

Systematic Review of AI-Based Early Warning System as the Basis for Preventing Dropouts and Improving Retention of History Education Students

Arwin Anin ^(1*) Malkisedek Taneo ⁽²⁾ Silvanus Jefrino Mali ⁽³⁾

^(1,2,3) History Education, FKIP Nusa Cendana University, Kupang, Indonesia

Received: 2026, 07,24 Accepted: 2026, 08,28
Available online: 2026, 09,28

* Corresponding author.

E-mail addresses: arwinanin324@gmail.com*, taneomelky67@undana.ac.id, silvanus_mali@staf.undana.ac.id

KEYWORDS	ABSTRACT
Keywords: Early Warning System; Artificial Intelligence; Machine Learning; Dropout; Student Retention. Copyright © 2026 Vifada Journal of Education. All rights reserved.	The phenomenon of student dropout poses a serious challenge for higher education institutions in achieving SDG 4 on quality education. This study aims to assess the effectiveness of an Artificial Intelligence (AI)-based Early Warning System (EWS) as a basis for developing an early detection system for dropout risk and improving student retention, particularly in the History Education Study Program. The method used was a Systematic Literature Review (SLR) with the PRISMA 2020 framework, through a search on Google Scholar between 2018 and 2026, resulting in 11 final articles that met the inclusion criteria. The study results indicate that the AI-based EWS is capable of detecting students at risk of dropping out with an accuracy of 82%-96.8% and an AUC of up to 0.9615. Initial Semester Grade Point Average (GPA) was consistently the most important predictor, and accuracy increased significantly when integrated with socio-economic data. The adaptive EWS proved superior to the static system. These findings provide a scientific basis for the History Education Study Program to develop a contextual, effective, and sustainable dropout prediction model.

1. Introduction

Higher education plays a central role in realizing the Sustainable Development Goals (SDGs), especially the fourth pillar that mandates quality, inclusive, and equitable education. A critical indicator of institutional quality is the success of students in completing their studies as well as a real contribution to human resource development. The high graduation rate also reflects a healthy academic ecosystem that is able to support students to complete their studies until graduation.

The phenomenon of dropping out of college or Drop out is still a serious problem in Indonesian universities, and even around the world. Students who fail to complete their studies on time have multiple impacts including personal losses for individuals, inefficiency of education budgets, decreased institutional credibility, and obstacles to achieving the SDGs targets nationally. Dropout is a big challenge for universities, funding agencies and also for the students themselves, so a deep understanding of the causative factors is absolutely necessary so that the response designed is right on target (Matz et al. 2023).

The History Education Study Program, in the fields of social sciences and humanities, also faces the same challenges. Many students consider that job opportunities after graduation are relatively limited, so the enthusiasm to finish college can decrease. In addition, difficulties in attending lectures, economic problems, lack of involvement in campus activities, and low support from family and friends can increase the likelihood of students dropping out of college. These various factors can make students choose to leave their studies before completing their education. Therefore, it is necessary to

make efforts to identify students who are at risk from the beginning so that they can be given the right support.

Higher education is now gaining various innovative opportunities in line with the rapid development of Artificial Intelligence (AI). One of the innovations for universities that is getting more attention is Early Warning System (EWS) is an AI-based early warning system that integrates various artificial intelligence algorithms as a processing engine to detect at-risk students Drop out as early as possible. The use of AI in education is supported by the development of the field Educational Data Mining (EDM) and Learning Analytics which relies on student data analysis to produce predictions and support appropriate academic decision-making (Lemay et al. 2021).

As a system, AI-based EWS works by comprehensively processing academic data starting from the track record of attendance, exam scores and activities in Learning Management System (LMS), to social interaction patterns in the campus environment by utilizing algorithms Machine Learning and deep learning as the analytical engine. High predictive capacity allows universities to take timely intervention steps. AI-based EWS is able to identify at-risk students from the early stages and support the provision of targeted interventions so that potential dropouts can be suppressed before they develop into more serious academic failures (Bañeres et al. 2023).

Although AI-based EWS has been widely applied in higher education, research findings are still diverse in terms of approaches, algorithms, and predictive variables. Therefore, a systematic review is needed to understand the effectiveness of EWS in preventing dropouts and increasing student retention. The results of this study are expected to be the basis for the development of an early detection system for at-risk students in the History Education Study Program.

Departing from this urgency, this study uses the Systematic Literature Review (SLR) approach to in-depth examine various findings related to AI-based EWS in preventing student dropouts. The main question to be answered is: how effective is the implementation of AI-based EWS in preventing dropouts, increasing student retention, and supporting the achievement of SDG 4 (Quality Education) in higher education. Through this systematic review, it is hoped that evidence-based recommendations will be prepared for stakeholders, especially the History Education Study Program in an effort to prevent dropouts and increase student retention rates and support the SDGs agenda in the field of education.

2. Research Design and Methodology

This study uses the Systematic Literature Review (SLR), which is a research method that is conducted in a structured and transparent manner and can be replicated to identify, evaluate, and synthesize scientific findings that are related to the research topic (Snyder 2019). The reporting process refers to the PRISMA 2020 framework Page et al. (2021) to ensure that the literature selection takes place in a transparent, systematic and traceable manner.

The study centered on the question of how effective is AI-based EWS in detecting dropout risk and increasing student retention in higher education? A search was conducted on Google Scholar which includes Scopus and SINTA indexed publications using the keywords "Early Warning System", "Artificial Intelligence", "Machine Learning", "student dropout", "academic retention", "artificial intelligence", and "academic retention", limited to the 2018-2026 edition.

Inclusion criteria include peer-reviewed articles in Indonesian or English with valid DOIs and direct relevance to the topic. Articles without access to complete text or outside the scope of the study were excluded. Of the initial 45 articles, 30 relevant articles were screened (15 national, 15 international). Furthermore, the results of the theme suitability check determined 20 relevant articles to be further reviewed. At the stage of verifying the completeness of the publication's identity, 9 articles were excluded because they did not have a valid DOI, resulting in 11 articles that were designated as the final material for the study. The selection flow is presented in the following diagram.

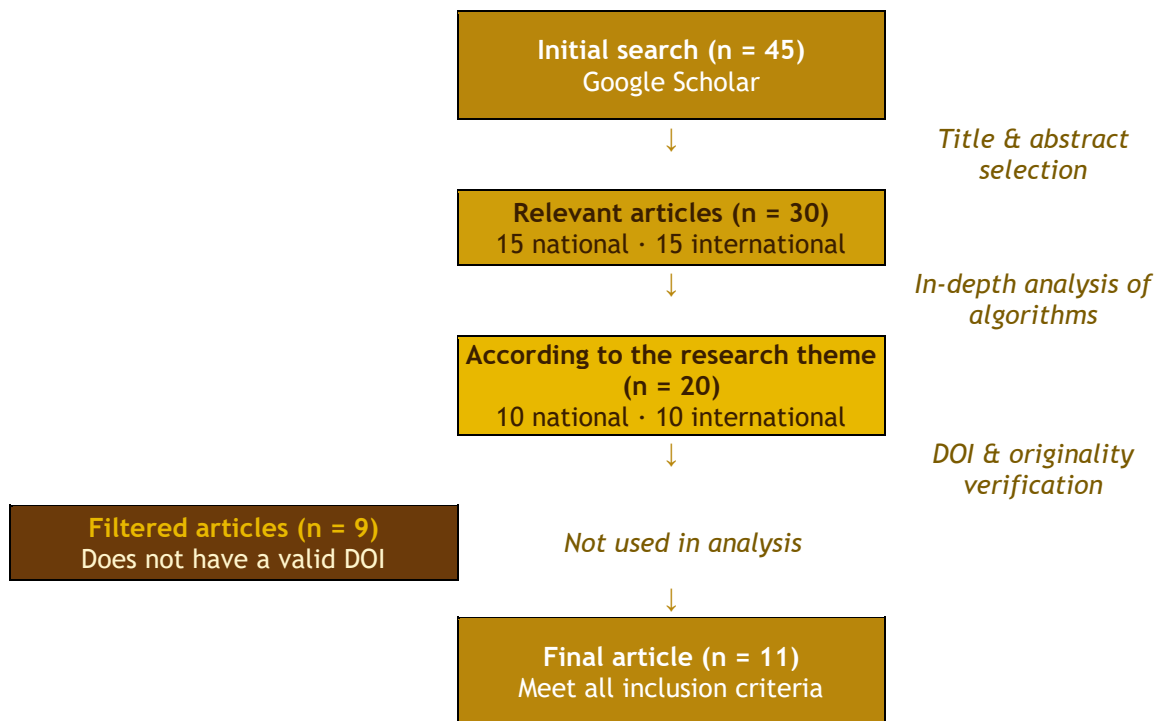


Figure 1. Flow Diagram of the Inclusion and Exclusion Process at the Systematic Literature Review Stage (n: number of articles)

3. Findings and Discussion

The AI-based Early Warning System (EWS) is an early warning system developed with the aim of identifying students at risk of dropouts by utilizing academic and non-academic data. Various studies show that the application of AI-based EWS is able to help universities detect at-risk students as early as possible and thus be given appropriate interventions before students actually leave their studies. In addition, EWS in its application also uses various machine learning algorithms that are proven to increase the accuracy of predictions and support efforts to increase student retention. Based on the results of the analysis of 11 articles that met the inclusion criteria, the information on the results of the research is presented in the following table.

Table 1. Research Findings Related to the Effectiveness of AI-Based EWS in Preventing *Dropouts* and Improving Student Retention.

Researcher and Year	Article Title	Research Results
F.A. Putra et al. (2025)	<i>Prediksi Dropout Mahasiswa Early-Warning Berbasis Enrollment dengan Machine Learning</i>	Two-phase EWS research based <i>Machine Learning</i> shows that Logistic Regression is most effective in the <i>enrollment-only</i> (recall 0.697; F1 0.651). The addition of semester 1 data increased recall to 0.792 and reduced false alarms by 45%, so that EWS is more operational and efficient in supporting student retention (F. A. Putra et al. 2025).
J. A. Putra & Septian (2025)	<i>Penerapan Algoritma Naive Bayes Dalam Memprediksi Tingkat Mahasiswa Drop out Pada Program Studi Pendidikan Informatika</i>	The Naive Bayes algorithm results in 82% accuracy, 87% recall, and 81% F1-score in predicting <i>Drop out</i> 1,181 UTS Informatics students. Of 274 cases <i>Drop out</i> Actual, 239 were identified as true. Django's web-based system allows for early monitoring and intervention by academic staff directly (J. A. Putra and Septian 2025).

Rahma & Budiman Kusdinar, (2024)	<i>Prediksi Mahasiswa Non-Aktif Menggunakan Algoritma C4.5 (Studi Kasus : Universitas Muhammadiyah Sukabumi)</i>	The C4.5 algorithm based on CRISP-DM produced an accuracy of 96.8%, accuracy of 98%, recall of 94.2%, and an F1-score of 96.5% in the prediction of non-active students of the University of Muhammadiyah Sukabumi. The integration of socio-economic data of parents with social studies in semesters 1-4 has been proven to significantly increase the accuracy of predictions compared to the use of academic data alone (Rahma et al. 2024).
Setiawan et al. (2024)	<i>Optimisasi Kesuksesan Akademis: Studi Sistem Deteksi Dini untuk Mengidentifikasi Potensi Drop Out Mahasiswa Tingkat Akhir dengan Protokol Intervensi Cepat pada Perguruan Tinggi</i>	The logistics regression-based EWS system successfully detects <i>Drop out</i> final year students with high accuracy using pre- and post-entry administrative data. The findings confirm that the relevance of predictors changes over time and varies between subgroups of students, so adaptive EWS is more effective than static systems (Setiawan and Wibowo 2024).
Irmayanti et al. (2021)	<i>Prediksi Mahasiswa Berpotensi Drop Out dengan Metode Iteratif Dichotomiser 3 (ID3)</i>	The ID3 algorithm generated a 50-node decision tree with IPS4 as the root node as well as 83.1% accuracy, 86.2% accuracy, and 73.2% recall. Of the 577 data of 2017 batch of students, 281 were identified as potential dropouts. Integration of predictions into academic information systems is recommended to support data-driven preventive coaching (Irmayanti et al. 2021).
Anwar et al. (2021)	<i>Model Prediksi Dropout Mahasiswa Menggunakan Teknik Data Mining</i>	The C4.5/J48 algorithm in WEKA produces 90.6% accuracy with IPS 1 and IPS 2 as the most determinant predictors of dropout. IPS 2 below 2.29 is an indicator <i>Drop out</i> strong, but a 4 GPA improvement above 2.64 can reverse the risk to graduation (Anwar et al. 2021).
de Oliveira et al. (2021)	<i>How Does Learning Analytics Contribute to Prevent Students' Dropout in Higher Education: A Systematic Literature Review</i>	Classification techniques based on <i>Decision Tree</i> and <i>Neural Network</i> is the most dominant in predicting dropouts. Academic and personal data of students are the main features. The prediction model is able to identify students at risk early, allowing institutions to take preventive measures before <i>Drop out</i> occurs (De Oliveira et al. 2021).
Villegas-Ch et al. (2023)	<i>Improving Student Retention in Institutions of Higher Education through Machine Learning: A Sustainable Approach</i>	Models Neural Network ML-based achieves 85% accuracy, Recall 80%, and F1 Score 82% in predicting student retention. Implementation in 500 students increased the retention rate from 75% to 81% (up 6.875%). This approach allows for early identification of at-risk students Drop out and the provision of measurable and sustainable personal interventions (Villegas-Ch et al. 2023).
Cannistrà et al. (2022)	<i>Early-predicting dropout of university students:</i>	Multilevel model based Generalised Linear Mixed Model (GLMM) and Random Forest reached AUC as high as 0.9615 in prediction

	an application of innovative multilevel machine learning and statistical techniques	Early Dropout engineering students of the Politecnico di Milano. First-semester credit count proved to be the most dominant predictor, with multilevel models outperforming non-hierarchical models in the accuracy of detection of at-risk students (Cannistrà et al. 2022).
Umendu et al. (2026)	Enhancing Student Retention in Higher Education Institutions (HEIs): Machine Learning Approach	Models Extra Trees yielded an AUC of 0.96 and an accuracy of 87.4% in multi-class classifications (Dropout, Enrolled, Graduate) on 4,424 student data. The SHAP analysis identified approved academic units and tuition payment status as the most influential predictors, confirming the central role of early academic performance and financial stability on student retention (Umendu et al. 2026).
Gonzalez-Nucamendi et al. (2023)	Predictive Analytics Study To Determine Undergraduate Student at Risk of Dropout	Random Forest by the Threshold Probability in 14,495 S1 students achieved precision Drop out up to 51% and retention of 81%. The average score in the first period of the first semester, the average score of the previous level, and the score of the entrance exam are the three most determinants in prediction Drop out First-year students (Gonzalez-Nucamendi et al. 2023).

A review of eleven articles showed that the Early Warning System (EWS) based on artificial intelligence is effective in detecting students at risk of dropout. This effectiveness is demonstrated by a high level of prediction accuracy, ranging from 82% to 96.8%, and an Area Under the Curve (AUC) value of 0.9615 in some more complex models. This level of accuracy allows universities to identify at-risk students early on and provide interventions before dropout decisions occur.

In terms of algorithms, the research studied shows the use of various machine learning methods. In national studies, the C4.5 and ID3 algorithms are the most widely used approaches. Rahma et al. (2024) reported that C4.5 resulted in an accuracy of 96.8%, especially when academic data was combined with student socio-economic data. Similar results were found by Anwar et al. (2021) which places the Semester Achievement Index (IPS) 1 and 2 as the main predictors of risk Drop out. Meanwhile, Irmayanti et al. (2021) showing that early semester social studies is the most dominant factor in the decision tree built using the ID3 algorithm.

At the international research level, it generally uses more complex methods, such as Random Forest and Generalised Linear Mixed Model (GLMM), and Extra Trees. Research by Cannistrà et al. (2022) indicates that the combination of GLMM and Random Forest resulting in an AUC value of 0.9615. In addition, Umendu et al. (2026) found that influencing factors Drop out not only related to academic achievement, but also the financial condition of students. The findings show that the problem Drop out influenced by various factors so that it cannot be explained only from the academic aspect.

Almost all studies showed consistency in two main groups of variables, namely early academic performance and multidimensional student data. Scores in the first and second semesters were consistently the strongest indicators. Gonzalez-Nucamendi et al. (2023) added that previous education scores and entrance exam results also contribute to predicting risk Drop out. Thus, the early signs of risk can be recognized from the beginning of the lecture period.

Another development that stands out is that EWS is moving into a more adaptive system, revealed by Setiawan and Wibowo (2024) that the relevance of predictors can change over time and become different in each group of students. Therefore, a system that can periodically update the model is considered more effective than a system that has fixed parameters. Putra et al. (2025) It also shows that the addition of first-semester data improves risk detection capabilities while reducing prediction errors.

In terms of Application, reported by Villegas-Ch et al. (2023) that the application of the model Neural Network can increase student retention rates from 75% to 81%. These findings are supported by various review studies that conclude that AI-based EWS makes a significant contribution to reducing dropout rates and increasing student study success.

Based on the findings obtained, AI-based EWS has potential and can be developed in the History Education Study Program as a detector for students at risk of dropout. There are various indicators that consistently appear in the research, such as the Semester Achievement Index (IPS) in the first semester, the level of attendance at lectures and digital learning activities, and the socio-economic factors of students can be used as a basis for the preparation of prediction models that are in accordance with the characteristics of students of the History Education Study Program. The use of these indicators can be a basis for the study program in identifying students who need academic assistance from an early age so that the potential for dropouts can be minimized.

Furthermore, the application of AI-based EWS also has the potential to support faster and more objective decision-making based on data during the student development process. This system actually does not replace the role of academic supervisors, but becomes a supporting tool that helps identify students before experiencing a more serious decline in academic performance. Although its implementation still requires empirical validation in the context of the History Education Study Program, thus the results of this systematic review show that AI-based EWS can be used as a basis for the development of strategies to prevent dropouts and increase student retention effectively and sustainably.

4. Conclusion

A systematic review of eleven articles confirmed that the Early Warning System (EWS) based on artificial intelligence has proven to be effective in detecting students at risk of dropping out early, with a prediction accuracy ranging from 82%-96.8% and an AUC of 0.9615. Early semester academic performance (IPS 1 and 2) has consistently been the most determinant predictor, and accuracy has increased significantly when integrated with students' socio-economic and financial data. The progress towards adaptive EWS has also proven to be superior to static systems, with an increase in recall of up to 0.792 and a reduction in false alarms of 45%.

These findings are relevant for the History Education Study Program: initial social studies indicators, attendance, digital activities, and socio-economic conditions can immediately be used as the basis for contextual prediction models. The next steps include structured data collection, pilot testing of the C4.5 or Random Forest algorithm, as well as the development of data governance and ethical frameworks before full implementation. Thus, AI-based EWS has the potential to become a strategic instrument for preventing dropouts as well as making a real contribution to the achievement of SDG 4 on inclusive and equitable quality education.

Statement on the Use of Generative Artificial Intelligence (AI)

During the preparation of this manuscript, the authors used generative artificial intelligence (AI) tools solely to support the scientific writing process. Specifically, Grammarly was used to improve grammar, language quality, readability, and writing clarity. The AI tools were not used to generate research ideas, analyze data, interpret findings, or draw scientific conclusions. All study design, data collection, analysis, interpretation of results, and conclusions were conducted and verified entirely by the authors. The authors take full responsibility for the accuracy, integrity, and originality of the content presented in this manuscript.

References

- Anwar, Muchamad Taufiq, Lucky Heriyanto, and Fadhla Fanini. 2021. "Model Prediksi Dropout Mahasiswa Menggunakan Teknik Data Mining." *Jurnal Informatika Upgris* 7 (1). <https://doi.org/10.26877/jiu.v7i1.8023>
- Bañeres, David, M. Elena Rodríguez-González, Ana-Elena Guerrero-Roldán, and Pau Cortadas. 2023. "An Early Warning System to Identify and Intervene Online Dropout Learners." *International Journal of Educational Technology in Higher Education* 20 (1): 3. <https://doi.org/10.1186/s41239-022-00371-5>.

- Cannistrà, Marta, Chiara Masci, Francesca Ieva, Tommaso Agasisti, and Anna Maria Paganoni. 2022. "Early-Predicting Dropout of University Students: An Application of Innovative Multilevel Machine Learning and Statistical Techniques." *Studies in Higher Education* 47 (9): 1935-56. <https://doi.org/10.1080/03075079.2021.2018415>.
- De Oliveira, Catarina Félix, Sónia Rolland Sobral, Maria João Ferreira, and Fernando Moreira. 2021. "How Does Learning Analytics Contribute to Prevent Students' Dropout in Higher Education: A Systematic Literature Review." *Big Data and Cognitive Computing* 5 (4): 64. <https://doi.org/10.3390/bdcc5040064>.
- Gonzalez-Nucamendi, Andres, Julieta Noguez, Luis Neri, Víctor Robledo-Rella, and Rosa María Guadalupe García-Castelán. 2023. "Predictive Analytics Study to Determine Undergraduate Students at Risk of Dropout." *Frontiers in Education* 8 (October): 1244686. <https://doi.org/10.3389/feduc.2023.1244686>.
- Irmayanti, Dede, Yusuf Muhyidin, and Dede Arif Nurjaman. 2021. "Prediksi Mahasiswa Berpotensi Drop Out Dengan Metode Iteratif Dichotomiser 3 (ID3)." *Jurnal Teknologi Informasi* 5 (2): 103-13. <https://doi.org/10.36294/jurti.v5i2.2054>.
- Lemay, David J., Clare Baek, and Tenzin Doleck. 2021. "Comparison of Learning Analytics and Educational Data Mining: A Topic Modeling Approach." *Computers and Education: Artificial Intelligence* 2: 100016. <https://doi.org/10.1016/j.caeai.2021.100016>.
- Matz, Sandra C., Christina S. Bukow, Heinrich Peters, Christine Deacons, Alice Dinu, and Clemens Stachl. 2023. "Using Machine Learning to Predict Student Retention from Socio-Demographic Characteristics and App-Based Engagement Metrics." *Scientific Reports* 13 (1): 5705. <https://doi.org/10.1038/s41598-023-32484-w>.
- Page, Matthew J., Joanne E. McKenzie, Patrick M. Bossuyt, et al. 2021. "The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews." *Journal of Clinical Epidemiology* 134 (June): 178-89. <https://doi.org/10.1016/j.jclinepi.2021.03.001>.
- Putra, Febri Andika, Syahisro Mirajdandi, Nandra Nandra, Bisma Okmarizal, and Sandy Mulyanda. 2025. "Prediksi Dropout Mahasiswa: Early-Warning Berbasis Enrollment Dengan Machine Learning." *JURNAL FASILKOM* 15 (3): 465-73. <https://doi.org/10.37859/jf.v15i3.10714>.
- Putra, Juniardi Akhir, and Bayu Septian. 2025. "Penerapan Algoritma Naive Bayes Dalam Memprediksi Tingkat Mahasiswa Drop Out Pada Program Studi Informatika." *Digital Transformation Technology* 4 (2): 1227-35. <https://doi.org/10.47709/digitech.v4i2.5460>.
- Rahma, Leonie Siti, Asep Budiman Kusdinar, and Prajoko Prajoko. 2024. "PREDIKSI MAHASISWA NON-AKTIF MENGGUNAKAN ALGRITMA C4.5: STUDI KASUS : UNIVERSITAS MUHAMMADIYAH SUKABUMI." *JATI (Jurnal Mahasiswa Teknik Informatika)* 8 (5): 10204-11. <https://doi.org/10.36040/jati.v8i5.10937>.
- Setiawan, Ismail, and Kresno Ario Tri Wibowo. 2024. "Optimisasi Kesuksesan Akademis: Studi Sistem Deteksi Dini Untuk Mengidentifikasi Potensi Dropout Mahasiswa Tingkat Akhir Dengan Protokol Intervensi Cepat Pada Perguruan Tinggi." *Jurnal Riset Sistem Dan Teknologi Informasi* 2 (1): 57-63. <https://doi.org/10.30787/restia.v2i1.1424>.
- Snyder, Hannah. 2019. "Literature Review as a Research Methodology: An Overview and Guidelines." *Journal of Business Research* 104 (November): 333-39. <https://doi.org/10.1016/j.jbusres.2019.07.039>.
- Umendu, Emeka Cajetan, Mustansar Ghanzanfar, Aaron Kans, and Md Atiqur Rahman Ahad. 2026. "Enhancing Student Retention in Higher Education Institutions (HEIs): Machine Learning Approach." *Electronics* 15 (4): 734. <https://doi.org/10.3390/electronics15040734>.
- Villegas-Ch, William, Jaime Govea, and Solange Revelo-Tapia. 2023. "Improving Student Retention in Institutions of Higher Education through Machine Learning: A Sustainable Approach." *Sustainability* 15 (19): 14512. <https://doi.org/10.3390/su151914512>.