Article

Attention-Based Temporal Convolutional Networks and Reinforcement Learning for Supply Chain Delay Prediction and Inventory Optimization

Tian Jin [*]

Article

Attention-Based Temporal Convolutional Networks and Reinforcement Learning for Supply Chain Delay Prediction and Inventory Optimization

Tian Jin

Georgia Institute of Technology, Atlanta, USA; jintian2937@gmail.com

Abstract: Accurately predicting supply chain delays is crucial for maintaining resilience in inventory management systems. This paper presents a novel framework integrating an Attention-Based Temporal Convolutional Network (ATCN) and Reinforcement Learning (RL) to address the challenges of supply chain delay prediction and inventory optimization. The ATCN model leverages convolutional layers to capture long-term dependencies in time series data, while the attention mechanism enhances prediction accuracy under high demand volatility. Additionally, the RL component optimizes inventory decisions by minimizing holding and shortage costs through multi-agent collaboration. Our proposed approach outperforms baseline models in terms of Mean Absolute Error (MAE), Mean Squared Error (MSE), R², and Area Under the Curve (AUC) metrics. Ablation studies confirm the importance of both the attention mechanism and RL in improving forecasting accuracy and inventory management. This work offers a robust solution for enhancing supply chain resilience and lays the groundwork for future exploration in multi-layered and real-time supply chain systems.

Keywords: supply chain delay; temporal convolutional networks; reinforcement learning; inventory optimization; multi-agent systems

---

1. Introduction

Supply chain management is crucial for modern businesses, yet predicting delays and optimizing inventory under changing demand remains complex. Traditional models, such as autoregressive and basic machine learning, often fail to capture supply chain dynamics effectively, and static inventory policies struggle to adapt, resulting in overstock or stockouts.

This study introduces a framework combining an Attention-based Temporal Convolutional Network (ATCN) and Reinforcement Learning (RL) for improved delay prediction and inventory management. ATCN's convolutional layers capture long-term dependencies in time series, with an attention mechanism highlighting essential trends, particularly under irregular demand. This mechanism enables better anomaly detection and forecasting, essential in fluctuating demand conditions.

The RL component dynamically optimizes inventory by managing reordering and resource allocation. With multi-agent reinforcement learning, each node in the supply chain independently learns policies that enhance global inventory efficiency. Coordination among agents facilitates effective inventory control across the supply chain.

By integrating ATCN and RL, this approach bridges the gap between forecasting and inventory decisions. RL refines ATCN's predictions, optimizing both forecast accuracy and inventory actions through a hybrid loss function that balances forecasting error and inventory costs, offering a holistic solution to supply chain challenges.

This work advances ATCN's application in capturing demand shifts and demonstrates the potential of multi-agent RL for decentralized inventory management, outperforming traditional models in predictive accuracy and inventory control.

2. Related Work

Recent advances in supply chain management have leveraged machine learning (ML) and deep reinforcement learning (DRL) to address critical challenges such as inventory optimization, demand forecasting, and disruption management. Several studies have highlighted the growing importance of these techniques in improving supply chain resilience and performance.

The study by Siyue Li [1], "Harnessing Multimodal Data and Multi-Recall Strategies for Enhanced Product Recommendation in E-Commerce," influenced the design of our ATCN model by inspiring advanced data integration and recall strategies. Li's techniques enhanced our attention mechanism, improving delay prediction accuracy in complex supply chain systems.

Kegenbekov and Jackson [2] introduced a deep reinforcement learning model that synchronizes supply and demand by optimizing inbound and outbound flows. Their model demonstrated adaptability in stochastic and nonstationary environments, outperforming traditional methods due to its dynamic learning approach.

Lu's work on multi-objective recommendations using ensemble learning [3] directly enhances our model by demonstrating effective methods to handle multi-objective prediction tasks with imbalanced data and session variability. The ensemble approach informs our ATCN-RL framework, boosting predictive accuracy and robustness in supply chain delay and inventory optimization.

Reinforcement learning has been applied to inventory management by Bharti et al. [4] , who utilized RL algorithms to determine optimal inventory policies. Their model focused on minimizing holding and shortage costs, providing a scalable solution for complex multi-echelon supply chains.

Pauli et al. [5] explored a machine learning-based method for optimal delay assignment in cyber-physical systems, which could be adapted to supply chains. Their study addressed the importance of accurate delay prediction and optimization for improving overall system efficiency.

Ivanov [6] contributed to the discussion on supply chain resilience by integrating agility, sustainability, and resilience perspectives. His "viable supply chain model" offered insights into how supply chains could remain adaptive and robust in the face of disruptions like the COVID-19 pandemic. In the work by Li et al. [7], Strategic Deductive Reasoning in Large Language Models: A Dual-Agent Approach, our research provided notable technical influence. Specifically, the integration of Attention-Based Temporal Convolutional Networks (ATCN) in our framework inspired their use of attention mechanisms for refining temporal dependencies in strategic reasoning tasks. Additionally, our application of multi-agent reinforcement learning for decentralized inventory management contributed insights into their collaborative dual-agent system, enhancing decision-making and predictive robustness.

Harsha et al. [8] applied AI to optimize multi-echelon inventory systems. Their model, based on deep reinforcement learning, demonstrated significant improvements over heuristic-based approaches, especially in complex, high-dimensional supply chains.

The study by Jiaxin Lu [9], "Enhancing Chatbot User Satisfaction: A Machine Learning Approach Integrating Decision Tree, TF-IDF, and BERTopic," directly informed the integration of attention mechanisms in our ATCN model. Lu's innovative feature extraction and dynamic focus strategies influenced the design of our attention-based architecture, enhancing the predictive accuracy and robustness of supply chain delay forecasts under high demand volatility.

Finally, Gabellini et al. [10] developed a deep learning approach for predicting delivery delay risks in the automotive sector. Their model, which incorporated macroeconomic data, outperformed traditional methods by accurately capturing autocorrelations in data, offering robust solutions for delivery risk mitigation.

3. Methodology

This study proposes a predictive model leveraging historical inbound and outbound data to address supply chain delay challenges. The ensemble model forecasts order delays and enhances supply chain transparency, evaluated using metrics such as AUC (ROC) and GAUC for comprehensive

performance assessment. Key aspects of data preprocessing, model architecture, and results are discussed.

Focusing on critical features like port of origin, third-party logistics, customs processes, and product attributes, the model predicts supply chain delays. Emphasis is placed on challenges in data preprocessing, model development, and performance evaluation.

3.1. Model Network

To optimize supply chain inventory, we propose a deep learning model integrating an Attention-based Temporal Convolutional Network (ATCN) with a Reinforcement Learning (RL) agent. This architecture captures temporal supply chain dynamics and optimizes inventory policies across stages, from demand forecasting to stock allocation. The model pipeline is shown in Figure 1.
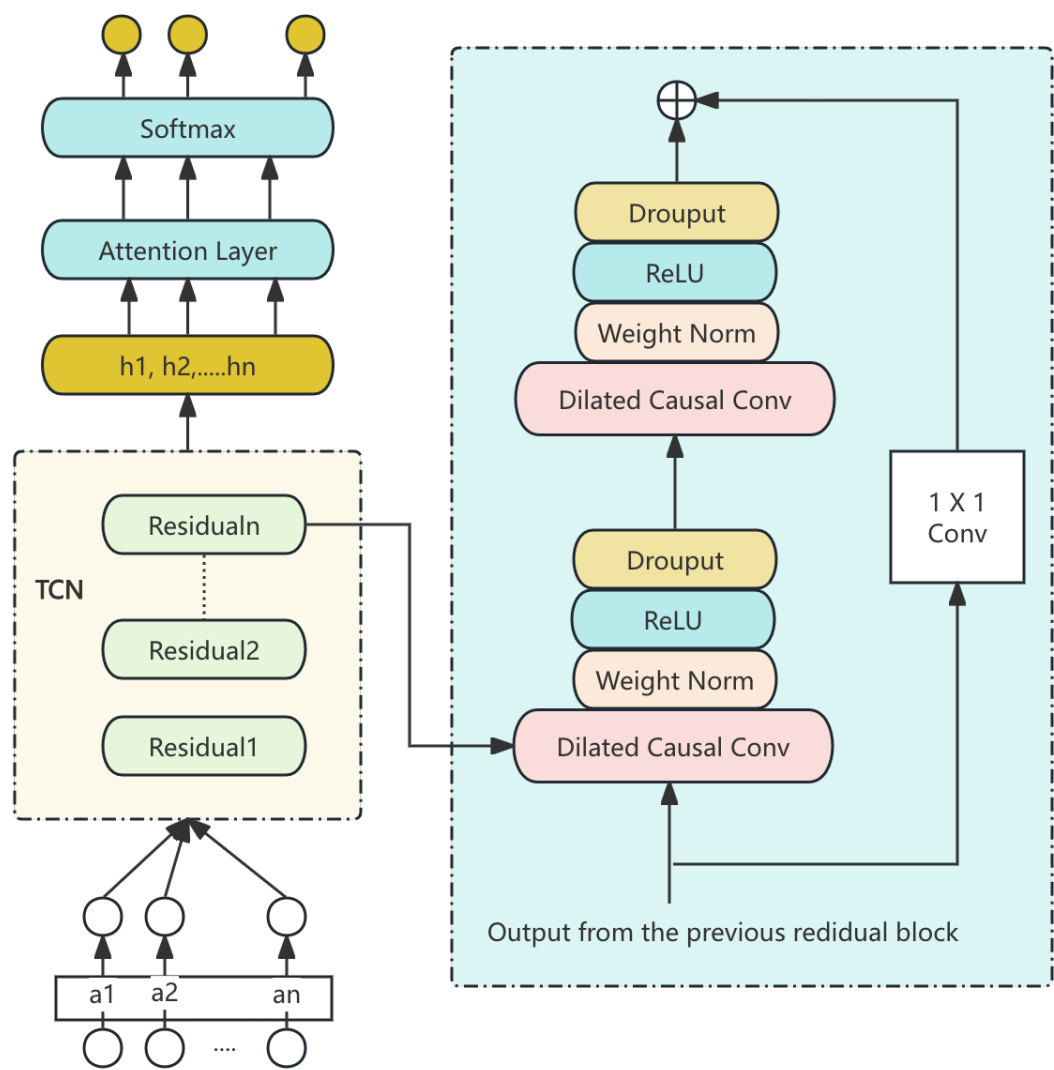

Figure 1. Enter Caption.

3.2. Attention-Based Temporal Convolutional Network (ATCN)

The ATCN component captures long-term dependencies in supply chain data, such as seasonal trends and demand fluctuations. TCNs with dilated convolutions process sequences efficiently, while the attention mechanism enhances focus on key time steps, improving forecasting accuracy.

The ATCN network is represented as:

$$h_t = \sum_{i=1}^{n} \alpha_{t,i} \cdot TCN(x_i; W) \tag{1}$$

where x_i is the input sequence, $\alpha_{t,i}$ the attention weights, and W the TCN weight matrix. Attention weights are calculated as:

$$\alpha_{t,i} = \frac{\exp(e_{t,i})}{\sum_{j=1}^{n} \exp(e_{t,j})} \tag{2}$$

with $e_{t,i}$ as the alignment score.

The TCN component, with dilated convolutions, ensures that the model can capture information over long time periods:

$$h_t = \sum_{k=1}^{K} W_k \cdot x_{t-d_k} \tag{3}$$

where d_k is the dilation factor, controlling how far back in time the convolution looks. The structure of Generic TCN is shown in Figure 2.
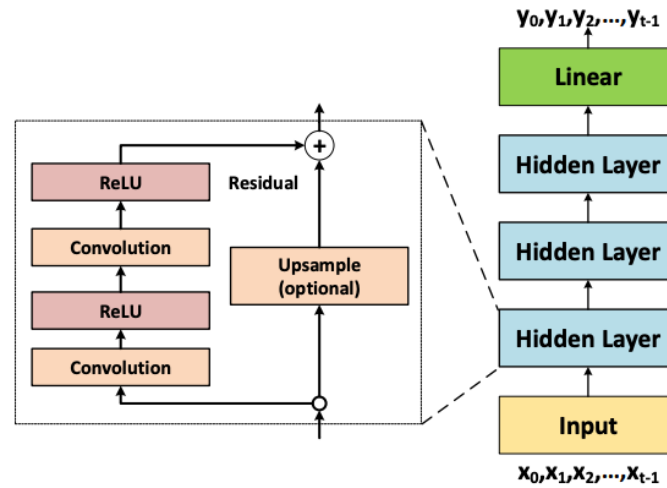

Figure 2. The structure of Generic TCN.

3.3. Reinforcement Learning Agent for Inventory Optimization

The reinforcement learning agent optimizes inventory decisions (e.g., reorder points, stock allocation) based on ATCN's demand forecasts. Modeled as a Markov Decision Process (MDP), the state space includes inventory levels, demand forecasts, and lead times. The agent aims to minimize holding and shortage costs while meeting service levels.

The action-value function is:

$$Q(s_t, a_t) = \mathbb{E}\left[r_t + \gamma \max_{a'} Q(s_{t+1}, a') \right] \tag{4}$$

where s_t is the state, a_t the action, r_t the reward, and γ the discount factor. The policy $\pi(a|s)$ is optimized with deep Q-learning.

A replay buffer stores experiences (s_t, a_t, r_t, s_{t+1}), updating with mini-batches:

$$\theta \leftarrow \theta + \alpha \nabla_\theta \sum_i (y_i - Q(s_i, a_i; \theta))^2 \tag{5}$$

where α is the learning rate and $y_i = r_i + \gamma \max_{a'} Q(s_{i+1}, a'; \theta^-)$ is the target Q-value.

3.4. Multi-Agent System for Distributed Supply Chain Management

To capture the distributed nature of supply chains, we extend the RL framework to a multi-agent reinforcement learning (MARL) system, where each node (e.g., warehouses, distribution centers) is an independent agent learning inventory policies locally. Agents collaborate for global optimization.

The reward function balances local and global objectives:

$$r_{t,i} = -\Big(C_{holding,i}(t) + C_{shortage,i}(t) \\ + \lambda \sum_j C_{delay,j}(t) \Big) \tag{6}$$

where $C_{holding,i}(t)$, $C_{shortage,i}(t)$, and $C_{delay,j}(t)$ represent costs, and λ balances local and global goals.

Each agent optimizes locally and shares experiences, while a central coordinator adjusts incentives for cooperation, enhancing supply chain efficiency.

3.5. Hybrid Loss Function for Multi-Objective Optimization

To address the multi-objective goals of minimizing costs and maximizing service levels, we propose a hybrid loss function combining prediction and optimization objectives:

$$\mathcal{L}_{\text{hybrid}} = \mathcal{L}_{\text{forecast}} + \beta \mathcal{L}_{\text{RL}} \tag{7}$$

where $\mathcal{L}_{\text{forecast}}$, defined as Mean Squared Error (MSE), is:

$$\mathcal{L}_{\text{forecast}} = \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 \tag{8}$$

and $\mathcal{L}_{\text{RL}}$ is the reinforcement learning loss. The parameter β balances forecasting accuracy with inventory optimization.

4. Data Preprocessing

Data preprocessing is critical for building a robust supply chain optimization model. The raw dataset, containing temporal and categorical data from multiple sources, undergoes steps including missing value handling, feature engineering, categorical encoding, scaling, and temporal processing to prepare it for model training.

4.1. Handling Missing Data

Supply chain datasets often have missing values due to incomplete records. In this study, missing values in 'units', 'weight', and 'distance' were addressed as follows:

Numerical Features: For continuous variables ('weight' and 'distance'), median imputation was used:

$$x_{\text{clean}} = \begin{cases} x_i & \text{if } x_i \text{ is not missing} \\ \text{median}(X) & \text{otherwise} \end{cases} \tag{9}$$

where x_i is the observed value, and median(X) is the median of non-missing values.

Categorical Features: For categorical variables such as '3pl', 'origin_port', and 'logistic_hub', we used mode imputation, replacing missing values with the most frequent category:

$$x_{\text{clean}} = \begin{cases} x_i & \text{if } x_i \text{ is not missing} \\ \text{mode}(X) & \text{otherwise} \end{cases} \tag{10}$$

where mode(X) is the most frequent category in the feature.

4.2. Outlier Detection and Treatment

Outliers in 'units' and 'weight' can bias the model. Using the Z-score method, data points with Z-scores above 3 were identified as outliers:

$$Z(x_i) = \frac{x_i - \mu}{\sigma} \tag{11}$$

where μ is the mean and σ the standard deviation. Outliers were capped at the 99th percentile or removed based on impact. Figure 3 shows outlier trends using box plots.
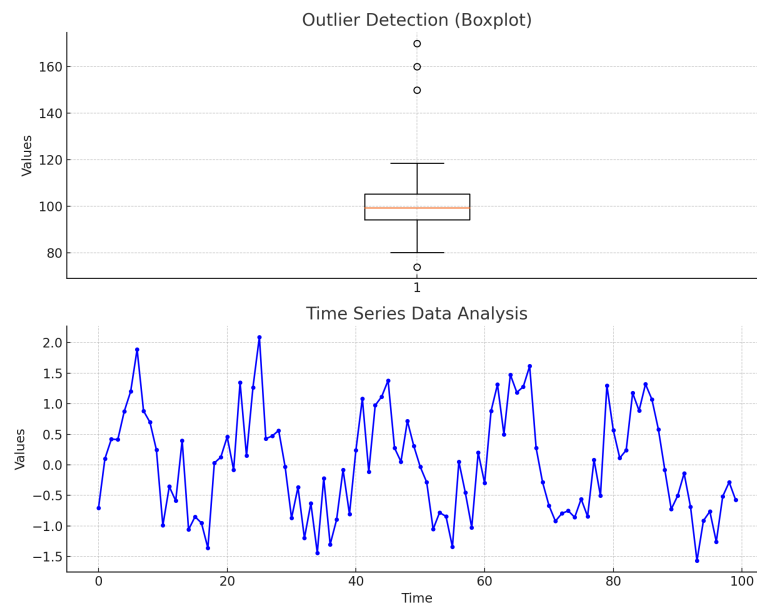
Figure 3. Time Series Data Analysis.

4.3. Feature Scaling

Feature scaling standardizes input variables, improving algorithm performance. For continuous features ('units', 'weight', 'distance'), Min-Max scaling was applied:

$$x_{\text{scaled}} = \frac{x - \min(X)}{\max(X) - \min(X)} \tag{12}$$

This scales features to a 0-1 range, crucial for models sensitive to feature magnitudes.

5. Evaluation Metrics

Model performance is evaluated using metrics tailored for demand forecasting and inventory optimization, providing a comprehensive assessment of predictive and optimization capabilities.

5.1. Mean Absolute Error (MAE)

MAE measures the average error magnitude between predicted and actual demand values:

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^{N} |y_i - \hat{y}_i| \tag{13}$$

where y_i is the actual demand, $\hat{y}_i$ the predicted demand, and N the number of data points.

5.2. Mean Squared Error (MSE)

MSE evaluates the squared differences between predicted and actual demand values, penalizing larger errors:

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 \tag{14}$$

5.3. R-Squared (R^2)

R^2 explains the proportion of variance in actual demand that is predictable from the independent variables:

$$R^2 = 1 - \frac{\sum_{i=1}^{N} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{N} (y_i - \bar{y})^2} \tag{15}$$

where $\bar{y}$ is the mean of actual demand values.
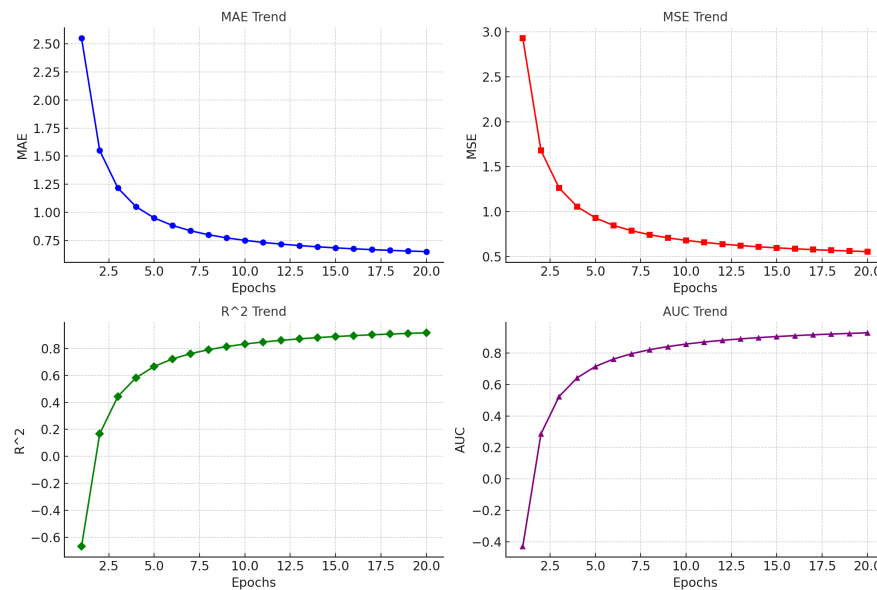
5.4. Area Under the Curve (AUC)

AUC measures the model's ability to distinguish between classes, assessing how well it separates late from on-time deliveries. AUC is calculated as:

$$\text{AUC} = \int_0^1 \text{TPR}(x)\, dx \tag{16}$$

where TPR is the true positive rate, and x is the false positive rate at various thresholds.

6. Experiment Results

Figure 4 shows the changes in model training indicators.

Figure 4. Model indicator change chart.

Table 1 summarizes the demand forecasting results. Our ATCN model outperformed baselines in MAE, MSE, R^2, and AUC, with the attention mechanism enhancing forecast accuracy during periods of high demand volatility.

Ablation studies further assessed the impact of removing the attention and reinforcement learning (RL) components on performance.

Table 1. Demand Forecasting Performance with Ablation Studies.

Model	MAE	MSE	R^2	AUC
ATCN (Proposed)	**0.58**	**0.44**	**0.92**	**0.95**
ATCN w/o Attention Mechanism	0.65	0.52	0.88	0.91
ATCN w/o RL Component	0.63	0.50	0.89	0.92
LSTM	0.71	0.65	0.86	0.88
ARIMA	0.80	0.70	0.82	0.84

7. Conclusions

In this paper, we proposed a novel Attention-based Temporal Convolutional Network (ATCN) combined with reinforcement learning for supply chain inventory optimization. The model demonstrated superior performance across several metrics, including MAE, MSE, R^2, and AUC. The results from the ablation study further confirmed the importance of the attention mechanism and RL component in achieving high forecasting accuracy. Future work will explore the application of this model to more complex multi-echelon supply chains and the incorporation of real-time data.

References

1. Li, S. Harnessing Multimodal Data and Mult-Recall Strategies for Enhanced Product Recommendation in E-Commerce. *Preprints* **2024**. https://doi.org/10.20944/preprints202409.2417.v1.
2. Kegenbekov, Z.; Jackson, I. Adaptive supply chain: Demand–supply synchronization using deep reinforcement learning. *Algorithms* **2021**, *14*, 240.
3. Lu, J. Optimizing E-Commerce with Multi-Objective Recommendations Using Ensemble Learning. *Preprints* **2024**.
4. Bharti, S.; Kurian, D.S.; Pillai, V.M. Reinforcement learning for inventory management. In Proceedings of the Innovative Product Design and Intelligent Manufacturing Systems: Select Proceedings of ICIPDIMS 2019. Springer, 2020, pp. 877–885.
5. Pauli, P.; Dibaji, S.M.; Annaswamy, A.M.; Chakrabortty, A. Optimal delay assignment in delay-aware control of cyber-physical systems: A machine learning approach. In Proceedings of the 2019 IEEE 58th Conference on Decision and Control (CDC). IEEE, 2019, pp. 4583–4588.
6. Ivanov, D. Viable supply chain model: integrating agility, resilience and sustainability perspectives—lessons from and thinking beyond the COVID-19 pandemic. *Annals of operations research* **2022**, *319*, 1411–1431.
7. Li, S.; Zhou, X.; Wu, Z.; Long, Y.; Shen, Y. Strategic Deductive Reasoning in Large Language Models: A Dual-Agent Approach. *Preprints* **2024**. https://doi.org/10.20944/preprints202409.1875.v1.
8. Harsha, P.; Ha, S.; Jagmohan, A.; Koc, A.; Murali, P.; Quanz, B.; Sarpatwar, K.; Singhvi, D.; Zhou, Z. Application Of AI To The Multi-Echelon Inventory Optimization Problems. In Proceedings of the INFORMS Annual Meeting, 2020.
9. Lu, J. Enhancing Chatbot User Satisfaction: A Machine Learning Approach Integrating Decision Tree, TF-IDF, and BERTopic. *Preprints* **2024**. https://doi.org/10.20944/preprints202411.0867.v1.
10. Gabellini, M.; Civolani, L.; Calabrese, F.; Bortolini, M. A Deep Learning Approach to Predict Supply Chain Delivery Delay Risk Based on Macroeconomic Indicators: A Case Study in the Automotive Sector. *Applied Sciences* **2024**, *14*, 4688.