International Journal of Information Technology and Computer Science Applications (IJITCSA)
p-ISSN: 2964-3139 e-ISSN: 2985-5330
Vol. 03, No. 01, page 1 - 10
Submitted 15/12/2024; Accepted 15/1/2025; Published 16/1/2025

Prediction Model for Product Stock Procurement Using the Naive Bayes Method

[1]Rafika Sari and [2]Muhammad Ashardi
[1]Informatics Department, Universitas Bhayangkara Jakarta Raya, Bekasi, INDONESIA
[2]Information Systems, Universitas Bhayangkara Jakarta Raya, Bekasi, INDONESIA
e-mail : [1]rafika.sari@dsn.ubharajaya.ac.id, [2]muhammad.ashardi18@mhs.ubharajaya.ac.id

Corresponding Autor: Rafika Sari

Abstract
Product sales circulation and product stock repurchase play a strategic role in meeting customer needs and increasing a company's profits. PT Rotaryana Engineering as a company that provides spare parts for electronic kitchen products in various brands is experiencing rapid dynamics in the sales of its various products. Therefore, a decision-making system is needed regarding spare parts stock policies that will support increased sales and customer satisfaction for the company. This research will model a product stock procurement prediction system using the Naive Bayes method. Sales data for one year will be used to categorize the availability of each product, namely the available category and the not available category. Furthermore, calculations using the Naive Bayes method produce likelihood probability values for each product item which are used to predict and recommend procurement of spare parts stock. This information can be used as a basis for determining priorities for procuring stock of the most popular goods and reducing stock of less popular goods so that the circulation of the company's product stock becomes more efficient and effective.

Keywords— predictions, spare parts, restock recommendations, naive bayes

1 Introduction

The manufacturing industry, especially those related to spare parts sales and service, plays an important role in supporting various economic sectors. PT. Rotaryana Engineering is a company engaged in service & maintenance of kitchen equipment. PT. Rotaryana Engineering currently provides spare parts and service services for many well-known brands such as Milnor, Electrolux, Masterchef, Mastercool, Scotsman, Bun, Fagor, Bezzera, Mazzer, Allbev, Allcoff, Vitamix, Cambro, Sankosha, Amercian Dryer Corporation, Frymaster, Hobart, HBS, ISI, Mastrena, etc. PT. Rotaryana Engineering customers, almost all 5-star hotels in Indonesia are regular customers PT. Rotaryana Engineering. Apart from that, PT Rotaryana Engineering also supplies kitchen equipment to many restaurant partners in Indonesia [1]. Therefore, as a company that focuses on selling spare parts, PT Rotaryana Engineering plays a strategic role in meeting customer needs regarding equipment maintenance and repair. In this case, the Electrolux brand is the brand with the most sales at PT Rotaryana Engineering. In the period May 2023 - May 2024, there were 609 different items sold for the Electrolux brand with total sales of more than 4 billion Rupiah.

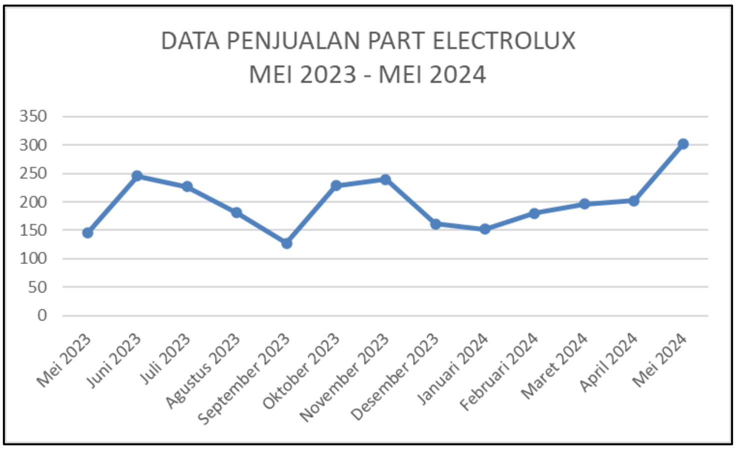

Figure 1. Electrolux Part Sales Data [1]

Figure 1 is a diagram that shows sales data for Electrolux spare parts every month. It can be seen from the diagram that sales of Electrolux spare parts are always above 100 items every month. Analysis of sales of kitchen equipment spare parts is very important to understand sales patterns, customer needs and effective marketing strategies. In this way, companies can increase spare parts sales, increase customer satisfaction and increase profits.

Stock management of kitchen equipment spare parts is also very important to ensure that the company has sufficient and timely inventory to meet customer needs. Insufficient spare parts stock can cause product shortages, increase shipping costs, and reduce customer satisfaction. On the other hand, excessive spare parts stock can cause unnecessary costs and reduce operational efficiency. Therefore, analysis related to sales of kitchen equipment spare parts and stock management of kitchen equipment spare parts is very important to ensure the sustainability and growth of the company.

This research analyzes sales of kitchen equipment spare parts and stock management of kitchen equipment spare parts at PT. Rotaryana Engineering using the Naive Bayes method. The primary data used was obtained directly from the company consisting of sales data & stock data. The Naive Bayes method was chosen in this research because the Naive Bayes classifier is a classification method rooted in probability and statistical theorems to predict future opportunities based on past experience. The Bayesian theorem calculates the posterior probability value $P(H|X)$ using the probabilities (H), $P(X)$, dan $P(X|H)$. The value X is testing data whose class is not yet known. The H value is the data hypothesis X which is a more specific class. The $P(X|H)$ value, also known as likelihood, is the probability of hypothesis Equation 1 is a formulation of Bayes' theorem [2], [3].

$$P(H|X) = \frac{P(X|H) \cdot P(H)}{P(X)} \tag{1}$$

Several studies used as references in this research include: (i) research conducted by Hamidah, D.A. (2024) Sentiment Analysis of TMLST Coffee Customer Reviews Using the Naïve Bayes Algorithm [4], the method used in this research is Naïve Bayes which aims as a classification method and is able to classify customer comments into positif or negatif; (ii) research by Alkarim, I. (2022) implementing spare part inventory control using the buffer stock and ROP methods [5]. Where in the stock control form there are service levels, lead times and forecasting so that buffer stock and ROP calculations can be obtained; (iii) Sinamon, S.D.C., (2020) created a spare parts stock application to provide information on items in inventory at CV Bumi Waras Heavy Equipment Division [6]; (iv) Sarmah. (2018) Spare Parts Inventory Management Literature and Direction Towards the Use of Data Mining Technique [7]. The industries have implemented integrated information systems or enterprise resource planning applications to automate their business process, reduce lead time, improve productivity and reduce cost; and (iv) Putra, W. (2022) predicted stock of goods in the future or next period with quite good model test values [8].

2 Research methods

This study employs a systematic approach to develop a prediction model for stock procurement using the Naive Bayes algorithm. The methodology consists of several sequential stages aimed at ensuring a comprehensive analysis and reliable results. These stages include data collection, preprocessing, model development, and evaluation [2], [9].

2.1 *Data Collection*

The first step involves gathering historical sales data for specific products. The dataset used in this research was sourced from company records of past transactions, covering variables such as product codes, product names, sales volume, stock availability, and purchase frequency. The timeframe for the data spans a complete business cycle, typically one year, to capture seasonality and demand trends. Additional data such as supplier lead times and customer demand patterns are also integrated to enrich the dataset and improve model accuracy.

2.2 *Data Prepocessing*

To ensure the dataset is ready for analysis, the data preprocessing phase includes the following steps:

- **Data Cleaning**: Erroneous, duplicate, or missing data entries are identified and corrected. Missing values for sales or stock availability are imputed using statistical methods such as mean substitution or interpolation.
- **Data Normalization**: Variables are transformed to a consistent format to facilitate uniformity. For instance, stock quantities are standardized to consistent units where discrepancies exist.
- **Categorization**: Data on stock availability is categorized into binary classes: "Available" and "Unavailable." The threshold for classification is determined based on business rules, such as minimum stock levels required to avoid stockouts.
- **Feature Selection**: Variables that significantly impact stock prediction, such as sales frequency, lead time, and product turnover rates, are selected for further analysis. Irrelevant or redundant attributes are excluded.

2.3 *Model Development*

The core of this methodology is the application of the Naive Bayes algorithm, a probabilistic classifier based on Bayes theorem. The development process includes [10]:

- **Training and Testing Split**: The dataset is divided into training and testing subsets, typically with a ratio of 80:20. The training set is used to develop the model, while the testing set evaluates its performance.
- **Class Probability Calculation**: The probabilities of each class ("Available" and "Unavailable") are computed based on the training data.

$$P(available) = \frac{amount\ of\ stock\ data\ available}{total\ data}$$
$$P(unavailable) = \frac{amount\ of\ stock\ data\ unavailable}{total\ data} \tag{2}$$

- **Likelihood Estimation**: Conditional probabilities for each feature are calculated with respect to the classes. For example, the likelihood of a product being unavailable given its sales frequency or historical stock trends is determined.

$$P(X_i|available) = \frac{the\ amount\ of\ data\ with\ feature\ X_i\ in\ the\ available\ class}{amount\ of\ stock\ data\ available}$$
$$P(X_i|unavailable) = \frac{amount\ of\ data\ with\ feature\ X_i\ in\ the\ unavailable\ class}{amount\ of\ stock\ data\ unavailable} \tag{3}$$

- **Model Integration**: The Naive Bayes formula combines prior probabilities and likelihoods to predict the class of new observations.

$$P(available|X) = \frac{P(available) \cdot P(X|available)}{P(X)}$$
$$P(unavailable|X) = \frac{P(unavailable) \cdot P(X|unavailable)}{P(X)} \tag{4}$$

Since $P(X)$ is the same for both classes (available/unavailable), it can be ignored in the comparison:

$$P(available|X) \propto P(available) \cdot P(X|available)$$
$$P(unavailable|X) \propto P(unavailable) \cdot P(X|unavailable) \tag{5}$$

2.4 *Evaluation and Validation*

The model is evaluated using performance metrics such as accuracy, precision, recall, and F1-score. The confusion matrix includes the following values [2], [3], [9]:

- True Positives (TP): Correct predictions for the positive class.
- True Negatives (TN): Correct predictions for the negative class.
- False Positives (FP): Incorrect predictions, where the model predicts positive while the actual class is negative.
- False Negatives (FN): Incorrect predictions, where the model predicts negative while the actual class is positive.

The testing dataset is used to generate predictions, which are compared against actual outcomes to assess model reliability. Cross-validation techniques are employed to ensure the robustness of the model, minimizing the risk of overfitting or underfitting.

$$\begin{aligned} accuracy &= \frac{TP + TN}{TP + TN + FP + FN} \\ Precision &= \frac{TP}{TP + FP} \\ Recall &= \frac{TP}{TP + FN} \\ F1 - Score &= 2 \cdot \frac{Precision \ \cdot Recall}{Precision + Recal} \end{aligned} \tag{5}$$

Cross-validation involves dividing the dataset into k equal parts (folds). Each fold is used as test data once, while the remaining folds are used as training data. This process is repeated k times.

$$Average\ accuracy = \frac{\sum_{i=1}^{k} Accuracy_i}{k} \tag{6}$$

$$Performance = \frac{1}{k}\sum_{i=1}^{k}[Accuracy_i, Precision_i, Recall_i, F1 - Score_i] \tag{7}$$

2.5 *Deployment and Recommendations*

The final stage involves deploying the model to predict future stock requirements. Predictions are used to generate actionable insights for procurement planning. Additionally, the study provides recommendations to improve stock management practices, such as identifying fast-moving items that require higher stock levels or optimizing reorder quantities for low-demand products [11]. This methodology ensures that the Naive Bayes model delivers accurate and actionable predictions, enabling businesses to enhance operational efficiency, reduce stockouts, and optimize inventory costs.

3 Results and Discussion

3.1 *Data Normalization*

Data preprocessing is critical for ensuring the accuracy of predictive models. In this study, raw sales data of spare parts from PT. Rotaryana Engineering was normalized to prepare it for analysis. The dataset initially included 12 columns with redundancies and irrelevant attributes, such as group information and customer codes. The normalization process involved: (i) **Eliminating Unnecessary Columns:** Non-essential columns like group, unit type, and customer details were excluded, and (ii) **Removing Redundancies:** Repeated records for the same item were consolidated, with quantities summed up to provide unique entries for each spare part (figure 2).

Grup Barang	Kode	Nama Barang	Satuan	Qty	Harga Satuan	No SO	Kode Customer	Perusahaan	Tanggal	Status
ELECTROLUX	002544	Ignition Cable; 700 mm (006514)	PC	2	288000	41100	NUSANT001	NUSANTARA SEJAHTERA RAYA,	01/07/2023	Selesai
ELECTROLUX	002544	Ignition Cable; 700 mm (006514)	PC	2	235600	43323	NUSANT001	NUSANTARA SEJAHTERA RAYA,	06/09/2023	Selesai
ELECTROLUX	002544	Ignition Cable; 700 mm (006514)	PC	2	313875	46546	PAWONB001	PAWON BOGA INTERNUSA, PT	01/11/2023	Selesai
ELECTROLUX	002544	Ignition Cable; 700 mm (006514)	PC	3	288000	55597	GRAHAAL001	GRAHA ALAM LESTARI, PT	03/05/2024	Dikirim
ELECTROLUX	002605	PIEZOELECTRIC IGNITER	PC	1	432432	56388	MDAABSK001	PT. MDA ABS KUNINGAN	03/05/2024	Pesanan
ELECTROLUX	002605	PIEZOELECTRIC IGNITER	PC	3	504000	54050	GRAHAAL001	GRAHA ALAM LESTARI, PT	21/03/2024	Selesai
ELECTROLUX	002626	LAMP HOLDER GASKET	PC	3	376448	37415	GUNUNG003	GUNUNG ANSA, PT	19/06/2023	Selesai
ELECTROLUX	002626	LAMP HOLDER GASKET	PC	2	376000	47963	CITRAK002	CITRA KIRANA PANTAI, PT	28/11/2023	Selesai
ELECTROLUX	002626	LAMP HOLDER GASKET	PC	1	396900	56056	PELITA001	PELITA BOGA SEJAHTERA, PT	25/04/2024	Pesanan
ELECTROLUX	002627	LAMP GLASS	PC	3	250965	37415	GUNUNG003	GUNUNG ANSA, PT	19/06/2023	Selesai
ELECTROLUX	002627	LAMP GLASS	PC	2	256000	47963	CITRAK002	CITRA KIRANA PANTAI, PT	28/11/2023	Selesai
ELECTROLUX	002627	LAMP GLASS	PC	3	256001	56191	CITRAK002	CITRA KIRANA PANTAI, PT	25/04/2024	Pesanan

Figure 2. Electrolux Spare Parts Sales Data Snapshot

The first step in data normalization is selecting existing data according to needs, the item group column is not needed because all goods have the same group, namely Electrolux, the unit column is also not needed because all the units are "pcs", the unit price column, so number, customer code , company, date and status are also not needed because the focus of the analysis is part expenditure prediction.

Kode	Nama Barang	Qty
002544	Ignition Cable; 700 mm (006514)	2
002544	Ignition Cable; 700 mm (006514)	2
002544	Ignition Cable; 700 mm (006514)	2
002544	Ignition Cable; 700 mm (006514)	3
002605	PIEZOELECTRIC IGNITER	1
002605	PIEZOELECTRIC IGNITER	3
002626	LAMP HOLDER GASKET	3
002626	LAMP HOLDER GASKET	2
002626	LAMP HOLDER GASKET	1
002627	LAMP GLASS	3
002627	LAMP GLASS	2
002627	LAMP GLASS	3

Figure 3. First data normalization

The second step in data normalization is to eliminate redundancies or repetitions. It can be seen in Figure 3 that there is still data redundancy. To eliminate redundancy we can add up the qty for each item with the same code as code 002544 Ignition Cable; 700 mm which appears 4 times with qty 2, 2, 2, 3, we can combine them into 1 data by adding the qty to 9 (figure 4).

Kode	Nama Barang	Qty
002544	Ignition Cable; 700 mm (006514)	9
002605	PIEZOELECTRIC IGNITER	4
002626	LAMP HOLDER GASKET	6
002627	LAMP GLASS	8

Figure 4. Second data normalization

Through this process, the dataset was refined to three key attributes: product code, product name, and quantity sold. This reduced data redundancy and enhanced clarity for subsequent analysis.

3.2 *Manual Naive Bayes Calculations*

The Naive Bayes algorithm was applied manually to classify stock availability based on sales data. Key steps included:

a. **Training Data Preparation:** A dataset of 609 unique items was categorized into two classes: "Available" (stock ≥ 5) and "Unavailable" (stock < 5). Only 4 items were classified as "Available," while the remaining 605 were "Unavailable."
b. **Testing Data Selection:** A subset of items was used to test the classification model.
c. **Class Probabilities:**
   - Probability of "Available": P(Available)=0.006568P
   - Probability of "Unavailable": P(Unavailable)=0.993432
d. **Likelihood Estimation:** For each item, probabilities were calculated using sales quantities and stock levels. The likelihood values helped determine the most likely classification for each item.

This analysis identified items requiring restocking, as their classification suggested a higher probability of being "Unavailable."

Table 1. Training data

No	Code	Item Products	Total Sales	Stock of Product	Status
1	2544	Ignition Cable; 700 mm (006514)	9	18	available
2	2605	PIEZOELECTRIC IGNITER	4	0	unavailable
3	2626	LAMP HOLDER GASKET	6	2	unavailable
4	2627	LAMP GLASS	8	0	unavailable
5	2672	KNOB	1	0	unavailable
6	2689	Heating Element Gasket	14	0	unavailable
7	2692	CAP; DESCALING LOAD IN THE BOILER	5	0	unavailable
8	2699	BEARING; FOR MOTOR SHAFT	3	0	unavailable
9	2730	Safety Thermostat; 338 C	1	0	unavailable
10	2873	O-RING; I=MM.28,17X3,53	4	0	unavailable
⋮	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮
608	0CB661	HALOGEN LAMP 12V 10W G4	14	8	available
609	0Q0002	O-RING; 8.73 X 1.78 SET 10 PCS	1	6	available
Sum Total Penjualan			**2639**	**104**	
Average Total Penjualan			**4,333**		

The Naïve Bayes algorithm is used to calculate the probability of each attribute in the training data. Primary data of 609 different items is shown in Table 1. From the stock data, groupings will be made with the provisions that if the stock of goods is ≥ 5 then it is classified as available class and if the stock of goods is < 5 then it is classified as not available class. With consideration to maintaining the stability of PT. Rotaryana Engineering due to the long indent process of around 2 months starting from the Purchase Request date.

Table 2. Testing data

Code	Item Products	Total Sales	Stock of Product	Status
0K8238	FINGERLINK;OPEN CONNECTION;DIST.120MM	80	0	unavailable
020094	Thermocople; M9 x1 L=120 cm	52	0	unavailable
20094	THERMOCOUPLE;M9X1 L=120CM	52	0	unavailable
0C3162	Interrupted Thermocouple; ORKLY 16400-35	42	0	unavailable
0C9240	PIEZOELECTRIC IGNITER REPL. 2605, 59251	42	0	unavailable
6879	Contactor; 220-240 V 50 Hz / 240-264 V 6	32	0	unavailable
0C0202	Thermocouple; M8 x 1 L= 750 mm (058332)	29	0	unavailable
87664	DOOR GASKET; 405X570 MM	28	0	unavailable
0L2583	PCB;DIWAP AZ60 repl. 0L1811	28	0	unavailable
0C7421	Thermocouple; DIA.33,17; L=1500mm	27	0	unavailable
0C3645	BICONO	2	51	available
2544	Ignition Cable; 700 mm (006514)	9	18	available
0Q0002	O-RING; 8.73 X 1.78 SET 10 PCS	1	6	available

From the classification results based on the conditions above, it was found that of the 609 data items there were 4 data that were in the available class and 605 data that were in the not available class (table 2).

3.3 Computational Implementation of Naive Bayes

To automate and validate the model, Python and Google Colab were used with libraries like Pandas and Scikit-learn. The steps included:

- **Data Import and Preparation:** Sales data was imported and preprocessed. Items were labeled based on whether they required restocking.
- **Model Training:** A Gaussian Naive Bayes model was trained on the normalized dataset, focusing on attributes like total sales and stock levels.
- **Prediction:** The trained model predicted stock statuses for items, flagging those that needed replenishment.

The computational approach mirrored the manual analysis, reinforcing the model's reliability.

```
   Avg Qty Penjualan  stok_barang         status
0           2.250000           18       Tersedia
1           2.000000            0  Tidak Tersedia
2           2.000000            2  Tidak Tersedia
3           2.666667            0  Tidak Tersedia
4           1.000000            0  Tidak Tersedia
5           2.000000            0  Tidak Tersedia
6           2.500000            0  Tidak Tersedia
7           1.500000            0  Tidak Tersedia
8           1.000000            0  Tidak Tersedia
9           1.333333            0  Tidak Tersedia
Probabilitas Prior:
Tidak Tersedia: 0.9934
Tersedia: 0.0066
```

Figure 5. Prior Probability Calculation Results for Each Class

```
Hasil Prediksi Pengeluaran Berdasarkan Data yang Sudah Terjual:
      kode                          nama_barang  total_penjualan  \
0     2544          Ignition Cable; 700 mm (006514)                9
1     2605                 PIEZOELECTRIC IGNITER                4
2     2626                    LAMP HOLDER GASKET                6
3     2627                            LAMP GLASS                8
4     2672                                  KNOB                1
..     ...                                   ...              ...
604  88361                                magnet                9
605  88477  CONTACTOR;220-240V 50HZ/240-264V 60HZ                3
606  88578                  MAGNETIC SWITCH CPL                1
607  93437                     Thermo Regulator                1
608  95194                                  PCB                1
```

```
     stok_barang  Pengeluaran  Prediksi Pengeluaran
0             18            1                     0
1              0            0                     0
2              2            1                     1
3              0            1                     1
4              0            0                     0
..           ...          ...                   ...
604            0            1                     1
605            0            0                     0
606            0            0                     0
607            0            0                     0
608            0            0                     0

[609 rows x 6 columns]
Hasil prediksi disimpan dalam file: hasil_prediksi_pengeluaran.xlsx
```

Figure 6. Calculation Results and Predictions

Figure 5 is the result of calculating the prior probability for each class. From the calculation results, the prior probability of Not Available is 0.9934 and the prior probability of Available is 0.0066. Figure 6 is the result of calculating and predicting part expenditure. If the predicted part expenditure is "1" then the item is recommended for stock, if the predicted part expenditure is "0" then the item is not recommended for stock.

3.4 Analysis and Recommendations

The prediction model identified 50 items that were consistently out of stock and required immediate replenishment. These items were prioritized based on their sales volumes and operational importance. The top recommendations included: (i) High-demand items like ignition cables and thermocouples; and (ii) Critical components such as heating elements and door gaskets. By aligning predictions with actual stock statuses, the model demonstrated high accuracy in identifying items requiring stock adjustments [11], [12].

Table 3. The final result of the calculation uses the Naive Bayes method

No.	Code	Item Products	Total Sales	Status
1	0K8238	FINGERLINK;OPEN CONNECTION;DIST.120MM	80	Stock is recommended
2	020094	Thermocople; M9 x1 L=120 cm	52	Stock is recommended
3	20094	THERMOCOUPLE;M9X1 L=120CM	52	Stock is recommended
4	0C3162	Interrupted Thermocouple; ORKLY 16400-35	42	Stock is recommended
5	0C9240	PIEZOELECTRIC IGNITER REPL. 2605, 59251	42	Stock is recommended
6	6879	Contactor; 220-240 V 50 Hz / 240-264 V 6	32	Stock is recommended
7	0C0202	Thermocouple; M8 x 1 L= 750 mm (058332)	29	Stock is recommended
⋮	⋮	⋮	⋮	⋮
49	49620	PRESSURE SWITCH; CALIB=250/130-390	13	Stock is recommended
50	7104	CONTACTOR; 220-240V 50HZ/240-264V 60HZ	12	Stock is recommended

Table 3 is the result of calculating and predicting part expenditure, there are the top 50 items displayed, the data is sorted based on the highest sales in the period May 2023 – May 2024. Almost all Electrolux brand parts are not available in stock, this is very disturbing in fulfilling customer orders, after carrying out analysis and calculations using the Naive Bayes algorithm the author succeeded in making a list of parts that need to be stocked, this list of parts will be very helpful in stock management in the future. Some products have higher sales rates than others, indicating the need to prioritize stock of the most popular items and reduce stock of less popular items.

The research highlights the importance of integrating machine learning algorithms like Naive Bayes into inventory management: (i) Efficiency, automating predictions reduced manual workload and increased decision-making speed; (ii) Accuracy, the model effectively identified stock deficiencies based on historical sales patterns; (iii) Scalability, the method can be adapted to larger datasets or different product categories. However, the analysis revealed certain limitations if: (i) Data Imbalance, the overwhelming majority of items were categorized as "Unavailable," which could skew results; and (ii) Simplified Assumptions, the Naive Bayes algorithm assumes attribute independence, which may not always hold in real-world scenarios [13], [14], [15].

In this study, the dataset exhibits a significant class imbalance, with the 'Unavailable' category dominating over the 'Available' category. This imbalance may skew the model's predictions towards the majority class, reducing its reliability. To mitigate this issue, we propose the application of resampling techniques. Specifically: (i) Oversampling, the Synthetic Minority Over-sampling Technique (SMOTE) can be applied to generate synthetic instances of the minority class; and (ii) Undersampling, randomly removing instances from the majority class to balance the dataset. Additionally, adjusting the decision threshold for classification may improve prediction reliability by giving more weight to the minority class. Future work can further explore these techniques to enhance the model's robustness.

The Naive Bayes classifier assumes that all features are independent, which may not hold in real-world datasets. This limitation can lead to suboptimal performance, particularly when features exhibit strong correlations. To address this, future studies could consider the following strategies: (i) Feature Engineering, creating new features that capture interactions between existing ones; and (ii) Algorithmic Alternatives, exploring algorithms that relax the independence assumption, such as Logistic Regression or Decision Trees. These strategies can help improve model performance and provide more accurate predictions.

To improve stock management and model performance: (i) Incorporate additional attributes, such as supplier lead times and demand seasonality; (ii) Address class imbalance by refining the classification thresholds or exploring alternative algorithms: and (iii) Continuously update the dataset to reflect real-time changes in sales trends and stock levels. This systematic application of the Naive Bayes method offers actionable insights for optimizing stock procurement, enhancing operational efficiency, and minimizing stockouts in inventory management systems [3], [16], [17].

4 Conclusion

The findings of this study underline significant insights for inventory management and decision-making in stock procurement at PT. Rotaryana Engineering: (i) Stock Availability Issues: The analysis revealed that almost all spare parts under the Electrolux brand were out of stock, significantly hindering the fulfillment of customer orders. By applying the Naive Bayes algorithm, the researcher successfully developed a prioritized list of parts requiring replenishment. This list serves as a practical tool for improving future stock management practices; (ii) Demand-Based Stock Prioritization: Certain products exhibited higher sales volumes compared to others, indicating the need to prioritize stocking popular items. Conversely, less popular items should be deprioritized to optimize inventory costs and reduce unnecessary overstock; and (iii) Limitations of Historical Data Dependence: While the sales analysis provides valuable insights into customer preferences and demand trends, it is inherently reliant on historical data. This dependence may overlook rapid market shifts and external factors influencing purchasing behavior, limiting its predictive capacity for unforeseen circumstances.

Overall, the implementation of a Naive Bayes-based prediction model offers a robust framework for addressing stock deficiencies and enhancing operational efficiency. However, integrating real-time data and external market indicators could further strengthen the model’s adaptability and relevance in dynamic market conditions. This study provides practical insights for industries engaged in inventory management, particularly in sectors such as manufacturing, retail, and services. By applying the Naive Bayes prediction model, companies can better anticipate stock requirements, minimize stockouts, and reduce excess inventory, thereby improving operational efficiency. The proposed model can be scaled and adapted for use in other industries facing similar inventory challenges, such as healthcare (medical equipment supplies) and automotive (spare parts inventory). Moreover, the integration of machine learning in decision-making processes represents a significant step towards smarter supply chain management, aligning with ongoing digital transformation trends in the industry.

5 Suggestion

From the research that has been carried out, suggestions for further research are: (i) Using a larger list of historical data parts for sales transactions and requests. This helps in optimizing stock levels by minimizing the risk of overstock and understock; and (ii) Conduct regular evaluations of stock management performance to identify the successes and weaknesses of the strategies implemented, as well as to adjust strategies based on changing market conditions.

6 Acknowledgments

The author would like to express sincere gratitude to PT Rotaryana Engineering for their invaluable support and contribution to this research. As a leading company specializing in providing spare parts and maintenance services for kitchen and laundry equipment in hotels, restaurants, and factories, PT Rotaryana Engineering graciously provided the primary data essential for this study.

BIBLIOGRAPHY

[1] PT Rotaryana Engineering, ‘Home - Rotaryana Engineering’. Accessed: Dec. 15, 2024. [Online]. Available: https://rotaryana-engineering.com/

[2] Jiawei Han, Jian Pei, and Hanghang Tong, *Data Mining: Concepts and Techniques*. United State: Elsevier Inc, 2023. Accessed: Dec. 15, 2024. [Online]. Available: https://books.google.co.id/books?hl=id&lr=&id=NR1oEAAAQBAJ&oi=fnd&pg=PP1&dq=Han,+J.,+Pei,+J.,+%26+Kamber,+M.+(2011).+Data+Mining:+Concepts+and+Techniques.+3rd+Edition.+Morgan+Kaufmann+Publishers&ots=_N4GOFsisX&sig=gfsomYdDyPjahFRFypyneupLHfs&redir_esc=y#v=onepage&q&f=false

[3] I. H. Witten Eibe Frank Mark A Hall ELSEVIER AMSTERDAM, L. New York, S. Diego, and M. Kaufmann, *Data Mining Practical Machine Learning Tools and Techniques*, 3rd ed. Elsevier.

[4] D. A. Hamidah, R. Salkiawati, and R. Sari, ‘Analisis Sentimen Ulasan Customer Kopi TMLST Menggunakan Algoritma Naïve Bayes’, *Journal of Students‘ Research in Computer Science*, vol. 5, no. 1, pp. 27–40, May 2024, doi: 10.31599/mrm89y71.

[5] I. Alkarim, ‘Sistem Pengendalian Ketersediaan Onderdil Traktor Dengan Teknik Buffer Stock Dan Reorder Point (Rop) Di Bangunan Agen Tanjung Karang (Studi Kasus CV. Karya Hidup Sentosa Lampung)’, *Teknologipintar.org*, vol. 2, no. 9, pp. 2022–2023, 2022.

[6] S. D. C. Sinamon, ‘Rancangan Aplikasi Stock Sparepart Alat Berat Menggunakan Metode Fifo Studi Kasus CV Bumi Waras Bandar Lampung’, *Teknologipintar.org*, vol. 2, no. 4, pp. 2022–2023, 2022.

[7] S. P. Sarmah and U. C. Moharana, ‘Spare Parts Inventory Management Literature and Direction Towards the Use of Data Mining Technique: A Review’, *https://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/978-1-5225-3232-3.ch028*, pp. 534–558, Jan. 2018, doi: 10.4018/978-1-5225-3232-3.CH028.

[8] P. W. Putra, D. Irawan, A. T. Martadinata, and A. Heryati, ‘Sistem Prediksi Persediaan Barang di Mini Market Mars Menggunakan Algoritma Naive Bayes Classifier dengan Pemrograman Phyton’, *Davit Irawan*, vol. 7, no. 2, p. 148, 2022.

[9] David L. Olson and Dursun Delen, *Advanced Data Mining Techniques*. Verlag Berlin: Springer, 2008. Accessed: Dec. 15, 2024. [Online]. Available: https://books.google.co.id/books?hl=id&lr=&id=2vb-LZEn8uUC&oi=fnd&pg=PA8&dq=Olson,+D.+L.,+%26+Delen,+D.+(2019).+Advanced+Data+Mining+Techniques.+Springer,+New+York&ots=zY4_Z33PmO&sig=jhL1ysvy28wAJ76uDysKocz0rr0&redir_esc=y#v=onepage&q&f=false

[10] K. P. Aronis, I. Magou, R. Dekker, and G. Tagaras, ‘Inventory control of spare parts using a Bayesian approach: A case study’, *Eur J Oper Res*, vol. 154, no. 3, pp. 730–739, May 2004, doi: 10.1016/s0377-2217(02)00837-8.

[11] R. A. P. Jayendran, P. Keikhosrokiani, and S. L. Chui, ‘Inventory Classification and Management System Using Machine Learning and Analytical Dashboard: A Case Study of a Manufacturing Industry’, *https://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/979-8-3693-1210-0.ch012*, pp. 299–318, Jan. 2024, doi: 10.4018/979-8-3693-1210-0.CH012.

[12] P. KumarBala, ‘An inventory replenishment model under purchase dependency in retail sale.’, *International Journal of Computers and Applications*, vol. 37, no. 10, pp. 43–48, Jan. 2012, doi: 10.5120/4648-6722.

[13] P. Suwignjo, L. Panjaitan, A. Raecky Baihaqy, and A. Rusdiansyah, ‘Predictive Analytics to Improve Inventory Performance: A Case Study of an FMCG Company’, *OPERATIONS AND SUPPLY CHAIN MANAGEMENT*, vol. 16, no. 2, pp. 293–310, 2023.

[14] D. Kurniawan, M. Sahata Sipayung, R. Ismayanti, M. Rivani Ibrahim, Y. Bintan, and S. Aulia Miranda, ‘Optimalisasi Strategi Pemenuhan Persediaan Stok Barang Menggunakan Algoritma Frequent Pattern Growth’, *METIK JURNAL*, vol. 6, no. 2, pp. 104–114, Dec. 2022, doi: 10.47002/metik.v6i2.387.

[15] A. Azis and Sutisna, ‘Penerapan Data Mining untuk Menentukan Ketersediaan Stok Barang Berdasarkan Permintaan Konsumen di PT Indonesia Thai Summit Plastech Menggunakan K-Means Clustering’, *Jurnal Indonesia : Manajemen Informatika dan Komunikasi*, vol. 5, no. 3, pp. 3099–3106, Sep. 2024, doi: 10.35870/JIMIK.V5I3.982.

[16] P. Kumar Bala, ‘An Inventory Replenishment Model under Purchase Dependency in Retail Sale’, 2012.

[17] R. Srikant and R. Agrawal, ‘Mining Sequential Patterns: Generalizations and Performance Improvements’, San Jose.