

Hybrid Machine Learning-based Optimization for Student Academic Performance Forecasting: An E-learning Data Mining Approach

Surapaneni Phani Praveen ¹, Naga Satya Koti Mani Kumar Tirumanadham ², Sri Krishna Adusumalli ³,
Deshinta Arrova Dewi ⁴, Tri Basuki Kurniawan ^{5,*}, and Sateesh Gorikapudi ⁶

¹ Department of Computer Science and Engineering, Prasad V Potluri Siddhartha Institute of Technology,
Kanuru, Vijayawada, India

² School of Computer Science and Engineering, VIT-AP University, Amaravathi, India

³ Department of Artificial Intelligence, Shri Vishnu Engineering College for Women, Bhimavaram India

⁴ Center for Data Science and Sustainable Technologies, INTI International University, Nilai, Malaysia

⁵ Post Graduate Program, Universitas Bina Darma, Palembang, Indonesia

⁶ Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation,
Green Fields, Vaddeswaram, Andhra Pradesh, India

Email: sppraveen@pvpsiddhartha.ac.in (S.P.P.); manikumar1248@gmail.com (N.S.K.M.K.T.);

srikrishna.au@gmail.com (S.K.A.); deshinta.ad@newinti.edu.my (D.A.D.);

tribasukikurniawan@binadarma.ac.id (T.B.K.); sateesh4u.325@gmail.com (S.G.)

*Corresponding author

Abstract—The rapid expansion of e-learning systems has transformed modern education by generating large volumes of student interaction data, creating new opportunities for monitoring and predicting academic performance. However, accurately measuring student achievement in digital learning environments remains challenging due to high-dimensional data, complex learning behaviors, and feature redundancy. This study proposes a hybrid machine learning-based framework for forecasting student academic performance using e-learning data mining. Student achievement is measured in terms of predicting and classifying learners' academic performance levels based on their behavioral, demographic, and academic interaction data. The proposed framework integrates data preprocessing, optimized feature selection, and predictive modeling to enhance accuracy and robustness. An optimization-driven feature selection strategy is employed to identify the most informative attributes, while a hybrid deep learning model is used to capture nonlinear and temporal learning patterns. Experimental evaluation conducted on a benchmark e-learning dataset demonstrates that the proposed approach achieves superior predictive performance compared to existing methods across standard evaluation metrics, including accuracy, precision, recall, F1-Score, and Area Under the Curve (AUC). Although the framework introduces additional computational complexity and is validated on a single dataset, the results indicate its strong potential for supporting early identification of at-risk students and enabling data-driven academic interventions in e-learning environments.

Keywords—educational access, process innovation, inclusive education, learning opportunities, e-learning, student

academic performance, hybrid cheetah-based cat swarm optimization

I. INTRODUCTION

The rapid expansion of e-learning platforms has fundamentally transformed modern education by enabling flexible access to learning resources and generating large volumes of student interaction data. While these digital environments offer significant opportunities for personalized learning and continuous assessment, educational institutions continue to face major challenges in accurately monitoring, analyzing, and predicting student academic performance in online settings [1, 2]. The inability to identify students at risk of poor performance at an early stage often limits the effectiveness of timely academic interventions and data-driven decision-making [3, 4].

Existing research in e-learning analytics has extensively explored machine learning and data mining techniques to predict student achievement using demographic, behavioral, and academic interaction data. Although these approaches have demonstrated promising results, several critical issues remain unresolved. Many studies rely on manually selected or generic feature sets, leading to feature redundancy, increased computational complexity, and unstable predictive performance [5]. Furthermore, traditional machine learning models often struggle to capture the nonlinear and temporal learning

behaviors inherent in e-learning data, resulting in limited generalization across diverse educational contexts.

Recent advances in deep learning have improved the ability to model complex learning patterns; however, deep learning-based approaches introduce new challenges, including overfitting, high training cost, and sensitivity to hyperparameter selection—particularly when applied to high-dimensional and imbalanced educational datasets. In parallel, optimization techniques have been applied to improve learning performance, but in most existing studies, optimization is treated as an auxiliary component rather than being fully integrated into the predictive learning pipeline [6, 7]. As a result, there remains a lack of a unified framework that simultaneously addresses feature selection, predictive accuracy, and convergence stability in e-learning performance analytics.

This study aims to bridge this gap by proposing a Hybrid Machine Learning-based Framework for Student Achievement Forecasting (HML-FSA) that integrates optimized feature selection and predictive modeling within a single architecture. The framework employs an optimization-driven feature selection strategy to identify the most informative academic and behavioral attributes, followed by an advanced deep learning-based prediction model capable of capturing nonlinear and temporal learning dynamics [8]. To enhance robustness and generalization, a hybrid optimization mechanism is incorporated into the learning process to mitigate overfitting and improve convergence behavior.

The main contributions of this work are threefold: (i) identification of key limitations in existing e-learning performance prediction approaches, (ii) development of an integrated hybrid framework that combines feature optimization and predictive learning, and (iii) comprehensive experimental validation on a benchmark e-learning dataset to demonstrate improved predictive performance over existing methods. By addressing these challenges, the proposed approach provides a scalable and reliable solution for early identification of at-risk students and supports effective, data-driven academic intervention in e-learning environments [9].

Despite these advances, several fundamental challenges remain unresolved in existing e-learning performance prediction research. Most current approaches address feature selection, predictive modelling, and optimization as isolated components, resulting in feature redundancy, unstable convergence, and limited robustness when applied to high-dimensional and imbalanced educational data. Furthermore, many predictive models fail to effectively capture the nonlinear and temporal nature of student learning behaviours, which is essential for accurate academic performance forecasting in dynamic e-learning environments. These limitations highlight a clear scientific gap: the absence of an integrated and unified framework that simultaneously optimizes feature relevance, predictive accuracy, and learning stability for student performance analysis.

A. Research Questions

The study is guided by the following research questions, which are directly addressed in the Results and Discussion section:

RQ1: Does the proposed hybrid machine learning framework improve the accuracy of student academic performance prediction using e-learning data?

RQ2: How does the Multi-verse Chaotic Krill Herd optimization algorithm affect feature selection efficiency and overall predictive performance?

RQ3: To what extent does Recurrent Quantum Diagonal Deep Neural Network-Hybrid Cheetah based Cat Swarm Optimization (RQDNN-HC2SO) model enhance robustness and reduce prediction loss compared with existing machine learning and deep learning methods?

RQ4: How does the proposed approach perform relative to state-of-the-art techniques in terms of standard evaluation metrics such as accuracy, precision, recall, F1-Score, AUC, sensitivity, and specificity?

B. Contributions of the Study

To address the identified research gap, this study proposes a hybrid machine learning-based framework for forecasting student academic performance in e-learning environments. The core contribution of this work lies in the unified integration of optimized feature selection and advanced predictive modeling within a single analytical pipeline. By employing an optimization-driven feature selection strategy, the framework reduces data dimensionality and mitigates feature redundancy, while the predictive model is designed to capture nonlinear and temporal learning behaviours inherent in e-learning data. In addition, the framework incorporates an adaptive optimization mechanism to improve convergence stability and reduce overfitting. Comprehensive experimental evaluation on a benchmark e-learning dataset demonstrates the effectiveness of the proposed approach when compared with existing state-of-the-art methods, highlighting its potential for scalable and data-driven learning analytics and early academic intervention.

The outline of the article is arranged as follows: The relevant works are reviewed in Section II; the proposed approach is detailed in Section III; in Section VI, the performance analysis is made and compared with state-of-the-art works. Finally, the work is concluded in Section V.

II. LITERATURE SURVEY

Research on e-learning analytics has evolved across multiple methodological dimensions, ranging from learner acceptance modeling and personalization strategies to machine learning-based performance prediction and optimization-driven intelligent systems. Although these studies collectively demonstrate the increasing role of artificial intelligence in digital education, the existing literature remains methodologically fragmented and often lacks integration between learner modeling, feature optimization, and predictive performance analysis. Moreover, many prior studies emphasize algorithmic

outcomes without sufficiently addressing robustness, generalizability, interpretability, or practical educational relevance. Therefore, rather than merely summarizing earlier work, this section critically analyzes the methodological strengths, limitations, and unresolved research gaps in existing studies in order to establish the need for the proposed framework.

A. E-Learning Adoption and Acceptance Models

There is a considerable amount of scholarship that dwells on the acceptance of learners of e-learning systems, which is mainly based on the extensions of the Technology Acceptance Model (TAM). A TAM-based framework created by Hanif *et al.* [10] in order to assess the perceptions of digital learners proved useful since it revealed that the perceptions of usefulness and the ease of use play a significant role in determining adoption intentions. Though the study used a fairly large sample and employed statistical modeling, the research was confined to attitudinal variables, and the issue of acceptance to concrete academic results of performance.

In the same manner, Salloum *et al.* [11] performed a massive systematic review of e-learning research studies based on TAM and was able to determine content quality, system quality and information quality as important drivers of adoption. Although this work is highly informative in terms of offering theoretical information, it is very descriptive without discussing prediction modeling and dynamics of learner behavior. Al-Rahmi *et al.* [12] also combined TAM with the theory of innovation diffusion to examine the intention to use e-learning systems, and once again, the model focused on perception and not data-driven academic prediction.

From a critical perspective, these studies are valuable for understanding learner acceptance and system usage intention; however, they contribute only indirectly to the problem of academic performance prediction. Their primary limitation lies in their dependence on attitudinal constructs rather than measurable learning outcomes. As a result, while such models are useful for adoption analysis, they are less effective for identifying at-risk students or supporting predictive intervention strategies in data-rich e-learning environments.

B. Personalization and Learning-Style Detection Approaches

Another significant stream of research emphasizes personalized learning through the identification of learning styles and adaptive content delivery. El Aissaoui *et al.* [13] proposed a learning-style detection framework based on web mining and fuzzy classification to reduce uncertainty in modeling learner behavior. While the approach demonstrated improved classification accuracy, it relied on predefined and non-standardized learning-style categories, thereby limiting its generalizability across diverse educational contexts.

Similarly, Azzi *et al.* [14] employed Fuzzy C-Means clustering to infer learning styles from student interaction logs, reporting enhancements in learner motivation and engagement. However, the proposed model was inherently static and lacked the ability to adapt to evolving

learner behaviors over time. Rajkumar and Ganapathy [15] introduced a bio-inspired learning framework incorporating brain-computer interfaces and chatbot-based interaction, achieving higher classification accuracy. Despite its effectiveness, the framework faced notable challenges related to scalability, system complexity, and deployment feasibility, particularly in mobile and resource-constrained learning environments.

Overall, while these studies underscore the importance of personalization in e-learning systems, they remain weakly connected to robust academic performance prediction. Most approaches prioritize classification accuracy in isolation and overlook critical factors such as feature redundancy, temporal learning dynamics, and real-world applicability, thereby limiting their effectiveness for scalable and data-driven academic performance forecasting.

A broader critical observation is that many personalization-oriented studies focus on adaptive content delivery or learner categorization rather than predictive academic analytics. Consequently, although personalization improves learner engagement, it does not necessarily translate into reliable forecasting of academic outcomes. This reveals a conceptual gap between adaptive learning system design and predictive performance modeling, which remains insufficiently addressed in the existing literature.

C. Machine Learning and Clustering-Based Performance Prediction

An increasing number of studies have applied machine learning techniques to predict student performance using behavioral and academic attributes. Kausar *et al.* [16] integrated clustering with big data analytics to personalize e-learning systems, reporting improvements in learner engagement and system efficiency. However, the proposed approach relied heavily on heuristic clustering strategies and lacked explicit mechanisms for modeling the progression of learners' academic capacity over time.

Similarly, Wan and Niu [17] introduced a hybrid filtering-based recommendation strategy to address cold-start and data sparsity issues by modeling learner influence. While effective for content recommendation, the approach primarily focused on resource suggestion and did not directly address academic performance prediction or learning outcome assessment. Mangaroska *et al.* [18] developed cross-platform learning analytics architectures using classification and regression models to analyze learning behaviours across multiple e-learning systems. Although their framework enhanced learning design and behavioural insight, it faced challenges related to feature extraction, data scalability, and consistent performance across heterogeneous datasets.

Overall, these studies reveal a persistent limitation in the reliance on manual or heuristic feature selection, which often leads to feature redundancy, increased computational complexity, and unstable predictive performance when applied across different e-learning contexts. This shortcoming underscores the need for automated and optimization-driven feature selection

mechanisms to support robust and scalable student performance prediction.

D. Bio-Inspired and Optimization-Driven Frameworks

In order to address the constraints of traditional machine learning systems, contemporary studies have delved into bio-inspired and optimization-based systems, especially in learning-style classification and adaptive e-learning system [17]. These strategies are used to improve feature representation, learning and classification strength. Nonetheless, most of the available frameworks use optimization as a facilitating tool but not as part of the learning pipeline.

In addition, optimization procedures in previous investigations often focus on accuracy of classification

and provide adequate examination of convergent behavior, balance of exploration and exploitation as well as high-dimensional or unbalanced e-learning information [16, 18]. Consequently, improved models of learning tend to be vulnerable to data variability and data scaling limitations.

Notably, the majority of the optimization-based researches focus on the effectiveness of algorithms rather than pedagogical appropriateness, and seldom consider the connection between selected aspects and learned representations and the existing learning theories of self-regulated learning or constructivist engagement models are illustrated in Table I.

TABLE I. LITERATURE ANALYSIS BASED ON THE E-LEARNING

Reference	Method/ Algorithm	Performance metrics	Limitations
Hanif <i>et al.</i> [10]	Technology Acceptance Model (TAM)	Convenient and high in performance	The external factors are implemented in any contexts
Salloum <i>et al.</i> [11]	Technology Acceptance Model (TAM)	The accuracy is increased	The present content does not sufficiently interact with the learners
Al-Rahmi <i>et al.</i> [12]	Technology Acceptance Model (TAM)	Quick and efficient	The reliability is determined to indicate the existence of e-learning
El Aissaoui <i>et al.</i> [13]	learning style detection approach	effective	There is an absence of a standardized learning style model
Azzi <i>et al.</i> [14]	Fuzzy C Means (FCM) algorithm	Increases the performance	unable to modify its own learning style
Rajkumar and Ganapathy [15]	Bio-Inspired learning style Brain Computing Interface (BIL-BCI)	Enhance accuracy	Smart device structure must be replaced
Kausar <i>et al.</i> [16]	adaptive data mining clustering approach	Efficient and robust	There is a lack of learning capacity
Wan and Niu [17]	Hybrid Filtering (HF) method	Effective, great adaptability, efficient	The information overload problems must be rectified
Mangaroska <i>et al.</i> [18]	machine learning method	Quality is increased	The size of the data is the weakness

III. PROPOSED METHODOLOGY

The work in this article is based on the hybrid optimization-based machine learning approach to analyze the performance of the students and therein sophisticatedly predict the nature of the respective students. The proposed methodology is illustrated well in Fig. 1.

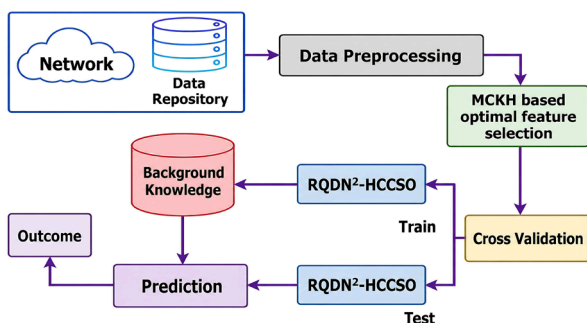


Fig. 1. Schematic diagram of the proposed methodology.

A. Data Collection

The student performance assessment is the major role of this study. An online learning management model and record the participation. For this study, the schools were unable to provide student dataset details due to confidentiality concerns. Hence, the data used for the

prediction of the performances of students were collected from an enormous number of websites, synthetic data, and real-time data. In this study, the data collected from <https://www.kaggle.com/datasets/aljarah/xAPI-Edu-Data> includes several data related to students' academic performances and students' data. Therefore, the collected data are in an imbalanced state and also possess wrongly classified instances. In context with this, it is ineluctable to normalize the data with the pre-processing steps. Students' Academic Performance Dataset (xAPI-Edu-Data) is a multivariate educational dataset that was collected in 2016 in the Kalboard 360 Learning Management System (LMS) with the help of the Experience API (xAPI) to follow the learning activities and interactions of students. The dataset consists of 480 student records with 16 attributes comprising both integer and nominal variables. It is primarily used for classification tasks in educational data mining and predictive modeling. The attributes are grouped into three categories: demographic features (e.g., gender and nationality), academic background features (e.g., educational level, grade level, and section), and behavioral features (e.g., classroom participation, resource utilization, attendance, parental survey responses, and school satisfaction). The dataset is complete and has students with varied countries, sex, semesters and so, which is fitted in the study of learning behaviors and forecasting performance of students in e-learning settings.

B. Data Pre-processing

Data pre-processing is carried out by Data Filtration. Preprocessing steps include removal of missing values, categorical encoding, normalization, and filtration of inconsistent records prior to feature selection. This step is effectuated to remove the empty values or null values and it can be achieved by the inbuilt function in MATLAB such as $y = \text{filter}(b, a, x)$, $y = \text{filter}(b, a, x, z_i)$, $y = \text{filter}(b, a, x, z_i, \text{dim})$, $[y, z_j] = \text{filter}(__)$.

C. Feature Selection Using MCKH Model

Next, we developed the Multi-Verse Chaotic Krill Herd Optimization (MCKHO) algorithm for optimal best feature selection among multiple features which reduces the data dimensionality problem. MCKHO algorithm is used for the selection of optimal features from the pre-processed data based on students' academic performances. The commonly researched marine species are Antarctic krill. The representative distribution regards the several uncertainties that characterize the krill herd [19]. Three aspects such as krill motion, foraging characteristics and random diffusion affect the individual krill position.

The integration of Krill Herd Optimization, Multi-verse Optimization, and chaotic maps within the proposed MCKHO algorithm is motivated by their complementary strengths. Krill Herd Optimization demonstrates strong capability in local exploitation, whereas Multi-verse Optimization excels in global exploration through the use of wormhole tunneling mechanisms. In addition, chaotic maps introduce deterministic randomness that helps prevent premature convergence and enhances search diversity. By synergistically combining these components, the MCKHO algorithm effectively balances the exploration-exploitation trade-off, which is critical for identifying the most informative features from high-dimensional e-learning datasets.

1) Krill motion

Each individual based on herd density maintenance is referred to by induced movement. Eq. (1) represents a mathematical model of krill motion

$$M_j(T+1) = M_{\max} \beta_j + \gamma_m M_j(T) \quad (1)$$

The maximum krill speed and inertial weights are $M_{\max}$ and β_j^{local} . The j -th individual based neighbors local effect is β_j^{local} . Eq. (2) calculates the best solution β_j^{target} from the j -th individuals.

$$B_j^{\text{target}} = D^{\text{best}} k_{j, \text{best}} Y_{j, \text{best}} \quad (2)$$

The target value for the j -th individual is β_j^{target} and the coefficient is D^{best} is illustrated in Eq. (3).

$$D^{\text{best}} = 2 \left(\frac{R+1}{\text{Max}_{\text{iteration}}} \right) \quad (3)$$

The random variable R tends from 0 to 1 interval.

2) Foraging characteristics

Eq. (4) formulates the mathematical expression of the foraging behavior of krill.

$$FR_j(T+1) = FS_f \alpha_j + \gamma_f FR_j(T) \quad (4)$$

The foraging inertia weight and speed are γ_f and FS_f . The best food attractions are A_j^{best} and $\alpha_j = \alpha_j^{\text{food}} + A_j^{\text{best}}$.

3) Wormhole tunnels

The global optimum is determined by requiring these two steps. It never ascertains the trade-off between exploration and exploitation [20]. Hence, we have used the wormhole tunnels step of the multi-verse algorithm to replace the random diffusion step of Chaotic Krill Herd Optimization (CKH) optimization. Then, the newly developed model is named as MCKHO algorithm. Wormholes enhance the inflation rate and it contains huge probability for each universe in order to offer local changes, which are explained as in Eq. (5):

$$Y_j^k = \begin{cases} Y_k + DRT \times (BU_k - BL_k) \times md + BL_k, & md < 0.5 \\ Y_k - DRT \times (BU_k - BL_k) \times md + BL_k, & md \geq 0.5 \end{cases} \quad (5)$$

where, DRT is the constant coefficient. The student's academic features based on k th parameter is Y_k . The upper and lower bounds are BU_k and BL_k in which the random number is rnd .

4) Chaotic map

We have selected 3 chaotic maps including sinusoidal, sine and cosine in order to enhance the chaotic krill optimization algorithm performance. The exploitation and exploration are two parts to divide the similar search scheme in all meta-heuristic algorithms [21]. Additional solutions are generated by evaluating search space during the phase of exploration. The global optimum with respect to rapid convergence is comprised of the exploitation phase. The Eq. (6) calculates the random value and the chaotic maps substitute the parameter R .

$$\text{Best}_c = 2 \left(C(T) + \frac{1}{\text{Max}_T} \right) \quad (6)$$

where, Max_t is the maximum number of iterations to attain the optimal best solution and Best_c is the optimal best solution to select the student academic performance data features from the pre-processing stage. The pseudocode of the MCKHO algorithm for selecting features is displayed in Algorithm 1. The feature selection is an important step to neglect the higher dimensionality data. For further process, this MCKHO algorithm selects the eight important features such as Gender, Nationality, "StageID", "SectionID", "Topic", "raised hands", "VisITedResources", "AnnouncementsView", "Discussion", "ParentAnsweringSurvey", "ParentschoolSatisfaction", "StudentAbsenceDays".

Algorithm 1. Pseudocode of MCKHO Algorithm

```

1 Initialize the parameters of MCKHO and maximum
  iteration with student academic features
2 Use a random way to set the chaotic maps
3 Update the inertia weights  $\beta_j^{local}$ 
4 Compute the fitness function
5 While  $m < M_{max}$ 
6   else
7     Stopping criterion not met
8   do
9     From best to worst initiate the population
10    For  $j = 1; N_p$ 
11    do
12      Update the krill movement using Eq. (1)
13      Update foraging behaviour using Eq. (4)
14      Use Eq. (5) to update the wormhole tunnel model
        based on multi-verse
15      Update the chaotic map model via Eq. (6)
16    end for
17    The new bets values are stored
18     $m = m + 1$ 
19  end while
20 Obtains less dimensionality and optimized features

```

D. Recurrent Quantum Diagonal Deep Neural Network (RQDNN)

After optimal feature selection, predicting student behaviour based on academic performance becomes a crucial task [22]. In this context, it is important to clarify that the proposed Recurrent Quantum Diagonal Deep Neural Network (RQDNN) is not a physical quantum computing system, but rather a quantum-inspired learning model designed for classical computation. The incorporation of quantum mechanical concepts such as superposition and rotation gates serves as a mathematical abstraction to enhance feature representation and introduce richer nonlinear decision boundaries when processing classical e-learning data. In particular, RQDNN encodes neuron activations as phase-based probability amplitudes, enabling the model to capture multiple learning states simultaneously. This capability allows the network to better respond to subtle variations in student behaviour and interaction patterns over time.

The theoretical motivation for adopting the RQDNN architecture lies in the need to model the highly nonlinear, heterogeneous, and interdependent relationships that exist among student demographic, behavioural, and academic features in e-learning environments. Traditional shallow classifiers and standard feedforward neural networks often face limitations in effectively capturing subtle interactions among these variables, particularly when student performance is influenced by multiple overlapping learning indicators. Therefore, a more expressive predictive architecture is required to represent these complex dependencies with greater discriminative capability.

The quantum-inspired component of the proposed model is introduced as a representational enhancement mechanism rather than a claim of physical quantum computation. In this framework, the concept of superposition is used as a mathematical abstraction to

encode multiple latent learning states within the same feature space, thereby enriching internal feature representation. This allows the model to preserve and process complex feature interactions more effectively than conventional scalar neuron activations. As a result, the RQDNN improves nonlinear separability between student performance classes and strengthens the model's ability to distinguish between poor, average, and excellent academic outcomes.

Furthermore, the recurrent structure of the RQDNN is theoretically justified by the interrelated and behaviour-dependent nature of student learning data, where academic outcomes are influenced by cumulative engagement indicators such as participation, resource access, discussion activity, and attendance-related behaviour. Although the dataset is tabular in form, the selected features collectively reflect progressive learning behaviour and can therefore benefit from recurrent dependency modeling. In addition, the diagonal transformation mechanism is incorporated to reduce parameter redundancy and improve learning stability by emphasizing feature-wise transformations while preserving computational efficiency. Consequently, the proposed RQDNN offers a balanced architecture with enhanced representation capacity, improved generalization, and stronger suitability for student academic performance prediction.

The enhanced representational capacity of the quantum-inspired layer improves the model's ability to discriminate between different performance classes, without claiming any physical quantum advantage. While conventional deep learning approaches, particularly recurrent neural networks, are commonly employed for student performance prediction, they may still be limited in representational flexibility when dealing with complex and overlapping educational features. In contrast, the integration of quantum-inspired mechanisms with a recurrent deep neural architecture enhances internal feature encoding and supports more accurate and robust prediction. The superior performance of the proposed RQDNN compared to a traditional recurrent deep neural network is demonstrated in the subsequent analysis. In quantum-inspired computation, the fundamental unit is the qubit, which serves as an abstract counterpart to the binary unit in conventional computation. Based on this abstraction, excitatory and inhibitory neuron states are represented as qubit states $|1\rangle$ and $|0\rangle$, respectively. According to the principle of superposition, the state of a qubit can be mathematically expressed as in Eq. (7),

$$|\theta\rangle = \beta|0\rangle + \chi|1\rangle \quad (7)$$

Probability amplitudes are defined as β and χ based on complex numbers and can be given as in Eq. (8),

$$|\beta|^2 + |\chi|^2 = 1 \quad (8)$$

Quantum gates are utilized for the general realization of n qubits attained by the quantum system-based quantum operation. Moreover, the computation efficiency can be

enhanced with the quantum parallelism based updated with quantum operation with simultaneous states. Generally, qubit Rotation gate H and NOT gate V with two controlled qubits. They are defined as in Eqs. (9) and (10):

$$H = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \quad (9)$$

$$V = \begin{bmatrix} \cos\left(\frac{\pi}{2} - 2\phi\right) & -\sin\left(\frac{\pi}{2} - 2\phi\right) \\ \sin\left(\frac{\pi}{2} - 2\phi\right) & \cos\left(\frac{\pi}{2} - 2\phi\right) \end{bmatrix} \quad (10)$$

To simplify the process, it is necessary to express the quantum state equation as follows in which the real and the imaginary parts denote the probability amplitude as illustrated in Eq. (11)

$$|\phi\rangle = f(\phi) = e^{i\phi} = \cos \phi + i \sin \phi \quad (11)$$

The quantum neuron can be evolved with the two qubits controlled-NOT gate and a single qubit Rotation gate is formulated as in Eqs. (12) and (13).

$$H|\phi\rangle = \begin{bmatrix} \cos(\theta + \phi) \\ \sin(\theta + \phi) \end{bmatrix} = f(\theta + \phi) = f(\theta)f(\phi) = e^{i(\theta+\phi)} \quad (12)$$

$$V|\phi\rangle = \begin{bmatrix} \cos\left(\frac{\pi}{2} - 2\phi\right) \\ \sin\left(\frac{\pi}{2} - 2\phi\right) \end{bmatrix} = f\left(\frac{\pi}{2} - 2\phi\right) = [\sin \phi + i \cos \phi]f(\theta - 1) \quad (13)$$

The architecture of the proposed RQDNN is illustrated in Fig. 2. From the architecture, the input samples and hidden layer correlation and relation between hidden and output layers are formulated as illustrated in Eqs. (14) and (15):

$$y_j = \sin\left(\frac{\pi}{2} f(y_j) - \arg(\sum_{i=1}^n H(\theta_j)|x_j)\right) \quad (14)$$

$$C = f\left(\sum_{j=1}^m y_{i,j} x_{i,j}\right) - f\left(\sum_{j=1}^m y_{i,j} \sin\left(\frac{\pi}{2} f(y_j) - \arg(\sum_{i=1}^n H(\theta_j)|x_j)\right)\right) \quad (15)$$

The optimally selected features obtained from the feature selection stage are subsequently forwarded to the Recurrent Quantum Diagonal Deep Neural Network (RQDNN) for performance prediction. The overall architecture of the proposed framework is depicted in Fig. 2. In the RQDNN model, the final layer of the recurrent deep neural network is transformed into a quantum-inspired phase representation, enabling enhanced modeling of complex and nonlinear learning patterns. By leveraging the principles of quantum superposition in a mathematical abstraction, the output layer is evaluated with improved representational capacity.

Furthermore, the RQDNN incorporates an intrinsic mechanism that modifies the conventional loss function by employing a squared reconstruction error, which is particularly effective for handling high-dimensional and complex educational data. This formulation theoretically

supports improved representation learning by penalizing deviations in reconstructed class-related patterns, thereby enabling more stable optimization and enhanced class discrimination. To further mitigate overfitting and enhance convergence stability during training, a Hybrid Cheetah-based Cat Swarm Optimization (HC2SO) algorithm is integrated into the learning process, as described in the following section.

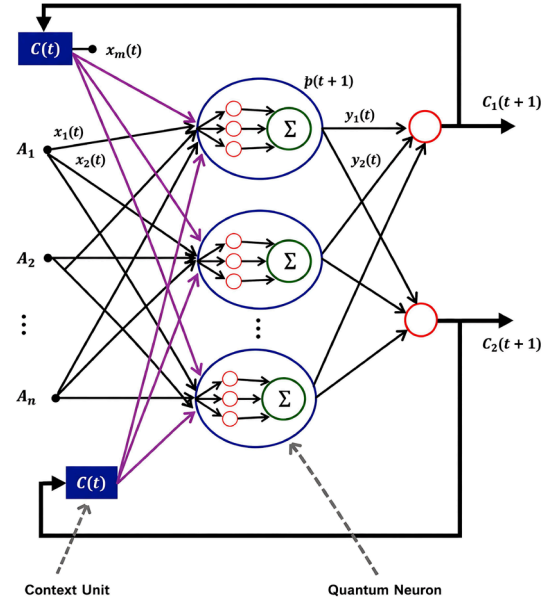


Fig. 2. Architecture of proposed RQDNN.

E. Hybrid Cheetah Based Cat Swarm Optimization (HC2SO) Algorithm

The Hybrid Cheetah-based Cat Swarm Optimization (HC2SO) algorithm is the amalgamation of both the Cheetah Optimization (CO) algorithm and Cat Swarm Optimization (CSO) algorithm in which the steps involved in the HC2SO algorithm are illustrated in the following section.

1) Cheetah Optimization (CO) algorithm

To effectuate prey detection, the Cheetah's scanning or patrolling behavior adopts the global search of node CO. There are three major steps such as an attack, waiting to reach prey closer, and watch and are included in the prey attacking [23]. The breakdown and energy limit reasons stop data transmission. The following section elaborates on attacking, sitting and waiting with searching for prey as the mathematical modeling of the CO algorithm.

- The scanning around the search space determines the prey.
- For the right situation, sit and watch the prey's determination.
- After prey finding, the maximum speed and the cheetah hop to the next position. Reach the position, the next position makes the cheetah's flexible speed.
- When the data cannot be forwarded to the subsequent access point with data transmission in a reasonable amount of time, interaction will be lost.

2) Searching mechanism

For data transmission, the particular data is determined and scanned. Switch on the activated mode and scatter the attendance data. The numerical model in Eq. (16) explains the search mechanism.

$$A_m = A_r + \beta \cdot \sigma \quad (16)$$

During the multi-hop, random parameters β and δ are the step length of the cheetah which A_r and A_m denotes the current and new hop of a cheetah.

3) Siting and waiting mechanism

The sit-and-wait mechanism makes a hop and data might have visible to the cheetah while effectuating a searching mode as in Eq. (17),

$$A_m = A_r \quad (17)$$

Update the current and new cheetah data are A_r and A_m . To avoid premature convergence, do not allow all cheetahs to change the attendance data.

4) Attack mechanism

Cheetahs jump onto the next data using two values such as pace and adaptability. When it is determined to hop to the next data then this is done so at a filled pace. Cheetah's attacking strategy is as in Eq. (18),

$$A_m = A_p + \kappa_1 \cdot \tau \quad (18)$$

For attendance data transmission, the current node and best positions are A_p , where, τ and κ determines the cheetah interaction and tuning factor.

F. Cat Swarm Optimization (CSO) Algorithm

To CSO algorithm adopts the global selection and it involves tracking and seeking modes. Both modes indicate the food searching and resting locations. These modes separate and the N-dimensional space initiated cats [24]. In the tracking mode, consider a few cats and less tracing mode of the cat. In sleeping mode, consider the student's academic performance and in tracing mode considers the active and sleeping modes.

1) Seeking mode (resting)

In the seeking mode (resting state), cats remain stationary while continuously monitoring their surroundings based on the student data representation. During this phase, the algorithm explores the local search space by generating candidate solutions around the current position. The cat's position and movement in the seeking mode are governed by the following parameters:

- Seeking Memory Pool (SMP): Determines the number of candidate copies generated for the current solution to explore nearby positions.
- Seeking Range of the Selected Dimensions (SRD): Defines the extent of variation applied to selected dimensions based on their previous and updated values.

- Counts of Dimension to Change (CDC): Specifies the number of dimensions to be modified during the local search process.
- Self-Position Consideration (SPC): Indicates whether the current position of the cat is retained as one of the candidate solutions.
- When SPC = 1, the current solution is included among the generated candidates, and the remaining positions are produced according to the SMP settings.
- The evaluation of these candidate solutions allows the algorithm to effectively explore the local neighborhood while maintaining convergence stability.

The Eq. (19) defines and evaluates the CDC depending upon the cat copies.

$$B_m = (1 \pm SRD \times \alpha) \times B_p \quad (19)$$

Under the range of 0 to 1 with α , B_p and B_m are the arbitrary number, current hop and next hop.

Depending upon probability properties select the cat when all cats display the same value, after the calculation of the fitness function as in Eq. (20),

$$Probability_i = \frac{|F_i - F_b|}{F_{max} - F_{min}} \quad (20)$$

From this, F_{max} and F_{min} are maximum and minimum fitness values, where, F_i is the fitness value and the current cat probability is $Probability_i$. Set $F_b = F_{min}$ and $F_b = F_{max}$ to a minimum and maximum optimization troubles and adopt the cat hop.

2) Tracing mode

This mode portrays the searching process of prey and the below equation formulates the cat velocity as in Eq. (21).

$$Velocity_i = v_i + \alpha_i \times r(best_B - B_i) \quad (21)$$

Depending upon the range of 0 to 1. α_i is the arbitrary number and v_i is the current cat velocity, where, $best_B$ and B_i are the optimal cat hop and the predefined value as illustrated in Eq. (22).

$$B_{new} = B_{old} + v_i \quad (22)$$

From this, B_{new} and B_{old} are the new and old locations, respectively.

3) Termination criteria

To achieve the optimal answer and yield a good convergence outcome, the termination or stopping condition is preferred. The number of iterations and the procedure runtime are two significant termination criteria.

G. Proposed RQDNN-HC2SO based Prediction of Behavior and Performance of Students

The amalgamation of both CO and CSO algorithms to form a Hybrid Cheetah Cat Swarm Optimization (HC2SO) algorithm. To handle the issues of exploration

and exploitation, the proposed HC2SO algorithm is used to deliver a globally optimal solution. Initially, both CO and CSO algorithm parameters are initialized and evaluate their fitness function [25]. Adaptively, the data location is updated and it has one control parameter. After that, the positions are updated and each function of CO and CSO is updated thereby obtaining the best results. And finally, both solutions are compared and we selected the optimal best solution for further process. Fig. 3 delineates the flow diagram based on the HC2SO algorithm.

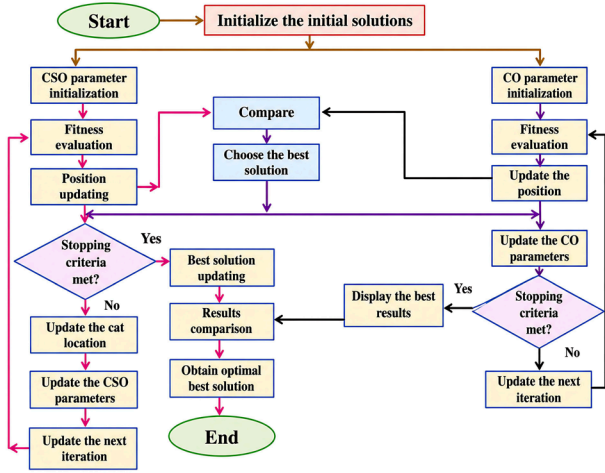


Fig. 3. Flow diagram based on HC2SO.

The proposed approach is based on the RQDNN-HC2SO [26] and is used to predict student performance as poor, good and excellent. It is described as the Grading of academic achievements. The quantum-based approach utilizes the extracted features, inter-behavior correlation features, and student demographic information. The adopted HC2SO algorithm is used to prevent overfitting and to mitigate issues related to loss and data imbalance. Thus, the performances of the students are evaluated based on their academic and behavioral data.

IV. EXPERIMENTAL ANALYSIS

To evaluate the effectiveness of the proposed approach for student performance prediction, multiple standard performance metrics were employed, including Area Under the Curve (AUC), loss, accuracy, precision, recall, sensitivity, specificity, and F1-Score [27]. Additionally, a stepwise complexity validation study was conducted to examine the individual contribution of each component in the proposed framework, namely MCKH-based feature selection, RQDNN-based prediction, and HC2SO-based optimization. The validation was performed using Accuracy, AUC, and training time in order to assess the trade-off between predictive performance and computational complexity. The experiments were conducted using the xAPI-Edu-Data dataset obtained from Kaggle [28], and all simulations were implemented in the MATLAB environment. To ensure a rigorous and fair assessment, the performance of the proposed framework was compared against representative and comparatively simpler baseline methods, namely HF, FCM, and

BIL-BCI. In addition, to directly validate the contribution of the proposed predictive architecture, an ablation-based comparative experiment was conducted between the conventional Recurrent Neural Network (RNN) and the proposed Recurrent Quantum Diagonal Deep Neural Network (RQDNN), with particular emphasis on Accuracy and Area Under the Curve (AUC). These approaches are widely adopted in e-learning analytics for tasks such as clustering, recommendation, and learning-style modeling. This comparative evaluation enables an objective assessment of whether the increased complexity of the proposed model leads to measurable and meaningful performance improvements over established and more interpretable techniques

A. Performance Metrics

The performance metrics are defined as follows:

Accuracy: It is defined as how accurately the student's performances are predicted from the set of data [29].

$$Precision = \frac{T_P + T_N}{T_P + T_N + F_P + F_N}$$

Precision: It is defined as the ratio of faulty identities to the total number of faulty information [30].

$$Precision = \frac{T_P}{T_P + F_N}$$

F-Score: The mean average of recall and precision with evaluated higher matching coefficients with expected efficiency [31].

$$F-Score = \frac{2 \times Precision \times Recall}{Recall + Precision}$$

Recall: It can be defined as the ratio of faulty identities to the total number of a dataset [32, 33].

$$Recall = \frac{T_P}{T_P + F_P}$$

where, T_P implies the true positive, T_N implies the true negative, F_P implies the false positive, and F_N implies the false negative.

B. Performance Assessments

For binary classification and detection tasks, a reasonable metric is an Area Under the Curve (AUC). In binary prediction, different threshold values are aggregated and it is the primary positive feature. From predictive power, the threshold-setting problems are separated. Fig. 4 explains the comparative analysis of AUC [34, 35]. The state-of-art studies such as HF [12], FCM [16] and BIL-BCI [17] with the proposed methodology are applied to demonstrate the effectiveness of AUC. From this plot, the proposed method achieves an AUC value of 0.98, whereas HF, FCM, and BIL-BCI achieve 0.91, 0.88, and 0.85, respectively. The higher AUC value of the proposed framework indicates its

stronger discriminative capability in distinguishing among student performance categories. This is particularly significant because AUC reflects the model's ability to maintain reliable classification performance across varying decision thresholds. Therefore, the superior AUC result suggests that the proposed approach is more dependable for identifying students at different levels of academic performance, including those who may be at risk of underachievement as illustrated in Table II.

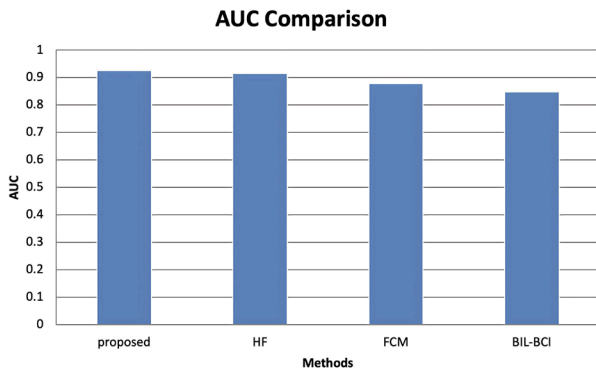


Fig. 4. Comparative study based on AUC analysis.

TABLE II. PERFORMANCE METRICS OF ANALYSIS BASED ON THE E-LEARNING

Metric	Value
AUC	0.98%
Loss	0.08%
Accuracy	0.98%
F-Score	0.98%
Precision	0.94%
Recall	0.93%
Sensitivity	0.95%
Specificity	0.97%

Fig. 5 presents the loss outputs with state-of-the-art studies. The comparison methods in terms of HF [12], FCM [16] and BIL-BCI [17] with the proposed methodology are applied to detect the loss results. From this plot, the proposed framework achieves the lowest loss value of 0.08, compared to 0.15, 0.48, and 0.76 obtained by HF, FCM, and BIL-BCI, respectively. The lower loss value indicates that the proposed model converges more effectively during training and learns more stable predictive representations from the educational dataset. This is significant because reduced loss not only reflects better optimization behaviour, but also suggests improved model reliability and lower prediction uncertainty in practical student performance forecasting scenarios.

Fig. 6 depicts the accuracy assessment comparative analysis. To show the efficacy of accuracy, state-of-the-art studies such as HF [12], FCM [16], and BIL-BCI [17] with the proposed methodology are used. The proposed framework achieves an accuracy of 0.98, outperforming HF (0.93), FCM (0.89), and BIL-BCI (0.86). This improvement is significant because higher accuracy directly reflects the model's effectiveness in correctly classifying student academic performance levels. In the context of educational analytics, such improvement supports more reliable academic monitoring

and increases confidence in deploying the model for real-world student assessment and intervention planning.

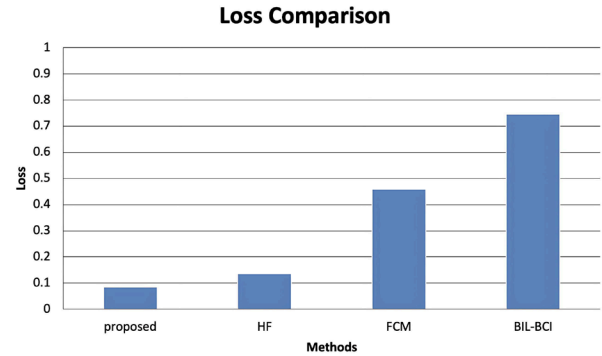


Fig. 5. Comparative study based on loss.

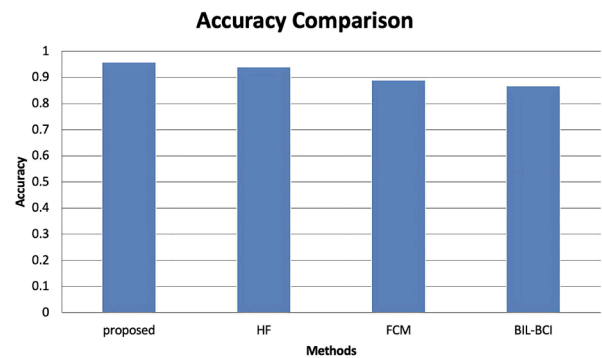


Fig. 6. Comparative study of accuracy evaluation.

The F-Score evaluation based on comparative analysis is depicted in Fig. 7. State-of-the-art research such as HF [12], FCM [16], and BIL-BCI [17] with the proposed methodology is used to demonstrate the adequacy of the F-Score. The F-Score is 0.98%, 0.93%, 0.89%, and 0.86% using the proposed methods, HF, FCM, and BIL-BCI strategies, respectively. The proposed method generates a 0.98% F-Score that is higher than previous methods such as HF, FCM, and BIL-BCI.

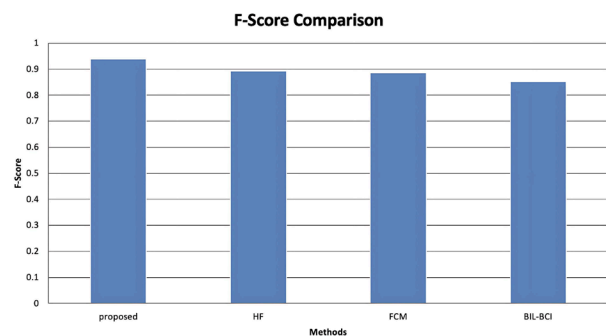


Fig. 7. Comparative study of F-Score evaluation.

Fig. 8 presents the precision-based performance comparison of the proposed approach with existing methods. The precision of the proposed model is evaluated against representative baseline techniques, including HF [12], FCM [16], and BIL-BCI [17]. The obtained precision values for the proposed method, HF, FCM, and BIL-BCI are 0.94, 0.91, 0.86, and 0.84,

respectively. As shown in the results, the proposed scheme achieves the highest precision value of 0.94, outperforming the conventional methods. This improvement indicates the superior capability of the proposed approach in correctly identifying relevant instances compared to existing techniques.

Recall estimation based on the comparative result is depicted in Fig. 9 and is effectuated with state-of-art works such as HF [12], FCM [16], and BIL-BCI [17]. The recall value of the proposed approach is about 0.93 and HF shows 0.89, FCM implies 0.87, and BIL-BCI shows 0.86. Thus, the recall of the proposed approach is greater than other approaches which show the prediction of student performance.

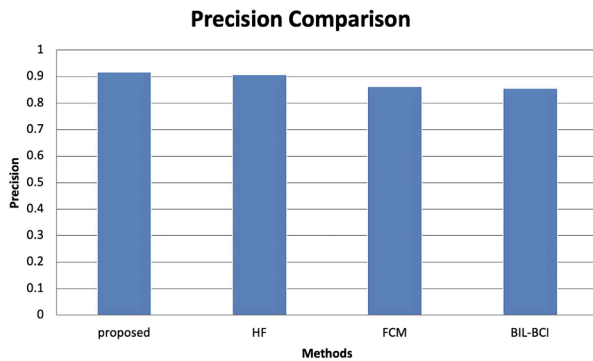


Fig. 8. Comparative study of precision evaluation.

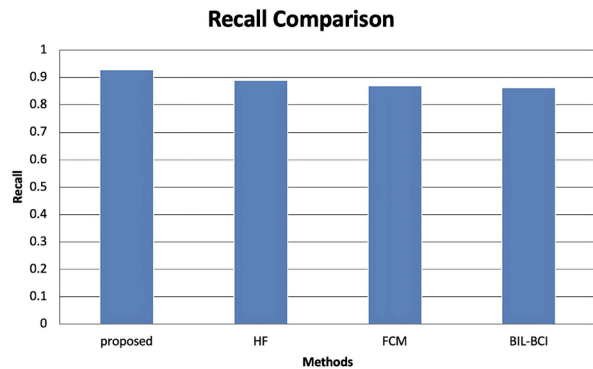


Fig. 9. Comparative study of recall evaluation.

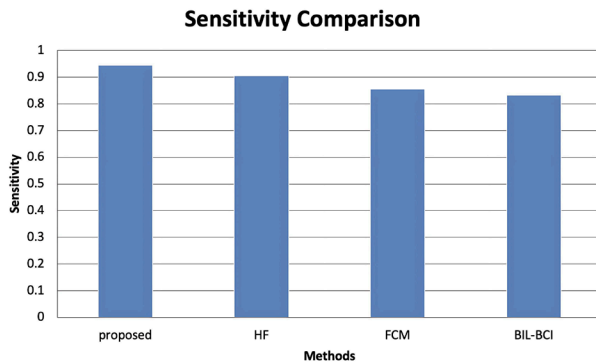


Fig. 10. Comparative study of Sensitivity evaluation.

The comparative comprehensive based on the sensitivity is demonstrated in Fig. 10 with state-of-art works such as HF [12], FCM [16], and BIL-BCI [17]. The sensitivity of the presented work is 0.95, FCM is 0.86, HF

is 0.91, and BIL-BCI is 0.82. Henceforth the comprehensive study based on the specificity is shown in Fig. 11. The proposed work depicts 0.97, HF shows 0.92, FCM shows 0.87, and BIL-BCI shows 0.85 specificity.

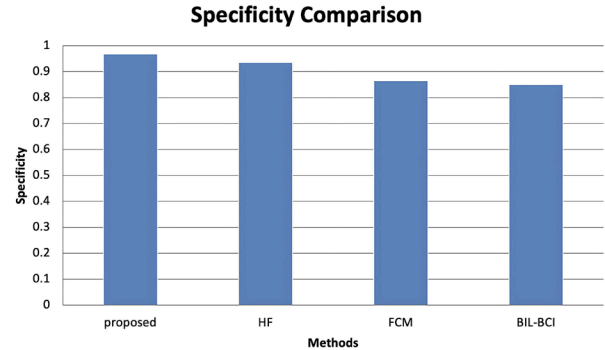


Fig. 11. Comparative study of comprehensive evaluation.

C. Comparative Analysis of RNN and Proposed RQDNN

To further validate the effectiveness of the proposed predictive architecture, an additional comparative experiment was conducted between a conventional Recurrent Neural Network (RNN) and the proposed Recurrent Quantum Diagonal Deep Neural Network (RQDNN). This comparison was performed using the same dataset, preprocessing pipeline, and experimental settings in order to ensure fairness and consistency. The purpose of this experiment is to directly assess whether the quantum-inspired diagonal representation contributes measurable performance improvements beyond a standard recurrent deep learning model. The comparison focused on two key evaluation metrics, namely Accuracy and Area Under the Curve (AUC), as these metrics provide a clear indication of classification effectiveness and discriminative capability. Table III presents the comparative results obtained for both models.

TABLE III. COMPARATIVE ANALYSIS OF PROPOSED APPROACH WITH RNN

Model	Accuracy	AUC
RNN	0.93	0.94
Proposed RQDNN	0.98	0.98

As shown in Table III, the proposed RQDNN outperforms the conventional RNN in terms of both Accuracy and AUC. The observed improvement indicates that the quantum-inspired diagonal representation enhances the model's ability to learn more discriminative feature relationships from student behavioral and academic data. In particular, the higher AUC value demonstrates that the proposed model provides better class separability and stronger predictive discrimination across student performance categories. Similarly, the improvement in Accuracy confirms the effectiveness of the RQDNN in correctly classifying student academic outcomes. These findings empirically support the superiority of the proposed architecture over the standard recurrent neural network baseline.

D. Stepwise Validation of Method Complexity

To further justify the architectural complexity of the proposed framework, a stepwise validation study was performed by incrementally introducing each major component of the model. The purpose of this analysis is to determine whether each added layer of methodological complexity contributes meaningful improvements in predictive performance relative to its computational cost.

Four experimental groups were considered in this validation:

- (i) a conventional baseline Recurrent Neural Network (RNN),
- (ii) the baseline model enhanced with MCKH-based feature selection,
- (iii) the MCKH-enhanced model integrated with the proposed RQDNN prediction architecture, and
- (iv) the complete proposed framework incorporating HC2SO-based optimization.

The performance of each group was evaluated using Accuracy, Area Under the Curve (AUC), and training time. This analysis provides an intuitive understanding of how each component affects both predictive quality and computational overhead.

As shown in Table IV, the stepwise validation clearly demonstrates that each additional component of the proposed framework contributes measurable improvements in predictive performance. The incorporation of MCKH-based feature selection improves both Accuracy and AUC over the baseline model by reducing feature redundancy and enhancing the quality of the input representation. The subsequent replacement of the conventional RNN with the proposed RQDNN architecture yields further performance gains, indicating that the quantum-inspired diagonal representation improves nonlinear feature learning and class discrimination. Finally, the inclusion of HC2SO-based optimization produces the best overall results, confirming its effectiveness in improving convergence behaviour and reducing optimization-related limitations during training.

Although each layer of complexity introduces an increase in training time, the observed performance gains justify the added computational cost. Overall, the results provide intuitive empirical evidence that the complexity of the proposed framework is not arbitrary, but rather functionally contributes to improved prediction performance in student academic forecasting.

TABLE IV. COMPARATIVE ANALYSIS OF PROPOSED APPROACH WITH OTHER BASELINE MODELS

Group	Model Configuration	Accuracy	AUC	Training Time (s)
1	Baseline (RNN)	0.90	0.91	18.6
2	RNN + MCKH	0.93	0.94	22.8
3	MCKH + RQDNN	0.96	0.97	29.5
4	MCKH + RQDNN + HC2SO (Proposed)	0.98	0.98	35.4

V. CONCLUSION

This study presented a hybrid machine learning-based framework for forecasting student academic performance

using e-learning data. The primary objective was to enhance predictive accuracy and robustness by integrating optimization-driven feature selection with an advanced predictive model capable of capturing nonlinear and temporal learning behaviors. The proposed framework was evaluated using a benchmark e-learning dataset, and the experimental results demonstrated strong predictive performance, achieving an accuracy of 0.98, AUC of 0.98, F1-Score of 0.98, precision of 0.94, recall of 0.93, sensitivity of 0.95, specificity of 0.97, and a low loss value of 0.08. These results indicate that the proposed approach provides high discriminative capability for classifying student performance levels. In particular, the strong AUC and recall values suggest that the framework is well suited for the early identification of students at risk of academic underperformance, thereby enabling timely and data-driven interventions in e-learning environments. Comparative analysis against representative baseline methods—including HF, FCM, BIL-BCI, and a conventional Recurrent Neural Network (RNN)—demonstrated consistent performance improvements across all evaluation metrics. These baselines were selected due to their widespread use in clustering-based, recommendation-oriented, and bio-inspired learning analytics, and the observed gains highlight the limitations of traditional approaches in handling feature redundancy and complex learning dynamics. Despite these promising outcomes, the present study has several important limitations that should be acknowledged. First, the experimental validation was conducted using only a single benchmark dataset, which may limit the generalizability of the findings across different educational institutions, learner populations, and e-learning platforms. Student behaviour and academic engagement patterns can vary significantly across contexts, and therefore the robustness of the proposed framework should be further examined using multiple and more diverse datasets. Second, although the proposed hybrid framework achieves strong predictive performance, its architectural complexity is relatively high due to the integration of feature selection, quantum-inspired prediction, and metaheuristic optimization components. This increased complexity may lead to higher computational cost, longer training time, and reduced practicality in large-scale or real-time educational analytics systems.

Third, the interpretability of the proposed model remains limited when compared with simpler and more transparent machine learning approaches. While the framework demonstrates strong predictive capability, it does not yet provide sufficiently explainable insights into which behavioural or academic factors most strongly influence the final prediction outcome. This may reduce its immediate usability for educators who require interpretable and actionable decision support. Finally, the current study primarily focuses on predictive performance and does not evaluate deployment-related factors such as real-time adaptability, integration with learning management systems, or user-centered educational impact. Therefore, although the proposed model shows

strong potential, further validation is necessary before large-scale practical adoption can be recommended.

Future work will therefore focus on validating the proposed framework across multiple and diverse e-learning datasets, improving computational efficiency for large-scale implementation, and integrating the model into real-world learning management systems. In addition, future research will investigate the incorporation of explainability mechanisms and interpretable learning analytics to improve model transparency and support informed decision-making by educators. Further studies may also explore lightweight model variants and real-time deployment strategies for practical educational applications.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

S.P.P.: Conceptualization, methodology, software implementation, experimentation, data curation, formal analysis, and original draft preparation. N.S.K.M.K.T.: Methodology refinement, validation, data analysis, interpretation of results, and manuscript review. S.K.A.: Investigation, software support, experimental validation, visualization, and manuscript editing. D.A.D.: Formal analysis, literature review, validation of results, and critical review of the manuscript. T.B.K.: Supervision, methodology review, interpretation of findings, and critical revision of the manuscript. S.G.: Conceptualization, supervision, project administration, funding/resource support, manuscript review, and final approval of the submitted version. All authors had approved the final version for publication.

REFERENCES

- [1] T. Favale, F. Soro, M. Trevisan *et al.*, "Campus traffic and e-learning during COVID-19 pandemic," *Computer Networks*, vol. 176, 107290, 2020.
- [2] A. Muhammad, A. Shaikh, Q. N. Naveed, N. *et al.*, "Factors affecting academic integrity in e-learning of Saudi Arabian Universities. An investigation using Delphi and AHP," *IEEE Access*, vol. 8, pp. 16259–16268, 2020.
- [3] A. Revadekar, S. Oak, A. Gadekar *et al.*, "Gauging attention of students in an e-learning environment," in *Proc. 2020 IEEE 4th Conference on Information & Communication Technology (CICT)*, 2020, pp. 1–6.
- [4] H. Rodrigues, F. Almeida, V. Figueiredo *et al.*, "Tracking e-learning through published papers: A systematic review," *Computers & Education*, vol. 136, pp. 87–98, 2019.
- [5] R. Cobos, F. Jurado, and A. Blázquez-Herranz, "A content analysis system that supports sentiment analysis for subjectivity and polarity detection in online courses," *IEEE Revista Iberoamericana de Tecnologías Del Aprendizaje*, vol. 14, no. 4, pp. 177–187, 2019.
- [6] A. A. Ali, G. A. Dafoulas, and J. C. Augusto, "Collaborative educational environments incorporating mixed reality technologies: A systematic mapping study," *IEEE Transactions on Learning Technologies*, vol. 12, no. 3, pp. 321–332, 2019.
- [7] F. Martínez-Plumed, L. Contreras-Ochando, C. Ferri *et al.*, "CRISP-DM twenty years later: From data mining processes to data science trajectories," *IEEE Transactions on Knowledge and Data Engineering*, vol. 33, no. 8, pp. 3048–3061, 2019.
- [8] H. Zhang, Z. Xiao, J. Wang *et al.*, "A novel IoT-perceptive Human Activity Recognition (HAR) approach using multihead convolutional attention," *IEEE Internet of Things Journal*, vol. 7, no. 2, pp. 1072–1080, 2019.
- [9] F. Salo, M. Injadat, A. B. Nassif *et al.*, "Data mining techniques in intrusion detection systems: A systematic literature review," *IEEE Access*, vol. 6, pp. 56046–56058, 2018.
- [10] A. Hanif, F. Q. Jamal, and M. Imran, "Extending the technology acceptance model for use of e-learning systems by digital learners," *IEEE Access*, vol. 6, pp. 73395–73404, 2018.
- [11] S. A. Salloum, A. Q. M. Alhamad, M. Al-Emran *et al.*, "Exploring students' acceptance of e-learning through the development of a comprehensive technology acceptance model," *IEEE Access*, vol. 7, pp. 128445–128462, 2019.
- [12] W. M. Al-Rahmi, N. Yahaya, A. A. Aldraiweesh *et al.*, "Integrating technology acceptance model with innovation diffusion theory: An empirical investigation on students' intention to use E-learning systems," *IEEE Access*, vol. 7, pp. 26797–26809, 2019.
- [13] O. El Aissaoui, Y. E. A. El Madani, L. Oughdir *et al.*, "A fuzzy classification approach for learning style prediction based on web mining technique in e-learning environments," *Education and Information Technologies*, vol. 24, no. 3, pp. 1943–1959, 2019.
- [14] I. Azzi, A. Jeghal, A. Radouane *et al.*, "A robust classification to predict learning styles in adaptive e-learning systems," *Education and Information Technologies*, vol. 25, no. 1, pp. 437–448, 2020.
- [15] R. Rajkumar and V. Ganapathy, "Bio-inspired learning style chatbot inventory using brain computing interface to increase the efficiency of e-learning," *IEEE Access*, vol. 8, pp. 67377–67395, 2020.
- [16] S. Kausar, X. Huahu, I. Hussain *et al.*, "Integration of data mining clustering approach in the personalized e-learning system," *IEEE Access*, vol. 6, pp. 72724–72734, 2018.
- [17] S. Wan and Z. Niu, "A hybrid e-learning recommendation approach based on learners' influence propagation," *IEEE Transactions on Knowledge and Data Engineering*, vol. 32, no. 5, pp. 827–840, 2019.
- [18] K. Mangaroska, B. Vesin, V. Kostakos *et al.*, "Architecting analytics across multiple e-learning systems to enhance learning design," *IEEE Transactions on Learning Technologies*, vol. 14, no. 2, pp. 173–188, 2021.
- [19] G. Wang *et al.*, "Incorporating mutation scheme into krill herd algorithm for global numerical optimization," *Neural Computing and Applications*, vol. 24, no. 3, pp. 853–871, 2014.
- [20] A. H. Gandomi and A. H. Alavi, "Krill herd: A new bio-inspired optimization algorithm," *Communications in Nonlinear Science and Numerical Simulation*, vol. 17, no. 12, pp. 4831–4845, 2012.
- [21] R. R. Rani and D. Ramyachitra, "Krill herd optimization algorithm for cancer feature selection and random forest technique for classification," in *Proc. 2017 8th IEEE International Conference on Software Engineering and Service Science (ICSESS)*, 2017, pp. 109–113. doi: 10.1109/ICSESS.2017.8342875
- [22] C. Romero and S. Ventura, "Educational data mining: A review of the state of the art," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 40, no. 6, pp. 601–618, 2010.
- [23] M. A. Akbari, M. Zare *et al.*, "The cheetah optimizer: A nature-inspired metaheuristic algorithm for large-scale optimization problems," *Scientific Reports*, vol. 12, no. 1, 10953, 2022. doi: 10.1038/s41598-022-14338-z
- [24] S. C. Chu, P. W. Tsai, and J. S. Pan, "Cat swarm optimization," in *Proc. Pacific Rim International Conference on Artificial Intelligence*, 2006, pp. 854–858. doi: 10.1007/978-3-540-36668-3_94
- [25] A. Tzanetos, and G. Dounias, "A comprehensive survey on the applications of swarm intelligence and bio-inspired evolutionary strategies," *Machine Learning Paradigms: Advances in Deep Learning-based Technological Applications*, pp. 337–378, 2020.
- [26] V. Dunjko and H. J. Briegel, "Machine learning & artificial intelligence in the quantum domain," *Reports on Progress in Physics*, vol. 81, no. 7, 2018. doi: 10.1088/1361-6633/aab406
- [27] M. R. Khurshid, S. Manzoor, T. Sadiq *et al.*, "Unveiling diabetes onset: Optimized XGBoost with Bayesian optimization for enhanced prediction," *PloS One*, vol. 20, no. 1, e0310218, 2025. doi: 10.1371/journal.pone.0310218
- [28] I. Aljarah. (2016). Students' academic performance dataset. *Kaggle*. [Online]. Available: <https://www.kaggle.com/datasets/aljarah/xAPI-Edu-Data>
- [29] N. S. K. M. K. Tirumanadham, T. Sekhar, S. B. Chandolu *et al.*, "ElimRidge-HFM: A mathematical integration of feature selection,

- boosting, and hyperparameter search for enhanced predictive performance in e-learning,” *Contemp. Math.*, vol. 6, no. 4, pp. 4459–4482, 2025. doi: 10.37256/cm.6420256289.
- [30] A. Luthra, P. Kaur, A. Kaur *et al.*, “Adaptalytics of Students’ to Online Learning Using Machine Learning Classifiers and Sequential Neural Network,” in *Recent Advances in Computational Intelligence Applications for Biometrics and Biomedical Devices*, Massachusetts, USA: Academic Press, 2026, pp. 321–341. doi: 10.1016/B978-0-443-33012-4.00008-2
- [31] F. Li, “Prediction of technical training efficiency using deep learning and real-time feedback systems in engineering education,” *Australian Journal of Electrical and Electronics Engineering*, pp. 1–16, 2025. doi: 10.1080/1448837X.2025.2607836
- [32] E. Elabd, H. M. Hamouda, M. A. M. Ali *et al.*, “Air quality index AQI classification based on hybrid particle swarm and grey wolf optimization with ensemble machine learning model,” *Scientific Reports*, vol. 16, 4152, 2026. doi: 10.1038/s41598-025-34278-8
- [33] M. Lemos, P. J. S. Cardoso, and J. M. F. Rodrigues, “From cues to engagement: A comprehensive survey and holistic architecture for computer vision-based audience analysis in live events,” *Multimodal Technologies and Interaction*, vol. 10, no. 1, p. 8, 2026. doi: 10.3390/mti10010008
- [34] J. A. Hanley and B. J. McNeil, “The meaning and use of the area under a Receiver Operating Characteristic (ROC) curve,” *Radiology*, vol. 143, no. 1, pp. 29–36, 1982. doi: 10.1148/radiology.143.1.7063747
- [35] T. Fawcett, “An introduction to ROC analysis,” *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874, 2006.

Copyright © 2026 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)).