Decision Support System for Determining the Feasibility of Material Utilization Using the Naïve Bayes Method

Aris Martono*[a,1], Padeli[b,2]

[a,b] Faculty of Science and Technology, UniversitasRaharja, Tangerang

e-mail: aris.martono@raharja.info*[1], padeli@raharja.info[2]

ABSTRACT

A company is engaged in the mechanical and electrical industry in providing goods and services, especially in medium and high power installations. In this company, the data management process carried out is still semi-computer, on a sheet of paper then copied into a computer with the Microsoft Excel application. Every work done by a technician in using material contains materials that are not used up or used materials. The material is then returned to the warehouse to be used as a report which is then carried out by the utilization process. The purpose of this research is to create a system that can assist in providing suggestions for used or used materials in the warehouse for utilization by applying the Naïve Bayes method. Naïve Bayes is a data mining technique with a classification function. This data mining application methodology uses the KDD (Knowledge Discovery in Database) stage starting from the selecting, preprocessing, transformation, data mining, and evaluation stages. The testing technique used for data validation uses the k-fold cross-validation technique accompanied by a Confusion Matrix. This research has produced a decision support system application and has been tested with an accuracy rate of 73.33%.

Keywords : Decision Support Systems, Materials, Warehouses, Naïve Bayes

1. INTRODUCTION

The development of information technology which is getting faster nowadays makes companies use this technology to produce a system capable of processing data in various ways to obtain quality information so that the work process can be more relevant, accurate, and timely in helping to improve performance and strengthen competitiveness company.

PT. Nusa Abadi Elektrika is a company engaged in the mechanical and electrical industry providing goods and services, especially in medium and high power installations. The first problem is that in this company the data management process carried out is still semi-computer, on a sheet of paper which is then copied by the officer into a computer with the Microsoft Excel application.

Every use of the material in doing work done by technicians usually contains materials that are not used up or used materials. The material will then be returned to the warehouse for recording which will be used as a material usage report then the utilization process is carried out.

Along with the large number of repair work, this creates a second problem in the part of the utilization process which runs slowly and has the potential to be inconsistent and often

miscommunication occurs, causing the accumulation of unused material items in the warehouse which makes the room narrow considering the limited space conditions.

With the existence of a decision support system to determine the feasibility of using the material using the Naive Bayes method, it will make it easier for officers at the warehouse to classify materials and provide suggestions on the feasibility of the material to be used.

The scope of this system is to act as decision support in the warehouse section in providing suggestions to determine the feasibility of used and unused materials in the warehouse of PT. Nusa Abadi Elektrika with decision-making analysis techniques in determining the feasibility of the material using the Naïve Bayes method. The purpose of this system is to provide suggestions for material utilization and maximize the use of materials in the work process. The benefits of this system are obtaining accurate and effective and efficient information in material management and minimizing errors in data collection that are still carried out conventionally.

In making this research, it is necessary to carry out a literature study to serve as a reference for existing studies, including:

a. Research conducted by DedenRustiana and Nina Rahayu (2017)at Raharja University with the title "Analysis of Car Automotive Market Sentiment: Twitter Tweets Using Naïve Bayes". This study analyzes data from Twitter with a search for the best-selling car brands based on data from GAIKINDO (Association of Indonesian Automotive Industries). By using three classes it can produce an accuracy value of 93%, and a positive precision level of 90%, a negative precision of 90%. and 100% neutral sentiment so that Naïve Bayes can analyze with high accuracy.[1]

b. Research conducted by BayuLesmana Putra (2016) at Raharja University with the title "Application of Analytical Hierarchy Process for Measurement Model of Employee Performance Development Plan as Decision Support System". This research is focused on helping human capital at PT Parador Management International in the process of developing employee performance plans using the Analytical Hierarchy Process (AHP) model which is expected to assist decision making in obtaining more objective information. [2]

c. Research conducted by DiasrinaDahri, et al (2017) at Mulawarman University with the title "The Naive Bayes Method for Determining Mulawarman University Bidikmisi Scholarship Recipients". This study aims to assist part of the selection process by making a decision support system software application for determining the recipients of the Mulawarman University Bidikmisi scholarship.[3]

d. Research conducted by PuspaEosina, et al. (2017) at Ibn Khaldun University Bogor with the title "Bayes Method in Determining Strategic Locations in the Housing Development Decision Support System". This study aims to optimize the selection of housing development locations against the predetermined criteria using the Bayes method in the Decision Support System (SPK) for Selecting a Strategic Location for Housing Development ".[4]

e. Research conducted by LitaAsyriati Latif, et al (2017) at Khairun University with the title "Application of Web-Based Decision Support Systems in Determining Tender Winners Using the Bayes Method and Group Technology". This study aims to develop a web-based decision support system application in determining tender winners using the Bayes method and group technology. The implementation of the Bayes criteria is by utilizing the weight of each category obtained based on the number of indicators used.[5]

f. Research conducted by EkaSugiyarti, et al (2018) from Malaysia in the International Journal of Pure and Applied Mathematics with the title "Decision Support System Of Scholarship Grantee Selection Using Data Mining". This study aims to select scholarship recipients where prospective recipients are prioritized to those who are less fortunate so that students can continue the selection study. Scholarship recipients will use data mining because processed variables have dynamic weights.[6]

g. Research conducted by VerryRiyanto, et al (2019) at STMIK Nusa Mandiri Jakarta in the Indonesian Journal of Artificial Intelligence and Data Mining (IJAIDM) with the title "Prediction of Student Graduation Time Using The Best Algorithm". This study aims to find a suitable architecture and significant knowledge to help and find out the most important factors in success between the several models tested with student graduation data.[7]

h. Research conducted by FirmanTempola, et al (2018) at Khairun Ternate University with the title "Comparison of Classification Between KNN and Naive Bayes on Determining Volcano Status with K-Fold Cross Validation". This study aims to compare two classification algorithms, namely the K-Nearest Neighbor and Naive Bayes Classifier on volcanic activity data in Indonesia.[8]

This research to implementing the naïve bayes method that measuring of feasibility utilization on decision support system.

2. METHOD

The model used in this study follows the stages in the KDD (Knowledge Discovery in Database) process. According to EforieBuulolo (2020: 1-5) [9], KDD is a process of discovery in a database or the entire process of extraction or identification of patterns, knowledge, and potential information from large data sets. Knowledge and information generated from KDD are valid, new, easy to understand, useful.

The stages of the Knowledge Discovery in Database process, see the figure 1 as bellow:
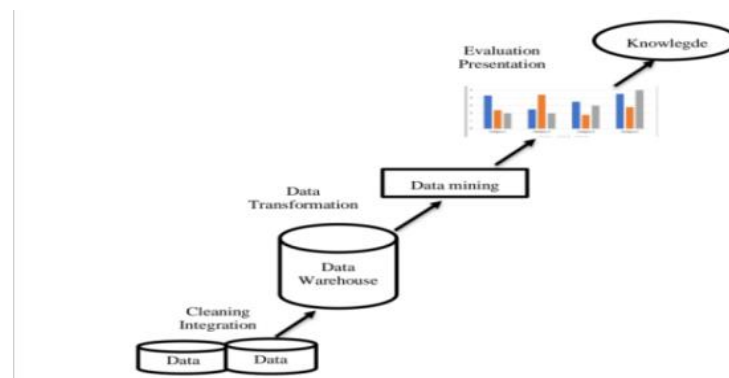

Figure 1. The Stages of the Knowledge Discovery in Database.

a. Selection

Not all existing data can be used. Therefore, the data selected. The data selection activity includes creating a data set on file storage.

b. Pre-processing / Cleaning

At this stage, the data that has been selected will be cleaned. The cleaning process includes removing duplicate data, correcting inconsistent data, and correcting data errors. In preprocessing/cleaning, the process of enriching data can also be carried out by adding other relevant information which is called enrichment.

c. Transformation

In data mining, there are so many algorithms/ methods/techniques that can be used. For the KDD process, it is changed first according to the algorithms/methods/techniques used in data mining.

d. Data Mining

Data Mining is the main stage in KDD. Data mining is the process of extracting and searching for useful knowledge and information using certain algorithms/methods/ techniques by the knowledge or information sought.

e. Interpretation / Evaluation

Knowledge or information generated from the data mining process will be presented or displayed in a form that is easily understood by interested parties such as information displayed in the form of graphics, decision trees, in the form of rules. The knowledge or information generated from the data mining process is examined whether or not it contradicts previous facts or hypotheses.

3. RESULT AND DISCUSSION

A. Selection

The source of the dataset used in this study is used material and used materials in the warehouse. The material comes from work residue under various conditions of damage. Meanwhile, the feasibility data comes from the warehouse head's decision that has been approved by the company leadership.

B. Pre-Processing Data

To find out the feasibility of each material, it is checked again manually in the warehouse after which it is evaluated with the leader for the next decision. As for some materials whose decisions have not been approved by the company leadership. Examples of data collection are arranged in table 1 Pre-processing Data.

C. Transformation

To make it easier for the next mining process, the data is transformed according to the type of data at the transformation stage, where the damage conditions will be grouped into categorical data. The following are the results of the transformed dataset in table 2 Transformation.

D. Data Mining

In data processing, the modeling stage is carried out for the classification process, namely applying the Naïve Bayes algorithm with Rapidminer software as shown in figure 2 below.
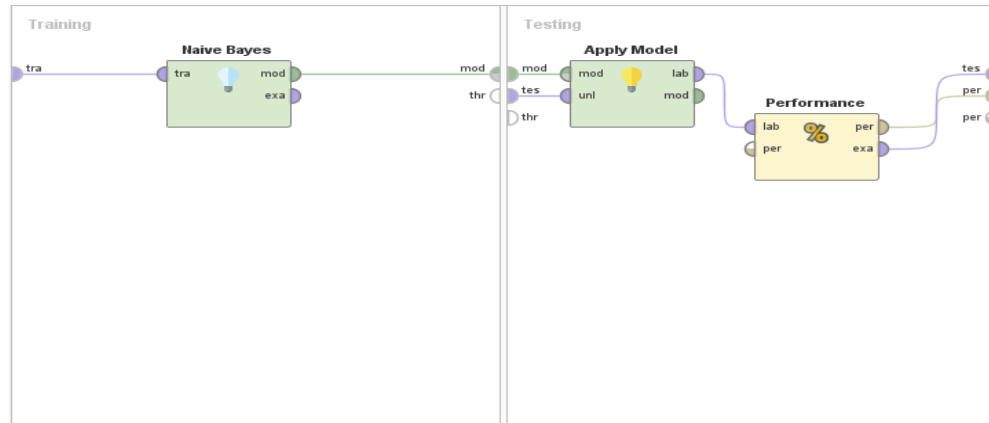
Figure. 2. The Naive Bayes Classification Model

This Naive Bayes calculation method is a method used in finding the probability where the calculation is carried out by the system. In calculating the naïve Bayes method, training data or system learning data are needed. The data amounted to 30 material data for calculations carried out by researchers. The following provides an example of the data to be tested (look at table 3 Data Testing).

Table 1. Data Testing

Data Testing / Data Uji	Damage	Handling	Budget
Box Panel 1,5x310x400x600mm (Bekas Pakai)	moderate	repaired	moderate

The steps taken from the Naïve Bayes method are as follows:

a. Counts the number of classes per label.

Table 2. Naive Bayes Label Results

feasible	13 / 30	0,43
not feasible	17 / 30	0,57

b. Calculating the Number of Cases Per Class can be looked at table 5, table 6, and table 7 as below:

Table 3. Damage Probability Results

P(Damage)	Moderate	
feasible	7 / 13	0,54
not feasible	5 / 17	0,29

Table 4. Handling Probability Results

P(Handling)	Repaired	
feasible	10 / 13	0,77
not feasible	5 / 17	0,29

Table 5. Budget Probability Results

P(Budget)	Moderate	
feasible	5 / 13	0,38
not feasible	5 / 17	0,29

c. Multiply all variables per class
    P (X | Eligible) = 0.54 * 0.77 * 0.38 = 0.159
    P (X | No) = 0.24 * 0.41 * 0.29 = 0.025

    Furthermore, the value of the above calculation is multiplied by the criteria for each category to get the final probability.

    P (X | Eligible) * P (Eligible) = 0.159 * 0.43 = 0.069
    P (X | No) * P (No) = 0.025 * 0.57 = 0.014

d. Compare results per class
    Eligible = 0.069
    No = 0.014

    So the final result of the probability value of the material is feasible to be used:

e. Evaluation
    Evaluation is carried out using the K-Fold Cross-validation technique (see Fig 3 below) with the value of k taken 5 fold so that from 30 data will be 6 data subsets with the same size, which is about 5 data. From each of the 6 subsets, 25 data becomes training data and 5 data becomes test data. The training and testing process model can be seen in figure 2 above.
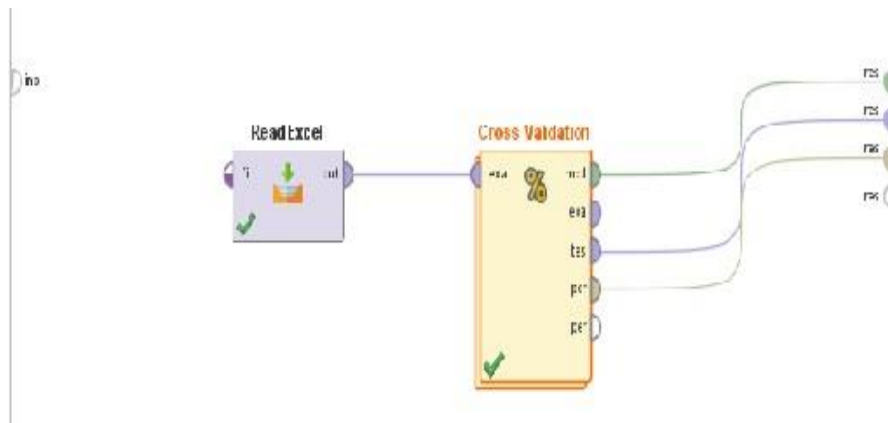
Figure 3. The K-Fold Cross-Validation Technique Model

The output of the performance results from this algorithm is a classification of materials that are included in the classification of the results "feasible" or "not" which is applied to the testing data. The amount of data that are predicted correctly by the Naïve Bayes algorithm is shown in table 8 Confusion Matrix:

Table 6. Confusion Matrix

Class Prediction	Feasible	Not feasible
Predictions worthy	9	4
Predictions are not feasible	4	13

Table 8 Confusion Matrix Naïve Bayes algorithm is:
  i.   The actual amount of DECENT and predicted DECENT data is 9
  ii.   The actual amount of data that DECIDES and predicts NOT is 4
  iii. The actual number of data that DID NOT and predictably WORTHY is 4
  iv.   The actual number of data that is NOT and predicted NOT is 13

So that the evaluation for this model uses the accuracy and error values.

The accuracy of the model is:

$$Accuracy = \frac{9 + 13}{9 + 4 + 4 + 13} x\ 100 = 73,33$$

$$Error = \frac{4 + 4}{9 + 4 + 4 + 13} x\ 100 = 26,67$$

The results of the implementation of the display interface for each menu of the decision support system for determining the feasibility of material used with the naïve Bayes method are as follows:

a. The Login Menu Display
    The login menu display (see figure 4) appears if you have logged in to the web browser. The user must enter a user name and password first.
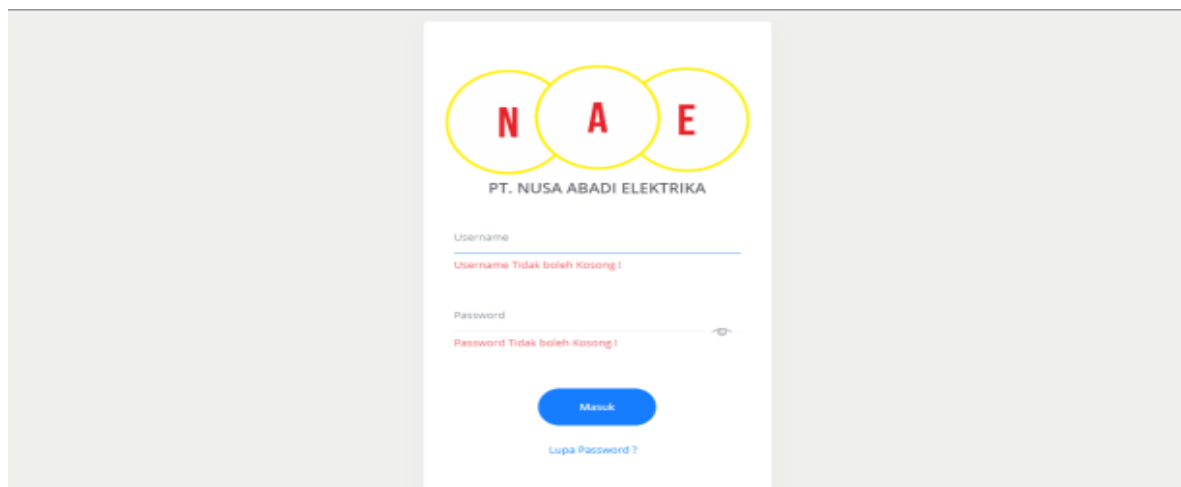
Figure 4. Login Page

b. Display Dashboard Menu

After a successful login, it will be directed to the Home or Dashboard menu (see figure 5) as the initial display of the system.
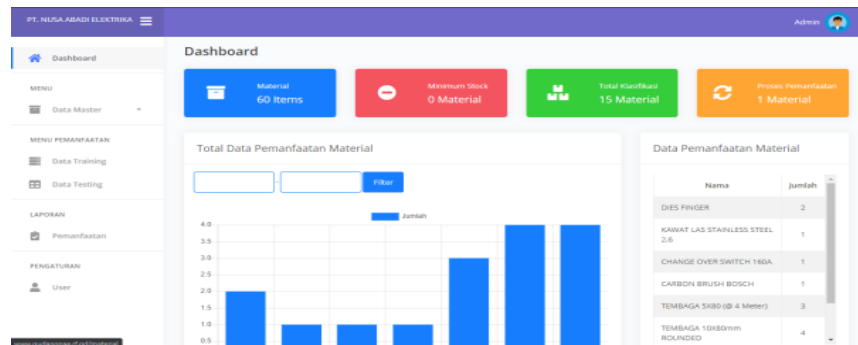

Figure 5. Dashboard Page

c. Display Data Training Menu

In this section of the menu (see figure 6) is a database or reference data that will be used as test material to classify testing data.

Figure 6. Display Data Training Menu

d. Display Data Testing Menu
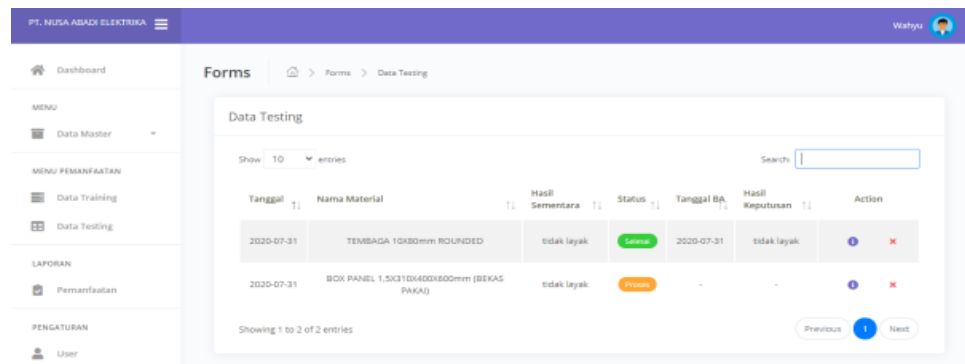
On the menu (see figure 7) where data testing is carried out using the naïve Bayes method.

Figure 7. Data Testing Menu Display

e. Report Menu Display
   This menu (see figure 8) is a report on the results of data testing that has been carried out on the data testing menu.

Figure 8. Report Menu Display

4. CONCLUSION

Based on the conclusions from the discussion topic of the Decision Support System for Determining the Feasibility of Material Utilization with the Naïve Bayes Method as follows:

1) Currently, the obstacles or problems faced by warehouse officers are in determining the appropriateness of used or used materials to be utilized.
2) Currently the process of determining the material feasibility of PT. Nusa Abadi Elektrika was carried out orally and saw the condition of the material with the naked eye and then recorded by the officer on a piece of paper which was later copied by the officer to the computer with the Microsoft Excel application.
3) With the creation of this decision support system, it is hoped that it can simplify the work process of warehouse officers in determining the feasibility of materials and minimize errors in managing material data.

REFERENCES

[1] Rustiana, D., & Rahayu, N. (2017). Automotive Market Sentiment Analysis: Twitter Tweet Using Naive Bayes. Simetris: Jurnal Teknik Mesin, Elektro dan Ilmu Komputer,, 113-120.

[2] Putra, B. L. (2016). Application of Analytical Hierarchy Process for Measurement Model of Employee Performance Development Plan as Decision Support System. Tangerang: Universitas Raharja.

[3] Dahri, D., Agus, F., & Khairina, D. M. (2016). The Naive Bayes Method for Determining Bidikmisi Scholarship Recipients. Jurnal Informatika Mulawarm, 29-36.

[4] Eosina, P., Lutfi, M., & Ridwan, M. (2017). Bayes Method in Determining Strategic Locations in a Housing Development Decision Support System. Journal Krea-TIF, 47-57.

[5] Latif, L. A., Abbas, S. H., & Jamil, M. (2017). Web-Based Decision Support System Application Development in Determining Tender Winners Using the Bayes Method and Group Technology. Jurnal Penelitian Pos dan Informatika, 73-82.

[6] Sugiyarti, E., Jasmi, K. A., & Basiron, B. (2018). Decision Support System Of Scholarship Grantee Selection Using Data Mining. International Journal of Pure and Applied Mathematics, 2239-2249.

[7] Riyanto, V., Hamid, A., & Ridwansyah, R. (2019). Prediction of Student Graduation Time Using The Best Algorithm. Indonesian Journal of Artificial Intelligence and Data Mining (IJAIDM), 1-9.

[8] Tempola, F., Muhammad, M., & Khairan, A. (2018). Comparison of Classification Between KNN and Naive Bayes in Determining Volcano Status with K-Fold Cross-Validation. Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK), 577-584.

[9] Buulolo, E. (2020). Data Mining For Higher Education. Yogyakarta: Deepublish.

[10] Hidayat, W. .(2020). Application of Material Feasibility Utilization System in Electrical Industry Companies. Tangerang : University of Raharja