International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(7s) 2026 pp. 201-215
https://doi.org/10.70917/ijcisim-2026-3067
Article

An LTC-LSTM Framework for Early Prediction of Course Completion Time and Student Dropout Risk

Devi S[1*], K. Maharajan[2], Jayalakshmi Murugan[3]

[1,2]Department of Computer Science and Engineering, Kalasalingam Academy of Research & Education, Tamil Nadu, India
[3]Department of Computer Science and Engineering-Cyber Security, Ramco Institute of Technology, Tamil Nadu, India
E-mail: [1]devinaveen74@gmail.com ,[2]maharajank@gmail.com , [3]jayalacsmi@gmail.com

Abstract: Ensuring student success in online and blended learning is a primary issue in education. One of the two most important are predicting the length of time that students will take to finish a course and identifying those that may be at risk of not finishing their study. These are especially challenging tasks. The reasons for this are complex, since the behaviour of students is determined by a variety of factors, including financial situation, Academic performance and engagement levels, mental well-being. Furthermore, these factors can vary considerably Predicting is a difficult problem because it is often changing, from one learner to another. Such non-linear and dynamic relationships are often not represented well by traditional predictive models; these are making it difficult for them to be effective in the real-world classroom. To tackle these issues this study, A Liquid Time Constant (LTC) network that is a unified Long Short-Term Memory (LSTM) network is introduced. Model for estimating course completion time and predicting dropping out. Unlike conventional models, the proposed approach introduces an adaptive liquid time constant into the LSTM framework to achieve an adaptive time constant. The network is to simulate continuous time dynamics and accommodate statistically irregular time periods between student-speech by using the network learning activities. The model can include the long-term temporal dependencies and evolving patterns in this design. More accurately describe behavioural patterns. In the data preprocessing step categorical attributes are converted into Label encoding is used for numerical representations, and min–max for numerical features normalization. The following are done to ensure consistent representation of features, stability of the model, and support: Convergence speed during training. The performance of the proposed LTC-LSTM model was tested on the three datasets that we use for educational benchmarking are KDD Cup 2015, Anon, and Dropout. Experimental results demonstrate the model attained accuracies of 92.10%, 84.65% and 94.48% respectively. In addition, LTC-LSTM achieved lower prediction errors than the others and needed less training to achieve higher predictive performance compared to the common approaches such as the Image Convolutional and Bi-directional Temporal methods, this one appears to be more efficient Convolutional Network (IC-BTCN). Conclusively, the results demonstrate the ability of LTC-LSTM to be a strong mathematically sound and efficient modelling approach for students' learning trajectories. By providing accurate, the proposed framework can help institutions with predictions of course completion and dropout risk. Effectively and timely putting into practice interventions that enhance student engagement, retention, and academic achievement based on the data.

Keywords: Student dropout prediction; course completion time; Liquid Time Constant; LSTM; temporal modelling; educational data mining; learning analytics; student retention

1. INTRODUCTION

In recent years, a growing number of students worldwide have turned to Massive Open Online Courses (MOOC) as learning platforms to bridge knowledge gaps or obtain new skills demanded by industry. These platforms support learner autonomy and offer a certain degree of personalization in the learning process. MOOCs are effective at managing learning activities often lack in supporting social interactions and collaborative behaviors of learners [18]. However, certain MOOCs platforms have set up Questions & Answers (Q&A), corresponding forums, and so on to assist learners in building appropriate interactive behaviours that provide knowledge exchange and problem feedback among learners or instructors [7] [9]. Innovative and effective higher education serves as the foundation of a successful and thriving technology as well as knowledge-based economy where production and services rely on knowledge-intensive and intellectual activities [28]. Time to course completion refers to the duration a student takes to complete an enrolled course. It depends on a variety of factors such as learning performance, studying skills, and interaction with learning materials.

This unique identification of this timeframe can enable institutions to provide timely interventions to support students to stay on track. Monitoring and tracking time to course completion will help to promote better academic scheduling, increase student retention and decrease dropouts. In Higher Education, the issue of student dropout and student delayed completion is an urgent problem that has a personal cost and a social cost, especially in Science, Technology, Engineering and Mathematics (STEM) programmes [26] [16]. Many studies have been conducted on the reasons for dropping out and for academic failure, which range from financial difficulties to family/ personal troubles to academic difficulty. All of these outcomes have a negative impact on a student's personal life, creating frustration, feelings of failure, self-doubt, anxiety and a limited choice of work opportunities and future earnings [20].High dropout rates also contribute to a less-educated workforce and lower aggregate economic productivity, with consequences for society as a whole [6]. The present work is motivated by the goal of reducing and anticipating student dropout in in-person learning environments and thereby limiting the associated financial losses [27] [25].

However, these developments have led to several unintended consequences:1) policymakers are forcing higher education institutions to rapidly supply labor market with skilled workers while also seeking for optimizing university funding; 2) universities are adjusting strategies for aligning with socio-political shifts like expansion of higher education and growing demand for qualified workforce, sometimes the requirements is lowered to acquire diploma; 3) as more students enroll in student education, a broader range of challenges and socio-political issue begin to emerge [22] [3].

Consequently, academic success and university dropout have evolved into complex and multi-dimensional issues influenced by various personal, institutional, and societal factors. Both national and international studies have distant outcomes in background variables of academic success and dropout because of constantly changing variables and background features [12]. The strategies for reducing dropouts are to identify early students at risk and implement effective retention interventions before they show symptoms of being at risk or disengaged [10].

Recently, Artificial Intelligence (AI) including Machine Learning (ML) have obtained promising results in predictive tasks over different fields. Many approaches have been proposed to predict a student's academic performance by using ML [11]. But, there are some drawbacks of using ML in education data management. Handcrafted features and inability to deal with high-dimensional, sequential, unstructured data are the issues traditional ML has been facing when addressing complex patterns or temporal dependencies within student behavior and performance records [2] [14]. To address this issue, Deep Learning (DL) emerged as an alternative method that automatically learn hierarchical feature representations from data without manual intervention [8]. DL are well-appropriate for analyzing non-linear, large-scale, and multi-modal data, which makes more accurate and robust predictions [13] [17]. Thus, the deployment of DL significantly enhances the predictive performance of academic performance.

Xinhong Zhang et al. [24] introduced an Image Convolutional and Bi-directional Temporal Convolutional Network (IC-BTCN) for predicting the student dropout for learners from MOOC courses. Initially, learning records of student data were transformed into 3-D learning behavior matrix. Later, the local behaviour feature matrix was extracted using convolutional approaches. Multi-Layer Perceptron (MLP) was employed to predict student dropout. However, the IC-BTCN model struggled with capturing long-term dependencies across widely spaced course activities because of limited temporal receptive fields.

Fatemeh Khoushehgir and Sadegh Sulaimany et al. [4] developed a bipartite link prediction method for determining students' dropout in the MOOC platform to identify the most probable links. Initially, the data was converted into a graph by mapping entities (courses and students) to relationships and nodes (enrollment data) to links. Later, graph-based method was utilized for predicting student dropout from enrollment data. Nevertheless, the bipartite link oversimplifies complex learners' behaviours by modelling only interactions among students and course elements. This lacks temporal dynamics and is less effective in capturing evolving engagement patterns over time.

Cong Jin [1] established an optimized Support Vector Regression (SVR) to predict students' dropout effectively. An Improved Quantum Particle Swarm Optimization (IQPSO) was used for optimizing three parameters of SVR for predicting students' dropout status each week. Based on the SVR method, prediction outcomes of SDP were refined to once a week, which enables prediction results to be easier and more detailed to identify changes in dropout trends.

However, SVR struggled to effectively capture the non-linear and dynamic nature of student engagement in MOOCs with sparse data that fail to accurately identify at-risk students.

Kowsar Talebi et al. [23] suggested an ensemble model depending on Convolutional Neural Network (CNN) and LSTM to predict dropout in MOOC. CNN was utilized to automatically extract the features, whereas LSTM considers time series data aspects to enhance early prediction. As an ensemble classifier, minimize variance of prediction error by utilizing an ensemble classifier in a neural network, increase the F1-score and accuracy without overfitting. This increases model accuracy and generates the possibility of early dropout prediction. Nevertheless, ensemble methods suffer from high computational complexity and training time which makes them less practical for large-scale MOOC platforms.

Ke Niu [21] et al. presented a Parallel Multiscale Convolutional Temporal (PMCT) to predict student dropout behaviour. PMCT was used for extracting multiscale behavioral features depending on various features. Stacked residual block of TCN was used for extracting correlation among scales from behavioral features. However, PMCT had difficulty in capturing long-range dependencies due to multi-scale convolutions focusing on localized patterns, which results in less accurate predictions. From the overall evaluation, existing methods had limitations like struggling with capturing long-term dependencies, lacking temporal dynamics which makes less effective in capturing evolving engagement patterns, suffering from high computational complexity and training time, and being affecting by diverse, interrelated factors like financial instability, mental health, and academic performance that vary across individuals leads to inaccurate identification and prediction.

To address the above issue, LTC-LSTM is proposed to predict the time to course completion and students' dropout effectively. This allows the model to adapt in irregular time intervals in student activity data by capturing long-term dependencies more effectively. This improves model's ability to identify changes in student engagement over time. Hence, LTC-LSTM performs more accurate, robust identification and prediction across diverse learning contexts.

Time to course completion and dropout prediction assist institutions in enhancing student retention by determining patterns associated with disengagement and academic struggles. Predicting time to course completion and student dropout in higher education, particularly in MOOC has become a significant issue worldwide, driven by complex and interrelated factors like personal, financial instability, academic challenges, and mental health concerns. These variables are often hidden, nonlinear, and evolve which makes it difficult to predict as discontinued or dropped.

One of the primary aims is to build a strong and accurate identification of time to course completion and student drop-out prediction model based on LTC-LSTM method. LTC-LSTM incorporates continuous time dynamics into LSTM model, making it adapt to arbitrary time interval between students' activities and capable of effectively modeling temporal dependencies. This allows for capturing of complex and variable sequences, which will improve predictive performance. Thus, LTC-LSTM not only makes the model have better performance on the long-term dependencies modelling of student behaviors but also has a great improvement on overall performance.

The primary contribution of this research is as follows:

- In traditional LSTM, LTC is incorporated, which allows the model to manage irregular time intervals in student activity data and capture intricate temporal dependencies that accurately predict time to course completion and dropout performance.
- In pre-processing, categorical variables are converted into a numeric form that enables to process of different input type effectively, which contribute to better predictive performance.
- Min-max normalization scales numerical features into a uniform range that enhances model's training stability and convergence.

2. PROPOSED METHODOLOGY

This research proposes LTC-LSTM to predict time to course completion and student dropout effectively. Initially, the data is obtained from KDD Cup 2015, dropout, and Anon datasets to evaluate model performance. Then, min-max and label encoding are used during the pre-processing stage to ensure all numerical features contribute equally and convert categorical variables into a numerical format that allows models to interpret and process the features effectively. At last, the proposed LTC-LSTM is used for student learning dropout prediction and time-to-course completion. Figure 1 depicts the workflow of the overall methodology process.

Problem formulation: Let a learner be represented by a time-ordered sequence of activity records $X = \{x_1, x_2, \ldots, x_t\}$, where each $x_t \in \mathbb{R}^d$ encodes the behavioural and demographic features observed at step t, and let Δt denote the generally non-uniform interval between consecutive records. Two related tasks are addressed. For dropout prediction, the objective is to learn a mapping $f\theta: X \rightarrow \hat{y} \in \{0, 1\}$, where $\hat{y} = 1$ indicates that the learner discontinues the course. For time-to-course-completion, the objective is to estimate a continuous progression target (cumulative credit hours) $\hat{y} \in \mathbb{R}$. Both tasks are cast as sequence-modelling problems in which the parameters θ are optimised to minimise a task-specific loss over the irregularly sampled sequence. The central difficulty addressed by the proposed model is that Δt varies across learners and over time, so a model with a fixed temporal resolution cannot represent the underlying engagement dynamics faithfully.

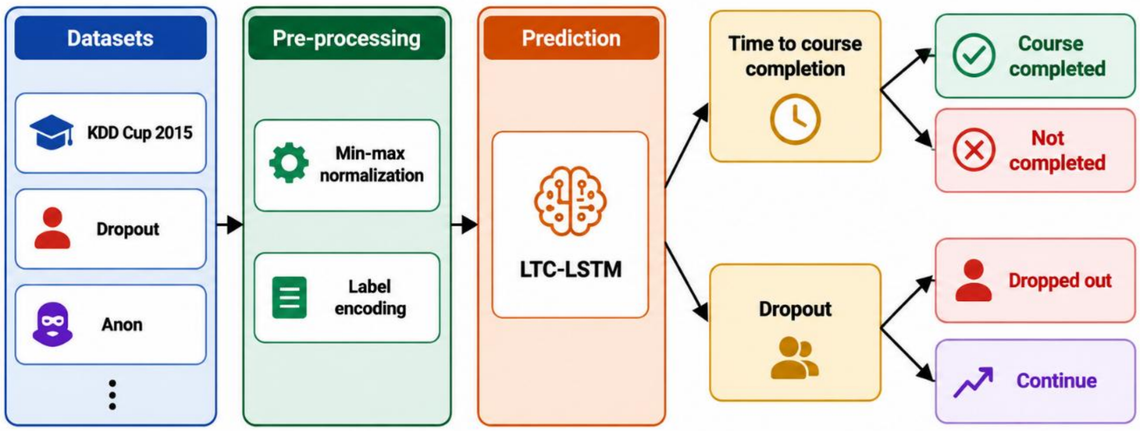

Figure 1: Workflow Of The Overall Methodology Process

2.1 Datasets

In this research, KDD Cup 2015 [29], dropout [30], and Anon [31] datasets are used to determine the effectiveness of model performance in predicting time to course completion and student dropout. These datasets encompass diverse educational environments that enable comprehensive performance across different learning contexts. A detailed description of this dataset is explained below.

KDD Cup 2015: This dataset originates from XuetangX, one of a largest MOOC platform in China and captures learning interactions from 2013 to 2014. KDD Cup 2015 comprises 120,542

records reflecting the activities of 79,186 students across 39 courses with a duration of 5 weeks. The dataset documents have seven distinct types of student learning behaviour on the platform.

Dropout: It involves user interaction logs to capture activities like session ID, timestamps, and actions for each (course, user) pair. Each entity is determined by linking users and courses to specific behaviour over time. The corresponding files provide dropout labels representing whether a user eventually dropped out (1) or continued (0), respectively.

Anon: The data comes from registrar information from a huge American research University. This institution is a highly ranked, midwestern, and large-enrollment University with nearly 50,000 students enrolled annually and 78% of graduation rate. This dataset contains timestamped course grades, degrees awarded, majors declared, and preparation information for 160,933 students from 1992 to 2012. The obtained data are passed through the pre-processing stage for further processing.

2.2 Pre-Processing

After obtaining data, label encoding [15] and min-max [5] are used, which ensures all numerical features contribute equally by scaling to a common range and enhance model training. Label encoding transforms categorical data into numeric form which enables model to process student attributes effectively for accurate dropout prediction. A detailed process for these two methods is explained below in detail.

Label encoding: It converts categorical variables like course type, gender, or education level into a numeric format that is appropriate for DL methods. Label encoder transforms all given nominal values into numeric positive integers from zero to feature's unique value range. One-hot encoding can also be used, but it creates multiple binary columns, whereas label encoding uses a single column which makes it more memory efficient. Hence, label encoding is used and this transformation allows models to interpret and process the features effectively which enhances model performance.

Min-max: Datasets have features with different ranges; if not normalized, features with larger scales dominate the learning process. Hence, min-max is used to scale student data into a uniform range in [0,1] that assists DL models to converge faster and perform better. Moreover, it preserves relationships between original data values and makes patterns related to dropout risk more detectable. Min-max employs minimum and maximum values in given features to evaluate the standard value in given features using (1),

$$x_{std} = \frac{x - x_{min}}{x_{max} - x_{min}} \qquad (1)$$

where min and max indicate minimum and maximum values of feature range, x represents original input value, $xstd$ denotes normalized value in (0,1) range respectively. This normalization minimizes influence of outliers, which ensures consistent feature contribution and enhances prediction accuracy. Then, the pre- processed data are passed as input into the prediction process using LTC-LSTM method.

3. PREDICTION USING LTC-LSTM

After pre-processing, LTC-LSTM is utilized to predict time to course completion and students' learning dropout effectively. LSTM capture temporal dependencies in sequential data like student performance over time. Unlike Recurrent Neural Network (RNN) and Gated Recurrent Unit (GRU), LSTM [19] remembers important patterns and behaviours from earlier time steps that assist in identifying gradual changes in dropout risk. Moreover, LSTM has the ability to manage time-series data with varying lengths, which makes it appropriate for modelling learning activities, engagement levels, or attendance across semesters. This results in more accurate, timely dropout predictions and enables the identification of time to course completion for proactive interventions.

In traditional LSTM, LTC enhance the ability to manage irregular and continuous time intervals in student learning data. Unlike LSTM assumes uniform time steps, LTC-LSTM dynamically adjusts to varying student activities across different periods. This assists the model to better capture the influence of time gaps on dropout risk and learning behaviors. This enhances prediction accuracy by aligning more closely with the temporal nature of educational data.

3.1 *Liquid Time Constant-Long Short-Term Memory (LTC-LSTM)*

In traditional LSTM, LTC is incorporated for establishing hidden time states and time-constant of learning system. This network provides a prediction model with a sharper learning curve for a long period. Another key benefits of employing LTC is processing differential graphs and obtain trained by different range of gradient-based methods. Figure 2 shows the structure of LTC-LSTM, comprising (a) the cell with its liquid time-constant mechanism and (b) the stacked network.

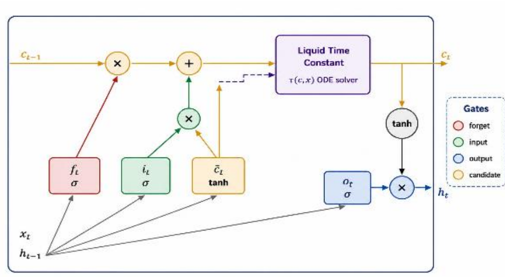

Figure 2(a): Structure Of LTC-LSTM cell with its liquid time-constant mechanism

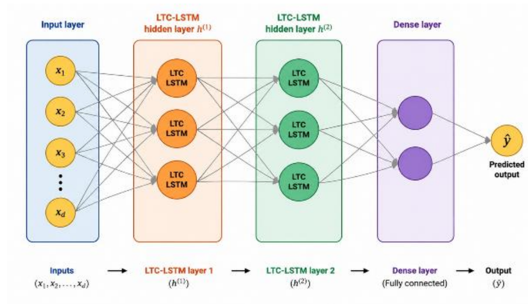

Figure 2(b): LTC-LSTM- Stacked Network

A dynamic behaviour of the LTC network is mathematically described by (2) and (3)

$$\frac{d}{dt}V_i = -\frac{1}{\tau_m}V_i + I_i(t) + \sum_j W_{ij}s_j(t) \qquad (2)$$

$$\frac{d}{dt}s_i = -\frac{1}{\tau_s}s_i + \sum_j Q_{ij}s_j(t)f\left(V_j(t-\Delta)\right) \qquad (3)$$

Where V_i indicates membrane neuron potential i, $I_i(t)$ denotes input current, $W_{ij}(t)$ represents synaptic weight among neurons i and j, τ_m refers time constant, $s_i(t)$ determines synaptic gating variable, τ_s illustrates synaptic time constants, $f(.)$ indicates firing rate function, and Δ refers axonal delay respectively. These equations determine behavior of synaptic gating and potential variables over time based on input current and synaptic connections. A synaptic weight W_{ij} (t) changes over time because of time-constant τ_w. LSTM has an ability for learning long-term dependencies and maintain longer time gaps in noisy input string. LSTM develops an effective model consist of gradients which avoids vanishing issue and forcing error backflow via internal interacting layers that has special functions for filtering unwanted information by integration of tanh and sigmoid cells.

RNN has a single unit and LSTM cells by containing 4 units. A cell state represents long-term memory of the network that allows it to store and retrieve information from events. The following are the inputs which are fed as a vector form: previous hidden state h_{t-1}, present input x_t, and previous cell state c_{t-1}. Forget gate is an initial block to identify which part of information is discarded or retained. The inputs are fed via sigmoid σ_g which obtain output values among zero and one. An activation vector of forget gate f_t is expressed in (4).

$$f_t = \sigma_g\left(W_f x_t + U_f h_{t-1} + b_f\right) \qquad (4)$$

An input gate updates cell status with two parts; initially, prior hidden state and present input are passed into second sigmoid function σg. Then, same inputs are fed into tan function σc for regulating network. The cell state vector ct in (7) indicates results of the candidate cell state (input cell activation vector) $\tilde{c}_t$as calculated in (5) and update gate activation vector it. An input gate formula is represented in (6).

$$\tilde{c}_t = \sigma_c(W_c x_t + U_c h_{t-1} + b_c) \qquad (5)$$

$$i_t = \sigma_g(W_i x_t + U_i h_{t-1} + b_i) \qquad (6)$$

$$c_t = f_t . c_{t-1} + i_t . \tilde{c}_t \qquad (7)$$

At last, output gate evaluates hidden state ht using (8) and (9) and hidden state involves information on prior inputs. A prior hidden state $ht-1$ and present input xt are fed into 3rd sigmoid σg and then modified cell state is fed into tangent σh. Those outputs are multiplied element-wise and enable network to identify hidden state carry. A parameter W ϵ $R^{h\times d}$, $U \epsilon R^{h\times h}$, and $b \epsilon R^{h}$indicates weights and bias vectors which are learned during training process.

$$o_t = \sigma_g(W_o x_t + U_o h_{t-1} + b_o) \qquad (8)$$

$$h_t = o_t . \sigma_h(c_t) \qquad (9)$$

Equations (4)–(9) describe the discrete LSTM gating mechanism, in which the cell state is updated under an implicit unit time step. The liquid time-constant formulation in (2)–(3) replaces this fixed step with a state- and input-dependent time constant, so that the continuous-time evolution of the cell state is governed by (10).

$$\frac{dc(t)}{dt} = -\left(\frac{1}{\tau} + f_t\right)c(t) + f_t \odot (i_t \odot \tilde{c}_t) \qquad (10)$$

Here τ is a trainable base time constant and the forget activation f_t acts as an input-dependent rate that contracts or dilates the effective system time constant τ_sys = τ / (1 + τf_t). Because f_t and the candidate state depend on the current input x_t and the previous hidden state $h_{(t-1)}$, the network continuously adapts its memory horizon to the irregular spacing of student activities. Equation (10) admits no closed-form solution for arbitrary inputs; it is therefore integrated using a fused explicit–implicit solver over an adaptive step Δt set to the observed inter-event interval, which yields the stable discrete update in (11)

$$c_t = \frac{c_{t-1} + \Delta t[f_t \odot (i_t \odot \tilde{c}_t)]}{1 + \Delta t\left(\frac{1}{T} + f_t\right)} \qquad (11)$$

The output gate and hidden state are then computed as in (8)–(9). When Δt = 1 and τ → ∞, the update in (11) reduces exactly to the conventional LSTM cell-state update, which confirms that LTC-LSTM is a strict generalisation of the LSTM in which the time constant is learned and modulated by the input rather than held fixed. This adaptive time constant is the source of the model's improved sensitivity to the non-uniform temporal structure of learner activity logs.

The computational cost of LTC-LSTM is governed by the gate projections and the fused-solver iterations. For an input dimension d, hidden size H, sequence length T, and K solver steps per time step, the per-layer time complexity is

$$O\big(TK(4Hd + 4H^2)\big) \qquad (12)$$

That is, a factor of K above a standard LSTM, while the parameter and memory footprint remain of order H^2 + Hd because the solver reuses the existing gate weights. In practice K is small (typically four to six), so the overhead is modest and is more than offset by faster convergence, which is consistent with the reduced training times reported in Table 5.Thus, LTC-LSTM enable the model to capture continuous time dynamics more effectively. It allows network to adaptively adjust time constant and enhance temporal sensitivity that leads to better modelling of irregular and time-varying dropout patterns.

4. RESULTS AND DISCUSSION

The proposed LTC-LSTM is simulated using Python 3.4 environment with windows 10 operating system, an i5 Intel processor, and 64 GB of Random Access Memory (RAM). Accuracy, F1-score, Recall, Precision, Root Mean Squared Error (RMSE), and Mean Square Error (MSE) are performance metrices used to determine model performance. The hyperparameters used for LTC-LSTM are a learning rate of 0.001, a dropout rate of 0.3, the tanh activation function, and the Adam optimizer. The mathematical expressions of these metrices are provided in (13) to (18) as:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \qquad (13)$$

$$F1 - Score = \frac{2TP}{2TP + FP + FN} \quad (14)$$

$$Recall = \frac{TP}{TP + FN} \quad (15)$$

$$Precision = \frac{TP}{TP + FP} \quad (16)$$

$$RMSE = \sum_{i=1}^{n} \sqrt{\frac{1}{n}|y_i - \hat{y}_i|^2} \quad (17)$$

$$MSE = \frac{1}{n}\sum(y_i - \hat{y}_i)^2 \quad (18)$$

Where FP denotes False Positive, TN denotes True Negative, FN represents False Negative, TP illustrates True Positive. The yi indicates the actual (ground-truth) value, n illustrates number of tested samples, and $\hat{y}_i$represents the value predicted by the model, and the metrics are reported per dataset in the following section.

4.1 Performance Analysis

Table 1 shows a performance analysis of different classifiers on three datasets. Compared to existing methods like LTC-RNN, LSTM, and LTC-Bidirectional LSTM (BiLSTM), the proposed LTC-LSTM achieves high accuracy of 92.10%, 84.65%, and 94.48% on KDD Cup 2015, Anon, and dropout datasets by effectively capturing complex temporal dependencies in student activity patterns. Unlike existing methods that have fixed time steps, LTC enable dynamic adjustment of time constant, which makes better management of asynchronous and irregular learning behaviours. This flexibility enhances the model's sensitivity to time-varying student engagement. This results in more precise dropout predictions that lead to high accuracy. The enhanced temporal modelling minimizes prediction errors which contributes to lower MSE.

Table 1: Performance Analysis Of Different Classifiers

Classifiers	Datasets	Accuracy (%)	Precision (%)	Recall (%)	F1- Score (%)	RMSE	MSE
LTC-RNN	KDD Cup 2015	86.25	91.40	91.00	91.20	0.95	0.072
	Anon	80.20	83.50	83.00	83.25	0.89	0.066
	Dropout	90.50	92.30	90.00	91.14	0.30	0.058
LSTM	KDD Cup 2015	88.50	93.10	92.75	92.92	0.80	0.064
	Anon	82.75	85.25	85.60	85.42	0.80	0.06
	Dropout	92.15	94.25	91.80	93.00	0.27	0.055
LTC-Bi LSTM	KDD Cup 2015	90.25	95.00	94.80	94.90	0.65	0.056
	Anon	83.95	86.00	86.50	86.25	0.76	0.057
	Dropout	93.35	95.25	92.20	93.70	0.25	0.054

LTC-LSTM	KDD Cup 2015	92.10	96.88	96.85	96.86	0.55	0.052
	Anon	84.65	86.59	86.92	86.69	0.73	0.054
	Dropout	94.48	96.32	93.15	94.71	0.22	0.052

Tables 2 to 4 represent the performance analysis of different intervals for the proposed LTC-LSTM. The proposed LTC-LSTM achieves a high accuracy of 92.10%, 84.65%, and 94.48% on KDD Cup 2015, Anon, and dropout datasets due to its ability in modeling continuous-time dynamics, which aligns effectively with the non-uniform nature of student interaction data. Moreover, this method captures subtle changes in learning behaviour than traditional LSTM. The liquid state nature enables richer internal dynamics that assist in enhancing prediction performance. This leads to more informed predictions and better predictions of at-risk students.

Table 2: Performance Analysis Of Different Intervals On KDD Cup 2015 Dataset

Time	Accuracy (%)	Precision (%)	Recall (%)	F1- Score (%)	RMSE	MSE
Week 1	85.75	90.80	90.50	90.65	0.062	0.062
Week 1-2	87.60	92.40	92.00	92.20	0.076	0.068
Week 1-3	89.45	94.10	93.85	93.97	0.068	0.060
Week 1-4	90.80	95.50	95.30	95.40	0.065	0.058
Week 1-5	92.10	96.88	96.85	96.86	0.55	0.052

Table 3: Performance Analysis Of Different Intervals On Dropout Dataset

Time	Accuracy%	Precision %	Recall%	F1- Score %	RMSE	MSE
Week 1	88.30	91.25	88.1	89.65	0.35	0.065
Week 1-2	90.75	93.1	89.9	91.47	0.3	0.062
Week 1-3	92.00	94.55	91.6	93.05	0.27	0.056
Week 1-4	93.25	95.4	92.4	93.85	0.24	0.054
Week 1-5	94.48	96.32	93.15	94.71	0.22	0.052

Table 4: Performance Analysis Of Different Intervals On Anon Dataset

Time	Accuracy%	Precision %	Recall%	F1- Score %	RMSE	MSE
Week 1	78.10	80.5	81.00	80.75	0.88	0.068
Week 1-2	80.25	82.25	82.90	82.57	0.83	0.063
Week 1-3	82.00	84	84.75	84.37	0.78	0.059
Week 1-4	83.50	85.25	86.10	85.67	0.75	0.056
Week 1-5	84.65	86.59	86.92	86.69	0.73	0.054

Table 5 indicates performance evaluation of time complexity and statistical analysis. Compared to existing methods like LTC-RNN, LSTM, and LTC-BiLSTM, proposed LTC-LSTM achieves a shorter training time of 783s, 668s, and 889s due to an adaptive time constant mechanism that makes the model temporal dynamics more effective. Unlike the LSTM unit is based on fixed time steps whereas LTC-LSTM adjust the time-scale dynamically that minimize unnecessary updates. This results in rapid convergence and a more compact representation of time-series data. Moreover, this method simplifies the network's structure by decreasing redundant temporal states. Therefore, the LTC-LSTM minimizes overall computational burden during training while maintaining prediction accuracy.

Table 5: Performance Evaluation Of Time Complexity And Statistical Analysis

Methods	Datasets	ANOVA p-value	Execution Time (s)	Memory Consumption (MB)	Inference Time (ms)	Training Time (s)
LTC-RNN	KDD Cup15	0.034	1386	1480	314	1219
	Anon	0.041	1123	1290	284	1037
	Dropout	0.029	1528	1570	348	1322
LSTM	KDD Cup15	0.027	1358	1020	187	1086
	Anon	0.031	1197	970	169	972
	Dropout	0.022	1482	1145	204	1224
LTC-Bi LSTM	KDD Cup15	0.019	1443	1350	263	1332
	Anon	0.023	1314	1275	241	1187
	Dropout	0.017	1578	1438	278	1412
LTC - LSTM	KDD Cup15	0.012	1205	850	124	783
	Anon	0.018	1086	795	109	668
	Dropout	0.01	1394	902	137	889

Performance analysis of Accuracy vs Epoch over 50 epochs for proposed LTC-LSTM is shown in Figure 3 (Dropout Dataset), Figure 4 (Anon Dataset) and Figure 5 (KDD Cup 2015 Dataset). In the dropout dataset, training accuracy consistently covers around 0.95 while validation accuracy stabilizes around 0.93 which provides strong generalization with less overfitting. In the Anon dataset, both training and validation accuracy rapidly increase by about 0.83 within the initial few epochs and remain stable, which represents effective learning and convergence. Minor fluctuations in the Anon training curve indicate sensitivity to data variability, but overall consistency with validation denotes robustness.

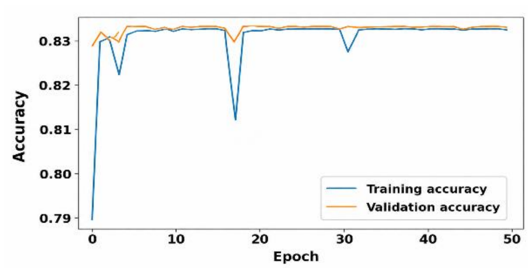

Figure 3: Performance Analysis Of Accuracy Vs Epoch For Dropout Dataset

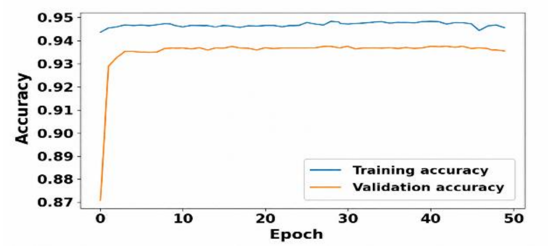

Figure 4: Performance Analysis Of Accuracy Vs Epoch For Anon Dataset

The graph from KDD Cup 2015 shows consistent enhancement in both training and validation, which provides better generalization and effective model learning.

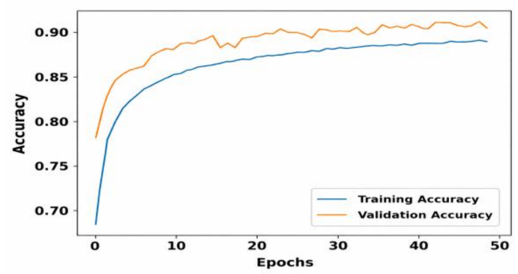

Figure 5: Performance Analysis Of Accuracy Vs Epoch For Kdd Cup 2015 Dataset

These results demonstrate that model's reliable performance across diverse datasets. Performance analysis of loss vs epochs for the proposed LTC-LSTM is shown in Figure 6 (Dropout Dataset), Figure 7 (Anon Dataset) and Figure 8 (KDD Cup 2015 Dataset). In dropout dataset, training loss remains consistently low over 0.04 whereas validation loss slightly stabilizes near 0.05 which represents effective learning with less overfitting. For Anon dataset, both validation and training loss rapidly converge around 0.42, which shows strong model generalization. Spikes in training loss, particularly in Anon dataset, model represent sensitivity to certain data breaches but do not significantly affect the overall trend. In KDD Cup 2015, the graph represents lower and consistently decreasing validation loss, suggesting the model is learning efficiently without overfitting. These confirm the stability and robustness of the proposed LTC-LSTM model across three datasets.

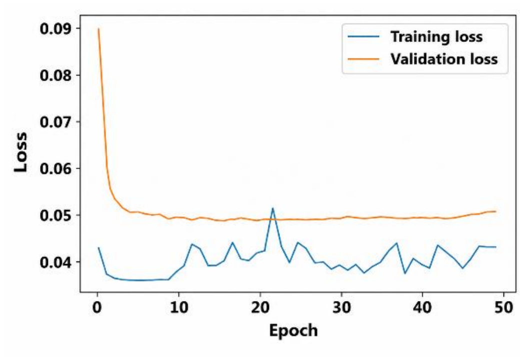

Figure 6: Performance Analysis Of Loss Vs Epoch For Dropout Dataset

Performance analysis of Receiver Operating Curve (RoC) for the proposed LTC- LSTM is shown in Figure 9 (Dropout Dataset), Figure 10 (Anon Dataset) and Figure 11 (KDD Cup 2015 Dataset). This graph demonstrates the model's robust prediction ability with an AUC of above 0.88.

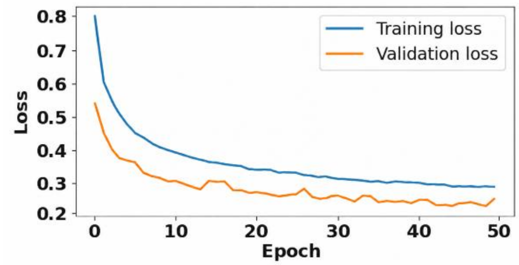

Figure 7: Performance Analysis Of Loss Vs Epoch For Anon Dataset

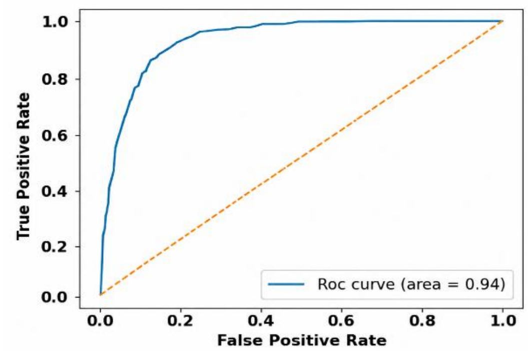

Figure 8: Performance Analysis Of Loss Vs Epoch For Kdd Cup 2015 Dataset

The curve represents a rise towards top-left corner that shows a high True Positive Rate (TPR) even at a low False Positive Rate (FPR). The KDD Cup 2015 dataset shows high TPR across a range of FPR with 0.92 AUC which provides strong performance in distinguishing between classes. This provides LTC-LSTM model is highly effective at correctly predicting student dropout while minimizing false alarms. A significant RoC curve validates the robustness and predictive strength of the model.

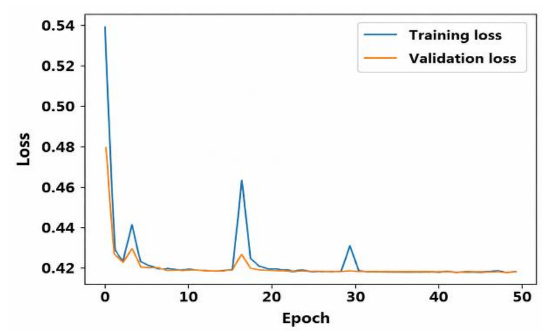

Figure 9: Performance Analysis Of TPR vs FPR For Dropout Dataset

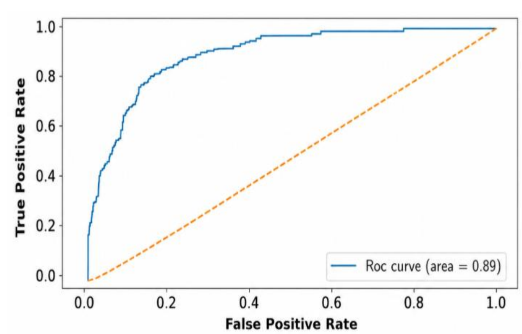

Figure 10: Performance Analysis Of TPR vs FPR For Anon Dataset

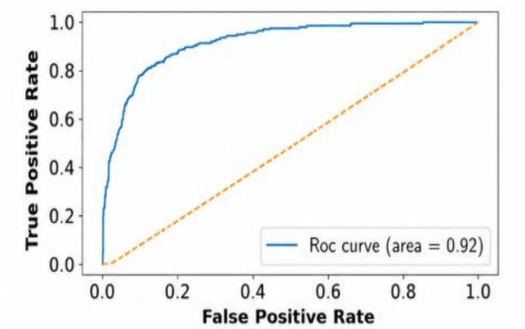

Figure 11: Performance Analysis Of TPR vs FPR For Kdd Cup Dataset

Figure 12 represents an analysis of actual vs predicted cumulative credit hours of 50 students, which is interpreted as an indicator of student's progression toward course completion.

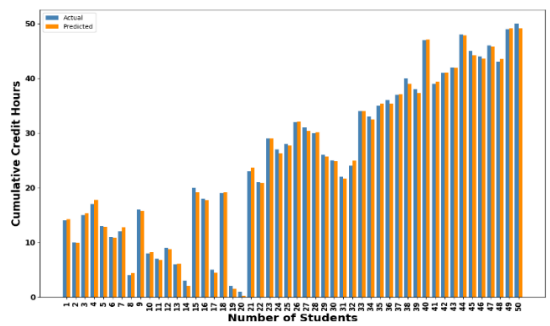

Figure 12: Analysis Of Actual Vs Predicted Cumulative Credit Hours

The near alignment between blue (actual) and orange (predicted) bars represents that the predictive model effectively captures the learning behaviours and patterns related to credit accumulation. As students accumulate credits at various rates, visual data represents both rapid and gradual progressions for individuals that reflect varying steps of course completion. This graph highlights the model's reliability in determining time to course completion.

4.2 Comparative Analysis

Existing techniques' comparative analysis on KDD Cup 2015 dataset is shown in table 6. In comparison with the other methods such as IC-BTCN [24], Supervised LP-Random Forest [4] and PMCT [21], the proposed LTC-LSTM achieves high accuracy of 92.10%, which adapts dynamically to the changing temporal patterns in student behavior, and is capable of capturing both short and long-term dependencies. This continuous time version enables small changes/granularity when modeling irregular engagement sequences which provides an accurate knowledge of student activity patterns and improves the prediction accuracy.

Table 6: Comparative Analysis Of Existing Methods Using The Kdd Cup 2015 Dataset

Methods	Accuracy (%)	F1-score (%)	Recall (%)	Precision (%)
IC-BTCN [24]	89.3	93.4	90.5	96.5
Supervised LP-Random Forest [4]	N/A	88.41	98.11	80.46
PMCT [21]	88	92.8	95.7	90
Proposed LTC-LSTM	92.10	96.86	96.85	96.88

4.3 Discussion

A benefit of proposed LTC-LSTM and existing technique limitation are explained in detail. The existing method limitations like IC-BTCN [24] model, struggled with capturing long-term dependencies across widely spaced course activities because of limited temporal receptive fields. Bipartite link [4] oversimplifies complex learners' behaviors by modelling only interactions among students and course elements. SVR [1] struggled to effectively capture the non-linear and dynamic nature of student engagement in MOOCs with sparse data that fail to accurately identify at-risk students. Ensemble methods [23] suffer from high computational complexity and training time which makes them less practical for large-scale MOOC platforms.

The proposed LTC-LSTM overcomes these existing methods' limitations by capturing irregular and complex temporal dynamics in learner behaviour. The LTC makes a model for adaptively adjusting to varying time intervals that is appropriate for student learning dropout prediction. This flexibility improves the identification of subtle patterns in dropout performance. Moreover, LTC-LSTM generate effective training and less training time compared to existing DL methods. Therefore, this process enhances both robustness and accuracy of dropout prediction systems.

5. CONCLUSION

This research proposed LTC-LSTM for student learning dropout predictions. By introducing continuous time dynamics, LTC-LSTM captures irregular and time-varying patterns in student learning behavior effectively that resulting in enhanced modelling of long-term dependencies and more robust performance. The deployment of label encoding and min-max during the pre-processing stage enhanced model training stability and interpretability of categorical features. These outcomes represent that LTC-LSTM not only enhance prediction accuracy but also minimizes errors and training time via temporal modelling. This enables a more effective and robust method for predicting dropout. The model is evaluated using KDD Cup 2015, Anon, and dropout datasets, which demonstrate superior performance in terms of accuracy, f1-score, recall, and precision compared to existing methods. Moreover, LTC-LSTM obtains high accuracy of 92.10%, 84.65%, and 94.48% on KDD Cup 2015, Anon, and dropout datasets compared to existing methods like IC- BTCN. In future, feature selection methods will be used to select the most appropriate features that enhance model performance.

References

1. C. Jin, "MOOC student dropout prediction model based on learning behavior features and parameter optimization", Interactive Learning Environments, vol. 31, no. 2, August 2020, pp. 714–732.
2. R. Mehdi and M. Nachouki, "A neuro-fuzzy model for predicting and analyzing student graduation performance in computing programs", Education and Information Technologies, vol. 28, no. 3, August 2022, pp. 2455–2484.
3. H. Pallathadka, A. Wenda, E. Ramirez-Asís, M. Asís-López, J. Flores-Albornoz, and K. Phasinam, "Classification and prediction of student performance data using various machine learning algorithms", Mater. Today Proc., vol. 80, January 2023, pp. 3782–3785.
4. F. Khoushehgir and S. Sulaimany, "Negative link prediction to reduce dropout in Massive Open Online Courses", Education and Information Technologies, vol. 28, no. 8, January 2023, pp. 10385–10404.
5. A. K. Pandey, P. Singh, D. Jain, A. K. Sharma, A. Jain, and A. Gupta, "Generative adversarial network and bayesian optimization in multi-class support vector machine for intrusion detection system", International Journal of Intelligent Engineering and Systems, vol. 16, January 2023, pp. 110–119.
6. Nurmalitasari, Z. Awang Long, and M. Faizuddin Mohd Noor, "Factors influencing dropout students in higher education", Education Research International, vol. 2023, no. 1, February 2023, Art. no. 7704142.
7. M. Nagy and R. Molontay, "Interpretable dropout prediction: towards XAI-based personalized intervention", Int. J. Artif. Intell. Educ., vol. 34, no. 2, March 2023, pp. 274–300.
8. J. A. Talamás-Carvajal and H. G. Ceballos, "A stacking ensemble machine learning method for early identification of students at risk of dropout", Education and Information Technologies, vol. 28, no. 9, March 2023, pp. 12169-12189.
9. M. V. Martins, L. Baptista, J. Machado, and V. Realinho, "Multi-class phased prediction of academic performance and dropout in higher education", Applied Sciences, vol. 13, no. 8, April 2023, Art. no. 4702.
10. I. Eegdeman, I. Cornelisz, M. Meeter, and C. van Klaveren, "Identifying false positives when targeting students at risk of dropping out", Education Economics, vol. 31, no. 3, April 2023, pp. 313–325.
11. B. Albreiki, T. Habuza, and N. Zaki, "Extracting topological features to identify at-risk students using machine learning and graph convolutional network models", International Journal of Educational Technology in Higher Education, vol. 20, no. 1, April 2023, Art. no. 23.
12. X. Xia and W. Qi, "Dropout prediction and decision feedback supported by multi temporal sequences of learning behavior in MOOCs", International journal of educational technology in higher education, vol. 20, no. 1, May 2023, Art. no. 32.

13. V. Kamakshamma and K. F. Bharati, “Adaptive-CSSA: adaptive-chicken squirrel search algorithm driven deep belief network for student stress-level and drop out prediction with MapReduce framework”, Social Network Anal. Min., vol. 13, no. 1, May 2023, Art. no. 90.
14. H. Liu, X. Chen, and F. Zhao, “Learning behavior feature fused deep learning network model for MOOC dropout prediction”, Education and Information Technologies, vol. 29, no. 3, June 2023, pp. 3257–3278.
15. M. Vishwakarma and N. Kesswani, “A new two-phase intrusion detection system with Naïve Bayes machine learning for data classification and elliptic envelop method for anomaly detection”, Decision Analytics Journal, vol. 7, June 2023, Art. no. 100233.
16. N. M. Alruwais, “Deep FM-based predictive model for student dropout in online classes”, IEEE Access, vol. 11, September 2023, pp. 96954-96970.
17. S. Malik and K. Jothimani, “Enhancing student success prediction with featurex: a fusion voting classifier algorithm with hybrid feature selection”, Education and Information Technologies, vol. 29, no. 7, September 2023, pp. 8741-8791.
18. C. Blundo, V. Loia, and F. Orciuoli, “A time-aware approach for MOOC dropout prediction based on Rule Induction and Sequential Three-Way Decisions”, IEEE Access, vol. 11, October 2023, pp.113189–113198.
19. E. Sarmas, E. Spiliotis, E. Stamatopoulos, V. Marinakis, and H. Doukas, “Short-term photovoltaic power forecasting using meta- learning and numerical weather prediction independent Long Short-Term Memory models”, Renewable Energy, vol. 216, November 2023, Art. no. 118997.
20. F. Scala, A. Ceschini, M. Panella, and D. Gerace, “A general approach to dropout in quantum neural networks”, Adv. Quantum Technol., December 2023, Art. no. 2300220.
21. K. Niu, Y. Zhou, G. Lu, W. Tai, and K. Zhang, “PMCT: Parallel Multiscale Convolutional Temporal model for MOOC dropout prediction”, Comput. Electr. Eng., vol. 112, December 2023, Art. no. 108989.
22. L. Vives, I. Cabezas, J. C. Vives, N. G. Reyes, J. Aquino, J. B. Cóndor, and S. F. S. Altamirano, “Prediction of students’ academic performance in the programming fundamentals course using long short-term memory neural networks,” IEEE Access, vol. 12, January 2024, pp. 5882–5898.
23. K. Talebi, Z. Torabi, and N. Daneshpour, “Ensemble models based on CNN and LSTM for dropout prediction in MOOC”, Expert Syst. Appl., vol. 235, January 2024, Art. no. 121187.
24. X. Zhang, X. Wang, J. Zhao, B. Zhang, and F. Zhang, “IC-BTCN: A Deep Learning Model for Dropout Prediction of MOOCs Students”, IEEE Trans. Educ., vol. 67, no. 6, June 2024, pp. 974–982.
25. R. Goran, L. Jovanovic, N. Bacanin, M. Stankovic, V. Simic, M. Antonijevic, and M. Zivkovic, “Identifying and understanding student dropouts using metaheuristic optimized classifiers and explainable artificial intelligence techniques”, IEEE Access, vol. 12, August 2024, pp. 122377–122400.
26. Á. Kocsis and G. Molnár, “Factors influencing academic performance and dropout rates in higher education”, Oxford Review of Education, vol. 51, no. 3, February 2025, pp. 414-432.
27. A. M. Rabelo and L. E. Zárate, “A model for predicting dropout of higher education students”, Data Sci. Manage., vol. 8, no. 1, March 2025, pp. 72–85.
28. S. Alghamdi, B. Soh, and A. Li, “ISELDP: An Enhanced Dropout Prediction Model Using a Stacked Ensemble Approach for In- Session Learning Platforms”, Electronics, vol. 14, no. 13, June 2025, Art. no. 2568.
29. KDD Cup dataset link: https://www.kaggle.com/datasets/sst2023/kdd-cup-2015
30. Dropout dataset link: http://moocdata.cn/data/user-activity
31. Anon dataset link: https://zenodo.org/records/4114005