

BAB IV

HASIL DAN PEMBAHASAN

4.1. Analisis Data

Pada tahapan analisis data merupakan tahapan yang dilakukan untuk menentukan data yang akan digunakan pada penelitian ini. Dengan adanya data, maka penelitian ini dapat dilakukan. Untuk data set yang akan digunakan yaitu terdapat 2 data set yang akan digunakan. Untuk data set yang pertama yaitu data training dan data set yang kedua yaitu data testing.

Data Training

Nama Penggugat	Usia Penggugat	Pekerjaan Penggugat	Lama Pernikahan	Jumlah Anak	Penyebab Perceraian
Laila Rahma	32	Dokter	10 Tahun	2	Masalah Pribadi
Andi Setiawan	35	Karyawan	10 Tahun	2	Ketidakkcocokan
Chandra Wijaya	37	Karyawan	12 Tahun	2	Kecurangan
Indra Gunawan	38	Karyawan	12 Tahun	2	Masalah Keuangan
Udin Jafar	39	Pengusaha	13 Tahun	2	Kecurangan
Putri Melati	40	Karyawan	13 Tahun	3	Ketidakkcocokan
Eka Putra	40	Pengusaha	14 Tahun	3	Perbedaan Pribadi
Budi Santoso	42	Pengusaha	15 Tahun	3	Masalah Keuangan
Tono Wibowo	42	Karyawan	16 Tahun	3	Ketidakkcocokan
Hadi Saputra	45	PNS	20 Tahun	4	Ketidakkcocokan
Made Surya	47	Karyawan	22 Tahun	3	Ketidakkcocokan
Rizki Fadillah	50	Pengusaha	25 Tahun	4	Masalah Keuangan

Nia Saraswati	29	Karyawan	4 Tahun	1	Perselisihan
Kiki Amelia	28	Ibu Rumah Tangga	5 Tahun	1	Ketidakkcocokan
Dian Purnama	29	Dokter	5 Tahun	1	Ketidakkcocokan
Sari Utami	30	Karyawan	6 Tahun	1	Perselisihan
Joko Susanto	31	Pengusaha	6 Tahun	1	Perselisihan
Gita Lestari	30	Karyawan	7 Tahun	1	Kecurangan
Fajar Hidayat	34	PNS	8 Tahun	2	Perselisihan
Oki Pratama	35	PNS	9 Tahun	2	Kecurangan

Pada table diatas merupakan data training yang digunakan. Untuk data training yang digunakan terdapat sebanyak 20 data training.

Data Testing

Nama Penggugat	Usia Penggugat	Pekerjaan Penggugat	Lama Pernikahan	Jumlah Anak	Penyebab Perceraian
Laila Rahma	32	Dokter	10 Tahun	2	Masalah Pribadi
Andi Setiawan	35	Karyawan	10 Tahun	2	Ketidakkcocokan
Chandra Wijaya	37	Karyawan	12 Tahun	2	Kecurangan
Indra Gunawan	38	Karyawan	12 Tahun	2	Masalah Keuangan
Udin Jafar	39	Pengusaha	13 Tahun	2	Kecurangan
Putri Melati	40	Karyawan	13 Tahun	3	Ketidakkcocokan
Eka Putra	40	Pengusaha	14 Tahun	3	Perbedaan Pribadi
Budi Santoso	42	Pengusaha	15 Tahun	3	Masalah Keuangan
Tono Wibowo	42	Karyawan	16 Tahun	3	Ketidakkcocokan
Hadi Saputra	45	PNS	20 Tahun	4	Ketidakkcocokan

Made Surya	47	Karyawan	22 Tahun	3	Ketidakcocokan
Rizki Fadillah	50	Pengusaha	25 Tahun	4	Masalah Keuangan
Nia Saraswati	29	Karyawan	4 Tahun	1	Perselisihan
Kiki Amelia	28	Ibu Rumah Tangga	5 Tahun	1	Ketidakcocokan
Dian Purnama	29	Dokter	5 Tahun	1	Ketidakcocokan
Sari Utami	30	Karyawan	6 Tahun	1	Perselisihan
Joko Susanto	31	Pengusaha	6 Tahun	1	Perselisihan
Gita Lestari	30	Karyawan	7 Tahun	1	Kecurangan
Fajar Hidayat	34	PNS	8 Tahun	2	Perselisihan
Oki Pratama	35	PNS	9 Tahun	2	Kecurangan

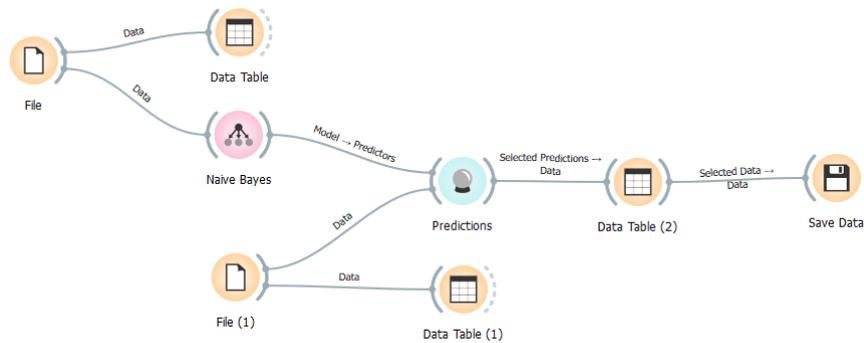
Pada table diatas merupakan tabel yang digunakan sebagai data testing ataupun data sampe Penelitian. Untuk data yang digunakan sebanyak 30 data testing.

4.2. Preprocessing Data

Pada tahap preprocessing data merupakan tahapan yang dilakukan untuk membersihkan data dari data yang tidak layak untuk digunakan. Jadi dengan adanya tahapan ini, maka tidak akan ada data yang salah dalam penelitian ini.

4.3. Perancangan Model Klasifikasi

Perancangan model klasifikasi merupakan model yang akan digunakan untuk melakukan klasifikasi data pada data mining. Dengan adanya model ini, data akana dengan mudah diklasifikasikan pada data mining.



Pada gambar diatas merupakan model yang digunakan untuk melakukan klasifikasi data menggunakan metode naive bayes.

4.4. Hasil Klasifikasi

nama Penggugat	Usia Penggugat	Pekerjaan Penggugat	Usia Tergugat	Pekerjaan Tergugat	Lama Pernikahan	Jumlah Anak	Penyebab Perceraian
Andi Setiawan	35	Karyawan	33	Ibu Rumah Tangga	10 Tahun	2	Ketidaccokokan
Budi Santoso	42	Pengusaha	39	Guru	15 Tahun	3	Masalah Keuangan
Chandra Wijaya	37	Karyawan	35	Karyawan	12 Tahun	2	Kecurangaan
Dian Purnama	29	Dokter	27	Perawat	5 Tahun	1	Ketidaccokokan
Eka Putra	40	Pengusaha	38	Ibu Rumah Tangga	14 Tahun	3	Perbedaan Pribadi
Fajar Hidayat	34	PNS	32	Karyawan	8 Tahun	2	Perselisihan
Gita Lestari	30	Karyawan	32	Karyawan	7 Tahun	1	Kecurangaan
Hadi Saputra	45	PNS	40	Guru	20 Tahun	4	Ketidaccokokan
Indra Gunawan	38	Karyawan	36	Ibu Rumah Tangga	12 Tahun	2	Masalah Keuangan

Joko Susanto	31	Pengusaha	29	Karyawan	6 Tahun	1	Perselisihan
Kiki Amelia	28	Ibu Rumah Tangga	30	Karyawan	5 Tahun	1	Ketidaccokokan
Laila Rahma	32	Dokter	34	Dokter	10 Tahun	2	Masalah Pribadi
Made Surya	47	Karyawan	44	Ibu Rumah Tangga	22 Tahun	3	Ketidaccokokan
Nia Saraswati	29	Karyawan	27	Karyawan	4 Tahun	1	Perselisihan
Oki Pratama	35	PNS	33	Ibu Rumah Tangga	9 Tahun	2	Kecurangan
Putri Melati	40	Karyawan	38	Pengusaha	13 Tahun	3	Ketidaccokokan
Rizki Fadillah	50	Pengusaha	48	Ibu Rumah Tangga	25 Tahun	4	Masalah Keuangan
Sari Utami	30	Karyawan	28	Karyawan	6 Tahun	1	Perselisihan
Tono Wibowo	42	Karyawan	39	Ibu Rumah Tangga	16 Tahun	3	Ketidaccokokan
Udin Jafar	39	Pengusaha	37	Karyawan	13 Tahun	2	Kecurangan

Pada tabel diatas merupakan tabel hasil klasifikasi yang dilakukan. Untuk hasil klasifikasi yang diperoleh yaitu dari 30 data sampel, terdapat 24 orang mengalami sakit dan sisanya 6 orang tidak sakit.

4.5. Evaluasi Metode Naive Bayes

Evaluasi dilakukan untuk mengetahui hasil kemampuan metode dalam melakukan klasifikasi. Untuk evaluasi yang akan dipeorleh yaitu Test and Score dan Confusion Matrix.

4.6. Hasil Evaluasi

Test and Score

Model	AUC	CA	F1	Precision	Recall	MCC
Naïve Bayes	1.000	1.000	1.000	1.000	1.000	1.000

Hasil evaluasi yang diperoleh metode naïve bayes yaitu sempurna. Hal ini karena hasil yang diperoleh mencapai nilai 100%.

Confusion Matrix

		Predicted						Σ
		Kecurangan	Ketidakkcocokan	Masalah Keuangan	Masalah Pribadi	Perbedaan Pribadi	Perselisihan	
Actual	Kecurangan	25.0 %	0.0 %	0.0 %	25.0 %	25.0 %	25.0 %	4
	Ketidakkcocokan	0.0 %	28.6 %	14.3 %	28.6 %	14.3 %	14.3 %	7
	Masalah Keuangan	33.3 %	0.0 %	66.7 %	0.0 %	0.0 %	0.0 %	3
	Masalah Pribadi	0.0 %	0.0 %	0.0 %	100.0 %	0.0 %	0.0 %	1
	Perbedaan Pribadi	0.0 %	0.0 %	0.0 %	0.0 %	100.0 %	0.0 %	1
	Perselisihan	0.0 %	0.0 %	0.0 %	25.0 %	0.0 %	75.0 %	4
Σ		2	2	3	5	3	5	20

Pada tabel di atas merupakan hasil evaluasi yang dilakukan pada metode Neural Network dengan menggunakan widget Confusion Matrix. Hasilnya menunjukkan bahwa terdapat 20 data yang termasuk dalam kategori True Positif (TP), 6 data dalam kategori True Negatif (TN), 0 data dalam kategori False Positif (FP), dan 0 data dalam kategori False Negatif (FN). Meskipun hasil ini tampak sangat baik, untuk mengukur nilai akurasi secara tepat, data tersebut harus dihitung terlebih dahulu menggunakan rumus yang ada pada confusion matrix, yaitu sebagai berikut.

$$Accuracy = \frac{24+6}{24+6+0+0} + 100\% \quad \text{Then the Accuracy value} = 100\%$$

$$Presisi = \frac{24}{24+0} + 100\% \quad \text{Then the Precision value} = 100\%$$

$$\mathbf{Recall} = \frac{20}{20+0} + 100\% \quad \text{Then the Recall value} = 100\%$$

Dari hasil confusion matrix bahwa metode naïve bayes memberikan hasil 100% yaitu hasil nya sempurna.