INTERNATIONAL JOURNAL OF APPLIED RESEARCH AND TECHNOLOGY
ISSN 2519-5115
IJART- Vol-8, Issue-6, December, 2023
Available Online at http://www.ijart.info/
DOI: https://doi.org./10.24163/ijart/2017/8(6):189-206

Machine Learning Applications for Recommender Systems in Higher Education: A Systematic Review

[1]Benard Ondiek
[2]Lucy Waruguru
[3]Stephen Njenga

[1]School of Technology, KCA University, Nairobi, Kenya

[2]School of Technology, KCA University, Nairobi, Kenya

[3]Department of Computer Science, Murang'a University of Technology, Muranga, Kenya

Benard Ondiek, School of Technology, KCA University, Nairobi, Kenya

Corresponding author:
benmacondiek@gmail.com

ABSTRACT

Envisioning an information system that can mitigate the wrong choice of courses for students in higher education is imperative because of the large pool of courses offered by numerous institutions of higher learning in developing countries. Thus, the main goal of this systematic literature review (SLR) was to explore the application of machine learning algorithms in modeling higher education enrolment and to understand the datasets, evaluation metrics, and validation techniques applied in the recommendation systems.

The current research reviews previous studies that have applied different machine learning algorithms to build recommender systems for use in higher education. The Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) method was adopted to select papers for review. The study reviewed 44 papers published between 2013 and 2023 alongside various criteria, such as applied algorithms, evaluation metrics, and validation techniques. Among the key findings was the realization that Naïve Bayes has been the most widely applied for building recommender systems in higher education. The results of this review pinpoint existing research gaps and provide recommendations for future research. This study can serve as a guide for future researchers on the trends of machine learning recommender systems in higher education. Practitioners, researchers, and policymakers could also benefit by understanding state-of-the-art algorithms, evaluations, and validation techniques while developing recommendation systems for higher education.

Received: November 01, 2023
Revised: December 21, 2023
Published: December 31, 2023

Keywords- Machine Learning, Recommender Systems, Higher Education, Systematic Literature Review.

1. INTRODUCTION

The university education market has continued to shrink due to various challenges, exposing students to a large pool of courses. It has become practically impossible to align students' capabilities with the courses they choose in Higher Education (Elahi, Starke, El Ioini, Lambrix, & Trattner, 2022). As presented by SDG 4, there is a push to have students select courses that would lead to more innovations (Kroll, Warchold, & Pradhan, 2019). It is important to have a system that will assist students by closely matching their interests with the vast amount of information online(Roy & Dutta, 2022).

Recommender systems apply artificial intelligence techniques for prediction (Assami, Daoudi, & Ajhoun, 2022; Girase, Powar, & Mukhopadhyay, 2017). These techniques range from machine learning to data mining (Singh, Kapoor, & Sohi, 2021). The main aim of recommender systems is to reduce information overload while personalizing suggestions to improve decision-making (Ferreira, Silva, Abelha, & Machado, 2020; Oliveira, Bernardini, & Viterbo, 2021). According to (Deng, 2019) several large companies including Netflix, LinkedIn, and Pandora, have attributed the increase in revenue to the use of recommender systems. A recommender system is an intelligent computer-based technique that predicts user adoption and usage. This allows the client to buy commodities from a vast range of online commodities (Burke, Felfernig, & Göker, 2011). Examples include Facebook's friend recommendations, YouTube video recommendations, and trip advisors (Burke et al., 2011).

There has been an increase in advances, particularly in the development of recommender systems for enrolment within the education domain. The techniques employed include the Markov chain model (Ezugwu & Ologun, 2017; Polyzou, Nikolakopoulos, & Karypis, 2019). Regression (Kumari & Yadav, 2018) and analytical hierarchy process (Liang, Ren, Gao, Dong, & Gao, 2017; Sangka & Muchsini, 2018).

Currently, researchers have engaged in new approaches to recommending course enrolment in higher education. These approaches include data mining (Hasan, Ahmed, Abdullah, & Rahman, 2016), Machine Learning (Angra & Ahuja, 2017; Uddin, Imran, Muhammad, Fayyaz, & Sajjad, 2021), and Deep Learning (Najafabadi et al., 2015; Vargas, Mosavi, & Ruiz, 2017). Recommendation systems can be classified into three types: content-based collaborative and hybrid (Gulzar, Leema, & Deepak, 2018; Guo et al., 2019; Portugal, Alencar, & Cowan, 2018; Sattar, Ghazanfar, & Iqbal, 2017).

Machine learning strives to generate algorithms based on "data trends and historical relationships between data" (Ismail & Yusof, 2022; Janiesch, Zschech, & Heinrich, 2021). According to (Barramuño, Meza-Narváez, & Gálvez-García, 2022) machine learning is the process through which a computer is given a task to perform. The computer imitates the task and continues to improve by learning from experience (Nichols, Herbert Chan, & Baker, 2019; Waheed et al., 2023) According to (Ray, 2019) machine learning is the process through which a computer is given a task to perform. The computer imitates the task and continues to improve through learning from experience. Machine learning can be classified as supervised or unsupervised (Kumari & Yadav, 2018). According to (Machado & Karray, 2022), supervised learning utilizes "labeled datasets", which conduct training for algorithms to classify data or predict outcomes accurately (Jiang, Gradus, & Rosellini, 2020; Miric, Jia, & Huang, 2023; Sridhar, Mootha, & Kolagati, 2020). According to (Goga, Kuyoro, & Goga, 2015) unsupervised learning involves using algorithms to" analyze and cluster unlabelled datasets" (Wang et al., 2022). According to (Choi, Coyner, Kalpathy-Cramer, Chiang, & Campbell, 2020) supervised learning dataset is divided into validation, training, and testing datasets. The moment the input data enter the model, the "weights are adjusted until the model has been fitted appropriately." The result of the "cross validation process" reduces "overfitting and under fitting" overfitting and under-fitting.

This study critically reviews studies that have applied machine learning to build recommender systems within the education sector. This study identifies popularly adopted machine learning algorithms, evaluation metrics, validation, and datasets. The primary objectives of this study are summarized as follows.

1). Identify trends in higher education's use or study of machine learning techniques for recommender systems.
2). Identify the dataset used while conducting research using machine learning algorithms.
3). Identify the evaluation metrics and validation techniques applied in higher education machine learning recommendation systems research.

2. RELATED STUDIES

Since reviews have not been conducted on Machine Learning applications in recommender systems in higher education, the study has identified studies within machine learning.

Aher and Lobo (2013) developed a model that combined different machine learning algorithms to recommend e-learning courses based on historical data. Clustering techniques that combined association rules and K-means clustering were implemented to generate the results.

Goga, Kuyoro, and Goga (2015) developed a recommender system based on machine learning to improve student performance. The system mainly relies on decision trees, association rules, and neural networks to predict the student's performance and recommend mitigations to the student to improve performance. A recommender system for university admission was developed. The model leveraged the strengths of multicriteria collaborative filtering (MC-CF) and Dimensionality Reduction techniques. The model was able to reduce computational costs while increasing the accuracy and efficiency.

Hasan, Ahmed, Abdullah, and Rahman (2016) developed a machine learning recommender system to help graduate prospective students obtain admissions to universities that match their profiles. A Massive open online course (MOOC) recommender system was developed. The system provides ease of accessibility to e-learning sites using machine learning algorithms.

Thangavel, Bkaratki, and Sankar (2017) developed a student-placement analyzer recommender system based on machine learning. The data were generated from both operational and historical data. The data were mainly used for training the model for rule identification and testing the classification. The model implemented several algorithms including a logistic regression classifier, naïve Bayes, and meta-bagging classifiers. Finally, the accuracy of the model was evaluated. (Sattar et al., 2017) built a hybrid recommender system to address the sparsity and cold start issues faced by previous recommender systems.

Baskota and Ng (2018) developed a graduate school recommender system that suggested the most appealing programs to students. (Obeid, Lahoud, El Khoury, & Champin, 2018) developed a recommendation model for selecting university courses. The model was based on the student's "vocational weaknesses, interests, and capabilities." (Zhou, Huang, Hu, Zhu, & Tang, 2018) improved the recommendation e-learning model using a machine learning model. The recommender system is based on long short-term memory (LSTM) clustering. The result was a more accurate and efficient recommendation of the learner's path.

Ezaldeen, Misra, Alatrash, and Priyadarshini (2019) developed a recommender system that suggested courses for students to take on the e-learning system. The system implements Naïve Bayes and the K-means algorithms. The selection and suggestions are based on the need for the course and the student's aptitude.

In 2020, (Yanes, Mostafa, Ezz, & Almuayqil, 2020) developed a recommender system based on machine learning to improve students' experiences at universities. The model incorporated educational, program, and course outcomes to predict appropriate recommendations for the faculty. Feedback was mapped to course specifications, academic records, and assessments.

In 2022, (Elahi et al., 2022) developed a university recommender system that could evaluate the recommender system's main role in capturing students' preferences. The preferences are then used to build predictions based on the recommendations of university rankings. Wang et al. (2022) developed a machine-learning recommender system to recommend courses to students in their upcoming semesters. The model is based on a hybrid recommender system using matrix factorization as the foundational algorithm.

Table 1 presents a summary of systematic literature reviews that have been conducted previously on recommender systems built using machine learning.

INTERNATIONAL JOURNAL OF APPLIED RESEARCH AND TECHNOLOGY
ISSN 2519-5115
IJART- Vol-8, Issue-6, December, 2023
Available Online at http://www.ijart.info/
DOI: https://doi.org./10.24163/ijart/2017/8(6):189-206

Table 1. Contribution of previous review studies

Ref	Title	Year	Type	Contribution
(Urdaneta-Ponte, Mendez-Zorrilla, & Oleagordia-Ruiz, 2021)	Recommendation Systems for Education: Systematic Review	2021	Systematic Review 2015-2020 (98 Articles)	Recommending educational resources Identified that there are gaps in the application of intelligent systems in resource recommendation.
(da Silva, Slodkowski, da Silva, & Cazella, 2023)	A systematic literature review on educational recommender systems for teaching and learning: research trends, limitations and opportunities	2023	Systematic Review 2015-2020 (16 Articles)	Identify the trends in developing recommender systems, evaluation techniques, research opportunities, and gaps.
(Deschênes, 2020)	Recommender systems to support learners' Agency in a Learning Context: a systematic Review	2020	Systematic Review 2008-2018 (56 Articles)	Built a system to define learners' goals in education.
(Sandoussi, Hnida, Daoudi, & Ajhoun, 2022)	Systematic Literature Review on Open Educational Resources Recommender Systems	2022	Systematic Review 2005-2013 (413 Articles)	Built a recommender system to recommend Open educational resources within the education sector. Adopted the deductiove approach using grounded theory. Applied both quantitative and qualitative analysis.
(Maphosa & Maphosa, 2023)	Fifteen Years of Recommender Systems Research in Higher Education: Current Trends and Future Direction	2023	Systematic Review 2007-20121 (272 Articles)	Applied bibliometric analysis to investigate the adoption of recommender systems within higher education. The Scopus database was selected as the source of data.
(Uddin et al., 2021)	A Systematic Mapping Review on MOOC Recommender Systems	2021	2013-2021 (116 Articles)	The paper reviewed several journals with then aim of finding solutions to online courses recommendations.

3. METHODOLOGY

This study aimed to understand the trends in research on recommender models in higher education based on machine learning. One of the complexities of these studies is the spread of papers in various journal databases. The literature is based on popular journal databases to achieve this objective.

3.1 Review Objectives and Research Questions

The objectives of this study were to (i) identify trends in higher education's use or study of machine learning techniques for recommender systems, (ii) identify the dataset used while conducting research using machine learning algorithms, and (iii) identify the evaluation metrics and validation techniques applied in machine

INTERNATIONAL JOURNAL OF APPLIED RESEARCH AND TECHNOLOGY
ISSN 2519-5115
IJART- Vol-8, Issue-6, December, 2023
Available Online at http://www.ijart.info/
DOI: https://doi.org./10.24163/ijart/2017/8(6):189-206

learning recommendation systems research in higher education. The following research question helps to achieve the set objectives.

RQ1: What are the most popular machine learning algorithms used in higher education recommender models?

RQ2: What are the most popular evaluation metrics used in higher-education recommender models?

RQ3: What are the most popular validation techniques used in higher-education recommender models?

RQ4: What are the most common dataset sources that help in the development of machine learning recommender models in higher education?

3.2 Search Strategies

The Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement guided the systematic literature review process, wherein several databases were in the spotlight (Page et al., 2021). The ACM digital library, ScienceDirect, and IEEE databases were used to search for the period from 2013 to 2023. Three terms were used to narrow the search of the specific literature which included “Recommender,” “Machine Learning,” and “Higher Education.” The three words were crafted to produce a string, which was then used for the search.

((((machine learning[Title/Abstract]) OR ("machine learning"[Title/Abstract])) OR (ML[Title/Abstract])) AND (((recommender systems[Title/Abstract]) OR ("recommender system"[Title/Abstract])) OR (recommendation[Title/Abstract]))) AND(((((((((higher education[Title/Abstract]) OR (university[Title/Abstract])) OR (college[Title/Abstract])) OR (graduate[Title/Abstract])) OR (undergraduate[Title/Abstract])) OR (postgraduate[Title/Abstract])) Selection Criteria

The following rules were applied to select the journals to be reviewed, as suggested by (Song & Wang, 2020)

Inclusion Criteria

1. The research papers should have been published in scientific peer-reviewed journals.
2. Studies should have implemented machine learning algorithms.
3. The choice of language is English.

Exclusion Criteria

1. The studies that have discussed recommender systems but not applied machine learning.
2. Studies that have both recommender systems and machine learning but discuss other domains.

The initial process involved removing duplicates from the selection process. This is achieved by conducting a screening process on the titles and abstracts so that we can obtain the most relevant literature. The above criteria should also be satisfied by individual authors. Figure 1 below presents a flowchart of the selection process.

INTERNATIONAL JOURNAL OF APPLIED RESEARCH AND TECHNOLOGY
ISSN 2519-5115
IJART- Vol-8, Issue-6, December, 2023
Available Online at http://www.ijart.info/
DOI: https://doi.org./10.24163/ijart/2017/8(6):189-206

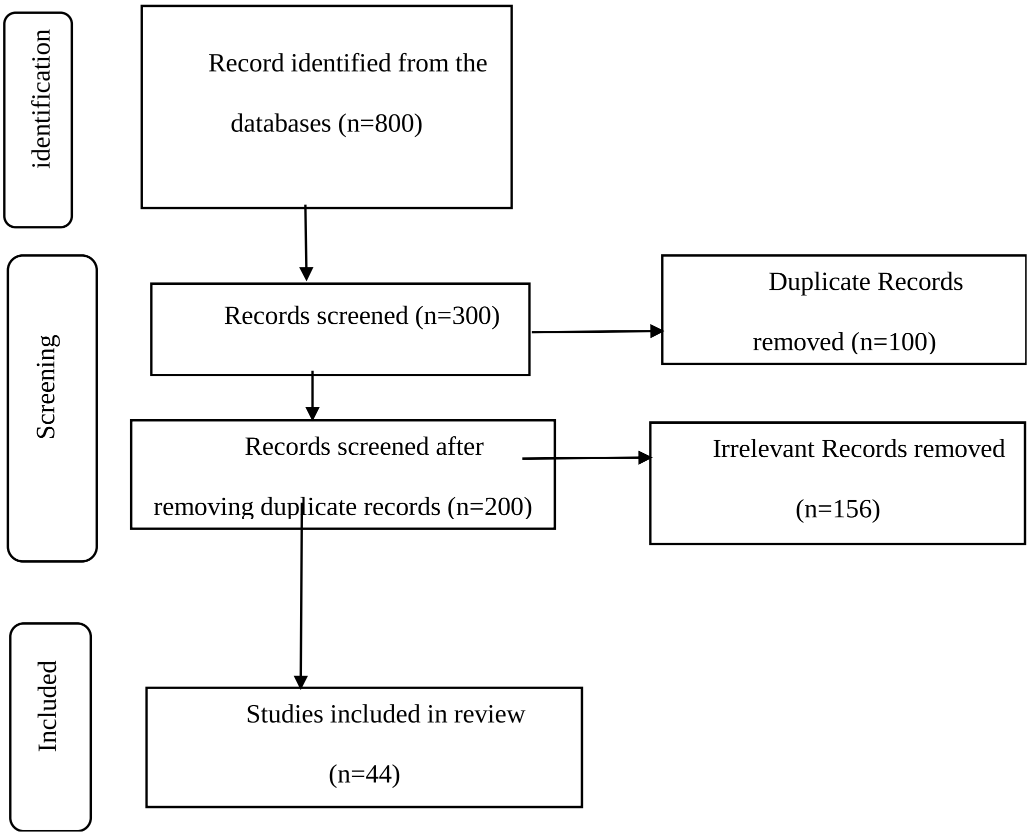

Fig.1. Selection procedure and data selection

Data extracted based on the below objectives

1. Topic of the study
2. Area of recommendation
3. Validation and evaluation of the studies
4. Dataset used
5. Machine learning models implemented

The selection process adopted both forward and backward searching techniques, resulting in the selection of 44 articles that represented the study domain. This exercise aims to ensure the validity and reliability of the process.

3.3 Chronological Review

This review is based on several publications conducted between 2013 and 2023 on the theme of the recommender model in education based on machine learning. The articles were categorized according to their types, as shown in Table 1.

INTERNATIONAL JOURNAL OF APPLIED RESEARCH AND TECHNOLOGY
ISSN 2519-5115
IJART- Vol-8, Issue-6, December, 2023
Available Online at http://www.ijart.info/
DOI: https://doi.org./10.24163/ijart/2017/8(6):189-206

Table 2. Chronological Review

Document Types	**Count**
Conference Paper	15
Article	29

The articles were classified according to the years of publication, as shown below. The number of publications increased from 2013 to 2019 and then decreased from 2020. That was the period when the world was hit by the Covid 19 pandemic hence, it could be a factor that led to the low number of publications. This publication demonstrates a steady increase from 2021, which could be attributed to a return to normalcy in research.

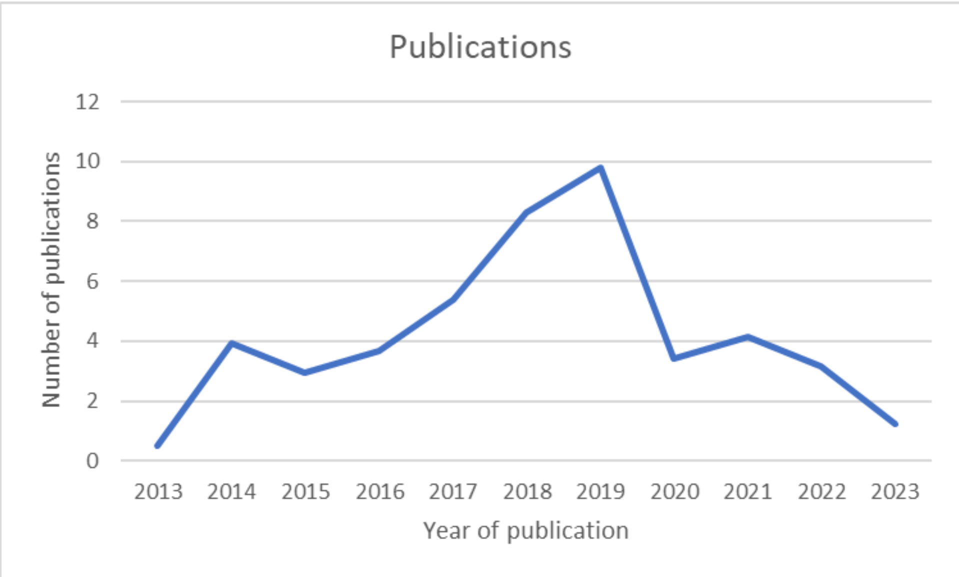

Figure 2. Chronology of Publications

4. RESULTS AND DISCUSSIONS

The guiding factors in the research were the machine learning algorithms, evaluation metrics, and datasets used in the literature review. The authors read the abstracts, introduction, approach used in the study, conclusion, and future work for the selected papers. A spreadsheet was created to score various elements, as described above. This presentation was then made possible by generating graphs.

4.1 Model Selection

This study reviewed several machine-learning algorithms used in building recommender systems for higher education. Below is a brief description of some of the reviewed machine-learning algorithms.

Srivastava and Bhambhu (2010) defined an (SVM) as a set of related supervised learning methods implemented in both classification and regression, and SVMs have been more instrumental in text classification based on their strength of "linear and nonlinear classification" (Gholami & Fakhari, 2017). Despite their low speed in terms of training time, they attain a high level of accuracy compared to other algorithms (Gaye, Zhang, & Wulamu, 2021; Pisner & Schnyer, 2020). According to (Ezaldeen et al., 2019) SVM has a high accuracy level compared to other classification algorithms.

Artificial Neural network (ANN) leverage their ability to self-learn and generate efficient results (Dastres & Soori, 2021; Jimoh, Abisoye, & Uthman, 2021; Jwo, Biswal, & Mir, 2023). It is independent of the data types, thereby being able to learn patterns independently (Farizawani, Puteh, Marina, & Rivaie, 2020). According to

INTERNATIONAL JOURNAL OF APPLIED RESEARCH AND TECHNOLOGY
ISSN 2519-5115
IJART- Vol-8, Issue-6, December, 2023
Available Online at http://www.ijart.info/
DOI: https://doi.org./10.24163/ijart/2017/8(6):189-206

(Hernández, Musso, Kyndt, Eduardo Cascallar, & Carlos Felipe, 2021) an ANN system processes information based on units referred to as neurons. The strength of ANNs is their ability to perform well in terms of metrics (Hernández et al., 2021). This was compared with other classification algorithms (Han, Kelley, & Knowles, 2021).

Naïve Bayes is widely accepted owing to its simplicity and speed of performing tasks (Alkubaisi, Kamaruddin, & Husni, 2018; Gan, Shao, Chen, Yu, & Jiang, 2021). Naive Bayes operates based on the theory of probability. The guiding principle is a "subjective probability of certain unknown states with insufficient information" (Yang, 2018). Figure 3 shows an analysis of the investigated papers.

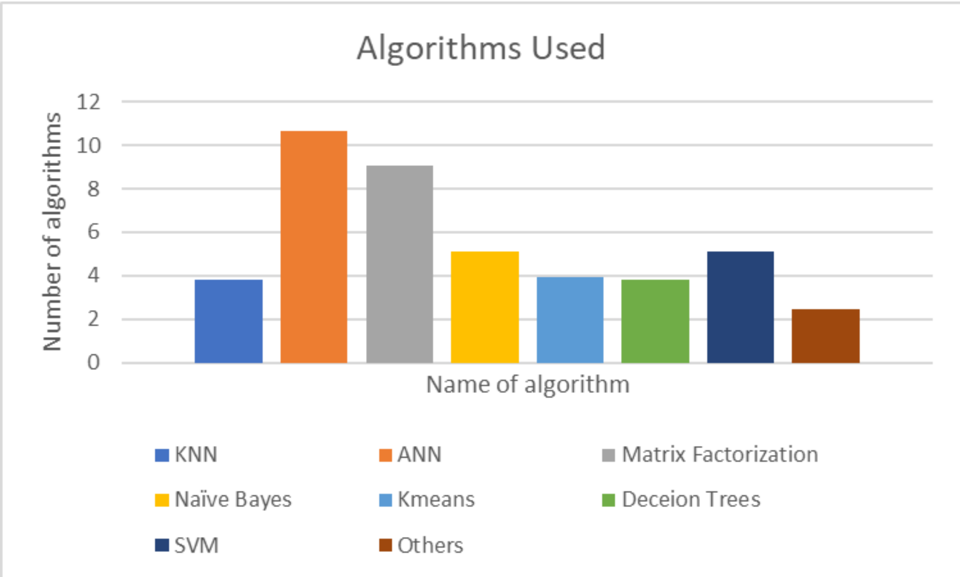

Fig. 3. Machine Learning Algorithms

According to the sampled papers, Naïve Bayes seemed to have a higher score than the others. The score was 24%. This can be attributed to the simplicity of the calculations. Decision trees also scored highly based on their simplicity in performing calculations. The Artificial Neural Network, despite being a great area of interest, scored 12%. This could be a result of the minimal knowledge of the domain. From the analysis above, much effort might be required to understand the individual model that would be sufficient on its own to produce robust results. Therefore, several studies have used the approach of combining several models to produce significant results.

4.2 Evaluation Metrics and Validation

The difference in the models produced necessitates different approaches for evaluation and validation (Schneider & Xhafa, 2022; Sekeroglu, Abiyev, Ilhan, Arslan, & Idoko, 2021). This was underpinned by the similarity in the strategy of dividing the data for training, which included data training and testing. Some evaluation metrics include the following.

Recall Measure

The Recall measure is a metric commonly used in machine learning within the education domain. Recall is used to measure the abilities of the model to detect specific output classes ((Fränti & Mariescu-Istodor, 2023; Hicks et al., 2022).

INTERNATIONAL JOURNAL OF APPLIED RESEARCH AND TECHNOLOGY
ISSN 2519-5115
IJART- Vol-8, Issue-6, December, 2023
Available Online at http://www.ijart.info/
DOI: https://doi.org./10.24163/ijart/2017/8(6):189-206

Precision Measure

Precision is a measure of the ratio of positives compared to all predictions (AlZoman & Alenazi, 2021). It can also be described as “An empirical assessment of the conditional likelihood of a classification that is accurate given the anticipated class 1’ (Hand, Christen, & Kirielle, 2021).

F1-Measure

Hand, Christen, and Kirielle (2021) argues that the F1 measure has gained popularity for measuring classification algorithms. The other term associated with the F1 measure is the F1-score. The F1 score measures average precision and recall (AlZoman & Alenazi, 2021). It can also be defined as the harmonic mean between the Precision and Recall (Sekeroglu et al., 2021).

Mean absolute error

Şekeroğlu, Abiyev, İlhan, Arslan, and Idoko (2021) described the Mean Absolute Error (MAE) as a regression-based evaluation technique, whose main strategy is to measure the error between the predicted and observed data. The changes in MAE are “linear, making it more intuitive’ (Schneider & Xhafa, 2022). MAE is the magnitude of error. The lower the MAE, the more accurate the model.

Accuracy

According to (Sekeroglu et al., 2021) accuracy is obtained by dividing the total number of samples in the test set by the total number of samples that were correctly identified.
Figure 4 below shows the model evaluation techniques from the reviewed papers.

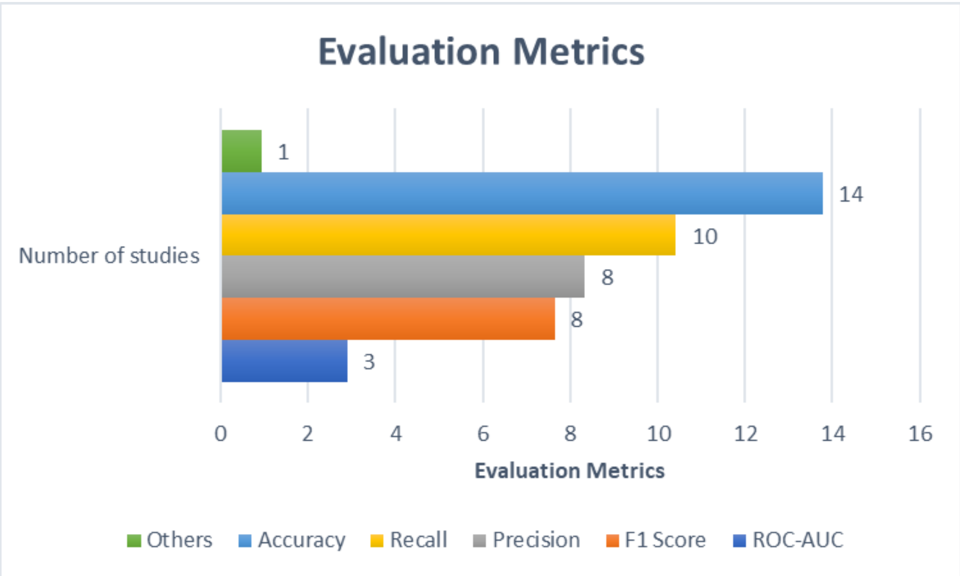

Fig.4. Model Evaluation

Accuracy was the most preferred metric for the evaluation of machine learning recommendation models in education, with a percentage score of 39%. This could be attributed to the ease of use of the formula. The least preferred was the RMSE metric, with a score of 4%. This can be attributed to the complexity of its use. Some of the reviewed papers have applied multiple evaluation techniques.

The validation technique also showed an increase in the use of the K-hold cross-validation technique. The percentage score for cross-validation was 32%. This may be due to the ease of access to the data. The data were split into training and testing datasets from the same population. Most researchers relied on one dataset that was used as both the test and training data. The hold-out method was the least preferred method (38%). Other researchers did not specify the validation technique they implemented while developing recommender models.

INTERNATIONAL JOURNAL OF APPLIED RESEARCH AND TECHNOLOGY
ISSN 2519-5115
IJART- Vol-8, Issue-6, December, 2023
Available Online at http://www.ijart.info/
DOI: https://doi.org./10.24163/ijart/2017/8(6):189-206

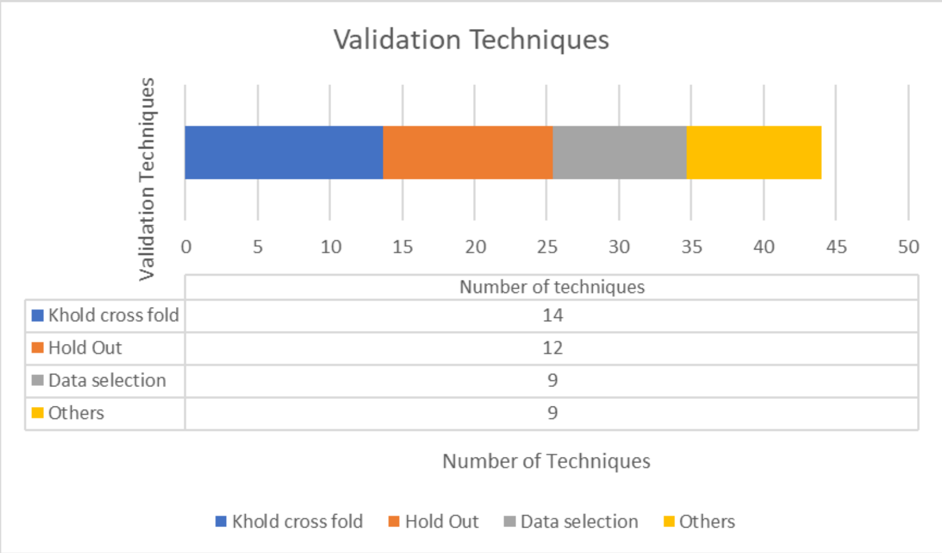

Fig.5. Model Validation

4.3 Dataset

There is a clear indication that several datasets have been used in the literature. According to the literature, most datasets are obtained from target universities or institutions of higher learning. University databases stood at 53% compared to the other data sources, which contributed to the remaining 47%. The use of datasets from university repositories can be attributed to the increased use of enterprise resource planners (ERP) to capture student data. The data are captured when the student is enrolled in the university, and tracking of the students is also performed within the ERP. According to (Ismail & Yusof, 2022) the use of open data is gaining popularity. This was motivated by the US government, the European Commission, and the Singaporean government, who developed legislation to have the open data public.

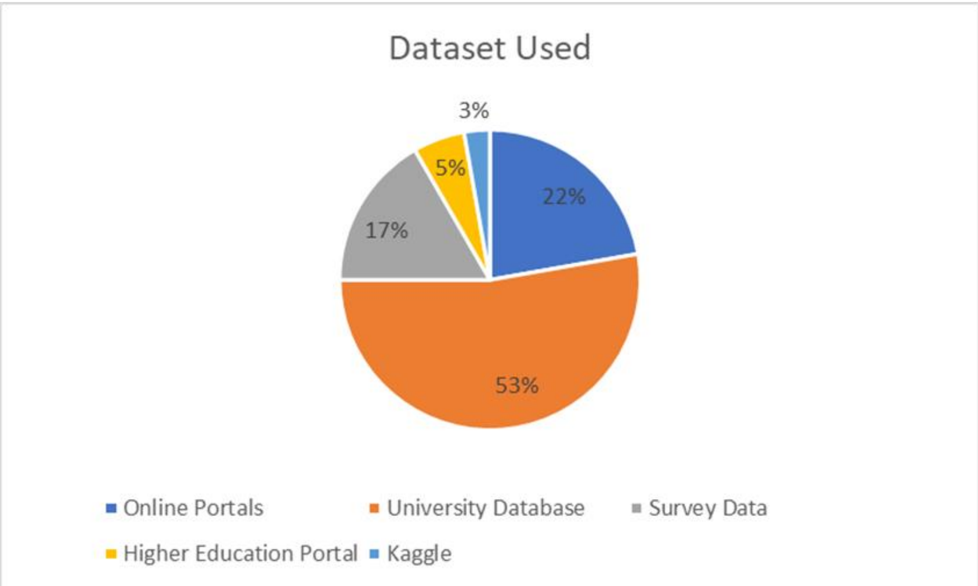

Fig. 6. Distribution of Data Sources.

INTERNATIONAL JOURNAL OF APPLIED RESEARCH AND TECHNOLOGY
ISSN 2519-5115
IJART- Vol-8, Issue-6, December, 2023
Available Online at http://www.ijart.info/
DOI: https://doi.org./10.24163/ijart/2017/8(6):189-206

4.4 Summary of key findings

Tab.2 Summary of key findings

Name of model	Author	Methodology/Technology	Key Findings	Usage statistics
Semantic Recommender system	(Obeid et al., 2018)	Semantic-based methods and machine learning techniques	The needs, interests, preferences, and capacities of students are determined via an ontology-based recommender system.	Accuracy 68%
Stacked Ensemble Learning	(Sridhar et al., 2020)	KNN, Random forest, decision tree,	The paper presented a system that predicts the eligibility of a student to be admitted to an institution. It leveraged the use of the GRE score	Decision tree- accuracy 65.5% Random forest- 62.5% KNN-57%
Attention Technique	(Yazdi, Chabok, & Kheirabadi, 2021)	LSTM, MLP, and BiLSTM with attention method	The paper presented a more robust model for selecting appropriate learning resources. The study leveraged user preferences and interests.	Accuracy 96%
Graduate school recommender system:	(Han et al., 2021)	KNN	The model was developed to propose to applicants potential Universities that offer to fund.	Accuracy 67%
E-Learning Course Recommender System	(Jena et al., 2022)	K-nearest neighbor (KNN), Singular Value Decomposition (SVD), and ANN	The model implements machine language techniques to recommend eLearning courses to participants based on their history and preferences.	Accuracy 70%
Grade prediction with models	(Polyzou et al., 2019)	Matrix factorization	The Model aimed at comparing two methods to find the one that generates the best grade prediction. The results showed that matrix factorization performed better than linear regression	Accuracy 62%
Improvement of student's learning experiences	(Yanes et al., 2020)	KNN	The model was developed to propose strategies for improving students' performance.	Accuracy 65%
Graduate school recommender system:	(Hasan et al., 2016)	KNN	The model was developed to propose to applicants potential Universities that offer funding.	Accuracy 67%
Graduate School Recommendation System Using the Multi-Class Support Vector	(Baskota & Ng, 2018)	Support Vector Machine and KNN Approaches	Applied the applicant's data and information from online portals to make recommendations for their enrollment.	Accuracy 58%

INTERNATIONAL JOURNAL OF APPLIED RESEARCH AND TECHNOLOGY
ISSN 2519-5115
IJART- Vol-8, Issue-6, December, 2023
Available Online at http://www.ijart.info/
DOI: https://doi.org./10.24163/ijart/2017/8(6):189-206

Machine and KNN Approaches				
Semantics aware intelligent system	(Ezaldeen et al., 2019)	Logistic Regression, Random Forest, SVM, and MLP	a novel method is put forth wherein ideas with graphs and additional contextual and semantic information are merged to infer the relative semantic linkages between terms and e-learning resources in order to construct the semantic matrix. With the use of this innovative method, learners' semantic datasets are created and used to classify the resources that are readily available and enhance suggestions.	Accuracy 84% F1 Score 73%

The findings of this study provide an overview of the use of machine learning recommender models within the context of higher education. The aim was to add to the previously conducted systematic literature reviews on recommender models, as presented in (Alyari & Navimipour, 2018; Portugal et al., 2018; Zhong, Xie, & Wang, 2019).

In addressing RQ1, the most popular machine learning algorithms for recommender models in higher education are Naïve Bayes and Decision trees. This is supported by (Cui, Chen, Shiri, & Fan, 2019; Portugal et al., 2018) who ranked the Naïve Bayes and Decision tree algorithms as the most popular models. Low complexity and ease of implementation were cited as ingredients for the rise in the use of the two algorithms. Artificial neural networks did not score well in usage because of the reduced knowledge gap in the algorithm. Despite their low popularity, Artificial Neural networks exhibit a greater performance (Maita, Martins, López Paz, Peres, & Fantinato, 2015).

To address RQ2, the author investigated the evaluation metrics applied in the development of machine learning recommender models in higher education. The popular evaluation metric used was accuracy, based on a literature review. Furthermore, the low uptake in the use of ANN resulted in a reduced usage of the F1 score and recall evaluation metrics. This agrees with (Hernández et al., 2021) who argues that ANN shows excellent performance with Recall and F1 scores. The RMSE was less popular in most of the reviewed studies.

In addressing RQ3, the validation technique also showed an increase in the use of the cross-validation technique. The percentage score for cross-validation was 32%. This can be attributed to the use of the same data for training and testing the model. Most researchers prefer this model because the data are collected once, and the test data are obtained from the dataset using a specific ratio.

Finally, RQ4 is addressed by presenting the most popular dataset. The most used dataset was the University dataset because of its ease of access. Most universities and institutions of higher education store admission and education data electronically; thus, it is easier to access and analyze the data. This can be attributed to the growing records of student data in higher institutions, as supported by (Christou et al., 2023). Publicly obtained datasets are not popular. This can be attributed to the data not having a local context(Sekeroglu et al., 2021).

5. CONCLUSIONS AND AREAS FOR FUTURE STUDIES

The goal of this research was to identify the trends in recommender systems built using machine learning algorithms within the higher education sector. Various practitioners, researchers, and policymakers will benefit from the results of this study. The developers of recommender systems are better placed to understand the most popular machine learning algorithms, evaluation techniques, validation methods, and sources of data. The systematic review included 44 publications published between 2013 and 2023, in line with the inclusion and exclusion criteria. All publications were reviewed by reading and drawing conclusions based on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement.

The trends showed a steady increase in Machine Learning recommendation systems within the education sector. The study identified Naïve Bayes and decision trees as the most popular algorithms used in machine learning studies based on their non-complexity nature. Most studies cited the use of accuracy as an evaluation metric for models in higher education. Nevertheless, some of the identified studies presented the use of a combination of evaluation techniques to better evaluate recommender systems.

Most studies implemented a cross-fold validation technique to validate the findings after developing the model. This could be attributed to the strength of splitting the data into training and testing sets. It is also worth noting that the source of the dataset that was widely used was the university repositories because of the ease of data availability.

This study serves as a basis for the investigation of machine-learning models in recommender systems in higher education. Future studies should not be limited to machine learning algorithms but should also focus on deep learning algorithms that are currently gaining popularity.

6. ACKNOWLEDGMENT

This research received no specific grants from any funding agency in the public, commercial, or not-for-profit sectors.

7. REFERENCES

Aher, S. B., & Lobo, L. M. R. J. (2013). Combination of machine learning algorithms for recommendation of courses in E-Learning System based on historical data. *Knowledge-Based Systems*, *51*, 1–14. https://doi.org/10.1016/j.knosys.2013.04.015

Alkubaisi, G. A. A. J., Kamaruddin, S. S., & Husni, H. (2018). Stock Market Classification Model Using Sentiment Analysis on Twitter Based on Hybrid Naive Bayes Classifiers. *Computer and Information Science*, *11*(1), 52. https://doi.org/10.5539/cis.v11n1p52

Alyari, F., & Navimipour, N. (2018). Recommender systems: A systematic review of the state of the art literature and suggestions for future research. *Kybernetes*, *47*(5), 985–1017. https://doi.org/10.1108/K-06-2017-0196

AlZoman, R. M., & Alenazi, M. J. F. (2021). A Comparative Study of Traffic Classification Techniques for Smart City Networks. *Sensors*, *21*(14), 4677. https://doi.org/10.3390/s21144677

Angra, S., & Ahuja, S. (2017). Machine learning and its applications: A review. *2017 International Conference on Big Data Analytics and Computational Intelligence (ICBDAC)*, 57–60. https://doi.org/10.1109/ICBDACI.2017.8070809

Assami, S., Daoudi, N., & Ajhoun, R. (2022). Implementation of a Machine Learning-Based MOOC Recommender System Using Learner Motivation Prediction. *International Journal of Engineering Pedagogy (IJEP)*, *12*(5), 68–85. https://doi.org/10.3991/ijep.v12i5.30523

Barramuño, M., Meza-Narváez, C., & Gálvez-García, G. (2022). Prediction of student attrition risk using machine learning. *Journal of Applied Research in Higher Education*, *14*(3), 974–986. https://doi.org/10.1108/JARHE-02-2021-0073

Baskota, A., & Ng, Y.-K. (2018). A Graduate School Recommendation System Using the Multi-Class Support Vector Machine and KNN Approaches. *2018 IEEE International Conference on Information Reuse and Integration (IRI)*, 277–284. Salt Lake City, UT: IEEE. https://doi.org/10.1109/IRI.2018.00050

Burke, R., Felfernig, A., & Göker, M. (2011). Recommender Systems: An Overview. *Ai Magazine*, *32*, 13–18. https://doi.org/10.1609/aimag.v32i3.2361

Choi, R. Y., Coyner, A. S., Kalpathy-Cramer, J., Chiang, M. F., & Campbell, J. P. (2020). Introduction to Machine Learning, Neural Networks, and Deep Learning. *Translational Vision Science & Technology*, *9*(2), 14–14. https://doi.org/10.1167/tvst.9.2.14

Christou, V., Tsoulos, I., Loupas, V., Tzallas, A. T., Gogos, C., Karvelis, P. S., … Giannakeas, N. (2023). Performance and early drop prediction for higher education students using machine learning. *Expert Systems with Applications*, *225*, 120079. https://doi.org/10.1016/j.eswa.2023.120079

Cui, Y., Chen, F., Shiri, A., & Fan, Y. (2019). Predictive analytic models of student success in higher education: A review of methodology. *Information and Learning Sciences*, *120*(3/4), 208–227. https://doi.org/10.1108/ILS-10-2018-0104

da Silva, F. L., Slodkowski, B. K., da Silva, K. K. A., & Cazella, S. C. (2023). A systematic literature review on educational recommender systems for teaching and learning: Research trends, limitations and opportunities. *Education and Information Technologies*, *28*(3), 3289–3328. https://doi.org/10.1007/s10639-022-11341-9

Dastres, R., & Soori, M. (2021). Artificial Neural Network Systems. *International Journal of Imaging and Robotics*, *21*, 13–25.

Deng, H. (2019, February). Towardsdatascience. Retrieved October 29, 2021, from Recommender Systems in Practice website: https://teams.microsoft.com/_#/pre-join-calling/19:meeting_OGU2Y2ZjNTMtMTcyYi00N2E1LWFhZGItMTA5NjU0ZjQ2MjJl@thread.v2

Deschênes, M. (2020). Recommender systems to support learners' Agency in a Learning Context: A systematic review. *International Journal of Educational Technology in Higher Education*, *17*(1), 50. https://doi.org/10.1186/s41239-020-00219-w

Elahi, M., Starke, A., El Ioini, N., Lambrix, A. A., & Trattner, C. (2022). Developing and Evaluating a University Recommender System. *Frontiers in Artificial Intelligence*, *4*, 796268. https://doi.org/10.3389/frai.2021.796268

Ezaldeen, H., Misra, R., Alatrash, R., & Priyadarshini, R. (2019). Machine Learning Based Improved Recommendation Model for E-learning. *2019 International Conference on Intelligent Computing and Remote Sensing (ICICRS)*, 1–6. https://doi.org/10.1109/ICICRS46726.2019.9555866

Ezugwu, V. O., & Ologun, S. (2017). Markov chain: A predictive model for manpower planning. *Journal of Applied Sciences and Environmental Management*, *21*(3), 557. https://doi.org/10.4314/jasem.v21i3.17

Farizawani, A., Puteh, M., Marina, Y., & Rivaie, A. (2020). A review of artificial neural network learning rule based on multiple variant of conjugate gradient approaches. *Journal of Physics: Conference Series*, *1529*(2), 022040. https://doi.org/10.1088/1742-6596/1529/2/022040

Ferreira, D., Silva, S., Abelha, A., & Machado, J. (2020). Recommendation System Using Autoencoders. *Applied Sciences*, *10*(16). https://doi.org/10.3390/app10165510

Fränti, P., & Mariescu-Istodor, R. (2023). Soft precision and recall. *Pattern Recognition Letters*, *167*, 115–121. https://doi.org/10.1016/j.patrec.2023.02.005

Gan, S., Shao, S., Chen, L., Yu, L., & Jiang, L. (2021). Adapting Hidden Naive Bayes for Text Classification. *Mathematics*, *9*(19), 2378. https://doi.org/10.3390/math9192378

Gaye, B., Zhang, D., & Wulamu, A. (2021). Improvement of Support Vector Machine Algorithm in Big Data Background. *Mathematical Problems in Engineering*, *2021*, 5594899. https://doi.org/10.1155/2021/5594899

Gholami, R., & Fakhari, N. (2017). Support Vector Machine: Principles, Parameters, and Applications. In *Handbook of Neural Computation* (pp. 515–535). Elsevier. https://doi.org/10.1016/B978-0-12-811318-9.00027-2

Girase, S., Powar, V., & Mukhopadhyay, D. (2017). A user-friendly college recommending system using user-profiling and matrix factorization technique. *2017 International Conference on Computing, Communication and Automation (ICCCA)*, 1–5. Greater Noida: IEEE. https://doi.org/10.1109/CCAA.2017.8229779

Goga, M., Kuyoro, S., & Goga, N. (2015). A Recommender for Improving the Student Academic Performance. *Procedia - Social and Behavioral Sciences*, *180*, 1481–1488. https://doi.org/10.1016/j.sbspro.2015.02.296

Gulzar, Z., Leema, A., & Deepak, G. (2018). PCRS: Personalized Course Recommender System Based on Hybrid Approach. *Procedia Computer Science*, *125*, 518–524. https://doi.org/10.1016/j.procs.2017.12.067

Guo, L., Liang, J., Zhu, Y., Luo, Y., Sun, L., & Zheng, X. (2019). Collaborative filtering recommendation based on trust and emotion. *Journal of Intelligent Information Systems*, *53*(1), 113–135. https://doi.org/10.1007/s10844-018-0517-4

Han, J., Kelley, T., & Knowles, J. G. (2021). Factors Influencing Student STEM Learning: Self-Efficacy and Outcome Expectancy, 21st Century Skills, and Career Awareness. *Journal for STEM Education Research*, *4*(2), 117–137. https://doi.org/10.1007/s41979-021-00053-3

Hand, D. J., Christen, P., & Kirielle, N. (2021). F*: An interpretable transformation of the F-measure. *Machine Learning*, *110*(3), 451–456. https://doi.org/10.1007/s10994-021-05964-1

Hasan, M., Ahmed, S., Abdullah, D. Md., & Rahman, Md. S. (2016). Graduate school recommender system: Assisting admission seekers to apply for graduate studies in appropriate graduate schools. *2016 5th International Conference on Informatics, Electronics and Vision (ICIEV)*, 502–507. Dhaka, Bangladesh: IEEE. https://doi.org/10.1109/ICIEV.2016.7760053

Hernández, R., Musso, M., Kyndt, E., Eduardo Cascallar, & Carlos Felipe. (2021). Artificial neural networks in academic performance prediction: Systematic implementation and predictor evaluation. *Computers and Education: Artificial Intelligence*, *2*, 100018. https://doi.org/10.1016/j.caeai.2021.100018

Hicks, S. A., Strümke, I., Thambawita, V., Hammou, M., Riegler, M. A., Halvorsen, P., & Parasa, S. (2022). On evaluation metrics for medical applications of artificial intelligence. *Scientific Reports*, *12*(1), 5979. https://doi.org/10.1038/s41598-022-09954-8

Ismail, N., & Yusof, U. K. (2022). RECENT TRENDS OF MACHINE LEARNING PREDICTIONS USING OPEN DATA: A SYSTEMATIC REVIEW. *Journal of Information and Communication Technology*, *21*. https://doi.org/10.32890/jict2022.21.3.3

Janiesch, C., Zschech, P., & Heinrich, K. (2021). Machine learning and deep learning. *Electronic Markets*, *31*(3), 685–695. https://doi.org/10.1007/s12525-021-00475-2

Jena, K. K., Bhoi, S. K., Malik, T. K., Sahoo, K. S., Jhanjhi, N. Z., Bhatia, S., & Amsaad, F. (2022). E-Learning Course Recommender System Using Collaborative Filtering Models. *Electronics*, *12*(1), 157. https://doi.org/10.3390/electronics12010157

Jiang, T., Gradus, J. L., & Rosellini, A. J. (2020). Supervised Machine Learning: A Brief Primer. *Behavior Therapy*, *51*(5), 675–687. https://doi.org/10.1016/j.beth.2020.05.002

Jimoh, R. G., Abisoye, O. A., & Uthman, M. M. B. (2021). Ensemble Feed-Forward Neural Network and Support Vector Machine for Prediction of Multiclass Malaria Infection. *Journal of Information and Communication Technology*, *21*. https://doi.org/10.32890/jict2022.21.1.6

Jwo, D.-J., Biswal, A., & Mir, I. A. (2023). Artificial Neural Networks for Navigation Systems: A Review of Recent Research. *Applied Sciences*, *13*(7), 4475. https://doi.org/10.3390/app13074475

Kroll, C., Warchold, A., & Pradhan, P. (2019). Sustainable Development Goals (SDGs): Are we successful in turning trade-offs into synergies? *Palgrave Communications*, *5*(1), 140. https://doi.org/10.1057/s41599-019-0335-5

Kumari, K., & Yadav, S. (2018). Linear regression analysis study. *Journal of the Practice of Cardiovascular Sciences*, *4*, 33. https://doi.org/10.4103/jpcs.jpcs_8_18

Liang, H., Ren, J., Gao, S., Dong, L., & Gao, Z. (2017). Chapter 8—Comparison of Different Multicriteria Decision-Making Methodologies for Sustainability Decision Making. In A. Scipioni, A. Manzardo, & J. Ren (Eds.), *Hydrogen Economy* (pp. 189–224). Academic Press. https://doi.org/10.1016/B978-0-12-811132-1.00008-0

Machado, M. R., & Karray, S. (2022). Assessing credit risk of commercial customers using hybrid machine learning algorithms. *Expert Systems with Applications*, *200*, 116889. https://doi.org/10.1016/j.eswa.2022.116889

Maita, A. R. C., Martins, L. C., López Paz, C. R., Peres, S. M., & Fantinato, M. (2015). Process mining through artificial neural networks and support vector machines: A systematic literature review. *Business Process Management Journal*, *21*(6), 1391–1415. https://doi.org/10.1108/BPMJ-02-2015-0017

Maphosa, V., & Maphosa, M. (2023). Fifteen Years of Recommender Systems Research in Higher Education: Current Trends and Future Direction. *Applied Artificial Intelligence*, *37*(1), 2175106. https://doi.org/10.1080/08839514.2023.2175106

Miric, M., Jia, N., & Huang, K. G. (2023). Using supervised machine learning for large-scale classification in management research: The case for identifying artificial intelligence patents. *Strategic Management Journal*, *44*(2), 491–519. https://doi.org/10.1002/smj.3441

Najafabadi, M. M., Villanustre, F., Khoshgoftaar, T. M., Seliya, N., Wald, R., & Muharemagic, E. (2015). Deep learning applications and challenges in big data analytics. *Journal of Big Data*, *2*(1), 1. https://doi.org/10.1186/s40537-014-0007-7

Nichols, J. A., Herbert Chan, H. W., & Baker, M. A. B. (2019). Machine learning: Applications of artificial intelligence to imaging and diagnosis. *Biophysical Reviews*, *11*(1), 111–118. https://doi.org/10.1007/s12551-018-0449-9

Obeid, C., Lahoud, I., El Khoury, H., & Champin, P.-A. (2018). Ontology-based Recommender System in Higher Education. *Companion of the The Web Conference 2018 on The Web Conference 2018 - WWW '18*, 1031–1034. Lyon, France: ACM Press. https://doi.org/10.1145/3184558.3191533

Oliveira, T. N. D., Bernardini, F., & Viterbo, J. (2021). An Overview on the Use of Educational Data Mining for Constructing Recommendation Systems to Mitigate Retention in Higher Education. *2021 IEEE Frontiers in Education Conference (FIE)*, 1–7. https://doi.org/10.1109/FIE49875.2021.9637207

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *PLOS Medicine*, *18*(3), e1003583. https://doi.org/10.1371/journal.pmed.1003583

Pisner, D. A., & Schnyer, D. M. (2020). Support vector machine. In *Machine Learning* (pp. 101–121). Elsevier. https://doi.org/10.1016/B978-0-12-815739-8.00006-7

Polyzou, A., Nikolakopoulos, A. N., & Karypis, G. (2019). *Scholars Walk: A Markov Chain Framework for Course Recommendation*.

Portugal, I., Alencar, P., & Cowan, D. (2018). The use of machine learning algorithms in recommender systems: A systematic review. *Expert Systems with Applications*, *97*, 205–227. https://doi.org/10.1016/j.eswa.2017.12.020

Ray, S. (2019). A Quick Review of Machine Learning Algorithms. *2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)*, 35–39. Faridabad, India: IEEE. https://doi.org/10.1109/COMITCon.2019.8862451

Roy, D., & Dutta, M. (2022). A systematic review and research perspective on recommender systems. *Journal of Big Data*, *9*(1), 59. https://doi.org/10.1186/s40537-022-00592-5

Sandoussi, R., Hnida, M., Daoudi, N., & Ajhoun, R. (2022). Systematic Literature Review on Open Educational Resources Recommender Systems. *International Journal of Interactive Mobile Technologies (IJIM)*, *16*(18), 44–77. https://doi.org/10.3991/ijim.v16i18.32197

Sangka, K. B., & Muchsini, B. (2018). Accommodating Analytic Hierarchy Process (AHP) for Elective Courses Selection. *IJIE (Indonesian Journal of Informatics Education)*, *2*(2). https://doi.org/10.20961/ijie.v2i2.24436

Sattar, A., Ghazanfar, M. A., & Iqbal, M. (2017). Building Accurate and Practical Recommender System Algorithms Using Machine Learning Classifier and Collaborative Filtering. *Arabian Journal for Science and Engineering*, *42*(8), 3229–3247. https://doi.org/10.1007/s13369-016-2410-1

Schneider, P., & Xhafa, F. (2022). Anomaly detection. In *Anomaly Detection and Complex Event Processing over IoT Data Streams* (pp. 49–66). Elsevier. https://doi.org/10.1016/B978-0-12-823818-9.00013-4

Sekeroglu, B., Abiyev, R., Ilhan, A., Arslan, M., & Idoko, J. B. (2021). Systematic Literature Review on Machine Learning and Student Performance Prediction: Critical Gaps and Possible Remedies. *Applied Sciences*, *11*(22), 10907. https://doi.org/10.3390/app112210907

Singh, K. J., Kapoor, D. S., & Sohi, B. S. (2021). All about human-robot interaction. In *Cognitive Computing for Human-Robot Interaction* (pp. 199–229). Elsevier. https://doi.org/10.1016/B978-0-323-85769-7.00010-0

Song, P., & Wang, X. (2020). A bibliometric analysis of worldwide educational artificial intelligence research development in recent twenty years. *Asia Pacific Education Review*, *21*(3), 473–486. https://doi.org/10.1007/s12564-020-09640-2

Sridhar, S., Mootha, S., & Kolagati, S. (2020). A University Admission Prediction System using Stacked Ensemble Learning. *2020 Advanced Computing and Communication Technologies for High Performance Applications (ACCTHPA)*, 162–167. https://doi.org/10.1109/ACCTHPA49271.2020.9213205

Srivastava, D., & Bhambhu, L. (2010). Data classification using support vector machine. *Journal of Theoretical and Applied Information Technology*, *12*, 1–7.

Thangavel, S. K., Bkaratki, P. D., & Sankar, A. (2017). Student placement analyzer: A recommendation system using machine learning. *2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS)*, 1–5. Coimbatore, India: IEEE. https://doi.org/10.1109/ICACCS.2017.8014632

Uddin, I., Imran, A. S., Muhammad, K., Fayyaz, N., & Sajjad, M. (2021). A Systematic Mapping Review on MOOC Recommender Systems. *IEEE Access*, *9*, 118379–118405. https://doi.org/10.1109/ACCESS.2021.3101039

Urdaneta-Ponte, M. C., Mendez-Zorrilla, A., & Oleagordia-Ruiz, I. (2021). Recommendation Systems for Education: Systematic Review. *Electronics*, *10*(14), 1611. https://doi.org/10.3390/electronics10141611

Vargas, R., Mosavi, A., & Ruiz, R. (2017). DEEP LEARNING: A REVIEW. *Advances in Intelligent Systems and Computing*, *5*.

Waheed, H., Hassan, S.-U., Nawaz, R., Aljohani, N. R., Chen, G., & Gasevic, D. (2023). Early prediction of learners at risk in self-paced education: A neural network approach. *Expert Systems with Applications*, *213*, 118868. https://doi.org/10.1016/j.eswa.2022.118868

Wang, X., Cui, L., Bangash, M., Bilal, M., Rosales, L., & Chaudhry, W. (2022). *A Machine Learning-based Course Enrollment Recommender System* (p. 443). https://doi.org/10.5220/0011109100003182

Yanes, N., Mostafa, A. M., Ezz, M., & Almuayqil, S. N. (2020). A Machine Learning-Based Recommender System for Improving Students Learning Experiences. *IEEE Access*, *8*, 201218–201235. https://doi.org/10.1109/ACCESS.2020.3036336

ISSN 2519-5115
IJART- Vol-8, Issue-6, December, 2023
Available Online at http://www.ijart.info/
DOI: https://doi.org./10.24163/ijart/2017/8(6):189-206

Yang, F.-J. (2018). An Implementation of Naive Bayes Classifier. *2018 International Conference on Computational Science and Computational Intelligence (CSCI)*, 301–306. Las Vegas, NV, USA: IEEE. https://doi.org/10.1109/CSCI46756.2018.00065

Yazdi, H. A., Chabok, S. J. S. M., & Kheirabadi, M. (2021). Dynamic Educational Recommender System Based on Improved Recurrent Neural Networks Using Attention Technique. *Applied Artificial Intelligence*, 1–24. https://doi.org/10.1080/08839514.2021.2005298

Zhong, J., Xie, H., & Wang, F. L. (2019). The research trends in recommender systems for e-learning: A systematic review of SSCI journal articles from 2014 to 2018. *Asian Association of Open Universities Journal*, *14*(1), 12–27. https://doi.org/10.1108/AAOUJ-03-2019-0015

Zhou, Y., Huang, C., Hu, Q., Zhu, J., & Tang, Y. (2018). Personalized learning full-path recommendation model based on LSTM neural networks. *Information Sciences*, *444*, 135–152. https://doi.org/10.1016/j.ins.2018.02.053