Enhance Student Well-being and Digital Literacy with Machine Learning and Spatial Analysis

Fabrizio Benelli[1,†], Erdet Këlliçi[2,†], Franco Maciariello[1,†], Claudio Salvadori[3,†] and Vittorio Stile[1,*,†]

[1]*Universitas Mercatorum, Piazza Mattei 10, Rome, 00186, Italy*
[2]*Tirana Business University College, Rruga Rezervat e Shtetit, Tiranë, 1023, Albania*
[3]*New Generation Sensors srl, Via Cisanello 38, Pisa (PI), 56124, Italy*

Abstract

Advanced data analytics and machine learning can reshape the way schools cultivate both digital skills and student well-being. We analysed data from three Italian high-school classes ($N = 64$), combining random-forest and neural-network predictors with Spatial Autoregressive and Geographically Weighted Regression models. The approach captures how individual attributes, classroom geography and peer interactions jointly influence learning. Average grades rose from 5.34 to 6.15 and well-being scores from 0.48 to 0.95 over one semester. Spatial estimates ($\rho = 0.31$, $p < 0.01$) show that sitting next to high achievers yields a mean gain of 0.38 grade points, while local pockets of well-being amplify the effect of digital-literacy growth on performance. The results may support that digital-literacy interventions, when delivered in spatially aware learning environments, produce measurable academic and affective benefits. The study offers a reproducible pipeline that blends machine-learning prediction with spatial econometrics and provides evidence to guide data-driven, equitable strategies for classroom design, teacher training and student support.

Keywords

Digital literacy, Machine learning, Student well-being, Spatial econometrics, Educational analytics,

1. Introduction

Numerous applications of Artificial Intelligence (AI) have emerged in education. For example, Khan Academy's *Khanmigo*, powered by GPT-4, is already being piloted as an AI tutor that delivers personalised feedback and Socratic scaffolding across subjects such as mathematics and language learning [1]. Similarly, *Duolingo* uses sophisticated AI systems to enhance language-learning experiences.

In educational robotics, SoftBank Robotics' *Nao* and *Pepper* robots are increasingly adopted as social tutors for language learning, highlighting the promise of robot-assisted language education [2]. The field of Artificial Intelligence in Education (AIED) is rapidly evolving, attracting significant investments. The global AIED market, valued at $1.82 billion in 2021, is projected to grow at a compound annual growth rate of 36% from 2022 to 2030 [3].

Recent studies show AI-enabled adaptive learning improves student test outcomes by 62%, and general AI use enhances performance by 30% while reducing anxiety by 20% [4]. Research on AIED has surged, exploring design, effectiveness, and outcomes [5].

These trends motivate continued research on effective, ethical, and equitable AI integration. Parallel evidence from innovative Small and Medium-sized Enterprise (SME)s shows that technology-driven cultures accelerate adoption curves and learning cycles [6].

2nd Workshop on Education for Artificial Intelligence (edu4AI 2025, https://edu4ai.di.unito.it/), Co-located with ECAI 2025, the 28th European Conference on Artificial Intelligence which will take place on October 26, 2025 in Bologna, Italy

*Corresponding author.
†These authors contributed equally.
fabrizio.benelli@gmail.com (F. Benelli); erdet.kellici@tbu.edu.al (E. Këlliçi); franco.maciariello@studenti.unimercatorum.it (F. Maciariello); claudio.salvadori@ngs-sensors.it (C. Salvadori); vittoriostile@gmail.com (V. Stile)
https://www.franco.cloud/ (F. Maciariello); https://linktr.ee/vstile (V. Stile)
0009-0001-2722-2223 (F. Benelli); 0000-0003-0681-1456 (E. Këlliçi); 0009-0006-9532-1609 (F. Maciariello); 0000-0003-1363-3379 (C. Salvadori); 0009-0007-6492-4201 (V. Stile)

The ongoing development of AI technologies highlights their potential to foster innovative pedagogical strategies and may improve educational outcomes, emphasizing the need for persistent research and investment in this dynamic and impactful field.

1.1. Research Objectives and Hypotheses

This study aims to explore how advanced data analytics and Machine Learning (ML) can inform educational settings, focusing on spatial statistical analysis. We investigate how digital literacy and student well-being propagate within student communities influenced by physical proximity and digital interactions.

By leveraging advanced data analytics and ML, we aim to provide empirical insights to guide educational strategies and interventions tailored to enhance digital literacy and well-being across varied educational landscapes.

- Improvements in digital literacy are associated with students' academic performance and overall well-being [7, 8].
- Spatial factors, such as classroom arrangement and digital interactions, crucially influence educational outcomes, reflecting spatial dependencies and spillover effects [9].

We employ a combination of traditional statistical techniques and ML models, including spatial econometric models like the Spatial Autoregressive (SAR) model, to examine these hypotheses (refer to methodology section for details). This integrated approach provides a comprehensive framework for understanding the dynamics of digital literacy and student well-being in educational settings, contributing to the growing body of knowledge in this field.

2. Literature Review

The integration of Big Data and AI in education is transforming how educational content is delivered, personalized, and assessed. These technologies enable the collection and analysis of vast amounts of data, offering deeper insights into student performance, engagement, and well-being [10, 11]. AI-driven analytics can enhance educational experiences by providing personalized learning paths and real-time feedback.

Big Data in education aggregates diverse data sources, such as student demographics, academic records, and interaction logs. This allows educators to identify at-risk students, tailor interventions, and predict performance outcomes [12]. Predictive analytics is crucial for developing strategies that address individual student needs, thereby improving overall educational outcomes [13].

Beyond personalized learning, AI can automate administrative tasks like grading and attendance tracking, improving operational efficiency [14]. AI and Big Data can also help bridge educational inequalities by providing scalable educational resources, reaching underserved and remote areas [15]. During the COVID-19 pandemic, AI-enabled online learning platforms maintained educational continuity, demonstrating their resilience [16].

AI and Big Data also contribute to educational research by enabling the analysis of large datasets, revealing complex educational phenomena and insights [9]. However, integrating AI in education requires addressing ethical considerations, including data privacy and the digital divide [17, 18]. Explainable AI has been proposed as a key enabler of transparency and fairness [19].

The role of Big Data and AI in education is multifaceted, enhancing personalized learning, administrative efficiency, and educational equity. Continued research and investment in these technologies are essential for realizing their full potential in transforming education.

2.1. Previous Research on Spatial Effects in Educational Settings

Research on spatial effects in education focuses on how physical and digital learning environments influence outcomes like student performance, engagement, and well-being. Using Geographic Information Systems (GIS) and spatial econometric models, such as the SAR model and Geographically Weighted Regression (GWR), provides insights into these spatial dependencies and interactions [9].

Studies have shown that classroom design elements, including lighting, acoustics, and seating arrangements, may impact academic performance and engagement [20]. Additionally, interactions within virtual classrooms can influence student engagement and learning outcomes [21].

Our study builds on this research by examining both physical and digital spatial effects in education. Using spatial econometric models, we analyze how proximity to peers and digital resources impacts academic performance and well-being. Findings indicate that peer interactions and spatial dynamics play crucial roles in educational outcomes, with high-performing peers positively influencing their classmates' performance [22].

Addressing the digital divide and geographic disparities in access to educational resources is also essential for minimizing inequalities and maximizing the benefits of spatially aware educational interventions. Our research contributes empirical evidence on these spatial dependencies and knowledge spillover effects, informing the design of effective and inclusive educational spaces.

3. Methodology

3.1. Data Collection: Integration of Big Data Techniques

Our data collection leveraged Big Data techniques to ensure comprehensive and detailed datasets. We gathered information from three high school classes, including demographics, academic performance records, digital literacy assess-ments, and well-being surveys. This multi-faceted approach provided a holistic view of students' educational experiences and outcomes.

We used Learning Management System (LMS) to track student interactions, participation in online activities, and assignment submissions. These systems provided detailed logs essential for analyzing digital literacy and its impact on educational outcomes. Standardized academic performance assessments conducted during the study period allowed us to track changes in student performance over time.

Digital literacy assessments evaluated students' proficiency in using digital tools, covering skills like basic computer operations and online safety. Well-being surveys measured students' emotional and psychological states, social interactions, and satisfaction with their learning environments at the beginning and end of the study period.

To ensure data integrity and reliability, we implemented rigorous validation and cleaning procedures, cross-referencing multiple sources and anonymizing data to protect student privacy. By integrating traditional data collection methods with digital tools, we captured a comprehensive picture of students' educational experiences, providing a solid foundation for our subsequent analysis using ML and spatial econometric models (refer to methodology section for details).

3.2. Analytical Methods: ML and Spatial Econometrics

Our study combines ML and spatial econometric techniques to explore the relationships between digital literacy, student well-being, and educational outcomes. This approach allows for a comprehensive analysis, capturing both predictive power and spatial dependencies.

We employed several ML algorithms, including random forests (RFs), neural networks, and K-means clustering. RFs predicted academic performance and well-being, identifying significant predictors [12]. Neural networks captured nonlinear relationships within the data [23]. K-means clustering segmented students into groups based on digital literacy, well-being, and performance, facilitating targeted interventions [24].

In addition, spatial econometric models examined spatial dependencies. The SAR model analyzed the impact of physical proximity to high-performing peers on academic performance, accounting for spatial autocorrelation [22]. GWR explored the spatial variability in relationships between digital literacy and educational outcomes, providing local insights [9].

By integrating these methods, we achieved a robust framework for understanding the multifaceted impacts of digital literacy on student well-being and performance. This comprehensive approach enhances the accuracy and depth of our findings, informing targeted educational interventions and policy decisions.

3.3. Variables and Model Specifications

The robustness and accuracy of our study depend on the careful selection of variables and precise model specifications. Our dataset comprises information from three high school classes, encompassing a total of 64 students. The dataset includes comprehensive details such as student demographics (age and gender), academic performance records, digital literacy assessments, and well-being surveys. This provides a useful basis for describing the dynamics influencing digital literacy, student well-being, and educational outcomes.

We utilized ML models such as RFs, neural networks, and K-means clustering. RFs, specified with 500 trees and optimized using cross-validation, were used to predict academic performance and well-being [12]. Neural networks, featuring a feedforward model with three hidden layers, captured nonlinear relationships within the data [23]. K-means clustering, with the optimal number of clusters determined using the Elbow method and set at three clusters, segmented students based on digital literacy, well-being, and performance [24, 25].

In addition to ML, we applied spatial econometric models to examine spatial dependencies. The SAR model analyzed the impact of physical proximity to high-performing peers on academic performance, accounting for spatial autocorrelation [22]. The GWR explored the spatial variability in relationships between digital literacy and educational outcomes, providing local insights [9].

Data preprocessing involved normalization, encoding categorical variables, and imputing missing values using multiple imputation techniques. Model performance was evaluated using R^2, Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE) for ML models, and goodness-of-fit measures such as Log-Likelihood and Akaike Information Criterion (AIC) for spatial econometric models.

4. Results

4.1. Analysis of Data Using ML Techniques

Students showed an increase in average grades from 5.34 to 6.15, and well-being scores improved from 0.48 to 0.95. These paired changes are descriptive and consistent with associations among baseline well-being and subsequent grades; no causal effect is inferred.

We applied various ML models to analyze the data. RFs and neural networks predicted academic performance and well-being, with RFs proving more effective based on evaluation metrics such as RMSE and R^2 [12]. K-means clustering identified three distinct student groups based on digital literacy levels, well-being scores, and academic performance, facilitating targeted interventions [26].

The RF model indicates moderate explanatory power, with an R^2 value of 0.55, meaning that 55% of the variance in final grades could be explained by the predictors. The neural network model, although less effective, provided valuable insights into the complex relationships between variables.

Overall, the ML analysis underscored the importance of digital literacy in enhancing educational outcomes and student well-being. These findings inform targeted strategies and interventions aimed at fostering digital literacy and improving student performance (refer to Section 3 for detailed model specifications).

The unsupervised algorithm identified three clusters among the students, each with distinct characteristics:

- **Cluster 0:** 9 students, younger, lower parental education, high extroversion and creativity, high school happiness, general well-being, and strong school performance.

- **Cluster 1:** 28 students, older, lower parental education, moderate personality traits, high general well-being but lower school happiness and performance.
- **Cluster 2:** 27 students, slightly younger, higher parental education, lower extroversion but consistent responsibility, high general well-being and school performance.

To predict end-of-term grades (2G), we used a RF model, evaluating its performance with Mean Squared Error (MSE) and the coefficient of determination (R^2). The RF model yielded an MSE of 0.8846 and an R^2 of 0.5499, indicating that it explained approximately 55% of the variance in 2G scores, leaving 45% of the variance unexplained.

We also developed a neural network model to predict 2G scores using the same dataset. This model's performance, with an MSE of 1.7162 and an R^2 of 0.1268, was less effective than the RF model. Therefore, the RF model was deemed more effective for predicting 2G scores with the available data.

4.2. Spatial Statistical Analysis and Interpretation

In our spatial statistical analysis, we investigated how classroom interactions and proximities affect students' academic performance, incorporating variables such as personality traits, well-being, and demographics. Using the SAR model, we examined the impact of seating arrangements and peer influences on academic outcomes.

The SAR model is expressed as:

$$y = \rho W y + X\beta + \epsilon \tag{1}$$

where:

- y represents end test grades,
- ρ is the SAR coefficient,
- W is the spatial weights matrix,
- X is the matrix of explanatory variables,
- ϵ is the error term.

The model showed strong explanatory power, with an R^2 value of 0.725, indicating that a substantial portion of the variance in end-of-test grades was explained by the predictors. Significant predictors included age, initial school happiness, initial general well-being, and median class grades at the beginning and end of the term.

These results highlight the importance of well-being and classroom dynamics in academic success.

To further explore spatial relationships, we employed GWR, which revealed variations in the impact of personality traits and academic performance across different spatial locations within the classroom. This analysis emphasized the nuanced effects of peer influences and seating arrangements on student outcomes.

Overall, our spatial statistical methods provided valuable insights into the role of spatial dependencies in education, suggesting that targeted interventions considering these factors can enhance student well-being and academic performance (refer to methodology section for detailed model specifications).

4.3. Effects of Digital Literacy on Educational Outcomes

Our analysis shows associations that improvements in digital literacy may enhance educational outcomes. Students with higher digital literacy showed better academic performance and greater overall well-being. This aligns with recent meta-analytic evidence on achievement [7] and with studies linking digital competence to psychological well-being [8].

The data indicated that students' academic performance and well-being improved as their digital literacy skills increased. This was particularly evident in the higher grades and well-being scores observed among students with better digital literacy. Our findings support the idea that digital literacy fosters a more engaging and effective learning environment, contributing to students' academic success and well-being.

These results emphasize the importance of integrating digital literacy into educational curricula to enhance student outcomes. By providing students with the necessary digital skills, educators can create a more inclusive and effective learning environment that supports both academic achievement and overall well-being (refer to methodology section for detailed analysis).

5. Discussion

5.1. Implications for Policy and Educational Practice

Our findings have significant implications for policy and educational practice. The positive correlation between digital literacy and both academic performance and student well-being suggests that educational policies should prioritize digital literacy initiatives. Integrating digital literacy into curricula can equip students with essential skills, enhancing their academic outcomes [7] and overall well-being [8].

Educational institutions should consider investing in digital tools and resources, as well as training programs for teachers to effectively incorporate digital literacy into their teaching methods. Policies that support the development and implementation of digital literacy programs can help bridge the digital divide and ensure equitable access to quality education [10].

Furthermore, our analysis of spatial effects in educational settings indicates that classroom arrangements and peer interactions play a crucial role in student outcomes. Policies aimed at optimizing classroom environments to foster positive peer interactions and support students' well-being can lead to improved academic performance [20].

Prioritizing digital literacy and considering spatial dynamics in educational settings can may enhance educational outcomes. These insights should guide policymakers and educators in designing strategies that promote effective learning environments and support students' holistic development.

5.2. How ML can Inform Educational Assessments and Interventions

ML may enhance educational assessments and interventions by providing detailed insights into student performance and identifying key predictors of success. Our study shows associations how ML models, such as RFs and neural networks, can predict academic outcomes and well-being, allowing for more personalized and effective educational strategies [12].

By analyzing large datasets, ML can identify patterns and trends that traditional methods might overlook. This enables educators to develop targeted interventions tailored to individual student needs, improving both academic performance and overall well-being. For instance, clustering techniques like K-means help segment students into groups based on their digital literacy and well being, facilitating customized support and resources [26, 24].

ML also improves the efficiency of educational assessments by automating data analysis and providing real-time feedback. This allows educators to quickly identify at-risk students and implement timely interventions, enhancing the overall effectiveness of educational practices.

ML can inform educational assessments and interventions by offering powerful tools for analyzing student data, predicting outcomes, and tailoring educational strategies to meet individual needs. These advancements contribute to a more personalized and effective learning environment.

5.3. Ethical Considerations in the Use of AI and Big Data in Education

The use of AI and Big Data in education raises important ethical considerations. Data privacy and security are paramount, as sensitive student information must be protected from unauthorized access and misuse. Robust data governance frameworks are essential to ensure that data is collected, stored, and used responsibly [18].

Transparency and fairness in AI algorithms are also critical [19]. It is important to ensure that AI systems do not perpetuate biases or inequalities. This requires ongoing monitoring and evaluation of AI tools to maintain fairness and equity in educational outcomes [27].

Additionally, there are concerns about the potential for AI to replace human interaction in education. While AI can provide valuable support, it should complement, not replace, the role of teachers. Human oversight is necessary to ensure that AI-driven decisions align with educational goals and values [18]. Legal and philosophical analyses further warn that AI and Business Intelligence (BI) deployments must respect institutional norms and learners' rights [28].

Ethical considerations also extend to the digital divide. Ensuring equitable access to AI and Big Data technologies is crucial to avoid exacerbating existing inequalities in education. Policies should aim to provide all students with the necessary digital resources and support [15].

The ethical use of AI and Big Data in education requires careful consideration of data privacy, fairness, transparency, and equity. These principles should guide the development and implementation of AI technologies in educational settings to ensure they benefit all students (refer to methodology section for detailed analysis).

6. Conclusions

6.1. Main Findings and Their Significance

Our research revealed several key findings with significant implications for education. Firstly, improvements in digital literacy were associated to enhanced academic performance and overall student well-being. Students with higher digital literacy skills showed better grades and increased well-being, underscoring the importance of integrating digital literacy into educational curricula [7, 8].

Secondly, spatial analysis indicated that classroom arrangements and peer interactions may influence educational outcomes. Proximity to high-performing peers positively impacted students' academic performance, highlighting the importance of considering spatial dynamics in educational settings [20].

ML models, particularly RFs, were effective in predicting academic performance and well-being, providing valuable insights for personalized educational strategies. The predictive power of these models suggests their potential in developing targeted interventions to support student success [12].

Finally, the ethical use of AI and Big Data in education requires careful consideration of data privacy, fairness, and equity. Ensuring that AI systems are transparent and do not perpetuate biases is crucial for maintaining educational integrity [17, 15].

These findings emphasize the importance of digital literacy, spatial dynamics, and ethical considerations in enhancing educational outcomes. They provide a foundation for developing effective policies and practices that support student achievement and well-being.

6.2. Contributions to the Field of Educational Big Data and AI

This study offers empirical evidence on associations between digital literacy and outcomes. By integrating advanced data analytics and ML techniques, we provided empirical evidence that highlights the importance of digital skills in modern education [7, 8].

The use of spatial econometric models in our analysis offers new insights into the role of classroom arrangements and peer interactions in influencing educational outcomes. This approach underscores the importance of considering spatial dynamics when designing learning environments [20].

Furthermore, our application of ML models, such as RFs and neural networks, illustrates their effectiveness in predicting academic performance and well-being. These models enable the development of personalized educational interventions, demonstrating the potential of AI to enhance teaching and learning practices [12].

Lastly, our research addresses critical ethical considerations in the use of AI and Big Data in education, emphasizing the need for data privacy, fairness, and equitable access to technology. These findings contribute to the ongoing discourse on the responsible use of AI in educational settings [17, 27, 15].

This study advances the understanding of how digital literacy, spatial dynamics, and AI technologies can be leveraged to improve educational outcomes, providing a valuable framework for future research and policy development.

6.3. Future Research Directions and Limitations

Future research should continue to explore the impact of digital literacy on educational outcomes, focusing on diverse educational settings and larger sample sizes to validate our findings. Additionally, investigating the long-term effects of digital literacy on student performance and well-being will provide deeper insights into its sustained benefits [7, 8].

Further studies should also examine the role of spatial dynamics in education, particularly how different classroom arrangements and peer interactions influence learning outcomes. This can help refine strategies to optimize learning environments for maximum student benefit [20].

While our study indicates that ML models may be effective in predicting academic performance and well-being, future research should explore other AI techniques and their applications in education. This could enhance the precision and applicability of predictive models in various educational contexts [12, 13].

Our research has limitations, including the specific demographic and geographic scope of the study, which may affect the generalizability of the results. Additionally, while we addressed key ethical considerations, ongoing evaluation of AI and Big Data use in education is essential to ensure fairness and equity [17, 27, 15].

Expanding research on digital literacy, spatial dynamics, and AI applications in education will further our understanding and help develop effective, equitable educational strategies. Future work will also explore federated-learning approaches to reconcile performance with privacy constraints in multi-school settings [29].

Acknowledgments This work was partially supported by the project 21-FIN/RIC (Fin. Comp. 2024 UM). The authors thank *Università Mercatorum*, *NGS Sensors*, and *Tirana Business University* for their institutional support. The anonymous reviewers provided valuable feedback that improved the manuscript's clarity and analytical quality. All names and organizations involved were anonymized in compliance with General Data Protection Regulation (GDPR) to ensure confidentiality and ethical integrity.

Disclosure of Interests The authors declare no financial conflicts of interest related to the concepts, technologies, or entities discussed. Affiliations with organizations in enterprise automation and AI did not influence the study design, analysis, or interpretation. All data were collected and analysed independently, following academic ethical standards. Institutional disclosures were made, and confidentiality was ensured by anonymizing collaborators, datasets, and proprietary tools in compliance with GDPR.

Declaration on Generative AI

During the preparation of this work, the authors used GPT-4o, o3 and *Mate Translate* in order to perform grammar and spelling check. After using these tools, the authors reviewed and edited the content as needed and take(s) full responsibility for the publication's content.

References

[1] S. Shetye, An Evaluation of Khanmigo, a Generative AI Tool, as a Computer-Assisted Language Learning App, Studies in Applied Linguistics and TESOL 24 (2024). URL: https://journals.library.columbia.edu/index.php/SALT/article/view/12869. doi:10.52214/salt.v24i1.12869.

[2] R. Van Den Berghe, Social robots in a translanguaging pedagogy: A review to identify opportunities for robot-assisted (language) learning, Frontiers in Robotics and AI 9 (2022) 958624. URL: https://www.frontiersin.org/articles/10.3389/frobt.2022.958624/full. doi:10.3389/frobt.2022.958624.

[3] Grand View Research, Artificial Intelligence in Education Market Size, Share & Trends Analysis Report, Technical Report, Grand View Research, 2021. URL: https://www.grandviewresearch.com/industry-analysis/artificial-intelligence-ai-education-market-report.
[4] Businessolution.org, AI in Education Statistics, 2023. URL: https://businessolution.org/ai-in-education-statistics/.
[5] T. K. Chiu, Q. Xia, X. Zhou, C. S. Chai, M. Cheng, Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education, Computers and Education: Artificial Intelligence 4 (2023) 100118. URL: https://linkinghub.elsevier.com/retrieve/pii/S2666920X2200073X. doi:10.1016/j.caeai.2022.100118.
[6] F. Benelli, F. Maciariello, C. Salvadori, The influence of technologies on organizational culture in innovative SMEs, Journal of Robotics and Automation Research 5 (2024) 01–11. URL: https://doi.org/10.33140/JRAR.
[7] F. Li, L. Cheng, X. Wang, L. Shen, Y. Ma, A. Y. M. A. Islam, The causal relationship between digital literacy and students' academic achievement: a meta-analysis, Humanities and Social Sciences Communications 12 (2025) 108. URL: https://www.nature.com/articles/s41599-025-04399-6. doi:10.1057/s41599-025-04399-6.
[8] X. Wang, R. Zhang, Z. Wang, T. Li, How Does Digital Competence Preserve University Students' Psychological Well-Being During the Pandemic? An Investigation From Self-Determined Theory, Frontiers in Psychology 12 (2021) 652594. URL: https://www.frontiersin.org/articles/10.3389/fpsyg.2021.652594/full. doi:10.3389/fpsyg.2021.652594.
[9] A. S. Fotheringham, C. Brunsdon, M. Charlton, Geographically weighted regression: the analysis of spatially varying relationships, nachdr. der ausg. 2002 ed., Wiley, Chichester, 2010.
[10] B. Daniel, Big Data and analytics in higher education: Opportunities and challenges, British Journal of Educational Technology 46 (2015) 904–920. URL: https://bera-journals.onlinelibrary.wiley.com/doi/10.1111/bjet.12230. doi:10.1111/bjet.12230.
[11] O. Zawacki-Richter, V. I. Marín, M. Bond, F. Gouverneur, Systematic review of research on artificial intelligence applications in higher education – where are the educators?, International Journal of Educational Technology in Higher Education 16 (2019) 39. URL: https://educationaltechnologyjournal.springeropen.com/articles/10.1186/s41239-019-0171-0. doi:10.1186/s41239-019-0171-0.
[12] G. James, D. Witten, T. Hastie, R. Tibshirani, An introduction to statistical learning: with applications in R, Springer texts in statistics, second edition ed., Springer, New York, 2021.
[13] D. Ifenthaler, J. Y.-K. Yau, Utilising learning analytics to support study success in higher education: a systematic review, Educational Technology Research and Development 68 (2020) 1961–1990. URL: https://link.springer.com/10.1007/s11423-020-09788-z. doi:10.1007/s11423-020-09788-z.
[14] McKinsey & Company, How Artificial Intelligence Will Impact K–12 Teachers, 2020. URL: https://www.mckinsey.com/industries/education/our-insights/how-artificial-intelligence-will-impact-k-12-teachers.
[15] D. Mhlanga, T. Moloi, COVID-19 and the Digital Transformation of Education: What Are We Learning on 4IR in South Africa?, Education Sciences 10 (2020) 180. URL: https://www.mdpi.com/2227-7102/10/7/180. doi:10.3390/educsci10070180.
[16] N. AlQashouti, M. Yaqot, R. E. Franzoi, B. C. Menezes, Educational System Resilience during the COVID-19 Pandemic—Review and Perspective, Education Sciences 13 (2023) 902. URL: https://www.mdpi.com/2227-7102/13/9/902. doi:10.3390/educsci13090902.
[17] S. Slade, P. Prinsloo, Learning Analytics: Ethical Issues and Dilemmas, American Behavioral Scientist 57 (2013) 1510–1529. URL: https://journals.sagepub.com/doi/10.1177/0002764213479366. doi:10.1177/0002764213479366.
[18] W. Holmes, K. Porayska-Pomsta, K. Holstein, E. Sutherland, T. Baker, S. B. Shum, O. C. Santos, M. T. Rodrigo, M. Cukurova, I. I. Bittencourt, K. R. Koedinger, Ethics of AI in Education: Towards a Community-Wide Framework, International Journal of Artificial Intelligence in Education 32 (2022) 504–526. URL: https://doi.org/10.1007/s40593-021-00239-1. doi:10.1007/s40593-021-00239-1.
[19] H. Khosravi, S. B. Shum, G. Chen, C. Conati, Y.-S. Tsai, J. Kay, S. Knight, R. Martinez-Maldonado,

S. Sadiq, D. Gašević, Explainable Artificial Intelligence in education, Computers and Education: Artificial Intelligence 3 (2022) 100074. URL: https://linkinghub.elsevier.com/retrieve/pii/S2666920X22000297. doi:10.1016/j.caeai.2022.100074.

[20] P. Barrett, F. Davies, Y. Zhang, L. Barrett, The impact of classroom design on pupils' learning: Final results of a holistic, multi-level analysis, Building and Environment 89 (2015) 118–133. URL: https://linkinghub.elsevier.com/retrieve/pii/S0360132315000700. doi:10.1016/j.buildenv.2015.02.013.

[21] A. Christopoulos, M. Conrad, M. Shukla, Increasing student engagement through virtual interactions: How?, Virtual Reality 22 (2018) 353–369. URL: https://doi.org/10.1007/s10055-017-0330-3. doi:10.1007/s10055-017-0330-3.

[22] L. Anselin, Under the hood Issues in the specification and interpretation of spatial regression models, Agricultural Economics 27 (2002) 247–267. URL: https://onlinelibrary.wiley.com/doi/10.1111/j.1574-0862.2002.tb00120.x. doi:10.1111/j.1574-0862.2002.tb00120.x.

[23] F. Okubo, T. Yamashita, A. Shimada, H. Ogata, A neural network approach for students' performance prediction, in: Proceedings of the Seventh International Learning Analytics & Knowledge Conference, ACM, Vancouver British Columbia Canada, 2017, pp. 598–599. URL: https://dl.acm.org/doi/10.1145/3027385.3029479. doi:10.1145/3027385.3029479.

[24] Y. L. Goh, Y. H. Goh, C.-C. Yip, C. H. Ting, Raymond Ling Leh Bin, K. P. Chen, Prediction of Students' Academic Performance by K-Means Clustering, Science Proceedings Series 2 (2020) 1–6. URL: http://readersinsight.net/SPS/article/view/1205. doi:10.31580/sps.v2i1.1205.

[25] M. Saqr, Using learning analytics to understand and support collaborative learning, Department of Computer and Systems Sciences, Stockholm University, Stockholm, 2018.

[26] R. F. Kizilcec, C. Piech, E. Schneider, Deconstructing disengagement: analyzing learner subpopulations in massive open online courses, in: Proceedings of the Third International Conference on Learning Analytics and Knowledge, ACM, Leuven Belgium, 2013, pp. 170–179. URL: https://dl.acm.org/doi/10.1145/2460296.2460330. doi:10.1145/2460296.2460330.

[27] H. Drachsler, W. Greller, Privacy and analytics: it's a DELICATE issue a checklist for trusted learning analytics, in: Proceedings of the Sixth International Conference on Learning Analytics & Knowledge - LAK '16, ACM Press, Edinburgh, United Kingdom, 2016, pp. 89–98. URL: http://dl.acm.org/citation.cfm?doid=2883851.2883893. doi:10.1145/2883851.2883893.

[28] V. Stile, V. Bonino, N. Cosmo, The impact BI and AI on traditional structures with legal and philosophical insights, in: Proceedings of the 21st Conference of the Italian Chapter of AIS (itAIS 2024), accepted for publication on AISeL, volume 21, Università Cattolica del Sacro Cuore (UCSC), Piacenza, Italy, 2024. URL: https://aisel.aisnet.org/itais2024/21.

[29] F. Liberti, D. Berardi, B. Martini, Federated Learning in Dynamic and Heterogeneous Environments: Advantages, Performances, and Privacy Problems, Applied Sciences 14 (2024) 8490. URL: https://www.mdpi.com/2076-3417/14/18/8490. doi:10.3390/app14188490.

Appendices

A. Python Scripts used for this paper

A.1. SAR Model Analysis of Student Well-being and Academic Performance

This Python script loads a dataset of student well-being and academic performance, processes spatial information related to student seating arrangements, constructs a spatial weights matrix, and applies a SAR model using the `statsmodels` library to analyze the influence of various predictors on student outcomes. The script concludes with visualizing the model's coefficients to understand the impact of each variable:

The detailed Python code for the spatial and ML models, including SAR and GWR, is available in a GitHub repository reachable here: `https://github.com/vstile/2025edu4ai`.

A.2. GWR Analysis of Student Well-being and Academic Performance

This Python script conducts a GWR analysis to investigate the spatially varying relationships between various predictors and student academic performance:

The complete code implementation is available as public GitHub repository reachable here: `https://github.com/vstile/2025edu4ai`.