# Leveraging Machine Learning to Predict Academic Success: A Comprehensive Literature Review and Trend Analysis

### Abstract

…

### Keywords

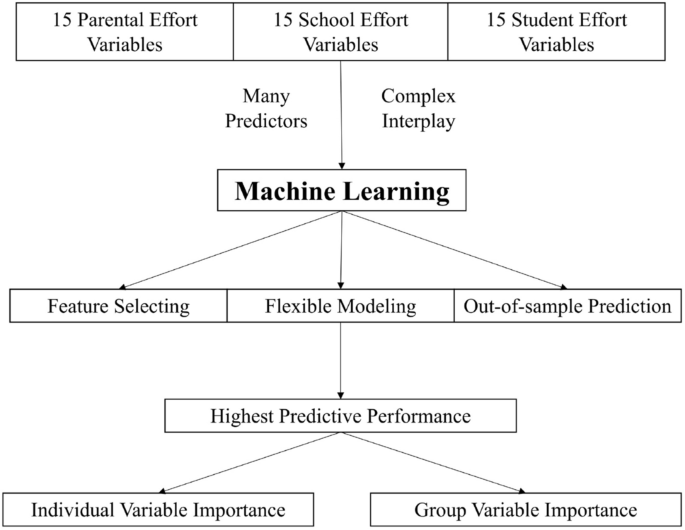
data mining; machine learning; academic success; educational analysis; systematic literature review; bibliographical analysis

## Introduction

In today’s complex world, predicting academic success has become a key focus in education research (Guanin-Fajardo et al., 2024). As the demand for skilled professionals grows, understanding what helps students succeed is more important than ever. This has led to a surge in studies aiming to predict academic outcomes, such as graduation rates and the likelihood of student dropouts.

One of the most promising tools in this area is machine learning (ML) (Balaji et al., 2021; Jin, 2023). ML allows researchers to analyze large amounts of data to find patterns and make predictions. This can help identify students who might struggle, giving educators the chance to offer support early on. By using data like academic history and student behavior, ML models can provide insights that help improve student outcomes.

Malo več o ML v edukaciji



Many studies have explored different factors that contribute to academic success, such as personal effort, family support, and the learning environment. These factors are crucial and have been widely researched, helping schools and educators develop strategies to support students. However, there is a noticeable gap in bibliometric analyses within this field. Understanding the research trends, publication patterns, and thematic developments through a bibliometric approach can provide valuable insights

The present review study is focused on research about machine learning statistics methods for the prediction of academic success of students and therefore opens the following research questions:

RQ1: What are the primary research themes and trends in the domain of academic success prediction using machine learning, based on the analysis of abstracts and keyword co-occurrence?

RQ2: How do publication patterns and cited reference counts vary across different publishers and over time in the research on academic success prediction?

RQ3: What is the distribution of publications over the years, and how has the focus on machine learning in predicting academic success evolved in terms of publication frequency and keyword prominence?

RQ4: Which are the most referenced papers, and how do their cited reference counts correlate with their impact and relevance in the field of academic success prediction?

RQ5: What are the characteristics and distribution of publications by language and publisher city, and how do these characteristics influence the citation patterns and research dissemination?

## Methods

Study conducted in the academic year 2025 within the subject of Raziskovalni seminar on PhD program Edukacijske vede.. …. To establish a rigorous, transparent, reproducible and adaptable review study, the search process was conducted based on the PRISMA protocol.

Preliminary analysis included SCOPUS, WOS, Taylor & Francis databases, and register Digitalni portal Univerze na Primorskem. A comparison between the databases and registries yielded high overlap for thematic of prediction of academic success using machine learning techniques. For the final bibliometric study WOS database was chosen, as it allows science citation index, which was one of the research questions. For the review articles from all selected databases and registries were accounted for, based on their importance.

For the final report in this article, the final search in the WOS core collection was conducted in January 2025. The search string applied was *("machine learning" OR "ML") AND ("predicting academic success" OR "academic performance prediction" OR "student success prediction") AND ("educational sciences" OR "education")*. As Prisma flowchart indicates in FIGURE, the search results in 286 matches. This was later reduced to article publications in the English language which resulted in X matches which were further screened based on a title and abstract. From X matches the final screening for the review was . While we acknowledge that there may be relevant articles not included in this review … Exclusion Criteria were:

Review articles;

Book chapters;

Factor analysis;

Articles not written in the English language.

The final selection of 286 articles was analysed using Python programming language to construct and visualize the bibliometric network

Records removed *before screening*:

Records removed for other reasons (n = 86)

Records identified from\*:

Databases (n = 286)

Registers (n = 86) (digitalni portal)

**Identification**

Records screened based on title, abstract and key words

(n = )

Records excluded\*\*

(n = )

Reports sought for retrieval

(n = )

Reports not retrieved

(n = )

**Screening**

Reports assessed for eligibility

(n = )

Reports excluded:

Reason 1 (n = )

Reason 2 (n = )

Reason 3 (n = )

etc.

Studies included in review

(n = 10)

Articles included in bibliographical analysis

(n = 286)

**Included**

## Results

…

A graph with blue rectangles

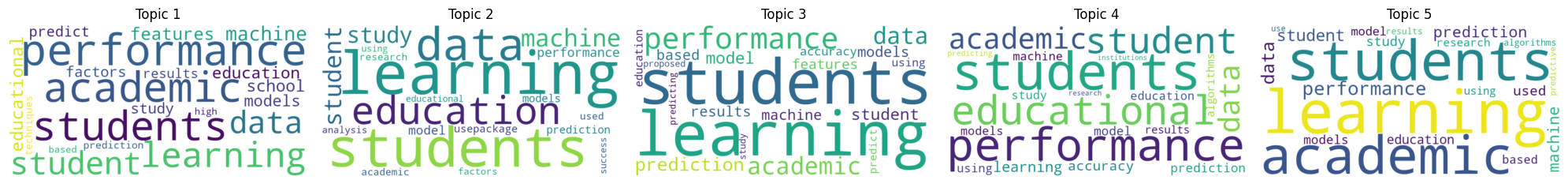
Description automatically generated

A graph of a number of cities

Description automatically generated

A graph of blue bars

Description automatically generated with medium confidence



A diagram of a box plot

Description automatically generated

A graph of blue bars

Description automatically generated with medium confidence

A graph of blue and orange bars

Description automatically generated

## Discussion

…

## Limitations and future directions

…

## References

Balaji, P., Alelyani, S., Qahmash, A., & Mohana, M. (2021). Contributions of Machine Learning Models towards Student Academic Performance Prediction: A Systematic Review. *Applied Sciences*, *11*(21), Article 21. https://doi.org/10.3390/app112110007

Guanin-Fajardo, J. H., Guaña-Moya, J., & Casillas, J. (2024). Predicting Academic Success of College Students Using Machine Learning Techniques. *Data*, *9*(4), 60. https://doi.org/10.3390/data9040060

Jin, X. (2023). Predicting academic success: Machine learning analysis of student, parental, and school efforts. *Asia Pacific Education Review*. https://doi.org/10.1007/s12564-023-09915-4