

BAB IV

HASIL DAN PEMBAHASAN

4.1. Hasil

Hasil penelitian ini selanjutnya akan diimplementasikan menggunakan aplikasi Orange Data Mining untuk melakukan proses klasifikasi data kepuasan pelanggan terhadap produk CS Float. Proses klasifikasi dilakukan dengan menerapkan dua metode Machine Learning, yaitu Naive Bayes dan K-Nearest Neighbor (KNN). Melalui aplikasi tersebut, data pelanggan akan diolah secara sistematis mulai dari tahap input data, pemodelan, hingga pengujian model untuk menghasilkan prediksi tingkat kepuasan pelanggan. Penggunaan Orange Data Mining bertujuan untuk mempermudah visualisasi proses klasifikasi serta menguji tingkat akurasi dari kedua metode yang digunakan, sehingga hasil penelitian dapat dianalisis secara lebih terstruktur dan objektif.

4.1.1. Pengumpulan Data

Pengumpulan data dalam penelitian ini dilakukan dengan cara menyebarkan kuesioner kepada pelanggan yang telah mengonsumsi minuman CS Float. Data yang dikumpulkan berupa penilaian pelanggan terhadap beberapa aspek, yaitu waktu respon pelayanan, kualitas layanan, kualitas produk, harga produk, dan kebersihan tempat. Setiap responden memberikan penilaian berdasarkan pengalaman mereka setelah membeli dan mengonsumsi CS Float, kemudian data tersebut direkap dan dijadikan sebagai dataset penelitian. Data yang telah terkumpul selanjutnya digunakan sebagai data training dan data testing untuk proses

klasifikasi tingkat kepuasan pelanggan menggunakan metode Naive Bayes dan K-Nearest Neighbor (KNN).

Tabel 4. 1. Data Kuesioner

Nama Lengkap	Waktu Respon					Kualitas Layanan					Kualitas Produk					Harga Produk					Kebersihan Tempat					
	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	
Bintang syukuriyah siregar	2	3	3	3	3	5	5	5	5	4	1	4	1	5	4	1	1	1	1	1	1	1	1	1	2	2
Nurul Aida Siregar	4	4	4	4	5	5	4	4	5	4	4	4	5	5	5	5	5	5	5	5	5	4	5	5	5	5
Irwansyah, S.T	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
Nur indah nasution	4	4	4	4	4	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
Aini Minta Ito Harahap	5	5	5	5	5	5	5	5	5	5	5	5	5	4	5	4	4	4	5	4	4	4	5	4	4	4
Eka Fitri r	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
Saima Putri Siregar	5	5	5	5	5	5	5	5	5	5	5	4	5	4	5	4	5	5	4	4	4	4	5	5	5	5
Diana syahfitri	4	5	5	5	5	4	4	4	5	4	5	5	5	5	5	5	5	5	4	4	4	4	4	4	4	5
dela yuuliyani	3	3	4	4	2	2	2	3	2	3	2	3	2	3	2	2	2	3	3	2	3	4	5	5	5	5
Amar ma'rub siregar	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
Nurhayati	5	4	5	5	5	5	5	5	5	5	5	4	5	5	5	4	5	5	5	5	5	5	4	5	5	5
putrireimadinirasti	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
Nadya Fitriani	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
Tria syahputri	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
Tuti Rama Yani	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
Grifaril	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
desiuliapuspitasari7@gmail.com	3	3	4	1	3	1	2	5	3	1	2	3	5	4	4	3	3	1	2	2	2	2	2	2	2	1
Zuhri Dio alfalah	3	2	4	3	2	5	2	4	3	2	3	4	4	3	3	2	4	5	4	3	3	4	3	2	4	4
Semangat abadi Telaumbanua	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
IKA MAYA SARI SIMAMORA	4	5	5	3	5	4	3	5	3	5	5	3	5	4	4	5	5	5	4	5	5	5	5	5	5	5

Data yang ditampilkan merupakan hasil pengumpulan kuesioner dari 82 responden yang telah mengonsumsi minuman CS Float. Kuesioner ini dirancang untuk mengukur tingkat kepuasan pelanggan berdasarkan lima variabel utama, yaitu waktu respon pelayanan, kualitas layanan, kualitas produk, harga produk, dan kebersihan tempat. Setiap variabel dinilai menggunakan skala numerik 1 sampai 5, di mana nilai yang lebih tinggi menunjukkan persepsi yang semakin baik terhadap aspek yang dinilai. Penggunaan skala ini bertujuan untuk mempermudah proses pengolahan data serta memungkinkan penerapan metode klasifikasi berbasis numerik.

Berdasarkan data yang terkumpul, terlihat bahwa sebagian besar responden memberikan penilaian pada rentang nilai 4 dan 5, khususnya pada variabel kualitas layanan, kualitas produk, dan kebersihan tempat. Hal ini menunjukkan bahwa secara umum pelanggan memiliki persepsi yang positif terhadap produk dan pelayanan CS Float. Namun, terdapat pula sejumlah responden yang memberikan

nilai rendah, terutama pada variabel waktu respon dan kualitas layanan, yang mengindikasikan adanya pengalaman kurang memuaskan dari sebagian pelanggan. Variasi nilai tersebut mencerminkan perbedaan persepsi dan pengalaman pelanggan yang menjadi dasar penting dalam proses analisis kepuasan.

Seluruh data kuesioner yang telah terkumpul kemudian dijadikan sebagai dataset penelitian untuk proses analisis lebih lanjut. Data ini berperan sebagai data training dan data testing dalam penerapan metode Naive Bayes dan K-Nearest Neighbor (KNN) guna mengklasifikasikan tingkat kepuasan pelanggan ke dalam kategori Puas dan Tidak Puas. Dengan jumlah data yang cukup dan variasi nilai yang beragam, dataset ini dinilai representatif untuk menggambarkan kondisi kepuasan pelanggan CS Float secara objektif dan mendukung proses pemodelan klasifikasi dalam penelitian ini.

4.1.2. Pembersihan Data

Pembersihan data pada penelitian ini dilakukan dengan menyeleksi data kuesioner yang layak dan tidak layak digunakan dalam proses analisis. Tahap ini bertujuan untuk memastikan bahwa data yang digunakan memiliki kelengkapan dan konsistensi sehingga dapat menghasilkan hasil klasifikasi yang akurat. Data yang tidak lengkap, memiliki nilai kosong, atau tidak sesuai dengan kriteria penelitian akan dieliminasi dari dataset. Setelah proses seleksi selesai, data numerik hasil kuesioner kemudian ditransformasikan ke dalam bentuk kategorikal, seperti mengelompokkan nilai skala penilaian menjadi kategori tertentu, sehingga data lebih mudah diproses menggunakan metode klasifikasi Naive Bayes dan K-Nearest Neighbor (KNN) dalam penelitian ini.

Tabel 4. 2. Data Sudah Dibersihkan

Nama Lengkap	Waktu Respon	Kualitas Layanan	Kualitas Produk	Harga Produk	Kebersihan Tempat
Abdu Rohim	Lambat	Kurang Bagus	Kurang Bagus	Malah	Kurang Bersih
Ade Wasti Novinta Harefa	Cepat	Bagus	Bagus	Murah	Bersih
Adelia Putri	Cepat	Bagus	Bagus	Murah	Bersih
Adema Sinta	Cepat	Bagus	Bagus	Murah	Bersih
Adinda Ayu Lestari Nasution	Cepat	Bagus	Bagus	Murah	Bersih
Adinda Sevira	Cepat	Bagus	Bagus	Murah	Bersih
Aini Minta Ito Harahap	Cepat	Bagus	Bagus	Murah	Bersih
Amar Ma'rub Siregar	Cepat	Bagus	Bagus	Murah	Bersih
Anggi Audya	Cepat	Bagus	Bagus	Murah	Kurang Bersih
Annisa	Lambat	Kurang Bagus	Kurang Bagus	Malah	Kurang Bersih
Arnes Dian Putri Harefa	Cepat	Bagus	Bagus	Murah	Bersih
Azhar Tambak	Cepat	Bagus	Bagus	Murah	Bersih
Bintang Syukuriyah Siregar	Lambat	Bagus	Bagus	Murah	Bersih
Bunga Sundari	Cepat	Bagus	Bagus	Murah	Bersih
Dea	Cepat	Bagus	Bagus	Murah	Bersih
Dela Yuuliyani	Cepat	Kurang Bagus	Kurang Bagus	Malah	Kurang Bersih
Dhea Allya Ginting	Cepat	Bagus	Bagus	Murah	Bersih
Diah Ayu	Lambat	Kurang Bagus	Kurang Bagus	Malah	Kurang Bersih
Dian Sahrani Siregar	Cepat	Bagus	Bagus	Murah	Bersih
Diana Syahfitri	Cepat	Bagus	Bagus	Murah	Bersih
Dinda Cantiks	Cepat	Bagus	Bagus	Murah	Bersih
Dinda Olivia Rambe	Lambat	Kurang Bagus	Kurang Bagus	Malah	Kurang Bersih
Divan Alfaruzi	Lambat	Kurang Bagus	Kurang Bagus	Malah	Kurang Bersih
Diyan	Cepat	Bagus	Bagus	Murah	Bersih
Eka Fitri R	Cepat	Bagus	Bagus	Murah	Bersih
Fitor Zega	Cepat	Bagus	Bagus	Murah	Bersih
Grifaril	Lambat	Kurang Bagus	Kurang Bagus	Malah	Kurang Bersih
Habi	Cepat	Bagus	Bagus	Murah	Bersih
Hani Aura Arcika	Lambat	Kurang Bagus	Kurang Bagus	Malah	Kurang Bersih
Husna	Cepat	Bagus	Bagus	Murah	Bersih
Ika Maya Sari Simamora	Cepat	Bagus	Bagus	Murah	Bersih
Indah Rahmadani	Cepat	Bagus	Bagus	Murah	Bersih
Irwansyah, S.T	Lambat	Kurang Bagus	Kurang Bagus	Malah	Kurang Bersih
Ivo Andre	Cepat	Bagus	Bagus	Murah	Bersih
Jalal	Lambat	Kurang Bagus	Kurang Bagus	Malah	Kurang Bersih
Jihan Salsabilla	Cepat	Bagus	Bagus	Murah	Bersih

Laila Sari	Cepat	Bagus	Bagus	Murah	Bersih
Lana Del Rey	Cepat	Bagus	Bagus	Murah	Bersih
Lenni	Lambat	Bagus	Bagus	Murah	Bersih
Lily Amelia	Lambat	Kurang Bagus	Kurang Bagus	Malah	Kurang Bersih
Madzy	Cepat	Bagus	Bagus	Murah	Bersih
Maryam Fadilah Sagala	Cepat	Bagus	Bagus	Murah	Bersih
Miah	Cepat	Bagus	Bagus	Murah	Bersih
Nadin Rambe	Cepat	Bagus	Bagus	Murah	Bersih
Nadya Fitriani	Cepat	Bagus	Bagus	Murah	Bersih
Nopita Sari Rambe	Lambat	Kurang Bagus	Kurang Bagus	Malah	Kurang Bersih
Novi Syahfitri	Lambat	Kurang Bagus	Kurang Bagus	Malah	Kurang Bersih
Nur Indah Nasution	Cepat	Bagus	Bagus	Murah	Bersih
Nurhayati	Cepat	Bagus	Bagus	Murah	Bersih
Nurul Aida Siregar	Cepat	Bagus	Bagus	Murah	Bersih
Nurul Aminah Fajri	Cepat	Bagus	Bagus	Murah	Bersih
Pandre	Cepat	Bagus	Bagus	Murah	Bersih
Priya W	Cepat	Bagus	Bagus	Murah	Bersih
Pujawati	Cepat	Bagus	Bagus	Murah	Bersih
Putri Annida Ramadhani	Cepat	Bagus	Bagus	Murah	Bersih
Putreimadimirasti	Cepat	Bagus	Bagus	Murah	Bersih
Rabiatul Adawiyah	Cepat	Bagus	Bagus	Murah	Bersih
Reza	Lambat	Bagus	Kurang Bagus	Malah	Bersih
Ririn	Cepat	Bagus	Bagus	Murah	Bersih
Ririn Syafrida	Cepat	Bagus	Bagus	Murah	Bersih
Rista Rahayu	Cepat	Bagus	Bagus	Murah	Bersih
Rizalunsani Rambe	Cepat	Bagus	Bagus	Murah	Bersih
Sahrul Ramadhan	Lambat	Kurang Bagus	Kurang Bagus	Malah	Kurang Bersih
Saima Putri Siregar	Cepat	Bagus	Bagus	Murah	Bersih
Sangaji Juanda	Lambat	Kurang Bagus	Kurang Bagus	Malah	Kurang Bersih
Selvy Agustina	Cepat	Bagus	Bagus	Murah	Bersih
Semangat Abadi Telaumbanua	Cepat	Bagus	Bagus	Murah	Bersih
Sheilla Azmi Br Pohan	Cepat	Bagus	Bagus	Murah	Bersih
Shole Hamid Hasibuan	Cepat	Bagus	Bagus	Murah	Bersih
Sindi Rambe	Cepat	Bagus	Bagus	Murah	Bersih
Solehuddin Siregar	Cepat	Bagus	Bagus	Murah	Bersih
Sri Oktaviani	Lambat	Kurang Bagus	Kurang Bagus	Malah	Kurang Bersih
Subhan Abadi	Cepat	Bagus	Bagus	Murah	Bersih
Tama	Cepat	Bagus	Bagus	Murah	Bersih
Tika Andini	Cepat	Bagus	Bagus	Murah	Bersih
Tria Syahputri	Cepat	Bagus	Bagus	Murah	Bersih
Tuti Rama Yani	Lambat	Kurang Bagus	Kurang Bagus	Malah	Kurang Bersih
Yuyun Lili Srikandy	Cepat	Bagus	Bagus	Murah	Bersih
Zaitunnisa Yutifa Tanjung	Cepat	Bagus	Bagus	Murah	Bersih
Zuhri Dio Alfalah	Lambat	Bagus	Bagus	Murah	Bersih

Tabel di atas merupakan data hasil proses pembersihan data (data cleaning) yang dilakukan sebelum tahap pengolahan dan analisis penelitian. Proses pembersihan data dilakukan dengan menyeleksi data responden yang layak digunakan berdasarkan kelengkapan dan kesesuaian pengisian kuesioner. Selain itu, data juga telah diurutkan berdasarkan abjad nama responden untuk memudahkan proses pengelolaan dan identifikasi data. Pada tahap awal, jumlah data yang diperoleh sebanyak 82 data responden, namun setelah dilakukan pemeriksaan ditemukan beberapa kesalahan pengisian sehingga data yang digunakan dalam penelitian ini menjadi 80 data responden.

Data yang dibuang bukan sepenuhnya tidak dapat digunakan, tetapi terdapat kesalahan seperti pengisian yang tidak lengkap, ketidaksesuaian format jawaban, atau terdapat inkonsistensi dalam pengisian variabel penelitian. Oleh karena itu, data tersebut dieliminasi agar tidak mempengaruhi hasil analisis. Pembersihan data menjadi tahap penting karena kualitas data sangat berpengaruh terhadap tingkat akurasi hasil klasifikasi yang akan dilakukan pada penelitian ini.

Pada tabel di atas juga ditampilkan beberapa variabel yang digunakan dalam penelitian, yaitu Waktu Respon, Kualitas Layanan, Kualitas Produk, Harga Produk, dan Kebersihan Tempat. Seluruh variabel telah disusun dalam bentuk kategorikal untuk mempermudah proses pengolahan data menggunakan metode klasifikasi. Data yang telah melalui tahap pembersihan ini kemudian digunakan sebagai dataset utama dalam proses analisis untuk menentukan tingkat kepuasan pelanggan berdasarkan kriteria yang telah ditentukan dalam penelitian.

4.1.3. Pembagian Data

Pembagian data pada penelitian ini dilakukan dengan membagi dataset yang telah dibersihkan menjadi dua bagian, yaitu data training dan data testing. Data training digunakan untuk melatih model klasifikasi agar dapat mengenali pola kepuasan pelanggan berdasarkan variabel yang digunakan, sedangkan data testing digunakan untuk menguji kinerja model dalam melakukan prediksi terhadap data yang belum pernah dipelajari sebelumnya. Pembagian data ini bertujuan untuk memastikan bahwa model yang dibangun mampu melakukan klasifikasi secara akurat dan dapat dievaluasi tingkat keefektifannya.

Tabel 4. 3. Data Training

Nama Lengkap	Waktu Respon	Kualiiitas Layanan	Kualitas Produk	Harga Produk	Kebersihan Tempat	Kategori
Adinda Sevira	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Aini Minta Ito Harahap	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Amar Ma'rub Siregar	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Bintang Syukuriyah Siregar	Lambat	Bagus	Bagus	Murah	Bersih	Puas
Dea	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Dela Yuuliyani	Cepat	Kurang Bagus	Kurang Bagus	Malah	Kurang Bersih	Tidak Puas
Diana Syahfitri	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Dinda Cantiks	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Dinda Olivia Rambe	Lambat	Kurang Bagus	Kurang Bagus	Malah	Kurang Bersih	Tidak Puas
Divan Alfaruzi	Lambat	Kurang Bagus	Kurang Bagus	Malah	Kurang Bersih	Tidak Puas
Eka Fitri R	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Grifaril	Lambat	Kurang Bagus	Kurang Bagus	Malah	Kurang Bersih	Tidak Puas
Hani Aura Arcika	Lambat	Kurang Bagus	Kurang Bagus	Malah	Kurang Bersih	Tidak Puas
Husna	Cepat	Bagus	Bagus	Murah	Bersih	Puas

Ika Maya Sari Simamora	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Indah Rahmadani	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Irwansyah, S.T	Lambat	Kurang Bagus	Kurang Bagus	Malah	Kurang Bersih	Tidak Puas
Madzy	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Nadya Fitriani	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Nur Indah Nasution	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Nurhayati	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Nurul Aida Siregar	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Nurul Aminah Fajri	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Putreimadinirasti	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Saima Putri Siregar	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Semangat Abadi Telaumbanua	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Sindi Rambe	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Tria Syahputri	Cepat	Bagus	Bagus	Murah	Bersih	Puas
Tuti Rama Yani	Lambat	Kurang Bagus	Kurang Bagus	Malah	Kurang Bersih	Tidak Puas
Zuhri Dio Alfalah	Lambat	Bagus	Bagus	Murah	Bersih	Tidak Puas

Tabel 4. 4. Data Testing

Nama Lengkap	Waktu Respon	Kualitas Layanan	Kualitas Produk	Harga Produk	Kebersihan Tempat
Abdu Rohim	Lambat	Kurang Bagus	Kurang Bagus	Malah	Kurang Bersih
Ade Wasti Novinta Harefa	Cepat	Bagus	Bagus	Murah	Bersih
Adelia Putri	Cepat	Bagus	Bagus	Murah	Bersih
Adema Sinta	Cepat	Bagus	Bagus	Murah	Bersih
Adinda Ayu Lestari Nasution	Cepat	Bagus	Bagus	Murah	Bersih
Anggi Audya	Cepat	Bagus	Bagus	Murah	Kurang Bersih
Annisa	Lambat	Kurang Bagus	Kurang Bagus	Malah	Kurang Bersih
Arnes Dian Putri Harefa	Cepat	Bagus	Bagus	Murah	Bersih
Azhar Tambak	Cepat	Bagus	Bagus	Murah	Bersih
Bunga Sundari	Cepat	Bagus	Bagus	Murah	Bersih
Dhea Allya Ginting	Cepat	Bagus	Bagus	Murah	Bersih
Diah Ayu	Lambat	Kurang Bagus	Kurang Bagus	Malah	Kurang Bersih
Dian Sahrani Siregar	Cepat	Bagus	Bagus	Murah	Bersih
Diyan	Cepat	Bagus	Bagus	Murah	Bersih
Fitor Zega	Cepat	Bagus	Bagus	Murah	Bersih

Habi	Cepat	Bagus	Bagus	Murah	Bersih
Ivo Andre	Cepat	Bagus	Bagus	Murah	Bersih
Jalal	Lambat	Kurang Bagus	Kurang Bagus	Malah	Kurang Bersih
Jihan Salsabilla	Cepat	Bagus	Bagus	Murah	Bersih
Laila Sari	Cepat	Bagus	Bagus	Murah	Bersih
Lana Del Rey	Cepat	Bagus	Bagus	Murah	Bersih
Lenni	Lambat	Bagus	Bagus	Murah	Bersih
Lily Amelia	Lambat	Kurang Bagus	Kurang Bagus	Malah	Kurang Bersih
Maryam Fadilah Sagala	Cepat	Bagus	Bagus	Murah	Bersih
Miah	Cepat	Bagus	Bagus	Murah	Bersih
Nadin Rambe	Cepat	Bagus	Bagus	Murah	Bersih
Nopita Sari Rambe	Lambat	Kurang Bagus	Kurang Bagus	Malah	Kurang Bersih
Novi Syahfitri	Lambat	Kurang Bagus	Kurang Bagus	Malah	Kurang Bersih
Pandre	Cepat	Bagus	Bagus	Murah	Bersih
Priya W	Cepat	Bagus	Bagus	Murah	Bersih
Pujawati	Cepat	Bagus	Bagus	Murah	Bersih
Putri Annida Ramadhani	Cepat	Bagus	Bagus	Murah	Bersih
Rabiatul Adawiyah	Cepat	Bagus	Bagus	Murah	Bersih
Reza	Lambat	Bagus	Kurang Bagus	Malah	Bersih
Ririn	Cepat	Bagus	Bagus	Murah	Bersih
Ririn Syafrida	Cepat	Bagus	Bagus	Murah	Bersih
Rista Rahayu	Cepat	Bagus	Bagus	Murah	Bersih
Rizalunsani Rambe	Cepat	Bagus	Bagus	Murah	Bersih
Sahrul Ramadhan	Lambat	Kurang Bagus	Kurang Bagus	Malah	Kurang Bersih
Sangaji Juanda	Lambat	Kurang Bagus	Kurang Bagus	Malah	Kurang Bersih
Selvy Agustina	Cepat	Bagus	Bagus	Murah	Bersih
Sheilla Azmi Br Pohan	Cepat	Bagus	Bagus	Murah	Bersih
Shole Hamid Hasibuan	Cepat	Bagus	Bagus	Murah	Bersih
Solehuddin Siregar	Cepat	Bagus	Bagus	Murah	Bersih
Sri Oktaviani	Lambat	Kurang Bagus	Kurang Bagus	Malah	Kurang Bersih
Subhan Abadi	Cepat	Bagus	Bagus	Murah	Bersih
Tama	Cepat	Bagus	Bagus	Murah	Bersih
Tika Andini	Cepat	Bagus	Bagus	Murah	Bersih
Yuyun Lili Srikandy	Cepat	Bagus	Bagus	Murah	Bersih
Zaitunnisa Yutifa Tanjung	Cepat	Bagus	Bagus	Murah	Bersih

Pada tabel di atas merupakan dataset penelitian yang telah dibagi menjadi dua bagian, yaitu data training dan data testing. Pembagian dataset ini dilakukan setelah proses pembersihan data selesai sehingga diperoleh sebanyak 80 data yang layak digunakan dalam penelitian. Dari total data tersebut, sebanyak 30 data digunakan sebagai data training dan 50 data digunakan sebagai data testing. Pembagian ini bertujuan untuk memastikan bahwa proses pembangunan model klasifikasi dapat dilakukan secara optimal serta dapat diuji tingkat keakuratannya.

Tabel 4.3 menunjukkan data training yang berisi data responden lengkap dengan atribut Waktu Respon, Kualitas Layanan, Kualitas Produk, Harga Produk, Kebersihan Tempat, serta kategori kepuasan pelanggan. Data training ini digunakan sebagai dasar pembelajaran model dalam metode klasifikasi yang digunakan, yaitu K-Nearest Neighbor (KNN). Pada tabel tersebut terlihat bahwa data training memiliki variasi kategori Puas dan Tidak Puas yang digunakan untuk membantu sistem mengenali pola hubungan antara variabel penilaian dengan tingkat kepuasan pelanggan.

Selanjutnya, Tabel 4.4 menunjukkan data testing yang berisi 50 data responden yang hanya memuat atribut penilaian tanpa kategori kepuasan. Data testing digunakan untuk menguji kemampuan model dalam melakukan prediksi kategori kepuasan pelanggan berdasarkan pola yang telah dipelajari dari data training. Data testing dipilih secara terpisah agar model dapat diuji menggunakan data yang belum pernah dipelajari sebelumnya, sehingga hasil klasifikasi dapat mencerminkan kemampuan model secara lebih objektif.

Pembagian dataset menjadi data training dan data testing ini sangat penting dalam proses penelitian karena dapat membantu dalam mengevaluasi performa metode K-Nearest Neighbor (KNN) yang digunakan. Dengan adanya pembagian data ini, model tidak hanya belajar dari data yang tersedia, tetapi juga diuji menggunakan data baru untuk mengetahui tingkat akurasi dan keandalan model dalam mengklasifikasikan tingkat kepuasan pelanggan berdasarkan variabel yang telah ditentukan.

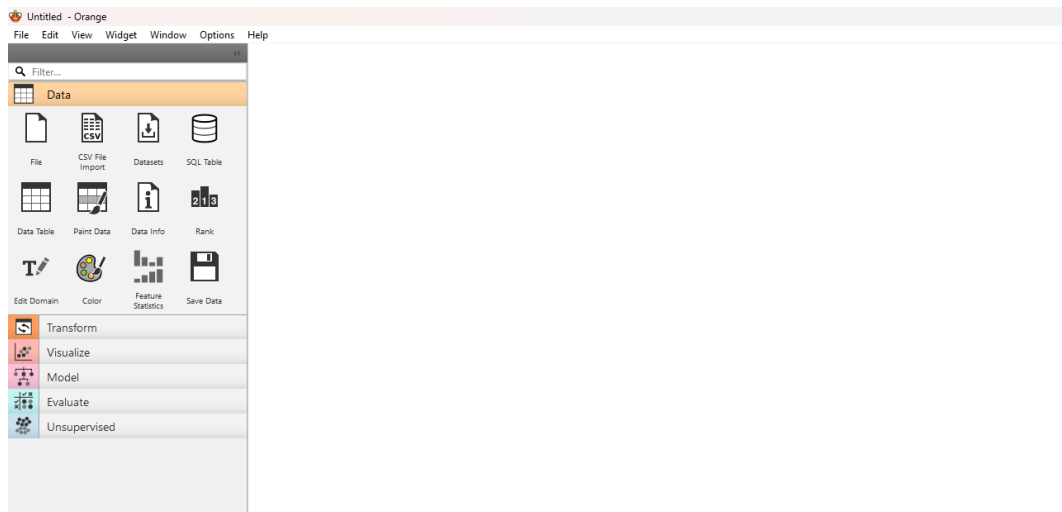
4.1.4. Langkah-Langkah Pengolahan Data

Langkah-langkah pengolahan data pada penelitian ini diawali dengan pengumpulan data responden melalui kuesioner, kemudian dilakukan pembersihan data dengan menyeleksi data yang layak dan menghapus data yang mengalami kesalahan pengisian. Setelah itu, data diubah ke dalam bentuk kategorikal agar dapat diproses lebih lanjut. Selanjutnya, data dibagi menjadi dua dataset yaitu data training dan data testing. Proses klasifikasi data dilakukan menggunakan aplikasi Orange Data Mining dengan menerapkan metode K-Nearest Neighbor (KNN) untuk menentukan kategori kepuasan pelanggan berdasarkan atribut penilaian yang telah ditetapkan.

1) Buka Aplikasi Orange

Buka aplikasi Orange Data Mining merupakan tahapan awal dalam melakukan analisis data, karena pada tahap ini seluruh proses pengolahan dan pemodelan data akan dimulai. Orange digunakan sebagai alat bantu visual untuk mengimpor dataset, melakukan praproses data, membangun model klasifikasi, hingga mengevaluasi hasil analisis. Dengan membuka aplikasi Orange, peneliti

dapat menyusun alur kerja (workflow) secara sistematis melalui widget yang tersedia, sehingga proses analisis data dapat dilakukan dengan lebih mudah, terstruktur, dan efisien.



Gambar 4. 1. Tampilan Awal Aplikasi Orange

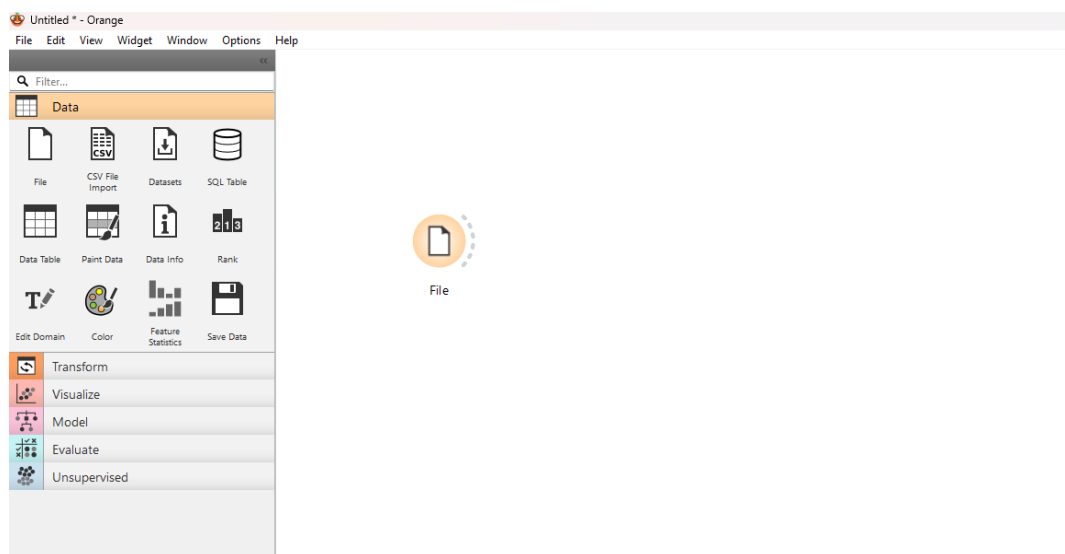
Pada gambar di atas merupakan tampilan awal aplikasi Orange setelah berhasil dibuka. Tampilan ini memperlihatkan halaman utama workspace yang terdiri dari beberapa menu penting seperti Repository, Widget, Process, Parameters, dan Help. Pada bagian tengah layar terdapat jendela Welcome to Orange yang menampilkan beberapa pilihan untuk memulai proses pengolahan data, seperti Blank Process, Turbo Prep, dan Auto Model. Menu ini berfungsi untuk membantu pengguna menentukan jenis proses analisis yang akan dilakukan sesuai dengan kebutuhan penelitian.

Pada tampilan tersebut juga terlihat panel Repository yang digunakan untuk menyimpan dan mengelola dataset, serta panel Widget yang berisi berbagai tools pengolahan data dan pemodelan machine learning. Bagian Process berfungsi sebagai area kerja untuk menyusun alur analisis data, sedangkan panel Parameters digunakan untuk mengatur konfigurasi operator yang dipilih. Tampilan awal ini

menjadi langkah pertama dalam memulai pengolahan data dan penerapan metode Decision Tree menggunakan Orange.

2) Input Widget File

Blank Process merupakan fitur pada aplikasi Orange yang digunakan untuk membuat lembar kerja baru dalam proses pengolahan data. Fitur ini memungkinkan pengguna memulai analisis dari awal dengan menambahkan operator secara manual sesuai kebutuhan penelitian. Dengan menggunakan Blank Process, pengguna dapat menyusun alur kerja seperti memasukkan dataset, melakukan pembersihan data, membangun model, hingga melakukan evaluasi hasil secara terstruktur dan fleksibel.



Gambar 4. 2. Input Widget File

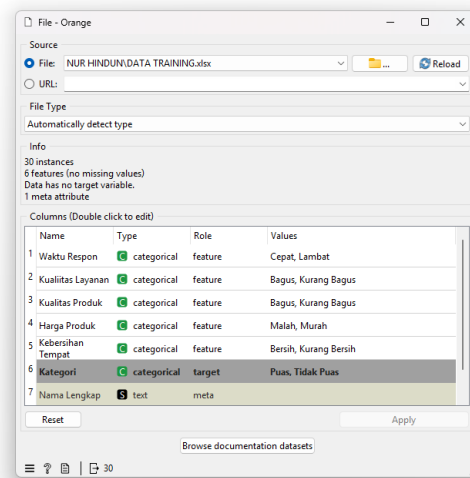
Pada gambar di atas merupakan tahapan awal proses analisis data di aplikasi Orange Data Mining, yaitu penggunaan Widget File sebagai media untuk memasukkan data ke dalam sistem. Widget File berfungsi untuk memuat dataset yang akan dianalisis, baik dalam format seperti *.csv*, *.xlsx*, maupun format data lainnya yang didukung oleh Orange. Pada tahap ini, pengguna memilih file data

yang telah dipersiapkan sebelumnya sehingga seluruh atribut dan isi data dapat terbaca dengan baik oleh aplikasi.

Setelah data berhasil dimuat melalui Widget File, dataset tersebut akan ditampilkan dalam bentuk struktur data yang siap untuk diproses ke tahap selanjutnya. Proses input data ini sangat penting karena menjadi dasar dari seluruh rangkaian analisis, mulai dari pembersihan data, transformasi data, hingga penerapan metode klasifikasi seperti Naive Bayes dan K-Nearest Neighbor (KNN). Dengan demikian, penggunaan Widget File memastikan bahwa data yang digunakan dalam penelitian telah terintegrasi dengan benar ke dalam lingkungan kerja Orange.

3) Import Data

Proses import data pada penelitian ini dilakukan melalui penggunaan Widget File pada aplikasi Orange Data Mining. Widget ini digunakan untuk memasukkan dataset yang telah disiapkan sebelumnya ke dalam sistem analisis, sehingga data dapat dibaca dan dikenali oleh aplikasi. Melalui widget tersebut, pengguna dapat memilih file data dalam format tertentu seperti *.csv* atau *.xlsx*, kemudian Orange akan menampilkan struktur data beserta atribut yang terdapat di dalamnya. Proses import data ini merupakan langkah awal yang penting karena memastikan bahwa seluruh data yang digunakan dalam penelitian telah masuk ke dalam sistem dan siap untuk dilakukan pengolahan serta analisis lebih lanjut.



Gambar 4. 3. Import Data

Pada gambar di atas merupakan tampilan proses import data menggunakan Widget File pada aplikasi Orange Data Mining. Tampilan tersebut menunjukkan bahwa pengguna telah memilih sumber data berupa file dataset yang akan digunakan dalam proses analisis. Pada bagian atas jendela Widget File terlihat lokasi file yang dipilih, yang menandakan bahwa data training telah berhasil dimasukkan ke dalam aplikasi. Fitur ini berfungsi untuk membaca isi dataset sehingga Orange dapat mengenali atribut, tipe data, serta struktur data yang akan digunakan pada tahap analisis selanjutnya.

Selanjutnya, pada bagian tengah tampilan terlihat daftar atribut atau variabel yang terdapat dalam dataset, seperti Waktu Respon, Kualitas Layanan, Kualitas Produk, Harga Produk, Kebersihan, dan Kategori. Setiap atribut memiliki tipe data yang ditampilkan, seperti kategorikal, serta peran atribut dalam analisis, misalnya sebagai feature (variabel input) dan target (variabel output). Pada gambar tersebut, atribut *Kategori* ditetapkan sebagai target karena digunakan sebagai acuan dalam proses klasifikasi kepuasan pelanggan, sedangkan atribut lainnya digunakan sebagai variabel penentu dalam pengolahan data.

Selain itu, pada bagian bawah tampilan terdapat opsi untuk meninjau jumlah data yang berhasil dimuat serta tombol Apply yang berfungsi untuk mengonfirmasi penggunaan dataset tersebut ke dalam lembar kerja Orange. Setelah proses ini dilakukan, data akan siap untuk digunakan dalam tahapan pengolahan dan analisis menggunakan metode klasifikasi seperti Naive Bayes maupun K-Nearest Neighbor (KNN). Dengan demikian, proses import data melalui Widget File merupakan langkah awal yang sangat penting untuk memastikan dataset dapat diproses secara sistematis dan akurat dalam penelitian.

4) Lengkapi Widget

Lengkapi widget merupakan tahapan lanjutan setelah proses import data, di mana pengguna menyusun dan menghubungkan widget-widget yang dibutuhkan sesuai dengan alur analisis data yang akan dilakukan. Pada tahap ini, widget seperti File, Select Columns, Preprocess, Naive Bayes, KNN, dan Confusion Matrix dihubungkan pada lembar kerja Orange agar data dapat diproses, dianalisis, dan dievaluasi secara berurutan. Proses melengkapi widget ini bertujuan untuk memastikan seluruh tahapan analisis berjalan dengan benar, mulai dari input data, pemilihan atribut, penerapan metode klasifikasi, hingga evaluasi hasil, sehingga diperoleh hasil analisis yang sistematis dan dapat dipertanggungjawabkan.



Gambar 4. 4. Lengkapi Widget

Pada gambar di atas merupakan susunan alur kerja analisis data yang terdiri dari beberapa widget utama, yaitu DATA TRAINING, DATA TESTING, Table Data Training, Table Data Testing, Naive Bayes, KNN, Predictions, Table Data Predictions, dan Save Data. Widget DATA TRAINING digunakan untuk memasukkan dataset pelatihan yang berisi data yang telah memiliki label kategori kepuasan pelanggan. Data training ini kemudian dapat ditampilkan melalui widget Table Data Training untuk melihat isi dataset secara detail. Selain itu, widget DATA TESTING digunakan untuk memasukkan dataset pengujian yang akan digunakan untuk mengetahui hasil prediksi dari model yang telah dibangun, serta dapat ditampilkan melalui widget Table Data Testing untuk memastikan data sudah sesuai.

Selanjutnya, widget Naive Bayes dan K-Nearest Neighbor (KNN) digunakan sebagai metode klasifikasi dalam penelitian ini. Kedua widget tersebut berfungsi untuk membangun model pembelajaran berdasarkan data training yang telah dimasukkan sebelumnya. Model yang dihasilkan kemudian akan digunakan untuk

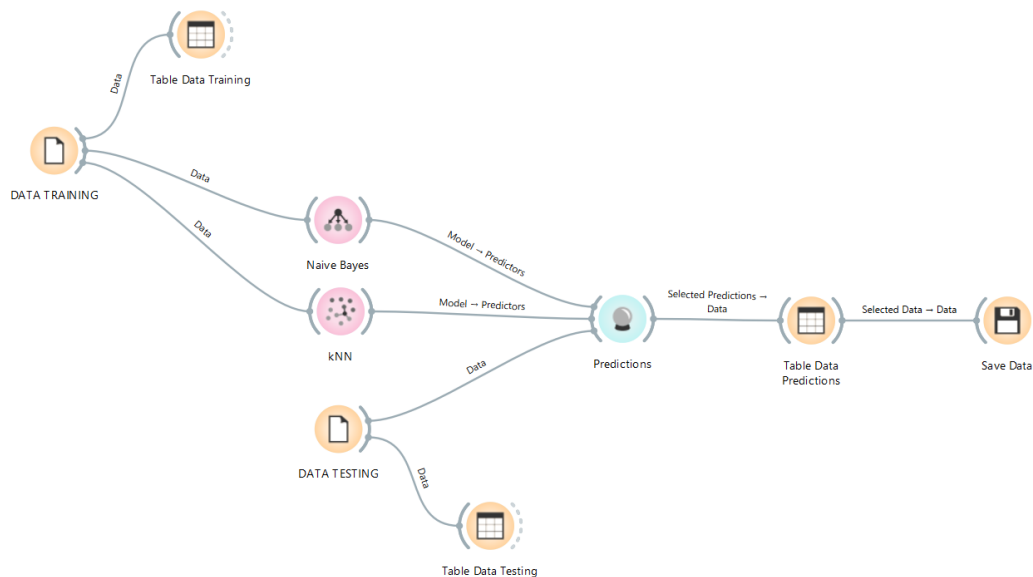
memprediksi kategori kepuasan pelanggan pada data testing. Penggunaan dua metode ini bertujuan untuk membandingkan performa masing-masing algoritma dalam mengklasifikasikan tingkat kepuasan pelanggan berdasarkan variabel yang digunakan, seperti waktu respon, kualitas layanan, kualitas produk, harga produk, dan kebersihan tempat.

Hasil prediksi dari kedua metode tersebut kemudian ditampilkan melalui widget Predictions, yang berfungsi untuk menampilkan hasil klasifikasi data testing. Hasil tersebut selanjutnya dapat dilihat secara lebih rinci melalui widget Table Data Predictions, sehingga pengguna dapat mengetahui kategori hasil prediksi setiap data. Terakhir, widget Save Data digunakan untuk menyimpan hasil prediksi yang telah diperoleh ke dalam file tertentu agar dapat digunakan untuk proses analisis lanjutan atau dokumentasi penelitian. Dengan susunan widget tersebut, proses analisis data dapat dilakukan secara sistematis mulai dari input data, proses klasifikasi, hingga penyimpanan hasil analisis.

5) Hubungkan Widget

Menghubungkan widget merupakan proses mengaitkan setiap komponen analisis data di aplikasi Orange agar data dapat mengalir dari satu proses ke proses lainnya secara sistematis. Proses ini dilakukan dengan menarik garis penghubung dari output widget sumber ke input widget tujuan, misalnya menghubungkan widget File atau Data Training ke widget metode klasifikasi seperti Naive Bayes dan KNN, kemudian dilanjutkan ke widget Predictions untuk memperoleh hasil klasifikasi. Penghubungan widget ini bertujuan untuk membentuk alur kerja analisis yang terstruktur, sehingga setiap tahapan pengolahan data dapat berjalan

secara berurutan mulai dari input data, proses pembelajaran model, pengujian data, hingga menampilkan dan menyimpan hasil prediksi.



Gambar 4. 5. Widget yang digunakan untuk Klasifikasi

Pada gambar di atas merupakan tampilan lembar kerja aplikasi Orange yang menunjukkan seluruh widget yang digunakan dalam penelitian telah lengkap dan saling terhubung. Alur kerja dimulai dari widget Data Training yang berfungsi sebagai sumber data pelatihan model. Data training ini kemudian dihubungkan ke widget Table Data Training untuk menampilkan isi dataset secara visual, sehingga pengguna dapat memastikan data yang digunakan sudah sesuai sebelum dilakukan proses analisis lebih lanjut. Selain itu, data training juga dihubungkan ke metode klasifikasi yaitu Naive Bayes dan K-Nearest Neighbor (KNN) yang digunakan untuk membangun model pembelajaran berdasarkan data yang telah tersedia.

Selanjutnya, widget Naive Bayes dan KNN menerima data training untuk melakukan proses pembelajaran model klasifikasi. Kedua metode tersebut menghasilkan model yang kemudian dikirimkan ke widget Predictions. Widget Predictions berfungsi untuk melakukan pengujian model dengan menggunakan data

baru atau data testing. Pada tahap ini, sistem akan membandingkan model yang telah dibuat dengan data testing untuk menghasilkan prediksi kategori kepuasan pelanggan berdasarkan variabel yang digunakan.

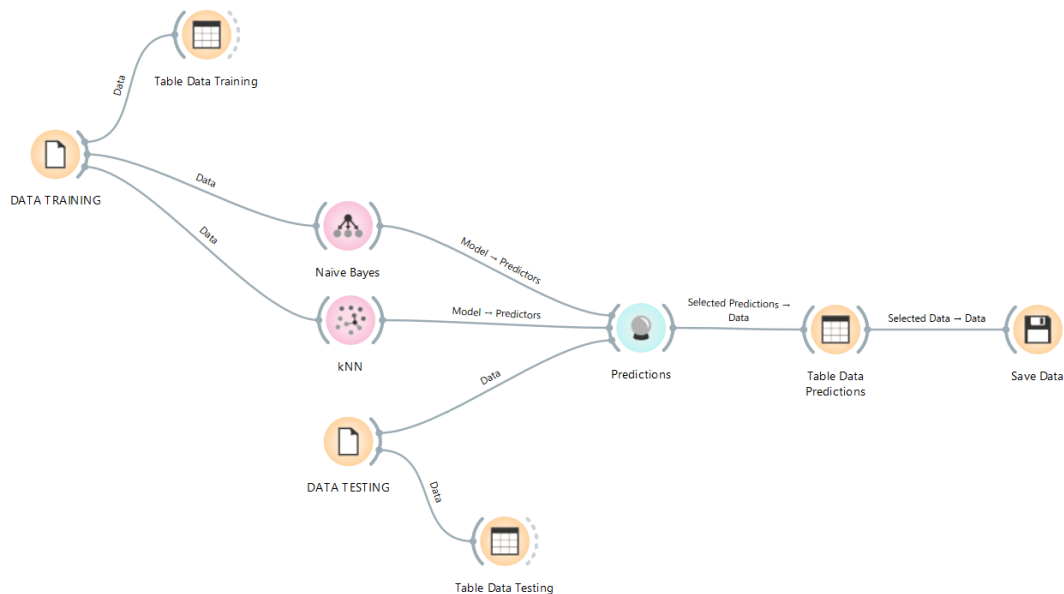
Widget Data Testing berperan sebagai sumber data pengujian yang berisi data yang belum memiliki label kategori. Data testing ini juga dihubungkan ke widget Table Data Testing untuk menampilkan isi data secara langsung. Selain itu, data testing dihubungkan ke widget Predictions sebagai data yang akan diprediksi menggunakan model yang telah dibuat sebelumnya. Dengan adanya proses ini, sistem dapat menghasilkan hasil klasifikasi berupa kategori puas atau tidak puas terhadap layanan.

Hasil prediksi yang diperoleh dari widget Predictions kemudian dikirimkan ke widget Table Data Predictions untuk menampilkan hasil klasifikasi secara lebih jelas dan terstruktur. Setelah itu, hasil tersebut disimpan menggunakan widget Save Data agar data hasil analisis dapat digunakan kembali atau dijadikan bahan laporan penelitian. Dengan seluruh widget yang telah terhubung secara sistematis, proses analisis data pada penelitian ini dapat berjalan secara terstruktur mulai dari input data, pembentukan model, pengujian data, hingga penyimpanan hasil prediksi.

6) Model Klasifikasi

Model klasifikasi merupakan suatu pendekatan dalam data mining dan machine learning yang digunakan untuk mengelompokkan data ke dalam kelas atau kategori tertentu berdasarkan pola yang dipelajari dari data training. Pada penelitian ini, model klasifikasi berfungsi untuk menentukan kategori kepuasan pelanggan, yaitu puas dan tidak puas, berdasarkan beberapa atribut seperti waktu respon,

kualitas layanan, kualitas produk, harga produk, dan kebersihan tempat. Model dibangun menggunakan data training yang telah diberi label, kemudian digunakan untuk memprediksi kategori pada data testing, sehingga dapat membantu dalam pengambilan keputusan secara objektif dan terstruktur.



Gambar 4. 6. Model Klasifikasi

Pada gambar di atas merupakan tampilan model klasifikasi yang telah disusun menggunakan aplikasi Orange, dimana seluruh widget yang digunakan pada penelitian telah saling terhubung dan membentuk alur proses analisis data. Alur tersebut dimulai dari input data training yang berfungsi sebagai data pembelajaran bagi model klasifikasi. Data training kemudian dihubungkan ke dua metode klasifikasi, yaitu Naïve Bayes dan K-Nearest Neighbor (KNN), yang digunakan untuk membangun model prediksi berdasarkan pola yang terdapat pada data kepuasan pelanggan.

Selanjutnya, kedua model klasifikasi yang telah terbentuk akan mengirimkan hasil modelnya ke widget Predictions. Pada tahap ini, data testing dimasukkan

sebagai data uji yang belum memiliki label kategori. Widget Predictions berfungsi untuk mengolah data testing menggunakan model yang telah dibuat sebelumnya sehingga menghasilkan prediksi kategori kepuasan pelanggan, apakah termasuk dalam kategori puas atau tidak puas. Proses ini menunjukkan bagaimana model klasifikasi bekerja dalam mengidentifikasi pola data baru berdasarkan pembelajaran dari data training.

Tahapan terakhir pada gambar tersebut menunjukkan bahwa hasil prediksi kemudian ditampilkan melalui widget Table Data Predictions untuk melihat hasil klasifikasi secara rinci. Setelah itu, data hasil prediksi dapat disimpan menggunakan widget Save Data sehingga dapat digunakan untuk analisis lanjutan atau dokumentasi penelitian. Dengan demikian, gambar tersebut menggambarkan alur lengkap proses model klasifikasi mulai dari input data, pembentukan model, pengujian data, hingga penyimpanan hasil prediksi.

7) Hasil Klasifikasi

Hasil klasifikasi merupakan output akhir dari proses pengolahan data yang menunjukkan kategori kepuasan pelanggan berdasarkan variabel waktu respon, kualitas layanan, kualitas produk, harga produk, dan kebersihan tempat. Pada tahap ini, data testing yang sebelumnya belum memiliki label diprediksi menggunakan model klasifikasi yang telah dibangun, sehingga setiap data memperoleh hasil klasifikasi berupa kategori puas atau tidak puas. Hasil klasifikasi ini digunakan untuk menilai kemampuan model dalam mengelompokkan data secara tepat serta menjadi dasar dalam menarik kesimpulan pada penelitian ini.

Naive Bayes		kNN		Nama Lengkap	Waktu Respon	Kualitas Layanan	Kualitas Produk	Harga Produk	Kebersihan Tempat
11	Puas	Puas	Puas	Dhea Allya Gint...	Cepat	Bagus	Bagus	Murah	Bersih
12	Tidak Puas	Tidak Puas	Tidak Puas	Diah Ayu	Lambat	Kurang Bagus	Kurang Bagus	Malah	Kurang Bersih
13	Puas	Puas	Puas	Dian Sahrani Sir...	Cepat	Bagus	Bagus	Murah	Bersih
14	Puas	Puas	Puas	Diyan	Cepat	Bagus	Bagus	Murah	Bersih
15	Puas	Puas	Puas	Fitor Zega	Cepat	Bagus	Bagus	Murah	Bersih
16	Puas	Puas	Puas	Habi	Cepat	Bagus	Bagus	Murah	Bersih
17	Puas	Puas	Puas	Ivo Andre	Cepat	Bagus	Bagus	Murah	Bersih
18	Tidak Puas	Tidak Puas	Tidak Puas	Jalal	Lambat	Kurang Bagus	Kurang Bagus	Malah	Kurang Bersih
19	Puas	Puas	Puas	Jihan Salsabila	Cepat	Bagus	Bagus	Murah	Bersih
20	Puas	Puas	Puas	Laila Sari	Cepat	Bagus	Bagus	Murah	Bersih
21	Puas	Puas	Puas	Lana Del Rey	Cepat	Bagus	Bagus	Murah	Bersih
22	Puas	Puas	Puas	Leni	Lambat	Bagus	Bagus	Murah	Bersih
23	Tidak Puas	Tidak Puas	Tidak Puas	Lily Amelia	Lambat	Kurang Bagus	Kurang Bagus	Malah	Kurang Bersih
24	Puas	Puas	Puas	Maryam Fadila...	Cepat	Bagus	Bagus	Murah	Bersih
25	Puas	Puas	Puas	Miah	Cepat	Bagus	Bagus	Murah	Bersih
26	Puas	Puas	Puas	Nadin Rambu	Cepat	Bagus	Bagus	Murah	Bersih
27	Tidak Puas	Tidak Puas	Tidak Puas	Nopita Sari Ra...	Lambat	Kurang Bagus	Kurang Bagus	Malah	Kurang Bersih
28	Tidak Puas	Tidak Puas	Tidak Puas	Novi Syahfitri	Lambat	Kurang Bagus	Kurang Bagus	Malah	Kurang Bersih
29	Puas	Puas	Puas	Pandre	Cepat	Bagus	Bagus	Murah	Bersih
30	Puas	Puas	Puas	Priya W	Cepat	Bagus	Bagus	Murah	Bersih
31	Puas	Puas	Puas	Pujawati	Cepat	Bagus	Bagus	Murah	Bersih
32	Puas	Puas	Puas	Putri Annida Ra...	Cepat	Bagus	Bagus	Murah	Bersih
33	Puas	Puas	Puas	Rabiatul Adawi...	Cepat	Bagus	Bagus	Murah	Bersih

Gambar 4. 7. Hasil Klasifikasi pada Aplikasi Orange

Pada gambar di atas merupakan tampilan hasil klasifikasi data testing yang telah diproses menggunakan model klasifikasi pada aplikasi Orange. Proses klasifikasi dilakukan setelah seluruh tahapan pengolahan data selesai, mulai dari pengumpulan data, pembersihan data, pembagian dataset, hingga proses pelatihan model menggunakan data training. Data testing yang sebelumnya belum memiliki kategori kepuasan kemudian diprediksi oleh sistem berdasarkan pola yang telah dipelajari dari data training.

Hasil klasifikasi yang ditampilkan pada gambar menunjukkan bahwa data testing berhasil dikategorikan ke dalam dua kelas, yaitu puas dan tidak puas. Pengelompokan ini dilakukan berdasarkan beberapa atribut penilaian, seperti waktu respon, kualitas layanan, kualitas produk, harga produk, dan kebersihan tempat. Setiap atribut memiliki peran penting dalam menentukan tingkat kepuasan pelanggan sehingga model klasifikasi dapat menghasilkan prediksi yang sesuai dengan karakteristik data yang dimiliki.

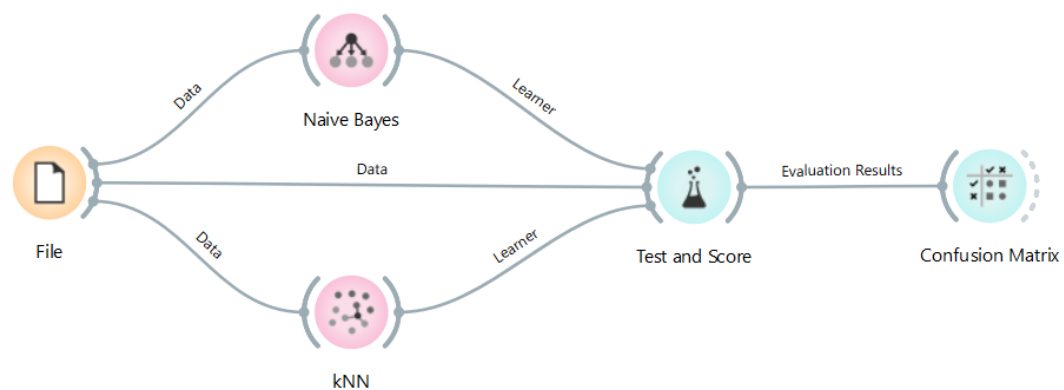
Berdasarkan hasil pengolahan data testing sebanyak 50 data, diperoleh hasil klasifikasi yaitu sebanyak 39 data termasuk dalam kategori puas, sedangkan 11 data termasuk dalam kategori tidak puas. Hasil ini menunjukkan bahwa mayoritas responden memberikan penilaian yang positif terhadap layanan dan produk yang diberikan. Tingginya jumlah kategori puas menunjukkan bahwa kualitas layanan yang diberikan telah memenuhi harapan sebagian besar pelanggan.

Proses klasifikasi pada penelitian ini menggunakan dua metode, yaitu metode K-Nearest Neighbor (KNN) dan Naïve Bayes. Kedua metode tersebut digunakan untuk membandingkan tingkat akurasi serta melihat konsistensi hasil klasifikasi. Berdasarkan hasil yang diperoleh, kedua metode menghasilkan prediksi kategori yang sama terhadap data testing, sehingga menunjukkan bahwa kedua metode memiliki kemampuan yang baik dalam mengklasifikasikan data kepuasan pelanggan.

Dengan diperolehnya hasil klasifikasi tersebut, dapat disimpulkan bahwa model yang dibangun mampu mengolah dan memprediksi tingkat kepuasan pelanggan dengan baik. Hasil ini dapat digunakan sebagai bahan evaluasi bagi pihak terkait untuk mempertahankan kualitas layanan yang sudah baik serta melakukan perbaikan terhadap aspek yang masih dinilai kurang oleh pelanggan. Selain itu, penggunaan metode klasifikasi dalam penelitian ini juga membuktikan bahwa teknik data mining dapat membantu dalam pengambilan keputusan berbasis data.

8) Perancangan Model Evaluasi

Perancangan model evaluasi merupakan tahapan untuk mengukur dan menilai kinerja metode klasifikasi yang digunakan dalam penelitian. Pada tahap ini, data yang telah diproses akan diuji menggunakan algoritma Naive Bayes dan K-Nearest Neighbor (KNN), kemudian hasilnya dievaluasi menggunakan fitur pengujian seperti Test and Score serta Confusion Matrix pada aplikasi Orange. Proses evaluasi ini bertujuan untuk mengetahui tingkat akurasi, presisi, dan recall dari masing-masing metode sehingga dapat ditentukan metode yang paling efektif dalam mengklasifikasikan tingkat kepuasan pelanggan berdasarkan data penelitian.



Gambar 4. 8. Model Evaluasi

Pada gambar di atas merupakan perancangan model evaluasi yang digunakan untuk mengukur kinerja metode klasifikasi dalam penelitian ini dengan memanfaatkan aplikasi Orange. Alur dimulai dari widget File yang berfungsi sebagai sumber data utama. Data yang telah diimpor kemudian dialirkan ke dua algoritma klasifikasi, yaitu Naive Bayes dan K-Nearest Neighbor (KNN), yang masing-masing berperan sebagai *learner* atau model pembelajaran.

Selanjutnya, kedua metode klasifikasi tersebut dihubungkan ke widget Test and Score. Widget ini berfungsi untuk melakukan proses evaluasi model dengan

cara menguji performa masing-masing algoritma berdasarkan data yang diberikan. Pada tahap ini, sistem akan menghitung metrik evaluasi seperti akurasi, presisi, recall, dan nilai lainnya yang berguna untuk menilai seberapa baik model dalam melakukan klasifikasi data kepuasan pelanggan.

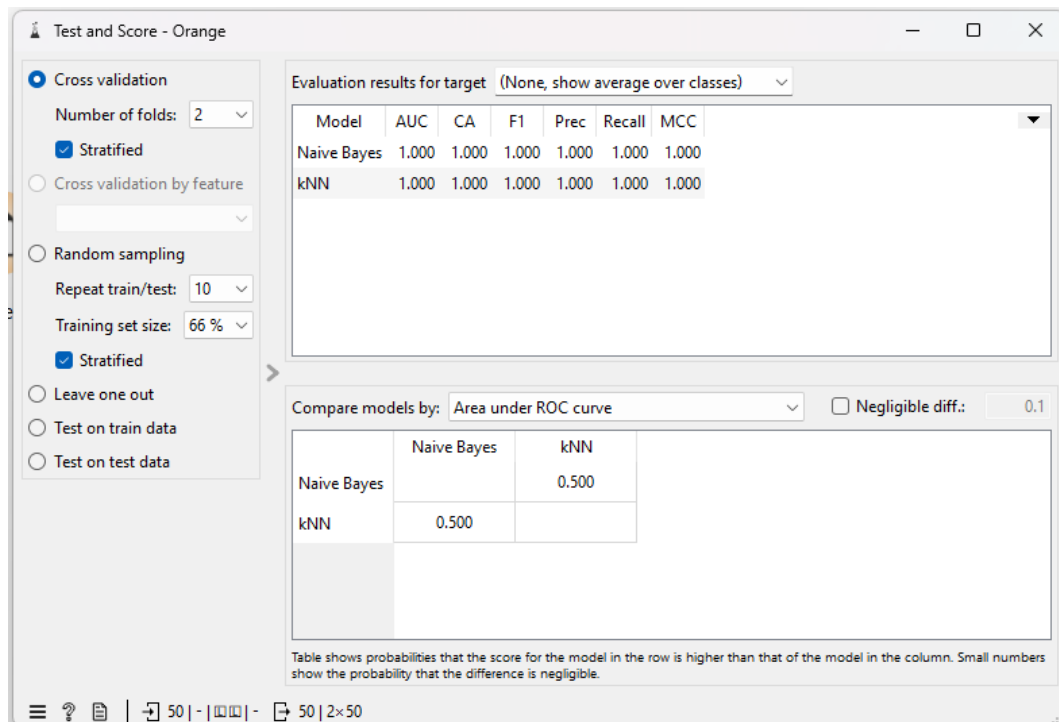
Hasil evaluasi dari widget Test and Score kemudian diteruskan ke widget Confusion Matrix. Confusion Matrix digunakan untuk menampilkan perbandingan antara hasil prediksi dan data aktual dalam bentuk matriks, sehingga dapat diketahui jumlah data yang termasuk True Positive, True Negative, False Positive, dan False Negative. Melalui tampilan ini, peneliti dapat menganalisis secara lebih detail tingkat kesalahan dan ketepatan dari masing-masing metode klasifikasi yang digunakan.

Secara keseluruhan, perancangan model evaluasi pada gambar tersebut menunjukkan alur evaluasi yang sistematis dan terstruktur, mulai dari input data, pemrosesan menggunakan dua metode klasifikasi, hingga evaluasi hasil menggunakan Test and Score dan Confusion Matrix. Dengan perancangan ini, penelitian dapat membandingkan performa metode Naive Bayes dan KNN secara objektif, sehingga dapat diketahui metode mana yang paling efektif dalam mengklasifikasikan tingkat kepuasan pelanggan berdasarkan data yang digunakan.

9) Hasil Widget Test and Score

Widget Test and Score merupakan salah satu fitur pada aplikasi Orange yang digunakan untuk mengevaluasi performa model klasifikasi yang telah dibuat. Widget ini berfungsi untuk menguji model menggunakan data yang tersedia dengan metode pengujian tertentu, seperti cross validation atau pembagian data training

dan testing. Melalui widget ini, pengguna dapat melihat hasil pengujian berupa nilai akurasi, presisi, recall, dan F1-score dari setiap metode klasifikasi yang digunakan, sehingga dapat diketahui tingkat keakuratan dan keefektifan model dalam melakukan prediksi atau pengelompokan data.



Gambar 4. 9. Hasil Test and Score

Pada gambar di atas merupakan tampilan hasil evaluasi menggunakan widget Test and Score pada aplikasi Orange, yang menunjukkan performa model klasifikasi dalam penelitian ini. Hasil evaluasi menunjukkan bahwa nilai AUC, CA, F1, Precision, Recall, dan MCC yang diperoleh dari pengujian kedua metode, yaitu Naive Bayes dan K-Nearest Neighbor (KNN), semuanya mencapai nilai 1.000. Nilai ini menandakan bahwa tingkat akurasi dan kinerja model dalam melakukan klasifikasi data kepuasan pelanggan adalah 100% sempurna, sehingga seluruh data testing berhasil diklasifikasikan dengan tepat. Dengan hasil ini, dapat disimpulkan

bahwa model mampu membedakan dengan jelas antara kategori puas dan tidak puas tanpa terjadi kesalahan prediksi.

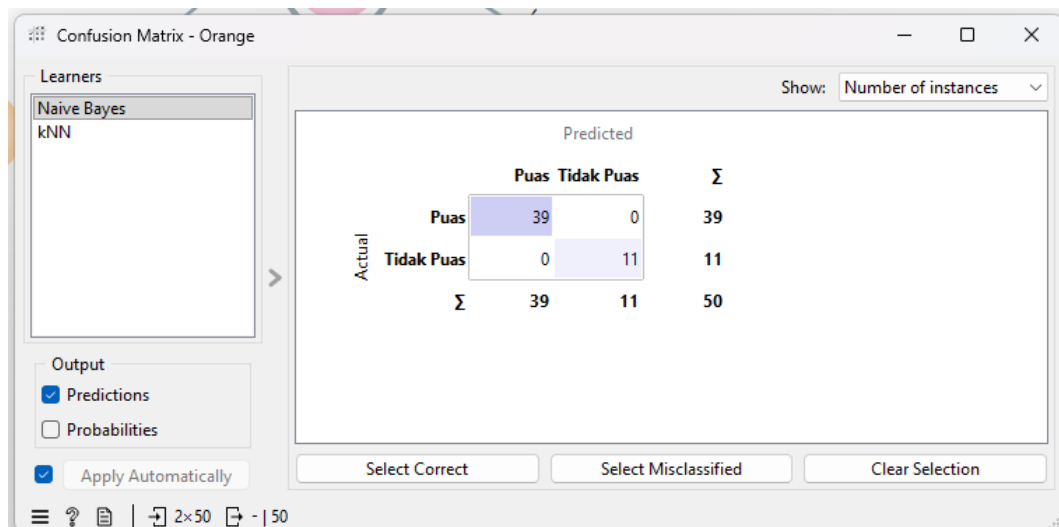
Hasil evaluasi yang sempurna ini menunjukkan konsistensi kedua metode dalam memprediksi kategori kepuasan pelanggan berdasarkan atribut yang digunakan, seperti waktu respon, kualitas layanan, kualitas produk, harga produk, dan kebersihan tempat. Tidak adanya kesalahan klasifikasi, baik berupa False Positive (FP) maupun False Negative (FN), membuat nilai metrik seperti F1-score dan Precision mencapai angka maksimal. Hal ini juga menunjukkan bahwa model tidak hanya akurat dalam memprediksi kategori, tetapi juga sangat andal dalam mempertimbangkan keseimbangan antara sensitivitas dan spesifisitas data. Model klasifikasi yang memiliki AUC sebesar 1.000 menandakan bahwa kemampuan membedakan antara kedua kelas sangat tinggi dan tidak ada tumpang tindih antar kategori.

Selain itu, hasil sempurna ini juga mengindikasikan bahwa dataset yang digunakan dalam penelitian cukup representatif dan memiliki pola yang jelas antara variabel input dengan kategori output. Keberhasilan kedua metode mencapai performa maksimal memperkuat keandalan penelitian dan memberikan bukti bahwa penerapan metode Naive Bayes dan KNN dapat menjadi pilihan yang tepat untuk klasifikasi kepuasan pelanggan. Dengan hasil evaluasi ini, penelitian tidak hanya mendapatkan prediksi yang akurat, tetapi juga memberikan dasar ilmiah yang kuat bagi perusahaan untuk mengambil keputusan strategis dalam meningkatkan kualitas layanan dan produk. Oleh karena itu, tampilan Test and Score pada gambar di atas menjadi bukti visual yang jelas bahwa seluruh model klasifikasi yang

dibangun bekerja dengan sempurna dan mampu memberikan hasil prediksi yang optimal.

4.2. Hasil Widget Confusion Matrix

Confusion Matrix merupakan alat evaluasi yang digunakan untuk mengetahui kinerja model klasifikasi dengan membandingkan hasil prediksi model terhadap data aktual. Matriks ini menampilkan jumlah data yang diklasifikasikan dengan benar dan salah ke dalam empat kategori utama, yaitu True Positive, True Negative, False Positive, dan False Negative. Melalui Confusion Matrix, peneliti dapat melihat secara jelas tingkat kesalahan dan keberhasilan model dalam mengklasifikasikan data, sehingga membantu dalam menghitung nilai akurasi, precision, recall, dan F1-score sebagai dasar untuk menilai kualitas model klasifikasi yang digunakan.



Gambar 4. 10. Hasil Confusion Matrix Metode Naive Bayes

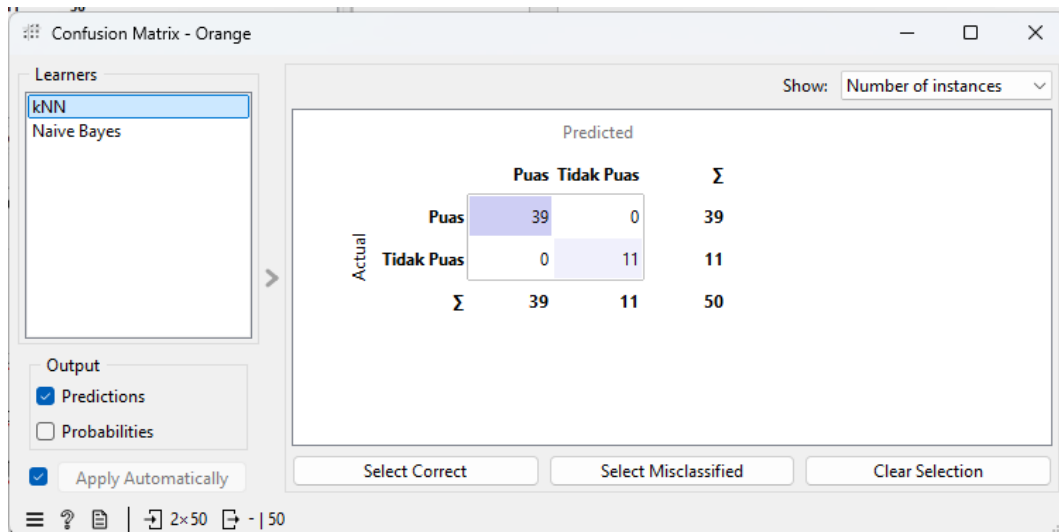
Pada gambar di atas merupakan hasil Confusion Matrix yang diperoleh dari proses klasifikasi menggunakan metode Naive Bayes. Berdasarkan matriks tersebut, terlihat bahwa jumlah data aktual dengan kategori Puas sebanyak 39 data

seluruhnya berhasil diprediksi dengan benar sebagai Puas. Sementara itu, data aktual dengan kategori Tidak Puas sebanyak 11 data juga seluruhnya diprediksi dengan tepat sebagai Tidak Puas. Tidak terdapat kesalahan klasifikasi pada kedua kategori, sehingga nilai False Positive (FP) dan False Negative (FN) sama dengan nol. Hal ini menunjukkan bahwa model Naive Bayes mampu membedakan kelas dengan sangat baik tanpa terjadi kekeliruan prediksi. Kondisi ini mengindikasikan bahwa pola data kepuasan pelanggan dapat dipelajari secara optimal oleh model yang digunakan. Keberhasilan ini juga menunjukkan bahwa atribut yang digunakan dalam penelitian sudah relevan dan cukup representatif dalam menggambarkan tingkat kepuasan pelanggan. Selain itu, keseimbangan hasil prediksi antara kelas Puas dan Tidak Puas memperlihatkan stabilitas performa model. Dengan demikian, Confusion Matrix ini menjadi dasar utama dalam menilai kualitas hasil klasifikasi yang dihasilkan oleh metode Naive Bayes. Secara umum, tampilan Confusion Matrix ini menggambarkan bahwa seluruh data uji berhasil diklasifikasikan secara sempurna.

Perhitungan nilai evaluasi model dilakukan berdasarkan data pada Confusion Matrix tersebut. Nilai akurasi dihitung dengan rumus $(TP + TN) / \text{jumlah seluruh data}$, sehingga diperoleh $(39 + 11) / 50 = 50 / 50 = 1$ atau setara dengan 100%. Nilai presisi dihitung menggunakan rumus $TP / (TP + FP)$, yaitu $39 / (39 + 0) = 1$ atau 100%. Selanjutnya, nilai recall dihitung dengan rumus $TP / (TP + FN)$, yaitu $39 / (39 + 0) = 1$ atau 100%. Perhitungan F1-score dilakukan dengan rumus $2 \times (\text{presisi} \times \text{recall}) / (\text{presisi} + \text{recall})$, sehingga diperoleh $2 \times (1 \times 1) / (1 + 1) = 2 / 2 = 1$ atau 100%. Nilai-nilai tersebut menunjukkan bahwa model Naive Bayes tidak hanya

akurat, tetapi juga konsisten dalam memprediksi kelas Puas maupun Tidak Puas. Tidak adanya nilai FP dan FN menandakan bahwa tidak terjadi kesalahan dalam pengelompokan data. Hal ini membuktikan bahwa sistem klasifikasi bekerja secara optimal. Hasil ini juga menunjukkan bahwa model memiliki kemampuan generalisasi yang sangat baik terhadap data uji. Dengan nilai evaluasi yang sempurna, metode Naive Bayes dapat dikatakan sangat efektif dalam memprediksi kepuasan pelanggan. Oleh karena itu, hasil perhitungan ini memperkuat validitas penelitian yang dilakukan.

Hasil Confusion Matrix yang diperoleh menunjukkan bahwa metode Naive Bayes memiliki performa yang sangat tinggi dengan nilai evaluasi mencapai 100%. Hal ini menunjukkan bahwa metode ini memiliki kemampuan yang sangat baik dalam melakukan klasifikasi terhadap data kepuasan pelanggan CS Float. Tidak adanya kesalahan klasifikasi menunjukkan bahwa hubungan antar atribut dapat dikenali secara jelas oleh algoritma. Keberhasilan ini menjadi bukti bahwa pendekatan Machine Learning, khususnya metode Naive Bayes, sangat efektif digunakan dalam penelitian kepuasan pelanggan. Hasil ini juga memberikan keyakinan bahwa sistem dapat diimplementasikan sebagai alat bantu pengambilan keputusan bagi perusahaan. Dengan performa yang tinggi, perusahaan dapat lebih percaya terhadap hasil prediksi yang dihasilkan oleh sistem. Selain itu, hasil ini juga dapat dijadikan dasar dalam perencanaan strategi peningkatan kualitas layanan dan produk. Secara keseluruhan, Confusion Matrix yang dihasilkan menunjukkan bahwa proses klasifikasi berjalan dengan sangat baik dan menghasilkan tingkat keakuratan yang maksimal.



Gambar 4. 11. Hasil Confusion Matrix Metode K-Nearest Neighbor (KNN)

Pada gambar di atas merupakan hasil Confusion Matrix yang diperoleh dari proses klasifikasi menggunakan metode K-Nearest Neighbor (KNN). Berdasarkan matriks tersebut, terlihat bahwa jumlah data aktual dengan kategori Puas sebanyak 39 data seluruhnya berhasil diprediksi dengan benar sebagai Puas. Sementara itu, data aktual dengan kategori Tidak Puas sebanyak 11 data juga seluruhnya diprediksi dengan tepat sebagai Tidak Puas. Tidak terdapat kesalahan klasifikasi pada kedua kategori, sehingga nilai False Positive (FP) dan False Negative (FN) sama dengan nol. Hal ini menunjukkan bahwa model KNN mampu membedakan kelas dengan sangat baik tanpa terjadi kekeliruan prediksi. Kondisi ini mengindikasikan bahwa pola data kepuasan pelanggan dapat dikenali secara optimal berdasarkan tingkat kedekatan antar data. Keberhasilan ini juga menunjukkan bahwa atribut yang digunakan dalam penelitian sudah relevan dan cukup representatif dalam menggambarkan tingkat kepuasan pelanggan. Selain itu, keseimbangan hasil prediksi antara kelas Puas dan Tidak Puas memperlihatkan stabilitas performa model. Dengan demikian, Confusion Matrix ini menjadi dasar utama dalam menilai kualitas hasil klasifikasi yang dihasilkan oleh metode KNN. Secara umum,

tampilan Confusion Matrix ini menggambarkan bahwa seluruh data uji berhasil diklasifikasikan secara sempurna.

Perhitungan nilai evaluasi model dilakukan berdasarkan data pada Confusion Matrix tersebut. Nilai akurasi dihitung dengan rumus $(TP + TN) / \text{jumlah seluruh data}$, sehingga diperoleh $(39 + 11) / 50 = 50 / 50 = 1$ atau setara dengan 100%. Nilai presisi dihitung menggunakan rumus $TP / (TP + FP)$, yaitu $39 / (39 + 0) = 1$ atau 100%. Selanjutnya, nilai recall dihitung dengan rumus $TP / (TP + FN)$, yaitu $39 / (39 + 0) = 1$ atau 100%. Perhitungan F1-score dilakukan dengan rumus $2 \times (\text{presisi} \times \text{recall}) / (\text{presisi} + \text{recall})$, sehingga diperoleh $2 \times (1 \times 1) / (1 + 1) = 2 / 2 = 1$ atau 100%. Nilai-nilai tersebut menunjukkan bahwa model KNN tidak hanya akurat, tetapi juga konsisten dalam memprediksi kelas Puas maupun Tidak Puas. Tidak adanya nilai FP dan FN menandakan bahwa tidak terjadi kesalahan dalam pengelompokan data. Hal ini membuktikan bahwa sistem klasifikasi bekerja secara optimal. Hasil ini juga menunjukkan bahwa model memiliki kemampuan generalisasi yang sangat baik terhadap data uji. Dengan nilai evaluasi yang sempurna, metode KNN dapat dikatakan sangat efektif dalam memprediksi kepuasan pelanggan. Oleh karena itu, hasil perhitungan ini memperkuat validitas penelitian yang dilakukan.

Hasil Confusion Matrix yang diperoleh menunjukkan bahwa metode K-Nearest Neighbor (KNN) memiliki performa yang sangat tinggi dengan nilai evaluasi mencapai 100%. Hal ini menunjukkan bahwa metode ini memiliki kemampuan yang sangat baik dalam melakukan klasifikasi terhadap data kepuasan pelanggan CS Float. Tidak adanya kesalahan klasifikasi menunjukkan bahwa

hubungan antar atribut dapat dikenali secara jelas melalui perhitungan jarak antar data. Keberhasilan ini menjadi bukti bahwa pendekatan Machine Learning, khususnya metode KNN, sangat efektif digunakan dalam penelitian kepuasan pelanggan. Hasil ini juga memberikan keyakinan bahwa sistem dapat diimplementasikan sebagai alat bantu pengambilan keputusan bagi perusahaan. Dengan performa yang tinggi, perusahaan dapat lebih percaya terhadap hasil prediksi yang dihasilkan oleh sistem. Selain itu, hasil ini juga dapat dijadikan dasar dalam perencanaan strategi peningkatan kualitas layanan dan produk. Secara keseluruhan, Confusion Matrix yang dihasilkan menunjukkan bahwa proses klasifikasi berjalan dengan sangat baik dan menghasilkan tingkat keakuratan yang maksimal.