Juma H. Shindo
Candidate in Masters of Science in Computer Science, College Tutor
The University of Dodoma, Department of Computer Science and Engineering, Dodoma, Tanzania
ORCID ID 0000-0003-4659-701X
shindostz@gmail.com

Mohamed M. Mjahidi
PhD in Computer Engineering, Lecturer
The University of Dodoma, Department of Computer Science and Engineering, Dodoma, Tanzania
ORCID ID 0000-0002-3962-6997
mohamedi.mjahidi@udom.ac.tz.

Mohamed D. Waziri
PhD in Information and Communication Science and Engineering
The University of Dodoma, Department of Computer Science and Engineering, Dodoma, Tanzania
ORCID ID 0000-0003-2668-3279
mohamed.waziri2@udom.ac.tz

DATA MINING ALGORITHMS FOR PREDICTION OF STUDENT TEACHERS' PERFORMANCE IN ICT: A SYSTEMATIC LITERATURE REVIEW

Abstract. Poor ICT performance in teacher training colleges makes it more difficult for the majority of teachers to successfully use ICT resources in their teaching and learning. When teachers can efficiently utilize ICT resources, it empowers them to update their knowledge through online learning, consequently enhancing the overall quality of teaching and learning. This positive outcome can be observed through improved ICT performance. The aim of this article is to identify the appropriate Data Mining algorithms for predicting student teachers' performance in ICT. The systematic literature review that was guided by the PRISMA statement 2020 served as the study methodology. It makes for clear reporting and offers a detailed checklist and flow diagram that direct the review procedure. On November 6, 2022, about 196 scholarly articles were downloaded from three digital libraries: Science Direct (38), ACM Digital Library (72), IEEE Xplore (51), and 35 from the Google Scholar search engine. After screening and eligibility checking, 28 scholarly articles were selected and analysed through content analysis in terms of the most commonly used algorithms, the year of publication, the study purposes, and the accuracy performance metrics. Considering the specific study findings represented quantitatively, Decision Trees and Naive Bayes were found to be the most commonly used Data Mining algorithms, with a count of 20.6% each. The most recently identified articles were published between 2014 and 2022. In terms of study purposes, a large number of studies focused on predicting student performance. Furthermore, about 6 out of 8 algorithms used in previous studies were found to score 80% or above in the average percentage of the highest and lowest accuracy metrics. Therefore, considering the general findings, the study identified five Data Mining algorithms as appropriate and most commonly used for prediction of student teachers' performance in ICT. They are Naive Bayes, K-Nearest Neighbour, Support Vector Machine, Random Forest, and Decision Tree. The findings of this study would assist the government, college tutors, and student teachers in making better decisions to improve ICT performance for pre-service and in-service teachers.

Keywords: data mining algorithms; educational data mining; student teachers; teacher training colleges.

1. INTRODUCTION

As Information and Communication Technologies (ICT) advanced, educational institutions used them to keep students' academic, financial, and social information [1]. Following the ICT project launched by the government of Tanzania in mutual aid with the Swedish International Development Cooperation Agency (SIDA) in 2005, teacher training colleges in Tanzania started teaching ICT as a subject to student teachers (STs) in 2007 [2]. A ST is a student pursuing a diploma or certificate in education at a teacher training college [3].

The governments, teacher training colleges, school authorities, and parents anticipated that the STs would graduate with excellent grades in ICT [4]. This is because ICT is the driving force for searching and preparing teaching and learning materials such as the schemes of work, lesson plans, and lesson notes for every subject. Additionally, firms are seeking top graduates to hire, especially in the fields of science and ICT [5].

However, due to a number of factors such as their demographic profile, financial situations, and educational background, the majority of STs perform poorly in ICT, which hinders them from acquiring the necessary ICT competence [6]. As a result, Tanzania's primary and secondary schools have a severe shortage of qualified and competent ICT teachers. This was evidenced in 2019 when the government of the United Republic of Tanzania (URT) sought to hire 273 ICT teachers for secondary schools, but only 72 (26.5%) were available [7].

The Tanzania teacher training colleges stores a huge amount of data collected from STs' social, financial, and academic records. The saved data has hidden information that can help students, tutors, and college administrators improve academic performance [8]. Despite the accumulation and availability of huge amounts of students' academic and social records, they are not often used to help students solve their academic problems [9], [10], [11]. Due to the latest developments in the field of Data Mining, it is now possible to mine educational data and provide information that may help students and teachers make better decisions [12]. The practice of uncovering hidden information from large sets of data is known as Data Mining [13]. When Data Mining is used in the field of education, it is known as Educational Data Mining (EDM) [14]. EDM is the name given to the technological tools that are used to anticipate students' performance on the provided dataset [15]. EDM algorithms are capable of mining the data and producing knowledge that may help STs, college tutors and educational managers in decision-making [5]. The reasons for students' failure or success can be established to aid in decision-making especially for policymakers, college academics, administrators, and the government at large [16]. So, this study identified the EDM algorithms that can be effectively used to analyze and predict the students' performance in ICT subjects at teacher training colleges.

The problem statement. The accurate prediction of a student's performance in ICT is impossible without identifying and using the appropriate Data Mining algorithms for machine learning. ICT performance is poor in teacher training colleges, which makes it hard for most teachers to use ICT resources effectively in the classroom and for administrative tasks. Good performance in ICT helps pre- and in-service teachers use ICT facilities efficiently, which enables them to keep their knowledge and skills up-to-date through online learning and searching the internet for teaching and learning materials [2]. As a result, teaching and learning procedures might become more standardized, which would lead to better performance in ICT as well as all other academic disciplines.

Knowledge of EDM algorithms is important for different educational stakeholders, such as the government, researchers, college tutors, parents, and student teachers. Understanding the proper predictive algorithms could enable researchers to come up with accurate findings that would prompt other stakeholders to figure out the best measures to improve students' performance in ICT.

Currently, there are different EDM algorithms used to predict student performance. Most of those algorithms have been widely used at higher education institutions in the fields of engineering, computer science, and health science. However, very few academic prediction studies have been done at teacher training colleges to identify the causes of students' failure and propose appropriate solutions [17]. That is why it is important to conduct a study that identifies appropriate Data Mining algorithms to predict STs' performance in ICT at teacher training colleges.

Analysis of recent studies and publications. Even though the same curriculum is implemented every year, Shingari and Kumar's study [18] found that it is hard to predict how well students will perform. The problem is that classroom activities like homework, quizzes, and tests, as well as students' cognitive abilities change annually. This makes it hard for the general prediction to be true for all students. The best solution to the problem is to use Data Mining algorithms that predict students' academic performance based on the pre-described features. Various studies on the prediction of students' performance using EDM algorithms have been conducted inside and outside Africa in recent years. The following are the reviewed studies that facilitate the attainment of the objective of this work.

The study by Hasan *et al*.[19] predicted students' performance in an e-commerce course in Oman's Private Higher Education Institutions using data from 22 undergraduate students. The study employed eight algorithms with WEKA software to create an academic prediction model. According to the study's findings, the Random Forest (RF), with an absolute error of 37.2857 percent, can forecast performance with a high degree of precision and accuracy. One of the researchers' successes was the use of eight algorithms to develop the prediction model, which ensures the reliability and validity of the results obtained.

In a similar vein, Cortez and Silva [20] conducted research using EDM approaches to predict secondary school students' success in mathematics and the Portuguese language in the Alentejo region of Portugal. During the 2006 examination, 788 student records were collected from two Portuguese public schools via documentary reviews and questionnaires. Five Data Mining algorithms were used to analyze the data: Decision Tree (DT), Night Vision (NV), RF, K-Nearest Neighbor (K-NN), and Support Vector Machine (SVM). The results show that the DT and K-NN algorithms outperformed the other algorithms in mathematics and Portuguese by 93.0 and 91.9 percent, respectively.

Also, Michael and Gold [21] conducted another study to find out the factors that have the highest impact on students' mathematics in Makundi, Nigeria. The DT algorithm was used in the study. A questionnaire was used to collect the 200 records that make up the dataset. The major drawback of the study is the researcher's decision to use only the DT algorithm, possibly due to its simplicity in analyzing the data. However, given the size of the datasets used in this investigation, the DT approach is less effective. The DT technique may be used with additional classification algorithms like the RF, Naive Bayes, and SVM for more trustworthy and realistic results.

Yehuala [22] applied the DT and NB algorithms to predict students' performance at Debre Markos University in Ethiopia. The dataset of 11,873 records with three attributes (students' background, perceptions, and study habits) of first-degree students was used to create and test the Data Mining model using the WEKA 3.7 software. The study employs a Cross-Industry Standard Process for Data Mining (CRISP) to assure the reliability and repeatability of the process. The research findings reveal that the maximum prediction correctness of low performing students (below 2.0 GPA) was 92.34%, provided by the DT algorithm by means of 10-fold cross validation.

Furthermore, Kavishe [15] conducted a study in Tanzania to predict the mathematics performance among the students specializing in management at Mzumbe University. The study utilized K-NN, RF, DT, SVM, and Multilayer Perceptron (ML) for data processing. According to the study's findings, the RF algorithm was the best one with 99 percent accuracy, and it can be used to create a model for predicting mathematical performance at the college.

The research goal. The main goal of the study was to identify the appropriate EDM algorithms that can be used to predict STs' performance in ICT.

To help reach the main goal of the study, a leading question was formulated: "What are the most common Data Mining algorithms used for predicting STs performance in the ICT subject?" To get the results from the reviewed journal articles, four themes were identified:

the most common Data Mining algorithms, the year of publications, the study purposes, and the evaluation metrics criteria.

2. RESEARCH METHODS

2.1. Systematic literature review

The systematic literature review (SLR) has been carried out to identify the most appropriate and commonly used Data Mining algorithms for predicting STs' performance in ICT. The Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA 2020 statement) was used to guide the SLR procedures [23]. The PRISMA approach facilitates clear reporting of SLR by providing a comprehensive checklist and flow diagram that guide the reviewing process [24]. Based on the Page et al. study, there are three basic review procedures: planning, conducting, and reporting, as shown in Figure1.

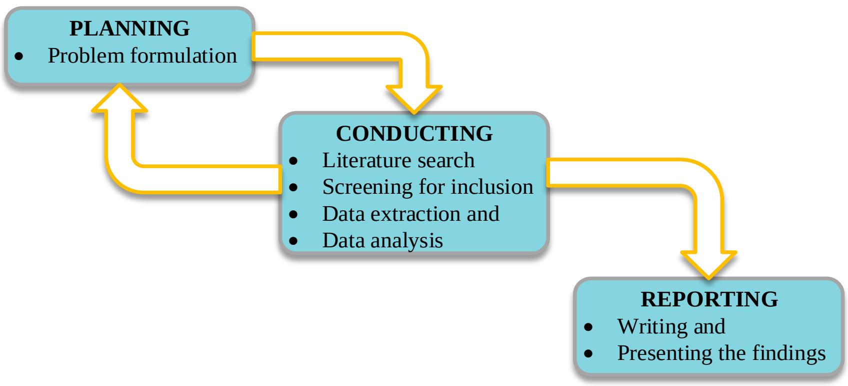

Figure 1. Systematic literature review methodology

2.2. Problem formulation

In the formulation of the study problem, the PICO model was used, as suggested by Namoun and Alshanqiti [25]. It was used in the identification of the research question, which focuses on four important elements: the population, the intervention, the context, and the outcome. Table 1 shows the PICO criteria and presentations that were used to formulate the study problem.

Table 1

Inclusion and exclusion criteria

PICO	Representation
Population	STs
Intervention	Data Mining algorithms
Context	Tanzania teacher training colleges
Outcome	ST's performance in ICT

2.3. Literature search

By using various keywords, on September 6, 2022, journal articles were searched and downloaded via Google Scholar and three scholarly digital databases, which include Science Direct, ACM Digital Library, and IEEE Xplore. That resulted in a total of 196 downloaded articles: 72 (36.73%) from the ACM Digital Library, 51 (26.02%) from IEEE Xplore, 38 (19.39%) from Science Direct, and 35 (17.86%) from Google Scholar. The key elements that facilitated article downloading are shown in Table 2.

Table 2

Search engines and keywords used in downloading journal articles

Digital databases	URL	Keywords
Science Direct	https://www.sciencedirect.com/	'Data Mining algorithms ' AND 'forecast' AND 'student performance'.
ACM Digital Library	https://dl.acm.org/	"Educational Data Mining" AND "prediction" AND "student performance."
IEEE Xplore	https://ieeexplore.ieee.org	"Educational Data Mining" AND "prediction" and "student performance."
Google Scholar	https://scholar.google.com/	('Data Mining algorithms ' OR 'Data Mining approach') AND ('predict') AND ('student performance').

2.4. Inclusion and exclusion criteria

Table 3 shows the criteria that were set up to decide whether a journal article should be included in or excluded from the analysis.

Table 3

Screening criteria

Screening	Criteria
Inclusion	Articles from the first five pages of scholarly databases
	Publication year between 2006 and 2022
	English-language articles
	Full text scholarly articles
Exclusion	Duplicated articles
	Disparities in the context compared to the current study's objective.

The exclusion eliminated 168 journal articles, of which 129 titles did not match the goal of the study; 2 were written in languages other than English; 2 were not in full text; 1 was a duplicate; and 1 was published before 2006. After the eligibility check, 33 articles were further rejected based on contextual disparity criteria, as shown in the PRISMA flow diagram in Figure 2.

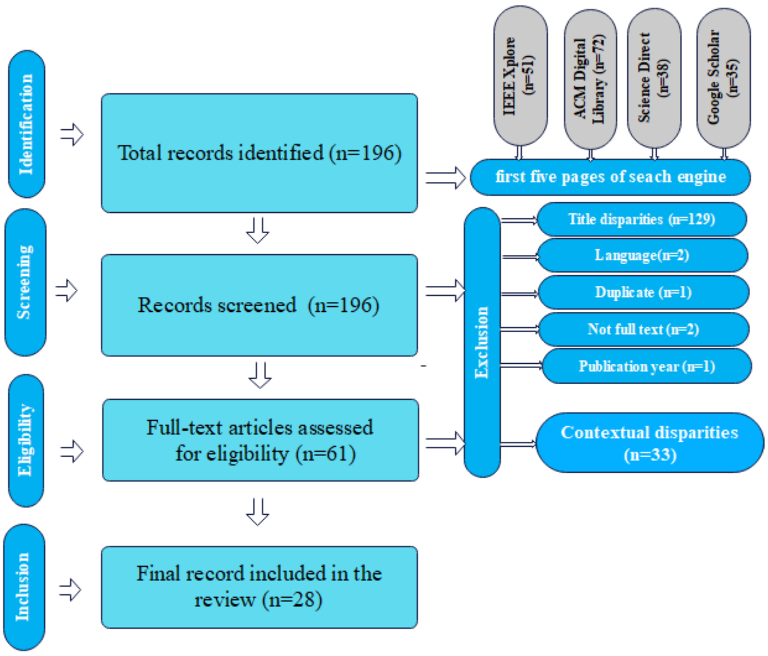

Figure 2. PRISMA flow diagram for inclusion and exclusion criteria

2.5. Data extraction

The 28 journal articles that were selected for review were: 9 (32%) from Google Scholar; 7 (25%) from IEEE Xplore; 5 (18%) from Science Direct; and 7 (25%) from ACM Digital Library. About 168 articles were excluded, of which 26 (15%) were from Google Scholar, 44 (26%) from IEEE Xplore, 33 (20%) from Science Direct, and 65 (39%) from ACM Digital Library. Table 4 shows the detailed information extracted from 28 articles and presented in a coded schema based on the references, study purposes, methodologies, Data Mining algorithms, evaluation metrics, and major findings.

Table 4

Detailed information of the major themes in a coded schema

Sn	Reference	Studies' Purpose	Methodology	DM Algorithms	Metrics	Finding
1	[26]	To find the relationship between EDM and SAP.	Experimental-Rapid Miner.	NB, DT, & ANN.	Accuracy.	ANN Accuracy of 95%.
2	[27]	To create SPPM	Experimental-edX platform.	LR, SVM, NB, K-NN & BN.	Accuracy, F1-score, Recall & Precision.	SVM Accuracy of 99.99.
3	[28]	To create SPPM in ILE.	Experimental, Cross Validation.	DT, NB, RF, SVM, & ANN.	Accuracy.	ANN Accuracy of 62.5% in Algorithm & Design.
4	[29]	To create SACCM.	Experimental.	SVM & NB.	Accuracy, Precision & F1-Score.	SVM Accuracy of 93%.
5	[30]	To use three MLA to PSP.	Experimental.	BP, SVR & LSTM.	Accuracy.	BP Accuracy of 87.78%.
6	[31]	To identify commonly used EDM in SPP.	SLR.	ANN, Regression, CL, DT, SVM, NB, & Association.	-	ANN mostly used in EDM.
7	[32]	To review commonly used DM technique.	SLR.	DT, ANN, NB, K-NN & SVM.	-	ANN & DT are highly used.
8	[33]	To review EDM on SAP.	Literature survey.	CL, Classification, Regression & Association.	-	Regression and classification are widely used on PSAP.
9	[34]	To compare recently used DM techniques.	Experimental.	NB, C4.5, IDE3, ADTree, LR, J48 NNge, & OneR.	Accuracy, Precision & GM.	ADTree Accuracy of 97.30%.
10	[35]	To find popular SPPM.	SLR.	DT, BN, RF, K-NN, NB & SVM.	-	Popular are: DT, BN, & RF.
11	[16]	To develop MLPM using classification algorithm.	Experimental-WEKA.	DT- C4.5, ID3 & CART.	Accuracy.	C4.5 Accuracy 67.78%.
12	[36]	To identify factors related to SAP.	SLR.	DT, ANN, NB, K-NN & SVM.	-	SVM & DT are mostly used. Factors-UEER, UTME &O-L.
13	[37]	To show DM	SLR.	DT, BC, NN, &	-	DT & CL are

Sn	Reference	Studies' Purpose	Methodology	DM Algorithms	Metrics	Finding
		techniques for better decision making.		Cl.		mostly used technique for PSAP.
14	[13]	To inform on mostly used DMM for PSAP.	SLR.	DT, NB, SVM, ANN, K-NN, LR, RF, & REPTree.	-	DT, NB, ANN, & K-NN are mostly used.
15	[38]	To identify slow learners using classification DM.	Experimental-WEKA.	N B, SMO, J48, REPTree & MP.	Accuracy.	MP Accuracy of 87.43%.
16	[39]	To use DM technique to analyze SAP.	Experimental-WEKA.	J48, PART, RF & BN.	Accuracy.	RF performs better by 99% Accuracy.
17	[40]	To determine mostly used DM technique to PSP.	Experimental.	DT, NN, NB, K-NN, & SVM.	Accuracy.	NN & DT have high Accuracy of 98% each.
18	[41]	To identify DM technique & tool mostly used in PSP.	Survey.	DT, K-NN, RF, SVM, ANN, & NB.	-	DT, ANN, NB, SVM, RF, KNN & WEKA are used.
19	[42]	To explore factors & MLM to PSP.	Experimental-Rapid Miner.	DT-C4.5, ID3, CART & NB.	Accuracy & Precision.	CART Accuracy of 40%.
20	[43]	To evaluate EDM to increase PSAP.	Experimental.	NB, LR, K-NN & RF.	Precision, recall, F-measure & Accuracy.	RF outperformed by 88% of Accuracy.
21	[44]	To assess MLA in PSAP.	Experimental.	RF, SVM, DT, NB, MP, & J48.	Accuracy.	MP Accuracy of 76.07%.
22	[45]	To enhance SAP using DMM.	Experimental-WEKA.	NB & SVM.	C-Matrix.	NB Accuracy of 89.74%.
23	[46]	To identify DMM for SPP.	Experimental-10-FCV.	DT NB RF, SMO, MLR, & LMT.	Accuracy, Precision, F-1 score & Recall.	RF Accuracy of 95.45%.
24	[47]	To provide overview on SAP using EDMM.	SLR.	RF, J48, SVM, NB, &LR.	-	Mostly used are J48, RF, SVM & NB.
25	[48]	To use C4.5 classification to analyze PSP.	Experimental.	C4.5 Classification.	Precision & Accuracy.	C4.5 had highest Accuracy 71.9%.
26	[49]	Use of Classification methods to PSP in DE.	Experimental-WEKA.	Classification-RF.	Accuracy.	RF Accuracy of 66%.
27	[50]	To determine factors ASAP.	Experimental.	SVM NB, C4.5 & ID3.	Accuracy & Error rate.	ST&I are linked to SAP SVM Accuracy of 95%.
28	[51]	Use DMM to assess SAP.	Experimental.	DT, RI, ANN, K-NN, NB & RF.	Accuracy.	NB Accuracy of 83.65%.

1. **Study purposes**: Student Performance Prediction Model (SPPM), Educational Data Mining (EDM), Student Anticipating Conversation Creativity Model (SACCM). Machine Learning Algorithms (MLA), Predict Student Performance (PSP), Data Mining Methods (DMM), Student Academic Performance (SAP), Predicting Student Academic Performance (PSAP), Student Performance Prediction (SPP), and Affecting Student Academic Performance (ASAP). Information Lacking Environment (ILE), Distributed Environment (DE).
2. **Methodology**: 10-Fold Cross Validation (10-FCV), Documentary Review (DR), Systematic Literature Review (SLR). Waikato Environment for Knowledge Analysis (WEKA).
3. **Data Mining algorithms**: Naïve Bayes (NB), Decision Tree (DT), Deep Learning (DL), Support Vector Machine (SVM) Bayes Network (BN), Clustering (CL) Artificial Neural Network (ANN), MLP (MP) Multinomial Linear Regression (MLR), Linear Regression (LR), Sequential Minimal Optimization (SMO), Classical Machine Learning (CML), Logistic Model Trees (LMT), Back-Propagation (BP), Rule Induction (RI), Bayesian classifier (BC).
4. **Metrics:** Confusion Matrix (C-Matrix), Geometric Mean (GM).
5. **Findings**: Machine Learning Predictive Model (MLPM), University Entrance Examinations Result (UEER), Unified Tertiary Matriculation Examination (UTME) Ordinary Level Result (O-L), Student Talent and Interest (ST&I).

2.6. Data analysis

The data extracted from the 28 articles was analyzed using a content analysis approach. Content analysis is an analytical method for analyzing text data and quantitatively representing the findings [52]. To provide an answer to the research question, four elements of component analysis were considered: mostly used Data Mining algorithms, the years of publication, study purpose, and performance metrics.

3. THE RESULTS AND DISCUSSION

3.1 Most Used Data Mining Algorithms

The findings from 28 articles summarized in Table 5 revealed the DT and NB as the most frequently used Data Mining algorithms, with 20.6% each. SVM counted 14 (13.7%), RF counted 11 (10.8%), and K-NN and ANN counted 9 (8.8%) each. The counts for BN, CL, MP, LR, and SMO were 4 (3.9%), 3 (2.9%), 2 (2.0%), 3 (2.9%), and 2 (2.0%), respectively. BP, Long Term-Short Memory (LTSM), and SVR were the least used Data Mining algorithms, with 1 count each.

Table 5

Most frequently used data mining algorithms

Sn	Algorithms	Count	Percentage (%)
1	NB	21	20.6
2	DT	21	20.6
3	ANN	9	8.8
4	BP	1	1.0
5	SVR	1	1.0
6	LSTM	1	1.0

7	SVM	14	13.7
8	K-NN	9	8.8
9	RF	11	10.8
10	CL	3	2.9
11	BN	4	3.9
12	SMO	2	2.0
13	MP	2	2.0
14	LR	3	2.9

3.2 Most Used Data Mining Algorithms Based on Years of Publication

The 28 reviewed articles were categorized by the year of their publication, from 2006 to 2009, 2010 to 2013, 2014 to 2017, and 2018 to 2022. There was no study published between 2006 and 2009. Only the study by Yadav and Pal [16] was published between 2010 to 2013. From 2014 to 2017, 10 studies were identified, while in 2018 to 2022, 17 studies were identified, as shown in Figure 3. Within the identified studies, different Data Mining algorithms were determined, as shown in Table 6.

Table 6

Data mining algorithms in relation to years of publication

Sn	**Data Mining algorithms**	**Number of Articles Reviewed**					
		Count	**Percent (%)**	**Count**	**Percent (%)**	**Count**	**Percent (%)**
		2018 to 2022		**2014 to 2017**		**2010 to 2013**	
1	NB	11	18.74	10	23.81	1	25.00
2	DT	14	22.58	8	19.05	1	25.00
3	ANN	6	9.68	5	11.90	0	0.00
4	BP	1	1.61	0	0.00	0	0.00
5	BN	1	1.61	2	4.76	0	0.00
6	LTSM	1	1.61	0	0.00	0	0.00
7	SVM	9	14.52	4	9.52	0	0.000
8	K-NN	6	9.68	7	16.67	1	25.00
9	RF	8	12.90	3	7.14	1	25.00
10	CL	3	4.84	2	4.76	0	0.00
11	MLP	1	1.61	0	0.00	0	0.00
12	LR	1	1.61	1	2.38	0	0.00

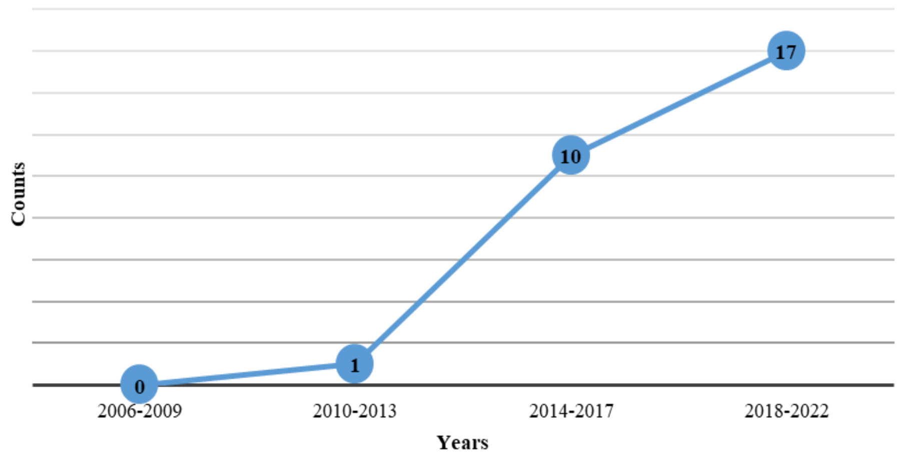

Figure 3. Data mining algorithms based on the year of publication

3.3 Most Used Data Mining Algorithms based on the Study Purposes

In the 28 articles that were examined, four main categories of study purposes were found. After analyzing the articles, it was discovered that about 12 articles focused on predicting student performance (PSP), 7 on predicting and classifying student performance (P&CSP), 6 on identifying factors that influence student performance (FISP), and 3 on determining the most commonly used Data Mining method (CDMM), as shown in Figure 4.

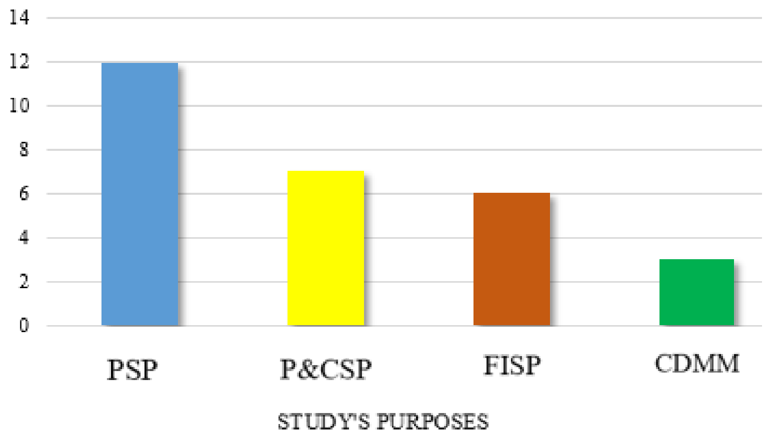

Figure 4. Data mining count based on study purposes

Moreover, the Data Mining algorithms identified across all four study purposes have been documented. The ANN, SVM, and RF scored 11 percent each, whereas NB and DT outperformed other algorithms with 21 percent each. The percentages for K-NN, CL, and LTSM were 9, 4, and 1, respectively, as detailed in Table 7.

Table 7

Data mining algorithms in relation to study purposes

Sn	Algorithms	PSP	P&CSP	FISP	CDMM	Total	Percentage (%)
1	NB	9	6	3	3	21	21
2	DT	11	2	3	5	21	21
3	ANN	3	4	1	3	11	11

4	BP	0	1	0	0	1	1
5	BN	2	0	1	0	3	3
6	LSTM	0	1	0	0	1	1
7	SVM	5	1	2	3	11	11
8	K-NN	3	1	2	3	9	9
9	RF	6	2	1	2	11	11
10	CL	1	1	0	2	4	4
11	MP	2	1	0	1	4	4
12	LR	1	2	0	0	3	3
	Total	43	22	13	22	100	100

3.4 Most Used Data Mining Algorithms Based on Evaluation Metrics

Academic prediction studies use a variety of evaluation metrics, such as recall, F1-score, accuracy, and precision, to measure student performance [53]. Over the others, the accuracy evaluation performance metric is more preferred [54]. About 19 (68%) out of the 28 reviewed studies have used the accuracy metrics to evaluate the effectiveness of a predictive models. As a result, accuracy was used in this section to identify the most commonly used EDM methods based on evaluation metrics.

In the 19 articles examined, DT had the highest accuracy of 98% and the lowest of 65% with a count of 11, while NB had the highest accuracy of 98% and the lowest of 63% with a count of 13. ANN got a score of 5, with the highest accuracy of 89% and the lowest accuracy of 59%. SVM received a count of 8, with the highest accuracy of 99% and the lowest accuracy of 69%, and KNN received a count of 6, with the highest accuracy of 99% and the lowest score of 63%. RF counted 6, with the highest accuracy of 99% and the lowest score of 76%. BN count 4 had the highest accuracy of 99% and the lowest of 59%, while MP count 3 had the highest accuracy of 87% and the lowest of 72%. BP count 1 has an accuracy of 88%, while CL, LR, and LTSM do not count in any study, as shown in Table 8. The accuracy count for each algorithm is shown in Figure 5.

Table 8

Data Mining algorithms in relation to accuracy evaluation metrics in percentage

Sn	Reference	NB	DT	RF	KNN	ANN	SVM	BP	BN	M P
1	[26]	88	93			59				
2	[29]	98			99		99		99	
3	[30]						83	88		
4	[34]	75	88			83	80			
5	[39]		74	99					65	
6	[44]									76
7	[45]	87								
8	[46]	84	79		85				69	72
9	[48]		72							
10	[50]	75	67				97			
11	[51]	84	68	89	63	74				
12	[38]	69					69			87
13	[27]	98		99	99		99		59	
14	[28]	75	65	82		62	75			
15	[16]		67							
16	[40]	94	98		83	89	83			
17	[55]	65		84	75					
18	[56]	63	80							
19	[49]			76						

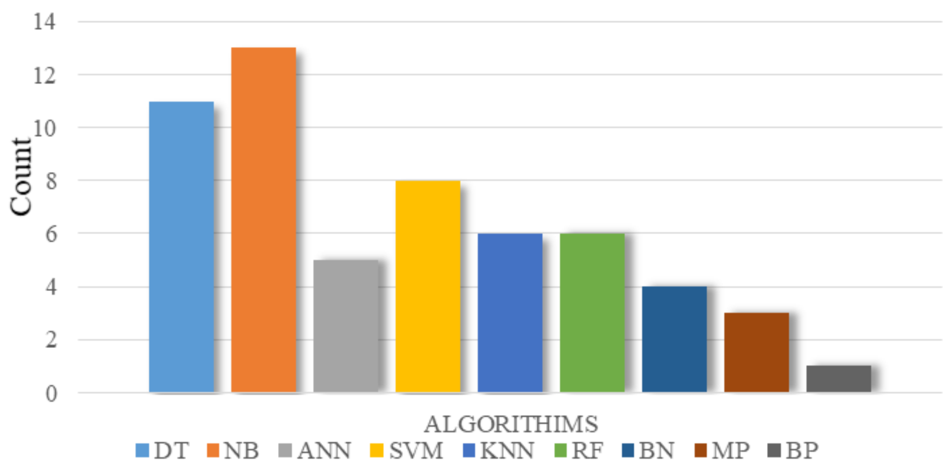

Figure 5. Accuracy counts of the data mining algorithms

Considering the average mean of the highest and lowest accuracy of every algorithm, the RF outperformed other algorithms by a score of 87.5%, followed by the SVM with a score of 84%. Other algorithms like DT, K-NN, and NB scored 81.5%, 81%, and 80.5%, respectively. In addition, MP, BP, ANN, and BN scored 79.5%, 79%, 74%, and 44%, respectively. Furthermore, LR, CL, and LTSM did not count in any analyzed article.

4. DISCUSSION

The results depicted in Table 5 show that DT and NB have the highest counts, followed by SVM, K-NN, and ANN, and are similar to those of Batool et al.[41], who used a survey to discover the most commonly used algorithms. Batool's study revealed that in the last ten years, DT, SVM, ANN, K-NN, NB, and RF were the most commonly used algorithms to predict student performance. Furthermore, the current findings are consistent with the Shahiri et al.[32] study, which used a systematic review from 2002 to 2015 in Malaysia. Shahiri's study discovered ANN, DT, SVM, K-NN, and NB to be the most used classification algorithms to predict student performance.

The current study's findings based on the years of publications show that a lot of research papers were published in recent years, between 2014 and 2022. The findings concur with those of Albreiki et al.[12], who revealed that a large number of predictive studies were published between 2014 and 2021. The Albreiki et al. study outlines approximately 26 research publications that were published from 2014 to 2017 and approximately 34 that were published from 2018 to 2022. Furthermore, the findings from Hellas et al.[57] in San Diego, USA, showed a gradual increase in the number of students' performance in predictive studies each year. Hellas described about 43 articles published in 2014, 53 in 2015, 70 in 2016, and 75 in 2017. The Albreiki and Hellas trends of rapid increases in publications are consistent with the current study, which found increases of 0, 1, 10, and 17 from 2006 to 2009, 2010 to 2013, 2014 to 2017, and 2018 to 2022, respectively. These similar results could be explained based on the fact that Data Mining is a rapidly expanding discipline utilizing different types of data for the prediction of student performance, which supports decision-making.

The current study identified four major themes in terms of the study purpose: PSP, P&CSP, FISP, and CDMM. About five Data Mining algorithms, including NB, DT, ANN, SVM, and RF, have been identified across all four study purposes. This finding contradicts a study done by Chaka [58] in South Africa, which has three purposes. The first purpose was about an educational Data Mining algorithm to predict student performance; the second purpose concerned a survey of educational Data Mining techniques and tools. The last purpose focuses on the student's dropout. These disparities in the study purpose can be caused by the different classifiers used in these academic performance prediction studies.

Considering the average mean of the highest and lowest accuracy, RF in the current study outperformed other algorithms with an accuracy of 87.5%. The findings concur with the study by Hussain et al. [39] in India, who analyzed data using WEKA. Husain's study reveals that RF has a higher correct classifier of 99% compared with other classifiers. This is due to the fact that RF is a supervised ensemble machine learning approach for classification, regression, and other tasks. It functions by constructing a number of DTs on various subsets of datasets during the training period and producing an average output of the class, which is the mode of the classes of the individual trees, to increase predictive accuracy. As a result, the RF forecasts the result rather than relying on a single DT by using predictions depending on each tree that received the majority of votes.

Therefore, considering the findings from the most commonly used algorithms, algorithms based on the year of publication, the study purpose, and evaluation metrics, the study identified five Data Mining algorithms that are mostly used. These algorithms include RF, SVM, NB, K-NN, and DT. They are adequate and could be mainly employed to predict student performance in ICT in teacher training colleges.

5. CONCLUSION

The main goal of the study was to find the most commonly used Data Mining algorithms for predicting student teachers' performance in ICT. Online scholarly digital databases such as Science Direct, the ACM Digital Library, IEEE Explore, and Google Search were used to find 196 scholarly articles. After screening and an eligibility check, 28 articles were retained for content analysis. The analysis was based on the percentage of accuracy evaluation metrics, year of publication, study purpose, and the most used Data Mining algorithms. Five algorithms emerged as appropriate and popular among the Data Mining algorithms used to predict student performance. These algorithms include RF, SVM, NB, K-NN, and DT. It is recommended that these algorithms be utilized to build an ICT subject performance prediction model that can produce incredibly accurate results to aid in decision-making at colleges.

REFERENCES (TRANSLATED AND TRANSLITERATED)

[1] D. Kabakchieva, "Student performance prediction by using data mining classification algorithms," *Int. J. Comput. Sci. Manag. Res.*, vol. 1, no. 4, pp. 686–690, 2012.

[2] B. Andersson, E. N. Nfuka, S. Sumra, P. Uimonen, and A. Pain, "Evaluation of implementation of ICT in teachers' colleges project in Tanzania," Stockholm, Sweden, 2014. [Online]. Available: http://www.sida.se/publication

[3] M. Chirwa, "Access and use of internet in teaching and learning at two selected teachers' colleges in Tanzania," *Int. J. Educ. Dev. using Inf. Commun. Technol.*, vol. 14, no. 2, pp. 4–16, 2018.

[4] A. R. Kimaro and M. Mohamed, "Access and Use of ICT among Tutors in Public Teachers Training Colleges: A Case of Moshi Rural District Tanzania," *JAET No. 22, 2019*, p. 46, 2019.

[5] Y. K. Saheed, T. O. Oladele, A. O. Akanni, and W. M. Ibrahim, "Student performance prediction based on data mining classification techniques," *Niger. J. Technol.*, vol. 37, no. 4, p. 1087, 2018, doi: 10.4314/njt.v37i4.31.

[6] F. Joel and H. Mungwabi, "Factors constraining effective application of ICTs in teachers' training colleges in Tanzania," *African Journals Online (AJOL)*, vol. 11, no. 1, pp. 53–70, 2015, [Online].

Available: https://www.ajol.info/index.php/udslj/article/view/162190

[7] M. Selemani, V. A. Ndume, and D. H. Kisanga, "Integrating ICT in Tanzania secondary schools : Experience of Tanzania as it grows to second world economy," *Int. Acad. J. Educ. Lit.*, vol. 2, no. 5, pp. 81–95, 2021, doi: 10.47310/iajel.2021.v02i05.010.

[8] A. Rajab and R. Ramadhan, "Application of data mining techniques in students' performance prediction and analysis big data analytics view project artificial intelligent field view project," *Int. J. Acad. Inf. Syst. Res. ISSN*, vol. 3, no. 3, pp. 23–36, 2020, [Online]. Available: www.ijeais.org/ijaisr

[9] K. Govindasamy and T. Velmurugan, "A Study on classification and clustering data mining algorithms based on students Academic performance prediction," *Int. J. Control theory Appl.*, vol. 10, no. 23, pp. 147–160, 2017.

[10] N. Louis, "ICT teaching and learning environment in teachers training : The case of diploma student-teachers in Tanzania," The University of Dodoma, 2015. [Online]. Available: http://hdl.handle.net/20.500.12661/796

[11] K. Masalu, "Challenges preventing student teachers and tutors from exploiting ICT in their learning and teaching in teachers Colleges," Open University of Tanzania, 2018. [Online]. Available: http://repository.out.ac.tz/id/eprint/2185

[12] B. Albreiki, N. Zaki, and H. Alashwal, "A systematic literature review of student' performance prediction using machine learning techniques," *Educ. Sci.*, vol. 11, no. 9, 2021, doi: 10.3390/educsci11090552.

[13] I. Shingari, D. Kumar, and M. Khetan, "A review of applications of data mining techniques for prediction of students' performance in higher education," *J. Stat. Manag. Syst.*, vol. 20, no. 4, pp. 713–722, 2017, doi: 10.1080/09720510.2017.1395191.

[14] K. Haruehansapong and S. Rungraungsilp, "Educational data mining applied for predicting students' ICT literacy," *Int. J. Sci. Technol. Res.*, vol. 10, no. 02, pp. 339–344, 2021, [Online]. Available: www.ijstr.org

[15] P. Kavishe, "Performance prediction in mathematics using educational data mining techniques: A case of Mzumbe University in Tanzania," University of Dodoma, 2020. [Online]. Available: http://hdl.handle.net/20.500.12661/2696

[16] S. K. Yadav and S. Pal, "Data mining: A prediction for performance improvement of engineering students using classification," vol. 2, no. 2, pp. 51–56, 2012, [Online]. Available: http://arxiv.org/abs/1203.3832

[17] C. Angeli, S. K. Howard, J. Ma, J. Yang, and P. A. Kirschner, "Data mining in educational technology classroom research: Can it make a contribution?," *Comput. Educ.*, vol. 113, pp. 226–242, 2017, doi: 10.1016/j.compedu.2017.05.021.

[18] I. Shingari and D. Kumar, "Predicting student performance using classification data mining techniques," *Int. J. Comput. Sci. Eng.*, vol. 6, no. 7, pp. 43–48, 2018, doi: 10.26438/ijcse/v6i7.4348.

[19] R. Hasan, S. Palaniappan, A. R. A. Raziff, S. Mahmood, and K. U. Sarker, "Student academic performance prediction by using decision tree algorithm," in *2018 4th International Conference on Computer and Information Sciences: Revolutionising Digital Landscape for Sustainable Smart Society, ICCOINS*, 2018, pp. 1–5. doi: 10.1109/ICCOINS.2018.8510600.

[20] P. Cortez and A. Silva, "Using data mining to predict secondary school student performance," *15th Eur. Concurr. Eng. Conf. 2008, ECEC 2008 - 5th Futur. Bus. Technol. Conf. FUBUTEC 2008*, no. January 2008, pp. 5–12, 2008.

[21] E. B. Michael and O. Gold, "Prediction of students' performance in general mathematics at Wassce using decision tree," *Am. J. Eng. Res.*, vol. 7, no. 6, pp. 336–343, 2018, [Online]. Available: www.ajer.org

[22] M. A. Yehuala, "Application of data mining techniques for student success and failure frediction (The case of Debre _ Markos University)," *Int. J. Sci. Technol. Res.*, vol. 4, no. 4, pp. 91–94, 2015, [Online]. Available: www.ijstr.org

[23] M. J. Page *et al.*, "The PRISMA 2020 statement: An updated guideline for reporting systematic reviews," *Int. J. Surg.*, vol. 88, no. 1, 2021, doi: 10.1016/j.ijsu.2021.105906.

[24] R. Alamri and B. Alharbi, "Explainable student performance prediction models: A systematic review," *IEEE Access*, vol. 9, pp. 33132–33143, 2021, doi: 10.1109/ACCESS.2021.3061368.

[25] A. Namoun and A. Alshanqiti, "Predicting student performance using data mining and learning analytics techniques: A systematic literature review," *Appl. Sci.*, vol. 11, no. 1, p. 237, 2020, doi: 10.3390/app11010237.

[26] M. V. Amazona and A. A. Hernandez, "Modelling student performance using data mining techniques: Inputs for academic program development," *ACM Int. Conf. Proceeding Ser.*, pp. 36–40, 2019, doi: 10.1145/3330530.3330544.

[27] C. Ma, B. Yao, F. Ge, Y. Pan, and Y. Guo, "Improving prediction of student performance based on

multiple feature selection approaches," *ACM Int. Conf. Proceeding Ser.*, vol. Part F1319, pp. 36–41, 2017, doi: 10.1145/3141151.3141160.
[28] H. Chanlekha and J. Niramitranon, "Student performance prediction model for early-identification of at-risk students in traditional classroom settings," *MEDES 2018 - 10th Int. Conf. Manag. Digit. Ecosyst.*, pp. 239–245, 2018, doi: 10.1145/3281375.3281403.
[29] X. Wu, Z. Shi, Y. Zhou, and H. Xing, "The application of three machine learning algorithms in student performance evaluation," in *In Proceedings of the 4th International Conference on Computer Science and Application Engineering*, 2020, pp. 1–5. doi: 10.1145/3424978.3425072.
[30] B. Sekeroglu, K. Dimililer, and K. Tuncal, "Student performance prediction and classification using machine learning algorithms," in *ACM International Conference Proceeding Series*, 2019, pp. 7–11. doi: 10.1145/3318396.3318419.
[31] S. Aydogdu, "Educational data mining studies in Turkey: A systematic review," *Turkish Online J. Distance Educ.*, vol. 21, no. 3, pp. 170–185, 2020, doi: 10.17718/TOJDE.762046.
[32] A. M. Shahiri, W. Husain, and N. A. Rashid, "A review on predicting student's performance using data mining techniques," *Procedia Comput. Sci.*, vol. 72, no. 1, pp. 414–422, 2015, doi: 10.1016/j.procs.2015.12.157.
[33] A. Khan and S. K. Ghosh, "Student performance analysis and prediction in classroom learning: A review of educational data mining studies," *Educ. Inf. Technol.*, vol. 26, no. 1, pp. 205–240, 2021, doi: 10.1007/s10639-020-10230-3.
[34] A. Ashraf, S. Anwer, and M. Gufran, "A comparative study of predicting student's performance by use of data mining techniques," pp. 122–136, 2018.
[35] M. H. bin Roslan and C. J. Chen, "Educational data mining for student performance prediction: A systematic literature review (2015-2021)," *Int. J. Emerg. Technol. Learn.*, vol. 17, no. 5, pp. 147–179, 2022, doi: 10.3991/ijet.v17i05.27685.
[36] A. Yusuf and A. Lawan, "Prediction of students' academic performance using educational datamining technique: Literature review," 2018, [Online]. Available: https://www.academia.edu/34009565/prediction_of_students_academic_performance_using_educational_datamining_technique_literature_review.
[37] T. Thilagaraj and N. Sengottaiyan, "A Review of educational data mining in higher education system," in *Proceedings of the Second International Conference on Research in Intelligent and Computing in Engineering*, 2017, vol. 10, pp. 349–358. doi: 10.15439/2017r87.
[38] M. Kumar, S. Shambhu, and P. Aggarwal, "Recognition of slow learners using classification data mining techniques," *Imp. J. Interdiscip. Res.*, vol. 2, no. 12, pp. 1–6, 2016, [Online]. Available: http://www.onlinejournal.in
[39] S. Hussain, N. A. Dahan, F. M. Ba-Alwib, and N. Ribata, "Educational data mining and analysis of students' academic performance using WEKA," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 9, no. 2, pp. 447–459, 2018, doi: 10.11591/ijeecs.v9.i2.pp447-459.
[40] V. Vijayalakshmi, K. Venkatachalapathy, and V. Ohmprakash, "A comparison of classification techniques on prediction of student performance in educational data mining," in *2017 International Conference On Intelligent Computing And Control(I2c2)*, 2017, no. 1, pp. 132–136.
[41] S. Batool, J. Rashid, M. Wasif, and N. Jungeun, *Educational data mining to predict students ' academic performance : A survey study*, no. 0123456789. Springer US, 2022. doi: 10.1007/s10639-022-11152-y.
[42] A. Saa, "Educational data mining and students' performance rediction," *Int. J. Adv. Comput. Sci. Appl.*, vol. 7, no. 5, pp. 212–220, 2016, doi: 10.14569/ijacsa.2016.070531.
[43] G. Ramaswami, T. Susnjak, and A. Mathrani, "On Developing Generic Models for Predicting Student Outcomes in Educational Data Mining," *Big Data Cogn. Comput.*, vol. 6, no. 1, 2022, doi: 10.3390/bdcc6010006.
[44] R. Raju, N. Kalaiselvi, M. Aathiqa Sulthana, I. Divya, and A. Selvarani, "Educational data mining: A comprehensive study," in *2020 International Conference on System, Computation, Automation and Networking, ICSCAN 2020*, 2020, pp. 1–5. doi: 10.1109/ICSCAN49426.2020.9262399.
[45] S. S. Athani, S. A. Kodli, M. N. Banavasi, and P. G. S. Hiremath, "Student academic performance and social behavior predictor using data mining techniques," *Proceeding - IEEE Int. Conf. Comput. Commun. Autom. ICCCA 2017*, vol. 2017-Janua, pp. 170–174, 2017, doi: 10.1109/CCAA.2017.8229794.
[46] N. Ndou, R. Ajoodha, and A. Jadhav, "Educational data-mining to determine student success at higher education institutions," in *2020 2nd International Multidisciplinary Information Technology and Engineering Conference, IMITEC 2020*, 2020, pp. 1–8. doi: 10.1109/IMITEC50163.2020.9334139.
[47] J. López-Zambrano, J. A. L. Torralbo, and C. Romero, "Early prediction of student learning performance through data mining: A systematic review," *Psicothema*, vol. 33, no. 3, pp. 456–465, 2021, doi: 10.7334/psicothema2021.62.

[48] L. Indi, P. Aji, U. A. Yogyakarta, A. Sunyoto, and U. A. Yogyakarta, "An implementation of C4 . 5 classification algorithm to analyze student's performance," in *Published in: 2020 3rd International Conference on Information and Communications Technology (ICOIACT)*, 2021, pp. 5–9. doi: 10.1109/ICOIACT50329.2020.9332088.

[49] K. Parmar, D. Vaghela, and P. Sharma, "Performance prediction of students using distributed data mining," *ICIIECS 2015 - 2015 IEEE Int. Conf. Innov. Information, Embed. Commun. Syst.*, no. Ddm, 2015, doi: 10.1109/ICIIECS.2015.7192860.

[50] H. Pallathadka, A. Wenda, E. Ramirez-Asís, M. Asís-López, J. Flores-Albornoz, and K. Phasinam, "Classification and prediction of student performance data using various machine learning algorithms," *Mater. Today Proc.*, 2021, doi: 10.1016/j.matpr.2021.07.382.

[51] R. Asif, A. Merceron, S. A. Ali, and N. G. Haider, "Analyzing undergraduate students' performance using educational data mining," *Comput. Educ.*, vol. 113, pp. 177–194, 2017, doi: 10.1016/j.compedu.2017.05.007.

[52] N. M. Masanja and H. Mkumbo, "The Application of Open Source Artificial Intelligence as an Approach to Frugal Innovation in Tanzania," *Int. J. Res. Innov. Appl. Sci.* |, vol. 5, no. 3, pp. 2454–6194, 2020, [Online]. Available: www.rsisinternational.org

[53] A. Khan, S. Ghosh, and S. K. Ghosh, "Measuring domain knowledge for early prediction of student performance: A semantic approach," *Proc. 2020 IEEE Int. Conf. Teaching, Assessment, Learn. Eng. TALE 2020*, pp. 444–451, 2020, doi: 10.1109/TALE48869.2020.9368439.

[54] R. Jidagam and N. Rizk, "Evaluation of predictive data mining algorithms in student academic performance," in *INTED2016 Proceedings*, 2016, vol. 1, no. 1, pp. 6314–6324. doi: 10.21125/inted.2016.0487.

[55] G. Ramaswami, T. Susnjak, A. Mathrani, J. Lim, and P. Garcia, "Using educational data mining techniques to increase the prediction accuracy of student academic performance," *Inf. Learn. Sci.*, vol. 120, no. 7–8, pp. 451–467, 2019, doi: 10.1108/ILS-03-2019-0017.

[56] S. Lemm, B. Blankertz, T. Dickhaus, and K. R. Müller, "Introduction to machine learning for brain imaging," *Neuroimage*, vol. 56, no. 2, pp. 387–399, 2011, doi: 10.1016/j.neuroimage.2010.11.004.

[57] A. Hellas *et al.*, "Predicting academic performance: A systematic literature review," in *Annual Conference on Innovation and Technology in Computer Science Education, ITiCSE*, 2018, no. 1, pp. 175–199. doi: 10.1145/3293881.3295783.

[58] C. Chaka, "Educational data mining, student academic performance prediction, prediction methods, algorithms and tools: An overview of reviews," pp. 1–17, 2021, doi: 10.20944/preprints202108.0345.v1.

Text of the article was accepted by Editorial Team 09.04.2023

АЛГОРИТМИ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ДАНИХ ДЛЯ ПРОГНОЗУВАННЯ УСПІШНОСТІ СТУДЕНТІВ ПЕДАГОГІЧНИХ СПЕЦІАЛЬНОСТЕЙ З ВОЛОДІННЯ ІКТ: СИСТЕМАТИЧНИЙ ОГЛЯД ЛІТЕРАТУРИ

Джума Х. Шиндо
кандидат в магістри комп'ютерних наук, викладач коледжу, факультет комп'ютерних наук та інженерії
Університет Додоми, м. Додома, Танзанія
ORCID ID 0000-0003-4659-701X
shindostz@gmail.com

Мохамед М. Мджахіді
PhD з комп'ютерної інженерії, викладач, факультет комп'ютерних наук та інженерії
Університет Додоми, м. Додома, Танзанія
ORCID ID 0000-0002-3962-6997
mohamedi.mjahidi@udom.ac.tz.

Мохамед Д. Вазірі
PhD в галузі інформаційно-комунікаційних наук та інженерії
Університет Додоми, факультет комп'ютерних наук та інженерії, м. Додома, Танзанія
ORCID ID 0000-0003-2668-3279
mohamed.waziri2@udom.ac.tz

Анотація. Низький рівень володіння ІКТ студентів педагогічних коледжів ускладнює для більшості з них успішне використання ресурсів ІКТ у викладанні та навчанні. Коли викладачі можуть ефективно використовувати ресурси ІКТ, це дає їм можливість оновлювати свої знання за допомогою онлайн-навчання, а отже, підвищує загальну якість викладання і навчання. Такий позитивний результат спостерігається у разі ефективного використання ІКТ. Метою цієї статті є визначення відповідних алгоритмів інтелектуального аналізу даних для прогнозування успішності студентів-педагогів у галузі ІКТ. Методологією дослідження слугував систематичний огляд літератури, проводячи який ми керувалися положенням PRISMA-2020. Цей документ визначає, як чітко представити дані, детальний контрольний список, а також блок-схему, які окреслюють процес огляду. Станом на 6 листопада 2022 року з трьох електронних бібліотек було завантажено близько 196 наукових статей: Science Direct (38), ACM Digital Library (72), IEEE Xplore (51) та 35 з пошукової системи Google Scholar. Після відбору та перевірки на відповідність критеріям відбору було відібрано 28 наукових статей, які було проаналізовано за допомогою контент-аналізу з точки зору найбільш поширених алгоритмів, року публікації, цілей дослідження, показників точності та ефективності. Розглядаючи конкретні результати досліджень, представлені кількісно, було виявлено, що найбільш часто використовуваними алгоритмами інтелектуального аналізу даних є "Дерева рішень" та "Наївний Байєс", частка кожного з них становить 20,6%. Обрані для аналізу статті були опубліковані між 2014 і 2022 роками. З точки зору цілей дослідження велика кількість робіт була зосереджена на прогнозуванні успішності студентів. Крім того, близько 6 з 8 алгоритмів, що використовувались у попередніх дослідженнях, отримали 80% або вище за середнім відсотком найвищої та найнижчої метрик точності. Отже, враховуючи загальні висновки у дослідженні визначено п'ять алгоритмів інтелектуального аналізу даних, які є найбільш придатними та найчастіше використовуються для прогнозування успішності студентів-педагогів у галузі ІКТ. Це "Наївний Байєс", "Метод К-найближчого сусіда", "Машина опорних векторів", "Випадковий ліс" і "Дерево рішень". Результати цього дослідження допоможуть уряду, викладачам коледжів і студентам-педагогам приймати кращі рішення для підвищення ефективності використання ІКТ викладачами під час підготовки та підвищення кваліфікації.

Ключові слова: алгоритми інтелектуального аналізу даних; інтелектуальний аналіз освітніх даних; студенти-педагоги; педагогічні коледжі.