Machine Learning Utilization to Predict Potential At-Risk Students in Higher Education: A Literature Review

Ahmad Yazid, Melva Sari Panjaitan, Fetty Fitriyanti Lubis, Yusep Romansyah

Abstract

A prosperous and functional society always requires capable and qualified human resources to operate. Ensuring its continuity demands the success of education as the foundation for fostering and nurturing future generations and enhancing human capacity. This study aims to provide insights into various state-of-the-art predictive models to improve student retention in higher education. This study examined 37 reputable papers selected through systematic literature review techniques and discussed 7 theoretical models, 7 relational models, and 14 predictive models. The best predictive models were able to attain 93.0% accuracy using linear regression and random forest algorithms. The study found that the development of predictive machine learning requires thorough planning and preparation suited to each educational institution, and its utilization is highly recommended to enhance student retention in higher education.

Keywords: Machine Learning, Predictive Models, Higher Education, Theoretical Models, Literature Review, Artificial Intelligence

Ahmad Yazid
School of Electrical Engineering and Informatics, Institut Teknologi Bandung, Indonesia
Fakultas Sains Komputasi dan Kecerdasan Digital, Universitas Pembangunan Panca Budi, Indonesia
e-mail: 33222303@mahasiswa.itb.ac.id

Melva Sari Panjaitan, Fetty Fitriyanti Lubis, Yusep Romansyah
e-mail: 33222309@mahasiswa.itb.ac.id, fettyfitriyanti@itb.ac.id, yusep@itb.ac.id
2nd International Conference on Islamic Community Studies (ICICS)
Theme: History of Malay Civilization and Islamic Human Capacity and Halal Hub in the Globalization Era
https://proceeding.pancabudi.ac.id/index.php/ICIE/index

Introduction

Education is a vital instrument to ensure the continuation of civilization and increase human capacity, uphold sovereignty, and prepare a global workforce [1] in the global community. Educational system and educational management are pivotal to foster lifelong learning [2], which contributes socially and economically to the country [3]. In modern societies, universities are viewed as the pillar for progress and responsible citizenship [4].

Student retention reflects the overall quality of learning behavior [5], is an enduring global institutional problem [6], and a major obstacle to Higher Educational Institution (HEI) [5] success. Each dropouts are associated with considerable personal and social costs [7].

How to increase student retention, reduce dropout rates, and augment academic performance is a growing concern [8] due to the increasing surge of dropout rates [1], [9]. Such a phenomenon occurred globally [10]–[13]. Dropouts cause stacking and lasting loss individually, institutionally, nationally, or globally [14]. A dropout individual is prone to miss chances for a better career [15] and psychological distress [16]. Dropouts may force HEI and the community to do extra reallocation and expend more resources and capital [17], [18].
Dropout and change of study can occur in the first year, and severe cases have been observed in higher educational institutes with more than 21.7%[19]. Failure to address this phenomenon produces lasting effects on individuals, society, institutions, and economics [20]. There are many approaches to tackle this problem and suggest the importance of solving the problem as early as possible [21]. Research has shown that it is possible to discover hidden patterns and features through technology [3], [22], such as Artificial Intelligence [7] (AI) to build profiles [23], [24].

Despite exhaustive research on the topic, the problem persists. Therefore, there is a need for a systematic literature review that provides insights into the models that can be applied to identify the hidden patterns to reduce dropout rates in higher education.

Literature Reviews

Dropout proves to be a complex problem influenced by numerous factors [25], and the need for ongoing research remains prevalent [26]. The use of technology, such as Artificial Intelligence (AI), Data Mining (DM), and Machine Learning (ML), to leverage big data [1], [23], requires further continuous research. The current development of Educational Data Mining (EDM) has 3 general approaches in the academic field [27]: (1) predict student chance for study completion, (2) predict student risk rate of dropout, and (3) forecast student academic performance. Forecasting, explaining, and preventing dropout in HEI is an urgent issue to predict critical moments, possible to intervene timely manner [28]. Success in early prediction enables the possibility of on-time intervention [29]. Previous studies also express the lack of clear criteria evaluation metric [30].

The current research approach highlights the need to develop a recommender system [31] to help implement support, retention strategy, and a proactive approach to better manage resource allocation and achieve success [4]. A combination and ensemble of models has been proven to produce better prediction performance and identification of at-risk students [29].

Research Methodology

This paper follows the protocol from Kitchenham and Charters' Systematic Literature Review [32]: Planning, Conducting, and Reporting. It is important to declare the component of the study: (1) the population is higher education, with (2) the intervention utilized is any prediction model, in (3) the context of student retention. This section explains the review protocol and process of conducting the review: statement of research question, the defined keywords and synonyms, search process such as basic search strings used to search related research, and selection of digital library sources or research database, criteria of inclusion and exclusion, quality assessment checklists, and data extractions.

As highlighted in the introduction, the components and objectives of the review highlight the research question: What methods are utilized to identify the likelihood of a dropout?. To obtain relevant research publications, three main keywords were selected to represent research objectives: “dropout”, “predict”, and “higher education”. Due to differences in each database, the exact search string may differ from one to another. The basic keywords defined and used for the automatic search were: ("higher education" OR university) AND drop* AND (MCDM OR "Neural Network" OR predict*). The systematic search was conducted on 5 established digital library databases: ACM Digital Library, IEEE Digital Library, ScienceDirect, Scopus, and Taylor and Francis Online. This search resulted in 595 relevant documents. Table I provides detailed information regarding the digital resources employed for this research.

Table I. Digital Library Sources

No.	Name	URL
1	ACM Digital Library	http://portal.acm.org/
2	IEEE Xplore	http://ieeexplore.ieee.org/
3	ScienceDirect	http://www.sciencedirect.com/
4	Scopus	http://www.scopus.com/
5	Taylor and Francis Online	http://tandfonline.com/

Inclusion and exclusion criteria were established in Table II to be used in the research. This is to guide the research to be more precise and relevant.

Table II. Inclusion and Exclusion Criteria

No.	Inclusion	Exclusion
1	Language: English	Did not address the research question
2	Open-Accessed	Inaccessible
3	Q1-Q2 based on ScimagoJR	From a non-reputable publisher
4	Theme: Dropout in HEI	Published locally
5	Published range 2018 - 2023	Too short

To ensure the quality of the selected paper, the criteria for the paper were as follows: (1) Does the research provide empirical or mathematical (simulation) usage of the given model?, and (2) Does the research provide a specific prediction model for the research?. At this stage, the research was conducted according to the protocol that had been established in the previous chapter.

The result yielded 595 potentially relevant papers with details: (1) ACM Digital Library: 1, (2) IEEE Digital Library: 159, (3) ScienceDirect: 19, (4) Scopus: 392, and (5) Taylor and Francis: 24. 21 articles were found to be duplicates. 574 papers were selected through the application of the inclusion and exclusion criteria, with 383 papers excluded from the study. 191 papers went through a further screening process with quality assessment, which left the study with 37 papers. The final accepted papers were: (1) ACM Digital Library: 1, (2) IEEE Digital Library: 7, (3) ScienceDirect: 2, (4) Scopus: 26, and (5) Taylor and Francis: 1. The search process and conducting phase are illustrated in Figure I.

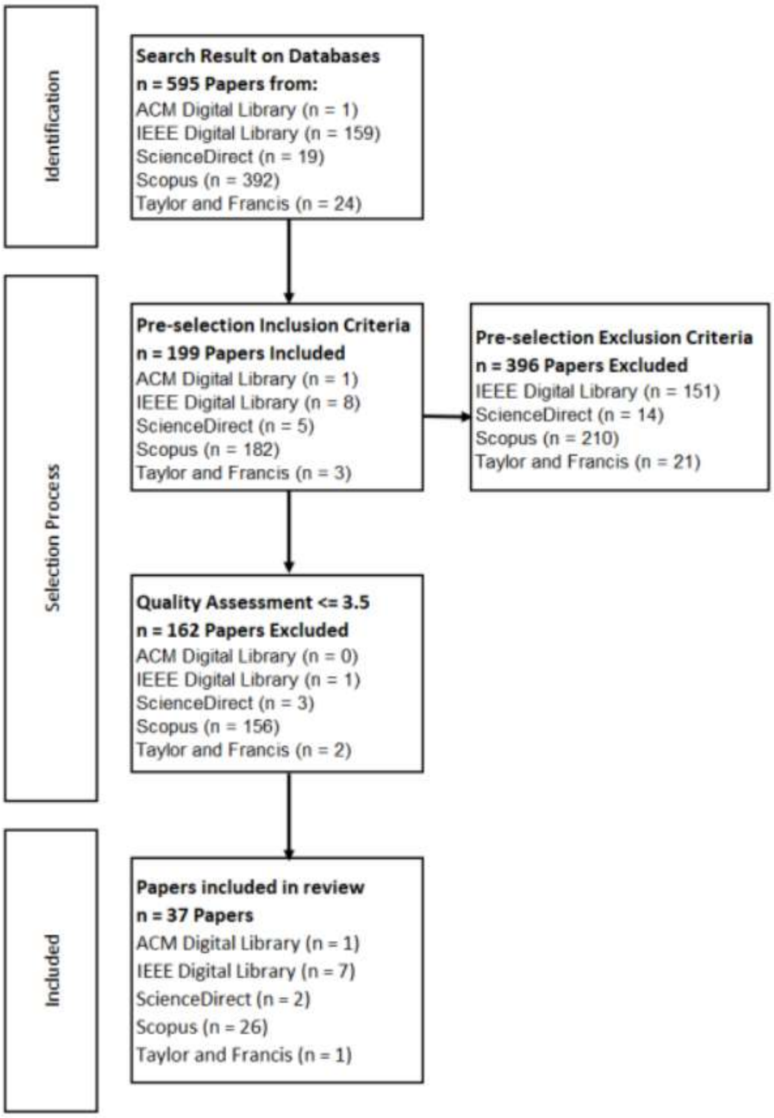

Figure I. Search Process and Conducting Phase

The details of the final included articles per year are 3 articles published in 2018, 3 articles in 2019, 13 articles in 2020, 12 articles in 2021, 6 articles in 2022, and no articles in 2023.

Results

This research examined the current methods utilized to predict and discover possible dropouts. This aspect was researched to gain insights into the current research on dropout identification and methods. The research highlights the approach methods to predict the chance of dropout in higher education.

There were many attempts to study and tackle dropout phenomena by using statistical data, educational data mining, machine learning, and deep learning to predict possible dropout in education. The recent research mentioned that the theoretical model utilized in the field has not been developed much [24]. The study reveals various methods previously used to study dropout in Table III.

Table III. Methods Utilized

No.	Name	Articles	Types	Result
1	Expectancy Value Model (EVM)	[12], [15]	Theoretical	Intent and career variance
2	Motivational Regulation Model	[9]	Theoretical	Intention correlation
3	Perceptions of Mindset Belief	[8]	Theoretical	Psychology vulnerability
4	Perceived Academic Control (PAC)	[33]	Theoretical	Early support for PAC
5	Self-Determination Theory (SDT)	[13], [39]	Theoretical	Intent variance
6	Theory of Planned Behavior (TPB)	[6]	Theoretical	Variance explanation
7	Socialization Questionnaire (SQ)	[41]	Theoretical	Socialization factor
8	Confirmatory Factor Analysis (CFA)	[6], [13], [15], [26], [37], [39], [41]	Relational	Variance explanation
9	Latent Change Model	[15]	Relational	Course variance
10	Multiple Regression Analysis	[40]	Relational	Burnout predictor
11	Survival Analysis	[34]	Relational	Viewing variance
12	Bayes	[28]	Relational	Correlation 62%
		[36]	Bayesian Multilevel	Accuracy 76.8%
		[5], [16], [18], [29], [35], [36]	Naïve Bayes (NB)	Accuracy 96.1%
13	Logistic Regression (LR)	[25], [38]	Relational	Wellbeing investment
		[11], [14]	Relational	Integrity variance
		[5], [16], [18], [24], [31], [35]	General	Accuracy 96%
		[3]	Binary	Accuracy 75.2%
		[27]	TM-OOC	Accuracy 78%
		[29]	HO	Accuracy 94.4%
14	AdaBoost Classifier	[23], [29]	HO	Accuracy 94.1%
15	Neural Network (NN)	[16], [35]	CNN	Accuracy 87.6%
		[7]	FCNN	Accuracy 72.4%
		[23], [36]	FFNN	Accuracy 72%
		[4], [24], [31]	MPNN	Accuracy 95%
		[5]	RNN	Accuracy 90%

No	Name	Articles	Types	Result
16	K-Nearest Neighbor (KNN)	[23], [29], [31], [36]	HO	Accuracy 89.6%
17	Decision Tree (DT)	[2], [5], [31]	General	Accuracy 90%
		[1]	C4.5	Accuracy 86.8%
		[35]	C5.0	Accuracy 62.2%
		[27]	TM-OOC	Accuracy 82%
		[29]	HO	Accuracy 87.6%
18	Extra Tree Classifier	[23]	General	Accuracy 91.9%
19	Regression Tree	[16], [36]	General	Accuracy 80.8%
20	Gradient Boosting (GB)	[10], [23], [31], [36]	General	Accuracy 91.7%
		[16]	GBDT	Accuracy 87.5%
		[7]	XGBoost	Accuracy 70.7%
21	Linear Discriminant Analysis (LDA)	[16]	General	Accuracy 83.9%
22	Random Forest (RF)	[5], [7], [10], [16], [18], [23], [31], [35], [36]	General	Accuracy 93%
		[17]	C5.0	Accuracy 65.5%
		[27]	TM-OOC	Accuracy 83%
		[29]	HO	Accuracy 93%
23	Radial Basis Function Network	[36]	General	Accuracy 64.6%
24	TabNet-based	[7]	General	Accuracy 71.5%
25	Support Vector Machine (SVM)	[5], [10], [16], [23], [29], [31], [35], [36]	General	Accuracy 93.9%
26	Ensemble	[10]	Ensemble	Accuracy 88.8%
27	Hybrid Ensemble	[29]	Stacking + Ensemble	Accuracy 94.8%
28	Recommender System	[31]	Ensemble	Accuracy 81.9%

TM-OOC: Temporal Multi-Objective Optimization Classifier, HO: Hyperparameter Optimization, CNN: Convolutional NN, FCNN: Fully Connected NN, FFNN: Feed Forward NN, MPNN: Multilayer Perceptron NN, RNN: Recurrent NN, GBDT: Gradient Boosted DT, XGBoost: eXtreme GB.

The results showed that 7 were on the theoretical models, 7 were on novel variables and variance explanation, and 14 studies discussed model development to increase prediction power. The most frequent model of prediction was LR with 13 times, RF 12 times, Tree-based models (DT, Extra Tree, Regression Tree) 10 times, NN 9 times, SVM 8 times, Bayesian and Naïve Bayes 6 times, KNN 4 times, AdaBoost twice, and the each of rest (LDA, Radial Basis, TabNet, Ensemble, Stacking, and Recommender system) once. The time of prediction is also important to enable early intervention and prevention; the best early prediction model was by the end of enrollment using SVM with 68.5% [10], the end of Semester 1 using ensemble with 80.7% [10], while the best accuracy was 93% using LR and RF by the end of the 2nd year [5] to mention a few notable models with clearly stated time of prediction.

Conclusion

The result of this study provides ample insight to answer the research questions on the current models utilized to face the problem of dropout in HEI. The result showed that 7 theoretical models can be adapted to approach dropout intention, 7 relational models explained the novel factors influencing such intention, and 14 studies with at least 84 models can be utilized to predict student at risk. There is no approach method that is better than another, even within the same institution [30]. Therefore, the development of any and each prediction model requires thorough planning and preparation. The development of predictive models should analyze and consider the collectability, accessibility, and availability of data that each HEI owns and stores.

Most research on the topic was conducted on the predictive power of the model, as many as 84 instances of prediction models were found within the selected studies, and the best accuracy reached was 96.1% [18]. In conclusion, the general development of a prediction model based on data mining or machine learning has shown different results based on the specific settings and backgrounds of the research. The study reveals the increasing trends of prediction models in HEI, although there was only little development on ensemble and combination models with decision support models to help decision makers in HEI.

Ensuring the achievement of optimum human capacity and the continuation of a civilization requires success in education. Reducing the risks of dropouts and increasing retention of students in HEI is essential to both causes. This study presents a systematic literature review that summarizes researched models adoptable by HEIs. Through the systematic process, 37 studies are selected and examined. The study has revealed 7 alternative theories, 7 correlational studies, and 14 studies on predictive models. This study contributes to providing information on current knowledge, possible theoretical and methodological advancements, and potential future research areas in education dropout and retention.

References

[1] A. M. Mariano, A. B. De Magalhães Lelis Ferreira, M. R. Santos, M. L. Castilho, and A. C. F. L. C. Bastos, "Decision trees for predicting dropout in Engineering Course students in Brazil," *Procedia Comput. Sci.*, vol. 214, no. C, pp. 1113–1120, 2022, doi: 10.1016/j.procs.2022.11.285.

[2] P. Nuankaew, "Dropout situation of business computer students, University of Phayao," *Int. J. Emerg. Technol. Learn.*, vol. 14, no. 19, pp. 117–131, Oct. 2019, doi: 10.3991/ijet.v14i19.11177.

[3] L. Beltrán, D. E. Gibaja-Romero, and R. M. Cantón-Croda, "Characterizing intercampus migration in a private university of Baja California, Mexico," *F1000Research*, vol. 11, p. 678, Jun. 2022, doi: 10.12688/f1000research.111079.1.

[4] M. F. Musso, C. F. R. Hernández, and E. C. Cascallar, "Predicting key educational outcomes in academic trajectories: a machine-learning approach," *High. Educ.*, vol. 80, no. 5, pp. 875–894, Nov. 2020, doi: 10.1007/s10734-020-00520-7.

[5] D. Uliyan, A. S. Aljaloud, A. Alkhalil, H. S. Al Amer, M. A. E. A. Mohamed, and A. F. M. Alogali, "Deep Learning Model to Predict Students Retention Using BLSTM and CRF," *IEEE Access*, vol. 9, pp. 135550–135558, 2021, doi: 10.1109/ACCESS.2021.3117117.

[6] C. Dewberry and D. J. R. Jackson, "An application of the theory of planned behavior to student retention," *J. Vocat. Behav.*, vol. 107, pp. 100–110, Aug. 2018, doi: 10.1016/j.jvb.2018.03.005.

[7] M. Baranyi, M. Nagy, and R. Molontay, "Interpretable Deep Learning for University Dropout Prediction," in *SIGITE 2020 - Proceedings of the 21st Annual Conference on Information Technology Education*, Oct. 2020, pp. 13–19. doi: 10.1145/3368308.3415382.

[8] K. Muenks *et al.*, "Does my professor think my ability can change? Students' perceptions of their STEM professors' mindset beliefs predict their psychological vulnerability,

engagement, and performance in class.," *J. Exp. Psychol. Gen.*, vol. 149, no. 11, pp. 2119–2144, Nov. 2020, doi: 10.1037/xge0000763.

[9] O. Kryshko, J. Fleischer, J. Waldeyer, J. Wirth, and D. Leutner, "Do motivational regulation strategies contribute to university students' academic success?," *Learn. Individ. Differ.*, vol. 82, p. 101912, Aug. 2020, doi: 10.1016/j.lindif.2020.101912.

[10] A. J. Fernandez-Garcia, J. C. Preciado, F. Melchor, R. Rodriguez-Echeverria, J. M. Conejero, and F. Sanchez-Figueroa, "A real-life machine learning experience for predicting university dropout at different stages using academic data," *IEEE Access*, vol. 9, pp. 133076–133090, 2021, doi: 10.1109/ACCESS.2021.3115851.

[11] J. M. Ryan, T. Potier, A. Sherwin, and E. Cassidy, "Identifying factors that predict attrition among first year physiotherapy students: a retrospective analysis," *Physiother. (United Kingdom)*, vol. 110, pp. 26–33, Mar. 2021, doi: 10.1016/j.physio.2017.04.001.

[12] C. Bargmann, L. Thiele, and S. Kauffeld, "Motivation matters: predicting students' career decidedness and intention to drop out after the first year in higher education," *High. Educ.*, vol. 83, no. 4, pp. 845–861, Apr. 2022, doi: 10.1007/s10734-021-00707-6.

[13] L. M. Jeno, J. Nylehn, T. N. Hole, A. Raaheim, G. Velle, and V. Vandvik, "Motivational Determinants of Students' Academic Functioning: The Role of Autonomy-support, Autonomous Motivation, and Perceived Competence," *Scand. J. Educ. Res.*, vol. 67, no. 2, pp. 194–211, Feb. 2021, doi: 10.1080/00313831.2021.1990125.

[14] S. L. Petersen, E. S. Erenrich, D. L. Levine, J. Vigoreaux, and K. Gile, "Multi-institutional study of GRE scores as predictors of STEM PhD degree completion: GRE gets a low mark," *PLoS One*, vol. 13, no. 10, p. e0206570, Oct. 2018, doi: 10.1371/journal.pone.0206570.

[15] D. K. Benden and F. Lauermann, "Students' Motivational Trajectories and Academic Success in Math-Intensive Study Programs: Why Short-Term Motivational Assessments Matter," *J. Educ. Psychol.*, vol. 114, no. 5, pp. 1062–1085, Jul. 2022, doi: 10.1037/edu0000708.

[16] Y. Wen, Y. Tian, B. Wen, Q. Zhou, G. Cai, and S. Liu, "Consideration of the local correlation of learning behaviors to predict dropouts from MOOCs," *Tsinghua Sci. Technol.*, vol. 25, no. 3, pp. 336–347, Jun. 2020, doi: 10.26599/TST.2019.9010013.

[17] A. Ortigosa, R. M. Carro, J. Bravo-Agapito, D. Lizcano, J. J. Alcolea, and Ó. Blanco, "From Lab to Production: Lessons Learnt and Real-Life Challenges of an Early Student-Dropout Prevention System," *IEEE Trans. Learn. Technol.*, vol. 12, no. 2, pp. 264–277, Apr. 2019, doi: 10.1109/TLT.2019.2911608.

[18] M. Jawthari and V. Stoffa, "Predicting At-Risk Students Using Weekly Activities and Assessments," *Int. J. Emerg. Technol. Learn.*, vol. 17, no. 19, pp. 59–73, Oct. 2022, doi: 10.3991/ijet.v17i19.31349.

[19] E. J. Lizarte Simón and J. Gijón Puerta, "Prediction of early dropout in higher education using the SCPQ," *Cogent Psychol.*, vol. 9, no. 1, p. 2123588, Dec. 2022, doi: 10.1080/23311908.2022.2123588.

[20] M. Morelli, A. Chirumbolo, R. Baiocco, and E. Cattelino, "Academic failure: Individual, organizational, and social factors," *Psicol. Educ.*, vol. 27, no. 2, pp. 167–175, May 2021, doi: 10.5093/psed2021a8.

[21] P. De-La-Cruz, R. Rojas-Coaquira, H. Vega-Huerta, J. Pérez-Quintanilla, and M. Lagos-Barzola, "A Systematic Review Regarding the Prediction of Academic Performance," *J. Comput. Sci.*, vol. 18, no. 12, pp. 1219–1231, Dec. 2022, doi: 10.3844/jcssp.2022.1219.1231.

[22] S. Alturki, I. Hulpuș, and H. Stuckenschmidt, "Predicting Academic Outcomes: A Survey from 2007 Till 2018," *Technol. Knowl. Learn.*, vol. 27, no. 1, pp. 275–307, Mar. 2022, doi: 10.1007/s10758-020-09476-0.

[23] M. Adnan *et al.*, "Predicting at-Risk Students at Different Percentages of Course Length for Early Intervention Using Machine Learning Models," *IEEE Access*, vol. 9, pp. 7519–7539, 2021, doi: 10.1109/ACCESS.2021.3049446.

[24] S. C. Tsai, C. H. Chen, Y. T. Shiao, J. S. Ciou, and T. N. Wu, "Precision education with statistical learning and deep learning: a case study in Taiwan," *Int. J. Educ. Technol. High. Educ.*, vol. 17, no. 1, p. 12, Dec. 2020, doi: 10.1186/s41239-020-00186-2.

[25] E. J. M. Bakker *et al.*, "Physical and mental determinants of dropout and retention among nursing students: Protocol of the SPRiNG cohort study," *BMC Nurs.*, vol. 17, no. 1, p. 27, Dec. 2018, doi: 10.1186/s12912-018-0296-9.

[26] L. Bardach, M. Lüftenegger, S. Oczlon, C. Spiel, and B. Schober, "Context-related problems and university students' dropout intentions—the buffering effect of personal best goals," *Eur. J. Psychol. Educ.*, vol. 35, no. 2, pp. 477–493, Jun. 2020, doi: 10.1007/s10212-019-00433-9.

[27] F. Jiménez, A. Paoletti, G. Sánchez, and G. Sciavicco, "Predicting the Risk of Academic Dropout with Temporal Multi-Objective Optimization," *IEEE Trans. Learn. Technol.*, vol. 12, no. 2, pp. 225–236, Apr. 2019, doi: 10.1109/TLT.2019.2911070.

[28] A. Kelava, P. Kilian, J. Glaesser, S. Merk, and H. Brandt, "Forecasting Intra-individual Changes of Affective States Taking into Account Inter-individual Differences Using Intensive Longitudinal Data from a University Student Dropout Study in Math," *Psychometrika*, vol. 87, no. 2, pp. 533–558, Jun. 2022, doi: 10.1007/s11336-022-09858-6.

[29] R. Z. Pek, S. T. Ozyer, T. Elhage, T. Ozyer, and R. Alhajj, "The Role of Machine Learning in Identifying Students At-Risk and Minimizing Failure," *IEEE Access*, vol. 11, pp. 1224–1243, 2022, doi: 10.1109/ACCESS.2022.3232984.

[30] K. Oqaidi, S. Aouhassi, and K. Mansouri, "Towards a Students' Dropout Prediction Model in Higher Education Institutions Using Machine Learning Algorithms," *Int. J. Emerg. Technol. Learn.*, vol. 17, no. 18, pp. 103–117, Sep. 2022, doi: 10.3991/ijet.v17i18.25567.

[31] A. J. Fernández-García, R. Rodríguez-Echeverría, J. C. Preciado, J. M. Conejero Manzano, and F. Sánchez-Figueroa, "Creating a recommender system to support higher education students in the subject enrollment decision," *IEEE Access*, vol. 8, pp. 189069–189088, 2020, doi: 10.1109/ACCESS.2020.3031572.

[32] K. Barbara and C. Stuart, "Guidelines for performing Systematic Literature Reviews in Software Engineering version 2.3," 2007. doi: 10.1541/ieejias.126.589.

[33] L. Respondek, T. Seufert, J. M. Hamm, and U. E. Nett, "Linking changes in perceived academic control to university dropout and university grades: A longitudinal approach," *J. Educ. Psychol.*, vol. 112, no. 5, pp. 987–1002, Jul. 2020, doi: 10.1037/edu0000388.

[34] Z. Xie, "Modelling the dropout patterns of MOOC learners," *Tsinghua Sci. Technol.*, vol. 25, no. 3, pp. 313–324, Jun. 2020, doi: 10.26599/TST.2019.9010011.

[35] Y. Zheng, Z. Gao, Y. Wang, and Q. Fu, "MOOC Dropout Prediction Using FWTS-CNN Model Based on Fused Feature Weighting and Time Series," *IEEE Access*, vol. 8, pp. 225324–225335, 2020, doi: 10.1109/ACCESS.2020.3045157.

[36] C. Imhof, I. S. Comsa, M. Hlosta, B. Parsaeifard, I. Moser, and P. Bergamin, "Prediction of Dilatory Behavior in eLearning: A Comparison of Multiple Machine Learning Models," *IEEE Trans. Learn. Technol.*, pp. 1–15, 2022, doi: 10.1109/TLT.2022.3221495.

[37] J. Marôco *et al.*, "Predictors of academic efficacy and dropout intention in university students: Can engagement suppress burnout?," *PLoS One*, vol. 15, no. 10 October, p. e0239816, Oct. 2020, doi: 10.1371/journal.pone.0239816.

[38] S. Langness, N. Rajapuram, M. Marshall, A. S. Rahman, and A. Sammann, "Risk factors associated with student distress in medical school: Associations with faculty support and availability of wellbeing resources," *PLoS One*, vol. 17, no. 4 April, p. e0265869, Apr. 2022, doi: 10.1371/journal.pone.0265869.

[39] N. Gillet, A. J. S. Morin, T. Huyghebaert-Zouagh, E. Alibran, S. Barrault, and C. Vanhove-Meriaux, "Students' Need Satisfaction Profiles: Similarity and Change over the Course of a University Semester," *Appl. Psychol.*, vol. 69, no. 4, pp. 1396–1437, Sep. 2020, doi: 10.1111/apps.12227.

[40] J. M. Dopmeijer *et al.*, "The role of performance pressure, loneliness and sense of belonging in predicting burnout symptoms in students in higher education," *PLoS One*, vol. 17, no. 12 December, p. e0267175, Dec. 2022, doi: 10.1371/journal.pone.0267175.

[41] M. L. Farnese, P. Spagnoli, and S. Livi, "Undergraduates' academic socialization. A cross-time analysis," *Br. J. Educ. Psychol.*, vol. 92, no. 4, pp. 1239–1255, Dec. 2022, doi: 10.1111/bjep.12497.

[42] F. Belloc, A. Maruotti, and L. Petrella, "University drop-out: An Italian experience," *High. Educ.*, vol. 60, no. 2, pp. 127–138, 2010, doi: 10.1007/s10734-009-9290-1.

[43] N. V. Ivanova and M. A. Vinogradova, "Plot-Didactic Game System Design in the Formation of a Digital Footprint in Preschool Children Career Guidance Based on Gamification," in *International Scientific Conference "Digitalization of Education: History, Trends and Prospects"(DETP 2020)*, 2020, pp. 850–858. doi: 10.2991/assehr.k.200509.151.

[44] A. Behr, M. Giese, H. D. Teguim Kamdjou, and K. Theune, "Dropping out of university: a literature review," *Rev. Educ.*, vol. 8, no. 2, pp. 614–652, 2020.

[45] A. Esteban, A. Zafra, and C. Romero, "Helping university students to choose elective courses by using a hybrid multi-criteria recommendation system with genetic optimization," *Knowledge-Based Syst.*, vol. 194, p. 105385, 2020, doi: 10.1016/j.knosys.2019.105385.

[46] E. G. Galimova, A. V Konysheva, O. A. Kalugina, and Z. M. Sizova, "Digital educational footprint as a way to evaluate the results of students' learning and cognitive activity in the process of teaching mathematics," *Eurasia J. Math. Sci. Technol. Educ.*, vol. 15, no. 8, 2019, doi: 10.29333/ejmste/108435.

[47] G. Conole, *Designing for learning in an open world*, vol. 4. Springer Science & Business Media, 2013. doi: 10.1007/978-1-4419-8517-0.