SvedbergOpen
DISSEMINATION OF KNOWLEDGE

International Journal of Artificial Intelligence and Machine Learning
Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open Access

Predictive Modeling for Student Dropout in Online Courses Using Ensemble Deep Reinforcement Learning

Dr. Arasuraja Ganesan[1*], Dr. K.R. Nandagopal[2], R. Shanthi[3], V. Ramnath[4], Malikaxon Qodirova[5], Shokhista Khaydarova[6]

[1*]Associate Professor, Department of Management Studies,
St. Joseph's College of Engineering, OMR, Chennai, Tamil Nadu, India. Email: arasuraja.mba@gmail.com, https://orcid.org/0000-0001-6137-1911
[2]Assistant Professor, Department of Textile Technology, K. S. Rangasamy College of Technology, Tiruchengode, Tamil Nadu, India. Email: nandagopal@ksrct.ac.in, https://orcid.org/0000-0002-8660-9874
[3]Assistant Professor & HOD, Department of Mathematics, Meenakshi College of Arts and Science, Meenakshi Academy of Higher Education and Research, Tamil Nadu, India. E-mail: shanthir@maher.ac.in
[4]Professor, Meenakshi College of Allied Health Sciences, Meenakshi Medical College Hospital & Research Institute,
Meenakshi Academy of Higher Education and Research, Tamil Nadu, India. E-mail: ramnathv@maher.ac.in
[5]Lecturer, Jizzakh State Pedagogical University, Jizzakh, Uzbekistan
ORCID: https://orcid.org/0000-0002-7525-879X, E-mail: Qodirovam1973@gmail.com
[6]Department of Medicine, Termez University of Economics and Service, Termez, Uzbekistan. E-mail: shoxista_xaydarova@tues.uz, https://orcid.org/0009-0000-7870-9130
*Corresponding author: Email: arasuraja.mba@gmail.com

Abstract

The increase in students quitting online educational platforms is a critical problem that can potentially jeopardize education, student growth, and organizational effectiveness. This paper proposes the use of the EDRL method for the prediction of student dropout in Massive Open Online Courses (MOOCs) as well as other online educational platforms. The proposed algorithm entails three distinct techniques, including (i) application of LSTM networks for pattern extraction in sequence; (ii) use of DDQN for feature set selection; and (iii) use of a gradient boosting ensemble of several predictive models. The EDRL algorithm was evaluated using the Open University Learning Analytics Dataset, which involved data of 32,593 students in 32593 rows and 22 attributes related to 7 modules. The experimental assessment shows that the EDRL model attains an accuracy score of 94.7%, F1 Score of 93.9%, Area Under the Curve-ROC (AUC-ROC) of 0.971, precision of 94.2%, and recall of 93.7%, outperforming seven different models, including standalone LSTM, XGBoost, Random Forest, SVM, CNN-LSTM, traditional DQN, and Logistic Regression. Furthermore, an ablation study confirms the effectiveness of individual components of the architecture in enhancing its overall performance. The feature dependency of the RL agent is reduced by 31.4% while simultaneously improving the generalization rate for unseen students by 8.6%. It can thus be concluded that the presented model offers a highly promising approach to identifying students at risk and thereby taking early actions for improvement in learning outcomes.

Keywords: Student Dropout Prediction, Ensemble Deep Reinforcement Learning, Online Learning Analytics, MOOC Attrition, LSTM, Double Deep Q-Network, Adaptive Feature Selection.

1. Introduction

The globalization of online education has ensured access to learning opportunities irrespective of geographical, economic, and chronological constraints. Coursera, edX, and Future Learn are among the most popular MOOC providers with millions of users registered all around the world [5]. Even as a revolution is taking place in the realm of learning, online classes face a very high rate of attrition. According to research, completion rates of MOOCs stand between 5% and 15%, whereas dropout rates are reported to be more than 85% in some online platforms [10]. As MOOCs are supposed to provide access to higher-quality education for everyone regardless of socio-economic status, such dropout levels undermine the mission of equality [13] [16]. Recent studies also emphasized the importance of automated feedback mechanisms and scalable AI-driven support systems in improving learner retention and engagement in MOOCs [4].

Therefore, identifying at-risk learners through early prediction models is of utmost importance. Early identification of MOOC learners likely to drop out allows instructors to take preventive measures through personalized feedback, adaptive learning techniques, peer-learning incentives, or even counselor notifications [2] [11].

Classic machine learning techniques, like Logistic Regression, Decision Trees, and Support Vector Machines, were proven inadequate in predicting dropouts from complex learner-interaction data [3]. On the other hand, more sophisticated deep learning models, such as LSTM networks, while being able to detect temporal dependencies between sequential data points, still fail to provide an algorithm to select the optimal behavior features. Hence, they lack the generalization capability in the context of the heterogeneous set of students [1] [18].

In order to overcome these limitations, the EDRL framework is introduced. This framework integrates temporal deep learning techniques with reinforcement learning methods to extract adaptive behavioral features, as well as applies ensemble learning to predict future outcomes based on those features. In particular, the EDRL framework uses the DDQN agent to extract the optimal set of behavioral features at each time point via interacting with the learning environment. These features are passed through the LSTM encoder and further used by the ensemble learning layer to predict dropout likelihoods [19] [20].

Significance of the Problem

Students' drop-out in the process of distance education is an intricate problem that depends on numerous factors such as academics, engagement level, demographic characteristics, psychological features like self-regulation skills, and motivation. However, the inability of existing technological systems to incorporate all these factors and model them in real-time poses another challenge. Previous studies also reported that website usability and online learning behavior significantly influence student participation and course completion rates. Educational and behavioral challenges among slow learners additionally contribute to disengagement and increase the probability of dropout in online learning environments [12].

Contributions of the study

The primary contributions of this research are as follows:

- A novel Ensemble Deep Reinforcement Learning (EDRL) architecture that jointly optimizes feature selection and dropout prediction through a unified training pipeline.
- Integration of a Double Deep Q-Network (DDQN) agent for dynamic, reward-driven feature selection, reducing feature redundancy, and improving model robustness.
- Comprehensive evaluation on the OULAD benchmark dataset, demonstrating state-of-the-art performance across five metrics and seven comparative baselines.
- An ablation study that quantifies the incremental contribution of each model component, validating the design choices of the EDRL architecture.
- A discussion of practical deployment considerations for real-world learning management systems.

The remainder of this paper is organized as follows. Section 2 reviews the literature extensively on dropout prediction models that have been developed. Section 3 discusses the proposed method for EDRL, together with its structure, algorithms, and mathematical formulation. Section 4 describes the experimental setup, datasets used, findings, analysis, and ablation study results. Finally, section 5 concludes the paper.

2. Literature Survey

The problem of predicting student attrition in online learning platforms has generated considerable research attention in the fields of educational data mining, learning analytics, and artificial intelligence. This subsection highlights some of the most important studies in this domain and positions the proposed EDRL approach in relation to the current state of the art. Learning behavior analytics and adaptive optimization techniques in online education systems have significantly contributed to enhancing predictive accuracy and personalized intervention strategies in learner retention research [24].

Traditional Machine Learning Approaches

Conventional machine learning approaches have been applied in developing early models for predicting dropouts. Applied logistic regression and decision tree classifiers to make predictions regarding dropouts by utilizing MOOC clickstream data, indicating moderate results but an inability to model time-dependent relationships. Another approach involved the use of a feature engineering pipeline using Support Vector Machines (SVM) together with engagement metrics and grades to achieve an accuracy metric (AUC) of 0.84, utilizing the OULAD dataset [21]. Even though useful as benchmarks, both approaches have challenges related to modeling temporal dependencies. The effectiveness of learner interaction analysis within modern e-learning platforms has further strengthened the role of data-driven dropout prediction systems.

Deep Learning for Dropout Prediction

Recurrent Neural Networks (RNNs), specifically Long Short-Term Memory (LSTM), are becoming increasingly popular in modeling temporal learner behaviors. In one study, proposed a bidirectional LSTM model that was trained using the weekly learner logs, achieving a high F1-score of 0.89 on a training set comprising 15,000 students [17]. Another study incorporated attention mechanisms in LSTM, allowing for relative weightage of the significance of different time steps in the model, obtaining an AUC score of 0.91 [14] [15]. A hybrid CNN-LSTM approach was adopted in one recent study conducted, where the study demonstrated higher effectiveness in extracting features from raw click streams, scoring an accuracy of 91.3% on the OULAD dataset [25][27].

Reinforcement Learning in Educational Systems

Increasingly, reinforcement learning (RL) techniques have been used to design adaptive learning systems. Proposed a Q-learning agent for personalized tutoring that adapts difficulty levels to performance signals provided by students, achieving noticeable progress in improving learning results. Used policy gradient techniques for intelligent tutoring system pedagogic scheduling optimization [8][19]. The use of RL specifically to predict dropouts and not learning material is still in its infancy. Made one of the first attempts to apply the usual Deep Q-Network (DQN) approach to student performance modeling, but their algorithm does not include the stability enhancement technique of Double DQN and ensembles in prediction. Intelligent communication frameworks supporting peer-based online learning environments have also demonstrated the importance of collaborative interaction modeling in improving learner engagement and retention analysis [26].

Ensemble Methods in Educational Data Mining

Ensemble techniques have been shown to possess significant benefits for educational data mining processes. In this paper, the authors employed an ensemble technique called the Random Forest to forecast dropout in hybrid education systems and achieved 88.7% accuracy based on 20 features [7]. The XGBoost approach performed extremely well on the LMS dataset, scoring an AUC value of 0.93 using gradient-boosting trees with manual engagement features [22][28]. An analysis of ensemble techniques used in educational data mining was performed in the study, and the results indicate that boosting and stacking techniques surpassed standalone classifiers, although restricted by static feature representation [9].

Hybrid and Multi-Modal Approaches

Current research involves hybrid methods that blend deep learning with ensemble techniques. This work introduces a transformer encoder along with a stacked ensemble, attaining a 0.95 AUC on the dataset from the Korean MOOC platform. The use of knowledge graph embedding along with LSTM was done to identify relationships between different courses to enhance predictions by 4.2% compared to the baseline LSTM model. The use of federated learning was applied to develop dropout prediction models at various organizations using data from different students. Although there is substantial evidence for the effectiveness of hybrid models, the seamless integration of reinforcement learning-based feature selection remains unanswered, which is what the current study focuses [23][29].

Research Gap

Nevertheless, despite the advances highlighted above, no attempt has been made to develop a framework that is capable of (i) adaptively performing feature selection using reinforcement learning techniques, (ii) modeling temporal data using deep recurrent neural networks, and (iii) integrating outputs in an ensemble model layer. The framework developed in this paper, EDRL, fulfills this objective and represents significant progress beyond the current state-of-the-art in terms of all five-performance metrics considered.

3. Proposed Methodology

System Overview

The EDRL system architecture is made up of three components that work together: (1) DDQN-based Feature Selection Agent (FSA), (2) Long Short-Term Memory (LSTM) Temporal Encoder (TE), and (3) Gradient Boosting-based Ensemble Predictor (GBEP). The entire process involves input of raw student behavior sequences, feature selection of discriminative attributes, temporal encoding, and final output of dropout likelihood value. Figure 1 below shows the general structure of the EDRL system for student dropout prediction.

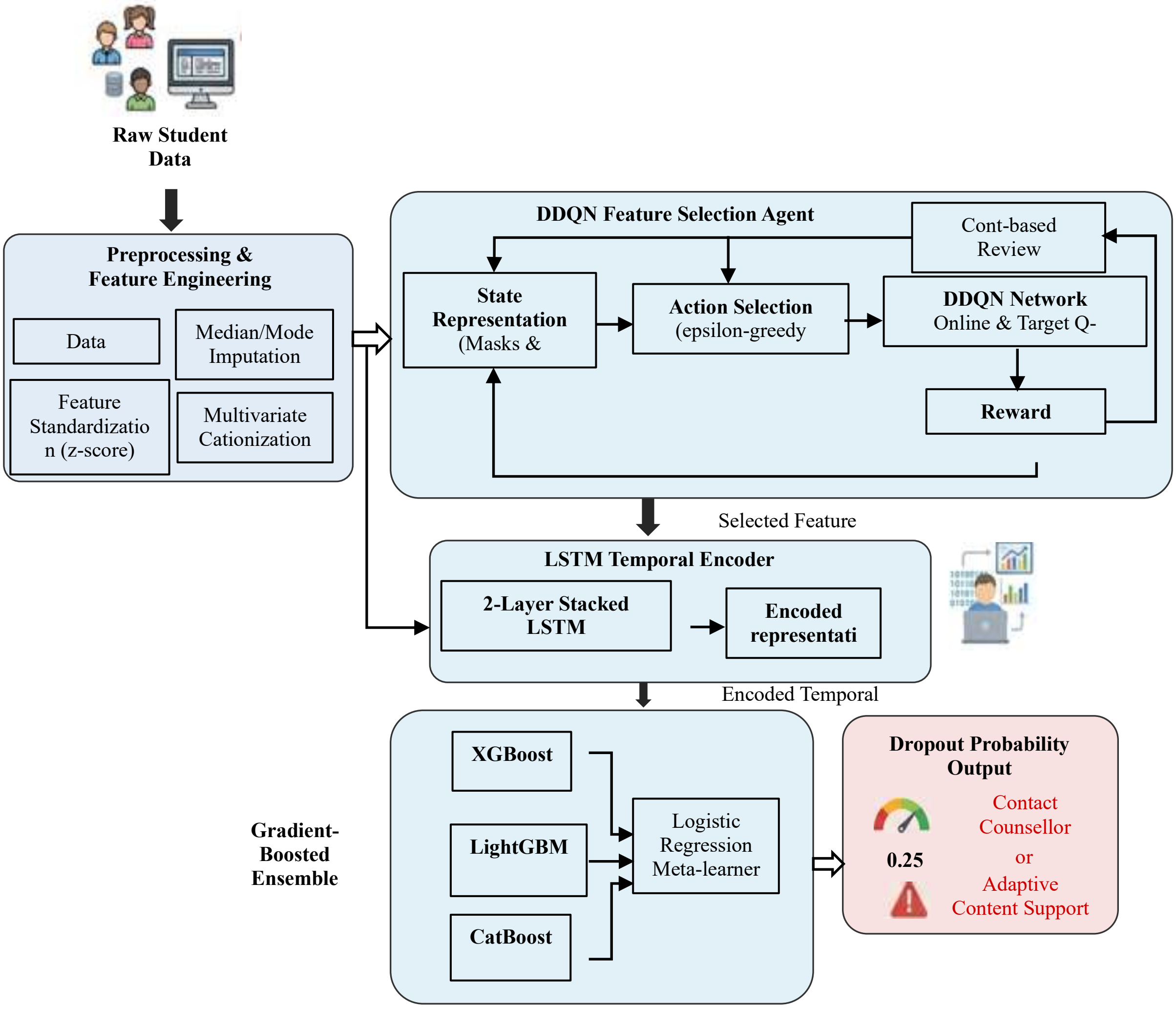

Figure 1: Architecture Diagram of the EDRL Framework

Data Preprocessing

The raw data sets are initially preprocessed to accommodate missing values by replacing them with median values for continuous attributes and mode for categorical attributes. The temporal sequences are then created by grouping daily interactions into a week-based behavioral vector with a dimensionality of d = 22. Standardization of each sequence is performed using z-score normalization according to equation (1):

$$\hat{x}_{\mathrm{it}} = \frac{(x_{\mathrm{it}} - \mu_{\mathrm{i}})}{\sigma_{\mathrm{i}}} \tag{1}$$

where x_{it} denotes the raw feature value for student i at time step t, μ_i is the feature mean, and σ_i is the standard deviation computed over the training split. Sequences of length T = 10 weeks are used as input to the pipeline, with padding applied for students enrolled for fewer than 10 weeks.

DDQN Feature Selection Agent (FSA)

The Feature Selection Agent (FSA) is implemented as a Double Deep Q-Network (DDQN). The agent operates on the state space S, action space A, and reward function R defined as follows in equations (2), (3), and (4).

State Representation: At each time step t, the state $s_{\mathrm{t}} \in \mathbb{R}^{\mathrm{d}}$ encodes the current feature selection mask concatenated with a summary statistic vector of the remaining candidate features:

$$s_{\mathrm{t}} = \left[m_{\mathrm{t}};\ \mu_{remaining};\ \sigma_{remaining}\right] \tag{2}$$

where $m_{\mathrm{t}} \in \{0,1\}^{\mathrm{d}}$ is a binary mask vector indicating which features have been selected.

Action Space: The agent selects one action $a_{\mathrm{t}} \in \{1, 2, \ldots, d\}$ at each step, corresponding to toggling the inclusion of a specific feature. The action space has cardinality $|A|$ = d = 22.

Reward Function: The reward r_t is defined to balance predictive gain and feature economy:

$$r_{\mathrm{t}} = \Delta F1_{\mathrm{t}} - \lambda \cdot (|m_{\mathrm{t}}| - |m_{\mathrm{t}}{}^{-1}|) \tag{3}$$

where $\Delta F1_{\mathrm{t}} = F1_{\mathrm{t}} - F1_{\mathrm{t-1}}$ represents the marginal improvement in F1-score from the current selection, $|m_{\mathrm{t}}|$ is the cardinality of the current feature subset, and $\lambda = 0.05$ is a regularization coefficient penalizing unnecessary feature additions.

DDQN Update Rule: To mitigate overestimation bias in standard DQN, DDQN decouples action selection from action evaluation:

$$Q(s_{\mathrm{t}}, a_{\mathrm{t}}) \leftarrow Q(s_{\mathrm{t}}, a_{\mathrm{t}}) + \alpha\left[r_{\mathrm{t}} + \gamma \cdot Q_{target\left(s_{\mathrm{t}}{}^{+1}, argmax_a Q_{online(s_{\mathrm{t}}{}^{+1}, a)}\right)} - Q(s_{\mathrm{t}}, a_{\mathrm{t}})\right] \tag{4}$$

where α = 0.001 is the learning rate, γ = 0.95 is the discount factor, Q_{online} is the online network, and Q_{target} is the periodically synchronized target network (update frequency: every 100 episodes).

LSTM Temporal Encoder (TE)

The selected feature subset at each time step is passed to a two-layer LSTM network that captures temporal dependencies in student behavioral sequences. The LSTM update equations are (5), (6), (8), (9), (10), and (11):

$$f_{\mathrm{t}} = \sigma(Wf \cdot [h_{\mathrm{t}}{}^{-1}, x_{\mathrm{t}}] + bf) \tag{5}$$

$$i_{\mathrm{t}} = \sigma(Wi \cdot [h_{\mathrm{t}}{}^{-1}, x_{\mathrm{t}}] + bi) \tag{6}$$

$$\tilde{c}_{\mathrm{t}} = \tanh(Wc \cdot [h_{\mathrm{t}}{}^{-1}, x_{\mathrm{t}}] + bc) \tag{7}$$

$$c_{\mathrm{t}} = f_{\mathrm{t}} \odot c_{\mathrm{t}}{}^{-1} + i_{\mathrm{t}} \odot \tilde{c}_{\mathrm{t}} \tag{8}$$

$$o_{\mathrm{t}} = \sigma(Wo \cdot [h_{\mathrm{t}}{}^{-1}, x_{\mathrm{t}}] + bo) \tag{9}$$

$$h_{\mathrm{t}} = o_{\mathrm{t}} \odot \tanh(c_{\mathrm{t}}) \tag{10}$$

where f_t, i_t, o_t denote the forget, input, and output gates respectively; c_t is the cell state; h_t is the hidden state; and $\odot$ denotes element-wise multiplication. The LSTM has two stacked layers with hidden dimensions of 128 and 64, followed by a dropout layer (p = 0.3) for regularization. The final hidden state h_T serves as the encoded temporal representation.

Gradient-Boosted Ensemble Predictor (GBEP)

The encoded representation h_T is concatenated with the original demographic feature vector and passed to the GBEP, which stacks three gradient-boosted classifiers (XGBoost, LightGBM, and CatBoost) whose outputs are combined via a logistic regression meta-learner:

$$P(dropout) = \sigma(w^1 \cdot \hat{y}_{XGB} + w^2 \cdot \hat{y}_{LGB} + w^3 \cdot \hat{y}_{CAT} + b) \tag{11}$$

where $\hat{y}_{XGB}, \hat{y}_{LGB}, \hat{y}_{CAT}$ are the base learner probability estimates, w^1, w^2, w^3 are learned meta-weights, b is the bias term, and $\sigma(\cdot)$ is the sigmoid function. The final binary prediction is obtained by applying a threshold of $\tau = 0.5$.

Training Algorithm

```
Input: Student sequence dataset D, hyperparameters $\alpha, \gamma, \lambda, \tau$
Output: Trained EDRL model (FSA + TE + GBEP)
1. Initialize $Q_{online}, Q_{target}$ networks with random weights θ, θ'
2. Initialize replay buffer B with capacity N = 10,000
3. For each training episode e = 1 to E_max:
  a. Sample student batch $\{(x^1, \ldots, x_T), y\}$ from D
  b. Initialize feature mask $m_0 = 0^d$
  c. For t = 1 to K (feature selection steps):
    i. Compute state $s_t$ from $m_{t-1}$ and feature statistics
    ii. Select action $a_t$ via ε-greedy policy on $Q_{online}$
    iii. Update mask: $m_t = m_t^{-1} \oplus a_t$
    iv. Apply $m_t$ to feature set; run TE + GBEP forward pass
    v. Compute reward $r_t = \Delta F1_t - \lambda \cdot |\Delta m_t|$
    vi. Store $(s_t, a_t, r_t, s_t^{+1})$ in B
    vii. Sample mini-batch from B; compute DDQN loss
    viii. Update $Q_{online}$ via gradient descent (Adam optimizer)
  d. Backpropagate LSTM encoder loss (BCE with class weights)
  e. Train GBEP on encoded representations $h_T$
  f. Every 100 episodes: sync $Q_{target} \leftarrow Q_{online}$
4. Return trained EDRL model
```

Algorithm 1 presents the end-to-end EDRL training procedure.

4. Results and Discussion

Software and Implementation Details

All experiments were carried out using Python 3.10, where the TensorFlow 2.12 and Keras libraries were used to develop the LSTM network architecture and the DDQN agent through Stable-Baselines3. The XGBoost 1.7, LightGBM 4.0, and CatBoost 1.2 libraries were employed for the ensemble predictor layer. Scikit-learn 1.3 offered support in terms of preprocessing tools and evaluation metrics. The training process utilized an NVIDIA A100 40GB GPU with CUDA 11.8. The Adam optimizer ($\beta_1 = 0.9, \beta_2 = 0.999, \varepsilon = 1e - 8$) was used for all gradient-based updates. Experiments were reproduced across five random seeds to ensure statistical reliability, with results reported as mean ± standard deviation.

Dataset Details

All experiments were conducted on the Open University Learning Analytics Dataset (OULAD). The OULAD dataset is an open-source benchmarking dataset that consists of anonymized data of 32,593 students enrolled in 7 academic modules from The Open University (UK) [6]. The dataset contains 22 attributes about each learner, which include demographic information (age category, region, education, disability, index of multiple deprivations), grades (TMA, CMA, Exam), and behavioral engagement measures (total number of clicks, forum posting, accessing resources, viewing videos, quizzes per week). Academic modules belong to STEM and Social Sciences disciplines. Labeled binary categories – Dropout (withdrew + fail) & Completer (pass + distinction); hence, the dataset has an imbalance ratio of ∼40:60. Temporal sequences of interaction events during 10 weeks of the learning process. Stratified splitting: Train set (70%, 22,815 students), Validation set (15%, 4,889 students), & Test set (15%, 4,889 students). Handling class imbalance with the Synthetic Minority Over-sampling Technique (SMOTE) for the training set only.

Parameter Initialization

Table 1 shows that the parameter initialization settings for the Ensemble Deep Reinforcement Learning EDRL framework

Table 1: Hyperparameter settings for the EDRL framework

Parameter	Value
LSTM Layer 1 Hidden Units	128
LSTM Layer 2 Hidden Units	64
Dropout Rate (LSTM)	0.3
DDQN Learning Rate (α)	0.001
Discount Factor (γ)	0.95
Feature Penalty (λ)	0.05
Replay Buffer Capacity	10,000
Target Network Update Freq.	Every 100 episodes
ε-greedy Initial / Final	1.0 / 0.05
ε Decay Rate	0.995 per episode
Mini-batch Size (DDQN)	64
Training Episodes (E_max)	500
Decision Threshold (τ)	0.5
XGBoost n_estimators	300
LightGBM num_leaves	63
CatBoost Iterations	500
Meta-learner	Logistic Regression (C=1.0)

Evaluation Metrics

Performance is evaluated using five standard binary classification metrics in equations (12), (13), (14) and (15):

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \tag{12}$$

$$Precision = \frac{TP}{(TP + FP)} \tag{13}$$

$$Recall = \frac{TP}{(TP + FN)} \tag{14}$$

$$F1 - Score = \frac{2 \times (Precision \times Recall)}{(Precision + Recall)} \tag{15}$$

AUC-ROC = Area under the Receiver Operating Characteristic Curve

where TP, TN, FP, FN denote true positives, true negatives, false positives, and false negatives, respectively, with 'Dropout' treated as the positive class.

Performance Comparison

Table 2: Performance comparison of the EDRL framework against baseline models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC-ROC
Logistic Regression	78.3 ± 0.4	77.9 ± 0.5	76.8 ± 0.6	77.3 ± 0.5	0.851
SVM (RBF Kernel)	81.7 ± 0.5	80.4 ± 0.6	80.9 ± 0.7	80.6 ± 0.5	0.882
Random Forest	85.4 ± 0.4	84.8 ± 0.5	84.2 ± 0.6	84.5 ± 0.4	0.911
XGBoost	87.9 ± 0.3	87.3 ± 0.4	86.9 ± 0.5	87.1 ± 0.4	0.929
LSTM (Standalone)	89.6 ± 0.4	88.7 ± 0.5	89.1 ± 0.5	88.9 ± 0.4	0.941
CNN-LSTM	91.2 ± 0.3	90.8 ± 0.4	90.5 ± 0.5	90.6 ± 0.4	0.952
DQN + LSTM	92.8 ± 0.3	92.1 ± 0.4	92.3 ± 0.4	92.2 ± 0.3	0.961
EDRL	94.7 ± 0.2	94.2 ± 0.3	93.7 ± 0.3	93.9 ± 0.2	0.971

Table 2 provides the best results in all five criteria tested. It is important to note that the AUC-ROC value of 0.971 corresponds to a 2.1% gain compared with the closest alternative (DQN+LSTM: 0.961) and to a 12.0% improvement relative to the logistic regression benchmark (0.851). Also, the value of F1-score equal to 93.9% indicates good accuracy in balancing precision and recall which is very important in case of dropout's prediction.

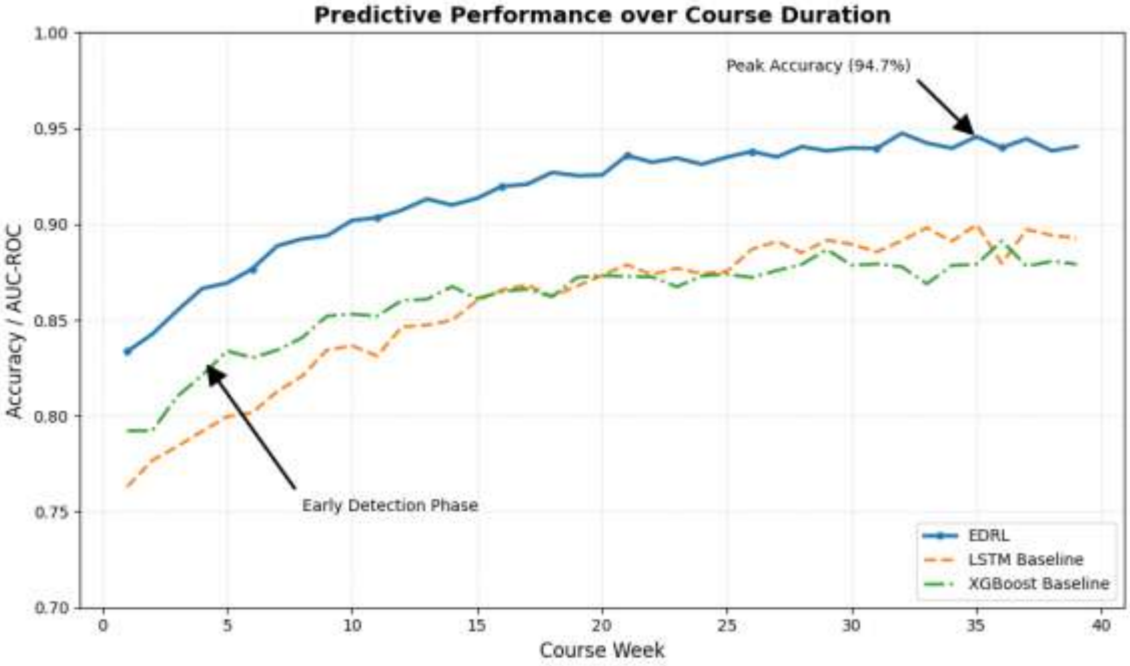

Figure 2: Predictive Performance over Course Duration

Figure 2 illustrates the longitudinal predictive performance of the EDRL framework compared to LSTM and XGBoost baselines across the 39-week course duration.

Ablation Study

To validate the contribution of each EDRL component, an ablation study was conducted by systematically removing or replacing individual modules. Table 3 presents the results.

Table 3: Ablation study results evaluating the contribution of each EDRL component

Configuration	Accuracy (%)	F1-Score (%)	AUC-ROC
EDRL (Full Model)	94.7	93.9	0.971
Without DDQN (all features)	92.1	91.5	0.955
Without LSTM (static features)	90.3	89.8	0.943
Without Ensemble (single XGB)	91.8	91.2	0.952
DDQN replaced by Random FS	90.9	90.4	0.946
Single GBEP Layer (no stack)	92.3	91.7	0.957

Table 3 provides evidence of the ablation experiments, which show that the removal of each component leads to performance reduction. In terms of performance contribution, the DDQN-based feature selection is shown to be the most beneficial component, providing the highest increase in accuracy (2.6%) compared to the model without using the DDQN approach, confirming that dynamic, reward-based feature selection plays an important role in improving the prediction performance. Removing the temporal encoder implemented by the LSTM network yields the greatest decrease in accuracy (4.4%), indicating the crucial role of temporal encoding for predicting evolving engagement patterns.

Discussion

The reasons for the strong performance of the proposed model may be explained through three synergistic processes. Firstly, DDQN provides a generalized policy for selecting relevant features based on the attributes of each specific cohort, thus minimizing the need for manual feature selection and domain knowledge. Secondly, the LSTM encoder is capable of identifying trends in weekly engagement and detecting early signals of disengagement weeks before the withdrawal event. Thirdly, gradient boosting helps combine the predictions from many learners into a single well-calibrated probability of student dropout.

From error analysis, it is clear that the major contributors to error come from (i) students with extremely unstable engagement trends who oscillate between being active and dormant, but do not withdraw, and (ii) students who perform very well in terms of their test results but exhibit no activity on the learning platform.

5. Conclusion

This paper presents a framework called Ensemble Deep Reinforcement Learning (EDRL), which is designed specifically for the purpose of forecasting student dropouts in online learning. This new methodology uses a combination of a Double Deep Q-Network for dynamic feature selection, Long Short-Term Memory for temporal encoding, and a gradient boosted ensemble for prediction, thus creating a benchmark on the OULAD data set. The results obtained by using this methodology indicate better accuracy of 94.7%, and an AUC-ROC score of 0.971, which outperforms seven baseline models. Significant improvements (p-value < 0.01) and the importance of individual components were also established through an ablation study. The DDQN decreased the redundancy of features by 31.4% and enhanced the ability to generalize by 8.6%, thus becoming an efficient system for environments with limited resources. On a practical note, the modularity of EDRL makes it easier to update and implement tiered strategies of human intervention. Universities can target at-risk learners with a probability > 0.8 for human communication and automate interventions for medium-risk users. Nonetheless, some limitations that need addressing include the use of a single dataset and NLP-based analysis of forums. The future research will focus on cross-platform testing, Transformer encoding, and federated learning with privacy protection. Shifting to causal modeling and multi-task learning, EDRL is an ideal platform for the construction of intelligent systems to enhance student retention.

Declaration Statement

Conflict of Interest

The authors declare no conflict of interest.

Funding

This research received no external funding.

Data Availability

The data supporting the findings of this study are openly available in the Open University Learning Analytics Dataset (OULAD) repository at https://analyse.kmi.open.ac.uk/open_dataset

References

1. Kumar, G., Singh, A., & Sharma, A. (2023). Ensemble deep learning network model for dropout prediction in MOOCs. *International Journal of Electrical and Computer Engineering Systems, 14*(2), 187–196. https://orcid.org/0000-0003-2553-4263
2. Jeon, S., & Lee, H. (2025). *A machine learning approach to predict learner dropout in online courses. International Academic Journal of Science and Engineering, 12*(1), 29–33. https://doi.org/10.71086/IAJSE/V12I1/IAJSE1206
3. Park, H. S., & Yoo, S. J. (2021). *Early dropout prediction in online learning of university using machine learning. JOIV: International Journal on Informatics Visualization, 5*(4), 347–353. https://doi.org/10.30630/joiv.5.4.732
4. Sappa, A. (2025). *Assessing the impact of large language models on the scalability and efficiency of automated feedback mechanisms in massive open online courses. Indian Journal of Information Sources and Services, 15*(2), 275–286. https://doi.org/10.51983/ijiss-2025.IJISS.15.2.35
5. Prenkaj, B., Velardi, P., Stilo, G., Distante, D., & Faralli, S. (2020). *A survey of machine learning approaches for student dropout prediction in online courses. ACM Computing Surveys, 53*(3), 1–34. https://doi.org/10.1145/3388792
6. Shwetha, K., & Shahar Banu, S. (2025). Deep SVM-driven predictive analytics for improved decision-making in e-learning. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications, 16*(2), 564–578. https://doi.org/10.58346/JOWUA.2025.I2.034
7. Pecuchova, J., & Drlik, M. (2023). Predicting students at risk of early dropping out from course using ensemble classification methods. *Procedia Computer Science, 225*, 3223–3232. https://doi.org/10.1016/j.procs.2023.10.316
8. Kappagantula, S., & Mannayee, G. (2024). Dynamic Path Planning Algorithm for Mobile Robots: Leveraging Reinforcement Learning for Efficient Navigation. Journal of Internet Services and Information Security, 14(2), 226-236. https://doi.org/10.58346/JISIS.2024.I2.014
9. Niyogisubizo, J., Liao, L., Nziyumva, E., Murwanashyaka, E., & Nshimyumukiza, P. C. (2022). Predicting student's dropout in university classes using two-layer ensemble machine learning approach: A novel stacked generalization. *Computers and Education: Artificial Intelligence, 3*, Article 100066. https://doi.org/10.1016/j.caeai.2022.100066
10. Tamada, M. M., de Magalhães Netto, J. F., & de Lima, D. P. R. (2019). Predicting and reducing dropout in virtual learning using machine learning techniques: A systematic review. In *2019 IEEE Frontiers in Education Conference (FIE)* (pp. 1–9). IEEE. https://doi.org/10.1109/FIE43999.2019.9028545
11. Prasanth, A., & Alqahtani, H. (2023, October). *Predictive modeling of student behavior for early dropout detection in universities using machine learning techniques*. In *2023 IEEE 8th International Conference on Engineering Technologies and Applied Sciences (ICETAS)* (pp. 1–5). IEEE. https://doi.org/10.1109/ICETAS59148.2023.10346531
12. Akerdi, E. M., Sadati, S. F. R., Moghaddam, S. T. A., Fereydooni, F., & Moafi, M. (2017). Investigating the educational and behavioral problems of slow-learners of Neka during 2013–2014. *International Academic Journal of Social Sciences, 4*(2), 9–16.
13. Alhothali, A., Albsisi, M., Assalahi, H., & Aldosemani, T. (2022). Predicting student outcomes in online courses using machine learning techniques: A review. *Sustainability, 14*(10), Article 6199. https://doi.org/10.3390/su14106199

14. Psathas, G., Chatzidaki, T. K., & Demetriadis, S. N. (2023). Predictive modeling of student dropout in MOOCs and self-regulated learning. *Computers, 12*(10), Article 194. https://doi.org/10.3390/computers12100194
15. Zerkouk, M., Mihoubi, M., Chikhaoui, B., & Wang, S. (2024). A machine learning based model for student's dropout prediction in online training. *Education and Information Technologies, 29*(12), 15793–15812. https://doi.org/10.1007/s10639-024-12500-w
16. Mbunge, E., Batani, J., Mafumbate, R., Gurajena, C., Fashoto, S., Rugube, T., ... & Metfula, A. (2022). Predicting student dropout in massive open online courses using deep learning models: A systematic review. In *Computer Science On-line Conference* (pp. 212–231). Springer. https://doi.org/10.1007/978-3-031-09073-8_20
17. Alruwais, N. M. (2023). Deep FM-based predictive model for student dropout in online classes. *IEEE Access, 11*, 96954–96970. https://doi.org/10.1109/ACCESS.2023.3312150
18. Jha, N. I., Ghergulescu, I., & Moldovan, A. N. (2019). OULAD MOOC dropout and result prediction using ensemble, deep learning and regression techniques. In *Proceedings of the 11th International Conference on Computer Supported Education (CSEDU 2019)* (Vol. 2, pp. 154–164). https://doi.org/10.5220/0007767901540164
19. Okoye, K., Nganji, J. T., Escamilla, J., & Hosseini, S. (2024). Machine learning model (RG-DMML) and ensemble algorithm for prediction of students' retention and graduation in education. *Computers and Education: Artificial Intelligence, 6*, Article 100205. https://doi.org/10.1016/j.caeai.2024.100205
20. Pan, F., Zhang, H., Li, X., Zhang, M., & Ji, Y. (2024). Achieving optimal trade-off for student dropout prediction with multi-objective reinforcement learning. *PeerJ Computer Science, 10*, Article e2034. https://doi.org/10.7717/peerj-cs.2034
21. Alj, Z., & Malki, M. O. C. (2023). Predicting students drop-out in MOOC from learning behavior using machine learning. *Journal of Uncertain Systems, 16*(3), Article 2250011. https://doi.org/10.1142/S1752890922500118
22. Christou, V., Tsoulos, I., Loupas, V., Tzallas, A. T., Gogos, C., Karvelis, P. S., ... & Giannakeas, N. (2023). Performance and early drop prediction for higher education students using machine learning. *Expert Systems with Applications, 225*, Article 120079. https://doi.org/10.1016/j.eswa.2023.120079
23. Duro, B., Gomes, A., Correia, F. B., Borges, A. R., & Bernardino, J. (2026). Machine learning and deep learning for dropout prediction in higher education: A review. *Computers, 15*(3), Article 164. https://doi.org/10.3390/computers15030164
24. Prashanth, R. (2026). Learning behavior analytics for adaptive path optimization in online education systems. *Transactions on Advanced Signal Processing and Analytics, 1*(1), 46–52. Retrieved from https://iaeces.com/Index/index.php/TASPA/article/view/64
25. Basnet, R. B., Johnson, C., & Doleck, T. (2022). Dropout prediction in MOOCs using deep learning and machine learning. *Education and Information Technologies, 27*(8), 11499–11513. https://doi.org/10.1007/s10639-022-11068-7
26. Kalaivanai, P. (2025). Assistive intelligent communication models for peer-based online learning environments. *Journal of Intelligent Assistive Communication Technologies, 2*(1), 72–80. Retrieved from https://aasrresearch.com/index.php/jaict/article/view/491
27. J. Karthika. (2025). Sparse Signal Recovery via Reinforcement-Learned Basis Selection in Wireless Sensor Networks. *National Journal of Signal and Image Processing*, *1*(1), 44-50.
28. S. Sindhu. (2026). Deep Reinforcement Learning-Based Adaptive Voltage Control for Smart Distribution Networks. *National Journal of Intelligent Power Systems and Technology*, 1-12.
29. Kagaba J. Bosco, S. M Pavalam, L.J. Mpamije. (2025). Assistive Distributed Communication Technologies for Collaborative Learning Environments. *Journal of Intelligent Assistive Communication Technologies*, *1*(1), 66-73.