INVENTORY PREDICTION SYSTEM USING THE LEAST SQUARE METHOD AT THE SERBA SERBI ONLINE SHOP TANJUNGPINANG

Hendi Setiawan 1*, Nanny Raras Setyoningrum 2, Danandjaya Saputra 3, Merry 4

1,2,4 Information Systems, Sekolah Tinggi Teknologi Indonesia Tanjung Pinang, Indonesia
3 Informatics Engineering, Sekolah Tinggi Teknologi Indonesia Tanjung Pinang, Indonesia
* Corresponding author: hendi@sttindonesia.ac.id

ARTICLE INFO
Received: 15 Dec 2023
Accepted: 30 Dec 2023

ABSTRACT
In this era of Industry 4.0, technology is advancing rapidly, and various applications are emerging to help traders optimize their product procurement. One such application is the forecasting or prediction system. Serba Serbi Online Shop Tanjungpinang is a shop that requires this system to manage its inventory successfully. As the shop offers a wide variety of goods, managing inventory has become challenging. The capital spent on inventory has increased, and unsold goods are piling up in the warehouse. To overcome this issue, the shop needs to calculate the amount of inventory required to reduce the stock build-up in the warehouse. The prediction method used for this purpose is the least square method. This method uses periodic or time-series data, which requires past sales data to predict future sales. Historical sales data is necessary to develop an inventory prediction system. The sales data for a single item from July 2022 to May 2023 was analyzed using the Least Square method. Based on this analysis, the predicted sales for June are 144.64, rounded to 145. The accuracy of the prediction results was tested, and the best accuracy was obtained from the prediction using 3 months of data, with an error rate of 5.47%. The prediction with the highest error rate was obtained from using 9 months of data, with an error rate of 10.65%. On average, the prediction accuracy is 0.02154%, which is considered very good.

Keywords: system prediction, inventory, online shop, least square method

1. Introduction

Inventory is one component that has an important role in a company. Every company usually has inventory to be able to carry out its company activities. Inventory can also be said to be a stock of an item or resource used in a company organization [1].

Information technology has brought many changes to organizations and business processes in the corporate world. Another benefit of Information Technology (IT) is that it provides support for administrative services and helps predict inventory. Information technology is a means of improving company and business performance [2].

Prediction is a process of systematically estimating something that is most likely to happen in the future based on past and present information that is available so that the error (the difference between something that happens and the predicted result) can be minimized [3].

One of the companies that requires this prediction system is the Serba Serbi Online Shop Tanjungpinang. This shop is an online shop located in Tanjungpinang, Riau Islands. The products sold by the Serba Serbi Online Shop Tanjungpinang are various kinds such as electric pans, digital scales, and power banks, and sell household appliances, cellphone accessories, and many more.

The Serba Serbi Online Shop Tanjungpinang has faced challenges in providing goods due to the need for a larger capital investment as the shop expands its range of products. Additionally, some products may not sell well and end up taking up space in the warehouse. To address this, the least squares method was used to predict sales forecasting. This method uses past sales data to forecast future sales and improve the accuracy of the forecasting process [4]. The outcome of this research is a goods inventory forecasting system that helps prioritize items and ensure that capital is being utilized efficiently.

2. Materials and Methods
a. Material
1) Information System
A system is a collection of components that are closely related and work together to achieve certain goals [5]. Information can be understood as the result of data processing which is transformed into a form that is more useful and meaningful for the recipient about actual events used for decision-making [6]. An information system is a system within an organization that connects daily data processing needs that support operational management functions with the strategic activities of an organization to provide the information needed to make decisions [7]. Information systems include several components
(humans, computers, information technology, and work procedures), something is processed (data becomes information), and is intended to achieve a goal [8].

2) Data

Data is a collection of information or facts made up of words, sentences, symbols, numbers, and so on. The data here was obtained through a search process and precise observations based on certain sources. Another definition of data is a collection of basic information or descriptions originating from objects or events [9].

3) Supply

Inventory refers to the merchandise that a company stores for future sale, use in production or for other purposes. It may consist of goods that are crucial for the company’s operations and are available for use in the production process or for sale. In the case of trading companies, inventory may include goods purchased for resale [10].

4) Prediction

Forecasting is a method of computational analysis that uses both qualitative and quantitative approaches to predict future events based on past data. The art and science of predicting future events is known as forecasting. Historical data is used to create a mathematical model that projects into the future [11]. Forecasting is an important tool for effective and efficient planning. Quantitative forecasting can be used if three conditions are met: first, information about the past is available; second, this information can be quantified in the form of numerical data; and third, it can be assumed that some aspects of past patterns will continue into the future [12].

5) Least Square Method

The least square method is a technique used to develop a periodic series or time series that has occurred in the past. This series is then used to forecast future events. This method is particularly useful for making large predictions for a specific variable with a certain time series. There are two different cases when using the least squares method to create periodic series or time series. The first case is for odd data, and the second case is for even data [13]. Here are the steps involved in performing calculations using the least squares method [14]:

1) Find the number of data (n)
2) Finding the value of parameter X, in determining parameter X, it can be divided into two cases according to the amount of data. If the amount of data is even then the value of variable X is -3, -1, 1, 3 and so on as many as the number of data with a difference of 2 for each data, whereas if the value of variable X has a difference of 1 then the amount of data is odd and the value of variable X namely -1, 0, 1, and so on as many times as there is data.
3) Find the sum of the values of $X^2$ and $XY$.
4) Find the coefficient value a can be seen in formula (1) below:
   \[ a = \frac{\sum y}{n} \]  
   \[ (1) \]
5) Find the coefficient value b can be seen in formula (2) below:
   \[ b = \frac{\sum xy}{\sum x^2} \]  
   \[ (2) \]
6) Calculating the Y forecast can be seen in formula (3) below:
   \[ Y = a + b(x) \]  
   \[ (3) \]

Note:
- $\sum y$ = the total of the actual number of Y variables
- n = the total of all actual data
- $\sum xy$ = the total of the sum of the products between variables x and y
- $\sum x^2$ = the total of the sum of the squares of the variables x
- a = trend variable at the beginning of the year
- b = average trend variable each year
- Y = value of the dependent variable
- X = value of the independent variable

6) Prediction Accuracy

MAPE (Mean Absolute Percentage Error) is a forecasting error testing method. It is calculated by dividing the absolute error for each period by the real observation value in that period and averaging the percentage of absolute error [15]. The MAPE equation is shown by formula (4) as follows:

\[ MAPE = \frac{\sum_{i=1}^{n} |Y_i - Y'_i|}{n} \times 100 \]  
\[ (4) \]

Note:
- Y = actual value
- Y’ = predicted value
- n = amount of data
Table 1. Interpretation MAPE value

<table>
<thead>
<tr>
<th>Range</th>
<th>Nilai MAPE</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&lt;10%</td>
<td>Very good</td>
</tr>
<tr>
<td></td>
<td>10-20%</td>
<td>Good</td>
</tr>
<tr>
<td></td>
<td>20-50%</td>
<td>Enough</td>
</tr>
<tr>
<td></td>
<td>&gt;50%</td>
<td>Not good</td>
</tr>
</tbody>
</table>

b. Methods

1) Method of Collecting Data

Selecting an appropriate method of data collection depends on the intended purpose and the source of the information. The success of a research project can be determined by the presence of objective and relevant data related to the chosen topic. Some examples of data collection methods are the techniques or tools utilized to gather primary or secondary data [16].

1. Literature Study
   In this research method, data is collected from various sources such as reports, notes, letters, books, etc. which are then analyzed for research purposes.

2. Observation
   In this research method, data is collected from various sources such as reports, notes, letters, books, etc. which are then analyzed for research purposes.

3. Interview
   This method involves direct communication between researchers and respondents. It can be achieved through face-to-face interaction, telephone, or video calls with structured or unstructured approaches.

2) Software Development Methods

The information system was built using the Waterfall (Classic Life Cycle) model, a software development methodology that follows a systematic and continuous approach. The development process starts from the top level of the system and moves on to the requirements, design, testing, and maintenance stages in a structured, dynamic and sequential manner. This method offers several advantages including better organization, flexibility, and a clear path for software development [17].

![Waterfall Model](image)

Figure 1. Model Waterfall

The research utilized the following system development method: requirements, design, testing, and maintenance [18]:

1. Requirements
   During the development phase, it's important for software developers to communicate with users and understand their expectations and limitations of the software. This can be achieved through interviews, discussions, or direct surveys. The information obtained from users is then analyzed to gather necessary data for improving user experience.

2. Design
   The developer creates a system design to determine hardware and system requirements and overall system architecture.

3. Coding
   The software is developed in stages, with each stage involving the integration of small programs called units. Each unit undergoes functionality testing before integration.

4. Testing
   System verification involves testing the code of a specific module. System testing is done to observe how the system behaves when all modules are integrated. Acceptance testing, on the other hand, is carried out either by or on behalf of the customer to ensure that all customer requirements are met.
5. Maintenance
   This is the final step of the waterfall method. The completed software is implemented and maintained, which includes fixing errors not found in previous steps.

3. Results
a. System Design
   System design is a crucial stage in the system development cycle that aims to provide a clear description of the system that needs to be created. It involves the integration of both hardware and software components to form a complete system that can meet the expectations of its users [19]. Generally, system design is divided into two parts, system design and detailed design. The former is concerned with the overall system architecture and setting the standards to be used during implementation, while the latter deals with designing each component to comply with the system architecture and standards [20].
b. Use Case Design
   Use case diagrams are a visual tool used to depict the expected functionality of a system. The emphasis is on "what" the system does, rather than "how" it does it. Each case represents the relationship between actors and the system [21]. Use cases are specific tasks that the system performs, such as logging into an account or creating a shopping list. An actor can be a human or machine entity that interacts with the system to perform a specific task.

![Use Case Diagram](image)

Figure 2. Use Case

c. Activity Diagrams Design
   Activity diagram is a UML diagram that models the flow from one activity to another within a system.

![Activity Diagram](image)

Figure 3. Activity Diagram
The explanation of the image above is as follows:
1) Superadmin/Admin selects the prediction menu.
2) The system displays the prediction page.
3) Superadmin/Admin selects goods and start-end date as data parameters and clicks the sales check button.
4) The system displays sales data for the selected items.
5) Superadmin/Admin clicks the sort data button.
6) The system displays the results of sales data calculations.
7) Superadmin/Admin clicks the prediction button.
8) The system carries out the prediction process and displays the results in the prediction results table.
9) Superadmin/Admin clicks the submit button to display the final prediction results.
10) The system displays the final prediction results.

**d. Problem Solving Prediction**

To create a system that can predict inventory needs, we require access to historical sales data. For our research, we will be using sales data for power banks sold by the Serba Serbi Online Shop Tanjungpinang between July 2022 and May 2023. As an example, we will be using the least square method to calculate sales data for one of the items sold during this time period.

<table>
<thead>
<tr>
<th>Month</th>
<th>Sales (Y)</th>
<th>X</th>
<th>XY</th>
<th>X²</th>
</tr>
</thead>
<tbody>
<tr>
<td>July</td>
<td>90</td>
<td>-5</td>
<td>-450</td>
<td>25</td>
</tr>
<tr>
<td>August</td>
<td>101</td>
<td>-4</td>
<td>-404</td>
<td>16</td>
</tr>
<tr>
<td>September</td>
<td>85</td>
<td>-3</td>
<td>-255</td>
<td>9</td>
</tr>
<tr>
<td>October</td>
<td>96</td>
<td>-2</td>
<td>-192</td>
<td>4</td>
</tr>
<tr>
<td>November</td>
<td>105</td>
<td>-1</td>
<td>-105</td>
<td>1</td>
</tr>
<tr>
<td>December</td>
<td>143</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>January</td>
<td>121</td>
<td>1</td>
<td>121</td>
<td>1</td>
</tr>
<tr>
<td>February</td>
<td>152</td>
<td>2</td>
<td>304</td>
<td>4</td>
</tr>
<tr>
<td>March</td>
<td>100</td>
<td>3</td>
<td>300</td>
<td>9</td>
</tr>
<tr>
<td>April</td>
<td>148</td>
<td>4</td>
<td>592</td>
<td>16</td>
</tr>
<tr>
<td>May</td>
<td>126</td>
<td>5</td>
<td>630</td>
<td>25</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>1,267</strong></td>
<td>0</td>
<td><strong>541</strong></td>
<td><strong>110</strong></td>
</tr>
</tbody>
</table>

**Table 2. Sample Data Prediksi Powerbank**

**Method of Completion:**
1) The first step that must be taken is to find the trend equation in the data, find the values of $a$ and $b$ first using equations (1) and (2).

\[ a = \frac{1267}{11} = 115.18 \]

\[ b = \frac{541}{110} = 4.91 \]

2) Once these two values are found, enter the values into the trend equation in equation (3)

\[ Y = 115.18 + 4.91x \]

3) Finally, calculate forecasting based on the data to be searched. For example, the predictions you want to know are for May and June. So, the value of $x$ from June is 6. Then enter this value into the trend equation previously formed:

\[ Y = 115.18 + 4.91 \times 6 = 144.64 \]

Based on the output produced, it can be concluded that the prediction for power bank sales for June is 144.64 with a rounding 145.

4) Next, to calculate the percentage error resulting from the prediction, find the difference between actual and predicted data and use equation (4) (see Table 3).

**Table 3. MAPE**

<table>
<thead>
<tr>
<th>Tahun</th>
<th>Bulan</th>
<th>Data Aktual (Yt)</th>
<th>Data Prediksi (Ft)</th>
<th>MAPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>2022</td>
<td>July</td>
<td>90</td>
<td>90</td>
<td>1.11</td>
</tr>
<tr>
<td>2022</td>
<td>August</td>
<td>101</td>
<td>95</td>
<td>5.94</td>
</tr>
<tr>
<td>2022</td>
<td>September</td>
<td>85</td>
<td>95</td>
<td>11.76</td>
</tr>
<tr>
<td>2022</td>
<td>October</td>
<td>96</td>
<td>92</td>
<td>4.16</td>
</tr>
<tr>
<td>2022</td>
<td>November</td>
<td>105</td>
<td>110</td>
<td>4.76</td>
</tr>
<tr>
<td>2022</td>
<td>December</td>
<td>143</td>
<td>115</td>
<td>19.58</td>
</tr>
<tr>
<td>2023</td>
<td>January</td>
<td>121</td>
<td>120</td>
<td>0.82</td>
</tr>
<tr>
<td>2023</td>
<td>February</td>
<td>152</td>
<td>125</td>
<td>17.76</td>
</tr>
</tbody>
</table>
MAPE is 10.52% which can be concluded that the Least Square method is included in the prediction model category which has very good forecasting capabilities.

e. Interface Implementation

According to the system design, the inventory prediction application interface implementation results of Serba Serbi Online Shop Tanjungpinang are as follows:

<table>
<thead>
<tr>
<th>Month</th>
<th>2023</th>
<th>2023</th>
<th>2023</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>March</td>
<td>April</td>
<td>May</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>148</td>
<td>126</td>
</tr>
<tr>
<td></td>
<td>130</td>
<td>135</td>
<td>140</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>8.78</td>
<td>11.11</td>
</tr>
</tbody>
</table>

\[
\text{MAPE} = 10.52\%
\]

<table>
<thead>
<tr>
<th>Figure 4. Login Menu</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 5. Menu Utama</td>
</tr>
<tr>
<td>Figure 6. Menu Kategori</td>
</tr>
<tr>
<td>Figure 7. Menu Barang</td>
</tr>
<tr>
<td>Figure 8. Menu Data Penjualan</td>
</tr>
</tbody>
</table>

Figure 9. Menu Prediksi
4. Discussion

Numerous studies have been conducted to solve the problem of predicting inventory, one of which is titled "Implementation of the Naive Bayes Algorithm for Predicting PT Goods Inventory: Dilmoni Citra Indonesian Furniture". The main objective of this research is to anticipate the demand for stock so that production activities can continue smoothly, and raw materials remain available. The stock prediction analysis is based on the sales data of the company for a specific period, which is then analyzed using the Naive Bayes algorithm [22].

The research paper titled "Creating a Goods Inventory Prediction System at Nabila Stores Using the Weighted Moving Average and Reorder Point Method" describes a method for predicting inventory levels and reorder points for Nabila Store. The system uses the weighted moving average method to forecast inventory levels and reorder points to determine when to order inventory. The study evaluates the accuracy of the system's predictions using mean absolute deviation (MAD), mean square error (MSE), and mean absolute percentage error (MAPE). The average values for MAD, MSE, and MAPE for 10 inventory prediction data at Nabila Store are 7.44, 77.99, and 31.90, respectively [23].

The following text describes a research study titled "Prediction of Goods Inventory Using the Weighted Moving Average, Exponential Smoothing and Simple Moving Average Methods", which aims to predict inventory levels to ensure that there is always enough stock available. The study focuses on predicting the supply of brown envelopes, which are in high demand at the F2 photocopy and office stationery shop where the research was conducted. The author uses three methods to estimate inventory levels: Weighted Moving Average, Exponential Smoothing, and Simple Moving Average. The data is processed using the Python programming language with Jupyter Notebook. The results show that the best method for the F2 photocopy and office stationery shop is Exponential Smoothing with $\alpha=0.1$. This method has the smallest RMSE value, which is 13.616, and predicts that there will be 57 brown envelopes available for the next period [24].

Based on the findings of previous studies, the weighted moving average method has been found to have limitations in accurately predicting inventory usage. On the other hand, research conducted using the Least Square Method suggests that it is possible to form a periodic series (time series) based on past data to guide future forecasting.

5. Conclusions

After conducting research, several conclusions can be drawn. A prediction system has been developed that can predict various types of goods, enabling the Serba Serbi Online Shop Tanjungpinang to prioritize which goods should be procured first. With this prediction system, it is hoped that the Tanjungpinang Online Shop can reduce the procurement of goods that are not selling well and prevent the accumulation of excess inventory. The prediction system was developed according to the needs of Serba Serbi Online Shop Tanjungpinang and utilizes sales data for a single item from July 2022 to May 2023. By implementing the Least Square method, it was determined that the sales prediction for June is 144.64, rounded up to 145. The prediction system has been tested for accuracy, and the highest accuracy was achieved with 3 months of data, resulting in a prediction accuracy of 5.47%. The prediction accuracy decreased with the increase in the amount of data used, with the highest error rate being 10.65% for 9 months of data. The overall average prediction accuracy is 0.02154%, which falls under the "very good" category.

References


