IT Journal Research and Development (ITJRD)
Vol.10, No.2, March 2026, E-ISSN : 2528-4053 | P-ISSN : 2528-4061
DOI : 10.25299/itjrd.2025.21431

DBCE: Dynamic Bipartite Contextual Embedding for Predicting National Science Olympiad Winners

Nasy'an Taufiq Al Ghifari[1], Luthfi Angger Ramdhani[2], Slamet Riyadi[3]
Faculty of Engineering, Universitas Muhammadiyah Yogyakarta[1,2,3]
nasyan.taufiq@umy.ac.id[1], luthfi.angger.ft22@mail.umy.ac.id[2], riyadi@umy.ac.id[3]

Article Info

Article history:

Received Feb 11, 2025
Revised Jul 21, 2025
Accepted Jan 14, 2026

Keyword:

Link Prediction
Dynamic Bipartite Graph
Contextual Graph Embedding
OSN Forecasting
DBCE

ABSTRACT

This study introduces Dynamic Bipartite Contextual Embedding (DBCE), a graph-based method for predicting school-level success in Indonesia's National Science Olympiad (OSN). DBCE combines three essential components: explicit bipartite graph structure, temporal representation learning, and contextual feature integration. The model is trained on a comprehensive OSN dataset from 2009 to 2024, which includes junior and senior high school participation records enriched with publicly available metadata such as school accreditation, teacher qualifications, and geographic location. Experimental results demonstrate that DBCE outperforms four baseline models across multiple evaluation metrics, including AUC, F1-score, NDCG, modularity, inference time, and training stability. Ablation studies confirm the importance of each architectural component. DBCE effectively identifies historically dominant schools and emerging institutions with high potential, offering practical insights for educational policy and strategic planning. The proposed framework provides a scalable and interpretable solution for forecasting academic achievement and monitoring institutional performance in educational settings.

Corresponding Author:

Nasy'an Taufiq Al Ghifari
Faculty of Engineering
Universitas Muhammadiyah Yogyakarta
Brawijaya Street, Tamantirto, Kasihan, Bantul, Yogyakarta, Indonesia
Email: nasyan.taufiq@umy.ac.id

1. INTRODUCTION

Education plays a fundamental role in shaping the future of a nation. In Indonesia, national-level competitions such as the National Science Olympiad (OSN) are strategic programs to identify, nurture, and recognize student talents in science and social sciences. The OSN is organized annually by the Ministry of Education, Culture, Research, and Technology as part of a broader effort to promote educational excellence and equal access to academic development. The selection process is multi-tiered, beginning at the elementary level and progressing through the district, province, and finally the national levels. Those students who do well at the OSN are often chosen to represent Indonesia at prestigious international competitions such as the IMO and IPhO.

Research suggests that participation in academic competitions can enhance student confidence, motivation, and engagement in learning, especially when supported by structured training environments and mentorship [1], [2], [3], [4], [5]. Schools that regularly excel on the OSN typically possess exceptional educators, targeted coaching initiatives, and robust institutional support [6], [7], [8]. These factors influence both immediate competitive outcomes and sustained academic development. It is therefore strategically important for policy-making, planning for mentorship, and fairly distributing educational resources to be able to guess which schools are likely to do well in future OSN events.

Graph-based machine learning is a strong way to model this predictive problem. A dynamic bipartite graph can show historical OSN data, with one set of nodes representing schools and the other representing medal categories. Edges between nodes show successful participation and are time-stamped to show how they change over time. This representation makes it possible to learn about complicated patterns in school performance over time. One of the previous methods to predict links in dynamic bipartite graphs is Dynamic Bipartite Graph Embedding (DBGE). This method uses horary random walks and skip-gram learning to put nodes into a latent space that shows structural and temporal relationships [9].

DBGE encounters certain issues, despite its contributions. It focuses exclusively on structural interactions and temporal sequences, disregarding non-structural contextual elements that are often crucial to educational achievement [10], [11], [12]. School accreditation, the quantity of skilled educators, and geographic disparities are critical elements influencing institutional performance. Prior studies have emphasized integrating contextual variables to improve model accuracy and interpretability in real-world applications [13], [14].

This paper presents Dynamic Bipartite Contextual Embedding (DBCE), a novel method designed to improve temporal bipartite graph learning using school-level contextual features. DBCE integrates temporal link data with external information such as accreditation status, instructor count, and historical participation records. The approach employs a context-sensitive graph neural architecture tailored for the educational sector. This enables the acquisition of structural patterns and semantic representations crucial for institutional success.

The primary contributions of this research can be encapsulated as follows:

- DBCE is introduced as a novel methodology that amalgamates structural and contextual data to forecast OSN victors using dynamic bipartite graphs.
- An exhaustive OSN dataset from 2009 to 2024, including both junior and senior high school levels, is supplemented with publicly available school metadata.
- We compare the DBCE method to four baseline models: DBGE [9], LPFormer[15], NBFNet [16], and HeLPormer [17]. We use standard evaluation metrics like AUC, F1-score, NDCG, modularity, inference time, and stability.
- Experimental results indicate that DBCE improves predictive accuracy and offers clear insights into the systemic aspects affecting academic success in OSN.

The remainder of this paper is organized as follows. Section 2 analyzes relevant literature on dynamic bipartite graph embedding and educational link prediction, highlighting the deficiencies of existing techniques. Section 3 delineates the proposed DBCE technique, encompassing its architectural components, contextual integration, and temporal modeling. Section 4 discusses the experimental configuration, baseline comparisons, evaluation metrics, and outcomes. Section 5 analyzes the findings, identifies the model's deficiencies, and proposes avenues for additional investigation. Section 6 concludes the research by examining the primary contributions and potential applications of DBCE in educational data analytics.

2. RELATED WORK

Graph-based machine learning has shown considerable promise in the modeling of intricate relational data [16], [18], [19]. Nonetheless, most graph neural network (GNN) models have been designed for static and homogeneous graphs, characterized by uniform node types and an absence of explicit temporal dependencies [18], [19]. On the other hand, educational datasets like OSN participation have different types of nodes and interactions that change over time. A better and more informative way to show this is to model schools and competition outcomes as nodes in a bipartite graph, with links showing OSN achievements over time.

2.1. Bipartite Graph Embedding

There are several methods that have been suggested just for learning representations in bipartite graphs. BiNE (Bipartite Network Embedding) is one of the first examples. It uses different embedding methods for each type of node and biased sampling to find proximity between partitions [20]. Li et al. (2021) discussed how motif-based link prediction adds subgraph structures that keep local interaction patterns the same, making link estimation better [21]. Bipartite Graph Convolutional Networks (BGCN) and other methods try to use message-passing directly on the bipartite topology [22], [23]. These methods work well in static settings, but they usually don't support temporal information and don't take into account node-level attributes. Because of this, they can't be used for real-world tasks that constantly change, like predicting OSN.

2.2. DBGE: Dynamic Bipartite Graph Embedding

Cai et al. (2020) came up with the Dynamic Bipartite Graph Embedding (DBGE) method to make link predictions in time-sensitive bipartite graphs [9]. DBGE uses Horary Random Walk to create node sequences that follow chronological order. Then, it uses a skip-gram language model to learn embeddings that show how time is structured. DBGE has done well in applications for analyzing employee retention and allocating healthcare resources.

DBGE has two major problems, despite doing many good things. First, it does not include contextual features at the node level. In education, institutional factors like accreditation, the number of qualified teachers, and the availability of funding affect how well students do [8], [24]. Second, DBGE does not distinguish semantic roles between schools and competition categories beyond structural divisions. Because of these limits, it is less useful for tasks that need accurate predictions and easy-to-understand results.

2.3. LPFormer: Transformer-Based Link Prediction

LPFormer, created by Shomer et al. (2024), uses a Transformer-based architecture to model how node pairs interact using self-attention mechanisms [15]. This method gets good results in link prediction by finding long-range dependencies and complicated relational patterns. But LPFormer was made for static, homogeneous graphs. It does not support bipartite topology natively or incorporate temporal sequences or node-level metadata. Its applicability to time-aware educational forecasting is therefore constrained.

2.4. NBFNet: Neural Bellman-Ford Network

The Neural Bellman-Ford Network (NBFNet) modifies the Bellman-Ford algorithm to acquire path-based representations for link prediction [16]. NBFNet sends messages along several paths to show how nodes are connected in more than one hop. It performs well in transductive and inductive settings and efficiently in large-scale graphs. Nevertheless, NBFNet does not account for bipartite constraints and lacks explicit mechanisms for modeling temporal evolution. These limitations make it less ideal for educational tasks involving sequential participation and categorical node interactions.

2.5. HeLPormer: Heterogeneous Link Prediction Transformer

Fan et al. (2025) created HeLPormer, a transformer architecture for heterogeneous link prediction [17]. The model uses attention-based multi-hop encoding to guess missing links between semantic roles and considers the node type. HeLPormer does a great job with knowledge graphs and recommendation systems. However, it does not enforce a strict bipartite structure during aggregation and focuses more on node-type encoding than external node attributes. For tasks that depend on both bipartite separation and context awareness, HeLPormer provides limited generalization.

2.6. Gap in Literature

A review of existing models reveals an important gap. To the best of current knowledge, no existing method integrates temporal dynamics, bipartite structure, and node-level contextual attributes in a single, unified architecture. DBGE models temporal bipartite graphs but lacks contextual inputs. LPFormer and NBFNet work well in homogeneous settings but don't support bipartite modeling or external features. HeLPormer includes different types of structures but doesn't include important contextual information. This gap is especially important in educational analytics because factors at the school level greatly impact how well students do and must be taken into account when making predictions [25], [26].

Integrating these dimensions (time, structure, and context) is critical for accurate and explainable prediction in domains such as OSN analysis. Without this unification, models may only partially understand the underlying dynamics that govern institutional success.

2.7. Position of DBCE

The proposed approach, DBCE (Dynamic Bipartite Contextual Embedding), addresses the deficiency by unifying structural learning, temporal modeling, and the incorporation of contextual data. DBCE maintains a bipartite topology, simulates the temporal formation of linkages, and utilizes empirical data from entities such as school accreditation and staffing metrics. This cohesive architecture enables DBCE to acquire more intricate and comprehensible node representations, enhancing the accuracy and generalizability of OSN winner predictions. Table 1 illustrates the comparative capabilities of DBCE and baseline techniques.

Table 1. Comparison of DBCE with Baseline Methods

Method	Bipartite Support	Temporal Modeling	Contextual Features	Graph Type
DBGE [9]	Yes	Yes	No	Dynamic Bipartite
LPFormer [15]	No	No	No	Homogeneous
NBFNet [16]	No	Multi-hop only	No	General
HeLPormer [17]	Partial	Yes	Partial	Dynamic Heterogeneous
DBCE	Yes	Yes	Yes	Dynamic Bipartite + Context

3. DBCE METHOD

This part talks about the DBCE method, which uses a dynamic bipartite graph structure to look at how people interact over time and in different situations to predict how well students will do in school in the OSN. DBCE unifies three critical aspects: bipartite graph modeling, temporal representation learning, and node-level contextual integration into a single framework. This unified approach distinguishes DBCE from prior work in dynamic link prediction, particularly in educational or institutional domains. To the best of the authors' knowledge, no prior model has proposed a mechanism that jointly addresses bipartite structure, temporal evolution, and contextual node attributes within a single predictive architecture. Figure 1 illustrates the comprehensive workflow of the DBCE approach.

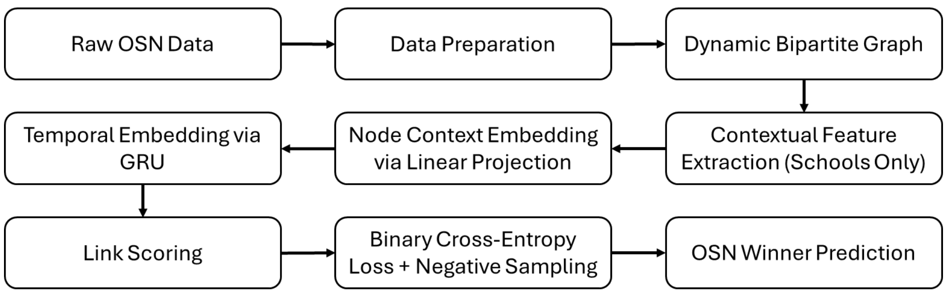

Figure 1. Workflow of the Dynamic Bipartite Contextual Embedding Method

3.1. Data Preparation

This subchapter describes transforming raw data into dynamic bipartite graph representations in prediction models. The OSN dataset initially contained attributes such as participant name, gender, school, province, district/city, field of competition, school level, class, medals, additional prizes, and year. From these attributes, three main attributes were selected to form the graph structure, namely:

- Nodes: school, and a combination of competition-medal fields.
- Edges: year of competition.

A combination of race and medal fields (e.g., "Mathematics-Gold") is used as a single node type to simplify the graph structure. Meanwhile, the school's attributes are chosen as the main node because students only have a limited opportunity to participate in the OSN (a maximum of three times). In contrast, the school can participate on an ongoing basis every year. Therefore, schools have become a more stable and relevant entity that can be used as the prediction object.

To enrich contextual information, public metadata about the school is added as an additional attribute to the school node. This process results in a dynamic bipartite graph that represents the interaction between schools and OSN medal categories from year to year. A visualization of the graph structure can be seen in Figure 2.

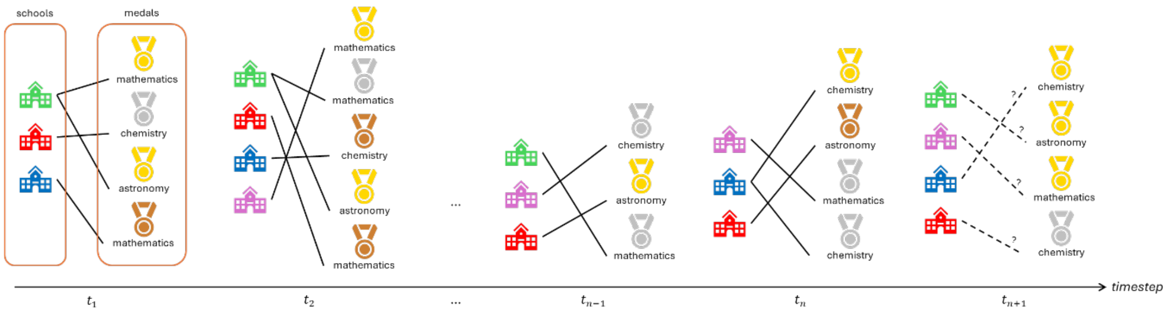

Figure 2. Example of a dynamic bipartite graph showing the interaction between the school and the OSN medal category. Each medal category node is only connected once at a time, while school nodes can be connected more than once or not at all. Schools that have no relationship will be hidden in the graph visualization. The main challenge in this link prediction task is to predict the last time interval based on the history of the previous graph structure.

3.2. Modeling Dynamic Bipartite Graph

A dynamic bipartite graph is one of the graphs that has special characteristics because it combines the structure of a bipartite graph and a dynamic graph structure. In a bipartite graph, the vertex type consists of two groups of vertices, where vertices within the same group cannot be directly connected, unlike in a heterogeneous graph. Meanwhile, dynamic graphs generally only have one type of node or are considered homogeneous graphs. Bipartite graphs are generally also static, so when dynamic characteristics are added to their structure, the treatment for the link prediction task cannot be the same. Special treatment is needed for this type of dynamic bipartite graph [27].

A Dynamic bipartite graph is $G = \{G_{s,m}^1, G_{s,m}^2, \dots, G_{s,m}^t\}$, where $G^t = (V^t, E^t)$ represents the t-th snapshot. Each $E_{s,m} \subseteq V_s \times V_m$ is a set of edges that connect a node in V_s representing schools to a node in V_m representing medal-category, which are defined by the combination of competition subject and medal type (e.g., Gold–Physics, Bronze–Chemistry). V^t is a set of nodes with two types of nodes. E^t is the set of edges that inform when relation $E_{s,m}$ occurs. T is the number of snapshots. So, in G^t, $V^t = \{V_s \cup V_m\}$ and $E^t = \{E_{s,m}\}$.

3.3. Integration of Contextual Features

The structural and temporal dimensions of OSN performance yield significant insights; however, institutional variables, including accreditation status, instructional capacity, and geographic location, also influence success. DBCE adds contextual attributes for school nodes to capture these non-structural effects. Each school node $v \in V_s$ is linked to a feature vector $x_v \in R^d$, which may encompass variables such as the number of teachers, type of region (urban or rural), previous OSN participation history, and additional school-level indicators.

A learnable linear transformation puts these contextual features in the same latent space as the structural representations:

$$c_v = W_c \cdot x_v + b_c \tag{1}$$

Where W_c and b_c are trainable parameters. The resulting vector c_v is integrated into the model's temporal update mechanism, ensuring that the evolution of a school's embedding is shaped by its graph neighborhood and unique institutional profile. Medal-category nodes do not receive contextual embeddings, as they do not possess associated real-world features or variation across time.

3.4. Temporal Representation Learning

DBCE uses a recurring message-passing system that updates node embeddings over time to track changing patterns in OSN participation. At each time step t, a node v updates its embedding $h_v^{(t)}$ based on its previous embedding $h_v^{(t-1)}$ and the aggregated messages from its neighbors in the current graph snapshot $G^{(t)}$. We use a Gated Recurrent Unit (GRU) to implement this temporal embedding mechanism. The GRU is good for modeling sequences and helps the network learn temporal dependencies without losing gradients.

The update rule is defined like this:

$$h_v^{(t)} = \mathrm{GRU}\left(h_v^{(t-1)}, \mathrm{AGG}\left(\{h_u^{(t-1)}: u \in \mathcal{N}(v)\}\right) + c_v\right) \tag{2}$$

The function AGG represents an aggregation operation over the embeddings of neighboring nodes. Depending on model complexity and available resources, this operation may be implemented as a simple mean pooling, summation, or as a learnable attention mechanism. The contextual vector c_v is included only when v is a school node, allowing the GRU to condition its update on both structural messages and school-specific context. The update for medal-category nodes is based solely on message aggregation from connected school nodes and their past embeddings.

Internally, the GRU learns gates determining how much of the past embedding to retain and how much of the new aggregated signal to incorporate. This design enables the model to capture long-term patterns such as consistent excellence or sudden changes in performance.

3.5. Link Scoring and Prediction

Once all temporal snapshots have been processed, DBCE uses the final embeddings $h_v^{(T)}$ to predict potential links in the next time step $T + 1$. The task is to estimate whether a school $s \in V_s$ will win a medal in category $m \in V_m$. This is framed as a binary classification task, where each possible school–category pair is scored using a link prediction function.

The score is computed using a bilinear projection followed by a sigmoid activation:

$$P(y_{s,m} = 1) = \sigma\left(w^{\top}\left[h_s^{(T)} | h_m^{(T)}\right] + b\right) \tag{3}$$

Here, $h_s^{(T)}$ and $h_m^{(T)}$ are the final embeddings of the school and medal-category nodes, respectively, and w, b are parameters of the prediction layer. The symbol || denotes vector concatenation. The sigmoid function ensures that the output is a number between 0 and 1, which lets you use thresholding to make binary decisions.

3.6. Supervised Learning and Optimization

The model is trained from start to finish using a binary cross-entropy loss on both the observed links and the sampled negative pairs. Negative sampling is necessary because the OSN dataset is sparse, and most combinations of schools and categories don't show real wins. For every positive instance in the training data, multiple negative instances are generated by associating the same school with medal categories that were not awarded that year.

The total loss across all sampled pairs is expressed as:

$$\mathcal{L} = -\sum_{(s,m)\in\mathcal{Y}} y_{s,m} \log(\hat{y}_{s,m}) + (1 - y_{s,m}) \log(1 - \hat{y}_{s,m}) \tag{4}$$

Where $y_{s,m} \in \{0,1\}$ is the true label, $\hat{y}_{s,m}$ is the predicted probability from the model, and $\mathcal{Y}$ denotes the set of all positive and negative pairs used for training. Depending on the size of the model and how it learns, optimization is done using mini-batch stochastic gradient descent or a variant like Adam.

This loss function makes the model give high probability to links that have been seen and low probability to links that haven't been seen, while still being able to generalize to school-category pairs that haven't been seen yet.

3.7. Scalability and Efficiency

The DBCE model is made to work well with computers and to be able to handle large datasets in education. A bipartite graph structure makes it easier to send messages by only allowing interactions between cross-type neighbors. This cuts down on the number of aggregation operations per node. Also, the temporal updates through GRU are localized and incremental, meaning embeddings can change without starting over from scratch. These design choices ensure that the framework scales linearly with the number of nodes and edges per snapshot and the number of time steps in the graph sequence. Such properties make DBCE practical for deployment in real-world educational data pipelines, accumulating new participation data yearly.

Algorithm 1 Dynamic Bipartite Contextual Embedding (DBCE)

Require: Sequence of bipartite graphs $\mathcal{G} = \{G^{(1)}, G^{(2)}, \ldots, G^{(T)}\}$, context vectors $\{\mathbf{x}_v\}$ for $v \in V_s$, learning rate η
Ensure: Predicted probabilities $\hat{y}_{s,m}$ for each school–medal pair in year $T+1$

Initialize node embeddings $\mathbf{h}_v^{(0)}$ for all $v \in V_s \cup V_m$
Compute context embeddings $\mathbf{c}_v = \mathbf{W}_c \cdot \mathbf{x}_v + \mathbf{b}_c$ for each $v \in V_s$
for $t = 1$ to T **do**
  for all nodes $v \in V_s \cup V_m$ **do**
    Aggregate neighbor messages:

$$\mathbf{m}_v^{(t)} = \mathrm{AGG}\left(\left\{\mathbf{h}_u^{(t-1)} : u \in \mathcal{N}(v)\right\}\right)$$

    if $v \in V_s$ **then**
      $\mathbf{z}_v^{(t)} \leftarrow \mathbf{m}_v^{(t)} + \mathbf{c}_v$
    else
      $\mathbf{z}_v^{(t)} \leftarrow \mathbf{m}_v^{(t)}$
    end if
    Update embedding using GRU:

$$\mathbf{h}_v^{(t)} = \mathrm{GRU}\left(\mathbf{h}_v^{(t-1)}, \mathbf{z}_v^{(t)}\right)$$

  end for
end for
for all school–medal pairs (s, m) **do**
  Predict probability of future link:

$$\hat{y}_{s,m} = \sigma\left(\mathbf{w}^\top \left[\mathbf{h}_s^{(T)} \| \mathbf{h}_m^{(T)}\right] + b\right)$$

end for
Compute binary cross-entropy loss over positive and negative samples
Update parameters using gradient descent

4. EXPERIMENTS AND EVALUATION

This chapter thoroughly assesses the DBCE method's efficacy in forecasting the connectivity between schools and medal categories at the OSN event. The evaluation was conducted to assess the model's effectiveness in capturing historical patterns, temporal dynamics, and institutional contexts, and to compare it with four current baseline methods. All analyses followed a strict experimental protocol using OSN data at the Indonesian junior and senior high school levels. The structure of this chapter includes a description of the dataset, exposure of comparative methods, experimental setup, evaluation metrics, key results, prediction interpretation, error analysis, and ablation studies to measure the contribution of each architectural component.

4.1. Dataset Description

This subchapter presents a statistical description and coverage of the OSN datasets used in the study. The dataset was acquired from public sources on Kaggle (https://www.kaggle.com/datasets/anakpindahan/indonesia-national-science-olympiad-osn) and encompasses information regarding participation and attainment of OSN medals at the junior and senior high school levels from 2009 to 2024. The dataset has information about each entry, such as the school's name, year, field of competition, type of medal (gold/silver/bronze), competition level, and education level.

This dataset is then turned into a temporal bipartite graph, with one set of nodes representing the school and the other representing a mix of race and medal fields. The relationship between nodes shows that a school won a certain medal in the same field of competition and year. Graphs for junior high and high school levels are built separately to maintain structural accuracy. A statistical summary of the OSN dataset is presented in Table 2. Details about the process of forming the graph and selecting attributes can be seen in Subchapter 3.1.

Table 2. Statistical properties of the OSN dataset

Property	SMP	SMA	Combined
Number of records	6,826	12,501	19,327
Number of participating schools	2,588	1,620	4,208
Number of medal-winning records	1,688	4,022	5,710
Number of medal categories (subjects)	3	9	12
Number of years	16	16	16

4.2. Baseline Methods

To objectively and measurably evaluate the DBCE method's performance, this study uses four widely known baseline methods to study link prediction on graphs. These four methods were chosen because they represent different approaches in terms of architecture, representation strategies, and the ability to handle dynamic and heterogeneous graph data. Although DBCE is a model that combines structural, temporal, and contextual representations in an integrated manner, the baselines used for each have certain relevant strengths as a comparator.

First, DBGE (Dynamic Bipartite Graph Embedding) is the method that comes closest to the initial architecture of DBCE. DBGE uses a random walk approach to form a sequence of nodes in a dynamic bipartite graph, which is then processed with skip-gram learning. Although it does not use contextual features, DBGE excels at explicitly representing structural and temporal linkages and is the main baseline in this study. Second, LPFormer (Link Prediction Transformer) is a transformer architecture-based model that utilizes a self-attention mechanism to study long-term relationships between nodes. Although designed for homogeneous graphs and does not specifically maintain bipartite structures, LPFormer is known to perform well in modeling network structures' temporal dynamics and evolutionary patterns. Third, NBFNet (Neural Bellman-Ford Network) is a propagation path-based approach that adapts the principles of the Bellman-Ford algorithm. This model calculates the connectivity between nodes by tracing the multi-hop paths available in the graph. Its advantages lie in its computational efficiency and flexibility in handling large graphs, although its main weakness is its inability to handle bipartite structures and contextual information simultaneously. Fourth, HeLPormer (Heterogeneous Link Prediction Transformer) is a transformer architecture developed specifically to handle heterogeneous graphs. This model considers different node types and supports semantic role-based aggregation. Despite being able to process more complex graphs, HeLPormer does not explicitly support bipartite constraints and has not thoroughly adopted contextual attributes.

In this study, the four baseline models were reimplemented and modified to be compatible with the OSN dataset, particularly in dealing with dynamic bipartite structures. However, contextual features such as school accreditation, number of qualified teachers, and geographic location are only granted to DBCE because the baseline architecture does not support the integration of such attributes without substantial modifications.

4.3. Experimental Setup

The experimental setup is designed to simulate real-world prediction scenarios, in which the model must project new relationships based on historical connectivity histories. For this reason, the experiment was formulated as an annual link prediction task, with a sliding window approach on temporal graph data.

In each iteration of the experiment, the model was trained using data from 2009 to t-th, then evaluated to predict new relationships in the $t+1$ year. For example, for the 2021 prediction, the model studied the entire graph snapshot from 2009 to 2020 and tested it on actual 2021 data. The scheme was repeated for the entire target year (2021, 2022, 2023, and 2024), resulting in four rounds of independent experiments whose results were then averaged to measure overall performance. To ensure consistency and relevance of the predictions, the model only gives information about nodes that have already appeared in the training data. This means that new schools that have never participated in the training period are excluded from the evaluation because they fall into a cold-start scenario that this model does not address. In the future, this scenario can be studied separately using a similarity-based transfer learning or initialization approach between schools.

Contextual features are provided exclusively on school nodes. Three features were used in this experiment: (1) the school's accreditation status, (2) the number of qualified teachers, and (3) the school's geographic location by province (with one-hot encoding). These three features are obtained from public data sources and embedded into the node representation through a linear transformation layer in the DBCE architecture. Each model (including the baseline) uses a 64-dimensional embedding representation. The model is optimized using an Adam optimizer with an initial learning rate determined through a search grid (the best value is found at 0.001), and the training process uses early stopping based on AUC metrics on the validation set. The validation set is taken from a snapshot one year before the target year. To overcome the imbalance in the number of links (since most node pairs are unrelated), a negative sampling strategy is applied with a ratio of 1:3, meaning that each positive link is matched with three randomly generated negative pairs (i.e., between schools and categories that are possible but never occur).

The entire experiment was repeated five times with different random initializations to ensure that the performance obtained was independent of a particular seed. Average values and standard deviations from key metrics are recorded to support performance stability analysis. The entire experiment was implemented using PyTorch and the Deep Graph Library (DGL). The training and inference process was performed on machines with the following specifications: AMD Ryzen 5 processor, 16 GB of RAM, and NVIDIA RTX 3080 GPU (10 GB VRAM). With this configuration, the training time for a single experimental fold (e.g., 2021 prediction) averages between 20 and 30 minutes depending on the model. Through this experimental setup, the DBCE model was tested in terms of prediction accuracy, computational efficiency, and resistance to training data variations. All baselines are evaluated in an identical scheme to ensure a fair comparison.

4.4. Evaluation Metrics

To fully assess the model's performance, six important metrics were used that show different aspects of prediction quality: accuracy of classification, quality of ranking, structure of representation, speed of computation, and stability of results. These metrics were chosen so that the evaluation results would not only focus on narrowly predictive performance but also consider the practical aspects and the quality of the representation produced by the model.

AUC (Area Under the ROC Curve) is used as a key metric to measure the model's ability to distinguish between positive links (school–medal category relationships that occur) and negative links (relationships that do not). The AUC value is in the range [0, 1], with higher values indicating better separation capability. This metric is not threshold sensitive and is suitable for probabilistic evaluations [28].

F1-score is the harmonic mean between precision and recall after the probabilistic prediction value is rounded up to a binary classification decision using a threshold of 0.5. This metric measures the balance between false positives and false negatives. It is especially useful in unbalanced datasets, such as graphs with a predominance of negative links. F1-score is also used to compare the performance of "hard" predictions between models.

Normalized Discounted Cumulative Gain (NDCG) is a ranking evaluation metric that values the model's predictive sequence. In this context, the model is not only required to predict the correct link but also to place the correct link at the top of the prediction list. NDCG is calculated based on the cumulative gain of the correct link with a logarithmic penalty for the lower position. An NDCG value close to 1 indicates that the model successfully sequenced high-quality predictions [18].

Different from other predictive metrics, modularity (Q) is used to assess the quality of the embedding representation generated by the model, particularly in the context of community structures in bipartite graphs. Modularity measures the extent to which the embedding structure separates the nodes into dense communities inside and rarely outside. In this context, a community is expected to form between schools and medal categories that are historically strongly interconnected. A modularity value greater than 0.5 indicates a strong community structure.

Inference time is the average time the model takes to make predictions based on one snapshot of the yearly graph (like 2021). These metrics are important for figuring out how well the model works in real-world situations, like making a recommendation system or keeping an eye on school performance in real time. In practice, using very accurate models may not be possible, but it may take a long time to make predictions.

Model stability was measured by calculating the standard deviation of AUC values from five iterations of the experiment. High stability (low deviation) indicates that the model performs consistently against initial weight initialization variations and training batches. This metric is also used to measure the training process's robustness and the model's sensitivity to noise in the data.

These six metrics complement each other and are used simultaneously in the analysis in subchapter 4.5 et seq. With this evaluation framework, the study assesses the model's accuracy in guessing connectivity and investigates the quality of the internal structures formed and the performance implications in real-world use.

4.5. Results

The experiment results showed that the DBCE method consistently excelled in all evaluation metrics compared to the four baseline methods in binary prediction and ranking quality. The evaluation was conducted for the last four years (2021–2024), and the results presented in Table 3 are the four years' averages. In terms of classification, DBCE recorded the highest AUC score of 0.928, far surpassing DBGE (0.894), LPFormer (0.861), and NBFNet (0.840). A similar trend was seen in the F1-score, where DBCE gained 0.741, demonstrating its ability to produce accurate and stable binary predictions. Regarding ranking quality, DBCE also excels with an NDCG value of 0.802, indicating its ability to prioritize relevant connectivity at the top.

Table 3. Performance Comparison (Average across 2021–2024)

Method	AUC	F1	NDCG	Q	Time (s)	Stability
DBCE	0.928	0.741	0.802	0.611	7.23	0.012
DBGE	0.894	0.703	0.766	0.577	6.98	0.018
LPFormer	0.861	0.685	0.751	0.549	12.80	0.025
HeLPormer	0.878	0.692	0.760	0.561	11.02	0.021
NBFNet	0.840	0.662	0.726	0.519	9.65	0.029

The DBCE representation is better because its modularity (Q) value is 0.611, which shows that the school and medal categories form a more coherent and meaningful embedding cluster. This number is higher than LPFormer's (0.549) and NBFNet's (0.519), which means that DBCE's community structure better represents how institutions have interacted in the past in OSN data. DBCE takes an average of 7.23 seconds to infer one annual snapshot, which is a little longer than DBGE (6.98 seconds) but much faster than LPFormer (12.80 seconds) and HeLPormer (11.02 seconds). This makes it a good choice for applications that need to check on things regularly.

Training stability is also an advantage of DBCE, with a standard deviation of AUC between repetitions of only 0.012, indicating consistent performance despite retraining with random initialization. NBFNet and LPFormer recorded deviations of 0.029 and 0.025, respectively. Overall, the combination of explicit bipartite structures, temporal dynamics, and contextual features makes DBCE superior in prediction accuracy, computational efficiency, and the formation of meaningful community representations. These findings form the basis for further analysis in the next subchapter, which discusses prediction interpretation, identification of institutional patterns, and systematic evaluation of model errors.

4.6. Predictions and Insights

After the training and validation process, the DBCE model is used to project the connectivity between the school and the gold medal category in 2024. The focus on gold medals was chosen because this category is considered the main indicator of institutional achievement in the OSN event. For each field of competition, probabilistic predictions were made on all school nodes that had appeared in the training data by calculating the possible connectivity between the school and the medal category. The predicted results are then sorted for each field. Tables 4 and 5 present the top three schools with the highest probability of winning gold medals, separately for junior and senior high school levels.

The prediction results analysis shows that historically dominant schools such as SMP Labschool Jakarta remain in the top ranking at the junior high school level. However, the model also managed to identify institutions such as MTsN 1 Yogyakarta, which showed an increasing trend despite not being dominant nationally. This indicates that contextual features in DBCE can capture nonlinear institutional growth signals. At the high school level, the dominance of flagship schools such as SMA Kristen BPK Penabur Gading Serpong and SMAN Unggulan M.H. Thamrin looks consistent. Still, the model also highlights schools such as SMA Semesta, SMA Kristen Petra, and MAN Insan Cendekia Gorontalo, which reflect DBCE's sensitivity to the performance dynamics of private and public schools outside the city center.

DBCE's ability to identify both dominant institutions and emerging schools demonstrates the effectiveness of contextual and historical attribute integration strategies in embedding formation. These findings validate the architectural approach of DBCE and provide potential practical uses for OSN organizers and educational coaching institutions in developing training programs based on institutional potential. Further evaluation of concrete predictions and fault analysis will strengthen the understanding of the strengths and limitations of this model.

Table 4. Three junior high schools have the highest probability of predicting gold medals per OSN competition.

Junior High School Competition	School 1 (Prob)	School 2 (Prob)	School 3 (Prob)
Mathematics	SMPN 2 Makassar (0.674)	SMP Labschool Jakarta (0.615)	SMPN 5 Bandung (0.615)
Natural Sciences	SMPN 8 Semarang (0.663)	MTsN 1 Yogyakarta (0.627)	SMPN 3 Palembang (0.610)
Social Sciences	SMPN 1 Denpasar (0.643)	SMPN 17 Palembang (0.638)	SMPN 2 Surabaya (0.628)

Table 5. Three senior high schools have the highest probability of predicting gold medals per field in the OSN competition.

Senior High School Competition	School 1 (Prob)	School 2 (Prob)	School 3 (Prob)
Mathematics	SMA Kristen BPK Penabur Gading Serpong (0.947)	SMA Kristen Petra 1 (0.882)	SMA Kharisma Bangsa (0.867)
Physics	SMAN Unggulan M.H. Thamrin (0.911)	SMA Semesta (0.870)	SMA Kharisma Bangsa (0.835)
Biology	SMA Kristen BPK Penabur Gading Serpong (0.935)	SMAN 8 Jakarta (0.916)	SMA Semesta (0.873)
Chemistry	SMA Kristen BPK Penabur Gading Serpong (0.935)	SMA Semesta (0.875)	SMA Kharisma Bangsa (0.840)
Astronomy	SMAN Unggulan M.H. Thamrin (0.936)	SMA Taruna Nusantara (0.900)	SMA Kharisma Bangsa (0.879)

Computer science	SMA Kristen BPK Penabur Gading Serpong (0.963)	SMA Semesta (0.926)	SMA Sutomo 1 Medan (0.862)
Economics	SMAN 8 Jakarta (0.948)	SMA Semesta (0.921)	SMA Kristen Petra 1 (0.910)
Earth	SMA Semesta (0.896)	SMA Kristen Petra 2 (0.881)	SMAN 1 Pati (0.790)
Geography	SMAN 8 Jakarta (0.908)	SMA Darul Ulum 2 Unggulan BPPT Jombang (0.859)	MAN Insan Cendekia Gorontalo (0.743)

4.7. Error Analysis

Although DBCE showed superior performance in predicting the connectivity between schools and OSN medal categories, error analysis is needed to identify model limitations and understand conditions where predictions are inaccurate. One of the main sources of error is cold-start cases, i.e., schools that have just joined or have historically had minimal participation. Because embedding is formed based on a multi-year history of interconnectedness, schools with limited data tend to obtain less informative representations. Although contextual features such as accreditation and geographic location contribute, historical signals remain dominant in embedding formation, so new institutions with high potential are often conservatively predicted.

In addition, the model shows a bias towards historically dominant schools, such as SMA Kristen BPK Penabur Gading Serpong and SMP Labschool Jakarta. This phenomenon is known as historical overconfidence, where models rely too much on past winning patterns and are less responsive to short-term performance fluctuations. The unequal distribution of resources among competition fields presents a significant challenge, particularly in disciplines characterized by a limited number of participants and victors, such as Astronomy or Social Studies. The sparse structure of the graph causes embedding to be prone to noise, so predictions in these fields tend to be unstable.

To overcome these limitations, several improvement strategies can be considered: (1) the integration of explicit temporal features such as the medal trend of the last three years, (2) regularization of historical dominance with the adjustment of medal frequency weights, (3) data augmentation for small fields through cluster-based sampling, and (4) two-stage training to improve the embedding of new schools based on geographical and institutional relationships. This error analysis confirms that while DBCE is structurally and predictively robust, an adaptive approach is still needed to address performance dynamics, unbalanced data distribution, and representation of entities with limited data.

4.8. Ablation Study

To evaluate the contribution of each component in DBCE architecture, an ablation study was conducted on three main elements: contextual features, temporal modeling, and explicit bipartite structures. Each component was omitted separately, and the remaining model was evaluated using the same metrics as in the main experiment. The three ablation variants tested included: (1) DBCE without contextual features, where school nodes are represented only by ID and historical connections; (2) DBCE without temporal modeling, with a statically combined node representation of the entire year; and (3) DBCE without an explicit bipartite structure, where the graph is treated as homogeneous without node-type separation. The evaluation was conducted for 2021–2024, and the results were averaged, as shown in Table 6.

Table 6. Results of the ablation study using the DBCE method.

Variant	AUC	F1	NDCG	Q
DBCE (Full)	0.928	0.741	0.802	0.611
– No Context	0.902	0.713	0.774	0.587
– No Temporal	0.889	0.694	0.763	0.566
– No Bipartite Struct.	0.874	0.681	0.749	0.534

The ablation results demonstrated that each component substantially enhanced the overall performance of the DBCE. Removing contextual features led to a decrease in AUC from 0.928 to 0.902 and a decrease in F1 and NDCG, indicating the importance of socio-geographic information such as accreditation and school location. When temporal modeling is eliminated, the performance decline is more pronounced, with AUC dropping to 0.889 and F1 to 0.694, confirming that historical dynamics between years are crucial in embedding formation. The elimination of the bipartite structure produced the most drastic impact, with the AUC decreasing to 0.874 and F1 to 0.681, and the modularity of Q falling to 0.534, suggesting that an explicit separation between the school and the medal category was necessary to maintain the interpretability and coherence of the representation.

These results show that DBCE architecture is not loosely modular but depends on tightly linking graph structure, temporal dynamics, and contextual attributes. Not a single component can be removed without substantially degrading performance. The combination of the three results in a model that is accurate, stable, and adaptive to the complexity of educational data. The implications of this study open up opportunities for similar architectural exploration for other domains, such as mapping regional competition coaching, predicting educational social networks, or dynamic graph-based cross-year achievement monitoring systems.

5. DISCUSSION

The results of the experiments presented in Chapter 4 show that DBCE architecture offers an effective and adaptive approach in predicting the connectivity between schools and the potential for medal acquisition at the OSN event. The advantages of DBCE are reflected in the dominance of evaluation metrics such as AUC, F1-score, and NDCG over the four baseline methods. This superiority stems from the synergy of three main components: explicit bipartite structures, dynamic temporal modeling, and integration of contextual features. The ablation study confirmed that removing one of these components led to a significant decrease in performance, confirming that the effectiveness of DBCE depends on the thorough integration between architectural elements.

The interpretation of the prediction results shows that DBCE can identify the dominance pattern of established institutions such as SMA BPK Penabur and SMAN 8 Jakarta, as well as detect the potential of new institutions such as MAN Insan Cendekia and SMA Taruna Nusantara. This capability suggests that the model relies on historical data and is sensitive to contextual signals and local dynamics. For education policy makers, this result opens up opportunities to develop a more inclusive and data-based OSN development strategy, taking into account institutions that show an improvement trend even though they do not have a long track record.

Although DBCE showed strong performance, some limitations still need to be noted, such as sensitivity to cold-start cases, sparsity in small race fields, and the lack of accommodation for sudden structural changes. To address this, further development can be directed towards integrating external data (e.g., curriculum and trainer track record), implementing transfer learning for new nodes, and designing multi-task learning architectures to handle multiple areas of the competition in parallel. Academically, DBCE makes an important contribution to the formulation of domain-based dynamic bipartite graph embedding. At the same time, it has the potential to be the foundation of educational intelligence systems for the continuous monitoring and fostering of institutional achievement.

6. CONCLUSION

This study proposes and evaluates DBCE architecture as a predictive approach to modeling the connectivity between educational institutions and medal acquisition in the OSN event. By integrating explicit bipartite graph structures, temporal dynamics, and institutional contextual features, DBCE demonstrated superior performance over four baseline methods in various evaluation metrics, including AUC, F1-score, NDCG, community modularity, inference efficiency, and training stability. The predicted results identify historically dominant institutions and capture the growth potential of previously underexposed schools, making DBCE practically relevant to support data-driven coaching policies.

In the future, model development can improve the representation of new schools through transfer learning, integration of non-structured external data such as trainer and curriculum track records, and multi-task learning architecture design to address correlations between competition fields. With the expansion of scope and increased adaptability, DBCE can become the foundation of a national education intelligence system that can detect patterns, map potential, and sustainably encourage equitable distribution of academic achievement.

REFERENCES

[1] M. Bilal, M. Omar, W. Anwar, R. H. Bokhari, and G. S. Choi, "The role of demographic and academic features in a student performance prediction," *Sci Rep*, vol. 12, no. 1, pp. 1–9, Dec. 2022, doi: 10.1038/S41598-022-15880-6.

[2] X. Du, J. Yang, J.-L. Hung, and B. Shelton, "Educational data mining: a systematic review of research and emerging trends," *Inf Discov Deliv*, vol. 48, no. 4, pp. 225–236, May 2020, doi: 10.1108/IDD-09-2019-0070.

[3] E. Kalita *et al.*, "Educational data mining: a 10-year review," *Discover Computing*, vol. 28, no. 1, p. 81, May 2025, doi: 10.1007/s10791-025-09589-z.

[4] C. Romero and S. Ventura, "Educational data mining and learning analytics: An updated survey," *Wiley Interdiscip Rev Data Min Knowl Discov*, vol. 10, no. 3, Feb. 2024, doi: 10.1002/widm.1355.

[5] W. Setyonugroho, S. Sesotya, I. Permana, T. Lestari, D. Mahendra, and H. Abda, "Enhancing Predictive Accuracy: Assessing the Effectiveness of SVM in Predicting Medical Student Performance," *E3S Web of Conferences*, vol. 465, p. 02028, Dec. 2023, doi: 10.1051/e3sconf/202346502028.

[6] C. H. Chen and N. T. T. Phan, "Development of the self-efficacy beliefs of engineering undergraduates preparing for an international contest," *Pertanika Journal of Social Sciences and Humanities*, vol. 29, no. 2, pp. 819–837, Jun. 2021, doi: 10.47836/pjssh.29.2.05.

[7] R. A. Young *et al.*, "Jumble judging: Cognitive and affective outcomes of intercollegiate collaboration at a soil judging competition," *Natural Sciences Education*, vol. 52, no. 1, Jun. 2023, doi: 10.1002/nse2.20104.

[8] A. Winata, B. Lian, and Mulyadi, "An Academic Supervision of School Principals and Teacher Performance and Its Implementation of Education Quality at State Junior High School in Cengal District," *Advances in Social Science, Education and Humanities Research*, vol. 565, 2021.

[9] X. Cai *et al.*, "DBGE: Employee Turnover Prediction Based on Dynamic Bipartite Graph Embedding," *IEEE Access*, vol. 8, pp. 10390–10402, 2020, doi: 10.1109/ACCESS.2020.2965544.

[10] I. B. Suleiman, O. A. Okunade, E. G. Dada, and U. C. Ezeanya, "Key factors influencing students' academic performance," *Journal of Electrical Systems and Information Technology 2024 11:1*, vol. 11, no. 1, pp. 1–18, Sep. 2024, doi: 10.1186/S43067-024-00166-W.

[11] T. Nema and S. Palwe, "Predicting Students' Academic Success using Machine Learning Algorithms," in *2023 7th International Conference On Computing, Communication, Control And Automation (ICCUBEA)*, IEEE, Aug. 2023, pp. 1–7. doi: 10.1109/ICCUBEA58933.2023.10392079.

[12] A. L. J. Martinez, K. Sood, and R. Mahto, "Early Detection of At-Risk Students Using Machine Learning," *Communications in Computer and Information Science*, vol. 2261 CCIS, pp. 396–406, 2025, doi: 10.1007/978-3-031-85930-4_36.

[13] Z. Wu and Z. Zhang, "Development Strategies for Higher Education Institutions Based on the Cultivation of Core Competitiveness," *International Journal of Emerging Technologies in Learning*, vol. 16, no. 21, pp. 163–176, 2021, doi: 10.3991/ijet.v16i21.26873.

[14] E. López-Meneses, P. C. Mellado-Moreno, C. Gallardo Herrerías, and N. Pelícano-Piris, "Educational Data Mining and Predictive Modeling in the Age of Artificial Intelligence: An In-Depth Analysis of Research Dynamics," *Computers*, vol. 14, no. 2, p. 68, Feb. 2025, doi: 10.3390/computers14020068.

[15] H. Shomer, Y. Ma, H. Mao, J. Li, B. Wu, and J. Tang, "LPFormer: An Adaptive Graph Transformer for Link Prediction," *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 2686–2698, Jun. 2024, doi: 10.1145/3637528.3672025.

[16] Z. Zhu, Z. Zhang, L. P. Xhonneux, and J. Tang, "Neural Bellman-Ford Networks: A General Graph Neural Network Framework for Link Prediction," *Adv Neural Inf Process Syst*, vol. 35, pp. 29476–29490, Jun. 2021, Accessed: Jul. 28, 2025. [Online]. Available: https://arxiv.org/pdf/2106.06935

[17] H. Fan, Z. Fan, Y. Lu, P. Jiao, and N. Ruan, "Helpormer: Heterogeneous Graph Transformer for Link Prediction," 2025, doi: 10.2139/SSRN.5297031.

[18] J. Li, C. Yang, G. Ye, and Q. V. H. Nguyen, "Graph neural networks with deep mutual learning for designing multi-modal recommendation systems," *Inf Sci (N Y)*, vol. 654, Jan. 2024, doi: 10.1016/j.ins.2023.119815.

[19] H. Liang, S. Wang, H. Li, L. Zhou, X. Zhang, and S. Wang, "BiGNN: Bipartite graph neural network with attention mechanism for solving multiple traveling salesman problems in urban logistics," *International Journal of Applied Earth Observation and Geoinformation*, vol. 129, May 2024, doi: 10.1016/j.jag.2024.103863.

[20] M. Gao, L. Chen, X. He, and A. Zhou, "BiNE: Bipartite Network Embedding," in *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*, New York, NY, USA: ACM, Jun. 2018, pp. 715–724. doi: 10.1145/3209978.3209987.

[21] C. Li, W. Wei, X. Feng, and J. Liu, "Research of Motif-Based Similarity for Link Prediction Problem," *IEEE Access*, vol. 9, pp. 66636–66645, 2021, doi: 10.1109/ACCESS.2021.3077016.

[22] Y. Jiang, H. Lin, Y. Li, Y. Rong, H. Cheng, and X. Huang, "Exploiting node-feature bipartite graph in graph convolutional networks," *Inf Sci (N Y)*, vol. 628, pp. 409–423, May 2023, doi: 10.1016/j.ins.2023.01.107.

[23] C. He *et al.*, "Cascade-BGNN: Toward Efficient Self-supervised Representation Learning on Large-scale Bipartite Graphs," vol. 12, 2020, doi: 10.1145/1122445.1122456.

[24] H.-C. Chen *et al.*, "Learning Performance Prediction with Imbalanced Virtual Learning Environment Students' Interactions Data," 2022, pp. 330–340. doi: 10.1007/978-3-030-79728-7_33.

[25] M. Yağcı, "Educational data mining: prediction of students' academic performance using machine learning algorithms," *Smart Learning Environments*, vol. 9, no. 1, p. 11, Dec. 2022, doi: 10.1186/s40561-022-00192-z.

[26] R. Tuanaya, S. Ali, S. Aslam, and A. Ajmal, "Machine Learning in Educational Data Mining: Current Trends and Emerging Gaps in Predicting Student Performance," *Journal of Technological Pedagogy and Educational Development*, vol. 10, no. 1, pp. 10–27, 2025, doi: 10.59247/jtped.v2i1.5.

[27] N. T. Al Ghifari, G. A. P. Saptawati, M. L. Khodra, and B. Sitohang, "The Evaluation of DyHATR Performance for Dynamic Heterogeneous Graphs," *Journal of ICT Research and Applications*, vol. 17, no. 2, pp. 231–248, Aug. 2023, doi: 10.5614/ITBJ.ICT.RES.APPL.2023.17.2.7.

[28] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognit Lett*, vol. 27, no. 8, pp. 861–874, Jun. 2006, doi: 10.1016/J.PATREC.2005.10.010.