Improve predictive model performance using feature selection techniques

**Abstract:**

STUDENT PERFORMANCE

The most important corner stone in the educational process is students’ performance. Therefore, early prediction of students’ performance aims to assist at-risk students by providing appropriate and early support and intervention. [3]

DROPOUT

Student dropout is an educational phenomenon studied for decades due to the diversity of its causes, whose effects fall on society's development. [9] It affects not only the individual who drops out but also the former school, family, and society in general. [8]

EDM

Applying data mining for improving the outcomes of the educational process has become one of the most significant areas of research. [3] Educational Data Mining is a field of study that aims to find patterns and information in educational institutions through mining educational data. [4] Among the educational data mining problems, the early prediction of the students' academic performance and chance of dropout are the most important tasks, so that timely and requisite support may be provided to the needy students. [2]

FEATURE SELECTION

PAST RESEARCH

Research is improving and growing fast in Educational Data Mining (EDM) due to many students' information. Researchers urge to invent valuable patterns about students' learning behavior using their data that needs to be adequately processed to transform it into helpful information. [11] Many researchers recently tackled this problem by proposing different frameworks for predicting students’ academic performance. [1]

GOALS

This paper is a literature study of the past research done in the field of EDM and Feature Selection being used as tools for predicting student performance and possibility of dropout. It’s main goal is to study the effects of various Feature Selection techniques and to conclude how they could be employed, in the future, in assisting Romanian Universities and the challenges they face regarding students performance and dropout.

**Keywords:**

Educational Data Mining (Exploatarea datelor educationale?), Feature Selection (Selectarea caracteristicilor?), Machine Learning (Invatare automata?), Students’ performance (Performanta studentilor?), University dropout prevention (Prevenirea abandonului universitar), Predictive System (Sistem predictiv?)

**1.** **Introduction**

STUDENT PERFORMANCE

To become a better teacher, teachers need to anticipate their pupils' performance patterns. Knowledge gained from it can be used in various ways, such as a strategic plan for delivering high-quality education. [4] Educators’ loss of ability to read students ’ comprehension level during the class through quick questions or nonverbal communication is one of the main challenges of online and blended learning. [1]

DROPOUT

Student dropout is a serious problem globally. [8] The dropout of university students is one of the topics of broad interest in higher education institutions and government education departments. Recent changes in education methods, socio-economic conditions, and the growing limitations of face to face interactions make it necessary to have tools that allow us to consider a broad set of factors related to the dropout phenomenon. [6]

EDM

Educational Data Mining (EDM) is gaining great importance as a new interdisciplinary research field related to some other areas. It is directly related to data mining (DM), the latter being a fundamental part of knowledge discovery in databases (KDD). This data is growing more and more and contains hidden knowledge that could be very useful for users (both teachers and students). It is convenient to identify such knowledge in the form of models, patterns, or any other representation scheme that allows better exploitation of the system. Data mining is revealed as the tool to achieve such discovery, giving rise to EDM. In this complex context, different techniques and learning algorithms are usually used to obtain the best results. Recently educational systems are adopting artificial intelligent systems, especially in the educational context, specific areas for extracting relevant information, such as EDM, which integrates numerous techniques that support the capture, processing, and analysis of these sets of records. The main technique associated with EDM is Machine Learning, which has been used for decades in data processing in different contexts, but with the advent of Big Data, there was an intensification in the application of this technique to extract relevant information from a huge amount of data. [10] Machine learning techniques may be used as an important tool for predicting low-performers in educational institutions. [2]

With the current development of science and technology, big data is emphasized as the most significant technology in data analysis. From the recorded educational data, efficient prediction of students’ dropout is currently a hot topic of research. [8]

FEATURE SELECTION

Many recent works in the study have proposed prediction models of students' academic performances with new data features, including student's behavioral features, Psychometric, family support, learning logs via e-learning management systems, and demographic information. [11]

PAST RESEARCH

GOALS

RESEARCH QUESTIONS

Lucrarea reprezinta o revizuire sistematica a literaturii ce se concentrează pe prezicerea performantei academice si a ratei de abandon in domeniul universitar.

Vor fi explorate următoarele întrebări de cercetare:

- Q1: Care este stadiul actual al domeniul predictiei performanței academice si a ratei de abandon în domeniul universitar?

– Q1A: Cum sunt definite performanta academica si abandonul in domeniul universitar?

– Q1B: Care sunt metricile folosite pentru a evalua predictia performantei academice si rata de abandon in domeniul universitar?

– Q1C: Ce caracteristici sunt folosite pentru a prezice performanta academica si rata de abandon?

– Q1D: Ce atribute sunt folosite pentru a prezice performanta academica si rata de ababdon?

- Q2: Care este calitatea lucrărilor actuale din domeniu?

- Q3: Cum sunt aplicate modelele de predictie si tehnicile de Feature Selection in cazuri specifice?

- Q4: Ce rol poate avea EDM in reducerea abandonului universitar si in cresterea promovabilitatii in domeniul universitar?

- Q5: Ce rol poate avea Feature Selection in reducerea abandonului universitar si in cresterea promovabilitatii in domeniul universitar?

- Q6: Ce rezultate pot fi obtinute, in viitor, aplicand EDM si Feature Selection asupra unui set de date ce cuprinde date despre studentii inmatriculati la universitati din Romania?

HOW IT CAN BE USED (WHY IT SHOULD BE USED) IN ROMANIAN UNIVERSITIES

<https://www.romania-insider.com/half-romanian-students-started-2015-dropped-out-dec-2022>

Roughly 48% of undergraduate students in Romania who started college in 2015 dropped out. A total of 42% of those in master's studies also quit, while 70% of doctoral students also abandoned their studies in the same period.

The data, provided by a recent study by the Ministry of Education’s Executive Unit for the Financing of Higher Education, Research, Development and Innovation (UEFISCDI), shows the worrying state of Romanian higher education.

The factors that influence the dropout rate of universities are multiple, ranging from the characteristics of the students and the field studied, to the baccalaureate grade and the high school course. For example, the dropout rate reaches 62.8% for those with average grades between 6 and 7 at the baccalaureate. Students who come from high schools with a technological profile have the highest dropout rates, 53.6%, according to the data presented in the study.

Of undergraduate students who drop out of college in Romania, 44% do so in the first year. Meanwhile, the dropout rate of master's students can vary – in some universities, only 30% of them graduate, while in others 85% manage to obtain their degree. Even so, master students who began their studies in 2015 also dropped out at an average rate of 42%. Meanwhile, 70% of doctoral students also drop out before obtaining their degree.

Civil engineering has the most dropouts, 62.44%, followed by mechanical engineering, 61.76%. Large degrees of variation exists between universities. In the case of social sciences, the dropout rate is 33% in Bucharest – compared to the 44% national average. The university centers in Timișoara and Iași have higher graduation rates in the field of engineering sciences. Cluj and Iași have higher graduation rates than the national average in fields such as mathematics and natural sciences.

Male students are also more likely to drop out than female ones. Nearly 55% of the former abandon their studies, while only 42% of the latter do so. The same is true for master programs.

Students over the age of 26 have a 57.4% dropout rate, compared to 46.19% for younger students. Those between 23 and 26, however, have the highest dropout rate, namely 70.93%.

Students who did not move cities for their studies are slightly more likely to drop out than those who did. Foreign students also edge out the national average when it comes to the dropout rate.

Students who come from high schools with a technological profile have the highest dropout rates, 53.6%, compared to 44.5% of those from the so-called theoretical high schools. 90% of students enrolled in 5- and 6-year programs have graduated from such theoretical high schools, and dropout rates are considerably lower than those from technological or vocational high schools.

Around three-in-four doctoral students in industrial engineering, materials, electrical, mechanical, or electronic engineering drop out. At the opposite end, the sciences of education, music, psychology, theology, biology, or mathematics have the lowest dropout rates – below 35%. Doctoral students who occupy a budgeted place have a higher dropout rate compared to those who pay for their studies.

<https://edu.ro/sites/default/files/_fi%C8%99iere/Minister/2020/Transparenta/Stare%20invatamant/Stare%20superior%202019-2020.pdf> also has some tables

1. Rezultate ale învăţării în învăţământul superior 3.1. Rezultate ale studenților din învățământul universitar de licență Din totalul de 402,7 mii de studenți înscriși la începutul anului 2018/2019 în învățământul universitar de licență, în evidențele de sfârșit de an s-au regăsit 363,8 mii (90,4%). Din rândul studenților înscriși la începutul anului universitar, au promovat 346,5 mii (86%), o pondere de 4,3%fiind declarați repetenți și/sau cu situația neîncheiată. Cea mai ridicată rată de promovare a fost înregistrată în cadrul învățământului de zi (86,8%). 23 Un indicator semnificativ este reprezentat de abandonul universitar (calculat ca diferența procentuală dintre efectivele de la începutși cele de la sfârșit de an universitar), care reprezintă 9,6%în 2018/2019 și are valori similare în învățământul de stat și în cel particular. Valoarea indicatorului crește pe parcursul perioadei analizate (de la 8,5% la 9,6%). Valorile cele mai mari se înregistrează în cadrul învățământului la distanță.

<https://www.edu.ro/sites/default/files/_fi%C8%99iere/Minister/2021/Transparenta/Stare%20invatamant/Raport_stare_invatamant_superior_RO_2020_2021.pdf> has some tables also

3. Rezultatele învăţării în învăţământul superior 3.1. Rezultatele studenților din învățământul universitar de licență În anul universitar 2019/2020 a fost înregistrată o creştere cu aproape 5 mii de studenţi înscrişi în învăţământul de licenţă, comparativ cu anul universitar anterior. Din totalul de 407,4 mii de studenţi înscrişi la început de an, în evidenţele de sfârşit al anului 2019/2020 s-au regăsit 372,8 mii. Astfel, din totalul celor înscrişi la începutul anului universitar: 87,9% au promovat, 3,6% au fost declarați repetenţi şi/sau cu situaţia neîncheiată şi 8,5% au abandonat studiile pe parcurs. Comparativ cu anul universitar anterior se constată o ameliorare a situaţiei din învăţământul de licenţă, caracterizată prin creşterea ponderii studenţilor promovaţi şi prin scăderea pierderilor şcolare (atât prin reducerea abandonului, cât şi prin reducerea procentajului de repetenţi). Un indicator semnificativ este reprezentat de abandonul universitar (calculat ca diferenţa procentuală dintre efectivele de la început şi cele de la sfârşit de an universitar), care a scăzut cu 1,1% la nivelul învăţământului universitar de licenţă per total, respectiv cu 1,4% la nivelul învăţământului public. O altă observaţie importantă este legată de menţinerea într-o proporție asemănătoare, începând cu anul 2016/2017, a distribuției studenţilor la final de an pe diferite forme de studiu: aproximativ 90% în învăţământul cu frecvenţă, 3,8% la frecvenţă redusă şi 6,4% în învăţământul la distanţă.

DESCRIPTION OF FURTHER SECTIONS

**2.** **Methodology**

Am selectat 11 articole pe tema aleasa de pe IEEE Explore, Scopus si Science Direct

Mai sunt in plus inca 4 articole de pe Scopus si Web Of Science pe care inca nu le-am parcurs in detaliu si care raman ca backup

Mai este un articol de pe Science Direct care nu avea foarte multe detalii despre Feature Selection care ramane ca backup

Proces de selectare:

Feature selection, educational data mining, last 5 years:

- Science Direct,<https://www-sciencedirect-com.am.e-nformation.ro/>, Advanced search

Query: (“educational data mining” OR “edm”) AND (“feature selection” OR “variable selection” OR “attribute selection” OR “variable subset selection”)

Filters applied: 2018-2023

Rezultate = 187 -> filtrate manual la 30 (24 eliminand articolele din 2018 si 2019) (22 eliminand apoi articolele din 2020) (15 eliminand apoi articolele din 2021)

- IEEE Explore,<https://ieeexplore.ieee.org/>, Advanced Search

Query: ("All Metadata":"educational data mining" OR "All Metadata":"edm") AND ("All Metadata":"feature selection" OR "All Metadata":"variable selection" OR “All Metadata”:”attribute selection” OR “All Metadata”:”variable subset selection”)

Filters Applied: 2018-2023

Rezultate = 32 -> filtrate manual la 20 (11 eliminand articolele din 2018 si 2019) (8 eliminand apoi articolele din 2020) (2 eliminand apoi articolele din 2021)

- Scopus,<https://www.scopus.com/home.uri>, Advanced search

Query: TITLE-ABS-KEY (“feature selection” OR “variable selection” OR “attribute selection” OR “variable subset selection”) AND TITLE-ABS-KEY (“educational data mining” OR “edm”) AND (LIMIT-TO(PUBYEAR, 2023) OR LIMIT-TO(PUBYEAR, 2022) OR LIMIT-TO(PUBYEAR, 2021) OR LIMIT-TO(PUBYEAR, 2020) OR LIMIT-TO(PUBYEAR, 2019) OR LIMIT-TO(PUBYEAR, 2018))

Rezultate = 139 -> filtrate manual la 25 (22 eliminand articolele din 2018 si 2019) (17 eliminand apoi articolele din 2020) (12 eliminand apoi articolele din 2021)

- Web Of Science,<https://www-webofscience-com.am.e-nformation.ro/>, Advanced search

Query: #4 OR #3 OR #2 OR #1

Refine By: Publication years: 2018 OR 2019 OR 2020 OR 2021 OR 2022 OR 2023

Where:

#1 Query: TI=((educational data mining OR edm) AND (feature selection OR variable selection OR attribute selection OR variable subset selection))

#2 Query: AB=((educational data mining OR edm) AND (feature selection OR variable selection OR attribute selection OR variable subset selection))

#3 Query: KP=((educational data mining OR edm) AND (feature selection OR variable selection OR attribute selection OR variable subset selection))

#4 Query: AK=((educational data mining OR edm) AND (feature selection OR variable selection OR attribute selection OR variable subset selection))

Rezultate = 143 -> filtrate manual la 11 (8 eliminand articolele din 2018 si 2019) (6 eliminand apoi articolele din 2020) (2 eliminand apoi articolele din 2021)

INITIAL

- Science Direct -> 187 articole

- IEEE Explore -> 32 articole

- Scopus -> 139 articole

- Web Of Science -> 143 articole

DUPA FILTRARE MANUALA

- Science Direct -> 30 articole

- IEEE Explore -> 20 articole

- Scopus -> 25 articole

- Web Of Science -> 11 articole

ELIMINAND ARTICOLELE DIN 2018 si 2019

- Science Direct -> 24 articole

- IEEE Explore -> 11 articole

- Scopus -> 22 articole

- Web Of Science -> 8 articole

ELIMINAND ARTICOLELE DIN 2020

- Science Direct -> 22 articole

- IEEE Explore -> 8 articole

- Scopus -> 17 articole

- Web Of Science -> 6 articole

ELIMINAND ARTICOLELE DIN 2021

- Science Direct -> 15 articole

- IEEE Explore -> 2 articole

- Scopus -> 12 articole

- Web Of Science -> 2 articole

DUPA INJUMATATIRE DUPA CITIRE ABSTRACT SI KEYWORDS

- Science Direct -> 8 articole

- IEEE Explore -> 1 articole

- Scopus -> 7 articole

- Web Of Science -> 2 articole

LISTA ARTICOLE INAINTE DE CITIRE INTEGRALA:

[1]

M. Kumar, C. Sharma, S. Sharma, N. Nidhi, and N. Islam, “Analysis of Feature Selection and Data Mining Techniques to Predict Student Academic Performance,” in *2022 International Conference on Decision Aid Sciences and Applications (DASA)*, Mar. 2022, pp. 1013–1017. doi: [10.1109/DASA54658.2022.9765236](https://doi.org/10.1109/DASA54658.2022.9765236).

[2]

A. Al-Zawqari, D. Peumans, and G. Vandersteen, “A flexible feature selection approach for predicting students’ academic performance in online courses,” *Computers and Education: Artificial Intelligence*, vol. 3, p. 100103, Jan. 2022, doi: [10.1016/j.caeai.2022.100103](https://doi.org/10.1016/j.caeai.2022.100103).

[3]

A. C. M. Yang, I. Y. L. Chen, B. Flanagan, and H. Ogata, “How students’ self-assessment behavior affects their online learning performance,” *Computers and Education: Artificial Intelligence*, vol. 3, p. 100058, Jan. 2022, doi: [10.1016/j.caeai.2022.100058](https://doi.org/10.1016/j.caeai.2022.100058).

[4]

S. Guzmán-Castillo *et al.*, “Implementation of a Predictive Information System for University Dropout Prevention,” *Procedia Computer Science*, vol. 198, pp. 566–571, Jan. 2022, doi: [10.1016/j.procs.2021.12.287](https://doi.org/10.1016/j.procs.2021.12.287).

[5]

H. Pallathadka, B. Sonia, D. T. Sanchez, J. V. De Vera, J. A. T. Godinez, and M. T. Pepito, “Investigating the impact of artificial intelligence in education sector by predicting student performance,” *Materials Today: Proceedings*, vol. 51, pp. 2264–2267, Jan. 2022, doi: [10.1016/j.matpr.2021.11.395](https://doi.org/10.1016/j.matpr.2021.11.395).

[6]

Y. Feldman-Maggor, R. Blonder, and I. Tuvi-Arad, “Let them choose: Optional assignments and online learning patterns as predictors of success in online general chemistry courses,” *The Internet and Higher Education*, vol. 55, p. 100867, Oct. 2022, doi: [10.1016/j.iheduc.2022.100867](https://doi.org/10.1016/j.iheduc.2022.100867).

[7]

J. Niyogisubizo, L. Liao, E. Nziyumva, E. Murwanashyaka, and P. C. Nshimyumukiza, “Predicting student’s dropout in university classes using two-layer ensemble machine learning approach: A novel stacked generalization,” *Computers and Education: Artificial Intelligence*, vol. 3, p. 100066, Jan. 2022, doi: [10.1016/j.caeai.2022.100066](https://doi.org/10.1016/j.caeai.2022.100066).

[8]

W. Rujuan and W. Lei, “Research on E-learning Behavior Evaluation of Students Based on Three-way Decisions Classification Algorithm,” *Procedia Computer Science*, vol. 208, pp. 367–373, Jan. 2022, doi: [10.1016/j.procs.2022.10.052](https://doi.org/10.1016/j.procs.2022.10.052).

[9]

L. S. Rodrigues, M. dos Santos, I. Costa, and M. A. L. Moreira, “Student Performance Prediction on Primary and Secondary Schools-A Systematic Literature Review,” *Procedia Computer Science*, vol. 214, pp. 680–687, Jan. 2022, doi: [10.1016/j.procs.2022.11.229](https://doi.org/10.1016/j.procs.2022.11.229).

[10]

S. Verma, R. K. Yadav, and K. Kholiya, “A Scalable Machine Learning-based Ensemble Approach to Enhance the Prediction Accuracy for Identifying Students at-Risk,” *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 8, pp. 185–192, 2022, doi: [10.14569/IJACSA.2022.0130822](https://doi.org/10.14569/IJACSA.2022.0130822).

[11]

M. F. Yacoub, H. A. Maghawry, N. A. Helal, T. F. Gharib, and S. Ventura, “An Enhanced Predictive Approach for Students’ Performance,” *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 4, pp. 879–883, 2022, doi: [10.14569/IJACSA.2022.01304101](https://doi.org/10.14569/IJACSA.2022.01304101).

[12]

S. Batool, J. Rashid, M. W. Nisar, J. Kim, H.-Y. Kwon, and A. Hussain, “Educational data mining to predict students’ academic performance: A survey study,” *Education and Information Technologies*, 2022, doi: [10.1007/s10639-022-11152-y](https://doi.org/10.1007/s10639-022-11152-y).

[13]

S. Begum and S. S. Padmannavar, “Prediction of Student Performance using Genetically Optimized Feature Selection with Multiclass Classification,” *International Journal of Engineering Trends and Technology*, vol. 70, no. 4, pp. 223–235, 2022, doi: [10.14445/22315381/IJETT-V70I4P219](https://doi.org/10.14445/22315381/IJETT-V70I4P219).

[14]

A. B. Urbina-Nájera and L. A. Méndez-Ortega, “Predictive Model for Taking Decision to Prevent University Dropout,” *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 7, no. 4, pp. 205–213, 2022, doi: [10.9781/ijimai.2022.01.006](https://doi.org/10.9781/ijimai.2022.01.006).

[15]

S. Begum and S. S. Padmannavar, “Student performance prediction with BPSO feature selection and CNN classifier,” *International Journal of Advanced and Applied Sciences*, vol. 9, no. 11, pp. 84–92, 2022, doi: [10.21833/ijaas.2022.11.010](https://doi.org/10.21833/ijaas.2022.11.010).

[16]

M. Q. Memon, Y. Lu, S. Yu, A. Memon, and A. R. Memon, “The Critical Feature Selection Approach using Ensemble Meta-Based Models to Predict Academic Performances,” *International Arab Journal of Information Technology*, vol. 19, no. Special Issue 3A, pp. 523–529, 2022, doi: [10.34028/iajit/19/3A/12](https://doi.org/10.34028/iajit/19/3A/12).

[17]

“Chicken Swarm-Based Feature Subset Selection with Optimal Machine Learning Enabled Data Mining Approach-Web of Science Core Collection.”<https://www-webofscience-com.am.e-nformation.ro/wos/woscc/full-record/WOS:000823686500001> (accessed Jan. 03, 2023).

[18]

“Feature Selection with Optimal Stacked Sparse Autoencoder for Data Mining-Web of Science Core Collection.”<https://www-webofscience-com.am.e-nformation.ro/wos/woscc/full-record/WOS:000779564300024> (accessed Jan. 03, 2023).

GREEN – CITITE SI SELECTATE

YELLOW – CITITE, NU TRATEAZA SUBIECTUL DESTUL DE IN DETALIU, PASTRATE CA BACKUP

RED – CITITE, PREA PUTINE DETALII DESPRE PROCESUL DE FEATURE SELECTION, ELIMINATE

WHITE – NECITITE, PASTRATE CA BACKUP

REZUMAT ARTICOLE:

ART 1 BAZAT PE SISTEMUL EDUCATIONAL PORTUGHEZ, SCOLI; FEATURE SELECTION (WRAPPER METHODS)

ART 2 ONLINE LEARNING; DETAILED FEATURE SELECTION

ART 3 ONLINE; SELF ASSESMENT; SOME DETAILS ON FEATURE SELECTION

ART 4 COLUMBIA; FEATURE SELECTION (RANDOM FOREST)

ART 5 UNIVERSITATE ITALIANA; TOO FEW DETAILS ON FEATURE SELECTION

ART 6 UNIVERSITATE DIN ISRAEL; ONLINE; VARIABLE SELECTION USED IN DECISION TREE ALGORITHM

ART 7 SLOVAKIA; FEATURE SELECTION DETAILED

ART 8 SUPERSTAR PLATFORM; ALMOST NO DETAILS ON FEATURE SELECTION

ART 9 SOUTH KOREA; MIDDLE AND HIGH SCHOOL; DETAILED METHODOLOGY; ALMOST NO DETAILS ON FEATURE SELECTION

ART 10 INDIA; FEATURE SELECTION ENOUGH DETAILED

ART 11 FEATURE SELECTION (SELECTKBEST)

ART 12

ART 13

ART 14 VERY DETAILED FEATURE SELECTION

ART 15 PORTUGUESE; FEATURE SELECTION BSPO

ART 16 BEIJING; FEATURE SELECTION (FILTER AND WRAPPER METHODS) – VERY DETAILED

ART 17

ART 18

SELECTATE IN FINAL:

11 articole

- 1 articol IEEE Explore

- 5 articole Science Direct

- 5 articole Scopus

**3.** **Results**

**4.** **Discussion**

**5.** **Conclusion**

References:

[1]

A. Al-Zawqari, D. Peumans, and G. Vandersteen, “A flexible feature selection approach for predicting students’ academic performance in online courses,” *Computers and Education: Artificial Intelligence*, vol. 3, p. 100103, Jan. 2022, doi: [10.1016/j.caeai.2022.100103](https://doi.org/10.1016/j.caeai.2022.100103).

[2]

S. Verma, R. K. Yadav, and K. Kholiya, “A Scalable Machine Learning-based Ensemble Approach to Enhance the Prediction Accuracy for Identifying Students at-Risk,” *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 8, pp. 185–192, 2022, doi: [10.14569/IJACSA.2022.0130822](https://doi.org/10.14569/IJACSA.2022.0130822).

[3]

M. F. Yacoub, H. A. Maghawry, N. A. Helal, T. F. Gharib, and S. Ventura, “An Enhanced Predictive Approach for Students’ Performance,” *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 4, pp. 879–883, 2022, doi: [10.14569/IJACSA.2022.01304101](https://doi.org/10.14569/IJACSA.2022.01304101).

[4]

M. Kumar, C. Sharma, S. Sharma, N. Nidhi, and N. Islam, “Analysis of Feature Selection and Data Mining Techniques to Predict Student Academic Performance,” in *2022 International Conference on Decision Aid Sciences and Applications (DASA)*, Mar. 2022, pp. 1013–1017. doi: [10.1109/DASA54658.2022.9765236](https://doi.org/10.1109/DASA54658.2022.9765236).

[5]

A. C. M. Yang, I. Y. L. Chen, B. Flanagan, and H. Ogata, “How students’ self-assessment behavior affects their online learning performance,” *Computers and Education: Artificial Intelligence*, vol. 3, p. 100058, Jan. 2022, doi: [10.1016/j.caeai.2022.100058](https://doi.org/10.1016/j.caeai.2022.100058).

[6]

S. Guzmán-Castillo *et al.*, “Implementation of a Predictive Information System for University Dropout Prevention,” *Procedia Computer Science*, vol. 198, pp. 566–571, Jan. 2022, doi: [10.1016/j.procs.2021.12.287](https://doi.org/10.1016/j.procs.2021.12.287).

[7]

Y. Feldman-Maggor, R. Blonder, and I. Tuvi-Arad, “Let them choose: Optional assignments and online learning patterns as predictors of success in online general chemistry courses,” *The Internet and Higher Education*, vol. 55, p. 100867, Oct. 2022, doi: [10.1016/j.iheduc.2022.100867](https://doi.org/10.1016/j.iheduc.2022.100867).

[8]

J. Niyogisubizo, L. Liao, E. Nziyumva, E. Murwanashyaka, and P. C. Nshimyumukiza, “Predicting student’s dropout in university classes using two-layer ensemble machine learning approach: A novel stacked generalization,” *Computers and Education: Artificial Intelligence*, vol. 3, p. 100066, Jan. 2022, doi: [10.1016/j.caeai.2022.100066](https://doi.org/10.1016/j.caeai.2022.100066).

[9]

A. B. Urbina-Nájera and L. A. Méndez-Ortega, “Predictive Model for Taking Decision to Prevent University Dropout,” *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 7, no. 4, pp. 205–213, 2022, doi: [10.9781/ijimai.2022.01.006](https://doi.org/10.9781/ijimai.2022.01.006).

[10]

S. Begum and S. S. Padmannavar, “Student performance prediction with BPSO feature selection and CNN classifier,” *International Journal of Advanced and Applied Sciences*, vol. 9, no. 11, pp. 84–92, 2022, doi: [10.21833/ijaas.2022.11.010](https://doi.org/10.21833/ijaas.2022.11.010).

[11]

M. Q. Memon, Y. Lu, S. Yu, A. Memon, and A. R. Memon, “The Critical Feature Selection Approach using Ensemble Meta-Based Models to Predict Academic Performances,” *International Arab Journal of Information Technology*, vol. 19, no. Special Issue 3A, pp. 523–529, 2022, doi: [10.34028/iajit/19/3A/12](https://doi.org/10.34028/iajit/19/3A/12).