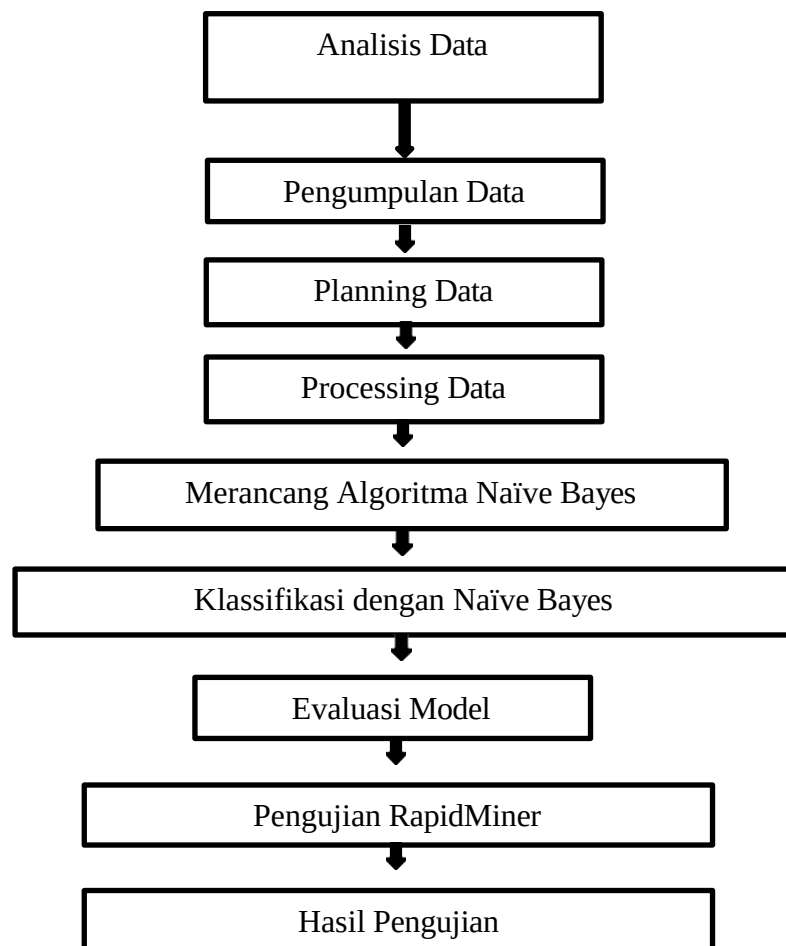


BAB III

METODE PENELITIAN

3.1 Tahap Analisa

Tahap analisa merupakan proses awal yang sangat penting dalam penelitian ini, karena bertujuan untuk memahami struktur data, mengidentifikasi kebutuhan sistem, serta merancang pendekatan algoritmik yang disajikan pada gambar berikut:



Gambar 3. 1 Diagram Penelitian

Keterangan :

1. Pengumpulan Data

Proses awal untuk mengumpulkan data akademik mahasiswa (misalnya: IPK, kehadiran, dll) dari sumber terpercaya seperti database kampus.

2. Planning Data

Menentukan data apa saja yang relevan, melakukan seleksi fitur yang akan digunakan, serta merancang strategi pembersihan dan transformasi data.

3. Processing Data

Melakukan preprocessing: normalisasi, pembersihan data dari duplikat/kesalahan, dan pengolahan agar data siap dianalisis oleh algoritma.

4. Perancangan Algoritma Naïve Bayes

Menyusun alur klasifikasi untuk memprediksi kemungkinan mahasiswa lulus atau tidak, berdasarkan probabilitas fitur-fitur data.

5. Klasifikasi dengan Naïve Bayes

Data yang sudah dikelompokkan (clustering) akan diproses lebih lanjut untuk dilakukan klasifikasi akhir (Lulus/Tidak Lulus) oleh algoritma Naïve Bayes.

6. Hasil Pengujian & RapidMiner

Menyajikan hasil akhir dari proses klasifikasi dan clustering yang telah diuji, dengan bantuan tool seperti RapidMiner untuk visualisasi dan validasi.

3.2 Proses Analisis Data

Proses analisis data merupakan tahapan penting dalam penelitian ini karena menjadi dasar dalam membangun model prediksi kelulusan mahasiswa. Proses ini

dilakukan secara sistematis mulai dari pengumpulan hingga persiapan data agar siap digunakan oleh algoritma Naïve Bayes. Adapun langkah-langkah dalam proses analisis data adalah sebagai berikut:

1) Pengumpulan Data

Data dikumpulkan dari sumber internal kampus, data yang diambil mencakup:

- a. Data mahasiswa
- b. Jenis Kelamin
- c. IPK
- d. Kehadiran
- e. Status Predikat (Lulus/Tidak Lulus)

Adapun tercantum dalam tabel berikut ini:

Tabel 3. 1 Data Penelitian Prodi Sistem Informasi

No	Nama Mahasiswa	Jenis Kelamin	IPK	Kehadiran (%)	Predikat Lulus/Tidak Lulus
1	Farhan Zam-Zami Nasution	Laki-laki	3.45	88	Lulus
2	Muhammad Zakuan Ritonga	Laki-laki	3.20	82	Lulus
3	Nadira Jannah Adeni Lubis	Perempuan	3.75	91	Lulus
4	Nova Tresia Pasaribu	Perempuan	3.60	87	Lulus
5	Roby Gusmawan Sirait	Laki-laki	2.95	74	Tidak Lulus
6	Sabrina Erika Kristina Sianturi	Perempuan	3.80	93	Lulus
7	Triya Agustina	Perempuan	2.85	72	Tidak Lulus
8	Nurul Fatma	Perempuan	3.30	84	Lulus
9	Sri Melyani	Perempuan	3.50	89	Lulus
10	Rahmayuni Syah	Perempuan	3.10	78	Lulus
11	Ardiansyah	Laki-laki	2.90	71	Tidak Lulus

	Hasibuan				
12	Diko Pradana Sihotang	Laki-laki	3.65	90	Lulus
13	Angga Prayoga Patriya	Laki-laki	3.40	85	Lulus
14	Yayuk Ariani	Perempuan	3.85	94	Lulus
15	Afrian Alfariz	Laki-laki	3.25	81	Lulus
16	Annisa Rizky Simangunsong	Perempuan	3.70	92	Lulus
17	Dinda Ajeng Lestari	Perempuan	2.80	70	Tidak Lulus
18	Fitri Andriani	Perempuan	3.55	86	Lulus
19	Hardiman Purba	Laki-laki	3.15	79	Lulus
20	Muhammad Alim	Laki-laki	3.90	95	Lulus
21	Nur Dilla	Perempuan	3.35	83	Lulus
22	Rasid Siddik	Laki-laki	3.75	90	Lulus
23	Sonia Sri Muliani	Perempuan	3.20	80	Lulus
24	Sri Rahmawati	Perempuan	2.95	75	Tidak Lulus
25	Suhaylah	Perempuan	3.60	88	Lulus
26	Veranica Sinaga	Perempuan	3.45	85	Lulus
27	Reza Puspita Sari Pohan	Perempuan	3.80	92	Lulus
28	Rikardo Lasroha Simamora	Laki-laki	2.80	70	Tidak Lulus
29	Arya Widana	Laki-laki	3.30	82	Lulus
30	Adam Wirayuda	Laki-laki	3.50	87	Lulus
31	Ade Eka Febriyanti	Perempuan	3.10	78	Lulus
32	Adek Santika Sari	Perempuan	2.90	72	Tidak Lulus
33	Adinda Puspita Sari	Perempuan	3.65	89	Lulus
34	Aditya Ruli Prayoga	Laki-laki	3.40	84	Lulus
35	Anggi Sofiani Harahap	Perempuan	3.85	93	Lulus
36	Arief Syahputra	Laki-laki	3.25	81	Lulus
37	Arvida	Perempuan	3.70	91	Lulus
38	Aulia Rahma	Perempuan	2.85	73	Tidak Lulus
39	Aulia Salsabila	Perempuan	3.55	86	Lulus
40	Cinrewina Purba	Perempuan	3.15	79	Lulus
41	Dana Rindiani	Perempuan	3.90	95	Lulus
42	Devi Lestari Hutagalung	Perempuan	3.35	83	Lulus
43	Diana Indriani	Perempuan	3.75	90	Lulus

	Rambe				
44	Dinda Julia Arfah	Perempuan	3.20	80	Lulus
45	Eka Nurjannah	Perempuan	2.95	75	Tidak Lulus
46	Elfy Rahmayani Lubis	Perempuan	3.60	88	Lulus
47	Eva Hermika	Perempuan	3.45	85	Lulus
48	Fitri Rahmadani	Perempuan	3.80	92	Lulus
49	Gunawan	Laki-laki	2.80	70	Tidak Lulus
50	Hana Tasya	Perempuan	3.30	82	Lulus
51	Khodijah Nasution	Perempuan	3.50	87	Lulus
52	Laila Sari	Perempuan	3.10	78	Lulus
53	Melisa	Perempuan	2.90	72	Tidak Lulus
54	Murni Zaliah Nurpa	Perempuan	3.65	89	Lulus
55	Nia Putri Panjaitan	Perempuan	3.40	84	Lulus
56	Nova Indayanti	Perempuan	3.85	93	Lulus
57	Novira Dwi Andini	Perempuan	3.25	81	Lulus
58	Nur Putri Andriani	Perempuan	3.70	91	Lulus
59	Nur'aini Indah Syahfitri	Perempuan	2.85	73	Tidak Lulus
60	Putri Anggraini Lestari	Perempuan	3.55	86	Lulus
61	Rahma Faradilla	Perempuan	3.15	79	Lulus
62	Rahmadani Ritonga	Perempuan	3.90	95	Lulus
63	Risky Amaldi Harahap	Laki-laki	3.35	83	Lulus
64	Rizaldi Ardiansyah	Laki-laki	3.75	90	Lulus
65	Rizki Amansyah	Laki-laki	3.20	80	Lulus
66	Salman Al Farisi	Laki-laki	2.95	75	Tidak Lulus
67	Surya Anggara Rambe	Laki-laki	3.60	88	Lulus
68	Syafira Eka Wardani	Perempuan	3.45	85	?
69	Tengku Irwan Aziz	Laki-laki	3.80	92	?
70	Tomi Hidayat	Laki-laki	2.80	70	?
71	Vina Tiara Mahani	Perempuan	3.30	82	?
72	Wulan Syafitri	Perempuan	3.50	87	?

73	Yuni Franata Sinurat	Perempuan	3.10	78	?
74	Yuni Saputri	Perempuan	2.90	72	?
75	Lianah	Perempuan	3.65	89	?
76	Fahreza Br Tumpul	Laki-laki	3.40	84	?
77	Ruswanto	Laki-laki	3.85	93	?
78	Aldimas Adi Anang	Laki-laki	3.25	81	?
79	Taufik Molid Hidayat Hasibuan	Laki-laki	3.70	91	?
80	Nelvi Nurrizqi M	Perempuan	2.85	73	?
81	Alwidah Yani Sipahutar	Perempuan	3.55	86	?
82	Irma Suriani Tambunan	Perempuan	3.15	79	?
83	Restu Fauzy Naibaho	Laki-laki	3.90	95	?
84	Nia Aulia Sari Lubis	Perempuan	3.35	83	?
85	Belinda Ray Silean	Perempuan	3.75	90	?
86	Putra Andika Ritonga	Laki-laki	3.20	80	?
87	Mirananda Ambarita	Perempuan	2.95	75	?
88	Nurhalimah Tambunan	Perempuan	3.60	88	?
89	Rifki Agustiawan Juhri	Laki-laki	3.45	85	?
90	Muhammad Adlin Hasibuan	Laki-laki	3.80	92	?
91	Shondy Raja Ferdinand Sinurat	Laki-laki	2.80	70	?
92	Syahrol Sulaiman Ritonga	Laki-laki	3.30	82	?
93	Rizka Nurfatni Sipahutar	Perempuan	3.50	87	?
94	Tongku Hamonangan Harahap	Laki-laki	3.10	78	?
95	Holong Marhula Sihite	Laki-laki	2.90	72	?
96	Eggi Ok Pernanda Gaja	Laki-laki	3.65	89	?

97	Nabila Syah Fitria	Perempuan	3.40	84	?
----	-----------------------	-----------	------	----	---

Sumber: Data Penelitian 2025

Tabel 3. 2 Data Penelitian Prodi Agroteknologi

No	Nama Mahasiswa	Jenis Kelamin	IPK	Kehadiran (%)	Predikat Lulus/Tidak Lulus
1	Agus Susanto	Laki-laki	3.45	85	Lulus
2	Andika Aldiana Agung	Laki-laki	3.20	80	Lulus
3	Arya Wiryansah	Laki-laki	3.75	90	Lulus
4	Ade Maya Sari E. Siregar	Perempuan	3.60	88	Lulus
5	Dedek Prayudha	Perempuan	2.95	75	Tidak Lulus
6	Berna Sirajudin	Perempuan	3.80	92	Lulus
7	Hari Ks Luar Adek	Laki-laki	2.80	70	Tidak Lulus
8	Julfan Efendi Nasution	Laki-laki	3.30	82	Lulus
9	Muhammad Riflan	Laki-laki	3.50	87	Lulus
10	Syar Yudi	Laki-laki	3.10	78	Lulus
11	Uli Amri	Laki-laki	2.90	72	Tidak Lulus
12	Zelvin Anggaka	Laki-laki	3.65	89	Lulus
13	Herliansyah Siregar	Laki-laki	3.40	84	Lulus
14	Bastian Fernandes Panggabean	Laki-laki	3.85	93	Lulus
15	Samuel Cristoper	Laki-laki	3.25	81	Lulus
16	Sri Wartini	Perempuan	3.70	91	Lulus
17	Pula Supar Rahmap	Perempuan	2.85	73	Tidak Lulus
18	Akbar Nur'aziz Nasution	Laki-laki	3.55	86	Lulus
19	Andy Ali Ridwan Nasution	Laki-laki	3.15	79	Lulus
20	Annis Hidayat Siregar	Perempuan	3.90	95	Lulus
21	Benni Janen Jaya Hutabarat	Laki-laki	3.35	83	Lulus
22	Bryang Michael Purba	Laki-laki	3.75	90	Lulus

23	Cayya Putra Febriansyahreza	Laki-laki	3.20	80	Lulus
24	Dicky Ariyapurna	Laki-laki	2.95	75	Tidak Lulus
25	Farid Rolli Siregar	Laki-laki	3.60	88	Lulus
26	Ferbita Damela Sinaga	Perempuan	3.45	85	Lulus
27	Herdiansyah Siregar	Laki-laki	3.80	92	Lulus
28	Iliham	Laki-laki	2.80	70	Tidak Lulus
29	Iul Arfandi	Laki-laki	3.30	82	Lulus
30	Khairil Anwar Nasution	Laki-laki	3.50	87	Lulus
31	M. Aziz Lubis	Laki-laki	3.10	78	Lulus
32	Malik Abdul Azis	Laki-laki	2.90	72	Tidak Lulus
33	Mhd. Fiksy Syahirana Purba	Laki-laki	3.65	89	Lulus
34	Muhammad Fadhil Rudi Kotonga	Laki-laki	3.40	84	Lulus
35	Muhammad Fauzi	Laki-laki	3.85	93	Lulus
36	Nico Armando	Laki-laki	3.25	81	Lulus
37	Nuryasmin	Perempuan	3.70	91	Lulus
38	Pasidon Siburan	Laki-laki	2.85	73	Tidak Lulus
39	Prio Aditya	Laki-laki	3.55	86	Lulus
40	Raimda Supraweteno	Laki-laki	3.15	79	Lulus
41	Raimdy Febayat	Laki-laki	3.90	95	Lulus
42	Ramadit Hidayat	Laki-laki	3.35	83	Lulus
43	Ramdhan	Laki-laki	3.75	90	Lulus
44	Reza Gilang Ramadhan	Laki-laki	3.20	80	Lulus
45	Rissi Deli Chandra	Perempuan	2.95	75	Tidak Lulus
46	Risna Dwiandri Yanti	Perempuan	3.60	88	Lulus
47	Saila Hassilian	Perempuan	3.45	85	Lulus
48	Saroh Zaitun Sayadah	Perempuan	3.80	92	Lulus
49	Suherman Prayuda	Laki-laki	2.80	70	?
50	Syafiir Azmi	Laki-laki	3.30	82	?

51	Tho Reafy	Laki-laki	3.50	87	?
52	Tho Pria Pratikno	Laki-laki	3.10	78	?
53	Tika Ali Rahmat Talipo	Perempuan	2.90	72	?
54	Verdy Anggriansyah	Laki-laki	3.65	89	?
55	Yohanna Rl Hutagalung	Perempuan	3.40	84	?
56	Herdi Prayoli	Laki-laki	3.85	93	?
57	Danang Aidil Rivanda	Laki-laki	3.25	81	?
58	Rolesbono Siregar	Laki-laki	3.70	91	?
59	Yuti Arfiyanti	Perempuan	2.85	73	?
60	Lisa Aklia Rambe	Perempuan	3.55	86	?
61	Intan Kurnia Dewi	Perempuan	3.15	79	?
62	Muhammad Sutryoso	Laki-laki	3.90	95	?
63	Hilman Sagani	Laki-laki	3.35	83	?
64	Rahmad Syaputra Siregar	Laki-laki	3.75	90	?
65	Tobias Alfred Lumban Batu	Laki-laki	3.20	80	?
66	Ardiansyah Putra Hasibuan	Laki-laki	2.95	75	?
67	Resti Fitriyanti	Perempuan	3.60	88	?
68	Dewi Raya Agpandri	Perempuan	3.45	85	?
69	Wira Adi Dharma	Laki-laki	3.80	92	?

Sumber: Data Penelitian 2025

Tabel 3. 3 Data Penelitian Prodi Manajemen Informatika

No	Nama Mahasiswa	Jenis Kelamin	IPK	Kehadiran (%)	Predikat Lulus/Tidak Lulus
1	Arini Mawaddah	Perempuan	3.85	96	Lulus
2	Diana Juni Triani	Perempuan	3.10	90	Lulus
3	Fitri Febriyani Hasibuan	Perempuan	2.85	82	Lulus
4	Nanda Fahrezi Munazhif	Laki-laki	2.45	79	Lulus

5	Putri Violita	Perempuan	2.60	87	Tidak Lulus
6	Putri Talia Cantika	Perempuan	2.75	73	Tidak Lulus
7	Zakia El Husna Siregar	Perempuan	3.25	91	Lulus
8	Rico Fadly Nasution	Laki-laki	3.70	94	Lulus
9	Sanny Khairani Lubis	Perempuan	3.90	97	Lulus
10	Selbi Amanda Aritonang	Perempuan	2.95	85	Lulus
11	Tiara Syavitri Rambe	Perempuan	3.50	93	Lulus
12	Zainatun Nisa	Perempuan	3.60	89	Lulus
13	Indri Cahaya Indah	Perempuan	3.80	92	Lulus
14	Sabdi Albi Hasibuan	Laki-laki	2.20	77	Lulus
15	Safrina Maizura	Perempuan	2.98	88	Lulus
16	Bella Rahmawati	Perempuan	3.25	90	Lulus
17	Mayang Sari	Perempuan	2.75	82	Lulus
18	Adinda Pratiwi Siregar	Perempuan	3.45	86	Lulus
19	Hubban Arfi Pratama	Laki-laki	2.10	79	Lulus
20	Fransisco Alexander Sinaga	Laki-laki	3.88	95	Lulus
21	Deni Mashuri	Laki-laki	2.95	84	Lulus
22	Diah Kusma Sari	Perempuan	1.85	80	?
23	Lorena Tanjung	Perempuan	3.20	88	?
24	Muhammad Syahrial Pane	Laki-laki	2.50	86	?
25	Sunardi	Laki-laki	3.15	91	?
26	Fitriani	Perempuan	2.00	75	?
27	Rizky Pratamasyah Putra	Laki-laki	3.90	98	?
28	Nurhizriani Hasibuan	Perempuan	2.65	85	?
29	Lenny Syahputri	Perempuan	3.55	93	?
30	Shinta Clara Sembiring	Perempuan	3.00	80	?

Sumber: Data Penelitian 2025

Tabel 3. 4 Data Penelitian Prodi Teknologi Informasi

No	Nama Mahasiswa	Jenis Kelamin	IPK	Kehadiran (%)	Predikat Lulus/Tidak Lulus
1	Ade Elyani Siregar	Perempuan	3.72	95	Lulus
2	Aditya Guna Dharma	Laki-laki	2.85	82	Lulus
3	Alwi Ardiansyah	Laki-laki	3.40	88	Lulus
4	Ega Mei Prianto	Laki-laki	2.55	76	Tidak Lulus
5	Eri Putra Pratama Hura	Laki-laki	2.40	68	Tidak Lulus
6	Fitrialdi Azhari	Laki-laki	3.15	90	Lulus
7	Reva Azhari Amirul Hafiz	Laki-laki	3.78	96	Lulus
8	Riski Febriyan Pasaribu	Laki-laki	2.65	84	?
9	Yogi Gunawan	Laki-laki	2.98	87	?
10	Tri Atmaja	Laki-laki	3.00	92	?
11	Yusnia Sari Ritonga	Perempuan	3.85	97	?

Sumber: Data Penelitian 2025

2) Seleksi dan Pembersihan Data (*Preprocessing*)

Pada tahap ini dilakukan:

- Seleksi atribut seperti hanya data yang relevan yang akan digunakan.
- Pembersihan data seperti menghapus data duplikat, memperbaiki nilai yang kosong atau tidak valid.
- Normalisasi seperti skala nilai seperti IPK dan kehadiran dinormalisasi agar bisa diproses bersama dalam algoritma K-Means.

3) Transformasi Data

Data mentah diubah menjadi format yang dapat digunakan oleh algoritma.

Misalnya:

- Nilai dikonversi menjadi format numerik atau kategorikal.

b. Label target dikodekan (misalnya: Lulus = 1, Tidak Lulus = 0).

2) Klasifikasi Data dengan Naïve Bayes

Setelah proses clustering, data digunakan untuk membangun model klasifikasi menggunakan algoritma Naïve Bayes. Algoritma ini menghitung probabilitas kelulusan berdasarkan fitur-fitur yang tersedia dan memberikan prediksi apakah seorang mahasiswa berpotensi lulus atau tidak.

3) Evaluasi Model menggunakan RapidMiner

Model yang dibangun dievaluasi menggunakan metrik seperti:

a. Akurasi

b. F1-score

3.2.1 Algoritma Naïve Bayes

Algoritma Naïve Bayes digunakan dalam tahap klasifikasi, yaitu untuk menentukan apakah seorang mahasiswa memiliki peluang untuk lulus atau tidak berdasarkan data fitur yang dimiliki.

Langkah-langkah implementasi:

1) Menghitung Probabilitas Awal (*Prior*):

Menentukan proporsi mahasiswa yang lulus dan tidak lulus. Yaitu proporsi jumlah mahasiswa yang lulus dan tidak lulus terhadap total data:

$$P(Lulus) = \frac{\text{Jumlah mahasiswa lulus}}{\text{Total Mahasiswa}}$$

$$P(Tidak Lulus) = \frac{\text{Jumlah mahasiswa tidak lulus}}{\text{Total Mahasiswa}}$$

2) Menghitung Probabilitas Kondisional (*Likelihood*)

Probabilitas setiap fitur (contoh: IPK, kehadiran) muncul pada masing-masing kelas (Lulus/Tidak). Jika fitur berupa data kontinu (seperti IPK), digunakan distribusi Gaussian (Normal):

$$P(x|C) = \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Keterangan:

- a. x = nilai fitur
- b. μ = rata-rata fitur pada kelas C
- c. σ = standar deviasi fitur pada kelas C

3) Menggunakan Teorema Bayes

Menghitung kemungkinan seorang mahasiswa termasuk kelas tertentu berdasarkan fitur-fiturnya:

$$P(C|X) \propto P(C) \cdot P(x_1|C) \cdot P(x_2|C) \cdot \dots \cdot P(x_n|C)$$

Pilih kelas C dengan nilai probabilitas terbesar sebagai hasil prediksi.

4) Prediksi

Setelah semua probabilitas dihitung, bandingkan:

$$P(Lulus|X) \text{ atau } P(Tidak Lulus|X)$$

Kelas dengan nilai lebih besar adalah hasil prediksi.