Journal of Artificial Intelligence and Engineering Applications

Website: https://ioinformatic.org/

15th February 2025. Vol. 4. No. 2; e-ISSN: 2808-4519

Optimizing the Classification Model for Plant Medicine Supplies Using the Decision Tree Algorithm at the Anugrah Tani Shop, Brebes Regency

Saeful Amri[1]*, Rudi Kurniawan[2], Saeful Anwar[3]

[1,2,3]STMIK IKMI Cirebon
Amrisaeful417@gmail.com[1], rudi226@gmail.com[2] , saeful.ikmi@gmail.com[3]*

Abstract

Retail businesses in the agricultural industry often face difficulties in estimating inventory needs, especially plant medicines which are important for protecting plants from pests and diseases. The lack of an accurate inventory prediction system can cause stock discrepancies, as happened at the Anugrah Tani Store, Brebes Regency, thereby disrupting operations and customer satisfaction. This research uses the Decision Tree classification technique to increase the accuracy of predicting the need for plant medicine supplies, with a clustering approach using the K-Means algorithm to determine the optimal K value through the Davies-Bouldin Index (DBI) calculation. A DBI value of -0.065 indicates good cluster quality with an optimal K of 2, where Cluster 0 has high inventory needs (1138 data) and Cluster 1 has low needs (4 data). The analysis results show that the accuracy level of the Decision Tree model is 98.25%, which is quite high. This model is not only able to predict inventory patterns accurately but also provides in-depth insights to support stock decision making. This research proves that the Decision Tree algorithm can help inventory management with a faster response to customer needs, while contributing to the development of machine learning-based classification models for the agricultural and retail sectors.

Keywords: *Decision Tree, classification, Knowledge Discovery in Database, Davies Bouldin Index, inventory management*

1. Introduction

Accurate predictions of the need for plant medicine supplies are very important for the Anugrah Tani Shop, Brebes Regency, to avoid problems with excess or shortage of stock which can be detrimental. Effective inventory management supports store operations and competitiveness. One method that can be used is the Decision Tree algorithm, which has been proven to be able to increase the accuracy of inventory predictions based on historical data [1], [2]. Decision Tree is a data mining and machine learning technique that builds a classification or prediction model based on a decision tree structure [3]. With advances in technology and the availability of big data, Decision Trees provide significant benefits for entrepreneurs in stock management and strategic planning, as well as supporting the efficiency of evidence-based decision making [4]. The implementation of this algorithm in the retail industry, including the Anugrah Tani Store, shows positive results in forecasting product needs and optimizing inventory management [5].

Inaccurate information regarding the inventory of plant medicines at the Anugrah Tani Shop, Brebes Regency, is a challenge in stock management, affecting business success and customer satisfaction. Stock shortages or excesses often occur due to the lack of effective methods for predicting demand. Classification algorithms such as Decision Tree offer solutions by identifying historical data trends and accurately predicting inventory needs [6]. The main benefit of this algorithm is its ability to predict the amount and time of stock availability based on previous data patterns. Research shows that Decision Trees can achieve a satisfactory level of accuracy, although the results depend on data quality, algorithm parameter settings, and other variables [7]. In small retail companies, this model is able to achieve accuracy of up to 85%, but implementation in small businesses still faces challenges, such as inconsistent data quality and selecting the right parameters [8]. Although this technology has great potential in improving inventory management, its use at the small business level, such as Toko Anugrah Tani, is still limited. Further research is needed to optimize the implementation of Decision Trees and maximize their benefits in this sector[9].

This research uses the Decision Tree algorithm to perfect the data classification model for plant medicine supplies at the Anugrah Tani Shop, Brebes Regency. The aim is to increase the accuracy of predicting inventory needs, thereby helping store management plan procurement more precisely, reducing the risk of stock shortages or excesses [10]. This technique is important for the agricultural sector, which has more complex dynamics than other sectors due to the influence of external factors such as season and crop varieties. This research closes the gap in the literature regarding inventory prediction in the agricultural sector. While the Decision Tree algorithm has been applied in other fields, such as pharmacy [11], the focus on the agricultural sector provides a new contribution. By increasing the accuracy of demand prediction and inventory management, this research offers practical solutions that can be directly applied to increase the efficiency of plant medicine stock management [12]. The results are not only relevant for stock management but also contribute to the development of informatics in the agricultural sector, helping to increase productivity and reduce waste.

This research uses the Knowledge Discovery in Database (KDD) approach which includes stages of data selection, pre-processing, data transformation, data mining, and evaluation to develop a prediction model for plant medicine supplies using the Decision Tree algorithm [13]. Data from Toko Anugrah Tani, such as units and product types, is used to train the model, while model performance is evaluated using quantitative metrics such as fusion matrices and other classification indicators [14]. This research aims to reduce the risk of excess or shortage of stock by accurately predicting inventory needs. By utilizing machine learning and structured data analysis, this approach provides practical solutions for inventory management in the agricultural sector as well as theoretical contributions in the application of classification algorithms in this industry [15].

2. Research Method

This research uses a quantitative method by predicting the supply of plant medicines at the Anugrah Tani shop using the Knowledge Discovery in Database (KDD) approach and the Decision Tree C4.5 algorithm as the main model. Because it can reveal hidden patterns in large amounts of data and help make better business decisions, the prediction development process includes data analyst steps using the Knowledge Discovery in Database method.

2.1. Data Analysis Techniques

The Decision Tree algorithm is used as a data analysis method during the data mining process. This technique aims to calculate the medicinal needs of Toko Anugrah Tani's plants by finding patterns and relationships in the data. Knowledge Discovery in Database (KDD) is an approach used to analyze data, with the Decision Tree algorithm as the main method.

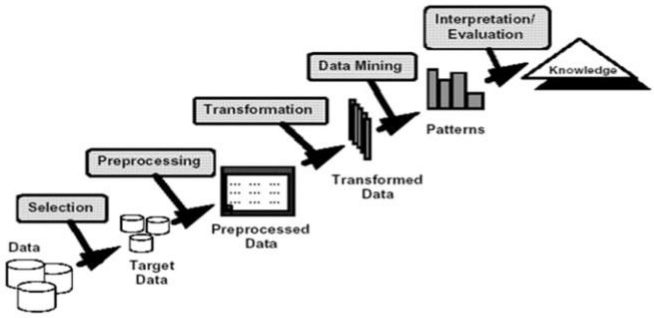


Fig. 1: Knowledge Discovery in Databases [16]

a. Seelection
At this point, relevant research data is collected in a systematic way. To ensure the correctness of the data, the Anugrah Tani shop owner was interviewed directly on October 9 2024. Product types, incoming stock, outgoing stock, and other attributes related to inventory needs analysis are included in 1,142 data which has 13 features. which form the data obtained. This technique was chosen primarily to aid inventory management and ensure that predictive models can be created with data that is accurate and relevant for research purposes.

b. Preprocesing
Carried out to improve the quality of data and make it available for use at a later stage. At this stage, the main tasks include
data cleansing : removes inconsistent, empty, or invalid data
encoding data: to make analysis easier, into numerical format this stage ensures that the data is anomaly free and in a format that can be used for further analysis.

c. Transformation
To ensure the data is ready to be used in the modeling process with the Decision Tree algorithm, two main operators are used at the data transformation stage of this research, namely Nominal to Numerical. Categorical type attributes such as "Product Type," "Unit," and "Entry Month" are converted to numeric format for algorithm compatibility using the Nominal to Numerical operator. To do this, each category is given a different numerical value, for example "Fungicide" is given a value of 1 and "Fertilizer" is given a value of 2. The purpose of this transformation is to simplify data patterns, increase the effectiveness of analysis, and support the classification process in building prediction models.

d. Data Mining
Data mining is the core stage in Knowledge Discovery in Databases (KDD), where important patterns are extracted from datasets. This research uses the Decision Tree algorithm to build a predictive model for the need for plant medicine supplies at the Anugrah Tani Store. This algorithm divides the data based on the most influential attributes, such as "Incoming Date" and "Incoming Stock," producing a decision tree that represents classification rules. This process aims to produce a model with high accuracy that makes stock management easier and provides insight into stock distribution patterns. By using quality data, the model can identify significant factors such as monthly stock patterns and product types. Performance evaluation is carried out using accuracy, precision and Davies-Bouldin Index (DBI) metrics.
Decision Trees were chosen because of their ability to produce models that are easy to understand and effective in handling numerical and categorical data. With this method, research is expected to make a practical contribution to inventory management in the agricultural sector.

e. Interpretation

interpretation and evaluation in the Knowledge Discovery in Database (KDD) process which aims to assess the quality and relevance of patterns or information found during the data mining process. At this stage, the results of the Decision Tree and K-Means algorithms are analyzed to ensure that the model built is able to meet the research objectives. Evaluation involves measuring model performance using metrics such as accuracy, precision, recall, and the Davies-Bouldin Index (DBI).

f. Knowledge
The final stage in the Knowledge Discovery in Database (KDD) process is extracting knowledge from plant medicine inventory data at the Anugrah Tani Store. The results of analysis using the Decision Tree method produce classification rules that can be used to predict inventory needs. The resulting decision tree identifies relationships between attributes, such as “Incoming Date,” “Incoming Stock,” and “Product Type,” with “Incoming Date” as the key attribute in determining inventory categories. This knowledge provides insight into the most influential factors in stock management.

3. Reuslt and Discussion

In this results and discussion section, the process of implementing Knowledge Discovery in Database (KDD) will be explained which consists of data selection, preprocessing, transformation, data mining, and evaluation. This research utilizes the K-Means algorithm to carry out clustering, which aims to group plant medicine inventory data into clusters with certain characteristics. The cluster labeling results from K-Means are used as input in a classification model using the Decision Tree algorithm to predict inventory needs more accurately. The discussion also includes analysis of clustering and classification results, including evaluation of model performance based on accuracy and patterns found in the data. This process is designed to support more effective decision making in stock management at the Anugrah Tani Store.

3.1. Data Selection

At the Data Selection stage, plant medicine inventory data from the Anugrah Tani Store is taken from an Excel file which contains information such as item code, item name, product type and unit. This data is imported into RapidMiner using Excel's Read operator, which allows direct use without the need to change the file format. This captured data becomes the main component in analysis with the Decision Tree algorithm, prepared for the next processing steps.

Table 1: Results of the Read Excel operator

No	Uraian	Keterangan
1	Record	**1142**
2	Spesial Attribut	**1**
3	Reguler Attribute	**13**
4	Attribute :	
	NO (id)	**Integer, missing 0**
	KODE BARANG	**Polynominal, missing 0**
	NAMA BARANG	**Polynominal, missing 0**
	JENIS PRODUK	**Polynominal, missing 0**
	SATUAN	**Polynominal, missing 1**
	HARGA MASUK	**Integer, missing 0**
	HARGA KELUAR	**Integer, missing 0**
	TANGGAL MASUK	**Integer, missing 0**
	BULAN MASUK	**Polynominal, missing 0**
	STOK MASUK	**Integer, missing 0**
	TANGGAL KELUAR	**Integer, missing 0**
	BULAN KELUAR	**Polynominal, missing 0**
	STOK KELUAR	**Integer, missing 0**

The Read Excel results show that the dataset used in this study consists of 1,142 records with no special attributes and 13 regular attributes. Each attribute has a specific data type: numerical elements (e.g., NO, HARGA MASUK, HARGA KELUAR, TANGGAL KELUAR, TANGGAL MASUK, STOK MASUK, STOK KELUAR) are of integer type, while categorical attributes (e.g., KODE BARANG, NAMA BARANG, SATUAN, JENIS PRODUK, BULAN MASUK, BULAN KELUAR) are of polynomial type. Missing values were found in the SATUAN attribute, while all other attributes contain complete data. The dataset's quality is sufficient for further analysis during the transformation and classification phases. With complete and well-structured data, this study can focus on predicting inventory needs using the Decision Tree technique.

Table 2: Results of the Set Role operator

No	Uraian	Keterangan
1	Record	**1142**
2	Spesial Attribut	**1**
3	Reguler Attribute	**13**
4	Attribute :	
	NO (id)	**Integer, missing 0**
	KODE BARANG	**Polynominal, missing 0**
	NAMA BARANG	**Polynominal, missing 0**
	JENIS PRODUK	**Polynominal, missing 0**
	SATUAN	**Polynominal, missing 1**
	HARGA MASUK	**Integer, missing 0**

HARGA KELUAR	**Integer, missing 0**
TANGGAL MASUK	**Integer, missing 0**
BULAN MASUK	**Polynominal, missing 0**
STOK MASUK	**Integer, missing 0**
TANGGAL KELUAR	**Integer, missing 0**
BULAN KELUAR	**Polynominal, missing 0**
STOK KELUAR	**Integer, missing 0**

The dataset contains 1142 records, each of which has 13 regular attributes and one special attribute, according to the Set Role results. The No property is used as an ID or special because it serves as a unique identifier for each row of data in the data set. To avoid redundancy and enable complete tracking of the analysis process, this option is important, and ensures that each data entry can be differentiated from others. Additionally, the integer data type in the NO attribute and the missing 0 value indicate that there is no blank data in this column, making it an excellent choice for primary keys in data management. Other variables such as PRODUCT TYPE, ITEM CODE, and ITEM NAME are used as descriptive attributes in the classification and analysis process. The structure of the data set is more organized when NO is used as the ID, and the Decision Tree method can process additional data more easily during the predictive analysis step.

3.2. Preprocesing

The preprocessing stage is an important initial stage in the data analysis process. One of the main tasks at this stage is data cleaning, namely the process of removing errors or inconsistencies from the data that may have an impact on the analysis findings. Cleaning data aims to remove empty values or data errors so that there are no problems when the data is processed. The results of preprocessing can be seen in the table 3.

Table 3: Results of the Missing Value operator process

NO	ATRIBUT	TIPE DATA	MISSING VALUE
1	NO	Integer	**0**
2	KODE BARANG	Polynominal	**0**
3	NAMA BARANG	Polynominal	**0**
4	JENIS PRODUK	Polynominal	**0**
5	SATUAN	Polynominal	**0**
6	HARGA MASUK	Integer	**0**
7	HARGA KELUAR	Integer	**0**
8	TANGGAL MASUK	Integer	**0**
9	BULAN MASUK	Polynominal	**0**
10	STOK MASUK	Integer	**0**
11	TANGGAL KELUAR	Integer	**0**
12	BULAN KELUAR	Polynominal	**0**
13	STOK KELUAR	Integer	**0**

indicates that the dataset no longer has missing items in the UNIT attribute.

3.3. Tranformation data

Data transformation is an important step in the data processing process, especially when attributes with nominal data types are found in raw data and must be converted to numeric so that they can be used with certain analysis techniques. The Nominal to Numeric operator is used to change the nominal (category) data type to a numeric data type for attributes. In the data mining process, this transformation is very important, especially if using the Decision Tree algorithm which can only handle numerical data. The Decision Tree algorithm used in research requires numerical data to process attributes. All data is guaranteed to be in the correct format by the Nominal to Numeric operator. In the Knowledge Discovery in Database (KDD) method, data transformation is an important step that changes data into the best analysis format. The results of the data transformation process can be seen in the table 4

Table 4: Results of the Nominal to Numerical operator transformation process

Atribut	Nominal	Numerik
JENIS PRODUK	Insektisida	**0**
	Perekat	**1**
	Pupuk Cair	**2**
	Fungisida	**3**
	Herbisida	**4**
	Pupuk Daun	**5**
	Pupuk Organik	**6**
	Pestisida	**7**
	Pupuk	**8**
	Pupuk Akar	**9**
	Nematisida	**10**
	Bakterisida	**11**
	Fumigan	**12**
	Akarisida	**13**
	Moluskisida	**14**
	Disinfektan	**15**
	Rodentisida	**16**
	Zat	**17**

SATUAN	Botol	**0**
	Pcs	**1**
	Kilogram	**2**
	Liter	**3**
BULAN MASUK	January	**0**
	February	**1**
	Maret	**2**
BULAN KELUAR	February	**0**
	Maret	**1**
	April	**2**
	Mei	**3**

the result of the Nominal to Numerical operator process which changes nominal type (category) variables to numeric type. This step is very important because most machine learning algorithms, including the Decision tree algorithm, require numerical data to carry out the analysis process.

3.4. Data Mining

Data mining is a series of processes used to collect valuable information from large amounts of data. Data on plant medicine supplies at the Anugrah Tani Store was analyzed in this research using data mining. Knowledge Discovery in Database (KDD) is used to discover patterns and relationships that drive optimal inventory control. To predict future inventory demand, data mining steps start with data collection and end with analysis of the resulting model. This data mining process uses several operators that have different objectives. The Clustering Operator (K-Means) is used to group data into a number of clusters based on certain similar patterns or characteristics. In this research, K-Means is used to identify patterns in Toko Anugrah Tani's plant medicine inventory data. By grouping the data into two clusters (high and low), research can focus more on the characteristics of each cluster.

Table 5: Cluster Model results

No	Cluster	Keterangan
1.	Cluster_0	**1138 items**
2.	Cluster_1	**4 items**
	Total	**1142 items**

shows that the data is separated into two clusters, cluster_0 which has 1138 items, and cluster_1 which has 4 items. The total amount of data is 1142. The K-Means algorithm, which groups data based on incoming patterns, outgoing stocks, and incoming dates, is used to perform this separation. Most of the data is covered by Cluster_0, which indicates that most items have comparable properties in terms of stock volume and transaction patterns. In contrast, Cluster_1 has fewer items, indicating low-volume transactions or things with more typical patterns. This separation reflects different patterns in the data, which is confirmed by the Davies-Bouldin Index (DBI) value of -0.065, indicating that the quality of the clustering is very good.

Table 6: operator performance results (cluster distance)

PerformanceVector
PerformanceVector
Avg. within centroid distance: -6962825819.280
Avg. within centroid distance_cluster_0: -6993364286.913
Avg. within centroid distance_cluster_1: -55198.880
Davies Bouldin: -0.065

The PerformanceVector results show how the average K-Means value is used to assess the quality of the clusters created during the clustering process. The average distance between the data points of each cluster and the cluster center is described by Within Centroid Distance of -6,962,825,819,280; negative values indicate converted distance-based estimates. The average distance of Cluster 0 is -6,993,364,286,913, much further than Cluster 1 which is only -55,198,880. This shows that Cluster 0 has a more complicated and wider data distribution, while Cluster 1 is more homogeneous and dense. The resulting cluster has good quality, as indicated by a Davies Bouldin Index (DBI) value of -0.065. because DBI values close to zero or negative indicate optimal cluster separation and high internal density.

Table 7: Result of the Multiply operator

No	AtributNO	Label	Keterangan
1	1	Cluster_0	**Callicron**
2	2	Cluster_0	**Brasso**
3	3	Cluster_0	**Josstik**
4	4	Cluster_0	**Joss trubus**
……	……….	………….	**……………**
26	26	Cluster_0	**Dargo**
27	27	Cluster_1	**Bescozeb**
28	28	Cluster_0	**Maher**
…..	……….	….	**………**
1141	1141	Cluster_0	**Zytrol**
1142	1142	Cluster_0	**Zytum**

Contains important details such as product name, product type, unit, month in, month out, price in, price out, date in, and stock in. The initial classification objective (ground truth), which is cluster_0, is displayed in the label column without any label modification. Each of the 1,145 rows in this dataset represents a different item entity, and variables such as PRICE IN, PRICE OUT, and STOCK IN show fluctuations in value that correspond to the data properties. Since there is no label change, additional analysis is needed to ensure this dataset is representative enough to support the analysis objectives, even though it is ready to be used in model training, evaluation, or visualization.

Table 8: Split Data Parameter Partition

NO	Parameter	Ratio
1	Data training	**0,8**
2	Data testing	**0,2**

Divided into two main parts training data and testing data, with a ratio of 0.8:0.2, with a ratio of 80% in the training data, this model reduces the possibility of underfitting, which occurs when the model fails to identify enough patterns in the data to carry out deep learning. This ratio is a compromise between having a large enough training data set and still having data available for testing. A ratio of 20% is considered adequate to achieve meaningful evaluation results while maintaining the size of the training data. This ratio helps ensure that assessment results accurately represent practical model potential, such as predictions on new data sets. An effective model training process and representative model performance evaluation is achieved by dividing the data in a ratio of 0.8:0.2, next is the Decision tree operator process to create a decision tree-based classification or regression model. This operator uses the dataset to be input to create a decision tree-based model that predicts target values from the processed dataset. By using attributes such as product type, incoming stock, outgoing stock, etc., a classification model was developed to predict the need for plant medicine supplies which can be seen in Figure 2

Fig. 2: Operator Decision Tree

A value of 4 indicates that separation can only be done if the node contains four pieces of data. Using the predictor variable EXIT DATE, produces a basic decision tree that divides the data into two branches according to a threshold value of 29,500, but all data is categorized as cluster 0. There is little or no data (thin line) in the left branch, which represents the EXIT DATE value greater than 29,500. On the right side, for values less than or equal to 29,500, there is more data (shown by the bold blue area), but the classification remains cluster_0. This shows that this variable is less effective in separating data into different clusters, because the classification results do not vary which can be seen in Figure 3

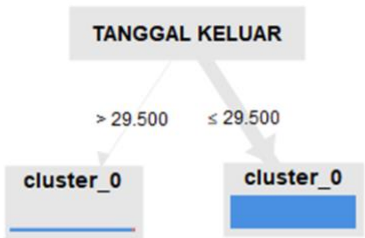
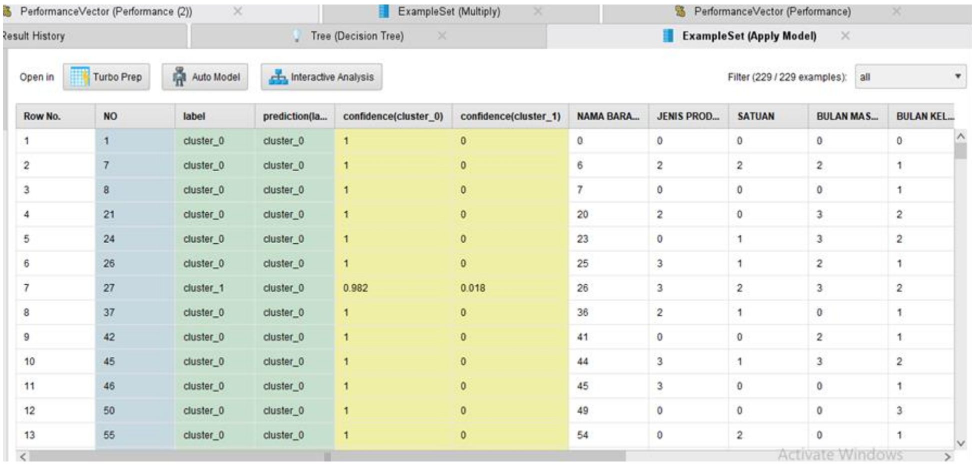


Fig. 3: Decision Tree Results

The Apply Model operator shows the strong tendency of the decision tree model in predicting data up to cluster_0, as evidenced by the prediction results and consistently high confidence values (confidence(cluster_0) = 1) in almost every row, except row 6, which shows a small amount of uncertainty with confidence(cluster_0) = 0.982 and confidence(cluster_1) = 0.018. The model prediction results that are all cluster_0 are displayed in the prediction (label), while the label column reflects the ground truth. Although other features such as ITEM NAME, PRODUCT TYPE, UNIT, MONTH IN, and MONTH OUT are used to help predictions, the model does not seem able to take advantage of differences in data enough to categorize them into other clusters. This indicates that the dataset may be imbalanced or the model rules are less complex, so further evaluation is needed, such as adding new features or adjusting model parameters, to improve classification performance. The results of the Apply model operator can be seen in Figure 4.



Row No.	NO	label	prediction(la...	confidence(cluster_0)	confidence(cluster_1)	NAMA BARA...	JENIS PROD...	SATUAN	BULAN MAS...	BULAN KEL...
1	1	cluster_0	cluster_0	1	0	0	0	0	0	0
2	7	cluster_0	cluster_0	1	0	6	2	2	2	1
3	8	cluster_0	cluster_0	1	0	7	0	0	0	1
4	21	cluster_0	cluster_0	1	0	20	2	0	3	2
5	24	cluster_0	cluster_0	1	0	23	0	1	3	2
6	26	cluster_0	cluster_0	1	0	25	3	1	2	1
7	27	cluster_1	cluster_0	0.982	0.018	26	3	2	3	2
8	37	cluster_0	cluster_0	1	0	36	2	1	0	1
9	42	cluster_0	cluster_0	1	0	41	0	0	2	1
10	45	cluster_0	cluster_0	1	0	44	3	1	3	2
11	46	cluster_0	cluster_0	1	0	45	3	0	0	1
12	50	cluster_0	cluster_0	1	0	49	0	0	0	3
13	55	cluster_0	cluster_0	1	0	54	0	2	0	1

Fig. 4: Result of the Apply model operator

The results of the performance (classification) operator can be seen in the table 9

Table 9: Performance (Classification) operator results

	True cluster_0	True cluster_1	Class precision
Pred.cluster_0	225	4	**98.25%**
Pred.cluster_1	0	0	**0.00%**
Class recall	100.00%	0.00%	
Accuracy: 98.25%			

The results of the Performance Cluster operator show that the model has an accuracy of 98.25%, with excellent performance in classifying data into cluster_0, where 225 data with the original label cluster_0 were classified correctly, resulting in 100.00% recall and 98.25% precision for this cluster. However, the model failed completely in predicting cluster_1, with no data correctly classified to cluster_1, resulting in a recall and precision of 0.00% for that cluster. This shows that there is a very strong model bias towards cluster_0, possibly due to data imbalance or model limitations in capturing the characteristics of cluster_1, so further evaluation is needed to improve model performance against non-dominant clusters.

3.5. Evaluation

This research conducted an evaluation to evaluate the performance of the Decision Tree algorithm-based classification model in predicting the need for plant medicine supplies at the Anugrah Tani Store. This evaluation method utilizes the Davies-Bouldin Index (DBI) accuracy, an accuracy of 98.25%, which shows that most of the predictions match the actual data. Precision is used to guarantee the accuracy of predictions, especially when differentiating between low and high stocks. In addition, a DBI of -0.065 indicates good clustering quality, with strong internal cohesion and excellent cluster separation. Data is classified into two groups: cluster 0 indicates high stock and cluster 1 indicates low stock.

Table 10: Evaluation Metrics

NO	Metrik evaluasi	Nilai
1	Akurasi	**98.25%**
2	Presisi	**-**
3	Davies Bouldin Index(DBI)	**-0,065%**
4	Cluster_0	**1138 data**
5	Cluster_1	**4 data**

The evaluation table above provides a comprehensive picture of the performance of the model used in the research. The accuracy value of 98.25% shows that the Decision Tree model has a very high level of reliability in predicting the need for plant medicine supplies at the Anugrah Tani Store. Although precision is not listed, the focus of the evaluation is directed at overall accuracy and clustering results. The Davies Bouldin Index (DBI) of -0.065 indicates that the clusters formed by the K-Means algorithm are of good quality, with fairly large distances between clusters and a tight distribution of data within the clusters. From the clustering results, Cluster_0 includes 1,138 data which represents high inventory needs, while Cluster_1 only consists of 4 data which shows low inventory needs. This information is very important in supporting inventory management, because it can help inventory managers prioritize stock needs more effectively and avoid overstock or understock. The combination of classification and clustering results shows that the approach used in this research is effective in supporting decision making at the Anugrah Tani Store.

4. Conclusions and Suggestions

4. 1. Conclusions

This research succeeded in creating a predictive model for managing plant medicine supplies at the Anugrah Tani store using the Decision Tree method guided by the K-Means clustering technique. This explanation can be explained as follows:

1. By using the Davies-Bouldin Index and accuracy values as evaluation criteria, this research succeeded in identifying the ideal number of clusters (K value) for grouping inventory data. The analysis findings support the more efficient use of the Tree

Algorithm in inventory prediction by showing that the Davies-Bouldin Index value a low one indicates very good cluster grouping quality.
2. From the results of data grouping, the clusters formed have different characteristics, especially in the context of the number of plant medicine supplies. Clusters with high inventory provide an indication of more stable stock needs, while clusters with low inventory identify more dynamic stock management needs. This information provides critical information for creating inventory management strategies that better match demand.
3. To categorize and predict plant medicine inventory data at the Anugrah Tani Store, this research succeeded in developing an efficient Decision Tree-based classification algorithm. This methodology supports more effective stock management and can make accurate predictions. This research increases store operational efficiency by integrating the K-Means clustering method with Decision Tree implementation. This not only speeds up response to changing stock requirements but also improves understanding of data trends.

4. 2. Suggestions

There are several recommendations that can become references for additional research in an effort to refine and expand the findings of this study, which aims to increase precision. The effectiveness, and applicability of the created prediction models, covering topics such as data diversity, model evaluation, technology integration, and application of various algorithms.

1. Other measures such as recall, F1 score, or ROC-AUC can also be used to improve model performance measurements. It is recommended that studies investigating these findings use more diverse datasets by combining data from additional similar repositories to provide a more thorough assessment of model performance.
2. Predictive models should be quickly integrated into the store's inventory management system so that end users can provide feedback and help determine how useful the model will be in practical situations. To improve results, it is planned that future research will compare the performance of the Decision Tree algorithm with other algorithms, including Random Forest, Gradient Boosting, or deep learning.
3. To speed up the decision-making process, it is recommended to incorporate Internet of Things (IOT) technology to obtain real-time stock data that is directly connected to predictive models. Additionally, more useful documentation and visualization such as time series graphs or heat maps are needed to help inventory managers understand increasingly complex data patterns.

References

[1] R. Rofiani, “Penerapan Metode Klasifikasi Decision Tree dalam Prediksi,” vol. 18, no. 1, pp. 126–139.
[2] 4 Rika Nursyahfitri1, Alfanda Novebrian Maharadja2, Riva Arsyad Farissa3, Yuyun Umaidah4 1,2,3, “Klasifikasi Penentuan Jenis Obat Menggunakan Algoritma Decision Tree,” pp. 53–60, 2021.
[3] D. Sartika, “Perbandingan Algoritma Klasifikasi Naive Bayes , Nearest Neighbour , dan Decision Tree pada Studi Kasus Pengambilan Keputusan Pemilihan Pola Pakaian,” vol. 1, no. 2, pp. 151–161, 2017.
[4] H. Priatmojo, F. Saputra, M. H. Prasetiyo, D. Puspitasari, and D. Nurlaela, “Perbandingan Klasifikasi Tingkat Penjualan Buah di Supermarket dengan Pendekatan Algoritma Decision Tree , Naive Bayes dan K-Nearest Neighbor,” vol. 3, no. 1, pp. 21–28, 2023.
[5] U. Arfan and N. Paraga, “The Comparison of K-Means , Naïve Bayes and Decision Tree Algorithm in Predicting Fuel Oil Sales Perbandingan Algoritma K-Means , Naïve Bayes dan Decision Tree dalam Memprediksi Penjualan Bahan Bakar Minyak,” vol. 4, no. October, pp. 1379–1389, 2024.
[6] Y. C. F. Sinta Amanad Pratiwi, Ahmad Fauzi*, Santi Arum Puspita Lestari, “Prediksi Persediaan Obat Pada Apotek Menggunakan Algoritma Decision Tree,” vol. 4, no. 4, pp. 2381–2388, 2024, doi: 10.30865/klik.v4i4.1681.
[7] Suwandi, “DALAM PENENTUAN PREDIKSI STOK BARANG,” vol. 3, no. 2, pp. 12–18, 2022.
[8] Y. Rizkita and Putri, “Prediksi pola kecelakaan kerja pada perusahaan non ekstraktif menggunakan algoritma decision tree: c4.5 dan c5.0,” vol. 2, no. 1, 2013.
[9] R. M. Simanjorang, “Penerapan Data Mining Dalam Menentukan Persediaan Obat Dengan,” vol. 22, pp. 414–421, 2023.
[10] A. Aziz, “Prediksi Penjualan Obat Dan Alat Kesehatan Terlaris Menggunakan,” vol. 6, no. 1, pp. 117–124, 2024.
[11] D. E. Sinaga, “Analisis Data Mining Algoritma Decision Tree Pada Prediksi Persediaan Obat (Studi Kasus : Apotek Franch Farma),” vol. 2, no. 4, pp. 123–131, 2022.
[12] T. Octaviany and A. Gunawan, “Mengoptimalkan Manajemen Persediaan Melalui Teknologi Rantai Pasokan,” vol. 01, no. 03, pp. 150–155, 2023.
[13] M. K. Al Fatach, “Perbandingan metode algoritma c4.5 dan naive bayes untuk memprediksi penjualan kosmetik pada toko jelita 1,2,” vol. 7, no. 2, pp. 220–225, 2024.
[14] M. F. Asshiddiqi, “1 , 2 1,” vol. 7, no. 3, pp. 9936–9948, 2020.
[15] S. A. Greysa, “Penerapan Teknologi AI dan Machine Learning dalam Manajemen TALENTA Conference Series Penerapan Teknologi AI dan Machine Learning dalam Manajemen Rantai Pasokan,” vol. 7, no. 1, 2024, doi: 10.32734/ee.v7i1.2303.
[16] M. P. Hilman Rifa’i1, Ryan Hamonangan2, Dian Ade Kurnia3 Kaslani4, “Implementasi Algoritma Decision Tree Dalam Klasifikasi Kompetensi Siswa,” vol. 06, no. 01, pp. 15–20, 2022.