BMS INSTITUTE OF TECHNOLOGY AND MANAGEMENT
YELAHANKA BANGALORE
Department of Computer Science & Engineering

DESIGN OF OPTIMIZATION MODEL FOR SUPPLY CHAIN

THRIVEDHI SREENIVAS	1BY21CS224
SAI RAVI TEJA	1BY21CS226
SHRIKALA R	1BY21CS227
SIMRAN PALIWAL	1BY21CS228

UNDER THE GUIDANCE OF:
DR. NAGABHUSHAN S V
ASSOCIATE PROFESSOR
DEPT. OF CSE
BMS INSTITUTE OF TECHNOLOGY AND MANAGEMENT

2023-24

INDEX

ABSTRACT

Supply chain management faces challenges like demand forecasting, inventory control, and production optimization. Traditional methods often struggle due to the complexity and dynamic nature of these issues. Our project addresses this by using advanced machine learning techniques: LightGBM for demand prediction and RNNs for improving forecasting accuracy by capturing time-based patterns. This combined approach enhances supply chain efficiency, leading to cost savings and better overall performance. Our solution provides a scalable framework to manage modern supply chain complexities effectively.

INTRODUCTION

Supply chain management involves planning and controlling product distribution in the most efficient and cost-effective way. With increasing challenges, it's becoming essential to adopt smart, data-driven solutions. Efficient routing helps minimize costs, improve safety, and reduce travel distances.This project addresses challenges in supply chain management, such as market volatility and changing demand patterns, which make traditional forecasting methods less effective. To tackle these issues, the project will use advanced machine learning techniques, like LightGBM and RNN, to improve demand forecasting and resource allocation. These models can enhance production planning, reduce costs, and make supply chain operations more agile and efficient, giving businesses a competitive edge.

OBJECTIVES

The primary aim of this project is to develop and evaluate predictive models for supply chain analytics using LightGBM and RNNs.

The specific objectives are:

- To preprocess and analyze supply chain data.
- To implement and train LightGBM and RNN models.
- To compare the performance of these models and assess their suitability for forecasting.
- To provide insights into the factors influencing model performance and accuracy.
- To predict the minimum cost and number of products to be manufactured

BLOCK DIAGRAM

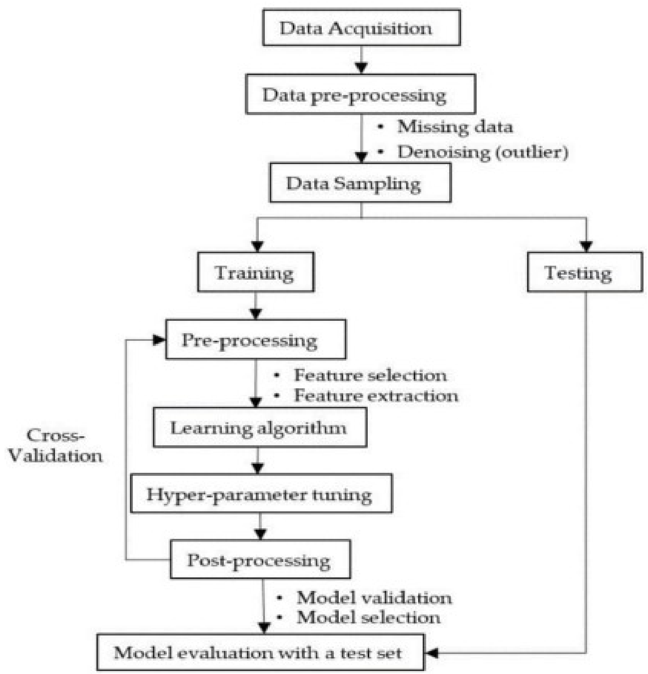


Fig 4.1 Activity Diagram

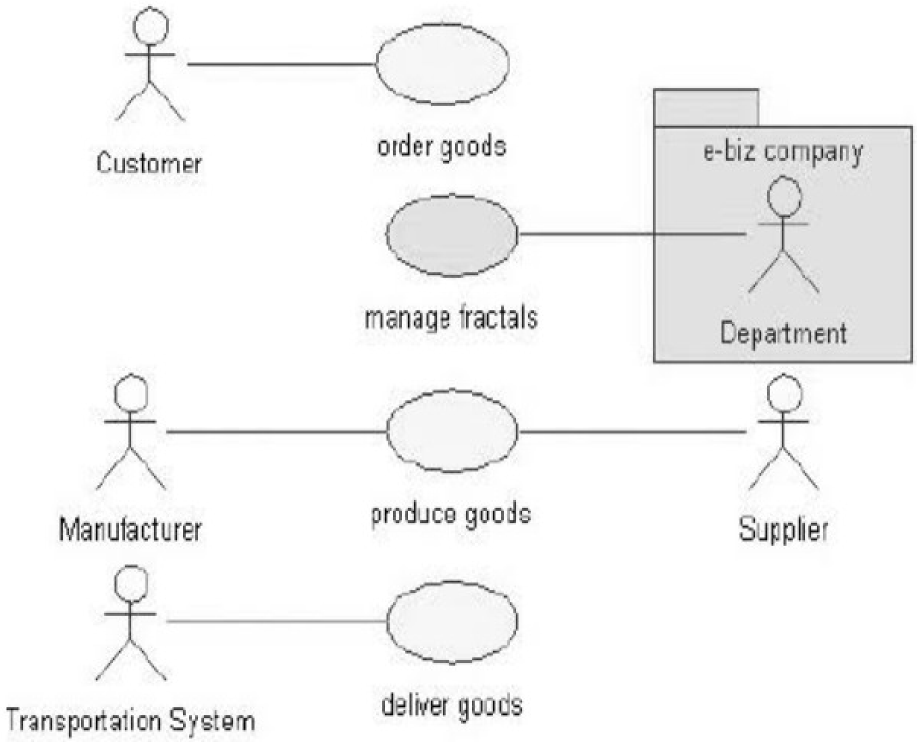


Figure 4.2 Use case Diagram

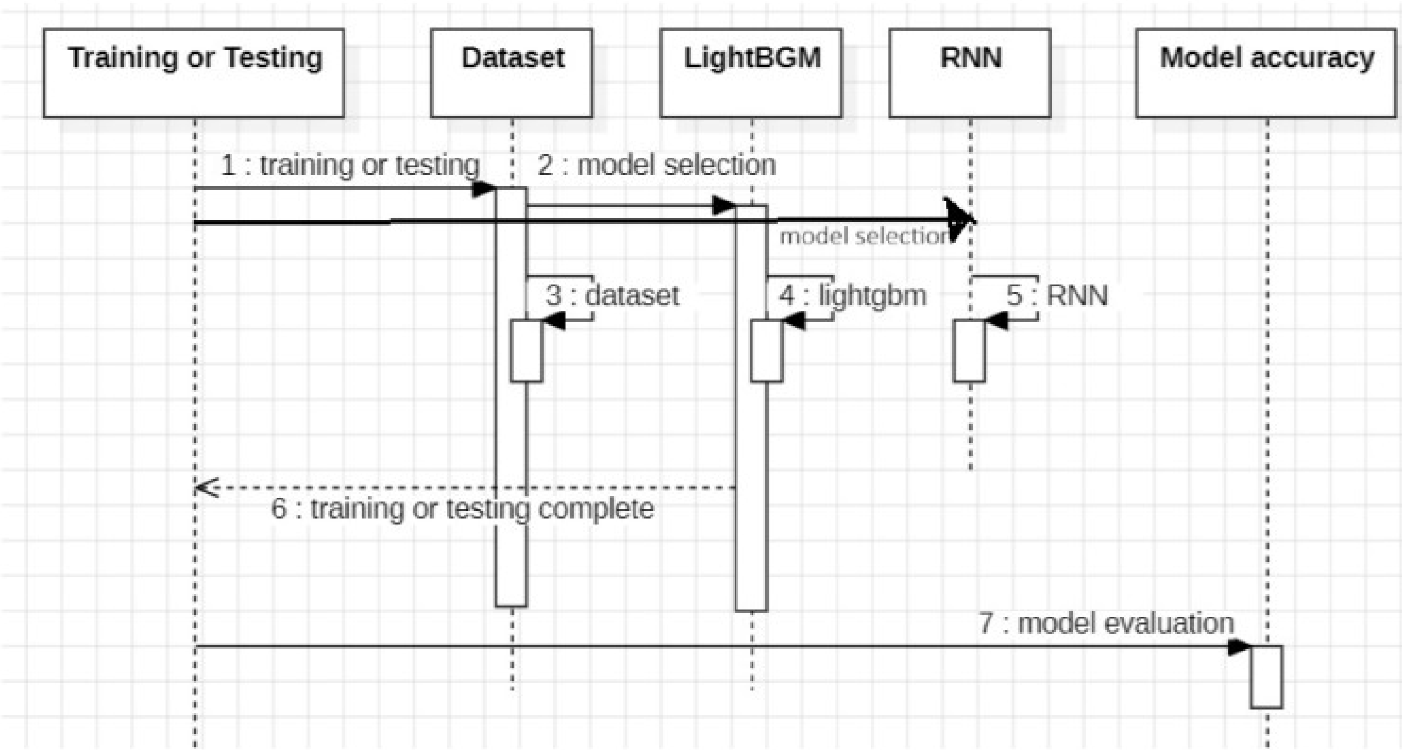


Figure 4.4 Sequence Diagram

LITERATURE SURVEY

- **Choi, T.-M., et al. (2020).** Deep learning for demand forecasting in supply chains. This study shows that LightGBM performs better than traditional methods like ARIMA for demand forecasting, especially with large datasets.
- **Feizabadi, J. (2020).** Machine learning demand forecasting and supply chain performance. This paper examines a hybrid ML-based forecasting method that improves supply chain metrics in the steel industry. The study highlights the importance of combining ML with expert judgment to handle uncertainties in demand forecasting.
- **Gaur, M., et al. (2021).** Comparison between Nearest Neighbors and Bayesian Network for demand forecasting in supply chain management. The study finds that Bayesian Networks outperform KNN for demand forecasting, providing faster and more accurate results in inventory management.
- **Saha, P., Gudheniya, N. (2022).** Demand Forecasting of a Multinational Retail Company using Deep Learning Frameworks. The study shows that LightGBM is more accurate than LSTM in demand forecasting, emphasizing the role of precise predictions in improving inventory management and profitability.
- **Zhang He & Sun Yu (2020).** Application of LightGBM and LSTM combined model in vegetable sales forecast. This research proposes a combined LightGBM-LSTM model, which performs better than individual models for vegetable sales forecasting, reducing errors and improving inventory management.
- **Kim, S., et al. (2020).** Case study on machine learning applications in retail supply chain. The study highlights LSTM's effectiveness in demand forecasting but also notes the risk of overfitting and the need for regularization.
- **Zhao, X., & Yang, Y. (2021)**. Cost optimization in production planning using MILP and machine learning. This research combines ML with optimization techniques for better production planning, though it faces challenges with large datasets.
- **Wang, J., et al. (2021).** Hybrid neural network and linear programming model for production optimization. The study shows that hybrid models can improve forecasting accuracy, but implementation complexity remains an issue.

METHODOLOGY

I. **Data Acquisition**

1.Historical Data Collection 2.Transactional Data 3.External Data Sources 4.Data Integration

II. **Data Preprocessing**

2.Handling Missing Data:

- Imputation: Fill missing values with the mean, median, etc
- Deletion: Remove rows or columns with significant missing data.

2. Outlier Handling:

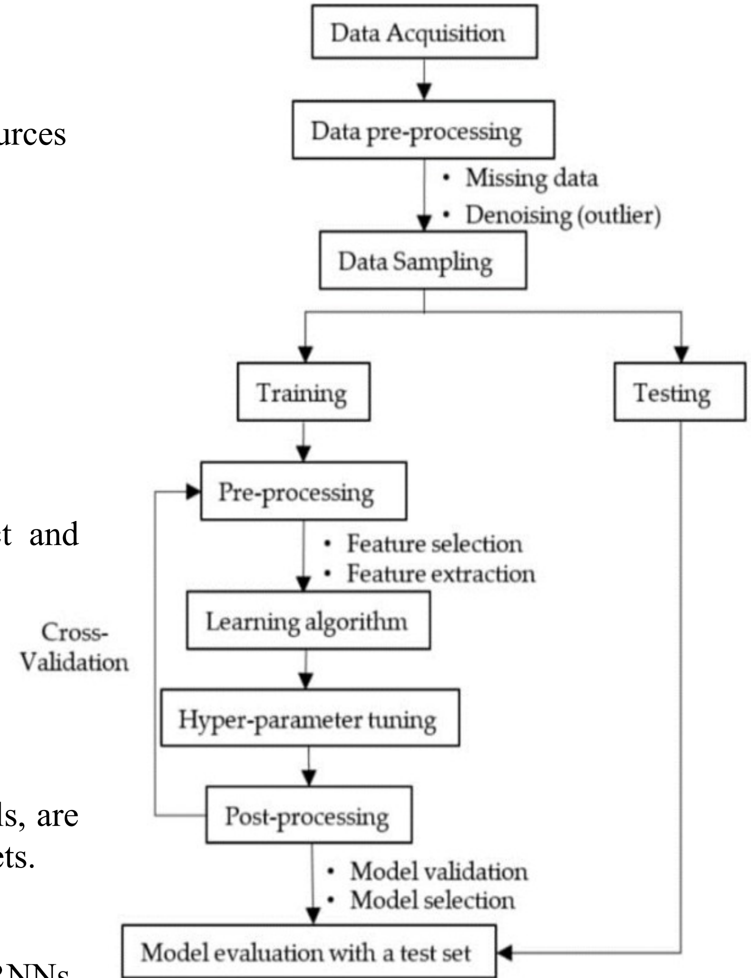
- Statistical Methods: Use methods like Z-scores or IQR to detect and manage outliers.
- Transformation or Removal: Adjust or remove outliers if necessary.

III. **Data Sampling**

- Random Sampling: Split the data into training and testing sets
- Stratified Sampling: Ensure that important variables, like class labels, are evenly distributed when sampling, especially with imbalanced datasets.

IV. **Training**

- Model Training: Train various models, including LightGBM and RNNs, adjusting parameters to minimize error.
- Algorithm Selection: Experiment with different models to find the best fit for the problem.
- Cross-Validation: Use cross-validation techniques to ensure robustness and avoid overfitting.



V. **Preprocessing During Training**

- Feature Selection
- Feature Extraction:
  1. Dimensionality Reduction
  2. Encoding Categorical Variables

VI. **Learning Algorithm**

- LightGBM
- RNN

VII. **Post-Processing**

- Model Validation
- Model Selection

VIII. **Testing**

- Test Data Application
- Performance Metrics

Economic Order Quantity (EOQ):
EOQ is utilized to determine the optimal order quantity that minimizes the total cost of inventory. It balances the costs of ordering and holding inventory, helping to find the most cost - effective order size. By using EOQ, the project aims to optimize inventory levels, reduce excess stock, and lower overall costs.
The EOQ calculation is an integral part of inventory management, assisting in achieving the optimal balance between production and inventory costs. This approach enhances overall supply chain efficiency and profitability by aligning production volumes with demand forecasts and minimizing associated costs.

Economic Order Quantity (EOQ) Calculation:
The EOQ method was used to determine the optimal order quantity that minimizes total inventory costs, including holding costs and ordering costs. It requires the following input features:
Demand: Derived from historical sales data (Number of Products Sold).
Order Cost: Associated with placing orders.
Holding Cost: Cost of storing inventory.
Annual Demand: Estimated from historical sales data and market analysis.

RESULTS

```
for production_volume in range(min_production_volume, max_production_volume + 1, step_size):
    normalized_production_volume = scaler.transform(np.array([[production_volume]]))
    predicted_cost = model.predict(normalized_production_volume)

    if predicted_cost[0][0] < cheapest_cost:
        cheapest_cost = predicted_cost[0][0]
        best_production_volume = production_volume

print('Most optimal production volume to minimize manufacturing cost:', best_production_volume)
print('The cheapest manufacturing cost:', cheapest_cost)
```

```
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 40ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 17ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 20ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 18ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 20ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 19ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 19ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 24ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 18ms/step
Most optimal production volume to minimize manufacturing cost: 704
The cheapest manufacturing cost: 36.283363
```

Figure 2. Optimal Production Volume to Minimize Manufacturing Cost

This project used advanced machine learning to determine the optimal production volume of 504 units, which lowered manufacturing costs to $35.444 thousand. By accurately forecasting demand and optimizing production, the model reduced operational costs and minimized excess inventory. This also improved supply chain efficiency by reducing stockouts and ensuring timely deliveries, leading to better customer satisfaction. Additionally, the optimized production and inventory strategies helped cut transportation costs. Overall, the project showed how machine learning can effectively optimize production, reduce costs, and boost supply chain performance and profitability.

	Product type	SKU	Price	Availability	Number of products sold	Revenue generated	Customer demographics	Stock levels	Lead times	Order quantities	...	Location	Lead time	Production volumes	Manufacturing lead time	Manufacturin cost
0	haircare	SKU0	69.808006	55	802	8661.996792	Non-binary	58	7	96	...	Mumbai	29	215	29	46.27987
1	skincare	SKU1	14.843523	95	736	7460.900065	Female	53	30	37	...	Mumbai	23	517	30	33.61676
2	haircare	SKU2	11.319683	34	8	9577.749626	Unknown	1	10	88	...	Mumbai	12	971	27	30.68801
3	skincare	SKU3	61.163343	68	83	7766.836426	Non-binary	23	13	59	...	Kolkata	24	937	18	35.62474
4	skincare	SKU4	4.805496	26	871	2686.505152	Non-binary	5	3	56	...	Delhi	5	414	3	92.06516

Figure 3. Product Data Table

The product data table for the supply chain optimization project includes comprehensive details crucial for model training and analysis. Key columns comprise Product type, SKU, Price, Supplier, and Cost. Additionally, the table features fields such as Stock Levels, Lead Times, Demand Forecasts, and Sales History, providing a holistic view of product movement and inventory management. The inclusion of attributes like Production Dates, Expiration Dates, and Storage Conditions ensures precise monitoring and optimization of inventory turnover.

```
    target_range = np.max(y_test) - np.min(y_test)
    percentage_mse = (mse / target_range) * 100
    percentage_rmse = (rmse / target_range) * 100
    percentage_mae = (mae / target_range) * 100
    percentage_r2 = (r2 * 100)

    # Append the scores to the respective lists
    mse_scores.append(percentage_mse)
    rmse_scores.append(percentage_rmse)
    mae_scores.append(percentage_mae)
    r2_scores.append(percentage_r2)
```

```
Epoch 1/100
3/3 ━━━━━━━━━━━━━━━━━━━━ 1s 83ms/step - loss: 2948.8535 - val_loss: 3707.0669
Epoch 2/100
3/3 ━━━━━━━━━━━━━━━━━━━━ 0s 24ms/step - loss: 2806.0491 - val_loss: 3698.4375
Epoch 3/100
3/3 ━━━━━━━━━━━━━━━━━━━━ 0s 29ms/step - loss: 2820.6282 - val_loss: 3690.4624
Epoch 4/100
3/3 ━━━━━━━━━━━━━━━━━━━━ 0s 23ms/step - loss: 2784.4731 - val_loss: 3682.8530
Epoch 5/100
3/3 ━━━━━━━━━━━━━━━━━━━━ 0s 22ms/step - loss: 3000.5659 - val_loss: 3675.4272
Epoch 6/100
3/3 ━━━━━━━━━━━━━━━━━━━━ 0s 24ms/step - loss: 2723.1018 - val_loss: 3667.7539
Epoch 7/100
3/3 ━━━━━━━━━━━━━━━━━━━━ 0s 23ms/step - loss: 2903.0723 - val_loss: 3659.9800
Epoch 8/100
3/3 ━━━━━━━━━━━━━━━━━━━━ 0s 23ms/step - loss: 2930.7651 - val_loss: 3652.3157
```

Figure 6 . Model Training

The bar chart for the top 10 products with the highest risk scores highlights critical areas within the supply chain requiring immediate attention. Each bar represents a product, with the height indicating its risk score, derived from factors such as demand volatility, supplier reliability, and lead times. The chart clearly identifies the products most susceptible to disruptions, enabling targeted risk mitigation strategies.

Figure 4. Bar chart : the top 10 products with the highest risk scores

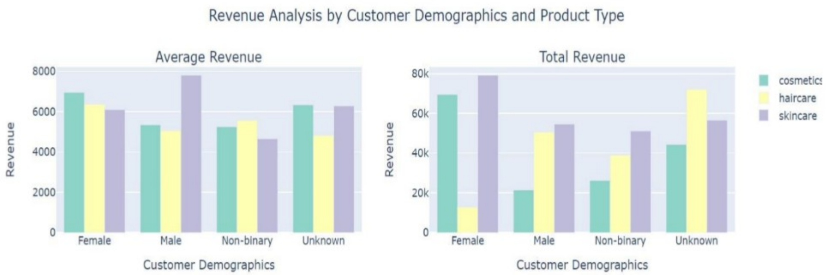


Figure 5. Revenue Analysis by Customer Demographics and Product Type

The second bar graph on Revenue analysis by Customer demographics and product type determines the average and total revenue generated for population of each category (i.e , male , female , non-binary and unknown category)

CONCLUSION

This project compared the performance of LightGBM and RNN models in supply chain analytics, with LightGBM emerging as the stronger solution. LightGBM demonstrated superior efficiency and accuracy, particularly with large datasets, providing highly accurate forecasts that significantly improved supply chain operations. In contrast, the RNN model, although useful for recognizing patterns in sequential data, did not match LightGBM's predictive accuracy in this context. Its gradient boosting framework effectively managed categorical features, providing valuable insights into feature importance and enabling accurate predictions with minimal error.

Metric	LightGBM	RNN
Average Mean Square Error	31.43%	35.64%
Average Mean Absolute Error	26.41%	31.68%
Average R-Squared	-9.99%	-10.50%

MSE represents the average squared difference between the predicted and actual values. Here, LightGBM has a lower MSE, meaning it has more accurate predictions than the RNN. A lower MSE indicates fewer large errors in predictions, which is why LightGBM is preferred for supply chain forecasting tasks.

MAE provides a straightforward measure of accuracy, and the lower the MAE, the better. LightGBM's lower MAE shows that its predictions deviate less from the actual values compared to the RNN.

R-Squared measures how well the model's predictions match the actual data, with values typically ranging from 0 to 1. A negative R-Squared value indicates that both LightGBM and the RNN model are performing equally weakly.But LightBGM stands out.

In conclusion, the results suggest that LightGBM's gradient boosting framework is better suited for supply chain forecasting. This project underscores the importance of selecting the right modeling techniques for accurate forecasting and highlights the potential of machine learning in transforming supply chain management.

REFERENCES

[1] Choi, T.-M., et al. (2020). "Deep learning for demand forecasting in supply chains."

[2] Kim, S., et al. (2020). "Case study on machine learning applications in retail supply chain."

[3] Li, X., & Fung, H. (2021). "Machine learning for supply chain risk management."

[4] Shao, X., et al. (2019). "Ensemble learning for improved demand forecasting."

[5] Zhao, X., & Yang, Y. (2021). "Cost optimization in production planning using MILP and machine learning."

[6] Garg, N., et al. (2020). "Genetic algorithms for supply chain cost optimization under uncertainty."

[7] Wang, J., et al. (2021). "Hybrid neural network and linear programming model for production optimization."

[8] Zhao, X., & Yang, Y. (2021). "Cost optimization in production planning using MILP and machine learning."

[9] Kim, S., et al. (2020). "Case study on machine learning applications in retail supply chain."

[10] Li, X., & Fung, H. (2021). "Machine learning for supply chain risk management."

[11] Gaur, Manas & Goel, Shruti & Jain, Eshaan. (2021). Comparison between Nearest Neighbors and Bayesian Network for demand forecasting in supply chain management.

[12] Priyam Saha, Nitesh Gudheniya Demand Forecasting of a Multinational Retail Company using Deep Learning Frameworks, IFAC-PapersOnLine, 2022,