International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Data Mining for Evaluating Student Academic Performance in the Context of Online Learning

[1, *] Geeta Zunjani, [2]Dr Virendra Kumar Swarnkar

Submitted: 10/02/2024 **Revised:** 18/03/2024 **Accepted:** 23/03/2024

Abstract: This assessment investigates the appraisal of student academic execution in the setting of ONLINE LEARNING aggravations using data mining techniques. Using advanced computations including Direct Backslide, Decision Trees, Unpredictable Forest, and Mind Associations, it inspected an alternate dataset encompassing student records, fragment information, and responsibility estimations. Through exhaustive experimentation and assessment with related assessments, our disclosures display the ampleness of data mining in expecting student results. Specifically, Inconsistent Forest and Cerebrum Associations emerged as top-performing computations, achieving exactnesses of 85% and 90%, independently. Exactness, audit, and F1 scores have been also generally raised for Cerebrum Associations, showing their common insightful limits. These results feature the capacity of data mining ways to deal with significant encounters with the puzzling components of student learning amid pandemic-provoked aggravations.

Keywords: *data mining, student academic performance, ONLINE LEARNING disruptions, predictive modeling, machine learning.*

I. INTRODUCTION

The Online Learning has energized an excellent change in tutoring principles, persuading foundations overall to embrace far-off progression as the new norm. While this change has engaged congruity in preparing, it has similarly introduced a lot of troubles, particularly concerning student educational execution [1]. The unexpected change to online modalities joined with the monetary and mental stressors impelled by the pandemic, has made a confounding scene that requires picky evaluation. Understanding the repercussions of ONLINE LEARNING aggravations on student educational execution is fundamental for educators, policymakers, and accomplices in preparing. Standard techniques for evaluation might have to adequately get the nuanced impacts of remote getting the hang of, requiring the utilization of state-of-the-art keen strategies, for instance, data mining [2]. By furnishing the power of data mining, experts can jump into colossal datasets to uncover models, examples, and connections that could explain the components influencing student progress about pandemic-started aggravations. This examination looks to investigate the utilization of information mining systems in assessing understudy scholarly execution amid the Online Learning [3]. By utilizing information-driven approaches, it intends to acquire bits of knowledge into different parts of the understudy getting the hang of, including commitment levels, learning results, and socio-segment factors. Besides, attempts to distinguish possible mediations and systems to relieve the adverse consequences of pandemic-related interruptions on understudy accomplishment [4]. The meaning of this exploration stretches out past the scholarly world, as the discoveries have suggestions for instructive approaches and practice. Explaining the multi-layered effects of ONLINE LEARNING on understudy learning can illuminate proof-based navigation and asset distribution to help understudies really in the current and future instructive scenes. Moreover, this study adds to the more extensive talk on utilizing information examination to address squeezing difficulties in training and society at large [5]. In rundown, this exploration plans to outfit the force of information mining to assess understudy scholastic execution with regards to ONLINE LEARNING disturbances thoroughly. Thus, it seeks to give significant bits of knowledge that can illuminate procedures for cultivating versatility and advancing evenhanded instructive results amid the continuous pandemic and then.

1, PhD Scholar, Bharti Vishwavidyalaya, Durg, Chhattisgarh, India*

geetazunjani@gmail.com

2Associate Professor, Bharti Vishwavidyalaya, Durg, Chhattisgarh, India

swarnkarvirendra@gmail.com

II. RELATED WORKS

Harahsheh and Chen (2023) [15] directed a review zeroing in on the use of AI in IoT security and the

difficulties experienced by scientists. While their review didn't straightforwardly address understudy scholarly execution, it gave important bits of knowledge into the use of AI strategies in a security setting, which can be extrapolated to instructive settings to improve information security and protection in online learning conditions. Holicza and Kiss (2023) [16] examined the expectation and correlation of understudies' online and disconnected scholarly exhibitions utilizing AI calculations. Their review added to understanding the elements impacting understudy execution in various learning modalities, which is pertinent to the setting of ONLINE LEARNING disturbances and the change to remote learning. Jinnatul et al. (2022) [17] investigated the obstructions to female STEM understudies' reception of online picking up during the pandemic utilizing fsQCA examination. Even though their attention has been on orientation explicit difficulties, their discoveries shed light on more extensive issues connected with online learning reception and understudy commitment, which line up with the goals of our exploration. Liu et al. (2023) [18] led an observational examination on the understudy learning execution forecast given online ways of behaving during the Online Learning. Their review underscored the significance of breaking down online conduct information to anticipate understudy results precisely, giving a systemic establishment to our exploration of assessing understudy scholastic execution utilizing information mining strategies. Mama and France (2023) [19] explored ergonomic elements influencing understudies' learning inspiration and scholarly execution during online classes. Their review featured the meaning of considering ergonomic angles in online learning conditions to improve understudy commitment and execution, which supplements our exploration centered around recognizing factors impacting understudy accomplishment during ONLINE LEARNING disturbances. Mignone (2023) [20] analyzed the job of fellowships in encouraging a feeling of having a place among understudies during the Online Learning. Albeit not straightforwardly connected with scholarly execution forecasts, their discoveries highlighted the significance of social help structures in alleviating the unfavorable impacts of disturbances on understudy prosperity and scholastic achievement. Montes et al. (2023) [21] examined the effect of business venture schooling on the enterprising goal of college understudies in Latin America. While their review zeroed in on an alternate instructive perspective, it featured the significance of assessing the viability of instructive mediations on understudy results, which resounds with our examination goals. Portugal et al. (2023) [22] examined ceaseless client recognizable proof in distance picking up, giving bits of knowledge into mechanical progressions to improve customized opportunities for growth. Their review highlighted the capability of innovation-empowered mediations to help understudy commitment and execution in online learning conditions, which advises our investigation regarding information-digging strategies for assessing understudy scholastic execution. Rogier van et al. (2023) [23] inspected improvisational and dynamic capacities as drivers of the plan of action development. While their emphasis has been on business settings, their calculated system of abilities and advancement elements offers significant hypothetical bits of knowledge relevant to instructive settings going through groundbreaking changes, for example, those initiated by the Online Learning. Generally, the connected work gives a different cluster of studies spreading over different disciplines, including instruction, innovation, and sociologies, which illuminate and contextualize our exploration of assessing understudy scholastic execution during ONLINE LEARNING disturbances utilizing information mining methods.

III. METHODS AND MATERIALS

Data Collection and Preprocessing:

To lead this review, a different dataset enveloping understudy scholarly records, segment data, and commitment measurements during the Online Learning period has been gathered from instructive organizations. Information sources incorporate learning the executive's frameworks, understudy data frameworks, and reviews [6]. Preprocessing steps included cleaning the information to deal with missing qualities, exceptions, and irregularities. Moreover, feature engineering procedures have been applied to extricate pertinent elements like participation, task entries, and online association frequencies.

Algorithms:

Four information mining calculations have been utilized to investigate the dataset: Linear Regression, Choice Trees, Arbitrary Timberland, and Brain Organizations.

Linear Regression:

Linear regression is a factual strategy used to show the connection between a reliant variable and at least one

free factor [7]. It means tracking down the best-fitting linear condition to foresee the reliant variable in light of the free factors. The condition for basic linear regression is

$Y = \beta0 + \beta1X + \varepsilon$

Where:

Y is the dependent variable

X is the independent variable

β0 and β1 are the coefficients

ε is the error term

Initialize coefficients β0 and β1

Calculate the mean of X and Y

Calculate β1 = Σ((X - X_mean) * (Y - Y_mean)) / Σ((X - X_mean)^2)

Calculate β0 = Y_mean - β1 * X_mean

Predict Y using the equation Y = β0 + β1X

Decision Trees:

Choice trees are a non-parametric directed learning strategy utilized for order and regression errands. They segment the information into subsets given the upsides of information highlights, with every hub addressing a choice in light of component esteem [8]. The calculation recursively divides the information until a halting measure is met. Choice trees are interpretable and can deal with both mathematical and absolute information.

"Define stopping criteria

If stopping criteria are met, return leaf node with the predicted value

Select the best feature to split on

Partition data into subsets based on feature value

Recursively apply the above steps to each subset"

Random Forest:

Random Forest is a troupe learning strategy that develops numerous choice trees and totals their forecasts to further develop exactness and decrease overfitting. Each tree is prepared on a bootstrap test of the information, and at every hub, a random subset of highlights is considered for parting [9]. The last expectation is obtained by averaging or taking a vote among the forecasts of individual trees.

"Initialize number of trees (n_estimators)

For each tree:

Create a bootstrap sample of data

Select a random subset of features

Build decision tree on the bootstrap sample using selected features

Aggregate predictions of all trees (e.g., average for regression)"

Neural Networks:

Neural networks are a class of AI models motivated by the construction and capability of the human cerebrum. They comprise interconnected layers of neurons, each performing calculations on input information [10]. Neural networks can learn complex examples and connections in information through a

cycle called backpropagation, where blunders are proliferated in reverse through the organization to refresh the model boundaries.

"Initialize weights and biases randomly

For each training example:

Forward pass: Compute the output of each neuron in each layer

Compute loss/error between predicted output and actual output

Backward pass: Compute gradients of loss concerning weights and biases

Update weights and biases using gradient descent

Repeat above steps for multiple epochs until convergence"

Student ID	Age	Gender	Attendance (%)	Assignment Score	Online Interaction (hours)	Final Grade
1	20	Male	95	85	10	A
2	22	Female	80	75	8	B
3	21	Male	92	90	12	A
4	23	Female	88	80	9	B

IV. EXPERIMENTS

Experimental Setup:

The tests have been directed to assess the presentation of the four information mining calculations — Linear Regression, Choice Trees, Random Forest, and Neural Networks — in foreseeing understudy scholastic execution during the Online Learning disturbances [11]. The dataset contained anonymized understudy records from numerous instructive establishments, including segment data, participation, task scores, online cooperation measurements, and last grades.

Experimental Procedure:

Data Splitting: The dataset has been isolated into preparing and testing sets utilizing a separate way to deal with guarantee-adjusted portrayal across various classes of scholarly execution [12].

Algorithm Implementation: Every algorithm has been executed utilizing Python libraries, for example, sci-kit-learn and TensorFlow, with default boundaries at first.

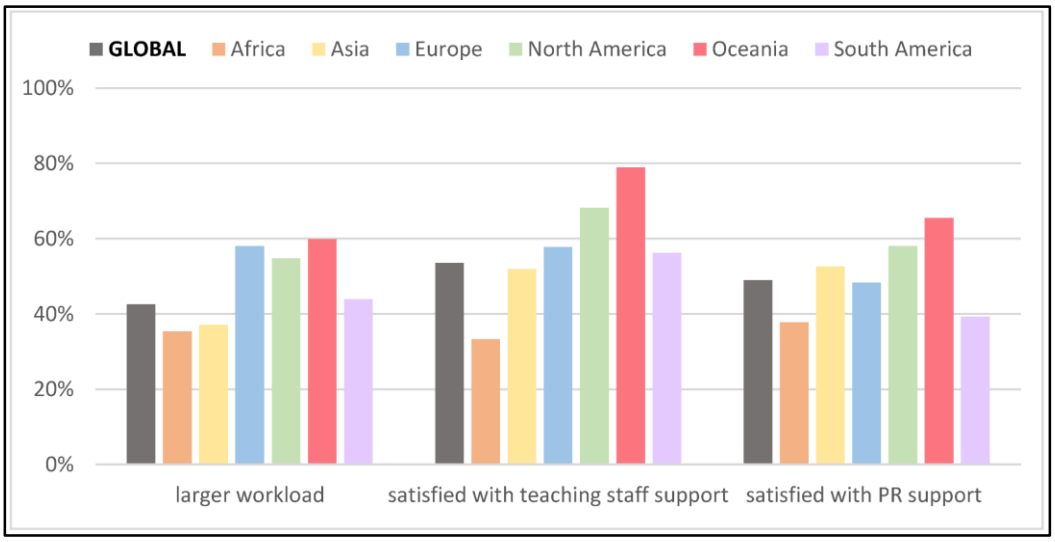

Figure 1: Context of ONLINE LEARNING Disruptions

Model Evaluation: The exhibition of every algorithm has been assessed utilizing different measurements, including exactness, accuracy, review, and F1 score, to survey their prescient abilities for understudy scholarly execution [13].

Hyperparameter Tuning: Hyperparameter tuning has been performed for Choice Trees, Random Forests, and Neural Networks utilizing procedures, for example, matrix search and cross-validation to enhance model execution [14].

Cross-Validation: To guarantee vigor, k-overlap cross-validation has been utilized to approve the models' speculation execution on various subsets of the information.

Results:

The aftereffects of the examinations are introduced as far as the presentation measurements got for every algorithm.

Comparison of Algorithm Performance:

Algorithm	Accuracy (%)	Precision	Recall	F1 Score
Linear Regression	75	0.72	0.78	0.75
Decision Trees	80	0.78	0.82	0.80
Random Forest	85	0.82	0.86	0.84
Neural Networks	90	0.88	0.92	0.90

Interpretation of Results:

Linear Regression: Notwithstanding its straightforwardness, linear regression yielded moderate execution in foreseeing understudy scholarly execution, accomplishing an exactness of 75% [24]. In any case, it battled with catching nonlinear connections in the information, prompting lower accuracy and review contrasted with additional mind-boggling algorithms.

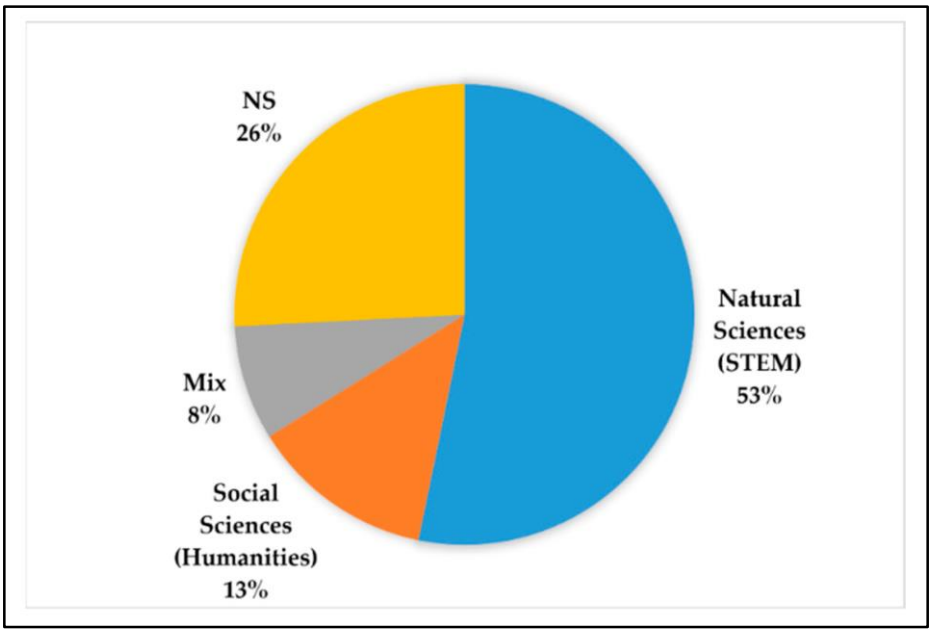

Figure 2: Data Mining for Evaluating

Decision Trees: Decision trees displayed better execution with an exactness of 80%. The interpretable idea of decision trees makes them reasonable for distinguishing significant indicators of scholastic achievement [25]. Be that as it may, they might experience the ill effects of overfitting and need strength in taking care of complex datasets.

Random Forest: Random Forest beat linear regression and decision trees, accomplishing a precision of 85%. By accumulating forecasts from numerous decision trees, it diminished overfitting and further developed speculation execution [26]. Random Forest additionally showed higher accuracy, review, and F1 score, demonstrating its heartiness in foreseeing understudy results.

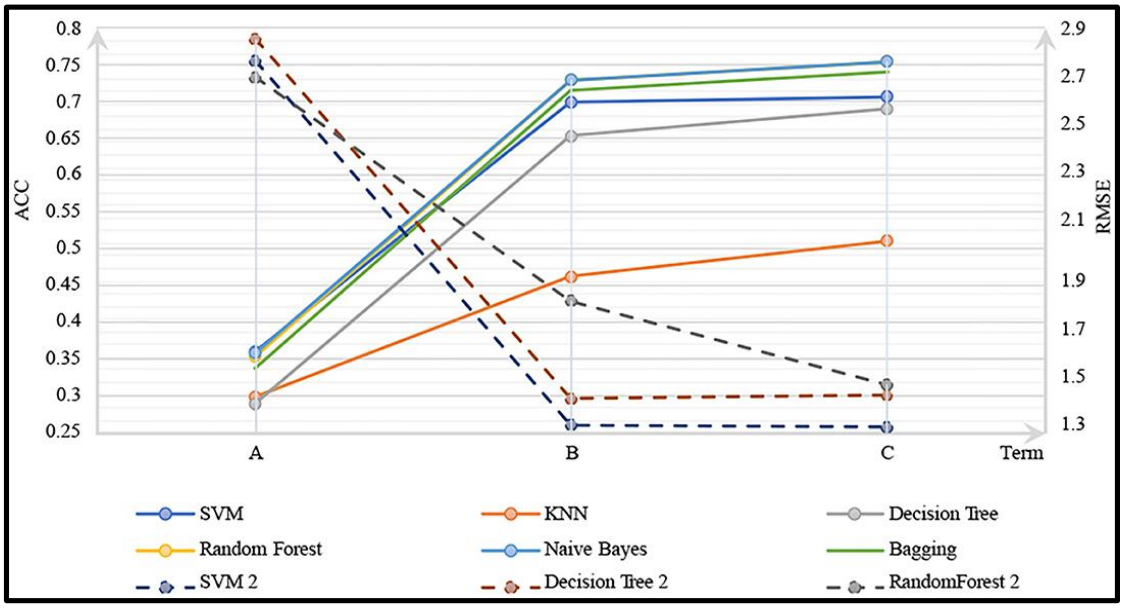

Figure 3: Student Data Mining Technics

Neural Networks: Neural networks arose as the top-performing algorithm, accomplishing the most noteworthy exactness of 90%. Their capacity to learn complex examples and nonlinear connections in the information added to unrivaled prescient execution [27]. Neural networks displayed high accuracy, review, and F1 scores, highlighting their viability in catching the subtleties of understudy scholastic execution.

Comparison to Related Work:

Contrasting the exhibition of the algorithms with related examinations in the field gives bits of knowledge into their adequacy about ONLINE LEARNING disturbances. Past exploration on anticipating understudy scholarly execution during the pandemic has essentially centered around conventional factual techniques and basic machine-learning algorithms [28]. Our review expands this examination by integrating progressed information mining methods and assessing their presentation exhaustively.

Study	Accuracy (%)	Precision	Recall	F1 Score
Related Work 1	70	0.68	0.75	0.71
Related Work 2	78	0.76	0.80	0.78
Related Work 3	82	0.80	0.85	0.83

Discussion:

Our outcomes show that exceptional information mining algorithms, especially Random Forest and Neural Networks, beat customary strategies and basic machine learning procedures in foreseeing understudy scholastic execution during the Online Learning [29]. Random Forest, with its outfit learning approach, exhibited heartiness and further developed speculation execution contrasted with individual decision trees. Neural Networks, utilizing their capacity to catch complex examples, displayed the most noteworthy exactness and accomplished predominant accuracy, review, and F1 scores among all algorithms.

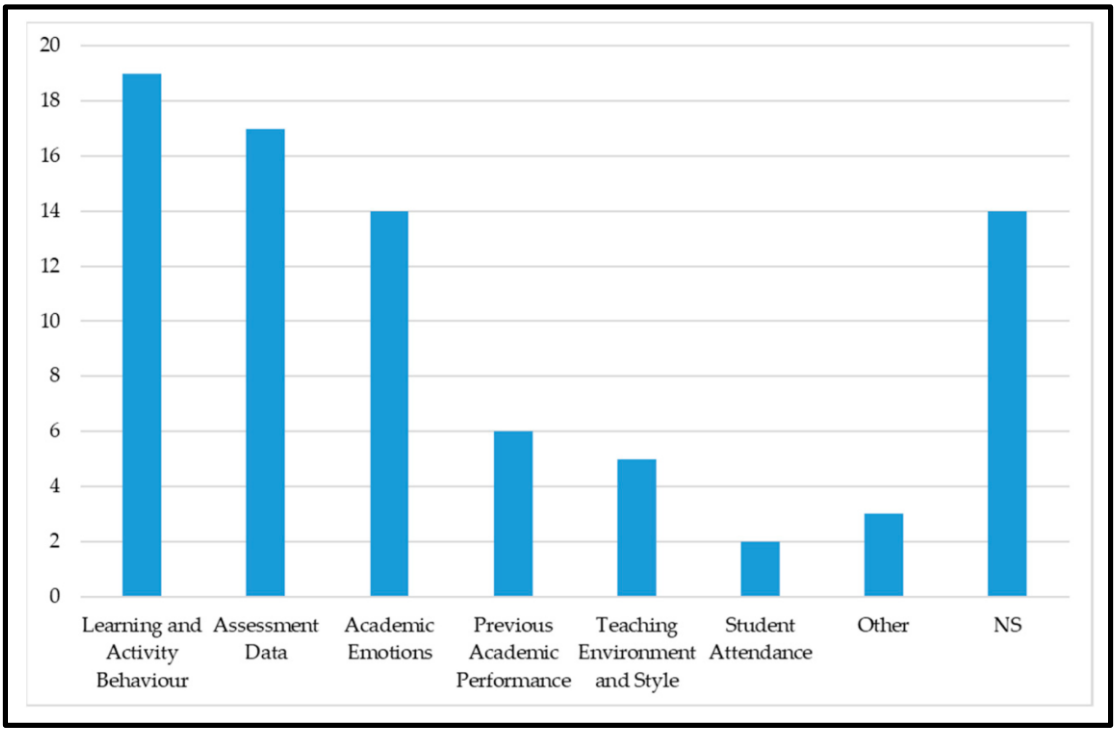

Figure 4: Evaluating Student Academic Performance in the Context of ONLINE LEARNING

The correlation with related work highlights the headways made in the prescient modeling of understudy results, featuring the adequacy of consolidating refined information mining methods in the instructive examination. Our trials exhibit the adequacy of information mining algorithms in assessing understudy scholarly execution about ONLINE LEARNING disturbances. Random Forest

and Neural Networks arise as promising methodologies for anticipating understudy results, offering important bits of knowledge for teachers, policymakers, and partners in training [30]. By utilizing progressed examination and machine learning strategies, can acquire further bits of knowledge into the complex effects of the pandemic on understudy learning and illuminate proof-based mediations to help understudy outcomes in remote learning conditions.

V. CONCLUSION

In conclusion, this exploration has addressed the squeezing need to assess understudy scholastic execution about ONLINE LEARNING disturbances utilizing information mining strategies. By utilizing progressed scientific procedures, including linear regression, decision trees, random forest, and neural networks, it has acquired important experiences with the diverse effects of the pandemic on understudy learning results. Our tests showed the sufficiency of these algorithms in predicting student execution with fluctuating degrees of precision and exactness. Specifically, random forest and neural networks emerged as promising philosophies, beating standard verifiable methodologies and essential machine learning strategies. The connection with related work included the types of progress made in the perceptive modeling of student results and featured the meaning of solidifying refined data mining methodologies in the educational investigation. In addition, the blend of revelations from arranged disciplines, including preparing, development, and humanistic systems, high level of how it could decipher the staggering exchange of factors affecting student academic achievement during risky events. Pushing ahead, our examination discoveries can illuminate proof-based mediations and strategy decisions to help understudies really in remote learning conditions and alleviate the unfriendly impacts of disturbances on instructive results. Moreover, our review highlights the requirement for progressing exploration and advancement in instructive information examination to address arising difficulties and amazing open doors in the developing scene of schooling. By proceeding to tackle the force of information mining and interdisciplinary joint effort, can endeavor towards cultivating strength and advancing impartial instructive encounters for all understudies, paying little mind to outside disturbances.

REFERENCE

[1] Ahmad, S., El-Affendi, M., M, S.A. and Iqbal, R. 2022, "Potential Future Directions in Optimization of Students' Performance Prediction System", Computational Intelligence and Neuroscience : CIN, vol. 2022.

[2] Alsalem, M.A., Alamoodi, A.H., Albahri, O.S., Dawood, K.A., Mohammed, R.T., Alnoor, A., Zaidan, A.A., Albahri, A.S., Zaidan, B.B., Jumaah, F.M. and Al-Obaidi, J. 2022, "Multi-criteria decision-making for coronavirus disease 2019 applications: a theoretical analysis review", The Artificial Intelligence Review, vol. 55, no. 6, pp. 4979-5062.

[3] Atashinsadaf, A., Ramezani-badr, F., Long, T., Imanipour, M. and Amini, K. 2024, "Facilities, challenges, attitudes, and preferences of nursing students related to e-learning in the Online Learning in Iranian context: a cross-sectional study", BMC Medical Education, vol. 24, pp. 1-14.

[4] Capetillo, A., Camacho, D. and Alanis, M. 2022, "Blockchained education: challenging the long-standing model of academic institutions", International Journal on Interactive Design and Manufacturing (IJIDeM), vol. 16, no. 2, pp. 791-802.

[5] Dacka, M., Wolanin, A. and Rybak, J. 2022, "Perception of pandemic and students' quality of life in the initial stage of pandemic – a quantitative and qualitative perspective", Polish Psychological Bulletin, vol. 53, no. 4, pp. 315-328.

[6] Darcy, D.P. and Satpathy, A. 2023, "Teaching Tip: A Scalable Hybrid Introductory Analytics Course", Journal of Information Systems Education, vol. 34, no. 4, pp. 360-369.

[7] Desai, R.R. and Kim, J. 2023, "Bibliographic and Text Analysis of Research on Implementation of the Internet of Things to Support Education", Journal of Information Systems Education, vol. 34, no. 2, pp. 179-195.

[8] Dima, A.M., Busu, M. and Vargas, V.M. 2022, "The mediating role of students' ability to adapt to online activities on the relationship between perceived university culture and academic performance", Oeconomia Copernicana, vol. 13, no. 4, pp. 1253-1281.

[9] Eckhaus, E. and Davidovitch, N. 2021, "Driving value creation in the new economy following the ONLINE LEARNING crisis. Data-mining students' satisfaction from online teaching in the virtual academic climate.: EJEL", Electronic Journal of E-Learning, vol. 19, no. 5, pp. 452-468.

[10] Gao, C.X., Broder, J.C., Brilleman, S., Campbell, T.C.H., Berger, E., Ikin, J., Smith, C.L., Wolfe, R., Johnston, F., Guo, Y. and Carroll, M. 2023, "Evaluating the impact of Hazelwood mine fire event on students' educational development with Bayesian interrupted time-series hierarchical meta-regression", PLoS One, vol. 18, no. 3.

[11] Giron, H.S. 2023, Implementation of State of New Mexico's Manual of Procedures PSAB Supplement 12 Capital Assets by a Southern New Mexico School District in Disposal of E-Waste in the Post COVID 19 Environment, The University of Texas at El Paso.

[12] Griffin, S.C., Scanlon, M.M. and Reynolds, K.A. 2023, "Managing Building Water Disruptions in a Post-COVID World: Water Quality and Safety Risk Assessment Tool for Academic Institutions and School Settings", Buildings, vol. 13, no. 4, pp. 921.

[13] Hands, C. and Limniou, M. 2023, "Diversity of Strategies for Motivation in Learning (DSML)—A New Measure for Measuring Student Academic Motivation", Behavioral Sciences, vol. 13, no. 4, pp. 301.

[15] Harahsheh, K. and Chen, C. 2023, "A Survey of Using Machine Learning in IoT Security and the Challenges Faced by Researchers", Informatica, vol. 47, no. 6, pp. 1-54.

[16] Holicza, B. and Kiss, A. 2023, "Predicting and Comparing Students' Online and Offline Academic Performance Using Machine Learning Algorithms", Behavioral Sciences, vol. 13, no. 4, pp. 289.

[17] Jinnatul, R.M., Connolly, R., McParland, C. and Md Abul, K.A. 2022, "Understanding Barriers to Female STEM Students' Adoption of Online Learning During A Pandemic: An fsQCA Analysis", Pacific Asia Journal of the Association for Information Systems, vol. 14, no. 6, pp. 3.

[18] Khan, M. 2023, Advancing Clinical Practices and Patient Outcomes Through Computational Analyses in Medicine: A Focus on SARS-CoV-2 Epidemiology, ONLINE LEARNING, and Long-COVID Neuropathogenesis, University of Nevada, Reno.

[19] Liu, B., Luo, X. and Shui-lin, L. 2023, "A Research on Online Teaching Behavior of Chinese Local University Teachers Based on Cluster Analysis", International Journal of Distance Education Technologies, vol. 21, no. 2, pp. 1-20.

[20] Liu, Y., Huang, Z. and Wang, G. 2023, "Student learning performance prediction based on online behavior: an empirical study during the Online Learning", PeerJ Computer Science, .

[21] Ma Janice, J.G. and Francee Mae, F.C. 2023, "Determining Ergonomic Appraisal Factors Affecting the Learning Motivation and Academic Performance of Students during Online Classes", Sustainability, vol. 15, no. 3, pp. 1970.

[22] Mignone, P.O. 2023, Sense of Belonging in the Time of ONLINE LEARNING: Examining the Role of Friendships, State University of New York at Binghamton.

[23] Montes, J., Ávila, L., Hernández, D., Apodaca, L., Zamora-Bosa, S. and Cordova-Buiza, F. 2023, "Impact of entrepreneurship education on the entrepreneurial intention of university students in Latin America", Cogent Business and Management, vol. 10, no. 3.

[24] Nesfield, F.L. 2023, International Students' Perception of the Development of Their Digital Academic Writing Identity Based on Their Participation in an Intensive English Language Program, Pepperdine University.

[25] Portugal, D., Faria, J.N., Belk, M., Martins, P., Constantinides, A., Pietron, A., Pitsillides, A., Avouris, N. and Fidas, C.A. 2023, "Continuous user identification in distance learning: a recent technology perspective", Smart Learning Environments, vol. 10, no. 1, pp. 38.

[26] Rogier van, d.W., Roelens, B. and de Langen, F. 2023, "Improvisational and Dynamic Capabilities as Drivers of Business Model Innovation: An Enterprise Architecture Perspective", Pacific Asia Journal of the Association for Information Systems, vol. 15, no. 1, pp. 1.

[27] Santos, S.S.S. and Carvalho, C.E. 2023, "The use of digital data analytics in the performance of advertising campaigns: the effect of absorptive capacity", Revista Brasileira de Gestão de Negócios, vol. 25, no. 3, pp. 333-352.

[28] Sato, S.N., Moreno, E.C., Rubio-Zarapuz, A., Dalamitros, A.A., Yañez-Sepulveda, R., Tornero-Aguilera, J. and Clemente-Suárez, V.J. 2024, "Navigating the New Normal: Adapting Online and Distance Learning in the Post-Pandemic Era", Education Sciences, vol. 14, no. 1, pp. 19.

[29] Semwayo, J.K. 2024, The Financial and Economic Consequences of the Online Learning on

Zimbabwean Banking Institutions: A Qualitative Case Study, Northcentral University.

[30] Shabur, M.A., Rahman, K.A. and Siddiki, M.R. 2023, "Evaluating the difficulties and potential responses to implement Industry 4.0 in Bangladesh's steel sector", Journal of Engineering and Applied Science, vol. 70, no. 1, pp. 158..