

Due to consent restrictions, the dataset is not publicly available but may be accessed upon reasonable request and additional ethical approval.

Instruments and measures

The School Maturity Test

The School Maturity Test (Novović et al., 2007) includes five subtests assessing children's cognitive and developmental abilities. *General Knowledge* evaluates practical knowledge and social adaptability, aiming to reduce socio-educational bias. *Visual Memory* measures retention and attention. *Block Assembly* tests visuomotor coordination, organization, and planning, and *Coding* assesses coordination, learning, and concentration. Lastly, the *Visual Vocabulary* subtest assesses linguistic competence and serves as an alternative measure specifically designed for children from socially disadvantaged backgrounds or those whose native language is not Serbian. Since the sample did not include children from such categories, the *Visual Vocabulary* test was not included in this study. In previous research, the reliability coefficients for the subtests of the School Maturity Test are as follows: *General Knowledge* (.52), *Visual Memory* (.80), *Block Assembly* (.82), and *Visual Vocabulary* (.77), while the overall reliability is .87 when including the *Visual Vocabulary* subtest and .85 when excluding it (Novović et al., 2007).

The Sociodemographic Questionnaire for Parents

The Sociodemographic Questionnaire for Parents gathered data on parental and household characteristics, including age, education, employment type, financial status, and involvement in the child's learning. Parental perceptions of the child's academic potential and school interest were assessed using two single-item questions, asking parents to rate their child's expected academic performance (satisfactory to outstanding) and to indicate whether they believe their child is interested in school.

The Student Data Questionnaire

The Student Data Questionnaire, developed for the purposes of this study, included questions about the student's gender and age, as well as tasks aimed at assessing cognitive and developmental indicators. *Spatial orientation* was assessed through a task that required students to identify

positions of objects relative to each other and follow spatial instructions. *Temporal orientation* was measured using sequencing tasks that required students to place daily activities in chronological order and interpret time-related concepts. *Word analysis and synthesis skills* were evaluated through phonological tasks involving breaking down words into individual phonemes (analysis) and combining phonemes into meaningful words (synthesis). Additionally, the questionnaire included questions about the student's GPA at the end of the 4th and 7th grades, as reported by school records.

Statistical methods

Descriptive statistics were first computed to summarize information about the dataset. Subsequently, linear regression, DT regressor, RF regressor, and GB regressor were applied to predict GPA. The DT regressor partitions the data using recursive binary splits, assigning each region the mean value of the target variable (Breiman et al., 1984). RF improves predictive accuracy and reduces overfitting by aggregating the outputs of multiple trees trained on bootstrapped samples and random subsets of features (Breiman, 2001). GB builds trees sequentially, with each tree correcting the residual errors of its predecessors through gradient-based optimization (Friedman, 2001). These models were selected to represent a progression in complexity and predictive power, enabling a comprehensive comparison in terms of performance, interpretability, and generalization ability.

All analyses were performed in *Python 3* (Van Rossum & Drake, 2009) using the *Jupyter Notebook* environment. The *Pandas 2.2.3 package* (McKinney, 2010) was used for descriptive statistics, while *Matplotlib* (Hunter, 2007) was used for data visualization.

Results

As the dataset contained no missing values nor measurement errors, descriptive statistical analysis was performed (Table 1).

Table 1*The Results of Descriptive Statistical Analysis (N = 218)*

Variables	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Skewness</i>	<i>Kurtosis</i>	<i>Max</i>	<i>K-S</i>	<i>p</i>
Mother's education level	4.12	0.81	3	-0.22	-1.45	5	.25	.00
Father's education level	3.95	0.78	3	0.08	-1.37	5	.21	.00
Financial situation	1.94	0.82	1	0.11	-1.51	3	.24	.00
Parental perceptions of the child's academic potential	3.67	0.97	2	-0.29	-0.87	5	.24	.00
General Knowledge	10.28	2.92	6	0.05	-1.30	15	.13	.00
Visual Memory	10.32	2.79	6	0.09	-1.10	15	.10	.01
Block Assembly	11.15	3.39	6	-0.11	-1.34	16	.13	.00
Coding	10.98	3.45	5	-0.16	-1.45	16	.12	.00
Spatial orientation	1.49	0.50	1	0.06	-2.02	2	.34	.00
Temporal orientation	1.53	0.50	1	-0.13	-2.01	2	.36	.00
Word analysis skills	1.50	0.50	1	-0.02	-2.02	2	.34	.00
Word synthesis skills	1.47	0.50	1	0.13	-2.01	2	.36	.00
4 th grade GPA	4.29	0.58	3	-0.52	-0.46	5	.12	.00
7 th grade GPA	3.92	0.79	2	-0.60	-0.66	5	.12	.00

Table 1 indicates that parents reported limited financial resources and moderate education levels, with mothers being slightly more educated. The parents generally held positive perceptions of their child's academic

potential. Children exhibited high and variable scores on *General Knowledge*, *Visual Memory*, *Block Assembly*, and *Coding*. *Spatial* and *Temporal Orientation*, *Word Analysis*, and *Word Synthesis* showed minimal variation, suggesting ceiling or binary effects. Academic performance was strong in 4th grade but declined slightly by 7th grade, with increased variability, potentially reflecting heightened academic demands or shifts in student engagement. Regarding the distributions of the study variables, the values of skewness and kurtosis indicated systematic deviations from normality. Specifically, the consistently negative kurtosis values pointed to flatter (platykurtic) distributions, while the presence of mild skewness suggested asymmetry. Moreover, the results of the One-Sample Kolmogorov–Smirnov Test (K–S) confirmed that the distributions of all variables significantly deviated from normality ($p \leq .01$). Therefore, the assumption of normality was violated, suggesting that caution is warranted in further statistical analysis.

4th grade GPA prediction

For the modeling phase, two *Python* packages were used: *NumPy* (Van der Walt et al., 2011) and *Scikit-learn* (Pedregosa et al., 2011), while *Matplotlib* (Hunter, 2007) was used for data visualization. The process began with encoding categorical features into dummy/indicator variables. Next, all predictors were standardized, and the dataset was split into training (80%) and testing (20%) subsets, following standard ML validation procedures. Each ML model was then run individually, with results presented in Table 2 and Table 3, as well as in Figure 1.

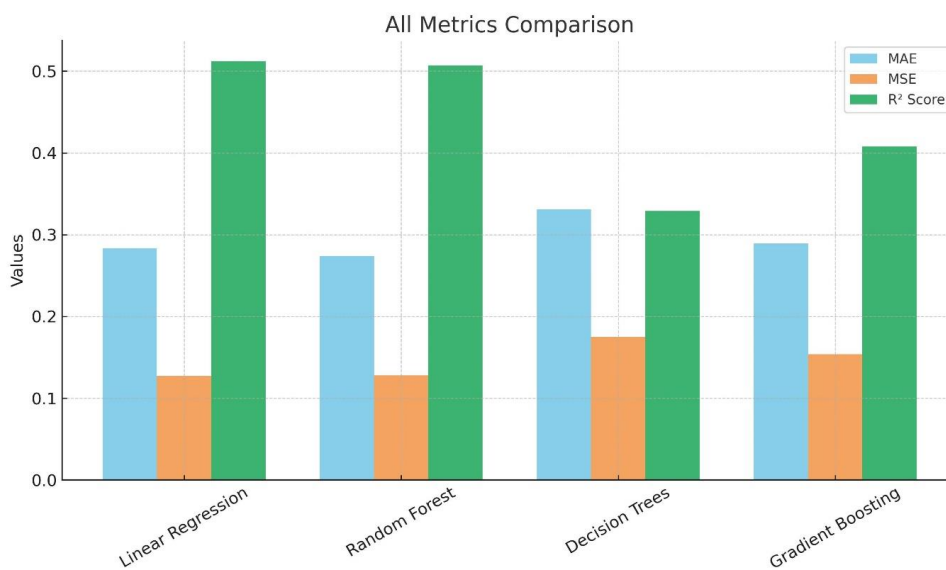
Table 2

The Results of All ML Models Used for 4th Grade GPA Prediction

Model	MAE	MSE	R ² Score
Linear Regression	.28	.13	.51
Decision Trees	.33	.17	.33
Random Forest	.27	.13	.51
Gradient Boosting	.29	.15	.41

Figure 1

Performance Comparison of All ML Models Used for 4th grade GPA Prediction



The results in Table 2 and Figure 1 indicate that all models had limited explanatory power. The RF model and the linear regression model demonstrated the highest explained variance, accounting for approximately 51% of the variance in the dependent variable (4th grade GPA). Those two models outperformed the other two (i.e., GB and DT). Additionally, the mean absolute error (MAE) was lowest in the RF model, while the mean squared error (MSE) was lowest in linear regression, suggesting that larger errors were more heavily penalized in the latter, indicating some variance in prediction errors between the two models. Similar trends were observed in the other two models (i.e., DT and GB), where lower R² values corresponded with higher MAE and MSE. The contribution of each predictor across all four models is presented in Table 3.

Table 3*Partial Contribution of Each Predictor in All ML Models for 4th Grade GPA Prediction*

Variables	Importance Score			
	Linear Regression	DT	RF	GB
Coding	.98	.29	.30	.30
Block Assembly	.69	.20	.16	.21
Visual Memory	.80	.17	.19	.20
General Knowledge	.67	.22	.18	.19
Parental perception of child's academic potential	.21	.04	.08	.06
Father's education level	.08	.03	.02	.01
Word synthesis skills	.16	.00	.01	.01
Mother's education level	.04	.01	.02	.01
Financial situation	.02	.00	.02	.01
Spatial orientation	.07	.00	.01	.01
Temporal orientation	.00	.01	.01	.00
Word analysis skills	.04	.00	.01	.00
Gender	.06	.00	.01	.00

Across all models (Table 3), *Coding* consistently emerged as the most important predictor of academic performance, highlighting its strong and stable predictive power. Other high-importance variables included *Visual Memory*, *Block Assembly*, and *General Knowledge*, all of which show substantial contributions across both linear and ensemble models, indicating the central role of cognitive abilities in predicting academic outcomes. It is important to note that linear regression shows higher partial contributions because it attributes importance based solely on direct linear effects without capturing complex interactions or nonlinearities that other ML models distribute across multiple variables (Molnar, 2020). On the other hand, the least important predictors across all models included *Temporal Orientation*, *Word Analysis Skills*, *Gender*, and *Financial Situation*, all of which showed minimal or no importance in 4th grade predictions. This suggests that demographic and socioeconomic variables contributed with little predictive value compared to direct cognitive skill measures. Notably,

while *Parental Perceptions of Child’s Academic Potential* showed moderate importance in linear regression, it was far less influential in tree-based models, indicating that subjective assessments were less robust predictors than objective cognitive measures in more complex models.

7th Grade GPA Prediction

The same *Python* packages used for the previous analysis were employed for the modeling section related to 7th grade GPA.

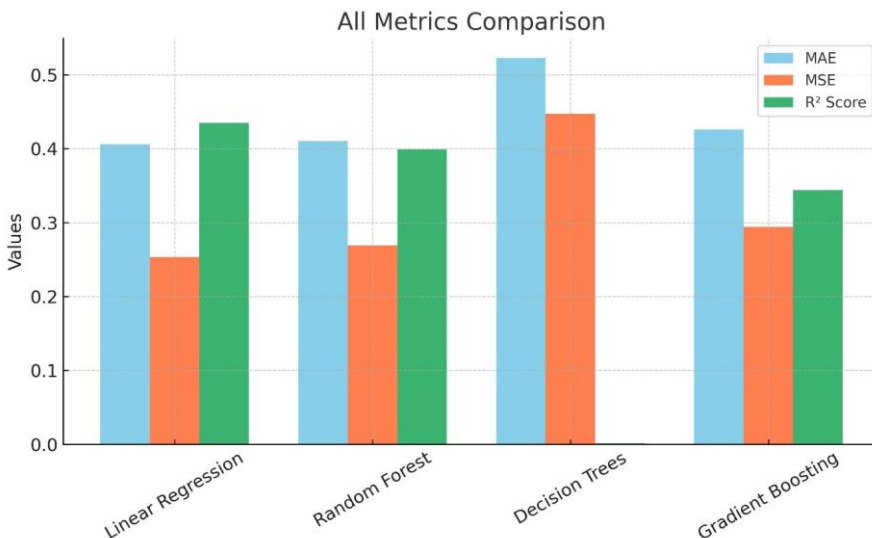
Table 4

The Results of All ML Models Used for 7th Grade GPA Prediction

Model	MAE	MSE	R ² Score
Linear Regression	.40	.25	.44
Decision Trees	.52	.45	.01
Random Forest	.40	.27	.40
Gradient Boosting	.43	.29	.34

Figure 2

Performance Comparison of All ML Models Used for 7th Grade GPA Prediction



The results presented in Table 4 and Figure 2 suggest that all ML models again had limited explanatory power (R^2 varied from .01 to .44). Again, the linear regression model explained the highest portion of the variance, followed by the RF model. The GB model performed similarly, while DT was the weakest model. A very low R^2 in a decision tree model may indicate that the model is too simple (e.g., too shallow tree) or that the data are non-linear for that structure. Furthermore, the MAE and MSE of linear regression model and RF model indicated that larger errors were penalized more. The GB model had similar MAE and MSE scores, but an identical pattern – lower MAE and MSE were followed by higher R^2 . The results of the impact of each feature in all models are listed in Table 5.

Table 5

Partial Contribution of Each Predictor in All ML Models for 7th Grade GPA Prediction

Variables	Importance Score			
	Linear Regression	DT	RF	GB
Coding	.99	.26	.27	.29
General Knowledge	.77	.24	.19	.22
Block Assembly	.70	.22	.16	.19
Visual Memory	.77	.17	.17	.16
Parental perceptions of child's academic potential	.33	.03	.10	.10
Financial situation	.02	.02	.02	.01
Father's education level	.12	.00	.02	.01
Mother's education level	.02	.02	.02	.01
Word Synthesis	.15	.01	.01	.01
Spatial orientation	.01	.00	.00	.00
Temporal orientation	.03	.00	.01	.00
Gender	.05	.00	.00	.00
Word analysis skills	.07	.00	.00	.00

The results presented in Table 5 suggest that *Coding* stood out as the most influential predictor across all models, indicating a consistently strong role in predicting academic outcomes. Other highly influential predictors included *General Knowledge*, *Visual Memory*, and *Block Assembly*, reinforcing the dominance of cognitive abilities as key predictors regardless of the model used. On the other hand, the least important predictors across all models included *Gender*, *Spatial Orientation*, *Word Analysis Skills*, and *Temporal Orientation*, all of which showed negligible or no contribution. Similarly, demographic variables such as *Father's Education Level*, *Mother's Education Level*, and *Financial Situation* had minimal predictive value. Interestingly, *Parental Perception of Child's Academic Potential* retained moderate importance in linear regression, but lost influence in all other models, suggesting subjective evaluations contribute less to performance prediction when more complex, non-linear relationships are modeled. As was the case in the prediction of 4th grade GPA, linear regression again showed higher partial contributions for several predictors, as it assigns importance based solely on direct linear relationships, without accounting for complex interactions or nonlinear effects that are more accurately captured and distributed across variables in machine learning models (Molnar, 2020).

Discussion

The comparative analysis revealed differences in predictive performance among the tested algorithms. In 4th grade, linear regression and RF achieved the same R^2 , although RF yielded a lower MSE. In 7th grade, linear regression slightly outperformed all models across all evaluation metrics, indicating that, despite its simplicity, it effectively captures relevant linear patterns. Nonetheless, this does not diminish the value of RF, particularly in situations where the assumptions of linear models may be violated. Linear regression assumes linearity, constant variance, and normally distributed residuals (Kutner et al., 2005), which often do not hold in real-world educational data - an issue also reflected in our findings, where skewness, kurtosis, and the Kolmogorov–Smirnov test indicated deviations from normality. Its higher accuracy may reflect data-specific factors, such as the train-test split or overfitting to linear trends. In contrast, RF, a nonparametric ensemble method, builds multiple decorrelated trees via bootstrap

aggregation and random feature selection (Breiman, 2001), enabling them to capture nonlinearities and interactions that linear models miss. Although they may have slightly higher variance in low-noise or small-sample settings, they reduce bias when data involve multicollinearity, threshold effects, or non-additive influences (Hastie et al., 2009). They're also more robust to outliers and offer internal out-of-bag error estimates, improving reliability. Given these strengths, RF offers a more robust, theoretically sound framework. Even if they do not always yield the lowest prediction error, their flexibility and robustness make them a valuable tool for educational prediction tasks involving nonlinear patterns and complex interactions.

4th Grade GPA prediction

Within all models related to the 4th grade GPA prediction, the *Coding* subtest had the strongest predictive value, underscoring the importance of processing speed and attention in academic performance (Fry & Hale, 2000). These results align with previous research, highlighting processing speed as a crucial factor in learning, particularly in symbol decoding and problem-solving tasks (Demetriou et al., 2014). *Visual Memory* was another significant predictor, reinforcing the role of working memory in academic achievement, which aligns with previous research suggesting that visuospatial working memory is critical for success in math and science (Alloway & Alloway, 2010; Bull et al., 2008). Additionally, *General Knowledge* was a strong predictor, supporting theories of crystallized intelligence that emphasize accumulated knowledge and experience as key factors in academic success (Cattell, 1987; Neisser et al., 1996). The predictive significance of *Block Assembly* highlights the role of spatial reasoning, a cognitive skill linked to problem-solving, in academic success (Wai et al., 2009; Verdine et al., 2014).

Spatial Orientation and *Temporal Orientation* had minimal or no impact on GPA. While relevant for daily functioning, these skills appear less crucial for structured academic tasks relying on working memory and reasoning (Diamond, 2013). Gender and parents' education level also had negligible influence, suggesting that cognitive and environmental factors outweigh biological sex differences in academic performance and offer limited predictive utility for GPA at this developmental stage (Halpern et al., 2007; Hyde, 2014). Similarly, *Word Analysis and Synthesis Skills* had limited

predictive values, likely because literacy skills are well-established by later elementary grades (Torgesen, 2002).

7th Grade GPA prediction

Results concerning the 7th grade GPA differed from those related to the 4th grade. While *Coding* remained the strongest predictor, its relative influence declined, and overall model R^2 scores were lower across all models. *General Knowledge*, *Block Assembly*, and *Visual Memory* retained some predictive power but were no longer dominant, which suggests that the explanatory power of early cognitive variables diminishes over time. This may be tentatively interpreted in light of Vygotsky's (1978) socio-cultural theory, which emphasizes the growing role of external factors such as instruction, environment, and motivation. Socioeconomic and demographic variables, such as *Mother's* and *Father's Education Level*, *Financial Situation*, and *Gender*, remained weak predictors across all models. The low importance of *Mother's Education Level* challenges prior findings that parental education strongly affects academic achievement (Davis-Kean, 2005), possibly indicating that school-based factors in this sample mitigate socioeconomic disadvantages.

Implications, limitations and future directions

The findings of this study have important implications for educational policy and intervention strategies. Given that cognitive readiness assessments were found to be strong predictors of GPA, early screening and targeted support programs should be implemented to help students with lower readiness levels. Schools could benefit from incorporating ML models into early warning systems (EWS) to identify at-risk students and provide timely interventions, such as personalized tutoring, cognitive skill development programs, and socio-emotional support (Bowers & Zhou, 2019; Jayaprakash et al., 2014). Furthermore, while RF did not outperform linear regression in this instance, its flexibility, robustness, and interpretability make it well-suited for complex educational data. It is especially valuable for identifying nuanced predictor interactions and offering actionable insights beyond what linear coefficients can provide.

This study has some limitations. While it included many predictors, factors like student motivation, peer influence, and teacher quality were not measured. Future research should address these gaps. Additionally, Serbia's centralized education system, with standardized curricula and GPA-based evaluations, may not be directly comparable to more flexible systems. Thus, findings may not fully apply to countries with different educational policies and socio-cultural influences. Cross-cultural studies could clarify whether key predictors of academic success are universal or system-dependent.

Nevertheless, this study underscores the potential of ML models, particularly RF, in predicting academic performance in primary school GPA. By identifying key predictors of GPA, this research contributes to the growing field of educational data mining and highlights the practical applications of advanced machine learning models in educational settings. The findings suggest that integrating various predictive models into school systems could facilitate early intervention strategies, ultimately improving personal learning, student outcomes, and informing data-driven policy decisions. Future research should build upon these results by incorporating additional predictors, expanding to diverse educational contexts, and exploring advanced ML techniques to further refine academic forecasting models.

Conflict of interest

We have no conflicts of interest to disclose.

Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

- Alloway, T. P., & Alloway, R. G. (2010). Investigating the predictive roles of working memory and IQ in academic attainment. *Journal of Experimental Child Psychology*, 106(1), 20–29. <https://doi.org/10.1016/j.jecp.2009.11.003>
- Baker, R. S. J. d., & Inventado, P. S. (2014). Educational data mining and learning analytics. In J. A. Larusson & B. White (Eds.), *Learning analytics: From research to practice* (pp. 61–75). Springer. https://doi.org/10.1007/978-1-4614-3305-7_4

- Blair, C. (2002). School readiness: Integrating cognition and emotion in a neurobiological conceptualization of children's functioning at school entry. *American Psychologist*, 57(2), 111–127. <https://doi.org/10.1037/0003-066X.57.2.111>
- Biro, M., Novović, Z. i Tovilović, S. (2006). Kognitivno funkcionisanje edukativno zapuštene dece predškolskog uzrasta [Cognitive functioning of educationally deprived pre-school children]. *Psihologija*, 39(2), 183–206. <https://doiserbia.nb.rs/Article.aspx?id=0048-57050602183B>
- Bowers, A. J., & Zhou, X. (2019). Receiver Operating Characteristic (ROC) Area under the curve (AUC): A diagnostic measure for evaluating the accuracy of predictors of education outcomes. *Journal of Education for Students Placed at Risk (JESPAR)*, 24(1), 20–46. <https://doi.org/10.1080/10824669.2018.1523734>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees*. Routledge. <https://doi.org/10.1201/9781315139470>
- Bucci Liddy, C. M., Brumariu, L. E., Diaconu-Gherasim, L. R., & Hunter, D. M. (2021). Maternal parenting and teaching strategies: relations with children's academic competence. *Educational Studies*, 1–16. <https://doi.org/10.1080/03055698.2021.1940870>
- Bull, R., Espy, K. A., & Wiebe, S. A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: Longitudinal predictors of mathematical achievement at age 7 years. *Developmental Neuropsychology*, 33(3), 205–228. <https://doi.org/10.1080/87565640801982312>
- Burke, M. A., & Sass, T. R. (2013). Classroom peer effects and student achievement. *Journal of Labor Economics*, 31(1), 51–82. <https://doi.org/10.1086/666653>
- Cattell, R. B. (1987). *Intelligence. Its structure, growth, and action*. North-Holland.
- Chen, X., & Ding, W. (2023). Early prediction of students' academic performance using machine learning techniques. *Journal of Educational Computing Research*, 61(2), 371–392. <https://doi.org/10.1177/07356331221075930>
- Davis-Kean, P. E. (2005). The influence of parent education and family income on child achievement: The indirect role of parental expectations and the home environment. *Journal of Family Psychology*, 19(2), 294–304. <https://doi.org/10.1037/0893-3200.19.2.294>

- Demetriou, A., Spanoudis, G., Shayer, M., van der Ven, S., Brydges, C. R., Kroesbergen, E., Podjarny, G., & Swanson, H. L. (2014). Relations between speed, working memory, and intelligence from preschool to adulthood: Structural equation modeling of 14 studies. *Intelligence*, 46, 107–121. <https://doi.org/10.1016/j.intell.2014.05.013>
- Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, 64(1), 135–168. <https://doi.org/10.1146/annurev-psych-113011-143750>
- D'Mello, S. K., & Graesser, A. C. (2012). AutoTutor and affective AutoTutor: Learning by talking with cognitively and emotionally intelligent computers that talk back. *ACM Transactions on Interactive Intelligent Systems*, 2(4), Article 23. <https://doi.org/10.1145/2395123.2395128>
- Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Pagani, L. S., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., & Japel, C. (2006). School readiness and later achievement. *Developmental Psychology*, 43(6), 1428–1446. <https://doi.org/10.1037/0012-1649.43.6.14284o>
- Eccles, J. S., & Roeser, R. W. (2011). Schools as developmental contexts during adolescence. *Journal of Research on Adolescence*, 21(1), 225–241. <https://doi.org/10.1111/j.1532-7795.2010.00725.x>
- Entwisle, Doris R., Alexander, Karl L., & Olson, L. (2005). First grade and educational attainment by Age 22: A New Story. *American Journal of Sociology*, 110(5), 1458–1502. <https://doi.org/10.1086/428444>
- Fan, X., & Chen, M. (2001). Parental involvement and students' academic achievement: A meta-analysis. *Educational Psychology Review*, 13(1), 1–22. <https://doi.org/10.1023/a:1009048817385>
- Fernández-Delgado, M., Cernadas, E., Barro, S., & Amorim, D. (2014). Do we need hundreds of classifiers to solve real world classification problems? *Journal of Machine Learning Research*, 15(1), 3133–3181. <https://doi.org/10.5555/2627435.2697065>
- Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, 29(5), 1189–1232. <https://doi.org/10.1214/aos/1013203451>
- Fry, A. F., & Hale, S. (2000). Relationships among processing speed, working memory, and fluid intelligence in children. *Biological Psychology*, 54(1-3), 1–34. [https://doi.org/10.1016/s0301-0511\(00\)00051-x](https://doi.org/10.1016/s0301-0511(00)00051-x)
- Gerasimović, M. (2012). *Prilog razvoju metodologija predviđanja i odlučivanja primenom veštačkih neuronskih mreža* [Contribution to development of

methodology for prediction and decision-making by applying artificial neural networks]. Doctoral dissertation, University of Belgrade, Faculty of Mechanical Engineering.

- Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S., & Gernsbacher, M. A. (2007). The science of sex differences in science and mathematics. *Psychological Science in the Public Interest*, 8(1), 1–51. <https://doi.org/10.1111/j.1529-1006.2007.00032.x>
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer. <https://doi.org/10.1007/978-0-387-84858-7>
- Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science & Engineering*, 9(3), 90–95. <https://doi.org/10.1109/MCSE.2007.55>
- Hyde, J. S. (2014). Gender similarities and differences. *Annual Review of Psychology*, 65(1), 373–398. <https://doi.org/10.1146/annurev-psych-010213-115057>
- Imose, R., & Barber, L. K. (2015). Using undergraduate grade point average as a selection tool: A synthesis of the literature. *The Psychologist-Manager Journal*, 18(1), 1–11. <https://doi.org/10.1037/mgr0000025>
- Ivić, I. D., Milinković, M., Rosandić, R. i Smiljanić-Čolanović, V. (1976). *Priručnik za NBS [Manual for NBS]*. Zavod za udžbenike i nastavna sredstva.
- Ivić, I., Milinković, M., Pešikan, A. i Bukvić, A. (1995). *Test za ispitivanje prvaka (TIP-1) – Priručnik [Test for Assessing First-Graders (TIP-1) – Manual]*. Društvo Psihologa Srbije.
- Jayaprakash, S. M., Moody, E. W., Lauría, E. J. M., Regan, J. R., & Baron, J. D. (2014). Early alert of academically at-risk students: An open source analytics initiative. *Journal of Learning Analytics*, 1(1), 6–47. <https://doi.org/10.18608/jla.2014.11.3>
- Jeynes, W. H. (2007). The relationship between parental involvement and urban secondary school student academic achievement. *Urban Education*, 42(1), 82–110. <https://doi.org/10.1177/0042085906293818>
- Jovanović, V., Tovilović, S., Novović, Z., & Biro, M. (2010). Prediction of school achievement and conative functioning among roma and non-roma children. *Primenjena Psihologija*, 3(3), 239–254. <https://doi.org/10.19090/pp.2010.3.239-254>
- Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2005). *Applied linear statistical models*. McGraw-hill.

- Lareau, A. (2018). *Unequal childhoods: Class, race, and family life*. In D. B. Grusky & J. Hill (Eds.), *Inequality in the 21st century* (pp. 444–451). Routledge. <https://doi.org/10.1525/9780520949904>
- Lawson, H., Lira, A., Lee, A., Lee, M., Linnenbrink-Garcia, L., Walton, S. P., & Briedis, D. (2020). Predicting engineering student success: An examination of college entrance exams, high school GPA, perceived competence, engineering achievement, and persistence. *2020 ASEE Virtual Annual Conference Content Access Proceedings*. <https://doi.org/10.18260/1-2--35072>
- Li, J., Xue, E., & You, H. (2024). Parental educational expectations and academic achievement of left-behind children in china: The mediating role of parental involvement. *Behavioral Sciences*, 14(5), 371–371. <https://doi.org/10.3390/bs14050371>
- Lipton, Z. C. (2016). The mythos of model interpretability. *Proceedings of the 2016 ICML Workshop on Human Interpretability in Machine Learning (WHI 2016)*, 96–100. <https://arxiv.org/abs/1606.03490>
- McClelland, M. M., Tominey, S. L., Schmitt, S. A., & Duncan, R. (2017). SEL interventions in early childhood. *The Future of Children*, 27(1), 33–47. <https://doi.org/10.1353/foc.2017.0002>
- McKinney, W. (2010). Data structures for statistical computing in Python. In: S. Van der Walt & J. Millman (Eds.). *Proceedings of the 9th Python in Science Conference* (pp. 51–56). <https://doi.org/10.25080/Majora-92bf1922-00a>
- Molnar, C. (2020). *Interpretable machine learning*. Mucbook Clubhouse.
- Mullis, I. V. S., Martin, M. O., Foy, P., & Arora, A. (2012). *TIMSS 2011 international results in mathematics*. TIMSS & PIRLS International Study Center, Boston College. <https://timss.bc.edu/timss2011/international-results-mathematics.html>
- Naite, I. (2021). Impact of parental involvement on children's academic performance at crescent international school, Bangkok, Thailand. *IOP Conference Series: Earth and Environmental Science*, 690(1), 012064. <https://doi.org/10.1088/1755-1315/690/1/012064>
- Neisser, U., Boodoo, G., Bouchard, T. J., Jr., Boykin, A. W., Brody, N., Ceci, S. J., Halpern, D. F., Loehlin, J. C., Perloff, R., Sternberg, R. J., & Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, 51(2), 77–101. <https://doi.org/10.1037/0003-066X.51.2.77>
- Novović, Z., Biro, M., Baucal, A., i Tovilović, S. (2007). *Test zrelosti za školu [School Maturity Test]*. Društvo Psihologa Srbije.

- Ojajuni, O., Ayeni, F., Akodu, O., Ekanoye, F., Adewole, S., Ayo, T., Misra, S., & Mbarika, V. (2021). Predicting student academic performance using machine learning. *Computational Science and Its Applications – ICCSA 2021*, 481–491. https://doi.org/10.1007/978-3-030-87013-3_36
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830. <http://www.jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf>
- Rajendran, S., Chamundeswari, S., & Sinha, A. A. (2022). Predicting the academic performance of middle- and high-school students using machine learning algorithms. *Social Sciences & Humanities Open*, 6(1), 100357. <https://doi.org/10.1016/j.ssaho.2022.100357>
- Reardon, S. F. (2018). The widening academic achievement gap between the rich and the poor. *Inequality in the 21st Century*, 177–189. <https://doi.org/10.4324/9780429499821-33>
- Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research*, 75(3), 417–453. <https://doi.org/10.3102/00346543075003417>
- Smith, J., & Lee, A. (2022). Advancing predictive modeling in education: From linear regression to ensemble methods. *Journal of Educational Data Science*, 10(3), 123–145. <https://doi.org/10.1234/jeds.2022.10307>
- Soltanlou, M., Artemenko, C., Dresler, T., Fallgatter, A. J., Ehlis, A.-C., & Nuerk, H.-C. (2019). Math anxiety in combination with low visuospatial memory impairs math learning in children. *Frontiers in Psychology*, 10. <https://doi.org/10.3389/fpsyg.2019.00089>
- Stanković, N. Lj. (2021). *Faktori učenja i predviđanje uspešnosti u programiranju primenom veštačkih neuronskih mreža* [Factors of learning and predicting success in programming using artificial neural networks]. Doctoral dissertation, University of Kragujevac, Faculty of Technical Sciences.
- Stock, P., Desoete, A., & Roeyers, H. (2009). Predicting arithmetic abilities. *Journal of Psychoeducational Assessment*, 27(3), 237–251. <https://doi.org/10.1177/0734282908330587>
- Tamayo Martinez, N., Xerxa, Y., Law, J., Serdarevic, F., Jansen, P. W., & Tiemeier, H. (2022). Double advantage of parental education for child educational achievement: the role of parenting and child intelligence. *European Journal*

- of Public Health*, 32(5), 690–695.
<https://doi.org/10.1093/eurpub/ckac044>
- Torgesen, J. K. (2002). The prevention of reading difficulties. *Journal of School Psychology*, 40(1), 7–26. [https://doi.org/10.1016/s0022-4405\(01\)00092-9](https://doi.org/10.1016/s0022-4405(01)00092-9)
- Tyler, J. H., & Lofstrom, M. (2009). Finishing high school: Alternative pathways and dropout recovery. *The Future of Children*, 19(1), 77–103.
<https://www.jstor.org/stable/27795036>
- Van der Walt, S., Colbert, S. C., & Varoquaux, G. (2011). The NumPy array: A structure for efficient numerical computation. *Computing in Science & Engineering*, 13(2), 22–30. <https://doi.org/10.1109/MCSE.2011.37>
- Van Rossum, G., & Drake, F. L. (2009). *Python 3*. Python Software Foundation.
<https://www.python.org/>
- Vapnik, V. N. (1995). *The nature of statistical learning theory*. Springer.
<https://doi.org/10.1007/978-1-4757-2440-0>
- Verdine, B. N., Irwin, C. M., Golinkoff, R. M., & Hirsh-Pasek, K. (2014). Contributions of executive function and spatial skills to preschool mathematics achievement. *Journal of Experimental Child Psychology*, 126, 37–51. <https://doi.org/10.1016/j.jecp.2014.02.012>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (Vol. 86). Harvard University Press.
- Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology*, 101(4), 817–835. <https://doi.org/10.1037/a0016127>
- Zhang, P., Wu, H.-N., Chen, R.-P., Dai, T., Meng, F.-Y., & Wang, H.-B. (2020). A critical evaluation of machine learning and deep learning in shield-ground interaction prediction. *Tunnelling and Underground Space Technology*, 106, 103593. <https://doi.org/10.1016/j.tust.2020.103593>

