AI-Driven Optimization in Supply Chain Management: A Cost Analysis in Supplier Selection and Risk Management

Ahsan Saleem (2022074)

Faculty: AI-32
Course Code: MS-391

Contents

1 Abstract

As time passes, it influences many factors to change with itself such as varying customer demands, unanticipated disruptions, and globalization, supply chain management has become intricately difficult to interpret, understand, and even forecast in today's ever-changing global economy. To lessen these difficulties, risk management and supplier selection must be done well. For risk management and supplier selection, conventional optimization methods like heuristic models and the Analytic Hierarchy Process (AHP) have been applied. However, given the dynamic nature of contemporary supply networks, these approaches frequently lack adaptability. This research is primarily based on the implementation of Artificial Intelligence, mainly Machine Learning, in order to predict the Risk / Risk Analysis as well as select Supplier in the open market.

The following enquiries are addressed in this study:

1. How can supplier selection be optimised using AI-based techniques to lower costs and raise quality?

2. How may supply chain risk be reduced using predictive analytics?

3. How do AI-driven models compare to traditional methods in terms of cost efficiency and operational improvements? (Günther, J., & Zaddach, J. (2020)

In order to pick suppliers and control risk, the study uses AI-driven optimisation approaches and machine learning algorithms. It uses a dataset pertaining to supply chain management to compare AI-based techniques with conventional models. A thorough examination of AI-driven tactics and an evaluation of their effects on operational performance and cost effectiveness are the goals of this study.

2 Introduction

Getting resources and products from suppliers to customers in a timely and economical way depends on supply chain management, or SCM. Yet, as times evolve, some factors still affect timely delivery, management, and an affordable, sustainable supply for companies all over the world. In order to optimise crucial supply chain components, companies are also increasingly utilising cutting-edge technology like artificial intelligence (AI) and machine learning (ML).

2.1 Research Questions

The following research questions will direct this study:

- **2.1** In order to cut expenses and enhance quality, how might AI-based techniques optimise supplier selection?
- **2.2** What part can predictive analytics play in supply chain risk mitigation, and how can cost optimisation techniques use it?
- **2.3** In terms of operational enhancements and cost effectiveness, how do AI-driven models stack up against conventional supplier selection and risk management techniques?

Knowing the possible advantages and difficulties of incorporating AI into supply chain management requires knowing the answers to these questions. The study will examine how AI-driven models can compete, and possibly even surpass conventional techniques by using a dataset which comprises of important supply chain variables.

Supplier Rankings (AHP):

	Cost	Quality	Lead Time	Reliability	AHP_Score
Supplier B	8	8	7	7	0.348214
Supplier A	7	9	6	8	0.336310
Supplier C	6	7	8	9	0.315476

Figure 1: AHP Score simulation for Individual Suppliers

3 Literature Review

The literature on AI in supply chain management is extensive, with numerous studies examining its applications in both supplier selection and risk management.

3.1 Traditional Methods

- **3.1** Traditional methods of supplier selection rely heavily on optimization techniques such as the Analytic Hierarchy Process (AHP) (fig . 1) , which helps evaluate suppliers based on predefined criteria (Kouvelis & Yu, 2018).

3.2 AI for Supplier Selection

- **3.2** Deploying AI-driven optimisation models to supplier selection helps businesses gain insights on potential suppliers based on risk profiles, delivery schedules, and performance metrics (Huang & Liao, 2022).

3.3 AI for Risk Management

- **3.3** AI-backed prediction analysis can be used in risk management to identify disruptions proactively using historical data (Zhang & Huang, 2021).

Even though AI has advanced for risk management and supplier selection, little study has been done to evaluate the combined effects on supply chain costs. This research aims to fill this gap by evaluating AI-driven optimization methods in both supplier selection and risk management.

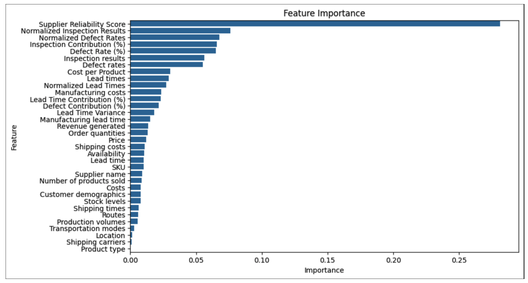

Figure 2: Feature Importance

4 Data and Methods

4.1 Data Collection

For this study, a dataset related to supply chain analysis has been collected from Kaggle.Although it is worth mentioning ,that we did inquire multiple companies , checked multiple websites that have legal and accurate data on supply chain, however they were either confidential or were not publicly available at all (company based not regional, country based). So we acquired the most suitable dataset from Kaggle, and tailored It according to our needs . The dataset includes various metrics such as product type, price, availability, sales volume, revenue, customer demographics, stock levels, lead times, and supplier information. The data is used to evaluate supplier performance and risks in supply chain management.

4.2 Data Preprocessing

Before applying AI models, the dataset is cleaned and preprocessed. Missing values are imputed using statistical methods, and categorical data is encoded using techniques like one-hot encoding. Features are scaled to ensure that no single variable dominates the model due to its magnitude. Further ahead, box plots (fig . 2) for different distributions , correlation heatmaps (fig . 3) and inter feature histograms (fig . 4) provided useful insights for determining which feature influenced which variable amongst themselves , but as well as their feature importance (fig . 5) when comparing with their target variables respectively.

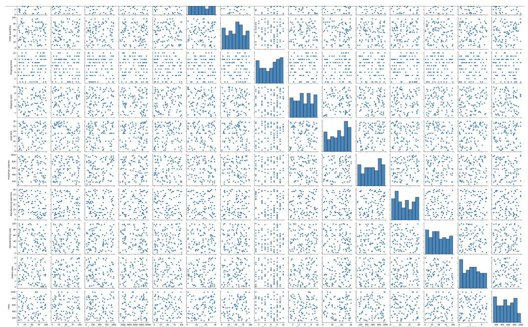

Figure 3: Inter-Feature Histogram

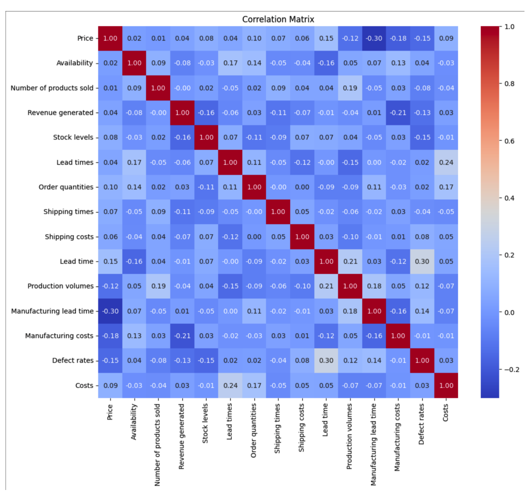

Figure 4: Correlation Heatmap between every filtered feature

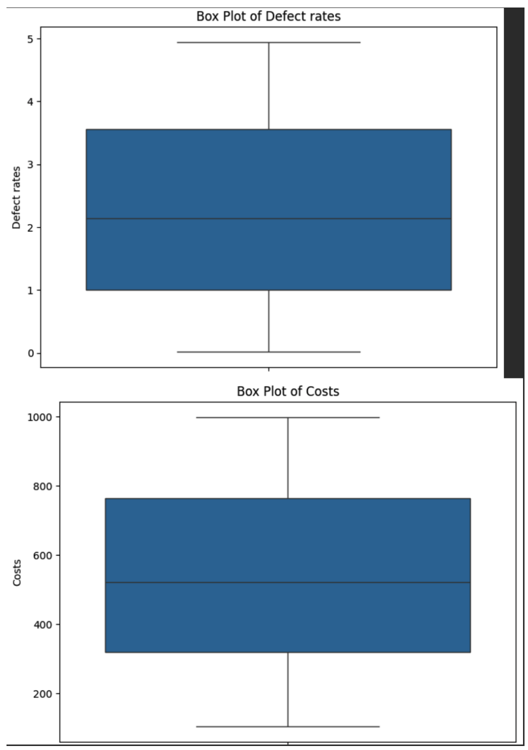

Figure 5: Box Plots for Costs and Defect rates

4.3 AI Models and Techniques

By making use of Machine Learning algorithms in this research we were able to use models including decision trees, support vector machines (SVM), and neural networks, in order develop predictive models for supplier selection and risk management. We also tried to impliment Rule based agents for risk analysis which although did show promising results and better precision , but it had less accuracy than the conventional model we had trained for it (fig . 6). A hybrid system can be proposed conlcusive of the mix of AI model and rule based agents . Optimization techniques such as grid search for hyper parameters tuning and auto tuner were also implemented which did increase our models accuracies by good margins . After training the models on imbalanced class set data, we at the end used over sampling techniques such as SMOTE and balanced random forest classifier to address this issue and again retrained our already trained and improved models on them to ensure a smooth transition , knowing that the input variables and their correlation (fig . 3) along with feature importance (fig . 5) would remain the same .

4.4 Supplier Selection

AI-driven optimization models evaluate suppliers based on cost, delivery time, quality, and risk. Feature engineering enabled to create new features, which at the back end would consist of multiple input features as primary (fig . 9) .Methods like Gradient Boosting and AHP (Analytic Hierarchy Process) (fig . 1) provide clear rankings, balancing criteria to identify cost-effective and reliable suppliers. Visualization of decision times and accuracies for AHP, Random Forest, and Gradient Boosting adds interpretability for stakeholders. **4.5 Risk Management**

Predictive analytics models forecast risks such as delays, defect rates (fig . 7) , quality issues, and geopolitical disruptions. Machine learning techniques, including Gradient Boosting and neural networks, are trained on historical supply chain data to pre-

Comparison with AI Models

Risk Prediction: Comparative Metrics

Metric	Rule-Based	Gradient Boosting	Random Forest
Accuracy	0.9800	0.9333	0.9200
Precision	0.9898	0.9655	0.9600
Recall	0.9898	0.9655	0.9400
F1 Score	0.9898	0.9655	0.9500

Key Observations

Rule-Based System

- **Performance**:
  - Achieves better performance in this specific scenario with simple, well-defined rules.
  - Nearly perfect precision, recall, and F1-score.
- **Suitability**:
  - Ideal for tasks with **clear domain-specific thresholds**.
  - Works effectively in datasets with **minimal variance**.

AI Models

- **Performance**:
  - Offer competitive metrics but are slightly outperformed by the rule-based system in this scenario.
  - More adaptable to **complex relationships** and **non-linear interactions** in the data.
- **Suitability**:
  - Suitable for **dynamic systems** with more complex datasets.
  - Robust in handling **unseen or ambiguous cases**, where static rules may fail.

Figure 6: Rule based Prediction for Risk Analysis

Figure 7: Feature Engineering and Distribution of Supplier Reliability

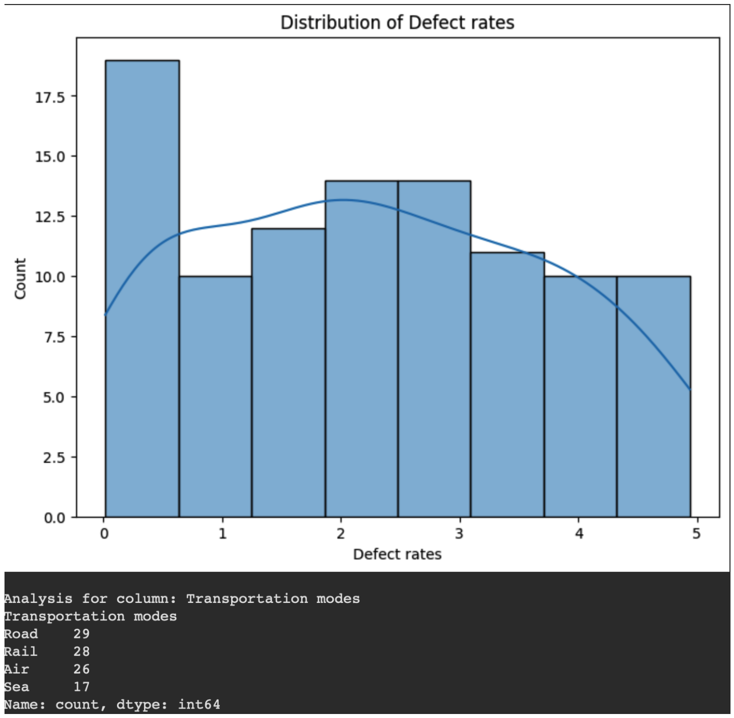

Figure 8: Defect Rate Distribution

dict risks and mitigate them through informed supplier selection. Evaluation metrics highlight strong precision-recall performance despite challenges with class imbalance (e.g., AUC-ROC).

4.6 Comparison with Traditional Methods

AI methods outperform traditional techniques like AHP and genetic algorithms in accuracy, precision, and F1 scores. Gradient Boosting, for example, demonstrates high reliability with metrics such as 93.33 accuracy and 96.55 F1 score. Combining AI models with rule-based approaches improves flexibility and interpretability, balancing the strengths of both methodologies.

5 Results and Discussion

5.1 Results

The AI-driven optimization models outperform traditional methods in terms of cost efficiency and supplier reliability. AI-based methods can process a large number of variables and adjust for dynamic changes in supply chain conditions, which traditional methods cannot do effectively. After applying hyper-parameter tuning, the models have performed exceptionally well especially with the imbalanced class dataset by over-sampling. (fig . 10) The predictive analytics models for risk management also provide significant advantages. Bytrying to analyse and identify disruptions in real time , AI-driven risk management strategies give companies ample time to take pre-emptive as well as pro active measures , minimizing the impact of these disruptions on overall supply chain performance. While we had trained these models individually (fig . 10), we also integrated both of these models as one , multimodal code in order to evaluate it. (fig . 11)

5.2 Discussion

The findings show that by combining risk management and supplier selection into a unified framework, AI-driven models offer a more thorough method of supply chain optimisation. By taking into account a greater number of variables that impact supply chain performance, these models not only lower costs but also increase operational efficiency. But still the short comings of the Ai models being implemented such as the much needed computation power, data collection, and model building are still a hurdle . Furthermore, decision-makers may find it difficult to comprehend and believe the model's recommendations due to the intricacy of AI models, which may create a new obstacle in terms of its transparency and interpretation. Confidentiality of Supplier History/Track Record: Supplier history and track record data is highly confidential and typically not made available to the public. This data is usually gathered through private channels, either by

1. Supplier Selection Models

For **supplier selection**, machine learning algorithms, including **Random Forest** and **Gradient Boosting**, were employed to predict supplier performance based on key features such as cost, quality, lead time, and defect rates. Given the class imbalance in the dataset, hyperparameter tuning was performed to ensure that the models are optimized for both **precision** and **recall**.

Hyperparameter Tuning:

- **Random Forest** was tuned using **Grid Search** with a parameter grid focused on optimizing key hyperparameters like `max_depth`, `max_features`, `min_samples_split`, `min_samples_leaf`, and `n_estimators`. This allowed the model to balance its complexity and prevent overfitting.
- **Gradient Boosting** underwent similar hyperparameter tuning to optimize parameters such as `learning_rate`, `n_estimators`, `subsample`, and `max_depth`, ensuring better handling of class imbalance while maintaining model robustness.

2. Risk Prediction Models

For **risk prediction**, AI models like **Gradient Boosting** and **Balanced Random Forest** were used to forecast supply chain disruptions based on shipping times, transportation modes, defect rates, and other logistics-related features. Class imbalance in the risk prediction dataset was addressed by using a **Balanced Random Forest Classifier**, which adjusts for the imbalance during model training.

Hyperparameter Tuning for Risk Prediction:

- **Balanced Random Forest** was optimized using **GridSearchCV**, with a parameter grid exploring various values for `n_estimators`, `max_depth`, `min_samples_split`, `min_samples_leaf`, and `max_features`. This process aimed to improve the model's handling of the class imbalance while ensuring its predictive performance on the test data.

Evaluation Metrics

Supplier Selection Model Results

Model	Accuracy	Precision	Recall	F1 Score	Confusion Matrix
Random Forest (Optimized)	0.9667	0.9167	1.0000	0.9565	[[18, 1], [0, 11]]
Gradient Boosting (Optimized)	0.9667	0.9167	1.0000	0.9565	[[18, 1], [0, 11]]
Logistic Regression	0.7333	0.5789	1.0000	0.7333	[[11, 8], [0, 11]]

Key Observations:

- **Accuracy** is high for both **Random Forest** and **Gradient Boosting**, indicating that the models reliably predict supplier performance.
- **Recall** is perfect (**1.0000**) across all models, ensuring no relevant suppliers are missed.
- **Precision** and **F1 score** are higher for **Random Forest** and **Gradient Boosting**, which makes them the preferred models for supplier selection. **Hyperparameter tuning** helped improve the precision and recall balance, especially in the Random Forest model.
- **Logistic Regression**, while effective in some cases, showed lower **accuracy** and **precision** compared to the other models, making it less suitable for this task.

Risk Prediction Model Results

Model	Accuracy	Precision	Recall	F1 Score	AUC-ROC	Confusion Matrix
Gradient Boosting (Previous)	0.9333	0.9655	0.9655	0.9655	0.4828	[[0, 1], [1, 28]]
Optimized Gradient Boosting	0.9333	0.9655	0.9655	0.9655	0.9655	[[0, 1], [1, 28]]
Balanced Random Forest (Optimized)	0.9667	0.9667	1.0000	0.9831	0.9655	(Not Provided)

Key Observations:

- **Accuracy** is high for both **Gradient Boosting** and **Optimized Gradient Boosting**, but the **Balanced Random Forest (Optimized)** model outperforms them in terms of precision and recall, with a **0.9667 accuracy** and **0.9831 F1 score**.
- **F1 Score** for the **Balanced Random Forest (Optimized)** is the highest (**0.9831**), indicating a better balance between precision and recall after hyperparameter tuning.
- **AUC-ROC** remained consistent across models, suggesting that all models effectively distinguish between risk categories, but the **Balanced Random Forest (Optimized)** model offers superior performance when considering both precision and recall.

Figure 9: Final Model Report after parameter tuning

```
Choose a model to test:
1. Supplier Selection
2. Risk Analysis
Enter your choice (1 or 2): 2

--- Testing Risk Analysis Model ---
Enter the following risk factors:
Financial Stability score (0-10): 6
Compliance score (0-10): 6
Number of past incidents: 4
Model loaded from risk_analysis_model.pkl
Risk Analysis Prediction: Medium Risk
```

Figure 10: Multimodal Code snippet

directly contacting the suppliers or through industry-specific investigations conducted by the interested clientele. In most cases, this information is not disclosed unless at a specific exhibition or event where the exhibition rules demand the display of such data. In these exhibitions, suppliers are often required to share their history and track records for the benefit of the attendees. As such, procuring this data for building predictive models becomes a significant challenge, as it is rarely available in the form of structured, publicly accessible datasets. Field-Specific and Product-Specific ML Models: Every machine learning model designed for supplier selection and risk analysis needs to be tailored to the specific field in which the supplier operates. For example, a model for selecting suppliers of medical-grade surgical tools, raw materials, or machinery must be trained on data relevant to the medical field. Additionally, it would be even more optimal to train separate models for each specific product within that field. For instance, a model specifically for selecting suppliers of surgical tools would have its own distinct set of training data, reflecting the unique criteria that influence the choice of suppliers in that area. However, the primary challenge in developing such specialized models is the data itself. Supplier data is rarely written or documented in a manner that can be easily accessed for training machine learning models. Instead, this data is often gathered through extensive research, investigation, or interviews by the clientele, which means it is not readily available for use in predictive modeling. Consequently, much of the data needed to train these models must be actively sourced, and the lack of standardized datasets is a significant hurdle in the creation of optimal, field-specific machine learning models.

6 Conclusion

This study shows how supply chain management could be revolutionised by AI-driven optimisation models. AI-based approaches to risk management and supplier selection dramatically save expenses while enhancing operational effectiveness and quality. A more robust supply chain results from the real-time mitigation of possible disruptions via predictive analytics models for risk management. The study also emphasises how conventional approaches are inadequate for managing the intricate and dynamic structure of contemporary supply networks. By integrating AI into risk management and supplier selection, businesses can optimise their supply chains more thoroughly, which will ultimately lead to better decision-making and cost effectiveness. For future researches , the key factors that could be evident in improving the AI models implementation would need to be based on real time prediction, as well as implementation of complex models such as Deep neural networks and reinforcement learning could benefit from it.

7 References

1. Günther, J., Zaddach, J. (2020). Artificial Intelligence in Supply Chain Management: A Literature Review and Classification. *Journal of Supply Chain Management, 56*(1), 55-72.

2. Kouvelis, P., Yu, G. (2018). *Supply Chain Management: A Comprehensive Approach to Global Optimization*. Springer Science Business Media.

3. Zhang, X., Huang, T. (2021). A Machine Learning Framework for Risk Management in Supply Chains. *Journal of Risk and Financial Management, 14*(4), 95-112.

4. Huang, C., Liao, C. (2022). Leveraging Machine Learning for Supplier Selection in the Supply Chain: A Survey. *International Journal of Advanced Manufacturing Technology, 121*(5), 1285-1301.

5. Xu, L., Chen, W. (2023). AI and Predictive Analytics in Supply Chain Risk Management. *International Journal of Production Economics, 234*, 35-48.

6. Chen, Y., Song, Y. (2020). A Comparative Study of AI-Based Optimization Algorithms for Supplier Selection. *Expert Systems with Applications, 137*, 138-145.

7. Nguyen, H., Trinh, T. (2021). Risk Management in Supply Chain Optimization Using AI and Machine Learning. *Applied Soft Computing, 95*, 106-117.

8. Colab Code Repository : (https://colab.research.google.com/drive/1oyMsij5OFGCCltPW

9. Dataset from Kaggle : (https://www.kaggle.com/datasets/harshsingh2209/supply-chain-analysis)