The 2nd International Conference on Multidisciplinary Engineering
(2nd ICOMDEN 2024)

The 2nd International Conference on Multidisciplinary Engineering: 00060 (2024) https://doi.org/10.29103/icomden.v2.xxxx
Volume 2, 2024
eISSN: xxxx-xxxx
Research Original Article

Implementation of Data Mining for Raw Material Stock Prediction in Clothing Production Using the C4.5 Algorithm Method

Nur Ismiza[1] **Nurdin**[2] **Lidya Rosnita**[3]
[1]Department of Informatics, Malikussaleh University, Bukit Indah, Lhokseumawe, 24355, Indonesia
Corresponding Author: nur.200170072@mhs.unimal.ac.id | **Phone : +6282376590573**

Abstract
Al-Fatih Convection is a business engaged in the textile industry, located in Baktiya, North Aceh Regency. This company produces various uniforms for schools and workwear. Raw material stock management is a crucial aspect that affects the smoothness of the production process. Currently, the purchase of raw material stock still relies on estimation methods, often leading to excessive or insufficient stock. Therefore, a raw material stock prediction system is needed to optimize stock management.This research aims to implement the C4.5 algorithm to predict raw material stock for clothing production. The method is chosen for its ability to build a predictive model based on attributes such as material type, price, availability, and demand. Using data mining, this study generates a decision tree that helps Al-Fatih Convection prioritize which raw materials should be purchased. The results from the implementation of the C4.5 algorithm show an accuracy rate of 93%, which is expected to help reduce excessive or insufficient stock and improve operational efficiency at Al-Fatih Convection.

Keywords: C4.5 Algorithm, Prediction, Stock, Convection.

INTRODUCTION

The development of science and information technology in the current era of globalization has been so significant and has had a rapid impact, creating major changes in all aspects of life. This is due to computer technology, which provides many conveniences and benefits in various fields, one of which is the industrial sector [1].
In the clothing convection industry, raw material stock management is a crucial aspect that affects the smooth production process. Maintaining adequate and accurate raw material stock is a challenge for convection companies, considering various factors that can influence demand and supply. So far, rawmaterial stock at Al-Fatih Convection still uses an estimation method, relying on details of the materials most in demand by consumers and reviewing old stock, which often leads to excessive raw material purchases. Over-purchasing raw materials results in stockpiling, causing the financial flow to not run well and smoothly.
Given this issue, a system is greatly needed to assist Al-Fatih Convection in predicting future stock requirements for production materials. This would help prevent both overstocking and shortages of raw materials. The aim of this research is to develop a system that can provide recommendations to Al-Fatih Convection regarding future raw material purchases. With the developed system, it will be possible to identify which raw materials should be prioritized for the upcoming purchases.
Data mining is the process of extracting important or interesting information from data stored in a database, resulting in highly valuable information [2]. In data mining, there are several methods, including prediction, classification, clustering, and association [3]. Prediction is a process used to forecast something that may happen in the future based on past and present data [4].
There are several studies referenced in this research, including the one conducted by Aziziah, Aldila Nur, Ade Irma Purnamasari, and Irfan Ali in 2024. This study used six attributes: material name, unit, initial stock, material outflow, final stock, and remarks. The results of the study showed that the C4.5 algorithm achieved an accuracy rate of 97.60%, providing evidence that this algorithm can be used for predicting beverage stock [5].
Another study was conducted by Pritalia in 2018. The C4.5 algorithm was used to analyze the timing of restocking depleting items by classifying which items needed to be replenished and which did not, ensuring that product availability remained stable and maintained. The results of the analysis using the C4.5 algorithm showed an accuracy rate of 98.9% in determining the timing for ensuring product availability [6].

Another study was conducted by Elisa in 2022. The method used in this research was the C4.5 algorithm. The results of the study produced a prediction decision tree, with the highest gain values being price, quality, and warranty, which were identified as the key factors influencing product availability [7].
Based on the description above, the author is interested in conducting further research on "The Implementation of Data Mining for Predicting Raw Material Stock for Clothing Production Using the C4.5 Algorithm Method."

LITERATURE REVIEW

Data Mining

Data mining can be described as the process of extracting information from a large amount of available data. The information obtained from the data mining process should be new, easily understandable, and beneficial [8]. In data mining, data is stored electronically and processed automatically by computers using specific techniques and calculations. The data mining process employs statistical methods, mathematics, and also utilizes artificial intelligence technology [9]. These complex techniques will then identify and extract relevant information from large datasets. [10]. Data mining can be divided into several groups based on the tasks that can be performed, namely: Description, Estimation, Prediction, Classification, Clustering, and Association.

Decision Tree

Decision tree is a powerful and popular method for classification and prediction. The decision tree method transforms large amounts of data into a decision tree that represents rules, making it easy to understand [11]
A decision tree is a flowchart-like structure that has a tree format, where each internal node represents a test attribute, each branch represents the outcome of that test, and each leaf node represents a class or class distribution [12].
The strategies that can be used for constructing a decision tree with a decision tree algorithm are as follows:

1. The tree starts as a single node (the root) that represents all the data.
2. After the root node is formed, the data at the root node will be measured using information gain to determine which attribute will be used as the splitting attribute.
3. A branch is created from the selected attribute, and the data is distributed to the respective branches.
4. This algorithm will continue to use the same process (recursively) to form a decision tree. Once an attribute has been selected as a splitting node/branch, it will not be included again in the calculation of information gain.
5. The recursive splitting process will stop when one of the following conditions is met:
   1. All data from the child branches belong to the same class.
   2. All attributes have been used, but there are still remaining data in different classes. In this case, the data representing the majority class will be taken as the class label for the leaf node.
   3. There is no data left in the new child branches. In this case, the leaf node will be chosen from the previous branch, and the data representing the majority class will be used as the class label.

C4.5 algorithm

The C4.5 algorithm is used to generate a decision tree and was developed by Ross Quinlan. The basic idea of this algorithm is to build a decision tree based on the selection of attributes with the highest priority, or in other words, those with the highest gain values, using the entropy of the attributes as the axis for classification attributes [13]. The C4.5 algorithm is one of the algorithms used for classification or predictive segmentation [14.

The C4.5 algorithm is considered highly effective in assisting with data classification because the characteristics of the data to be classified are clearly defined, both in the depiction of the decision tree (rules) and in the if-then rules. This clarity makes it easier for users to observe information related to the data [15].
There are four steps in the process of creating a decision tree using the C4.5 algorithm:

a. Select an attribute as the root, based on the highest gain value among the available attributes.
b. Create branches for each value, meaning that branches are created according to the number of values of the highest gain variable.
c. Split each case into branches based on the highest gain calculations. This calculation is done after the initial highest gain calculation, and then the process of calculating the highest gain is repeated without including the values of the initial gain variable.
d. Repeat the process in each branch until all cases in the branches belong to the same class.

To select an attribute as the root, it is based on the highest gain value among the available attributes. To calculate the entropy value, the following formula is used:
$Entropy\ (S) = \sum_{i=1}^{n} -p_i * \log_2 p_i$ .... (1)
Explanation:
S = Set of cases
n = Number of partitions of S
p_i = Proportion of Si relative to S
Meanwhile, to calculate the entropy value, it can be seen in the following Equation 2:
$Gain\ (S, A) = Entropy(s) - \sum_{i=1}^{n} \frac{|S_i|}{|S|} * Entropy(S_i)$ .... (2)
Explanation:

S = Number of Cases (Sampling)
A = Attribute
N = Number of Partitions of S
|S| = Number of Cases in the i-th Partition
|S| = Number of Cases in S

The advantages of the C4.5 algorithm are as follows:
1. It can handle missing attribute values.
2. It can manage continuous attributes.
3. It can prune the decision tree to address overfitting.

Despite its many strengths, the C4.5 algorithm also has limitations, such as:
1. Computational complexity; calculating gain for each attribute in large datasets can be time-consuming.
2. Difficulty in handling very large data sets, making it less effective when applied to extremely large datasets, especially with many attributes.

MATERIALS & METHODS

Research Steps

The steps taken by the author in the process of this final project research can be seen in the image below:

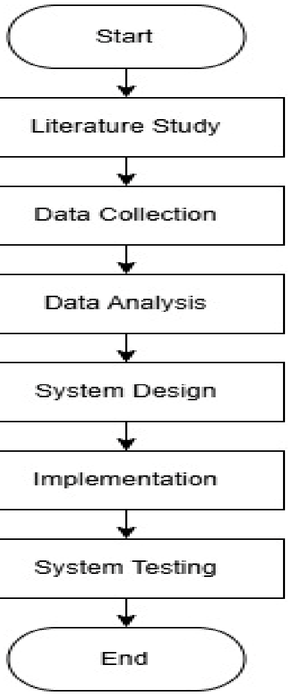

Figure 1 Research Steps

System Schema

The schematic of the raw material stock prediction system for Al-Fatih Convection using the C4.5 algorithm can be seen in the image below:

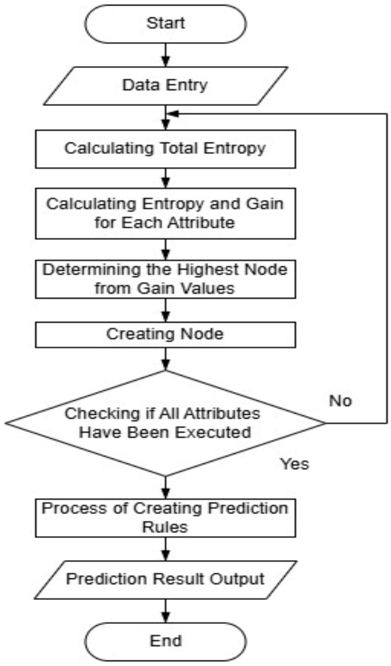

Figure 2 System Schema

Explanation:

1. Start.
2. Enter the production raw material stock data, including material quality data, type of material, inventory, demand, and remarks.
3. Calculate the total entropy using the formula in Equation 1.
4. Calculate the entropy and gain for each attribute using the formula in Equation 2.
5. Determine the highest node based on the gain values.
6. Create a node from the attribute with the highest gain as the root node.
7. Check whether each attribute has been executed; if yes, it will produce a decision tree, and if not, repeat the process for each branch until all branches have the same class.
8. Create the prediction rules.
9. Output the prediction results, which include the decision tree model indicating which stock items are approved and pending.
10. Finish.

RESULTS AND DISCUSSION

The results of this study show that the C4.5 algorithm can be implemented in a prediction system for raw material stock requirements in clothing production, aiming to reduce the risks of stock shortages or surpluses. This system operates by inputting various values or criteria from historical sales data, raw material stock, and orders obtained from Al-Fatih Convection, then performing calculations that help predict future raw material stock needs. Thus, this system has the potential to enhance stock management efficiency and optimize the production process.

Data Transformation

The data transformation process is carried out to modify the attributes according to the format that can be processed in the program. Based on the cleaned data for clothing production materials, data transformation is performed on the price, inventory, and demand attributes. The transformation involves changing the attribute values from numerical to categorical. The transformed data can be seen in Table below:

Table 1. Data Transformation

No	Attribute	*Class*
1	Price	Rp. 18,000 - Rp. 30,000 (Cheap)

		Rp. 31,000 - Rp. 50,000 (Expensive) Rp. 51,000 - Rp. 70,000 (Very Expensive)
2	Inventory	5 - 50 (Low) 51-167 (High)
3	Demand	10 - 40 (Decrease) 4 – 283 (Increase)
4	Result	Demand > Inventory (Approved)) Demand < Inventory (Pending)

The data that has undergone the conversion of attribute values from numerical to categorical can be seen in Table 4.8 below:

Table 2. Normalized Dataset

No.	Type	Price	Inventory	Demand	Remarks
1	Toyobo	Expensive	Low	Increase	Approved
2	Drill	Cheap	Low	Increase	Approved
3	Tissue	Expensive	High	Increase	Pending
4	Adidas	Very Expensive	Low	Increase	Approved
5	Lotto	Cheap	Low	Decrease	Pending
6	Higet	Cheap	High	Decrease	Pending
7	Tissue	Expensive	Low	Increase	Approved
8	PE	Expensive	High	Decrease	Pending
9	Drill	Cheap	Low	Increase	Approved
10	TC	Very Expensive	High	Increase	Approved
11	Cotton	Very Expensive	Low	Decrease	Pending
12	Higet	Cheap	High	Increase	Pending
13	Toyobo	Expensive	High	Increase	Approved
14	PE	Expensive	Low	Increase	Approved
15	Paragon	Very Expensive	High	Decrease	Pending
16	Lotto	Cheap	Low	Decrease	Pending
17	Higet	Cheap	High	Decrease	Pending
18	TC	Very Expensive	Low	Increase	Approved
19	Cotton	Very Expensive	Low	Decrease	Approved
20	Tissue	Expensive	Low	Increase	Approved
...					
341	Higet	Cheap	Low	Increase	Approved
342	Carded	Very Expensive	Low	Decrease	Approved
343	Paragon	Very Expensive	Low	Increase	Approved
344	Drill	Cheap	Low	Increase	Approved
345	Carded	Very Expensive	Low	Decrease	Approved
346	Swessi	Cheap	High	Decrease	Pending
347	PE	Expensive	Low	Increase	Approved
348	Adidas	Very Expensive	Low	Decrease	Approved
349	Carded	Very Expensive	Low	Increase	Approved
350	Higet	Cheap	Low	Decrease	Approved
351	Toyobo	Expensive	Low	Increase	Approved
352	TC	Very Expensive	Low	Decrease	Approved
353	Cotton	Very Expensive	Low	Increase	Approved
354	Higet	Cheap	Low	Decrease	Pending
355	Toyobo	Expensive	Low	Increase	Approved
356	Paragon	Very Expensive	Low	Decrease	Approved

C4.5 Calculation

To create a decision tree using the C4.5 algorithm, the dataset, which includes attributes such as type, price, demand, and inventory, is divided into 80% for training and 20% for testing. Next, the number of cases with the labels "Approved" and "Pending" is calculated, followed by calculating the entropy of all cases divided by class based on the attributes type, price, demand, and inventory. Then, the entropy and gain for each attribute that will serve as the root in the decision tree formation are calculated.

Entropy (Total)

$$\text{Entropy Total} = ((-\frac{\text{total approved}}{\text{total cases}}) \times \log 2\left(\frac{\text{total approved}}{\text{total cases}}\right)$$
$$+(-\frac{\text{total pending}}{\text{total cases}} \times \log 2\left(\frac{\text{total pending}}{\text{total cases}}\right))$$
$$= ((-\frac{213}{285}) \times \log 2\left(\frac{213}{285}\right) + (-\frac{71}{285}) \times \log 2\left(\frac{71}{285}\right))$$
$$= 0.81348$$

Entropy Type

$$\text{Entropy [Drill]} = ((-\frac{15}{22}) \times \log 2\left(\frac{15}{22}\right) + (-\frac{7}{22}) \times \log 2\left(\frac{7}{22}\right)) = 0.90239$$

$$\text{Entropy [Toyobo]} = ((-\frac{15}{21}) \times \log 2\left(\frac{15}{21}\right) + (-\frac{6}{21}) \times \log 2\left(\frac{6}{21}\right)) = 0.86312$$

$$\text{Entropy [Swessi]} = ((-\frac{13}{19}) \times \log 2\left(\frac{13}{19}\right) + (-\frac{6}{19}) \times \log 2\left(\frac{6}{19}\right)) = 0.899774$$

$$\text{Entropy [Tissue]} = ((-\frac{21}{26}) \times \log 2\left(\frac{21}{26}\right) + (-\frac{5}{26}) \times \log 2\left(\frac{5}{26}\right)) = 0.70627$$

$$\text{Entropy [Cotton]} = ((-\frac{14}{21}) \times \log 2\left(\frac{14}{21}\right) + (-\frac{7}{21}) \times \log 2\left(\frac{7}{21}\right)) = 0.9183$$

$$\text{Entropy [PE]} = ((-\frac{15}{26}) \times \log 2\left(\frac{15}{26}\right) + (-\frac{11}{26}) \times \log 2\left(\frac{11}{26}\right)) = 0.98286$$

$$\text{Entropy [TC]} = ((-\frac{18}{23}) \times \log 2\left(\frac{18}{23}\right) + (-\frac{5}{23}) \times \log 2\left(\frac{5}{23}\right)) = 0.75538$$

$$\text{Entropy [Carded]} = ((-\frac{16}{16}) \times \log 2\left(\frac{16}{16}\right) + (-\frac{0}{16}) \times \log 2\left(\frac{0}{16}\right)) = 0$$

$$\text{Entropy [Higet]} = ((-\frac{15}{22}) \times \log 2\left(\frac{15}{22}\right) + (-\frac{7}{22}) \times \log 2\left(\frac{7}{22}\right)) = 0.90239$$

$$\text{Entropy [Diadora]} = ((-\frac{14}{19}) \times \log 2\left(\frac{14}{19}\right) + (-\frac{5}{19}) \times \log 2\left(\frac{5}{19}\right)) = 0.83147$$

$$\text{Entropy [Lotto]} = ((-\frac{18}{24}) \times \log 2\left(\frac{18}{24}\right) + (-\frac{6}{24}) \times \log 2\left(\frac{6}{24}\right)) = 0.81128$$

$$\text{Entropy [Adidas]} = ((-\frac{21}{24}) \times \log 2\left(\frac{21}{24}\right) + (-\frac{3}{24}) \times \log 2\left(\frac{3}{24}\right)) = 0.54356$$

$$\text{Entropy [Paragon]} = ((-\frac{19}{22}) \times \log 2\left(\frac{19}{22}\right) + (-\frac{3}{22}) \times \log 2\left(\frac{3}{22}\right)) = 0.57464$$

Entropy Price

$$\text{Entropy [Cheap]} = ((-\frac{61}{87}) \times \log 2\left(\frac{61}{87}\right) + (-\frac{26}{87}) \times \log 2\left(\frac{26}{87}\right)) = 0.87988$$

$$\text{Entropy [Ekspensive]} = ((-\frac{65}{92}) \times \log 2\left(\frac{65}{92}\right) + (-\frac{27}{92}) \times \log 2\left(\frac{27}{92}\right)) = 0.87317$$

$$\text{Entropy [Very Ekspensive]} = ((-\frac{88}{106}) \times \log 2\left(\frac{88}{106}\right) + (-\frac{18}{106}) \times \log 2\left(\frac{18}{106}\right)) = 0.65727$$

Entropy Inventory

$$\text{Entropy [Low]} = ((-\frac{190}{233}) \times \log 2\left(\frac{190}{233}\right) + (-\frac{43}{233}) \times \log 2\left(\frac{43}{233}\right)) = 0.68993$$

$$\text{Entropy [High]} = ((-\frac{24}{52}) \times \log 2\left(\frac{24}{52}\right) + (-\frac{28}{52}) \times \log 2\left(\frac{28}{52}\right)) = 0.99573$$

Entropy Demand

$$\text{Entropy [Decrease]} = ((-\frac{65}{133}) \times \log 2\left(\frac{65}{133}\right) + (-\frac{68}{133}) \times \log 2\left(\frac{68}{133}\right)) = 0.99963$$

$$\text{Entropy [Increase]} = ((-\frac{149}{152}) \times \log 2\left(\frac{149}{152}\right) + (-\frac{3}{152}) \times \log 2\left(\frac{3}{152}\right)) = 0.1399$$

The next step is to calculate the gain value using Equation 2.

$$\textbf{Gain (Total, type)} = (\text{entropy total}) - \left(\left(\frac{22}{285}\right) \times 0.90239\right) - \left(\left(\frac{21}{285}\right) \times 0.86312\right) - \left(\left(\frac{19}{285}\right) \times 0.89974\right) - \left(\left(\frac{26}{285}\right) \times 0.70627\right) - \left(\left(\frac{21}{285}\right) \times 0.9183\right) - \left(\left(\frac{26}{285}\right) \times 0.98286\right) - \left(\left(\frac{23}{285}\right) \times 0.75538\right) - \left(\left(\frac{16}{285}\right) \times 0\right) - \left(\left(\frac{22}{285}\right) \times 0.90239\right) - \left(\left(\frac{19}{285}\right) \times 0.83147\right) - \left(\left(\frac{24}{285}\right) \times 0.81128\right) - \left(\left(\frac{24}{285}\right) \times 0.54356\right) - ((\frac{22}{285}) \times 0.57464) = 0.05398$$

$$\textbf{Gain (Total, Prince)} = (\text{entropy total}) - \left(\left(\frac{87}{285}\right) \times 0.87988\right) - \left(\left(\frac{92}{285}\right) \times 0.87317\right) - \left(\left(\frac{106}{285}\right) \times 0.65727\right) = 0.01856$$

$$\textbf{Gain (Total, Inventory)} = (\text{entropy total}) - \left(\left(\frac{233}{285}\right) \times 0.68993\right) - \left(\left(\frac{52}{285}\right) \times 0.99573\right) = 0.06776$$

$$\textbf{Gain (Total, Demand)} = (\text{entropy total}) - \left(\left(\frac{133}{285}\right) \times 0.99963\right) - \left(\left(\frac{152}{285}\right) \times 0.13996\right) = 0.2723$$

Table 3. Results of Root Node

		Total	Approved	Pending	Entropy	Gain
Total		285	213	71	0.813485	
Type						0.05398
	Drill	22	15	7	0.902393	
	Toyobo	21	15	6	0.863121	
	Swessi	19	13	6	0.899744	
	Tissue	26	21	5	0.706274	
	Cotton	21	14	7	0.918296	
	PE	26	15	11	0.982859	
	TC	23	18	5	0.755375	
	Carded	16	16	0	0	
	Higet	22	15	7	0.902393	
	Diadora	19	14	5	0.831474	
	Lotto	24	18	6	0.811278	
	Adidas	24	21	3	0.543564	
	Paragon	22	19	3	0.574636	
Price	Cheap	87	61	26	0.879881	0.018564
	Ekspensive	92	65	27	0.873172	
	Very Ekspensive	106	88	18	0.657273	
Inventory	Low	233	190	43	0.689929	0.067761
	High	52	24	28	0.995727	
Demand	Decrease	133	65	68	0.999633	0.272344
	Increase	152	149	3	0.13996	

Based on the calculations above, the highest gain value is selected for forming the root. The table shows that the highest gain value comes from the Demand attribute, with a value of 0.272344, thus the Demand attribute will become the root node.

In the next stage, calculations are performed at node level 1. The gain value from the highest attribute calculation at the root is no longer included. The process begins by finding node 1.1, which is Permintaan = Decrease. Below are the calculation results for the class attribute Demand = Decrease, as shown in the table below:

Table 4. Results of Node 1.1 Search

Note			Total	Approved	Pending	Entropy	Gain
	Total		133	65	68	0.999633	
	Type						0.402878
1.1		Drill	7	0	7	0	
		Toyobo	11	5	6	0.99403	
		Swessi	6	0	6	0	

		Tissue	10	6	4	0.970951	
		Cotton	14	7	7	1	
		PE	11	0	11	0	
		TC	14	9	5	0.940286	
		Carded	9	9	0	0	
		Higet	10	5	5	1	
		Diadora	5	0	5	0	
		Lotto	6	0	6	0	
		Adidas	13	10	3	0.77935	
		Paragon	17	14	3	0.672295	
	Price	Cheap	29	5	24	0.663197	0.187859
		Ekspensive	37	11	26	0.877962	
		Very Ekspensive	67	49	18	0.83953	
	Inventory	Low	108	65	43	0.969857	0.212079
		High	25	0	25	0	

From the calculation results in the table, the attribute that becomes the branch node from Demand = Decrease is Type, with the highest gain value of 0.402878.

The next step is to calculate the entropy for the next node, where at this stage we calculate the remaining attributes (the entropy of the attributes that have not been valued at 0) based on the previous level 1 calculations.

Table 5. Results of Node 1.1.2 Search

Node			Total	Approved	Pending	Entrophy	Gain
1.1.2	Total		11	5	6	0.99403	
						0	
	Price	Cheap	0	0	0	0	0
		Ekspensive	11	5	6	0.99403	
		Very Ekspensive	0	0	0	0	
	Inventory	Low	10	5	5	1	0.084939
		High	1	0	1	0	

The highest gain value in the table is Price, with a value of 0, which becomes the root of the decision tree. There are three sub-attributes of Price: Cheap, Expensive, and Very Expensive. The sub-attributes Cheap and Very Expensive result in a Pending decision, while the sub-attribute Expensive requires further calculations.

The next stage in the calculation process for node 3 involves a detailed calculation of the entropy and gain values from the remaining attributes. The attributes to be calculated at this stage include various components such as 1.1.2.2, 1.1.4.2, 1.1.5.3, 1.1.7.3, 1.1.9.1, 1.1.12.3, 1.1.13.1, and 1.2.2.9. In this case, the first step will begin by focusing on node 1.1.2.2. The following table shows the calculation results for node 1.1.4.2:

Table 6. Results of Node 1.1.4.2 Search

Node			Total	Approved	Pending	Entrophy	Gain
	Total		11	5	6	0.99403	
						0	
1.1.2 .2	Inventory	Low	10	5	5	1	0.084939
		High	1	0	1	0	

Based on the table above, the value of the sub-attribute Few results in a Pending decision, while the value of the sub-attribute Many results in an Approved decision. The calculations are carried out until all nodes have their respective decisions. Thus, the calculation process is complete, and the decision tree can be seen in the image below:

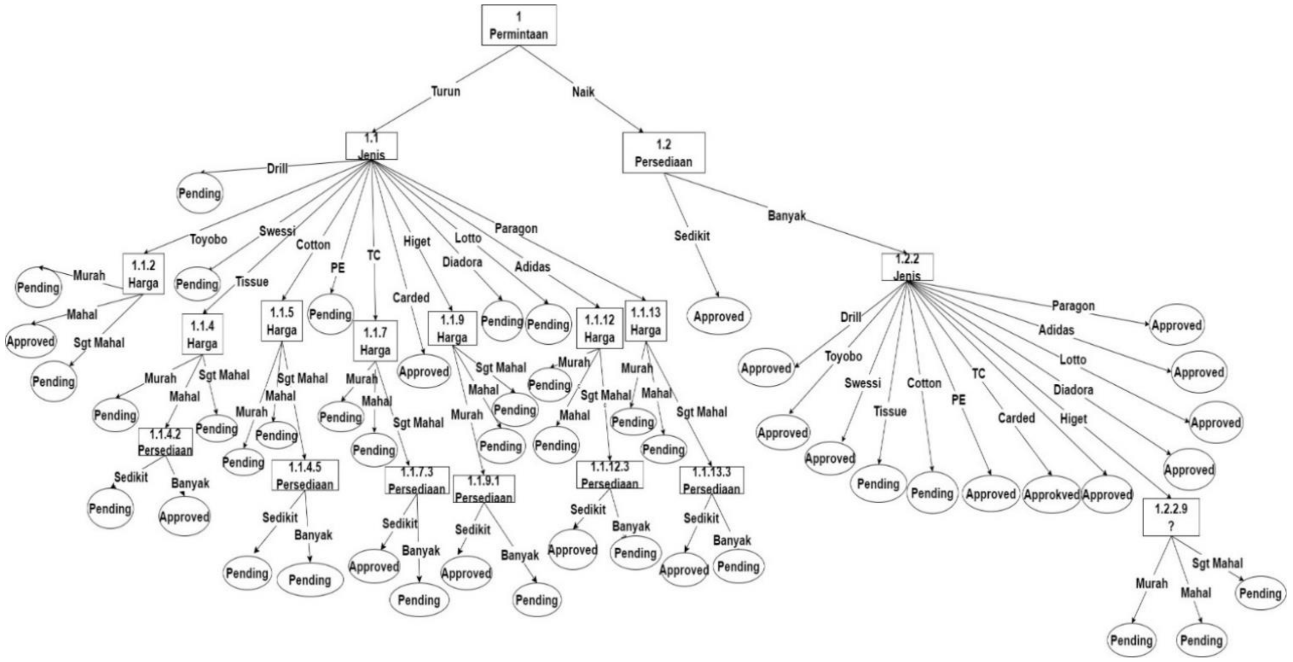

Figure 3 Final Decision Tree Results

Testing the Method

At this stage, a comprehensive evaluation of the performance of the C4.5 algorithm implemented in the system is conducted. The testing process uses 20% of the data from the dataset, specifically taken from the total of 365 available datasets. This test data is selected to measure how well the C4.5 algorithm can make predictions on new data, as well as to validate the accuracy of the algorithm in processing data that the system has not encountered before.

$$Acuracy = \frac{\text{Number of Correct Predictions}}{\text{Total Number of Predictions}}$$
$$= \frac{66}{71} \times 100$$
$$= 93\%$$
$$Error = \frac{\text{Number of Incorrect Predictions}}{\text{Total Number of Predictions}}$$
$$= \frac{5}{71} \times 100$$
$$= 7\%$$

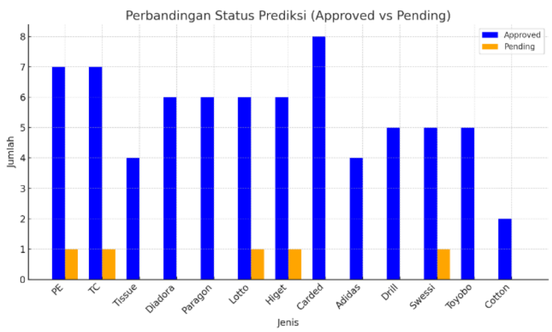

Figure 4 Test results graph

From the graph, it is evident that most types of goods have a high predicted quantity for the "Approved" status, indicating that the demand for these items is quite high and meets the approval criteria. Meanwhile, some items show a "Pending'"status, indicating the need for further evaluation to determine the eligibility of the request.

CONCLUSIONS

Based on the results that have been researched, the following conclusions can be drawn:

1. This study produces a prediction system for raw material stock for clothing production to assist garment owners in future raw material purchases. The attributes used are Type, Price, Inventory, and Demand.

2. The implementation of the C4.5 method in data mining has proven effective for predicting raw material stock for clothing production. This algorithm successfully identifies patterns and relationships in complex data, resulting in an accurate model for predicting raw material needs.
3. The use of the C4.5 method results in an accuracy rate of 93% in determining the status of raw materials, both approved and pending, with a 7% error rate. This allows for better accuracy in identifying stock needs.
4. Based on the results of this method's implementation, it provides clear guidance for production managers in making decisions related to ordering and managing raw materials. This not only enhances operational efficiency but also supports better data-driven production strategies.

Acknowledgments

We Would Like To Express Our Gratitude To The Manager Of Al-Fatih Konveksi In Baktiya, North Aceh, For Granting Permission To Collect The Necessary Data For This Research. We Also Extend Our Appreciation To All Parties Who Provided Support, Making The Publication Of This Research Possible.

REFERENCES

[1] Surajiyo, Rina Wahyu, and Winarni, "Hubungan Ilmu Pengetahuan Dengan Teknologi Relationship Of Science With Technology," SEMINASTIKA , vol. 3, no. 1, pp. 182–187, 2021. https://doi.org/10.47002/seminastika.v3i1.263

[2] N. Nurdin and D. Astika, "Penerapan Data Mining Untuk Menganalisis Penjualan Barang Dengan Menggunakan Metode Apriori Pada Supermarket Sejahtera Lhokseumawe," TECHSI-Jurnal Teknik Informatika, vol. 7, no. 1, pp. 132–155, 2019. https://doi.org/10.29103/techsi.v7il.184

[3] M. Fikry, R. Tjut Adek, F. Fadlisyah, N. Nurdin, D. Hamdana, and M. Qamal, "Data Mining for Processing of Research and Community Service by Lecturer Using Decision Tree Method," International Journal of Psychosocial Rehabilitation, vol. 24, no. 2, pp. 367–371, 2020. https://www.psychosocial.com/article/PR201323/1564

[4] N. Nurdin, F. Fajriana, M. Maryana, and A. Zanati, "Information System for Predicting Fisheries Outcomes Using Regression Algorithm Multiple Linear," Journal Of Informatics And Telecommunication Engineering, vol. 5, no. 2, pp. 247–258, 2022. https//doi.org/10.31289/jite.v5i2.6023

[5] A. N. Aziziah, A. I. Purnamasari, and I. Ali, "Penerapan Algoritma C4. 5 Untuk Prediksi Stok Bahan Minuman Di Cafe Semanis," JATI (Jurnal Mahasiswa Teknik Informatika), vol. 8, no. 1, pp. 292–295, 2024. https://doi.org/10.36040/jati.v8i1.8347

[6] G. L. Pritalia, "Penerapan Algoritma C4. 5 untuk Penentuan Ketersediaan Barang E-commerce," Indonesian Journal of Information Systems, vol. 1, no. 1, pp. 47–56, 2018. https://doi.org/10.24002/ijis.v1i1.1727

[7] E. Elisa, "Penerapan Algoritma C4. 5 Untuk Memprediksi Penjualan Barang Pada Pt Batam Bangun Prathama," Computer and Science Industrial Engineering (COMASIE), vol. 7, no. 1, pp. 127–133, 2022. https://ejournal.upbatam.ac.id/index.php/comesiejournal

[8] Dwita Elisa Sinaga, Agus Perdana Windarto, and Rizki Alfadillah Nasution, "Analisis Data Mining Algoritma Decision Tree Pada Prediksi Persediaan Obat (Studi Kasus : Apotek Franch Farma)," KLIK: Kajian Ilmiah Informatika dan Komputer, vol. 2, no. 4, pp. 123–131, 2022, https://doi.org/10.30865/klik.v2i4.328

[9] H. Hendri, "Implementasi Data Mining Dengan Metode C4. 5 Untuk Prediksi Mahasiswa Penerima Beasiswa," Indonesian Journal of Computer Science, vol. 10, no. 2, pp. 312–321, 2021. https://doi.org/10.33022/ijcs.v10i2.3013

[10] B. G. Sudarsono, M. I. Leo, A. Santoso, and F. Hendrawan, "Analisis Data Mining Data Netflix Menggunakan Aplikasi Rapid Miner," JBASE - Journal of Business and Audit Information Systems, vol. 4, no. 1, 2021, https://doi.org/ 10.30813/jbase.v4i1.2729

[11] Kusrini and T. L. Emha, Algoritma Data Mining. Yogyakarta: Andi Offset, 2009.

[12] B. Q. Husaini and J. Jemakmun, "Penerapan Algoritma Decision Tree C45 untuk Klasifikasi Penjurusan Siswa," Jurnal Teknologi Informatika dan Komputer, vol. 9, no. 1, pp. 455–470, 2023, https://doi.org/10.37012/jtik.v9i1.1512

[13] A. R. Sukma, R. Halfis, and A. Hermawan, "Klasifikasi Channel Youtube Indonesia Menggunakan Algoritma C4. 5," Jurnal Teknik Komputer, vol. 5, no. 1, pp. 21–28, 2019. https://doi.org/10.31294/jtk.v5i1.4823

[14] Asmaul Husnah Nasrullah, "Implementasi Algoritma Decision Tree Untuk Klasifikasi Produk Laris ," Jurnal Ilmiah Ilmu Komputer, vol. 7, no. 2, pp. 45–51, 2021. https://doi.org/10.35329/jiik.v7i2.203

[15] S. Yunita and V. N. Alaeyda, "Penerapan Algoritma C4.5 Untuk Prediksi Penerimaan Beasiswa di SD 4 Pelangsian," ICIT Journal, vol. 8, no. 2, pp. 181–193, 2022, https://doi.org/ 10.33050/icit.v8i2.2408