The Hybrid QGNN–QSVM (H-QGQS) Framework for the Prediction of University Publication Performance

Lukman Anas
Doctoral Program of Information Systems, Postgraduate School, Universitas Diponegoro, Semarang, Indonesia | Department of Informatics, Universitas Muhammadiyah Makassar, Makassar, Indonesia
lukmananas@unismuh.ac.id (corresponding author)

Aghus Sofwan
Department of Electrical Engineering, Universitas Diponegoro, Semarang, Indonesia
asofwan@elektro.undip.ac.id

Iwan Setiawan
Department of Electrical Engineering, Universitas Diponegoro, Semarang, Indonesia
iwansetiawan@live.undip.ac.id

Received: 19 January 2026 | Revised: 7 March 2026 and 20 March 2026 | Accepted: 21 March 2026

DOI: https://doi.org/10.48084/etasr.17650

ABSTRACT

Assessing university publication performance using scientometric indicators is challenging due to the relational and non-linear nature of scholarly data. This study proposes a hybrid Quantum Graph Neural Network and Quantum Support Vector Machine (QGNN–QSVM) framework for graph-based classification of university publication performance into Low, Medium, and High categories. The framework integrates graph representation learning and quantum-enhanced classification through multiple experimental scenarios, progressing from classical machine learning models to a fully quantum-enhanced learning approach. Experimental results show that the classical Support Vector Machine (SVM) baseline achieved an accuracy of 0.92 on a balanced dataset of 3,818 records, whereas the proposed QGNN–QSVM framework achieved an accuracy of 0.73. The results indicate that while the classical SVM achieves higher overall accuracy, incorporating graph structures and quantum-based embeddings provides more expressive representations of the complex scientometric relationships. In particular, the proposed framework demonstrates strong capability in identifying high-performing universities, which is the most critical category in institutional performance evaluation. These findings suggest that quantum-enhanced graph learning offers a promising alternative for modeling complex scientometric data under current near-term quantum computing constraints. The study also highlights existing limitations and suggests future research directions, including architectural optimization, feature enrichment, and scalability improvements.

Keywords-Quantum Machine Learning (QML); Graph Neural Networks (GNNs); scientometric analysis; Quantum Support Vector Machine (QSVM); institutional performance

I. INTRODUCTION

The performance of scientific publications is widely recognized as a key indicator of a university's research capacity, scientific productivity, and global reputation, particularly in the era of digital transformation. Various scientometric indicators, including publication counts, citation metrics, institutional h-index values, journal quartile distribution (Q1–Q4), and Field-Weighted Citation Impact (FWCI), are commonly used to evaluate scientific output and institutional competitiveness within the global research ecosystem [1-3]. Beyond measuring research productivity, these indicators also reflect collaboration patterns, the structure of academic networks, and the strategic direction of research development within universities. As research governance increasingly relies on data-driven evaluation, the ability to analyze and predict publication performance has become essential for institutional planning, strategic decision-making, and research management.

Various computational approaches have been applied to scientometric analysis, including classical machine learning and deep learning techniques [4, 5]. These approaches have demonstrated promising capability in modeling patterns within publication data and predicting research impact. However, most existing models treat scientometric indicators as independent numerical variables, ignoring the inherent relational structure among entities such as authors, institutions, citations, and collaborations. In practice, scientometric data naturally form interconnected networks where relationships between entities play an essential role in determining research performance. Modeling these relational dependencies using traditional feature-based learning approaches remains challenging, as classical models often struggle to capture nonlinear interactions and complex structural patterns within such networks.

Recent advances in Quantum Machine Learning (QML) provide new opportunities to address these limitations by enabling the exploration of high-dimensional feature spaces and more expressive representations of complex patterns [6-8]. In particular, Quantum Support Vector Machines (QSVMs) extend classical Support Vector Machine (SVM) models by employing quantum kernels that map data into richer Hilbert spaces, potentially improving classification capability for complex datasets [9-11].

Meanwhile, Graph Neural Networks (GNNs) have emerged as a powerful framework for modeling relational data structures by explicitly learning from node–edge interactions within graph representations [12, 13]. Building upon this concept, Quantum Graph Neural Networks (QGNNs) incorporate quantum computing principles such as superposition and entanglement to represent graph structures in quantum states, enabling more expressive modeling of nonlinear relationships and complex topological dependencies that are difficult to capture using classical graph learning methods [14, 15].

Earlier research in scientometrics has largely relied on traditional bibliometric techniques such as publication counting, citation analysis, and collaboration mapping using statistical and network analysis approaches [16, 17]. While these methods are effective for describing historical research performance, they often fail to capture hidden patterns and nonlinear interactions within large-scale heterogeneous scientific datasets. Consequently, machine learning algorithms such as logistic regression, SVM, naive Bayes, decision trees, and neural networks have increasingly been used to model relationships between scientometric indicators and scientific impact [18, 19]. Although these approaches improve predictive capability, most of them rely on vector-based representations and therefore cannot fully exploit the relational structure embedded within scientometric networks.

Recent studies have also explored hybrid quantum–classical learning architectures, particularly in applications such as image classification using Quantum Convolutional Neural Networks (QCNN) [20]. These approaches demonstrate the potential of integrating classical preprocessing with quantum circuits to enhance feature representation and classification performance. However, existing studies primarily focus on image-based datasets and do not explicitly address graph-structured data such as scientometric networks, where relational dependencies among entities play a crucial role.

In response to this research gap, this study introduces the Hybrid QGNN–QSVM (H-QGQS) framework as a predictive method specifically designed to classify and predict the publication performance of higher education institutions. In this study, the classification task refers to categorizing Muhammadiyah and 'Aisyiyah Higher Education Institutions (PTMA) into three performance levels: Low, Medium, and High, based on their scientometric indicators.

This classification is not only a technical modeling task but also serves as a decision-support mechanism for evaluating institutional research performance. In real-world applications, the proposed framework can assist policymakers, university administrators, and research managers in identifying high-performing institutions, detecting underperforming institutions, and formulating strategies for improving research productivity and resource allocation.

To achieve this objective, the proposed framework integrates QGNNs for learning relational representations of scientometric data and QSVMs for classification in a high-dimensional quantum feature space. By combining graph-based learning with quantum-enhanced classification, this study aims to provide a more expressive and robust predictive model while offering meaningful insights for institutional research management and policy development.

II. METHODOLOGY

A. Research Framework Overview

This study proposes a H-QGQS framework as a predictive method to examine and project university publication performance using scientometric data. This framework combines quantum graph learning and quantum classification within a hybrid quantum–classical architecture, aiming to reveal complex correlations in publication data and enhance prediction accuracy in high-dimensional feature spaces. Overall, the research process consists of several main stages, including data collection and preprocessing, scientometric graph construction, representation learning using a QGNN, publication performance classification using QSVM, and predictive model evaluation, as illustrated in Figure 1.

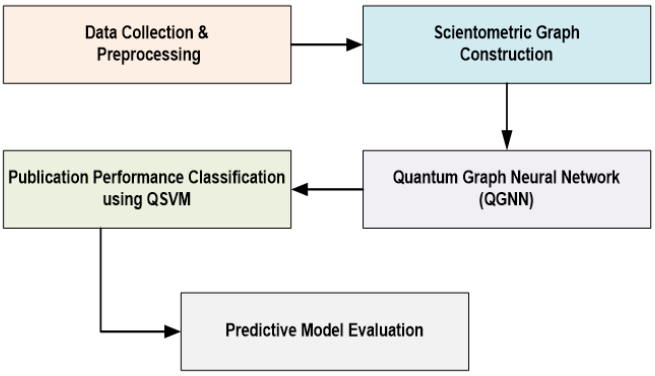

Fig. 1. Research stages.

B. Data Collection and Preprocessing

Data collection was carried out using the Indonesian scientific indexing portal, Science and Technology Index (SINTA), which offers integrated scientometric information from various global databases, including Scopus, Web of Science (WoS), and Google Scholar. Through this portal, the research obtained verified bibliographic data that reflect the publication performance of higher education institutions on an institutional level. The data collected included key scientometric indicators that show publication activity at PTMA.

In this study, a total of 19,088 scientific publication records were collected and used as the initial dataset, representing publication activities at PTMA. The dataset was extracted from the SINTA portal on 15 October 2025. Each record represents an institution-level scientometric profile associated with a unique SINTA ID, containing aggregated publication indicators such as SINTA score, citation counts, and h-index values from multiple databases.

To ensure data quality and consistency, several preprocessing steps were applied. First, duplicate institutional records were removed to ensure that each PTMA was uniquely represented. Second, records with incomplete or missing key scientometric indicators were excluded. Third, institutions with zero or non-informative values across all indicators were removed, as they do not contribute meaningful patterns for classification.

Finally, a class balancing process based on the tertile distribution of SINTA scores was applied to ensure equal representation of Low, Medium, and High categories. As a result of these filtering and preprocessing steps, the dataset was reduced from 19,088 to 3,818 valid and balanced samples used for subsequent modeling and evaluation.

Table I shows a sample of scientometric data from the SINTA portal used in this study. Each row represents a single institutional entity (PTMA), identified by its SINTA ID and affiliation code, with publication performance indicators from various international databases. The SINTA score indicates the institution's overall research performance, whereas the h-index from Scopus, Google Scholar, and WoS shows the impact of publications based on each database's scope and characteristics. The presence of indicators from various sources creates heterogeneous data characteristics, both in score scales and in the distribution of publication impacts across institutions. Moreover, the table shows a wide range of performance, from institutions with high impact scores to those with zero scores across all indicators. This situation highlights the importance of data preprocessing before using them in further modeling.

Next, during the preprocessing stage, the scientometric data were cleaned to remove duplicate entities and standardize institutional affiliations so that each PTMA was uniquely identified. Then, indicator values were normalized to minimize bias caused by differences in scale between data sources and institutions. This normalization ensures that no single indicator dominates the learning process due to varying value ranges.

TABLE I. SAMPLE DATA OF SCIENTIFIC PUBLICATIONS OF PTMA

No	**Sinta ID**	**Affiliation code**	**Sinta score**	**Scopus h-index**	**Google h-index**	**WoS h-index**
1	23026	51013	11,061.7	26	35	13
2	6672813	51013	4,433.25	20	26	7
3	23016	51013	6,707.19	11	30	4
4	4547	61004	6,333.37	16	20	8
5	5981297	91059	3,223.57	11	20	8
6	6176471	51013	2,749.51	17	18	16
7	5974494	51013	6,947.32	10	31	1
8	5986807	101002	3,629.02	15	33	8
9	23059	51013	6,263.12	5	23	4
10	6774551	51013	2,910.5	5	9	3
11	5989087	61008	4,383.05	11	20	10
12	257453	31011	5,024.99	13	20	7
13	6691487	91065	7,897.65	3	30	0
14	6001481	41087	2,743.63	6	11	10
15	257094	41061	6,516.3	30	32	20
16	6000001	51007	2,178.12	12	14	4
17	23037	51013	3,980.55	8	14	6
…	…	…	…	…	…	…
19,088	18083	6941440	0	0	0	0

The results from the preprocessing stage were then compiled into initial features (node attributes) that represent the publication performance of each PTMA. These features form the foundation for building a scientometric graph where collaborative relationships between institutions are clearly modeled.

C. Publication Performance Labeling

To enable the classification of institutional publication performance, each institution in the dataset was assigned a performance label based on the distribution of the institutional SINTA score. The institutions were categorized into three performance levels: Low, Medium, and High.

A tertile-based thresholding strategy was applied to ensure balanced class distribution. The institutional SINTA scores were sorted in ascending order and divided into three groups. Based on the score distribution in the dataset, institutions with SINTA scores in the range 0–3,200 were labeled as Low, those with scores in the range 3,201–6,400 were labeled as Medium, and institutions with scores greater than 6,400 were labeled as High. This labeling strategy allows the classification model to capture different levels of institutional research performance while maintaining a balanced multi-class dataset suitable for machine learning training and evaluation.

D. Scientometric Graph Construction

The next step after collecting and pre-processing the data was to build a scientometric graph that illustrates the relational structure of higher education publication performance. At this stage, the scientometric data, initially in table form and feature vectors, were converted into graph-structured data to clearly represent the complex relationships among academic entities. In this study, the scientometric graph is represented as:

$$G = (V, E) \qquad (1)$$

where V represents the set of nodes and E the set of edges. Each node signifies an institutional entity, specifically PTMA, whereas the edges represent the scientific collaboration relationships between institutions established through joint publications, citation links, and cross-institutional author affiliations. Each node $v_i \in V$ corresponds to a PTMA institution and is associated with a scientometric feature vector x_i. This feature vector consists of normalized quantitative indicators from the previous stage, including:

- Number of publications,
- Number of citations,
- h-index (Scopus, Google Scholar, and WoS) values,
- Institutional SINTA score.

Edges between nodes represent scientific collaboration relationships between institutions. An edge e_{ij} is created between two institutions v_i and v_j if the institutions share at least one collaborative publication, citation relationship, or cross-institutional author affiliation within the dataset. In this study, the graph is represented as an undirected network, where the presence of an edge indicates the existence of a collaboration relationship between institutions. The resulting graph structure enables the model to capture relational patterns among institutions, which cannot be effectively represented using independent feature vectors. By modeling institutional interactions as a graph, the subsequent learning process can exploit both node attributes and topological relationships, allowing more comprehensive representation learning for scientometric analysis.

E. Representation Learning Using Quantum Graph Neural Network

After constructing the scientometric graph, the next step is representation learning to extract latent structural information from the graph. This process is performed using a QGNN, which integrates graph-based learning with QML to model complex relational patterns among institutions.

The scientometric graph $G = (V, E)$, together with the node feature matrix X and adjacency matrix A, serves as the input to the QGNN model. Each node v_i is associated with a feature vector x_i derived from normalized scientometric indicators. These classical features are mapped into quantum states through a quantum feature encoding process. In this study, angle encoding is employed to transform classical node features into quantum states. Each feature component is encoded as a rotation angle applied to a qubit using parameterized rotation gates. The quantum circuit consists of 6 qubits with a circuit depth of four layers, employing a hardware-efficient ansatz composed of parameterized rotation gates and entangling Controlled NOT (CNOT) operations.

Figure 2 illustrates the conceptual stages of the quantum learning process adopted in the proposed framework. The process consists of four main components. The first component, Quantum Encoding, converts classical scientometric features into quantum states, enabling the representation of high-dimensional publication data in the quantum domain. In this study, the node feature vectors are encoded using an angle encoding strategy, where each normalized scientometric feature is mapped to the rotation angle of a quantum gate. The quantum circuit employs 6 qubits, corresponding to the encoded feature dimensions of each institutional node.

Fig. 2. Stages of the QGNN framework.

The second component, Quantum Message Passing, facilitates the propagation of information between interconnected nodes in the graph through quantum operations. In this stage, relational dependencies among institutions are modeled by applying parameterized quantum gates that allow information exchange between qubits representing neighboring nodes in the graph structure. This mechanism enables the model to capture structural relationships embedded within the scientometric collaboration network.

The third component, Quantum Aggregation and Transformation, aggregates quantum messages received from neighboring nodes and applies variational quantum transformations to produce enriched representations. The Parameterized Quantum Circuit (PQC) consists of three variational layers (circuit depth = 3), each composed of rotation gates (Ry and Rz) followed by linear CNOT entanglement between adjacent qubits. This entanglement structure allows the circuit to capture nonlinear dependencies and correlations among node features in the graph. In total, the variational circuit contains 48 trainable parameters distributed across the rotation gates within the circuit layers.

Finally, the fourth component, Measurement and Embedding, performs quantum measurement to project the learned quantum states back into a classical feature space. The expectation values of the measured qubits form the final node embeddings z_i, which represent the learned structural characteristics of institutional publication performance within the scientometric graph.

The QGNN model is trained using the Adam optimizer with a learning rate of 0.001 for 100 training epochs. All quantum circuits were implemented using the Qiskit Machine

Learning framework and executed on the Qiskit Aer state-vector simulator, which enables efficient simulation of quantum circuits under current Noisy Intermediate-Scale Quantum (NISQ) constraints. The resulting node embeddings z_i are subsequently used as input for the classification stage using the QSVM model.

F. *Publication Performance Classification Using Quantum Support Vector Machine*

At this stage, the institutional graph embeddings z_i produced by the QGNN serve as input for classifying publication performance. This classification is performed using a QSVM, which leverages a quantum kernel to project the embeddings into a high-dimensional quantum feature space, enabling the model to capture complex nonlinear relationships among institutions. The QSVM process begins by encoding each embedding vector into a quantum state through a quantum feature map implemented using PQCs. In this study, angle encoding is applied to transform classical embedding features into quantum states. The encoding process can be represented as:

$$|\phi(z_i)\rangle = U(z_i)|0\rangle \tag{2}$$

where z_i is the institutional embedding resulting from the QGNN, $U(z_i)$ is the PQC, and $|0\rangle$ is the ground state of the qubit. The similarity between two institutional embeddings is then calculated using the quantum kernel as:

$$K(z_i, z_i) = \left|\langle \phi(z_i) | \phi(z_j) \rangle\right|^2 \tag{3}$$

Based on the kernel, QSVM constructs a decision function to identify the optimal hyperplane that separates the university publication performance classes, which is formulated as:

$$f(z) = \sum_{i=1}^{N} a_i y_i K(z_i, z) + b \tag{4}$$

where a_i represents the support vector coefficients, y_i is the publication performance class label (Low, Medium, and High), and b is the bias. Through this mechanism, QSVM performs classification.

G. *Evaluation of Model Predictions*

Model performance evaluation was conducted to assess the effectiveness of the proposed H-QGQS framework in predicting institutional publication performance. The dataset was divided into training and testing subsets using an 80:20 split strategy, where 80% of the data were used for model training and the remaining 20% were reserved for testing. This split ensures that the evaluation reflects the model's ability to generalize to unseen data.

Prediction performance was evaluated using standard multi-class classification metrics, including accuracy, precision, recall, and F1-score. These metrics were calculated based on the confusion matrix derived from the predicted and actual class labels.

Accuracy measures the proportion of correctly classified instances among all observations:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{5}$$

where TP represents true positives, TN represents true negatives, FP denotes false positives, and FN denotes false negatives.

Precision evaluates the proportion of correctly predicted positive instances among all predicted positives:

$$\text{Precision} = \frac{TP}{TP+FP} \tag{6}$$

Recall measures the proportion of correctly identified positive instances among all actual positives:

$$\text{Recall} = \frac{TP}{TP+FN} \tag{7}$$

Finally, the F1-score provides a harmonic mean of precision and recall:

$$\text{F1} = \frac{2\times\text{Precision}\times\text{Recall}}{\text{Precision}+\text{Recall}} \tag{8}$$

These evaluation metrics enable a comprehensive assessment of the classification model, particularly in multi-class prediction tasks involving Low, Medium, and High publication performance categories.

III. RESULTS AND DISCUSSION

This section presents experimental results and a discussion on the performance of the H-QGQS framework in predicting university publication success. Evaluations were carried out to assess the model's ability to recognize complex scientometric patterns and the effectiveness of combining graph-based learning and quantum classification.

A. *Experimental Scenarios*

To assess the effectiveness of the H-QGQS framework, multiple experimental scenarios were created to examine the contribution of each model component and compare them with traditional approaches. All scenarios used the same scientometric dataset and identical training and testing schemes to ensure consistent evaluation. After preprocessing and class balancing, a total of 3,818 scientometric records were used.

Four experimental scenarios (S1–S4) were designed to evaluate different stages of the proposed research framework illustrated in Figure 2. Scenario S1 evaluates baseline performance using classical feature representations and a standard SVM, corresponding to the conventional learning stage. Scenario S2 examines the contribution of graph-based representation learning using a GNN. Scenario S3 evaluates a hybrid approach by integrating QGNN-based representation learning with classical SVM classification. Scenario S4 assesses the complete hybrid framework by integrating QGNN-based representation learning with QSVM classification.

The first scenario (S1) used a traditional machine learning method by representing scientometric indicators as feature vectors and classifying them with a standard SVM as a baseline.

Table II indicates that the baseline SVM model achieved an accuracy of 0.92, demonstrating strong overall classification performance. The High class performed best with a precision of 0.98 and an F1-score of 0.96, showing that high-performing institutions can be identified accurately. The Low class also maintained stable performance, with a recall of 0.95, reflecting the model's consistent ability to detect low-performing institutions. In contrast, the Medium class had the lowest recall (0.86), indicating difficulty in distinguishing middle-tier institutions due to overlapping scientometric features. This is supported by Figure 3 (confusion matrix), which reveals that most misclassification errors occurred within the Medium class, primarily being misclassified as Low. This pattern highlights the limitations of feature vector-based SVM in capturing complex relationships between scientometric indicators, indicating a need for a more advanced method in future analyses.

TABLE II. PERFORMANCE EVALUATION OF SVM

Class	Precision	Recall	F1-score	Support
High	0.98	0.95	0.96	1,272
Low	0.89	0.95	0.92	1,274
Medium	0.89	0.86	0.88	1,272
Accuracy			**0.92**	3,818

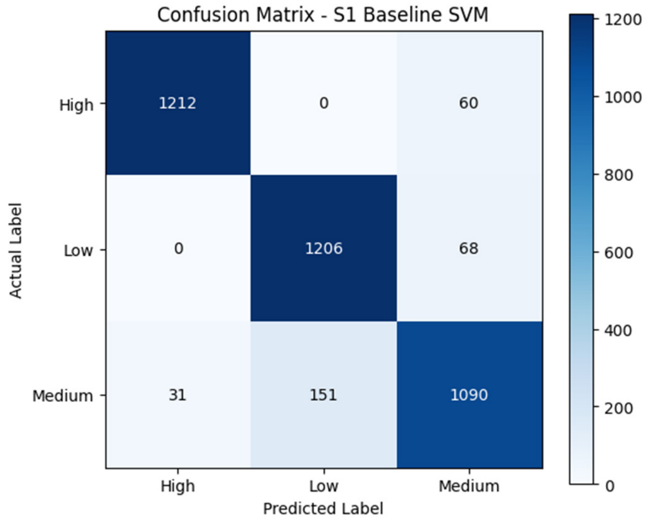

Fig. 3. Confusion matrix of the baseline SVM model.

Meanwhile, in the second scenario (S2), publication performance analysis was conducted using a scientometric graph representation to capture the relationships among institutions. Unlike S1, which used independent feature vectors, in S2, scientometric data were first formed into a graph where nodes represent institutions and edges represent collaboration or citation links. Graph representation learning was performed using a traditional GNN, which aims to combine node attribute information with the graph's topological structure. The node embeddings produced by the GNN were then used as input for a standard SVM classifier. This scenario is designed to assess the contribution of graph-based learning without involving quantum learning components.

Figure 4 shows a scientometric graph representation. Visually, the graph indicates that nodes with high publication performance (High) occupy more central positions and have denser connections, reflecting strong connectivity and similarity with other institutions. Conversely, nodes with low performance (Low) are scattered at the periphery with fewer connections. This pattern suggests a natural hierarchical structure in the scientometric network, where high-performing institutions serve as the central hubs of relationships. Nodes with medium performance (Medium) occupy a transitional role, connecting Low and High clusters. This distribution suggests that medium-performing institutions share features with both extremes, making them harder to distinguish using feature vectors alone, as in the baseline scenario (S1). The presence of closely grouped color clusters demonstrates that graph-based learning can capture relational patterns and topological structures that cannot be modeled by conventional machine learning methods. Therefore, this visualization supports the quantitative finding that graph representations add relevant context in analyzing higher education publication performance.

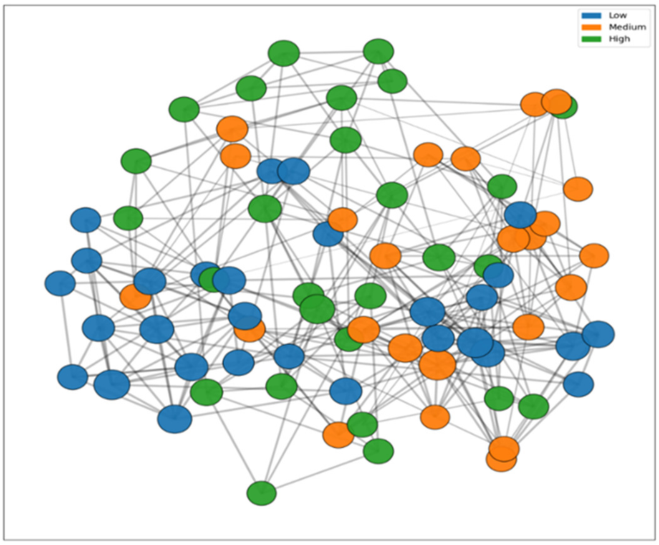

Fig. 4. Scientometric graph representation.

Moreover, in the third scenario (S3), a QGNN was used to learn scientometric graph embeddings, whereas classification was performed with a classical SVM. Unlike S2, which used a conventional GNN, S3 adopts QML principles to generate graph representations. The same institutional graph is used as input, with each node encoded into a quantum state and processed through quantum message passing and quantum aggregation using PQCs. The resulting quantum-based graph embedding then serves as input to the SVM, creating a hybrid quantum–classical architecture.

Table III shows that the QGNN embedding–SVM model achieved 0.77 accuracy, confirming that quantum-based graph embeddings effectively capture relational structure. The Low class demonstrated the most consistent performance with precision, recall, and F1-score values of 0.80 each, reflecting the model's reliable ability to identify that class. The High class had a relatively high recall (0.80) but lower precision (0.73), suggesting some false positives. The Medium class showed the lowest performance with an F1-score of 0.74, indicating overlapping feature characteristics among classes. Overall, these results confirm that the H-QGQS framework is stable on small, balanced datasets and shows potential for further development with larger datasets. Figure 5 provides a visualization of QGNN embeddings.

TABLE III. PERFORMANCE EVALUATION OF QGNN EMBEDDING–SVM

Class	Precision	Recall	F1-score	Support
High	0.73	0.80	0.76	1,272
Low	0.80	0.80	0.80	1,272
Medium	0.78	0.70	0.74	1,272
Accuracy			**0.77**	3,818

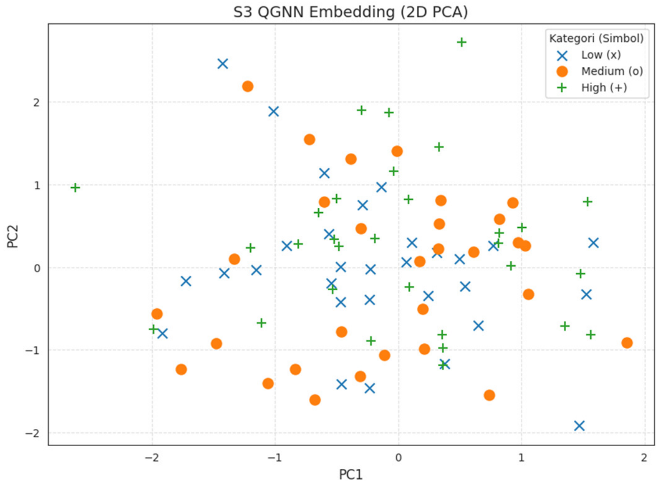

Fig. 5. QGNN embedding visualization.

The QGNN embedding visualization in Figure 5, reduced to two dimensions with Principal Component Analysis (PCA), shows that the Low class forms a relatively dense and consistent cluster, aligning with the stable classification performance. The High class displays a wider spread, indicating greater graph structure diversity, whereas the Medium class overlaps with other classes. This pattern explains the differences in precision, recall, and F1-score values between classes and shows that the QGNN embedding can meaningfully represent the graph's relational structure, even though class separation is not clear.

The process then proceeds to the fourth scenario (S4), which integrates a QGNN as a learning module for scientometric graph embeddings with a QSVM as a classifier. Unlike S3, which still used a classical SVM for classification, S4 implements a fully quantum-based classification to the embeddings generated by the QGNN. In this scenario, the same institutional graph serves as model input. The QGNN first produces quantum-based graph embeddings through quantum encoding, quantum message passing, and quantum aggregation using PQCs. These embeddings are then mapped into a quantum feature space and classified using a QSVM with a quantum kernel. This approach forms a full quantum-enhanced learning architecture, where both representation learning and classification utilize QML principles. Scenario S4 is designed to evaluate the potential for improved classification performance when the entire learning pipeline operates within the quantum feature space.

Table IV presents the performance evaluation of the QGNN embedding–QSVM model in the fourth experimental scenario (S4). The evaluation used 30 test samples evenly distributed across three classes—Low, Medium, and High—with 10 samples per class. The model achieved an overall accuracy of 0.73, correctly classifying 22 out of 30 instances. This result demonstrates competitive classification effectiveness, particularly in balanced class discrimination.

TABLE IV. PERFORMANCE EVALUATION OF QGNN EMBEDDING–QSVM

Class	Precision	Recall	F1-score	Support
High	0.900	0.600	0.721	1,272
Low	0.538	0.700	0.609	1,272
Medium	0.733	0.906	0.808	1,272
Accuracy			**0.73**	3,818

For the Low class, the model obtained a precision of 0.900, showing that predictions assigned to this class were highly reliable with minimal false positives. However, the recall of 0.600 indicates that only six out of ten Low instances were correctly identified, suggesting that some Low samples were still confused with adjacent classes. The resulting F1-score of 0.721 reflects a reasonable balance between precision and recall for this class.

The Medium class shows improved performance compared to previous results, achieving a precision of 0.538 and a recall of 0.700, which corresponds to seven correctly classified instances. The F1-score of 0.609 indicates moderate classification stability. Despite this improvement, the Medium class remains the most challenging to distinguish, likely due to overlapping feature representations between the Low and High classes within the quantum embedding space.

The High class demonstrates the strongest performance. With a recall of 0.906, nearly all High instances were correctly identified, confirming the model's effectiveness in capturing dominant graph-structured patterns associated with high scientometric performance. Although the precision of 0.733 suggests the presence of some false positives, the class achieved the highest F1-score of 0.808, indicating robust and consistent classification.

Overall, these findings confirm that the integration of QGNN as a graph embedding learner with QSVM as a quantum kernel-based classifier provides strong discriminatory capability, particularly for the High class, which is often the most critical category in scientometric performance evaluation. The Medium class remains the most challenging to distinguish, highlighting opportunities for future improvements such as refining labeling strategies, incorporating additional indicators, or increasing the representational depth of the QGNN. Figure 6 displays the confusion matrix for the QGNN embedding–QSVM model.

B. Discussion

The experimental results provide a comparative evaluation of different modeling strategies for classifying scientometric performance. In the first scenario (S1), the classical SVM baseline relies on numerical scientometric indicators such as citation counts, h-index values, and SINTA scores. Despite using only feature-based representations, the SVM model demonstrates strong predictive capability, indicating that conventional scientometric indicators already contain substantial information for distinguishing institutional publication performance. However, this approach treats each institution independently and does not explicitly capture structural relationships such as collaboration patterns or institutional connections.

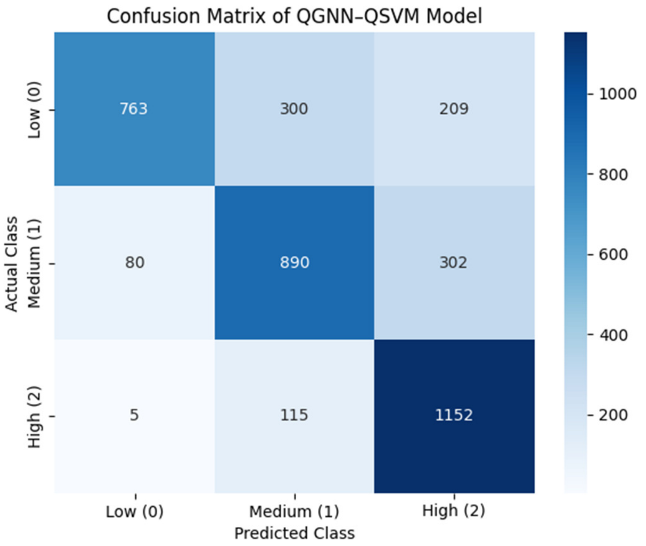

Fig. 6. Confusion Matrix of the QGNN embedding–QSVM model.

In the second scenario (S2), the incorporation of a GNN introduces relational learning by modeling interactions among institutions through graph structures. By leveraging graph embeddings, the model is able to represent contextual relationships between nodes in the scientometric network. Such representations allow the learning process to incorporate structural information beyond individual features. Previous studies have shown that GNN-based models are particularly effective in handling heterogeneous and network-based data structures where relationships among entities play an important role [21, 22].

The third scenario (S3) further extends the relational modeling capability by employing a QGNN as the graph embedding learner. In this configuration, classical graph representations are transformed into quantum feature spaces using quantum encoding mechanisms. Through quantum superposition and interference principles, QGNN aims to capture complex nonlinear dependencies within the scientometric graph structure. Although the embedding stage becomes more expressive in the quantum representation space, the classification stage still relies on a classical SVM. As a result, the overall predictive performance reflects a hybrid configuration where quantum-based representation learning is combined with classical decision boundaries. Similar hybrid approaches have been explored in recent studies to bridge classical graph learning and emerging QML techniques [23, 24].

In the fourth scenario (S4), the model integrates QGNN-based embeddings with a QSVM, enabling the classification stage to operate directly in a quantum feature space through the use of a quantum fidelity kernel. In this configuration, the model achieves an overall accuracy of 0.73 on the evaluation dataset. The strongest performance is observed in the High class, with a recall of 0.906 and an F1-score of 0.808, indicating that the model is particularly effective in identifying institutions with strong scientometric performance. This capability is important in institutional evaluation contexts where identifying top-performing entities is often a primary objective. Prior research has also suggested that QSVM can be advantageous when dealing with complex and nonlinearly separable data distributions in high-dimensional feature spaces [25, 26].

Despite these promising results, the Medium class remains the most challenging category to classify accurately. The moderate recall value of 0.700 indicates that instances within this class frequently overlap with the neighboring Low and High classes in the learned representation space. This phenomenon can be attributed to the transitional characteristics of institutions whose publication performance lies between lower and higher tiers, resulting in ambiguous feature patterns within the scientometric dataset.

Overall, the experimental analysis highlights the different roles of feature-based learning, graph-based relational modeling, and quantum-enhanced representations in scientometric classification. While classical machine learning methods provide strong baseline performance, graph-based and quantum-inspired approaches offer alternative perspectives for capturing relational structures and complex dependencies within scientometric data. These findings suggest that hybrid graph–quantum learning frameworks remain a promising research direction for future scientometric analysis. Further improvements may be achieved by refining the labeling strategy, incorporating additional scientometric indicators, and increasing the representational capacity of the QGNN architecture.

IV. CONCLUSION

This research develops and evaluates a graph-based scientometric performance classification framework by integrating graph learning and quantum computing techniques. Using a dataset consisting of 3,818 scientometric records, four experimental scenarios were designed to investigate the roles of classical machine learning, graph-based learning, and quantum-enhanced classification. The experimental results show that the combination of a Quantum Graph Neural Network (QGNN) and a Quantum Support Vector Machine (QSVM) is capable of classifying institutional publication performance with moderate accuracy. Although the achieved accuracy is not yet superior to classical baselines, the results demonstrate the feasibility of applying quantum-enhanced graph learning to model complex relational patterns in scientometric data.

The main contribution of this research lies in the development and empirical evaluation of a hybrid QGNN–QSVM framework for graph-based scientometric analysis, a topic that remains relatively underexplored in the literature.

The study also provides insights into the practical implementation of Quantum Machine Learning (QML) under current Noisy Intermediate-Scale Quantum (NISQ) constraints. These findings open opportunities for future work, including architectural optimization of the QGNN model, incorporation of richer scientometric indicators, and experimentation with more advanced quantum computing environments.

DECLARATION OF COMPETING INTERESTS

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

ACKNOWLEDGMENT

The authors would like to thank Universitas Diponegoro and Universitas Muhammadiyah Makassar for supporting this research. The authors also acknowledge the SINTA portal for providing access to scientometric data used in this study.

DATA AVAILABILITY

The scientometric data used in this study were obtained from the publicly accessible SINTA portal (https://sinta.kemdiktisaintek.go.id). The processed dataset and preprocessing scripts used in this research are available from the authors upon reasonable request.

AI USE AND DECLARATION OF GENERATIVE AI USE

The authors declare that generative Artificial Intelligence (AI) tools were used in this study to assist in language editing and improving the clarity of the manuscript. The authors take full responsibility for the content of the publication.

REFERENCES

[1] F. Mayta-Tovalino *et al.*, "Scientometric evaluation of trends and global characteristics of published research on occupational public health," *Heliyon*, vol. 8, no. 12, Dec. 2022, Art. no. e12165, https://doi.org/10.1016/j.heliyon.2022.e12165.

[2] N. Scelles and J. A. Teixeira da Silva, "Making the impact of publications within a field comparable by improving the field-weighted citation impact (FWCI): the case of sport management," *Scientometrics*, vol. 130, no. 3, pp. 1571–1586, Mar. 2025, https://doi.org/10.1007/s11192-025-05268-4.

[3] C. García-Villar and J. M. García-Santos, "Bibliometric indicators to evaluate scientific activity," *Radiología (English Edition)*, vol. 63, no. 3, pp. 228–235, May 2021, https://doi.org/10.1016/j.rxeng.2021.01.002.

[4] F. Emmert-Streib, S. Tripathi, A. Farea, and O. Yli-Harja, "Advanced Scientometric Analysis of Scientific Machine Learning and PINNs: Topic Modeling and Trend Analysis," *IEEE Access*, vol. 12, pp. 153253–153272, 2024, https://doi.org/10.1109/ACCESS.2024.3481671.

[5] M. Daradkeh, L. Abualigah, S. Atalla, and W. Mansoor, "Scientometric Analysis and Classification of Research Using Convolutional Neural Networks: A Case Study in Data Science and Analytics," *Electronics*, vol. 11, no. 13, June 2022, Art. no. 2066, https://doi.org/10.3390/electronics11132066.

[6] P. Lamichhane and D. B. Rawat, "Quantum Machine Learning: Recent Advances, Challenges, and Perspectives," *IEEE Access*, vol. 13, pp. 94057–94105, 2025, https://doi.org/10.1109/ACCESS.2025.3573244.

[7] F. Rodríguez-Díaz, D. Gutiérrez-Avilés, A. Troncoso, and F. Martínez-Álvarez, "A Survey of Quantum Machine Learning: Foundations, Algorithms, Frameworks, Data and Applications," *ACM Computing Surveys*, vol. 58, no. 4, Oct. 2025, Art. no. 91, https://doi.org/10.1145/3764582.

[8] D. Ranga, A. Rana, S. Prajapat, P. Kumar, K. Kumar, and A. V. Vasilakos, "Quantum Machine Learning: Exploring the Role of Data Encoding Techniques, Challenges, and Future Directions," *Mathematics*, vol. 12, no. 21, Oct. 2024, Art. no. 3318, https://doi.org/10.3390/math12213318.

[9] L. Anas, A. Sofwan, and I. Setiawan, "Talented Researcher Identification Model Based on Scientometric Indicators with Quantum Machine Learning (QML) Approach," in *International Conference on Data Analytics and Management*, London, UK, 2025, pp. 340–350, https://doi.org/10.1007/978-3-032-03072-6_27.

[10] L. Xu, X. Zhang, M. Li, and S. Shen, "Quantum support vector machine for multi classification," *Communications in Theoretical Physics*, vol. 76, no. 7, June 2024, Art. no. 075105, https://doi.org/10.1088/1572-9494/ad48fc.

[11] D. Alaminos, M. B. Salas, and M. A. Fernández-Gámez, "Quantum Computing and Deep Learning Methods for GDP Growth Forecasting," *Computational Economics*, vol. 59, no. 2, pp. 803–829, Feb. 2022, https://doi.org/10.1007/s10614-021-10110-z.

[12] J. Zhou *et al.*, "Graph neural networks: A review of methods and applications," *AI Open*, vol. 1, pp. 57–81, Jan. 2020, https://doi.org/10.1016/j.aiopen.2021.01.001.

[13] X.-M. Zhang, L. Liang, L. Liu, and M.-J. Tang, "Graph Neural Networks and Their Current Applications in Bioinformatics," *Frontiers in Genetics*, vol. 12, July 2021, Art. no. 690049, https://doi.org/10.3389/fgene.2021.690049.

[14] G. Verdon, T. McCourt, S. Leichenauer, E. Luzhinca, V. Singh, and J. Hidary, "Quantum Graph Neural Networks," in *Second Workshop on Machine Learning and the Physical Sciences*, Vancouver, Canada, 2019, https://doi.org/10.48550/arXiv.1909.12264.

[15] S. Piperno, A. Ceschini, S. Y. Chang, M. Grossi, S. Vallecorsa, and M. Panella, "A study on quantum graph neural networks applied to molecular physics," *Physica Scripta*, vol. 100, no. 6, May 2025, Art. no. 065126, https://doi.org/10.1088/1402-4896/add8e9.

[16] N. Donthu, S. Kumar, D. Mukherjee, N. Pandey, and W. M. Lim, "How to conduct a bibliometric analysis: An overview and guidelines," *Journal of Business Research*, vol. 133, pp. 285–296, Sept. 2021, https://doi.org/10.1016/j.jbusres.2021.04.070.

[17] R. Kumar, "Bibliometric Analysis: Comprehensive Insights into Tools, Techniques, Applications, and Solutions for Research Excellence," *Spectrum of Engineering and Management Sciences*, vol. 3, no. 1, pp. 45–62, Jan. 2025, https://doi.org/10.31181/sems31202535k.

[18] Y. Dong, R. A. Johnson, and N. V. Chawla, "Can Scientific Impact Be Predicted?," *IEEE Transactions on Big Data*, vol. 2, no. 1, pp. 18–30, Mar. 2016, https://doi.org/10.1109/TBDATA.2016.2521657.

[19] Y. A. Alohali, M. S. Fayed, T. Mesallam, Y. Abdelsamad, F. Almuhawas, and A. Hagr, "A Machine Learning Model to Predict Citation Counts of Scientific Papers in Otology Field," *BioMed Research International*, vol. 2022, July 2022, Art. no. 2239152, https://doi.org/10.1155/2022/2239152.

[20] V. M. Venkatappa and Venkateshappa, "A Novel Quantum Convolution Neural Network for Image Classification Applications," *Engineering, Technology & Applied Science Research*, vol. 15, no. 5, pp. 27046–27051, Oct. 2025, https://doi.org/10.48084/etasr.12416.

[21] T. Hoang, "GNN: Graph Neural Network and Large Language Model for Data Discovery." arXiv, Aug. 27, 2024, https://doi.org/10.48550/arXiv.2408.13609.

[22] B. Khemani, S. Patil, K. Kotecha, and S. Tanwar, "A review of graph neural networks: concepts, architectures, techniques, challenges, datasets, applications, and future directions," *Journal of Big Data*, vol. 11, no. 1, Jan. 2024, Art. no. 18, https://doi.org/10.1186/s40537-023-00876-4.

[23] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, "Quantum machine learning," *Nature*, vol. 549, no. 7671, pp. 195–202, Sept. 2017, https://doi.org/10.1038/nature23474.

[24] M. Schuld, I. Sinayskiy, and F. Petruccione, "An introduction to quantum machine learning," *Contemporary Physics*, vol. 56, no. 2, pp. 172–185, Apr. 2015, https://doi.org/10.1080/00107514.2014.964942.

[25] T. Yin, "Quantum support vector machines: theory and applications," *Theoretical and Natural Science*, vol. 51, pp. 34–42, Nov. 2024, https://doi.org/10.54254/2753-8818/51/2024CH0158.

[26] D. Sharma, V. B. Sabale, P. Singh, and A. Kumar, "Harnessing quantum support vector machines for cross-domain classification of quantum states," *Quantum Machine Intelligence*, vol. 7, no. 1, Apr. 2025, Art. no. 49, https://doi.org/10.1007/s42484-025-00274-4.