Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control
Journal homepage: http://kinetik.umm.ac.id
ISSN: 2503-2267
Vol. 7, No. 1, February, Pp. 35-44

Review of technique and algorithm for educational data mining: trend and challenge in games design

Ulfatun Nadifa*[1], Fitra Abdurracman Bachtiar[2], Ahmad Afif Supianto[3], Herman Tolle[4]
Universitas Brawijaya, Indonesia[1,2,3]
Lembaga Ilmu Pengetahuan Indonesia, Indonesia[4]

Article Info
Keywords:
Educational Data Mining, Game Design, Technique, Algorithm, Machine Learning, Student Performance

Article history:
Received: October 26, 2021
Accepted: January 15, 2022
Published: February 28, 2022

Cite:
U. Nadifa, F. A. Bachtiar, A. A. Supianto, and H. Tolle, "Review of Technique and Algorithm for Educational Data Mining: Trend and Challenge in Games Design", *KINETIK*, vol. 7, no. 1, pp. 35-44, Feb. 2022.
https://doi.org/10.22219/kinetik.v7i1.1349

*Corresponding author.
Ulfatun Nadifa
E-mail address:
Ulfatunnadifa4@gmail.com

Abstract
This study reviews techniques and algorithm models often used in the analysis of educational data mining. The review in this study is based on previous studies to provide researchers knowledge about trends and challenges analysis educational data mining in game design meaningful. However, there is a lot of games design developed without analysis educational data mining which then will not answer the student problem. The analysis needed periodic data and developing the game required actual student conditions, this is a combination inseparable. Determine research questions, search terms, and filtering for the selection and analysis of the article review. There are some student problems on analysis review, namely prediction student performance, student behavior, student at-riks, and student dropout. The number of Articles in the study was 33 with 21 articles of research and 12 of article review. The number of studies 8 with percent 38% used techniques confusion matric with 33% percent used algorithms Decision Tree in 7 of studies. The section in this study consists of techniques evaluation, model selection, outcome, subject, and algorithm method. Which are recommended techniques and algorithms for analysis educational data mining and in ideal game design to further research.

1. Introduction

Game design is a major concern in its development, a good game will depend on its design. But the best design of a game has not been able to increase interest of students who do not like learning with the media of gaming. This causes problems for researchers of game development that will not meet improvements, because the comparison of effort to developed a game and the user of the game has not had a big effect. For this reason, in this study, there is an idea to solve the problem by combining the field of data mining analysis in designing games. The ideas are based on an analysis of trends and challenges in previous researchers using statistical data analysis techniques, then the results can be proven.

Analysis data mining is finding patterns in data with the prediction and classification to provide information [1]. According to Iswarya, e*ducational data mining* (EDM) is new learning from the data mining and knowledge discovery databases (KDD) field that focuses on the usage patterns of mining and the use of learning knowledge from the education information systems, such as admissions systems, registration systems, course management systems, and other additional systems. Some of the results of analysis *educational data mining* from the system, that which educators take for the understanding problem of student, for improving acceptance standard by educational institutions, and taking of major of the student. However, the analysis will not answer the student problem of misconceptions, motivation, improved skills, and knowledge without further research such as developing tools learning, namely e-learning, games, e-books, system, and others. The analysis *educational data mining* needs periodic data from tools learning and system for meaningful analysis, and there are many techniques and algorithms used in the analysis [2]. The right of techniques and algorithms in analysis improved accuracy and time of computational in prediction and classification.

Games are one of the educational media expected to be the ideal idea for the education of field improvement. According to Atmaja *et al*., the rapid growth of educational games is hindered by various problems, including their development cost [3]. Development game needs preparation support game, design, content, subject, system, interface, e.g., and it takes time and a lot of costs. Nevertheless, the implementation is a field not helped student problems as the result of the research review shows improvement of student knowledge before and after using games is 8.7% percent. These things happen because the design of games is not analysis education data mining based.

In this study, we will review the article of research to know growth e*ducational data mining* of analysis, which then figure out to design game comparing. Trends explained about techniques and algorithms used often in analysis and challenges will recommend to researchers. The formulation of the problem in this study is how to help the problem of students and make meaningful design games with analysis of *Educational data mining* results. Here are some related studies.

Rambola *et al*. reviews techniques and algorithms in *Educational data mining* [4]. The purpose of their research is to provide empirical evidence from the years 2012 to 2017. But in this study, there is no mention of the source and amount of review literature and does not discuss trends and challenges. The algorithms Decision tree, Apriori, K-Means, and followed by Classification techniques, Association, and Clustering are the most popular in this study. It should be attention that techniques and algorithms are very influential on research, so we need consider to selecting.

In addition, Awad *et al*. also conducted a systematic literature review discussing Mining in Educational data: Review and future directions [5]. The purpose of this study is to examine data mining in educational trends and Machine Learning used. The reviewed literature previous research from the years 2016 to 2019. However, this research does not include the database and number of the *Article* used. From this, we know the right choice of techniques will be the effect of the resulting research.

Fairuzabadi *et* al. review learning support factors on mathematic games, the purpose of this study is to find trends subject, game type, and evaluation to develop a game [6]. Several online databases use, namely: BioMed Central, ASSIA, ERIC, IEEE, Ingenta, and Science Direct. The different studies are the factor design games with analysis EDM, comparing of design will show in this study to see how efficiency from one of.

Therefore, this study attempts to provide empirical evidence about trends and challenges in *Educational data mining* analysis that affect the design game. Collecting *Article*s in this study from several databases; ERIC, Microsoft Academic, NATURE, Springer Link, Wiley Online Library, IJSR, Taylor and Francis, KINETIK, IJEECS, IJEE, and IEEE. We got 33 of several *Article*s at the search terms, passing by screening, review, and filtering. In this study, we found that the best algorithms are not always suitable for new research. This study explained section, for the first is the purpose of research trends and challenges the next section about research method and the last section is a result with the conclusion that is discussed answers our research question.

2. Research Method

In this research, we used and determine a method to handle bias or Non-random error caused analysis process. The analysis process is a step to get an answer to the research question. To handle bias by the error collection *Article,* with determining some of the tier answers by Search Terms method, and Research Question solved bias caused over or underestimation from purposed research. Limitation of source *Article* also determines in filtering method with used only reputation *Article*. This method does by systematic to reduce bias in this research and get evidence trends and challenges for support increasing analysis EDM and apply in the ideal design of the game.

Figure 1. Research Method

In Figure 1 is the step of the method of research to achieve objectives following research method consists of:

1. Determine the research boundaries over the research question (RQ). The formulation of RQ is under the point of research purposes. We have four research questions for this study it will explain in the next section.
2. The search terms we created to answer the problem formulation. That's helped to easy in selection *Article* and article we have. Divide three search terms over the question of algorithms, subject, and outcome.
3. Determine the online database, namely the source, keywords, limited. The criteria of the selected database are credible which is recognized, and limited on years of publication.
4. Screening *Article*s from an online database a done in two stages. The first stage we call filtering B1 based on abstract, title, and years, and filtering B2 analysis content of the *Article* to answer the research question.
5. After obtaining the relevant articles and *Article*s for this research, then analysis of data calculating of a percentage data. The result of the data analysis percentage is in the table to explain the review.

2.1 Research Question

This section contains questions related to problem formulation and the purpose of the study. Four research questions will answer a suitable *Article*. Namely of research question: (RQ1) *What algorithms method was that used in Educational data mining?* (RQ2) *What is the technique selection used of algorithms method in analysis Educational data mining?* (RQ3) *What evaluation model in analysis Educational data mining?* (RQ4) *what is the difference in the game design based on analysis of Educational data mining?*

2.2 Search Terms and Online Database

Search terms are some terms that we create from the problem formulation, which defines several terms that help answer and select suitable *Article*s. In this study, we divide three search terms types of algorithms, subject, outcome, i.e., (Types of Algorithms (Machine Learning OR Feature Selection)) (A1), then the subject of learning terms: AND (Data

OR Data Mining OR EDM OR Techniques OR Methods) (A2). Then for possible results or effects: AND (Outcome of impact: (Student performance OR Student at-riks OR Behavior OR Student dropout OR Recommend system)) (A3), AND (Outcome of effect :(Techniques OR Evaluation OR Selection)) (A4) AND (Conclusion:(Tools Name OR Games OR Trends OR Challenges)) (A5).

The online database criteria used to search scientific papers is an international database of *Article* sites indexed by Scopus and others. The sites are as follows: ERIC, Microsoft Academic, NATURE, Springer Link, Wiley Online Library, IJSR, Taylor and Francis, KINETIK, IJEECS, IJEE, and IEEE. The limitation on publications in this study from 2016 to 2021.

2.3 Filter

This section is screened with certain restrictions to get the appropriate *Article*. The screening has two limited processes. Here are the limits in screening.

1. Title and Abstract (B1). In the first limit, screenings are the title and abstract of the research suitable with problem formulation and requirements of years. Each research title and abstract must relate to *Educational data mining* OR Techniques selection OR Algorithms machine learning OR Design game OR Tools system.
2. The answer to the research question (B2). The second filter is to review the content of an *Article* with the Research question (RQ). The result of this step is finding data about the content of the *Article*.

2.4 Data Analysis

Where $x_d = \sum_{i=i_n} i.1(i_1, i_2, i_3, i_n \dots), x_{d_i} = \sum_{d=1} d.d_1 + d.d_n \dots (d_1, d_2, d_3, d_n \dots)$. That's Equation 1 used to calculate the percent of all of the *Article*s. Analysis of percent research *Article* used the Equation 2 with a maximum number of $x_{d_i} = 21$.

$$\% = \sum \frac{x_d}{x_{d_i}} \times 100\% \quad (1)$$

$$\% = \sum \frac{x_d}{x_{d_{i=21}}} \times 100\% \quad (2)$$

3. Results and Discussion

Table 1. Article Search Results Based on Search Terms

Database	*A1*	*A2*	*A3*	*A4*	*A5*	Percentage
IEEE	24	24	23	23	23	45%
ERIC	7	10	9	9	8	14%
Taylor And Francis	4	6	5	4	3	6%
IJSR	5	5	5	5	5	10%
Microsoft Academic	3	3	3	3	2	4%
Wiley Online Library	2	2	2	2	2	4%
Springer Link	3	3	3	3	3	6%
NATURE	2	2	2	2	2	4%
KINETIK	2	2	2	2	2	4%
IJEECS	1	1	1	1	1	2%
IJEE	1	1	1	1	1	2%
Number	54	59	56	55	**52**	100%

This Table 1 section will explain the result analysis of the review in this study. The result presents in the table and figures with information support on previous research.

3.1 Papers identified by search terms

Table 1 presents the result identified by search terms, and the number included is 52 *Article*s. The *Article* has does do a review under search terms, namely: A1(*Type Algorithms*), A2 (*Subject*), A3-A4(*Outcome*), A5(*Conclusion*),

and then will be screened on the next step. The following are the results of the search terms and the percentage of *Article*s and articles obtained in each database. In Table 1, it can see IEEE has first ranked with 23 studies (45%). Then *Articles* in the database IJEECS and IJEE successively occupied the fewest position with a number percentage of 2% and the total number 1 of *Article*s.

3.2 Filtered Papers

The next stage is doing screening to get *Article*s that are suitable for this study. The screening process was carried out with two steps, namely filtering B1 and B2. The first stage of screening is by title and abstract of an *Article* (B1), the second stage is reviewing the contents and answering RQ that we determine before (B2). The result of this section is 33 *Article*s with 21 research *Article*s and 12 for *Article* review. Table 2 contains data number of *Article*s and their percentages. The number of *Article*s decreased at this stage from 52 to 33. The high percentage of databases is IEEE with 41% percent. That's show information publication *Article* about analysis *Educational data mining* a lot found in database IEEE.

Table 2. The Number of Filtered Articles

Database	Search Terms	*B1*	*B2*	*Article* Review	*Article* Research	Percentage
IEEE	23	20	14	3	11	41%
ERIC	8	8	3	0	3	9%
Taylor And Francis	3	6	2	1	1	6%
IJSR	5	3	4	2	2	12%
Microsoft Academic	2	3	2	0	2	6%
Wiley Online Library	2	2	1	1	0	3%
Springer Link	3	3	2	1	1	6%
NATURE	2	2	1	0	1	3%
KINETIK	2	2	2	2	0	6%
IJEECS	1	1	1	1	0	6%
IJEE	1	1	1	1	0	3%
Number	52	51	**33**	**12**	**21**	100%

3.3 Subject Area

Table 3. Subject Areas Machine Learning Model

Machine Learning Model	Research	Database	Number of *Article*	Percentage
Classification/ Prediction/ sequential patterns	[7], [8], [9], [10], [11], [12], [13], [14], [15], [16], [17], [18], [19], [20], [21], [22], [23], [24], [25], [26], [27]	IEEE ERIC Taylor and Francis NATURE IJSR IJEECS Microsoft Academic	21	100%
Clustering	[27], [25]	IEEE ERIC NATURE	2	10%
Association	[8], [12], [24], [28]	IEEE IJSR	3	14%

Parmar *et al*., there are several major data mining techniques: *association, classification, clustering, prediction,* and *sequential patterns* [29]. The subject area of this research is the machine learning model. Kurdi *et al*. [30], also says data mining techniques consist of *predictive* and *descriptive*. Predictive techniques include *classification*,

prediction, and *time series analysis*. Then the *descriptive* Data Mining techniques are *Association, Clustering*, and *Summarization*. Table 3 shows techniques, number of research, and database. The review classification and prediction model have a high percentage till 100% with 21 studies. This is following Anoopkumar *et al*. [2] classification is still the subject with the most from years 2005 to 2015.

3.4 Outcomes of Study

Based on the results of a review from problem formulation is analysis prediction problem of the student. Table 4 is the outcome of the study on previous research. These 15 studies are focused on predicting student performance with the highest percentage of 71% and course institute with a percentage studies 19%, the least of percentage research is GPA, Student Dropout and Behavior with 5%.

Table 4. Outcome of Analysis EDM

Outcome Analysis EDM	Research	Database	Number of *Articles*	Percentage
Student Performance	[9], [10], [11], [12], [27], [15], [16], [17], [18], [19], [23], [24], [25], [26], [28]	IEEE ERIC IJSR IJEECS Microsoft Academic	15	71%
Student Dropout	[10]	IEEE	1	5%
Student At-Riks,GPA/SAP	[7], [13], [20]	IEEE ERIC	3	14%
Institution OR Courses	[8], [12], [14], [21],	IEEE Taylor and Francis	4	19%
Behavior	[22]	IEEE NATURE	1	5%

3.5 Articles that answer the research questions

Review of 33 *Article*s that we have been by answering research questions that we determine before. The following is a discussion of the answer.

RQ.1 *What algorithms method was that used in Educational data mining?*

The answer (RQ1) is the Algorithms type used in the analysis of *educational data mining*. Table 5 shows algorithms type with data analysis. The algorithm Decision tree is the most algorithm used in 7 studies with a percentage of 33%, 4 of them became the algorithm with the best accuracy level of 1, and 3 other studies became the best accuracy of 2 and 3. Radha K.R [13] has the results of an algorithm's Decision tree are popular.

Table 5. Technique Selection of Algorithms Model

Technique Selection	Research	Database	Numbero of *Article*	Percentage
Confusion Matric	[7], [8], [18], [19], [20], [22], [25], [26]	IEEE ERIC NATURE Springer Link Wiley Online Library	8	38%
Fitur Selection/ Sampling/Subset	[9], [14], [17], [27], [28]	IEEE IJEECS	5	23%
Classification Model WEKA	[11], [12], [21]	IEEE Taylor and Francis	3	14%
Other	[15], [16], [23], [24]	IEEE IJSR	4	19%

RQ.2 *What is the technique selection used of algorithms method in analysis Educational data mining?*

Results of the review show there are discussed techniques selection algorithm. The effect is improving accuracy in prediction and classification. The confusion Matric figures out the threshold to select the best accuracy algorithm

method. This technique can be used with data balance or unbalance considerations. Table 6 shows Confusion matrix has a high percentage of 38% in 8 studies.

RQ.3 *What evaluation model in analysis Educational data mining?*

Analysis *Educational data mining* needed an evaluation model to confirm the extent of the research. The evaluation model follows the method and techniques used. Combination techniques or algorithms needed exhaustive evaluation. In Table 6 shows the popular evaluation model is accuracy evaluation with 80% percent and 17 studies from 21.

Table 6. Evaluation of analysis EDM

Evaluation	Research	Database	Number of *Article*	Percentage
Accuracy	[7], [8], [10], [11], [12], [13], [14], [15], [16], [18], [19], [20], [21], [23], [24], [25], [26]	IEEE ERIC Taylor and Francis IJSR Springer Link Wiley Online Library	17	80%
Precision	[9], [17]	IEEE	2	10%
G-mean/ Iteration	[14][27]	IEEE IJEECS	2	10%

RQ.4 *What is the difference in the game design based on analysis of Educational data mining?*

Indispensable of further application from analysis education data mining, namely developing system student admissions, tools media learning as games, e-learning, e-book,.i.e. and other. This section focuses on different games design with an analysis of *Educational data mining*. According to Vandercruysse *et al*., every educational game consists of game content, e.g., game mechanics and interfaces, and learning content, e.g., lessons and assessments [31]. Fairuzabadi *et* al. support factor of design games education is content, subject, type & element of games, and evaluation method [6]. Nevertheless, this study explains the difference in support factor game design. Table 7 is differences in the game design based on analysis of *Educational data mining*.

Table 7. Different of Game Design

Game Design	Content	Interface	Element	Outcome
With EDM	Based on Problem Solved: (created to look for student learning patterns)	Layout/ Rules/ Description game: (Adapted to answer student problems)	Type Games: (puzzle or simulation, for detection of early and advanced misconceptions e.g.,)	Data and focus on solving the student problem
Without EDM	Based on support factors, namely: trends type game	Adapted to the type of game and made it attractive for students	Type Games: (interactive game more interesting, e.g., RTS, FPS, Life Simulation)	Satisfaction (improved: motivated, Data)

From Table 7 the developing game design with EDM focused on solving student problems to get data and patterns. Predictions and classifications in analysis *Educational data mining* will be more meaningful with periodic data games. Furthermore, it was used to improve the design of the game and develop a system from analysis *Educational data mining*. Generally, game design without EDM is not considered to student problem. That is the reason for insignificant improvements in students' knowledge and skills even though using the game, this is following in the review, Akhrian *et al*. discussed trends that impact and affect using games is to improve motivation learning [32]. Analysis EDM is useful to improve in game design and data periodic game make analysis meaningful. This area needs a combination to support each other in future research.

3.6 Trends and challenges

To achieve the maximum accuracy of algorithm method in EDM research, we must know the supporting factors. Knowing Trends and Challenges are the most important factors in the latest research. This study discusses trends in EDM research and future research in challenges.

Baek, Clare Doleck, Tenzin [33] conducted a review of the 2015-2019 *Article*, the fewer studies data mining than learning analytics in the last of 5 years. There was an increase in the publication of *Educational data mining* even though the research was passive before. Such as research of developing learning media by the result of EDM is game, e-learning, and other tools for solve student problems. On the other hand, e-learning or games provide data that be analyzed with EDM more periodically.

Table 8. Algorithms Method in Analysis EDM

Algorithms Method	Research	Database	Number of *Article*	Percentage
Decision Tree	[8], [12], [27], [13], [18], [23], [25]	IEEE ERIC IJSR IJEECS Microsoft Academic	7	33%
Logistic Regression	[7], [17], [20], [22], [25]	IEEE ERIC NATURE Microsoft Academic	5	23%
SVM	[12], [14], [16], [19]	IEEE ERIC	4	19%
J48	[11], [12], [17], [21]	IEEE Taylor and Francis	4	19%
Random Forest	[9], [13], [25]	IEEE Microsoft Academic	3	14%
Naïve Bayesian/	[7], [10], [11], [23]	IEEE	4	19%
MLP	[7], [5]	IEEE	2	10%
MP and JRip	[12] [11]	IEEE	2	10%
NBTree	[15], [23]	IEEE	2	10%
AFSA	[21]	Taylor and Francis	1	5%
KNN and NN	[18] [24]	IJSR	2	10%
ANN	[26]	Microsoft Academic	1	5%
Clustering	[25]	IEEE	1	5%
Apriori	[8], [24], [28]	IEEE Microsoft Academic	3	14%

From Table 8, we can see that most research applies classification algorithms. However, machine learning has a lot of algorithms classification, and previous studies are limited in comparing the accuracy of each model without attention to proper technique selection, evaluation, and sample preparation techniques. Table 6 also describes the current popular model selection technique using the confusion matrix and classification using the WEKA tools. Sampling techniques are also widely used by above combining two techniques. The algorithms decision tree is the best one with high accuracy based on previous research. The challenge means in this study is variations in model selection techniques. The trends and challenges in this research are knowledge information researchers do not make the same mistakes and increase the probability of success and provide an overview of the limitations of current education data mining research.

Based on the main idea in this study, it can be known that it is necessary to apply after knowing the techniques and algorithms that are popular in data mining analysis. Next from the results of the analysis, develop a game that is in design based on it. And for further development of trends and challenges above is to build tools such as systems, games, and other tools.

4. Conclusion

The results of the reviews and discussions trends and challenges will explain the limited analysis of EDM and recommend technique, algorithm as well as game design. In the latest of 6 years, analysis EDM has increased in publication topics. Techniques selection, evaluation, and algorithms type this is a popular technique are control variables in research.

The result of the review shows an algorithm in analysis *Educational data mining* is essential to find patterns and prediction or classification data. Currently, the classification model is a trend in research because it shows positive results and a good evaluation, a popular algorithm is Decision Tree with 33% percent of 7 studies, and Logistic Regression with 23% percent of 5 studies, and techniques evaluation are popular is Confusion matric with 38% percent of 8 studies. Then the challenge is developing a new model to get the best accurate and less time computation in prediction or classification.

The trends and challenges are important areas to apply analysis *Educational data mining* is more meaningful. This research is limited to a review of analysis techniques and algorithms for Educational Data Mining, and in the future, we can implement this review to develop tools like games education in solving student problems with the best games design based on analysis data mining.

Notation

Data analysis in this study used percentages with the following notation:

x_d : The number of *Article* studies.
x_{d_i} : The number of the *Article* in on the whole databased.
i_n : The number of *Article*s in.
i : *Article* in is (i=1) and the *Article* out is (i=0).
d : Scores *Article* in (d=1).

References

[1] C. Iswarya, "Educational Data Mining Using Analysis Student Learning Process," vol. 10, no. 6, pp. 407–410, 2021.
[2] Z. R. Anoopkumar M, "A Review on Data Mining techniques and factors used in Educational Data Mining to predict student amelioration," *Proc. 2016 Int. Conf. Data Min. Adv. Comput. SAPIENCE 2016*, pp. 122–133, 2016. https://doi.org/10.1109/SAPIENCE.2016.7684113
[3] S. Pratama Wirya Atmaja, "Balancing Entertainment, Cost, and Educational Strength: A Design Framework for Medium-Coupling Educational Games," *Kinet. Game Technol. Inf. Syst. Comput. Network, Comput. Electron. Control*, vol. 4, pp. 27–40, 2021. https://doi.org/10.22219/kinetik.v6i1.1158
[4] S. H. Dr. Radha Krishna Rambola, Mrunmayee Inamke, "Literature review- techniques and algorithms used for various applications of educational data mining (EDM).," *2018 4th Int. Conf. Comput. Commun. Autom. ICCCA 2018*, pp. 1–4, 2018. https://doi.org/10.1109/CCAA.2018.8777556
[5] K. S. Said A. Salloum, Muhammad Alshurideh, Ashraf Elnagar, "Mining in Educational Data: Review and Future Directions," *AICV*, vol. 2, no. 1153, pp. 92–102, 2020. https://doi.org/10.1007/978-3-030-44289-7_9
[6] A. A. S. Ahmad Fairuzabadi, "An Overview Of Learning Support Factors On Mathematic Games," *Kinet. Game Technol. Inf. Syst. Comput. Network, Comput. Electron. Control*, vol. 4, no. 2, pp. 169–178, 2019. http://dx.doi.org/10.22219/kinetik.v4i2.761
[7] J. M. Steven Lehr, Hong Liu, Sean Klinglesmith, Alex Konyha Natalia Robaszewska, "Use Educational Data Mining to Predict Undergraduate Retention," no. 1, pp. 2–4, 2016. https://doi.org/10.1109/ICALT.2016.138
[8] P. Rojanavasu, "Educational data analytics using association rule mining and classification," *ECTI DAMT-NCON 2019 - 4th Int. Conf. Digit. Arts, Media Technol. 2nd ECTI North. Sect. Conf. Electr. Electron. Comput. Telecommun. Eng.*, pp. 142–145, 2019. https://doi.org/10.1109/ECTI-NCON.2019.8692274
[9] M. Zaffar, M. A. Hashmani, and K. S. Savita, "Performance analysis of feature selection algorithm for educational data mining," *2017 IEEE Conf. Big Data Anal. ICBDA 2017*, vol. 2018-Janua, pp. 7–12, 2018. https://doi.org/10.1109/ICBDAA.2017.8284099
[10] M. V. H. Ms.Tismy Devasia, Ms.Vinushree T P, "Prediction of Students Performance using Educational Data Mining," *IEEE*, vol. 1, no. 3, pp. 266–279, 2016. https://doi.org/10.1109/SAPIENCE.2016.7684167
[11] K. I. M. Ramaphosa, T. Zuva, and R. Kwuimi, "Educational Data Mining to Improve Learner Performance in Gauteng Primary Schools," *2018 Int. Conf. Adv. Big Data, Comput. Data Commun. Syst. icABCD 2018*, pp. 1–6, 2018. https://doi.org/10.1109/ICABCD.2018.8465478
[12] C. Jalota and R. Agrawal, "Analysis of Educational Data Mining using Classification," in *Proceedings of the International Conference on Machine Learning, Big Data, Cloud and Parallel Computing: Trends, Prespectives and Prospects, COMITCon 2019*, 2019, pp. 243–247. https://doi.org/10.1109/COMITCon.2019.8862214
[13] K. J. O. De Santos, A. G. Menezes, A. B. De Carvalho, and C. A. E. Montesco, "Supervised learning in the context of educational data mining to avoid university students dropout," *Proc. - IEEE 19th Int. Conf. Adv. Learn. Technol. ICALT 2019*, vol. 2161–377X, pp. 207–208, 2019. https://doi.org/10.1109/ICALT.2019.00068
[14] Y. Pristyanto, I. Pratama, and A. F. Nugraha, "Data level approach for imbalanced class handling on educational data mining multiclass classification," *2018 Int. Conf. Inf. Commun. Technol. ICOIACT 2018*, vol. 2018-Janua, pp. 310–314, 2018. https://doi.org/10.1109/ICOIACT.2018.8350792
[15] anoushka panwar tanvi gera, "AFSA: A comprehensive analysis of educational big data using the advanced feature selection algorithm," *IEEE*, vol. 7, pp. 5–9, 2021. https://doi.org/10.1109/ICACITE51222.2021.9404745
[16] B. Guo, R. Zhang, G. Xu, C. Shi, and L. Yang, "Predicting Students Performance in Educational Data Mining," *Proc. - 2015 Int. Symp. Educ. Technol. ISET 2015*, pp. 125–128, 2016. https://doi.org/10.1109/ISET.2015.33
[17] C. C. Kiu, "Data Mining Analysis on Student's Academic Performance through Exploration of Student's Background and Social Activities," *Proc. - 2018 4th Int. Conf. Adv. Comput. Commun. Autom. ICACCA 2018*, pp. 1–5, 2018. https://doi.org/10.1109/ICACCAF.2018.8776809

[18] M. Mimis, M. El Hajji, Y. Es-saady, A. Oueld Guejdi, H. Douzi, and D. Mammass, "A framework for smart academic guidance using educational data mining," *Educ. Inf. Technol.*, vol. 24, no. 2, pp. 1379–1393, 2019. https://doi.org/10.1007/s10639-018-9838-8
[19] T. Doleck, D. J. Lemay, R. B. Basnet, and P. Bazelais, "Predictive analytics in education: a comparison of deep learning frameworks," *Educ. Inf. Technol.*, vol. 25, no. 3, pp. 1951–1963, 2020. https://doi.org/10.1007/s10639-019-10068-4
[20] Q. Hu and H. Rangwala, "Towards Fair Educational Data Mining: A Case Study on Detecting At-risk Students," Hu, Q., Rangwala, H. (2020). Towar. Fair Educ. Data Min. A Case Study Detect. At-risk Students. Proc. 13th Int. Conf. Educ. Data Min., no. Edm, pp. 431–437, 2020.
[21] O. Moscoso-Zea, P. Saa, and S. Luján-Mora, "Evaluation of algorithms to predict graduation rate in higher education institutions by applying educational data mining," *Australas. J. Eng. Educ.*, vol. 24, no. 1, pp. 4–13, 2019. https://doi.org/10.1080/22054952.2019.1601063
[22] V. Lampos, J. Mintz, and X. Qu, "An artificial intelligence approach for selecting effective teacher communication strategies in autism education," *npj Sci. Learn.*, vol. 6, no. 1, 2021. https://doi.org/10.1038/s41539-021-00102-x
[23] M. Rajathi and R. Murugesh, "Comparative Study of Binary Classification Algorithms to Analyze the Students ' Performance on Virtual Machine," vol. 10, no. 4, pp. 2017–2021, 2021.
[24] A. Abdulrahman Al-Noshan, M. Abdullah Al-Hagery, H. Abdulaziz Al-Hodathi, and M. Sulaiman Al-Quraishi, "Performance Evaluation and Comparison of Classification Algorithms for Students at Qassim University," Int. J. Sci. Res., vol. 8, no. 11, pp. 1277–1282, 2018.
[25] R. Hasan, S. Palaniappan, S. Mahmood, A. Abbas, K. U. Sarker, and M. U. Sattar, "Predicting student performance in higher educational institutions using video learning analytics and data mining techniques," *Appl. Sci.*, vol. 10, no. 11, 2020. https://doi.org/10.3390/app10113894
[26] T. Elaf Abu Amrieh, I. Hamtini, and Aljarah, "mining educational data to predict student's academic performance using esemble methods," *Int. J. Database Theory Appl.*, vol. 9, no. 8, pp. 119–136, 2016. http://dx.doi.org/10.14257/ijdta.2016.9.8.13
[27] F. Jauhari and A. A. Supianto, "Building student's performance decision tree classifier using boosting algorithm," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 14, no. 3, pp. 1298–1304, 2019. http://doi.org/10.11591/ijeecs.v14.i3.pp1298-1304
[28] A. A. Yahya and A. Osman, "A data-mining-based approach to informed decision-making in engineering education," *Comput. Appl. Eng. Educ.*, vol. 27, no. 6, pp. 1402–1418, 2019. https://doi.org/10.1002/cae.22158
[29] K. P. H. Khan, "A Survey on Analysis the Students Mind in Different Area," Int. J. Sci. Res., vol. 7, no. 12, pp. 633–639, 2018.
[30] I. E. Moustafa M. Kurdi, Hatim Al-Khafagi, "Mining educational data to analyze students' behavior and performance," *Proc. 2018 JCCO Jt. Int. Conf. ICT Educ. Training, Int. Conf. Comput. Arab. Int. Conf. Geocomputing, JCCO TICET-ICCA-GECO 2018*, pp. 171–175, 2018. https://doi.org/10.1109/ICCA-TICET.2018.8726203
[31] S. Vandercruysse and J. Elen, "Instructional Techniques to Facilitate Learning and Motivation of Serious Games," *Instr. Tech. to Facil. Learn. Motiv. Serious Games*, pp. 17–35, 2017. https://doi.org/10.1007/978-3-319-39298-1_2
[32] A. A. Syahidi, A. A. Supianto, T. Hirashima, and H. Tolle, "Learning Models in Educational Game Interactions : A Review," vol. 3, no. June, pp. 11–29, 2021.
[33] C. Baek and T. Doleck, "Educational Data Mining versus Learning Analytics: A Review of Publications From 2015 to 2019," *Interact. Learn. Environ.*, vol. 0, no. 0, pp. 1–23, 2021. https://doi.org/10.1080/10494820.2021.1943689