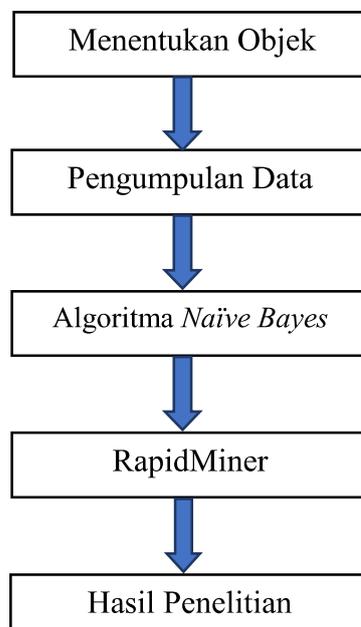


BAB III

METODE PENELITIAN

3.1 Desain Penelitian

Desain penelitian ini suatu sistematis yang dirancang untuk menentukan arah penelitian. Tujuan utama pembuatan desain penelitian adalah untuk mempermudah peneliti dalam menyusun penelitiannya. Tahapan-tahapan atau langkah-langkah dalam penelitian ini dilakukan dengan tujuan untuk menghasilkan hasil yang dapat diidentifikasi secara konseptual oleh semua pihak yang terlibat, serta untuk mengembangkan model penelitian yang sesuai dan menghasilkan data yang akurat. Berdasarkan deskripsi desain penelitian yang telah disebutkan, pembahasan mengenai setiap tahap dalam penelitian dapat dijelaskan sebagai berikut.



Gambar 3. 1 Desain Penelitian
Sumber: (Data Penelitian, 2024)

Keterangan:

1. Menentukan Objek

Langkah awal yang penting dalam penelitian adalah menentukan objek penelitian dengan jelas. Dalam hal ini, peneliti berfokus pada penghuni *Dormitory* Batamindo Batamindo, Muka Kuning, untuk memastikan bahwa sumber yang relevan dapat diperoleh selama penelitian berlangsung.

2. Pengumpulan Data

Data diambil dari *Dormitory* Batamindo Batamindo melalui observasi, kuesioner, dan studi pustaka yang mencakup bidang Data Mining dan Algoritma *Naïve Bayes*. Referensi yang berasal dari jurnal-jurnal nasional maupun internasional.

3. Algoritma *Naïve Bayes*

Dalam pengumpulan data peneliti memanfaatkan Algoritma *Naïve Bayes* dalam Data Mining untuk melakukan pengolahan data guna prediksi dampak penggunaan aplikasi *GetContact* terhadap keamanan pengguna.

4. RapidMiner

Dalam penelitian ini, peneliti menggunakan perangkat lunak RapidMiner sebagai aplikasi pendukung. Penggunaan *Software* ini memfasilitasi proses penelitian untuk mencapai tingkat akurasi data yang tinggi.

5. Hasil penelitian

Peneliti akan menyajikan hasil dan kesimpulan dari analisis yang telah dilakukan menggunakan teknik Data Mining dengan metode *Naïve Bayes*, serta memanfaatkan aplikasi RapidMiner untuk memprediksi dampak

penggunaan aplikasi *GetContact* terhadap keamanan pengguna di *Dormitory Batamindo Batamindo*.

3.2. Teknik Pengumpulan Data

Teknik pengumpulan data, atau *data collection*, adalah proses mengumpulkan informasi dalam penelitian dengan menggunakan metode ilmiah untuk mendapatkan data secara terstruktur, yang kemudian akan dianalisis sesuai dengan konteks masalah penelitian yang sedang diteliti. Teknik pengumpulan data yang digunakan dalam penelitian ini adalah sebagai berikut.:

1. Studi Pustaka

Peneliti menggunakan metode studi pustaka untuk mengumpulkan berbagai informasi yang berkaitan dengan topik penelitian, seperti KDD, data mining, teknik pengumpulan data, algoritma *Naïve Bayes*, dan aplikasi *GetContact*. Pendekatan ini dilakukan untuk mempelajari dan memperoleh wawasan dari penelitian sebelumnya, yang termasuk jurnal internasional, jurnal nasional, skripsi, serta situs web yang membahas topik serupa.

2. Observasi

Pada teknik pengumpulan data ini peneliti melakukan kunjungan langsung ke *Dormitory Batamindo Batamindo blok Q17* untuk memperoleh data yang dibutuhkan.

3. Wawancara

Wawancara dalam penelitian ini dilakukan secara langsung pada penghuni *Dormitory Batamindo* dengan tujuan memperoleh informasi mengenai indikator-indikator yang telah ditetapkan untuk mendapatkan data pengaruh

penggunaan aplikasi *GetContact* terhadap keamanan pengguna dalam menerima panggilan dan pesan.

4. Kuesioner

Untuk data Kuesioner pada penelitian ini dengan membagikan kuesioner berbasis link, kuesioner berbentuk google form.

3.3 Operasional Variabel

Dalam analisis data, penting untuk menggunakan apa yang disebut sebagai variabel operasional. Variabel operasional yang digunakan dalam studi ini adalah pengaruh penggunaan aplikasi *GetContact* terhadap keamanan pengguna. Peneliti mengumpulkan tanggapan dari penghuni *Dormitory* Batamindo terkait pernyataan mengenai dampak penggunaan aplikasi *GetContact* terhadap keamanan pengguna.

Tabel 3. 1 Data Kuisioner Setelah Penyeleksian

Responden	Frekuensi Penggunaan Aplikasi	Frekuensi Penggunaan Fitur	Pemahaman Fitur	Tingkat Keamanan	Pengaruh Aplikasi
0	3	3	4	4	Ya
1	3	3	4	4	Ya
2	1	1	3	3	Tidak
3	3	2	4	4	Ya
4	3	4	5	5	Ya
5	2	3	3	4	Ya
6	2	2	4	5	Ya
7	3	3	4	3	Tidak
8	2	2	3	3	Tidak
9	4	3	2	4	Ya
10	4	5	4	3	Tidak
11	3	2	4	3	Tidak
12	3	3	4	5	Ya
13	2	2	3	3	Tidak
14	2	3	4	4	Ya
15	2	2	4	4	Ya
16	1	1	4	3	Tidak
17	4	4	4	5	Ya

18	4	4	5	4	Ya
19	3	3	4	3	Tidak
20	5	5	5	5	Ya
21	3	3	4	3	Tidak
22	3	3	4	3	Tidak
23	1	1	4	3	Tidak
24	3	3	4	4	Ya
25	1	1	4	4	Ya
26	5	5	4	5	Ya
27	3	3	4	3	Tidak
28	1	1	3	3	Tidak
29	2	3	3	4	Ya
30	3	2	3	3	Tidak
31	3	3	4	5	Ya
32	5	5	4	5	Ya
33	4	4	4	4	Ya
34	3	1	4	5	Ya
35	3	3	4	4	Ya
36	3	2	4	3	Tidak
37	4	4	4	4	Ya
38	3	2	3	3	Tidak
39	2	1	2	4	Ya
40	3	3	3	4	Ya
41	5	5	5	3	Tidak
42	2	2	5	5	Ya
43	3	4	3	4	Ya
44	3	2	4	4	Ya
45	3	3	3	3	Tidak
46	3	1	2	5	Ya
47	3	3	3	3	Tidak
48	3	3	4	3	Tidak
49	3	3	3	4	Ya
50	3	3	4	4	Ya
51	3	2	3	3	Tidak
52	2	2	4	4	Ya
53	4	4	4	3	Tidak
54	3	3	4	5	Ya
55	2	2	4	3	Tidak
56	4	4	4	3	Tidak
57	3	3	3	4	Ya
58	3	3	4	3	Tidak

59	3	3	4	3	Tidak
60	3	1	3	4	Ya
61	3	1	3	3	Tidak
62	1	1	1	4	Ya
63	1	1	4	3	Tidak
64	2	1	3	3	Tidak
65	3	4	5	3	Tidak
66	1	1	4	3	Tidak
67	3	3	4	5	Ya
68	3	2	3	5	Ya
69	3	4	4	4	Ya
70	5	5	5	5	Ya
71	3	1	3	4	Ya
72	3	3	4	4	Ya
73	3	3	4	3	Tidak
74	4	3	3	4	Ya
75	4	4	4	4	Ya

Sumber: (Data Penelitian, 2024)

3.4 Algoritma Data Mining

Algoritma dapat digunakan dalam penyusunan prosedur yang jelas untuk menyelesaikan persoalan dengan urutan langkah-langkah tertentu dan terbatas, mengubah input data menjadi output yang mengandung informasi. Algoritma ini mengandalkan teorema *Bayes* untuk menghitung probabilitas kelas yang mungkin dari sebuah data berdasarkan nilai atribut-atributnya. Meskipun sederhana, *Naïve Bayes* sering kali memberikan hasil yang baik dalam berbagai kasus aplikasi, terutama ketika jumlah atribut relatif besar dibandingkan dengan jumlah data yang terbatas. Dalam algoritma ini, proses klasifikasi terdiri dari dua fase, yaitu *fase training* dan *fase testing*. *Fase training* atau sering disebut *fase pembelajaran*, melibatkan penggunaan sebagian data yang sudah diketahui kelasnya untuk membangun model prediksi. Selanjutnya pada *fase testing*, atau *fase*

pengklasifikasian, model yang telah dibangun diuji dengan sebagian data lainnya untuk mengevaluasi tingkat akurasi model yang telah terbentuk. Berikut persamaan-persamaan yang dapat dilakukan dalam Algoritma *Naïve Bayes* :

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

Rumus 3. 1 Algoritma Naïve Bayes.

Keterangan:

$P(H|X)$: Probabilitas hipotesis H berdasarkan kondisi X

$P(X|H)$: Probabilitas X berdasarkan kondisi pada hipotesis

$P(H)$: Probabilitas hipotesis H (prior probabilitas)

X : Data dengan kelas yang belum diketahui.

H : Hipotesis data merupakan suatu kelas spesifik

Dalam kasus setiap fitur menghitung kemungkinan probabilitas tersebut dalam setiap kelas yang ada. Berikut rumus prediksi probabilitas:

$$P(C_i|X) \propto P(X|C_i).P(C_i)$$

Rumus 3. 2 Prediksi Probabilitas.

Misal akan dilakukan perhitungan *Naive Bayes* dari data pelatihan pada Tabel 45, maka langkah pertama yang dilakukan adalah menghitung prior dari kelas “Ya” dan kelas “Tidak”. Berdasarkan 61 data tersebut diperoleh bahwa kelas Ya terdiri dari 34 data, dan kelas Tidak terdiri dari 27 data. Sehingga proses perhitungan priornya adalah sebagai berikut.

$$P(Ya) = \frac{34}{61} = 0.557$$

$$P(Tidak) = \frac{27}{61} = 0.442$$

Ketika prior data latih telah diperoleh, selanjutnya akan dilakukan perhitungan probabilitas masing-masing dari setiap kriteria yang digunakan. Kriteria yang digunakan adalah skala linkert 5 yaitu 1, 2, 3, 4, dan 5. Proses perhitungannya adalah sebagai berikut:

Tabel 3. 2 Hasil dari Kriteria Frekuensi Penggunaan Aplikasi

Variabel	Jumlah Kejadian		Probabilitas	
	Ya	Tidak	Ya	Tidak
5	2	1	0,0588235	0,037037
4	7	3	0,2058824	0,1111111
3	17	14	0,5	0,5185185
2	7	4	0,2058824	0,1481481
1	1	5	0,0294118	0,1851852
Total	34	27	1	1

Sumber: (Data Penelitian, 2024)

Probabilitas pada kriteria jumlah kejadian yaitu pada kategori 5 dengan skala "Ya" memiliki probabilitas 0.0588, 4 dengan skala "Ya" memiliki probabilitas 0.2059, 3 pada skala "Ya" memiliki probabilitas 0.5, 2 pada skala "Ya" memiliki probabilitas 0.2059, 1 pada skala "Ya" memiliki probabilitas 0.0294. Pada kategori 5 dengan skala "Tidak" memiliki probabilitas 0.0370, 4 dengan skala "Tidak" memiliki probabilitas 0.1111, 3 pada skala "Tidak" memiliki probabilitas 0.5185, 2 pada skala "Tidak" memiliki probabilitas 0.1481, 1 pada skala "Tidak" memiliki probabilitas 0.1852.

Tabel 3. 3 Hasil dari Kriteria Frekuensi Penggunaan Fitur

Variabel	Jumlah Kejadian		Probabilitas	
	Ya	Tidak	Ya	Tidak
5	2	2	0,0588235	0,0740741
4	8	3	0,2352941	0,1111111
3	13	9	0,3823529	0,3333333
2	7	6	0,2058824	0,2222222
1	4	7	0,1176471	0,2592593
Total	34	27	1	1

Sumber: (Data Penelitian, 2024)

Berdasarkan tabel yang diberikan, berikut penjelasan probabilitas pada kriteria "Frekuensi Penggunaan Aplikasi": Pada kategori 5, jumlah kejadian "Ya" adalah 2 dan "Tidak" adalah 2, dengan probabilitas "Ya" sebesar 0.0588 dan "Tidak" sebesar 0.0741. Pada kategori 4, jumlah kejadian "Ya" adalah 8 dan "Tidak" adalah 3, dengan probabilitas "Ya" sebesar 0.2353 dan "Tidak" sebesar 0.1111. Pada kategori 3, jumlah kejadian "Ya" adalah 13 dan "Tidak" adalah 9, dengan probabilitas "Ya" sebesar 0.3824 dan "Tidak" sebesar 0.3333. Pada kategori 2, jumlah kejadian "Ya" adalah 7 dan "Tidak" adalah 6, dengan probabilitas "Ya" sebesar 0.2059 dan "Tidak" sebesar 0.2222. Pada kategori 1, jumlah kejadian "Ya" adalah 4 dan "Tidak" adalah 7, dengan probabilitas "Ya" sebesar 0.1176 dan "Tidak" sebesar 0.2593. Penjelasan terperinci ini menunjukkan bahwa kategori 5 memiliki jumlah kejadian 2 untuk "Ya" dan 2 untuk "Tidak", dengan probabilitas masing-masing 0.0588 dan 0.0741. Kategori 4 memiliki jumlah kejadian 8 untuk "Ya" dan 3 untuk "Tidak", dengan probabilitas masing-masing 0.2353 dan 0.1111. Kategori 3 memiliki jumlah kejadian 13 untuk "Ya" dan 9 untuk "Tidak", dengan probabilitas masing-masing 0.3824 dan 0.3333. Kategori 2 memiliki jumlah kejadian 7 untuk "Ya" dan 6 untuk "Tidak", dengan probabilitas masing-masing 0.2059 dan 0.2222. Kategori 1 memiliki jumlah kejadian 4 untuk "Ya" dan 7 untuk "Tidak", dengan probabilitas masing-masing 0.1176 dan 0.2593.

Tabel 3. 4 Hasil dari Kriteria Pemahaman Fitur

Variabel	Jumlah Kejadian		Probabilitas	
	Ya	Tidak	Ya	Tidak
5	5	2	0,1470588	0,0740741

4	18	17	0,5294118	0,6296296
3	8	8	0,2352941	0,2962963
2	2	0	0,0588235	0
1	1	0	0,0294118	0
Total	34	27	1	1

Sumber: (Data Penelitian, 2024)

Berdasarkan tabel yang diberikan, berikut penjelasan probabilitas pada kriteria "Pemahaman Fitur": Pada kategori 5, jumlah kejadian "Ya" adalah 5 dan "Tidak" adalah 2, dengan probabilitas "Ya" sebesar 0.1471 dan "Tidak" sebesar 0.0741. Pada kategori 4, jumlah kejadian "Ya" adalah 18 dan "Tidak" adalah 17, dengan probabilitas "Ya" sebesar 0.5294 dan "Tidak" sebesar 0.6296. Pada kategori 3, jumlah kejadian "Ya" adalah 8 dan "Tidak" adalah 8, dengan probabilitas "Ya" sebesar 0.2353 dan "Tidak" sebesar 0.2963. Pada kategori 2, jumlah kejadian "Ya" adalah 2 dan "Tidak" adalah 0, dengan probabilitas "Ya" sebesar 0.0588 dan "Tidak" sebesar 0. Pada kategori 1, jumlah kejadian "Ya" adalah 1 dan "Tidak" adalah 0, dengan probabilitas "Ya" sebesar 0.0294 dan "Tidak" sebesar 0.

Tabel 3.5 Hasil dari Kriteria Tingkat Keamanan

Variabel	Jumlah Kejadian		Probabilitas	
	Ya	Tidak	Ya	Tidak
5	10	0	0,2941176	0
4	24	0	0,7058824	0
3	0	27	0	1
2	0	0	0	0
1	0	0	0	0
Total	34	27	1	1

Sumber : (Data Penelitian, 2024)

Berdasarkan tabel yang diberikan, berikut penjelasan probabilitas pada kriteria "Tingkat Keamanan": Pada kategori 5, jumlah kejadian "Ya" adalah 10

dan "Tidak" adalah 0, dengan probabilitas "Ya" sebesar 0.2941 dan "Tidak" sebesar 0. Pada kategori 4, jumlah kejadian "Ya" adalah 24 dan "Tidak" adalah 0, dengan probabilitas "Ya" sebesar 0.7059 dan "Tidak" sebesar 0. Pada kategori 3, jumlah kejadian "Ya" adalah 0 dan "Tidak" adalah 27, dengan probabilitas "Ya" sebesar 0 dan "Tidak" sebesar 1. Pada kategori 2 dan 1, jumlah kejadian "Ya" dan "Tidak" adalah 0, sehingga probabilitas untuk kedua kategori ini adalah 0.

Setelah masing-masing probabilitas kriteria telah didapatkan selanjutnya dilakukan perhitungan nilai data uji yang diberikan responden untuk menentukan klasifikasi. Berdasarkan data yang akan ditampilkan pada Tabel 4.6, data uji dengan jumlah 15 data akan diklasifikasikan ke dalam kelas “Ya” atau “Tidak”, untuk menjadi contoh akan digunakan data dari responden 40 dengan nilai yang diberikan pada kriteria Frekuensi Penggunaan Aplikasi = 3, Frekuensi Penggunaan Fitur = 3, Pemaahaman Fitur = 3, dan Tingkat Keamanan = 4.

$$\begin{aligned}
 P(40|Ya) &= P(\text{Frekuensi Penggunaan Aplikasi} = S|Ya) \times P(\text{Frekuensi} \\
 &\quad \text{Penggunaan Fitur} = S|Yes) \times P(\text{Pemahaman Fitur Keamanan} = \\
 &\quad S|Ya) \times P(\text{Tingkat Keamanan} = S|Ya) \\
 &= 0.5 \times 0.5 \times 0.5 \times 0.205 \\
 &= 0.0257
 \end{aligned}$$

$$\begin{aligned}
 P(40|Tidak) &= P(\text{Frekuensi Penggunaan Aplikasi} = S|Tidak) \times P(\text{Frekuensi} \\
 &\quad \text{Penggunaan Fitur} = S|Tidak) \times P(\text{Pemahaman Fitur Keamanan} = \\
 &\quad S|Tidak) \times P(\text{Tingkat Keamanan} = S|Tidak) \\
 &= 0.518 \times 0.518 \times 0.518 \times 0.111 \\
 &= 0.015
 \end{aligned}$$

Setelah nilai Ya dan Tidak dari responden 40 diperoleh, selanjutnya dilakukan perhitungan nilai maksimum dari data responden 40 sebagai berikut

$$\begin{aligned} P(\text{Ya}|C) &= P(\text{Rn}|C) * P(\text{Ya}) \\ &= P(40C) * P(\text{Ya}) \\ &= 0.0257 \times 0.557 \\ &= 0.014 \end{aligned}$$

$$\begin{aligned} P(\text{Tidak}|C) &= P(\text{Rn}|C) * P(\text{Ya}) \\ &= P(40C) * P(\text{Ya}) \\ &= 0.015 \times 0.442 \\ &= 0.006 \end{aligned}$$

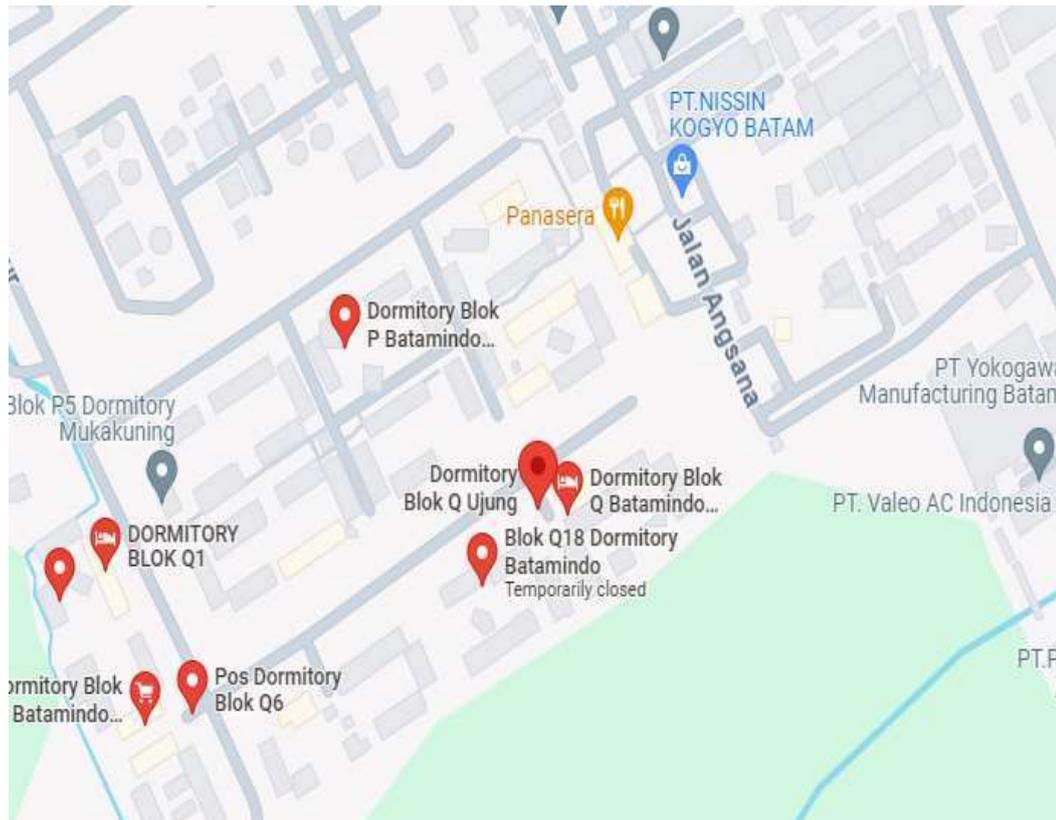
Langkah ini juga diterapkan hingga 30 data uji selesai dilakukan perhitungan nilai maksimum. Setelah nilai maksimal dari semua skala telah diperoleh, langkah selanjutnya membandingkan probabilitas maksimal yang muncul dari skala tersebut. Sehingga data uji responden 40 dapat dimasukkan ke dalam kelas Ya atau Tidak.

Pada data responden 40 diketahui perhitungan maksimal kelas Ya = 0.014, dan Tidak = 0.006. Sehingga, $0.014 > 0.006$, dapat dinyatakan bahwa data responden 40 merupakan pengguna yang merasa bahwa *GetContact* memberikan pengaruh terhadap keamanan pengguna dalam menerima panggilan dan pesan.

3.6 Lokasi dan Jadwal Penelitian

3.6.1 Lokasi Penelitian

Penelitian ini dilakukan di Kota Batam tepatnya di *Dormitory* blok Q17 yang berada di Kawasan Batamindo Industrial Park, Muka Kuning.



Gambar 3. 2 Lokasi Penelitian.
(Sumber: Data Penelitian, 2024).

3.6.2 Jadwal Penelitian

Penelitian ini dilakukan dari bulan Maret 2024 sampai bulan Juli 2024, berikut jadwal dari penelitian yang dilakukan.

Tabel 3. 6 Jadwal Penelitian.

No	Kegiatan	Maret				April				Mei				Juni				Juli					
		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4		
1	Pengajuan Judul	■	■	■																			
2	Penyusunan Bab I				■	■	■	■	■														
3	Penyusunan Bab II							■	■	■	■	■	■										
4	Penyusunan Bab III										■	■	■	■	■	■	■	■					
5	Penyusunan Bab IV										■	■	■	■	■	■	■	■	■	■			
6	Penyusunan Bab V																			■	■	■	
7	Pengumpulan Skripsi																					■	

Sumber: (Data Penelitian, 2024)