CHAPTER 4
IMPLEMENTATION AND RESULTS

4.1. Experiment Setup

In this Project the computer environment that I use is :

- Ryzen AMD 3
- RAM 4 GB
- Windows 10

4.2. Implementation

```
1. <html>
2. <body>
3. <p>Menampilkan Hasil Prediksi Penyakit Diabetes Mellitus Dari Algoritma
   Naive Bayes </p><br>
4. <p>Data Ujicoba Naive Bayes </p><br>
5.
6. <?php
7. $con=mysqli_connect("localhost","root","","pjdiabetes");              /*
   menghubungkan ke database yang sudah dibuat */
8. $no=1;
9. $qu=mysqli_query($con,"SELECT * FROM diabetesmelitus");
10.    while($dat=mysqli_fetch_array($qu)) {
11.    $no++;
12.    }
13.    $normal=0;
14.    $sakit=1;
15.    $x1=$_POST['x1'];
16.    $x2=$_POST['x2'];
17.    $x3=$_POST['x3'];
18.    $x4=$_POST['x4'];
19.    $x5=$_POST['x5'];
20.    $x6=$_POST['x6'];
21.    $x7=$_POST['x7'];
22.    $x8=$_POST['x8'];
23.
24.    $norm=mysqli_query($con,"SELECT * FROM diabetesmelitus WHERE y=0");
25.    $normal=mysqli_num_rows($norm);
26.    $sakt=mysqli_query($con,"SELECT * FROM diabetesmelitus WHERE y=1");
27.    $sakit=mysqli_num_rows($sakt);
28.    $v1c1=mysqli_query($con,"SELECT * FROM diabetesmelitus WHERE y=0 AND
   x1=$x1");
29.        $vr1c1=mysqli_num_rows($v1c1);
```

```
30.       $v1c2=mysqli_query($con,"SELECT * FROM diabetesmelitus WHERE y=1
   AND x1=$x1");
31.       $vr1c2=mysqli_num_rows($v1c2);
32.
33.       $v2c1=mysqli_query($con,"SELECT * FROM diabetesmelitus WHERE y=0
   AND x2=$x2");
34.       $vr2c1=mysqli_num_rows($v2c1);
35.       $v2c2=mysqli_query($con,"SELECT * FROM diabetesmelitus WHERE y=1
   AND x2=$x2");
36.       $vr2c2=mysqli_num_rows($v2c2);
37.
38.       $v3c1=mysqli_query($con,"SELECT * FROM diabetesmelitus WHERE y=0
   AND x3=$x3");
39.       $vr3c1=mysqli_num_rows($v3c1);
40.       $v3c2=mysqli_query($con,"SELECT * FROM diabetesmelitus WHERE y=1
   AND x3=$x3");
41.       $vr3c2=mysqli_num_rows($v3c2);
42.
43.       $v4c1=mysqli_query($con,"SELECT * FROM diabetesmelitus WHERE y=0
   AND x4=$x4");
44.       $vr4c1=mysqli_num_rows($v4c1);
45.       $v4c2=mysqli_query($con,"SELECT * FROM diabetesmelitus WHERE y=1
   AND x4=$x4");
46.       $vr4c2=mysqli_num_rows($v4c2);
47.
48.       $v5c1=mysqli_query($con,"SELECT * FROM diabetesmelitus WHERE y=0
   AND x5=$x5");
49.       $vr5c1=mysqli_num_rows($v5c1);
50.       $v5c2=mysqli_query($con,"SELECT * FROM diabetesmelitus WHERE y=1
   AND x5=$x5");
51.       $vr5c2=mysqli_num_rows($v5c2);
52.
53.       $v6c1=mysqli_query($con,"SELECT * FROM diabetesmelitus WHERE y=0
   AND x6=$x6");
54.       $vr6c1=mysqli_num_rows($v6c1);
55.       $v6c2=mysqli_query($con,"SELECT * FROM diabetesmelitus WHERE y=1
   AND x6=$x6");
56.       $vr6c2=mysqli_num_rows($v6c2);
57.
58.       $v7c1=mysqli_query($con,"SELECT * FROM diabetesmelitus WHERE y=0
   AND x7=$x7");
59.       $vr7c1=mysqli_num_rows($v7c1);
60.       $v7c2=mysqli_query($con,"SELECT * FROM diabetesmelitus WHERE y=1
   AND x7=$x7");
61.       $vr7c2=mysqli_num_rows($v7c2);
62.
63.       $v8c1=mysqli_query($con,"SELECT * FROM diabetesmelitus WHERE y=0
   AND x8=$x8");
64.       $vr8c1=mysqli_num_rows($v8c1);
65.       $v8c2=mysqli_query($con,"SELECT * FROM diabetesmelitus WHERE y=1
   AND x8=$x8");
66.       $vr8c2=mysqli_num_rows($v8c2);
67.    $has1=($normal/$no)*($vr1c1)/($normal)            *($vr2c1)/($normal)
   *($vr3c1)/($normal)
   *($vr4c1)/($normal)*($vr5c1)/($normal)*($vr6c1)/($normal)*($vr7c1)/($no
   rmal)*($vr8c1)/($normal);
```

```
68.       $has2=($sakit/$no)      *($vr1c2)/($sakit)      *($vr2c2)/($sakit)
   *($vr3c2)/($sakit)
   *($vr4c2)/($sakit)*($vr5c2)/($sakit)*($vr6c2)/($sakit)*($vr7c2)/($sakit
   )*($vr8c2)/($sakit);
69.    if($has1>$has2) {
70.       $out="Normal atau Tidak Terkena Diabetes"; }
71.       else {
72.       $out="Sakit atau Terkena Diabetes"; }
73.    ?>
74.    <br>
75.    <label> Jumlah Hamil: </label> <?php echo $x1; ?> <br>
76.    <label> Glukosa: </label> <?php echo $x2; ?> <br>
77.    <label> Tekanan Darah: </label> <?php echo $x3; ?> <br>
78.    <label> Ketebalan Kulit: </label> <?php echo $x4; ?> <br>
79.    <label> Insulin : </label> <?php echo $x5; ?> <br>
80.    <label> BMI : </label> <?php echo $x6; ?> <br>
81.    <label> Silsilah Fungsi DIabetes : </label> <?php echo $x7; ?> <br>
82.    <label> Usia : </label> <?php echo $x8; ?> <br>
83.    <label>    Hasil    Perhitungan    1    :    </label>    <?php    echo
   $has1."=(".$normal."/".$no.")*(".$vr1c1.")/".$normal.")*(".$vr2c1.")/".
   $normal.")*(".$vr3c1.")/".$normal.")*(".$vr4c1.")/".$normal.")*(".$vr5c
   1.")/".$normal.")*(".$vr6c1.")/".$normal.")*(".$vr7c1.")/".$normal.")*(
   ".$vr8c1.")/".$normal.")" ?> <br>
84.    <label>    Hasil    Perhitungan    2    :       </label>    <?php    echo
   $has2."=(".$sakit."/".$no.")*(".$vr1c2.")/".$sakit.")*(".$vr2c2.")/".$s
   akit.")*(".$vr3c2.")/".$sakit.")*(".$vr4c2.")/".$sakit.")*(".$vr5c2.")/
   ".$sakit.")*(".$vr6c2.")/".$sakit.")*(".$vr7c2.")/".$sakit.")*(".$vr8c2
   .")/".$sakit.")" ?> <br>
85.    <p>Hasil Klasifikasi Naive Bayes adalah : <?php echo $out; ?> </p>
   <br>
86.
87.
88.    </body>
89.    </html>
```

In this first view, it contains the source code when calculating using the Naïve Bayes method. Line 7 is a function that is used to connect to the database that has been created. Line 9 is a function to run queries to the mysql database. Line 24 contains the function norm variable mysqli_query to calculate the number of 0s that are different in the y variable in the diabetes mellitus table. Line 25 contains normal variables with the mysqli_num_rows command which functions to find out the number of 0 values in the y variable in the diabetes mellitus database. The 28th line contains commands to find and calculate the number of values 0 and 1 in column x1 in the diabetes mellitus table. Line 67 is a variable that contains commands to calculate the total number of columns x1 – x8 that have searched for a value of 0 or normal. Line 68 is a variable that contains commands to calculate the total number of columns x1 – x8 that have searched for a value of 1 or sick. Line 69 is a process used to determine whether the results of the input data have diabetes or not.

4.3. Results

In this section the authors conducted trials into 3 stages of trials. That is :

1. Try the influence of the amount of data used to detect diabetes mellitus and see the results of accuracy, precision, and recall.
2. Try using new data and old data to get diabetes mellitus detection results and see the results of accuracy, precision, and recall.
3. Looking for the dominant factors that affect the value of accuracy, precision, and recall.

1.1 Experiment by 769 Data Training, 580 Data Testing

Table 1 The variable table uses 769 Data Training, 580 Data Testing

Hasil/Label	F	T
F	**341**	**40**
T	**33**	**166**

TP : 166 FP : 33 FN : 40 TN : 341

ACCURACY : $\frac{TP+TN}{TP+TN+FP+FN} * 100\%$

$\frac{166+341}{166+33+40+341} * 100\% = 87\%$

RECALL : $\frac{TP}{FN+TP} * 100\%$

$\frac{166}{40+166} * 100\% = 81\%$

PRECISION: $\frac{TP}{FP+TP} * 100\%$

$\frac{166}{33+166} * 100\% = 83\%$

1.2 Experiment by 769 Data Training, 384 Data Testing

Table 2 The variable table uses 769 Data Training, 384 Data Testing

Hasil/Label	F	T
F	216	23
T	24	121

TP : 121 FP : 24 FN : 23 TN : 216

ACCURACY : $\frac{TP+TN}{TP+TN+FP+FN} * 100\%$

$\frac{121+216}{121+33+23+216} * 100\% = 86\%$

RECALL : $\frac{TP}{FN+TP} * 100\%$

$\frac{121}{121+23} * 100\% = 84\%$

PRECISION : $\frac{TP}{FP+TP} * 100\%$

$\frac{121}{24+121} * 100\% = 83\%$

1.3 Experiment by 769 Data Training, 50 Data Testing

Table 3 The variable table uses 769 Data Training, 50 Data Testing

Hasil/Label	F	T
F	18	7
T	2	23

TP : 23 FP : 2 FN : 7 TN : 18

ACCURACY : $\frac{TP+TN}{TP+TN+FP+FN} * 100\%$

$\frac{23+18}{23+2+7+18} * 100\% = 82\%$

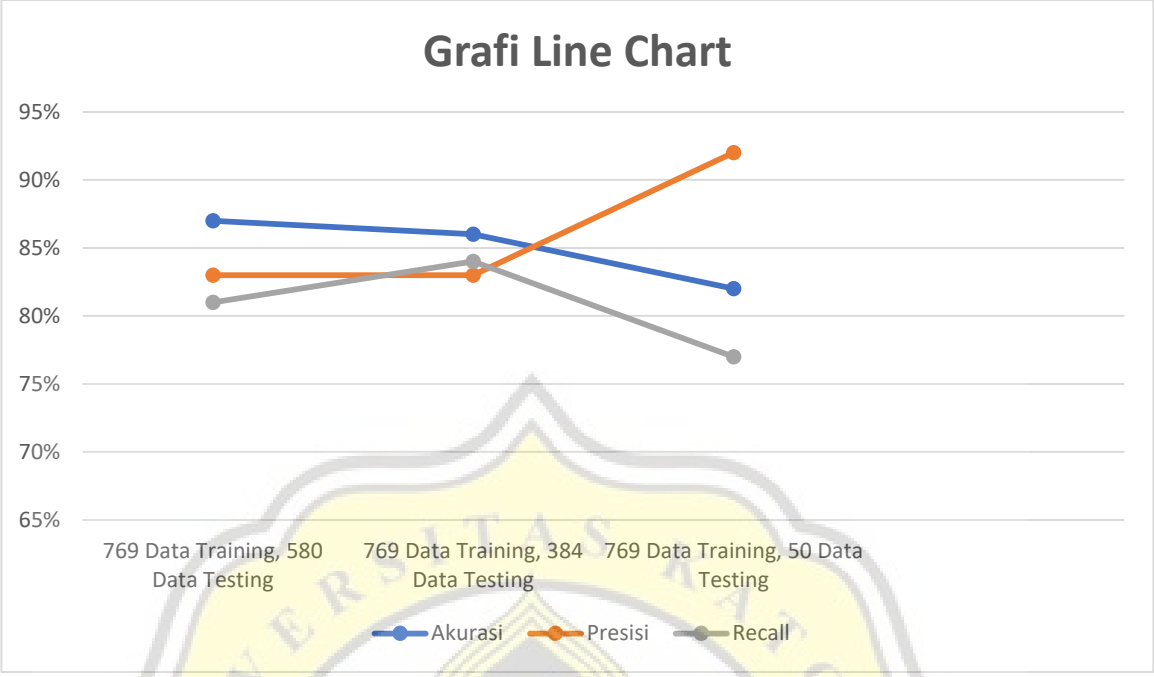
RECALL : $\frac{TP}{FN+TP} * 100\%$

$\frac{23}{7+23} * 100\% = 77\%$

PRECISION : $\frac{TP}{FP+TP} * 100\%$

$\frac{23}{2+23} * 100\% = 92\%$

No	Kasus Uji	Akurasi	Presisi	Recall
1	**769 Data Training, 580 Data Testing**	**87%**	**83%**	**81%**
2	**769 Data Training, 384 Data Testing**	**86%**	**83%**	**84%**
3	**769 Data Training, 50 Data Testing**	**82%**	**92%**	**77%**



- From the experimental results obtained, it can be seen that the lowest level of accuracy is when testing using 50 testing data with a result of 82%. Or the Naïve Bayes method can produce good accuracy results when more testing data is tested.
- From the experimental results obtained, it can be seen that the lowest level of precision is when testing using 580 testing data and 384 testing data with a result of 83%. Or the Naïve Bayes method can produce good precision results when the testing data being tested is small.
- From the experimental results obtained, it can be seen that the lowest recall rate is when testing using 50 data testing with a result of 77%. Or the Naïve Bayes method can produce good recall results when more testing data is tested.

2.1 Experiment by 769 Data Training, 50 data testing or old data

Table 4 The variable table uses 769 Data Training, 50 data testing or old data

Hasil/Label	F	T
F	18	7
T	2	23

TP : 23 FP : 2 FN : 7 TN : 18

ACCURACY : $\frac{TP+TN}{TP+TN+FP+FN} * 100\%$

$\frac{23+18}{23+2+7+18} * 100\% = 82\%$

RECALL : $\frac{TP}{FN+TP} * 100\%$

$\frac{23}{7+23} * 100\% = 77\%$

PRECISION : $\frac{TP}{FP+TP} * 100\%$

$\frac{23}{2+23} * 100\% = 92\%$

2.2 Experiment by 719 Data Training, 50 data testing or new data

Table 5 The variable table uses 719 Data Training, 50 data testing or new data

Hasil/Label	F	T
F	12	13
T	5	20

TP : 20 FP : 5 FN : 13 TN : 12

ACCURACY : $\frac{TP+TN}{TP+TN+FP+FN} * 100\%$

$\frac{20+12}{20+5+13+12} * 100\% = 64\%$

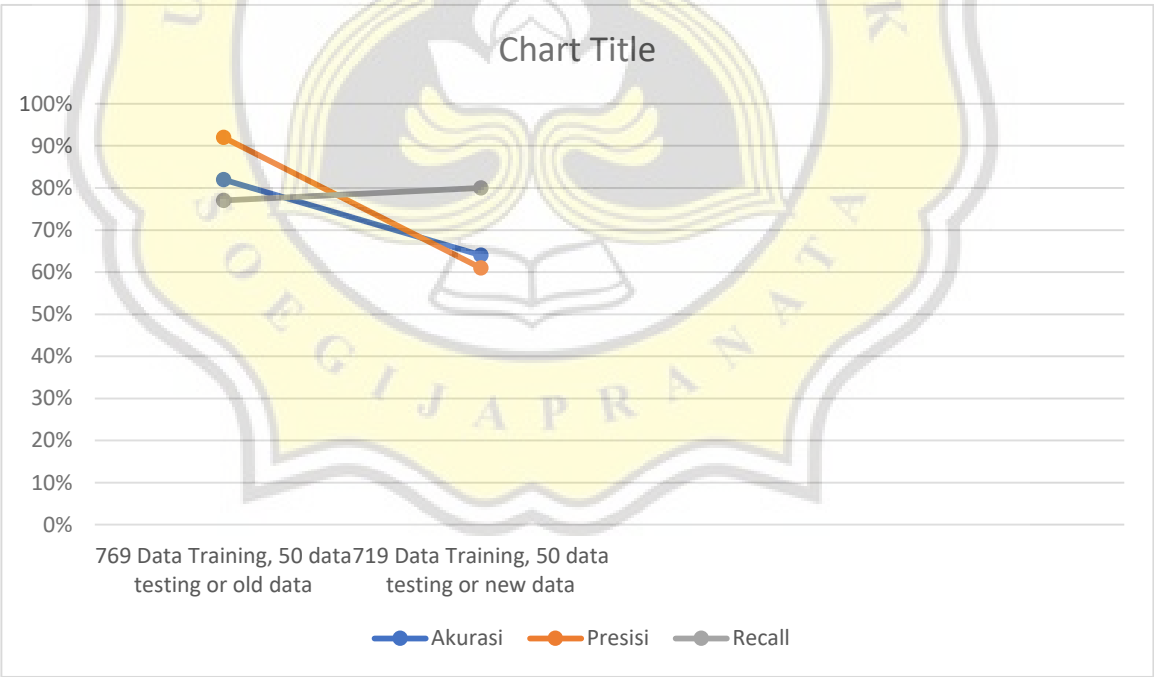
RECALL : $\frac{TP}{FN+TP} * 100\%$

$\frac{20}{13+20} * 100\% = 61\%$

PRECISION : $\frac{TP}{FP+TP} * 100\%$

$\frac{20}{5+20} * 100\% = 80\%$

No	Kasus Uji	Akurasi	Presisi	Recall
1	**769 Data Training, 50 data testing or old data**	**82%**	**92%**	**77%**
2	**719 Data Training, 50 data testing or new data**	**64%**	**61%**	**80%**



- From the experimental results obtained, it can be seen that the lowest level of accuracy is when testing using 50 testing data derived from new data with a result of 64%. Or the

Naïve Bayes method can produce good accuracy results when the testing data being tested comes from old data or data that already exists in the core data.

- From the experimental results obtained, it can be seen that the lowest level of precision is when testing using 50 testing data derived from new data with a result of 61%. Or the Naïve Bayes method can produce good precision results when the testing data being tested comes from old data or data that already exists in the core data.
- From the experimental results obtained, it can be seen that the lowest recall rate is when testing using 50 testing data that comes from existing core data or old data with a result of 77%. Or the Naïve Bayes method can produce good recall results when the testing data being tested comes from new data.

3.1 Experiment by eliminating the variable x1

Table 6 The deleted x1 variable table

Hasil/Label	F	T
F	450	51
T	47	221

TP : 221 FP : 47 FN : 51 TN : 450

ACCURACY : $\frac{TP+TN}{TP+TN+FP+FN} * 100\%$

$\frac{221+450}{221+450+47+51} * 100\% = 89\%$

RECALL : $\frac{TP}{FN+TP} * 100\%$

$\frac{221}{51+221} * 100\% = 81\%$

PRECISION : $\frac{TP}{FP+TP} * 100\%$

$\frac{221}{47+221} * 100\% = 83\%$

3.2 Experiment by eliminating the variable x2

Table 7 Table of deleted variable x2

Hasil/Label	F	T
F	**428**	**73**
T	**63**	**205**

TP : 205 FP : 63 FN : 73 TN : 428

ACCURACY : $\frac{TP+TN}{TP+TN+FP+FN} * 100\%$

$\frac{205+428}{205+428+63+73} * 100\% = 82\%$

RECALL : $\frac{TP}{FN+TP} * 100\%$

$\frac{205}{73+205} * 100\% = 74\%$

PRECISION : $\frac{TP}{FP+TP} * 100\%$

$\frac{205}{63+205} * 100\% = 77\%$

3.3 Experiment by eliminating the variable x3

Table 8 Table of deleted variable x3

Hasil/Label	F	T
F	**448**	**53**
T	**49**	**219**

TP : 219 FP : 49 FN : 53 TN : 448

ACCURACY : $\frac{TP+TN}{TP+TN+FP+FN} * 100\%$

$$\frac{219 + 448}{219+53 +448 +49} * 100\% = 87\%$$

RECALL : $\frac{TP}{FN+TP} * 100\%$

$$\frac{219}{53+219} * 100\% = 81\%$$

PRECISION : $\frac{TP}{FP+TP} * 100\%$

$$\frac{219}{219 + 49} * 100\% = 82\%$$

3.4 Experiment by eliminating the variable x4

Table 9 Table of deleted variable x4

Hasil/Label	F	T
F	453	48
T	49	219

TP : 219 FP : 49 FN : 48 TN : 453

ACCURACY : $\frac{TP+TN}{TP+TN+FP+FN} * 100\%$

$$\frac{219+453}{219+49+48+453} * 100\% = 87\%$$

RECALL : $\frac{TP}{FN+TP} * 100\%$

$$\frac{219}{219+48} * 100\% = 82\%$$

PRECISION : $\frac{TP}{FP+TP} * 100\%$

$\frac{219}{49+219} * 100\% = 82\%$

3.5 Experiment by eliminating the variable x5

Table 10 Removed x5 variable table

Hasil/Label	F	T
F	**438**	**63**
T	**60**	**208**

TP : 208 FP : 60 FN : 63 TN : 438

ACCURACY : $\frac{TP+TN}{TP+TN+FP+FN} * 100\%$

$\frac{208+438}{208+60+63+438} * 100\% = 84\%$

RECALL : $\frac{TP}{FN+TP} * 100\%$

$\frac{208}{63+208} * 100\% = 77\%$

PRECISION : $\frac{TP}{FP+TP} * 100\%$

$\frac{208}{60+208} * 100\% = 78\%$

3.6 Experiment by eliminating the variable x6

Table 11 Table of deleted variable x6

Hasil/Label	F	T

F	**450**	**51**
T	**52**	**216**

TP : 216 FP : 52 FN : 51 TN : 450

ACCURACY : $\frac{TP+TN}{TP+TN+FP+FN} * 100\%$

$\frac{216+450}{216+52+51+450} * 100\% = 87\%$

RECALL : $\frac{TP}{FN+TP} * 100\%$

$\frac{216}{51+216} * 100\% = 81\%$

PRECISION : $\frac{TP}{FP+TP} * 100\%$

$\frac{216}{216+52} * 100\% = 81\%$

3.7 Experiment by eliminating the variable x7

Table 12 Table of deleted variable x7

Hasil/Label	F	T
F	**445**	**56**
T	**46**	**222**

TP : 222 FP : 46 FN : 56 TN : 445

ACCURACY : $\frac{TP+TN}{TP+TN+FP+FN} * 100\%$

$\frac{222+445}{222+46+56+445} * 100\% = 87\%$

RECALL : $\frac{TP}{FN+TP} * 100\%$

$$\frac{222}{56+222} * 100\% = 80\%$$

PRECISION : $\frac{TP}{FP+TP} * 100\%$

$$\frac{222}{46+222} * 100\% = 83\%$$

3.8 Experiment by eliminating the x8 variable

Table 13 Table of deleted x8 variables

Hasil/Label	F	T
F	451	50
T	50	218

TP : 218 FP : 50 FN : 50 TN : 451

ACCURACY : $\frac{TP+TN}{TP+TN+FP+FN} * 100\%$

$$\frac{218+451}{218+50+50+451} * 100\% = 87\%$$

RECALL : $\frac{TP}{FN+TP} * 100\%$

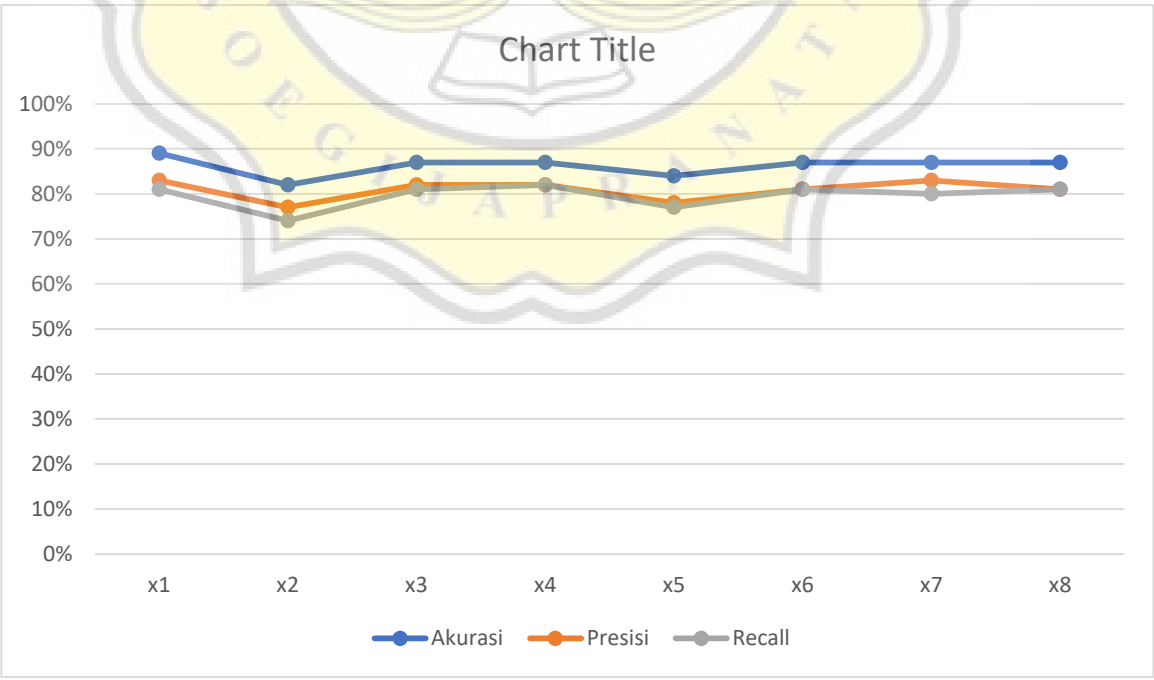
$$\frac{218}{50+218} * 100\% = 81\%$$

PRECISION : $\frac{TP}{FP+TP} * 100\%$

$$\frac{218}{50+218} * 100\% = 81\%$$

No	Kasus Uji	Akurasi	Presisi	Recall

1	**769 Data Training, 769 data testing without x1**	**89%**	**83%**	**81%**
2	**769 Data Training, 769 data testing without x2**	**82%**	**77%**	**74%**
3	**769 Data Training, 769 data testing without x3**	**87%**	**82%**	**81%**
4	**769 Data Training, 769 data testing without x4**	**87%**	**82%**	**82%**
5	**769 Data Training, 769 data testing without x5**	**84%**	**78%**	**77%**
6	**769 Data Training, 769 data testing without x6**	**87%**	**81%**	**81%**
7	**769 Data Training, 769 data testing without x7**	**87%**	**83%**	**80%**
8	**769 Data Training, 769 data testing without x8**	**87%**	**81%**	**81%**



y

- From the experimental results obtained, it can be seen that the lowest level of accuracy is when eliminating the variable x2 with an accuracy value of 82%.
- From the experimental results obtained, it can be seen that the lowest level of precision is when eliminating the variable x2 with an accuracy value of 74%.
- From the experimental results obtained, it can be seen that the lowest recall rate is when eliminating the variable x2 with an accuracy value of 77%.

4.4. Discussion

From the three experiments that have been carried out, it can be concluded that:

- The smaller the amount of data tested, the lower the accuracy value of using the Naïve Bayes method.
- If you use new data as part of a test, the accuracy value will be lower than old data or existing data.
- The most dominant factor in this test is the x2 variable. When the x2 variable is removed from the data column, the accuracy, precision, and recall values will automatically decrease.