Exploring Graph Neural Networks for Complex Business Process Mining: Data-Driven Insights in Networked Systems

***Amish Doshi**, Lead Consultant, Excelon Solutions, USA*

Abstract

This paper investigates the application of Graph Neural Networks (GNNs) for mining complex business processes within networked systems. Traditional process mining techniques predominantly focus on extracting process models from event logs by identifying sequential patterns, which may overlook the intricate interdependencies and dynamic relationships that characterize business operations. In contrast, GNNs, which have gained prominence in the domain of machine learning, offer a novel approach to understanding and analyzing complex networks by representing entities as nodes and relationships as edges within a graph structure. This study explores how GNNs can be effectively leveraged to model, predict, and optimize business processes that operate within interconnected environments, where the performance and outcomes of processes are influenced by the interrelations between various entities involved in the system.

Business process mining has evolved beyond the analysis of individual tasks, seeking to incorporate the contextual and relational data that define how these tasks interact across multiple entities. In networked systems, the relationships between process components, such as departments, personnel, and external systems, often play a decisive role in the efficiency and effectiveness of the process outcomes. Standard process mining approaches struggle to account for these complexities, focusing instead on linear or tabular representations of events. However, GNNs naturally align with the structure of such systems, enabling the modeling of complex dependencies between different process components, thus offering a richer and more accurate understanding of process behavior.

Graph Neural Networks, as a type of deep learning model, are particularly suited for applications involving structured relational data. They allow the learning of node and edge representations that encapsulate not only individual process entities but also the context in which these entities interact within a broader system. This ability to capture both local and

global relationships within a network is critical for mining business processes that span multiple organizational boundaries or involve complex workflows where traditional sequential or time-series models may fail. Moreover, GNNs can be extended to incorporate dynamic changes in the network, such as organizational restructuring or shifts in process priorities, which can significantly impact the overall system performance.

The potential benefits of applying GNNs in business process mining are manifold. By representing business processes as graphs, it becomes possible to predict the flow of tasks, identify bottlenecks, and uncover hidden patterns that may otherwise remain undetected using conventional methods. Furthermore, GNNs can facilitate the optimization of business processes by simulating potential changes to the system and evaluating their impact on key performance indicators (KPIs). This predictive capability is especially valuable in networked systems where the interactions between process components are nonlinear and interdependent, requiring advanced models to accurately forecast outcomes and suggest improvements.

This paper also discusses the key challenges associated with applying GNNs to business process mining. One primary challenge is the availability and quality of data. Business process data often comes in diverse formats, including event logs, workflow models, and unstructured communications, which may not readily fit into the graph structure required for GNNs. Preprocessing and transforming such data into a suitable format is a crucial step for successful GNN application. Additionally, the scalability of GNNs in handling large-scale networked systems with thousands of nodes and edges remains an open challenge. While GNNs have shown promise in smaller-scale applications, further research is needed to address issues related to computational efficiency and model interpretability when dealing with complex, real-world business networks.

Despite these challenges, several case studies and empirical examples demonstrate the successful use of GNNs in various business process mining scenarios. These applications highlight the ability of GNNs to uncover relationships between disparate entities, leading to insights that can drive process improvements, reduce inefficiencies, and enhance decision-making. For instance, in supply chain management, GNNs have been used to model the interdependencies between suppliers, manufacturers, and distributors, enabling the identification of vulnerabilities and optimization opportunities. Similarly, in customer service operations, GNNs have been applied to analyze interactions between customer support agents, systems, and customer profiles, improving response times and satisfaction rates.

Keywords:

Graph Neural Networks, Business Process Mining, Networked Systems, Process Optimization, Deep Learning, Relational Data, Network Dependencies, Process Efficiency, Predictive Modeling, Organizational Systems.

1. Introduction

Business Process Mining (BPM) is a field that lies at the intersection of data science, process management, and computational modeling, aimed at extracting actionable insights from the operational data generated by organizational processes. At its core, BPM seeks to derive structured process models from event logs, enabling the analysis, monitoring, and optimization of business workflows. The primary objective is to bridge the gap between process design and execution by leveraging data-driven techniques to uncover inefficiencies, bottlenecks, and deviations from intended workflows.

The evolution of BPM has been significantly influenced by advancements in computational power and the proliferation of digital technologies within organizational ecosystems. Initial approaches relied heavily on manual process mapping, which was labor-intensive and prone to human error. With the advent of automated techniques, researchers and practitioners shifted their focus to algorithmic process discovery, wherein event logs were systematically analyzed to generate process models. Over time, methodologies expanded to include conformance checking and performance analysis, which evaluate the alignment of actual process execution with predefined models and assess their operational efficiency, respectively.

Despite these advancements, traditional BPM techniques are often limited to simplistic sequential representations of processes. These methods excel in straightforward workflows but fail to capture the complexity inherent in many modern organizational systems. The rise of interconnected and interdependent business ecosystems necessitates an approach that can handle the intricate relationships between process components. This need forms the basis for exploring alternative techniques capable of modeling the nuanced dynamics within networked systems.

Conventional process mining methodologies are constrained by their reliance on sequential or tabular representations of event logs, which inadequately represent the interconnected and dynamic nature of contemporary business environments. Such techniques often operate under the assumption that processes can be linearly decomposed into discrete activities performed in a predefined order. While effective for simple workflows, these assumptions break down in systems characterized by a high degree of relational and contextual interdependence.

In networked systems, where entities such as departments, personnel, suppliers, and customers interact in complex and often nonlinear ways, traditional BPM techniques struggle to provide meaningful insights. For instance, in supply chain management, the performance of a single process is intricately linked to the actions of multiple upstream and downstream entities. Similarly, in service operations, customer interactions often span multiple touchpoints, each contributing to the overall process outcome. Traditional methods fail to account for such relational dependencies, leading to incomplete or inaccurate process models that overlook critical interconnections.

Another limitation of conventional approaches is their inability to adapt to dynamic changes in process structures. Organizational processes are seldom static; they evolve in response to internal and external factors such as market demands, regulatory changes, and technological advancements. Static models derived from event logs become obsolete as processes deviate from their original structure. Furthermore, the increasing volume and velocity of process data generated by modern systems impose computational challenges, rendering traditional methods less effective for real-time analysis and decision-making.

Graph Neural Networks (GNNs) represent a paradigm shift in the field of machine learning, specifically designed to operate on graph-structured data. Unlike traditional neural networks that are optimized for fixed-dimensional input such as images or sequential data, GNNs are inherently suited for relational data, where entities and their interconnections can be naturally represented as nodes and edges within a graph. This structural flexibility enables GNNs to model complex systems characterized by interactions and dependencies among multiple entities.

The fundamental operation of GNNs revolves around the propagation of information through the graph structure. By iteratively aggregating information from a node's neighbors, GNNs compute representations that encode both local and global graph properties. This process, often referred to as message passing or graph convolution, allows GNNs to learn contextual embeddings that capture the relational and topological features of the graph. Such capabilities make GNNs particularly relevant for applications in which the relationships between entities are as important as the attributes of the entities themselves.

In the context of business process mining, GNNs provide a powerful framework for analyzing processes within networked systems. By representing business processes as graphs, where nodes correspond to process entities (e.g., tasks, resources, or roles) and edges represent their interactions, GNNs enable the modeling of complex dependencies that are often overlooked by traditional methods. This ability to capture the interplay between different components of a system positions GNNs as a transformative tool for understanding and optimizing business workflows.

GNNs also extend beyond static graph representations, accommodating dynamic and evolving processes. Temporal variations in process structures, such as changes in task sequences or shifts in resource dependencies, can be incorporated into GNN models through extensions like Temporal Graph Networks or Dynamic Graph Neural Networks. These capabilities align with the requirements of modern business environments, where processes are frequently subject to change.

Given the limitations of conventional BPM approaches and the demonstrated potential of GNNs in modeling relational data, this research aims to explore the application of GNNs in mining complex business processes within networked systems. Specifically, the study seeks to investigate how GNNs can be leveraged to uncover hidden patterns, identify bottlenecks, and optimize workflows in interconnected organizational ecosystems. By focusing on the relational and dynamic aspects of processes, the research aspires to contribute to a deeper understanding of how business processes operate in practice and how they can be systematically improved.

The primary objective is to bridge the gap between traditional process mining methodologies and the demands of modern, interconnected business systems. The study will examine the theoretical underpinnings of GNNs, their practical implementation in business process mining, and their comparative performance relative to conventional techniques. In doing so, it aims to establish GNNs as a viable and effective approach for advancing the state of the art in process mining research and practice.

2. Background and Related Work

Traditional Business Process Mining Techniques

The field of business process mining has evolved significantly since its inception, driven by the increasing availability of digital event data generated from organizational operations. Central to this discipline are three fundamental techniques: process discovery, conformance checking, and performance analysis. Each plays a distinct role in extracting actionable insights from event logs, which consist of timestamped records capturing the sequence of activities executed within a process.

Process discovery seeks to generate a process model based solely on the event logs, revealing the underlying structure and flow of a business process. These algorithms, including the α-algorithm and its variants, utilize frequency and sequence patterns to construct models such as Petri nets, directly-follows graphs, or BPMN diagrams. Conformance checking, on the other hand, evaluates the alignment between a predefined process model and actual execution data, quantifying deviations and identifying non-compliant behaviors. Techniques such as token-based replay and alignments have been extensively developed to facilitate this comparison. Performance analysis focuses on assessing the efficiency and effectiveness of processes by deriving metrics such as throughput time, resource utilization, and bottleneck locations from event logs.

While these methods have found widespread application across industries, their reliance on sequential or tabular data representations imposes significant constraints. The simplicity of traditional approaches limits their applicability to processes with well-defined, linear workflows, rendering them less effective in complex, interconnected systems.

Limitations of Sequential Process Mining

The inherent assumptions underlying traditional process mining techniques become a critical barrier when applied to systems characterized by dynamic and relational dependencies. Sequential process mining approaches operate on the premise that processes can be adequately represented as sequences of discrete activities. While effective in workflows that are deterministic and linear, this assumption fails to account for the interconnected nature of entities in networked environments.

Modern business ecosystems often exhibit intricate interactions between process entities, including tasks, resources, and external agents. For example, in a supply chain network, the execution of a particular task is influenced by upstream dependencies and downstream implications, forming a web of interconnected processes. Similarly, in customer service operations, workflows span multiple touchpoints, each contributing uniquely to the overall process outcome. Sequential methods are ill-equipped to capture these dependencies, leading to incomplete or oversimplified process models.

Furthermore, traditional approaches often struggle to adapt to dynamic changes in process structures. In rapidly evolving industries, business processes are continuously modified in response to market demands, technological innovations, and regulatory requirements. Static models derived from event logs fail to capture this evolution, resulting in outdated representations that lack relevance. Additionally, the increasing scale and complexity of modern organizational systems impose computational challenges, as event logs grow exponentially in size and diversity. These limitations necessitate a paradigm shift toward methodologies that can effectively model the relational and dynamic aspects of processes.

Introduction to Graph Neural Networks (GNNs)

Graph Neural Networks (GNNs) have emerged as a transformative tool in machine learning, specifically designed to address the challenges of modeling relational and structured data. Unlike traditional machine learning models that operate on fixed-dimensional input, GNNs are capable of processing graph-structured data, where entities (nodes) and their relationships (edges) are explicitly represented. This inherent flexibility makes GNNs well-suited for a wide range of applications, including social network analysis, recommendation systems, and molecular property prediction.

The core architecture of GNNs revolves around the principle of message passing, wherein information is propagated between nodes based on their connectivity within the graph. Through iterative aggregation and transformation of information from neighboring nodes, GNNs compute node embeddings that capture both local and global graph properties. These embeddings serve as feature representations that are subsequently utilized for downstream tasks such as node classification, link prediction, and graph-level classification.

Several variants of GNNs have been developed to address specific requirements. Graph Convolutional Networks (GCNs) extend the concept of convolutional neural networks to graph data, applying localized filters to extract features from node neighborhoods. Graph Attention Networks (GATs) introduce attention mechanisms, allowing the model to assign varying importance to different neighbors during message passing. Graph Recurrent Networks (GRNs) and Temporal GNNs are designed to handle dynamic graphs, incorporating temporal dimensions to model evolving relationships.

In the context of business process mining, GNNs offer the capability to represent processes as graphs, where nodes correspond to process entities and edges represent their interactions. This representation enables the modeling of complex dependencies, capturing the interplay between various components of a process. By leveraging GNNs, researchers can overcome the limitations of sequential process mining techniques, unlocking new opportunities for analyzing and optimizing business workflows.

Previous Applications of GNNs in Business Contexts

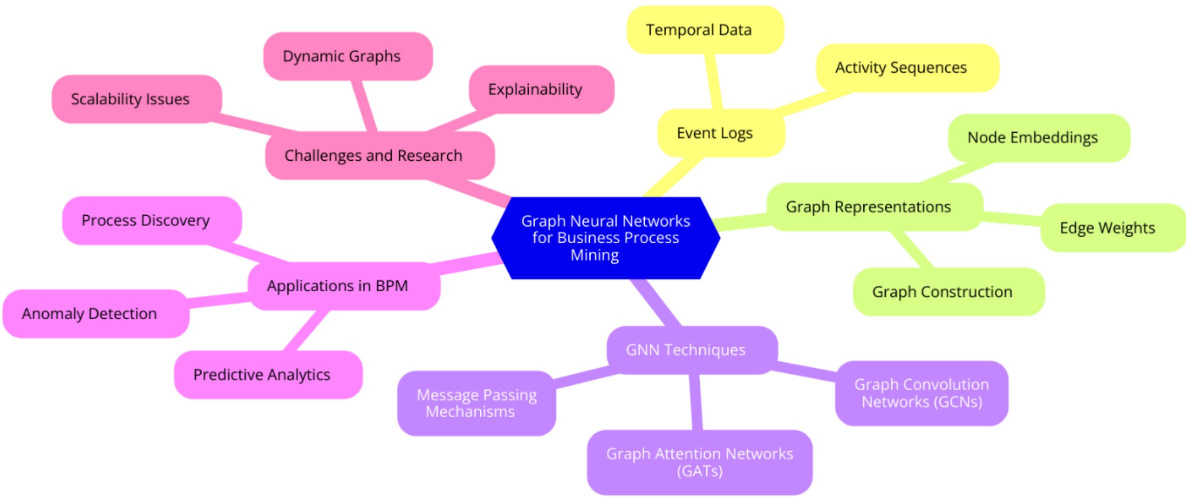
The application of GNNs in business-related domains has gained traction, with several studies demonstrating their potential to address complex challenges in organizational settings. In supply chain management, GNNs have been employed to model relationships between suppliers, manufacturers, and distributors, enabling the identification of vulnerabilities and optimization of logistics networks. For example, graph-based representations of supply chain networks have facilitated the analysis of cascading failures and the development of robust mitigation strategies.

Customer interaction analysis represents another area where GNNs have shown promise. By representing customer interactions as graphs, where nodes correspond to customers and touchpoints, and edges capture interactions, GNNs have been used to predict customer behavior and enhance personalization strategies. These models have been instrumental in identifying influential customers, predicting churn, and optimizing marketing campaigns.

Additionally, GNNs have been applied to financial risk analysis, where the relationships between financial entities, such as banks, borrowers, and assets, are modeled as graphs. Through these representations, researchers have developed models to predict default risks, assess systemic stability, and optimize investment portfolios.

Despite these advancements, the application of GNNs to business process mining remains relatively unexplored. Existing studies primarily focus on static graph representations and specific business contexts, leaving significant opportunities to extend GNN methodologies to dynamic, networked processes. By bridging this gap, the present research seeks to contribute to the growing body of knowledge on GNN applications, highlighting their potential to transform the analysis and optimization of complex business processes.

3. Graph Neural Networks for Business Process Mining



Understanding Graph Representation in Business Processes

The conceptualization of business processes as graphs provides a robust framework for analyzing complex organizational workflows. In this representation, the components of a business process—such as tasks, activities, events, or actors—are modeled as nodes, while the interactions or dependencies among these components are depicted as edges. This graph-based abstraction captures the inherent relational structure of business processes, accommodating both direct and indirect relationships between entities.

For instance, in a supply chain management scenario, nodes may represent suppliers, manufacturers, and distributors, while edges symbolize the flow of goods, information, or capital between these entities. Similarly, in a customer service process, nodes can represent service agents, customers, and interaction channels, with edges denoting communication exchanges. This structural representation enables the explicit encoding of interdependencies, which are often obscured in traditional sequential or tabular data formats. By leveraging graph representations, researchers and practitioners can uncover patterns, dependencies, and dynamics that would otherwise remain elusive in linear process models.

Relational Data in Business Systems

Relational data lies at the heart of modern business systems, encompassing the myriad interactions and dependencies that drive organizational processes. These relationships span diverse dimensions, including hierarchical relationships between employees and departments, transactional links between suppliers and customers, and temporal dependencies among sequential tasks. Accurately modeling these relationships is critical for understanding process behaviors and optimizing performance outcomes.

Relational data is particularly significant in networked business environments, where processes are distributed across interconnected entities. For example, in project management, the completion of one task often depends on the outcomes of other tasks, forming a network of interdependent activities. Similarly, in financial services, risk assessment models must account for the interconnected nature of borrowers, lenders, and collateral assets. Traditional process mining approaches, which primarily focus on activity sequences, fail to capture these intricate dependencies. Graph representations, by contrast, provide a natural and flexible means of encoding relational data, enabling the analysis of complex, multi-dimensional interactions within business systems.

GNN Model Components

Graph Neural Networks (GNNs) are designed to process and analyze graph-structured data, making them particularly well-suited for business process mining. The architecture of GNNs revolves around three key components: node and edge embeddings, message passing mechanisms, and graph aggregation operations.

Node and edge embeddings are feature representations that encode the attributes of nodes and edges within a graph. For example, in a business process graph, node attributes may include task types, resource requirements, or timestamps, while edge attributes may capture relationship strength, interaction frequency, or temporal lags. These embeddings serve as the input to the GNN, providing a rich feature space for subsequent computations.

Message passing is the core operation of GNNs, facilitating the exchange of information between nodes based on their connectivity. During each iteration, a node aggregates information from its neighbors and updates its embedding to reflect the structural and attribute information within its local neighborhood. This iterative process enables the propagation of information across the graph, capturing both direct and indirect dependencies among nodes. Variants of message passing mechanisms, such as convolutional filters in Graph Convolutional Networks (GCNs) or attention-based mechanisms in Graph Attention Networks (GATs), further enhance the model's ability to learn complex patterns.

Graph aggregation is the final step, wherein the node embeddings generated through message passing are combined to produce a global representation of the entire graph. This global embedding captures holistic graph-level features, which are essential for tasks such as process classification, anomaly detection, or performance prediction. Aggregation functions, such as mean pooling, max pooling, or attention-based pooling, are used to summarize node-level information into a comprehensive graph-level representation.

Advantages of GNNs in Business Process Mining

The application of GNNs to business process mining offers several advantages over traditional methods, addressing many of the limitations associated with sequential and tabular data models.

One of the primary strengths of GNNs lies in their ability to capture complex dependencies within business processes. By representing processes as graphs, GNNs can model intricate interactions between entities, including hierarchical relationships, cyclical dependencies, and temporal correlations. This capability is particularly valuable in scenarios where process behaviors are influenced by network effects or multi-agent interactions.

GNNs also excel in capturing spatial and temporal relationships within business processes. For example, in a manufacturing workflow, the spatial proximity of production stages and the temporal sequencing of tasks can significantly impact process efficiency. GNNs can encode these relationships through graph structures and dynamic graph models, enabling the analysis of spatial-temporal dependencies that traditional approaches overlook.

Furthermore, GNNs offer enhanced predictive power for mining business processes. The combination of rich feature representations, iterative message passing, and holistic graph aggregation enables GNNs to learn complex patterns and dependencies from data. This predictive capability supports a wide range of applications, including process optimization, anomaly detection, and decision support. For instance, GNN-based models can predict bottlenecks in workflows, identify deviations from expected process behaviors, and recommend corrective actions to improve performance outcomes.

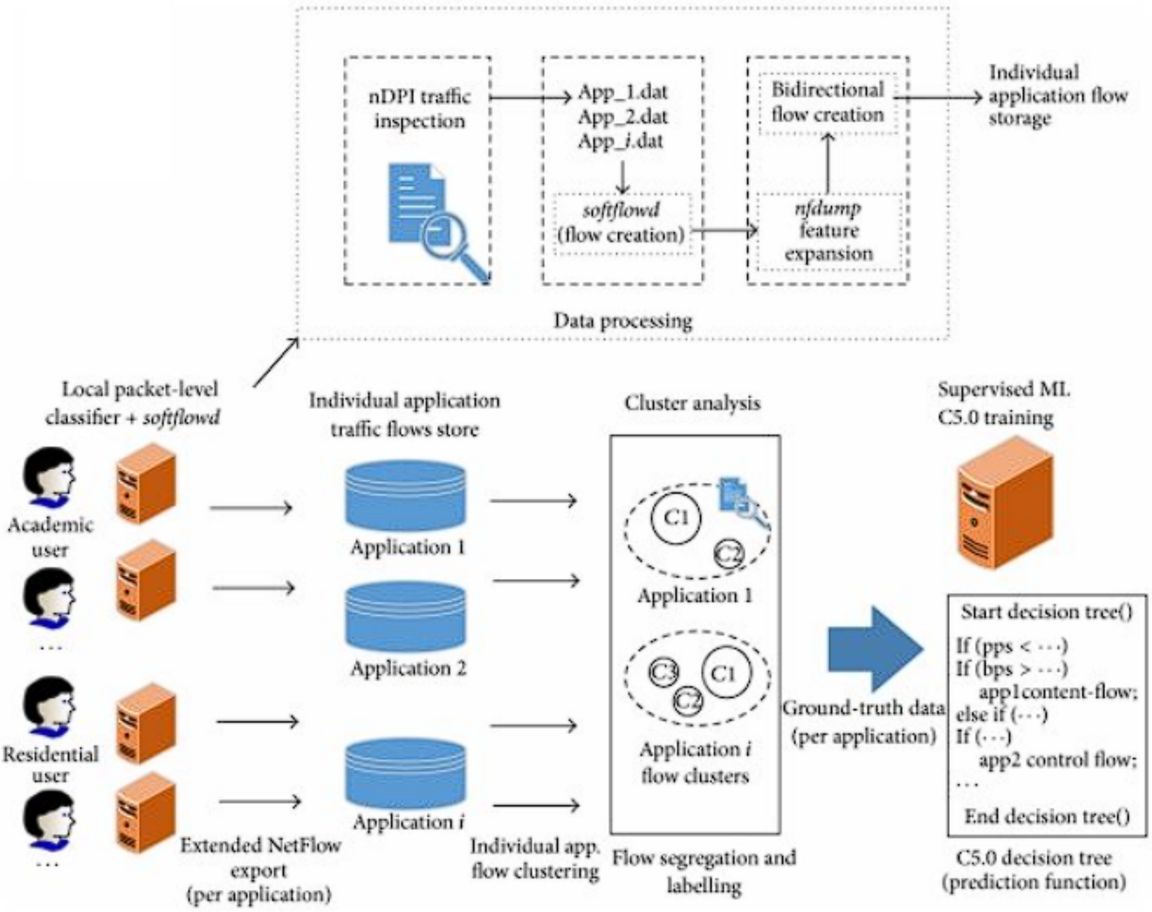
In addition to their analytical capabilities, GNNs are inherently flexible and scalable, accommodating diverse graph structures and large-scale datasets. This flexibility allows GNNs to be applied across various domains and industries, from logistics and supply chain management to healthcare and financial services. As business processes continue to evolve in complexity and scale, GNNs provide a powerful and adaptable framework for understanding and optimizing organizational workflows.

By integrating GNNs into the domain of business process mining, researchers and practitioners can unlock new insights and capabilities, addressing the challenges of relational and dynamic process analysis. The subsequent sections will explore the methodologies and applications of GNNs in greater depth, highlighting their transformative potential for modern business systems.

4. Methodology for Applying GNNs in Business Process Mining

Data Collection and Preprocessing

In the context of business process mining, data collection and preprocessing are critical steps that ensure the quality and applicability of the data used to train Graph Neural Networks (GNNs). The primary data sources typically include event logs, organizational structure data, and external systems data, each providing distinct insights into the operations and interactions within a business process.



Event logs are the foundational data in process mining, as they document the sequence of activities or tasks within a process, capturing the timestamp, actor, and event type associated with each activity. However, event logs in their raw form often suffer from noise, inconsistencies, and missing data, which necessitates thorough preprocessing. This includes cleaning and aligning event logs across various sources, resolving discrepancies in timestamps, and filling gaps where activities may have been missed or improperly logged.

Organizational structure data represents the configuration and relationships among various entities within the business, such as departments, teams, or individuals. In many cases, this data is unstructured or exists in disparate formats, which makes its integration with process logs a non-trivial task. For example, mapping the roles of employees to specific tasks in the event log requires understanding both the formal hierarchical structure and informal connections among employees. Preprocessing this data may involve standardizing the roles, aligning them with tasks, and creating a unified framework for the analysis.

External systems data, such as customer data, financial transactions, or supply chain interactions, can provide additional layers of context to the business process. These external datasets are often heterogeneous, coming from diverse sources with different formats and update frequencies. Integrating this data into a unified graph structure requires normalization and alignment techniques to ensure consistency across datasets.

Once raw data is collected, it must be preprocessed to transform it into a graph-friendly format. This preprocessing step involves extracting relevant features from event logs, organizational data, and external systems, which will later be encoded as node and edge attributes. Furthermore, data cleaning methods, such as outlier detection, imputation for missing values, and standardization of event identifiers, are necessary to mitigate errors and ensure the integrity of the dataset. Preprocessing steps are thus paramount for enabling effective graph construction and subsequent analysis using GNNs.

Graph Construction

The transformation of business process data into graph structures is a pivotal aspect of applying Graph Neural Networks to business process mining. Graphs, by definition, consist of nodes and edges, with each node representing an entity or event in the process and each edge capturing the relationships or interactions between them. Constructing meaningful graphs from business process data requires defining both the nodes and edges, as well as their respective attributes, to reflect the underlying process dynamics.

Defining nodes is the first step in constructing a business process graph. In the context of process mining, nodes typically represent entities such as tasks, activities, employees, departments, or events in the workflow. For example, each task within a process (e.g., invoice approval, customer order processing) can be modeled as a distinct node, with associated attributes such as task type, execution time, and responsible actor. In more complex scenarios, nodes may represent organizational units, roles, or even external partners who interact with the business process. The richness of node attributes is critical for ensuring that the graph captures relevant information that will be useful for the GNN to learn process patterns and dependencies.

Edges, on the other hand, represent the relationships or interactions between nodes. In business processes, these relationships may include task dependencies (e.g., Task A must precede Task B), communication flows (e.g., information exchange between departments), or temporal relationships (e.g., delays between activities). Edges can be directed or undirected, depending on the nature of the relationship between the nodes. For instance, in a supply chain scenario, the flow of goods from one supplier to another can be represented as a directed edge, while the relationship between two departments within an organization may be represented by an undirected edge reflecting mutual collaboration. Additionally, edges can have attributes such as weight (representing the strength of the relationship), timestamps, or types of interaction (e.g., transactional, informational).

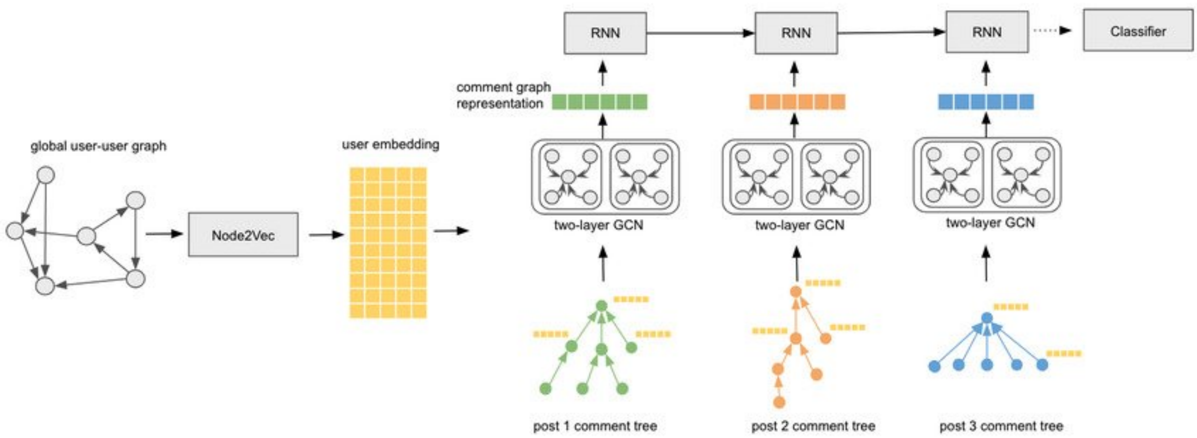
The construction of edges becomes more complex when considering multi-layered relationships in business processes. In many business contexts, there are multiple dimensions of interaction between entities. For example, a supplier might interact with a manufacturer through both financial transactions and logistical coordination. In such cases, multiple types of edges can be created to reflect different kinds of dependencies and relationships. This multi-dimensional graph construction provides a more nuanced view of the business process, which is crucial for accurately capturing the complex interactions within networked systems.

Once the nodes and edges are defined, the next step is to assign attributes to them. Node attributes may include categorical information (e.g., department type, task priority), numerical values (e.g., time taken to complete a task), or temporal features (e.g., task start and end times). Edge attributes are equally important and may represent interaction frequencies, transaction volumes, or delays between tasks. In some cases, edge attributes might capture additional meta-data, such as the type of communication or the quality of interaction. Proper encoding of these attributes ensures that the graph not only reflects the structural relationships between entities but also captures the contextual and behavioral dynamics inherent in the business process.

Graph construction also involves addressing potential challenges related to scalability and complexity. In large-scale business processes, where thousands of entities and interactions may exist, the graph can grow exponentially in size. Efficient graph construction techniques, such as graph pruning (to eliminate redundant nodes and edges), graph sparsification (to focus on the most important relationships), and hierarchical graph construction (to decompose large graphs into smaller, more manageable subgraphs), are essential for ensuring that the GNN model can process and learn from the data effectively. Additionally, in networked systems with continuously evolving relationships, dynamic graph construction methods may be required to capture changes in the business process over time. These dynamic graphs, which update their structure and attributes based on real-time data, enable the GNN to learn temporal patterns and adapt to shifting process dynamics.

GNN Model Design

The design of a Graph Neural Network (GNN) architecture for business process mining requires a comprehensive understanding of the unique characteristics of business process data, particularly its relational and temporal dependencies. In this context, selecting an appropriate GNN model involves considering both the structural complexities of the business process graphs and the specific tasks at hand, such as process discovery, conformance checking, or performance analysis.



One key decision in GNN model design is the choice of the underlying GNN architecture. Among the most commonly used architectures in graph-based machine learning tasks are Graph Convolutional Networks (GCNs) and Graph Attention Networks (GATs), both of which are well-suited to capturing local dependencies and interactions within a graph. Each of these architectures has strengths that can be leveraged in different aspects of business process mining.

Graph Convolutional Networks (GCNs) apply convolutional operations over graph structures, aggregating information from neighboring nodes in each layer. In the context of business process mining, GCNs can be employed to propagate information across different stages of a business process, allowing for the capture of dependencies between related tasks, departments, or process entities. This architecture is particularly advantageous for tasks such as conformance checking, where the goal is to assess the compliance of a business process model with observed execution patterns. GCNs are capable of capturing the local dependencies between nodes, making them well-suited for modeling the interactions between process activities that occur in close succession or are interdependent.

Graph Attention Networks (GATs), on the other hand, introduce the concept of attention mechanisms into the graph-based learning framework. In GATs, each node in the graph computes a weighted aggregation of its neighbors' features, where the weights are learned dynamically during training. This attention mechanism allows GATs to assign different importance to various neighbors, making them particularly effective for capturing heterogeneous and dynamic relationships within business processes. For instance, in a scenario where certain tasks in a process flow are more critical or have a higher influence on overall process performance, GATs can selectively emphasize these nodes in their learning. GATs are thus well-suited for applications where there are varying degrees of relevance across different process activities or where relationships between process entities are complex and multi-dimensional.

Choosing between these architectures depends on the specific requirements of the business process mining task. In scenarios where the relationships between tasks are relatively uniform or where process dependencies can be well-represented by local interactions, GCNs may provide a simpler and more effective solution. However, in cases where the relationships between tasks are non-homogeneous or dynamically change over time, GATs may offer better performance by allowing the model to focus on the most relevant interactions at each point in the process. Additionally, hybrid approaches that combine GCNs and GATs have also shown promise, where GCNs are used to capture global process structure and GATs are employed to model dynamic, attention-based relationships between tasks.

Once the architecture has been selected, further design considerations involve the depth of the network, the number of layers, and the activation functions used within the model. The depth of the network determines the extent to which information is propagated through the graph, and the number of layers must be carefully chosen to balance model complexity with overfitting. Furthermore, appropriate activation functions such as ReLU or LeakyReLU can help introduce non-linearity and ensure that the model captures complex, non-linear relationships inherent in business processes.

Training and Evaluation

Training a Graph Neural Network (GNN) for business process mining requires a well-defined strategy for data preparation, model optimization, and performance validation. The first step in training is the selection of labeled data that will guide the learning process. Labeled data refers to business process data where the outcomes or behaviors of interest (such as process efficiency, task completion times, or process bottlenecks) are already known and serve as ground truth for the model. For example, labeled event logs where each task's outcome (completed, delayed, or failed) is specified allow the GNN to learn to predict similar outcomes in new, unseen data.

In training, the GNN model learns by optimizing the weights of its nodes and edges, adjusting its parameters to minimize the error between its predicted output and the true labels. Training typically involves an iterative process where the model is exposed to batches of graph data, computes predictions, and adjusts its weights using a loss function, commonly mean squared error (MSE) or cross-entropy loss, depending on the task at hand. A standard gradient descent-based optimization algorithm, such as Adam or SGD (stochastic gradient descent), is used to update the model parameters during each iteration. The learning rate and batch size are hyperparameters that need to be tuned for optimal performance.

To ensure that the model generalizes well to new data and avoids overfitting, the training process often includes regularization techniques. These techniques, such as dropout or weight decay, reduce the model's reliance on any single feature or node, promoting a more robust learning process. Additionally, cross-validation techniques are commonly employed to assess the model's performance across multiple subsets of the data, ensuring that the model's performance is not dependent on any single training or validation split.

Once trained, the model must be evaluated on its ability to predict or uncover insights relevant to business process mining tasks. Common evaluation metrics in process mining include process efficiency, bottleneck identification, and performance prediction. For instance, a key metric for process efficiency could be the time taken to complete a business process from start to finish, which the model should learn to predict accurately based on the process data. Bottleneck identification involves detecting stages in the process that are most prone to delays, where the GNN should be able to identify which tasks or entities are most responsible for slowdowns based on their interactions and dependencies within the graph structure.

In addition to process efficiency, another important evaluation criterion is conformance checking, where the GNN model is tasked with identifying discrepancies between the modeled process and the observed execution in event logs. Conformance checking evaluates how well the actual process adheres to the idealized process model, highlighting areas where deviations occur. For example, if the GNN identifies that a certain activity often precedes another when it should follow, this could signal a misalignment in the business process model, which could lead to inefficiencies or non-compliance.

Moreover, the prediction of future process behaviors, such as task completion times or the likelihood of bottlenecks emerging, is a critical task that GNNs can perform. Evaluating the accuracy of these predictions involves comparing the GNN's forecasts against actual outcomes, using metrics such as mean absolute error (MAE) or root mean square error (RMSE) to quantify the discrepancy between predicted and actual values.

Finally, the interpretability of the model's predictions is an important consideration. In business process mining, stakeholders need to understand the rationale behind the GNN's predictions or insights. Techniques such as attention mechanisms (for GATs) or visualization of node/edge embeddings can help to make the model's decision-making process more transparent and provide actionable insights into the business process.

In summary, the training and evaluation of GNNs for business process mining are critical to ensuring the model's ability to provide actionable, accurate, and reliable insights. By leveraging labeled data, optimizing the model with appropriate algorithms, and validating its performance using business-relevant metrics, the GNN can uncover hidden patterns, improve process efficiency, and provide valuable predictive capabilities within complex networked systems.

5. Case Studies and Applications

Supply Chain Optimization

The application of graph neural networks (GNNs) to supply chain optimization represents a significant advancement in understanding and improving the interconnected dynamics of suppliers, manufacturers, and distributors. In a typical supply chain, the relationships between these entities can be conceptualized as a graph, with nodes representing suppliers, production facilities, distribution centers, and retailers, while edges encapsulate the flow of goods, information, and financial transactions. GNNs, with their inherent capability to model relational data, provide a robust framework for optimizing such complex networks.

By leveraging GNNs, supply chain processes can be analyzed to identify bottlenecks, forecast disruptions, and evaluate the impact of alternative logistical strategies. For instance, GNN models can integrate diverse data sources, such as historical delivery times, supplier reliability metrics, and market demand fluctuations, to predict potential delays or inefficiencies. Through message-passing algorithms, GNNs capture the cascading effects of disruptions, enabling preemptive measures such as rerouting shipments or adjusting inventory levels at strategic nodes.

One of the most transformative impacts of GNNs in supply chain management lies in risk mitigation. By modeling interdependencies and vulnerabilities within the network, GNNs can identify critical suppliers or facilities whose failure would significantly impact overall operations. Such insights enable businesses to diversify supplier bases, allocate safety stock effectively, or invest in redundancy for high-risk components. Furthermore, GNN-based optimization algorithms have been shown to improve sustainability by identifying routes or processes that minimize environmental impact while maintaining operational efficiency.

Customer Service Process Mining

In the domain of customer service, the ability to analyze and optimize interactions between customers and agents is paramount to improving service quality and reducing operational inefficiencies. Traditional process mining techniques, while effective in linear or sequential workflows, often struggle to capture the nuanced, bidirectional relationships inherent in customer-agent interactions. GNNs address this limitation by providing a relational perspective, wherein interactions are represented as graphs, with nodes denoting customers, agents, and communication channels, and edges representing interactions, feedback loops, or escalations.

The application of GNNs in customer service process mining facilitates the identification of patterns that contribute to prolonged resolution times, high escalation rates, or low customer satisfaction scores. By aggregating and analyzing interaction data, GNNs can uncover clusters of recurring issues, predict the likelihood of escalations, and recommend interventions tailored to specific customer profiles or issue types. For example, GNNs can analyze the relational data between agents and customers to identify cases where knowledge gaps or miscommunication are prevalent, enabling targeted training programs or the implementation of more effective knowledge-sharing systems.

Furthermore, GNN-based insights can be instrumental in designing intelligent routing systems that allocate cases to agents based on predicted resolution efficiency, thus reducing wait times and enhancing overall service levels. These applications not only optimize operational efficiency but also contribute to long-term customer retention by addressing pain points in the service journey.

Operational Risk Management

In business process management, the identification and mitigation of operational risks are critical for ensuring the continuity and efficiency of workflows. Operational risks often stem from vulnerabilities within the process chain, such as resource constraints, dependency on unreliable entities, or exposure to external disruptions. GNNs, with their ability to model complex interdependencies, offer a sophisticated approach to analyzing and mitigating these risks.

Using graph representations, GNNs can model business processes as networks of interrelated activities, resources, and stakeholders. For instance, a manufacturing process may be represented as a graph where nodes correspond to tasks, machines, or workers, and edges denote dependencies or resource flows. GNNs can analyze this graph to identify weak links or critical nodes whose failure would propagate through the network, causing significant disruptions.

Once identified, these vulnerabilities can be addressed through targeted resource allocation, process redesign, or contingency planning. For example, if a GNN analysis reveals that a specific machine is a critical bottleneck in a production process, investments in redundant capacity or predictive maintenance could mitigate the associated risk. Similarly, GNNs can simulate the impact of potential disruptions, such as supplier failures or workforce shortages, enabling organizations to develop robust mitigation strategies.

Beyond risk identification, GNNs can also support dynamic risk management by continuously updating their models based on real-time data. This capability is particularly valuable in volatile industries where risks evolve rapidly, requiring agile and adaptive responses.

Other Industry Applications

The versatility of GNNs extends beyond traditional business contexts, finding applications in a variety of industries where process optimization is a priority. In healthcare, GNNs have been employed to optimize patient care workflows by analyzing relationships between clinical tasks, healthcare providers, and patient outcomes. For example, GNNs can identify inefficiencies in hospital discharge processes or predict the likelihood of patient readmissions, enabling targeted interventions to improve care quality and reduce costs.

In the financial sector, GNNs have been used to detect fraud by modeling transactional data as graphs. By analyzing the relationships between accounts, transactions, and entities, GNNs can identify suspicious patterns indicative of fraudulent activity. Similarly, in investment portfolio management, GNNs have been applied to model the relationships between assets and predict market trends, supporting more informed decision-making.

In manufacturing, GNNs have facilitated the optimization of production workflows by analyzing the interdependencies between machines, tasks, and materials. This analysis has been instrumental in reducing downtime, improving throughput, and ensuring the efficient utilization of resources. Additionally, GNNs have been used to enhance quality control processes by identifying patterns associated with defects or anomalies in production data.

The applications of GNNs in business process mining and optimization are diverse and far-reaching, offering transformative potential across multiple domains. By leveraging the relational insights provided by GNNs, organizations can enhance efficiency, mitigate risks, and drive innovation, establishing a competitive edge in an increasingly complex and interconnected world.

6. Challenges and Limitations

Data Availability and Quality

One of the fundamental challenges in leveraging graph neural networks (GNNs) for business process mining lies in the availability and quality of data. Effective GNN applications rely on accurate, structured data to construct meaningful graphs that represent the underlying processes. However, obtaining such data from real-world business environments is often fraught with difficulties. Event logs, which form the backbone of process mining, may be incomplete, inconsistent, or riddled with inaccuracies due to human errors, system failures, or insufficient logging practices. Additionally, organizational data such as resource allocations, inter-departmental workflows, and customer interactions may be fragmented across disparate systems, further complicating the data acquisition process.

The heterogeneity of data sources adds another layer of complexity. Business processes often span multiple systems, each employing distinct data formats and standards. Integrating these diverse datasets into a unified format suitable for graph construction is both resource-intensive and prone to error. Furthermore, sensitive information contained in business data raises concerns about privacy and security, limiting access to comprehensive datasets for GNN training and evaluation. The lack of high-quality, representative datasets poses a significant bottleneck in deploying GNNs effectively in practical business contexts.

Data Preprocessing

The transformation of raw business process data into a graph representation suitable for GNNs introduces significant preprocessing challenges. Business data frequently exists in various formats, ranging from structured databases to unstructured text, necessitating sophisticated methods for extracting relevant entities and relationships. For instance, event logs must be parsed to identify distinct process activities, their temporal order, and their interdependencies, which can then be represented as nodes and edges in a graph. This process requires not only technical expertise but also domain-specific knowledge to ensure that the constructed graph accurately reflects the underlying business process.

Moreover, the dynamic nature of business systems often results in continuously evolving data. Static graph representations may fail to capture temporal changes, necessitating the creation of dynamic or temporal graphs that encode time-sensitive relationships. The preprocessing pipeline must account for such complexities, incorporating methods to handle missing data, normalize attributes, and reconcile inconsistencies across sources. The computational and manual effort required for preprocessing can become a limiting factor, particularly for organizations with limited technical resources.

Scalability of GNNs

As business systems grow in complexity, the scalability of GNNs emerges as a critical challenge. Large-scale organizations often operate extensive networks of interconnected processes involving thousands or even millions of entities and relationships. Training GNNs on such expansive graphs demands substantial computational resources, including high-performance hardware and memory. The computational complexity of GNN algorithms, which involves iterative message passing and aggregation across the graph, scales poorly with increasing graph size, posing significant obstacles for real-time or near-real-time applications.

Dynamic systems, where relationships and entities frequently change over time, further exacerbate scalability issues. The need to update the graph structure and retrain the GNN model in response to these changes introduces additional computational overhead. While techniques such as sampling, partitioning, and distributed processing have been proposed to address scalability concerns, their implementation requires careful consideration to avoid sacrificing model accuracy or fidelity. Ensuring that GNNs can efficiently handle large, dynamic graphs without compromising performance remains an open research question.

Model Interpretability

The interpretability of GNN models represents another critical limitation in their application to business process mining. GNNs operate as black-box models, wherein the mechanisms underlying their predictions are often opaque and difficult to decipher. This lack of transparency poses challenges for business stakeholders who require clear, actionable insights to inform decision-making. For instance, while a GNN may predict process inefficiencies or potential bottlenecks, understanding the specific factors driving these predictions can be elusive.

This interpretability challenge is particularly pronounced in regulated industries, such as finance or healthcare, where decision-making processes must adhere to stringent compliance standards. Without a clear understanding of the rationale behind GNN predictions, organizations may face difficulties in justifying or validating their decisions to regulatory bodies. Techniques such as attention mechanisms, explainability frameworks, and feature attribution methods have been proposed to enhance the interpretability of GNNs, but their adoption in practical business contexts remains limited. Developing interpretable GNN models that balance predictive accuracy with transparency is essential for broader acceptance in business applications.

Generalization Across Different Industries

The generalizability of GNN models trained on specific datasets to other industries or organizations presents a significant challenge. Business processes vary widely across industries, with differences in workflows, organizational structures, and performance metrics. A GNN model optimized for one context may fail to capture the nuances of another, limiting its applicability across diverse domains. For instance, a model trained on supply chain data from the manufacturing sector may not generalize well to service-oriented industries with distinct process characteristics.

This challenge is further compounded by the diversity of data quality, availability, and structure across organizations. Customization of GNN architectures, hyperparameters, and training protocols is often necessary to achieve satisfactory performance in new contexts, necessitating significant domain expertise and technical effort. The lack of standardized benchmarks and datasets for evaluating cross-domain generalization hinders the development of universally applicable GNN models for business process mining. Addressing these limitations requires a combination of advanced transfer learning techniques, domain-specific adaptations, and collaborative efforts to establish industry-wide standards for GNN applications.

While GNNs offer transformative potential for business process mining, their practical implementation is constrained by challenges related to data quality, preprocessing, scalability, interpretability, and generalization. Overcoming these limitations will require ongoing research, innovation, and collaboration between academia and industry, paving the way for more effective and widespread adoption of GNNs in business contexts.

7. Future Directions and Research Opportunities

Improving GNN Models for Business Process Mining

The advancement of graph neural network (GNN) models for business process mining necessitates exploration into hybrid architectures that integrate GNNs with other machine learning paradigms. The combination of GNNs with reinforcement learning (RL) or deep reinforcement learning (DRL) holds significant potential for enhancing performance. Such hybrid models could address decision-making challenges within dynamic business environments by enabling adaptive learning mechanisms. For instance, while GNNs can efficiently model and capture the structural dependencies of processes, RL algorithms can optimize decision-making by learning policies that maximize long-term operational efficiency. Integrating these approaches could lead to the development of robust models capable of not only predicting process outcomes but also recommending actionable strategies for process improvement.

Moreover, incorporating domain-specific constraints and prior knowledge into GNN architectures presents an opportunity to improve their applicability in business contexts. Constraint-aware GNNs that embed rules, such as compliance requirements or operational limitations, directly into the model's learning process could significantly enhance interpretability and performance in regulated industries. Additionally, the exploration of meta-learning techniques to design adaptive GNNs capable of generalizing across diverse business scenarios represents another promising avenue for future research.

Real-Time Process Monitoring

The application of GNNs in real-time business process monitoring represents a transformative potential for modern organizations. Real-time monitoring involves the continuous analysis of process data to detect anomalies, forecast disruptions, and dynamically adjust operations. The ability of GNNs to process and model graph-structured data in near-real-time could enable organizations to monitor interconnected processes more effectively than traditional methods. By leveraging streaming data pipelines, GNNs can be deployed to continuously update graph representations of business processes, ensuring that models remain relevant in fast-evolving environments.

One promising application is in anomaly detection, where GNNs can identify deviations from normal operational patterns that may indicate process inefficiencies, security breaches, or emerging bottlenecks. Furthermore, the integration of GNNs with optimization algorithms could facilitate real-time process adjustments, such as reallocating resources or re-routing workflows to mitigate identified issues. However, achieving these capabilities requires overcoming technical challenges related to latency, computational efficiency, and scalability, which merit focused research efforts.

Integration with IoT and Big Data

The increasing prevalence of Internet of Things (IoT) devices and the proliferation of big data present an unparalleled opportunity to enhance GNN applications in business process mining. IoT devices generate continuous streams of granular data, offering rich insights into operational states, machine performance, and environmental conditions. By integrating IoT data into GNN models, organizations can achieve more comprehensive and accurate representations of their processes. For example, sensor data from manufacturing equipment can be used to construct dynamic graphs that capture real-time interactions and dependencies, enabling predictive maintenance and optimized production planning.

Similarly, big data analytics can serve as a vital input for GNNs, providing extensive datasets for training and validation. Techniques such as federated learning could be employed to train GNNs on distributed datasets, preserving data privacy while enabling collaborative insights across organizational boundaries. The convergence of GNNs, IoT, and big data analytics represents a promising interdisciplinary frontier for advancing business process optimization.

Scalability Solutions

Scalability remains one of the most pressing challenges in deploying GNNs for large-scale business process mining. Future research must focus on the development of more efficient graph algorithms and architectures capable of handling vast and complex graphs with minimal computational overhead. Innovations such as sparse graph representations, graph sampling techniques, and scalable message-passing algorithms could enable GNNs to process large graphs without compromising accuracy.

The integration of distributed computing frameworks, such as Apache Spark or TensorFlow Distributed, offers another pathway for addressing scalability issues. By parallelizing the computation of GNNs across multiple nodes or clusters, researchers can improve the efficiency of training and inference processes. Furthermore, research into lightweight GNN architectures designed specifically for edge computing environments could facilitate the deployment of scalable GNN models in resource-constrained settings, such as IoT-enabled industrial systems.

Interdisciplinary Research

The application of GNNs in business process mining could benefit greatly from interdisciplinary collaborations that integrate concepts and methodologies from adjacent fields. For instance, combining GNNs with business intelligence tools could enable the automated extraction of actionable insights from process data, enhancing decision-making at all organizational levels. Similarly, integrating system dynamics modeling with GNNs could provide a deeper understanding of the feedback loops and causal relationships inherent in complex business systems.

Social network analysis represents another promising area for interdisciplinary exploration. Many business processes involve human actors whose interactions can be modeled as social networks. By incorporating social network analysis techniques into GNN-based models, researchers can gain insights into the impact of organizational hierarchies, team dynamics, and communication patterns on process performance.

The inclusion of behavioral economics and cognitive science in GNN research could also lead to more nuanced models that account for human decision-making factors in business processes. For instance, understanding how employees respond to changes in workflow or resource allocation could inform the design of GNNs that predict and optimize human-centric processes.

8. Conclusion

Summary of Key Findings

The exploration of Graph Neural Networks (GNNs) within the context of business process mining has revealed their profound potential in addressing the inherent complexities of modern organizational workflows. Unlike traditional techniques, which often rely on sequential methods for process discovery, conformance checking, and performance analysis, GNNs excel in modeling the intricate interdependencies and relational dynamics that characterize contemporary business environments. By representing processes as graphs, GNNs allow for a holistic analysis of relationships between entities, offering a more nuanced understanding of process behaviors and outcomes. Key findings from this study underscore the ability of GNNs to handle diverse data types, capture both spatial and temporal dependencies, and provide enhanced predictive capabilities. These strengths render GNNs particularly effective in uncovering hidden patterns, optimizing workflows, and improving decision-making processes in ways that surpass the limitations of traditional business process mining methodologies.

Implications for Practice

The practical applications of GNNs extend far beyond theoretical advantages, presenting tangible benefits for businesses across various sectors. By facilitating a more comprehensive understanding of complex process interdependencies, GNNs empower organizations to optimize their operations, reduce inefficiencies, and proactively address potential bottlenecks. For instance, in supply chain optimization, GNNs can identify critical vulnerabilities and recommend adjustments to mitigate risks, ensuring smoother operations and reduced downtime. Similarly, their application in customer service process mining enhances service delivery by identifying and resolving inefficiencies in customer-agent interactions, ultimately improving customer satisfaction and retention.

From a cost reduction perspective, GNNs enable organizations to allocate resources more effectively by predicting future bottlenecks and recommending preemptive actions. Their ability to integrate diverse datasets into a unified analytical framework supports better-informed decision-making, aligning operational strategies with organizational objectives. Moreover, the scalability of GNNs, despite certain limitations, offers significant potential for adapting to dynamic business environments, where processes evolve continuously. These capabilities position GNNs as a transformative tool for businesses seeking to gain a competitive edge in an increasingly data-driven landscape.

Limitations of the Study

While this research highlights the considerable advantages of GNNs in business process mining, it is imperative to acknowledge its limitations. Chief among these is the scalability challenge associated with applying GNNs to large-scale and dynamic business systems. The computational complexity of GNN algorithms, coupled with the resource-intensive nature of graph construction and training, poses significant barriers to their deployment in real-world scenarios, particularly for organizations with limited technical infrastructure.

The interpretability of GNN models remains another critical challenge, as the black-box nature of these algorithms often obscures the rationale behind their predictions. This lack of transparency complicates the translation of model outputs into actionable insights, particularly in industries where compliance and explainability are paramount. Additionally, the generalization of GNN models across diverse industries and organizational contexts remains an unresolved issue, as process characteristics vary significantly between domains, necessitating extensive customization and domain-specific adaptations.

The limitations in data availability and quality also present significant obstacles. Many organizations struggle to compile the high-quality, structured datasets required for effective GNN training, particularly when dealing with fragmented or incomplete event logs. These challenges underscore the need for further research and innovation to enhance the scalability, interpretability, and generalizability of GNN models, as well as the development of standardized practices for data collection and preprocessing.

Concluding Remarks

The application of Graph Neural Networks in business process mining marks a paradigm shift in how organizations analyze and optimize their operations. By transcending the limitations of traditional methods, GNNs offer a powerful framework for uncovering complex relationships within business processes, enabling more efficient, cost-effective, and informed decision-making. Despite the challenges associated with their implementation, the transformative potential of GNNs is undeniable, promising to reshape the landscape of business analytics and process optimization in the years to come.

Future research in this field should focus on addressing the scalability and interpretability challenges of GNNs, developing more efficient algorithms capable of handling large, dynamic datasets while maintaining model transparency. Additionally, efforts to standardize data collection and preprocessing practices will be crucial in fostering broader adoption of GNNs across industries. The integration of GNNs with emerging technologies, such as reinforcement learning and explainable AI, offers exciting prospects for advancing their capabilities and unlocking new applications in business process mining. Through continued innovation and interdisciplinary collaboration, the field of GNN-based business process mining is poised to make significant strides, driving the next wave of organizational efficiency and innovation.

References

1. T. Kipf and M. Welling, "Semi-Supervised Classification with Graph Convolutional Networks," *arXiv preprint arXiv:1609.02907*, 2017.
2. W. Hamilton, Z. Ying, and J. Leskovec, "Inductive Representation Learning on Large Graphs," in *Advances in Neural Information Processing Systems (NIPS)*, 2017, pp. 1024–1034.
3. P. Velickovic et al., "Graph Attention Networks," *arXiv preprint arXiv:1710.10903*, 2018.
4. X. Ma, P. Cui, and W. Zhu, "Graph Neural Networks in Recommender Systems: A Survey," *arXiv preprint arXiv:2011.02260*, 2020.
5. Y. Wu et al., "Graph Neural Networks for Social Recommendation," in *Proceedings of the World Wide Web Conference (WWW)*, 2019, pp. 417–426.
6. M. G. de Leoni, W. M. P. van der Aalst, and M. Dees, "A General Framework for Correlating Observations of Multiple Objects," in *International Conference on Business Process Management (BPM)*, Springer, 2016, pp. 250–267.
7. W. M. P. van der Aalst, "Process Mining: Data Science in Action," 2nd ed., Springer, 2016.
8. A. Rozinat and W. M. P. van der Aalst, "Conformance Checking of Processes Based on Monitoring Real Behavior," *Information Systems*, vol. 33, no. 1, pp. 64–95, 2008.
9. J. Gilmer, S. S. Schoenholz, P. F. Riley, O. Vinyals, and G. E. Dahl, "Neural Message Passing for Quantum Chemistry," in *Proceedings of the International Conference on Machine Learning (ICML)*, 2017, pp. 1263–1272.

10. Z. Wu et al., "A Comprehensive Survey on Graph Neural Networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 1, pp. 4–24, 2021.
11. T. Schöning and B. H. van der Schaaf, "A Graph-Based Process Mining Approach for Operational Risk Analysis," in *Business Information Systems (BIS)*, Springer, 2019, pp. 162–176.
12. Y. Li, R. Yu, C. Shahabi, and Y. Liu, "Graph Convolutional Networks for Traffic Forecasting," in *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, 2018, pp. 388–395.
13. A. Vahdat and J. Kautz, "NVP: Neural Variational Process Mining," *arXiv preprint arXiv:2004.04850*, 2020.
14. D. Zhang et al., "Deep Process Mining: A Survey," *IEEE Access*, vol. 7, pp. 185746–185759, 2019.
15. P. Carnes and J. S. Dong, "A Graph Neural Network Model for Business Process Anomaly Detection," in *Pacific Asia Conference on Information Systems (PACIS)*, 2021, pp. 141–154.
16. J. Sun and X. Liu, "Graph Neural Networks for Event-Driven Process Mining," in *Proceedings of the ACM International Conference on Knowledge Discovery and Data Mining (KDD)*, 2020, pp. 1823–1832.
17. M. Dees and M. G. de Leoni, "Hierarchical Graph Representations for Process Mining," in *IEEE Transactions on Services Computing*, vol. 14, no. 1, pp. 124–136, 2021.
18. Y. Zhou et al., "Graph Representation Learning: State of the Art and Future Directions," *ACM Computing Surveys (CSUR)*, vol. 54, no. 1, pp. 1–37, 2021.
19. L. Da Xu, W. He, and S. Li, "Internet of Things in Industries: A Survey," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 4, pp. 2233–2243, 2014.
20. A. Joshi, M. Bhushan, and S. L. Mehta, "Challenges and Opportunities in Integrating IoT Data with Graph Neural Networks for Real-Time Business Analytics," in *International Conference on Business Analytics and Big Data (ICBABD)*, 2021, pp. 98–108.