REPORT 18

Profiles of Students from the Military School System

Approved in Military and Civil Institutions:
A Speculative Approach Using Machine Learning

[ANONYMIZED REPORT — ALL IDENTIFYING INFORMATION REMOVED]

Abstract

This work investigates the application of machine learning techniques to identify profiles of students from [omitted] military schools who were approved in military institutions and civil universities, using the speculative design methodology. The objective is to group similar profiles and suggest improvements in study plans, aiming to increase success in entrance examinations. With the growth of e-learning, large educational data repositories require automated machine learning techniques. Challenges include transforming data into insights, ensuring user understanding of decisions, and generalizing models to different contexts. The speculative design methodology helps identify future scenarios and develop AI solutions that serve society as a whole. The article covers theoretical foundations, related studies, speculation of futures, and proposals for AI solutions.

1. Introduction

The use of Machine Learning techniques to identify student profiles originating from the [omitted] Military School System (SCMB), approved in military institutions and civil universities using the speculative design methodology, is the central theme of this work.

The study aims to group similar student profiles and suggest improvements in study plans, providing targeted pathways for military entrance examinations and civil higher education institutions, in order to increase student performance from Military Schools (CM), enhancing success in examinations for Military Schools, Institutes, and Academies, as well as for Civil Universities.

The field of machine learning in educational contexts is gaining prominence. The models created have a notable characteristic in supporting informed decision-making; they are applied to information systems as decision-making components, add value to an institution's business processes, and can bring benefits to society by increasing students' opportunities to access higher education institutions, whether military or civil.

The significant growth in the use of e-learning, web-based systems, and the internet in education, combined with distributed databases containing student information, has resulted in the creation of large educational data repositories. These web systems are highly relevant in storing enormous quantities of data from multiple sources and formats. Analyzing all this data manually becomes impractical [Romero and Ventura 2020].

Machine learning techniques are especially important for analyzing large volumes of data, making the use of automated techniques mandatory.

However, the use of machine learning alone presents challenges for the field. According to [Romero and Ventura 2020], educational data analysis must result in the transformation of data into valuable insights for students, teachers, and administrators. Another challenge is ensuring that human users understand the decisions made by analytical learning models. Additionally, there is a need to generalize models so they can be applied across different population contexts. Finally, there are specific challenges in Learning Analytics, such as transparency of objectives and a clear definition of how collected information will be used.

This work proposes the use of the speculative design methodology for the construction of a solution to the proposed theme that serves all segments of society. Speculative design offers the necessary tools for futurology by speculating on the causes and effects of identified scenarios.

The remainder of this article is structured as follows: a theoretical foundation section is presented, bringing important concepts for understanding the content, related studies in the field, and the identification of trends that position the topic

in the current scenario, presenting the main challenges. Next, possible futures are speculated, indicating positive and negative scenarios with their consequences. In the following section, an AI solution is proposed with necessary actions to address positive scenarios and mitigate negative ones. Finally, a conclusion and references are presented.

2. Theoretical Foundation

This section presents the theoretical concepts that support the development of this work. It addresses fundamental aspects related to machine learning, educational data mining, learning analytics, and speculative design, which together provide the basis for the proposed approach.

2.1. Machine Learning

Machine Learning is a subfield of Artificial Intelligence that focuses on the development of algorithms capable of learning patterns from data and making predictions or decisions without being explicitly programmed. These algorithms identify relationships in data sets and use them to improve performance on specific tasks over time.

In educational contexts, machine learning has been widely applied to analyze student performance, predict academic outcomes, and identify learning patterns. Classification, clustering, and regression techniques are commonly used to group students with similar characteristics, forecast success rates, and recommend personalized learning paths.

Supervised learning techniques rely on labeled data to train models, while unsupervised learning techniques, such as clustering, aim to identify structures or patterns in unlabeled data. In this work, clustering techniques are particularly relevant, as they enable the identification of student profiles based on shared attributes.

Despite their potential, machine learning models present challenges related to interpretability and transparency. Complex models may produce accurate predictions but offer limited insight into how results are generated, making it difficult for users to trust and understand their outputs.

2.2. Educational Data Mining and Learning Analytics

Educational Data Mining (EDM) and Learning Analytics (LA) are research areas focused on analyzing educational data to improve learning processes and educational outcomes. EDM emphasizes the development of methods and algorithms to extract patterns from educational data, while LA focuses on measuring, collecting, analyzing, and reporting data about learners and learning contexts.

According to Romero and Ventura (2020), these areas aim to transform raw educational data into actionable insights that can support decision-making by students, teachers, and educational administrators. The growth of digital learning environments has significantly increased the availability of educational data, making EDM and LA increasingly important.

However, both areas face challenges related to data quality, privacy, and ethical use. Educational data often include sensitive personal information, requiring careful governance and compliance with data protection regulations. Additionally, ensuring that analytical results are understandable and useful to stakeholders remains a critical concern.

2.3. Transparency and Explainability

Transparency and explainability are essential requirements for the adoption of machine learning models in educational contexts. Transparency refers to the availability of information about how models are built, trained, and used, while explainability concerns the ability to provide understandable explanations for model outputs.

In learning analytics applications, transparency and explainability are crucial to ensure trust and acceptance among users. Students and educators need to understand why certain recommendations or predictions are made, especially when these outputs influence important decisions such as study plans or admission strategies.

The lack of transparency may lead to resistance to the use of machine learning systems or to the misuse of analytical results. Therefore, incorporating explainability mechanisms into machine learning solutions is a key challenge addressed by this work.

2.4. Speculative Design

Speculative design is a methodological approach that explores possible futures through the creation of speculative artifacts, scenarios, and narratives. Rather than predicting the future, speculative design aims to question current assumptions and examine the social, ethical, and technological implications of emerging trends.

By envisioning alternative futures, speculative design allows researchers and designers to anticipate challenges, identify opportunities, and consider the consequences of technological choices. In the context of education and machine learning, speculative design supports reflection on how data-driven technologies may reshape learning experiences and institutional practices.

In this work, speculative design is used to guide the exploration of future scenarios involving machine learning-based student profiling and to propose solutions that balance technological innovation with social responsibility.

3. Related Studies and Trends

This section presents related studies and identifies trends relevant to the application of machine learning and learning analytics in educational contexts, particularly in scenarios involving student profiling and performance prediction.

Several studies have explored the use of machine learning techniques to analyze educational data and support decision-making processes. According to Romero and Ventura (2020), educational data mining has been applied to tasks such as predicting student dropout, identifying learning difficulties, and recommending personalized learning resources. These studies

demonstrate the potential of data-driven approaches to improve educational outcomes.

In the context of military education, research has focused on evaluating academic performance, physical preparedness, and psychological factors that influence student success. However, there is still limited work integrating machine learning techniques to comprehensively analyze student profiles and support strategic planning for entrance examinations in both military and civil institutions.

Learning Analytics has also gained prominence as a means of transforming educational data into actionable insights. Dashboards and analytical tools have been developed to visualize student performance and support interventions. Despite these advances, challenges remain regarding the interpretability of results and the alignment of analytics with pedagogical objectives.

A significant trend identified in the literature is the growing emphasis on transparency and explainability in machine learning models. As analytical systems increasingly influence educational decisions, there is greater demand for models that provide understandable explanations and support human oversight.

Another trend is the integration of multiple data sources to build richer student profiles. Combining academic records, behavioral data, and contextual information enables more comprehensive analyses but also raises concerns related to data privacy and governance.

The use of speculative design in educational research is an emerging trend that complements empirical studies by exploring possible futures and examining the broader implications of technological adoption. By integrating speculative design with machine learning research, it becomes possible to reflect not only on what is technically feasible but also on what is socially desirable.

These related studies and trends inform the development of the speculative scenarios and solution proposed in this work, highlighting both opportunities and challenges associated with the use of machine learning in education.

4. Speculation of Possible Futures

Based on the theoretical foundation and the analysis of related studies and trends, this section speculates on possible future scenarios involving the use of machine learning to identify student profiles from the [omitted] Military School System and support decision-making related to entrance examinations for military and civil institutions.

4.1. Positive Future Scenario

In a positive future scenario, machine learning models are effectively integrated into educational information systems and used to identify student profiles in a transparent and explainable manner. These models support students and educators by providing insights into learning patterns, strengths, and areas for improvement.

Educational institutions use these insights to design personalized study plans that increase students' chances of success in entrance examinations. Students receive clear recommendations aligned with their profiles, allowing them to focus on specific subjects and competencies. Teachers and advisors use analytical outputs as decision-support tools rather than prescriptive mechanisms.

In this scenario, transparency and explainability mechanisms are embedded in machine learning systems, ensuring that users understand how profiles are generated and how recommendations are derived. Ethical guidelines and data governance frameworks protect students' privacy and ensure responsible use of data.

4.2. Negative Future Scenario

In a negative future scenario, machine learning systems are adopted without adequate transparency, governance, or human oversight. Student profiles are generated using opaque models, and recommendations are presented as objective truths without explanation.

This lack of transparency leads to mistrust among students and educators, who may feel constrained or unfairly categorized by

algorithmic decisions. Biases present in data or models may reinforce inequalities, disadvantaging certain groups of students.

In addition, the misuse of student data and the absence of clear accountability mechanisms may result in ethical and legal issues. Students may be pressured to follow algorithmic recommendations without understanding their limitations, reducing autonomy and agency.

4.3. Discussion of Future Scenarios

The speculative scenarios illustrate the importance of responsible design and implementation of machine learning systems in educational contexts. While positive scenarios highlight opportunities for personalization and improved performance, negative scenarios reveal risks related to opacity, bias, and loss of trust.

These futures emphasize that technological outcomes are not predetermined but depend on choices made by designers, institutions, and policymakers. Speculative design enables reflection on these choices and supports the development of solutions that align technological innovation with educational and social values.

5. Proposed AI Solution

Based on the speculative analysis of possible futures, this section presents a proposed Artificial Intelligence solution aimed at supporting students from the [omitted] Military School System in preparing for entrance examinations for military and civil institutions. The solution is designed to address positive scenarios and mitigate risks identified in negative scenarios.

The proposed solution consists of a machine learning-based system integrated into educational information systems used by military schools. Its primary objective is to identify student profiles through clustering techniques and provide personalized recommendations for study plans.

The system collects and processes data related to students' academic performance, exam results, and learning behaviors. These data are used to train machine learning models capable of grouping students with similar characteristics. The resulting clusters represent different student profiles, each associated with specific strengths, weaknesses, and learning needs.

Based on the identified profiles, the system generates recommendations for study plans tailored to each group. These recommendations may include suggested subjects for focused study, allocation of study time, and identification of relevant learning resources. The goal is to increase students' chances of success in entrance examinations by aligning preparation strategies with individual profiles.

Transparency and explainability are central features of the proposed solution. The system provides explanations of how student profiles are generated and how recommendations are derived. Visualizations and descriptive summaries are used to make analytical results understandable to students, teachers, and advisors.

Human oversight is maintained throughout the process. Teachers and educational advisors can review, adjust, or override system recommendations based on their professional judgment and contextual knowledge. This ensures that the system supports decision-making rather than replacing human expertise.

Data governance and ethical considerations are explicitly addressed in the proposed solution. Clear policies define how data are collected, stored, and used, ensuring compliance with data protection regulations. Mechanisms for auditing and monitoring system performance are included to detect biases or errors and promote continuous improvement.

As a speculative proposal, this AI solution does not represent a finalized implementation but serves as a conceptual model for how machine learning can be responsibly applied to educational decision-making. It illustrates how technological innovation can be aligned with transparency, accountability, and social responsibility.

6. Conclusion

This work explored the application of machine learning techniques to identify student profiles from the [omitted] Military School System and support preparation for entrance examinations for military and civil institutions, using speculative design as a guiding methodology.

The analysis demonstrated that machine learning has significant potential to transform educational data into actionable insights, enabling personalized study plans and informed decision-making. However, the responsible use of these technologies requires careful attention to transparency, explainability, data governance, and ethical considerations.

Speculative design proved to be a valuable approach for examining possible futures and anticipating the consequences of technological adoption. By exploring both positive and negative scenarios, it became possible to identify key factors that influence the success and social acceptability of machine learning-based solutions in education.

The proposed AI solution illustrates how clustering techniques and decision-support mechanisms can be integrated into educational information systems in a way that balances technological innovation with human oversight. By emphasizing transparency and accountability, the solution seeks to enhance students' opportunities while preserving trust and autonomy.

As limitations, this work is based on speculative scenarios that may not fully capture real-world complexities. Future research could involve empirical validation of the proposed approach, interdisciplinary collaboration, and participatory design processes involving students, teachers, and policymakers.

In conclusion, the integration of machine learning into educational contexts holds promise for improving educational outcomes, but its success depends on intentional design choices, ethical governance, and continuous reflection on social impacts.

References ***(kept verbatim from the original document)***

(The reference list is reproduced exactly as in the source document, without translation or normalization.)

Alyahyan, E. and D¨usterg¨or, D. (2020). Predicting academic success in higher education:

literature review and best practices. International Journal of Educational Technology

in Higher Education.

Bittencourt, I. I. and Isotani, S. (2018). Evidence-based computers in education: A

manifesto (inform´atica na educac, ão baseada em evidˆencias: Um manifesto). Brazilian

Journal of Computers in Education (Revista Brasileira de Inform´atica na Educac, ão -

RBIE).

Dunne, A. and Raby, F. (2013). Speculative Everything: Design, Fiction, and Social

Dreaming. MIT Press (MA).

Futures-wheel (2014). Futures wheel. https://futurescouting.com.au/

futures-wheel/. [online: acesso em 02-julho-2024].

of possibilities, C. (2014). Cone of possibilities. https://futurescouting.com.

au/futures-wheel/. [online: acesso em 02-julho-2024].

Raja, P. S. and Thangavel, K. (2019). K. missing value imputation using unsupervised

machine learning techniques. Soft Comput.

Romero, C. and Ventura, S. (2020). Educational data mining and learning analytics: An

updated survey. Wires Data Mining and knowledge discovery.

Trends, G. (2014a). Analytics learning. https://trends.google.com.br/

trends/explore?date=today%205-y&q=analytics%20learning&

hl=pt. [online: acesso em 02-julho-2024].

Trends, G. (2014b). Learning personalization. https://trends.google.

com.br/trends/explore?date=today%205-y&q=learning%

20personalization&hl=pt. [online: acesso em 02-julho-2024].

Trends, G. (2014c). Machine learning in education. https://trends.

google.com.br/trends/explore?date=today%205-y&q=machine%

20learning%20in%20education&hl=pt. [online: acesso em 02-julho-2024].

Trends, G. (2014d). Predictive analytics. https://trends.google.com.br/

trends/explore?date=today%205-y&q=predictive%20analytics&

hl=pt. [online: acesso em 02-julho-2024].

Yang, C. C. Y., Chen, I. Y. L., and Ogata, H. (2021). Toward precision education:

Educational data mining and learning analytics for identifying students' learning

patterns with ebook systems. Educational Technology & Society.