International Journal of Basic and Applied Sciences, 14 (5) (2025) 753-764

International Journal of Basic and Applied Sciences

Website: www.sciencepubco.com/index.php/IJBAS
https://doi.org/10.14419/jnrqjz39
Research paper

Machine Learning Application for Sales Forecasting and Inventory Optimization in Wholesale Trade

Pujan Shailesh Kakkad

Senior Systems Development Engineer, Amazon Robotics, Business Applications and Solutions Engineering Team, Austin, Texas
**Corresponding author E-mail: pujankakkad@gmail.com*

Received: July 28, 2025, Accepted: September 4, 2025, Published: September 20, 2025

Abstract

This work explores the use of machine learning in optimizing inventory management processes, using the case study of Vadodara, Gujarat, the largest wholesaler of plumbing and heating products in the Czech Republic and Slovakia. The work combines financial analysis with predictive analytics to identify opportunities for cost reduction and inventory optimization. A linear regression model was applied to historical sales data of corded and cordless power tools to forecast future demand. The findings suggest a clear trend of increasing sales of cordless tools and stagnating demand for corded tools. Based on this, inventory turnover was analyzed, and an ABC analysis was conducted to identify high-priority stock. Recommendations were made to rebalance inventory investments, enhance stock efficiency, and implement strategic purchasing decisions. The application of machine learning in this context demonstrated measurable benefits in predicting sales patterns and improving inventory planning, ultimately contributing to better cost management.

Keywords: *Machine Learning; Inventory Optimization; Predictive Analytics; Linear Regression; ABC Analysis; Demand Forecasting*

1. Introduction

In today's dynamic and competitive market, efficient inventory management is essential to maximize profits and customer satisfaction. The purpose of this study is to investigate Ptachek India Pvt. Ltd., a well-known distributor of plumbing and aerospace equipment in India, which has been growing steadily since its inception in 1992.

The focus of this study is Vadodara, Gujarat, with a special focus on the company's portfolio of tools – electric (corded) and battery-operated (cordless). As the market shifts to wireless devices based on growing volatility, technological advancements, and global preferences, accurate forecasting of future sales is essential for sustainable inventory management. To answer this question, this study uses linear regression, a key machine learning technique, to predict monthly sales trends for both categories.

In addition to forecasting, the study also examines the company's current inventory configuration and turnover performance. Historical sales data, inventory levels, and tied-up capital showed that the company's inventory management was ineffective, particularly in the supply of slow-flowing capacity with disproportionate investment. Based on the ABC analysis, the company categorized key suppliers based on inventory contribution, which enabled a more targeted strategic approach to sourcing and managing suppliers.

The primary purpose of this study was to provide recommendations for inventory improvement. By maintaining inventory levels in line with planned demand, optimizing turnover, and restructuring capital allocation, the company can significantly reduce operating costs. This approach has enabled more intelligent and data-driven decision-making, and Vadodara, Gujarat, the Role of LLC in a Changing Market Environment.

2. Literature Review

Recent studies have highlighted the transformative potential of emerging technologies in supply chain optimization.

Charles et al. (2023) [10] critically analyzed the integration of blockchain and artificial intelligence (AI) in the supply chain, noting that blockchain ensures transparency and secure data exchange, thereby enabling intelligent prediction and decision-making in the era of AI. Together, these technologies provide powerful tools for real-time visualization and automation, but also bring challenges such as scalability and cost.

Curcio and Longo (2009) [11] emphasized the importance of inventory and internal logistics management and showed through simulation that inefficiencies in these areas can significantly affect the efficiency of the entire supply chain.

Sharma and Singh (2018) [12] complemented this by using the quality function deployment (QFD) approach to show how vendor managed inventory (VMI) can shorten delivery times, reduce errors, and improve service quality by supporting collaborative inventory control.

Based on the role of advanced analytics, Mahraz et al. (2022) [13] conducted a systematic literature review on the application of machine learning (ML) in supply chain management, focusing on its application in demand forecasting, inventory optimization, and logistics planning. They emphasize that to realize the full potential of machine learning, strong data and domain-specific models are required.
Bolgasemi et al. (2020) [14] focus on demand forecasting in the wild and point out that traditional models often fail to consider advertising tasks, so adaptive and advertising-sensitive forecasting techniques are needed.
In the context of Supply Chain 4.0, Salamai et al. (2021) [15] propose a dynamic voice classifier (DVC), a hybrid AI model that can improve risk detection and support proactive risk mitigation strategies.
Finally, Souvindjo et al. (2023) [16] show how predictive analytics can be used to improve inventory performance in fast-moving consumer goods (FMCG) companies and demonstrate how data-driven insights can help reduce inventory costs and increase productivity. Overall, these studies establish the key role of digital technologies, predictive models, and collaborative strategies in building efficient, sustainable, and future-proof supply chains.
Mahraz, Benabbu, and Berrado (2022) [17] conducted a comprehensive systematic review of machine learning applications in supply chain management. Their study found that predictive models, classification algorithms, and optimization approaches are increasingly adopted to improve demand forecasting, inventory management, and distribution planning. Importantly, the review stated that integrating machine learning techniques can enable firms to reduce uncertainty in matching supply and demand and improve the accuracy of decision-making.
Based on this perspective, Charles, Emruznejad, and Herman (2023) [18] critically examined the convergence of blockchain and artificial intelligence within supply chains. Their research shows that artificial intelligence improves forecasting accuracy and decision-making, while blockchain adds transparency, traceability, and security to supply chain transactions. This integration is particularly beneficial for businesses with a diverse supplier base and high-value inventory, as it helps ensure accountability while supporting predictive analytics.
Based on these insights, the reviewed research provides practical implications for the challenges facing Vadodara, Gujarat. For example, Mahraz et al. (2022) [13] emphasize the importance of robust data and domain-specific models in realizing the full potential of machine learning; this actually applies to Vadodara, where disaggregated supplier and customer data often limit forecast accuracy. Similarly, Charles et al. (2023) [10] show that blockchain integration improves transparency and secure data sharing, which can fulfill Vadodara's need for more reliable inventory visibility in its nationwide distribution network. Furthermore, Curcio and Longo's (2009) [11] emphasis on the impact of internal logistics inefficiencies corresponds to Vadodara's complex warehousing and distribution operations, where simulation-based optimization can provide measurable benefits. Together, these connections illustrate how theoretical advances in analytics, blockchain, and collaborative inventory management can be adapted to address Vadodara's unique operational constraints and opportunities.

3. Company Description

Established in 1975, Indian Agencies is a leading distributor of industrial tools in Vadodara, Gujarat. The company supplies a wide range of power tools, cordless tools, welding equipment, and accessories. It is an authorized partner for global brands such as Bosch, Makita, KPT, and Esab, while also operating service centers for after-sales support. Known for reliable delivery and strong customer service, Indian Agencies has built a solid reputation as a one-stop solution for industrial and construction tool needs.
The company offers a full suite of services, including:

- Technical consultation
- Product training and demonstrations
- Custom delivery solutions
- Government subsidy advisory for alternative energy products

3.1 Description of Financial Statements

The consolidated financial performance of the Vadodara-based company in FY2022 highlights steady growth and operational efficiency. Total assets reached ₹6,237 crore, reflecting a healthy increase of about ₹437 crore compared to the previous year. Equity capital rose to ₹3,925 crore, underscoring strong capitalization and limited reliance on borrowings. Net revenue crossed ₹4,000 crore, marking an expansion of nearly ₹500 crore year-on-year, largely driven by robust infrastructure activity and rising demand for PVC and CPVC pipes and fittings in the housing sector. Operating profit improved significantly to ₹740 crore, up by more than ₹120 crore from FY2021, pointing to efficiency gains in procurement and distribution. Based on these results, the return on equity (ROE) stands at around 19%, reaffirming the company's financial strength and its ability to generate strong shareholder value.

3.2 Financial Statement Evaluation

Overall, Vadodara demonstrates strong financial health:

- A high equity ratio confirms low dependency on external borrowing, which reduces financial risk.
- Consistent revenue growth amid market challenges reflects a strong customer base and effective expansion strategy.
- An 18% ROE suggests efficient use of equity in generating profits, assuming the industry benchmark is lower than or close to 12–15%.
- Stable profit margins further reinforce the company's strong market positioning and effective procurement.

For continued financial health, the company must:

- Benchmark its net profit margins against key Indian competitors (e.g., Astral or Supreme Industries).
- Stay updated on government subsidy policies, import duties, and green energy incentives.
- Regularly evaluate inventory turnover and cash flow ratios.

4. Applications for Machine Learning in The Inventory Management Process

After receiving data from Vadodara, a graph was created that consisted of the number of sales volumes depending on the category of goods. It was found that the sales of battery-powered tools are increasing, while the sales volume of corded tools has been stagnating over the years. This interesting trend had already been noticed earlier in the company itself, as shown in the graph in Figure 1. The graph shows the period from January 2020 to April 2023, with individual breaks representing data for each month. This finding suggests the potential to more accurately determine the volume of inventory for this assortment in a way that is as advantageous as possible for the company. A machine learning algorithm and predictive analytics knowledge were used for this. The assumption is that it would be appropriate in the future to reduce the inventory ratio for products whose sales will decline or stagnate and increase the inventory ratio for products whose sales will increase. The goal is to reduce costs by minimizing capital tied up in inventory. Based on the results, the company proposed cost-saving measures for the future period. The business process in which these savings are sought is the inventory management process and the related goods purchasing process. Such output can be used as an auxiliary point for the company's buyer when deciding on the quantities of goods to order.

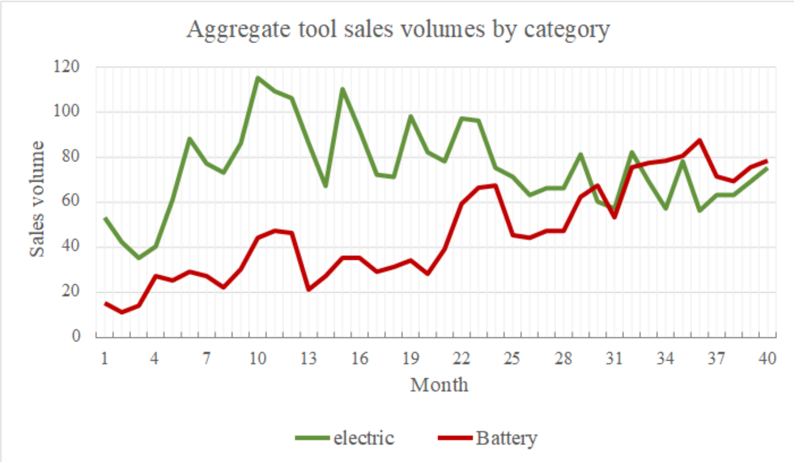

Fig. 1: Graph of aggregate tool sales volumes by category (own processing)

4.1 Received Dataset

In Table 1, we can see a small sample of the data that I received from the company's analytical department. The table contains all 243 different types of corded and 229 types of cordless tools that the company offers to its customers. Electric cordless means that these are tools powered by a battery, and electric means that these are tools powered by a cable from the electrical distribution network. For a more specific idea, the cordless tools that the company offers are, for example, drills, saws, and hammers. Corded tools in the table also include drills, saws, radios, vacuum cleaners, and grinders. In the header of the table on the left side, we see: Product ID, Goods - description, Property (electric/electric battery), Unit price, Stock - quantity, and finally, the table lists the months. Below the list of individual months in the table are numbers indicating the number of products sold. The numbers indicate how many were sold for each calendar month from January 2020 to April 2023. The product ID is just a number, probably used to identify the goods using a barcode. Goods - description contains a brief description and basic characteristics of the goods. The product property says whether it was an electric or an electric battery tool. The data also contains the recommended prices of individual products, and how much the company offers them for approximately. Other information is the number of pieces and how many individual tools the company had in stock as of May 1, 2023. The exact purchase prices could not be published, as this is sensitive internal company information and would damage the future negotiating position with suppliers. Therefore, this valuable information, which could have worked well, could not be obtained.
The appendix to the thesis contains this described dataset, which we used as a basis, called Export_dat.csv. Not all information from this file is important for the purposes of the analysis in the practical part.

Table 1: Received data (own processing)

ID	Product Description	Property	Unit Price	Stock - Quantity	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2080090	Bosch Electric Tools	electric	€20,544.40	0	1											
2080090	Bosch Electric Tools	electric	€10,468.40	0		1										
2080090	Bosch Electric Tools	electric	€10,468.40	0			1									
2080090	Bosch Electric Tools	electric	€19,833.09	0				1								
2080090	Bosch Electric Tools	electric	€21,509.64	0					1							
2080090	Bosch Electric Tools	electric	€16,325.00	0						1						
2080090	Bosch Electric Tools	electric	€15,251.65	0							1					
2080090	Bosch Electric Tools	electric	€10,468.40	0								1				
2080090	Bosch Electric Tools	electric	€37,261.80	0									1			

2080090	Bosch Electric Tools	electric	€4,908.02	0										1		
2080090	Bosch Electric Tools	electric	€8,011.60	0											1	
2080090	Bosch Electric Tools	electric	€17,671.04	0												1
2080090	Bosch Electric Tools	electric	€21,807.38	0												1
2080090	Bosch Electric Tools	electric	€6,609.13	0												

5. Sales Prediction using Machine Learning

Based on the theoretical background obtained from the theoretical part, from the chapter "General procedure for data processing using machine learning", the procedure for predictive analysis in the practical part is built as follows.

5.1 Tool selection

In the case of a larger amount of data that the table contains, it would be complicated to calculate the resulting numbers on a calculator. Another option is to use a function in Excel, but in the case of larger volumes of data, Excel becomes slow compared to machine learning methods. Excel is further limited by the row limit, which is a problem with large data files. Using a machine learning tool is more universal and, above all, more advantageous if we want to automate the model. When changing the input data, it is then faster to obtain a new result if the input data is complicated.

5.2 Data collection

The data was exported from the internal ERP system in the .xlsx file format, which can be opened in Microsoft Excel. I do not have access to information on how the ERP system records and processes data in the company. The data received is genuine and was generated directly by the company's analytical department.

5.3 Data description

In the received data, it is possible to find all the important information that will be needed for the prediction. The data that will be used in the prediction are: Property (electric or electric battery) and sales volumes of tools for individual months. The remaining data will not be needed at all, or will be used only in the next step, when the results will be analyzed. For now, the data will be postponed so that it can be returned to later.

5.4 Data editing

First, the data must be prepared in a suitable table format. The next step is to edit the data into the necessary form suitable for the application of machine learning. No extreme values were found in the data, which could be an error and should be deleted. Only one file in one format was received, so there was no problem converting the data into one common format.

For this step, you first need to open Jupyter Notebook and import and install the libraries that will be needed. There are several tools, but the most important is the scikit-learn library called LinearRegression, which contains the machine learning model itself. Using the scipy.stats library, the 95 percent confidence interval of this prediction will ultimately be determined, and thanks to the matplotlib.pyplot library, this graph will be displayed (Figure 2). The data is uploaded to this environment, and the tools are divided into two categories according to their properties, i.e., "electric" and "electric battery". At the same time, the .sum command adds up the sales volumes for each month, according to the product properties.

```
# Import libraries
import pandas as pd
import numpy as np
import openpyxl
import datetime
import statsmodels.api as sm
import scipy.stats as stats
import matplotlib.pyplot as plt
from sklearn.linear_model import LinearRegression
from sklearn.model_selection import train_test_split
from sklearn.metrics import mean_squared_error, r2_score
```

```
# Import data and pre-processing → dividing the sales volume into two categories by months
df = pd.read_excel('Export_dat.xlsx', sheet_name='Sheet1')
df = df.groupby('Property').sum()
```

Fig. 2: Importing libraries and data (own processing)

As part of data editing, it is necessary to get rid of data that will not be needed. The command gets rid of unnecessary table columns, specifically 'ID', 'Goods - description', 'Unit price', 'Stock - quantity', and only those columns that are needed for this prediction will remain. Next, the table axes are flipped for clarity, and a more appropriate date format is selected. After these modifications, the original Excel table has become much smaller. Figure 3 shows the first 10 rows of the new table, which contains the total sales volumes for each month, divided into categories.

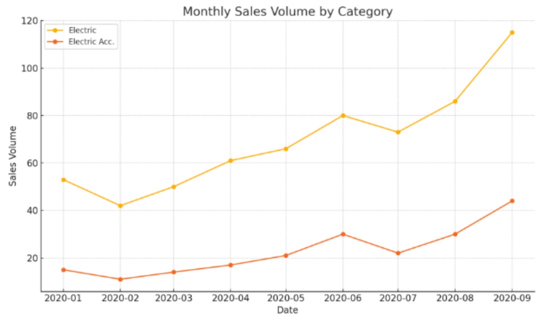

Fig. 3: Data editing (own processing)

Now you need to select one of the categories in which you are interested in predicting sales development. This is shown in Table 2. Since you only worked with two categories, it is easier to repeat the entire process for the second category than to create a special script that would automate this process. However, in real use in practice, this is usually necessary. Therefore, the following procedure will be repeated for the second category, which is an easy and quick process in the Jupyter Notebook environment.

Table 2: Selecting a product category (own processing)

Output Interpreted	
The DataFrame electric_df shows the first 10 rows:	
Date	electric batteries
2020-01-01	15
2020-02-01	11
2020-03-01	14
2020-04-01	27
2020-05-01	30
2020-06-01	29
2020-07-01	27
2020-08-01	30
2020-09-01	31
2020-10-01	44

Next, the date needs to be converted to an index, because the linear regression model can only work with numeric data types (Figure 4. Now the variable X contains serial numbers from 0 to 39, because it has data for 40 months. The variable Y contains the total sales volumes for each month of the selected category.

```
# We convert the date to an index because linear regression can only work with numbers, not dates
X = np.array(range(len(electric_df)))[:, np.newaxis]
y = electric_df['electric batteries'].values
```

Fig. 4: Selecting a tool category, converting the date to an index (own processing)

5.5 Data division

Now the data can be divided into two sets. The first set is the training set. It contains 80% of the items. The second set is the test set and contains the remaining 20% of the items. This is ensured by multiplying by a coefficient of 0.2 in the “test_size=0.2” command, as shown in Figure 5. The random_state parameter affects the process of randomly dividing the data into training and test sets.

```
# Splitting the data into training and testing sets, I chose a ratio of 80% and 20%
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=41)
```

Fig. 5: Data division (own processing)

5.6 Algorithm selection and learning

Now comes the model selection phase. Several models can be used for this type of problem, such as LSTM, XGBoost, or a Decision tree. In the case of a larger amount of data, they could be more accurate, but at the same time, it would be necessary to adjust their so-called hyperparameters for these models. Even if the neural network were more accurate, the results would then be very difficult to explain. A machine learning model called linear regression will be used in the work. Thanks to its easy interpretability and universality, it is suitable for this problem. Another essential feature that played a role in the selection was the enormous simplicity of this method, thanks to the absence of hyperparameters and its ease of use. These are all the reasons why this algorithm was chosen, as can be seen in Figure 6. Next, the model is applied to the training data set and allowed to “learn” our data.

```
# Application of a machine learning model, linear regression
model = LinearRegression()
model.fit(X_train, y_train)
```

Fig. 6: Machine learning model selection (own processing)

5.7 Testing the algorithm

The selected model is now tested and evaluated. Figure 7 shows how the algorithm is tested on the second data set. The lower the values, the more accurate the model. An R-squared, or coefficient of determination, of 0.89 is a good result, and the model can be evaluated as suitable for use.

```
# Create prediction on the test set
y_pred = model.predict(X_test)

# Evaluate model accuracy
mse = mean_squared_error(y_test, y_pred)
r2 = r2_score(y_test, y_pred)

print("Mean Squared Error: {:.2f}".format(mse))
print("R-squared: {:.2f}".format(r2))
```

Fig. 7: Model testing and evaluation (own processing)

5.8 Applying the algorithm

The model application produces an output of 12 numbers, representing the sales volume prediction for the next 12 months (Figure 8). This is the last step for this tool category, as the prediction is complete. Now the entire sequence of these steps must be repeated for the second tool category. If necessary, you can override the number "+ 12" in this step to obtain a prediction for any number of months. However, it is usually recommended to predict only volumes for the nearest subsequent period, because then the inaccuracy increases significantly, as will be shown by the confidence interval graph.

```
# Future prediction
future_X = np.array(range(len(electric_df), len(electric_df) + 12))[:, np.newaxis]
future_y = model.predict(future_X)
print("Future Predictions:", future_y)
```

Fig. 8: Model application and result (own processing)

Finally, it is necessary to create an Excel file where the obtained predictions can be entered. This step is optional because the prediction is already done. Generally, this is considered the correct, if not necessary, procedure, because the output is displayed directly in the Excel file and no mistake is made by manual rewriting (Figure 9). In practice, this is required so that the data is automatically prepared for interpretation.

```
# Creating an Excel file where data will be saved
workbook = openpyxl.Workbook()

# Selecting the active worksheet
worksheet = workbook.active

# Setting column headers
worksheet.cell(row=1, column=1, value="Product")
worksheet.cell(row=1, column=2, value="Month")
worksheet.cell(row=1, column=3, value="Predicted Value")

# Loop to write the predicted values for each month
now = datetime.datetime.now()
for i in range(len(future_y)):
    next_month = now + datetime.timedelta(days=30 * i)
    value = future_y[i]
    worksheet.cell(row=i+2, column=1, value='electric tools')
    worksheet.cell(row=i+2, column=2, value=next_month.strftime("%Y-%m"))
    worksheet.cell(row=i+2, column=3, value=value)

# Save the .xlsx file to the computer
workbook.save("Prediction_for_electric_tools.xlsx")
```

Fig. 9: Creating an Excel file (own processing)

Table 3: Resulting table in Excel (own processing)

Product	Month	Predicted Value
electric tools	2023-05	79.08814726
electric tools	2023-06	80.68759517
electric tools	2023-07	82.28704308
electric tools	2023-08	83.88649099
electric tools	2023-09	85.48593889
electric tools	2023-10	87.08538678
electric tools	2023-11	88.68483471
electric tools	2023-12	90.28428262
electric tools	2024-01	91.88373051
electric tools	2024-02	93.48317843
electric tools	2024-03	95.08262634
electric tools	2024-04	96.68207425

The resulting table that was generated is shown in Table 3. It shows how many pieces of battery-powered tools will be sold in the next 12 months. Finally, the results are rounded to whole numbers. Table 4 is created by combining the predicted volume of pieces of electric, or corded, tools and battery-powered cordless tools.

Table 4: Rounded prediction results (own processing)

Month, Year	May 2023	Jun 2023	Jul 2023	Aug 2023	Sep 2023	Oct 2023	Nov 2023	Dec 2023	Jan 2024	Feb 2024	Mar 2024	Apr 2024
Electric (pcs)	74	74	74	74	74	74	74	73	73	73	73	73
Battery (pcs)	79	81	82	84	85	87	89	90	92	93	95	97

5.9 Calculation of the confidence interval and description of the graphs

In this phase, the 95% confidence interval of the prediction of both categories was calculated using the script shown in Figure 10. This procedure produces the graphs in Figures 11 and 12.

- Trains a **linear regression model** on past sales data.
- Predicts **future sales** for electric tools.
- Calculates a **95% confidence interval** around the predictions.
- Plots:
  - **Blue line** = Original sales
  - **Red line** = Predicted sales
  - **Gray area** = Confidence interval

Fig. 10: Creating a confidence interval graph (own processing)

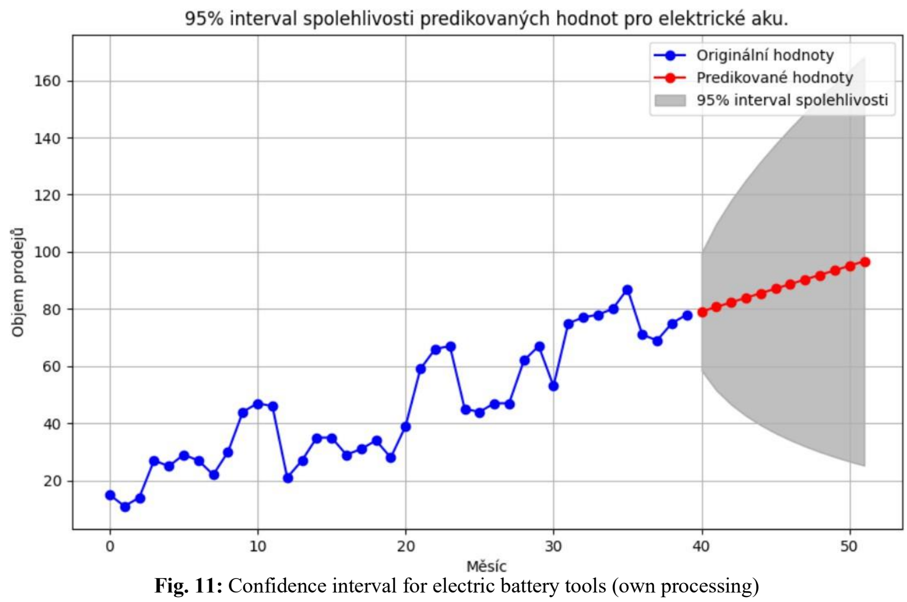

Fig. 11: Confidence interval for electric battery tools (own processing)

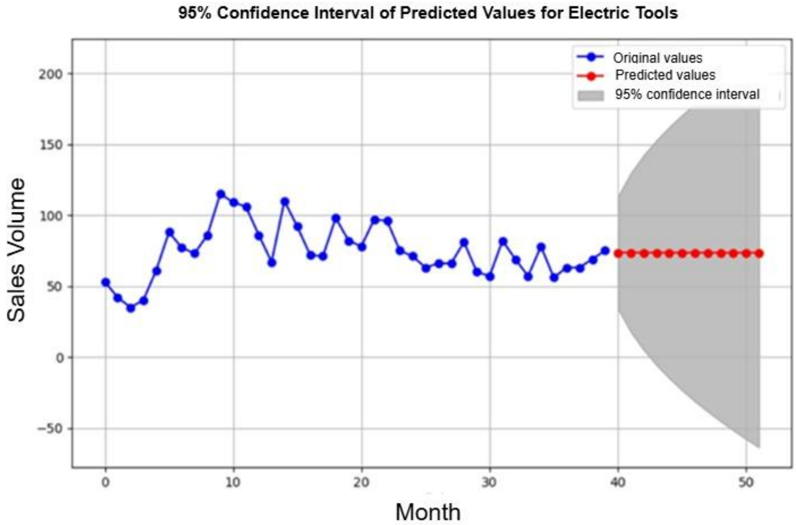

Fig. 12: Confidence interval for electric tools (own processing)

Figures 11 and 12 show the prediction from Table 1 displayed in graphs. The red points represent the predicted values, and the blue points represent the actual original values. The graphs also display the 95% confidence interval of the predicted values. It can be expected with 95% confidence that the actual sales volume will be within this range. This is a statistical term, which in this case gives the range of values represented by a gray area, i.e., the degree of this certainty is shown by a gray field of a funnel shape. With a more distant prediction in time, it can be seen that the inaccuracy increases very quickly and the informative value decreases. Therefore, it is generally recommended to predict sales only for the next closest necessary time period.

6. Interpretation of Results

Figure 13 visualizes the development of the volumes of goods sold by category for each month starting in the past and ending in the future. The graph shows previous sales, starting in January 2020, up to sales for the month of April 2023. These data are followed by the curves from the prediction, which are visually separated by a vertical line, starting at the values of 74 electrical goods and 79 battery goods in May 2023. The curves end in April 2024 at the values of 73 electrical goods and 97 battery goods. The total sales volume of the cordless tools category has visibly grown, and the prediction results show that this trend will most likely continue. The total sales of the power tools category appeared to be declining, and the prediction showed that they would stagnate or decline very slowly.

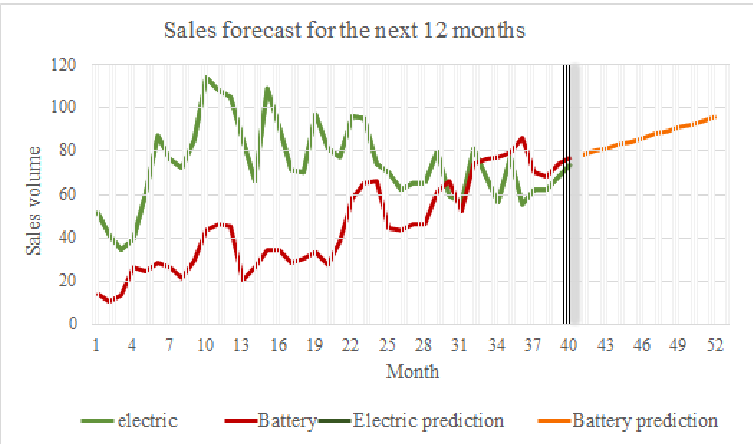

Fig. 13: Sales forecast for the next 12 months (own processing)

The graph shows that in February 2023, the volume of sales of cordless tools was higher than that of corded tools for the first time, and from the information received by the company, this global trend of growth in sales of cordless tools should not change. The company must prepare for this change in demand for various types of goods promptly and implement appropriate measures to prevent a drop in sales. Trading and other companies that can adapt to the market based on this knowledge will have a competitive advantage over others.
It is worth noting the cyclical nature of cordless tool sales in this graph, which had an increase every pre-Christmas period and a steep decline of tens of percent every January. This is probably because the company's largest customers are primarily companies that can reduce their tax base by making this purchase before the end of the accounting period, i.e., before the financial statements, and thus use the opportunity to optimize tax.

6.1 Possible limitations of the prediction

A limitation of the prediction calculation was that it was based on data starting only in 2020. This is because the company previously only had a small number of tools. It was only in 2020 that the company decided to expand this type of goods in its stores. However, tools

are not a priority in 2023 either, and the company does not focus on them much compared to other products. The low sales volumes then result in lower forecast accuracy.
Another drawback is that longer-term forecasts do not make much sense, since the limited input data was sufficient only for the nearest future time period. With each subsequent month, the prediction inaccuracy then rapidly increases, as shown in Figures 15 and 16.
Another limitation is that we cannot know from the prediction what specific brands and goods will be in demand, because it is a prediction of aggregate sales per category. It would certainly be possible to create a more complicated and accurate model using many machine learning methods, taking into account many more factors, but this would be too demanding in terms of the appropriate difficulty of the work. Therefore, the linear regression method was chosen.
The limitation of the forecast calculation is that it is based only on data from 2020. This is because in the past, the company had only limited tools. It was not until 2020 that the company decided to expand the sale of such products in stores. But in 2023, the tools were still not a priority, and the company did not give them much importance compared to other products. The forecast accuracy is low due to the small number of transactions. Another drawback is that long-term forecasts are not meaningful, since the input data was only sufficient for a short period of time. For each subsequent month, the forecast error increases rapidly, as shown in Figures 15 and 16. Another limitation of the forecast is that we cannot know which brands and products are in demand, since it is a forecast of overall sales by category. Of course, a more complex and accurate model can be created using different machine learning methods, taking into account many factors, but this will be quite challenging given the complexity of the task required. Therefore, the linear regression method was chosen. Future research could explore advanced models such as LSTM networks to better capture seasonal fluctuations in wireless device sales (e.g., pre-Christmas spikes) or potential nonlinear trends and correlations in the XGBoost dataset. These approaches could improve long-term forecast accuracy and provide deeper insights into demand shifts.

7. Optimization of Warehouse Stock

This chapter deals with the analysis and optimization of the current state of stocks of the selected assortment. The chapter is based not only on the created forecast, but also on the received sales data. It is already known from the previous chapter how sales will develop in the future, and that one product category will probably increase in sales, while the other will stagnate. Therefore, it is possible to recommend how to prepare for this change from the company's perspective. Next, another indicator of inventory turnover time will be calculated. It will also be calculated how the company should work with inventory in order to remain as efficient as possible. Next, the suppliers will be segmented according to the ABC method. Finally, appropriate measures will be proposed on how the company should approach these changes in the market and what steps to take.

7.1 Analysis of the current state of the inventory of the tool assortment

As of May 1, 2023, the company had a total of 1,776 items of power tools and 1,078 items of cordless tools in stock. Since publishing the exact purchase prices of the goods would damage the future negotiating position with suppliers, these prices were not provided to me. The purchase prices will therefore be calculated using the previously calculated average margin. The same 25% margin will therefore be assumed for all products, even though this is an inaccurate assumption.
As of May 1, 2023, the total inventory value of the tool assortment, comprising both power tools and cordless tools, amounts to approximately ₹90.88 lakh. Out of this, around ₹63.18 lakh is tied up in power tools, while ₹27.70 lakh is allocated to cordless tools. This reveals a noticeable imbalance in capital distribution within the inventory, with a significantly larger share invested in power tools. The calculated values highlight how the company’s capital is disproportionately tied up in one segment, which may impact inventory turnover and overall financial efficiency. The graph (Figure 15) shows the prediction of the ratio of product sales by category in April 2024. Comparing the graphs in Figures 14 and 15, we see that the ratios differ significantly, which may be the problem. The company still has significantly more cash tied up in corded tools, but sales of battery tools are already higher this year and will continue to increase.

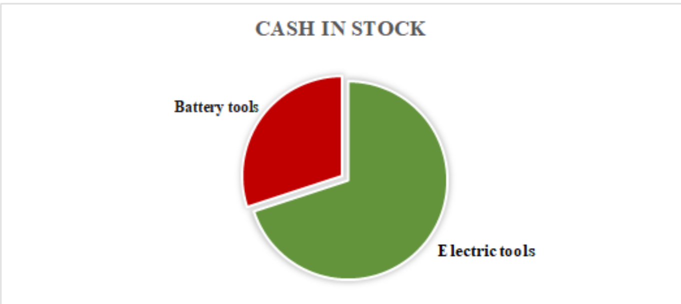

Fig. 14: Capital allocation in tool inventory as of May 1, 2023, higher investment in power tools (own processing)

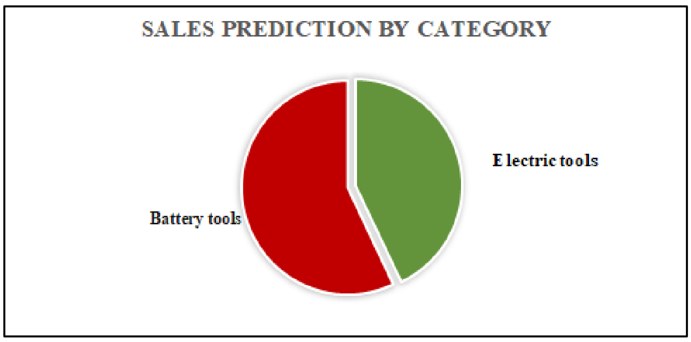

Fig. 15: Forecasted sales distribution in April 2024, showing a shift toward higher demand for cordless tools (own processing)

8. Application of ABC Analysis

This ABC analysis aims to differentiate the tool inventory (Table 5) into three categories to show which brands of stocked tools have the most capital tied up. This division allows for more efficient inventory management. The classification criterion is the cumulative purchase price. The categories are A, B, and C. To determine the limits, the data was sorted in descending order by the total purchase price, which consists of the number of stocked tools from a given supplier multiplied by their price. The previously calculated margin of 25% is deducted from the price, since the prices in the table are sales, not purchase prices. Each tool is priced differently, and therefore, only the total is shown here. Category A will consist of brand items that make up 60% of the funds. Category B will include brand items that make up 20
% and the remaining items will belong to category C. The distribution of inventory into individual categories according to the selected criterion and the set limits is shown in Table 6.

Table 5: Data for ABC analysis (own processing)

Rank	Brands in Stock	Number of Items	Cumulative Price [INR]	Cumulative Price Without Margin [INR]
1.	Cordless Stanley	207	₹3,423,715	₹2,739,374
2.	Cordless Makita	77	₹10,072,249	₹8,057,800
3.	Cordless Bosch	11	₹976,204	₹780,962
4.	Cordless Milwaukee	195	₹8,033,526	₹6,426,821
5.	Cordless DeWALT	351	₹7,325,636	₹6,180,509
6.	Cordless Bosch	237	₹3,424,999	₹2,739,207
7.	Stanley	245	₹2,404,332	₹1,923,466
8.	Bosch	55	₹6,491,196	₹5,192,957
9.	Bosch	578	₹53,463,182	₹42,770,545
10.	DeWALT	403	₹7,803,353	₹6,242,681
11.	Bosch	495	₹8,222,025	₹6,577,621

Table 6: Application of ABC analysis (own processing)

Rank	Brands in Stock	Cumulative Price Without Margin [INR]	Cumulative Price Without Margin [%]	Cumulative Total [%]	Category
9.	Bosch	₹42,770,545	47.72%	47.7%	A
2.	Cordless Makita	₹8,059,801	8.99%	56.7%	A
11.	Bosch	₹6,577,621	7.34%	64.1%	B
4.	Cordless Milwaukee	₹6,426,821	7.17%	71.2%	B
10.	DeWALT	₹6,242,681	6.96%	78.2%	B
5.	Cordless DeWALT	₹6,180,509	6.90%	85.1%	B
8.	Bosch	₹5,192,957	5.79%	90.9%	C
6.	Cordless Bosch	₹2,739,207	3.06%	93.9%	C
1.	Cordless Stanley	₹2,739,000	3.06%	97.0%	C
7.	Stanley	₹1,923,466	2.15%	99.1%	C
3.	Cordless Bosch	₹780,962	0.86%	100.0%	C
Σ		₹89,633,570	100%		

Calculation of the percentage of items by category:
Items A: (578 + 77) / 2,854 x 100 ≐ 23%
Items B: (495 + 195 + 403 + 351) / 2,854 x 100 ≐ 51%
Items C: (55 + 237 + 207 + 245 + 11) / 2,854 x 100 ≐ 26%
The ABC analysis for the Vadodara-based company shows that although Category A items represent only about 23% of the total inventory, they account for nearly 60% of the financial capital tied up in stock. Category A is primarily made up of high-value brands such as Bosch and Makita, which dominate the tool assortment in terms of capital allocation. Category B items, which form the majority at 51% of the stock, account for roughly 20% of the tied-up capital. Meanwhile, Category C items make up the remaining 26% of inventory but represent just 10% of the capital.
This distribution highlights the importance of closely managing Category A brands like Bosch and Makita to optimize financial resources. These suppliers directly influence overall inventory efficiency and profitability. Category B brands such as DeWALT, Stanley, and Milwaukee play a supporting role, requiring balanced stock levels to maintain availability without overinvestment. Category C items, although numerous, contribute relatively little to capital usage and should be managed through leaner stocking strategies.

8.1 Inventory Turnover Rate

From the data received, the number of units sold in the last month is first determined. In this case, in April 2023, it is 75 units of power tools and 78 cordless tools sold. Now enough values are known to calculate an important indicator in inventory management, the inventory turnover time. The inventory turnover time in this case is calculated as: inventory/sales for this month.
Calculation for power tools:

$$\frac{17{,}418{,}687}{18{,}260 \times 75} = 12.7 \text{ months}$$

Calculation for cordless tools:

$$\frac{7{,}479{,}186}{11{,}941 \times 78} = 8 \text{ months}$$

The results show that the turnover time for the power tools category is worse than for the cordless tools category. The result of 12.7 months is not very good. Assuming a margin range of 20–30%, inventory turnover for corded tools could vary between 11.5 and 13.5 months, highlighting the need for precise margin data. The inventory turnover time also changes as sales develop over time.

In April 2023, there were more corded tools in stock than cordless tools. The average purchase price of corded tools that were in stock was also higher than the average purchase price of corded tools. The inventory is significantly biased towards power tools, although the sales volumes of both categories are similar, and we can assume that the higher turnover will be mainly battery tools in the future. The significantly lower inventory turnover time of battery tools means that it is more profitable for the company. The company should look for a solution to reduce the inventory turnover time for power tools. An inventory turnover time of 12.7 months means a large amount of inventory is held, which can lead to high storage costs and the risk of inventory obsolescence. Battery tools with an inventory turnover time of 8 months represent a much better result, which could improve even further if the sales of these goods increase.

8.2 Proposal for recommendations

The last step is to propose measures that would help the company reduce costs and optimize its inventory. Based on the assessment of the current stock position, this section outlines strategies for cost savings through better management of tied-up capital in inventory. Since demand for tools is evolving, the company must proactively adapt to these changes.

Maintaining both electric and cordless tools in the assortment is beneficial, as this strengthens the company's position as a one-stop provider and improves customer satisfaction. Demand for both product categories remains consistent, with cordless tools showing stronger growth.

From a sales perspective, the company should reduce or at least carefully regulate new tool purchases unless there are surplus funds available. If cash flow is tight, the company may prioritize selling off existing stock through promotions, targeted campaigns, or special discounts. Additionally, there is a need to balance capital distribution across corded and cordless tools. Increasing investment in battery-powered tools, which show better turnover, would align inventory with future sales trends.

Considering inventory turnover, the funds tied up in electric (corded) tools should be reduced. With turnover time for these tools being relatively high, the recommendation is to decrease their stock levels and instead boost sales through marketing initiatives. Simultaneously, the company should improve demand forecasting (at the brand level), negotiate shorter lead times with suppliers, and improve storage and distribution efficiency to avoid overstocking.

The ultimate goal is to reduce the overall inventory without risking stockouts, thereby improving liquidity and reducing carrying costs.

Finally, the ABC analysis indicates that to remain competitive, focus should be placed on Category A brands — Bosch and Makita — which account for the largest share of tied-up funds. Strengthening partnerships with these suppliers (e.g., through bulk purchasing discounts or negotiated lead times) will further improve cost efficiency.

Category A suppliers such as Bosch and Makita should be managed with a focus on cost efficiency and supply reliability. The company can negotiate shorter lead times to reduce excess stockholding and improve turnover. At the same time, bulk purchase agreements or long-term contracts could help secure discounts and ensure a steady supply. Collaborative demand planning with these suppliers would further align procurement with sales trends, lowering storage costs and strengthening competitiveness.

9. Conclusion

This work focused on handling large volumes of data and applying machine learning techniques to predict sales of selected products for a company, reimagined in the context of a large Indian wholesale enterprise. The practical part shows an overview of the company's profile and a brief financial analysis. It then progressed into building and deploying a predictive model to estimate the number of units expected to be sold in the following year. The methodology, model selection, and implementation steps were explained clearly, emphasizing relevance to real-world Indian inventory and sales dynamics.

Using these forecasts, an in-depth analysis of the current inventory—especially focused on tool assortments—was performed. The results were then compared to actual stock levels, allowing for data-driven insights and actionable recommendations aimed at improving inventory efficiency and reducing unnecessary costs. The study concluded with an ABC inventory classification, a method particularly effective for Indian businesses managing a wide range of Stock Keeping Units with varying importance and movement rates.

References

[1] Odimarha, A. C., Ayodeji, S. A., & Abaku, E. A. (2024). Machine learning's influence on supply chain and logistics optimization in the oil and gas sector: A comprehensive analysis. Computer Science and IT Research Journal, 5, 725–740.

[2] Pasupuleti, V., Thuraka, B., Kodete, C. S., & Malisetty, S. (2024). Enhancing supply chain agility and sustainability through machine learning: Optimization techniques for logistics and inventory management. Logistics, 8(3), 73. https://doi.org/10.3390/logistics8030073

[3] Akbari, M., & Do, T. N. A. (2021). A systematic review of machine learning in logistics and supply chain management: Current trends and future directions. Benchmarking: An International Journal, 28, 2977–3005.

[4] Yang, M., Lim, M. K., Qu, Y., Ni, D., & Xiao, Z. (2023). Supply chain risk management with machine learning technology: A literature review and future research directions. Computers & Industrial Engineering, 175, 108859.

[5] Daios, A., Kladovasilakis, N., Kelemis, A., & Kostavelis, I. (2025). AI applications in supply chain management: A survey. Applied Sciences, 15(5), 2775. https://doi.org/10.3390/app15052775

[6] Sayyad, J., Attarde, K., & Yilmaz, B. (2024). Improving machine learning predictive capacity for supply chain optimization through domain adversarial neural networks. Big Data and Cognitive Computing, 8(8), 81. https://doi.org/10.3390/bdcc8080081

[7] Dey, P. K., Chowdhury, S., Abadie, A., Vann Yaroson, E., & Sarkar, S. (2024). Artificial intelligence-driven supply chain resilience in Vietnamese manufacturing small-and medium-sized enterprises. International Journal of Production Research, 62, 5417–5456.

[8] Nimmagadda, V. S. P. (2020). AI-powered predictive analytics for retail supply chain risk management: Advanced techniques, applications, and real-world case studies. Distributed Learning and Broad Applications in Science Research, 6, 152–194.

[9] Pal, S. (2023). Integrating AI in sustainable supply chain management: A new paradigm for enhanced transparency and sustainability. International Journal of Research in Applied Science and Engineering Technology, 11, 2979–2984.
[10] Charles, V., Emrouznejad, A., & Gherman, T. (2023). A critical analysis of the integration of blockchain and artificial intelligence for supply chain. Annals of Operations Research, 327, 7–47.
[11] Curcio, D., & Longo, F. (2009). Inventory and internal logistics management as critical factors affecting the supply chain performances. International Journal of Simulation and Process Modelling, 5, 278–288.
[12] Sharma, N., & Singhi, R. (2018). Logistics and supply chain management quality improvement of supply chain process through vendor managed inventory: A QFD approach. Journal of Supply Chain Management Systems, 7, 23–33.
[13] Mahraz, M. I., Benabbou, L., & Berrado, A. (2022). Machine learning in supply chain management: A systematic literature review. International Journal of Supply and Operations Management, 9, 398–416.
[14] Abolghasemi, M., Beh, E., Tarr, G., & Gerlach, R. (2020). Demand forecasting in supply chain: The impact of demand volatility in the presence of promotion. Computers & Industrial Engineering, 142, 106380.
[15] Salamai, A. A., El-Kenawy, E. S. M., & Abdelhameed, I. (2021). Dynamic voting classifier for risk identification in supply chain 4.0. Computers, Materials & Continua, 69(3).
[16] Suwignjo, P., Panjaitan, L., Baihaqy, A., & Rusdiansyah, A. (2023). Predictive analytics to improve inventory performance: A case study of an FMCG company. Operations and Supply Chain Management: An International Journal, 16, 293–310.
[17] Mahraz, M.-I., Benabbou, L., & Berrado, A. (2022). Machine learning in supply chain management: A systematic literature review. International Journal of Supply and Operations Management, 9(4), 398–416.
[18] Charles, V., Emrouznejad, A., & Gherman, T. (2023). A critical analysis of the integration of blockchain and artificial intelligence for supply chain. Annals of Operations Research, 327, 7–47.