CoreID JOURNAL

Vol. 1 No. 2, July 2023 (75-82)
ISSN: 2987-6990
DOI: https://doi.org/10.60005/coreid.v1i1.10

Classification of Non-Civil Servant Performance Appraisal Using Naïve Bayes Classifier Algorithm

Sofia Dewi
[1,2] Information System Department, Ma'soem University, Indonesia

Article Info

Article history:

Received April 17, 2023
Revised May 06, 2023
Accepted July 25, 2023

Keywords:

Classification
Naïve Bayes Classifier
Performance Assessment
Rapid Miner

ABSTRACT

Employee performance assessment is a way to measure the level of employee productivity. In the process of assessing the performance of Non-Civil Servants (non-PNS) employees at the Regional Technical Implementation Unit of Education and Training of Cooperatives and Entrepreneurs (UPTD P3W) at this time, it is required to classify data based on several factors to find out whether the employee fits into the eligible category or not as the best employee to become a civil servant (PNS) candidate. The purpose of this research is to make it easier to determine the classification of the performance assessment of non-PNS employees at UPTD P3W using the Naïve Bayes Classifier Algorithm and to determine the level of accuracy in the classification of the performance assessment. In this study, the authors used 498 data as training data and 105 data as testing data for manual testing in Excel and for testing using RapidMiner tools. Based on the analysis in the study, the result of the predictions determines the best employees to become candidates for civil servants quickly and accurately, while from the tests performed by comparing training data and with data testing using RapidMiner tools, the accuracy rate is 84.76%.

Corresponding Author:

Sofia Dewi
Department of Information Systems, Ma'soem University, Indonesia
Jl. Raya Cipacing No.22, Cipacing, Kec. Jatinangor, Kabupaten Sumedang, Jawa Barat 45363
Email: sofiadewi.job@gmail.com

1. INTRODUCTION

In a company, both government and private companies, employees have a very important role in the company [1], [2]. In this very important role, employee performance needs to be managed properly [3], [4]. Employee performance needs to be measured to determine the maximum level of employee professionalism that has been achieved. Employee performance assessments like this have been performed by various organizations in the past.

Employee performance assessment is an effort to measure the level of employee productivity [1], [5]–[7]. Criteria that greatly influence performance assessment include length of service, age, gender, division, status, education, type of work, work behavior, personality aspects, technical work aspects, non-technical aspects, and goal attainment [8], [9]. The assessment standard that is applied is the attitude of employees and superiors when facing this assessment.

The Regional Technical Implementation Unit of Education and Training of Cooperatives and Entrepreneurs (UPTD P3W) is a government institution under the auspices of the Office of Cooperatives and Small Enterprises (KUK) - West Java Province [10], [11], located in Jl. Soekarno-Hatta No.708, Babakan Penghulu, Cinambo, Bandung City, West Java 40924. There are 55 employees at UPTD P3W, 24 civil servants (PNS), and 31 non-civil servants (non-PNS). 31 non-PNS employees are aged between 25 years and over 50 years with a working period of between 3 months and 14 years. Then, they will be classified as civil servant candidates, with a total of 498 data taken for data testing of 105 data. The problems in this research are:

a) In the process of assessing the performance of non-PNS employees currently at the UPTD, they still use manuals and already have assessment indicators and then it is necessary to classify data based on factors to determine whether the employee fits into the appropriate "eligible" or inappropriate "not eligible" category as a candidate for civil servants.
b) To find out the level of accuracy in classifying employee performance assessments.

2. METHOD

The method used is the Naïve Bayes Classifier. Naïve Bayes Classifier is one of the statistical classifiers [12]–[17], where this classifier can predict the probability of class membership of a data tuple that will enter a certain class according to probability calculations. The data mining method for classifying employee data is performed in several stages. The flow of the classification process can be seen in figure 1:

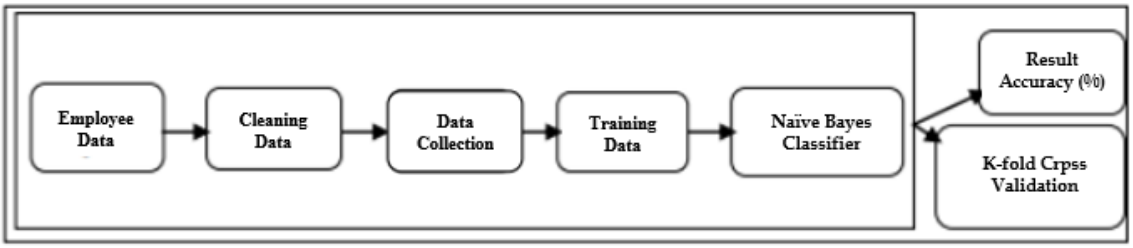

Figure 1. Process of Naïve Bayes Classifier Method

The flow of data mining using the Naïve Bayes Classifier method is shown in figure 2, which consists of several stages, including:

2.1. Data Collection

The data used in this research is using primary data. Primary data is the source of research data obtained directly from the source in the form of interviews. The data collected is data on daily reports on the activities of non-PNS employees at UPTD P3W - West Java, with a total of 498 data.

2.2. Preliminary Data Processing

Data preparation is taking data identified in the previous stage or preparing it for analysis using data mining methods. This is expected to be compared with other stages in CRISP-DM [18]–[20].

a) First, determining the data to be processed from the data that has been obtained, where not all data will be processed because the research to be carried out has limitations on the data used.
b) Second, handling missing value data. The missing value is incomplete data because the attribute is not recorded or the attribute is not owned. Handling of missing values is performed by deleting empty records.
c) Third, determining the attributes that will be used from the first stage.

2.3. The Proposed Method

After performing all stages of data processing, training data will then be generated. Training data is the data used to perform the calculation process [21], [22] using the Naïve Bayes Classifier method.

2.4. Final Testing/Validation

The test results will be validated and then evaluated using the RapidMiner software. RapidMiner is open-source software. RapidMiner is a solution for analyzing data mining, text mining, and predictive analysis. RapidMiner has approximately 500 data mining operators, including input, output, data preprocessing, and visualization operators.

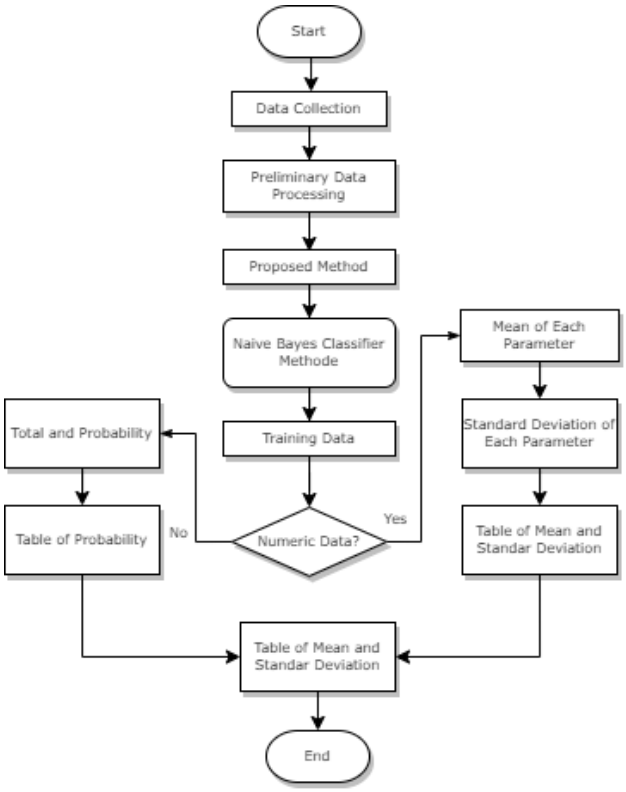

Figure 2. Work Procedure

3. RESULTS AND DISCUSSION

3.1. RESULTS

3.1.1. Initial Data

Data processing requires supporting data to determine whether the employee will be eligible or not as a candidate for civil servant employees. This data is used as a reference to determine the assessment of each employee, including:

a) Performance Assessment Support
b) Work Behavior
c) Personality Aspects
d) Technical Work Aspects
e) Non-Technical Aspects
f) Goal Attainment

Examples of data used in this study are data on non-civil servants at UPTD P3W West Java as shown in Figure 3.

No	Name	Official Positor	Date	Day	In (AM)	Out (PM)	Range	Activity	Information	Criteria
1	Muh. Rizky Cahyadi	Techincal Staff	04/01/2023	Monday	7.30	4.30	2	filled	approved	eligible
2	Muh. Rizky Cahyadi	Techincal Staff	05/01/2023	Tuesday	7.30	5.00	3	filled	approved	eligible
3	Muh. Rizky Cahyadi	Techincal Staff	06/01/2023	Wednesday	7.30	5.00	3	filled	approved	eligible
4	Muh. Rizky Cahyadi	Techincal Staff	07/01/2023	Thursday	7.30	5.00	3	filled	approved	eligible
5	Muh. Rizky Cahyadi	Techincal Staff	08/01/2023	Friday	7.30	6.00	3	filled	approved	eligible
6	Muh. Rizky Cahyadi	Techincal Staff	11/01/2023	Monday	7.30	5.00	3	filled	approved	eligible
7	Muh. Rizky Cahyadi	Techincal Staff	12/01/2023	Tuesday	7.30	5.00	3	filled	approved	eligible
8	Muh. Rizky Cahyadi	Techincal Staff	13/01/2023	Wednesday	7.30	5.00	3	filled	approved	eligible
9	Muh. Rizky Cahyadi	Techincal Staff	14/01/2023	Thursday	7.30	5.00	3	filled	approved	eligible
10	Muh. Rizky Cahyadi	Techincal Staff	15/01/2023	Friday	7.30	5.00	3	filled	approved	eligible
11	Muh. Rizky Cahyadi	Techincal Staff	18/01/2023	Monday	7.30	5.00	3	filled	approved	eligible
12	Muh. Rizky Cahyadi	Techincal Staff	19/01/2023	Tuesday	7.30	5.00	3	filled	approved	eligible
13	Muh. Rizky Cahyadi	Techincal Staff	20/01/2023	Wednesday	7.30	5.00	3	filled	approved	eligible
14	Muh. Rizky Cahyadi	Techincal Staff	21/01/2023	Thursday	7.30	5.00	3	filled	approved	eligible
15	Muh. Rizky Cahyadi	Techincal Staff	22/01/2023	Friday	7.30	5.00	3	filled	approved	eligible
16	Muh. Rizky Cahyadi	Techincal Staff	25/01/2023	Monday	7.30	5.00	3	filled	approved	eligible
17	Muh. Rizky Cahyadi	Techincal Staff	01/02/2023	Monday	7.30	5.00	3	filled	approved	eligible
18	Muh. Rizky Cahyadi	Techincal Staff	02/02/2023	Tuesday	7.30	5.00	3	filled	approved	eligible
19	Muh. Rizky Cahyadi	Techincal Staff	03/02/2023	Wednesday	7.30	5.00	3	filled	approved	eligible
20	Muh. Rizky Cahyadi	Techincal Staff	04/02/2023	Thursday	7.30	5.00	3	filled	approved	eligible
21	Muh. Rizky Cahyadi	Techincal Staff	05/02/2023	Friday	7.30	5.00	3	filled	approved	eligible
22	Muh. Rizky Cahyadi	Techincal Staff	08/02/2023	Monday	7.30	5.00	3	filled	approved	eligible
23	Muh. Rizky Cahyadi	Techincal Staff	09/02/2023	Tuesday	7.30	5.00	3	filled	approved	eligible
24	Muh. Rizky Cahyadi	Techincal Staff	10/02/2023	Wednesday	7.30	5.00	3	filled	approved	eligible

Figure 3. Preliminary Data

3.1.2. Cleaning Data

Cleaning data is performed to remove noise and inconsistent data or irrelevant data, as shown in Figure 4.

No	Day	In (AM)	Out (PM)	Range	Activity	Information	Criteria
1	Monday	7.30	4.30	2	filled	approved	eligible
2	Tuesday	7.30	5.00	3	filled	approved	eligible
3	Wednesday	7.30	5.00	3	filled	approved	eligible
4	Thursday	7.30	5.00	3	filled	approved	eligible
5	Friday	7.30	6.00	3	filled	approved	eligible
6	Monday	7.30	5.00	3	filled	approved	eligible
7	Tuesday	7.30	5.00	3	filled	approved	eligible
8	Wednesday	7.30	5.00	3	filled	approved	eligible
9	Thursday	7.30	5.00	3	filled	approved	eligible
10	Friday	7.30	5.00	3	filled	approved	eligible
11	Monday	7.30	5.00	3	filled	approved	eligible
12	Tuesday	7.30	5.00	3	filled	approved	eligible
13	Wednesday	7.30	5.00	3	filled	approved	eligible
14	Thursday	7.30	5.00	3	filled	approved	eligible
15	Friday	7.30	5.00	3	filled	approved	eligible
16	Monday	7.30	5.00	3	filled	approved	eligible
17	Monday	7.30	5.00	3	filled	approved	eligible
18	Tuesday	7.30	5.00	3	filled	approved	eligible
19	Wednesday	7.30	5.00	3	filled	approved	eligible
20	Thursday	7.30	5.00	3	filled	approved	eligible
21	Friday	7.30	5.00	3	filled	approved	eligible
22	Monday	7.30	5.00	3	filled	approved	eligible
23	Tuesday	7.30	5.00	3	filled	approved	eligible
24	Wednesday	7.30	5.00	3	filled	approved	eligible

Figure 4. Cleaning Data

3.1.3. Data Collection

Data collection is conducted to combine data from other sources. is shown in Figure 5.

No	Gende	Age (year	Status	Education	Work Status	ngth of Serv	Division	ork Behav	rsonality	Asphical Work A	Technical A	pal Attainme	Activity	Information	Criteria
1	Male	36-40	Married	Senior High School	Under Contract	2 years	Park Staff	5	5	5	5	5	filled	approved	eligible
2	Male	25-30	Not married	Senior High School	Under Contract	2 years	Cleaning Staff	5	5	5	5	5	filled	approved	eligible
3	Male	25-30	Married	Undergraduate	Under Contract	7 years	Cleaning Staff	5	5	5	5	5	filled	approved	eligible
4	Male	41-45	Married	Senior High School	Temporary Worker	14 years	Administration Staff	5	5	5	4	5	filled	approved	eligible
5	Male	25-30	Married	Undergraduate	Under Contract	1 year	Administration Staff	4	2	4	3	4	filled	approved	eligible
6	Male	31-35	Not married	Undergraduate	Under Contract	7 years	Administration Staff	4	4	4	4	4	filled	approved	eligible
7	Male	31-35	Married	Undergraduate	Under Contract	8 years	Administration Staff	5	4	5	4	4	filled	approved	eligible
8	Male	25-30	Not married	Undergraduate	Under Contract	1 year	Administration Staff	4	4	5	4	4	filled	approved	eligible
9	Male	31-35	Not married	Senior High School	Temporary worker	14 years	Driver	4	5	5	5	5	filled	approved	eligible
10	Male	25-30	Not married	Undergraduate	Under Contract	1 year	Cleaning Staff	5	5	5	5	5	not filled	not approved	eligible
11	Male	25-31	Not married	Senior High School	Under Contract	1 year	Front Office Staff	5	5	4	4	4	filled	approved	eligible
12	Male	25-32	Not married	Senior High School	Under Contract	1 year	Front Office Staff	4	5	4	4	5	filled	approved	eligible
13	Male	25-33	Not married	Senior High School	Under Contract	3 months	Front Office Staff	5	5	5	5	5	filled	approved	eligible
14	Male	41-45	Divorced	Senior High School	Under Contract	8 years	Park Staff	5	4	4	5	4	filled	approved	eligible
15	Male	31-35	Not married	Senior High School	Under Contract	14 years	Cleaning Staff	5	2	2	2	2	filled	approved	eligible
16	Male	41-45	Not married	Senior High School	Under Contract	4 years	Cleaning Staff	2	3	3	3	3	filled	approved	not eligible
17	Male	36-40	Married	Senior High School	Under Contract	14 years	Cleaning Staff	4	5	3	5	4	filled	approved	eligible
18	Male	36-40	Married	Senior High School	Under Contract	14 years	Cleaning Staff	5	4	4	4	4	filled	approved	eligible
19	Male	41-45	Married	Senior High School	Temporary Worker	10 years	Cleaning Staff	4	5	4	3	5	filled	approved	eligible
20	Male	31-35	Married	Undergraduate	Under Contract	3 years	Cleaning Staff	4	3	4	3	3	filled	approved	eligible
21	Male	36-40	Married	Senior High School	Under Contract	14 years	Security Guard	4	5	5	5	5	filled	approved	eligible
22	Male	41-45	Married	Senior High School	Under Contract	14 years	Security Guard	4	5	4	4	4	filled	approved	eligible
23	Male	over 50	Married	Senior High School	Under Contract	9 years	Security Guard	4	5	5	4	4	filled	approved	eligible
24	Male	36-40	Married	Senior High School	Under Contract	7 years	Security Guard	5	5	5	4	4	filled	approved	eligible
25	Male	36-41	Married	Senior High School	Under Contract	2 years	Security Guard	5	4	4	4	5	filled	approved	eligible
26	Male	36-42	Married	Senior High School	Under Contract	7 years	Security Guard	5	5	5	5	4	filled	approved	eligible
27	Male	25-30	Not married	Undergraduate	Under Contract	6 years	Administration Staff	5	5	5	5	5	filled	approved	eligible
28	Male	41-45	Married	Senior High School	Under Contract	7 years	Administration Staff	5	4	4	5	5	filled	approved	eligible
29	Male	41-45	Married	Senior High School	Under Contract	12 years	Administration Staff	5	5	4	4	4	filled	approved	eligible

Figure 5. Data Collection

3.1.4. Training Data

This training data uses data from a combined data of 498 data. Figure 6 below is a sample of 26 data.

No	Gende	Age (year	Status	Education	Work Status	ngth of Serv	Division	ork Behav	rsonality	Asphical Work A	Technical A	pal Attainme	Activity	Information	Criteria
1	Male	36-40	Married	Senior High School	Under Contract	2 years	Park Staff	5	5	5	5	5	not filled	not approved	not eligible
2	Male	25-30	Not married	Senior High School	Under Contract	2 years	Cleaning Staff	5	5	5	5	5	not filled	not approved	not eligible
3	Male	25-30	Married	Undergraduate	Under Contract	7 years	Cleaning Staff	5	5	5	5	5	not filled	not approved	not eligible
4	Male	41-45	Married	Senior High School	Temporary Worker	14 years	Administration Staff	5	5	5	4	5	filled	approved	eligible
5	Male	25-30	Married	Undergraduate	Under Contract	1 year	Administration Staff	4	2	4	3	4	filled	approved	eligible
6	Male	31-35	Not married	Undergraduate	Under Contract	7 years	Administration Staff	4	4	4	4	4	filled	approved	eligible
7	Male	31-35	Married	Undergraduate	Under Contract	8 years	Administration Staff	5	4	5	4	4	filled	approved	eligible
8	Male	25-30	Not married	Undergraduate	Under Contract	1 year	Administration Staff	4	4	5	4	4	filled	approved	eligible
9	Male	31-35	Not married	Senior High School	Temporary worker	14 years	Driver	4	5	5	5	5	not filled	not approved	not eligible
10	Male	25-30	Not married	Undergraduate	Under Contract	1 year	Cleaning Staff	5	5	5	5	5	not filled	not approved	not eligible
11	Male	25-31	Not married	Senior High School	Under Contract	1 year	Front Office Staff	5	5	4	4	4	not filled	not approved	not eligible
12	Male	25-32	Not married	Senior High School	Under Contract	1 year	Front Office Staff	4	5	4	4	5	not filled	not approved	not eligible
13	Male	25-33	Not married	Senior High School	Under Contract	3 months	Front Office Staff	5	5	5	5	5	not filled	not approved	not eligible
14	Male	41-45	Divorced	Senior High School	Under Contract	8 years	Park Staff	5	4	4	5	4	not filled	not approved	not eligible
15	Male	31-35	Not married	Senior High School	Under Contract	14 years	Cleaning Staff	5	2	2	2	2	not filled	approved	eligible
16	Male	41-45	Not married	Senior High School	Under Contract	4 years	Cleaning Staff	2	3	3	3	3	not filled	not approved	not eligible
17	Male	36-40	Married	Senior High School	Under Contract	14 years	Cleaning Staff	4	5	3	5	4	not filled	not approved	not eligible
18	Male	36-40	Married	Senior High School	Under Contract	14 years	Cleaning Staff	5	4	4	4	4	not filled	not approved	not eligible
19	Male	41-45	Married	Senior High School	Temporary Worker	10 years	Cleaning Staff	4	5	4	3	5	not filled	not approved	not eligible
20	Male	31-35	Married	Undergraduate	Under Contract	3 years	Cleaning Staff	4	3	4	3	3	not filled	not approved	not eligible
21	Male	36-40	Married	Senior High School	Under Contract	14 years	Security Guard	4	5	5	5	5	not filled	not approved	not eligible
22	Male	41-45	Married	Senior High School	Under Contract	14 years	Security Guard	4	5	4	4	4	not filled	not approved	not eligible
23	Male	over 50	Married	Senior High School	Under Contract	9 years	Security Guard	4	5	5	4	4	not filled	not approved	not eligible
24	Male	36-40	Married	Senior High School	Under Contract	7 years	Security Guard	5	5	5	4	4	not filled	not approved	not eligible
25	Male	36-41	Married	Senior High School	Under Contract	2 years	Security Guard	5	4	4	4	5	not filled	not approved	not eligible
26	Male	36-42	Married	Senior High School	Under Contract	7 years	Security Guard	5	5	5	5	4	not filled	not approved	not eligible
27	Male	25-30	Not married	Undergraduate	Under Contract	6 years	Administration Staff	5	5	5	5	5	filled	approved	eligible
28	Male	41-45	Married	Senior High School	Under Contract	7 years	Administration Staff	5	4	4	5	5	not filled	not approved	not eligible
29	Male	41-45	Married	Senior High School	Under Contract	12 years	Administration Staff	5	5	4	4	4	not filled	not approved	not eligible

Figure 6. Training Data

3.1.5. Testing Data

This testing data uses 105 data. Figure 7 below is the sample of 30 data.

No	Gende	Age (year	Status	Education	Work Status	ngth of Serv	Division	ork Behav	rsonality	Asphical Work A	Technical A	pal Attainme	Activity	Information	Criteria
1	Male	31-35	Not married	Senior High School	Under Contract	14 years	Cleaning Staff	2	2	2	2	2	not filled	not approved	not eligible
2	Male	41-45	Not married	Senior High School	Under Contract	4 years	Cleaning Staff	4	3	3	3	3	not filled	not approved	not eligible
3	Male	36-40	Married	Senior High School	Under Contract	14 years	Cleaning Staff	5	5	3	5	4	not filled	not approved	not eligible
4	Male	36-40	Married	Senior High School	Under Contract	14 years	Cleaning Staff	4	4	4	4	4	not filled	not approved	not eligible
5	Male	41-45	Married	Senior High School	Temporary Worker	10 years	Cleaning Staff	4	5	4	3	5	not filled	not approved	not eligible
6	Male	31-35	Married	Undergraduate	Under Contract	3 years	Cleaning Staff	4	2	4	3	3	not filled	not approved	not eligible
7	Male	36-40	Married	Senior High School	Under Contract	14 years	Security Guard	4	5	5	5	5	not filled	not approved	not eligible
8	Male	41-45	Married	Senior High School	Under Contract	14 years	Security Guard	5	5	4	4	4	filled	approved	not eligible
9	Male	over 50	Married	Senior High School	Under Contract	9 years	Security Guard	5	5	5	4	4	filled	approved	eligible
10	Male	36-40	Married	Senior High School	Under Contract	7 years	Security Guard	5	5	5	4	4	filled	approved	eligible
11	Male	36-40	Married	Senior High School	Under Contract	2 years	Security Guard	5	4	4	4	5	filled	approved	eligible
12	Male	36-40	Married	Senior High School	Under Contract	7 years	Administration Staff	5	5	5	5	4	filled	approved	eligible
13	Male	41-45	Married	Senior High School	Under Contract	7 years	Administration Staff	5	4	4	5	5	filled	approved	eligible
14	Male	41-45	Married	Senior High School	Under Contract	12 years	Administration Staff	5	5	4	4	4	filled	approved	eligible
15	Male	41-45	Married	Senior High School	Under Contract	5 years	Administration Staff	5	5	5	4	5	filled	approved	eligible
16	Male	25-30	Not married	Senior High School	Under Contract	2 years	Administration Staff	5	5	5	5	4	filled	approved	eligible
17	Male	36-40	Married	Senior High School	Under Contract	2 years	Administration Staff	5	5	5	5	5	filled	approved	eligible
18	Male	25-30	Not married	Undergraduate	Under Contract	2 years	Administration Staff	5	5	5	5	5	filled	approved	eligible
19	Male	25-30	Married	Senior High School	Under Contract	7 years	Administration Staff	4	5	5	4	5	filled	approved	eligible
20	Male	41-45	Married	Senior High School	Temporary Worker	14 years	Administration Staff	4	5	5	4	5	filled	approved	eligible
21	Male	25-30	Not married	Undergraduate	Under Contract	1 year	Administration Staff	5	4	5	5	5	filled	approved	eligible
22	Male	31-35	Married	Senior High School	Under Contract	14 years	Administration Staff	5	5	5	5	4	filled	approved	eligible
23	Male	25-30	Not married	Undergraduate	Under Contract	1 year	Administration Staff	4	5	4	4	5	filled	approved	eligible
24	Male	25-31	Not married	Senior High School	Under Contract	1 year	Administration Staff	5	5	4	4	5	filled	approved	eligible
25	Male	25-32	Not married	Senior High School	Under Contract	1 year	Administration Staff	5	5	5	5	4	filled	approved	eligible
26	Male	25-33	Not married	Senior High School	Under Contract	3 months	Administration Staff	5	5	4	5	2	filled	approved	eligible
27	Male	41-45	Married	Senior High School	Under Contract	8 years	Administration Staff	2	4	4	2	4	filled	approved	eligible
28	Male	31-35	Not married	Senior High School	Under Contract	14 years	Administration Staff	4	2	2	3	2	filled	approved	not eligible
29	Male	41-45	Not married	Senior High School	Under Contract	4 years	Administration Staff	5	3	3	5	3	filled	approved	eligible
30	Male	36-40	Married	Senior High School	Under Contract	14 years	Administration Staff	4	5	3	4	4	filled	approved	eligible
31	Male	3640	Married	Senior High School	Under Contract	14 years	Administration Staff	4	4	4	4	4	filled	approved	eligible

Figure 7. Testing Data

3.2. Calculation of Naïve Bayes

$$P(Ci|X) = \frac{P(X|Ci).P(Ci)}{P(X)} \quad (1)$$

$$Ci = \frac{P(X|Ci)P(Ci)}{P(X)} \quad (2)$$

Where [23]–[25]:

X = Data with *unknown* classes
Ci = The X data hypothesis is a *specific* class
P(Ci|X) = is the posterior probability of classy given input features X (*posteriori prob*)
P(Ci) = Probability of hypothesis Ci (*prior prob*)
P(X|Ci) = is the likelihood of observing the input features X given class Ci
P(X) = is the marginal likelihood of the input features X (*tuple* X has the same probability of entering any class, so the maximized is P(X|Ci)P(Ci)

The first stage of calculation to determine the best non-PNS employee to become the PNS candidate with the Naïve Bayes method is to find the probability of each class. In determining the best employee, 2 classes will be determined: the "Eligible" and "Not Eligible" categories. The calculation method is to find out how much data is in the "Eligible" class and how much is in the "Not Eligible" class, from the total training data, then divide it by the total data. The results of the calculations can be seen in the figure 8 below:

- P(Ci)
- P(Eligible) = 437/498 = 0.88
  The total of "Eligible" data in the "Criteria" column is divided by the amount of data.
- P(Not Eligible) = 61/498 = 0.12
  The total of "Not Eligible" data in the "Criteria" column is divided by the amount of data.

Figure 8. Class Probability

3.3. Manual Calculation of Naïve Bayes

a) Calculation result (Example Set)

Sample 6 of data. As shown in Table 1 below:

Table 1. Calculation Result (Example Set)

Row No.	Criteria	Prediction (Criteria)	Confidence (Not Eligible)	Confidence (Eligible)
1	Not Eligible	Not Eligible	1.000	0.000
2	Not Eligible	Not Eligible	0.998	0.002

Row No.	Criteria	Prediction (Criteria)	Confidence (Not Eligible)	Confidence (Eligible)
3	Eligible	Eligible	0.016	0.984
4	Eligible	Eligible	0.008	0.992
5	Eligible	Eligible	0.009	0.991
6	Eligible	Eligible	0.031	0.991

b) The calculation result of Accuracy

From the results of Figure 4.11, it can be seen that the level of accuracy of the data used is 84.76% and the class precision "Eligible" is 84.72%, and "Not Eligible" is 84.85%. As for the class recall itself, for the "Eligible" class it is 92.42% and for the "Not Eligible" class it is 71.79%. In general, precision, recall, and accuracy can be formulated as follows:

i. For “Eligible” class

Precision = 61/(61+11) = 61/72 = 0.8472 = 84.72%
Recall = 61/(61+5) = 61/66 = 0.9242 = 92.42%

ii. For “Not Eligible” class

Precision = 28/(28+5) = 28/33 = 0.8485 = 84.85%
Recall = 28/(28+11) = 28/39 = 0.7179 = 71.79%
Accuracy = (61+28)/(61+5+11+28) = 89/105 = 0.8476 = 84.76%
Error = (5+11)/(61+5+11+28) = 16/105 = 0.1524 = 15.24%

3.4. Comparison of Accuracy

a) Comparison of accuracy

A comparison of overall accuracy with validation accuracy using Cross Validation is shown in Table 2.

Table 2. Comparison of accuracy

	Accuracy	*Precision*	*Recall*
Accuracy %	84.76%	84.72%	92.42%
Cross Validation	80.00%	83.82%	84.52%
	+/- 14.13%	+/-13,97%	+/- 17.14%

b) Percentage of “Eligible” and “Not Eligible”

The percentage of Eligible and Not Eligible obtained from the results of testing 150 data with the Naïve Bayes Classifier, is shown in table 3.

Table 3. Percentage of Eligible and Not Eligible

	Eligible	Not Eligible
Percentage	63%	37%

4. CONCLUSION

Based on the results of research that has been conducted on the classification of non-PNS performance assessments using the Naïve Bayes Classifier Algorithm method at UPTD P3W - West Java Province, it can be concluded as follows:

a) This performance assessment has 15 indicators, such as gender, age, status, education, employment, length of service, division, work behavior, personality aspects, technical work aspects, non-technical aspects, goal attainment, activities, information, and criteria. The classification process using the Naïve Bayes method uses training data to generate probabilities for each criterion in different classes. These probability values can be optimized to determine the best employee to become a civil servant (PNS) candidate at UPTD P3W - West Java Province.

b) Based on the results of the analysis in the research, it is found that the prediction results determine the best employees become candidates for civil servants quickly and accurately. From the tests performed by comparing training data with testing data using RapidMiner supporting tools, it is obtained that the accuracy rate is 84.76%.

REFERENCES

[1] M. P. Sari, "The Influence of Leadership on Work Discipline: A Perceptual Study of Employees at the Regional People's Credit Bank LPK Jalancagak Subang Company," Indonesia University of Education, 2015.
[2] W. W. Sari and B. Arianto, "Kinerja Karyawan PT. Loka Refractories Wira Jatim Ditinjau dari Stres Kerja dan Lingkungan Kerja," *Mak. J. Manaj.*, vol. 8, no. 1, pp. 45–55, Jul. 2022.
[3] W. Widodo, "PENGARUH KOMPENSASI DAN MOTIVASI TERHADAP KINERJA KARYAWAN SEKRETARIAT NASIONAL PEMBERDAYAAN PEREMPUAN KEPALA KELUARGA," *J. Manaj. Bisnis Krisnadwipayana*, 2017.
[4] Y. Mulyani, "Penerapan Absensi Online Berbasis Android pada Peningkatan Kedisiplinan dan Kinerja Guru Pegawai Negeri Sipil pada Bidang PAI," *J. Educ. FKIP UNMA*, vol. 6, no. 1, pp. 205–208, Jul. 2020.
[5] P. W. Kastawan, D. M. Wiharta, and M. Sudarma, "Implementasi Algoritma C5.0 pada Penilaian Kinerja Pegawai Negeri Sipil," *Maj. Ilm. Teknol. Elektro*, vol. 17, no. 3, p. 371, 2018.
[6] M. R. Hamdalah, K. Muhammad, and telma anis safitri, "Analisis Kinerja Perusahaan Sebagai Upaya Meningkatkan Kinerja Perusahaan Menggunakan Metode Balanced Scorecard (Studi Kasus: PT. XYZ)," *J. Pendidik. dan Teknol. Indones.*, vol. 1, no. 1, pp. 27–33, Jan. 2021.
[7] M. E. Widyaningrum and M. E. (Design & L. Widiana, "EVALUASI KINERJA:Untuk Meningkatkan Produktifitas Karyawan Perusahaan," 2020.
[8] S. dan R. Damayanti, "KEPUASAN KERJA DAN KINERJA KARYAWAN Tinjauan dari Dimensi Iklim Organisasi, Kreativitas Individu, dan Karakteristik Pekerjaan," *Yogyakarta: Indomedia Pustaka*, pp. 1–85, Feb. 2017.
[9] A. Utaminingsih, "Perilaku Organisasi: Kajian Teoritik & Empirik terhadap Budaya Organisasi ... - Alifiulahtin Utaminingsih - Google Books," *UB Press*, 2014. [Online]. Available: https://books.google.co.id/books?hl=en&lr=&id=-wVRDwAAQBAJ&oi=fnd&pg=PA119&dq=Kriteria+yang+sangat+mempengaruhi+penilaian+kinerja+antara+lain:+masa+kerja,+usia,+jenis+kelamin,+divisi,+status,+pendidikan,+pekerjaan,+perilaku+kerja,+aspek+kepribadian,+aspek. [Accessed: 08-Aug-2022].
[10] A. D. Y. Umaran, "Kampanye PR dalam pengembangan wirausaha: Studi terhadap program wirausaha baru (WUB) UPTD P3W Dinas Koperasi dan Usaha Kecil (KUK) Provinsi Jawa Barat," UIN Sunan Gunung Djati Bandung, 2019.
[11] N. Engkus, Suparman and D. F. Sulistia, "Peranan Pembinaan Dinas Koperasi Dan Usaha Kecil Jawa Barat Dalam Memajukan Perkoperasian Di Daerah Koperasi sebagai sokoguru dan Masalah Lemahnya kesadaran Lemahnya partisipasi Lemahnya kedisiplinan Lemahnya jiwa kewirausahaan Lemahnya keterampilan Lemah," *J. Ilmu Adm. Publik*, vol. 7, no. 2, pp. 58–70, 2019.
[12] M. Irfan, W. Uriawan, O. T. Kurahman, M. A. Ramdhani, and I. A. Dahlia, "Comparison of Naive Bayes and K-Nearest Neighbor methods to predict divorce issues," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 434, no. 1, p. 012047, Dec. 2018.
[13] W. B. Zulfikar and N. Lukman, "Comparison of Naïve Bayes Classifier and Nearest Neighbor for Eye Disease Identification," *J. Online Inform.*, vol. 1, no. 2, Dec. 2016.
[14] O. Arifin and T. B. Sasongko, "Comparative Analysis of Performance Levels between Support Vector Machine and Naïve Bayes Classifier Methods," *Semin. Nas. Teknol. Inf. dan Multimed. 2018*, vol. 6, no. 1, pp. 67–72, 2018.
[15] E. Rohadi, A. Amalia, J. D. Bagaskara, B. Harijanto, and S. Adhisuwignjo, "GIS for coffee shops classification and routing using Naive Bayes method," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 732, no. 1, p. 012079, Jan. 2020.
[16] D. Jurafsky and J. Martin, "Naive Bayes and Sentiment Classification," *Speech Lang. Process.*, pp. 1–19, 2017.
[17] A. F. Hidayatullah, C. I. Ratnasari, and S. Wisnugroho, "Analysis of Stemming Influence on Indonesian Tweet Classification," *Telkomnika (Telecommunication Comput. Electron. Control.*, 2016.
[18] I. Rahmayuni, "Perbandingan Performansi Algoritma C4.5 dan Cart Dalam Klasifikasi Data Nilai Mahasiswa Prodi Teknnik Komputer Politeknik Negeri Padang," *J. TeknoIf*, vol. 2, no. 1, May 2014.
[19] D. Kurniawan and D. M. Yasir, "Optimization Sentimen Analysis using CRISP-DM and Naive Bayes Methods Implemented on Social Media," *Cybersp. J. Pendidik. Teknol. Inf.*, vol. 6, no. 2, pp. 74–85, Oct. 2022.
[20] M. Berdasarkan Dosen, dan Hasil Belajar Eka Sabna, S. Hang Tuah Pekanbaru Jl Mustafa Sari No, and T. Selatan Pekanbaru, "Penerapan Data Mining Untuk Memprediksi Prestasi Akademik Mahasiswa Berdasarkan Dosen, Motivasi, Kedisiplinan, Ekonomi, dan Hasil Belajar," *J. CoreIT J. Has. Penelit. Ilmu Komput. dan Teknol. Inf.*, vol. 2, no. 2, pp. 41 – 44, Dec. 2016.
[21] D. H. Kamagi and S. Hansun, "Implementasi Data Mining dengan Algoritma C4.5 untuk Memprediksi

Tingkat Kelulusan Mahasiswa," *J. Ultim.*, vol. 6, no. 1, pp. 15–20, Jun. 2014.
[22] N. Lukman, B. Subaeki, H. N. Abdullah, A. R. Atmadja, and M. Wildanuddin, "Prediction of national examination question using C4.5 algorithm," in *Journal of Physics: Conference Series*, 2019, vol. 1402, no. 7.
[23] W. Zhang and F. Gao, "An Improvement to Naive Bayes for Text Classification," *Procedia Eng.*, vol. 15, pp. 2160–2164, 2011.
[24] M. M. Saritas and A. Yasar, "Performance Analysis of ANN and Naive Bayes Classification Algorithm for Data Classification," *Int. J. ofIntelligent Syst. Appl. Eng.*, vol. 7, no. 2, 2019.
[25] J. Chen, H. Huang, S. Tian, and Y. Qu, "Feature selection for text classification with Naïve Bayes," *Expert Syst. Appl.*, vol. 36, no. 3, pp. 5432–5435, Apr. 2009.