

## **BAB III**

### **METODE PENELITIAN**

#### **3.1. Jenis Penelitian**

Penelitian ini menggunakan desain eksperimen komparatif dengan pendekatan kuantitatif untuk membandingkan performa dua algoritma machine learning, Naïve Bayes dan Support Vector Machine (SVM), dalam mengklasifikasikan perceraian. Pendekatan ini memungkinkan peneliti untuk mengukur efektivitas kedua algoritma dalam memprediksi perceraian, memberikan wawasan objektif mengenai kelebihan dan kekurangan masing-masing algoritma.

##### **3.1.1. Penelitian Kuantitatif dan Desain Eksperimen Komparatif**

Penelitian ini mengadopsi pendekatan kuantitatif untuk mengumpulkan dan menganalisis data numerik, dengan fokus pada data perceraian dari Pengadilan Agama Kabupaten Labuhanbatu. Pendekatan ini memungkinkan peneliti untuk mengukur pengaruh faktor-faktor seperti ekonomi, narkoba, judi KDRT, selingkuh, komunikasi dan penyebab perceraian terhadap keputusan perceraian, serta menguji hipotesis yang ada. Desain eksperimen komparatif digunakan untuk membandingkan efektivitas dua algoritma machine learning, *Naïve Bayes* dan *Support Vector Machine (SVM)*, dalam mengklasifikasikan perceraian. Kedua algoritma diuji pada data yang sama untuk menilai keakuratan dan kemampuan masing-masing dalam memprediksi perceraian, memungkinkan perbandingan yang objektif. Eksperimen ini juga mengevaluasi kinerja algoritma dalam berbagai

kondisi, seperti variabilitas data dan data yang tidak terstruktur, untuk memberikan wawasan mendalam tentang kelebihan dan kekurangan masing-masing algoritma dalam konteks perceraian.

### **3.1.2. Metode Analisis dalam Penelitian**

Metode analisis dalam penelitian ini bertujuan untuk mengevaluasi akurasi dan performa dua algoritma machine learning dalam mengklasifikasikan perceraian, menggunakan metrik evaluasi seperti *akurasi*, *presisi*, *recall*, dan *F1-score*.

#### **1. Akurasi**

Mengukur proporsi prediksi yang benar dibandingkan dengan total data. Meskipun penting, akurasi tidak cukup untuk data yang tidak seimbang (misalnya lebih banyak pasangan yang tidak bercerai).

#### **2. Presisi**

Mengukur ketepatan model dalam memprediksi pasangan yang bercerai. Penting untuk menghindari prediksi perceraian yang salah, mengingat dampak sosial dan emosionalnya.

#### **3. Recall**

Mengukur seberapa banyak pasangan yang bercerai dapat terdeteksi. Recall penting untuk menghindari kesalahan false negative, yang bisa menghambat intervensi dalam perceraian yang sebenarnya.

#### 4. F1-Score

Menggabungkan presisi dan recall dengan rata-rata harmonis, memberikan gambaran seimbang tentang kinerja model, khususnya ketika ada *trade-off* antara keduanya.

5. Mengukur kemampuan model dalam membedakan antara pasangan yang bercerai dan yang tidak, tanpa bergantung pada threshold tertentu. Ini sangat penting untuk data yang tidak seimbang.

Dengan menggunakan metrik ini, penelitian ini dapat mengevaluasi dan membandingkan kinerja kedua algoritma dalam mengklasifikasikan perceraian secara objektif.

### 3.2. Variabel Penelitian

Penelitian ini bertujuan untuk menganalisis faktor-faktor yang mempengaruhi perceraian di Kabupaten Labuhanbatu dengan menggunakan algoritma *Naïve Bayes* dan *Support Vector Machine (SVM)*. Variabel-variabel yang digunakan dalam penelitian ini terbagi menjadi variabel dependen dan variabel independen. Pemilihan variabel yang tepat dan pengoperasionalisasiannya sangat penting untuk memastikan model *machine learning* dapat berfungsi secara optimal dalam mengklasifikasikan status perceraian.

#### 3.2.1. Variabel Dependen

Variabel dependen yang akan diprediksi dalam penelitian ini adalah status perceraian, yang mencerminkan keputusan apakah pasangan tersebut bercerai atau tidak. Variabel ini dibagi menjadi dua kategori utama, yaitu:

1. Bercerai: Pasangan yang telah bercerai berdasarkan keputusan yang tercatat di Pengadilan Agama Kabupaten Labuhanbatu.
2. Tidak Bercerai: Pasangan yang tetap menikah meskipun menghadapi berbagai permasalahan dalam hubungan mereka.

Status perceraian ini adalah variabel target yang akan digunakan dalam model *machine learning* untuk memprediksi hasil perceraian berdasarkan data variabel independen yang ada.

### **3.2.2. Variabel Independen**

Variabel independen dalam penelitian ini adalah faktor-faktor yang diduga memengaruhi terjadinya perceraian. Variabel-variabel ini digunakan sebagai atribut masukan (input) dalam proses klasifikasi status perceraian menggunakan algoritma Naive Bayes dan Support Vector Machine (SVM). Adapun variabel independen yang digunakan dalam penelitian ini meliputi:

1. Ekonomi

Variabel ekonomi menggambarkan kondisi finansial dalam rumah tangga, seperti ketidakmampuan memenuhi kebutuhan hidup, pengangguran, atau konflik akibat masalah keuangan. Permasalahan ekonomi sering menjadi faktor dominan yang memicu pertengkaran dan ketidakharmonisan dalam pernikahan.

2. Narkoba

Variabel narkoba menunjukkan adanya keterlibatan salah satu pasangan dalam penyalahgunaan narkoba atau zat adiktif lainnya. Penggunaan narkoba dapat berdampak negatif terhadap stabilitas emosional, ekonomi,

dan tanggung jawab dalam keluarga, sehingga meningkatkan risiko perceraian.

### 3. Judi

Variabel judi merepresentasikan kebiasaan berjudi yang dilakukan oleh salah satu pasangan. Perilaku ini sering menimbulkan masalah finansial, hilangnya kepercayaan, serta konflik berkepanjangan dalam rumah tangga.

### 4. Kekerasan Dalam Rumah Tangga (KDRT)

Variabel KDRT mencakup tindakan kekerasan fisik, psikis, verbal, atau ekonomi yang terjadi dalam rumah tangga. KDRT merupakan salah satu faktor serius yang secara langsung mendorong terjadinya perceraian demi keselamatan dan kesejahteraan korban.

### 5. Selingkuh

Variabel selingkuh menunjukkan adanya hubungan di luar pernikahan yang dilakukan oleh salah satu pasangan. Perselingkuhan dapat merusak kepercayaan dan komitmen dalam pernikahan, sehingga menjadi penyebab utama perceraian.

### 6. Komunikasi

Variabel komunikasi menggambarkan kualitas komunikasi antara pasangan, termasuk kemampuan menyampaikan pendapat, menyelesaikan konflik, dan memahami satu sama lain. Komunikasi yang buruk dapat memperbesar konflik dan mempercepat keretakan hubungan pernikahan.

## 7. Mabuk

Variabel mabuk menunjukkan kebiasaan konsumsi alkohol berlebihan oleh salah satu pasangan. Perilaku ini dapat menurunkan kontrol diri, memicu konflik, serta mengganggu keharmonisan rumah tangga sehingga meningkatkan risiko perceraian.

### 3.3. Teknik Pengumpulan Data

Pengumpulan data dalam penelitian ini menggunakan data sekunder dari Pengadilan Agama Kabupaten Labuhanbatu, mencakup semua kasus perceraian dalam lima tahun terakhir. Data ini valid, dapat diandalkan, dan relevan dengan topik penelitian yang menganalisis faktor-faktor perceraian menggunakan algoritma machine learning seperti *Naïve Bayes* dan *Support Vector Machine (SVM)*.

#### 3.3.1. Jenis Data

Penelitian ini menggunakan data sekunder dari Pengadilan Agama Kabupaten Labuhanbatu, yang mencakup dua jenis data utama: Jenis data yang digunakan dalam penelitian ini mencakup dua kategori utama:

##### 1. Data Numerik

Meliputi variabel yang diukur dengan angka, seperti usia menikah, lama pernikahan, dan jumlah anak, yang dapat dianalisis menggunakan teknik statistik atau *machine learning*.

## 2. Data Kategorikal

Meliputi variabel yang terklasifikasi dalam kategori, seperti tingkat pendidikan, pekerjaan, dan penyebab perceraian, yang membantu memahami karakteristik pasangan yang bercerai dan tidak.

### 3.3.2. Sumber Data

Sumber data utama penelitian ini adalah Pengadilan Agama Kabupaten Labuhanbatu, yang mencatat semua kasus perceraian secara sistematis. Keunggulannya antara lain:

#### 1. Keabsahan dan Validitas Data

Data yang tercatat resmi dan sah menurut hukum, memastikan kredibilitas dan keakuratan penelitian.

#### 2. Kelengkapan Data

Mencakup berbagai faktor penyebab perceraian spesifik seperti ekonomi, narkoba, judi, KDRT, selingkuh, dan komunikasi, memungkinkan analisis yang mendalam melalui model machine learning.

#### 3. Aksesibilitas Data

Peneliti dapat mengakses database pengadilan setelah izin, menghemat waktu dalam pengumpulan data.

#### 4. Keterjangkauan dan Efisiensi

Menggunakan data sekunder menghemat waktu dan biaya, memungkinkan fokus pada analisis dan penerapan algoritma machine learning.

### 3.3.3. Proses Pengumpulan Data

Proses pengumpulan data dilakukan melalui beberapa tahap:

#### 1. Permohonan Akses Data

Peneliti mengajukan permohonan akses ke Pengadilan Agama Kabupaten Labuhanbatu, mencakup tujuan penelitian dan cara pengelolaan data untuk menjaga kerahasiaan.

#### 2. Verifikasi Data

Data yang diperoleh diverifikasi untuk memastikan keakuratan dan konsistensinya.

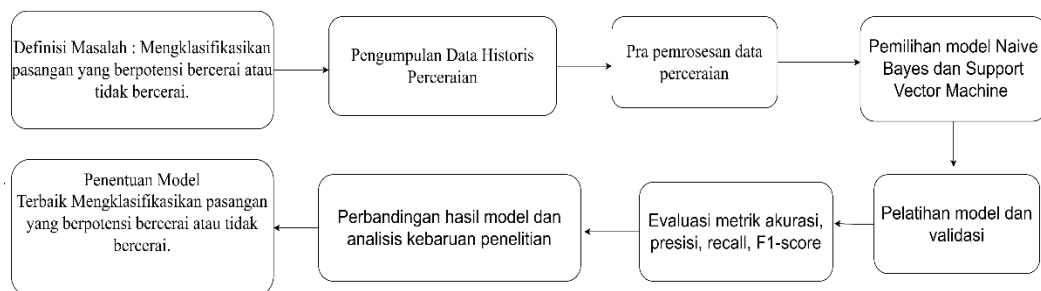
#### 3. Penyusunan dan Pengolahan Data

Data yang diverifikasi dibersihkan dan diproses menggunakan teknik preprocessing seperti normalisasi, pengkodean, dan seleksi fitur.

#### 4. Analisis dan Evaluasi

Data yang telah diproses dianalisis menggunakan algoritma *Naïve Bayes* dan *Support Vector Machine (SVM)*, dengan evaluasi metrik *akurasi*, *presisi*, dan *recall* untuk menilai performa kedua algoritma.

### 3.4. Desain Penelitian



**Gambar 3. 1** Kerangka Penelitian

### **3.4.1. Pendekatan Penelitian**

Penelitian ini menggunakan pendekatan kuantitatif dengan desain eksperimen komparatif untuk membandingkan dua algoritma, *Naïve Bayes* dan *Support Vector Machine (SVM)*, dalam memprediksi perceraian berdasarkan data Pengadilan Agama Kabupaten Labuhanbatu. Metode eksperimen ini mengontrol variabel perceraian dan menguji efektivitas kedua algoritma dalam klasifikasi. Dengan menggunakan data yang sama untuk kedua model, penelitian ini mengevaluasi akurasi, presisi, recall, dan metrik lainnya untuk menentukan algoritma yang lebih efektif dalam analisis perceraian.

### **3.4.2. Langkah-langkah Penelitian**

Langkah-langkah dalam penelitian ini mengikuti urutan yang terstruktur untuk memastikan konsistensi dan integritas dalam pengumpulan dan analisis data. Berikut adalah tahapan-tahapan yang akan dilalui dalam penelitian ini:

1. Pengumpulan Data

Pengumpulan data dilakukan pada perkara perceraian di Pengadilan Agama Kabupaten Labuhanbatu selama lima tahun terakhir, dengan mencatat faktor ekonomi, narkoba, judi, KDRT, selingkuh, dan komunikasi sebagai atribut utama penyebab perceraian.

2. Pra-pemrosesan Data

Data akan diproses untuk mengatasi nilai hilang, normalisasi data numerik, dan pengkodean data kategorikal. Teknik imputasi digunakan untuk nilai hilang, normalisasi dilakukan untuk skala data, dan one-hot encoding diterapkan pada variabel kategorikal.

### 3. Pemilihan Model dan Pelatihan Model

Dua algoritma machine learning, *Naïve Bayes* dan *Support Vector Machine (SVM)*, dipilih untuk mengklasifikasikan perceraian. *Naïve Bayes* efisien untuk data dengan banyak fitur, sedangkan *Support Vector Machine (SVM)* cocok untuk pola data yang kompleks dan *non-linear*. Pelatihan model dilakukan untuk mencapai performa terbaik.

### 4. Evaluasi Model

Model dievaluasi menggunakan data pengujian dan metrik seperti *akurasi*, *presisi*, *recall*, dan *F1-score* untuk menilai kinerjanya dalam mengklasifikasikan perceraian.

### 5. Perbandingan Hasil Model

Hasil dari *Naïve Bayes* dan *Support Vector Machine (SVM)* dibandingkan berdasarkan akurasi, kecepatan, efisiensi, dan interpretasi, untuk menentukan algoritma terbaik dalam analisis perceraian di Kabupaten Labuhanbatu.

## 3.5. Instrumen Penelitian

Instrumen penelitian memainkan peran kunci dalam proses penelitian karena alat-alat yang digunakan akan langsung memengaruhi kualitas data dan hasil yang diperoleh. Pada penelitian ini, instrumen utama yang digunakan adalah data yang dikumpulkan dari Pengadilan Agama Kabupaten Labuhanbatu, yang mencakup berbagai faktor penyebab perceraian. Selain itu, perangkat pengolahan data yang digunakan dalam penelitian ini adalah perangkat lunak yang mendukung

penerapan algoritma *machine learning* untuk analisis data perceraian, dengan tools seperti *Orange* yang digunakan untuk implementasi algoritma *Naïve Bayes* dan *Support Vector Machine (SVM)*.

### 3.5.1. Keunggulan dan Kelemahan Penggunaan Orange

Keunggulan Orange sebagai perangkat pengolahan data untuk penelitian ini meliputi beberapa aspek penting yang mendukung efektivitas penelitian:

1. Antarmuka Visual yang Mudah Digunakan

Orange menawarkan antarmuka *drag-and-drop* yang memungkinkan peneliti merancang alur kerja tanpa menulis kode, ideal untuk peneliti tanpa latar belakang pemrograman..

2. Integrasi Berbagai Algoritma

Mendukung berbagai algoritma machine learning seperti *Naïve Bayes*, *Support Vector Machine (SVM)*, dan *Random Forest*, memberi fleksibilitas untuk menguji dan membandingkan model dalam satu platform.

3. Open-Source dan Gratis

Sebagai *platform open-source*, *Orange* dapat digunakan secara gratis, menjadikannya pilihan tepat bagi peneliti dengan anggaran terbatas.

Kelemahan: Orange kurang fleksibel untuk algoritma lanjutan dibandingkan pustaka seperti *Scikit-learn* atau *TensorFlow*, meskipun masih efektif untuk algoritma dasar seperti *Naïve Bayes* dan *Support Vector Machine (SVM)*.

### 3.6. Perhitungan Rumus Algoritma dan Teknik Evaluasi

#### 3.6.1. Pembersihan Data

Pembersihan data adalah tahap penting dalam mempersiapkan dataset untuk analisis. Proses ini meliputi penghapusan duplikat, penanganan nilai hilang, normalisasi variabel numerik, dan *encoding* variabel kategorikal. Tujuannya adalah memastikan data bersih, konsisten, dan siap untuk pelatihan model, sehingga menghasilkan prediksi yang akurat dalam mengklasifikasikan pasangan yang berpotensi bercerai.

**Tabel 3. 1** Data Bersih Perceraian

<b>Ekonomi</b>	<b>Narkoba</b>	<b>Judi</b>	<b>KDRT</b>	<b>Selingkuh</b>	<b>Komunikasi</b>	<b>Mabuk</b>
Cukup baik	Tidak	Tidak	Ya	Tidak	Buruk	Tidak
Rendah	Ya	Ya	Ya	Ya	Buruk	Ya
Cukup baik	Tidak	Tidak	Tidak	Tidak	Baik	Ya
Cukup baik	Tidak	Tidak	Tidak	Tidak	Baik	Tidak
Cukup baik	Tidak	Tidak	Tidak	Tidak	Baik	Tidak
Rendah	Ya	Tidak	Ya	Tidak	Buruk	Tidak
Cukup baik	Tidak	Tidak	Tidak	Tidak	Baik	Tidak
Cukup baik	Tidak	Tidak	Ya	Tidak	Baik	Tidak
Cukup baik	Tidak	Tidak	Tidak	Tidak	Baik	Tidak
Cukup baik	Tidak	Tidak	Tidak	Tidak	Baik	Tidak
Cukup baik	Ya	Tidak	Ya	Tidak	Buruk	Tidak
Cukup baik	Tidak	Tidak	Tidak	Tidak	Baik	Tidak
Rendah	Ya	Tidak	Ya	Ya	Buruk	Tidak
Cukup baik	Ya	Ya	Ya	Tidak	Buruk	Ya

Cukup baik	Tidak	Tidak	Tidak	Tidak	Baik	Tidak
Cukup baik	Ya	Tidak	Ya	Ya	Buruk	Ya
Cukup baik	Ya	Ya	Tidak	Tidak	Buruk	Tidak

Tabel tersebut menggambarkan variasi kondisi faktor-faktor penyebab perceraian, yaitu ekonomi, narkoba, judi, KDRT, selingkuh, komunikasi dan mabuk, melalui kategori seperti cukup baik, rendah, ya, dan tidak. Kombinasi nilai pada setiap baris menunjukkan tingkat kerentanan rumah tangga terhadap perceraian yang selaras dengan TemanHukum.id. (2024). *Alasan Perceraian di Indonesia: Sebuah Analisis Mendalam*. (Mengulas faktor ekonomi dan KDRT sebagai alasan penting perceraian).

### 3.6.2. Pembagian Data

Pembagian data dibagi menjadi 2 jenis data yaitu data latih (training data) dan data uji (testing data), dengan pembagian 80:20 (48 data untuk latih dan 12 data untuk uji).

#### a. Data Training

**Tabel 3. 2** Data Training Perceraian

Ekonomi	Narkoba	Judi	KDRT	Selingkuh	Komunikasi	Mabuk	Status
Cukup baik	Tidak	Tidak	Ya	Tidak	Buruk	Tidak	Cerai
Rendah	Ya	Ya	Ya	Ya	Buruk	Ya	Cerai
Cukup baik	Tidak	Tidak	Tidak	Tidak	Baik	Ya	Masih menikah
Cukup baik	Tidak	Tidak	Tidak	Tidak	Baik	Tidak	Masih menikah
Cukup baik	Tidak	Tidak	Tidak	Tidak	Baik	Tidak	Masih menikah
Rendah	Ya	Tidak	Ya	Tidak	Buruk	Tidak	Cerai

Cukup baik	Tidak	Tidak	Tidak	Tidak	Baik	Tidak	Masih menikah
Cukup baik	Tidak	Tidak	Ya	Tidak	Baik	Tidak	Cerai
Cukup baik	Tidak	Tidak	Tidak	Tidak	Baik	Tidak	Masih menikah
Cukup baik	Tidak	Tidak	Tidak	Tidak	Baik	Tidak	Masih menikah
Cukup baik	Ya	Tidak	Ya	Tidak	Buruk	Tidak	Cerai
Cukup baik	Tidak	Tidak	Tidak	Tidak	Baik	Tidak	Masih menikah
Rendah	Ya	Tidak	Ya	Ya	Buruk	Tidak	Cerai
Cukup baik	Ya	Ya	Ya	Tidak	Buruk	Ya	Cerai
Cukup baik	Tidak	Tidak	Tidak	Tidak	Baik	Tidak	Masih menikah
Cukup baik	Ya	Tidak	Ya	Ya	Buruk	Ya	Cerai
Cukup baik	Ya	Ya	Tidak	Tidak	Buruk	Tidak	Cerai
Cukup baik	Tidak	Tidak	Tidak	Tidak	Baik	Tidak	Masih menikah
Cukup baik	Tidak	Tidak	Tidak	Tidak	Baik	Tidak	Masih menikah
Cukup baik	Ya	Tidak	Ya	Ya	Baik	Tidak	Cerai
Cukup baik	Ya	Ya	Ya	Tidak	Baik	Ya	Cerai
Cukup baik	Tidak	Tidak	Tidak	Ya	Buruk	Tidak	Cerai
Rendah	Ya	Ya	Ya	Ya	Buruk	Tidak	Cerai
Rendah	Tidak	Tidak	Ya	Tidak	Buruk	Tidak	Cerai
Rendah	Ya	Tidak	Tidak	Tidak	Buruk	Tidak	Cerai
Cukup baik	Tidak	Tidak	Tidak	Tidak	Baik	Tidak	Masih menikah
Cukup baik	Tidak	Tidak	Ya	Ya	Buruk	Tidak	Cerai
Rendah	Ya	Ya	Tidak	Tidak	Buruk	Tidak	Cerai
Cukup baik	Ya	Tidak	Ya	Ya	Buruk	Tidak	Cerai
Rendah	Tidak	Ya	Ya	Tidak	Baik	Tidak	Cerai
Rendah	Ya	Ya	Tidak	Ya	Buruk	Tidak	Cerai
Cukup baik	Tidak	Tidak	Ya	Tidak	Baik	Tidak	Cerai
Cukup baik	Tidak	Tidak	Tidak	Tidak	Baik	Tidak	Masih menikah
Rendah	Ya	Tidak	Ya	Ya	Buruk	Tidak	Cerai
Cukup baik	Tidak	Tidak	Tidak	Tidak	Baik	Tidak	Masih menikah
Rendah	Ya	Ya	Ya	Ya	Baik	Ya	Cerai

Cukup baik	Ya	Tidak	Ya	Tidak	Buruk	Tidak	Cerai
Cukup baik	Ya	Ya	Tidak	Ya	Buruk	Tidak	Cerai
Rendah	Ya	Tidak	Ya	Tidak	Buruk	Tidak	Cerai
Cukup baik	Tidak	Tidak	Ya	Tidak	Buruk	Tidak	Cerai
Rendah	Ya	Ya	Ya	Ya	Baik	Tidak	Cerai
Cukup baik	Tidak	Tidak	Tidak	Tidak	Baik	Tidak	Masih menikah
Cukup baik	Tidak	Tidak	Ya	Tidak	Buruk	Tidak	Cerai
Rendah	Ya	Tidak	Ya	Ya	Buruk	Tidak	Cerai
Cukup baik	Ya	Ya	Ya	Tidak	Baik	Tidak	Cerai
Rendah	Ya	Ya	Ya	Tidak	Baik	Tidak	Cerai
Rendah	Tidak	Tidak	Tidak	Ya	Buruk	Tidak	Cerai
Cukup baik	Ya	Ya	Tidak	Tidak	Buruk	Tidak	Cerai

Data training yang digunakan mencakup 80% dari keseluruhan dataset dan terdiri dari variabel (1) Ekonomi, (2) Narkoba, (3) Judi, (4) KDRT, (5) Selingkuh, (6) Mabuk dan (7) Komunikasi. Variabel-variabel ini merepresentasikan kondisi rumah tangga terkait aspek finansial, perilaku berisiko, kekerasan dalam rumah tangga, kesetiaan, serta kualitas interaksi pasangan yang berpotensi memicu perceraian. Data training digunakan untuk melatih model machine learning agar mampu mengenali pola pada kombinasi faktor-faktor tersebut dan mengklasifikasikan atau memprediksi status perceraian secara lebih akurat. Pembagian data menjadi 80% data training dan 20% data testing memungkinkan evaluasi kemampuan model dalam menggeneralisasi pola ke data baru yang belum pernah dilihat sebelumnya.

## b. Data Testing

**Tabel 3. 3** Data Testing Perceraian

Ekonomi	Narkoba	Judi	KDRT	Selingkuh	Komunikasi	Mabuk	Status
Rendah	Ya	Tidak	Ya	Ya	Buruk	Tidak	Cerai
Cukup baik	Tidak	Tidak	Tidak	Tidak	Baik	Ya	Masih menikah
Rendah	Ya	Ya	Tidak	Tidak	Buruk	Ya	Cerai
Cukup baik	Ya	Ya	Ya	Tidak	Buruk	Tidak	Cerai
Rendah	Tidak	Tidak	Ya	Ya	Buruk	Tidak	Cerai
Rendah	Ya	Ya	Ya	Ya	Buruk	Tidak	Cerai
Cukup baik	Tidak	Tidak	Tidak	Tidak	Baik	Tidak	Masih menikah
Cukup baik	Tidak	Tidak	Tidak	Tidak	Baik	Tidak	Masih menikah
Cukup baik	Tidak	Tidak	Tidak	Tidak	Baik	Tidak	Masih menikah
Cukup baik	Ya	Ya	Tidak	Tidak	Buruk	Tidak	Cerai
Rendah	Ya	Ya	Ya	Tidak	Baik	Tidak	Cerai
Cukup baik	Ya	Ya	Tidak	Tidak	Buruk	Tidak	Cerai

Data testing, yang mencakup 20% dari keseluruhan dataset, digunakan untuk menguji model machine learning yang telah dilatih. Data ini berisi variabel yang sama dengan data training, yaitu (1) Ekonomi, (2) Narkoba, (3) Judi, (4) KDRT, (5) Selingkuh, (6) Mabuk dan (7) Komunikasi. Tujuan data testing adalah untuk mengevaluasi seberapa baik model dapat mengklasifikasikan kasus perceraian yang belum pernah dilihat sebelumnya, sehingga akurasi dan kinerja model dapat diukur secara objektif.

### 3.6.3. Perhitungan Algoritma Naïve Bayes

#### 1. Seleksi Data

Seleksi data adalah tahap pemilihan data yang akan digunakan dalam penelitian ini. Data yang dikumpulkan akan diproses terlebih dahulu sebelum digunakan dalam analisis.

#### 2. Preprocessing Data

Pada tahap preprocessing, data yang tidak lengkap atau tidak valid dibuang, sementara data relevan diproses lebih lanjut. Variabel numerik dikelompokkan, dan variabel kategorikal diubah menjadi format numerik untuk digunakan dalam algoritma Naïve Bayes. Setelah itu, data pelatihan dan pengujian dipisahkan untuk evaluasi model.

**Tabel 3. 4** Menghitung Jumlah Data

<b>MENGHITUNG JUMLAH DATA</b>	
Cerai	42
Masih Menikah	18
Total	60

#### a. Menghitung Probabilitas Awal

Rumus :

$$P(\text{Bercerai}) = \frac{\text{Jumlah Bercerai}}{\text{Total Data}}$$

$$P(\text{Masih Menikah}) = \frac{\text{Jumlah Masih Menikah}}{\text{Total Data}}$$

**Tabel 3. 5** Probabilitas Klasifikasi

<b>PROBABILITAS KLASIFIKASI</b>	
Cerai	0,690
Masih Menikah	0,310
Total	1

**Tabel 3. 6** Probabilitas Ekonomi

<b>Ekonomi</b>	<b>Cerai</b>	<b>Masih menikah</b>
Rendah	0,5	0
Cukup baik	0,5	1
Total	1	1

**Tabel 3. 7** Probabilitas Narkoba

<b>Narkoba</b>	<b>Cerai</b>	<b>Masih menikah</b>
Ya	0	0
Tidak	0,262	1
Total	0,262	1

**Tabel 3. 8** Probabilitas Judi

<b>Judi</b>	<b>Cerai</b>	<b>Masih menikah</b>
Ya	0,167	0
Tidak	0,524	1
Total	0,690	1

**Tabel 3. 9** Probabilitas KDRT

<b>KDRT</b>	<b>Cerai</b>	<b>Masih menikah</b>
Ya	0,262	0
Tidak	0,262	1
Total	0,524	1

**Tabel 3. 10** Probabilitas Selingkuh

Selingkuh	Cerai	Masih menikah
Ya	0,310	0
Tidak	0,571	1
Total	0,881	1

**Tabel 3. 11** Probabilitas Komunikasi

Komunikasi	Cerai	Masih menikah
Baik	0,238	1
Buruk	0,762	0
Total	1	1

**Tabel 3. 12** Probabilitas Mabuk

Mabuk	Cerai	Masih menikah
Ya	0,119	0,111
Tidak	0,881	0,889
Total	1	1

**Tabel 3. 13** Menghitung Prediksi

Kelas Prediksi	Cerai	Masih menikah
Cerai	0,003	0
Masih menikah	0,005	0,276
Cerai	0,002	0
Masih menikah	0,002	0,034

**Tabel 3. 14** Confusion Matrix

Confusion Matrix	
TP	2
TN	2
FP	0
FN	0

b. Menghitung Evaluasi Model

Rumus :

1. Accuracy :

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)}$$

Dengan nilai :

$$Accuracy = \frac{(2 + 2)}{(2 + 2 + 0 + 0)} = \frac{4}{4} = 1$$

2. Precision :

Rumus :

$$Precision = \frac{TP}{(TP + FP)}$$

Dengan nilai :

$$Precision = \frac{2}{(2 + 0)} = \frac{2}{2} = 1$$

3. Recall :

Rumus :

$$Recall = \frac{TP}{(TP + FN)}$$

Dengan nilai :

$$Recall = \frac{2}{(2 + 0)} = \frac{2}{2} = 1$$

4. F1-Score :

Rumus :

$$F1\text{-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Dengan nilai :

$$F1\text{-Score} = 2 \times \frac{1 \times 1}{1 + 1} = 2 \times \frac{1}{2} = 1$$

**Tabel 3. 15** Hasil Evaluasi Model

Hasil Evaluasi Model	
Akurasi	1
Precision	1
Recall	1
F1-Score	1

Tabel yang baru menunjukkan hasil evaluasi model dengan skor yang seragam untuk semua metrik: akurasi, presisi, recall, dan F1-score, masing-masing mendapat skor 1. Ini mengindikasikan bahwa model memiliki performa yang konsisten pada semua metrik, namun skor 1 mungkin menunjukkan kinerja yang relatif rendah atau perlu ditingkatkan.

#### 3.6.4. Perhitungan Algoritma SVM

1. Parameter yang di gunakan :
  - a.  $w_1 = 0.5$  Untuk Faktor Ekonomi
  - b.  $w_2 = 0.5$  Untuk Faktor Narkoba
  - c.  $w_3 = 0.5$  Untuk Faktor Judi
  - d.  $w_4 = 0.5$  Untuk Faktor KDRT
  - e.  $w_5 = 0.5$  Untuk Faktor Selingkuh
  - f.  $w_6 = 0.5$  Untuk Faktor Komunikasi
  - g.  $b = (1)$  Bias

2. Rumus Fungsi Keputusan:

Fungsi keputusan  $f(x)$  akan dihitung dengan rumus:

$$a. f(x) = w_1 \cdot \text{Ekonomi} + w_2 \cdot \text{Narkoba} + w_3 \cdot \text{Judi} + w_4 \cdot \text{KDRT} + w_5 \cdot \text{Selingkuh} + w_6 \cdot \text{Komunikasi} + b$$

3. Langkah 1: Menghitung Fungsi Keputusan

Contoh Perhitungan:

- a. Ekonomi = 1, Narkoba = -1, Judi = -1, KDRT = 1, Selingkuh = -1, Komunikasi = -1, Bias = (1)

$$f(x) = 1 \cdot 0.5 + (-1) \cdot 0.5 + (-1) \cdot 0.5 + 1 \cdot 0.5 + (-1) \cdot 0.5 + (-1) \cdot 0.5 + 1$$

$$f(x) = 0.5 - 0.5 - 0.5 + 0.5 - 0.5 - 0.5 + 1 = 0$$

Prediksi: Karena  $f(x) = 0$ , maka prediksi = Masih menikah

- b. Ekonomi = -1, Narkoba = 1, Judi = 1, KDRT = 1, Selingkuh = 1, Komunikasi = -1, Bias = (1)

$$f(x) = (-1) \cdot 0.5 + 1 \cdot 0.5 + 1 \cdot 0.5 + 1 \cdot 0.5 + 1 \cdot 0.5 + (-1) \cdot 0.5 + 1$$

$$f(x) = (-0.5) + 0.5 + 0.5 + 0.5 + 0.5 - 0.5 = 2$$

Prediksi: Karena  $f(x) > 0$ , maka prediksi = Cerai

**Tabel 3. 16** Hasil Perhitungan SVM

Ekonomi	Narkoba	Judi	KDRT	Selingkuh	Komunikasi	Mabuk	Label	W1	W2	W3	W4	W5	W6	b	Akhir	Hasil
1	-1	-1	1	-1	-1	-1	1	0,5	0,5	0,5	0,5	0,5	0,5	1	-0,5	-1
-1	1	1	1	1	-1	1	1	0,5	0,5	0,5	0,5	0,5	0,5	1	2,5	1
1	-1	-1	-1	-1	1	1	-1	0,5	0,5	0,5	0,5	0,5	0,5	1	0,5	1
1	-1	-1	-1	-1	1	-1	-1	0,5	0,5	0,5	0,5	0,5	0,5	1	-0,5	-1

1	-1	-1	-1	-1	1	-1	-1	0,5	0,5	0,5	0,5	0,5	0,5	1	-0,5	-1
-1	1	-1	1	-1	-1	-1	1	0,5	0,5	0,5	0,5	0,5	0,5	1	-0,5	-1
1	-1	-1	-1	-1	1	-1	-1	0,5	0,5	0,5	0,5	0,5	0,5	1	-0,5	-1
1	-1	-1	1	-1	1	-1	1	0,5	0,5	0,5	0,5	0,5	0,5	1	0,5	1
1	-1	-1	-1	-1	1	-1	-1	0,5	0,5	0,5	0,5	0,5	0,5	1	-0,5	-1
1	-1	-1	-1	-1	1	-1	-1	0,5	0,5	0,5	0,5	0,5	0,5	1	-0,5	-1
1	1	-1	1	-1	-1	-1	1	0,5	0,5	0,5	0,5	0,5	0,5	1	0,5	1
1	-1	-1	-1	-1	1	-1	-1	0,5	0,5	0,5	0,5	0,5	0,5	1	-0,5	-1
-1	1	-1	1	1	-1	-1	1	0,5	0,5	0,5	0,5	0,5	0,5	1	0,5	1
1	1	1	1	-1	-1	1	1	0,5	0,5	0,5	0,5	0,5	0,5	1	2,5	1
1	-1	-1	-1	-1	1	-1	-1	0,5	0,5	0,5	0,5	0,5	0,5	1	-0,5	-1
1	1	-1	1	1	-1	1	1	0,5	0,5	0,5	0,5	0,5	0,5	1	2,5	1
1	1	1	-1	-1	-1	-1	1	0,5	0,5	0,5	0,5	0,5	0,5	1	0,5	1
1	-1	-1	-1	-1	1	-1	-1	0,5	0,5	0,5	0,5	0,5	0,5	1	-0,5	-1
1	-1	-1	-1	-1	1	-1	-1	0,5	0,5	0,5	0,5	0,5	0,5	1	-0,5	-1
1	1	-1	1	1	1	-1	1	0,5	0,5	0,5	0,5	0,5	0,5	1	2,5	1
1	1	1	1	-1	1	1	1	0,5	0,5	0,5	0,5	0,5	0,5	1	3,5	1
1	-1	-1	-1	1	-1	-1	1	0,5	0,5	0,5	0,5	0,5	0,5	1	-0,5	-1
-1	1	1	1	1	-1	-1	1	0,5	0,5	0,5	0,5	0,5	0,5	1	1,5	1
-1	-1	-1	1	-1	-1	-1	1	0,5	0,5	0,5	0,5	0,5	0,5	1	-1,5	-1
-1	1	-1	-1	-1	-1	-1	1	0,5	0,5	0,5	0,5	0,5	0,5	1	-1,5	-1
1	-1	-1	-1	-1	1	-1	-1	0,5	0,5	0,5	0,5	0,5	0,5	1	-0,5	-1
1	-1	-1	1	1	-1	-1	1	0,5	0,5	0,5	0,5	0,5	0,5	1	0,5	1
-1	1	1	-1	-1	-1	-1	1	0,5	0,5	0,5	0,5	0,5	0,5	1	-0,5	-1
1	1	-1	1	1	-1	-1	1	0,5	0,5	0,5	0,5	0,5	0,5	1	1,5	1
-1	-1	1	1	-1	1	-1	1	0,5	0,5	0,5	0,5	0,5	0,5	1	0,5	1
-1	1	1	-1	1	-1	-1	1	0,5	0,5	0,5	0,5	0,5	0,5	1	0,5	1
1	-1	-1	1	-1	1	-1	1	0,5	0,5	0,5	0,5	0,5	0,5	1	0,5	1
1	-1	-1	-1	-1	1	-1	-1	0,5	0,5	0,5	0,5	0,5	0,5	1	-0,5	-1
-1	1	-1	1	1	-1	-1	1	0,5	0,5	0,5	0,5	0,5	0,5	1	0,5	1
1	-1	-1	-1	-1	1	-1	-1	0,5	0,5	0,5	0,5	0,5	0,5	1	-0,5	-1
-1	1	1	1	1	1	1	1	0,5	0,5	0,5	0,5	0,5	0,5	1	3,5	1

Tabel menampilkan data biner faktor penyebab perceraian (Ekonomi, Narkoba, Judi, KDRT, Selingkuh, Komunikasi) beserta label kelas dan parameter model seperti bobot W1–W6, bias b, nilai akhir, dan hasil klasifikasi. Tabel ini

menggambarkan proses perhitungan model machine learning berbasis pembobotan untuk memprediksi status perceraian.

“Struktur tabel ini disusun mengikuti konsep model klasifikasi biner berbasis bobot dan bias pada machine learning (mis. Goodfellow dkk., 2016).”

Keterangan :

1. Ekonomi : Rendah (-1), Cukup Baik (1)
2. Narkoba : Ya (1), Tidak (-1)
3. Judi : Ya (1), Tidak (-1)
4. KDRT : Ya (1), Tidak (-1)
5. Selingkuh : Ya (1), Tidak (-1)
6. Komunikasi : Baik (1), Buruk (-1)
7. Mabuk : Ya (1), Tidak (-1)

Langkah 1: Menghitung Confusion Matrix SVM

**Tabel 3. 17** Hasil Confusion Matrix

<b>Confusion Matrix</b>	
TP	30
FP	0
FN	12
TN	16

Tabel 3.17 menyajikan hasil confusion matrix model klasifikasi perceraian, dengan nilai True Positive (TP) 30, False Positive (FP) 0, False Negative (FN) 12, dan True Negative (TN) 16. Nilai-nilai ini digunakan untuk menghitung metrik kinerja seperti akurasi, presisi, recall, dan f1-score model.

## 2: Menghitung Evaluasi Model SVM

a. Accuracy :

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN}$$

Dengan nilai :

$$\text{Accuracy} = \frac{30 + 18}{30 + 0 + 18 + 12} = \frac{48}{60} \approx 0.8 \text{ Atau } 80\%$$

b. Precision :

$$\text{Precision} = \frac{TP}{TP + FP}$$

Dengan nilai :

$$\text{Precision} = \frac{30}{30 + 0} = \frac{30}{30} \approx 1 \text{ Atau } 100\%$$

c. Recall :

$$\text{Recall} = \frac{TP}{TP + FN}$$

Dengan nilai :

$$\text{Recall} = \frac{30}{30 + 12} = \frac{30}{42} \approx 0.714 \text{ Atau } 71,4\%$$

d. F1-Score :

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Dengan nilai :

$$\text{F1-Score} = 2 \times \frac{1 \times 0.714}{1 + 0.714} \approx 0.833 \text{ Atau } 83,3\%$$

**Tabel 3. 18** Hasil Evaluasi Model

Evaluasi Model	
AKURASI	0,767
PRECISION	0,938
RECALL	0,714
F1-SCORE	0,811

Tabel evaluasi model menunjukkan bahwa kinerja klasifikasi yang dihasilkan cukup baik. Nilai akurasi sebesar 0,767 menunjukkan bahwa model mampu memprediksi data dengan tingkat ketepatan yang cukup tinggi. Nilai precision sebesar 0,938 menunjukkan bahwa sebagian besar prediksi positif yang dihasilkan model adalah benar. Sementara itu, recall sebesar 0,714 dan F1-score sebesar 0,811 menggambarkan keseimbangan yang cukup baik antara ketepatan prediksi dan kemampuan model dalam menemukan data positif.