

BAB IV

HASIL DAN PEMBAHASAN

3.1. Pengumpulan Data

Pengumpulan data pada penelitian ini dilakukan dengan melakukan observasi secara langsung dan penyebaran kuesioner kepada pelanggan layanan Indihome, yang bertujuan untuk memperoleh informasi mengenai kepuasan mereka terhadap layanan yang diberikan. Data yang dikumpulkan meliputi beberapa variabel penting yaitu kekuatan sinyal, penggunaan layanan (individu atau usaha), harga berlangganan, jumlah kapasitas user, serta kategori kepuasan pelanggan (Puas atau Tidak Puas). Proses pengumpulan data ini dilakukan secara langsung kepada responden yang merupakan pelanggan aktif, sehingga data yang diperoleh bersifat primer dan merefleksikan kondisi aktual yang relevan untuk dianalisis menggunakan metode klasifikasi Naive Bayes dan k-Nearest Neighbor (kNN).

Tabel 4. 1. Data Sampel Penelitian

Responden	Kekuatan Sinyal	Penggunaan	Harga Berlangganan	Jumlah Kapasitas User	Kategori
R1	Normal	Usaha	Normal	Normal	Puas
R2	Normal	Individu	Normal	Normal	Puas
R3	Normal	Individu	Normal	Normal	Puas
R4	Normal	Usaha	Normal	Normal	Puas
R5	Normal	Usaha	Normal	Normal	Puas
R6	Normal	Individu	Normal	Normal	Puas
R7	Tinggi	Individu	Normal	Normal	Puas
R8	Normal	Individu	Normal	Normal	Puas
R9	Normal	Individu	Normal	Normal	Puas
R10	Tinggi	Individu	Normal	Tinggi	Puas
R11	Tinggi	Usaha	Normal	Normal	Puas

R12	Tinggi	Individu	Normal	Normal	Puas
R13	Normal	Individu	Normal	Normal	Puas
R14	Normal	Individu	Normal	Normal	Puas
R15	Normal	Individu	Normal	Normal	Tidak Puas
R16	Rendah	Usaha	Tinggi	Tinggi	Tidak Puas
R17	Rendah	Usaha	Tinggi	Tinggi	Tidak Puas
R18	Normal	Individu	Tinggi	Normal	Puas
R19	Tinggi	Individu	Normal	Tinggi	Puas
R20	Tinggi	Individu	Normal	Normal	Puas
R21	Normal	Individu	Normal	Rendah	Puas
R22	Rendah	Usaha	Normal	Normal	Puas
R23	Normal	Individu	Normal	Normal	Puas
R24	Tinggi	Individu	Tinggi	Tinggi	Puas
R25	Normal	Individu	Normal	Normal	Puas
R26	Tinggi	Individu	Normal	Normal	Puas
R27	Normal	Usaha	Normal	Normal	Puas
R28	Rendah	Individu	Tinggi	Normal	Tidak Puas
R29	Normal	Usaha	Normal	Normal	Puas
R30	Tinggi	Individu	Tinggi	Normal	Puas
R31	Normal	Usaha	Tinggi	Normal	Puas
R32	Rendah	Usaha	Normal	Tinggi	Tidak Puas
R33	Normal	Individu	Normal	Normal	Puas
R34	Tinggi	Individu	Normal	Normal	Puas
R35	Rendah	Individu	Tinggi	Normal	Tidak Puas
R36	Normal	Individu	Tinggi	Normal	Tidak Puas
R37	Normal	Individu	Normal	Normal	Puas
R38	Tinggi	Usaha	Tinggi	Normal	Puas
R39	Normal	Individu	Normal	Tinggi	Puas
R40	Normal	Usaha	Normal	Tinggi	Puas
R41	Rendah	Usaha	Normal	Normal	Tidak Puas
R42	Normal	Individu	Normal	Normal	Puas
R43	Normal	Usaha	Tinggi	Normal	Puas
R44	Normal	Individu	Normal	Normal	Puas
R45	Rendah	Individu	Normal	Normal	Tidak Puas
R46	Normal	Individu	Tinggi	Normal	Tidak Puas
R47	Normal	Individu	Normal	Tinggi	Puas
R48	Tinggi	Individu	Normal	Normal	Puas
R49	Normal	Individu	Normal	Rendah	Tidak Puas
R50	Normal	Individu	Normal	Normal	Puas
R51	Normal	Usaha	Normal	Normal	Puas
R52	Normal	Individu	Tinggi	Normal	Tidak Puas

R53	Normal	Individu	Tinggi	Normal	Puas
R54	Normal	Usaha	Normal	Tinggi	Puas
R55	Normal	Individu	Normal	Rendah	Puas
R56	Rendah	Individu	Tinggi	Normal	Tidak Puas
R57	Normal	Individu	Rendah	Normal	Puas
R58	Normal	Individu	Normal	Rendah	Puas
R59	Tinggi	Individu	Normal	Normal	Puas
R60	Normal	Individu	Tinggi	Rendah	Tidak Puas
R61	Normal	Individu	Normal	Normal	Puas
R62	Normal	Individu	Tinggi	Rendah	Tidak Puas
R63	Normal	Individu	Normal	Normal	Puas
R64	Tinggi	Individu	Normal	Normal	Puas
R65	Normal	Individu	Tinggi	Normal	Tidak Puas
R66	Rendah	Usaha	Normal	Normal	Tidak Puas
R67	Normal	Individu	Normal	Tinggi	Puas
R68	Normal	Individu	Normal	Normal	Puas
R69	Tinggi	Individu	Normal	Normal	Puas
R70	Normal	Individu	Tinggi	Normal	Puas
R71	Normal	Individu	Normal	Rendah	Puas
R72	Normal	Individu	Normal	Rendah	Puas
R73	Tinggi	Individu	Normal	Rendah	Puas
R74	Rendah	Individu	Tinggi	Normal	Tidak Puas
R75	Normal	Usaha	Tinggi	Normal	Tidak Puas
R76	Normal	Individu	Tinggi	Normal	Tidak Puas
R77	Normal	Individu	Normal	Tinggi	Puas
R78	Rendah	Usaha	Normal	Normal	Tidak Puas
R79	Normal	Individu	Rendah	Normal	Puas
R80	Normal	Individu	Normal	Tinggi	Puas
R81	Normal	Individu	Rendah	Rendah	Puas
R82	Normal	Usaha	Normal	Normal	Puas
R83	Tinggi	Individu	Normal	Tinggi	Puas
R84	Tinggi	Individu	Tinggi	Normal	Puas
R85	Rendah	Individu	Normal	Normal	Tidak Puas
R86	Tinggi	Individu	Tinggi	Tinggi	Puas
R87	Normal	Usaha	Normal	Normal	Puas
R88	Rendah	Individu	Normal	Tinggi	Tidak Puas
R89	Normal	Individu	Tinggi	Normal	Tidak Puas
R90	Normal	Usaha	Normal	Normal	Puas
R91	Tinggi	Individu	Tinggi	Normal	Puas
R92	Normal	Individu	Normal	Normal	Puas
R93	Rendah	Individu	Tinggi	Normal	Tidak Puas

R94	Normal	Individu	Normal	Normal	Puas
R95	Tinggi	Individu	Tinggi	Normal	Puas
R96	Normal	Individu	Normal	Normal	Puas
R97	Normal	Usaha	Tinggi	Normal	Puas
R98	Normal	Individu	Rendah	Normal	Puas
R99	Rendah	Individu	Normal	Normal	Tidak Puas
R100	Normal	Individu	Normal	Normal	Puas
R101	Normal	Usaha	Tinggi	Rendah	Puas
R102	Tinggi	Individu	Tinggi	Normal	Puas
R103	Normal	Individu	Tinggi	Normal	Tidak Puas
R104	Normal	Individu	Tinggi	Normal	Tidak Puas
R105	Normal	Usaha	Normal	Normal	Puas
R106	Tinggi	Individu	Normal	Normal	Puas
R107	Normal	Individu	Normal	Normal	Puas
R108	Normal	Usaha	Tinggi	Tinggi	Puas
R109	Normal	Usaha	Normal	Normal	Puas
R110	Normal	Individu	Tinggi	Normal	Puas
R111	Normal	Individu	Normal	Normal	Puas
R112	Tinggi	Individu	Normal	Tinggi	Puas
R113	Normal	Individu	Tinggi	Normal	Puas
R114	Normal	Individu	Normal	Normal	Puas
R115	Tinggi	Individu	Normal	Tinggi	Puas
R116	Tinggi	Usaha	Tinggi	Normal	Puas
R117	Tinggi	Individu	Normal	Normal	Puas
R118	Normal	Individu	Normal	Normal	Puas
R119	Normal	Individu	Normal	Normal	Puas
R120	Normal	Individu	Normal	Normal	Puas
R121	Normal	Usaha	Tinggi	Tinggi	Puas
R122	Rendah	Usaha	Tinggi	Tinggi	Puas
R123	Normal	Individu	Tinggi	Normal	Puas
R124	Tinggi	Individu	Normal	Tinggi	Puas
R125	Tinggi	Individu	Normal	Normal	Puas
R126	Normal	Individu	Normal	Rendah	Puas
R127	Normal	Usaha	Normal	Normal	Puas
R128	Normal	Usaha	Normal	Normal	Tidak Puas
R129	Tinggi	Individu	Tinggi	Tinggi	Puas
R130	Normal	Individu	Normal	Normal	Puas
R131	Tinggi	Individu	Normal	Normal	Puas
R132	Normal	Usaha	Normal	Normal	Puas
R133	Rendah	Individu	Tinggi	Normal	Puas
R134	Normal	Usaha	Normal	Tinggi	Puas

R135	Tinggi	Individu	Tinggi	Normal	Puas
R136	Normal	Usaha	Tinggi	Normal	Puas
R137	Rendah	Usaha	Normal	Tinggi	Tidak Puas
R138	Normal	Individu	Normal	Normal	Puas
R139	Tinggi	Individu	Normal	Normal	Puas
R140	Rendah	Individu	Tinggi	Normal	Tidak Puas
R141	Normal	Individu	Tinggi	Normal	Tidak Puas
R142	Normal	Individu	Normal	Normal	Puas
R143	Tinggi	Usaha	Tinggi	Normal	Puas
R144	Normal	Individu	Normal	Tinggi	Puas
R145	Normal	Usaha	Normal	Tinggi	Puas
R146	Rendah	Usaha	Normal	Normal	Tidak Puas
R147	Normal	Individu	Normal	Normal	Puas
R148	Normal	Usaha	Tinggi	Normal	Puas
R149	Normal	Individu	Normal	Normal	Puas
R150	Normal	Individu	Normal	Normal	Puas

Pada tabel di atas merupakan hasil pengumpulan data dari 150 responden pelanggan Indihome yang mencakup empat variabel utama, yaitu kekuatan sinyal, jenis penggunaan layanan (individu atau usaha), harga berlangganan, dan jumlah kapasitas user. Setiap responden diidentifikasi dengan kode unik (R1 hingga R150) dan memberikan informasi kategorikal untuk masing-masing variabel. Data ini menggambarkan kondisi layanan yang dirasakan oleh pelanggan, seperti kekuatan sinyal yang diklasifikasikan menjadi rendah, normal, dan tinggi, Jenis penggunaan yang terbagi antara individu dan usaha, Serta harga dan kapasitas yang juga dikelompokkan ke dalam beberapa tingkatan. Informasi tersebut menjadi dasar dalam proses klasifikasi tingkat kepuasan pelanggan menggunakan metode machine learning, yaitu Naive Bayes dan k-Nearest Neighbor.

3.2. Pembagian Data

Pada penelitian ini, data yang telah dikumpulkan sebanyak 150 data responden dibagi menjadi dua bagian, yaitu data training dan data testing. Pembagian ini dilakukan untuk memfasilitasi proses pelatihan dan pengujian model klasifikasi yang digunakan, yaitu algoritma Naive Bayes dan k-Nearest Neighbor. Data training digunakan sebagai dasar untuk membentuk model prediksi dengan mempelajari pola dari variabel-variabel yang ada, seperti kekuatan sinyal, penggunaan, harga berlangganan, dan jumlah kapasitas user. Sedangkan data testing digunakan untuk mengukur sejauh mana model yang telah dilatih mampu memprediksi tingkat kepuasan pelanggan secara akurat pada data baru yang tidak termasuk dalam proses pelatihan. Pembagian ini menjadi langkah penting dalam memastikan model yang dibangun memiliki kinerja yang baik dan mampu melakukan generalisasi dengan tepat.

Tabel 4. 2. Data Training

Responden	Kekuatan Sinyal	Penggunaan	Harga Berlangganan	Jumlah Kapasitas User	Kategori
R1	Normal	Usaha	Normal	Normal	Puas
R2	Normal	Individu	Normal	Normal	Puas
R3	Normal	Individu	Normal	Normal	Puas
R4	Normal	Usaha	Normal	Normal	Puas
R5	Normal	Usaha	Normal	Normal	Puas
R6	Normal	Individu	Normal	Normal	Puas
R7	Tinggi	Individu	Normal	Normal	Puas
R8	Normal	Individu	Normal	Normal	Puas
R9	Normal	Individu	Normal	Normal	Puas
R10	Tinggi	Individu	Normal	Tinggi	Puas
R11	Tinggi	Usaha	Normal	Normal	Puas
R12	Tinggi	Individu	Normal	Normal	Puas
R13	Normal	Individu	Normal	Normal	Puas
R14	Normal	Individu	Normal	Normal	Puas
R15	Normal	Individu	Normal	Normal	Tidak Puas
R16	Rendah	Usaha	Tinggi	Tinggi	Tidak Puas

R17	Rendah	Usaha	Tinggi	Tinggi	Tidak Puas
R18	Normal	Individu	Tinggi	Normal	Puas
R19	Tinggi	Individu	Normal	Tinggi	Puas
R20	Tinggi	Individu	Normal	Normal	Puas
R21	Normal	Individu	Normal	Rendah	Puas
R22	Rendah	Usaha	Normal	Normal	Puas
R23	Normal	Individu	Normal	Normal	Puas
R24	Tinggi	Individu	Tinggi	Tinggi	Puas
R25	Normal	Individu	Normal	Normal	Puas
R26	Tinggi	Individu	Normal	Normal	Puas
R27	Normal	Usaha	Normal	Normal	Puas
R28	Rendah	Individu	Tinggi	Normal	Tidak Puas
R29	Normal	Usaha	Normal	Normal	Puas
R30	Tinggi	Individu	Tinggi	Normal	Puas
R31	Normal	Usaha	Tinggi	Normal	Puas
R32	Rendah	Usaha	Normal	Tinggi	Tidak Puas
R33	Normal	Individu	Normal	Normal	Puas
R34	Tinggi	Individu	Normal	Normal	Puas
R35	Rendah	Individu	Tinggi	Normal	Tidak Puas
R36	Normal	Individu	Tinggi	Normal	Tidak Puas
R37	Normal	Individu	Normal	Normal	Puas
R38	Tinggi	Usaha	Tinggi	Normal	Puas
R39	Normal	Individu	Normal	Tinggi	Puas
R40	Normal	Usaha	Normal	Tinggi	Puas
R41	Rendah	Usaha	Normal	Normal	Tidak Puas
R42	Normal	Individu	Normal	Normal	Puas
R43	Normal	Usaha	Tinggi	Normal	Puas
R44	Normal	Individu	Normal	Normal	Puas
R45	Rendah	Individu	Normal	Normal	Tidak Puas
R46	Normal	Individu	Tinggi	Normal	Tidak Puas
R47	Normal	Individu	Normal	Tinggi	Puas
R48	Tinggi	Individu	Normal	Normal	Puas
R49	Normal	Individu	Normal	Rendah	Tidak Puas
R50	Normal	Individu	Normal	Normal	Puas
R51	Normal	Usaha	Normal	Normal	Puas
R52	Normal	Individu	Tinggi	Normal	Tidak Puas
R53	Normal	Individu	Tinggi	Normal	Puas
R54	Normal	Usaha	Normal	Tinggi	Puas
R55	Normal	Individu	Normal	Rendah	Puas
R56	Rendah	Individu	Tinggi	Normal	Tidak Puas
R57	Normal	Individu	Rendah	Normal	Puas

R58	Normal	Individu	Normal	Rendah	Puas
R59	Tinggi	Individu	Normal	Normal	Puas
R60	Normal	Individu	Tinggi	Rendah	Tidak Puas
R61	Normal	Individu	Normal	Normal	Puas
R62	Normal	Individu	Tinggi	Rendah	Tidak Puas
R63	Normal	Individu	Normal	Normal	Puas
R64	Tinggi	Individu	Normal	Normal	Puas
R65	Normal	Individu	Tinggi	Normal	Tidak Puas
R66	Rendah	Usaha	Normal	Normal	Tidak Puas
R67	Normal	Individu	Normal	Tinggi	Puas
R68	Normal	Individu	Normal	Normal	Puas
R69	Tinggi	Individu	Normal	Normal	Puas
R70	Normal	Individu	Tinggi	Normal	Puas
R71	Normal	Individu	Normal	Rendah	Puas
R72	Normal	Individu	Normal	Rendah	Puas
R73	Tinggi	Individu	Normal	Rendah	Puas
R74	Rendah	Individu	Tinggi	Normal	Tidak Puas
R75	Normal	Usaha	Tinggi	Normal	Tidak Puas
R76	Normal	Individu	Tinggi	Normal	Tidak Puas
R77	Normal	Individu	Normal	Tinggi	Puas
R78	Rendah	Usaha	Normal	Normal	Tidak Puas
R79	Normal	Individu	Rendah	Normal	Puas
R80	Normal	Individu	Normal	Tinggi	Puas
R81	Normal	Individu	Rendah	Rendah	Puas
R82	Normal	Usaha	Normal	Normal	Puas
R83	Tinggi	Individu	Normal	Tinggi	Puas
R84	Tinggi	Individu	Tinggi	Normal	Puas
R85	Rendah	Individu	Normal	Normal	Tidak Puas
R86	Tinggi	Individu	Tinggi	Tinggi	Puas
R87	Normal	Usaha	Normal	Normal	Puas
R88	Rendah	Individu	Normal	Tinggi	Tidak Puas
R89	Normal	Individu	Tinggi	Normal	Tidak Puas
R90	Normal	Usaha	Normal	Normal	Puas
R91	Tinggi	Individu	Tinggi	Normal	Puas
R92	Normal	Individu	Normal	Normal	Puas
R93	Rendah	Individu	Tinggi	Normal	Tidak Puas
R94	Normal	Individu	Normal	Normal	Puas
R95	Tinggi	Individu	Tinggi	Normal	Puas
R96	Normal	Individu	Normal	Normal	Puas
R97	Normal	Usaha	Tinggi	Normal	Puas
R98	Normal	Individu	Rendah	Normal	Puas

R99	Rendah	Individu	Normal	Normal	Tidak Puas
R100	Normal	Individu	Normal	Normal	Puas
R101	Normal	Usaha	Tinggi	Rendah	Puas
R102	Tinggi	Individu	Tinggi	Normal	Puas
R103	Normal	Individu	Tinggi	Normal	Tidak Puas
R104	Normal	Individu	Tinggi	Normal	Tidak Puas
R105	Normal	Usaha	Normal	Normal	Puas

Pada tabel di atas merupakan data training yang digunakan dalam penelitian ini dengan jumlah sebanyak 105 data responden. Data training ini memuat lima atribut, yaitu kekuatan sinyal, penggunaan, harga berlangganan, jumlah kapasitas user, dan kategori kepuasan pelanggan sebagai label kelas. Data tersebut digunakan untuk membangun dan melatih model klasifikasi menggunakan algoritma Naive Bayes dan k-Nearest Neighbor. Dengan menganalisis pola dari variabel-variabel yang tersedia, model diharapkan mampu mengidentifikasi dan memprediksi tingkat kepuasan pelanggan secara akurat berdasarkan karakteristik layanan yang diterima oleh penggunaan.

Tabel 4. 3. Data Testing

Responden	Kekuatan Sinyal	Penggunaan	Harga Berlangganan	Jumlah Kapasitas User	Kategori
R1	Tinggi	Individu	Normal	Normal	Puas
R2	Normal	Individu	Normal	Normal	Puas
R3	Normal	Usaha	Tinggi	Tinggi	Puas
R4	Normal	Usaha	Normal	Normal	Puas
R5	Normal	Individu	Tinggi	Normal	Puas
R6	Normal	Individu	Normal	Normal	Puas
R7	Tinggi	Individu	Normal	Tinggi	Puas
R8	Normal	Individu	Tinggi	Normal	Puas
R9	Normal	Individu	Normal	Normal	Puas
R10	Tinggi	Individu	Normal	Tinggi	Puas
R11	Tinggi	Usaha	Tinggi	Normal	Puas
R12	Tinggi	Individu	Normal	Normal	Puas

R13	Normal	Individu	Normal	Normal	Puas
R14	Normal	Individu	Normal	Normal	Puas
R15	Normal	Individu	Normal	Normal	Puas
R16	Normal	Usaha	Tinggi	Tinggi	Puas
R17	Rendah	Usaha	Tinggi	Tinggi	Puas
R18	Normal	Individu	Tinggi	Normal	Puas
R19	Tinggi	Individu	Normal	Tinggi	Puas
R20	Tinggi	Individu	Normal	Normal	Puas
R21	Normal	Individu	Normal	Rendah	Puas
R22	Normal	Usaha	Normal	Normal	Puas
R23	Normal	Usaha	Normal	Normal	Tidak Puas
R24	Tinggi	Individu	Tinggi	Tinggi	Puas
R25	Normal	Individu	Normal	Normal	Puas
R26	Tinggi	Individu	Normal	Normal	Puas
R27	Normal	Usaha	Normal	Normal	Puas
R28	Rendah	Individu	Tinggi	Normal	Puas
R29	Normal	Usaha	Normal	Tinggi	Puas
R30	Tinggi	Individu	Tinggi	Normal	Puas
R31	Normal	Usaha	Tinggi	Normal	Puas
R32	Rendah	Usaha	Normal	Tinggi	Tidak Puas
R33	Normal	Individu	Normal	Normal	Puas
R34	Tinggi	Individu	Normal	Normal	Puas
R35	Rendah	Individu	Tinggi	Normal	Tidak Puas
R36	Normal	Individu	Tinggi	Normal	Tidak Puas
R37	Normal	Individu	Normal	Normal	Puas
R38	Tinggi	Usaha	Tinggi	Normal	Puas
R39	Normal	Individu	Normal	Tinggi	Puas
R40	Normal	Usaha	Normal	Tinggi	Puas
R41	Rendah	Usaha	Normal	Normal	Tidak Puas
R42	Normal	Individu	Normal	Normal	Puas
R43	Normal	Usaha	Tinggi	Normal	Puas
R44	Normal	Individu	Normal	Normal	Puas
R45	Normal	Individu	Normal	Normal	Puas

Pada tabel di atas merupakan data testing yang digunakan dalam penelitian ini, yaitu data sampel penelitian dengan jumlah sebanyak 45 data. Data ini terdiri dari empat atribut utama, yaitu kekuatan sinyal, penggunaan, harga berlangganan, dan jumlah kapasitas user. Data testing ini berfungsi untuk menguji performa dan

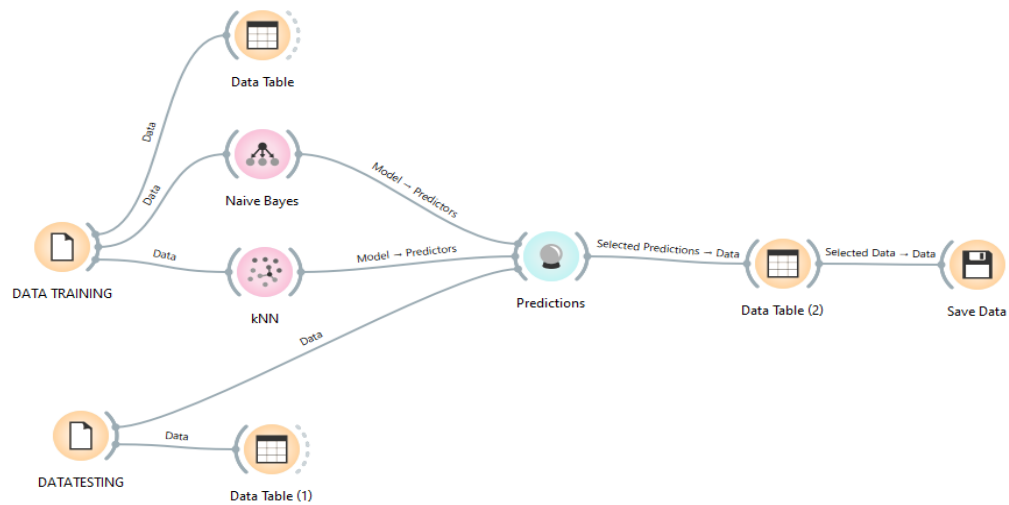
akurasi model klasifikasi yang telah dibangun sebelumnya menggunakan data training. Melalui data testing ini, dapat diketahui sejauh mana model mampu mengklasifikasikan data baru secara tepat berdasarkan pola yang telah dipelajari dari data pelatihan.

3.3. Pembersihan Data

Pembersihan data merupakan proses penting dalam tahap pra-pemrosesan data yang bertujuan untuk menghilangkan atau memperbaiki data yang tidak lengkap, tidak konsisten, atau mengandung kesalahan. Tahapan ini mencakup identifikasi dan penanganan terhadap nilai kosong (missing values), duplikasi data, kesalahan penulisan, serta penyimpangan data yang tidak relevan. Pembersihan data sangat diperlukan agar hasil analisis atau model yang dibangun menjadi lebih akurat dan dapat diandalkan. Dalam penelitian ini, proses pembersihan data telah dilakukan dengan seksama, dan hasilnya menunjukkan bahwa data yang diperoleh sudah dalam kondisi baik, konsisten, serta layak untuk digunakan dalam proses klasifikasi menggunakan algoritma yang telah ditentukan.

3.4. Perancangan Model Klasifikasi

Perancangan model klasifikasi dilakukan menggunakan aplikasi Orange dengan mengimpor data bersih ke dalam canvas, kemudian dilakukan proses pengkodean data kategorikal dan pembagian data menjadi data latih dan data uji. Algoritma Naïve Bayes dan k-Nearest Neighbor (kNN) digunakan sebagai model klasifikasi, lalu dievaluasi menggunakan *Test & Score* untuk mengukur akurasi, presisi, recall, dan F1-score. Perancangan ini memastikan model dapat mengklasifikasikan tingkat kepuasan pelanggan secara optimal.



Gambar 4. 1. Perancangan Model Klasifikasi

Gambar di atas menunjukkan alur perancangan model klasifikasi menggunakan aplikasi Orange. Data dibagi menjadi dua bagian, yaitu Data Training dan Data Testing. Data training dimasukkan ke dalam dua algoritma klasifikasi, yaitu Naive Bayes dan k-Nearest Neighbor (kNN), untuk membentuk model prediksi. Kedua model tersebut kemudian dikoneksikan ke widget Predictions, yang juga menerima input dari Data Testing. Output dari Predictions menunjukkan hasil klasifikasi yang kemudian ditampilkan dalam Data Table (2) dan disimpan menggunakan widget Save Data. Proses ini bertujuan untuk membandingkan hasil klasifikasi dari kedua algoritma berdasarkan data uji yang telah disiapkan

3.5. Hasil Klasifikasi

Data hasil klasifikasi dalam penelitian ini akan disajikan dalam bentuk tabel untuk mempermudah analisis dan interpretasi. Tabel tersebut menampilkan prediksi kategori dari masing-masing data uji berdasarkan model yang digunakan.

Tabel 4. 4. Data Hasil Klasifikasi

Responden	Kekuatan Sinyal	Penggunaan	Harga Berlangganan	Jumlah Kapasitas User	Naïve Bayes	KNN
R1	Tinggi	Individu	Normal	Normal	Puas	Puas
R2	Normal	Individu	Normal	Normal	Puas	Puas
R3	Normal	Usaha	Tinggi	Tinggi	Puas	Tidak Puas
R4	Normal	Usaha	Normal	Normal	Puas	Puas
R5	Normal	Individu	Tinggi	Normal	Puas	Puas
R6	Normal	Individu	Normal	Normal	Puas	Puas
R7	Tinggi	Individu	Normal	Tinggi	Puas	Puas
R8	Normal	Individu	Tinggi	Normal	Puas	Puas
R9	Normal	Individu	Normal	Normal	Puas	Puas
R10	Tinggi	Individu	Normal	Tinggi	Puas	Puas
R11	Tinggi	Usaha	Tinggi	Normal	Puas	Puas
R12	Tinggi	Individu	Normal	Normal	Puas	Puas
R13	Normal	Individu	Normal	Normal	Puas	Puas
R14	Normal	Individu	Normal	Normal	Puas	Puas
R15	Normal	Individu	Normal	Normal	Puas	Puas
R16	Normal	Usaha	Tinggi	Tinggi	Puas	Tidak Puas
R17	Rendah	Usaha	Tinggi	Tinggi	Tidak Puas	Tidak Puas
R18	Normal	Individu	Tinggi	Normal	Puas	Puas
R19	Tinggi	Individu	Normal	Tinggi	Puas	Puas
R20	Tinggi	Individu	Normal	Normal	Puas	Puas
R21	Normal	Individu	Normal	Rendah	Puas	Puas
R22	Normal	Usaha	Normal	Normal	Puas	Puas
R23	Normal	Usaha	Normal	Normal	Puas	Puas
R24	Tinggi	Individu	Tinggi	Tinggi	Puas	Puas
R25	Normal	Individu	Normal	Normal	Puas	Puas
R26	Tinggi	Individu	Normal	Normal	Puas	Puas
R27	Normal	Usaha	Normal	Normal	Puas	Puas
R28	Rendah	Individu	Tinggi	Normal	Tidak Puas	Tidak Puas
R29	Normal	Usaha	Normal	Tinggi	Puas	Puas
R30	Tinggi	Individu	Tinggi	Normal	Puas	Puas
R31	Normal	Usaha	Tinggi	Normal	Puas	Puas

R32	Rendah	Usaha	Normal	Tinggi	Tidak Puas	Tidak Puas
R33	Normal	Individu	Normal	Normal	Puas	Puas
R34	Tinggi	Individu	Normal	Normal	Puas	Puas
R35	Rendah	Individu	Tinggi	Normal	Tidak Puas	Tidak Puas
R36	Normal	Individu	Tinggi	Normal	Puas	Puas
R37	Normal	Individu	Normal	Normal	Puas	Puas
R38	Tinggi	Usaha	Tinggi	Normal	Puas	Puas
R39	Normal	Individu	Normal	Tinggi	Puas	Puas
R40	Normal	Usaha	Normal	Tinggi	Puas	Puas
R41	Rendah	Usaha	Normal	Normal	Tidak Puas	Tidak Puas
R42	Normal	Individu	Normal	Normal	Puas	Puas
R43	Normal	Usaha	Tinggi	Normal	Puas	Puas
R44	Normal	Individu	Normal	Normal	Puas	Puas
R45	Normal	Individu	Normal	Normal	Puas	Puas

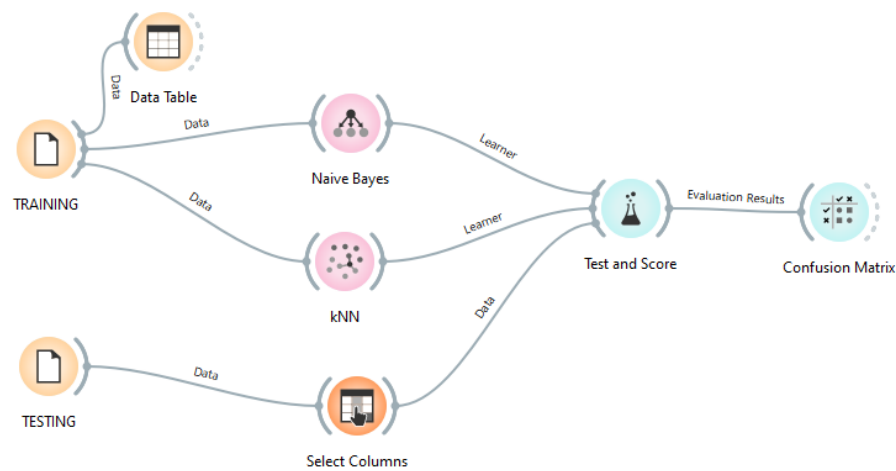
Tabel di atas menunjukkan hasil klasifikasi tingkat kepuasan pelanggan Indihome terhadap layanan yang diberikan, berdasarkan empat variabel utama yaitu kekuatan sinyal, jenis penggunaan (individu atau usaha), harga berlangganan, dan jumlah kapasitas user. Proses klasifikasi ini dilakukan menggunakan dua algoritma machine learning, yaitu Naïve Bayes dan k-Nearest Neighbor (kNN). Masing-masing responden (R1 hingga R45) dianalisis berdasarkan atribut yang dimilikinya, kemudian diprediksi tingkat kepuasannya oleh kedua metode. Hasil klasifikasi terbagi dalam dua kategori, yaitu Puas dan Tidak Puas, yang mencerminkan persepsi pelanggan terhadap kualitas layanan yang diterima.

Dari total 45 data responden yang diuji, metode Naïve Bayes mengklasifikasikan sebanyak 40 pelanggan ke dalam kategori Puas dan 5 pelanggan sebagai Tidak Puas. Sementara itu, metode kNN menghasilkan prediksi 38 pelanggan Puas dan 7 pelanggan Tidak Puas. Perbedaan jumlah prediksi ini

menunjukkan bahwa metode Naive Bayes lebih banyak memprediksi pelanggan sebagai puas dibandingkan KNN. Selanjutnya, perbandingan akurasi antara kedua metode akan menjadi bagian penting dalam mengevaluasi efektivitas model dalam mengklasifikasikan data baru secara tepat berdasarkan pola yang telah dipelajari dari data training.

3.6. Perancangan Model Evaluasi

Perancangan model evaluasi dalam penelitian ini dilakukan untuk mengukur kinerja dari algoritma klasifikasi yang digunakan, yaitu Naive Bayes dan k-Nearest Neighbor (KNN), dalam memprediksi tingkat kepuasan pelanggan Indihome. Tahapan evaluasi melibatkan pembagian data menjadi data latih dan data uji, serta penerapan confusion matrix sebagai dasar penilaian. Dari confusion matrix tersebut, dihitung metrik evaluasi seperti akurasi, presisi, recall, dan F1-score untuk masing-masing metode. Proses ini bertujuan untuk mengetahui sejauh mana masing-masing algoritma mampu mengklasifikasikan data dengan benar, serta membandingkan efektivitas kedua metode dalam mengolah data responden berdasarkan fitur-fitur yang tersedia, yaitu kekuatan sinyal, penggunaan, harga berlangganan, dan jumlah kapasitas user.



Gambar 4. 2. Perancangan Model Evaluasi

Gambar di atas menunjukkan alur perancangan model evaluasi yang dilakukan menggunakan perangkat lunak Orange. Proses dimulai dengan pembagian data menjadi dua bagian utama, yaitu Training dan Testing. Data Training dikirim ke dua jalur, pertama menuju Data table untuk melihat pratinjau data dan memastikan keseluruhan data, dan kedua langsung menuju algoritma pembelajaran Naive Bayes dan kNN, yang masing-masing berperan sebagai *learner* atau pembelajar dari data latih tersebut.

Sementara itu, data Testing dikirim ke widget Select Columns untuk memastikan bahwa hanya atribut-atribut relevan yang digunakan dalam pengujian. Hasil dari pemilihan kolom ini kemudian disalurkan ke widget Test and Score, bersama dengan dua model pembelajaran (Naive Bayes dan kNN) dari Training sebelumnya. Widget Test and Score berfungsi sebagai alat evaluasi yang menguji performa kedua model menggunakan data uji (Testing) berdasarkan metrik tertentu seperti Akurasi, Presisi, Recall dan f1-Score.

Hasil evaluasi tersebut diteruskan ke Confusion Matrix yang memperlihatkan jumlah prediksi benar dan salah untuk masing-masing kelas.

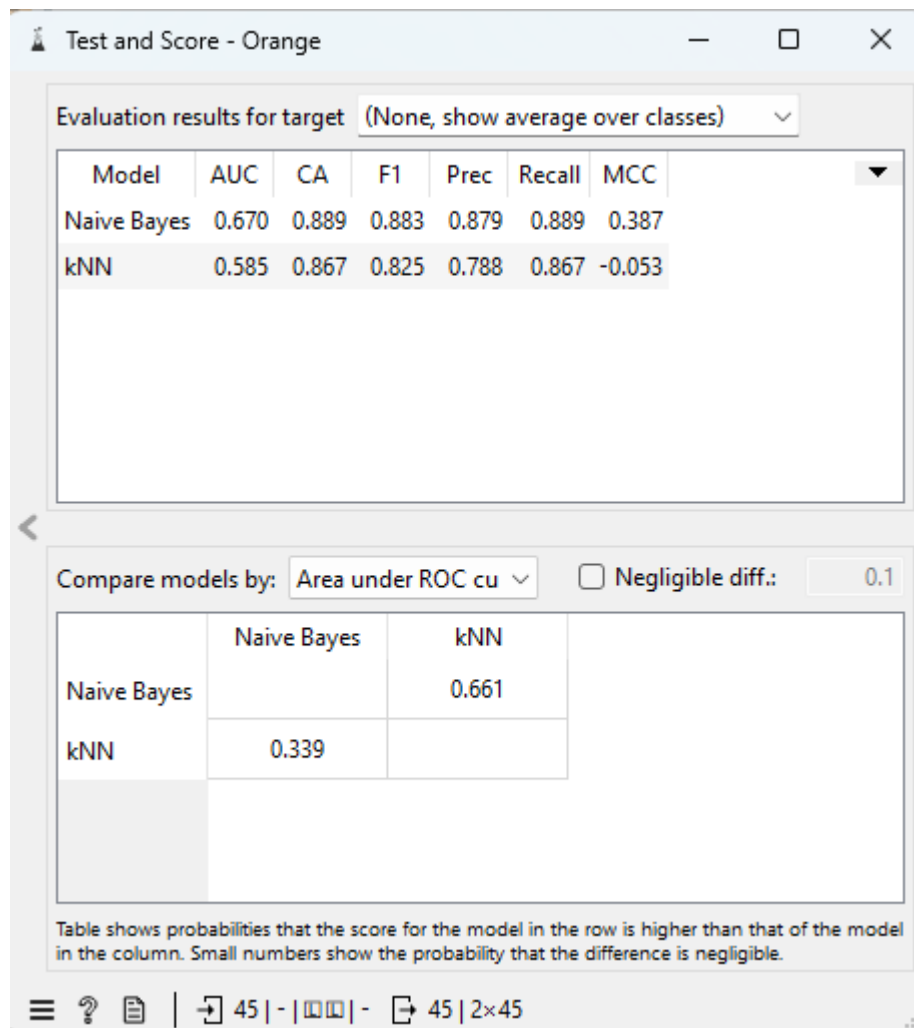
Sehingga memungkinkan peneliti untuk membandingkan efektivitas kedua algoritma dalam mengklasifikasikan tingkat kepuasan pelanggan.

3.7. Hasil Evaluasi

Hasil evaluasi dari proses perancangan model ditampilkan melalui dua widget, yaitu Test and Score dan Confusion Matrix. Kedua widget ini digunakan untuk melihat performa dari masing-masing algoritma dalam mengklasifikasikan data berdasarkan kategori kepuasan pelanggan. Evaluasi dilakukan secara otomatis oleh sistem setelah data diproses melalui model Naive Bayes dan k-Nearest Neighbor (kNN).

4.7.1. Hasil Test and Score

Test and Score adalah komponen evaluasi yang digunakan untuk menguji performa model dengan membandingkan hasil prediksi terhadap data sebenarnya. Komponen ini menghasilkan metrik evaluasi seperti akurasi, presisi, recall, dan F1-score untuk setiap model yang diuji.



Gambar 4. 3. Hasil Test and Score

Tabel 4 5. Hasil Perbandingan Metode Naïve Bayes dan kNN

Metrik Evaluasi	Penjelasan	Naive Bayes	k-NN
AUC (Area Under Curve)	Kemampuan membedakan antar kelas.	0.670	0.585
CA (Classification Accuracy)	Presentase prediksi yang benar terhadap total data.	0.889	0.867

F1 Score	Rata-rata harmonis antara Precision dan Recall	0.883	0.825
Precision	Proporsi prediksi positif yang benar	0.879	0.788
Recall	Proporsi actual positif yang berhasil diprediksi dengan benar	0.889	0.867
MCC (Matthews Corr. Coef.)	Indikator keseimbangan antar kelas. (Nilai negative menunjukkan ketidakseimbangan atau performa buruk)	0.387	-0.053

Berdasarkan hasil evaluasi performa model yang diperoleh dari perangkat lunak Orange Data Mining, diperlihatkan bahwa model Naive Bayes secara signifikan lebih unggul dibandingkan k-Nearest Neighbor (k-NN) dalam seluruh metrik evaluasi. Model Naive Bayes menunjukkan nilai AUC sebesar 0,670, sedangkan k-NN hanya mencapai 0,585. Ini menunjukkan bahwa Naive Bayes lebih mampu membedakan antara kelas target secara umum. Dalam hal akurasi (Classification Accuracy), Naive Bayes memperoleh skor 0,889, sedikit lebih tinggi dibandingkan k-NN yang mencetak 0,867.

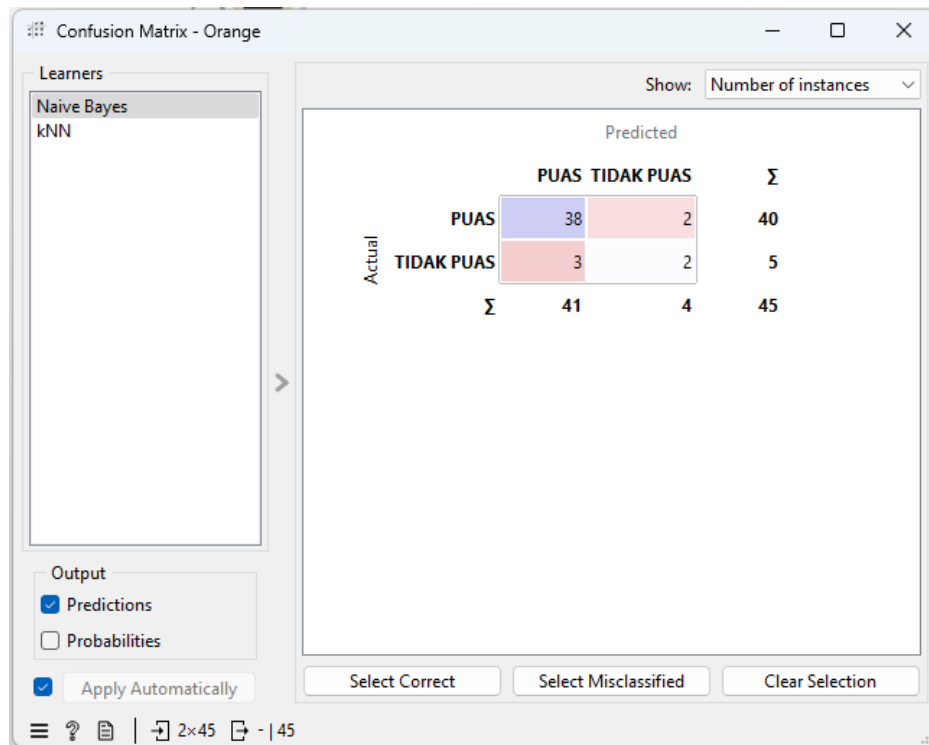
Kemudian F1 Score, yang menggabungkan nilai presisi dan recall dalam satu ukuran harmonis, juga lebih tinggi pada Naive Bayes (0,883) dibandingkan dengan k-NN (0,825). Ini menunjukkan bahwa Naive Bayes tidak hanya membuat prediksi yang lebih tepat (Precision: 0,879) tetapi juga lebih sensitif dalam mengenali data positif (Recall: 0,889). Sementara itu, nilai Matthews Correlation

Coefficient (MCC) pada Naive Bayes adalah positif (0,387), menandakan adanya korelasi yang kuat antara hasil prediksi dan kenyataan. Sebaliknya, MCC k-NN bernilai negatif (-0,053), yang merupakan indikasi bahwa prediksi model tersebut bahkan lebih buruk dari tebakan acak.

Perbandingan lebih lanjut berdasarkan Area Under ROC Curve menunjukkan bahwa Naive Bayes memiliki probabilitas sebesar 66,1% untuk menghasilkan skor yang lebih tinggi dibandingkan k-NN, menguatkan bukti bahwa model ini lebih handal.

4.7.2. Hasil Confusion Matrix

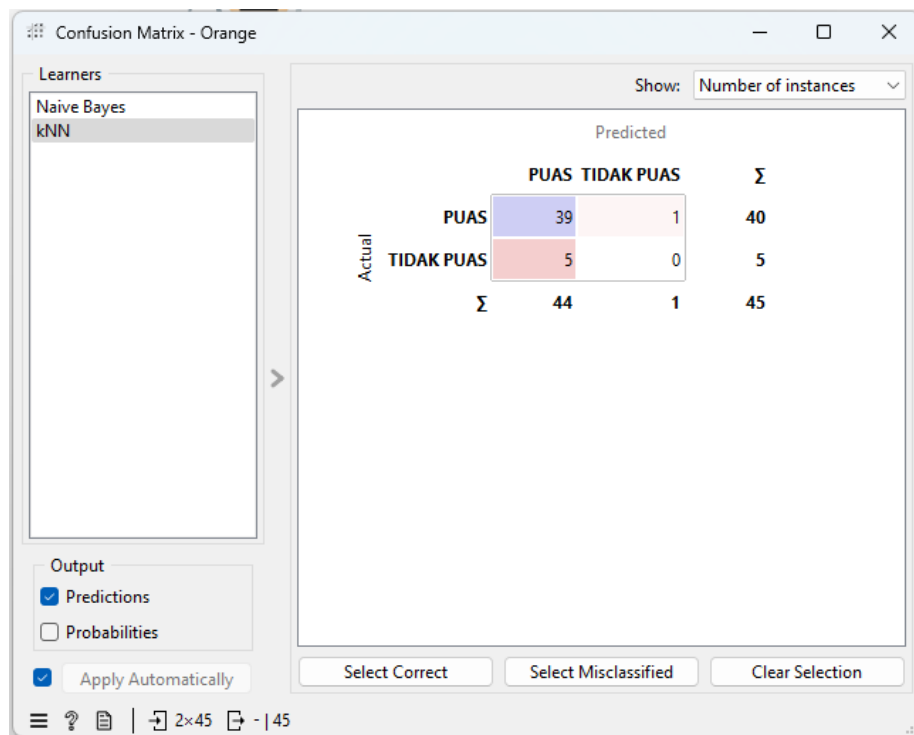
Confusion Matrix adalah alat evaluasi yang digunakan untuk melihat kinerja model klasifikasi dengan membandingkan hasil prediksi terhadap data aktual. Matriks ini menunjukkan jumlah prediksi yang benar dan salah untuk setiap kelas, sehingga memudahkan dalam menghitung metrik seperti akurasi, presisi, recall, dan F1-score.



Gambar 4. 4. Hasil Evaluasi Confusion Matrix Metode Naïve Bayes

Hasil evaluasi Confusion Matrix untuk metode Naïve Bayes diatas dapat dilihat bahwa terdapat 45 data yang diuji. Dari jumlah tersebut, sebanyak 40 data memiliki label puas , dan sisanya 5 data memiliki label Tidak Puas. Model berhasil mengklasifikasikan 38 data Puas dengan benar (true positives) dan salah mengklasifikasikan 2 data Puas sebagai Tidak Puas (false negatives). Sementara itu, dari 5 data yang sebenarnya Tidak Puas , 2 data diklasifikasikan dengan benar (true negatives) dan 3 data diklasifikasikan salah sebagai Puas (false positives).

Dengan demikian, total prediksi benar yang dilakukan model adalah 38 (Puas) + 2 (Tidak Puas) = 40 prediksi benar dari 45 total prediksi, yang menghasilkan akurasi sekitar 88.89%.



Gambar 4. 5. Hasil Evaluasi Confusion Matrix Metode kNN

Hasil Confusion Matrix pada metode k-Nearest Neighbor (kNN) menunjukkan bahwa dari 45 data pelanggan, Model berhasil mengklasifikasikan 39 data dengan benar (true positives), sementara 1 data Puas salah diklasifikasikan sebagai Tidak Puas (false negative). Sementara itu, seluruh 5 data yang sebenarnya Tidak Puas justru semuanya salah diklasifikasikan sebagai Puas(false positives), dan tidak ada satupun data Tidak Puas yang diprediksi dengan benar (true negatives = 0).

Dengan hasil tersebut, model *kNN* menghasilkan total prediksi yang benar sebanyak 39 (Puas) dari 45 data, yang berarti akurasi model adalah sekitar 86.67%. Meskipun akurasi terlihat cukup tinggi, model ini memiliki kelemahan signifikan dalam mengenali kelas Tidak Puas, yang terlihat dari seluruh 5 data

Tidak Puas salah diklasifikasikan. Hasil ini menunjukkan bahwa model kNN lebih sering menebak data sebagai (Puas), Sehingga tidak terlalu baik dalam mengenali data yang sebenarnya (Tidak Puas).