ISSN: 2460 – 7223

Naïve Bayes Classifier Algorithm For Predicting Non-Participation Of Elections in Lampung Province

1st Fitria
Faculty of Computer Science
Departement of Informatics Technology
Institut Informatika dan Bisnis Darmajaya
Bandar Lampung, Indonesia
fitria@darmajaya.ac.id

2nd Rifad Sobah
Faculty of Computer Science
Departement of Informatics Technology
Institut Informatika dan Bisnis Darmajaya
Bandar Lampung, Indonesia
rifadsobah@gmail.com

3rd Chairani Fauzi
Faculty of Computer Science
Departement of Informatics Technology
Institut Informatika dan Bisnis Darmajaya
Bandar Lampung, Indonesia
chairani@darmajaya.ac.id

4th Septilia Arfida
Faculty of Computer Science
Departement of Informatics Technology
Institut Informatika dan Bisnis Darmajaya
Bandar Lampung, Indonesia
septilia@darmajaya.ac.id

5th Suci Mutiara
Faculty of Computer Science
Departement of Informatics Technology
Institut Informatika dan Bisnis Darmajaya
Bandar Lampung, Indonesia
sucimutiara@darmajaya.ac.id

6th Siti Nurlaila
Faculty of Computer Science
Departement of Informatics Technology
Institut Informatika dan Bisnis Darmajaya
Bandar Lampung, Indonesia
sitinurlaila@darmajaya.ac.id

Abstract—Several problems related to the DPT (Permanent Voter List) including the KPU (General Election Commission) it is difficult to get the NIK (Population Identification Number) of people who are in correctional institutions or prisoners, beginner voters who do not have an ID card (Kartu Tanda Sipil) who are currently in prison. study in student dormitories, Islamic boarding schools, and others who are outside the city, the number of which is 3-5% of invalid NIK, voters who do not have a resident identity, voters with KTP (Kartu Identity Card)/old Family Card and NIK (Population Identification Number)) invalid around 7-19% and voters are difficult to find around 5-8% so that the KPU must-visit houses as regulated in the legislation. This could allow not all DPT (Permanent Voters List) to be registered. Naïve Bayes Classifier is one of the classification methods used in Data Mining which is based on the Bayes theorem. Bayes is a simple probability-based prediction technique based on the application of Bayes' theorem (or Bayes' rule) with strong (naive) independent (independence) assumptions. Naive Bayes is only a method for analyzing, it takes other media to display information that is easy to understand the results of the Naïve Bayes Classifier calculations. Pentaho data integration is a tool that integrates large amounts of data, calls from Excel, MySQL, and provides instructions for existing data. Tableau is an application that will improve, tableau can call data that has been integrated and display the data in the form of diagrams, text, spatial data, and points from locations.

Keywords: Naïve Bayes Classifier, Algorithm, Election Commission

I. INTRODUCTION

A general election (election) is not a foreign thing for the people of Indonesia today. Among them are elections (general elections) which include, pilkada (regional head elections), pileg (legislative elections), and pilpres (presidential elections). In essence, the general election is a way or means to find out the wishes of the people regarding the direction and policies of the state in the future. There are at least three kinds of general election objectives, namely enabling a safe and orderly transition to exercise people's sovereignty in the context of implementing the human rights of citizens [6].

General Elections are held to elect the People's Representatives. The state of Indonesia itself held the first Indonesian general election after its independence in 1955, and this 2019 election is the 12th. Elections for members of the executive body have only been held since 2004, therefore there are still many problems that occur in the election in terms of the Permanent Voters List. The KPU (General Election Commission) Regulation No.26 of 2013 explains that the DPT (Permanent Voters List) is a registered citizen to vote at a designated TPS (Polling Place) [12].

Some of the problems related to the DPT (Permanent Voters List) include the KPU (General Election Commission) finding it difficult to get the NIK (Population Identification Number) of people who are in correctional institutions or detainees, beginner voters who do not have an ID card (Kartu Identity Card) who are studying in prison. student dormitories, Islamic boarding schools, and others who are outside the city, the number of which is 3-5% of invalid NIK, voters who do not have a resident identity, voters with old ID cards (KTP)/family cards and invalid NIK (Population Identification Numbers). around 7-19% and voters are difficult to find around 5-8% so the KPU must-visit houses as regulated in the legislation. This could allow not all DPT (Permanent Voters List) to be registered [3].

these various problems, resulted in a lack of public participation to take part in this election. If the possibility of DPT (Permanent Voters List) who will not vote is known earlier by the KPU (General Election Commission), then it will be easier to determine the target of election socialization and to increase the percentage of community participation. Therefore, analysis is needed to predict participation in the upcoming elections.

Naïve Bayes Classifier is one of the classification methods used in Data Mining which is based on the Bayes theorem. Bayes is a simple probability-based prediction technique based on the application of Bayes' theorem (or Bayes' rule) with strong (naive) independent (independence) assumptions [7]. Naive Bayes is only a method for analyzing, it takes other media to display information that is easy to understand the results of the Naïve Bayes Classifier calculations. Pentaho data integration is a tool that integrates large amounts of data, calls from excel, MySQL and provides instructions for existing data. Tableau is an application that will improve, tableau can call data that has been integrated and display the data in the form of diagrams, text, and spatial data points from locations.

II. LITERATURE REVIEW

DPT (Permanent Voters List)

According to KPU Regulation No. 26 of 2013 Chapter II Article 6 stipulates that there are 4 voters in the election, namely DPT (Permanent Voters List), DPT (Additional Permanent Voters List), DPK (Special Voters List), and DPTKTb (Additional Special Voters List). DPT (Permanent Voters List) itself is a citizen who has been registered to vote at a designated TPS (Polling Place). The non-participating DPT (Permanent Voters List) is a DPT that has been registered but did not vote during the election. The number of non-participating DPT (Permanent Voters List) affects the percentage of public participation in elections [5].

Naïve Bayes Classifier (NBC)

The Naive Bayes algorithm is a classification algorithm based on the Bayesian theorem in statistics. The Naive Bayes algorithm can be used to predict the probability of membership in a class.

The Bayesian theorem calculates the posterior probability value P(H|X) using the probabilities P(H), P(X), and P(X|H), where the value of X is testing data whose class is unknown. The value of H is the hypothesis of X data which is a more specific class. The value of P(X|H) or also called likelihood, is the probability X based on the condition H. The value of P(H) or also called the prior probability is the probability hypothesis H. While the value of P(X) is also called the predictor prior probability , is the probability X.

$$P = \frac{P(X|H).P(H)}{P(X)}$$

"Nominal Equation"

The Naive Bayes algorithm is very suitable for classifying nominal-type datasets. For datasets the nominal type.

The steps of the Naive Bayes algorithm are as follows:

1. Prepare the dataset.
2. Count the number of classes in the training data.
3. Count the number of cases with the same class.
4. Multiply all the results according to the testing data that the class will look for.
5. Compare the results per class, the highest score is assigned to the new class.

a. Xampp

XAMPP is an apache web server software in which a My SQL server database is available and can support installation on Linux and Windows. Another advantage is that it only installs once. Apache Web Server, MySQL, PHP Database Server Support (PHP 4 and PHP 5) and several other modules are available [14].

b. SQLyog

SQLyog is a very popular MySQL client system used in Indonesia. This system has many features that make it easier for users to administer and perform MySQL data processing,

SQLyog is a system used to connect one computer to another, so that one computer with a computer can access the other, it can be said that the system is a system commonly used to implement a client- server[10].

c. Pentaho Data Integration (PDI) Kettle

Pentaho Data Integration (PDI) Kettle is an Open Source software from Pentaho that can be used to integrate data. The kettle provides ETL (Extraction, Transformation and Loading) facilities. Kettles can be used for cleaning data, loading from files to databases or vice versa, and migrating between large volumes of applications. The commercial version of Kettle is Pentaho Data Integration (PDI).

d. Tableau

A Tableau is a tool that can analyze/describe a collection of data to be presented in an attractive form. This tableau is a business intelligence research result from the Gartner Report February 2016. Tableau occupies the chart or the top quadrant in the BI (Business Intelligence) platform along with 2 similar platforms. It has three main products, namely tableau desktop, tableau server, and tableau online. This application has the following specifications:

1. Ability to connect and retrieve data from various data sources, dynamic and multidimensional data analysis capabilities.
2. Ability to analyze data on integrated maps (no need for additional software), Ability to create information dashboards).
3. Provide a means of collaborative Business Intelligence.

4. Access information anywhere, anytime, and using a mobile device.
5. Ability to add calculations from existing data and parameters.
6. Having internal data facilities that have high performance.
7. Business Intelligence platform with enterprise scale (High Availability, Multi-tenancy, performance with scale up and scale across, proxy data).
8. Print, and export analysis results and data. Sharing template files. Fast implementation.

e. Microsoft Excel

Excel or Microsoft Office Excel is a spreadsheet application program created and distributed by Microsoft Corporation for the Microsoft Windows operating system. This application has calculation and graphing features that use Microsoft's aggressive marketing strategy, making Microsoft Excel one of the most popular computer programs used as one of the most popular computer programs used in microcomputers to date. this program is currently the most widely used spreadsheet program by many parties. [8].

III. RESEARCH METHOD

3.1 Data Collection Techniques

The data collection technique is a form of data collection that aims to obtain the information needed to achieve the research objectives. The data source referred to in the study is the subject from which the data can be obtained and has clear information about how to retrieve the data and how the data is collected. is processed. And the source of the data in this study was obtained from the KPU to analyze the DPT (Permanent Voters List). Data sources are divided into 2, namely:

1. Premier Data
2. Secondary Data

3.2 Research Flow

In conducting Non-Participate Predictions using the Naïve Bayes Classifier Algorithm, it can be seen in the flowchart diagram in Figure 1. below:

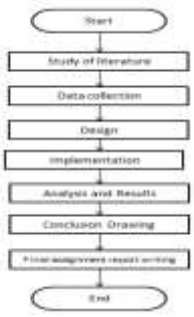

Figure 1. Flowchart of research

In the flowchart diagram above, it explains the flow/stages of the research in this research;

1. Study Literature
   After finding the problem, here the researcher begins to collect material to broaden understanding or as a reference/reference whether it is studying the object, the method and the program, which are obtained from digital and traditional information media.
2. Data Collection Process
   After the problems and literature are known, the data needed here begins to be collected, either from interviews or sampling.
3. System Design
   Here the design is made starting from the criteria, methods, input data, data integration methods and the resulting output.
4. Implementation
   After the data is obtained, the data will be processed and it will add how much additional data to make the development of the problem of the researcher. Then the data will be implemented in the nave Bayes classifier algorithm.
5. Analysis and Results
   This stage, several experimental data samples are analyzed and the results are determined.
6. Drawing Conclusion
   After the data is implemented on the nave Bayes algorithm, it will be tested and analyzed for results and validation.
7. Report Writing this stage, the final stage of this research, here the preparation of reporting from this research begins to be done by researchers.

3.3 Naïve Bayes Classifier

i. Determination of Variables to be used
   The variables used as data that have values to be calculated using the Naive Bayes algorithm can be seen in table 1.

Table 1. Research Data Variables

Place of birth	Variables that inform Place of Birth from DPT
Marital Status B/S/P	This variable informs the marital status where B (Not yet), S (Already), P (Never) from DPT
Gender	Variables that inform the Gender of the DPT
Address (Street/ Dukuh, RT, RW)	This variable informs the Address From DPT
Information	Variables that inform Participate and Non-Participate Information.

ii. Determination of Data Set
   The data used to find a result in Predicting Non-Participating DPT (Permanent Voters List) in the Election, data sets to analyze existing DPT data with Non Participate results after being analyzed through the input variables above if Place of Birth = Outside Voter Area / Voter Area, Marital Status = Not yet/Already/Already, Gender = Male/Female, Address = Outside Voter

Area/Voting Area, Possibility of Participating or Non-Participating.

iii. Calculation of the Number of Cases in Variables
Count the Number of Cases with the same class
P (Place of Birth = "VOCATE REGION" | C="Non Participate")
P (Place of Birth = "OUTSIDE OF THE VOTING REGION" | C="Non Participate")
P(Status = "B" | C="Non Participate")
P(Status = "S" | C="Non Participate")
P(Status = "P" | C="Non Participate")
P (Gender = "P" | C="Non Participate")
P (Gender = "L" | C="Non Participate")
P (Address = "VOCTOR REGION"| C="Non Participate")
P (Address = "OUTSIDE THE VOTING REGION" | C="Non Participate")

iv. Multiply all the results according to the testing data that the class will look for

P(X| C = "Non Participate") = VALUE OF BIRTH * VALUE OF STATUS * VALUE OF GENDER * VALUE OF ADDRESS * DESCRIPTION

v. Compare Results Per Class

From the calculation of the NON-PARTICIPATE probability in the previous step, it can be concluded that the PLACE OF BIRTH, STATUS, GENDER, ADDRESS, DESCRIPTION of greater value.

3.5 Framework of Thought

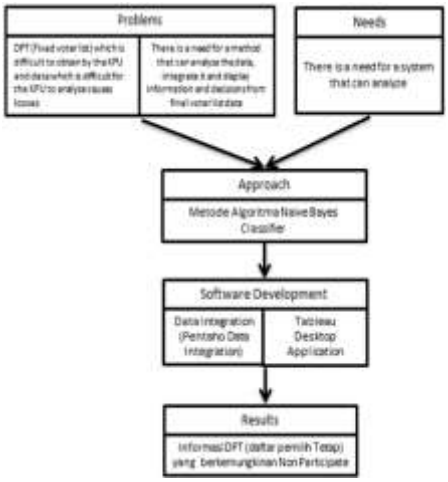

Figure 2. Thinking Framework

The current problem is that the DPT which is difficult to obtain by the KPU and the data that is difficult to analyze by the KPU causes losses. However, data alone is not enough, a method is needed to analyze the data, integrate and display the data properly.
The approach used is the Naïve Bayes Classifier method to determine. The existing data is then analyzed using the Naïve Bayes Classifier to analyze the data to make a decision on the DPT (Permanent Voters List) which is likely to be Non-Participate.

This time the software development uses pentaho data integration for data integration that has been developed using the Naïve Bayes Classifier and then displayed with a desktop tableau to display the results of data analysis in the form of a Non-Participate DPT graph.

3.6. Architecture

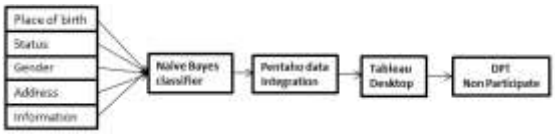

Figure 3. Data flow architecture and tools

3.6.1. Variable

The variables used as data that have values to be calculated using the nave Bayes classifier algorithm are Place of Birth, Marital Status B/S/P, Gender, Address, Information

3.6.2 Naive Bayes Classifier

Data analysis method to analyze existing data by analyzing existing DPT data with Non Participate results after being analyzed through the input variables above if Place of Birth = Outside Voter Area / Voter Area, Marital Status = Not yet / Already / Ever, Gender = Male Male/Female, Address = Outside Voter Area / Voter Area. So, it is possible to Participate or Non Participate.

3.6.3. Pentaho data integration

After obtaining the results of data analysis using the nave Bayes classifier algorithm, the data is integrated using pentaho, connecting the data through the primary key and giving rules.

3.6.4. Desktop Tableau

The last stage in the analysis and decision making in this research is to display the data using a desktop tableau in the form of a graph chart, as well as spatial data from DPT that has the potential to be non-participant.

IV. RESULTS AND DISCUSSION

The following will explain the results of the implementation of the Naiva Bayes Classifier algorithm to predict non-participant DPT (permanent voter lists) using Microsoft Excel. Pentaho and tableau are applications that display the results of the classification per kelurahan, per sub-district and per district, which can be seen as follows:

4.1. Results Per Village

This display describes the results of the calculation of the data set from the Naive Bayes algorithm per Banyumas Village which can be seen in the following figure:

Figure 4. Prediction Results of Banjar Rejo and Banyu Urip Villages

The picture above is the result of each probability calculation from TPS in Banjar Rejo and Banyu Urip villages.

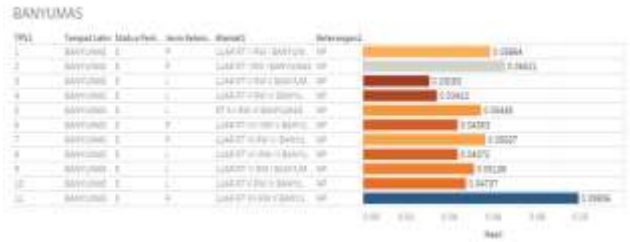

Figure 5. Prediction Results of Banyumas Village

The picture above is the result of each calculation with probability from TPS in Banyumas Village.

Figure 6. Prediction Results of Banyuwangi Village

The picture above is the result of each calculation with probability from TPS in Banyuwangi kelurahan.

Figure 7. Prediction Results of Waya Krui Village

The picture above is the result of each calculation with probability from TPS in Waya Krui village

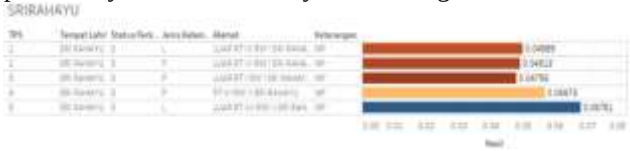

Figure 8. Prediction Results of Sri Rahayu Village, Sri Wungu, Suka Mulya

The picture above is the result of each calculation with probability from TPS in Sri Rahayu, Sri Wungu and Suka Mulya villages.

Figure 9. Prediction Results of the Villages of Mulyorejo, Sinar Mulia, and Nusa Wungu

The picture above is the result of each calculation with probability from TPS in the villages of Mulyorejo, Sinar Mulia, and Nusa Wungu.

The proof in the manual count with training data at TPS 1 Banjar Rejo can be seen as follows:

1. Calculation of the Number of Cases on Attributes
   The class in the training data consists of 177 categories, so the probability for non-participation is as follows;
   Number of Non-Participate Classes = 26
   So : P(C = "Non Participate") = 26/177 = 0.146893

2. Count the same number of cases with the same class
   Count the Number of Cases with the same class.
   P (Place of Birth = "Banjar Rejo"| C="Non Participate") = 14/26 = 0.538462
   P (Place of Birth = "Outside Banjar Rejo" | C="Non Participate") = 12/26 = 0.461538
   P(Status = "B"| C="Non Participate") = 19/26 = 0.730769
   P(Status = "S"| C="Non Participate") =6/26 = 0.230769
   P(Status = "P"| C="Non Participate") = 1/26 = 0.038462
   P (Gender = "P"| C="Non Participate") = 12/26 = 0.461538
   P (Gender = "L"| C="Non Participate") = 14/26 = 0.538462
   P (Address = "RT II RW I BANJAREJO"| C=" Non Participate") 6/26 = 0.230769
   P (Address = "OUTSIDE BANJAREJO"| C="Non Participate") 20/26 = 0.769231

3. Non Participate Prediction Testing Data
   a. Data Testing 1

Place of Birth: Outside Banjar Rejo
Status : S
Gender : P
Address : OUTSIDE RT II RW I BANJAREJO
Description: NP

b. Data Testing 2
Place of Birth: Outside Banjar Rejo
Status : S
Gender : L
Address : OUTSIDE RT II RW I BANJAREJO
Description: NP

c. Data Testing 3
Place of Birth: Outside Banjar Rejo
Status : B
Gender : P
Address : OUTSIDE RT II RW I BANJAREJO
Description: NP

d. Testing Data 4
Place of Birth: Outside Banjar Rejo
Status : B
Gender : L
Address : OUTSIDE RT II RW I BANJAREJO
Description: NP

e. Testing Data 5
Place of Birth: Outside Banjar Rejo
Status : P
Gender : P
Address : OUTSIDE RT II RW I BANJAREJO
Description: NP

f. Testing Data 6
Place of Birth: Outside Banjar Rejo
Status : P
Gender : L
Address : OUTSIDE RT II RW I BANJAREJO
Description: NP

g. Testing Data 7
Place of Birth: Outside Banjar Rejo
Status : S
Gender : P
Address : RT II RW I BANJAREJO
Description: NP

h. Testing Data 8
Place of Birth: Outside Banjar Rejo
Status : S
Gender : L
Address : RT II RW I BANJAREJO
Description: NP

i. Data Testing 9
Place of Birth: Outside Banjar Rejo
Status : B
Gender : P
Address : RT II RW I BANJAREJO
Description: NP

j. Testing Data 10
Place of Birth: Outside Banjar Rejo
Status : B
Gender : L
Address : RT II RW I BANJAREJO
Description: NP

k. Testing Data 11
Place of Birth: Outside Banjar Rejo
Status : P
Gender : P
Address : RT II RW I BANJAREJO
Description: NP

l. Testing Data 12
Place of Birth: Outside Banjar Rejo
Status : P
Gender : L
Address : RT II RW I BANJAREJO
Description: NP

m. Testing Data 13
Place of Birth: Banjar Rejo
Status : S
Gender : P
Address : RT II RW I BANJAREJO
Description: NP

n. Test Data 14
Place of Birth: Banjar Rejo
Status : S
Gender : L
Address : RT II RW I BANJAREJO
Description: NP

o. Data Testing 15
Place of Birth: Banjar Rejo
Status : B
Gender : P
Address : RT II RW I BANJAREJO
Description: NP

p. Test Data 16
Place of Birth: Banjar Rejo
Status : B
Gender : L
Address : RT II RW I BANJAREJO
Description: NP

q. Data Testing 17
Place of Birth: Banjar Rejo
Status : P
Gender : P
Address : RT II RW I BANJAREJO
Description: NP

r. Data Testing 18
Place of Birth: Banjar Rejo
Status : P
Gender : L
Address : RT II RW I BANJAREJO
Description: NP

s. Testing Data 19
Place of Birth: Banjar Rejo
Status : S
Gender : P
Address : OUTSIDE RT II RW I BANJAREJO
Description: NP

t. Testing Data 20
Place of Birth: Banjar Rejo
Status : S
Gender : L
Address : OUTSIDE RT II RW I BANJAREJO
Description: NP

u. Testing Data 21
Place of Birth: Banjar Rejo
Status : B
Gender : P
Address : OUTSIDE RT II RW I BANJAREJO
Description: NP

v. Test Data 22
Place of Birth: Banjar Rejo
Status : B
Gender : L
Address : OUTSIDE RT II RW I BANJAREJO
Description: NP

w. Testing Data 23
Place of Birth: Banjar Rejo
Status : P
Gender : P
Address : OUTSIDE RT II RW I BANJAREJO
Description: NP

x. Testing Data 24
Place of Birth: Banjar Rejo
Status : P
Gender : L
Address : OUTSIDE RT II RW I BANJAREJO
Description: NP

4. Multiply all the results according to the testing data that will be looking for the class

a. Testing Data 1 = 0.461538*0.230769*0.461538*0.769231*0.14689 = 0.0055546
b. Testing Data 2 = 0.461538*0.230769*0.538462*0.769231*0.14689 = 0.0064803
c. Testing Data 3 = 0.461538*0.730769*0.461538*0.769231*0.14689 = 0.0175895
d. Testing Data 4 = 0.461538*0.730769*0.538462*0.769231*0.14689 = 0.0205210
e. Testing Data 5 = 0.461538*0.038462*0.461538*0.769231*0.14689 = 0.0009258
f. Testing Data 6 = 0.461538*0.038462*0.538462*0.769231*0.14689 = 0.0010801
g. Testing Data 7 = 0.461538**0.230769*0.461538*0.230769*0.14689 = 0.0016664
h. Testing Data 8 = 0.461538**0.230769*0.538462*0.230769*0.14689 = 0.0019441
i. Testing Data 9 = 0.461538**0.730769*0.461538*0.230769*0.14689 = 0.0052768
j. Testing Data 10 = 0.461538**0.730769*0.538462*0.230769*0.14689 = 0.0061563
k. Testing Data 11 = 0.461538**0.038462*0.461538*0.230769*0.14689 = 0.0002777
l. Testing Data 12 = 0.461538**0.038462*0.538462*0.230769*0.14689 = 0.0003240
m. Testing Data 13 = 0.538462*0.230769*0.461538*0.230769*0.14689 = 0.0019441
n. Testing Data 14 = 0.538462*0.230769*0.538462*0.230769*0.14689 = 0.0022681
o. Testing Data 15 = 0.538462*0.730769*0.461538*0.230769*0.14689 = 0.0061563
p. Testing Data 16 = 0.538462*0.730769*0.538462*0.230769*0.14689 = 0.0071824
q. Testing Data 17 = 0.538462*0.038462*0.461538*0.230769*0.14689 = 0.0003240
r. Testing Data 18 = 0.538462*0.038462*0.538462*0.230769*0.14689 = 0.0003780
s. Testing Data 19 = 0.538462*0.230769*0.461538*0.769231*0.14689 = 0.0064803
t. Testing Data 20 = 0.538462*0.230769*0.538462*0.769231*0.14689 = 0.0075604
u. Testing Data 21 = 0.538462*0.730769*0.461538*0.769231*0.14689 = 0.0205210
v. Testing Data 22 = 0.538462*0.730769*0.538462*0.769231*0.14689 = 0.0239412
w. Testing Data 23 = 0.538462*0.038462*0.461538*0.769231*0.14689 = 0.0010801
x. Testing Data 24 = 0.538462*0.038462*0.538462*0.769231*0.14689 = 0.0012601

5. Compare Results Per Class

From the calculation of the Non Participate probability in the previous step, it can be concluded that Place of Birth = Banjar Rejo, Status = B, Gender = L, Address = OUTSIDE RT II RW I BANJAREJO, Information = NP, Has a Higher Value of 0.0239412.

A. District Results

Figure 10. Prediction Results of Banyumas District

The picture above is data from the Calculation Results of All Villages in Banyumas District, the data is recalculated with the results based on the calculation of the Naive Bayes Algorithm showing results in categories:

Place of birth: Voter District
Status : Already
Gender : Male
Address : Outside Voter Area
Description: NP
Value : 0.7272727

B. Results per District

Figure 11. Results Per Distric

Gender : Male
Address : Outside Voter Area
Description: NP
Value : 1

4.2. Discussion

The experiment in this study used testing data from Pringsewu Regency with a total of 9 Districts, 131 Kelurahan, 1,416 TPS (Polling Stations), 295,934 DPT (Permanent Voters List) and the algorithm used was the Naive Bayes Classifier. The results of the experiments that have been carried out using the naive Bayes model, it can be seen that the data that has gone through the process of selecting data, selecting attributes and also converting data from data with 5 variables produces non-participatory DPT characteristics that are obtained, namely Place of Birth: Voter Region, Status : Yes, Gender: Male, Address: Outside Voter Area, Information: NP, Value: 1.

The Naive Bayes Classifier Calculation Result Data is then uploaded to the results database.

Figure 12. Upload Data Into The Result Database

The data that has been uploaded is then displayed on the dashboard of the results of the urban village.

Figure 13. Connector Database Tableau

From the results obtained in this study using the Naïve Bayes Classifier algorithm to help make decisions for the processed data can help us analyze the final voter list data, the data used is the Pringsewu district permanent voter list data with a total of 9 sub-districts, 131 sub-districts , 1,416 TPS (Polling Points), 295,934 DPT (Permanent Voters List) to predict Non Participate DPT obtained the following results:

Place of birth: Voter District
Status : Already
Gender : Male
Address : Outside Voter Area
Description: NP
Value : 1

V. CONCLUSION

Based on the research that has been done by researchers, it can be concluded several things as follows:

a. From the research that has been done, it can be concluded that to predict public participation in General Elections (PEMILU) in Lampung Province, it can be done using the Naïve Bayes Algorithm with 5 predetermined variables, namely Place of Birth, Status, Gender, Address, Information. Prediction results of election participation from datasets taken from all DPT (Permanent Voters List) Pringsewu Regency with a total of 9 Districts, 131 Kelurahan, 1,416 TPS (Polling Points), 295,934 DPT (Permanent Voters List).
b. From the variables used in the study, it is known that the characteristics of Non-Participating DPT from the data obtained by researchers are Place of Birth (Voter

Region), Status (Already), Gender (Male), Address (Outside Voter Area), Information (NP), Value (1).

REFERENCES

[1] Akbar, D. O. (2017). Implementasi Bussinesee Intelligence Pada Analisis Peningkatan Sarana Perairan Kota Padang Tahun 2013 – 2015 Menggunakan Aplikasi Tableau. 1-4.

[2] Sabita, H., Fitria, F. and Herwanto, R., 2021. Analisa Dan Prediksi Iklan Lowongan Kerja Palsu Dengan Metode Natural Language Programing Dan Machine Learning. *Jurnal Informatika*, *21*(1), pp.14-22.

[3] Fitriani. (2019). Penerapan Data Mining Menggunakan Metode Klasifikasi Naive Bayes untuk Memprediksi Partisipasi Pemilihan Gubernur. *Jurnal teori*, 1-7.

[4] Wibowo, H. and Indriyani, F., 2018, October. K-Nearest Neighbor Method For Monitoring Of Production And Preservation Information (Treatment) Of Rubber Tree Plant. In *Proceeding International Conference on Information Technology and Business* (pp. 29-44).

[5] Listyaningrum. (2014). Penerapan Data Mining Untuk Analisis Karakteristik DPT Non-Participate sebagai Prediksi Partisipan Pemilu dengan Menggunakan Naive Bayes Classifier. 1-7.

[6] Morissan. (2018, Juli 07). *Pengertian Pemilu*. Retrieved from www.silontong.com: http://www.silontong.com/2018/07/07/pengertian-pemilu/#

[7] Muktamar. (2015). Analisis Perbandingan Tingkat Akurasi Algoritma Naive Bayes Classifier dengan Correlated Naiva Bayes Classifier. 1-6.

[8] Pramana, S. P. (2013). Aplikasi Microsoft Office Excel 2010 untuk Menganalisis Butir Soal Pilihan Ganda. 1-3.

[9] Rancang Bangun Sistem Perpustakaan untuk Jurnal Elektronik. (2012). 1-9.

[10] Rosida, N. (2018). Perancangan Perangkat Lunak Bantu Sistem Penjualan Berbasis Aplikasi Dekstop Pada Cafe Instamie Pangandaran. 1-10.

[11] Saleh. (2015). Implementasi Metode Klasifikiras Naive Bayes Dalam Memprediksi Besarnya Pengunaan Listrik Rumah Tangga. 1-11.

[12] Purwati, N. (2015). e-voting . Kata Kunci:pemilu,e-voting, pilkada, 3(1).

[13] Sugiyono. (2017). *Metode Penelitian Kuantitatif, Kualitatif*, dan R&D. Bandung :Alfabeta, CV.

[14] Herny, F., & Zuliarso, E. (2012). Rancang Bangun Sistem Perpustakaan untuk Jurnal Elektronik. *Jurnal Teknologi Informasi DINAMIK*, *17*(2), 124–132.

[15] Subuh, Darmawan, Desviantina Rakhmiati, and Djoko Harsono. 2019. “Business Intelligence Untuk Mengukur Tingkat Kinerja Cabang Pada PT . XYZ Dengan Menggunakan Tools Pentaho Data Integration Dan Tableau Dekstop.” 38–46.