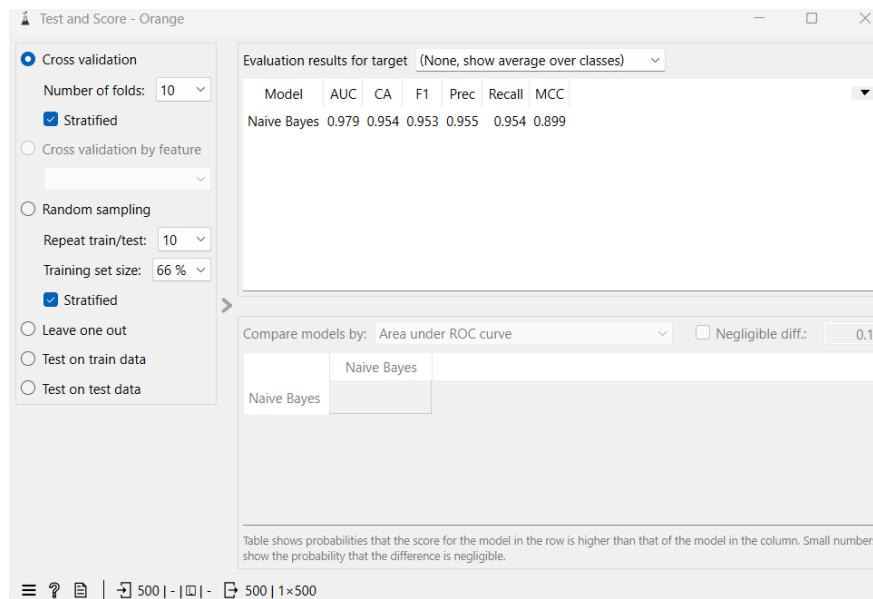


## BAB IV

### HASIL DAN PEMBAHASAN

#### 4.1. Performa Algoritma Naïve Bayes

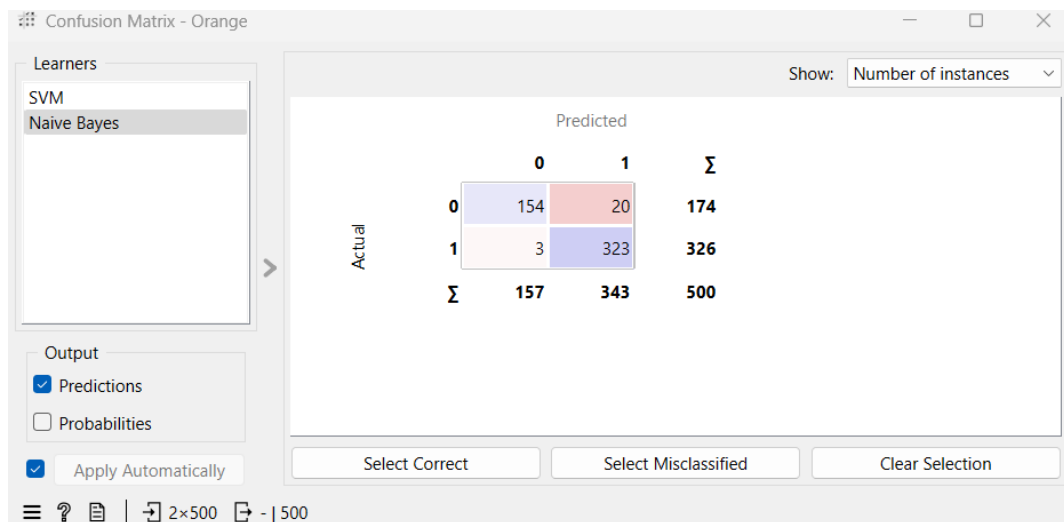
Pengujian model menggunakan algoritma Naïve Bayes dilakukan untuk mengetahui kemampuan model dalam mengklasifikasikan status perceraian berdasarkan faktor-faktor yang telah ditentukan. Hasil evaluasi diperoleh melalui widget *Test & Score* pada Orange menggunakan metode *percentage split* dengan pembagian 80% data training dan 20% data testing.



**Gambar 4. 1** Hasil Evaluasi Model Naïve Bayes pada Widget *Test & Score*

**Tabel 4. 1** Hasil Evaluasi Model Naïve Bayes

METRIK	NILAI
AUC	0.979 (97.9%)
ACCURACY	0.954 (95.4%)
PRECISION	0.955 (95.5%)
RECALL	0.954 (95.4%)
F1-SCORE	0.953 (95.3%)
MCC	0.899 (89.9%)



**Gambar 4. 2** Confusion Matrix Model Naïve Bayes

Penjelasan :

Total Data : 500

True Negatif (TN) : 154

False Positive (FP) : 20

False Negative (FN) : 3

True Positive (TP) : 323

Berdasarkan confusion matrix yang ditunjukkan pada Gambar 4.2, model Naïve Bayes berhasil mengklasifikasikan 154 data tidak cerai dengan benar (True Negative) dan 323 data cerai dengan benar (True Positive).

Namun, terdapat 20 data yang diprediksi sebagai cerai padahal sebenarnya tidak cerai (False Positive), serta 3 data cerai yang tidak terdeteksi oleh model (False Negative). Jumlah kesalahan prediksi yang relatif kecil menunjukkan bahwa model Naïve Bayes memiliki kemampuan klasifikasi yang baik dalam membedakan status perceraian.

Nilai False Negative yang rendah menunjukkan bahwa model cukup efektif dalam mendeteksi kasus perceraian yang sebenarnya.

Dataset yang digunakan dalam penelitian ini terdiri dari 500 data training yang digunakan untuk melatih model klasifikasi. Untuk menguji performa model yang telah dibangun, digunakan 100 data testing yang berfungsi sebagai data uji dalam mengevaluasi kemampuan model dalam mengklasifikasikan status perceraian.

**Tabel 4. 2** Data Training Naïve Bayes

No	Ekonomi	Narkoba	Judi	KDRT	Selingkuh	Komunikasi	Mabuk	Status
1	1	0	0	1	0	0	0	1
2	0	1	1	1	1	0	1	1
3	1	0	0	0	0	1	1	0
4	1	0	0	0	0	1	0	0
5	1	0	0	0	0	1	0	0
6	0	1	0	1	0	0	0	1
7	1	0	0	0	0	1	0	0

<b>8</b>	1	0	0	1	0	1	0	1
<b>9</b>	1	0	0	0	0	1	0	0
<b>10</b>	1	0	0	0	0	1	0	0
<b>11</b>	1	1	0	1	0	0	0	1
<b>12</b>	1	0	0	0	0	1	0	0
<b>13</b>	0	1	0	1	1	0	0	1
<b>14</b>	1	1	1	1	0	0	1	1
<b>15</b>	1	0	0	0	0	1	0	0
.....	.....	.....	.....	.....	.....	.....	.....	.....
<b>500</b>	1	0	1	0	0	1	0	0

Berdasarkan Tabel 4.2, data training yang digunakan dalam penelitian ini berjumlah 500 data yang merepresentasikan berbagai faktor penyebab perceraian, yaitu ekonomi, narkoba, judi, KDRT, selingkuh, komunikasi, dan mabuk sebagai variabel independen, serta status perceraian sebagai variabel dependen. Setiap variabel dinyatakan dalam bentuk biner (0 dan 1) untuk menunjukkan ada atau tidaknya faktor tersebut. Data ini digunakan sebagai input dalam proses pelatihan (training) model klasifikasi menggunakan algoritma *Naive Bayes*. Dengan adanya variasi kombinasi faktor dalam data training, model diharapkan mampu mengenali pola dan hubungan antar variabel sehingga dapat memprediksi status perceraian secara akurat.

**Tabel 4. 3** Penjelasan Nilai Atribut pada Dataset Naïve Bayes

<b>Atribut</b>	<b>Nilai 1</b>	<b>Nilai 0</b>
Ekonomi	Ada masalah ekonomi	Tidak ada masalah ekonomi
Narkoba	Ada penggunaan narkoba	Tidak ada penggunaan narkoba
Judi	Ada kebiasaan berjudi	Tidak berjudi
KDRT	Terjadi KDRT	Tidak terjadi KDRT
Selingkuh	Terjadi perselingkuhan	Tidak selingkuh
Komunikasi	Komunikasi buruk	Komunikasi baik
Mabuk	Ada kebiasaan mabuk	Tidak mabuk
Status	Bercerai	Masih menikah

**Tabel 4. 4** Data Testing Naïve Bayes

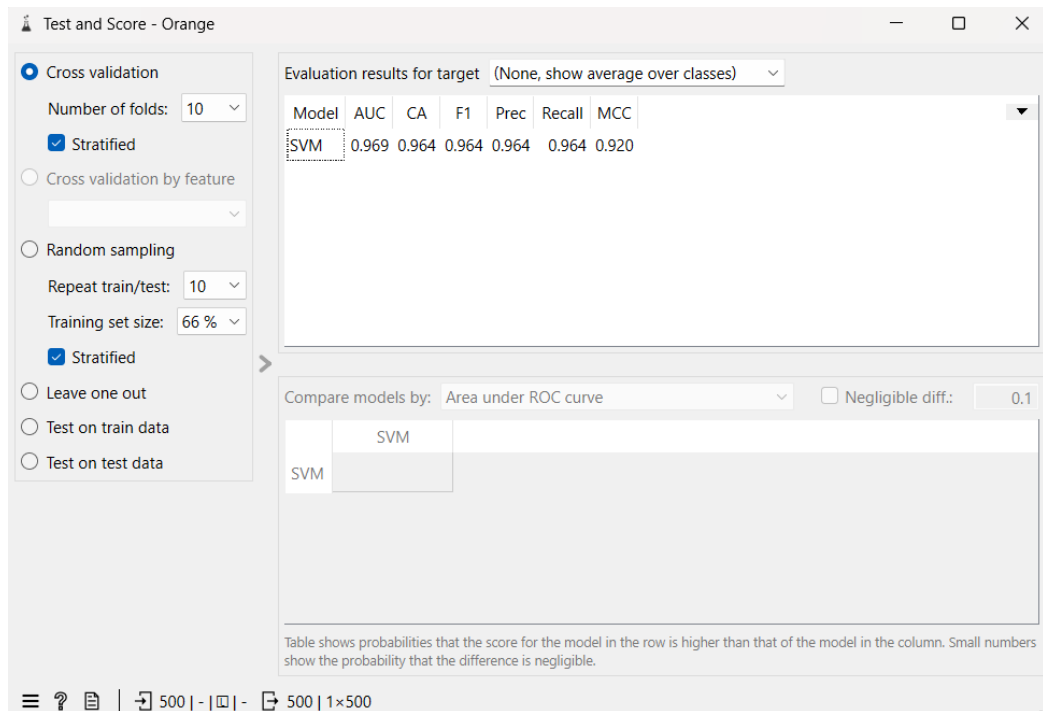
<b>No</b>	<b>Ekonomi</b>	<b>Narkoba</b>	<b>Judi</b>	<b>KDRT</b>	<b>Selingkuh</b>	<b>Komunikasi</b>	<b>Mabuk</b>	<b>Status</b>
<b>1</b>	1	0	1	0	0	1	1	0
<b>2</b>	1	0	0	1	1	1	0	1
<b>3</b>	0	0	0	0	0	1	0	0
<b>4</b>	1	0	0	1	1	0	0	1
<b>5</b>	1	1	0	1	1	1	0	1
<b>6</b>	1	0	0	0	0	1	0	0
<b>7</b>	1	0	0	1	1	1	0	1
<b>8</b>	0	0	0	1	0	0	0	1
<b>9</b>	0	0	0	1	1	1	0	1

<b>10</b>	0	0	0	0	0	1	0	1
.....	.....	.....	.....	.....	.....	.....	.....	.....
<b>100</b>	1	1	0	1	1	0	1	1

Data testing dalam penelitian ini berjumlah 100 data yang memiliki struktur variabel yang sama dengan data training, yaitu ekonomi, narkoba, judi, KDRT, selingkuh, komunikasi, dan mabuk sebagai variabel independen, serta status sebagai variabel dependen. Data ini digunakan untuk menguji kinerja model yang telah dilatih sebelumnya menggunakan algoritma *Naive Bayes*. Melalui data testing, dapat diketahui tingkat akurasi dan kemampuan model dalam memprediksi status perceraian pada data baru yang belum pernah dipelajari sebelumnya.

#### 4.2. Performa Algoritma SVM

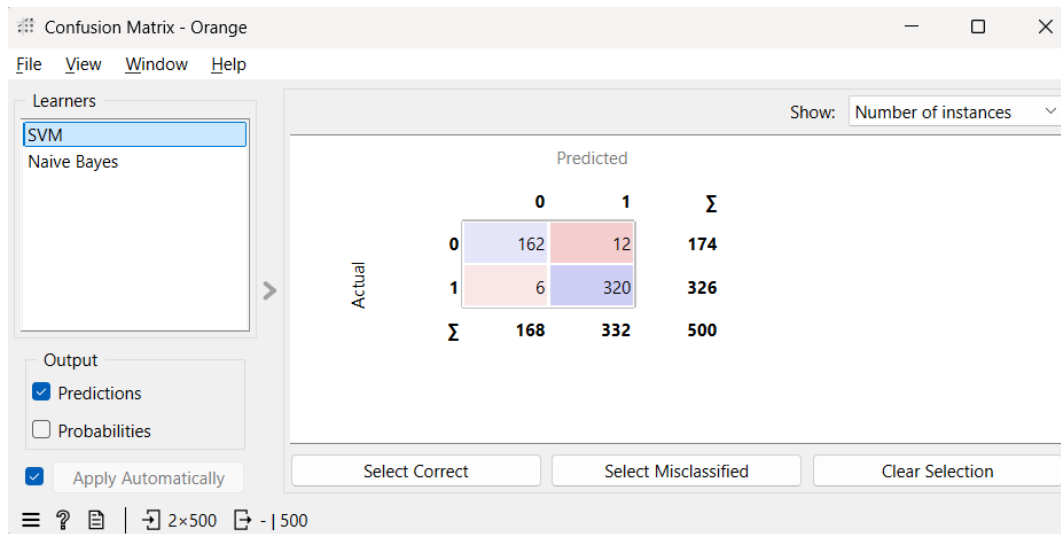
Pengujian model menggunakan algoritma SVM dilakukan untuk mengetahui kemampuan model dalam mengklasifikasikan status perceraian berdasarkan faktor-faktor yang telah ditentukan. Hasil evaluasi diperoleh melalui widget *Test & Score* pada Orange menggunakan metode *percentage split* dengan pembagian 80% data training dan 20% data testing.



**Gambar 4.3** Hasil Evaluasi Model SVM pada Widget Test & Score

**Tabel 4.5** Hasil Evaluasi Model SVM

METRIK	NILAI
AUC	0.969 (96.9%)
ACCURACY	0.964 (96.4%)
PRECISION	0.964 (96.4%)
RECALL	0.964 (96.4%)
F1-SCORE	0.964 (96.4%)
MCC	0.920 (92%)



**Gambar 4.4** *Confusion Matrix Model SVM*

Penjelasan :

Total Data : 500

True Negatif (TN) : 162

False Positive (FP) : 12

False Negative (FN) : 6

True Positive (TP) : 320

Berdasarkan confusion matrix yang ditunjukkan pada Gambar 4.4, model SVM berhasil mengklasifikasikan 162 data tidak cerai dengan benar (True Negative) dan 320 data cerai dengan benar (True Positive).

Namun, masih terdapat 12 data yang diprediksi sebagai cerai padahal sebenarnya tidak cerai (False Positive), serta 6 data cerai yang tidak terdeteksi oleh model (False Negative).

Jumlah kesalahan prediksi yang relatif kecil menunjukkan bahwa algoritma SVM memiliki kemampuan klasifikasi yang sangat baik dalam membedakan status perceraian. Nilai False Positive yang lebih rendah dibandingkan Naïve

Bayes menunjukkan bahwa SVM lebih berhati-hati dalam memprediksi kelas cerai.

Dataset yang digunakan dalam penelitian ini terdiri dari 500 data training yang digunakan untuk melatih model klasifikasi. Untuk menguji performa model yang telah dibangun, digunakan 100 data testing yang berfungsi sebagai data uji dalam mengevaluasi kemampuan model dalam mengklasifikasikan status perceraian.

**Tabel 4. 6 Data Training SVM**

No	Ekonomi	Narkoba	Judi	KDRT	Selingkuh	Komunikasi	Mabuk	Status
1	1	0	0	1	0	0	0	1
2	0	1	1	1	1	0	1	1
3	1	0	0	0	0	1	1	0
4	1	0	0	0	0	1	0	0
5	1	0	0	0	0	1	0	0
6	0	1	0	1	0	0	0	1
7	1	0	0	0	0	1	0	0
8	1	0	0	1	0	1	0	1
9	1	0	0	0	0	1	0	0
10	1	0	0	0	0	1	0	0
11	1	1	0	1	0	0	0	1
12	1	0	0	0	0	1	0	0
13	0	1	0	1	1	0	0	1
14	1	1	1	1	0	0	1	1
15	1	0	0	0	0	1	0	0
.....	.....	.....	.....	.....	.....	.....	.....	.....
500	1	0	1	0	0	1	0	0

Berdasarkan Tabel 4.6, data training yang digunakan dalam penelitian ini berjumlah 500 data yang merepresentasikan berbagai faktor penyebab perceraian, yaitu ekonomi, narkoba, judi, KDRT, selingkuh, komunikasi, dan mabuk sebagai variabel independen, serta status perceraian sebagai variabel dependen. Setiap variabel dinyatakan dalam bentuk biner (0 dan 1) untuk menunjukkan ada atau tidaknya faktor tersebut. Data ini digunakan sebagai input dalam proses pelatihan (training) model klasifikasi menggunakan algoritma *Support Vector Machine*. Dengan adanya variasi kombinasi faktor dalam data training, model diharapkan mampu mengenali pola dan hubungan antar variabel sehingga dapat memprediksi status perceraian secara akurat.

**Tabel 4. 7** Penjelasan Nilai Atribut pada Dataset SVM

<b>Atribut</b>	<b>Nilai 1</b>	<b>Nilai 0</b>
Ekonomi	Ada masalah ekonomi	Tidak ada masalah ekonomi
Narkoba	Ada penggunaan narkoba	Tidak ada penggunaan narkoba
Judi	Ada kebiasaan berjudi	Tidak berjudi
KDRT	Terjadi KDRT	Tidak terjadi KDRT
Selingkuh	Terjadi perselingkuhan	Tidak selingkuh
Komunikasi	Komunikasi buruk	Komunikasi baik
Mabuk	Ada kebiasaan mabuk	Tidak mabuk
Status	Bercerai	Masih menikah

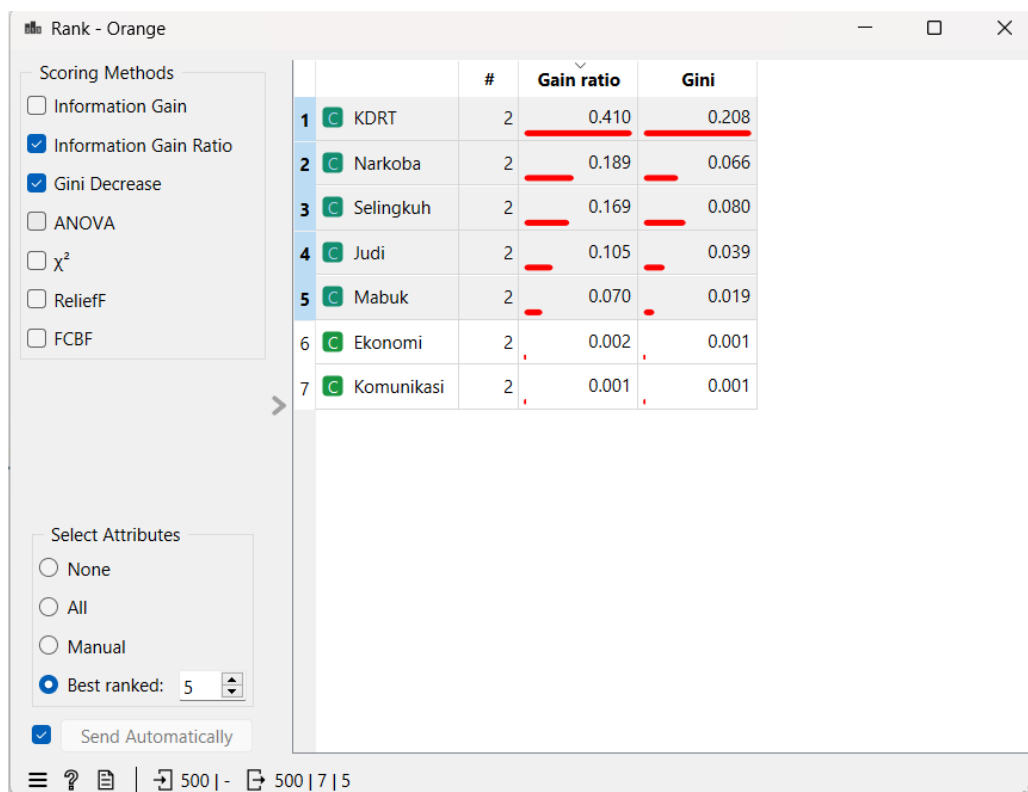
**Tabel 4. 8** Data Testing SVM

No	Ekonomi	Narkoba	Judi	KDRT	Selingkuh	Komunikasi	Mabuk	Status
1	1	0	1	0	0	1	1	0
2	1	0	0	1	1	1	0	1
3	0	0	0	0	0	1	0	0
4	1	0	0	1	1	0	0	1
5	1	1	0	1	1	1	0	1
6	1	0	0	0	0	1	0	0
7	1	0	0	1	1	1	0	1
8	0	0	0	1	0	0	0	1
9	0	0	0	1	1	1	0	1
10	0	0	0	0	0	1	0	1
.....	.....	.....	.....	.....	.....	.....	.....	.....
100	1	1	0	1	1	0	1	1

Data testing dalam penelitian ini berjumlah 100 data yang memiliki struktur variabel yang sama dengan data training, yaitu ekonomi, narkoba, judi, KDRT, selingkuh, komunikasi, dan mabuk sebagai variabel independen, serta status sebagai variabel dependen. Data ini digunakan untuk menguji kinerja model yang telah dilatih sebelumnya menggunakan algoritma *Support Vector Machine*. Melalui data testing, dapat diketahui tingkat akurasi dan kemampuan model dalam memprediksi status perceraian pada data baru yang belum pernah dipelajari sebelumnya.

### 4.3. Analisis Faktor Dominan Penyebab Perceraian

Analisis faktor dominan dalam penelitian ini dilakukan menggunakan widget *Rank* pada Orange dengan metode Gain Ratio untuk mengetahui atribut yang paling berpengaruh terhadap status perceraian pada dataset penelitian. Hasil perankingan atribut ditunjukkan pada Gambar 4.5.



		#	Gain ratio	Gini
1	KDRT	2	0.410	0.208
2	Narkoba	2	0.189	0.066
3	Selingkuh	2	0.169	0.080
4	Judi	2	0.105	0.039
5	Mabuk	2	0.070	0.019
6	Ekonomi	2	0.002	0.001
7	Komunikasi	2	0.001	0.001

**Gambar 4.5** Hasil Perankingan Atribut Menggunakan Gain Ratio pada Orange

Berdasarkan Gambar 4.5, faktor Kekerasan Dalam Rumah Tangga (KDRT) memperoleh nilai Gain Ratio tertinggi sebesar 0.410, diikuti oleh faktor Narkoba dan Selingkuh. Nilai tersebut menunjukkan bahwa KDRT merupakan faktor paling dominan dalam memengaruhi terjadinya perceraian pada dataset yang digunakan dalam penelitian ini.

#### 4.4. Perbandingan Kinerja Naive Bayes dan Support Vector Machine

Berdasarkan hasil evaluasi yang telah diperoleh pada subbab sebelumnya, dilakukan perbandingan kinerja antara algoritma Naïve Bayes dan Support Vector Machine (SVM). Perbandingan dilakukan berdasarkan indikator AUC, accuracy, precision, recall, F1-score, serta analisis confusion matrix.

**Tabel 4. 9** Perbandingan Kinerja Model Naïve Bayes dan Support Vector Machine

<b>METRIK</b>	<b>NAÏVE BAYES</b>	<b>SVM</b>
AUC	97.9%	96.9%
ACCURACY	95.4%	96.4%
PRECISION	95.5%	96.4%
RECALL	95.4%	96.4%
F1-SCORE	95.3%	96.4%
MCC	89.9%	92%

Berdasarkan Tabel 4.9, algoritma SVM menunjukkan nilai accuracy sebesar 96.4%, sedikit lebih tinggi dibandingkan Naïve Bayes yang memperoleh 95.4%. Hal ini menunjukkan bahwa SVM memiliki tingkat ketepatan yang lebih baik dalam mengklasifikasikan status perceraian.

Dari sisi precision dan recall, SVM juga menunjukkan nilai yang lebih tinggi dibandingkan Naïve Bayes. Hal ini mengindikasikan bahwa SVM memiliki keseimbangan yang lebih baik dalam memprediksi dan mendeteksi kasus perceraian.

Namun, berdasarkan analisis confusion matrix, Naïve Bayes memiliki jumlah False Negative yang lebih kecil dibandingkan SVM, yang berarti Naïve Bayes sedikit lebih baik dalam mendeteksi seluruh kasus cerai yang sebenarnya. Sebaliknya, SVM memiliki jumlah False Positive yang lebih rendah, sehingga lebih akurat dalam menghindari kesalahan prediksi cerai pada kasus yang sebenarnya tidak cerai.

Secara keseluruhan, meskipun kedua algoritma menunjukkan performa yang sangat baik dengan tingkat akurasi di atas 95%, algoritma Support Vector Machine (SVM) dapat dikatakan lebih unggul berdasarkan nilai accuracy, precision, recall, F1-score, dan MCC yang lebih tinggi.

Berdasarkan hasil seleksi fitur menggunakan metode *Gain Ratio* dan *Gini*, diketahui bahwa variabel yang paling dominan dalam mempengaruhi perceraian adalah KDRT, dengan nilai Gain Ratio sebesar 0,410 dan Gini sebesar 0,208. Hal ini menunjukkan bahwa KDRT memiliki kontribusi paling besar dalam proses klasifikasi dibandingkan variabel lainnya. Selanjutnya, variabel narkoba, selingkuh, judi, dan mabuk juga memberikan pengaruh yang cukup signifikan, meskipun nilainya lebih rendah dibandingkan KDRT. Sementara itu, variabel ekonomi dan komunikasi memiliki nilai yang sangat kecil, sehingga kontribusinya dalam mempengaruhi perceraian tergolong rendah. Dengan demikian, dapat disimpulkan bahwa faktor kekerasan dalam rumah tangga merupakan faktor yang paling dominan dalam menentukan terjadinya perceraian berdasarkan hasil analisis yang dilakukan.

Selain pengujian pada dataset penelitian, dilakukan juga perbandingan dengan penelitian terdahulu pada kasus prediksi prestasi akademik. Berdasarkan penelitian tersebut, algoritma Support Vector Machine (SVM) cenderung memiliki performa yang lebih baik dibandingkan Naïve Bayes, terutama dalam hal akurasi dan kestabilan model. Namun, Naïve Bayes tetap menunjukkan kinerja yang kompetitif pada kondisi data tertentu. Dengan demikian, hasil penelitian ini sejalan dengan penelitian terdahulu yang menunjukkan bahwa SVM lebih unggul dibandingkan Naïve Bayes.