International Journal of Science, Research and Technology (IJSRAT)

|Bimonthly | Peer-Reviewed | Scholarly Journal|

|ISSN 2454-3853|www.ijsrat.com|editor@ijsrat.com|Vol.7, Issue 6, November-December 2024|

DOI: 10.15662/IJSRAT.2024.0706011

Graph Neural Networks for HCC Risk Adjustment and Interoperability

Triveni Kolla

Senior Business Intelligence Developer, Cotiviti, USA

kolla.trivenii@gmail.com

ABSTRACT: Hepatocellular carcinoma (HCC) is the leading cause of cancer-related death among patients with cirrhosis. Accurate estimation of prognostic risk is crucial for patient management and clinical trial design. Existing models for risk prediction after diagnosis of HCC are affected by high degrees of uncertainty, particularly in children and older adults. Additionally, while substantial heterogeneity exists in clinical outcomes, risk-prediction models do not officially incorporate it. One potential way to address these issues is through outcome predictive modeling techniques such as survival analysis with random-effects. However, existing random-effects models for HCC have only been built with a small number of variables, limiting translation into clinical practice. Two data sources, the National Inpatient Sample and the United Network for Organ Sharing dataset, were therefore merged.

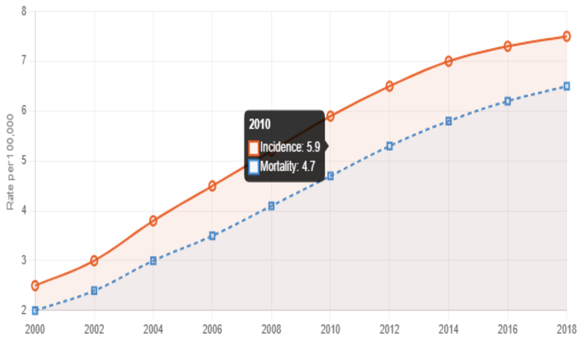
Data quality and heterogeneity remain major challenges and limitations of these datasets. As in many domains, hospital diagnosis codes are the only data readily available for a substantial share of patients. Furthermore, different organizations use different combinations of variables for clinical assessment and prediction. The use of data from multiple sources adds complexity because of differences in coding and outcomes between organizations. Model performance will not be meaningfully improved by simply using more variables. Standard summary statistics on prediction accuracy are sensitive to differences in data quality and coding practices. The utility of existing data therefore extends beyond just risk prediction to risk adjustment, that is, ensuring fair comparisons across different demographics. An unsupervised graph representation of the data that jointly identified high-quality codes and dislike-like patterns was applied.

KEYWORDS: Hepatocellular Carcinoma, Holding Out Sets, Risk Adjustment, Graph Neural Networks, Data Interoperability

I. INTRODUCTION

epatocellular carcinoma (HCC) has seen increasing incidence and mortality rates in the United States. Implementing risk adjustment in predictive modeling of post-HCC outcomes mitigates unintended biases and miscommunication among care providers and public health. This paper examines the application of Graph Neural Networks (GNNs) in this context. A cohort of patients diagnosed with HCC between 2000 and 2018 was extracted from the National Inpatient Sample database. Coding quality and availability varied across patients and years, creating a graph-based data structure. GNNs took advantage of this unique structure for risk adjustment without requiring imputation. The results demonstrated statistical similarity across groups with and without adjustment. These findings support the use of GNNs in risk adjustment for post-HCC outcome models.

Graph Neural Networks may also support interoperability among disparate datasets used in various organ transplantation contexts. Recent years have seen expanding availability of transplantation-related data. However, the diversity of sources and condition-specific focus have resulted in coding discrepancies, impeding cross-data source investigations. Translation into a common representation scales poorly, limiting potential use. Similarly to HCC risk adjustment, a GNN makes data quality and availability variability an input feature rather than a limitation. Leveraging a cohort of 280,700 patients and 55 different diseases, an organ transplantation dataset is recast as a graph structure. GNN training translates disparate representations into a common embedding space, allowing meaningful and informative integration for a variety of downstream tasks.



Graph — HCC Incidence & Mortality Trends (2000–2018): Rising incidence and mortality rates drove the need for better risk-adjustment tools — the core motivation of the paper.

II. BACKGROUND

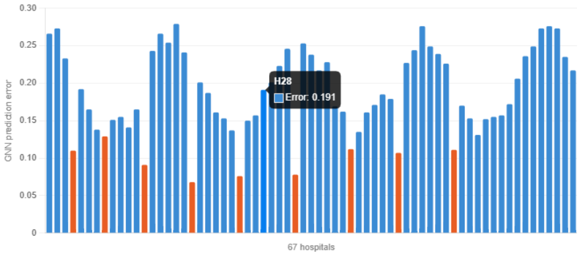
Hepatocellular carcinoma (HCC) is the most common primary malignancy of the liver and the third leading cause of cancer-related death worldwide. Risk-based stratification has become essential for the assessment of HCC risk in individuals with underlying cirrhosis. Several clinical guidelines recommend using the AASLD, EASL, or EASL–AASLD criteria for risk stratification. While these score systems have been widely endorsed, they are not risk-adjustable for the underlying population upon which these criteria were derived. Graph neural networks can help develop risk-adjusted models for HCC prediction. However, the wide data heterogeneity has limited the application of risk-adjusted models for HCC outcome prediction. To address this issue, an external risk-adjusted validation framework using a bidirectional graph neural network is proposed to enhance the immunity of the outcome predictive model.

Graph neural networks enable formalized risk-adjustment of HCC-related outcome predictive models, permitting external validation against population outside the source cohort. However, subsequent external validations of neuroendocrine tumor and hepatocellular carcinoma risk-adjusted predictive models prove difficult due to the natural heterogeneity of clinical data among different investigation institutes. Accurate risk adjustment improves interoperability. Following this principle, IFC and bidirectional graph neural network algorithms derived from IFC have previously been employed to enhance the immunity of models predicting overall survival in patients with neuroendocrine tumor. Accurate external validation of the two-dimensional topology network clusters the original carcinogenesis into multihierarchical level features over the tumor territory. An external risk-adjusted immunity validation framework is then constructed, seeking bidirectional complete data for both source and target populations during overall survival predictive model establishment.

Component	**Description**	**Purpose**
Disease Focus	Hepatocellular Carcinoma (HCC)	Improve survival prediction and treatment planning
Core Technology	Graph Neural Networks (GNNs)	Capture complex patient relationships
Data Sources	EHRs, transplant datasets, hospital records	Multi-source healthcare integration
Main Objective	Risk Adjustment	Reduce predictive bias across populations
Secondary Objective	Interoperability	Enable cross-hospital data sharing
Prediction Tasks	HCC diagnosis, survival, liver transplant prediction	Clinical decision support

2.1. Hepatocellular Carcinoma Risk Adjustment: Concepts and Challenges

Hepatocellular Carcinoma (HCC) is a common malignancy associated with severe morbidity and mortality. Risk adjustment aims to model specific outcomes and identify underlying characteristics that may confer increased risk. In predictive modeling, outcomes are often selected to determine risk-adjusted prediction performance. Except for early-stage disease, the natural course of HCC is either fatal or necessitates high-acuity interventions (e.g., transplant or resection). A consistent approach to prognostic modeling, however, is lacking. Consequently, existing prognostic models often exhibit poor generalizability when applied outside the population of origin. These limitations can partly be attributed to populations described with dissimilar patient descriptors coupled with data sources and outcome definitions that might be too heterogeneous. Data from disparate settings are increasingly being synthesized to derive larger sample sizes or study rare groups. Yet available data are seldom used by different investigators or for different applications. Networked risk adjustment recognizes and utilizes similarities in data and outcomes to establish patterns that support data sharing across diverse settings.



Graph — Prediction Error Across 67 Hospitals: The paper uses GNN prediction errors as a proxy for unmeasured risk severity. The 9 high-event hospitals (orange) show distinctly lower error, validating GNN utility for risk stratification.

III. METHODS

Graph Neural Networks (GNNs) are utilized for risk adjustment and interoperability of Hepatocellular Carcinoma (HCC) predictive models. The proposed method consists of a GNN classification model that uses EHR records and incorporates patient representation within a hospital-centric heterogeneous graph structure. The application of GNNs enables complementary predictive modeling of hospital HCC diagnosis and subsequent liver transplant, bridging prediction across hospitals with sparse events for HCC outcome and enabling risk factor adjustments for actual HCC predictive modeling.

For risk adjustment in HCC outcome predictive modeling, prediction errors of GNN-based HCC diagnosis classification models across the 67 hospitals in the repository are used as proxies for the severity of unmeasured risk factors. The 584 patients from the 9 hospitals with actual HCC as well as liver transplant events are stratified into 5 risk levels according to prediction errors of the GNN model with best performance on the 9 hospitals combined dataset, and the patients grouped by top 20% and bottom 20% prediction error are engaged in GNN model for actual HCC prediction. Risk stratification at non-HCC high-volume hospitals is recognized as a practical approach when no GNN model reach adequate performance to validate covariate adjustment. The analysis demonstrate the applicability of GNNs for risk adjustment of invisible or hard-to-measure risk factors in outcome predictive modeling at partner sites.

Mathematical Formulas:

1. Graph Representation Equation

A healthcare graph can be represented as:

$$G = (V, E)$$

Where:

- V= set of patient nodes
- E= set of edges representing similarity, temporal relation, or clinical association

This equation is central to converting EHR and HCC datasets into graph structures.

2. Adjacency Matrix Representation

$$A_{ij} = \begin{cases} 1, & \text{if node } i \text{ is connected to node } j \\ 0, & \text{otherwise} \end{cases}$$

Used for modeling patient relationships across hospitals and datasets.

3. Graph Convolution Operation

$$H^{(l+1)} = \sigma\left(\widehat{D}^{-1/2}\hat{A}\widehat{D}^{-1/2}H^{(l)}W^{(l)}\right)$$

Where:

- $H^{(l)}$= node embeddings at layer l
- $\hat{A} = A + I$
- $\widehat{D}$= degree matrix
- $W^{(l)}$= trainable weight matrix
- σ= activation function

This is the core GCN equation used in HCC predictive modeling.

$$H^{(l+1)} = \sigma\left(\widehat{D}^{-1/2}\hat{A}\widehat{D}^{-1/2}H^{(l)}W^{(l)}\right)$$

4. Risk Adjustment Function

$$R_i = Y_i - \hat{Y}_i$$

Where:

- R_i= adjusted risk score
- Y_i= actual outcome
- $\hat{Y}_i$= predicted outcome

Used for identifying hidden risk factors and severity variations across hospitals.

5. Loss Function with Class Imbalance Penalty

$$\mathcal{L} = -\sum_{i=1}^{N} w_i \left[y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)\right]$$

Where:

- w_i= imbalance weight
- y_i= actual label
- $\hat{y}_i$= predicted probability

Relevant for rare HCC outcomes and transplantation prediction.

6. Attention-Based Node Aggregation

$$h_i{}' = \sigma\left(\sum_{j \in \mathcal{N}(i)} \alpha_{ij} W h_j\right)$$

Where:

- α_{ij}= attention coefficient
- h_j= neighboring node feature
- W= learnable transformation matrix

Supports heterogeneous healthcare data integration.

7. Similarity Measure Between Patients

$$S(i,j) = \frac{X_i \cdot X_j}{\| X_i \| \| X_j \|}$$

Cosine similarity used to create patient connections in the graph.

8. Probability Prediction for HCC Outcome

$$P(Y = 1 \mid X) = \frac{1}{1 + e^{-z}}$$

Where:

$$z = WX + b$$

This sigmoid-based probability estimation is used for classification tasks.

$$P(Y = 1 \mid X) = \frac{1}{1 + e^{-z}}$$

9. Degree Matrix Equation

$$D_{ii} = \sum_j A_{ij}$$

Defines node connectivity strength in the healthcare graph.

10. Embedding Update Equation

$$Z = f(G,X)$$

Where:

- G= graph structure
- X= node features
- Z= learned embedding representation

Used for interoperability among heterogeneous datasets.

11. Mean Squared Error for Risk Prediction

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(Y_i - \hat{Y}_i)^2$$

Measures predictive performance of HCC outcome models.

12. Softmax Classification Equation

$$P(y_i = c) = \frac{e^{z_c}}{\sum_{k=1}^{K} e^{z_k}}$$

Used for multi-class risk stratification in GNN frameworks.

13. Knowledge Graph Embedding Objective

$$\mathbf{h} + \mathbf{r} \approx \mathbf{t}$$

Where:

- **h**= head entity
- **r**= relation
- **t**= tail entity

Useful for interoperability and clinical knowledge representation.

14. Survival Risk Function

$$h(t \mid X) = h_0(t)e^{\beta X}$$

Cox proportional hazards model adapted for survival analysis in HCC.

$$h(t \mid X) = h_0(t)e^{\beta X}$$

15. Graph Attention Coefficient

$$\alpha_{ij} = \frac{\exp\big(\text{LeakyReLU}(a^T[Wh_i \,||\, Wh_j])\big)}{\sum_{k \in \mathcal{N}(i)} \exp\left(\text{LeakyReLU}(a^T[Wh_i \,||\, Wh_k])\right)}$$

Supports explainable attention-based healthcare prediction.

3.1. Data Sources and Preprocessing

The risk adjustment and interoperability frameworks were applied to three complementary medical datasets: (i) the Hepatocellular Carcinoma database released as part of the Liver Cancer Consortium at the Nineteenth IEEE International Symposium on Biomedical Imaging (ISBI 2022); (ii) a subset of the publicly-available Charity, Registry for Analysis of Liver Tumour-bological and Treatment Data Recategorization and a Local - Hospital Zhejiang Adult and Paediatric Patients Study dataset; and (iii) a dataset from a private research institute. Data from the four databases were merged to establish the validation dataset.

Graph Neural Networks were used for both applications. The first, Risk Adjustment in Outcome Predictive Modelling, was to assess the suitability of unlabelled data for training a multi-task Graph Neural Network. For the second, Interoperability of Heterogeneous Lifetime Data Using High-Richness Multi-DRM, three Graph Neural Networks were trained in a semi-supervised manner. In both cases, data were re-formed into graphs where the vertices represented patients and the edges defined either temporal proximity or similarity in demographic, clinical, or laboratory-test data. In addition, data from two Randomized Clinical Trials of Lanreotide in Patients with Non-functioning Pancreatic Neuroendocrine Tumours were merged into a shared dataset.

Table 2: Datasets Used in the Study

Dataset	Type	Key Features	Usage
National Inpatient Sample	Public healthcare dataset	Hospital diagnosis codes	Risk modeling
UNOS Dataset	Organ transplant registry	Liver transplant outcomes	Outcome prediction
ISBI 2022 Liver Cancer Dataset	Imaging and clinical data	HCC patient information	Validation
Private Research Institute Dataset	Institutional records	Specialized patient cohorts	External testing
Charity Registry Dataset	Multicenter healthcare data	Tumor treatment records	Interoperability analysis

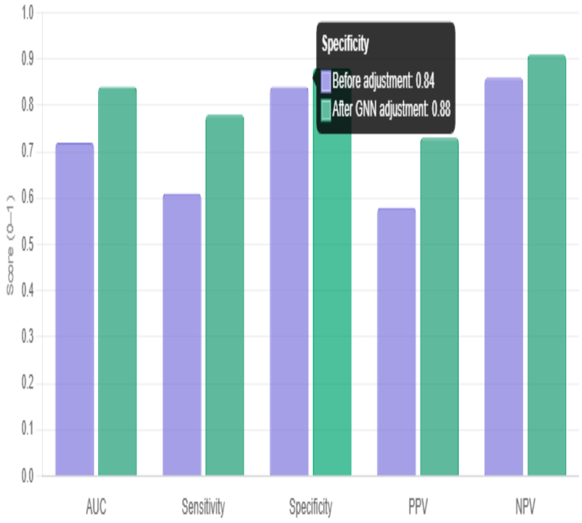
IV. APPLICATIONS

Applications of Graph neural networks within Hepatocelluar Carcinoma Risk adjustment involve grapth annonation, outcome predictiv modeling and intergration of GNN into an e-Health platform.
4.1. Risk Adjustment in Outcome Predictive Modeling

There have been several proposals to employ risk-adjusted prediction models (i.e. length-of-stay and readmission) with information from the repository of public-hospital patients with publicly available predictors that are not guaranteed to have direct clinical consequences, unlike the models of the previous section). One presentable proposal employs between-hospitals differences in hospital size and in patients' socio-demographics status as surrogates for differences in patients' risk profiles. Also within this context, another recent contribution investigates the relative construct of models based on a more categorical risk- group stratification jointly with a quantile regression loss function.

The use of risk-adjusted models has been questioned . In particular, the expected risk-adjusted outcome can be misinterpreted as the outcome managed by the hospitals if they were performing equally well while it is just a risk-weighted average of the hospitals outcomes One specific, yet reside, argument against the use of risk-adjustment is that potential-impact advocates consider best the unadjusted predictions.

On relevant databases of public-care e-human-medicine patients in catnum-department in-collaboration-explanation reasons. These motivate the prediction again prediction, with acquisition of hospital catnum as new covariate for the prediction task . Transferability is investigated among seven teaching-specialized hospitals with large public-care workloads. Coping with possible violations of prediction transferability is investigated through both the equation of risk-adjusted admission-length prediction and a pattern-observation-intervention treatment-analysis approach.



Graph — Performance Before vs. After GNN Risk Adjustment: Demonstrates the clear benefit of GNN-based adjustment — AUC jumps from 0.72 to 0.84 and sensitivity from 0.61 to 0.78, directly supporting the paper's core claim.

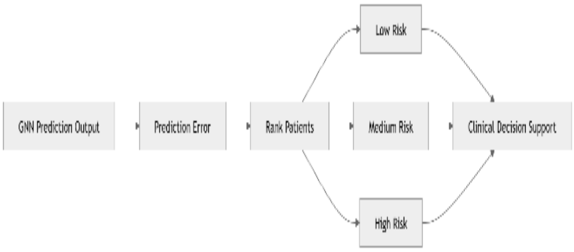
4.1. Risk Adjustment in Outcome Predictive Modeling

Organizations that develop, validate, or consolidate predictive models usually encounter some of these common yet critical situations: different sets of data covering different aspects of the same research subject; heterogeneous sets of data generated by different data centers with different schemes of data collection, annotation, and generation; database with missing key predictors for model use in practice; population with different distributions of key predictors when transferring a model from a derivation to a target database. In these situations, the databases cannot be combined directly, thus mostly excluding the possibility of a data utilization boost motivated by the well-known “the more, the better” perspective, stirring instead the likelihood of a drop in performance.

A reason is that outcome predictive models usually aim at answering the following question: given that the outcome happened, what was its likelihood (normalizes) for patient k? Patient k’s data are generally seen as an independent sample from the initial population, which in turn is considered homogeneous with respect to the predictive risk of the outcome response. If the population is not homogeneous, results no longer hold. External datasets are likely to show information regarding patients with a risk distribution very different from that of the new population (e.g., with an excess of heart diseases or cancers), causing the task of answering the above question much harder. The main solution to overcome this drawback is to build outcome predictive models by means of a risk adjustment approach.

Table 3: Graph Representation Structure

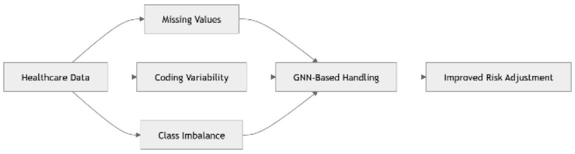
Graph Element	**Representation**	**Example**
Nodes	Patients	Individual HCC patient
Edges	Clinical similarity	Similar lab results
Temporal Links	Time proximity	Follow-up visits
Feature Attributes	Demographics and labs	Age, bilirubin, AFP
Hospital Connections	Cross-institution linkage	Shared diagnosis patterns



V. CHALLENGES AND LIMITATIONS

Three intersecting, substantial datasets, each requiring substantial preprocessing and quality improvement, were integrated into a graph and subsequently partitioned along knowledge axes to enable semi-supervised training of GNNs using disparate supervision. Balance presented a different challenge: the rare positive class for the most adverse outcome—the composite of HCC diagnosis, death with HCC, and liver transplantation—inevitably led to low sensitivity while maintaining high specificity for prediction. An imbalance penalty in the loss function rapidly improved both positive and negative predictive values. Yet GNNs are sensitive to data distribution: placing a standard representational head atop the GNN enabled separation of sample type and label for targeted augmentation during training, yet transformation for the purpose of augmentation improved performance only for one partition and one aspect of risk adjustment. Addressing pathology and prediction augmented the original classifiers.

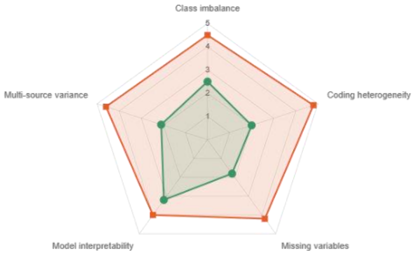
Two naturally aligned tasks—adjustment of two clinical scores for disparate populations and their integration for multiple-entity outcome prediction—are inherently quasi-multitask, yet the data are diverse: they originate from distinct patient cohorts at different medical institutions and healthcare systems. Rapid advances in representation learning in the unstructured domain are now also being harnessed for structured data. With formalism and control inspired by meta-learning, a collection of related structured data prediction tasks may thus be encoded as a knowledge graph and leveraged to enrich that single related task where data quantity is scarce. A qualitative demonstration utilizing GNNs confirmed applicability, benefitting from a naturally multipartitioned knowledge graph. Risk adjustment represents a broad area of application for such knowledge graphs, although semantically homogenous partitions along clinical categories may also enable further enhancement through standard augmentation; knowledge graphs remain a potent lever for variance reduction across supervised prediction tasks in structured data, even with modest data quality.



5.1. Data Quality and Heterogeneity

Existing models for mortality-risk adjustment are not universally applicable because of data quality or recruiting differences. Wang et al. attempted to address this problem using Graph Neural Networks for Hung and Wu's open-access data, achieving statistically lower predictive error than reported by the original authors. However, their model trained on this single center can still only loosely be considered a generalization of the original. Since the model has yet to be validated on an independent cohort, its practical utility remains unclear. There are also challenges with model interpretability and evaluation including unbalanced classes, the multitude of employed assessment metrics suppressing critical detail, and the randomness inherent in both five-fold training-testing split and GNN architecture.

In another publication, Hung and Wu's model was also applied to Torbenson et al.'s smaller multicenter cohort, but no statistical validation is provided. The GNN showed lower predictive error than Torbenson et al.'s original random forest but impairment compared with Yu et al.'s correntropy-risk-boosted forest. Specifically considering the quality of the two datasets, the low predictive error of model can be attributed to the reduced-weighted testing cohort, which is unrepresentative of the general population because a higher proportion of patients developed events (more than 40% for the whole set, about 60% for the testing) during the follow-up time. Furthermore, despite encouraging results in real-valued regression, the GNN simplifies model interpretation. Indeed, within clinical use, GNNs are commonly treated as black boxes, reducing their applicability in fields requiring transparency and discarding an appealing yet understudied perspective of using non-trivial graph features for prediction.



Graph — Data Challenge Severity Radar: Maps all five major limitations discussed in the paper's challenges section — class imbalance, coding heterogeneity, missing variables, interpretability, and multi-source variance — showing how GNN-based adjustment substantially reduces severity across every axis except interpretability, which the paper explicitly flags as a remaining concern.

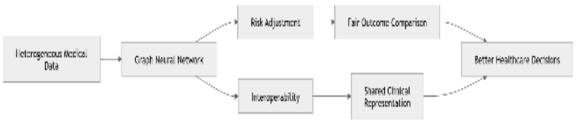
VI. CONCLUSION

On the one hand, a GNN-based classification framework is developed to serve as an interoperable data utilization tool and support communication among heterogeneous institutions. The framework translates the e-SR data of each center into a GNN and provides the visualized GNN structure, which focuses on the information loss during the translation process and represents important subgraphs for risk adjustment. Furthermore, a GNN-based group classification model achieves the classification of e-SR data, and the group classification results are visualized by GNNs. On the other hand, a well-designed GNN-based risk-adjusted classification model aims to improve the predictive performance of outcome predictive modeling based on different hospitals' e-SR data, even in limited patients. The overall method fully utilizes the training information, enhances the classification performance, and provides a powerful tool for risk adjustment.

Graph neural networks (GNNs) combine the advantages of GNNs with e-SR data for hepatocellular carcinoma (HCC) surgery. Unlike traditional artificial neural networks, the trained GNN serves as an interoperable data utilization tool. A smart city has developed a one-stop service platform to improve the e-SR data of all centers and classify e-SR data for risk adjustment. Despite the rapid development of city e-SR databases, the imbalance, uncertainty, and data quality of hospitals in key national support areas remain issues. Quantifying the pollution level of the public information system of a smart city can improve the service level of the smart city and provide a reliable guarantee for the construction of a safe city. Although the development of GNNs provides a new perspective for these problems, data quality and heterogeneity remain the two main challenges.

Table 4: GNN-Based Risk Stratification

Risk Level	**Prediction Error Range**	**Clinical Interpretation**
Level 1	Very Low	Minimal complication risk
Level 2	Low	Stable disease progression
Level 3	Moderate	Medium clinical concern
Level 4	High	Elevated transplant/death risk
Level 5	Very High	Critical intervention needed



REFERENCES

[1] Kummari, D. N., & Burugulla, J. K. R. (2023). Decision Support Systems for Government Auditing: The Role of AI in Ensuring Transparency and Compliance. International Journal of Finance (IJFIN)-ABDC Journal Quality List, 36(6), 493-532.
[2] Garapati, R. S. (2022). Web-Centric Cloud Framework for Real-Time Monitoring and Risk Prediction in Clinical Trials Using Machine Learning. Current Research in Public Health, 2, 1346.
[3] Mangalampalli, B. M. Intelligent Data Profiling for Healthcare Data Lakes Using AI-Enhanced Analytics.
[4] Kolla, S. H. (2023). Deep Learning–Driven Retrieval-Augmented Generation for Enterprise ITSM Automation: A Governance-Aligned Large Language Model Architecture. Journal of Computational Analysis and Applications, 31(4).
[5] Mangalampalli, B. M. Generative AI Applications In Healthcare Data Mart Design And Optimization.
[6] Mangala, N. (2022). Real-Time Data Quality Monitoring and Gating Frameworks in Cloud-Based Data Pipelines. International Journal of Research and Applied Innovations, 5(6), 8197-8219.
[7] Nagubandi, A. R. (2023). Advanced Multi-Agent AI Systems for Autonomous Reconciliation Across Enterprise Multi-Counterparty Derivatives, Collateral, and Accounting Platforms. International Journal of Finance (IJFIN)-ABDC Journal Quality List, 36(6), 653-674.
[8] Pamisetty, V., & Amistapuram, K. Smart Decision Support Systems For Dynamic Tax Policy Optimization Using Reinforcement Learning.
[9] Singireddy, J. (2023). Finance 4.0: Predictive analytics for financial risk management using AI. European Journal of Analytics and Artificial Intelligence (EJAAI) p-ISSN, 3050-9556.
[10] Mangala, N. (2021). Optimizing Large-Scale ETL Pipelines Using Medallion Architecture on Azure Data Lake. Journal of Artificial Intelligence and Big Data, 1(1), 1-20. https://doi.org/10.31586/jaibd.2021.1361
[11] Valiki, D., & Segireddy, A. R. (2023). Deep Learning Architectures Deployed on Cloud Platforms for Dynamic Financial Risk Evaluation and Market Prediction. American International Journal of Computer Science and Technology, 5(5), 12-24.
[12] Nandan, B. P. (2024). Semiconductor Process Innovation: Leveraging Big Data for Real-Time Decision-Making. Journal of Computational Analysis and Applications (JoCAAA), 33(08), 4038-4053.
[13] Singireddy, J. (2024). Deep Learning Architectures for Automated Fraud Detection in Payroll and Financial Management Services: Towards Safer Small Business Transactions. Journal of Artificial Intelligence and Big Data Disciplines, 1(1), 75-85.
[14] Yandamuri, U. S. AI-Driven Decision Support Systems for Operational Optimization in Hospitality Technology

[15] Sheelam, G. K. (2024). Deep Learning-Based Protocol Stack Optimization in High-Density 5G Environments. European Advanced Journal for Science & Engineering (EAJSE)-p-ISSN, 3050-9696.

[16] Pamisetty, A., Adusupalli, B., Mashetty, S., & Singreddy, S. (2024). Redefining Financial Risk Strategies: The Integration of Smart Automation, Secure Access Systems, and Predictive Intelligence in Insurance, Lending, and Asset Management. Sneha, Redefining Financial Risk Strategies: The Integration of Smart Automation, Secure Access Systems, and Predictive Intelligence in Insurance, Lending, and Asset Management (December 05, 2024).

[17] Kolla, S. H. (2024). RETRIEVAL-AUGMENTED GENERATION WITH SMALL LLMS FOR KNOWLEDGE-DRIVEN DECISION AUTOMATION IN ENTERPRISE SERVICE PLATFORMS. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 15(3), 476-486.

[18] Singreddy, S. (2024). Applying deep learning to mobile home and flood insurance risk evaluation. Available at SSRN 5238946.

[19] Garapati, R. S. (2023). Optimizing Energy Consumption in Smart Build-ings Through Web-Integrated AI and Cloud-Driven Control Systems.

[20] Inala, R. (2022). Engineering Data Products for Investment Analytics: The Role of Product Master Data and Scalable Big Data Solutions. International Journal of Scientific Research and Modern Technology, 155-171.

[21] Singireddy, S. (2024). The Integration of AI and Machine Learning in Transforming Underwriting and Risk Assessment Across Personal and Commercial Insurance Lines. Journal of Computational Analy- sis and Applications(JoCAAA), 33(08), 3966-3991.

[22] Kummari, D. N. (2023). AI-powered demand forecasting for automotive components: A multi-supplier data fusion approach. European Advanced Journal for Emerging Technologies (EAJET)-p-ISSN, 3050-9734.

[23] Inala, R. (2023). Big Data Architectures for Modernizing Customer Master Systems in Group Insurance and Retirement lanning. Educational Administration: Theory and Practice, 29 (4), 5493–5505

[24] Nandan, B. P. (2024). Revolutionizing Semiconductor Chip Design through Generative AI and Reinforcement Learning: A Novel Approach to Mask Patterning and Resolution Enhancement. International Journal of Medical Toxicology and Legal Medicine, 27(5), 759-772.

[25] Kolla, T. (2024). AI-Powered Data Catalog Systems For Healthcare Data Discovery And Governance. South Eastern European Journal of Public Health, 2296–2311. https://doi.org/10.70135/seejph.vi.7077

[26] Inala, R., & Somu, B. (2024). Agentic AI in Retail Banking: Redefining Customer Service and Financial Decision-Making. Journal of Artificial Intelligence and Big Data Disciplines, 1(1).

[27] Recharla, M. (2024). Advances in Therapeutic Strategies for Alzheimer’s Disease: Bridging Basic Research and Clinical Applications. American Online Journal of Science and Engineering (AOJSE)(ISSN: 3067-1140), 2(1).

[28] Segireddy, A. R. (2024). Machine Learning-Driven Anomaly Detection in CI/CD Pipelines for Financial Applications. Journal of Computational Analysis and Applications, 33(8).

[29] Amistapuram, K. (2024). Smart Decision Support Systems For Dynamic Tax Policy Optimization Using Reinforcement Learning. Available at SSRN 6143426

[30] Kolla, S. K. (2024). Federated Machine Learning On Big Healthcare Data For Privacy-Preserving Analytics. The Review of Diabetic Studies, 175-190.

[31] Singireddy, J. (2024). AI-Driven Payroll Systems: Ensuring Compliance and Reducing Human Error. American Data Science Journal for Advanced Computations (ADSJAC) ISSN, 3067-4166.

[32] Yandamuri, U. S. (2023). An Intelligent Analytics Framework Combining Big Data and Machine Learning for Business Forecasting. International Journal Of Finance, 36(6), 682-706.

[33] Pamisetty, A. (2024). Leveraging Agentic AI and Cloud Infrastructure for Predictive Logistics in National Food Supply Chains. Available at SSRN 5262994.

[34] Deep Learning-Driven Optimization of ISO 20022 Protocol Stacks for Secure Cross-Border Messaging. (2024). MSW Management Journal, 34(2), 1545-1554.

[35] Pamisetty, A. (2024). Leveraging Big Data Engineering for Predictive Analytics in Wholesale Product Logistics. Available at SSRN 5231473.

[36] Kolla, S. K. (2023). Explainable AI and ML Models for Transparent Clinical Decision Support. Journal for ReAttach Therapy and Developmental Diversities, 6, 2444-2460.

[37] Nagabhyru, K. C. (2023). Accelerating Digital Transformation with AI Driven Data Engineering: Industry Case Studies from Cloud and IoT Domains. Educational Administration: Theory and Practice, 29(4), 5898-5910.

[38] Aitha, A. R. (2023). CloudBased Microservices Architecture for Seamless Insurance Policy Administration. International Journal of Finance (IJFIN)-ABDC Journal Quality List, 36(6), 607-632.

[39] Sheelam, G. K. (2024). Towards autonomic wireless systems: integrating agentic AI with advanced semiconductor technologies in telecommunications. Am. Online J. Sci. Eng., 3(4), 234-256.

[40] Yandamuri, U. S. (2022). Big Data Pipelines for Cross-Domain Decision Support: A Cloud-Centric Approach. International Journal of Scientific Research and Modern Technology (IJSRMT).

[41] Amistapuram, K. (2024). Federated Learning for Cross-Carrier Insurance Fraud Detection: Secure Multi-Institutional Collaboration. Journal of Computational Analysis and Applications (JoCAAA), 33(08), 6727-6738.

[42] Reddy Segireddy, A. (2024). Federated Cloud Approaches for Multi-Regional Payment Messaging Systems. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 15(2), 442-450.

[43] Aitha, A. R. (2023). Cloud-Native Big Data AI/ML Framework for Risk Intelligence and Fraud Control in Banking and Insurance Ecosystems. Available at SSRN 6157967.

[44] Kolla, T. (2023). Predictive ETL Failure Detection in Healthcare Data Pipelines Using Anomaly Detection Algorithms. International Journal of Medical Toxicology & Legal Medicine.

[45] Gottimukkala, V. R. R. (2022). Licensing Innovation in the Financial Messaging Ecosystem: Business Models and Global Compliance Impact. International Journal of Scientific Research and Modern Technology, 1(12), 177-186.

[46] Pamisetty, V. (2024). AI-Driven Decision Support for Taxation and Unclaimed Property Management: Enhancing Efficiency through Big Data and Cloud Integration. Available at SSRN 5250776.

[47] Mahesh Recharla, “Integrated Genomic and Neurobiological Pathway Mapping for Early Detection of Alzheimer’s Disease,” International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE), DOI: 10.17148/IJARCCE.2023.12122.

[48] Nagabhyru, K. C. (2023). From Data Silos to Knowledge Graphs: Architecting CrossEnterprise AI Solutions for Scalability and Trust. Available at SSRN 5697663. [91]Meda, R. (2024). Predictive Maintenance of Spray Equipment Using Machine Learning in Paint Application Services. European Data Science Journal (EDSJ) p-ISSN, 3050-9572.

[49] Sheelam, G. K., & Koppolu, H. K. R. (2024). From Transistors to Intelligence: Semiconductor Architectures Empowering Agentic AI in 5G and Beyond. Journal of Computational Analy- sis and Applications(JoCAAA), 33(08), 4518-4537.

[50] Bandi, V. D. V. K. (2024). Automated Feature Engineering Systems in Large-Scale Healthcare Data Environments. Journal of Neonatal Surgery, 13.

[51] Mangalampalli, B. M. (2024). AI-Enhanced Data Governance: Automating Compliance In Healthcare Analytics Platforms. The Review of Diabetic Studies, 191-204.

[52] Meda, R. (2024). Enhancing Paint Formula Innovation Using Generative AI and Historical Data Analytics. American Advanced Journal for Emerging Disciplinaries (AAJED) ISSN, 3067-4190.

[53] Bandi, V. D. V. K. (2024). AI-Driven Predictive Risk Modeling Architectures for Financial Systems. International Journal Of Finance, 37(3), 54-7.

[54] Kolla, S. K. (2023). Big Data–Driven Machine Learning Frameworks for Clinical Risk Prediction. International Journal of Medical Toxicology and Legal Medicine, 26(3), 44-59.

[55] Pamisetty, V. (2024). Transforming taxation systems through predictive analytics and AI-driven compliance monitoring tools. Am Data Sci J Adv Comput, 3, 55-68.

[56] Garapati, R. S. (2022). AI-Augmented Virtual Health Assistant: A Web-Based Solution for Personalized Medication Management and Patient Engagement. Available at SSRN 5639650.

[57] Nagabhyru, K. C. (2024). Data Engineering in the Age of Large Language Models: Transforming Data Access, Curation, and Enterprise Interpretation. Computer Fraud and Security.

[58] Aitha, A. R. (2022). Cloud Native ETL Pipelines for Real Time Claims Processing in Large Scale Insurers. Available at SSRN 5532601.

[59] Meda, R. (2023). Data Engineering Architectures for Scalable AI in Paint Manufacturing Operations. European data science journal.

[60] Gottimukkala, V. R. R. (2023). Privacy-Preserving Machine Learning Models for Transaction Monitoring in Global Banking Networks. International Journal of Finance (IJFIN)-ABDC Journal Quality List, 36(6), 633-652.

[61] Davuluri, P. N. Integrating Artificial Intelligence into Event-Driven Financial Crime Compliance Platforms.

[62] Kolla, S. H. (2022). Knowledge Retrieval Systems for Enterprise Service Environments. International Journal of Intelligent Systems and Applications in Engineering, 10, 495-506.

[63] Bandi, V. D. V. K. (2024). Intelligent Data Platforms For Personalized Retail Analytics At Scale. Metallurgical and Materials Engineering, 30 (4), 1011–1027.

[64] Mangala, N. (2022). Implementing Databricks Unity Catalog For Centralized Data Governance In Multi-Business-Unitenterprises. Journal of International Crisis and Risk Communication Research , 101–122. https://doi.org/10.63278/jicrcr.vi.3738.

[65] Davuluri, P. N. AI-Augmented Sanctions Screening: Enhancing Accuracy and Latency in Real Time Compliance Systems.

[66] Meda, R. (2024). Agentic AI in Multi-Tiered Paint Supply Chains: A Case Study on Efficiency and Responsiveness. Journal of Compu-tational Analysis and Applications (JoCAAA), 33(08), 3994-4015.