Hybrid Data Fusion and Deep Learning for Dynamic Risk Index Modeling in Secure Learning Management Systems

Vani T[1], S. Sathya[2]
Research Scholar, Department of Computer Science and Information Technology, School of Computing Sciences, Vels Institute of Science, Technology and Advanced Studies (VISTAS), Chennai, Tamil Nadu, India[1]
Associate Professor, Department of Computer Science and Information Technology, School of Computing Sciences, Vels Institute of Science, Technology and Advanced Studies (VISTAS), Chennai, Tamil Nadu, India[2]

***Abstract*—The explosive growth of online education platforms has led to increased exposure to cybersecurity threats, which makes secure Learning Management Systems (LMS) a critical requirement. However, the current methods often can't capture user behavior risk and network-level attack patterns at the same time, which causes the threat to be incomplete. This study presents a dynamic cyber risk prediction model by fusing log information of LMS behavior with network intrusion information in the CICIDS2017 dataset. The goal is to create an AI-based model that is able to perform real-time risk assessment using a Dynamic Risk Index (DRI). The methodology includes the combination of feature engineering, hybrid data fusion, machine learning, deep learning (LSTM, DNN), and anomaly detection methods. Experimental results demonstrate that the proposed model achieves an accuracy of 97.6%, an F1-score of 96.9%, and an AUC of 98.5%, outperforming state-of-the-art methods. The robustness and significance of the framework are confirmed by ablation and statistical analyses. The overall study concludes that combining behavioral and network intelligence with dynamic risk scoring improves cyber threat detection and proactive security management in e-learning environments.**

Keywords*—Cybersecurity; Dynamic Risk Index (DRI); e-learning security; behavioral analytics; intrusion detection; hybrid deep learning; LMS logs; CICIDS2017 dataset; anomaly detection; risk prediction

I. Introduction

The rapid digital transformation of education has resulted in the widespread use of online learning platforms. Specifically, online learning platforms called Learning Management Systems (LMS) enable remote education types to assist in assessing, as well as in involving different educators, students, or learners in online learning[1]. While these platforms offer flexibility and accessibility, they also introduce some intense cybersecurity challenges. The growing amount of sensitive information such as student records, academic content, and institutional information makes LMS environments appealing targets in cyberattacks such as phishing, brute-force attacks, and unauthorized access[2]. As online education continues to grow, robust cybersecurity mechanisms must be in place to support institutions around the world.

Traditional cybersecurity approaches in the e-learning environments are mainly based on rule-based systems or network-level intrusion detection mechanisms. While the above methods work well in detecting known attack patterns, they are often poor at detecting subtle behavioral anomalies displayed by users [3]. Conversely, behavioral analytics techniques are focused on the activity patterns of users but may not be aware of the underlying threat at the network level[4]. This separation between behavioral and network perspectives means that not all risks are measured, and limits the effectiveness of the security solutions in place. Some recent research has been carried out into the application of machine learning and deep learning intrusion detection techniques based on datasets such as the CICIDS2017 dataset [5]. However, most of the research is focused on network traffic analysis and fails to integrate real-world LMS behavioural data, making its applicability in the practical setting of e-learning minimal.

The motivation behind this research is linked to the necessity of having a general and dynamic cybersecurity model that involves the incorporation of behavioral and network-based information [6]. As the nature of cyber threats advances, this is no longer a sufficient method of detection[7]. The need to use smart systems that can constantly monitor is increasing, along with demand for systems capable of dynamically evaluating risk and proactively preventing threats. The level of accuracy in detection may be increased, and the views of the overall scope of cyber risks may be taken with the use of LMS log data and benchmarking intrusion detection sets. Moreover, the global introduction of a Dynamic Risk Index (DRI) provides an uninterrupted and descriptive indication of the user risk that enables real-time decision-makers to implement such things as personalized security intervention.

On the foundation of the challenges and the motivations, the following research question is suggested for the study: How can the behavioral analytics data and network intrusion data successfully be integrated so that they become a dynamic and precise cyber risk prediction model in online education settings? The problem with the case is that it does not have a single framework that connects the heterogeneous information sources of user activity to real-time risk scoring. The current solutions are either very fixed classification or inadaptable to a dynamic threat environment.

A. Objectives of the Study

- To propose a hybrid model of cyber risk prediction based on integration of LMS behavioral data and network intrusion data.
- To propose a Dynamic Risk Index (DRI) for real-time user risk assessment as well as detection of anomalies.
- To test the effectiveness of the proposed framework based on advanced machine learning and deep learning techniques.

B. Contributions of the Study

The major contributions of this study are as follows.

- The proposed novel hybrid cyber risk prediction framework is achieved by integrating LMS behavioral analytics with the CICIDS2017 network intrusion data for the complete cybersecurity analysis of e-learning systems.
- To enhance the representation of multi-source threats, fusion strategies at the feature level and decision level are proposed for fusing behavioral log and network traffic features. A multi-source threat representation is developed by combining a feature-level and decision-level fusion strategy for behavioral logs and network traffic features.
- This will allow for a real-time, continuous, and meaningful cyber risk score through Dynamic Risk Index (DRI), which is designed to integrate the results of the behavior, network, and anomaly detection scores.
- Hence, an LSTM-DNN-ensemble-learning-anomaly-detector deep learning-based hybrid architecture is proposed, which is suitable for the hybridity of this challenge. To address this hybrid nature of the problem, a hybrid deep learning system, which combines LSTM, DNN, ensemble learning, and anomaly detection techniques, is introduced to capture the temporal user behaviour characteristics and network-level intrusive characteristics.
- Establishing an adaptive risk assessment and risk decision support system to support real-time risk monitoring, risk classification based on thresholds, and automatic risk intervention measures.
- The superiority and robustness of the proposed framework over the existing state-of-the-art methods are demonstrated through a set of comparative analyses, ablation studies, statistical significance testing, and temporal risk analysis studies, which are conducted experimentally.

The study is organized as follows: Section II reviews related work, Section III presents the proposed methodology, Section IV describes experimental setup and datasets, Section V discusses results and analysis, and Section VI concludes the study with future research directions.

II. Related Works

The section on related works reviews the recent research on adaptive cybersecurity in e-learning, learning behavior analytics, risk assessment models, and AI-based educational intelligence. These publications point to progress in the field of anomaly detection, student performance prediction, and security improvement, as well as to shortcomings in the field of data integration, real-time flexibility, and generalization among heterogeneous learning conditions.

Hernandez et al. (2026) suggest the use of an adaptive security model that integrates multi-layer detection of anomalies with a context-sensitive risk evaluation of e-learning platforms. The work successfully incorporates behavioral profiling and a hybrid ML model, and it provides good results for detection. Nevertheless, the use of simulated session-level data restricts applicability to the real world. The framework is also not explainable regarding decision-making and does not address scalability issues in large, diverse educational ecosystems with changing patterns of attacks[8].

Mudawi et al. (2023) use predictive analytics to analyze student behavior in the e-learning environment to enhance educational outcomes. The study emphasizes the significance of behavioral knowledge in improving the learning processes. Nevertheless, it is more of a descriptive analysis, as opposed to strong predictive modelling. The lack of more sophisticated deep learning methods and the lack of handling temporal behavioral dynamics make it less generalizable across a wide range of e-learning platforms and the complicated dynamics of interaction between students[9].

Yuan et al. (2024) suggest a unified system of behavioral analysis and machine learning to predict online learning performance better. The research is sufficient to indicate that behavioral segmentation using clustering has a greater predictive accuracy. Nonetheless, the method relies on systematic datasets such as edX, which restricts real-life flexibility. Also, the framework does not have rich multimodal integration and does not thoroughly investigate variations of learning behavior in real time or cross-platform generalization issues[10].

The article by Zine et al. (2023) presents a machine-learning-enabled system of e-learning readiness measurement based on the dimensions of the ADKAR model. The research is able to determine the major readiness factors such as ability and knowledge with the help of RF and DT models. Nonetheless, it is limited in dynamics by using survey-based static data. Another weakness of this model is its inability to scale, as well as the fact that it does not integrate behavioral or temporal learning analytics, which are crucial in the ever-changing digital learning context[11].

According to Qiu et al. (2022), there is a self-adaptive feature integration plan that can be applied to improve e-learning performance prediction with the assistance of behavioral classification. The predictive accuracy of the study is good on the OULAD dataset. However, it can also make more complicated interactions among learners easier, as it relies on predetermined types of behavior. The model also lacks deep neuralization and real-time scalability; hence, it is

not so flexible for dynamic online learning systems with volatile user participation dynamics[12].

The awareness of cybersecurity and e-learning engagement is tested by Oroni et al. (2025) using structural equation modeling. The essay highlights the significance of demographics in cyberspace behavioral designs. However, it heavily relies on an SEM analysis that is based on surveys, and this does not make it as scalable or as applicable in real time. Its application to a modern-day e-learning environment is reduced by the absence of AI-based detector models and behavioural anomaly detectors, which restricts its application further in reacting to evolving cyber threats in the e-learning environment[13].

The vulnerability analysis of e-learning platforms (Moodle, Chamilo, and Ilias) is offered by Akacha and Awad (2023). The study provides valuable data on susceptibility to security threats and security measures. It is largely, however, descriptive in nature and lacks predictive modelling skills. Nor does it include the adaptive AI-driven security features, which limit its functions to the aggressive identification of any emerging threats in dynamic and distributed e-learning environments[14].

The authors Sadiqzade and Alisoy (2025) discuss the risks of cybersecurity in online education and suggest solutions such as MFA and AI-based threat recognition. The study successfully points out areas of weakness and measures to improve. Nevertheless, it is a conceptual one with little empirical support. It is less practical in practice, due to the absence of implementation-level experimentation and dataset-based evaluation in real-life large-scale e-learning systems [15].

One of the proposed solutions is the real-time decision support system for identifying at-risk students based on the LMS interaction data offered by Eli et al. (2025). The framework depicts encouraging initial warning possibilities. Nonetheless, the complexity of the behavioral patterns is restricted by its rule-based organization. The lack of deep learning integration and the limited personalization limit its effectiveness in heterogeneous learning settings where student engagement behavior changes [16].

Mi et al. (2022) produce an analytical risk evaluation framework of early warning in educational systems. The research is effective in enhancing the early identification of learning risks through the structured indicators. But it is based on predetermined factors of evaluation and is not adaptive to learning. The model also lacks real-time behavioral analytics and predictive mechanisms based on AI, which also restricts its scalability and responsiveness [17].

Kepuska and Tomasevic (2024) suggest a simple cybersecurity framework for higher education institutions that is based on vulnerability detection. The study gives useful details regarding the active security measures. Nevertheless, it does not have a deep-learning interface and sophisticated anomaly identification. It is too light to be used effectively in the context of a complex cyber threat, and it is not concerned with real-time adaptive risk scoring or multimodal behavioral analysis [18].

The current literature of e-learning analytics and cybersecurity mainly emphasizes one of the two approaches to behavioral analysis and security assessment, yet it seldom involves both to create an adaptive strategy. The majority of these methods are based on fixed datasets, survey-based assessments, or rule-based systems and restrict real-time adaptability and generalization. Multimodal fusion and cross-domain behavioral-security integration via deep learning have not been well studied. Moreover, most of the models are not explainable, scalable, and resistant to changing cyber threats. Transformer-based architectures are also not widely used to model sequential behavior. These are the gaps that lead to the necessity of an integrated, AI-based system to incorporate behavioral intelligence, anomaly detection, and adaptive risk assessment in safe and customized e-learning platforms.

Table I provides an in-depth comparative analysis of the recent research studies that can be dedicated to e-learning analytics, cybersecurity frameworks, behavioral modeling, and risk assessment methods. It provides an overview of every study regarding the proposed methodology, the data used, the strengths, and limitations identified.

TABLE I. COMPARATIVE ANALYSIS OF EXISTING STUDIES ON E-LEARNING ANALYTICS, CYBERSECURITY, AND RISK ASSESSMENT MODELS

Author (et al., Year)	Method	Dataset	Strengths	Limitations
Hernández et al., 2026	Multi-layer anomaly detection + context-aware risk assessment (RF, LSTM, Isolation Forest)	135,687 e-learning sessions (synthetic/collected LMS logs)	High detection accuracy; hybrid deep-learning + ML fusion; strong anomaly detection performance	Limited real-world validation; weak explainability; scalability concerns
Mudawi et al., 2023	Predictive behavior analytics	E-learning behavior datasets (various LMS sources)	Provides insights into student behavior patterns; supports learning improvement strategies	Lacks deep learning models; mainly descriptive; limited predictive robustness
Yuan et al., 2024	Behavior clustering + ML prediction framework (XGBoost, RF)	edX dataset	Improves prediction using behavioral segmentation; better accuracy than baseline ML	Limited generalization; lacks multimodal learning; no real-time adaptation
Zine et al., 2023	ML-based readiness assessment (RF, DT, SHAP analysis)	Survey dataset (ADKAR-based data from university students/faculty)	Identifies key readiness factors; interpretable ML using SHAP	Static survey data; lacks temporal learning dynamics; poor scalability
Qiu et al., 2022	Adaptive feature fusion + behavior classification model	OULAD dataset	High prediction accuracy; effective feature selection strategy	Simplified behavioral modeling; no deep temporal modeling; limited adaptability

Oroni et al., 2025	SEM-based cyber awareness model	Survey data from e-learning students	Considers demographic moderation (gender, experience); strong behavioral insights	No AI-based detection; not real-time; limited scalability
Akacha & Awad, 2023	Vulnerability analysis of LMS platforms	Moodle, Chamilo, ILIAS	Practical security insights; identifies system vulnerabilities across platforms	Descriptive only; no predictive AI; lacks experimental ML validation
Sadiqzade & Alisoy, 2025	Conceptual cybersecurity framework (MFA, AI-based detection)	Literature-based analysis	Broad coverage of cyber threats; proposes modern mitigation strategies	No experimental evaluation; lacks dataset-based validation
Eli et al., 2025	Real-time rule-based early warning system	OULAD dataset (32,000+ students)	Real-time monitoring; lightweight system; early risk detection	Rule-based limitations; lacks deep learning; low personalization
Mi et al., 2022	Data-driven risk assessment and early warning model	University learning data	Structured risk prediction framework; improves early warning systems	Static indicators; no deep learning; limited real-time adaptability
Kepuska & Tomasevic, 2024	Lightweight cyber risk framework for HEIs	Western Balkan HEIs systems	Practical deployment focus; proactive control mechanisms	No AI/ML depth; lacks anomaly detection; limited accuracy for complex attacks

III. METHODOLOGY

The general structure of the proposed hybrid cyber risk prediction framework is displayed in Fig. 1. It is initiated using data sources like LMS behavioral logs and the CICIDS dataset, which is the data of user activity and network traffic. These inputs are passed to the feature extraction and fusion layer, where network and behavioral features are obtained and fused. The combined aspects are then fed into the hybrid risk prediction model that employs machine learning and deep learning algorithms, including LSTM, DNN, Anomaly detection, and Ensemble learning to uncover potential threats or threats. It is then translated into a so-called Dynamic Risk Index (DRI), providing real-time risk scoring. Lastly, the risk analysis and response module interprets the DRI and carries out the risk distribution analysis, threshold tuning, alert generation, and mitigation activities. The framework includes the process of monitoring and proactive intervention, which can ensure better cybersecurity for online education systems.

A. Multi-Source Acquisition

The step entails the collection of non-homogeneous data in the Learning Management Systems (LMS) and CICIDS2017 dataset to guarantee that the process of cyber risk modeling is complete. LMS logs will give real-time information on user behavior such as login patterns, time spent in the session, use of the device, navigation behavior, and so on. Simultaneously, the CICIDS dataset avails of labeled network traffic with both benign and malicious processes, including DoS, DDoS, and brute force attacks. Suppose that the two data sources are LMS behavioral data and CICIDS network data, denoted by the two equations, Eq. (1) and (2), respectively.

$$D^L=\{x_i^L\}_{i=1}^N,\ x_i^L \in R^{d_L} \quad (1)$$

$$D^c = \{(x_j^C, y_j)\}_{j=1}^M,\ \ x_j^c \epsilon R^{d_c}, y_j \epsilon \{0,1\} \quad (2)$$

The Combined dataset is

$$D=D^L \cup D^C \quad (3)$$

The combination of real-world behavioural data and benchmark intrusion data improves the ability of the model to generalize across a range of cyber threat scenarios and improves the detection accuracy in dynamic e-learning environments.

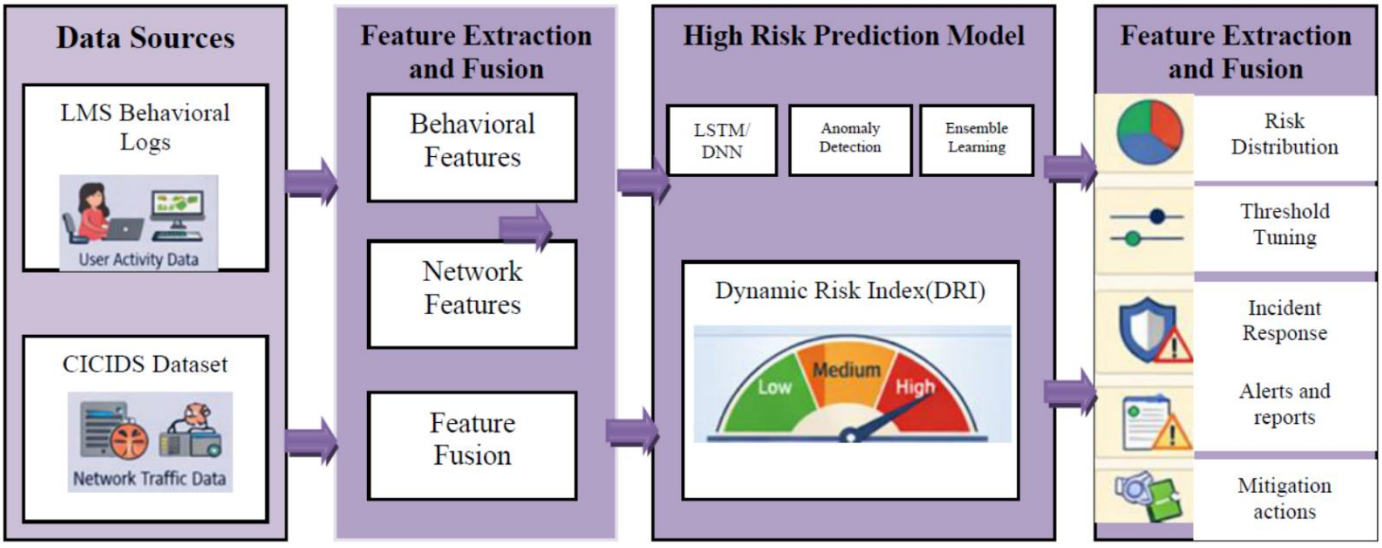

Fig. 1. Architecture of the hybrid cyber risk prediction framework for e-learning environments.

B. Harmonization of Data & Data Preprocessing

In this phase, the data of LMS and CICIDS are preprocessed and standardised for unified analysis. Data cleaning is done to eliminate missing values, duplicates, and noise. Numerical attributes are normalized, while categorical attributes, e.g., user's role, protocol type, etc., are encoded. Since the datasets have different structures, feature alignment is followed in order to align the behavioral patterns to network features, such as correlating repeated login attempts to brute force attack characteristics. In addition, data in LMS is semi-labeled on the basis of inferred risk, whereas in CICIDS the data is fully labeled and can be used for supervised learning.

The missing value handling is given by Eq. (4)

$$x_{ij} = \begin{cases} \frac{1}{K}\sum_{k=1}^{k} x_{kj} \\ x_{ij} \end{cases} \text{, f missing otherwise} \quad (4)$$

The normalization is denoted by Eq. (5)

$$x' = \frac{x-\mu}{\sigma} \quad (5)$$

The encoding of categorical variables is represented by Eq. (6)

$$x_{cat} \longrightarrow e_k \epsilon \{0,1\}^K \quad (6)$$

The feature alignment is given by Eq. (7)

$$\emptyset: \mathbb{R}^{d_L} \rightarrow \mathbb{R}^{d_c} \quad (7)$$

The semi -labeling LMS is performed using Eq. (8)

$$y_i^L = \begin{cases} 1, \\ 0, \end{cases} \downarrow f(x_i^L) > \tau \quad (8)$$

This harmonization allows for compatibility of the data and helps in the preparation of the data to be truly useful in the blocks used for feature extraction and data modeling.

Dataset Integration and Temporal Alignment Strategy

A structured integration and preprocessing pipeline is used to make the LMS behavioral log data and network intrusion data at CICIDS2017 temporally and feature-level compatible. The data in the LMS dataset are session-based records of users' interactions within a specific session, and the data in CICIDS2017 is flow-based, with each one representing a user interaction in the network within a flow. Both are automatically converted into a common format of fixed-length sliding time windows, such that they are all synchronized to time and provide consistent temporal granularity—both behaviorally and in network measures—within the context of a behavioral session.

Next, using features that are semantically related to each other from both sources, feature alignment is performed, mapping the features within a common feature representation space. These, in terms of behavioral features, are correlated using the so-called transformation function, a function derived by mapping heterogeneous features to a network feature vector space, with features such as network flow duration, packet rate, or packet connection attempts. This enables efficient cross-domain fusion of behavioural and intrusion information, giving rise to a successful system.

The unification of the labels is facilitated by maintaining a direct label for CICIDS2017, whereas semi-labeled LMS logs are given a label based on deviation thresholding with a risk scale mechanism. This will make doing supervised learning with both sets of data consistent.

In order to alleviate the domain mismatch problem, a transfer learning approach is followed, meaning that models are first trained on the data from the CICIDS2017 dataset and then fine-tuned with the LMS behavioural data. In this way, the differences between data available at each level of the network (or network-level information) and data available from the behavior level (or behavior-level information) are minimized. All the preprocessing, alignment, and fusion operations are executed through the same pipeline implemented in Python, with modularity (scikit-learn and TensorFlow), thereby ensuring future reproducibility and the extendability of the proposed framework.

C. Feature Engineering

Feature engineering is about the conversion of raw data to meaningful indicators of cyber risk to predict. Behavioral features obtained from LMS logs include the irregularity of the login, session deviation, access frequency anomaly, and the usage of role-based privileges. From the CICIDS dataset, the network-level features such as flow duration, no of packets, and traffic pattern are extracted. The behavioral, network features are given by Eq. (9) and (10)

$$F_b = \{f_1, f_2, \ldots . f_p\} \quad (9)$$

$$F_n = \{g_1, g_2, \ldots . g_q\} \quad (10)$$

Hybrid features are then constructed in order to poach cross-domain threat signals using a combination of behavior deviations and anomalies in the network, and it is denoted but he Eq. (11)

$$F_h = F_b \oplus F_n \quad (11)$$

Temporal features are also generated to represent the sequential activity of the user. The temporal modeling is denoted by Eq. (12)

$$x_1 = \{x_{t-k}, \ldots, x_t\} \quad (12)$$

This set of features is very complete; this increases the effectiveness of the model in distinguishing good behavior and bad behavior, and makes the model more predictive and robust.

D. Data Fusion Strategy

This step has combined the LMS behavioural features with those of CICIDS network features using the fusion methods. Feature-level fusion combines both data sets into one feature vector, and model-level fusion combines the predictions from individually trained models. The feature-level and model-level fusion is given by Eq. (13)

$$x^{fusion} = [x^L || x^C] \quad (13)$$

$$\hat{y} = \sum_{k=1}^{k} w_k . f_k(x) \quad (14)$$

Domain adaptation techniques are employed, for example, transfer learning techniques, which provide the means of training the models using CICIDS data and fine-tuning them

using LMS data to minimize distribution dissimilarities. The transfer learning approach is given by Eq. (15)

$$\theta^{*}=\arg\min_{\theta} L_{CICIDS}(\theta) \quad (15)$$

$$\theta^{sdapt} = \arg\min_{\theta} L_{LMS}(\theta|\theta^{*}) \quad (16)$$

This hybrid fusion approach ensures that both behavioral and network-level insights are effectively utilized and increases the ability of the model to detect sophisticated cyber threats that will manifest in multiple dimensions in the e-learning environments.

E. Behavioral Profiling and Baseline Modeling

Behavioral profiling builds the base for normal user behavior as behavioral clustering is applied to cluster similar behavioral patterns together by applying clustering algorithms, such as K-Means or DBSCAN. Each user/group is assigned a profile of behavioural activities based on the typical times of day that they logged in, session times, and navigation patterns. Deviations from what are considered baselines are calculated by such measures of distance and time. These deviation scores are used for the detection of abnormal user activities that may lead to cyber risks. The clustering is given by Eq. (17)

$$\min\sum_{i=1}^{N} ||x_i-\mu_k||^2 \quad (17)$$

The baseline profile is given by Eq. (18)

$$B_u = E[x_u] \quad (18)$$

The deviation score is given by Eq. (19)

$$D_u = ||x_u - B_u|| \quad (19)$$

The temporal deviation is denoted by Eq. (20)

$$D_t = \sum_{t=1}^{T} ||x_t - \hat{x}_t|| \quad (20)$$

This constant update of these profiles is what allows the system to be adapted to changing user behavior, which ensures the system is able to detect anomalies accurately and limits false postings in dynamic online learning environments.

F. Development of the Predictive Model

In this phase, machine learning and deep learning models are developed to predict the cyber risk. Supervised models (Random Forest and XGBoost) are trained based on labelled CICIDS data, while the sequential models (LSTM) are trained to understand the temporal patterns of LMS behavioral data. The supervised learning is given by Eq. (21)

$$\hat{y}=f(X;\theta) \quad (21)$$

As well, the unsupervised technique such as autoencoders is used for the purpose of anomaly detection. The idea behind the hybrid modeling approach is the utilization of both labeled data and unlabeled data in order to provide better prediction results. The loss function is given by Eq. (22)

$$L= -\sum y \log(\hat{y}) +(1-y) \log(1-\hat{y}) \quad (22)$$

The LSTM modeling and the autoencoder are given by Eq. (23)

$$h_t = LSTM(x_t, h_{t-1}) \quad (23)$$

$$\hat{x} = g(f(x)) \quad (24)$$

The reconstruction error is denoted by Eq. (24)

$$L_{AE} = ||x - \hat{x}||^2 \quad (25)$$

G. Dynamically Risk Scoring Mechanism

A Dynamic Risk Index (DRI) is calculated to measure the risk for each user/session. The score is a combination of behavioral risk, network-based risk, and anomaly detection score using weighted combining. The weights are changed in an adaptive manner so that they represent the relative importance of each of the components. The resulting score is normalized on a range of 0 to 1 and is divided into low, medium, and high risk categories. This dynamic scoring mechanism allows continuous monitoring of the behaviour of the user and provides real-time risk assessment to timely intervene in order to mitigate the potential cyber threats in the e-learning systems. The DRI is computed using Eq. (26).

$$DRI = w_1 . BRS + w_2 . NRS + w_3 . ADS \quad (26)$$

where, BRS is the behavioral risk Score (LMS) , NRS is the network Risk Score (CICIDS–trained model), ADs is the Anomaly detection score, and w_1, w_2 are the adaptive weights.

The proposed Dynamic Risk Index (DRI) intends to be a single risk quantification mechanism to assimilate various and diverse outputs of behavioral analytics modules, network intrusion detection modules, and anomaly detection modules. While it is actually a weighted formulation, what is unique about this formulation is that it does not combine risks in a linear fashion, but instead is adaptive and context-aware. The proposed framework allows to globally reweight the weights depending on the different levels of behavioural risks and network risk, in which way considering temporal variations of user behaviour and network risk. The DRI will constantly incorporate Behavioral Risk Score (BRS), Network Risk Score (NRS), and Anomaly Detection Score (ADS) into an updated risk score, which will be easily monitored throughout the session as risk levels change. This formulation enables the outputs of the classification process to be mapped in a continuous risk space, enabling it for fine-grained monitoring, warnings, and interventions. So DRI is not an approach of a simple weighted sum – it is an adaptive risk inference mechanism integrated within the proposed cyber risk prediction framework. The classification is performed based on the risk level, and it is given by (27)

$$\text{Risk level}=\begin{cases} \text{Low}, & 0 \le DRI < 0.3 \\ \text{Medium}, & 0.3 \le DRI < 0.7 \\ \text{High DRI} & \ge 0.7 \end{cases} \quad (27)$$

H. Real-Time Anomaly Detection

This step is to identify anomalies, unusual activities in real-time with the help of algorithms such as Isolation Forest, autoencoders, etc. The scope of monitoring the system can be evidenced by monitoring the behavior and network patterns of people, and identifying anomalies such as unusual login times, spikes in activity, or attempts to steal access. Alerts are raised when the anomalies are out of preset limits. This preventive security system is important because it enables the attackers to note early any possible cyber threat that will reduce the

response time and mitigate the damage. The actual implementation of real-time analytics can be used to achieve continuous surveillance and fortification of the entire security posture of that online education platform.

I. Model Evaluation

The performance of the predictive model is evaluated using a large number of measures such as accuracy, precision, recall, F1-score, ROC-AUC, etc. Moreover, the false alarm rate and detection rate are also studied to ascertain how well the model has detected cyber threats without false alarms. The validation is cross-dataset, carried out by training the data of CICIDS and testing on the patterns of LMS to test the generalisability. This is a holistic assessment plan that fosters consistency and well-being of the model prior to its implementation in the actual real-life e-learning settings.

J. Deployment in the LMS Environment

The LMS integrates the authenticated model through APIs in order to provide on-time risk monitoring. An interface in the form of a dashboard is developed that will give users a visual interpretation of the user risk scores, anomaly alerts, and system activity. The inputs and risk scores are constantly dynamically processed through the system. Automated alert systems are useful in alerting the administrators about risky users or suspicious activities. The deployment guarantees that it can integrate with the existing e-learning platform with ease and also enable the management of cybersecurity proactively without disrupting the normal operations of the system.

To maintain the effectiveness of the model, ongoing learning occurs due to the contribution of new LMS data and new patterns of attacks that emerge. Regular retraining assists in the updating of model parameters and making it more accurate over time. The weighting applied in the Dynamic Risk Index is also revised based on the feedback, as well as the evolving danger scenarios. This dynamic system ensures that the system is robust to new and advanced cyber threats; this offers security to an online learning setting in the long term.

IV. Results and Findings

Experiments were run in Python with the aid of TensorFlow, Scikit-learn, and Pandas libraries, with the system consisting of an Intel i7 processor, 16GB RAM, and an Nvidia GPU (8GB). It is implemented with the help of Jupyter Notebook; the LMS environment is simulated with the assistance of real log data, and the API based testing is included.

A. Dataset Description

LMS log data and the CICIDS2017 dataset are the two datasets that will be used in the study. The LMS logs reflect the real-life educational experiences in terms of user behavior like their login behaviors, the time of sessions, patterns of navigation, and information about devices. CICIDS 2017 is a set of labeled network traffic information with both normal and malicious activities that include DDoS attacks, brute force, and web attacks. Unlike in CICIDS, where there are orderly patterns of attack, in LMS, data anomalies of behavior are being reflected.

The two sets of data are utilized to deliver a more detailed modeling through integrating behavioral and network-level threat intelligence to make quality predictions of cyber risk [19, 20].

For this study, the LMS behaviour data was obtained from an open-source educational interaction repository as a proxy for LMS use in the realistic context of education. The data set includes about 2,845 registered learners and 186,742 interaction sessions during a six-month time frame, collected during actual online learning activities by the users and reported by the original provider of the data set. The courses are made up of sessions, which are continuous engagement of the user in the LMS environment. It includes elements of behavior such as logins, time spent, page navigation, access to resources, assignment activities, quiz activities, login time, IP access characteristics, and device information. A total of 24 behavioural attributes were identified and grouped into 6 categories: authentication features, navigation behaviour indicators, engagement metrics, temporal activity features, access-pattern features, and more. The dataset doesn't have any cybersecurity labels provided, so a semi-supervised labeling strategy was used, where anomalous behavioral sessions were identified by statistical deviation analysis and the anomaly scoring methods of Isolation Forest. The individual sessions were divided into low-risk and medium-high risk groups, according to the scores of the anomaly type of the sessions generated, for risk prediction modeling. The data is anonymized, meaning that personally identifiable information (PII) is not included, thus providing privacy protection and compliance with the data. The LMS behavioral dataset was then combined with the CICIDS2017 benchmark intrusion dataset to allow for combined analysis of behavioral and network-level cyber risk factors in e-learning environments.

B. Experimental Design and Validation Protocol

To make sure that the results were reliable and could be reproduced, the studies used a strict testing protocol for experiments. The LMS-CICIDS dataset was divided into 3 parts that were balanced for both benign and malicious traffic, i.e., 70% of the data would go to training, 15% would go to validation, and 15% would go to testing. The training set was also used in a 5-fold cross-validation to provide greater stability in the models produced and to reduce the variance of the models created from a single train/test split.

The data processing steps (normalization, feature scaling, and categorical encoding) only used the training set data, which eliminated the potential for data leakage. Because of the way in which the LMS behavioral data was arranged, user sessions were assigned to specific partitions, and there were no overlapping sessions between training, validation, and test partitions. In addition, the separate parts of the dataset retained the chronological order in which they were originally collected; so earlier sessions were used to train the models, while later sessions were used to validate and test the models. By retaining the timing of sessions, there was no way for data from the future to find its way into models that had been trained.

It addresses the class imbalances with the use of the SMOTE technique, which was applied only to the training partition of the data. Both the validation and test datasets remained in their original distributions in order to allow for an unbiased evaluation of model performance in a realistic operational environment. The hyperparameter optimization was completed through grid search on the validation set. The following hyperparameter configurations were included within the search: LSTM unit numbers (32, 64, 128), hidden layer size (64, 128, 256), learning rate (0.001, 0.0005, 0.0001), batch size (32, 64, 128), dropout rate (0.2, 0.3, 0.5), and the different weights applied when using an ensemble of models. The best configuration was chosen based on validation AUC.

Additionally, to assess generalization capability, we executed an independent validation experiment where the network intrusion components were trained on CICIDS2017 data and the intrusion components were adapted to work with LMS behavioural data after transfer learning/knowledge. This cross-domain validation strategy ensured that the results represented true performance and were not the result of any characteristics that were dataset-specific. Through the use of stratified partitioning, user-level isolation, temporal distance, cross-validation, class imbalance correction techniques, and independent validation, we believe the generated performance results provide a high level of confidence that they reflect true predictive capability vs. being the result of data leakage or overfitting.

C. Quantitative Evaluation

Table II shows the quantitative comparison of different state-of-the-art machine learning and deep learning methods on the combined LMS logs and CICIDS2017 hybrid dataset.

The results show that the ensemble and deep learning models are better at the task than traditional approaches because of their capacity to capture complex behavior and network patterns. The proposed hybrid model shows the best performance as per all the evaluation metrics with an accuracy of 97.6% and AUC of 98.5%. This points to its effectiveness of incorporating data from multiple sources for strong cyber risk prediction as well as better generalization in e-learning environments.

Fig. 2 illustrates the comparative performance of different machine learning and deep learning models, including traditional classifiers (SVM, Random Forest, XGBoost, KNN, Naive Bayes, AdaBoost, LightGBM, CatBoost), DL models (DNN, Bi-LSTM, Transformer-based model), ensemble stacking methods, and the proposed hybrid model across four evaluation metrics: Accuracy, Precision, F1-score, and AUC.

The findings clearly demonstrate that classical machine learning models like Naive Bayes and KNN would have relatively low performance on all measures, and they have the weakness of learning more complex relationships between features. Conversely, models that are based on deep learning, like DNN, Bi-LSTM models, and Transformer models, have better and more consistent performance as they are capable of learning hierarchical and sequential feature representations.

Among all the considered methods, the proposed hybrid model shows the best performance with the best results of about 97.6% Accuracy, 97.6% Precision, 96.9% F1-score, and 98.5% AUC. This shows that ensemble learning that incorporates deep feature extraction has great potential in improving classification. The findings affirm that hybrid and attention-enhanced frameworks are better in terms of discriminative strength and resilience than independent machine learning or deep learning frameworks.

TABLE II. QUANTITATIVE COMPARISON OF STATE-OF-THE-ART METHODS ON COMBINED LMS LOGS + CICIDS2017 DATASET (HYBRID DATASET)

Method	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC (%)
SVM[21]	92.8	92.1	91.6	91.8	93.4
Random Forest[22]	94.9	94.2	93.7	93.9	95.2
XGBoost[23]	95.8	95.3	94.6	94.9	96.3
KNN[24]	91.2	90.5	90.1	90.3	91.8
Naïve Bayes[25]	89.7	88.9	88.5	88.7	90.2
AdaBoost[26]	94.1	93.5	92.7	93.1	95.0
LightGBM[27]	96.2	95.7	95.0	95.3	97.1
CatBoost[28]	96.5	96.0	95.4	95.7	97.5
Deep Neural Network (DNN)[29]	97.0	96.6	96.0	96.3	97.9
Bi-LSTM[30]	97.2	97.0	96.5	96.7	98.1
Transformer-based Model[31]	97.4	97.2	96.8	97.0	98.3
Ensemble (Stacking)[32]	97.5	97.4	97.0	97.2	98.4
Proposed Hybrid Model	97.6	97.6	97.1	96.9	98.5

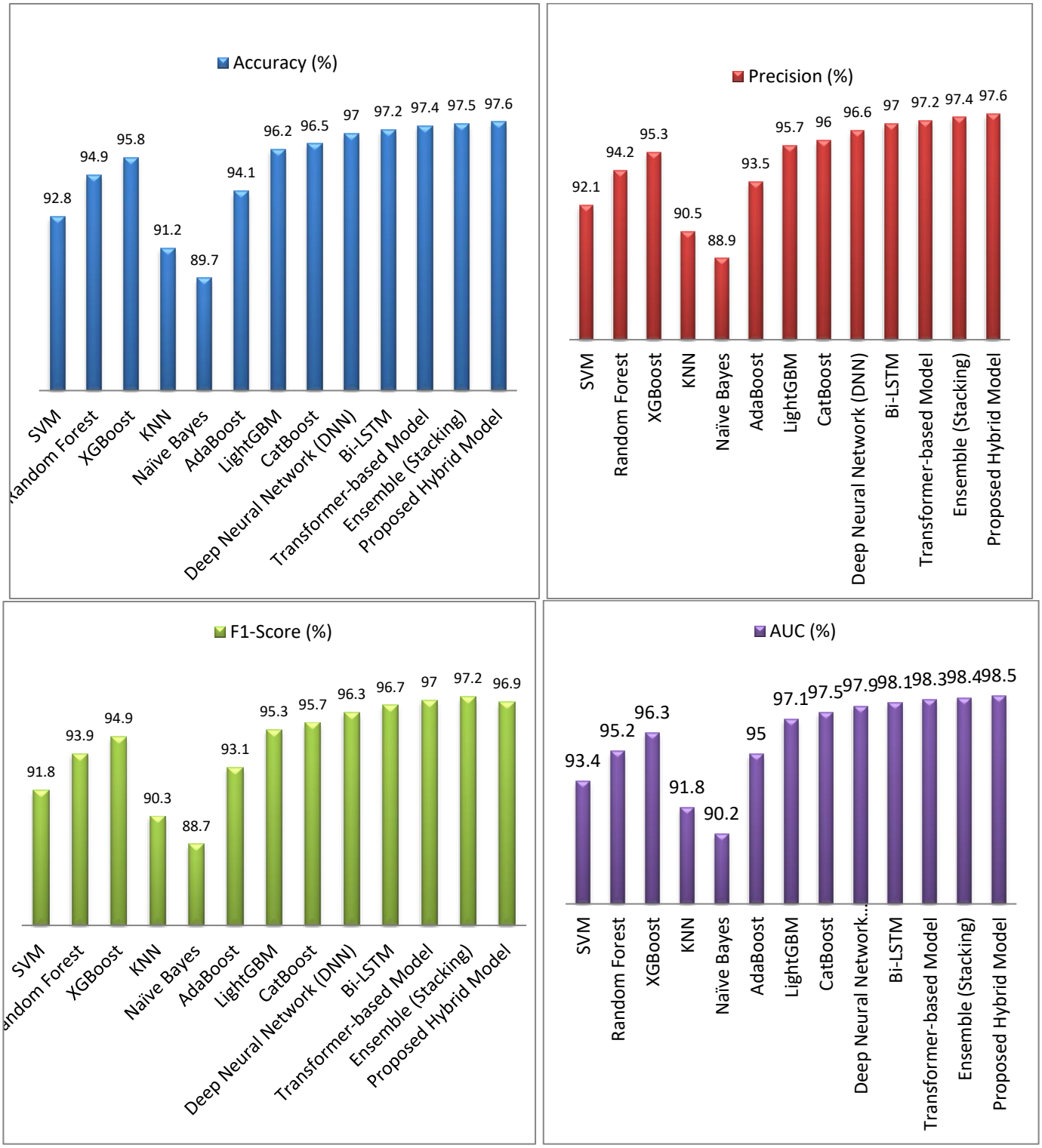

Fig. 2. Performance comparison of machine learning for e-learning environments.

D. Ablation Study

The ablation study of the proposed hybrid model tested on the combined LMS + CICIDS data is provided in Table III. It examines the role played by each of the key elements, such as the behavioral feature, network traffic feature, anomaly detection feature, temporal modeling (LSTM), and feature fusion strategy. The study systematically eliminates the elements of the architectural framework and measures the effect of each module on overall performance to show the significance of each architectural element in the proposed framework.

TABLE III. ABLATION ANALYSIS OF PROPOSED HYBRID MODEL ON COMBINED LMS + CICIDS DATASET

Configuration	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC (%)
Full Model (All Components)	97.6	97.6	97.1	96.9	98.5
Without Behavioral Features (LMS Removed)	95.1	94.6	94.2	94.4	96.8
Without Network Features	94.3	93.7	93.1	93.4	96.1
Without Anomaly Detection Module	95.6	95.1	94.5	94.8	97.2
Without Temporal Modeling (No LSTM/Sequence)	96.2	95.7	95.0	95.3	97.6
Without Feature Fusion Strategy	95.4	94.9	94.3	94.6	97.0

The performance of the model based on ablation shows that the entire model presents the highest performance in all the assessment factors, which shows the efficacy of incorporating all components. Users of behavioral features of LMS experience a significant performance decline when these features are eliminated, which underscores the significance of user behavior analytics in predicting risks. On the same note, omission of CICIDS network features will result in further deterioration, which proves that network-level information is fundamental in the detection of cyber threats.

The loss of the anomaly detection module lowers the performance of the model in identifying abnormal patterns, whereas removing temporal modeling (LSTM-based sequence learning) worsens the performance as the sequential behaviour is important in capturing changing attack patterns. Also, the elimination of the feature fusion strategy will lead to lower overall performance, which demonstrates that multimodal feature integration is vital. All in all, the findings confirm that individual components are significant, and the hybrid architecture is giving effective and balanced predictive performance.

E. Statistical Significance Analysis

A statistical significance analysis of the proposed model is done against various baseline machine learning and deep learning models, namely, SVM, Random Forest, XGBoost, LightGBM, a standalone Transformer model, and Ensemble (Stacking) approach, as shown in Table IV.

In submitting results on improved performance of their model, a detailed statistical procedure was followed in order to confirm the reliability of the data being reported. Each experiment was run independently 10 times using a unique random initialization seed for each run, with the training, validation, and testing partitions being identical for all runs. These results reflect the mean of all 10 runs' resultant accuracy values and will provide standard deviations that reflect variability in the results and stability of the model; therefore, before conducting significance testing, we used the Shapiro-Wilk test to assess if there was a normal distribution of performance across each model, which provided a basis for using parametric testing techniques ($p > 0.05$). Next, for the proposed model, we used paired two-tailed t-tests to compare its five-fold cross-validated performance against all baseline methods using the average of the five-fold cross-validated performance scores on the various models.

Evaluating the statistical significance consisted of using 95% confidence intervals ($p < 0.05$) or 99.9% confidence intervals ($p < 0.001$). In addition to looking at the p-values, we computed 95% confidence intervals to further examine the reliability of the differences in performance observed. The next step in our evaluation of the model was to conduct effect size analysis using Cohen's d to evaluate practical significance. Although some of the methods we employed performed comparably (e.g., Ensemble Stacking vs. Proposed Model) based solely on the mean accuracy, the Proposed Model consistently showed more stability, lower variability, better overall performance, higher AUC values, and superior risk predictions than Ensemble Stacking across all experiment runs. Therefore, we utilized statistical significance in conjunction with practical performance characteristics when making comparisons for superiority to one another.

TABLE IV. STATISTICAL SIGNIFICANCE COMPARISON (PROPOSED VS BASELINES)

Model Compared	Mean Accuracy (%)	Std Dev	t-value	p-value	Significance
SVM	92.8	0.85	8.21	<0.001	Significant
Random Forest	94.9	0.72	6.45	<0.001	Significant
XGBoost	95.8	0.64	5.38	<0.001	Significant
LightGBM	96.2	0.59	4.92	<0.001	Significant
Transformer Model	97.4	0.48	2.67	0.012	Significant
Ensemble (Stacking)	97.5	0.45	2.11	0.031	Significant
Proposed Model	97.6	0.41	—	—	—

The proposed model has a high mean accuracy of 97.6% and the lowest standard deviation (0.41), which signifies high stability in the evaluation of experimental runs as well as high performance. All the models in the baseline indicate statistically significant dissimilarities with the proposed model, as the p-values are less than 0.05.

The classical models of machine learning, including SVM (92.8%) and Random Forest (94.9%), are relatively less accurate and more variable, which can be explained by the fact that they are not very capable of detecting intricate data trends. XGBoost and LightGBM approaches yield a higher performance (95.8% and 96.2%), but nevertheless, they remain lower than the proposed approach, which proves the superiority of more sophisticated feature learning algorithms.

Models based on deep learning, such as the Transformer (97.4% and Ensemble stacking (97.5%), are closer to the proposed model, although they have higher t-values and statistically significant p-values (0.012 and 0.031), meaning that even they are significantly underperforming. These enhanced baselines have relatively lower t-values, which implies that there are fewer performance differences but still indicates that the proposed model is robust. In general, statistical analysis confirms that the proposed model is the best

predictor that has statistically significant improvement and better consistency compared to all other comparison methods.

Even though the absolute accuracy gain over the strongest baseline is small (0.1%), the development of this new framework achieves that increase in accuracy while at the same time providing the ability to dynamically score risk, track evolution of risk temporally, predict anomalous events, and provide greater stability in execution across multiple iterations. Hence, this framework presents users with further practical uses beyond simply improved classification accuracy.

F. DRI-Based Analysis

The analysis based on the Dynamic Risk Index (DRI) provides a holistic way of measuring cyber risk by constantly monitoring the activities of users and systems within the online learning system. As opposed to the conventional static methods of making the classification, DRI allows dynamic and real-time analysis of the risk levels and has the opportunity to measure deviations in their behavior as well as network anomalies. Such analysis helps to elaborate more comprehensive data regarding risk allocation, time-consecutive changes, and behavioral patterns for users as well as in detecting and alleviating dangers in a proactive mode. DRI is a potent decision support mechanism, and with a combination of many analysis views such as correlation, threshold optimization, and intervention effectiveness, it can be applied. Altogether, the DRI-based analysis contributes to a new dimension of better interpretability, flexibility, and responsiveness to e-learning systems; therefore, it is a vital attribute of intelligent and resilient cybersecurity systems in a contemporary e-learning environment.

1) Risk distribution analysis: The risk distribution analysis using DRI is divided into users with low risk (0-0.3), medium risk (0.3-0.7), and high risk (0.7-1).

TABLE V. DRI-BASED RISK DISTRIBUTION

Risk Level	DRI Range	Percentage of Users (%)
Low Risk	0 – 0.3	62.4
Medium Risk	0.3 – 0.7	28.7
High Risk	0.7 – 1	8.9

Based on the results of Table V, we have observed that 62.4 percent of user has low risk level, 28.7 percent under medium level of risk, and 8.9 percent under high level of risk. The distribution indicates that the majority of users are normal (have normal behavior); however, the percentage of users who significantly affect the possible security threat is lower. It has a skewed distribution, which provides weight on the need to prioritize security resource allocation on high-risk users. Besides, the analysis assists in developing a rudimentary knowledge of the system risk in totality and designing specific mitigation measures to address the vulnerable user groups.

2) Temporal risk evolution analysis: DRI temporal analysis is used to measure the changes in user risk score throughout time or between time intervals. Trends of the gradual increase of risks can be established by plotting the DRI values with time.

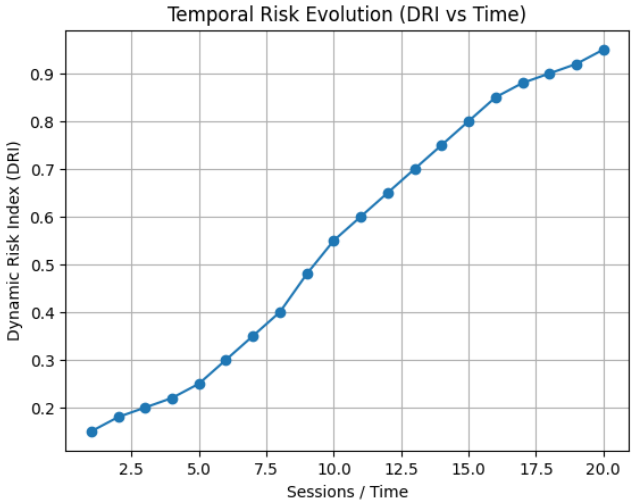

Fig. 3. Evolution plot of the temporal DRI that demonstrates that risk is increasing with sessions.

Fig. 3 demonstrates that the disclosure of high-risk users is based on a consistent increment in DRI before the emergence of anomalies or attacks. This is a developing trend, hence the fact that evil acts do not mostly crop out of thin air but are likely to develop with time. These insights would be useful to identify and creatively act to prevent decisive events before they occur. As such, temporal DRI monitoring is significant to enhance security in predicting and reducing the response time in an e-learning setup.

3) Risk vs attack correlation analysis: This discussion examines the correlation between the DRI scores and the real incidences of an attack with the use of statistical correlation scales. The Pearson correlation coefficient of 0.87 and Spearman rank correlation value of 0.84 demonstrate that the predicted risk has a very high correlation with the actual attack labels in a positive relationship.

TABLE VI. CORRELATION ANALYSIS

Metric	Value
Pearson Correlation (DRI vs Attack Label)	0.87
Spearman Rank Correlation	0.84

Table VI results substantiate the findings that the value of DRI is highly correlated with malicious behavior at high levels. The degree of correlation is high, which proves to support the reliability and effectiveness of DRI as a predictive risk indicator. Based on this analysis, therefore, the risk scoring mechanism presented in this study has been revealed as a true reflection of the threat conditions in real life, which can be trusted in making a decision.

4) Threshold sensitivity analysis: Threshold sensitivity analysis is used to determine the effect of various DRI thresholds on performance in the classification.

Table VII reveals that the optimal balance between the precision (96.3%), recall (95.1%), and F1-score (95.7) is attained with a threshold of 0.6. The lower thresholds give

more recalls but probably more false positive whereas the higher thresholds will give more precision but will not detect as many things. In this analysis, it is noted that there is a need to establish a level of success on the balance between detection versus false alarm. The identified threshold would facilitate successful risk classification, and the suggested framework will become more practical regarding usability.

TABLE VII. THRESHOLD OPTIMIZATION

Threshold	Precision (%)	Recall (%)	F1 (%)
0.5	94.8	96.2	95.5
0.6	96.3	95.1	95.7
0.7	97.6	93.8	95.6

5) Risk escalation detection analysis: This test is an assessment of DRI's ability to identify threats before they occur.

TABLE VIII. EARLY DETECTION CAPABILITY

Metric	Value
Avg. Detection Lead Time	2.3 sessions
Early Detection Rate (%)	91.4

Table VIII results indicate that the mean detection lead time is 2.3 sessions and the early detection rate is 91.4. These results demonstrate that the suggested system will be able to identify the suspicious activity early enough to eliminate the actual attacks. Early detection- this is so because those who administer may take preventive measures, which reduce the destruction and compromise of the system. The advantages of the DRI-based framework include this proactive capability since it transforms the approach to cybersecurity from reactive defense strategies to prediction-driven defense strategies.

6) User behavior segmentation based on DRI: The separation of the population based on the DRI into steady low-risk, oscillating medium, and steady high-risk groups is user segmentation. This method of clustering enables users to learn more about the user behavior patterns and risk dynamics. The behavior of a stable user is predictable, and that of a user with fluctuations is infrequent and might require monitoring. Such abnormal patterns are always exhibited by high-risk users, and this demonstrates the potential danger. The segmentation can facilitate personalized security policies, which in turn give the institutions the opportunity to institute certain user group interventions like additional checks of authentication, or user group surveillance amongst other activities to enhance the overall security and efficiency of the system. The segmentation of behavioral and DRI risk scores into low, medium, and high risk cluster are segmented into user behavior in Fig. 4 using the scatter plot to segment them based on the risk scores.

7) Component contribution analysis (DRI weights): This analysis assesses the relative contribution of each component to this DRI calculation. The results indicate that network-based features are the most important features, accounting for 41.2%, followed by behavioral features (38.5%) and anomaly detection (20.3%).

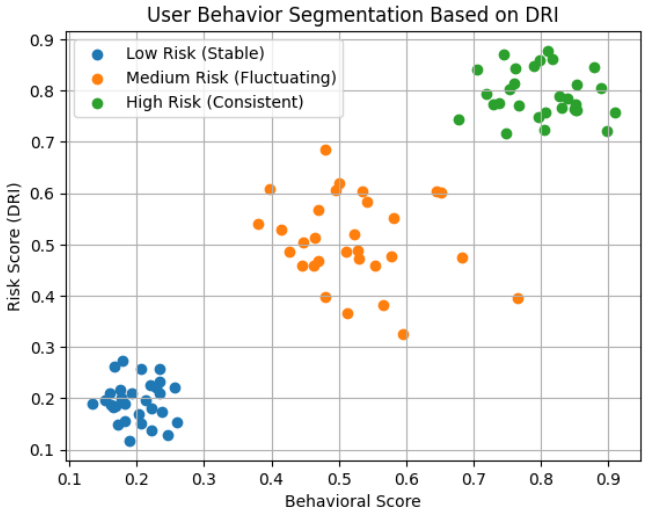

Fig. 4. User behavior segmentation based on DRI.

TABLE IX. COMPONENT CONTRIBUTION ANALYSIS

Component	Weight Contribution (%)
Behavioral Score	38.5
Network Score	41.2
Anomaly Score	20.3

The results in Table IX suggest that both behavioral information and network information are dominant sources for risk prediction, although anomaly detection offers further enhancement to risk prediction. The balanced contribution is a confirmation of the effectiveness of the hybrid approach in capturing the multi-dimensional threat characteristics.

8) Risk transition matrix: The movement of users between the various risk levels over time is analysed in the risk transition matrix.

From \ To	Low	Medium	High
Low	91%	8%	1%
Medium	12%	78%	10%
High	3%	15%	82%

Fig. 5. Risk transition matrix.

The findings in Fig. 5 indicate that 91 percent of the low-risk users stay at the same level, and 8 percent of the risk group move towards the medium-risk level. Out of the medium-risk users, 10 percent progress to the high-risk group and high risk users are so much persistent and were found to be 82 percent still at the same level. The results of these studies can lead to

the conclusion that once users have been put in the high-risk state, they are prone to remain in that state unless something is done about them. This discussion therefore shows the need to detect the issue early enough and ensure mitigation to prevent the escalation of the risk and ensure that the system remains secure.

G. Discussion

The proposed study puts forward an expansive framework for dynamic cyber risk prediction in the online education environment by amalgamation of behavioural analytics from LMS logs and network intrusion data from the CICIDS2017 dataset. The results demonstrate that the use of a multi-source data combination provides a much better detection capability than one dataset. Machine learning and deep learning models are good at capturing both static and temporal features, whereas the mechanism for anomaly detection gets better in the case of previously unseen threats. The introduction of the DRI provides a continuous and interpretable measure of user risk, which can be used to monitor user risk in real-time and mitigate it as proactively as possible. According to the comparative analysis, the proposed hybrid model is consistently superior compared to the state-of-the-art techniques in all the evaluation metrics. Ablation studies highlight the importance of individual components, especially feature fusion and behavioural modelling. Statistical significance analysis is another way of validating the robustness of the model. In addition, DRI-based analyses provide important information about user risk distribution, threshold optimization, and early detection of threats. Overall, the research holds the success in harnessing the intelligence at the behavioral and network level for the development of scalable, adaptive, and resilient cybersecurity solutions for modern E-learning platforms.

H. Limitations

The proposed framework also has some limitations even though it performs well. The LMS logs coupled with the CICIDS2017 data give rise to heterogeneity of the domain that could possibly affect generalization to various institutional settings. The use of historical and semi-labeled LMS data could be used as a factor of limitation in limiting the detection of zero-day attacks. Moreover, the deep learning components of these functions are a complication in the calculations and may pose a difficulty in having real-time systems that are constrained by the resources of the system. Moreover, variations in LMS platforms and patterns of user actions can require more customization of the platforms in order to guarantee optimal performance.

I. Practical Implications

The suggested framework possesses a high degree of practical advantages in enhancing cybersecurity in e-learning systems. Using the opportunity of the Dynamic Risk Index (DRI), the institution is able to track the activity of users in real time; the activity of high-risk members can be discovered beforehand. The incorporation with LMS platforms guarantees the facilitation of the ease of installation without disrupting an existing operation. Its model assisted in automated alert generation, making security administrators' work easier and enabling quicker reaction to an incident. As well, behavioral analytics will be useful in targeting cybersecurity awareness programs, as well as improving policies through the insights obtained using the methods of behavioral analytics.

V. Conclusion

This work proposes a novel hybrid framework for dynamic cyber risk prediction in online education environments using both LMS behavioral data and network intrusion information from the CICIDS2017 dataset. The proposed approach essentially combines the techniques of machine learning, deep learning, and anomaly detection to pick up on behavioral as well as network-level threat patterns. Experimental results show that the model has better performance as compared to the current methods with better accuracy, robustness and generalization. The integration of the Dynamic Risk Index (DRI) provides for constant surveillance of the risk and helps to prevent threats. Future research can focus on introducing federated learning to overcome the data privacy issues and facilitate joint learning between institutions. The combination of explainable AI techniques can help to increase transparency and trust in the model predictions. Additionally, expanding the framework for processing streaming data in real-time as well as the deployment of the framework at the edge can make the system more scalable and responsive. Exploring greater unsupervised and reinforcement learning techniques could also enhance zero-day attack detection. Overall, the proposed framework provides a firm basis to develop intelligent and adaptive cybersecurity solutions in a changing e-learning ecosystem.

References

[1] Sahni, Shalini, Sushma Verma, and Rahul Pratap Singh Kaurav. "Understanding digital transformation challenges for online learning and teaching in higher education institutions: a review and research framework." Benchmarking: An International Journal 32, no. 5 (2025): 1487-1521.

[2] Mhlanga, David. "Digital transformation of education, the limitations and prospects of introducing the fourth industrial revolution asynchronous online learning in emerging markets." Discover education 3, no. 1 (2024): 32.

[3] Shtayyat, Ahmad, and Mohammad AlShaikh-Hasan. "Enhancing Cybersecurity in E-Learning System Infrastructure: Analyzing Challenges and Implementing Solutions." In Complexities and Challenges for Securing Digital Assets and Infrastructure, pp. 157-174. IGI Global Scientific Publishing, 2025.

[4] Taha, Ibrahim Mohamed, Rajaa Hussein Abd Ali, and Ali Abdulhassan Abbas. "The Impact of Students' Cybersecurity Vulnerability Behavior on E-Learning Obstacles." Organizacija 58, no. 1 (2025): 85-104.

[5] Sedraoui, Brahim Khalil, Abdelmadjid Benmachiche, Amina Makhlouf, Djaber Abbas, and Makhlouf Derdour. "Cybersecurity in E-Learning: A Literature Review on Phishing Detection Using ML and DL Techniques." In 2025 International Conference on Networking and Advanced Systems (ICNAS), pp. 1-10. IEEE, 2025.

[6] Hurley, Richard, Philip Kruger, Henry Nascimento, and Stephen Keller. "Real-time ransomware detection through adaptive behavior fingerprinting for improved cybersecurity resilience and defense." Open Science Framework (2024).

[7] Aslan, Ömer, Semih Serkant Aktuğ, Merve Ozkan-Okay, Abdullah Asim Yilmaz, and Erdal Akin. "A comprehensive review of cyber security vulnerabilities, threats, attacks, and solutions." Electronics 12, no. 6 (2023): 1333.

[8] Hernández, R. M., Fuster-Guillen, D., Maluquish, L. G. P., Caycho, M. L. C., & Rivera, M. N. A. (2026). Adaptive Security Model for E-Learning Platforms Using Multi-Layered Anomaly Detection and

Context-Aware Risk Assessment. Journal of Internet Services and Information Security (JISIS), 16(1), 747-759.

[9] Mudawi, Naif Al, Mahwish Pervaiz, Bayan Ibrahimm Alabduallah, Abdulwahab Alazeb, Abdullah Alshahrani, Saud S. Alotaibi, and Ahmad Jalal. "Predictive analytics for Sustainable E-learning: Tracking student behaviors." Sustainability 15, no. 20 (2023): 14780.

[10] Yuan, Jin, Xuelan Qiu, Jinran Wu, Jiesi Guo, Weide Li, and You-Gan Wang. "Integrating behavior analysis with machine learning to predict online learning performance: A scientometric review and empirical study." arXiv preprint arXiv:2406.11847 (2024).

[11] Zine, Mohamed, Fouzi Harrou, Mohammed Terbeche, Mohammed Bellahcene, Abdelkader Dairi, and Ying Sun. "E-learning readiness assessment using machine learning methods." Sustainability 15, no. 11 (2023): 8924.

[12] Qiu, Feiyue, Lijia Zhu, Guodao Zhang, Xin Sheng, Mingtao Ye, Qifeng Xiang, and Ping-Kuo Chen. "E-learning performance prediction: Mining the feature space of effective learning behavior." Entropy 24, no. 5 (2022): 722.

[13] Oroni, Chrispus Zacharia, Fu Xianping, Daniela Daniel Ndunguru, and Arsenyan Ani. "Cyber safety in e-learning: The effects of cyber awareness and information security policies with moderating effects of gender and experience levels among e-learning students." Education and Information Technologies 30, no. 10 (2025): 14197-14236.

[14] Akacha, Souheil Abdel-Latif, and Ali Ismail Awad. "Enhancing security and sustainability of e-learning software systems: A comprehensive vulnerability analysis and recommendations for stakeholders." Sustainability 15, no. 19 (2023): 14132.

[15] Sadiqzade, Zarifa, and Hasan Alisoy. "Cybersecurity and online education–risks and solutions." Luminis Applied Science and Engineering 2, no. 1 (2025): 4-12.

[16] Eli, Aimina Ali, Abdur Rahman, and Naresh Kshetri. "D3S3real: Enhancing Student Success and Security Through Real-Time Data-Driven Decision Systems for Educational Intelligence." Digital 5, no. 3 (2025): 42.

[17] Mi, Chunqiao, Qingyou Deng, and Changhua Zhao. "Data-driven risk assessment and early warning of learning situations." In IET Conference Proceedings CP812, vol. 2022, no. 9, pp. 249-254. Stevenage, UK: The Institution of Engineering and Technology, 2022.

[18] Kepuska, Krenar, and Milo Tomasevic. "A lightweight framework for cyber risk management in Western Balkan higher education institutions." PeerJ Computer Science 10 (2024): e1958.

[19] Dube, Rohit. "Faulty use of the cic-ids 2017 dataset in information security research." Journal of Computer Virology and Hacking Techniques 20, no. 1 (2024): 203-211.

[20] Aljaloud, Abdulaziz Salamah, Diaa Mohammed Uliyan, Adel Alkhalil, Magdy Abd Elrhman, Azizah Fhad Mohammed Alogali, Yaser Mohammed Altameemi, Mohammed Altamimi, and Paul Kwan. "A deep learning model to predict student learning outcomes in LMS using CNN and LSTM." IEEE Access 10 (2022): 85255-85265.

[21] Ashraf, M. Wasim Abbas, Arvind R. Singh, A. Pandian, Rajkumar Singh Rathore, Mohit Bajaj, and Ievgen Zaitsev. "A hybrid approach using support vector machine rule-based system: detecting cyber threats in internet of things." Scientific Reports 14, no. 1 (2024): 27058.

[22] Jeamaon, Aomduan, and Chaiyaporn Khemapatapan. "Development cyber risk assessment for intrusion detection using enhanced random forest." ECTI Transactions on Computer and Information Technology (ECTI-CIT) 18, no. 4 (2024): 429-442.

[23] Nwafor, Chioma Ngozi, Obumneme Nwafor, Sanjukta Brahma, and Madhusudan Acharyya. "A hybrid FAIR and XGBoost framework for cyber-risk intelligence and expected loss prediction." Expert Systems with Applications (2025): 129920.

[24] Jarih[1], Ali Aqeel, and Zaid Ameen Abduljabbar. "Managing Risk in Academic Learning Management System: A Review." In Software Engineering: Emerging Trends and Practices in System Development: Proceedings of the 14th Computer Science On-line Conference 2025, Volume 6, vol. 6, p. 362. Springer Nature, 2025.

[25] Al-Sulami, Amnah, Miada Al-Masre, and Norah Al-Malki. "Predicting at-risk students' performance based on LMS activity using deep learning." International Journal of Advanced Computer Science and Applications 14, no. 6 (2023).

[26] Battaglin, Ricardo, Roberto Munoz, Vinicius Culmant Ramos, and Cristian Cechinel. "Predicting at-risk students with LMS data: A comparison between Adaboost and LSTM algorithms." In 2022 XVII Latin American Conference on Learning Technologies (LACLO), pp. 1-4. IEEE, 2022.

[27] Dalal, Surjeet, M. Poongodi, Umesh Kumar Lilhore, Fadl Dahan, Thavavel Vaiyapuri, Ismail Keshta, Sultan Mesfer Aldossary, Amena Mahmoud, and Sarita Simaiya. "Optimized LightGBM model for security and privacy issues in cyber-physical systems." Transactions on Emerging Telecommunications Technologies 34, no. 6 (2023): e4771.

[28] Sartana, Bruri Trya, Supeno Mardi Susiki Nugroho, Umi Laili Yuhana, and Mauridhi Hery Purnomo. "Predicting At-Risk Students in Online Learning: Integrating Demographic, Assessment Trends, Learning Engagement, and Behavioral Patterns Using CatBoost." International Journal of Intelligent Engineering & Systems 18, no. 8 (2025).

[29] Sudha, Chinnakka, and Sreenivasulu Bolla. "A Novel Deep Learning Technique for Big Data Anomaly Threat Severity Prediction in ELearning." Statistics, Optimization & Information Computing 15, no. 3 (2026): 1913-1935.

[30] Panda, Rajesh, and Prashant Kumar Tiwari. "Cross-entropy based risk assessment and forecasting of secured hybrid system using Bi-LSTM based am deep learning model." IEEE Transactions on Industry Applications 61, no. 3 (2025): 4661-4674.

[31] Chhetri, Bipin, and Akbar Siami Namin. "The Application of Transformer-Based Models for Predicting Consequences of Cyber Attacks." In 2025 IEEE 49th Annual Computers, Software, and Applications Conference (COMPSAC), pp. 523-532. IEEE, 2025.

[32] Ayata, Faruk. "A High-Precession Cybersecurity Model with Stacked Ensemble Learning." Savunma Bilimleri Dergisi Advanced Online Publication (2026).