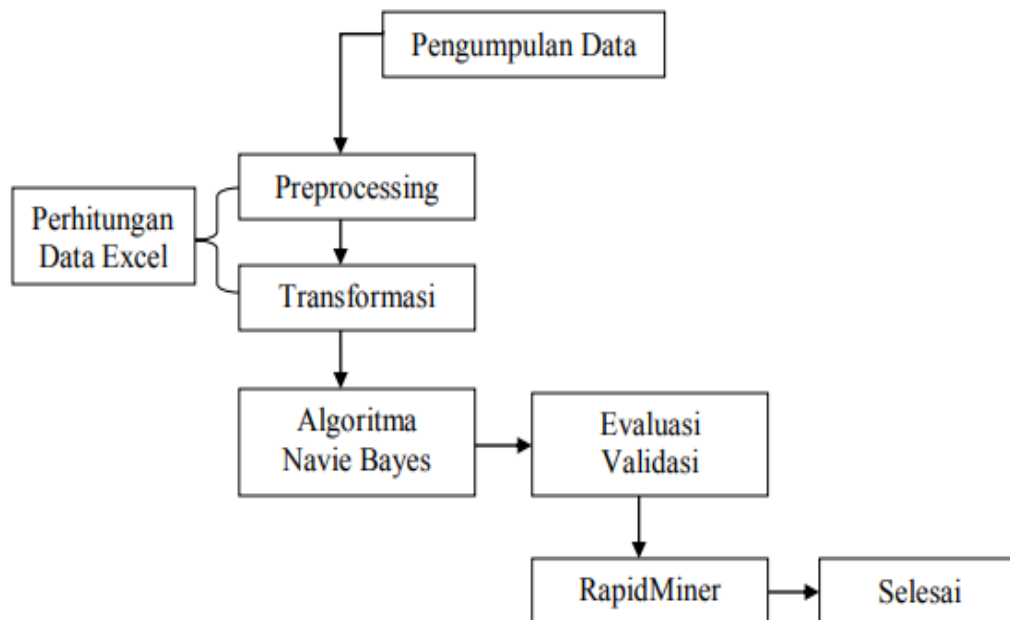


BAB III ANALISA DAN PERANCANGAN

3.1 Kerangka Kerja Penelitian

Kerangka kerja penelitian ini mencakup beberapa langkah penting yang dirancang untuk memastikan kelancaran dan keberhasilan implementasi sistem dalam mengklasifikasikan penerima bantuan pendidikan. Setiap langkah dalam penelitian memiliki tujuan yang jelas untuk memastikan bahwa data yang digunakan akurat dan relevan, dan hasil analisis dapat diterapkan secara efektif untuk pengambilan keputusan. Langkah-langkah utama dalam kerangka kerja penelitian ini bisa kita lihat pada gambar di bawah ini sebagai berikut:



Gambar 3.1 Karangan Kerja

Adapun penjelasan dari gambar Langkah-langkah untuk memeperjelas penelitian yang akan di lakutama dalam kerangka kerja penelitian ini adalah sebagai berikut:

1. Pengumpulan Data

Pengumpulan data adalah langkah penting yang mencakup pengambilan data siswa dan nilai akademik. Data ini akan digunakan untuk menganalisis kelayakan penerima bantuan.

2. *Preprocessing*

Proses pembersihan data mencakup penghapusan duplikasi, pengisian nilai yang hilang, dan pengoreksian kesalahan dalam data yang ada, guna memastikan bahwa data yang digunakan dalam analisis adalah akurat dan relevan.

3. Transformasi

Pada tahap ini, data yang telah dibersihkan diubah ke dalam format yang sesuai untuk analisis lebih lanjut, seperti pengkodean data kategori menjadi data numerik.

4. Penerapan Algoritma

Pada tahap ini, algoritma seperti Naïve Bayes.diterapkan untuk menganalisis data dan mengklasifikasikan siswa berdasarkan kriteria yang telah ditetapkan.

5. Evaluasi dan Validasi

Evaluasi dilakukan untuk menilai akurasi model yang dikembangkan, dengan menggunakan metrik seperti *Akurasi*, *Presisi*, *Recall*, dan *F1-Score* untuk mengukur performa model dalam memprediksi penerima bantuan.

3.2 Langkah-Langkah Naïve Bayes

Metode Naïve Bayes merupakan salah satu algoritma dalam teknik data mining yang didasarkan pada prinsip Teorema Bayes. Algoritma ini digunakan untuk melakukan klasifikasi dengan pendekatan *probabilistic* berdasarkan hubungan antar *variabel* dalam dataset. Keunggulan utama dari metode ini adalah kemampuannya dalam menangani dataset dengan jumlah data yang besar dan proses perhitungan yang relatif cepat. Dalam penelitian ini, penerapan metode Naïve Bayes dilakukan melalui beberapa tahapan yang mengacu pada konsep *Knowledge Discovery in Database* (KDD). KDD merupakan proses sistematis dalam mengekstraksi informasi yang bermanfaat dari data yang tersedia. Adapun tahapan yang dilakukan dalam Metode Naïve Bayes meliputi:

3.2.1 Pengumpulan Data

Pada tahapan ini merupakan tahapan utama yang akan dilakukan yaitu mengumpulkan data yang akan digunakan pada penelitian ini. Dengan adanya data, penelitian ini dapat dilakukan. Pada data penelitian yang akan digunakan terdapat 2 dataset yang pertama data sampel atau data *Testing* dan yang kedua data *Training*. Pada Tabel 3.1 dibawah ini merupakan data mentah yang belum di ubah sebagai berikut:

Tabel 3.1 Data Mentah Siswa Prestasi

Nama	Kelas	Nilai akhir	Tahun	Status Bantuan
x1	10	1292	2020	Tidak Dapat
x2	11	1292	2020	Tidak Dapat
x3	11	1290	2020	Tidak Dapat
x4	10	1342	2020	Dapat
x5	12	1426	2020	Dapat
x6	11	1295	2020	Tidak Dapat
x7	12	1194	2020	Tidak Dapat
x8	12	1450	2020	Dapat
x9	11	1340	2020	Dapat
x10	10	1284	2020	Tidak Dapat
x11	11	1420	2020	Dapat
x12	10	1299	2020	Tidak Dapat
x13	12	1403	2020	Dapat
x14	10	1365	2020	Dapat
x15	12	1346	2020	Dapat
x16	11	1410	2020	Dapat
x17	10	1290	2020	Tidak Dapat
x18	11	1362	2020	Dapat
x19	11	1289	2020	Tidak Dapat
x20	12	1292	2020	Tidak Dapat
x21	10	1193	2020	Tidak Dapat
x22	11	1199	2020	Tidak Dapat
x23	12	1468	2020	Dapat
x24	10	1199	2020	Tidak Dapat
x25	12	1440	2021	Dapat
x26	11	1190	2021	Tidak Dapat
x27	12	1345	2021	Dapat
x28	10	1460	2021	Dapat
x29	11	1285	2021	Tidak Dapat
x30	12	1189	2021	Tidak Dapat
x31	11	1289	2021	Tidak Dapat
x32	12	1410	2021	Dapat
x33	12	1402	2021	Dapat
x34	11	1195	2021	Tidak Dapat
x35	12	1298	2021	Tidak Dapat
x36	10	1394	2021	Tidak Dapat
x37	12	1456	2021	Dapat
x38	12	1325	2021	Dapat
x39	12	1420	2021	Dapat
x40	10	1430	2021	Dapat
x41	11	1342	2021	Dapat
x42	12	1195	2022	Tidak Dapat
x43	11	1297	2022	Tidak Dapat
x44	10	1292	2022	Tidak Dapat
x45	11	1362	2022	Dapat
x46	11	1491	2022	Tidak Dapat
x47	12	1403	2022	Dapat
x48	11	1381	2022	Dapat

Nama	Kelas	Nilai Akhir	Tahun	Status Bantuan
x49	10	1278	2022	Tidak Dapat
x50	11	1288	2022	Tidak Dapat
x51	10	1456	2022	Dapat
x52	11	1245	2022	Tidak Dapat
x53	12	1235	2022	Tidak Dapat
x54	10	1197	2022	Tidak Dapat
x55	12	1440	2022	Dapat
x56	10	1257	2022	Tidak Dapat
x57	11	1263	2022	Tidak Dapat
x58	11	1404	2022	Dapat
x59	10	1324	2022	Dapat
x60	11	1242	2022	Tidak Dapat
x61	12	1199	2022	Tidak Dapat
x62	10	1424	2022	Dapat
x63	11	1240	2022	Tidak Dapat
x64	12	1313	2022	Dapat
x65	12	1250	2023	Tidak Dapat
x66	10	1367	2023	Dapat
x67	12	1255	2023	Tidak Dapat
x68	10	1453	2023	Dapat
x69	11	1427	2023	Dapat
x70	10	1134	2023	Tidak Dapat
x71	12	1420	2023	Dapat
x72	11	1374	2023	Dapat
x73	11	1371	2023	Dapat
x74	12	1259	2023	Tidak Dapat
x75	11	1456	2023	Dapat
x76	12	1399	2023	Dapat
x77	11	1278	2023	Tidak Dapat
x78	11	1386	2023	Dapat
x79	10	1221	2023	Tidak Dapat
x80	12	1335	2023	Dapat
x81	11	1277	2023	Tidak Dapat
x82	11	1286	2023	Tidak Dapat
x83	11	1419	2023	Dapat
x84	11	1391	2023	Dapat
x85	10	1440	2023	Dapat
x86	11	1330	2023	Dapat
x87	12	1196	2023	Tidak Dapat
x88	12	1404	2024	Dapat
x89	11	1175	2024	Dapat
x90	12	1266	2024	Tidak Dapat
x91	11	1455	2024	Dapat
x92	10	1448	2024	Dapat
x93	10	1281	2024	Tidak Dapat
x94	11	1202	2024	Tidak Dapat
x95	10	1478	2024	Dapat
x96	11	1254	2024	Tidak Dapat
x97	11	1458	2024	Dapat

Nama	Kelas	Nilai Akhir	Tahun	Status Bantuan
x98	12	1396	2024	Dapat
x99	12	1441	2024	Dapat
x100	10	1198	2024	Tidak Dapat
x101	11	1442	2024	Dapat
x102	10	1382	2024	Dapat
x103	11	1279	2024	Tidak Dapat
x104	11	1289	2024	Tidak Dapat
x105	12	1278	2024	Tidak Dapat
x106	10	1310	2024	Dapat

Pada Tabel 3.1 merupakan data mentah yang diperoleh dari sekolah SMA Swasta Has Sepakat Bilih Hilir. Data diatas nantinya akan di *Preprocessing* untuk ditentukan data yang akan digunakan untuk dapat diolah.

2.8.4 *Preprocessing*

Pada tahapan ini merupakan proses yang bertujuan untuk melakukan seleksi dan pembersihan data, jika ada data yang *Invalid* dan tidak layak digunakan. Tetapi karena data yang diperoleh sudah sesuai, maka data sudah dapat digunakan.

Tabel 3.2 Data *Training* Penelitian Metode Naïve Bayes

Nama	Kelas	Nilai akhir	Tahun	Status Bantuan
x1	X	<1300	2020	Tidak Dapat
x2	X	<1300	2020	Tidak Dapat
x3	XI	<1300	2020	Tidak Dapat
x4	X	>1300	2020	Dapat
x5	XII	>1300	2020	Dapat
x6	XI	<1300	2020	Tidak Dapat
x7	XII	<1300	2020	Tidak Dapat
x8	XII	>1300	2020	Dapat
x9	XI	>1300	2020	Dapat
x10	X	>1300	2020	Tidak Dapat
x11	XI	>1300	2020	Dapat
x12	X	<1300	2020	Tidak Dapat
x13	XII	>1300	2020	Dapat
x14	X	>1300	2020	Dapat
x15	XII	>1300	2020	Dapat
x16	XI	>1300	2020	Dapat
x17	X	>1300	2020	Tidak Dapat

Nama	Kelas	Nilai Akhir	Tahun	Status Bantuan
x18	XI	>1300	2020	Dapat
x19	XI	<1300	2020	Tidak Dapat
x20	XII	<1300	2020	Tidak Dapat
x21	X	<1300	2020	Tidak Dapat
x22	XI	<1300	2020	Tidak Dapat
x23	XII	>1300	2020	Dapat
x24	X	<1300	2020	Tidak Dapat
x25	XII	>1300	2021	Dapat
x26	XI	<1300	2021	Tidak Dapat
x27	XII	>1300	2021	Dapat
x28	X	>1300	2021	Dapat
x29	XI	<1300	2021	Tidak Dapat
x30	XII	>1300	2021	Dapat
x31	XI	<1300	2021	Tidak Dapat
x32	XII	>1300	2021	Dapat
x33	XII	>1300	2021	Dapat
x34	XI	<1300	2021	Tidak Dapat
x35	X	<1300	2021	Tidak Dapat
x36	X	<1300	2021	Tidak Dapat
x37	XII	>1300	2021	Dapat
x38	XII	>1300	2021	Dapat
x39	XII	>1300	2021	Dapat
x40	X	>1300	2021	Dapat
x41	XI	>1300	2021	Dapat
x42	XII	<1300	2022	Tidak Dapat
x43	XI	<1300	2022	Tidak Dapat
x44	X	<1300	2022	Tidak Dapat
x45	XI	>1300	2022	Dapat
x46	XI	<1300	2022	Tidak Dapat
x47	XI	>1300	2022	Dapat
x48	X	>1300	2022	Dapat
x49	X	<1300	2022	Tidak Dapat
x50	XI	<1300	2022	Tidak Dapat
x51	X	>1300	2022	Dapat
x52	XI	<1300	2022	Tidak Dapat
x53	XII	<1300	2022	Tidak Dapat
x54	X	<1300	2022	Tidak Dapat
x55	XII	>1300	2022	Dapat
x56	X	<1300	2022	Tidak Dapat
x57	XI	<1300	2022	Tidak Dapat
x58	XI	>1300	2022	Dapat

Nama	Kelas	Nilai Akhir	Tahun	Status Bantuan
x59	X	>1301	2022	Dapat
x60	XI	<1300	2022	Tidak Dapat
x61	XII	<1301	2022	Tidak Dapat
x62	X	>1300	2022	Dapat
x63	XI	<1300	2022	Tidak Dapat
x64	XII	>1300	2022	Dapat
x65	X	<1300	2023	Tidak Dapat
x66	X	>1300	2023	Dapat
x67	XII	<1300	2023	Tidak Dapat
x68	X	>1300	2023	Dapat
x69	XI	>1300	2023	Dapat
x70	X	<1300	2023	Tidak Dapat
x71	XI	>1300	2023	Dapat
x72	XI	>1300	2023	Dapat
x73	XI	>1300	2023	Dapat
x74	XI	<1300	2023	Tidak Dapat
x75	XI	>1300	2023	Dapat
x76	XII	>1300	2023	Dapat
x77	XI	<1300	2023	Tidak Dapat
x78	XI	>1300	2023	Dapat
x79	X	<1300	2023	Tidak Dapat
x80	XII	>1300	2023	Dapat
x81	XI	<1300	2023	Tidak Dapat
x82	XI	<1300	2023	Tidak Dapat
x83	XI	>1300	2023	Dapat
x84	XI	>1300	2023	Dapat
x85	X	>1300	2023	Dapat
x86	XI	>1300	2023	Dapat
x87	XII	<1300	2023	Tidak Dapat
x88	XII	>1300	2024	Dapat
x89	XI	>1300	2024	Dapat
x90	XII	<1300	2024	Tidak Dapat
x91	XI	>1300	2024	Dapat
x92	X	>1300	2024	Dapat
x93	X	<1300	2024	Tidak Dapat
x94	XI	<1300	2024	Tidak Dapat
x95	X	>1300	2024	Dapat
x96	XI	<1300	2024	Tidak Dapat
x97	XI	>1300	2024	Dapat
x98	XII	>1300	2024	Dapat
x99	XII	>1300	2024	Dapat

Nama	Kelas	Nilai Akhir	Tahun	Status Bantuan
x100	X	<1300	2024	Tidak Dapat
x101	XI	>1300	2024	Dapat
x102	XI	>1300	2024	Dapat
x103	XI	<1300	2024	Tidak Dapat
x104	X	<1300	2024	Tidak Dapat
x105	XII	<1300	2024	Tidak Dapat
x106	X	>1300	2024	Dapat

Tabel 3.2 menunjukkan data sampel penelitian, di mana siswa dengan nilai >1300 berhak menerima bantuan, sedangkan siswa dengan nilai <1300 tidak menerima bantuan. Data ini akan digunakan dan dianalisis menggunakan metode Naïve Bayes untuk perhitungan lebih lanjut.

Tabel 3.3 Data *Testing*

Nama	Kelas	Nilai	Tahun	Status Bantuan
X1	XI	>1300	2022	Dapat
X2	XII	<1300	2023	Tidak Dapat
X3	XII	>1300	2024	Dapat
X4	XI	>1300	2024	Dapat
X5	XII	<1300	2024	Tidak Dapat
X6	XI	>1300	2024	Dapat
X7	X	>1300	2024	Dapat
X8	X	<1300	2024	Tidak Dapat
X9	XI	<1300	2024	Tidak Dapat
X10	X	>1300	2024	Dapat
X11	XI	<1300	2024	Tidak Dapat
X12	XI	>1300	2024	Dapat
X13	XII	>1300	2024	Dapat
X14	XII	>1300	2024	Dapat
X15	X	<1300	2024	Tidak Dapat
X16	XI	>1300	2024	Dapat
X17	XI	>1300	2024	Dapat
X18	XI	<1300	2024	Tidak Dapat
X19	X	<1300	2024	Tidak Dapat
X20	XII	<1300	2024	Tidak Dapat

Data pada Tabel 3.3 diatas merupakan data *Testing* yang akan digunakan untuk membantu proses pengolahan data. Pada Tabel 3.2 diatas nantinya akan

dipisah berdasarkan setiap atribut yang ada pada Tabel diatas. Untuk perubahannya akan diubah pada tahapan Transformasi.

3.2.3 Transformasi

Pada tahap ini akan digunakan untuk merubah data *Training* dari 1 Tabel menjadi beberapa Tabel berdasarkan dari atribut yang digunakan. untuk Tabelnya dapat dilihat di bawah ini.

Tabel 3.5 Atribut Kelas

Atribut	Partisi	Dapat	Tidak Dapat	P(Dapat)	P(Tidak Dapat)
Kelas	X	14	18	14/56	18/50
	XI	22	23	22/56	23/50
	XII	20	9	20/56	9/50
	Jumlah	56	50	100%	100%

Tabel 3.5 menunjukkan bahwa probabilitas penerimaan bantuan adalah 14/56 untuk kelas X, 22/56 untuk kelas XI, dan 20/56 untuk kelas XII. Sementara itu, probabilitas tidak menerima bantuan adalah 18/50 untuk kelas X, 23/50 untuk kelas XI, dan 9/50 untuk kelas XII.

Tabel 3.6 Atribut Nilai Akhir

Atribut	Partisi	Dapat	Tidak Dapat	P(Dapat)	P(Tidak Dapat)
Nilai Akhir	>1300	54	0	56/56	0/50
	<1300	0	50	0/56	50/50
	Jumlah	56	50	100%	100%

Tabel 3.6 menunjukkan bahwa semua siswa dengan nilai akhir di atas >1300 memiliki probabilitas 56/56 atau 100% untuk menerima bantuan, sedangkan siswa dengan nilai di bawah <1300 memiliki probabilitas 50/50 untuk tidak menerima bantuan. Hal ini menegaskan bahwa nilai akhir merupakan faktor utama dalam menentukan kelayakan penerimaan bantuan.

Tabel 3.7 Atribut Tahun

Atribut	Partisi	Dapat	Tidak Dapat	P(Dapat)	P(Tidak Dapat)
Tahun	2020	11	13	11/56	13/50
	2021	11	6	11/56	6/50
	2022	9	14	9/56	14/50
	2023	14	9	14/56	9/50
	2024	11	8	12/56	8/50
	Jumlah	56	50	100%	100%

Tabel 3.7 ini menunjukkan distribusi penerimaan bantuan berdasarkan tahun. Dari 56 siswa yang menerima bantuan, (11 berasal dari tahun 2020), (11 dari tahun 2021), (9 dari tahun 2022), (14 dari tahun 2023), dan (11 dari tahun 2024). Sementara itu, dari 50 siswa yang tidak menerima bantuan, (13 berasal dari tahun 2020), (6 dari tahun 2021), (14 dari tahun 2022), (9 dari tahun 2023), dan (8 dari tahun 2024).

Probabilitas penerimaan bantuan dihitung sebagai (11/56 untuk tahun 2020), (11/56 untuk tahun 2021), (9/56 untuk tahun 2022), (14/56 untuk tahun 2023), dan (12/56 untuk tahun 2024). Sedangkan probabilitas tidak menerima bantuan adalah (13/50 untuk tahun 2020), (6/50 untuk tahun 2021), (14/50 untuk tahun 2022), (9/50 untuk tahun 2023), dan (8/50 untuk tahun 2024).

Tabel 3.8 Atribut Kategori

Atribut			P(Dapat)P(Tidak Dapat)
Kategori	Dapat	54	56/106
	Tidak Dapat	50	50/106
	Jumlah	104	100%

Tabel 3.8 menggambarkan distribusi penerimaan bantuan berdasarkan kategori. Dari total 106 siswa, sebanyak 54 siswa menerima bantuan dengan probabilitas 54/106, sementara 50 siswa tidak menerima bantuan dengan probabilitas 50/106. Jumlah ini mencerminkan keseluruhan populasi yang dianalisis dalam penelitian.

3.2.4 Data Mining

Pada tahapan ini merupakan pengolahan ataupun perhitungan data yang akan dilah pada *Data Mining* dengan menggunakan metode Naïve Bayes. Untuk perhitungannya, penulis akan menggunakan data sampel dari data testing yaitu data X1.

$$P(\text{Kategori}) = P(\text{Kelas}|\text{XII}) \times P(\text{Nilai Akhir} > 13000) \times P(\text{Tahun}|\text{2022}) \times P(\text{Kategori}|\text{Dapat})$$

$$\begin{aligned} P(\text{Dapat}) &= P(\text{Laki-Laki}|\text{Dapat}) \times P(\text{XII}|\text{Dapat}) \times P(>13000|\text{Dapat}) \times \\ &\quad P(\text{2022}|\text{Dapat}) \times P(\text{Kategori}|\text{Dapat}) \\ &= \left(\frac{20}{56}\right) \times \left(\frac{56}{56}\right) \times \left(\frac{9}{56}\right) \times \left(\frac{56}{106}\right) \\ &= 0,030323 \text{ (Nilai Dapat)} \end{aligned}$$

$$P(\text{Kategori}) = P(\text{Kelas}|\text{XII}) \times P(\text{Nilai Akhir} < 13000) \times P(\text{Tahun}|\text{2022}) \times P(\text{Kategori}|\text{Tidakdapat})$$

$$\begin{aligned} P(\text{Tidak Dapat}) &= P(\text{XII}|\text{TidakDapat}) \times P(>13000|\text{TidakDapat}) \times \\ &\quad P(\text{2022}|\text{TidakDapat}) \times P(\text{Kategori}|\text{TidakDapat}) \\ &= \left(\frac{9}{50}\right) \times \left(\frac{0}{50}\right) \times \left(\frac{14}{50}\right) \times \left(\frac{50}{106}\right) \\ &= 0 \text{ (Nilai Tidak Dapat)} \end{aligned}$$

Dari hasil perhitunga data pada Metode Naïve Bayes dengan menghitung data dari X1, bahwasanya dari hasil yang di peroleh nilai yang (Dapat) yaitu, 0,03052800 dan yang (Tidak Dapat) sebesar 0, jadi hasil pengolahan data bahwasanya data X1 (Dapat Bantuan).

3.2.5 Evaluasi Validasi

Pada data ini merupakan tahapan Evaluasi ataupun Uji Performa Metode Naïve Bayes pada data mining. Untuk evaluasinya dapat dilihat Tabel dibawah ini.

3.2.6 Uji Performa Dengan Counfusion Matrix

Tabel3.9 Performa Dengan Counfusion Matrix

		Dapat	Tidak Dapat	Σ
Actual	Dapat	10	1	11
	Dapat Tidak	1	8	9
	Σ	11	9	20

Hasil *True Positive (TP)* adalah 11 dan *False Negative(FN)* adalah 9.

Maka Nilai *Akurasi, Presisi Dan Recall* adalah sebagai berikut:

$$Accuracy = \frac{10+8}{10+8+1+1} \times 100\% \quad \text{Then the Accuracy value} = 90,00\%$$

$$Presisi = \frac{10}{11+1} \times 100\% \quad \text{Then the Precision value} = 88,89\%$$

$$Recall = \frac{10}{10+1} \times 100\% \quad \text{Then the Recall value} = 88,89\%$$

Hasil akurasi yang diperoleh sangat bagus, dan dapat di kesimpulan bahwa Metode Naïve Bayes dapat digunakan untuk melakukan pengolahan data siswa perstasi yang dapat bantuan.