Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Issue 1, January 2023
DOI : https://doi.org/10.33395/sinkron.v8i1.11968
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Data Mining Selection of Prospective Government Employees with Employment Agreements using Naive Bayes Classifier

Yosep Bustomi[1)*], Anwar Nugraha[2)], Christina Juliane[3)], Sri Rahayu[4)]
[1,2,3)]STMIK Likmi, Bandung, Indonesia [4)]Institut Teknologi Garut, Garut, Indonesia
[1)]yosepbustomi@gmail.com, [2)]anugrahcahaya@gmail.com, [3)] christina.juliane@likmi.ac.id, [4)]srirahayu@itg.ac.id

Submitted : Dec 5, 2022 | **Accepted** : Dec 8, 2022 | **Published** : Jan 1, 2023

Abstract: The implementation of health in Indonesia is still marked by problems in managing the workforce, especially among honorary health workers. The State Personnel Agency seeks to improve service quality by selecting good human resources to enhance community services. Several Ministries agreed to change the team member recruitment system from Prospective Civil Servants to Government Employees with Work Agreements. In the field, in the Government Employees with Work Agreements acceptance process, there are still pros and cons, both from the rules and the appointment process. The community hopes that the appointment process can be objective and open so that no group is disadvantaged. To achieve these expectations, researchers used data mining to classify health workers who would become Government Employees with Work Agreements. The data mining process uses probability with the Naive Bayes Classifier algorithm from historical data on Government Employees with Work Agreements receipts for 2021. Data on the history of Government Employees with Work Agreements acceptance of health workers as many as 1078 data have been filtered and cleaned. The results of testing the data are 0.00012 for the worthy assumptions and 0.0032 for the unworthy assumptions. Data processing results will be visually displayed using bubble diagrams and Python programming. The researcher concluded that the process of classifying prospective first-aid team members for medical personnel could be done by data mining using the Naïve Bayes algorithm. The results of this classification can be used as a reference for the following year's Government Employees with Work Agreements revenue classification process.

Keywords: Data Mining; Managing the Workforce; Naïve Bayes; Government Employees with Work Agreements; Python Programming

INTRODUCTION

The government recognizes the existence of Government Employees with Work Agreements as one of the elements of office employees in the Government environment. The government is trying to implement PP No. 49 of 2018 concerning Government Employees with Work Agreements by filling in the empty formations by recruiting employees from Non-Permanent Employees and related agencies to meet human resource needs, even though the regional budget is minimal(Kurniawan et al., 2020). Government Employees with Work Agreements has the same rights and obligations as Government employees. The only difference lies in the benefits of old age. The government has not fully regulated Government Employees with Work Agreements because the regulations have not been explained in detail regarding legal protection. For example, applicants who have served a long time in the agency are not distinguished from new applicants through the general route.

The government held Government Employees with Work Agreements to overcome the problem of structuring the civil servant management system in government, where there has been no legal certainty for honorary workers. With Government Employees with Work Agreements, agencies or institutions can get professional, transparent, competent staff according to their expertise(Qomarani, 2020). The civil servant procurement mechanism is carried out by the central or regional governments throughout Indonesia because there are technical guidelines for the leadership of the National Civil Service Agency for both Prospective Civil Servants and Government Employees with Work Agreements procurement. The difference lies in the technical or local content listed in the respective Regional Work Unit rules. In Head Regulation of the National Civil Service Agency No.1 of 2019 concerning technical Government Employees with Work Agreements

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Issue 1, January 2023
DOI : https://doi.org/10.33395/sinkron.v8i1.11968
e-ISSN : 2541-2019
p-ISSN : 2541-044X

procurement, it consists of several stages, including preparation, publication, application, administrative selection, Field Competency Selection, Field Competency Selection Results, appointment as Government Employees with Work Agreements candidate, and final appointment as Government Employees with Work Agreements, these stages must be passed by all prospective Government Employees with Work Agreements applicants(Hanamunika et al., 2021).

The government in the process of appointing Government Employees with Work Agreements has various problems, for example, the implementation of Government Employees with Work Agreements procurement for local government starting with the issuance of PP No.49 of 2018, that non-Government employees workers can be appointed through Government Employees with Work Agreements procurement with administrative and competency selection stages. This makes non-Government employees workers who have been working for a long time cannot be appointed directly like the previous policy(Fitriani et al., 2022). Researchers identified a problem related to Government Employees with Work Agreements, namely how to classify Government Employees with Work Agreements candidates using data mining so that the government could classify Government Employees with Work Agreements candidates who were eligible to be appointed earlier.

Many researchers have researched Government Employees with Work Agreements, but there are no data mining cases with the Naïve Bayes Classifier algorithm. Because of this, researchers are looking for the same research, but the object is similar, for example, related to civil servants. The object of research on civil servants with data mining processing with the Naïve Bayer Classification algorithm can be used as a reference in this research process. The researcher has obtained a research title used as a reference, namely the implementation of data mining using the Naïve Bayes Classifier method for data on promotions for the Medan City Employment Service(Simanjuntak et al., 2022). The research is the same as hiring the workforce and the methods used. The difference is between the data objects' location and the dataset attributes' details. Researchers in this study aim to classify Government Employees with Work Agreements applicants with data mining so that they become one of the government's references in determining eligible Government Employees with Work Agreements candidates. Researchers in this study aim to classify Government Employees with Work Agreements applicants with data mining so that they become one of the government's references in determining eligible Government Employees with Work Agreements candidates.

METHOD

The research method is a process flow of how to obtain data and information up to executing data to find accurate data. This process flow will assist in the research process. Figure 1 illustrates the stages of this research.

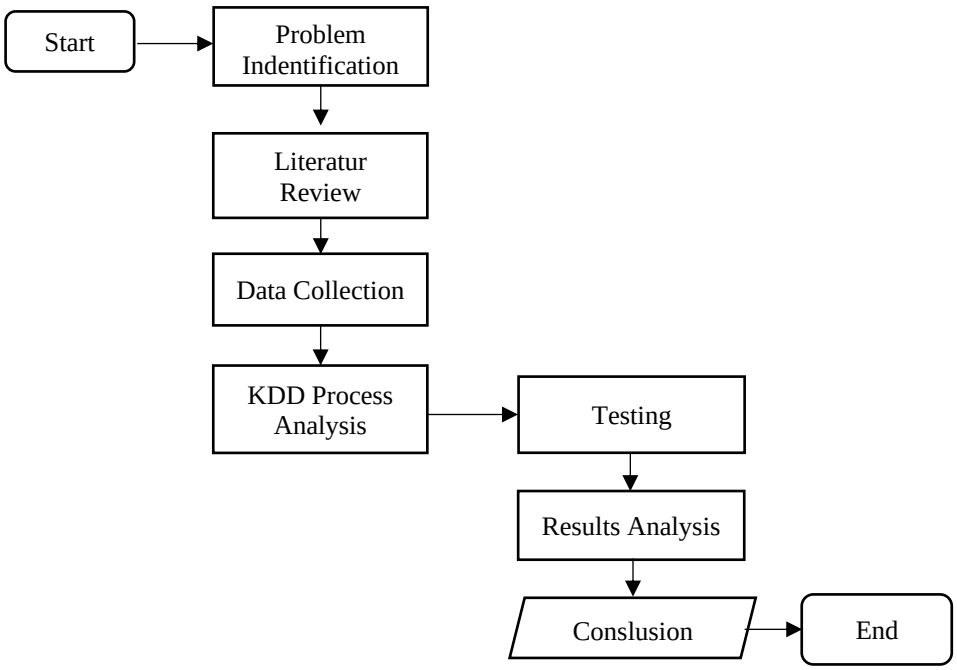

Figure 1. Research Stages

Figure 1 is a research flow starting from problem identification, literature study, data collection, KDD (Knowledge Discovery In Database) analysis, testing, analysis of results, and conclusions(Hant & Hendry, 2022).

*name of corresponding author

e-ISSN : 2541-2019
p-ISSN : 2541-044X

Problem Identification

Research begins with problem identification, which is helpful for researchers to focus on the problem under study. Based on the explanation in the introductory section, several problem points can be concluded: 1. How to classify prospective Government Employees with Work Agreements employees, especially health workers? 2. How can the Naïve Bayes Classifier Algorithm help classify prospective Government Employees with Work Agreements employees and see the level of accuracy of Naïve Bayes algorithm? 3. Can the Naïve Bayes Algorithm help government agencies to see the classification of prospective Government Employees with Work Agreements employees?

Literatur Review

At this stage, collecting the information needed for the research will be carried out. Researchers can get information from various sources, such as journals, books, or other references, of course, related to research using the Naïve Bayes algorithm so that they can be used as relevant references.

Data Collection

Researchers collected data by observing and interviewing Regional Personnel Agency officers. In the observations and interviews, researchers obtained data related to the recruitment process for Government Employees with Work Agreements employees in the form of an excel file dataset listing prospective Government Employees with Work Agreements employees for medical personnel in 2021. Participant data registered in the excel file totaled 1078 participants.

Data Analysis Training and Data Testing

Researchers in classifying Government Employees with Work Agreements team member candidates use the Naïve Bayes Classification algorithm. The author had to analyze the data obtained. Not all of the data was analyzed, only part of it, out of 1078 data, only 10 sample data were used as training data and 1 for testing data. Then for the overall data, a visualization of the results of the classification will be carried out using Google Colab so that the results of the classification are more informative.

Process Analysis of Knowledge Discovery in Database

Researchers use KDD to gain knowledge from existing data. The stages of KDD in this writing are 1. Data Selection is a process of selecting the data to be used. The selected data is then divided into training data and testing data. The type of data used in this study included primary data obtained directly from Regional Personnel Agency officers. The data in the database is not used, only part of it, but it still represents the original data. 2. Data Cleaning removes some attributes that are not needed in the data classification process. Data cleaning is done so that data consistency is maintained and redundant or missing values do not occur. In the data cleaning process, the researcher manages the dataset that will be used by cleaning several attributes in the classification process: Education, Position, Age, GPA, and Years Graduated. The data obtained from Regional Personnel Agency employees are clean and can be directly used for the classification process. 3. In the data transformation process, the author performs the preprocessing process on the data. The data results will be stored in a form ready to be used in data mining tools. In this study, the dataset was processed using SPSS (Statistical Product and Service Solutions) software from the IBM company from excel to CSV format(Jayal et al., 2018; Munandar et al., 2022).

Naïve Bayesian Classification Model

The method used in the Government Employees with Work Agreements classification process uses the Naïve Bayes Classifier algorithm method. The Naïve Bayes Classifier Algorithm is a classification method with a simple process. The basis of the Naïve Bayes Classifier method is a probability method by calculating the number and combination of values from the dataset. Compared to why use this algorithm, the Naïve Bayes algorithm is the most widely used because of its excellent efficiency and ability to combine evidence from data. This algorithm can predict the future based on past historical data. The Naïve Bayes method does not require extensive data in the classification process but a high level of accuracy so that data can be processed quickly(Afdhaluzzikri et al., 2022; Alwan et al., 2022). The researcher cites the Thomas Bayes formula(Gupitha, 2018; Prakoso et al., 2019), while the general formula is P (H|X) = P (X|H) P (H) ……. P (X), the formula explains that X is a data sample that has an unknown class (label), H is the hypothesis that X is class label data, P (H) is the probability of the H hypothesis, P (X) is the probability from the observed sample data, P (X|H) is the probability of X sample data if it is assumed that the hypothesis is true, so the Naïve Bayes Classifier can also be defined as a classification method based on probability theory and Bayesian theorem assuming that each

*name of corresponding author

variable or parameter determines the decision is independent so that the existence of each variable has nothing to do with other attributes.

Testing with Python

Python is an application program that can be used to manage data mining(Adila, 2022; Hamid et al., 2022). Python is a solution for analyzing data mining combined with Google Colab. At Google Colab, the features for managing data mining are complete, starting from storing datasets, processing data, and visualizing it in the form of diagrams(Harmaja et al., 2022).

RESULT

Researchers applied a data mining implementation approach to determine the classification of Government Employees with Work Agreements appointments based on specific criteria. The Naïve Bayes Classifier method is used from the sample data obtained then to process the data classification. The researcher will later process the data using SPSS, Python, and Google Colab so that the results of the Government Employees with Work Agreements acceptance classification can be visually seen.

Research Process Flow

Researchers in the research process follow the research flow as follows: 1. Researchers come to the Regional Personnel Agency office and conduct observations and interviews to obtain the required datasets related to Government Employees with Work Agreements, 2. perform data selection of the datasets obtained which are only used for Government Employees with Work Agreements classification, 3. perform cleaning data, for this stage, the author did not do it because the data was clean when received, 4. the results of the calculation were worthy and unworthy then processed using the Naïve Bayes method formula, 5. the results of data processing with the Naïve Bayes formula were visualized using Python and Google Colab.

Data Processing Process

The data used by the researchers were obtained from Regional Personnel Agency officers, where data were obtained from 1078 Government Employees with Work Agreements participants for Health Workers. The attributes used are 5, consisting of Name, Education, Position, Age, GPA, and Graduation Year. Of the 1078 applicants, 140 participants were accepted or eligible.

Data processing with the Naïve Bayes Classification Algorithm

Researchers perform data processing from 1078 data results from primary data, and researchers will test manually, namely 10 data, and the rest will be processed with a system designed as training data. The following shows the training data in table 1.

Table 1 Government Employees with Work Agreements Training Data

NO	NAME	EDUCATION	POSITION	AGE	GPA	GRADUATION YEAR	DESCRIPTION
1	SRI MULYANI	D-III	SKILLED - NURSE	34	2.96	2009	WORTHY
2	CECEP SUDIRMAN	D-III	SKILLED - NURSE	44	2.47	2000	WORTHY
3	PERA ANDANI	D-III	SKILLED - MIDWIFE	28	3.45	2015	WORTHY
4	DEA SINTA OKTASARI	S-1	FIRST EXPERT - HEALTH ADMINISTRATOR	30	3.34	2014	WORTHY
5	ANNISHA FATHONAH	S-1	FIRST EXPERT - HEALTH ADMINISTRATOR	31	3.41	2019	WORTHY
6	HERDIANSYAH SUKMANA	S-1	FIRST EXPERT - HEALTH ADMINISTRATOR	38	2.81	2007	UNWORTHY
7	NUR ISMI KUSNADI AMD KEP	D-III	SKILLED - NURSE	34	3.33	2009	UNWORTHY
8	YANI DEWI SARTIKA	D-III	SKILLED - NURSE	40	3.34	2003	UNWORTHY
9	SELI SELPIANI	D-III	SKILLED - NURSE	42	2.97	2004	UNWORTHY
10	LINDA HERLINDA	D-III	SKILLED - NURSE	41	2.95	2004	UNWORTHY

Table 1 shows ten training data, each of which has five eligible to become Government Employees with Work Agreements and five not eligible to become Government Employees with Work Agreements. This class is taken from the history of Government Employees with Work Agreements acceptance data in 2021. the Naïve Bayes

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Issue 1, January 2023
DOI : https://doi.org/10.33395/sinkron.v8i1.11968
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Classifier method is used to calculate the training data, namely by using several stages where the stages This is a calculation using a predetermined formula(Simanjuntak et al., 2022).

The first stage calculates the number of classes/labels using the total data for each class in each eligibility column divided by the total number of data.

P(Y=WORTHY) =5/10 = 0,5
P(Y=UNWORTHY) = 5/10 =0,5
Information : P = The probability of the hypothesis, Y = Class goals

The second stage calculates the same number of cases using the amount of data X divided by the amount of data Y, where X is a variable and Y is a class goal. For details related to the second stage process in table 2.

Table 2 Calculation of The Number of Cases

EDUCATION D-III	EDUCATION S-1
Based on AGE P (AGE = >30 |Y= WORTHY) =3/10 = 0,3 P (AGE = >30 |Y= UNWORTHY) =3/10 = 0,3	Based on AGE P (AGE = >30 |Y= WORTHY) =2/10 = 0,2 P (AGE = >30 |Y= UNWORTHY) =4/10 = 0,4
Based on AGE P (AGE < 30 |Y= WORTHY) =1/10 = 0,1 P (AGE < 30 |Y= UNWORTHY) =0/10 = 0	Based on AGE P (AGE < 30 |Y= WORTHY) =0/10 = 0 P (AGE < 30 |Y= UNWORTHY) =0/10 = 0
Based on GPA P (GPA => 3 |Y= WORTHY) =1/10 = 0,1 P (GPA => 3 |Y= UNWORTHY) =2/10 = 0,2	Based on GPA P (GPA => 3 |Y= WORTHY) =2/10 = 0,2 P (GPA => 3 |Y= UNWORTHY) =0/10 = 0
Based on EDUCATION D-III P(EDUCATION D-III | Y=WORTHY) = 3/10 = 0,3 P(EDUCATION D-III | Y= UNWORTHY) = 4/10 = 0,4	Based on EDUCATION S-1 P(EDUCATION S-1 | Y=WORTHY) = 2/10 = 0,2 P(EDUCATION S-1 | Y= UNWORTHY) = 1/10 = 0,1
Based on POSITION P(SKILLED - NURSE|Y= WORTHY) = 2/10 = 0,2 P(SKILLED - NURSE|Y=UNWORTHY)=4/10 = 0,4 Based on POSITION P(SKILLED - MIDWIFE |Y= WORTHY) = 1/10 = 0,1 P(SKILLED - MIDWIFE |Y=UNWORTHY)=0/10 = 0	Based on POSITION P(FIRST EXPERT - HEALTH ADMINISTRATOR|Y= WORTHY) = 2/10 = 0,2 P(FIRST EXPERT - HEALTH ADMINISTRATOR|Y= WORTHY |Y=UNWORTHY) =1/10 = 0,1
Based on GPA P (GPA < 3 |Y= WORTHY) =2/10 = 0,2 P (GPA < 3 |Y= UNWORTHY) =2/10 = 0,2	Based on GPA P (GPA < 3 |Y= WORTHY) =0/10 = 0 P (GPA < 3 |Y= UNWORTHY) =1/10 = 0,1
Based on GRADUATION YEAR P (GRADUATION YEAR >2005 |Y= WORTHY) = 2/10 =0,2 P (GRADUATION YEAR >2005 |Y=UNWORTHY)= 1/10 =0,1	Based on GRADUATION YEAR P (GRADUATION YEAR >2005 |Y= WORTHY) = 2/10 =0,2 P (GRADUATION YEAR >2005 |Y= UNWORTHY)=1/10 =0,1

The third stage is multiplying all the variables. In this process, the author multiplies the results of all variables. The detailed data is in table 3.

Table 3 Multiplication of All Variable Results from Calculations in Table 2

EDUCATION D-III	EDUCATION S-1
P(EDUCATION D-III | Y=WORTHY) * P(SKILLED - NURSE|Y= WORTHY) * P (AGE =>30 |Y= WORTHY) * P (GPA => 3 |Y= WORTHY) * P (GRADUATION YEAR >2005 |Y= WORTHY) = 0,3*0,2*0,3*0,1*0,2 = 0,00036	P(EDUCATION S-1 | Y=WORTHY) * ¬ P(FIRST EXPERT - HEALTH ADMINISTRATOR |Y= WORTHY) * P (AGE = >30 |Y= WORTHY) * P (GPA => 3 |Y= WORTHY) * P (GRADUATION YEAR >2005 |Y= WORTHY) = 0,2 * 0,2 * 0,2 *0,2 * 0,2 = 0,00032

Table 3 Multiplication of All Variable Results from Calculations in Table 2 (Continue)

EDUCATION D-III	EDUCATION S-1
P(EDUCATION D-III | Y=WORTHY) * P(SKILLED - MIDWIFE |Y= WORTHY) * P (AGE =>30 |Y= WORTHY) *	P(EDUCATION D-III | Y=UNWORTHY) * P(SKILLED - NURSE|Y= UNWORTHY) * P (AGE = >30 |Y= UNWORTHY) *

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Issue 1, January 2023
DOI : https://doi.org/10.33395/sinkron.v8i1.11968
e-ISSN : 2541-2019
p-ISSN : 2541-044X

P (GPA => 3 |Y= WORTHY) * P (GRADUATION YEAR >2005 |Y= WORTHY) = 0,3*0,1*0,3*0,1*0,2 = 0,00018	P (GPA => 3 |Y=UNWORTHY) * P (GRADUATION YEAR > 2005 |Y= UNWORTHY) = 0,4* 0,4* 0 ,3 *0,2 *0,1 = 0,00096

The fourth stage is to compare the class results. From the calculation results in table 3, it can be concluded that the probability data worthy as Government Employees with Work Agreements candidates based on the "SKILLED - NURSE" Position with education "S-1" is worthy 0.00036 while for the "SKILLED - MIDWIFE" education "D-III" is worthy 0. 00018. Furthermore, the results of data calculations for the education "S-1" category with the First Expert Position - Health Administrator are worthy 0.00032. Table 4 is the result of each data.

Table 4 Result Data

NO	NAME	EDUCATION	POSITION	AGE	GPA	GRADUATION YEAR	SCORE	DESCRIPTION
1	SRI MULYANI	D-III	SKILLED - NURSE	34	2.96	2009	0,00036	WORTHY
2	CECEP SUDIRMAN	D-III	SKILLED - NURSE	44	2.47	2000	0,00036	WORTHY
3	PERA ANDANI	D-III	SKILLED - MIDWIFE	28	3.45	2015	0,00018	WORTHY
4	DEA SINTA OKTASARI	S-1	FIRST EXPERT - HEALTH ADMINISTRATOR	30	3.34	2014	0,00032	WORTHY
5	ANNISHA FATHONAH	S-1	FIRST EXPERT - HEALTH ADMINISTRATOR	31	3.41	2019	0,00032	WORTHY
6	HERDIANSYAH SUKMANA	S-1	FIRST EXPERT - HEALTH ADMINISTRATOR	38	2.81	2007	0	UNWORTHY
7	NUR ISMI KUSNADI AMD KEP	D-III	SKILLED - NURSE	34	3.33	2009	0	UNWORTHY
8	YANI DEWI SARTIKA	D-III	SKILLED - NURSE	40	3.34	2003	0	UNWORTHY
9	SELI SELPIANI	D-III	SKILLED - NURSE	42	2.97	2004	0	UNWORTHY
10	LINDA HERLINDA	D-III	SKILLED - NURSE	41	2.95	2004	0	UNWORTHY

The fifth stage is testing with data testing, and the following data testing is used to test whether the data falls into the worthy or unworthy category, whether the data is the name Nurul Rojiah, "D-III" education, "SKILLED - NURSE" position, age 28, GPA 3.01, 2015 graduation year. Multiply based on the data with two assumptions whether one assumption is worthy and the two assumptions are unworthy, then take the highest result, then if the highest value from the results of the calculation is worthy, the data is in the worthy category, and otherwise, then the data is in the unworthy category. At the same time, the value multiplied is the result of the calculation in stage two, which is adjusted to the characteristics of the testing data being tested.

Table 5 Calculation of Testing Data for Nurul Rojiah's Data

ASSUMPTION WORTH	ASSUMPTION UNWORTHY
P(EDUCATION D-III | Y=WORTHY) * P(SKILLED - NURSE|Y= WORTHY) * P (AGE < 30 |Y= WORTHY)* P (GPA => 3 |Y= WORTHY)* P (GRADUATION YEAR >2005 |Y= WORTHY) = 0,3*0,2*0,1*0,1*0,2 = 0.00012	P(EDUCATION D-III | Y= UNWORTHY) * P(SKILLED - NURSE|Y= UNWORTHY) * P (AGE < 30 |Y= UNWORTHY)* P (GPA => 3 |Y= UNWORTHY)* P (GRADUATION YEAR >2005 |Y= UNWORTHY) = 0,4*0,4*0*0,2*0,1= 0.0032

The data testing in table 5 concludes that the assumptions are worthy to produce a value of 0.00012, and the assumptions are unworthy to produce a value of 0.0032 when compared. The data testing is in the ineligible category because the calculation is unworthy and is greater than the calculation is worthy.

Data Processing With Python and Google Colab

At the data processing stage, the author uses Python programming and Google Colab to process datasets, so that information related to data classification can be more easily understood. The flow of this process is shown in figure 2.

Figure 2. Google Colab Import Flow

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Issue 1, January 2023
DOI : https://doi.org/10.33395/sinkron.v8i1.11968
e-ISSN : 2541-2019
p-ISSN : 2541-044X

In figure 2, the dataset preparation stages that will be processed are in the form of files in CSV format. Then the files are imported into Google Colab and processed using script programming so that later the results can be visualized using bubble diagrams like figure 3.

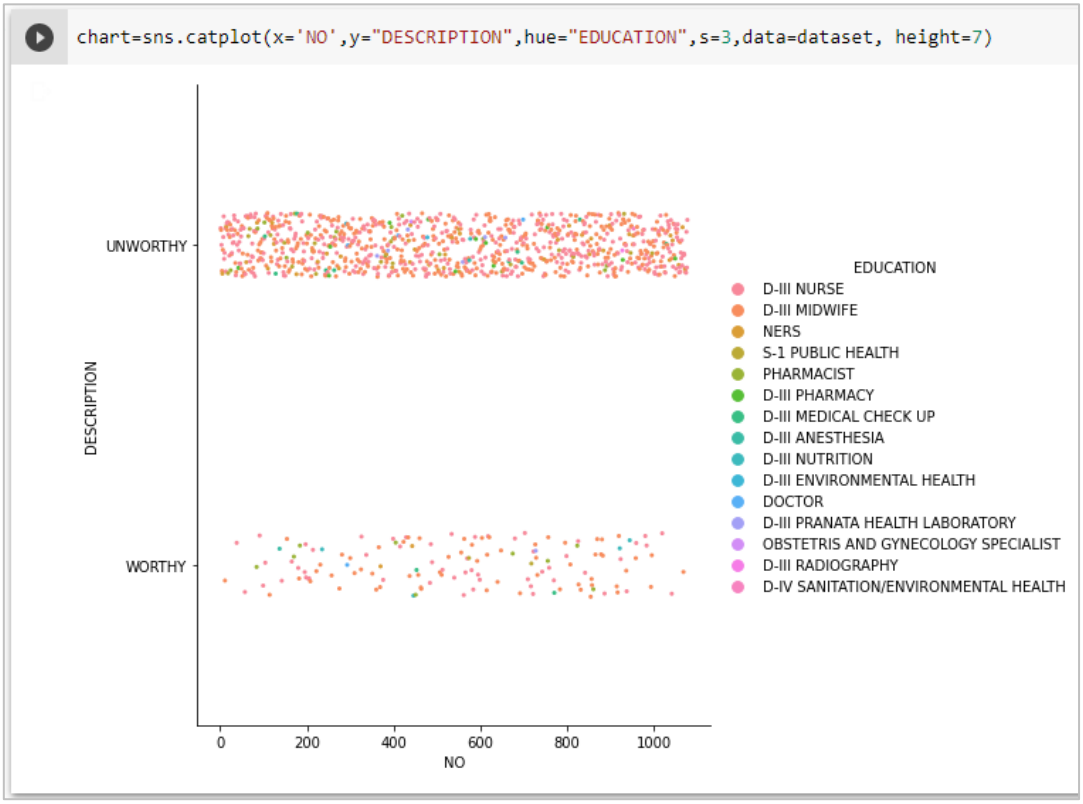

Figure 3. Dataset Visualization with a Bubble Data Diagram

In Figure 3, We can see the classification division between prospective Government Employees with Work Agreements employees who are worthy and unworthy to be appointed. The distribution of unworthy dominates the distribution of worthy. One hundred forty participant data are worthy of becoming Government Employees with Work Agreements employees based on the dataset received, and the rest fall into the unworthy class.

DISCUSSIONS

Researchers used the Naive Bayes algorithm to classify Government Employees with Employment Agreements datasets. The first process is knowing the flow of research from data collection to issuing calculation results, which can be in the form of values or data visualization. When processing data with the Naive Bayes algorithm, 10 data are used as training data from 1078 data. The 10 data were processed using the naive Bayes algorithm, resulting in a classification for the worthy category of 0.00012 and 0.0032 for the unworthy category. The results of these calculations can be used for testing other data or called data testing. The data used for testing use registrant data named Nurul Rojiah. The data is entered into the data multiplication formula for testing and produces 0.00012 for the assumption that it is worthy and 0.0032 for the assumption that it is unworthy. Of the two values when compared, Nurul Rojian is in the ineligible category because the value is greater than the worthy test value.

When compared with other studies, the calculation results are different due to differences in places and datasets processed. The title of the comparative research is "DATA MINING IMPLEMENTATION USING THE NAIVE BAYES CLASSIFIER METHOD FOR DATA RANK ADJUSTMENTS FOR THE MEDAN CITY EMPLOYMENT SERVICE," written by Adelin Yoseva Simanjuntak in 2022. The calculation results for the probability of data eligible for promotion based on class III with a Bachelor's degree education are worthy 0.0162. In contrast, for group IV, S1 education is worthy 0.027. Furthermore, the results of data calculations for the Master's education category with groups III and IV are worthy 0.008.

The limitations of this research study are that it is limited to only classifying worthy and unworthy candidates for Government Employees with Work Agreements.

*name of corresponding author

CONCLUSION

The researcher concluded this study with the following results: 1. The process of classifying prospective Government Employees with Work Agreements employees for medical personnel can be done using data mining, 2. The Naïve Bayes Classifier algorithm can help classify prospective Government Employees with Work Agreements employees. As for the classification results for classes, the worthy class is 0.00012, and the unworthy class is 0.0032.3. The results of this classification can be used as a reference for the classification of the Government Employees with Work Agreements acceptance process for the following year. The testing machine is the same as testing for data testing. It just needs to be adjusted to the following year's data.

REFERENCES

Adila, N. (2022). Implementation of Web Scraping for Journal Data Collection on the SINTA Website. *Sinkron*, *7*(4), 2478–2485. https://doi.org/10.33395/sinkron.v7i4.11576

Afdhaluzzikri, A., Mawengkang, H., & Sitompul, O. S. (2022). Perfomance analysis of Naive Bayes method with data weighting. *SinkrOn*, *7*(3), 817–821. https://doi.org/10.33395/sinkron.v7i3.11516

Alwan, J. K., Jaafar, D. S., & Ali, I. R. (2022). Diabetes diagnosis system using modified Naive Bayes classifier. *Indonesian Journal of Electrical Engineering and Computer Science*, *28*(3), 1766. https://doi.org/10.11591/ijeecs.v28.i3.pp1766-1774

Fitriani, F., Utami, E., & Hartanto, A. D. (2022). ANALISIS SENTIMEN MASYARAKAT TERHADAP PELAKSANAAN P3K GURU DENGAN ALGORITMA NAIVE BAYES DAN DECISION TREE. *TEKNIMEDIA: Teknologi Informasi Dan Multimedia*, *3*(1), 23–30. https://doi.org/10.46764/TEKNIMEDIA.V3I1.53

Gupitha, R. (2018). PENERAPAN KLASIFIKASI STATUS PEGAWAI MENGGUNAKAN METODE NAÃ• VE BAYES DI RSU H. SYAIFUL ANWAR. *Global*, *5*(1), 28–38. http://ejournal.unsub.ac.id/index.php/FASILKOM/article/view/418

Hamid, M. A., Aditama, D., Permata, E., Kholifah, N., Nurtanto, M., & Abdul Majid, N. W. (2022). Simulating the Covid-19 epidemic event and its prevention measures using python programming. *Indonesian Journal of Electrical Engineering and Computer Science*, *26*(1), 278. https://doi.org/10.11591/ijeecs.v26.i1.pp278-288

Hanamunika, G., Sari, D. S., & Dai, R. M. (2021). IMPLEMENTASI KEBIJAKAN PENGADAAN PEGAWAI PEMERINTAH DENGAN PERJANJIAN KERJA (PPPK) DI LINGKUNGAN PEMERINTAH KABUPATEN BANDUNG. *Responsive*, *3*(3), 143. https://doi.org/10.24198/responsive.v3i3.31921

Hant, M. I. P., & Hendry, H. (2022). DATA MINING TECHNIQUE USING NAÏVE BAYES ALGORITHM TO PREDICT SHOPEE CONSUMER SATISFACTION AMONG MILLENNIAL GENERATION. *Jurnal Teknik Informatika (Jutif)*, *3*(4), 829–838. https://doi.org/10.20884/1.JUTIF.2022.3.4.295

Harmaja, O. J., Purba, W., Siregar, M. P., Manurung, H. T., & Sirait, F. A. (2022). ANALISIS TINGKAT KEPUASAN PELANGGAN PADA RUDANG HOTEL BERASTAGI MENGGUNAKAN METODE CUSTOMER SATISFACTION INDEX (CSI). *Jurnal Teknik Informasi Dan Komputer (Tekinkom)*, *5*(1), 147. https://doi.org/10.37600/tekinkom.v5i1.511

Jayal, A., McRobert, A., Oatley, G., & O'Donoghue, P. (2018). Knowledge Discovery in Databases and Data Mining. In *Sports Analytics* (pp. 190–207). Routledge. https://doi.org/10.4324/9781315222783-10

Kurniawan, R., Rasyidin, R., Muhammad, M., & Aruni, F. (2020). Upaya Pemerintah dalam Menerapkan Peraturan Pemerintah No. 49 Tahun 2018 Tentang Pegawai Pemerintah Dengan Perjanjian Kerja (P3K) dI Kabupaten Aceh Utara. *Jurnal Public Policy*, *6*(1), 41–44. https://doi.org/10.35308/JPP.V6I1.1899

Munandar, D., Ruchjana, B. N., & Abdullah, A. S. (2022). PRINCIPAL COMPONENT ANALYSIS-VECTOR AUTOREGRESSIVE INTEGRATED (PCA-VARI) MODEL USING DATA MINING APPROACH TO CLIMATE DATA IN THE WEST JAVA REGION. *BAREKENG: Jurnal Ilmu Matematika Dan Terapan*, *16*(1), 099–112. https://doi.org/10.30598/barekengvol16iss1pp099-112

Prakoso, B. S., Rosiyadi, D., Utama, H. S., & Aridarma, D. (2019). Klasifikasi Berita Menggunakan Algoritma Naive Bayes Classifer Dengan Seleksi Fitur Dan Boosting. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, *3*(2), 227–232. https://doi.org/10.29207/resti.v3i2.1042

Qomarani, L. N. (2020). ANOMALI KEHADIRAN PEGAWAI PEMERINTAH DENGAN PERJANJIAN KERJA (PPPK) DALAM CAKRAWALA KEPEGAWAIAN DI INDONESIA. *Cepalo*, *4*(2), 95–110. https://doi.org/10.25041/cepalo.v4no2.1979

Simanjuntak, A. Y., Simatupang, I. S. E. S., & Anita, A. (2022). IMPLEMENTASI DATA MINING MENGGUNAKAN METODE NAÃVE BAYES CLASSIFIER UNTUK DATA KENAIKAN PANGKAT DINAS KETENAGAKERJAAN KOTA MEDAN. *JOURNAL OF SCIENCE AND SOCIAL RESEARCH*, *5*(1), 85. https://doi.org/10.54314/jssr.v5i1.804

*name of corresponding author