RESEARCH ARTICLE

Graph Neural Networks in Supply Chain Analytics and Optimization: Concepts, Perspectives, Dataset and Benchmarks

Azmine Toushik Wasi[ac], MD Shafikul Islam[ac], Adipto Raihan Akib[a],
and Mahathir Mohammad Bappy[bc]

[a]Department of Industrial and Production Engineering, Shahjalal University of Science and Technology, Sylhet, Bangladesh;
[b]Department of Mechanincal and Industrial Engineering, Louisiana State University, Baton Rounge, LA, 70803, USA;
[c]Computational Intelligence and Operations Laboratory, Sylhet, Bangladesh

ABSTRACT
Graph Neural Networks (GNNs) have recently gained traction in transportation, bioinformatics, language and image processing, but research on their application to supply chain management remains limited. Supply chains are inherently graph-like, making them ideal for GNN methodologies, which can optimize and solve complex problems. The barriers include a lack of proper conceptual foundations, familiarity with graph applications in SCM, and real-world benchmark datasets for GNN-based supply chain research. To address this, we discuss and connect supply chains with graph structures for effective GNN application, providing detailed formulations, examples, mathematical definitions, and task guidelines. Additionally, we present a multi-perspective real-world benchmark dataset from a leading FMCG company in Bangladesh, focusing on supply chain planning. We discuss various supply chain tasks using GNNs and benchmark several state-of-the-art models on homogeneous and heterogeneous graphs across six supply chain analytics tasks. Our analysis shows that GNN-based models consistently outperform statistical Machine Learning and other Deep Learning models by around 10-30% in regression, 10-30% in classification and detection tasks, and 15-40% in anomaly detection tasks on designated metrics. With this work, we lay the groundwork for solving supply chain problems using GNNs, supported by conceptual discussions, methodological insights, and a comprehensive dataset.

1. Introduction

Supply chain is a dynamic network of organizations that participate in the various processes and activities that produce value in the form of products and services for consumers via upstream and downstream linkages (Mentzer et al. 2001), entailing a continuous flow of information, goods, and money among its various stages (Higuchi and Troutt 2004). As the supply chain consists of interconnected entities in a complex network, it involves intricate interdependencies and complex decision-making pro-
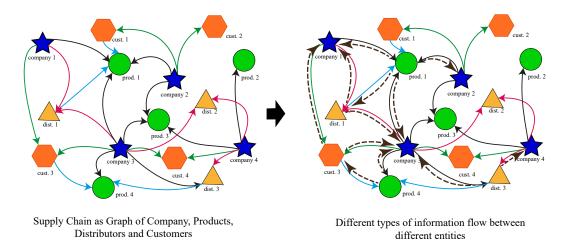
arXiv:2411.08550v1 [cs.LG] 13 Nov 2024

Supply Chain as Graph of Company, Products, Distributors and Customers

Different types of information flow between different entities

Figure 1. Supply Chain as a Graph of Interconnected Company, Products, Distributors and Customers.

cesses (Surana et al. 2005). Additionally, the modern supply chain generates enormous data, with relationships and dependencies between entities requiring sophisticated models to capture (Sadeghiamirshahidi et al. 2014). The potential benefits of using computational methods to solve supply chain problems include improved coordination, efficient logistics, and effective supply chain solutions (Chaovalitwongse and Furman 2013).

Graph Neural Networks (GNNs) are widely used across various sectors due to their ability to model complex relationships in data. In social networks, GNNs help analyze user connections for recommendations, information analysis and community detection Wasi et al. (2024a). In biology, they assist in drug discovery by modelling molecular interactions (Johnson et al. 2024; Wasi et al. 2024b). GNNs are used in finance for fraud detection (Wu, Chao, and Li 2024), in HR for job matching Wasi (2024), and in research for knowledge graph reasoning Kosasih et al. (2022) and anomaly detection Sudrich, Borges, and Beigl (2017); Zhou et al. (2023). Their strength lies in effectively capturing dependencies in non-Euclidean data, making them versatile for tasks involving connected entities. Connecting to supply chain, GNNs can enable the modelling of complex relationships and dependencies within the supply chain, facilitating tasks such as sales predictions, production planning, risk assessment, and uncovering hidden risks (Zhou et al. 2023; Kosasih et al. 2022). By leveraging GNN methodologies, it is possible to optimize supply chain operations, enhance risk management, and improve decision-making processes by extracting relevant information from graph data and inferring multiple types of hidden relationship risks (Kosasih et al. 2022).

Production planning plays a pivotal role in supply chain management by forecasting future product or service demand, aiding organizations in optimizing inventory levels, production schedules, and resource allocation (Aamer, Eka Yani, and Alan Priyatna 2020; Zougagh, Charkaoui, and Echchatbi 2020). The accuracy of demand prediction significantly impacts company revenue, prompting the exploration of diverse deep learning and machine learning models (Alves, Silva, and Mateus 2021; Pirhooshyaran and Snyder 2020; Pacella and Papadia 2021). While traditional models have shown promise, GNNs offer a unique advantage in modelling the network-like structures inherent in supply chains, such as global trade flows or social networks (Kosasih and Brintrup 2022), as shown in Figure 1. Despite limited prior studies on GNNs

2

in supply chains, recent research has demonstrated their utility in tasks like hidden link prediction to mitigate risk and uncover hidden dependencies (Aziz et al. 2021). However, several challenges remain: there is a lack of comprehensive conceptual foundations and formulations specific to supply chain applications of GNNs, and researchers often lack awareness of the diverse tasks that GNNs can address in this domain. Moreover, the scarcity of publicly available datasets and proper benchmarks hinders thorough evaluation and development of GNN models for supply chains. These gaps underscore the need for detailed methodologies and robust datasets to advance research and practical applications of GNNs in supply chain optimization.

To address these challenges, we take several key steps. First, we thoroughly discuss and connect supply chains with graph structures, formulating them as graphs for effective GNN application. We provide detailed definitions and explanations of how supply chain components map to graph properties, illustrating how GNNs can utilize these properties to tackle supply chain problems efficiently. Second, we introduce a new, multi-perspective benchmark dataset tailored for GNN applications in supply chain planning, accompanied by a comprehensive exploratory study of the dataset. Third, we explore a variety of supply chain tasks from different graph perspectives and benchmark state-of-the-art models on both homogeneous and heterogeneous graphs across six distinct analytics and modelling tasks. Our experiments demonstrate that GNN-based models significantly outperform traditional statistical and machine learning methods in various supply chain applications, underscoring the vital role of GNNs in advancing supply chain analysis and modelling.

The contributions of this research are summarized into these key aspects:

- We comprehensively discuss and connect supply chains with graphs, and formulate supply chains as graphs for effective application of GNNs. We define and explain how various supply chain components correspond to graph properties, and demonstrate how GNNs can leverage these properties to solve supply chain problems efficiently.
- We introduce a new, multi-perspective benchmark dataset for GNN problem-solving in supply chain planning, paving the way for extensive research and analytics using GNNs in supply chain management. Additionally, we present a comprehensive exploratory study of the dataset. The dataset is publicly available at `DOI: 10.5281/zenodo.13652826` under the CC BY 4.0 Licence.
- We describe and explore various tasks of supply chain from different graph perspectives like homogeneous graph, heterogeneous graph and hypergrphs; bench-marking the performance of state-of-the-art models on both homogeneous and heterogeneous graphs across six different supply chain analytics and modelling tasks on the dataset.
- Our dataset analysis and modelling experiments show that graph-based models outperform traditional statistical and machine learning methods in various supply chain applications, highlighting the critical role of GNNs in effective supply chain analysis and modelling.
- We also discuss real-life implications of the work, addressing their societal impacts and offering insights into potential future research directions; along with limitations of the current study and supply chain as graph methods.

With these contributions, we aid to lay the groundwork for advancing the application

of GNNs in supply chain contexts. The following sections will review the related works and provide a detailed analysis of machine learning and GNN approaches in supply chain management.

2.   Related Works

Supply chain management has seen significant advancements through the application of machine learning and graph neural networks. This section reviews the existing literature on machine learning approaches and the emerging role of GNNs in supply chain management, identifying research gaps and potential areas for future exploration.

2.1.   *Machine Learning (ML) Approaches in Supply Chain*

ML models aid in demand forecasting, inventory management, route optimization, supplier risk assessment, and quality control, leading to improved operational efficiency and cost reduction (Polo-Triana, Gutierrez, and Leon-Becerra 2024). By applying AI algorithms to the supply chain management system, organizations can achieve visualization, automation, and intelligent management of all supply chain links, ultimately improving responsiveness to market demands and reducing operating costs (Lin, Lin, and Wang 2022). Extensive research has been dedicated to harnessing ML (Feizabadi 2022; Zhu et al. 2021; Aamer, Eka Yani, and Alan Priyatna 2020). Numerous studies aim to improve demand prediction and optimize production processes through these methodologies (Nitsche et al. 2021; Younis, Sundarakani, and Alsharairi 2022; Filali, Filali et al. 2021).

Deep learning techniques, including ANNs and CNNs, have been widely explored in supply chains. ANNs are effective in modelling non-linear relationships and adapting to historical data, though their performance is influenced by parameters like neuron count and learning rate (Lunardi and Lima 2021; Mrad and Boujelbene 2019). While Random Forest classifiers often outperform ANNs in demand forecasting, ANNs remain close in accuracy (Vairagade et al. 2019). CNNs have been used for various supply chain tasks, with studies showing LSTM models generally outperform CNNs for grocery sales forecasting (Husna, Amin, and Shah 2021). CNNs are useful for improving inventory forecasts and reducing costs by enhancing supplier preparation (Tang and Ge 2022). Comparative analyses indicate that CNNs slightly outperform LSTMs in some datasets (Bousqaoui, Slimani, and Achchab 2021). Extreme gradient boosting models also achieve high precision with enhanced features (Lingelbach et al. 2021).

While no single model consistently demonstrates superior performance across all scenarios, various models excel in different situations based on the problem type and specific context. This underscores the potential of GNNs to provide robust solutions in scenarios where traditional models show variable performance.

2.2.   *Graph Neural Networks in Supply Chain*

Recent advancements in GNNs have enhanced demand forecasting and operational resilience, leading to more efficient and adaptive supply chain management. Graph representation learning has further improved link prediction by uncovering hidden dependencies within supply networks (Aziz et al. 2021; Kosasih and Brintrup 2022)

and building on previous GNN link prediction research (Zhang and Chen 2018, 2017; Cai et al. 2020; Cai and Ji 2020). These developments highlight the effectiveness of machine learning in refining demand forecasting and production planning. Notable contributions include GNN-based hidden link prediction for risk mitigation and integrating GNNs with knowledge graph reasoning to identify latent risks and extract insights (Aziz et al. 2021; Kosasih and Brintrup 2022; Kosasih et al. 2022).

GNNs have been applied for supplier recommendations by analysing network data to suggest alternative suppliers during disruptions (Tu et al. 2024). The Hierarchical Knowledge Transferable Graph Neural Network (HKTGNN) simplifies complex supply chains and uses a centrality-based knowledge transfer module for risk assessment (Zhou et al. 2023). GNNs have also enhanced the accuracy of industry classification based on supply chain data (Wu, Wang, and Olson 2023) and been used in heterogeneous graph neural networks for fraud detection and explanation in supply chain finance, leveraging multi-view information (Wu, Chao, and Li 2024).

2.3. *Research Gap Analysis*

While the potential of machine learning in supply chain management is widely acknowledged, the research focus has only recently shifted to Graph Neural Networks (GNNs) due to their unique capabilities in modelling supply chain networks (Song et al. 2021). However, the application of GNNs within supply chain research remains relatively nascent. Piraveenan et al. (2020) undertook an analysis of supply chains from a graph-theoretical standpoint, employing measures such as node centrality, clustering, and other statistical graph approaches. Notably absent from their study were any methods based on GNNs. Similarly, Wagner and Neshat (Wagner and Neshat 2010) pursued a comparable analysis using graph-theoretic techniques, yet they also did not incorporate GNN-based methodologies in their work.

Despite the growing recognition of GNNs' potential, progress in GNN-based supply chain research has been slow. This highlights the significance of specialized datasets and cutting-edge methodologies for fully utilizing GNNs' advantages in addressing complex demand prediction and production planning issues in the supply chain domain. Various supply chain problems have been addressed using GNNs, such as hidden link prediction tasks (Aziz et al. 2021; Kosasih and Brintrup 2022), supplier recommendation (Tu et al. 2024), risk mitigation (Kosasih et al. 2022), supply chain risk assessment (Zhou et al. 2023), industry classification (Wu, Wang, and Olson 2023), and fraud detection and explanation in supply chain finance (Wu, Chao, and Li 2024). However, production planning, one of the most critical tasks in supply chain management that ensures profitability and is influenced by numerous internal and external factors, has not yet been effectively addressed using GNNs. This represents a significant untapped potential for GNN-based methodologies to provide superior solutions in this domain.

To address this gap, we propose SCG, a benchmark dataset designed specifically for GNN-based supply chain analysis. This dataset aims to explore the potential of GNN models in production planning, demonstrating their capability to handle complex supply chain structures and improve decision-making processes. The introduction of this dataset is intended to spur further research and development in this promising area, leveraging GNNs to achieve more efficient and adaptive supply chain operations.

Following the review of related works, the terminology and foundational concepts

5

of graph theory are introduced. In the section 4, titled *Supply Chain Problem Formulation in Graphs*, the supply chain problem is rigorously defined and formulated within a graph-based framework. Subsequently, the *Data Collection* section details the process of gathering and organizing data from a leading FMCG company in Bangladesh, ensuring a robust dataset for analysis. This is followed by a detailed description and statistical analysis of the dataset. The *Dataset Potential and Practical Applications* section explores the dataset's versatility in addressing various supply chain challenges. Experimental results demonstrating the effectiveness of GNNs in this domain are showcased in the *Experiments* section. The study concludes with a discussion of the findings, their implications, and future research directions and possibilities.

3. Graph Neural Networks

Graphs can be categorized into different types, including homogeneous graphs, where all nodes and edges are of the same type, and heterogeneous graphs, which involve multiple types of nodes and edges. Each type is tailored for specific applications, depending on the structure and relationships represented. Homogeneous graphs are ideal for simpler relationships, such as a social network where nodes represent people and edges represent friendships. Heterogeneous graphs, however, feature multiple types of nodes and edges, capturing more complex interactions, like in a business network with nodes representing employees, departments, and projects. Hypergraphs extend this complexity further by allowing hyperedges to connect more than two nodes, useful for modelling multi-party interactions. Different types of GNNs designed for different types of graphs can be used in supply chain applications.

Here, we define and discuss graphs and different types of GNNs, along with their pros and cons, in relation to supply chain optimization applications.

3.1. *Definition of Graphs and GNNs*

A *graph* $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is a collection of *nodes* $\mathcal{V}$ and *edges* $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ between pairs of nodes. For the purpose of the following discussion, we will further assume the nodes to be endowed with s-dimensional *node features*, denoted by $\mathbf{x}_u$ for all $u \in \mathcal{V}$.

We consider a graph to be specified with an adjacency matrix $\mathbf{A}$ and node features $\mathbf{X}$. In the generic setting $\mathcal{E} \neq \emptyset$, the graph connectivity can be represented by the $n \times n$ *adjacency matrix* $\mathbf{A}$, defined as

$$a_{uv} = \begin{cases} 1 & (u, v) \in \mathcal{E} \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

Here, the adjacency and feature matrices $\mathbf{A}$ and $\mathbf{X}$ are synchronized, in the sense that a_{uv} specifies the adjacency information between the nodes described by the uth and vth rows of $\mathbf{X}$.

A *Graph Neural Network (GNN)* is a type of neural network designed to operate on graph structures. In a graph, nodes represent entities, and edges represent relationships between these entities. GNNs leverage the connections and relationships within the graph to learn node representations, enabling tasks like node classification, link prediction, graph classification, and many more. By aggregating information

from neighbouring nodes, GNNs can capture both local and global graph structure, making them effective for complex relational data (Bronstein et al. 2021; Hamilton 2020; Scardapane 2024).

3.2. *Classification of GNNs*

As discussed by Bronstein et al. (2021), GNNs can be described into three main types or, flavours. They are: Convolutional, Attentional, and Message-passing.

3.2.1. *Convolutional GNNs*

Convolutional GNNs (Kipf and Welling 2017; Defferrard, Bresson, and Vandergheynst 2016; Wu et al. 2019) aggregate features from neighbouring nodes using fixed weights. In this class, the features of the neighbourhood nodes are directly aggregated with fixed weights,

$$\vec{h}_u = \phi \left(\vec{x}_u, \bigoplus_{v \in \mathcal{N}_u} c_{uv} \psi(\vec{x}_v) \right). \tag{2}$$

Here, c_{uv} specifies the *importance* of node v to node u's representation. It is a constant that often directly depends on the entries, and $\mathbf{A}$ representing the structure of the graph.

This method is straightforward and computationally efficient, making it suitable for large-scale graphs. However, it often overlooks the varying importance of different neighbours, potentially leading to suboptimal performance in scenarios where certain relationships are more significant than others. In supply chain analytics, Convolutional GNNs can be used for tasks like demand forecasting and inventory optimization, where the relationships between various entities (e.g., suppliers and distributors) are relatively uniform.

3.2.2. *Attentional GNNs*

Attentional GNNs address the limitations of Convolutional GNNs by introducing learnable self-attention mechanisms that assign varying importance to different neighbours. In this flavour (Veličković et al. 2018; Monti et al. 2017; Zhang et al. 2018), the interactions are implicit:

$$\vec{h}_u = \phi \left(\vec{x}_u, \bigoplus_{v \in \mathcal{N}_u} a(\vec{x}_u, \vec{x}_v)\psi(\vec{x}_v) \right). \tag{3}$$

Here, a is a learnable *self-attention mechanism* that computes the importance coefficients $\alpha_{uv} = a(\vec{x}_u, \vec{x}_v)$ implicitly. They are often softmax normalised across all neighbours. When $\bigoplus$ is the summation, the aggregation is still a linear combination of the neighbourhood node features, but now the weights are feature-dependent.

This approach allows the model to focus on more relevant connections, improving prediction accuracy. The downside is the increased computational cost and complexity. In supply chain analytics, Attentional GNNs are particularly useful for anomaly detection and risk management, where the ability to prioritize critical relationships can lead to more accurate insights.

3.2.3. Message-passing GNNs

Message-passing GNNs (Gilmer et al. 2017; Battaglia et al. 2018) take a more flexible approach by allowing arbitrary messages to be passed across edges. It amounts to computing arbitrary vectors (*messages*) across edges,

$$\vec{h}_u = \phi \left(\vec{x}_u, \bigoplus_{v \in \mathcal{N}_u} \psi(\vec{x}_u, \vec{x}_v) \right).$$

(4)

Here, ψ is a learnable *message function*, computing v's vector sent to u, and the aggregation can be considered as a form of message passing on the graph.

This method is highly expressive and can capture complex dependencies within the graph. However, it can be computationally intensive and may require substantial training data to achieve good performance. In supply chain analytics, Message-passing GNNs can be applied to production forecasting and dynamic routing, where the ability to model intricate interactions is crucial.

3.3. *Temporal GNNs*

Another interesting and effective GNN model for supply chain analytics is the Temporal Graph Neural Network (Temporal GNN) (Zhao et al. 2018; Rossi et al. 2020; Longa et al. 2023). Temporal GNNs extend traditional GNNs by incorporating the dimension of time, allowing them to model dynamic graphs where relationships and features change over time. A general equation for Temporal GNNs can be expressed as:

$$\mathbf{h}_v^{(t+1)} = \sigma \left(\sum_{u \in \mathcal{N}(v)} \mathbf{A}_{uv}^{(t)} \mathbf{h}_u^{(t)} \mathbf{W}^{(t)} + \mathbf{B}^{(t)} \mathbf{h}_v^{(t)} \right)$$

(5)

where $\mathbf{h}_v^{(t)}$ represents the hidden state of node v at time t, $\mathcal{N}(v)$ denotes the neighbors of v, $\mathbf{A}_{uv}^{(t)}$ is the temporal adjacency matrix, $\mathbf{W}^{(t)}$ and $\mathbf{B}^{(t)}$ are learnable weight matrices, and σ is an activation function. Temporal GNNs are crucial for modelling dynamic systems where relationships and features evolve over time. In supply chain analytics, they enable the prediction of time-dependent phenomena such as demand fluctuations, shipment delays, and inventory levels. By capturing temporal patterns, these models provide insights into future trends, helping businesses optimize operations and anticipate disruptions.

Spatio-Temporal GNNs, a subset of Temporal GNNs, combine spatial and temporal information, making them exceptionally powerful for supply chain analytics. These models can capture not only the evolving relationships between entities but also their spatial dependencies (Zhang et al. 2018; Rozemberczki et al. 2021). In supply chains, Spatio-Temporal GNNs can significantly enhance predictive accuracy for tasks such as demand forecasting, dynamic route optimization, and real-time inventory management. Using both spatial and temporal dynamics, these models enable more informed decision-making, leading to improved efficiency and resilience in supply chain operations.
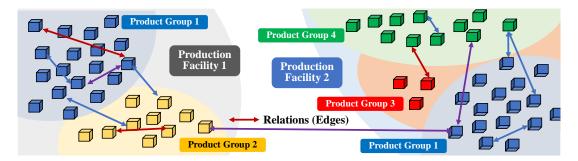
Figure 2. **Supply Chain Problem Formulation in Homogeneous Graph.** Boxes represent various product types, with colour indicating different groups. They are closely located based on product groups and production facilities. Different relational connections denote shared raw material requirements, interdependence between products, and other impacts.

4. Supply Chain Problem Formulation in Graphs

Supply chains consist of numerous interconnected components, including production facilities, products, and raw materials. There are various production facilities in an industry, each responsible for producing different types of products or products of different stock keeping units (SKUs). Each component, whether a different production facility or a different product group within the same production facility, can be represented as a node in a graph. Connections (edges) within this graphical model represent the relationships among the nodes. Products within the same group or produced in the same facility share connections due to common factors such as production capacity, raw material requirements, and demand trends. Additionally, production facilities may be interrelated due to shared resources, production dependencies, or logistical connections. By employing a graph-based approach, these complex relationships can be systematically analysed. This graphical representation allows for the integration of diverse data attributes for each node, including demand, production capacity, and sales metrics, thereby enabling a holistic analysis of the supply chain network.

To illustrate the concept, consider a potato chip manufacturing company. This company operates two production facilities and produces four types of chips: normal chips, triangular chips, ring chips, and string chips, each represented by different colours in Figure 2. Each type of chip is available in various packaging sizes, such as 30g, 50g, and 100g, depicted as different boxes in the figure. Within this graphical model, connections represent the relationships among the chips, facilities, and production processes. For instance, products within the same group or produced in the same facility share connections due to common factors such as production capacity, raw material requirements, and demand trends. If the production capacity for a chip group is limited to 500kg per day, increasing the production of 500g packets may necessitate reducing the output of 30g packets, demonstrating an internal group dependency. Moreover, if a machine can only produce one flavour at a time, producing multiple flavours on the same day would be inefficient due to the time lost in changeovers. If one facility specializes in a particular chip type, producing different types in the same facility may lead to inefficiencies. These interdependencies are depicted as connections within the same facility in Figure 2. Inter-facility relationships also exist, where producing a specific chip type in one facility may impact the production capabilities of another facility. Shared raw materials between product groups can create dependencies, where the production of one group affects the other.

By utilizing a graph-based approach, these complex relationships can be systemati-

9

cally analysed. GNNs leverage this structured representation to enhance production forecasting and planning, offering a comprehensive solution to supply chain management challenges.Formulating supply chain problems as graphs enables the visualization and analysis of intricate relationships among supply chain elements. Figure 3 illustrates a heterogeneous graph representing the supply chain network. It includes nodes for products (beige), plants (cyan), and storage locations (purple), with edges depicting the relationships between these entities. Solid lines indicate connections between products and plants, while dashed lines represent links between products and storage locations, showcasing the complexity and interdependencies within the supply chain.

4.1. *Defining Supply Chain as Graph*

Following discussions on supply chain as graph in Section 4 and general GNN formulations in Section 3.1, we can formulate a supply chain graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ by considering each product as a node $u \in \mathcal{V}$, and the connections between products—such as being manufactured in the same factory or belonging to the same product group or subgroup as edges $(u, v) \in \mathcal{E}$. The nodes are equipped with s-dimensional feature vectors $\mathbf{x}_u$, which capture the production and demand history of each product. This means that for each product (node), we have an associated feature vector $\mathbf{x}_u$ that contains historical data on its production quantities and demand patterns. The graph connectivity is represented by an $n \times n$ adjacency matrix $\mathbf{A}$. a_{uv} from $\mathbf{A}$ specifies the connectivity between the products represented by the uTh and vTh rows of $\mathbf{X}$. This representation helps in understanding and analysing the relationships and dependencies between different products in the supply chain, providing insights into production efficiency and demand forecasting.

5. SCG Dataset

Here, we discuss the development of the SCG dataset, detailing the processes of dataset collection, curation, feature selection, and exploratory statistics.

5.1. *Data Collection*

Data was collected from the central database system of one of the largest and most prominent FMCG (Fast Moving Consumer Goods) companies in Bangladesh. The dataset was reorganized for temporal graph utilization, extracting relevant nodes and features. The actual product names, product codes, and the name of the company are not disclosed to maintain the competitive standing of the concerned company. Table 1 summarizes the dataset information.

Here, we provide information on the quality control of the dataset, its coverage, and availability.

Quality Control. Each individual node, edge, and node feature undergoes a thorough manual examination and validation process. This involves scrutinizing the data for any irregularities, such as anomalies or missing information. Additionally, even zero values are meticulously reviewed to ensure their accuracy and legitimacy.

Temporal & Location Coverage. The occurrences in this database cover the period from January 1, 2023, to August 9, 2023, and include four features: production, sales
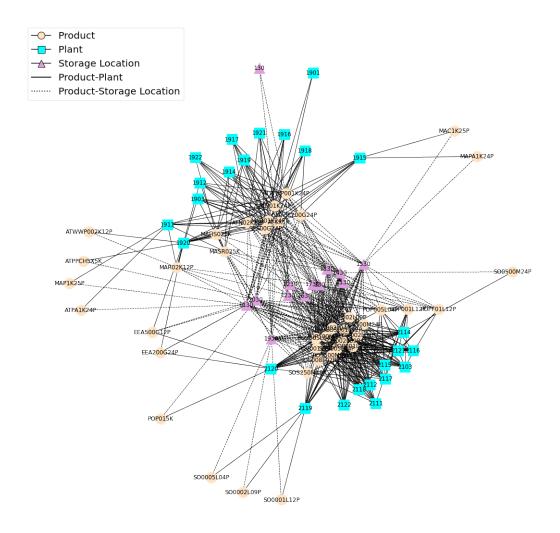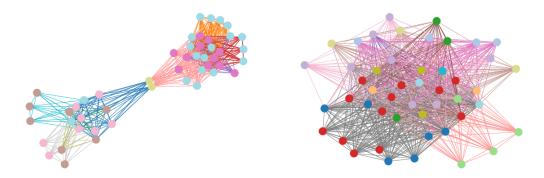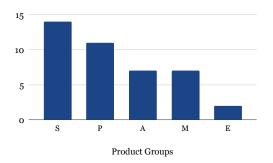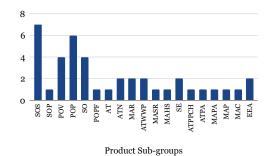
Figure 3. Heterogeneous Graph Example using SCG dataset. Products, plants, and storage locations are nodes and their relations are edges.

(a) Here, the nodes are sub-group products, and plants are edges. Colors denote different types of nodes and edges.

(b) Here, nodes plant-products, and storage locations are edges. Colors denote different types of nodes and edges.

Figure 4. Homogeneous Graph Examples using SCG dataset.

(a) Number of products in each product group.

(b) Number of products in each product sub-group.

Figure 5. Dataset Statistics.

Table 1. Dataset Edges Information.

Edge Classes		Edge Count	
Total Edge Types	62	Total Unique Edges	684
Class (Group)	5	Count (Group)	188
Class (Sub-group)	19	Count (Sub-group)	52
Class (Plant)	25	Count (Plant)	1647
Class (Storage)	13	Count (Storage)	3046

orders, delivery to distributors, and factory issues in two metrics: unit and weight produced in metric tons. The company operates across the entirety of Bangladesh.

Dataset Availability. . The dataset is publicly avail- able on GitHub at *https://github.com/CIOL-SUST/SCG* under the LGPL-2.1 Licence.

Graphs in Dataset. The dataset contains graphs in both homogeneous (see examples in Figure 4) and heterogeneous (see example in Figure 3) formats, with four temporal data points as node features. The construction, implementation, feature selection, and use of different graph properties depend on the specific application and requirements. In this paper, we provide a brief description of these aspects.
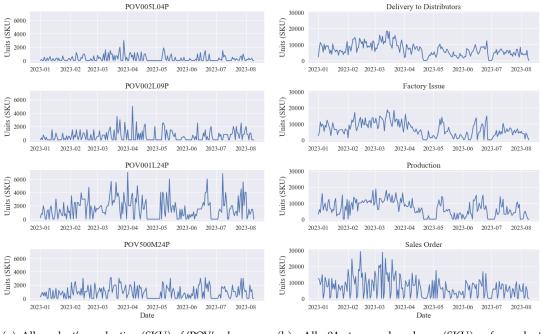
5.2. *Data Feature Description*

Within SCG, **nodes** pertains to distinct products, while **edges** represent various connections linking these products: the same product group or subgroup, the same plant or storage location. Figure 4 presents some examples of homogeneous graphs.

In the temporal data, node features include production, sales orders, delivery to distributors, and factory issues. We have all the temporal data in two modes: number of units produced (example: 500 units) and total weight (example: 10 metric tons) produced.

Definition of temporal features:

- *Production*, which quantifies product output considering sales orders, customer demand, vehicle fill rate, and delivery urgency. This quantity is typically measured in units or Metric Tons. For example, a production value might be 500 units or 10 metric tons produced in a given timeframe.
- *Sales Order* signifies distributor-requested quantities, pending approval from the accounts department. It reflects overall product demand. For example, a sales order might indicate a request for 300 units or 6 metric tons.

12

(a) All product's production (SKU) of 'POV' sub-group. (b) All 04 temporal values (SKU) of product 'SOS008L02P'.
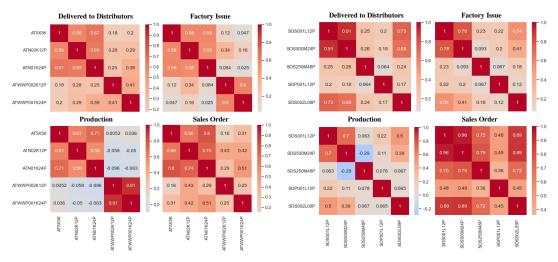
Figure 6. Data Samples.

- *Delivery to Distributors* denotes dispatched products aligning with orders, impacting company revenue significantly. For example, delivery data might show that 450 units or 9 metric tons have been sent to distributors.
- *Factory Issue* covers total products shipped from manufacturing facilities, with some going to distributors and the rest to storage warehouses. For example, factory issues might report that 700 units or 14 metric tons have been dispatched, with 500 units going to distributors and the remaining 200 units going to warehouses.

5.3. *Data Statistics*

5.3.1. *Product Group and Sub-groups*

Figure 5(a) shows that product group "S" has the highest number of products, reflecting its highest variety, and product group "E" has the lowest number of products. Notably, a majority of the product groups exhibit a range of around 8 to 10 products each. Figure 5(b) reveals the diversity within product subgroups. As previously noted, the "s" product group stands out for its remarkable variety, and this characteristic is consistently maintained within its corresponding subgroup. Among our product subgroups, "SOS" emerges as the one encompassing the largest array of product categories. Interestingly, despite the elevated diversity exhibited by the product group, the corresponding subgroup maintains a relatively stable count.

(a) Temporal correlation of some products from Product Group A.

(b) Temporal correlation of some products from Product Group A.

Figure 7. Temporal correlations.

5.3.2. *Analyzing Temporal Trends.*

Figure 6 (a), A particularly attractive pattern emerges: the production curve exhibits distinct and sharp fluctuations. Notably, the production unit of the company occasionally closes, which can be attributed to factors such as scheduled holidays, adjustments in the product's MRP rate, demand variations, and optimizations in transportation policies. In contrast, the other aspects of the graph, encompassing sales orders, delivery to distributors, and factory issuance, remain notably stable and consistent. Their minimal deviations stand in contrast to the variabilities characterizing the production curve, providing an insight into the dynamics of the supply chain activities.

In, Figure 6 (b), we can notice a very interesting pattern. The product is manufactured in large batches, with orders being accumulated over several days before production begins. The processes of delivering to distributors, handling factory issues, and fulfilling sales orders appear to be running smoothly and consistently. The company has opted to operate the production plant intermittently and maintain product inventory. When you observe straight lines in the production graph, it indicates that the plant is not actively engaged in production during those periods.

5.3.3. *Analysing Temporal Correlations.*

The correlation plot portrayed in Figure 7(a) serves to demonstrate the ways in which diverse subgroups of product "A" are interconnected in a multitude of facets. It is worth noting that the relationship between delivery to distributors and sales orders is particularly robust, whereas the production unit exhibits a distinct and prominent correlation pattern. Furthermore, a noteworthy observation pertains to the correlation between the production quantities of two subgroups, specifically "ATWWPOO2K12P" and "ATWWPOO1K24P". Additionally, it is worth highlighting that factory issues and sales orders similarly display correlation trends, albeit with some discernible variations.

Upon examining the correlation plot of the "S" subgroup in Figure 7(b), a distinct pattern becomes apparent: there is a significant and consistent correlation between

the "SOSOO1L12P" and "SOS500M24P" subgroups across all variables. Notably, an intriguing observation arises in the context of sales orders, where a majority of the products within the "S" subgroup demonstrate a relatively higher correlation compared to delivery to distributors, factory issuances, and production.

The comprehensive analysis of correlations and temporal activities strongly indicates a substantial opportunity for the application of graph-based machine learning methodologies in the realm of supply chains. Leveraging the inherent structures and relationships within supply chain networks, graph-based machine learning holds great promise in enhancing decision-making processes, optimizing resource allocation, and fostering adaptability in the dynamic landscape of supply chain operations.

6. Practical Applications of Supply Chain as Graphs

In this section, we explore the practical applications of representing supply chains as graphs. We illustrate how graph-based approaches can optimize various supply chain functions and improve overall efficiency and decision-making.

6.1. *Homogeneous Graphs*

In homogeneous form, our SCG dataset's structural design models products as nodes, while the interrelationships—such as shared product groups, subgroups, production facilities, and storage locations—are represented as edge connections. This setup enables the application of GNNs to tackle complex supply chain issues, such as:

- **Supply Chain Planning** involves the utilization of historical data pertaining to production and demand as node attributes, with the integration of time-related features to capture seasonality and trends. To enhance demand forecasting accuracy, related entities' influence can be harnessed through neighbourhood nodes. The hierarchical structure of nodes, including product groups, subgroups, plants, and storage locations, can be used for hierarchy-aware forecasting. In data-scarce scenarios, improved forecasting can be achieved by applying transfer learning across nodes of the same classification.
- **Product Classification** involves classifying products by product groups, subgroups, facilities; and also by economic profitability and production similarities for economic decision-making.
- **Product Relation Classification** involves classifying or predicting product relations in the supply chain graph. Product Relation detection involves detecting a missing edge or relation in the supply chain graph by a binary prediction objective.
- **Supply Chain Optimization** involves modelling goods flow between nodes, considering demand, lead times, and reorder points. Optimal routes and quantities can be recommended using edges as transportation links. Nodes representing plants can also aid in suggesting production adjustments based on demand projections, constraints, and capacity.
- **Anomaly Detection** is achieved by contrasting predicted demand with actual data to identify disruptions, stock outs, or unusual demand spikes. Deviations in demand from aggregated neighbourhood demand can signal inconsistencies within neighbourhoods, which can be detected as anomalies.
- **Event Classification** involves the training of GNNs to classify events based on

edges and nodes, such as production capacity changes, new product launches, and disruptions.

- **Fluctuation Detection** entails training GNNs to identify and classify fluctuations in the network based on impacts of global price hikes, supply chain issues, and disruptions. Fluctuations in supply chains refer to disruptions that can occur due to various factors and can significantly impact the operations of businesses. These disruptions can amplify negative shocks, affecting not just the firm experiencing the failure, but also its suppliers and customers, and even firms in other parts of the production network (Acemoglu and Tahbaz-Salehi 2020). Using temporal data, the model can recognize patterns in these fluctuations over time and determine if any demand and supply fluctuation can occur in the near future for better planning.
- **Combinatorial** (Cappart et al. 2022) and **Constrained** (Kotary et al. 2021) **Optimization** involves utilizing GNNs to optimize complex decision-making within the supply chain, considering various factors and constraints. GNNs can effectively model the intricate relationships between nodes and edges, enabling efficient allocation of resources, route planning, and inventory management, ultimately leading to improved performance.

6.2. *Heterogeneous Graphs*

All these tasks discussed can be executed using both heterogeneous and homogeneous graphs, but heterogeneous graphs excel in modelling complex systems. By incorporating diverse nodes (e.g., products, storage locations, production facilities) and edges representing their intricate relationships, heterogeneous graphs capture the complex interactions within supply chain networks. This approach enables a comprehensive representation of varying supply chain components and their interdependencies, addressing challenges and enhancing decision-making. Ultimately, it provides deeper insights into supply chain dynamics, allowing for optimized strategies and improved performance and resilience.

6.3. *Hypergraphs*

A hypergraph generalizes a graph by allowing edges, called hyperedges, to connect multiple vertices, rather than just two. This capability is useful for modelling complex relationships where interactions involve several entities. In supply chain management, hypergraphs can represent intricate relationships among products, storage locations, and production facilities through hyperedges. This approach captures the complexity of supply chain networks more effectively, offering a dynamic framework to manage the diverse characteristics and interdependencies of supply chain components. Utilizing hypergraphs can enhance understanding, performance, and resilience within supply chain management.

7. Experiments and Benchmarking

Here, we model our data with several benchmark graph-based models using PyTorch Geometric and PyTorch Geometric Temporal libraries (Rozemberczki et al. 2021). We use plant connections as edges in the graph to get predictions for all products,

together with 8:2 train-test split of the temporal dataset for forecasting, planning, and anomaly detection. In product classification, product relation detection, and classification, demands are used as node features. Product classification encompasses the significant task of categorizing products, while edge classification, or detection, involves discerning storage relationships. Our models underwent training for 200 epochs, with the possibility of early stopping if they converged ahead of time. Different metrics were employed depending on the specific task. It's important to note that we included non-neural network-based methods in our comparative analysis.

7.1. *Experimental Details*

Here, we discuss all the experimental details of our bench-marking process for each task on both homogeneous and heterogeneous graphs.

7.1.1. *Supply Chain Planning (Demand) and Forecasting*

In these tasks, we use a temporal approach where the node features display temporal and dynamic characteristics, while nodes and edges maintain static attributes to predict demand (Sales Order). The edges represent the similarity between different plants or production facilities, and the products are represented as nodes. In the case of traditional models, we independently model each product and subsequently aggregate the results. Conversely, with temporal GNNs, all predictions are generated simultaneously. It's worth noting that we have maintained the use of default hyper-parameters, as mentioned previously. MSE loss is used. Among the models, ARMA (Reinsel 1993) is a traditional statistical model, GRU (Chung et al. 2014) and LSTM (Hochreiter and Schmidhuber 1997) are deep learning based models, DCRNN (Li et al. 2018), TGCN (Zhao et al. 2018), GConvGRU (Seo et al. 2016) are GNN-based models. In heterogeneous GNNs, we use THGNN (Xiang et al. 2022), HTGNN (Fan et al. 2021) and HSTGCN (Zhong et al. 2021) with default settings. Our approach for production forecasting model is the same to demand prediction model, with the distinction being that the primary output objective was production rather than demand.

7.1.2. *Product Classification*

In the context of product classification, we harnessed the temporal features as static node attributes to forecast the product group, with the edges denoting the similarities between production facilities. The loss function was cross-entropy loss. Our Artificial Neural Network (ANN) model featured two linear layers, the K-Nearest Neighbors (KNN) algorithm incorporated a setting of 5 closest neighbours, and the Graph Attention Network (GAT) employed 4 attention heads with 3 layers in each GNN. Among the models, Logistic Regression is a statistical model, XGB Classifier (Chen and Guestrin 2016) is also a traditional boosting based model, KNN (Mucherino, Papajorgji, and Pardalos 2009) is a statistical clustering based model, ANN is a deep learning based model, GCN (Kipf and Welling 2017) and GAT (Veličković et al. 2018) are GNN-based models. In the heterogeneous task, we used temporal features as static node attributes to predict product groups, where edges represented similarities between production facilities.

17

7.1.3. Anomaly Detection

For anomaly detection, we employed sales orders as dynamic temporal node characteristics, applying a model with MSE loss to identify anomalies within time series data. Increased data fluctuations across the graph during a specific time frame are anomaly as an anomaly and warrant communication with the company. It is a supervised approach; following GDN (Deng and Hooi 2021). Among the models, PCA is a statistical model; ANN, AE (Aggarwal 2015) and LSTM-VAE (Park, Hoshi, and Kemp 2017) are deep learning based models; GANF (Dai and Chen 2022) and GDN (Deng and Hooi 2021) are GNN-based temporal anomaly detection models. In the heterogeneous task, we utilized sales order data as dynamic temporal node characteristics. Significant data fluctuations across the graph within a specific time frame are flagged as anomalies, requiring communication with the company. Ad-EHG (Sudrich, Borges, and Beigl 2017) and HRGCN (Li et al. 2023) are the heterogeneous GNN-based models here.

7.1.4. Product Relationship Detection and Classification

Product relation detection entailed a binary classification task focused on discerning the presence or absence of edges, while product relation classification involved a multi-class classification task. The loss function utilized in this context was the cross-entropy loss. As before, the ANN model featured two linear layers, the KNN algorithm incorporated a setting of 5 closest neighbours, and the GAT employed 4 attention heads with 3 layers in each GNN. In the heterogeneous task, it is also designed as a binary classification objective aimed at distinguishing the presence or absence of edges. In contrast, the product relation classification involved a multi-class classification goal. The loss function employed here was cross-entropy loss.

7.2. Results

Table 2 reveals that graph-based models consistently achieve higher scores compared to other models in the experiments. The experiments were performed multiple times and an average is taken.

7.2.1. Supply Chain Planning (Demand)

In demand prediction (see Table 2(a)), we observe a range of models, including traditional time series methods like ARMA and more advanced recurrent neural networks such as GRU and LSTM. Notably, the newer graph-based Graph Convolutional GRU (GConvGRU) achieves competitive performance. It's evident that the recurrent models, particularly DCRNN and TGCN, outperform ARMA, demonstrating the benefits of leveraging sequential information with GNNs. GNN-based models like GConvGRU also showcase their potential in modelling supply chain demand patterns. Heterogeneous graph segment also follows the same pattern.

7.2.2. Production Forecasting

Similar to demand prediction (see Table 2(c)), production forecasting shows a similar trend with respect to model performance. The recurrent models, DCRNN and TGCN, continue to outperform ARMA, emphasizing the superiority of GNN-based methods for this task. LSTM stands out as the best-performing individual model, closely followed by DCRNN and TGCN, signifying their utility in forecasting production levels

Table 2. Benchmark scores and Comparison of Various Models on Different Tasks.

Models and Type		(a) SC Planning (Demand) [Hm]		(b) SC Planning (Demand) [Ht]	
Type	Model	RMSE ($\downarrow$)	MAE ($\downarrow$)	RMSE ($\downarrow$)	MAE ($\downarrow$)
Statistical	ARMA	37.9875	486.8858	37.7776	486.6877
Deep Learning	GRU	33.2277	359.8285	32.3457	358.7857
Deep Learning	LSTM	30.7782	304.6777	30.2656	304.1456
GNN-based	DCRNN[Hm]/THGNN[Ht]	**27.9811**	**253.1135**	28.1633	253.3335
GNN-based	TGCN[Hm]/HTGNN[Ht]	28.0145	253.9384	**26.2424**	**251.2342**
GNN-based	GConvGRU[Hm]/HSTGCN[Ht]	28.1814	272.6487	29.2522	281.3533

Models and Type		(c) Production Forecasting [Hm]		(d) Production Forecasting [Ht]	
Type	Model	RMSE ($\downarrow$)	MAE ($\downarrow$)	RMSE ($\downarrow$)	MAE ($\downarrow$)
Statistical	ARMA	37.4542	483.5387	37.1332	480.2552
Deep Learning	GRU	32.3532	354.2312	33.5222	356.2322
Deep Learning	LSTM	30.2142	300.2942	30.1322	297.3422
GNN-based	DCRNN[Hm]/THGNN[Ht]	27.9434	**252.6459**	28.4745	256.6222
GNN-based	TGCN[Hm]/HTGNN[Ht]	**27.9244**	253.2390	**27.1232**	**249.3252**
GNN-based	GConvGRU[Hm]/HSTGCN[Ht]	28.0414	270.5629	29.9235	276.3662

Models and Type		(e) Product Cls. [Hm]		(f) Product Cls. [Ht]	
Type	Model	Accuracy	Precision	Accuracy	Precision
Statistical	Logistic Reg.	66.67%	76.67%	69.32%	73.24%
Boosting-based	XGB	65.56%	61.48%	71.45%	61.42%
Statistical	KNN	64.44%	64.07%	72.23%	74.07%
Deep Learning	ANN	66.37%	67.78%	74.67%	75.18%
GNN-based	GCN	71.23%	72.18%	81.35%	78.34%
GNN-based	GAT	**75.68%**	**75.32%**	**82.24%**	**79.56%**

Models and Type		(g) Product Relation Cls. [Hm]		(h) Product Relation Cls. [Ht]	
Type	Model	Precision	Recall	Accuracy	Precision
Statistical	Logistic Reg.	62.73%	66.78%	66.12%	67.18%
Boosting-based	XGB	71.78%	71.86%	71.22%	71.13%
Statistical	KNN	74.72%	74.75%	74.23%	73.73%
Deep Learning	ANN	76.58%	77.72%	77.23%	77.79%
GNN-based	GCN	**91.36%**	**82.27%**	88.23%	87.14%
GNN-based	GAT	89.43%	81.22%	**90.18%**	**91.02%**

Models and Type		(i) Product Relation Det. [Hm]		(j) Product Relation Det. [Ht]	
Type	Model	Accuracy	Precision	Accuracy	Precision
Statistical	Logistic Reg.	68.63%	75.15%	67.24%	68.37%
Boosting-based	XGB	71.23%	81.12%	71.25%	75.32%
Statistical	KNN	74.63%	78.26%	75.38%	75.73%
Deep Learning	ANN	78.74%	81.62%	78.26%	80.15%
GNN-based	GCN	89.32%	88.28%	**92.12%**	89.51%
GNN-based	GAT	**91.45%**	**91.64%**	91.27%	**90.15%**

Models and Type		(k) Anomaly Det. [Hm]		(l) Anomaly Det. [Ht]	
Type	Model	Precision	Recall	Precision	Recall
Statistical	PCA	31.33%	27.34%	27.24%	28.43%
Deep Learning	ANN	55.24%	47.75%	53.43%	49.24%
Deep Learning	AE	75.78%	71.78%	71.22%	69.32%
Deep Learning	LSTM-VAE	91.87%	86.58%	89.23%	87.24%
GNN-based	GANF[Hm]/Ad-EHG[Ht]	85.45%	76.33%	90.34%	89.13%
GNN-based	GDN[Hm]/HRGCN[Ht]	**94.15%**	**87.76%**	**92.43%**	**89.34%**

[Hm] Homogeneous Graph modelling, [Ht] Heterogeneous Graph modelling, Cls. = Classification, Det. = Detection

using GNNs. GNNs also outperforms other models in heterogeneous setup, too.

7.2.3. Product Classification

In product classification demonstrated in Table 2(e), traditional logistic regression achieves a respectable accuracy of 66.67%, but GNN-based models like GCN and GAT exhibit competitive performance with an accuracy of 71.23% and 75.68%, respectively. The performance of these models showcases their adaptability for supply chain product classification tasks.

In heterogeneous product classification (Table 2(f)), we observe improved performance compared to the homogeneous version. Traditional logistic regression achieves 69.% accuracy, while GNN-based models like heterogeneous GCN and GAT demonstrate competitive performance, achieving accuracies of 81.35% and 82.24%, respectively; showing effectiveness of GNN models.

7.2.4. Product Relation Classification

In product relation classification demonstrated in Table 2(g), showcases the efficiency of GNNs, with GCN and GAT achieving remarkable accuracy levels of 92.12% and 91.27%, respectively; better than homogeneous counterparts. They significantly outperform traditional methods and even other machine learning algorithms, demonstrating their prowess in classifying complex product relationships.

In product relation classification for heterogeneous graphs (Table 2(h)), GNNs like GCN and GAT achieve notable accuracy levels of 88.23% and 90.18%, respectively, though slightly lower than their homogeneous counterparts. But, these models excel over traditional methods and other machine learning algorithms, demonstrating their efficacy in classifying intricate product relationships.

7.2.5. Product Relation Detection

The task of detecting relationships between products is approached effectively by various models. Table 2(i) shows that, GCN and GAT outperform the traditional approaches, achieving an accuracy of 89.32% and 91.45%, respectively. These results underscore the capacity of GNNs to model intricate product relationships in the supply chain.

In product relation detection for heterogeneous graphs (Table 2(j)), GNNs such as GCN and GAT achieve impressive accuracy levels of 92.12% and 91.27%, respectively, surpassing their homogeneous counterparts. They outperform traditional methods and other machine learning algorithms, showcasing their effectiveness in classifying complex product relationships.

7.2.6. Anomaly Detection

Anomaly detection involves identifying deviations from expected patterns in time series. In this task, Table 2(k) shows that LSTM-VAE and GDN lead the way in terms of precision and recall, underlining their effectiveness in pinpointing anomalies. The GANF model also demonstrates strong performance, emphasizing the potential of GNN-based approaches for anomaly detection.

In heterogeneous anomaly detection, the goal is to detect deviations from expected patterns in dynamic heterogeneous graphs. Table 2(l) highlights that GNN-based HRGCN (Li et al. 2023) and Ad-EHG (Sudrich, Borges, and Beigl 2017) achieve

20

superior precision and recall, showcasing their effectiveness in identifying anomalies and highlighting the potential of GNN-based methods for anomaly detection.

7.2.7. *Comparison between Homogeneous and Heterogeneous Tasks*

The benchmark scores presented in Table 2 reveal that GNN-based models consistently outperform statistical and deep learning models across all tasks. Comparing heterogeneous and homogeneous graph modelling, heterogeneous models generally achieve better performance on average. For Supply Chain planning and production forecasting, GNN-based models like TGCN and HTGNN show lower RMSE and MAE in the heterogeneous versions. In product classification and relation classification, GNN-based models (GAT) achieve higher accuracy and precision in the heterogeneous setting. Similarly, for product relation detection and anomaly detection, heterogeneous models (GAT and HRGCN) outperform their homogeneous counterparts. This suggests that incorporating heterogeneous graph structures can capture more complex relationships, leading to improved predictive accuracy and anomaly detection capabilities.

8. Discussion

We believe this dataset can be a foundational cornerstone for using GNNs in supply chain analysis and problem-solving. The current landscape suffers from a lack of datasets tailored to the complexities of supply chain management. Our contribution addresses this gap by providing a comprehensive benchmark graph dataset specifically designed for supply chain applications, covering various tasks and challenges. This dataset enables the modelling of intricate network relationships, leading to more robust decision-making and optimization. By establishing standardized evaluation metrics and methodologies, we lay the groundwork for solving supply chain problems using GNNs, promoting advancements and innovation in the field.

Societal Impact. The societal impact of our work is significant. By improving supply chain analysis and optimization through GNNs, we can enhance efficiency in critical areas such as global trade, manufacturing, and logistics. This can lead to more reliable and resilient supply chains, reducing disruptions and ensuring timely delivery of goods. Enhanced supply chain management also has implications for environmental sustainability, as optimized logistics can reduce waste and lower carbon footprints. Furthermore, our dataset and methodologies can spur further research and development in both academia and industry, fostering innovation and creating new opportunities for economic growth. By addressing the current gaps in data and methodology, our work supports more informed decision-making and drives progress towards more sustainable and efficient supply chain practices.

Managerial Insights for Decision Makers in Industry. Our research provides valuable insights for decision-makers in the industry. By leveraging GNNs, managers can gain a deeper understanding of complex supply chain dynamics and make data-driven decisions that enhance efficiency and resilience. The benchmark dataset enables precise modelling of supply chain networks, facilitating better risk management and identification of hidden dependencies. Additionally, the improved accuracy in forecasting and anomaly detection aids in proactive problem-solving and resource optimization. Ultimately, adopting GNN-based solutions can lead to significant cost savings and improved operational performance.

Limitations.   We acknowledge that the dataset's temporal scope is limited, which constrains its ability to capture long-term supply chain dynamics. However, its concise timeframe is well-suited for analysing immediate responses and short-term intricacies in supply chain management. Despite this limitation, the dataset remains a valuable tool for addressing diverse supply chain challenges. The utility of the dataset depends on the chosen model or problem and varies based on how different information is represented as nodes or edges. It is up to researchers to determine the most effective way to leverage its potential.

Future Research Directions.   Future research directions in this field could explore the integration of more diverse and granular data sources to enhance the robustness and applicability of GNN models in supply chain management. Expanding the temporal scope and incorporating real-time data can provide deeper insights into long-term trends and immediate responses. Additionally, investigating the potential of heterogeneous modelling by including varied types of nodes, such as storage locations and production facilities, could lead to more comprehensive and impactful analyses. Exploring the intersection of GNNs with other advanced technologies, like reinforcement learning and AI-driven optimization, could further elevate decision-making processes. Finally, developing standardized benchmarks and evaluation metrics will be crucial for facilitating consistent and comparable advancements across different research efforts in this domain.

9.   Conclusion

In conclusion, this study presents a pioneering exploration of GNNs in the domain of supply chain management, addressing the current gap in research by connecting the inherently graph-like structure of supply chains with GNN methodologies. By formulating supply chains as graphs and leveraging GNNs, we have demonstrated significant performance improvements across multiple supply chain tasks, including regression, classification, detection, and anomaly detection. We introduce a new, multi-perspective benchmark dataset from a leading FMCG company in Bangladesh, which serves as a crucial resource for further research and analytics. This dataset provides a solid foundation for future studies in the application of GNNs to supply chain management. Our benchmarking demonstrates that GNN-based approaches consistently outperform traditional statistical and machine learning methods by 10-30% in regression and classification tasks, and by 15-40% in anomaly detection, underscoring their potential to enhance supply chain efficiency and decision-making. Furthermore, we discuss the real-world implications of applying GNNs to supply chains, considering their societal impacts and the limitations of our current study. This work not only lays the groundwork for future research but also highlights the transformative potential of GNNs in solving complex supply chain problems, ultimately contributing to the advancement of supply chain analytics and modelling.

Conflict of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper. All research procedures followed ethical guidelines, and the study was conducted with integrity and transparency. The authors have no financial, personal,

or other relationships that could inappropriately influence or bias the content of this work.

Data Availability

Data related to this work is publicly available at `DOI: 10.5281/zenodo.13652826` under the CC BY 4.0 Licence.

Author Contributions

ATW and MSI conceived the core idea, developed the methodology and theoretical aspects, prepared and processed raw data, conducted formal analysis, wrote and edited the paper, and led the project. ATW was responsible for designing and performing the experiments, conducting all analyses, and preparing visualizations. MSI handled the background study, data collection, and provided project supervision. ARA contributed to formal analysis and data collection. MMB provided guidance on manuscript preparation, writing, and revisions.

References

Aamer, Ammar, LuhPutu Eka Yani, and IMade Alan Priyatna. 2020. "Data analytics in the supply chain management: Review of machine learning applications in demand forecasting." *Operations and Supply Chain Management: An International Journal* 14 (1): 1–13.

Acemoglu, Daron, and Alireza Tahbaz-Salehi. 2020. *Firms, Failures, and Fluctuations: The Macroeconomics of Supply Chain Disruptions*. Working Paper 27565. National Bureau of Economic Research.

Aggarwal, Charu C. 2015. "Outlier Analysis." In *Data Mining*, 237–263. Springer International Publishing.

Alves, Julio César, Diego Mello da Silva, and Geraldo Robson Mateus. 2021. "Applying and comparing policy gradient methods to multi-echelon supply chains with uncertain demands and lead times." In *International Conference on Artificial Intelligence and Soft Computing*, 229–239. Springer.

Aziz, Ajmal, Edward Elson Kosasih, Ryan-Rhys Griffiths, and Alexandra Brintrup. 2021. "Data considerations in graph representation learning for supply chain networks." *arXiv preprint arXiv:2107.10609* .

Battaglia, Peter W, Jessica B Hamrick, Victor Bapst, Alvaro Sanchez-Gonzalez, Vinicius Zambaldi, Mateusz Malinowski, Andrea Tacchetti, et al. 2018. "Relational inductive biases, deep learning, and graph networks." *arXiv:1806.01261* .

Bousqaoui, Halima, Ilham Slimani, and Said Achchab. 2021. "Comparative analysis of short-term demand predicting models using ARIMA and deep learning." *International Journal of Electrical and Computer Engineering* 11 (4): 3319.

Bronstein, Michael M., Joan Bruna, Taco Cohen, and Petar Velickovic. 2021. "Geometric Deep Learning: Grids, Groups, Graphs, Geodesics, and Gauges." *CoRR* abs/2104.13478.

Cai, Lei, and Shuiwang Ji. 2020. "A Multi-Scale Approach for Graph Link Prediction." In *AAAI Conference on Artificial Intelligence*, .

Cai, Lei, Jundong Li, Jie Wang, and Shuiwang Ji. 2020. "Line Graph Neural Networks for Link Prediction." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 44: 5103–5113.

Cappart, Quentin, Didier Chételat, Elias Khalil, Andrea Lodi, Christopher Morris, and Petar Veličković. 2022. "Combinatorial optimization and reasoning with graph neural networks." .

Chaovalitwongse, Wanpracha Art, and Kevin C Furman. 2013. "Computational optimization and logistics challenges in industrial applications." *Annals of Operations Research* 203: 1–2.

Chen, Tianqi, and Carlos Guestrin. 2016. "XGBoost: A Scalable Tree Boosting System." In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, New York, NY, USA, 785–794. ACM.

Chung, Junyoung, Caglar Gulcehre, KyungHyun Cho, and Yoshua Bengio. 2014. "Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling." .

Dai, Enyan, and Jie Chen. 2022. "Graph-Augmented Normalizing Flows for Anomaly Detection of Multiple Time Series." In *International Conference on Learning Representations*, .

Defferrard, Michaël, Xavier Bresson, and Pierre Vandergheynst. 2016. "Convolutional neural networks on graphs with fast localized spectral filtering." *NIPS* .

Deng, Ailin, and Bryan Hooi. 2021. "Graph neural network-based anomaly detection in multivariate time series." In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 35, 4027–4035.

Fan, Yujie, Mingxuan Ju, Chuxu Zhang, Liang Zhao, and Yanfang Ye. 2021. "Heterogeneous Temporal Graph Neural Network." .

Feizabadi, Javad. 2022. "Machine learning demand forecasting and supply chain performance." *International Journal of Logistics Research and Applications* 25 (2): 119–142.

Filali, Aicha EL, Sanaa EL Filali, et al. 2021. "Exploring applications of Machine Learning for supply chain management." In *2021 Third International Conference on Transportation and Smart Technologies* (*TST*), 46–52. IEEE.

Gilmer, Justin, Samuel S Schoenholz, Patrick F Riley, Oriol Vinyals, and George E Dahl. 2017. "Neural message passing for quantum chemistry." *arXiv:1704.01212* .

Hamilton, William L. 2020. *Graph Representation Learning*. Springer International Publishing.

Higuchi, Toru, and Marvin D Troutt. 2004. "Dynamic simulation of the supply chain for a short life cycle product—Lessons from the Tamagotchi case." *Computers & Operations Research* 31 (7): 1097–1114.

Hochreiter, Sepp, and Jürgen Schmidhuber. 1997. "Long short-term memory." *Neural computation* 9 (8): 1735–1780.

Husna, Asma, Saman Hassanzadeh Amin, and Bharat Shah. 2021. "Demand forecasting in supply chain management using different deep learning methods." In *Demand forecasting and order planning in supply chains and humanitarian logistics*, 140–170. IGI Global.

Johnson, Ruth, Michelle M. Li, Ayush Noori, Owen Queen, and Marinka Zitnik. 2024. "Graph Artificial Intelligence in Medicine." *Annual Review of Biomedical Data Science* https://doi.org/10.1146/annurev-biodatasci-110723-024625.

Kipf, Thomas N., and Max Welling. 2017. "Semi-Supervised Classification with Graph Convolutional Networks." In *ICLR*, .

Kosasih, Edward Elson, and Alexandra Brintrup. 2022. "A machine learning approach for predicting hidden links in supply chain with graph neural networks." *International Journal of Production Research* 60 (17): 5380–5393.

Kosasih, Edward Elson, Fabrizio Margaroli, Simone Gelli, Ajmal Aziz, Nick Wildgoose, and Alexandra Brintrup. 2022. "Towards knowledge graph reasoning for supply chain risk management using graph neural networks." *International Journal of Production Research* 1–17.

Kotary, James, Ferdinando Fioretto, Pascal Van Hentenryck, and Bryan Wilder. 2021. "End-to-End Constrained Optimization Learning: A Survey." .

Li, Jiaxi, Guansong Pang, Ling Chen, and Mohammad-Reza Namazi-Rad. 2023. "HRGCN: Heterogeneous Graph-level Anomaly Detection with Hierarchical Relation-augmented Graph Neural Networks." .

Li, Yaguang, Rose Yu, Cyrus Shahabi, and Yan Liu. 2018. "Diffusion Convolutional Recurrent Neural Network: Data-Driven Traffic Forecasting." .

Lin, Haifeng, Ji Lin, and Fang Wang. 2022. "An innovative machine learning model for supply chain management." *Journal of Innovation & Knowledge* 7 (4): 100276.

Lingelbach, Katharina, Yannick Lingelbach, Sebastian Otte, Michael Bui, Tobias Künzell, and Matthias Peissner. 2021. "Demand forecasting using ensemble learning for effective schedul-

ing of logistic orders." In *Advances in Artificial Intelligence, Software and Systems Engineering: Proceedings of the AHFE 2021 Virtual Conferences on Human Factors in Software and Systems Engineering, Artificial Intelligence and Social Computing, and Energy, July 25-29, 2021, USA*, 313–321. Springer.

Longa, Antonio, Veronica Lachi, Gabriele Santin, Monica Bianchini, Bruno Lepri, Pietro Lio, Franco Scarselli, and Andrea Passerini. 2023. "Graph Neural Networks for temporal graphs: State of the art, open challenges, and opportunities." .

Lunardi, Antonio Ricardo, and Francisco Rodrigues Lima. 2021. "Comparison of artificial neural networks learning methods to evaluate supply chain performance." *Gestão & Produção* 28.

Mentzer, John T, William DeWitt, James S Keebler, Soonhong Min, Nancy W Nix, Carlo D Smith, and Zach G Zacharia. 2001. "Defining supply chain management." *Journal of Business logistics* 22 (2): 1–25.

Monti, Federico, Davide Boscaini, Jonathan Masci, Emanuele Rodola, Jan Svoboda, and Michael M Bronstein. 2017. "Geometric deep learning on graphs and manifolds using mixture model cnns." In *CVPR,* .

Mrad, Mariem, and Younes Boujelbene. 2019. "An Improved Neural Approaches for Forecasting Demand in Supply Chain Management." *International Journal of Computer Applications* 975: 8887.

Mucherino, Antonio, Petraq J. Papajorgji, and Panos M. Pardalos. 2009. *k-Nearest Neighbor Classification*, 83–106. New York, NY: Springer New York.

Nitsche, Anna-Maria, Christian-Andreas Schumann, Bogdan Franczyk, and Kevin Reuther. 2021. "Mapping supply chain collaboration research: a machine learning-based literature review." *International Journal of Logistics Research and Applications* 1–29.

Pacella, Massimo, and Gabriele Papadia. 2021. "Evaluation of deep learning with long short-term memory networks for time series forecasting in supply chain management." *Procedia CIRP* 99: 604–609.

Park, Daehyung, Yuuna Hoshi, and Charles C. Kemp. 2017. "A Multimodal Anomaly Detector for Robot-Assisted Feeding Using an LSTM-based Variational Autoencoder." .

Piraveenan, Mahendra, Hongze Jing, Petr Matous, and Yasuyuki Todo. 2020. "Topology of International Supply Chain Networks: A Case Study Using Factset Revere Datasets." *IEEE Access* 8: 154540–154559. https://doi.org/10.1109/ACCESS.2020.3015910.

Pirhooshyaran, Mohammad, and Lawrence V Snyder. 2020. "Simultaneous decision making for stochastic multi-echelon inventory optimization with deep neural networks as decision makers." *arXiv preprint arXiv:2006.05608* .

Polo-Triana, Sonia, Juan Camilo Gutierrez, and Juan Leon-Becerra. 2024. "Integration of machine learning in the supply chain for decision making: A systematic literature review." *Journal of Industrial Engineering and Management* 17 (2): 344–372.

Reinsel, Gregory C. 1993. *Vector ARMA Time Series Models and Forecasting*, 21–51. New York, NY: Springer US.

Rossi, Emanuele, Ben Chamberlain, Fabrizio Frasca, Davide Eynard, Federico Monti, and Michael Bronstein. 2020. "Temporal Graph Networks for Deep Learning on Dynamic Graphs." .

Rozemberczki, Benedek, Paul Scherer, Yixuan He, George Panagopoulos, Alexander Riedel, Maria Astefanoaei, Oliver Kiss, et al. 2021. "PyTorch Geometric Temporal: Spatiotemporal Signal Processing with Neural Machine Learning Models." In *Proceedings of the 30th ACM International Conference on Information and Knowledge Management*, 4564–4573.

Sadeghiamirshahidi, Narjes, Jafar Afshar, Ali Reza Firouzi, and SAHS Hassan. 2014. "Improving the efficiency of manufacturing supply chain using system dynamic simulation." *Jurnal Teknologi* 69 (2): 10–11113.

Scardapane, Simone. 2024. *Alice's Adventures in a Differentiable Wonderland – Volume I, A Tour of the Land*.

Seo, Youngjoo, Michaël Defferrard, Pierre Vandergheynst, and Xavier Bresson. 2016. "Structured Sequence Modeling with Graph Convolutional Recurrent Networks." .

Song, Xiang, Runjie Ma, Jiahang Li, Muhan Zhang, and David Paul Wipf. 2021. "Network in graph neural network." *arXiv preprint arXiv:2111.11638* .

Sudrich, Simon, Julio Borges, and Michael Beigl. 2017. "Anomaly Detection in Evolving Heterogeneous Graphs." In *2017 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*, 1147–1149.

Surana, Amit, Soundar Kumara*, Mark Greaves, and Usha Nandini Raghavan. 2005. "Supply-chain networks: a complex adaptive systems perspective." *International Journal of Production Research* 43 (20): 4235–4265.

Tang, Zheng, and Yuemeng Ge. 2022. "CNN model optimization and intelligent balance model for material demand forecast." *International Journal of System Assurance Engineering and Management* 13 (Suppl 3): 978–986.

Tu, Yuchun, Wenxin Li, Xiao Song, Kaiqi Gong, Lu Liu, Yunhao Qin, Songsong Liu, and Ming Liu. 2024. "Using graph neural network to conduct supplier recommendation based on large-scale supply chain." *International Journal of Production Research* 1–14.

Vairagade, Navneet, Doina Logofatu, Florin Leon, and Fitore Muharemi. 2019. "Demand forecasting using random forest and artificial neural network for supply chain management." In *Computational Collective Intelligence: 11th International Conference, ICCCI 2019, Hendaye, France, September 4–6, 2019, Proceedings, Part I 11*, 328–339. Springer.

Veličković, Petar, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò, and Yoshua Bengio. 2018. "Graph Attention Networks." In *ICLR*, .

Wagner, Stephan M., and Nikrouz Neshat. 2010. "Assessing the vulnerability of supply chains using graph theory." *International Journal of Production Economics* 126 (1): 121–129. Improving Disaster Supply Chain Management – Key supply chain factors for humanitarian relief, https://doi.org/https://doi.org/10.1016/j.ijpe.2009.10.007.

Wasi, Azmine Toushik. 2024. "HRGraph: Leveraging LLMs for HR Data Knowledge Graphs with Information Propagation-based Job Recommendation." In *Proceedings of the 1st Workshop on Knowledge Graphs and Large Language Models (KaLLM 2024)*, Bangkok, Thailand, August, 56–62. Association for Computational Linguistics.

Wasi, Azmine Toushik, Taki Hasan Rafi, Raima Islam, and Dong-Kyu Chae. 2024a. "BanglaAutoKG: Automatic Bangla Knowledge Graph Construction with Semantic Neural Graph Filtering." In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, Torino, Italia, May, 2100–2106. ELRA and ICCL.

Wasi, Azmine Toushik, Taki Hasan Rafi, Raima Islam, Serbetar Karlo, and Dong-Kyu Chae. 2024b. "CADGL: Context-Aware Deep Graph Learning for Predicting Drug-Drug Interactions." .

Wu, Bin, Kuo-Ming Chao, and Yinsheng Li. 2024. "Heterogeneous graph neural networks for fraud detection and explanation in supply chain finance." *Information Systems* 121: 102335.

Wu, Desheng, Quanbin Wang, and David L Olson. 2023. "Industry classification based on supply chain network information using Graph Neural Networks." *Applied Soft Computing* 132: 109849.

Wu, Felix, Amauri Souza, Tianyi Zhang, Christopher Fifty, Tao Yu, and Kilian Weinberger. 2019. "Simplifying graph convolutional networks." In *ICML*, .

Xiang, Sheng, Dawei Cheng, Chencheng Shang, Ying Zhang, and Yuqi Liang. 2022. "Temporal and Heterogeneous Graph Neural Network for Financial Time Series Prediction." In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*, CIKM '22, October. ACM.

Younis, Hassan, Balan Sundarakani, and Malek Alsharairi. 2022. "Applications of artificial intelligence and machine learning within supply chains: systematic review and future research directions." *Journal of Modelling in Management* 17 (3): 916–940.

Zhang, Jiani, Xingjian Shi, Junyuan Xie, Hao Ma, Irwin King, and Dit-Yan Yeung. 2018. "Gaan: Gated attention networks for learning on large and spatiotemporal graphs." *arXiv:1803.07294* .

Zhang, Muhan, and Yixin Chen. 2017. "Weisfeiler-Lehman Neural Machine for Link Prediction." *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* .

Zhang, Muhan, and Yixin Chen. 2018. "Link prediction based on graph neural networks." *Advances in neural information processing systems* 31.

Zhao, Ling, Yujiao Song, Min Deng, and Haifeng Li. 2018. "Temporal Graph Convolutional Network for Urban Traffic Flow Prediction Method." *ArXiv* abs/1811.05320.

Zhong, Weida, Qiuling Suo, Xiaowei Jia, Aidong Zhang, and Lu Su. 2021. "Heterogeneous Spatio-Temporal Graph Convolution Network for Traffic Forecasting with Missing Values." In *2021 IEEE 41st International Conference on Distributed Computing Systems* (*ICDCS*), 707–717.

Zhou, Zhanting, Kejun Bi, Yuyanzhen Zhong, Chao Tang, Dongfen Li, Shi Ying, and Ruijin Wang. 2023. "HKTGNN: Hierarchical Knowledge Transferable Graph Neural Network-based Supply Chain Risk Assessment." In *2023 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking* (*ISPA/BDCloud/SocialCom/SustainCom*), 772–782. IEEE.

Zhu, Xiaodan, Anh Ninh, Hui Zhao, and Zhenming Liu. 2021. "Demand forecasting with supply-chain information and machine learning: Evidence in the pharmaceutical industry." *Production and Operations Management* 30 (9): 3231–3252.

Zougagh, Nisrine, Abdelkabir Charkaoui, and Abdelwahed Echchatbi. 2020. "Prediction models of demand in supply chain." *Procedia computer science* 177: 462–467.