

BAB III

ANALISIS DAN PERANCANGAN

3.1. Arsitektur Sistem

Arsitektur sistem dalam penelitian ini dirancang secara modular untuk mendukung proses analisis kepuasan terhadap opini masyarakat mengenai layanan rehabilitasi sosial yang diberikan oleh Dinas Sosial Rantauprapat. Sistem terdiri dari beberapa komponen utama, yaitu modul pengumpulan data melalui kuesioner digital, preprocessing data (pembersihan data tidak relevan dan agregasi skor), transformasi data menjadi kategori (Baik/Tidak Baik), dan modul klasifikasi menggunakan algoritma *Naive Bayes*. Setiap komponen terintegrasi secara berurutan untuk memastikan proses analisis berjalan efisien, mulai dari input data hasil kuesioner hingga output berupa klasifikasi kepuasan masyarakat dalam dua kategori, yaitu "Puas" dan "Tidak Puas". Arsitektur ini juga mencakup modul evaluasi yang menggunakan confusion matrix untuk menilai kinerja model klasifikasi, sehingga hasil analisis dapat dijadikan acuan dalam memahami persepsi masyarakat dan meningkatkan kualitas pelayanan secara berkelanjutan.

3.1.1. Lokasi dan Waktu Penelitian

Penelitian ini dilaksanakan di Dinas Sosial Kabupaten Labuhanbatu yang berlokasi di wilayah Rantauprapat, Sumatera Utara, sebagai instansi yang menjadi fokus utama dalam pengumpulan dan analisis opini masyarakat terkait layanan rehabilitasi sosial. Pemilihan lokasi ini didasarkan pada peran strategis Dinas Sosial dalam menyelenggarakan program kesejahteraan dan pelayanan publik kepada masyarakat setempat. Adapun waktu pelaksanaan penelitian dimulai pada bulan

Februari hingga Maret 2025, yang mencakup tahapan pengumpulan data, preprocessing, pelatihan model, evaluasi, hingga penyusunan laporan akhir. Rentang waktu ini dipilih untuk memastikan seluruh proses penelitian berjalan secara sistematis dan komprehensif.

3.1.2. Populasi dan Sampel

Populasi dalam penelitian ini adalah seluruh opini masyarakat yang diunggah di media sosial dan platform digital yang berkaitan dengan layanan Dinas Sosial Kabupaten Labuhanbatu di wilayah Rantauprapat. Karena jumlah opini yang tersedia secara daring sangat besar dan tidak terstruktur, penelitian ini menggunakan teknik sampling purposive, yaitu memilih data opini yang secara eksplisit menyebutkan pengalaman atau tanggapan terhadap pelayanan Dinas Sosial. Sebagai sampel, diambil sebanyak 500 opini masyarakat dari media sosial seperti Twitter dan komentar dari forum publik selama periode pengamatan Februari hingga Maret 2025. Sampel ini dianggap mewakili kepuasan masyarakat secara umum karena telah melalui proses seleksi berdasarkan relevansi dan kejelasan isi opini.

3.1.3. Teknik Pengumpulan Data

Dalam penelitian ini, teknik pengumpulan data dilakukan melalui penyebaran kuesioner digital berbasis Google Form yang dirancang untuk menjangkau opini masyarakat terhadap layanan rehabilitasi sosial oleh Dinas Sosial Rantauprapat. Kuesioner disusun secara terstruktur dan terdiri dari pertanyaan tertutup menggunakan skala penilaian serta pertanyaan terbuka yang memberikan ruang bagi responden untuk menyampaikan opini atau pengalaman secara langsung.

Isi kuesioner mencakup beberapa bagian penting, yaitu data demografis responden (usia, jenis kelamin, pekerjaan), penilaian terhadap aspek layanan (seperti kecepatan, keramahan, dan keadilan pelayanan), serta kolom opini terbuka yang menjadi sumber utama dalam analisis kepuasan. Data teks dari opini terbuka inilah yang digunakan sebagai bahan utama untuk pelatihan dan pengujian model klasifikasi menggunakan algoritma *Naive Bayes*.

Pengumpulan data dilakukan selama periode Februari hingga Maret 2025, dengan metode penyebaran melalui media sosial, grup komunitas lokal, dan jaringan komunikasi masyarakat Rantauprapat. Untuk menjaga kualitas data, respon yang kosong atau tidak relevan dieliminasi pada tahap preprocessing. Pendekatan ini dinilai efektif dalam menjangkau masyarakat secara luas dan efisien, serta mendukung pengolahan data opini menjadi dataset terstruktur yang siap dianalisis secara otomatis melalui teknik machine learning.

Tabel 3. 1. Data Kuesioner

No	Kategori	Pertanyaan
1	KUALITAS PELAYANAN	Petugas rehabilitasi melayani dengan ramah dan profesional?
		Proses rehabilitasi berjalan dengan cepat dan efisien ?
		Saya mendapatkan informasi yang jelas mengenai layanan rehabilitasi?
2	AKSESIBILITAS LAYANAN	Lokasi layanan rehabilitasi mudah dijangkau?
		Persyaratan administrasi untuk mendapatkan layanan mudah dipenuhi?
		Saya dapat mengakses layanan rehabilitasi tanpa hambatan yang berarti?
3	TRANSPARANSI INFORMASI	Informasi tentang layanan rehabilitasi tersedia secara jelas dan lengkap?
		Saya mengetahui biaya yang harus dikeluarkan (jika ada) dengan transparan?
		Ada media resmi (website, sosial media, atau brosur) yang menyediakan informasi layanan rehabilitasi?
4	KEPUASAN MASYARAKAT	Saya merasa puas dengan pelayanan rehabilitasi yang diberikan oleh Dinas Sosial?

		Layanan rehabilitasi telah memberikan manfaat yang sesuai dengan harapan saya?
		Saya akan merekomendasikan layanan rehabilitasi ini kepada orang lain yang membutuhkan?

Pada tabel di atas ditampilkan struktur pertanyaan kuesioner yang digunakan untuk mengukur persepsi masyarakat terhadap layanan rehabilitasi sosial yang diselenggarakan oleh Dinas Sosial Rantauprapat. Pertanyaan dibagi ke dalam empat kategori utama, yaitu Kualitas Pelayanan, Aksesibilitas Layanan, Transparansi Informasi, dan Kepuasan Masyarakat. Masing-masing kategori memuat tiga pertanyaan yang disusun dalam bentuk pernyataan dan dijawab oleh responden menggunakan skala Likert, dengan tujuan untuk menangkap tingkat persetujuan atau pengalaman pribadi mereka terhadap setiap aspek pelayanan yang diberikan. Pendekatan ini memungkinkan peneliti untuk melakukan penilaian yang lebih terukur dan sistematis terhadap berbagai dimensi layanan publik.

Kategori Kualitas Pelayanan menyoroti interaksi antara petugas dan masyarakat serta efisiensi dalam proses pelayanan. Aksesibilitas Layanan berfokus pada kemudahan dalam menjangkau layanan serta kejelasan dan kelengkapan persyaratan yang harus dipenuhi oleh masyarakat. Transparansi Informasi mencakup ketersediaan dan kejelasan informasi terkait prosedur, biaya, serta media penyampaian informasi. Sementara itu, Kepuasan Masyarakat mengukur persepsi akhir responden terhadap manfaat layanan yang diterima serta sejauh mana mereka bersedia merekomendasikan layanan tersebut kepada pihak lain. Seluruh pertanyaan ini menjadi bagian penting dalam analisis kuantitatif, dan bila

dikombinasikan dengan data opini terbuka, akan memperkuat hasil analisis kepuasan secara keseluruhan dalam penelitian ini.

3.2. Desain Aktifitas Sistem

Desain aktivitas sistem pada penelitian ini menggambarkan alur kerja proses klasifikasi kepuasan opini masyarakat terhadap layanan rehabilitasi sosial yang disediakan oleh Dinas Sosial Rantauprapat. Aktivitas dimulai dengan tahapan pengumpulan data, yaitu opini atau komentar masyarakat yang diperoleh melalui kuesioner online dan platform digital lainnya. Data yang terkumpul kemudian masuk ke tahap preprocessing, yang mencakup pembersihan teks, tokenisasi, penghapusan stopwords, dan stemming untuk memastikan data dalam format teks siap dianalisis. Setelah itu, data teks yang telah bersih diekstraksi fiturnya menggunakan metode Term Frequency-Inverse Document Frequency (TF-IDF) untuk mengubah teks menjadi bentuk numerik. Data yang sudah dalam bentuk fitur ini kemudian digunakan untuk pelatihan model menggunakan algoritma *Naive Bayes*, dengan tujuan membentuk model klasifikasi kepuasan.

Tahap selanjutnya adalah proses pengujian model, yaitu menguji kemampuan model dalam mengklasifikasikan opini baru ke dalam kategori kepuasan positif, negatif, atau netral. Hasil klasifikasi ini menjadi output akhir dari sistem, yang memberikan gambaran umum mengenai persepsi masyarakat terhadap layanan rehabilitasi sosial di Dinas Sosial Rantauprapat. Informasi ini selanjutnya dapat digunakan oleh pihak instansi sebagai dasar evaluasi dan perumusan strategi peningkatan mutu layanan yang lebih tepat sasaran dan sesuai dengan harapan masyarakat.

3.3. Langkah-Langkah Perhitungan Data

Langkah-langkah perhitungan data dalam penelitian ini dilakukan dengan menggunakan pendekatan *Naive Bayes*, yang berfokus pada perhitungan probabilitas kemunculan kata-kata dalam teks untuk menentukan kategori kepuasan. Tahap awal dimulai dengan pra-pemrosesan data, yaitu mengubah opini masyarakat yang bersifat kualitatif menjadi bentuk teks bersih yang dapat diproses secara matematis. Proses ini mencakup penghapusan tanda baca, simbol, dan stopwords, serta dilakukan tokenisasi dan stemming untuk mengubah kata menjadi bentuk dasarnya. Setiap entri opini masyarakat kemudian diubah menjadi vektor fitur dengan menggunakan metode Term Frequency-Inverse Document Frequency (TF-IDF), sehingga sistem dapat mengenali bobot penting dari setiap kata dalam korpus.

Setelah itu, data diberi label kepuasan berdasarkan kategori Puas dan Tidak Puas sebagai target klasifikasi. Dalam proses pelatihan model, algoritma *Naive Bayes* akan menganalisis distribusi kata-kata yang muncul dalam masing-masing kategori dan menghitung probabilitas setiap kata untuk setiap kelas. Model akan mengasumsikan bahwa setiap fitur (kata) bersifat independen satu sama lain, dan akan menghitung probabilitas total sebuah opini termasuk dalam kategori tertentu menggunakan rumus Teorema Bayes. Model yang telah dilatih ini kemudian digunakan untuk mengklasifikasikan opini baru dengan cara menghitung kemungkinan terbesar berdasarkan kata-kata yang muncul dalam teks. Pendekatan ini dinilai efisien dan akurat, terutama untuk data teks berbahasa Indonesia yang

memiliki kompleksitas tersendiri, serta mampu menangani opini dalam jumlah besar dengan waktu komputasi yang relatif cepat.

3.3.1. Pengumpulan Data

Pengumpulan data dalam penelitian ini diperoleh langsung dari masyarakat melalui penyebaran kuesioner digital yang disebarakan secara online, khususnya kepada warga yang pernah berinteraksi dengan layanan rehabilitasi Dinas Sosial Rantauprapat. Kuesioner ini dirancang untuk menggali persepsi masyarakat terhadap kualitas layanan rehabilitasi sosial yang mereka terima. Data yang dikumpulkan mencakup lima atribut utama, yaitu Nama Responden, Kualitas Pelayanan, Aksesibilitas Layanan, Transparansi Informasi, dan Kepuasan Masyarakat. Dari kelima atribut tersebut, empat digunakan sebagai variabel independen, sementara atribut Kepuasan Masyarakat berfungsi sebagai variabel dependen atau kelas target dalam proses klasifikasi. Data ini kemudian dijadikan dasar dalam penerapan algoritma *Naive Bayes* untuk menentukan tingkat kepuasan masyarakat terhadap pelayanan rehabilitasi sosial yang diberikan oleh Dinas Sosial.

Tabel 3. 2. Data Hasil Kuesioner

Nama	Usia	Jenis Kelamin	Pekerjaan	Kualitas Pelayanan			Aksesibilitas Layanan			Transparansi Informasi			Kepuasan Masyarakat			Total	Kategori
				P1X1	P1X2	P1X3	P2X1	P2X2	P2X3	P3X1	P3X2	P3X3	P4X1	P4X2	P4X3		
Alfin	22	Pria	jual goreng	4	4	4	4	4	4	4	4	4	4	4	4	12	Puas
Amanda utrk br siregar	22	wanita	admin	4	3	4	4	4	3	4	4	4	4	4	4	11	Puas
Andi pandiangan	56	Pria	Pedagang	5	3	4	4	1	3	3	3	4	4	3	4	12	Puas
Anggie Caniago	35	Pria	Angkat galon	5	5	5	5	5	5	5	5	5	5	5	5	15	Puas
Annisa Sepreni	22	wanita	Mahasiswa	3	4	3	5	5	3	3	4	4	4	4	3	10	Puas
Aprelia marita devi	22	wanita	Ibu Rumah tangga	4	4	4	4	4	4	4	4	4	4	4	4	12	Puas
Ayau	22	wanita	Ibu Rumah tangga	4	4	4	4	4	4	4	4	4	4	4	4	12	Puas
Azary	22	wanita	Admin	5	4	5	5	5	5	5	5	5	5	5	5	14	Puas
Berta ngia	21	wanita	Mahasiswa	4	4	3	4	4	5	4	3	3	4	4	4	11	Puas
Biang sitompul	25	Pria	Sales	5	5	5	5	5	5	5	5	5	5	5	5	15	Puas
Bobby	23	Pria	PT PKS	5	5	5	5	5	5	5	5	5	5	5	5	15	Puas
Dedek	25	wanita	msl	5	5	5	5	5	5	5	5	5	5	5	5	15	Puas
Devi enjes marbun	28	wanita	Penjahit	4	3	4	4	4	4	4	4	4	4	4	4	11	Puas
Devi lestari	23	wanita		5	3	3	4	3	3	5	3	3	4	5	5	11	Puas
Devi vana	23	wanita	Guru	4	3	4	4	3	3	3	3	4	3	3	3	11	Tidak Puas
Devi kusuma	25	wanita	Guru	4	4	2	4	4	3	4	4	4	4	4	3	10	Puas
Donna Be Sinurat	44	wanita	Ibu Rumah tangga	4	3	4	3	4	3	4	3	4	3	3	3	11	Puas
Elfredo	20	Pria	Usaha	3	3	3	3	3	3	3	3	3	3	3	3	9	Tidak Puas
Zeyhann Se ga	18	Pria	abang bengkel	5	3	4	3	3	3	3	4	4	3	4	5	12	Puas

Pada tabel di atas merupakan data hasil kuesioner yang telah dikumpulkan dari responden masyarakat terkait persepsi mereka terhadap layanan rehabilitasi sosial yang diberikan oleh Dinas Sosial Rantauprapat. Setiap baris merepresentasikan satu responden yang memberikan penilaian terhadap empat aspek utama pelayanan, yaitu Kualitas Pelayanan, Aksesibilitas Layanan, Transparansi Informasi, dan Kepuasan Masyarakat, yang masing-masing terdiri dari tiga pernyataan (item). Penilaian dilakukan menggunakan skala Likert 1–5, yang kemudian dijumlahkan dalam kolom masing-masing aspek. Total keseluruhan skor dari keempat aspek dihitung untuk mendapatkan gambaran umum tentang tingkat kepuasan tiap responden terhadap layanan yang diterima.

Berdasarkan total skor yang diperoleh, ditambahkan pula kolom "Kategori" sebagai klasifikasi manual atau aktual dari tingkat kepuasan masyarakat. Dalam hal ini, apabila total nilai responden ≤ 40 , maka dikategorikan sebagai "Tidak Puas", sedangkan jika > 40 , maka dikategorikan sebagai "Puas". Kategori ini merupakan label aktual yang dihasilkan langsung dari perhitungan terhadap data kuesioner dan akan digunakan sebagai acuan (ground truth) dalam pengujian model klasifikasi berbasis *Naive Bayes*. Dengan adanya label kategori ini, proses analisis data dapat diarahkan pada upaya memprediksi kepuasan atau kepuasan masyarakat secara otomatis melalui pendekatan machine learning.

Selanjutnya, data ini akan digunakan sebagai dataset latih dan uji dalam penerapan algoritma *Naive Bayes* untuk klasifikasi kepuasan. Hasil prediksi dari model akan dibandingkan dengan kategori aktual untuk menilai seberapa akurat model dalam mengenali pola kepuasan responden. Evaluasi kinerja model akan

dilakukan dengan menghitung akurasi, presisi, dan recall menggunakan confusion matrix. Langkah ini bertujuan untuk mengetahui sejauh mana model mampu mengklasifikasikan opini masyarakat secara tepat dan menjadi dasar untuk merekomendasikan perbaikan layanan berbasis data nyata.

3.3.2. Seleksi Data

Tahapan seleksi data merupakan proses penting dalam tahap awal pengolahan data, di mana peneliti melakukan pemilihan dan penyaringan terhadap data kuesioner yang telah dikumpulkan. Dalam penelitian ini, seleksi dilakukan dengan memilih kolom-kolom yang relevan, yaitu atribut Kualitas Pelayanan, Aksesibilitas Layanan, Transparansi Informasi, dan Kepuasan Masyarakat, serta kolom Kategori sebagai label kelas. Sementara itu, data yang tidak berkaitan langsung dengan proses klasifikasi seperti nama, usia, jenis kelamin, dan pekerjaan responden diabaikan agar tidak memengaruhi hasil analisis. Setiap atribut dihitung berdasarkan skor total dari tiga indikator per kategori, kemudian dijumlahkan untuk memperoleh skor keseluruhan. Skor total tersebut digunakan untuk menentukan klasifikasi tingkat kepuasan masyarakat ke dalam dua kategori, yaitu "Puas" jika nilai lebih dari 40 dan "Tidak Puas" jika nilai kurang dari atau sama dengan 40. Data yang telah terseleksi inilah yang kemudian digunakan sebagai input dalam proses klasifikasi menggunakan algoritma *Naive Bayes*, guna memprediksi kepuasan masyarakat secara lebih akurat dan terstruktur.

Tabel 3. 3. Data Sampel Penelitian

Nama	Kualitas Pelayanan	Akseibilitas Layanan	Transparansi Informasi	Kepuasan Masyarakat
Alfan	12	12	12	Puas
Amanda Srtk Br.Siregar	11	11	12	Puas
Andi Pandiangan	12	8	10	Puas
Anggik Caniago	15	15	15	Puas

Annisya Sepyanti	10	13	11	Puas
Aprilia Nurlita Dewi	12	12	12	Puas
Ayuu	12	12	12	Puas
Azurey	14	15	15	Puas
Berta Nglñ	11	13	10	Puas
Biang Sitompul	15	15	15	Puas
Bobby	15	15	15	Puas
Dedek	15	15	15	Puas
Devi Enjes Marbun	11	12	12	Puas
Devi Lestari	11	10	11	Puas
Devi Yunita	11	10	9	Tidak Puas
Dewi Kusuma	10	11	12	Puas
Donna Br Sinurat	11	10	11	Puas
Elpredo	9	9	9	Tidak Puas
Eprida Silaban	14	15	15	Puas
Erika Fatrecia Silaban	12	12	13	Puas
Erlita Pakpahan	10	13	9	Puas
Febriana Yolanda	11	10	12	Puas
Finky Surenda Ritonga	12	11	10	Puas
Flowerin Elsheva	9	12	12	Puas
Gracia Anggita Silaban	13	12	11	Puas
Henry Tio Dora Br Simanjuntak	11	10	10	Puas
Inayah	11	12	12	Puas
Indri Khairani Br Dalimunthe	15	15	12	Puas
Irna Ritongga	10	7	7	Tidak Puas
Jaini	12	12	12	Puas
Jojo Christine Simanjuntak	9	10	10	Tidak Puas
Kasmila	10	10	8	Tidak Puas
Kevin Sanjaya	13	10	9	Puas
Khairunnisa	13	12	11	Puas
Lanni Marlina Pakpahan	13	12	12	Puas
Lia	12	12	12	Tidak Puas
Luvfihasibuan	15	15	15	Puas
Marapuli	13	11	13	Puas
Mariani Nainggolan	10	10	11	Puas
Maryo Pasaribu	12	9	9	Tidak Puas
Maulidina Harahap	9	9	6	Tidak Puas
Minawaroh	9	9	9	Tidak Puas
Nike	10	12	12	Puas
Nora	15	13	15	Puas
Nur Aini	10	9	9	Tidak Puas
Nurjannah	12	13	13	Puas
Nurmaidha Hasibuan	12	12	12	Puas
Pudan	11	9	10	Puas
Putra Afrika	15	15	15	Puas
Raja Pasaribu	11	11	13	Puas
Ridho Ramadhan	12	11	12	Puas
Rina Nasution	11	11	10	Puas
Rini	9	10	7	Tidak Puas
Rio Siahaan	9	9	10	Tidak Puas
Ririn Hasibuan	8	9	8	Tidak Puas
Riska Ulanda	12	12	12	Puas
Rizky	12	12	14	Puas

Rizky Pratama Harahap	11	13	10	Puas
Rosmawati Angelina Hutabarat	11	11	10	Puas
Roy Virnando Situmorang	10	15	12	Puas
Rudi Nainggolan	9	8	10	Puas
Saifuljamil Pasaribu	15	15	15	Puas
Sasan	12	10	9	Tidak Puas
Sayba	13	12	13	Puas
Septi Ananda	11	12	13	Puas
Sintia Ninggolan	13	12	11	Puas
Siswadi	9	9	9	Tidak Puas
Siti	12	12	12	Puas
Siti Hasanah Rambe	15	15	15	Puas
Siti Nur	15	15	15	Puas
Sri Berganti	6	9	10	Tidak Puas
Sri Indah Sari	9	11	6	Tidak Puas
Sri Rahmatika	13	12	12	Puas
Sriwati Hasibuan	9	9	9	Tidak Puas
Tasya Putri	12	12	13	Puas
Vira	12	12	12	Puas
Wahyuli Pasaribu	12	14	12	Puas
Wima	12	6	7	Tidak Puas
Windi Listiana	12	10	12	Puas
Yesi	10	13	15	Puas
Zahra Saragi	12	13	13	Puas
Zahrotussaidah	12	13	13	Puas
Zeylvanus Se`Ga	12	9	11	Puas

Pada tabel di atas merupakan hasil dari seleksi data yang dilakukan terhadap seluruh data kuesioner yang telah terkumpul sebelumnya. Seleksi ini bertujuan untuk menyaring dan hanya mempertahankan atribut-atribut yang relevan dengan proses klasifikasi tingkat kepuasan masyarakat terhadap layanan rehabilitasi sosial. Dari data lengkap kuesioner, yang digunakan dalam penelitian ini hanyalah empat atribut utama yaitu: Nama, Kualitas Pelayanan, Aksesibilitas Layanan, Transparansi Informasi, dan Kategori Kepuasan sebagai label. Setiap nilai pada atribut layanan merupakan hasil penjumlahan dari tiga indikator dalam masing-masing aspek, sehingga memberikan representasi nilai total persepsi responden terhadap setiap dimensi layanan.

Variabel lain seperti usia, jenis kelamin, pekerjaan, serta nilai-nilai per item indikator tidak lagi ditampilkan dalam tabel ini karena dianggap tidak berkontribusi langsung terhadap proses klasifikasi dalam algoritma yang digunakan, yaitu *Naive Bayes*. Fokus penelitian diarahkan pada data numerik yang sudah teragregasi untuk masing-masing atribut, yang lebih sederhana dan efisien untuk pengolahan model machine learning. Label “Kategori” yang berisi klasifikasi "Puas" atau "Tidak Puas" juga merupakan hasil klasifikasi manual berdasarkan skor total yang dihitung sebelumnya. Data ini akan digunakan untuk pelatihan dan pengujian model klasifikasi, serta menjadi dasar dalam mengevaluasi performa model melalui pengukuran akurasi, presisi, dan recall.

3.3.3. Metode *Naive Bayes*

Naive Bayes merupakan metode klasifikasi yang menggunakan prinsip dasar dari Teorema Bayes, dengan anggapan bahwa setiap atribut atau fitur dalam data saling tidak bergantung satu sama lain. Algoritma ini bekerja dengan cara menghitung peluang kemunculan setiap kelas berdasarkan kombinasi nilai atribut yang dimiliki oleh suatu data. Setelah probabilitas dari masing-masing kelas dihitung, algoritma akan memilih kelas dengan nilai probabilitas terbesar sebagai hasil klasifikasinya.

1. Analisis Data

Analisis data merupakan tahapan penting dalam penelitian ini yang bertujuan untuk memahami pola dan hubungan antara variabel-variabel yang telah ditentukan, yaitu kualitas pelayanan, aksesibilitas layanan, dan transparansi informasi terhadap tingkat kepuasan masyarakat terhadap layanan rehabilitasi

sosial. Pada tahap ini, data yang telah melalui proses seleksi dan validasi akan diolah menggunakan metode klasifikasi *Naive Bayes* untuk memprediksi kategori kepuasan, apakah responden merasa puas atau tidak puas. Analisis dilakukan untuk melihat kontribusi masing-masing atribut dalam menentukan kepuasan masyarakat, serta untuk mengevaluasi sejauh mana model dapat mengenali pola dari data yang ada. Hasil dari analisis ini diharapkan dapat memberikan gambaran yang akurat mengenai faktor-faktor utama yang memengaruhi persepsi masyarakat, dan menjadi landasan dalam merumuskan strategi peningkatan mutu pelayanan di Dinas Sosial Rantauprapat.

Tabel 3. 4. Data Hasil Kuesioner

Nama	Kualitas Pelayanan	Akseibilitas Layanan	Transparansi Informasi	Kepuasan Masyarakat
Alfan	12	12	12	Puas
Amanda Srtk Br.Siregar	11	11	12	Puas
Andi Pandiangan	12	8	10	Puas
Anggik Caniago	15	15	15	Puas
Annisya Sepyanti	10	13	11	Puas
Aprilia Nurlita Dewi	12	12	12	Puas
Ayuu	12	12	12	Puas
Azurey	14	15	15	Puas
Berta Nglñ	11	13	10	Puas
Biang Sitompul	15	15	15	Puas
Bobby	15	15	15	Puas
Dedek	15	15	15	Puas
Devi Enjes Marbun	11	12	12	Puas
Devi Lestari	11	10	11	Puas
Devi Yunita	11	10	9	Tidak Puas
Dewi Kusuma	10	11	12	Puas
Donna Br Sinurat	11	10	11	Puas
Elpredo	9	9	9	Tidak Puas
Eprida Silaban	14	15	15	Puas
Erika Fatrecia Silaban	12	12	13	Puas
Erlita Pakpahan	10	13	9	Puas
Febriana Yolanda	11	10	12	Puas
Finky Surenda Ritonga	12	11	10	Puas
Flowerin Elsheva	9	12	12	Puas
Gracia Anggita Silaban	13	12	11	Puas
Heny Tio Dora Br Simanjuntak	11	10	10	Puas
Inayah	11	12	12	Puas
Indri Khairani Br Dalimunthe	15	15	12	Puas
Irna Ritongga	10	7	7	Tidak Puas

Jaini	12	12	12	Puas
Jojo Christine Simanjuntak	9	10	10	Tidak Puas
Kasmila	10	10	8	Tidak Puas
Kevin Sanjaya	13	10	9	Puas
Khairunnisa	13	12	11	Puas
Lanni Marlina Pakpahan	13	12	12	Puas
Lia	12	12	12	Tidak Puas
Luvfihasibuan	15	15	15	Puas
Marapuli	13	11	13	Puas
Mariani Nainggolan	10	10	11	Puas
Maryo Pasaribu	12	9	9	Tidak Puas
Maulidina Harahap	9	9	6	Tidak Puas
Minawaroh	9	9	9	Tidak Puas
Nike	10	12	12	Puas
Nora	15	13	15	Puas
Nur Aini	10	9	9	Tidak Puas
Nurjannah	12	13	13	Puas
Nurmaidha Hasibuan	12	12	12	Puas
Pudan	11	9	10	Puas
Putra Afrika	15	15	15	Puas
Raja Pasaribu	11	11	13	Puas
Ridho Ramadhan	12	11	12	Puas
Rina Nasution	11	11	10	Puas
Rini	9	10	7	Tidak Puas
Rio Siahaan	9	9	10	Tidak Puas
Ririn Hasibuan	8	9	8	Tidak Puas
Riska Ulanda	12	12	12	Puas
Rizky	12	12	14	Puas
Rizky Pratama Harahap	11	13	10	Puas
Rosmawati Angelina Hutabarat	11	11	10	Puas
Roy Virnando Situmorang	10	15	12	Puas
Rudi Nainggolan	9	8	10	Puas
Saifuljamil Pasaribu	15	15	15	Puas
Sasan	12	10	9	Tidak Puas
Sayba	13	12	13	Puas
Septi Ananda	11	12	13	Puas
Sintia Ninggolan	13	12	11	Puas
Siswadi	9	9	9	Tidak Puas
Siti	12	12	12	Puas
Siti Hasanah Rambe	15	15	15	Puas
Siti Nur	15	15	15	Puas
Sri Berganti	6	9	10	Tidak Puas
Sri Indah Sari	9	11	6	Tidak Puas
Sri Rahmatika	13	12	12	Puas
Sriwati Hasibuan	9	9	9	Tidak Puas
Tasya Putri	12	12	13	Puas
Vira	12	12	12	Puas
Wahyuli Pasaribu	12	14	12	Puas
Wima	12	6	7	Tidak Puas
Windi Listiana	12	10	12	Puas
Yesi	10	13	15	Puas
Zahra Saragi	12	13	13	Puas
Zahrotussaidah	12	13	13	Puas

Zeylvanus Se'Ga	12	9	11	Puas
-----------------	----	---	----	------

Pada tabel di atas merupakan data hasil seleksi yang digunakan dalam tahap analisis data penelitian ini. Tabel tersebut berisi nilai total dari tiga atribut utama, yaitu Kualitas Pelayanan, Aksesibilitas Layanan, dan Transparansi Informasi, yang dikumpulkan melalui kuesioner masyarakat terkait layanan rehabilitasi sosial Dinas Sosial Rantauprapat. Setiap nilai merepresentasikan penjumlahan dari tiga indikator pada masing-masing atribut, sehingga memberikan gambaran yang ringkas dan jelas terhadap persepsi responden. Kolom Kategori Kepuasan Masyarakat berfungsi sebagai label klasifikasi yang akan menjadi target dalam pengolahan data menggunakan algoritma *Naive Bayes*. Data dalam tabel inilah yang kemudian diolah untuk mempelajari hubungan antar variabel, membentuk model klasifikasi, serta mengevaluasi performa model melalui perhitungan akurasi, presisi, dan recall.

2. Transformasi Data

Transformasi data merupakan tahap penting dalam proses pengolahan data yang bertujuan untuk menyesuaikan struktur atau format data agar selaras dengan kebutuhan algoritma yang digunakan. Dalam penelitian ini, proses transformasi dilakukan dengan cara mengubah data numerik dari atribut Kualitas Pelayanan, Aksesibilitas Layanan, dan Transparansi Informasi menjadi bentuk kategori diskrit, seperti "Baik" dan "Tidak Baik". Transformasi ini diperlukan karena metode *Naive Bayes* yang digunakan dalam klasifikasi lebih optimal bekerja pada data kategorikal dibandingkan numerikal. Rentang nilai untuk masing-masing kategori ditentukan berdasarkan distribusi data, sehingga hasil klasifikasi dapat lebih akurat dan

mencerminkan karakteristik asli dari respon masyarakat terhadap layanan rehabilitasi yang diberikan oleh Dinas Sosial Rantauprapat.

Tabel 3. 5. Data Konversi

Atribut	Partisi	Nilai
Kualitas Pelayanan	Baik	> 11
	Tidak Baik	≤ 10
Aksesibilitas Layanan	Baik	> 11
	Tidak Baik	≤ 10
Transparansi Informasi	Baik	> 11
	Tidak Baik	≤ 10

Tabel di atas menunjukkan hasil transformasi data dari bentuk numerikal ke kategorikal yang digunakan dalam penelitian ini. Proses ini dilakukan untuk menyesuaikan data dengan kebutuhan algoritma *Naive Bayes*, yang lebih efektif bekerja dengan data kategorikal. Masing-masing atribut utama, yaitu Kualitas Pelayanan, Aksesibilitas Layanan, dan Transparansi Informasi, dibagi ke dalam dua kategori, yaitu "Baik" dan "Tidak Baik", berdasarkan nilai total responden pada setiap indikator. Jika total nilai pada suatu atribut kurang dari atau sama dengan 10, maka diklasifikasikan sebagai "Baik", sedangkan jika lebih dari 11, maka dikategorikan sebagai "Tidak Baik". Klasifikasi ini bertujuan untuk menyederhanakan pemodelan dan membantu dalam mengidentifikasi pengaruh masing-masing variabel terhadap kepuasan masyarakat.

Tabel 3. 6. Data Hasil Konversi

Nama	Kualitas Pelayanan	Aksesibilitas Layanan	Transparansi Informasi	Kepuasan Masyarakat
Alfan	Baik	Baik	Baik	Puas
Amanda srtk br.siregar	Baik	Baik	Baik	Puas
Andi pandiangan	Baik	Tidak Baik	Tidak Baik	Puas
Anggik Caniago	Baik	Baik	Baik	Puas
Annisya Sepyanti	Tidak Baik	Baik	Baik	Puas
Aprilia nurlita dewi	Baik	Baik	Baik	Puas
Ayuu	Baik	Baik	Baik	Puas
Azurey	Baik	Baik	Baik	Puas

Berta ngl	Baik	Baik	Tidak Baik	Puas
Biang sitompul	Baik	Baik	Baik	Puas
Bobby	Baik	Baik	Baik	Puas
Dedek	Baik	Baik	Baik	Puas
Devi enjes marbun	Baik	Baik	Baik	Puas
Devi lestari	Baik	Tidak Baik	Baik	Puas
Devi yunita	Baik	Tidak Baik	Tidak Baik	Tidak Puas
Dewi kusuma	Tidak Baik	Baik	Baik	Puas
Donna Br Sinurat	Baik	Tidak Baik	Baik	Puas
Elpredo	Tidak Baik	Tidak Baik	Tidak Baik	Tidak Puas
Eprida silaban	Baik	Baik	Baik	Puas
Erika Fatrecia silaban	Baik	Baik	Baik	Puas
Erlita pakpahan	Tidak Baik	Baik	Tidak Baik	Puas
Febriana yolanda	Baik	Tidak Baik	Baik	Puas
Finky Surenda Ritonga	Baik	Baik	Tidak Baik	Puas
Flowerin Elsheva	Tidak Baik	Baik	Baik	Puas
Gracia Anggita Silaban	Baik	Baik	Baik	Puas
Heny Tio Dora Br Simanjuntak	Baik	Tidak Baik	Tidak Baik	Puas
Inayah	Baik	Baik	Baik	Puas
Indri Khairani br dalimunthe	Baik	Baik	Baik	Puas
Irna Ritongga	Tidak Baik	Tidak Baik	Tidak Baik	Tidak Puas
Jaini	Baik	Baik	Baik	Puas
Jojo Christine simanjuntak	Baik	Tidak Baik	Tidak Baik	Tidak Puas
Kasmila	Tidak Baik	Tidak Baik	Baik	Tidak Puas
Kevin Sanjaya	Baik	Tidak Baik	Tidak Baik	Puas
Khairunnisa	Baik	Baik	Baik	Puas
Lanni Marlina Pakpahan	Baik	Baik	Baik	Puas
Lia	Baik	Baik	Baik	Tidak Puas
Luvfihasibuan	Baik	Baik	Baik	Puas
Marapuli	Baik	Baik	Baik	Puas
Mariani nainggolan	Tidak Baik	Tidak Baik	Baik	Puas
Maryo Pasaribu	Baik	Tidak Baik	Tidak Baik	Tidak Puas
Maulidina Harahap	Tidak Baik	Tidak Baik	Tidak Baik	Tidak Puas
Minawaroh	Tidak Baik	Tidak Baik	Tidak Baik	Tidak Puas
Nike	Tidak Baik	Baik	Baik	Puas
Nora	Baik	Baik	Baik	Puas
Nur Aini	Tidak Baik	Tidak Baik	Tidak Baik	Tidak Puas
Nurjannah	Baik	Baik	Baik	Puas
Nurmaidha Hasibuan	Baik	Baik	Baik	Puas
Pudan	Baik	Tidak Baik	Tidak Baik	Puas
Putra Afrika	Baik	Baik	Baik	Puas
Raja Pasaribu	Baik	Baik	Baik	Puas
Ridho Ramadhan	Baik	Baik	Baik	Puas
Rina Nasution	Baik	Baik	Tidak Baik	Puas
Rini	Tidak Baik	Tidak Baik	Baik	Tidak Puas
Rio Siahaan	Baik	Tidak Baik	Tidak Baik	Tidak Puas
Ririn Hasibuan	Tidak Baik	Baik	Tidak Baik	Tidak Puas
Riska Ulanda	Baik	Baik	Baik	Puas
Rizky	Baik	Baik	Baik	Puas
Rizky Pratama Harahap	Baik	Baik	Tidak Baik	Puas
Rosmawati Angelina Hutabarat	Baik	Baik	Tidak Baik	Puas
Roy Virnando Situmorang	Tidak Baik	Baik	Baik	Puas
Rudi Nainggolan	Tidak Baik	Tidak Baik	Tidak Baik	Puas

Saifuljamil Pasaribu	Baik	Baik	Baik	Puas
Sasan	Baik	Tidak Baik	Tidak Baik	Tidak Puas
Sayba	Baik	Baik	Baik	Puas
Septi Ananda	Baik	Baik	Baik	Puas
Sintia Ninggolan	Baik	Baik	Baik	Puas
Siswadi	Tidak Baik	Tidak Baik	Tidak Baik	Tidak Puas
Siti	Baik	Baik	Baik	Puas
Siti Hasanah Rambe	Baik	Baik	Baik	Puas
Siti Nur	Baik	Baik	Baik	Puas
Sri Berganti	Tidak Baik	Tidak Baik	Tidak Baik	Tidak Puas
Sri Indah Sari	Tidak Baik	Baik	Tidak Baik	Tidak Puas
Sri rahmatika	Baik	Baik	Baik	Puas
Sriwati hasibuan	Tidak Baik	Tidak Baik	Tidak Baik	Tidak Puas
Tasya Putri	Baik	Baik	Baik	Puas
Vira	Baik	Baik	Baik	Puas
Wahyuli Pasaribu	Baik	Baik	Baik	Puas
Wima	Baik	Tidak Baik	Tidak Baik	Tidak Puas
Windi Listiana	Baik	Tidak Baik	Baik	Puas
Yesi	Tidak Baik	Baik	Baik	Puas
Zahra Saragi	Baik	Baik	Baik	Puas
Zahrotussaidah	Baik	Baik	Baik	Puas
Zeylvanus Se`ga	Baik	Tidak Baik	Baik	Puas

Tabel di atas merupakan data sampel yang telah digunakan dalam penelitian ini dan telah melalui proses transformasi data dari bentuk numerikal ke kategorikal. Setiap atribut utama—Kualitas Pelayanan, Aksesibilitas Layanan, dan Transparansi Informasi—telah dikonversi ke dalam dua kategori, yaitu "Baik" dan "Tidak Baik", berdasarkan batas nilai tertentu yang telah ditentukan sebelumnya. Transformasi ini dilakukan agar data lebih sesuai dengan karakteristik algoritma klasifikasi seperti *Naive Bayes*, yang dapat bekerja lebih optimal dengan data kategorikal. Kategori "Kepuasan Masyarakat" yang merupakan label atau kelas target juga telah disesuaikan berdasarkan hasil aktual dari kuesioner. Data dalam tabel ini selanjutnya digunakan sebagai input utama dalam proses pelatihan dan pengujian model untuk mengidentifikasi pola dan menentukan tingkat kepuasan masyarakat terhadap layanan yang diberikan.

3. Pembagian Data

Dalam penelitian ini, data yang telah dikumpulkan kemudian dibagi menjadi dua kelompok utama, yaitu data training dan data testing. Pembagian ini dilakukan untuk memisahkan data yang digunakan dalam proses pelatihan model dengan data yang digunakan untuk menguji keakuratan hasil prediksi dari model yang telah dibentuk. Sebanyak 33 data pertama digunakan sebagai data training, yaitu data yang akan dianalisis oleh algoritma *Naive Bayes* untuk mempelajari pola hubungan antara atribut-atribut input seperti Kualitas Pelayanan, Aksesibilitas Layanan, dan Transparansi Informasi dengan kategori Kepuasan Masyarakat. Sementara itu, 50 data sisanya dijadikan sebagai data testing, yang bertujuan untuk menguji kemampuan model dalam mengklasifikasikan data baru yang belum pernah dikenali sebelumnya. Meskipun pembagian ini tidak mengikuti rasio umum, penyesuaian dilakukan berdasarkan ketersediaan data yang ada, agar tetap memberikan hasil evaluasi yang relevan terhadap performa model klasifikasi yang dikembangkan.

Tabel 3. 7. Data Training Penelitian

Nama	Kualitas Pelayanan	Akseibilitas Layanan	Transparansi Informasi	Kepuasan Masyarakat
Alfan	Baik	Baik	Baik	Puas
Amanda srtk br.siregar	Baik	Baik	Baik	Puas
Andi pandiangan	Baik	Tidak Baik	Tidak Baik	Puas
Anggik Caniago	Baik	Baik	Baik	Puas
Annisya Sepyanti	Tidak Baik	Baik	Baik	Puas
Aprilia nurlita dewi	Baik	Baik	Baik	Puas
Ayuu	Baik	Baik	Baik	Puas
Azurey	Baik	Baik	Baik	Puas
Berta ngln	Baik	Baik	Tidak Baik	Puas
Devi yunita	Baik	Tidak Baik	Tidak Baik	Tidak Puas
Dewi kusuma	Tidak Baik	Baik	Baik	Puas
Donna Br Sinurat	Baik	Tidak Baik	Baik	Puas
Irna Ritongga	Tidak Baik	Tidak Baik	Tidak Baik	Tidak Puas
Jaini	Baik	Baik	Baik	Puas
Jojo Christine simanjuntak	Baik	Tidak Baik	Tidak Baik	Tidak Puas
Kasmila	Tidak Baik	Tidak Baik	Baik	Tidak Puas

Minawaroh	Tidak Baik	Tidak Baik	Tidak Baik	Tidak Puas
Nike	Tidak Baik	Baik	Baik	Puas
Nora	Baik	Baik	Baik	Puas
Nur Aini	Tidak Baik	Tidak Baik	Tidak Baik	Tidak Puas
Nurjannah	Baik	Baik	Baik	Puas
Nurmaidha Hasibuan	Baik	Baik	Baik	Puas
Pudan	Baik	Tidak Baik	Tidak Baik	Puas
Putra Afrika	Baik	Baik	Baik	Puas
Raja Pasaribu	Baik	Baik	Baik	Puas
Ridho Ramadhan	Baik	Baik	Baik	Puas
Rina Nasution	Baik	Baik	Tidak Baik	Puas
Rini	Tidak Baik	Tidak Baik	Baik	Tidak Puas
Rio Siahaan	Baik	Tidak Baik	Tidak Baik	Tidak Puas
Ririn Hasibuan	Tidak Baik	Baik	Tidak Baik	Tidak Puas
Riska Ulanda	Baik	Baik	Baik	Puas
Rizky	Baik	Baik	Baik	Puas
Rizky Pratama Harahap	Baik	Baik	Tidak Baik	Puas

Tabel di atas merupakan data training yang digunakan dalam penelitian ini, yang terdiri dari 33 data responden. Data tersebut telah melalui proses transformasi dari numerik ke kategorikal, dengan tiga atribut utama yaitu Kualitas Pelayanan, Aksesibilitas Layanan, dan Transparansi Informasi, serta satu atribut target yaitu Kepuasan Masyarakat. Setiap baris mewakili hasil penilaian individu terhadap layanan rehabilitasi sosial yang diberikan oleh Dinas Sosial Rantauprapat. Kategori "Puas" dan "Tidak Puas" pada kolom Kepuasan Masyarakat digunakan sebagai label kelas dalam proses pelatihan model klasifikasi menggunakan algoritma *Naive Bayes*. Model ini akan mempelajari pola hubungan antar atribut dari data training tersebut agar dapat digunakan untuk memprediksi kepuasan masyarakat pada data testing yang belum dikenali sebelumnya.

Tabel 3. 8. Data Testing Penelitian

Nama	Kualitas Pelayanan	Akseibilitas Layanan	Transparansi Informasi	Kepuasan Masyarakat
Biang sitompul	Baik	Baik	Baik	Puas
Bobby	Baik	Baik	Baik	Puas
Dedek	Baik	Baik	Baik	Puas
Devi enjes marbun	Baik	Baik	Baik	Puas
Devi lestari	Baik	Tidak Baik	Baik	Puas
Elpredo	Tidak Baik	Tidak Baik	Tidak Baik	Tidak Puas
Eprida silaban	Baik	Baik	Baik	Puas

Erika Fatrecia silaban	Baik	Baik	Baik	Puas
Erlita pakpahan	Tidak Baik	Baik	Tidak Baik	Puas
Febriana yolanda	Baik	Tidak Baik	Baik	Puas
Finky Surenda Ritonga	Baik	Baik	Tidak Baik	Puas
Flowerin Elsheva	Tidak Baik	Baik	Baik	Puas
Gracia Anggita Silaban	Baik	Baik	Baik	Puas
Henry Tio Dora Br Simanjuntak	Baik	Tidak Baik	Tidak Baik	Puas
Inayah	Baik	Baik	Baik	Puas
Indri Khairani br dalimunthe	Baik	Baik	Baik	Puas
Kevin Sanjaya	Baik	Tidak Baik	Tidak Baik	Puas
Khairunnisa	Baik	Baik	Baik	Puas
Lanni Marlina Pakpahan	Baik	Baik	Baik	Puas
Lia	Baik	Baik	Baik	Tidak Puas
Luvfihasibuan	Baik	Baik	Baik	Puas
Marapuli	Baik	Baik	Baik	Puas
Mariani nainggolan	Tidak Baik	Tidak Baik	Baik	Puas
Maryo Pasaribu	Baik	Tidak Baik	Tidak Baik	Tidak Puas
Maulidina Harahap	Tidak Baik	Tidak Baik	Tidak Baik	Tidak Puas
Rosmawati Angelina Hutabarat	Baik	Baik	Tidak Baik	Puas
Roy Virnando Situmorang	Tidak Baik	Baik	Baik	Puas
Rudi Nainggolan	Tidak Baik	Tidak Baik	Tidak Baik	Puas
Saifuljamil Pasaribu	Baik	Baik	Baik	Puas
Sasan	Baik	Tidak Baik	Tidak Baik	Tidak Puas
Sayba	Baik	Baik	Baik	Puas
Septi Ananda	Baik	Baik	Baik	Puas
Sintia Ninggolan	Baik	Baik	Baik	Puas
Siswadi	Tidak Baik	Tidak Baik	Tidak Baik	Tidak Puas
Siti	Baik	Baik	Baik	Puas
Siti Hasanah Rambe	Baik	Baik	Baik	Puas
Siti Nur	Baik	Baik	Baik	Puas
Sri Berganti	Tidak Baik	Tidak Baik	Tidak Baik	Tidak Puas
Sri Indah Sari	Tidak Baik	Baik	Tidak Baik	Tidak Puas
Sri rahmatika	Baik	Baik	Baik	Puas
Sriwati hasibuan	Tidak Baik	Tidak Baik	Tidak Baik	Tidak Puas
Tasya Putri	Baik	Baik	Baik	Puas
Vira	Baik	Baik	Baik	Puas
Wahyuli Pasaribu	Baik	Baik	Baik	Puas
Wima	Baik	Tidak Baik	Tidak Baik	Tidak Puas
Windi Listiana	Baik	Tidak Baik	Baik	Puas
Yesi	Tidak Baik	Baik	Baik	Puas
Zahra Saragi	Baik	Baik	Baik	Puas
Zahrotussaidah	Baik	Baik	Baik	Puas
Zeylvanus Se`ga	Baik	Tidak Baik	Baik	Puas

Tabel di atas merupakan data testing yang digunakan dalam penelitian ini, yang terdiri dari 50 data responden. Data ini telah melalui proses transformasi dari bentuk numerik menjadi kategorikal dan mencakup tiga atribut utama, yaitu Kualitas Pelayanan, Aksesibilitas Layanan, dan Transparansi Informasi, serta satu

atribut target yaitu Kepuasan Masyarakat. Data testing berperan penting dalam mengevaluasi performa model klasifikasi *Naive Bayes* yang sebelumnya telah dilatih menggunakan data training. Tujuannya adalah untuk mengetahui sejauh mana model mampu mengklasifikasikan data baru yang belum dikenali, berdasarkan pola-pola yang telah dipelajari sebelumnya. Dengan menggunakan data testing ini, peneliti dapat mengukur akurasi, presisi, dan kemampuan generalisasi model terhadap data yang bersifat independen, sehingga validitas dan keandalan hasil analisis dapat dipertanggungjawabkan.

4. Pembersihan Data

Transformasi data merupakan proses penting dalam tahap pra-pemrosesan data yang bertujuan untuk mengubah format, struktur, atau nilai data mentah menjadi bentuk yang lebih sesuai dan siap digunakan dalam proses analisis atau pelatihan model klasifikasi. Dalam penelitian ini, transformasi dilakukan dengan cara mengubah data numerik menjadi data kategorikal berdasarkan rentang nilai tertentu, sehingga atribut seperti Kualitas Pelayanan, Aksesibilitas Layanan, dan Transparansi Informasi diklasifikasikan menjadi dua kategori, yaitu "Baik" dan "Tidak Baik". Proses ini membantu algoritma *Naive Bayes* dalam memahami dan mengolah data dengan lebih optimal, karena algoritma ini lebih efektif bekerja dengan data dalam bentuk diskrit atau label.

Tabel 3. 9. Atribut Kualitas Pelayanan

Atribut	Partisi	Puas	Tidak Puas	P (Puas)	P (Tidak Puas)
Kualitas Pelayanan	Baik	21	3	21/24	3/9
	Tidak Baik	3	6	3/24	6/9
	Total	24	9	100%	100%

Tabel 3. 10. Atribut Aksesibilitas Layanan

Atribut	Partisi	Puas	Tidak Puas	P (Puas)	P (Tidak Puas)
Aksesibilitas Layanan	Baik	21	1	21/24	1/9
	Tidak Baik	3	8	3/24	8/9
	Total	24	9	100%	100%

Tabel 3. 11. Atribut Transparansi Informasi

Atribut	Partisi	Puas	Tidak Puas	P (Puas)	P (Tidak Puas)
Transparansi Informasi	Baik	20	2	20/24	2/9
	Tidak Baik	4	7	4/24	7/9
	Total	24	9	100%	100%

Tabel 3. 12. Atribut Kategori

Kategori		P (Puas) dan P (Tidak Puas)
Puas	24	24/33
Tidak Puas	9	9/33
Total	33	100%

5. Perhitungan Data

Pada perhitungan metode *Naïve Bayes* terdapat perhitungan yang dilakukan secara manual yaitu sebagai berikut.

Rumus Metode *Naïve Bayes*,
$$P(A | B) = \frac{P(B | A) P(A)}{P(B)}$$

Information:

A : hipotesis data A (kelas tertentu)

B : data dengan kelas yang tidak diketahui

$P(A | B)$: Probabilitas hipotesis berdasarkan kondisi B

$P(A)$: Kemungkinan hipotesis A

$P(B | A)$: Probabilitas B ketika kondisi A

$P(B)$: Probabilitas

Perhitungan

Pada perhitungan yang akan dilakukan, pertama penulis akan menghitung data Biang Sitompul. Adapun perhitungannya sebagai berikut.

$$P(\text{Kategori}) = P(\text{Kualitas Pelayanan}|\text{Baik}) \times P(\text{Akseibilitas Layanan}|\text{Baik}) \times P(\text{Transparansi Layanan}|\text{Baik}) \times P(\text{Kategori}|\text{Puas})$$

$$\begin{aligned} P(\text{Puas}) &= P(\text{Baik}|\text{Puas}) \times P(\text{Baik}|\text{Puas}) \times P(\text{Baik}|\text{Puas}) \times P(\text{Kategori}|\text{Puas}) \\ &= \left(\frac{21}{24}\right) \times \left(\frac{21}{24}\right) \times \left(\frac{20}{24}\right) \times \left(\frac{24}{33}\right) \\ &= 0,464015 \text{ (Nilai Puas)} \end{aligned}$$

$$\begin{aligned} P(\text{Tidak Puas}) &= P(\text{Baik}|\text{Tidak Puas}) \times P(\text{Baik}|\text{Tidak Puas}) \times P(\text{Baik}|\text{Tidak Puas}) \times P(\text{Kategori}|\text{Tidak Puas}) \\ &= \left(\frac{3}{9}\right) \times \left(\frac{1}{9}\right) \times \left(\frac{2}{9}\right) \times \left(\frac{9}{33}\right) \\ &= 0,002244 \text{ (Nilai Tidak Puas)} \end{aligned}$$

Hasil yang diperoleh dari data Biang Sitompul bahwasanya nilai puas yang diperoleh yaitu sebesar 0,464015 dan untuk nilai tidak puas yang diperoleh yaitu sebesar 0,002244. Jadi dari hasil yang sudah diperoleh bahwa Biang Sitompul Puas pada Layanan Rehabilitasi Dinas Sosial Rantaupraptat. Untuk sisa hasil selanjutnya akan dipaparkan dalam bentuk tabel dibawah ini.

Tabel 3. 13. Hasil Perhitungan

Nama	Kualitas Pelayanan	Akseibilitas Layanan	Transparansi Informasi	Kepuasan Masyarakat
Biang sitompul	Baik	Baik	Baik	Puas
Bobby	Baik	Baik	Baik	Puas
Dedek	Baik	Baik	Baik	Puas
Devi enjes marbun	Baik	Baik	Baik	Puas
Devi lestari	Baik	Tidak Baik	Baik	Puas
Elpredo	Tidak Baik	Tidak Baik	Tidak Baik	Tidak Puas
Eprida silaban	Baik	Baik	Baik	Puas
Erika Fatrecia silaban	Baik	Baik	Baik	Puas
Erlita pakpahan	Tidak Baik	Baik	Tidak Baik	Puas
Febriana yolanda	Baik	Tidak Baik	Baik	Puas
Finky Surenda Ritonga	Baik	Baik	Tidak Baik	Puas
Flowerin Elsheva	Tidak Baik	Baik	Baik	Puas
Gracia Anggita Silaban	Baik	Baik	Baik	Puas
Heny Tio Dora Br Simanjuntak	Baik	Tidak Baik	Tidak Baik	Puas
Inayah	Baik	Baik	Baik	Puas
Indri Khairani br dalimunthe	Baik	Baik	Baik	Puas
Kevin Sanjaya	Baik	Tidak Baik	Tidak Baik	Puas

Khairunnisa	Baik	Baik	Baik	Puas
Lanni Marlina Pakpahan	Baik	Baik	Baik	Puas
Lia	Baik	Baik	Baik	Puas
Luvfihasibuan	Baik	Baik	Baik	Puas
Marapuli	Baik	Baik	Baik	Puas
Mariani nainggolan	Tidak Baik	Tidak Baik	Baik	Puas
Maryo Pasaribu	Baik	Tidak Baik	Tidak Baik	Tidak Puas
Maulidina Harahap	Tidak Baik	Tidak Baik	Tidak Baik	Tidak Puas
Rosmawati Angelina Hutabarat	Baik	Baik	Tidak Baik	Puas
Roy Virnando Situmorang	Tidak Baik	Baik	Baik	Puas
Rudi Nainggolan	Tidak Baik	Tidak Baik	Tidak Baik	Tidak Puas
Saifuljamil Pasaribu	Baik	Baik	Baik	Puas
Sasan	Baik	Tidak Baik	Tidak Baik	Tidak Puas
Sayba	Baik	Baik	Baik	Puas
Septi Ananda	Baik	Baik	Baik	Puas
Sintia Ninggolan	Baik	Baik	Baik	Puas
Siswadi	Tidak Baik	Tidak Baik	Tidak Baik	Tidak Puas
Siti	Baik	Baik	Baik	Puas
Siti Hasanah Rambe	Baik	Baik	Baik	Puas
Siti Nur	Baik	Baik	Baik	Puas
Sri Berganti	Tidak Baik	Tidak Baik	Tidak Baik	Tidak Puas
Sri Indah Sari	Tidak Baik	Baik	Tidak Baik	Tidak Puas
Sri rahmatika	Baik	Baik	Baik	Puas
Sriwati hasibuan	Tidak Baik	Tidak Baik	Tidak Baik	Tidak Puas
Tasya Putri	Baik	Baik	Baik	Puas
Vira	Baik	Baik	Baik	Puas
Wahyuli Pasaribu	Baik	Baik	Baik	Puas
Wima	Baik	Tidak Baik	Tidak Baik	Tidak Puas
Windi Listiana	Baik	Tidak Baik	Baik	Puas
Yesi	Tidak Baik	Baik	Baik	Puas
Zahra Saragi	Baik	Baik	Baik	Puas
Zahrotussaidah	Baik	Baik	Baik	Puas
Zeylvanus Se`ga	Baik	Tidak Baik	Baik	Puas

6. Evaluasi Metode

Evaluasi metode *Naive Bayes* pada penelitian ini menggunakan Confusion Matrix. Hal ini dilakukan agar nantinya dapat diperoleh hasil akurasi presisi dan recall dari Confusion Matrix. Hasil tersebut yang nantinya akan memberikan nilai efektivitas dari metode yang digunakan.

Tabel 3. 14. Evaluasi Model
Predicted

		Predicted		Σ
		Puas	Tidak Puas	
Actual	Puas	39	1	40
	Tidak Puas	1	9	10

$$\Sigma \quad 40 \quad 10 \quad 50$$

Hasil True Positive (TP) adalah 39. True Negative (TN) adalah 9, False Positive (FP) adalah 1 dan False Negative (FN) adalah 1. Maka Nilai akurasi, presisi dan recall adalah sebagai berikut:

$$\text{Accuracy} = \frac{39+9}{39+9+1+1} \times 100\% \quad \text{Then the Accuracy value} = 96\%$$

$$\text{Presisi} = \frac{39}{39+1} \times 100\% \quad \text{Then the Precision value} = 97\%$$

$$\text{Recall} = \frac{39}{39+1} \times 100\% \quad \text{Then the Recall value} = 97\%$$

Evaluasi terhadap performa metode *Naive Bayes* dalam penelitian ini dilakukan menggunakan pendekatan Confusion Matrix, yang bertujuan untuk mengetahui efektivitas klasifikasi model dalam mengelompokkan data ke dalam kategori Puas dan Tidak Puas. Berdasarkan confusion matrix yang dihasilkan, terdapat 39 data diklasifikasikan dengan benar sebagai "Puas" (True Positive), 9 data diklasifikasikan dengan benar sebagai "Tidak Puas" (True Negative), sementara masing-masing terdapat 1 kesalahan klasifikasi, yakni False Positive (dinyatakan "Puas" padahal "Tidak Puas") dan False Negative (dinyatakan "Tidak Puas" padahal "Puas"). Hasil ini menunjukkan bahwa model memiliki kemampuan klasifikasi yang sangat baik terhadap data testing yang digunakan.

Berdasarkan nilai-nilai tersebut, diperoleh hasil evaluasi berupa akurasi sebesar 96%, yang berarti bahwa dari 50 data testing, sebanyak 48 data berhasil diklasifikasikan dengan benar oleh model. Selain itu, nilai presisi sebesar 97% menunjukkan bahwa model sangat jarang melakukan kesalahan dalam

mengklasifikasikan data ke dalam kategori Puas. Begitu juga dengan nilai recall sebesar 97%, yang menggambarkan bahwa model sangat sensitif dalam mendeteksi data yang benar-benar Puas. Dengan nilai-nilai evaluasi yang tinggi ini, dapat disimpulkan bahwa metode *Naive Bayes* yang digunakan dalam penelitian ini sangat efektif dan andal dalam melakukan klasifikasi tingkat kepuasan masyarakat terhadap layanan rehabilitasi sosial Dinas Sosial Rantauprapat.