BAB IV

HASIL DAN PEMBAHASAN

4.1. Pengumpulan Data

Pada tahap pengumpulan data, proses ini bertujuan untuk menentukan data yang akan digunakan dalam penelitian ini. Data yang diperlukan akan dikumpulkan secara sistematis agar penelitian dapat dilaksanakan dengan baik dan tepat.

4.2. Preprocessing Data

Proses preprocessing data merupakan tahap penting dalam analisis data yang melibatkan pembersihan, pengorganisasian, serta persiapan dataset untuk meningkatkan kualitas sebelum dilakukan analisis. Tahap ini mencakup penanganan data yang hilang, normalisasi, pengkodean variabel kategorikal, serta penghapusan outlier. Dengan melakukan preprocessing yang baik, kualitas data dapat ditingkatkan, sehingga mendukung keakuratan dan konsistensi hasil analisis.

4.3. Pembagian Data

Pada tahap pembagian daya, data akan dibagi menjadi 2 dataset yaitu sebagai berikut.

Data Training

Data training merupakan data yang akan digunakan untuk membantu proses klasifikasi data. Jadi untuk data training yang akan digunakan yaitu sebanyak 30 data. Untuk data training akan disajikan dalam bentuk table seperti berikut.

Tabel 4. 1. Data Training

Nama Lengkap	Akses Menuju Lokasi	Fasilitas Sekolah	Pemahaman TIK	Dukungan Orang Tua	Biaya Sekolah	Kategori
Alma Kirana	Sulit	Lengkap	Bagus	Setuju	Murah	Minat
Bayu Damar	Mudah	Lengkap	Bagus	Tidak Setuju	Mahal	Minat
Caca Rania	Mudah	Lengkap	Bagus	Setuju	Murah	Minat
Deni Raditya	Mudah	Lengkap	Kurang Bagus	Setuju	Murah	Minat
Echa Mahira	Mudah	Lengkap	Bagus	Tidak Setuju	Mahal	Minat
Fiko Rizky	Sulit	Lengkap	Bagus	Setuju	Murah	Minat
Gibran Pratama	Mudah	Tidak Lengkap	Bagus	Setuju	Mahal	Minat
Hani Najwa	Mudah	Lengkap	Bagus	Setuju	Murah	Minat
Ilham Raka	Mudah	Lengkap	Bagus	Setuju	Murah	Minat
Juno Prasetya	Mudah	Lengkap	Bagus	Tidak Setuju	Mahal	Minat
Kezia Melani	Mudah	Lengkap	Kurang Bagus	Tidak Setuju	Mahal	Tidak Minat
Lani Pramesti	Sulit	Lengkap	Bagus	Setuju	Murah	Minat
Lian Syafiq	Mudah	Lengkap	Kurang Bagus	Tidak Setuju	Mahal	Tidak Minat
Miko Farrel	Mudah	Lengkap	Bagus	Tidak Setuju	Mahal	Minat
Naya Safira	Mudah	Lengkap	Kurang Bagus	Setuju	Murah	Minat
Oki Adrian	Sulit	Lengkap	Bagus	Setuju	Murah	Minat
Putri Zahira	Mudah	Tidak Lengkap	Bagus	Tidak Setuju	Mahal	Tidak Minat
Raya Fauzan	Mudah	Lengkap	Bagus	Setuju	Murah	Minat
Sinta Maharani	Sulit	Lengkap	Bagus	Setuju	Murah	Minat
Tio Bintang	Mudah	Lengkap	Bagus	Tidak Setuju	Mahal	Minat
Umar Dani	Sulit	Lengkap	Kurang Bagus	Setuju	Mahal	Tidak Minat
Uta Alvin	Mudah	Lengkap	Bagus	Tidak Setuju	Mahal	Minat
Vika Widyanti	Mudah	Lengkap	Kurang Bagus	Tidak Setuju	Mahal	Tidak Minat
Vino Althaf	Mudah	Tidak Lengkap	Bagus	Setuju	Mahal	Minat
Wira Damar	Mudah	Lengkap	Kurang Bagus	Tidak Setuju	Mahal	Tidak Minat
Wulan Larasati	Mudah	Lengkap	Kurang Bagus	Setuju	Murah	Minat
Xavi Dimas	Sulit	Lengkap	Bagus	Setuju	Murah	Minat
Yara Nadhira	Mudah	Lengkap	Kurang Bagus	Tidak Setuju	Mahal	Tidak Minat
Zaki Ilham	Mudah	Lengkap	Bagus	Setuju	Murah	Minat
Zidan Alif	Mudah	Lengkap	Bagus	Setuju	Murah	Minat

Tabel di atas menunjukkan data training yang akan digunakan dalam proses klasifikasi, yang terdiri dari 30 data siswa dan siswi yang dikumpulkan melalui distribusi kuesioner terkait minat mereka pada Jurusan TKJ.

Data Testing

Data testing merupakan data yang nantinya akan menjadi sampel Penelitian pada data mining menggunakan metode Naive Bayes. Untuk data testing yang akan digunakan yaitu sebagai berikut.

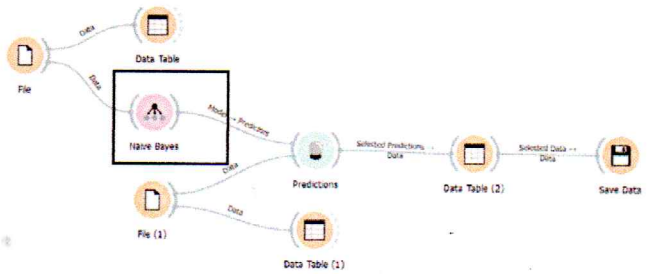
Tabel 4. 2. Data Testing

Nama Lengkap	Akses Menuju Lokasi	Fasilitas Sekolah	Pemahaman TIK	Dukungan Orang Tua	Biaya Sekolah
Agus Suryono	Sulit	Lengkap	Kurang Bagus	Setuju	Mahal
Aldi Supriyadi	Mudah	Tidak Lengkap	Bagus	Setuju	Mahal
Alif Syahrul	Sulit	Lengkap	Kurang Bagus	Setuju	Mahal
Arif Rahman	Mudah	Lengkap	Kurang Bagus	Setuju	Murah
Ayu Lestari	Sulit	Lengkap	Bagus	Setuju	Murah
Bima Aryo	Mudah	Tidak Lengkap	Bagus	Setuju	Mahal
Bunga Melati	Mudah	Lengkap	Kurang Bagus	Tidak Setuju	Mahal
Citra Aulia	Mudah	Lengkap	Kurang Bagus	Setuju	Murah
Daffa Arjuna	Sulit	Lengkap	Bagus	Setuju	Murah
Desi Ambarwati	Mudah	Lengkap	Bagus	Setuju	Murah
Dina Kartika	Sulit	Lengkap	Bagus	Setuju	Murah
Eka Purnama	Mudah	Lengkap	Bagus	Tidak Setuju	Mahal
Eko Priyono	Mudah	Lengkap	Bagus	Setuju	Murah
Elma Salsabila	Mudah	Lengkap	Kurang Bagus	Tidak Setuju	Mahal
Farel Aditya	Mudah	Lengkap	Bagus	Setuju	Murah
Fitri Handayani	Mudah	Lengkap	Kurang Bagus	Setuju	Murah
Gita Zahra	Sulit	Lengkap	Bagus	Setuju	Murah
Hani Rizky	Mudah	Lengkap	Bagus	Tidak Setuju	Mahal
Hendra Saputra	Mudah	Lengkap	Bagus	Tidak Setuju	Mahal
Icha Lestari	Mudah	Lengkap	Bagus	Setuju	Murah
Imam Subekti	Sulit	Lengkap	Bagus	Setuju	Murah
Iqbal Firdaus	Mudah	Tidak Lengkap	Bagus	Setuju	Mahal
Juna Abimanyu	Mudah	Lengkap	Kurang Bagus	Setuju	Murah
Keyla Nadira	Mudah	Lengkap	Bagus	Tidak Setuju	Mahal
Kiki Hermansyah	Mudah	Lengkap	Bagus	Setuju	Murah
Lala Putri	Sulit	Lengkap	Bagus	Setuju	Murah
Maya Anindya	Mudah	Lengkap	Bagus	Setuju	Murah
Maya Rizqia	Mudah	Lengkap	Bagus	Tidak Setuju	Mahal
Mika Rahman	Mudah	Tidak Lengkap	Bagus	Setuju	Mahal
Nina Rahmawati	Mudah	Lengkap	Kurang Bagus	Tidak Setuju	Mahal
Nisa Amalia	Mudah	Lengkap	Bagus	Setuju	Murah
Nurul Aisyah	Mudah	Lengkap	Kurang Bagus	Tidak Setuju	Mahal
Ojan Firdaus	Mudah	Lengkap	Bagus	Setuju	Murah
Puti Anindya	Mudah	Lengkap	Bagus	Tidak Setuju	Mahal
Putra Mahardika	Mudah	Lengkap	Bagus	Tidak Setuju	Mahal
Qila Annisa	Mudah	Lengkap	Kurang Bagus	Tidak Setuju	Mahal
Raka Nugraha	Mudah	Lengkap	Kurang Bagus	Tidak Setuju	Mahal
Ratna Dewi	Mudah	Lengkap	Kurang Bagus	Setuju	Murah
Ridwan Syahputra	Sulit	Lengkap	Bagus	Setuju	Murah
Rudi Hartono	Mudah	Tidak Lengkap	Bagus	Tidak Setuju	Mahal
Salsabila Putri	Mudah	Lengkap	Bagus	Tidak Setuju	Mahal
Sari Ningsih	Mudah	Lengkap	Bagus	Setuju	Murah
Sari Wulandari	Sulit	Lengkap	Bagus	Setuju	Murah
Susi Marlina	Mudah	Lengkap	Bagus	Tidak Setuju	Mahal
Taufik Hidayat	Mudah	Lengkap	Bagus	Tidak Setuju	Mahal
Tiara Maharani	Mudah	Lengkap	Kurang Bagus	Tidak Setuju	Mahal
Tiara Melati	Mudah	Lengkap	Kurang Bagus	Setuju	Murah
Toni Wibowo	Mudah	Lengkap	Kurang Bagus	Tidak Setuju	Mahal
Winda Kusuma	Mudah	Lengkap	Bagus	Setuju	Murah
Yoga Setiawan	Sulit	Lengkap	Bagus	Setuju	Murah

Pada tabel di atas merupakan data sampel penelitian yang akan digunakan dalam penelitian ini terkait minat siswa dan siswi pada jurusan TKJ (Teknik Komputer dan Jaringan) menggunakan metode data mining. Jumlah data sampel yang digunakan sebanyak 50 siswa dan siswi, yang diperoleh melalui pembagian kuesioner kepada mereka.

4.4. Pembentukan dan Pengujian Model Metode Naive Bayes

Pembentukan model ini merupakan perancangan sistem yang akan digunakan untuk menganalisis minat siswa dan siswi pada Jurusan TKJ. Sistem ini nantinya akan digunakan untuk melakukan klasifikasi minat tersebut dengan menggunakan metode Naive Bayes.



Gambar 4. 1. Widget Patterns in Data Mining

Gambar di atas menunjukkan sebuah model-model yang digunakan untuk proses klasifikasi menggunakan algoritma Naive Bayes. Alur kerja ini dimulai dengan dua widget File, yang masing-masing digunakan untuk mengimpor dataset ke dalam sistem. Widget File pertama mengarahkan datanya ke Data Table, di mana data yang diimpor dapat dilihat dan diverifikasi sebelum diproses lebih lanjut. File yang kedua langsung dihubungkan dengan Naive Bayes. Widget Naive Bayes berfungsi sebagai inti dari proses klasifikasi. Di sini, data yang diimpor dilatih untuk membuat model klasifikasi berdasarkan fitur yang ada. Hasil dari pelatihan

model ini diteruskan ke widget Predictions, yang menghasilkan prediksi atau klasifikasi berdasarkan model yang telah dibangun.

Prediksi yang dihasilkan oleh Naive Bayes kemudian diolah lebih lanjut. Hasilnya dikirimkan ke Data Table (2) untuk dilihat dan diverifikasi. Widget Data Table (2) menampilkan data prediksi yang dipilih untuk memastikan bahwa hasilnya sesuai dengan yang diharapkan. Terakhir, data hasil prediksi ini disimpan menggunakan widget Save Data. Widget ini memungkinkan pengguna untuk menyimpan data hasil analisis dalam format yang diinginkan, sehingga bisa digunakan untuk keperluan lebih lanjut. Secara keseluruhan, workflow ini menggambarkan proses yang terstruktur dari mengimpor data, melatih model, menghasilkan prediksi, dan menyimpan hasil akhir untuk analisis lebih lanjut.

4.5. Hasil Klasifikasi Model Metode Naive Bayes

Pada tahapan ini merupakan hasil klasifikasi dalam data mining menggunakan metode Naive Bayes. Dari 50 data yang digunakan, sebanyak 39 data menunjukkan minat siswa dan siswi terhadap Jurusan TKJ di SMKS Al-Washliyah 2 Merbau, sementara 11 data menunjukkan tidak minat terhadap Jurusan TKJ. Hasil ini menjelaskan bahwa mayoritas siswa dan siswi menunjukkan minat terhadap Jurusan TKJ di SMKS Al-Washliyah 2 Merbau. Tabel hasil klasifikasi dapat dilihat pada tabel di bawah ini.

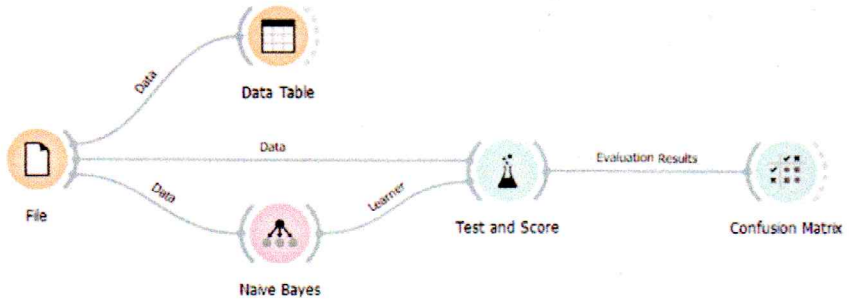
Tabel 4. 3. Classification Results and Model Predictions

Nama Lengkap	Akses Menuju Lokasi	Fasilitas Sekolah	Pemahaman TIK	Dukungan Orang Tua	Biaya Sekolah	Kategori
Agus Suryono	Sulit	Lengkap	Kurang Bagus	Setuju	Mahal	Tidak Minat
Aldi Supriyadi	Mudah	Tidak Lengkap	Bagus	Setuju	Mahal	Minat
Alif Syahrul	Sulit	Lengkap	Kurang Bagus	Setuju	Mahal	Tidak Minat
Arif Rahman	Mudah	Lengkap	Kurang Bagus	Setuju	Murah	Minat
Ayu Lestari	Sulit	Lengkap	Bagus	Setuju	Murah	Minat
Bima Aryo	Mudah	Tidak Lengkap	Bagus	Setuju	Mahal	Minat

Bunga Melati	Mudah	Lengkap	Kurang Bagus	Tidak Setuju	Mahal	Tidak Minat
Citra Aulia	Mudah	Lengkap	Kurang Bagus	Setuju	Murah	Minat
Daffa Arjuna	Sulit	Lengkap	Bagus	Setuju	Murah	Minat
Desi Ambarwati	Mudah	Lengkap	Bagus	Setuju	Murah	Minat
Dina Kartika	Sulit	Lengkap	Bagus	Setuju	Murah	Minat
Eka Purnama	Mudah	Lengkap	Bagus	Tidak Setuju	Mahal	Minat
Eko Priyono	Mudah	Lengkap	Bagus	Setuju	Murah	Minat
Elma Salsabila	Mudah	Lengkap	Kurang Bagus	Tidak Setuju	Mahal	Tidak Minat
Farel Aditya	Mudah	Lengkap	Bagus	Setuju	Murah	Minat
Fitri Handayani	Mudah	Lengkap	Kurang Bagus	Setuju	Murah	Minat
Gita Zahra	Sulit	Lengkap	Bagus	Setuju	Murah	Minat
Hani Rizky	Mudah	Lengkap	Bagus	Tidak Setuju	Mahal	Minat
Hendra Saputra	Mudah	Lengkap	Bagus	Tidak Setuju	Mahal	Minat
Icha Lestari	Mudah	Lengkap	Bagus	Setuju	Murah	Minat
Imam Subekti	Sulit	Lengkap	Bagus	Setuju	Murah	Minat
Iqbal Firdaus	Mudah	Tidak Lengkap	Bagus	Setuju	Mahal	Minat
Juna Abimanyu	Mudah	Lengkap	Kurang Bagus	Setuju	Murah	Minat
Keyla Nadira	Mudah	Lengkap	Bagus	Tidak Setuju	Mahal	Minat
Kiki Hermansyah	Mudah	Lengkap	Bagus	Setuju	Murah	Minat
Lala Putri	Sulit	Lengkap	Bagus	Setuju	Murah	Minat
Maya Anindya	Mudah	Lengkap	Bagus	Setuju	Murah	Minat
Maya Rizqia	Mudah	Lengkap	Bagus	Tidak Setuju	Mahal	Minat
Mika Rahman	Mudah	Tidak Lengkap	Bagus	Setuju	Mahal	Minat
Nina Rahmawati	Mudah	Lengkap	Kurang Bagus	Tidak Setuju	Mahal	Tidak Minat
Nisa Amalia	Mudah	Lengkap	Bagus	Setuju	Murah	Minat
Nurul Aisyah	Mudah	Lengkap	Kurang Bagus	Tidak Setuju	Mahal	Tidak Minat
Ojan Firdaus	Mudah	Lengkap	Bagus	Setuju	Murah	Minat
Puti Anindya	Mudah	Lengkap	Bagus	Tidak Setuju	Mahal	Minat
Putra Mahardika	Mudah	Lengkap	Bagus	Tidak Setuju	Mahal	Minat
Qila Annisa	Mudah	Lengkap	Kurang Bagus	Tidak Setuju	Mahal	Tidak Minat
Raka Nugraha	Mudah	Lengkap	Kurang Bagus	Tidak Setuju	Mahal	Tidak Minat
Ratna Dewi	Mudah	Lengkap	Kurang Bagus	Setuju	Murah	Minat
Ridwan Syahputra	Sulit	Lengkap	Bagus	Setuju	Murah	Minat
Rudi Hartono	Mudah	Tidak Lengkap	Bagus	Tidak Setuju	Mahal	Tidak Minat
Salsabila Putri	Mudah	Lengkap	Bagus	Tidak Setuju	Mahal	Minat
Sari Ningsih	Mudah	Lengkap	Bagus	Setuju	Murah	Minat
Sari Wulandari	Sulit	Lengkap	Bagus	Setuju	Murah	Minat
Susi Marlina	Mudah	Lengkap	Bagus	Tidak Setuju	Mahal	Minat
Taufik Hidayat	Mudah	Lengkap	Bagus	Tidak Setuju	Mahal	Minat
Tiara Maharani	Mudah	Lengkap	Kurang Bagus	Tidak Setuju	Mahal	Tidak Minat
Tiara Melati	Mudah	Lengkap	Kurang Bagus	Setuju	Murah	Minat
Toni Wibowo	Mudah	Lengkap	Kurang Bagus	Tidak Setuju	Mahal	Tidak Minat
Winda Kusuma	Mudah	Lengkap	Bagus	Setuju	Murah	Minat
Yoga Setiawan	Sulit	Lengkap	Bagus	Setuju	Murah	Minat

4.6. Evaluasi Model Metode Naive Bayes

Evaluasi model adalah langkah penting untuk menentukan tingkat akurasi metode yang digunakan, dengan tujuan menilai sejauh mana metode tersebut efektif dalam mengklasifikasikan data.



Gambar 4. 2. Evaluation of the C4.5 Algorithm Classification Model

Pada gambar di atas terdapat sistem yang digunakan untuk mengevaluasi metode yang diterapkan. Akurasi hasil dievaluasi menggunakan beberapa widget, yaitu test and score serta confusion matrix. Hasil akurasi yang diperoleh adalah sebagai berikut.

Tabel 4. 4. Result of Test and Score

Model	AUC	CA	F1	Precision	Recall
Naive Bayes	1.000	1.000	1.000	1.000	1.000

Pada tabel di atas, ditampilkan hasil akurasi dari metode Naive Bayes pada widget Test and Score, yang mencapai 100%. Hasil ini menunjukkan bahwa metode Naive Bayes sangat efektif dalam mengklasifikasikan minat siswa dan siswi terhadap jurusan Teknik Komputer Jaringan (TKJ), menjadikannya alat yang sangat baik untuk analisis ini.

Tabel 4. 5. Result of Confusion Matrix

		Predicted		
		Minat	Tidak Minat	∑
Actual	Minat	39	0	39
	Tidak Minat	3	8	11
	∑	42	8	50

Figure 12. The True Positive (TP) result is 39. True Negative (TN) is 8, False Positive (FP) is 0 and False Negative (FN) is 3. Then the values for accuracy, precision and recall are as follows:

$$\boldsymbol{Accuracy} = \frac{39+8}{39+8+0+3} + \mathbf{100\%}$$ Then the Accuracy value= 94%

$$\boldsymbol{Presisi} = \frac{39}{39+0} + \mathbf{100\%}$$ Then the Presisi value = 100%

$$\boldsymbol{Recall} = \frac{39}{39+3} + \mathbf{100\%}$$ Then the Recall value = 92%

Pada widget Confusion Matrix, hasil akurasi dari metode Naive Bayes dalam mengklasifikasikan minat siswa terhadap jurusan Teknik Komputer Jaringan (TKJ) mencapai 94%. Hasil ini menunjukkan bahwa metode Naive Bayes sangat efektif dalam menentukan minat siswa terhadap jurusan TKJ.