Soft Computing
https://doi.org/10.1007/s00500-021-05925-9

FOCUS

A novel flexible data analytics model for leveraging the efficiency of smart education

Karamath Ateeq[1] · Beenu Mago[1] · Manas Ranjan Pradhan[1]

Accepted: 29 May 2021

Abstract

Conventional educational systems have been uplifted for their efficiency using information and communication technologies in a pervasive manner. The information accumulated from the students, environments, and observations increases data exchanged in the smart education platform. The challenging aspect is the data correlation and its correctness in delivering interactive educational services. Because of addressing the correctness issue, this article proposes a Flexible Observation Data Analytics Model (FODAM). The proposed model relies on the session requirement for extracting useful information. The correctness of the information is verified at the interaction level without losing any tiny observation data. In this model, regressive learning is used for the progressive identification of required data. This learning relies on session requirements and the discreteness of the observed data. The progressive training thwarts the discreteness to provide reliable and non-interrupting educational data for the interacting members. The proposed model's performance is verified using the metrics delay, efficiency, interrupts, and information rate.

Keywords Big data analytics · Data correctness · Discrete event · Linear regression · Smart education

1 Introduction

A radical change in the environment amends the education system according to demand in the society. The modern schooling system replaces agrarian education to face skilled labor force demand, and there is a need for change in the education system in our knowledge-based society also (Kim et al. 2018). Information and communication technology (ICT) has come up with education management in schools throughout the world. Smart education uses recent technology to allow students to use up to date technology for learning and use various materials according to their intellectual and aptitude levels (Manogaran and Lopez 2018; Freigang et al. 2018). The teacher-centered model, followed by the existing educational system, allows knowledge and information flow in a single direction from teachers to students. Students from the top and bottom classes are unsatisfied with single direction education system (Raizada et al. 2020). The quality of output is managed, and teacher productivity is raised by information and communication technology. In education management, using ICT is greatly under accentuated. Education managers make resource allocation decisions according to the needs such as quality of teaching and teacher and student flow (Kumar et al. 2020).

Smart education is promoted by intelligent technologies like cloud computing, big data analytics, wearable technology, learning analytics, the Internet of things, etc. The data can be collected, analyzed, and focused on improving teaching and learning (Akhrif et al. 2020). Big data analytics supports the development of adaptive learning and personalized learning. With adaptive learning technologies, personalized learning is carried out efficiently for the learners according to their intellectual level (Nieto et al. 2019). Using big data analytics in educational institutions makes decision-making and strategic planning effectively by collecting the data's richness. The amount of data

Communicated by Vicente Garcia Diaz.

✉ Manas Ranjan Pradhan
manasranjanpradhan123@yahoo.com

Karamath Ateeq
karamath.ateeq@skylineuniversity.ac.ae

Beenu Mago
beenu.mago@skylineuniversity.ac.ae

[1] School of Information Technology, Skyline University College, Sharjah, United Arab Emirates

generated by education increases, and solutions for managing data will also increase (Shen et al. 2020). To improve education, government uses Educational Big Data (EBD). EBD value becomes visible, and its benefits are provided to education management systems to adopt new strategies in education management (Sathishkumar et al. 2020). Educational Big Data enables educational institutes to access data from various sources and respond to the education sector changes and progress in real-time students' behavior to provide solutions (Shorfuzzaman et al. 2019; Prathik et al. 2016).

Artificial intelligence is a valuable tool to develop and imitate the decision-making process that is accepted by the people. Different artificial intelligence techniques such as neural networks, genetic algorithms, Bayesian networks, decision trees, Markov model, and fuzzy logic are used (Baig et al. 2020; Wang et al. 2015). Searching relevant artificial intelligence techniques and learning theory for applying in a learning environment have no standard approach. Traditional educational environment and e-learning environment use learning style artificial intelligence techniques to convey adaptive education (Prathik et al. 2016; Lnenicka et al. 2020). The learning model and student model are defined by fuzzy logic and the neural network used to interconnect the systems to process the data collected from students (Williamson 2018; Lara et al. 2020). Artificial neural networks identify the learning style, and a decision tree is used to provide the learning patterns (Lister 2018). The Bayesian network is used to model the complete learning process. Markov model identifies the student behavior and regulates the similarity, and then, genetic algorithm provides an optimal learning path for students according to their intellectual level or behavior (Hernandez-de-Menendez et al. 2020).

Artificial intelligence is a helpful tool for developing and imitating the people's decision-making process. A variety of artificial intelligence approaches are utilized, including neural network, genetic algorithms, Bayesian networks, decision-making trees, Markov's model, and fuzzy logic. There is no standard strategy for the search for suitable artificial intelligence and learning theory for application in a learning environment. The artificial intelligence learning strategies use traditional educational environments and e-learning surroundings to transmit adaptive learning.

Strategies based on Educational Big Data are used to improve education. Further, Educational Big Data value becomes obvious and beneficial for education management systems to embrace new educational management techniques. Various strategies on the Big Data education system allow schools to accede to data from many sources and to adapt to changes in the educational sector, and to improve their behavior. Artificial intelligence strategies are a helpful tool for developing and imitating the people's decision-making process. A variety of artificial intelligence approaches are utilized, including neural network, genetic algorithms, Bayesian networks, decision-making trees, Markov's model, and fuzzy logic.

2 Related works

Knowledge aware learning analytics is proposed by Chen (Şerban and Todericiu 2020) for smart learning. A data-driven approach optimizes the learning environment. The framework relies on environmental and look-up factors aggregated.

Serban et al. (2020) proposed Q-Learn Framework for smart learning environment to provide easy identification of learning style. The new learning design learns the collaborative manner that the students' are involved. Q-Learn provides a smart learning environment for students. The transfer of knowledge from teacher to students is ensured effectively and efficiently.

Smart education with an artificial intelligence-based determination of learning styles is used in Bajaj and Sharma (2018). The proposed tool framework is used to provide flexibility for selecting and implementing suitable learning model software. For a particular environment, the tool is used to select the most relevant learning style and other learning models. A scalable solution in a cloud environment is provided by the tool to offer rapid and easy identification of learning styles.

IoT-based student interaction framework is used by Farhan et al. (2018). Computer vision library and Visual C# programming language is implemented to collect data. A video camera is used to capture the movement and status of the eyes. Better teaching performance and learning experience is obtained through hidden meaning and correlation of data. Location awareness, social behavior, helping hand, and accessibility of students are provided by IoT-based infrastructure.

Learning analytics tasks as services in smart classrooms are suggested by Aguiler et al. (2018). In the smart classroom, teaching and learning requirement is adapted and responds to the academic services for non-intelligent and intelligent agents through combined multi-agent paradigms and cloud paradigms. For a smart classroom, the learning process is improved by the knowledge feedback loop by defining a set of learning analytics services.

Wongthongtham et al. (2018) reported State-of-the-art personalized teaching and learning strategies for smart learning fuzzy intelligence system. The individual learning process is promoted by identifying relevant teaching resources. The fuzzy technique is used to direct smart learning, and multiple intelligence is used to evaluate the proposed framework.

In Fang et al. (2019), to build a smart lecture-recording system MK-CPN network is used for heterogeneous data sources. Event tracking, searching, and detection are included in the virtual cameraman and counter propagation neural (CPN) network to characterize the machine learning techniques. Multiple kernel learning (MKL) is applied to increase the accuracy. Oration record is provided by the proposed system and extends it for real human teams. The proposed system is configured with suitable training materials, so it will not limit live speeches.

To review the ontology-based recommender systems for e-learning, Tarus et al. (2018) designed a knowledge-based recommendation system. Various recommendation techniques in ontology-based e-learning recommenders are categorized. Knowledge representation technique, ontology representation language, and ontology types of the proposed system are classified. The future trends are discussed for the proposed approach in the context of e-learning. The proposed framework improves the equality of the proposed system.

Integration of data mining clustering approach in the personalized e-learning system is implemented by Kasar et al. (2018). According to students learning capabilities, the teaching contents are detected and responded to by the personalized e-learning system framework. The proposed work will improve learning capabilities by mining-based analysis using clustering methods. The proposed technique will analyze big data effectively to provide vigorous education systems.

Lin et al. (2018) considered an intelligent recommendation system for course selection a smart education environment. Course enrolment datasets are collected to implement the proposed work. A top-N recommendation is generated by sparse linear methods (SLIM). The performance of the proposed work is compared with the state of art methods. The result shows that the proposed approach is accurate and efficient.

User perceptions of smart class services in teaching and learning interactions are recommended by Jin et al. (2019). The proposed work is used to provide a feasible tool for university students. Correlation is analyzed, and the quantitative effects on user technology are also examined. Student's perceptions are examined by the ten variables from already validated measurement scales. Total students from the university are selected, and in return, valid observations are provided by data analysis.

A collaborative filtering recommendation algorithm based on the influence sets of the e-learning group's behavior is used (Liu 2019). A recent e-learning group improves the evaluation density, computation, ring method, and prediction of resources. Comparing to other existing techniques proposed method will effective solve the sparse data problem. The recommendation quality is improved.

A data-driven approach is suggested by Yang et al. (2020) in the global learning factory for quality analytics of the screwing process. The proposed work is validated by the artificial intelligence innovation factory, 14.0 innovation center, and learning global factory production. Across the global production network, the data transformation is modeled for cloud and edge-based analytics.

For self-regulated learning assessment, Cerezo et al. (2020) proposed a process mining-based solution. Optimal models are developed by the inductive miner algorithm for a pass and fail students. The self-regulated learning process is used to find students who are not following the instructors. The proposed models examine specific actions. In the pass group and fail group, absenteeism used proposed work to support collaborative learning.

2.1 Proposed flexible observation data analytics model

The most demanding component of the interactive educational services is the correlation of data and its accuracy. For obtaining usable information, the FODAM relies on the session requirement. The accuracy of the information is checked without losing minute observational data on the degree of contact. Regressive learning is utilized to gradually identify the needed data in this model to improve educational in schools and universities. This education depends on the requirements of the session and discretion of the data seen. The progressive training contradicts the discretion to supply the interacting members with dependable and continuous education data.

Smart education is used to deploy better communication and exchange information between end-to-end users. In this smart education platform, the delay and session expire addressed by developing the FODAM method. By performing this, correctness of the information is achieved and improves the information rate. The non-interruption service is obtained from the proposed work by identifying the observed data in big data analytics. The observed data are large, so it reflects the session's activity and degrades the continuity of data.

This correlation factor matches the session data in the smart education and replaces them to enhance the data correctness. Thus, FODAM is used to sort out the delay and interruption and improve the efficiency and information rate. Here, the data is examined from the data analytics in smart education and provides an efficient information rate. The following equation is used to evaluate the data from observations and identify the large observed data and needs more activity sessions.

$$x_a = (d_t + \beta) * \left(\frac{v_i/_{e_s}}{b_0} \right) + (e_n - j') * \sum_{e_s} (u' + n_0) - \partial \tag{1}$$

In Eq. (1), the examination of observed data is equated. It is done from the big data analytics; here, the interaction and communication between the end-to-end users are detected. Here, the detection is represented as ∂ , the data is denoted as d_t, and the session and activity are termed as $e_s$andv_i. The identification of data is denoted as β, and here, the observed data is referred to as b_0, and the relationship is termed as n_e. The interactive session $u^{'}$ is estimated for the end-to-end users, it is represented as n_0, and it is performed on time that is denoted as j'.

The identification of data from the big data is used to allocate the data to the required smart education session that deploys the correctness of information among the users. The data is delivered to the end-user and examine the session processing. It provides the correct information without any delay, and it is represented as $\left(\frac{v_i/_{e_s}}{b_0} \right) + (e_n - j')$. Here, e_n is denoted as the number of sessions, and examination is represented as x_a that deploys the discrete data. This work aims to address the discrete data and improve the continuity of the session without data loss in the smart education system. The following section is used to detect the discrete data in the session and avoid further processing.

2.2 Detection of discrete data sequence

The discrete data is detected in the preliminary step and addresses the observed data activity during the education session that leads to interruption. This work aims to address the interruption between the end-users associated with reliable interaction in the smart education system. The following equation is used to evaluate the detection of discrete data from the session, and avoid the data, decreasing the session processing. Smart education is used to better communicate with end-to-end users and to share information. The delay and sitting expired with the development of the FODAM approach in this intelligent training platform. This accuracy of the information is obtained, and the information pace is improved. The non-interruption service is achieved from the work proposed in large-scale analytics by recognizing the observed data. The data seen are substantial and therefore represent the activity of the session and decrease the data continuity.

$$\partial(c_r) = \left(\frac{d_t + x_a}{\beta/_{v_i}} \right) * \prod_{b_0} (a_0 - e_s) + \left(\frac{d_t * n_0}{t_e} \right) - (q' - j') \tag{2}$$

The discrete data is detected in the above equation, and it is denoted as $\partial(c_r)$; here, the examination of data analytics is determined for the sessions, and it is termed as $\left(\frac{d_t + x_a}{\beta/_{v_i}} \right)$. The detection is done from the previous state of data action in the smart education system and avoids them from further processing. For the number of sessions, the discrete data occurs, and it is addressed in the initial stage and rectified for the upcoming processing. The interruption a_0 occurs if there are discrete data resides in the session; in this step, the previous state of matching is performed and avoids the data delay, and it is denoted as q'. In big data analytics, the discrete data is identified from the previous state of action and addresses it; here, the correctness t_e is improved and it is denoted as $\left(\frac{d_t * n_0}{t_e} \right)$.

2.3 Similar data replacement

The interaction is carried out between the end-users, and the continuity is maintained promptly, and the correctness is enhanced. Here, the session's activity is identified, avoids the discreteness, and allocates the average data to the particular session to improve the data processing. The discrete data is detected in the above equation; a check is done by replacing the data in a continuous form from this similarity. The following equation is used to replace similar data from the previous processing and improve smart education sessions.

$$l_c = \sum_{v_m}^{d_t} (e_s + \partial) * \left[\left(\frac{\beta}{c_r/_{u'}} \right) + \left(\partial + d_t/_{a_0} \right) \right] * e_s - v_m + \prod (j' - b_0) * \beta \tag{3}$$

Replacing the similar data is done by matching with the previous state of action associated with the identification process and improving the interaction. A similar data replacement process is illustrated in Fig. 1.

As illustrated in Fig. 1, the observed, aggregated, and activity data are exploited for the interaction session in the data replacement process. The data is differentiated for its discreteness/non-discreteness. For discrete data, similarity from the last-known session is verified, and appropriate data is replaced. The session detection is used to rectify the discrete data and find similar data and replaces them to improve the session that enhances the interaction. The identification of data is made for interaction between the end-users, and it is termed as $\left(\frac{\beta}{c_r/_{u'}} \right)$, in this interruption is sort out. Here, the previous state is used to replace similar

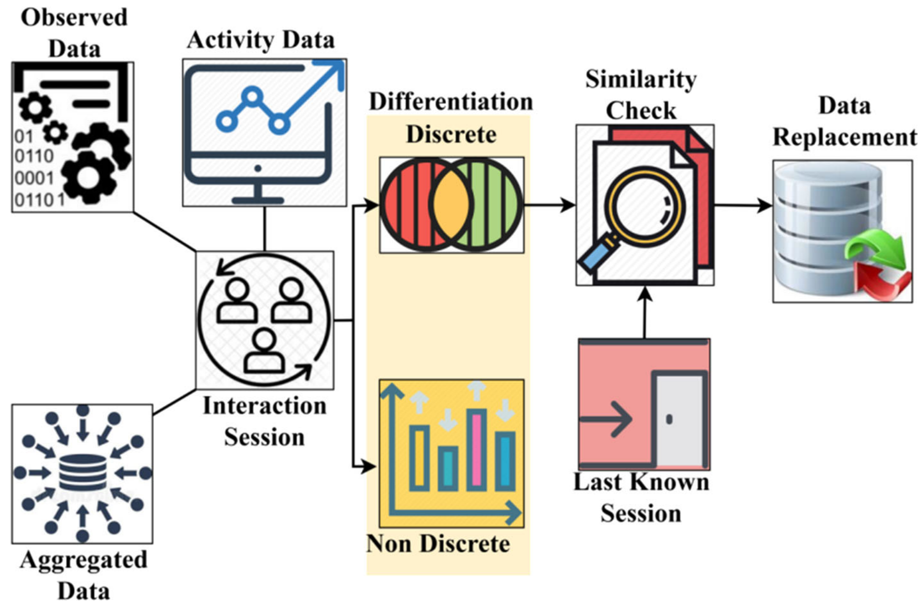


Fig. 1 Similar data replacement

data from the detected discrete data and improves the session reliably.

Smart education is used to better communicate with end-to-end users and to share information. The delay and sitting expired with the development of the FODAM approach in this intelligent training platform. This accuracy of the information is obtained, and the information pace is improved. The non-interruption service is achieved from the work proposed in large-scale analytics by recognizing the observed data. The data seen are substantial and therefore represent the activity of the session and decrease the data continuity.

The replacement is referred to as l_c, by deploying this correctness of data, and it is examined promptly in the smart education associated with the information rate. The objective of this work is to improve the information rate. For this, interruption in between the sessions is addressed and it is equated as $\left(\partial + d_t/_{a_0}\right)$. The previous data is denoted as v_m; it is performed for every step of computation that relates to finding discrete data. The identification of data is made promptly, and it is termed as $\prod(j' - b_0) * \beta$. The below equation is equated to the data correctness process that is performed post to replace similar data.

$$t_e(x_a) = \begin{cases} \prod_{u'}(d_t * n_0) + \left(\partial/_{v_m}\right) * u' \\ = \left(\dfrac{l_c * v_i}{u' + d_t/_\partial}\right) * n_0 - (q' - j') \end{cases} \tag{4}$$

The data correctness is examined in Eq. (4); here, the interaction between the end-users is evaluated and determines the activity. For the end-user, data delivery is carried out promptly, which is associated with better interaction. Here, the detection is evaluated to replace similar data, and it is denoted as $\left(\partial/_{v_m}\right) * u'$. A similar data replacement is done for the varying session that deploys discrete data in the initial state and improves the correctness efficiently.

Here, the correctness is evaluated in the session activity and determines the interaction between the end-users, which matches the previous state. This information rate is used to deploy the observed data from the big data and provides efficient data processing that is associated with the interaction, and it is termed as $\left(\frac{l_c * v_i}{u' = d_t/_\partial}\right)$. In this processing, the observed data's activity is identified and needs more allocation of the session that leads to degrading the information rate. The below division is used to evaluate the correlation factor that is performed to replace similar data to produce the complete session.

2.4 Correlation

The correlation is used to match the previous state of data and provides an efficient session and activity continuously; thus, processing the previous data is necessary for mapping. The correlation is done to achieve the correctness of data that relates to similar data matching. Figure 2 presents the correlation process illustration.

The observed data is analyzed for activity and continuity throughout the session. The activity data is mapped based on correlation, after which further differentiation is performed. On the other hand, the non-discrete data is verified for its discreteness, and correlation is performed in the aggregated instance only (Refer to Fig. 2). If discrete data

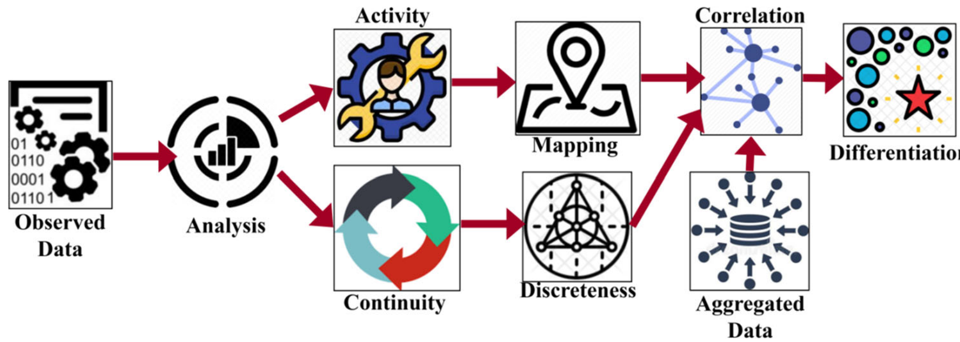


Fig. 2 Correlation process illustration

occurs during the processing state, the activity and session continuity degrades, which leads to data loss in analytics. For overcoming this, for every computation step, the correlation factor is used to replace the similar data and maintain the continuous session to the end-users. The below equation is evaluated to find the correlation between the previous data and provides an efficient session by allocating the data to the particular session. Artificial neural networks define the study style and are used to give learning patterns through a decision tree. The Bayesian network models the whole process of learning. Markov's model recognizes the behavior of the student and regulates the similarity and then gives a genetic algorithm for students to learn according to their intellectual or behavioral level.

$$o_e = \prod_{d_t}^{\partial}(v_m + \beta) * (u' + a_0) + \left(\frac{x_a}{\sum_{n_0} d_t}\right) * \delta - (j' + \beta) \tag{5}$$

In the above equation, the correlation is equated by matching with the previous state, and it is deploying the identification of observed data from the session and avoids them. Here, the identification is carried out for the varying set of processing associated with the examination of data from big data analytics. The correlation is used to map the similar data from the previous state and performs the replacement efficiently, that is done by examining, and it is denoted as $(u' + a_0) + \left(\frac{x_a}{\sum_{n_0} d_t}\right)$.

The correlation is termed as o_e; it is evaluated for every step of detecting observed data in the session that leads to more activity from the user. Here, the matching with the previous state is done to address the discrete data associated with the detection process, and it includes the activity of the session continuously. The information rate is improved by correlating similar data promptly, and it is represented as $\delta - (j' + \beta)$; the information rate is defined as δ. Thus, replacing similar data is carried out in Eq. (5), from that differentiation of explanatory and dependent data is performed by introducing linear regression. The accuracy in the session activity is evaluated, and the interaction between the end-user is determined to correspond to the preceding condition. This rate of information is utilized to deploy the observation data from large amounts of data and to offer efficient interaction data processing. In this procedure, the data activity seen is detected and the session has to be allocated to the extent that the information rate is degraded. The following division is used for assessing the correlation factor done for replacing comparable data throughout the entire session.

2.5 Linear regression (LR)

LR is used to perform post to the correlation of similar data; here, it is used to determine the data points and their position and produces the results by identifying the relationship. The explanatory is defined as the independent data; it is defined as the non-discrete data, whereas dependent states the discrete data. If it is non-discrete, the session is carried out continuously without any delay; in another case, the session interruption occurs. To address this issue, LR is developed and determines the smart education data points and improves the information rate reliably. The following equation is used to evaluate the differentiation of explanatory and dependent data from big data analytics.

$$f_i = \left.\begin{array}{ll} (\beta + d_t) * \left(\dfrac{s_k + a'}{v_m}\right) + t_e * u' - j', & \forall \text{ Explanatory} \\ \left(\dfrac{a_0 + d_t}{\prod_{e_s} v_m}\right) * \sum_{q'}(\beta + c_r) - v_i, & \forall \text{ Dependent} \end{array}\right\} \tag{6}$$

In the above equation, differentiation of data is evaluated, and it is represented as f_i; here, the first case states the explanatory data, whereas the second case is denoted as a dependent. In the first case, the independent data is evaluated that deploys the continuous and non-interrupt session, and it is denoted as s_k and a$'$. Here, the previous state of detection evaluated for this correlation is used and maps similar data and finds whether it is independent or dependent data. In this processing, the interaction between the end-users is carried out promptly $\left(\frac{s_k+a\prime}{v_m}\right) = t_e * u' - j'$; in this, continuous data flow is estimated. In Fig. 3(a) and (b), the regression's discrete and interruption states are represented.

The second case is used to analyze the dependent data; here, the discrete and interruption are stated from the big data analytics that degrades the interaction. This part of the computation is addressed in this LR, which is associated with the FODAM method and satisfies this proposed work's objective. Here, the delay occurs during data processing at the session's activity, and it is identified and rectified by evaluating the upcoming solutions. The delay and discrete data is identified and analyzed the activity, and it is denoted as $\sum_{q'} (\beta + c_r) - v_i$, and thus, the differentiation of data is evaluated in the above equation and identifies the explanatory and dependent data. The following equation is used to estimate the prediction to address the dependent data and sort out the issues during the session.

$$k_0 = \sum_{d_t} (v_m + \beta) * \left(\frac{\partial + a_0 - a\prime}{o_e - v_m}\right) + (x_a * u') + \prod_{x_p - p_d} (h' + \partial) - j' \quad (7)$$

The prediction is made by equating Eq. (7); here, the matching data is evaluated by previous data and forthcoming data and provides a better information rate. Here, the identification of discrete data is detected that is associated with the interrupt and non-interrupt data, and it is denoted as $\left(\frac{\partial+a_0-a\prime}{o_e-v_m}\right)$. The prediction is termed as k_0, that is used to match with similar data and improves the correctness of the session. The examination of discrete data in the session is determined in the LR and provides a better correlation for further data in the education system.

From the differentiation process, explanatory and dependent data is determined, and it is referred to as $x_p$andp_d; by evaluating this, the interaction is estimated. Here, the allocation of data to the session is done by deploying this prediction step associated with the data point selection and represented as h'. In this processing, similar data are correlated and replaced to evaluate the education system's continuous session activity. The data points are used to differentiate the explanatory and dependent data, and it is associated with the LR model, for this best fit is evaluated in the following equation.

$$g_0 = (e_n + \partial) * \left(\frac{\sum_{n_0} c_r}{\delta}\right) (x_p - p_d) - j' + \left(\frac{f_i}{q' - v_m}\right) * l_c(o_e) - [h'(v_m + o_e)] * k_0 \quad (8)$$

In Eq. (8), the best-fit data is detected for the number of sessions in a smart education system related to the differentiation of data. Here, the best fit is evaluated by determining the explanatory and dependent data that is represented as $\left(\frac{\sum_{n_0} c_r}{\delta}\right)(x_p - p_d)$ in this; it matches the data and replaces it with similar data. The prediction is estimated to find the best fit in LR associated with the explanatory and dependent data. The replacement of similar data is used to evaluate the best fit in LR that is denoted as $[h\prime(v_m + o_e)] * k_0$; here, the prediction is processed to find the data point. Thus, the best fit is identified by equating the above equation by performing the prediction and correlation factor. The following equation is used to find the error function that decreases the efficiency of the proposed work.

$$r' = \prod_{e_s}^{e_n} d_t + \partial * \left(\frac{o_e - k_0}{g_0}\right) - (q' + v_m) \quad (9)$$

In Eq. (9), the error function is detected that is used to achieve a better fit solution for the LR model that relates to the prediction data. The error function is denoted as r' ; in this, detection of observed data is identified that deploys the prediction process and finds the best fit that is represented as $\partial * \left(\frac{o_e-k_0}{g_0}\right)$. For every step of detection, discrete

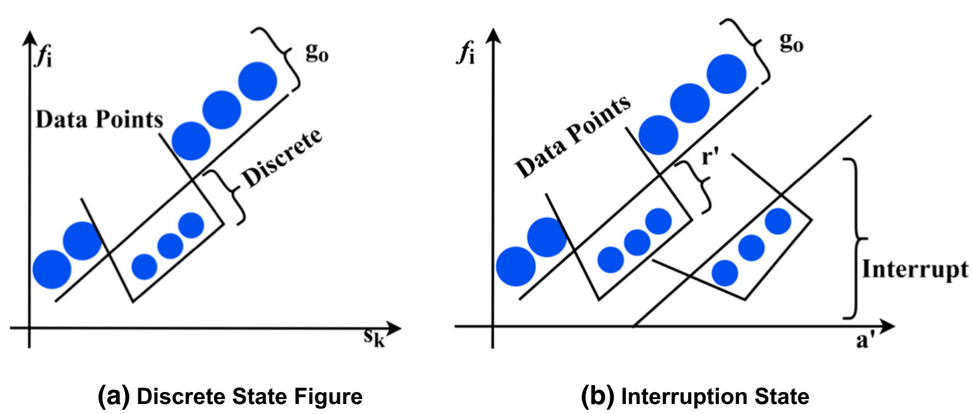


Fig. 3 **a** Discrete state figure. **b** Interruption state

data in the big data analytics is identified that deploys differentiated data and improves the interaction. The prediction process is evaluated to match the previous computation step that is associated with identifying the observed data. The examination of data is used to find the error function used to deploy the information rate and the best fit. The following equation is used to determine the relationship between the explanatory and dependent data based on weight. It is observed that the student model and the learning model are defined by fluid logic and the neural network that link the systems in order to handle student input. Artificial neural networks recognize styles of learning and utilize a decision tree to generate models of learning. The Bayesian network models the whole process of learning. The Markov model identifies and regulates the student's behavior, and then, the evolutionary algorithm offers pupils with an optimum educational route according to their intellectual levels.

A model Markov is a stochastic one in probability theory, used for modeling pseudo-random systems that change. Future states are considered to depend exclusively on the present state, not on preceding occurrences. This assumption generally allows the rationale and calculation of the model that would otherwise be unwieldy. This is why it is desired for a given model to display the Markov property in the areas of predictive modeling and probabilistic forecasting.

$$y_e(n_e) = \prod_{e_n} o_e * w_0(e_n) + d_t * v_i * \left(\frac{v_m + h\prime}{n_0}\right) + (g_0 + k_0) \tag{10}$$

In Eq. (10), the determination of relationship is evaluated, and it is denoted as $y_e(n_e)$ that deploys to detect the correlation of data from the previous state. Here, the relationship is determined between the explanatory and dependent data in LT and provides the continuous and data correctness to the number of sessions in smart education. For varying sessions and activities, the prediction model is used to find the plane's data points and maps the best fit. In this stage, the error data is identified and decreases the delay factor by equating Eq. (10) related to the identification of observed data. In Table 1, the best-fit solutions for different sessions are tabulated.

For the varying sessions, the instances are evaluated that is associated with the discrete and non-discrete data. The discrete data shows low to high value, whereas non-discrete decreases if discrete data is detected and vice versa. The correlation percentage shows a higher range if non-discrete data is detected. The best fit increases if the correlation factor is improved and shows better session interaction (Table 1).

In this, best fit is deselected by deploying prediction that shows better interaction among the end-users in the smart education, and it is formulated as $\left(\frac{v_m+h\prime}{n_0}\right)$. The previous state of matching is carried out by evaluating the set of data by assigning weight, and it is denoted as w_0; for every data processing, the weights are provided. In this processing, the data's determination is identified, and its weights identify the relationship among the data. The following equation is used to identify the interaction; by performing this, the session's activity is improved. It is done by detecting the relationship factor for the varying data that deploys better interaction in smart education.

$$\beta = \frac{1}{e_n} * \sum_{r\prime}(d_t + v_m) * \left(\frac{y_0}{g_0/h\prime}\right) + n_e(d_t) * \left(f_i + \frac{t_e}{o_e}\right) + v_i \tag{11}$$

The interaction is identified by equating Eq. (11) that deploys the number of sessions and detects the observed data. This prediction is used to map the previous data detection step and address the dissimilarity of data. For this processing, differentiation of data is used in LR and determines the error function and improves data points' detection. In this identification step, the relationship between the explanatory and dependent data is differentiated and finds the data points, and it is equated as

$$\left(\frac{y_0}{g_0/h\prime}\right) + n_e(d_t)$$

.

Here, differentiation of data is used to deploy the data, improve the correctness efficiently, and relate to the better correlation. The data correlation is used to replace the similar data that shows better processing for the proposed work. Here, FODAM is used to detect the availability of data and it helps to improve the correctness of the particular session. In this identification step, the activity in the session is detected and improves the data allocation to the particular session represented as $\left(f_i + \frac{t_e}{o_e}\right) + v_i$. Thus, the interaction is improved by equating the above equation, and the following equation is used to detect the session interruption.

$$\partial(a_0) = v_m + \left(d_t * \frac{\beta}{s_k}\right) * \delta + (w_0 * o_e) + \left(y_e * \frac{l_c}{g_0}\right) \tag{12}$$

In Eq. (12), the interrupt data is detected in the session, and it is avoided for the upcoming activity in education and shows better interaction. Here, the detection of interruption is done from the differentiation of data that utilizes the LR, and similar data replacement is done by evaluating the relationship's determination. The best fit is detected that

Table 1 Best-fit solutions for different sessions

Sessions	Instances		Correlation (%)	Best-fit solutions
	Discrete	Non-discrete		
1	31	103	93.28	59
2	52	85	83.23	42
3	65	97	68.68	38
4	58	28	71.16	45
5	56	20	72.6	39
6	31	103	88.9	48
7	152	20	76.35	35
8	129	15	93.13	58
9	115	25	75.71	62
10	156	12	68.46	19

deploys the interaction and correctness, and here, the delay factor is decreased by identifying the observed data, and it is denoted as $v_m + \left(d_t * \frac{\beta}{s_k}\right)$. The regression process for interrupt classification is illustrated in Fig. 4(a) and (b) for the s_k and a', respectively.

The continuous data processing is evaluated that is associated with finding the data points and improves the session. Here, the interrupt will not occur and shows a better information rate by addressing the initial step's discrete data. Equation (13) is used to decrease the delay factor, for this information rate is improved that is done by mapping the previous state of data processing and identifies the observed data. By evaluating this, the delay factor is decreased for the session and shows reliable interaction.

$$\delta(\beta) = d_t * s_k + \prod_{y_e}(v_m + r') + \partial * \left(\frac{x_a}{r\prime/x_a}\right) - q' + (o_e * h') + l_c \tag{13}$$

The information rate is improved and shows lesser delay for the smart education session that deploys reliable interaction for the varying users. This continuous flow of activity is provided that deploys the data allocation to the necessary sessions that relate to finding the discrete data. Here, the previous state of matching is done to address the data correctness by detecting the observed data, and it is denoted as $\left(\frac{x_a}{r\prime/x_a}\right) - q'$; in this, error function is used to decrease the interruption. Thus, the delay factor is decreased for the proposed work and shows a better information rate; from this, the efficiency is enhanced in the following equation.

$$\mu = \frac{(u' + a\prime) * l_c(t_e)}{e_n + d_t} \tag{14}$$

In Eq. (14), the data's efficiency is improved by detecting the observed data that deploys the prediction process and matches the data and improves the interaction. Here, the interruption of data is used to deploy the data correctness and shows better efficiency, and it is denoted as μ. The efficiency is done by estimating the number of sessions and the associated data by replacing similar data from the previous state of action and providing the un-

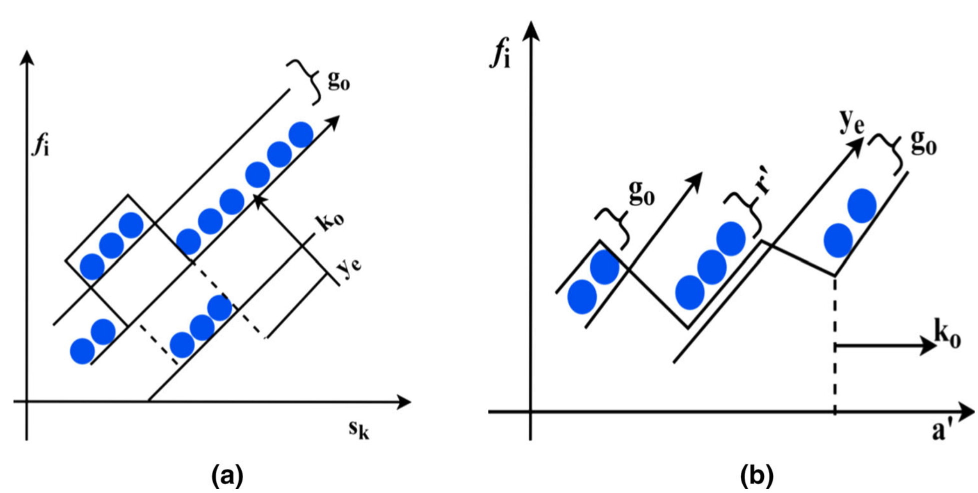


Fig. 4 Interrupt classification for **a** s_k and **b** a

interruption session. It is done by evaluating FODAM and LR method and addresses the objective that deploys the big data analytics in smart education. In Table 2, the efficiency % for different information rates is presented.

The information rate ranges from 0.1 to 1 and shows a varying error that relies on a random value. If the error function is higher, the session's interruption also increases, whereas the best fit decreases if the interruption increases. The efficiency percentage increases if interruption shows a lesser range of values (Table 2).

The prediction varies for correctness and replacement; here, correctness ranges from low to high, whereas replacement shows a high to a low value. If the prediction is improved, the correctness value also increases. The replacement is performed if the similar data is mapped from the prediction method [Refer to Fig. 5(a)]. The prediction varies for best-fit solutions associated with the error function and decreases for the session activity. If the prediction increases, the best fit is improved and shows better detection of discrete data. The error factor for 0.04 shows a higher best fit than 0.08, which deploys better prediction [Refer to Fig. 5(b)].

Differentiation varies for factors associated with the error and information and deploys the discrete data detection. The error factor shows a high to low range of value, whereas information relies from low to high. If the error rate increases, the information rate also decreases and relies on sessions' interruption (Refer to Fig. 6).

3 Results and discussion

This section presents the performance of the proposed FODAM using experimental analysis. A smart classroom with 20 students interaction is designed for observing their activity and preferred information monitoring. The activities of the students are monitored, and questionnaires are observed from different students from 10 sessions. The linear information differentiation takes place in 50 instances. The performance is measured using the metrics delay, efficiency, interrupt factor, and information rate. A comparative analysis is presented using the existing methods RTDMA (Hernandez-de-Menendez et al. 2020), CFRA (Tarus et al. 2018), and DMCA (Aguilar et al. 2018).

3.1 Delay

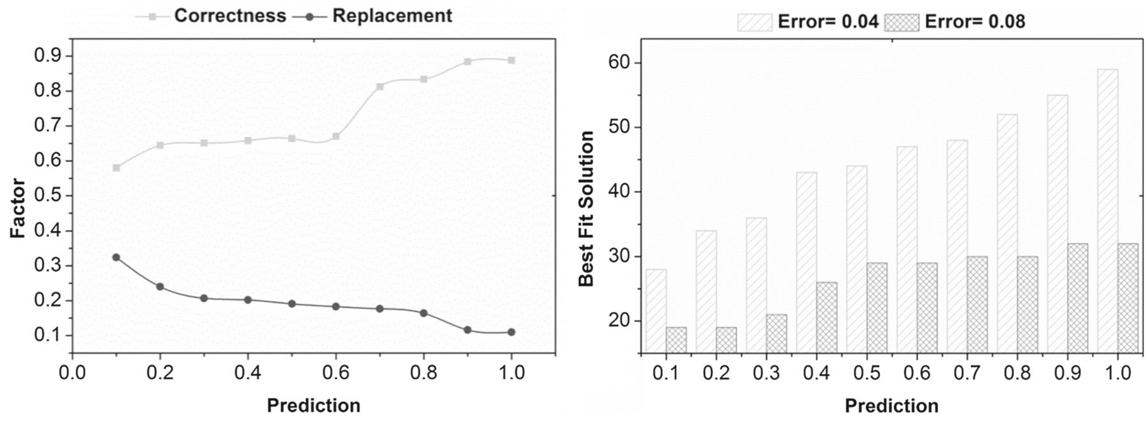
The delay for the proposed work decreases for varying sessions, and the information rate deploys after the session expires on time. If the delay factor decreases, the session rate is improved associated with the non-interrupt of service equated as $\left(\frac{v_i/e_s}{b_0}\right) + (e_n - j')$. Here, the number of sessions is analyzed to deploy a better prediction for the varying data. The data is allocated to the improved sessions, and replacement is done to evaluate the best fit. In this evaluation, the error function is detected that relates to the correctness of the data. Here, the data points are used to deploy better efficiency and show a better information rate. The delay factor is decreased for the different types of sessions and shows reliable processing. In this, discrete data is detected in the session and finds continuous data evaluation. The process's determination is used to allocate the data to the varying sessions in big data analytics. The interrupt of the session is determined that deploys the interaction between the session, and it is represented as $\left(\frac{d_t+x_a}{\beta/v_i}\right)$. The discrete data is detected, decreasing the delay factor; this prediction is used to address the data. The information rate is improved and finds a better session continuously [Refer to Fig. 7(a) and (b)].

3.2 Efficiency

In Fig. 8(a) and (b), the proposed work's efficiency increases for varying sessions and information rate. Here, the error function is addressed by developing a linear regression method associated with classifying the explanatory and

Table 2 Efficiency % for different information rates

Information rate	Error	Interrupt instances	Best-fit solutions	Efficiency %
0.1	0.056	17	56	94.25
0.2	0.058	19	58	94.19
0.3	0.043	12	59	94.39
0.4	0.066	22	51	90.16
0.5	0.069	22	36	89.99
0.6	0.069	24	32	89.08
0.7	0.077	35	31	89
0.8	0.087	37	19	84.63
0.9	0.079	33	26	85.12
1	0.082	35	23	84.74



(a) Correctness and Replacement Factor **(b)** Best Fit Solution for Error

Fig. 5 **a** Correctness and replacement factor. **b** Best-fit solution for error

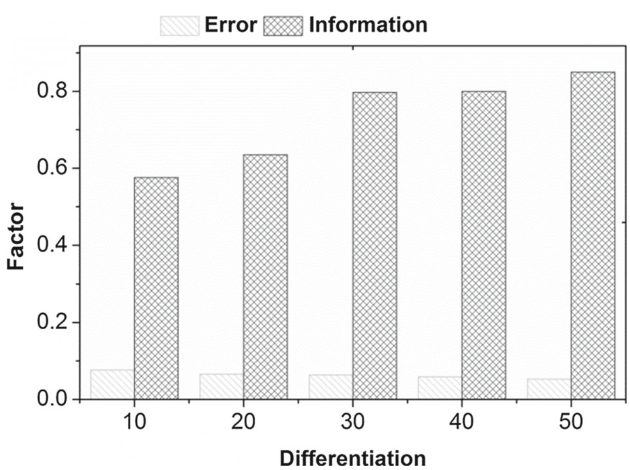
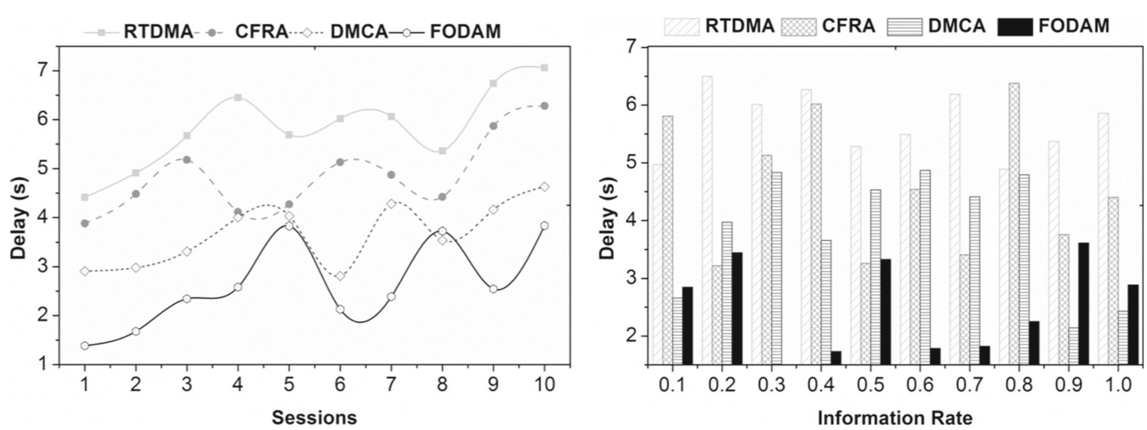


Fig. 6 Error and information factor

dependent data. The data is allocated to the session improvement, and it is formulated as $\left(\partial + {d_t}/{a_0}\right)$. Here, the detection of discrete data is identified as associated with finding the best fit in the plane. It is carried out by using the data without interruption among the session. If the information rate increases, the classification for the varying data relies on identifying the discrete data. Here, the proposed work's efficiency shows a higher value compared to the other three existing methods. In this method, co-relation is used to replace the data with similarity and decrease the computation step. The efficiency is defined by decreasing the delay factor that deploys the varying session interval and information rate. If the information rate increases, the error function is decreasing. In this approach, both the information rate and delay are addressed and the efficiency is improved. Here, the correlation factor is used to replace the data in the session that decreases the activity. The data points are selected in LR and address the error function, and it is denoted as $\left(\frac{x_a}{\sum_{n_0} d_t}\right) * \delta$.



(a) Sessions **(b)** Information Rate

Fig. 7 Delay for **a** sessions, **b** information rate

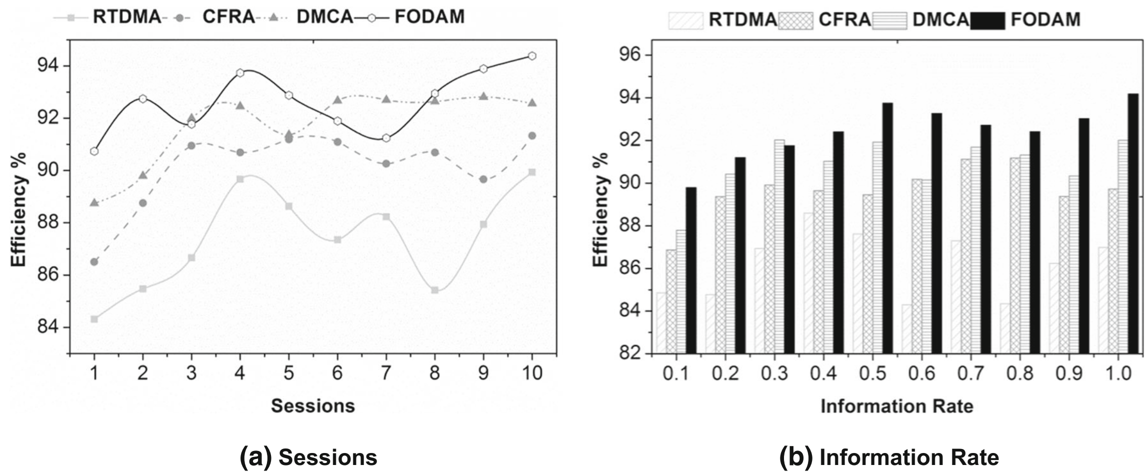


Fig. 8 Efficiency % for **a** sessions, **b** information rate

3.3 Interrupt factor

The interrupt factor decreases for the varying sessions and information rate that deploys by performing prediction. Here, the prediction is used to find the discrete data in the smart education system and addresses them, and it is denoted as $\left(\frac{f_i}{q' - v_m}\right) * l_c(o_e)$. The end-user is used to deploy the session interaction and decreases the interruption of data. By performing this, the continuity is carried out and shows better efficiency, and in turn, the interruption is decreased. Here, the relationship's determination is addressed to detect the observed data and reduce the discrete data. In this processing, the interaction among the session is evaluated better and deploys the communication. Here, the delay factor is decreased related to assigning weights to the data to differentiate the data. This dependent data provides lesser data allocation compared to explanatory data, and it is computed as $v_i * \left(\frac{v_m + h\prime}{n_0}\right)$. The sessions' activity is used to define the better efficiency and provides the relationship among the sessions. Here, the information rate is improved and shows better efficiency for the proposed work. For this evaluation, the prediction model is developed that deploys the replacement of similar data to decrease the interruption for the varying sessions [Refer to Fig. 9(a) and (b)].

3.4 Information rate

Figure 10 shows that the proposed work's information rate shows a higher value than the existing three methods. Here, the correctness is improved that deploys the prediction of discrete data, and it is denoted as $\left(f_i + \frac{t_e}{o_e}\right) + v_i$. The activity in the sessions is analyzed and identifies the data points by performing correlation. If the correlation satisfies the data, then the replacement is done for similar data that reduces the processing step. This information rate is associated with the LR data points and provides data allocation to the sessions. Here, the discrete data is identified and continuously deploys reliable sessions. In this information, the rate is used to evaluate better efficiency and decrease the interruption. If the interruption is decreased, then the information rate is improved for the proposed work's varying sessions. Here, the relationship is used to address the explanatory and dependent data related to identifying the discrete data. The weights are assigned to the data and provide better processing of the relationship that deploys the prediction. Here, the preceding data is matched with the forthcoming data and evaluates the better information rate. In this processing, the data's correctness is improved and shows reliable correlation and differentiation of sessions. The comparative analysis is presented in Tables 3 and 4 for sessions and information rate.

The proposed model improves efficiency by 9.3% with the information rate of 8.68%, further it reduces delay by 11.98% and interrupt factor by 11.59%.

FODAM improves efficiency by 13.83% and reduces delay and interrupt factor by 10.59% and 11.08%.

4 Conclusion

This paper introduces a flexible observation data analytics model to improve smart education data concentric sessions' efficiency. The significance of the observed and aggregated data is verified for providing better correctness. The precise session information is examined for its actual requirement, and the interaction is identified. In this interaction identification process, linear regression learning

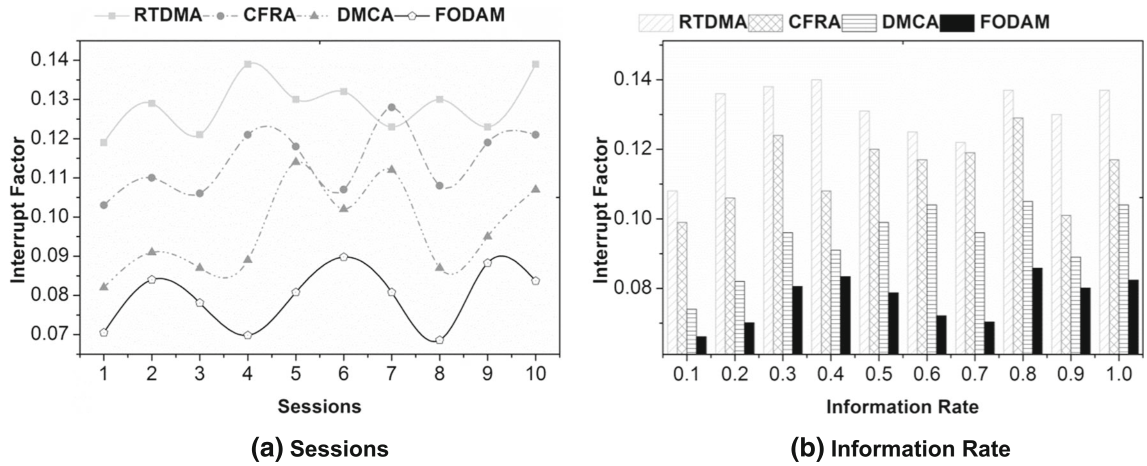


Fig. 9 Interrupt factor for **a** sessions, **b** information rate

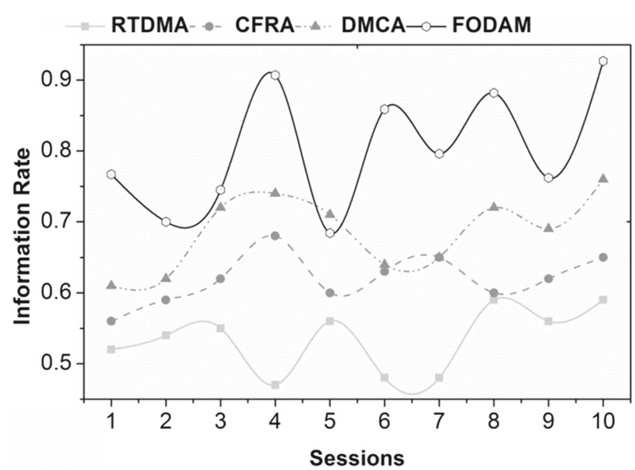


Fig. 10 Information rate for sessions

Table 3 Comparative analysis for sessions

Metrics	RTDMA	CFRA	DMCA	FODAM
Delay (s)	7.06	6.28	4.63	3.837
Efficiency %	89.93	91.33	92.56	94.384
Interrupt factor	0.139	0.121	0.107	0.0837
Information rate	0.59	0.65	0.76	0.927

Table 4 Comparative analysis for information rate

Metrics	RTDMA	CFRA	DMCA	FODAM
Delay (s)	5.86	4.4	2.43	2.886
Efficiency %	86.99	89.72	92.01	94.183
Interrupt factor	0.137	0.117	0.104	0.0824

is used. The linear regression process predicts the need for interaction and data requirement based on its correctness. Depending on this factor, the best-fit solutions are obtained by mitigating the interrupt data. The non-discrete data is exploited based on the relationship between the data and hence the error functions. The replacement without error reduces the impact of delay and improves the information rate in different sessions. The performance assessment shows that the proposed model improves efficiency and information rate by reducing delay and interrupts.

Authors contributions KA contributed to conception and design of study. BM was involved in acquisition of data. KA contributed to analysis and/or interpretation of data. MRP was involved in drafting of the manuscript.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

Aguilar J, Sánchez M, Cordero J, Valdiviezo-Díaz P, Barba-Guamán L, Chamba-Eras L (2018) Learning analytics tasks as services in smart classrooms. Univ Access Inf Soc 17(4):693–709

Akhrif O, Benfares C, Hmina N (2020) Collaborative approaches in smart learning environment: a case study. Procedia Comput Sci 175:710–715

Baig MI, Shuib L, Yadegaridehkordi E (2020) Big data in education: a state of the art, limitations, and future research directions. Int J Educ Technol High Educ 17(1):1–23

Bajaj R, Sharma V (2018) Smart education with artificial intelligence based determination of learning styles. Procedia Comput Sci 132:834–842

Cerezo R, Bogarín A, Esteban M, Romero C (2020) Process mining for self-regulated learning assessment in e-learning. J Comput High Educ 32(1):74–88

Fang CY, Luo AC, Deng YS, Lu CJ, Chen SW (2019) Building a smart lecture-recording system using MK-CPN network for heterogeneous data sources. Neural Comput Appl 31(8):3759–3777

Farhan M, Jabbar S, Aslam M, Ahmad A, Iqbal MM, Khan M, Maria MEA (2018) A real-time data mining approach for interaction analytics assessment: IoT based student interaction framework. Int J Parallel Programm 46(5):886–903

Freigang S, Schlenker L, Köhler T (2018) A conceptual framework for designing smart learning environments. Smart Learn Environ 5(1):27

Hernandez-de-Menendez M, Díaz CAE, Morales-Menendez R (2020) Engineering education for smart 4.0 technology: a review. Int J InteractDes Manuf (IJIDeM) 14(3):789–803

Jin N, Yang F, Yan M, Feng Y, Zhuang Y, Liu H, Kwok PK, Wen K (2019) User perceptions of smart class services in teaching and learning interactions. Procedia CIRP 83:785–788

Kausar S, Huahu X, Hussain I, Wenhao Z, Zahid M (2018) Integration of data mining clustering approach in the personalized E-learning system. IEEE Access 6:72724–72734

Kim Y, Soyata T, Behnagh RF (2018) Towards emotionally aware AI smart classroom: current issues and directions for engineering and education. IEEE Access 6:5308–5331

Kumar KS, Anbarasi M, Shanmugam GS, Shankar A (2020) Efficient predictive model for utilization of computing resources using machine learning techniques. In: 2020 10th International conference on cloud computing, data science & engineering (confluence) 2020 Jan 29 (pp. 351–357). IEEE

Lara JA, De Sojo AA, Aljawarneh S, Schumaker RP, Al-Shargabi B (2020) Developing big data projects in open university engineering courses: lessons learned. IEEE Access 8:22988–23001

Lin J, Pu H, Li Y, Lian J (2018) Intelligent recommendation system for course selection in smart education. Procedia Comput Sci 129:449–453

Lister PJ (2018) A smarter knowledge commons for smart learning. Smart Learn Environ 5(1):8

Liu X (2019) A collaborative filtering recommendation algorithm based on the influence sets of e-learning group's behavior. Clust Comput 22(2):2823–2833

Lnenicka M, Kopackova H, Machova R, Komarkova J (2020) Big and open linked data analytics: a study on changing roles and skills in the higher educational process. Int J Educ Technol High Educ 17(1):1–30

Manogaran G, Lopez D (2018) A Gaussian process based big data processing framework in cluster computing environment. Clust Comput 21(1):189–204

Nieto Y, García-Díaz V, Montenegro C, Crespo RG (2019) Supporting academic decision making at higher educational institutions using machine learning-based algorithms. Soft Comput 23(12):4145–4153

Prathik A, Uma K, Anuradha J (2016) Particulate matter on human health and their feasibility study using machine learning algorithms. J Chem Pharm Res 8(9):260–264

Raizada S, Mala S, Shankar A (2020). Vector borne disease outbreak prediction by machine learning. In: 2020 International conference on smart technologies in computing, electrical and electronics (ICSTCEE) 2020 Oct 9 (pp. 213–218). IEEE

Sathishkumar VE, Rahman AB, Park J, Shin C, Cho Y (2020). Using machine learning algorithms for fruit disease classification. In: Basic & clinical pharmacology & toxicology 2020 Apr 1 (Vol. 126, pp. 253–253). 111 River ST, Hoboken 07030-5774, NJ USA: Wiley

Şerban C, Ioan L (2020) QLearn: towards a framework for smart learning environments. Procedia Comput Sci 176:2812–2821

Şerban C, Todericiu IA (2020) Alexa, What classes do I have today? The use of artificial intelligence via smart speakers in education. Procedia Comput Sci 176:2849–2857

Shen Y, Heng R, & Qian D (2020) Smart classroom learning atmosphere monitoring based on FPGA and convolutional neural network. Microprocess Microsyst p. 103488

Shorfuzzaman M, Hossain MS, Nazir A, Muhammad G, Alamri A (2019) Harnessing the power of big data analytics in the cloud to support learning analytics in mobile learning environment. Comput Hum Behav 92:578–588

Tarus JK, Niu Z, Mustafa G (2018) Knowledge-based recommendation: a review of ontology-based recommender systems for e-learning. Artif Intell Rev 50(1):21–48

Wang S, Zhou A, Hsu CH, Xiao X, Yang F (2015) Provision of data-intensive services through energy-and QoS-aware virtual machine placement in national cloud data centers. IEEE Trans Emerg Top Comput 4(2):290–300

Williamson B (2018) The hidden architecture of higher education: building a big data infrastructure for the "smarter university." Int J Educ Technol High Educ 15(1):1–26

Wongthongtham P, Chan KY, Potdar V, Abu-Salih B, Gaikwad S, Jain P (2018) State-of-the-art ontology annotation for personalised teaching and learning and prospects for smart learning recommender based on multiple intelligence and fuzzy ontology. Int J Fuzzy Syst 20(4):1357–1372

Yang S, Liu H, Zhang Y, Arndt T, Hofmann C, Häfner B, Lanza G (2020) A data-driven approach for quality analytics of screwing processes in a global learning factory. Procedia Manuf 45:454–459

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.