

BAB IV IMPLEMENTASI SISTEM

4.1. Implementasi Sistem

Implementasi sistem dalam penelitian ini diawali dengan pemanfaatan atribut-atribut utama yang dianggap berpengaruh dalam menentukan kelayakan penerima bantuan BPJS, yaitu usia, pekerjaan, pendapatan per bulan, jumlah tanggungan, dan status tempat tinggal. Data tersebut diperoleh melalui hasil wawancara langsung kepada masyarakat di wilayah Rantauprapat. Tahap berikutnya adalah pembersihan dan transformasi data agar sesuai dengan karakteristik metode klasifikasi yang digunakan, yakni Naïve Bayes. Data yang awalnya masih berupa nilai mentah diolah menjadi bentuk kategorikal, misalnya dengan melabelkan hasil klasifikasi ke dalam dua kategori utama, yaitu “Dapat” dan “Tidak Dapat”.



RapidMiner

Proses ini mencakup penyesuaian nilai pendapatan, jumlah tanggungan, serta status rumah agar dapat dikelompokkan secara sistematis. Melalui analisis ini, pola-pola hubungan antar atribut dapat diidentifikasi, seperti apakah pendapatan yang rendah selalu berkorelasi dengan kemungkinan menerima bantuan, atau apakah status rumah yang menumpang atau sewa menjadi faktor yang berpengaruh signifikan. Hasil dari analisis ini kemudian digunakan untuk membangun model klasifikasi yang mampu memprediksi status penerimaan bantuan masyarakat secara akurat dan objektif, berdasarkan atribut yang tersedia.

4.2. Persiapan Data

Persiapan data merupakan tahap awal yang sangat penting dalam proses pembangunan model klasifikasi menggunakan algoritma Naïve Bayes. Pada tahap ini, data mentah yang telah dikumpulkan dari responden dipersiapkan agar siap diolah dan dianalisis. Proses persiapan mencakup penggabungan data dari berbagai sumber, pengecekan kelengkapan data, serta penyesuaian format agar seragam dan sesuai dengan kebutuhan sistem. Selain itu, dilakukan pula identifikasi terhadap atribut yang relevan, seperti usia, pekerjaan, pendapatan per bulan, jumlah tanggungan, dan status rumah, untuk memastikan hanya data yang berkaitan langsung dengan tujuan klasifikasi yang digunakan. Dengan adanya persiapan data yang baik, maka proses selanjutnya seperti preprocessing, pelatihan model, hingga pengujian dapat berjalan lebih efektif dan menghasilkan model klasifikasi yang akurat.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
26		Nama	Usia	Pekerjaan	Pendapatan/Bulan	Jumlah Tanggungan	Status Rumah	Label						
27		D1	36	Kuli Pasar	Rp 850,000.00	3	Sewa	Dapat						
28		D2	35	PNS	Rp 850,000.01	1	Milik Sendiri	Tidak Dapat						
29		D3	30	Tidak Bekerja	Rp 850,000.02	4	Menumpang	Dapat						
30		D4	28	Pedagang	Rp 850,000.03	2	Sewa	Dapat						
31		D5	36	Buruh	Rp 500,000.00	4	Menumpang	Dapat						
32		D6	34	Tidak Bekerja	Rp 400,000.00	4	Menumpang	Dapat						
33		D7	42	Tidak Bekerja	Rp 350,000.00	3	Menumpang	Dapat						
34		D8	40	Tidak Bekerja	Rp 400,000.00	5	Menumpang	Dapat						
35		D9	21	Buruh	Rp 900,000.00	3	Sewa	Dapat						
36		D10	52	Buruh	Rp 500,000.00	4	Menumpang	Dapat						
37		D11	37	PNS	Rp 2,600,000.00	1	Milik Sendiri	Tidak Dapat						
38		D12	66	Tidak Bekerja	Rp 500,000.00	3	Sewa	Dapat						
39		D13	63	Tidak Bekerja	Rp 300,000.00	4	Menumpang	Dapat						
40		D14	33	Buruh	Rp 850,000.00	3	Menumpang	Dapat						
41		D15	32	Karyawan Swasta	Rp 2,200,000.00	0	Milik Sendiri	Tidak Dapat						
42		D16	70	Tidak Bekerja	Rp 250,000.00	4	Menumpang	Dapat						
43		D17	63	Petani	Rp 700,000.00	4	Sewa	Dapat						
44		D18	35	Pedagang	Rp 2,000,000.00	1	Sewa	Dapat						
45		D19	50	Tidak Servis	Rp 900,000.00	3	Sewa	Dapat						
46		D20	35	Tidak Bekerja	Rp 350,000.00	3	Menumpang	Dapat						
47		D21	59	Tidak Bekerja	Rp 300,000.00	5	Menumpang	Dapat						
48		D22	30	Buruh	Rp 450,000.00	3	Menumpang	Dapat						
49		D23	38	PNS	Rp 2,700,000.00	1	Milik Sendiri	Tidak Dapat						
50		D24	42	Petani	Rp 800,000.00	3	Sewa	Dapat						

Data Testing

Selain itu, pada bagian kanan tabel terlihat proses perhitungan probabilitas secara manual yang digunakan sebagai contoh penerapan rumus Naïve Bayes dalam menghitung peluang klasifikasi. Nilai-nilai probabilitas bersyarat (conditional probability) untuk masing-masing atribut dihitung berdasarkan frekuensi kemunculan setiap kategori dalam data training, kemudian dikalikan dengan peluang kelas (prior probability) untuk menentukan hasil prediksi. Pendekatan ini membantu memverifikasi kebenaran logika klasifikasi sebelum data dimasukkan ke dalam perangkat lunak seperti RapidMiner untuk proses pelatihan model secara otomatis. Dengan demikian, data Excel ini tidak hanya berfungsi sebagai sumber data utama penelitian, tetapi juga sebagai sarana awal untuk memahami dan menguji konsep perhitungan dalam metode Naïve Bayes.

4.3. Pengumpulan Data

Pengumpulan data pada penelitian ini dilakukan melalui wawancara langsung kepada masyarakat di wilayah Rantauprapat yang menjadi target penerima bantuan

sosial (BPJS). Proses wawancara ini dilakukan secara terstruktur dengan panduan pertanyaan yang telah disusun sebelumnya, untuk memperoleh informasi yang relevan terkait usia, pekerjaan, pendapatan per bulan, jumlah tanggungan, status rumah, serta status penerimaan bantuan. Data yang berhasil dikumpulkan sebanyak 100 responden, di mana masing-masing data mencerminkan kondisi sosial dan ekonomi masyarakat yang beragam. Informasi ini memberikan gambaran yang representatif terhadap faktor-faktor yang memengaruhi kelayakan seseorang dalam menerima bantuan. Data tersebut kemudian digunakan sebagai dasar dalam membangun model klasifikasi menggunakan algoritma Naive Bayes, guna mengidentifikasi pola-pola yang memengaruhi keputusan pemberian bantuan serta mengelompokkan data ke dalam kategori “Dapat” atau “Tidak Dapat” secara objektif dan sistematis.

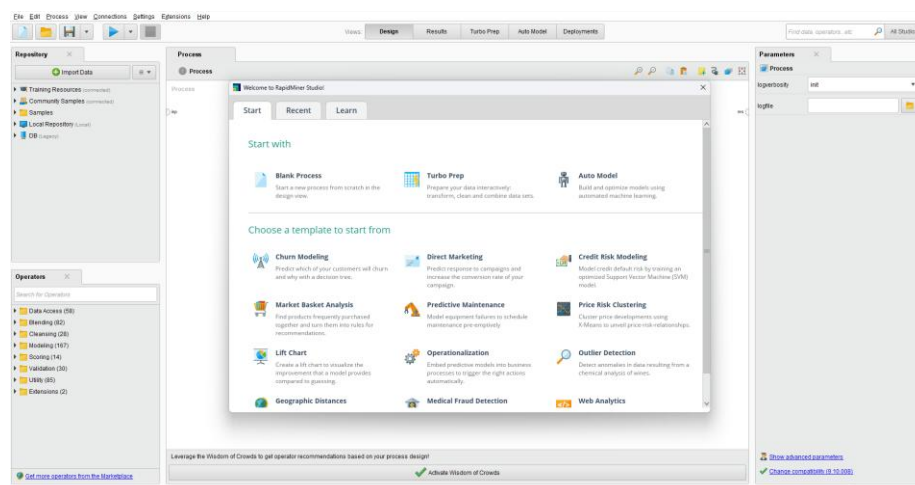
Pengumpulan data pada penelitian ini dilakukan melalui wawancara langsung kepada masyarakat yang berdomisili di wilayah Rantauprapat dan menjadi calon atau penerima manfaat program bantuan sosial berupa BPJS. Wawancara dilakukan menggunakan panduan terstruktur untuk mendapatkan informasi yang akurat mengenai usia, pekerjaan, pendapatan per bulan, jumlah tanggungan, status kepemilikan rumah, serta status apakah individu tersebut menerima bantuan (BPJS) atau tidak. Sebanyak 100 data responden berhasil dikumpulkan dalam proses ini. Data tersebut mencerminkan kondisi sosial dan ekonomi masyarakat yang beragam dan relevan dengan kriteria pemberian bantuan. Seluruh data yang diperoleh kemudian digunakan sebagai dasar dalam membangun model klasifikasi menggunakan algoritma Naive Bayes. Tujuan dari proses ini adalah untuk

mengidentifikasi pola dan hubungan antar atribut yang berpengaruh terhadap kemungkinan seseorang menerima bantuan, serta mengklasifikasikan masyarakat ke dalam kategori “Dapat” atau “Tidak Dapat” secara sistematis dan berbasis data.

4.4. Pembersihan Data

Pada tahapan pembersihan data, dilakukan proses seleksi untuk memastikan bahwa hanya data yang layak dan relevan yang digunakan dalam analisis klasifikasi penerima bantuan di wilayah Rantauprapat. Proses ini mencakup penghapusan data yang tidak lengkap, tidak konsisten, atau tidak sesuai dengan format yang telah ditentukan, seperti data yang memiliki nilai kosong pada atribut penting—yakni usia, pekerjaan, pendapatan per bulan, jumlah tanggungan, dan status kepemilikan rumah. Selain itu, data yang tidak memenuhi kriteria wawancara juga disaring. Langkah ini penting untuk menjaga kualitas dataset dan memastikan bahwa proses pelatihan dan pengujian model klasifikasi menggunakan algoritma Naive Bayes dilakukan dengan data yang bersih, representatif, dan valid, sehingga hasil klasifikasi menjadi lebih akurat dan andal.

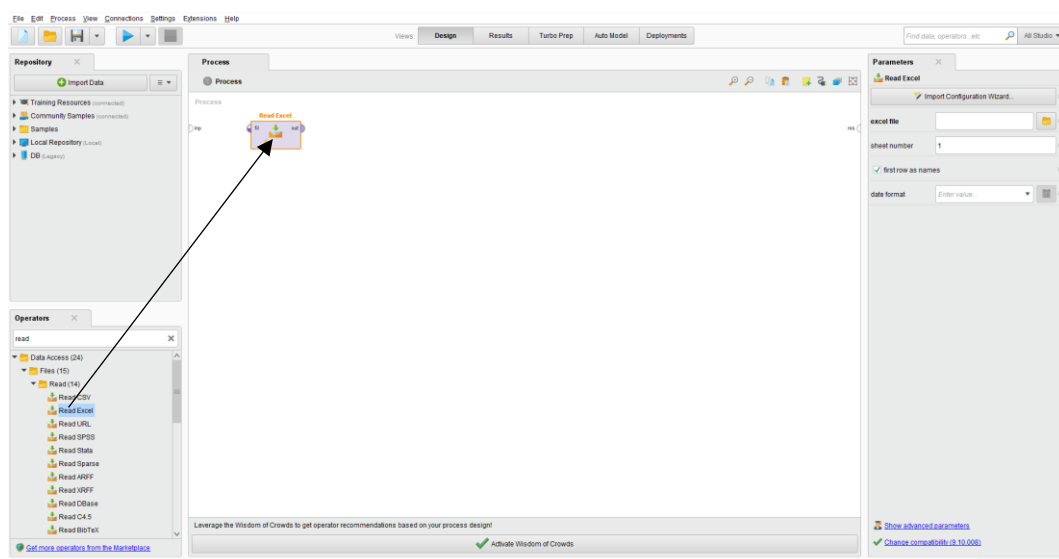
4.5. Buka Aplikasi RapidMiner



Gambar 4. 1. Tampilan Awal Aplikasi RapidMiner

Gambar di atas memperlihatkan tampilan awal dari aplikasi RapidMiner yang digunakan dalam penelitian ini untuk membangun model klasifikasi penerima bantuan berdasarkan data hasil wawancara masyarakat di Rantauprapat. Antarmuka utama RapidMiner menyediakan berbagai menu seperti New Process, Open, dan Recent, yang berfungsi untuk membuat proyek baru, membuka proyek yang sudah ada, serta mengakses proses terakhir yang digunakan. Selain itu, terdapat juga panel penting seperti Operators, Repositories, dan Design View yang memudahkan pengguna dalam merancang alur pemrosesan data secara visual. Dalam penelitian ini, proses analisis dimulai dengan memilih New Process untuk membuka lembar kerja baru, di mana peneliti menyusun tahapan mulai dari input data hasil wawancara, pembersihan data, pengolahan atribut, hingga klasifikasi menggunakan algoritma Naive Bayes. RapidMiner membantu menyederhanakan seluruh proses tersebut melalui antarmuka yang intuitif dan fitur drag-and-drop yang efisien.

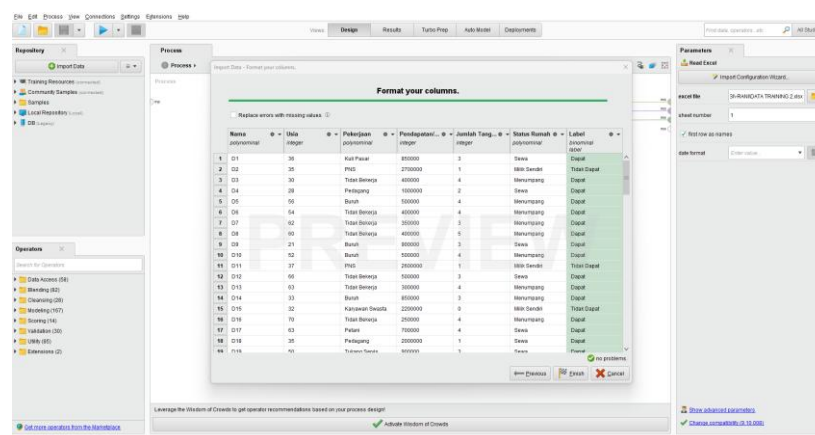
4.6. Tambahkan Widget File



Gambar 4. 2. Widget Read Excel

Gambar di atas menunjukkan tahap awal pemuatan dan visualisasi data hasil wawancara masyarakat dalam aplikasi RapidMiner yang digunakan dalam penelitian ini. Pada sisi kiri antarmuka terlihat panel Operators, sementara di bagian tengah area kerja ditampilkan alur proses utama yang menggunakan dua operator penting, yaitu Read Excel dan Result Viewer. Operator Read Excel digunakan untuk mengimpor dataset eksternal yang berisi data wawancara masyarakat Rantauprapat terkait penerimaan bantuan (BPJS), yang telah disusun dalam format Excel (.xlsx). Setelah data berhasil dimuat, alurnya diteruskan ke operator Result Viewer untuk menampilkan isi data secara visual. Tahapan ini memungkinkan peneliti untuk melakukan verifikasi langsung terhadap struktur dan isi data sebelum melanjutkan ke proses pembersihan data, seleksi atribut, pelatihan model klasifikasi menggunakan algoritma Naive Bayes, hingga evaluasi performa model. Dalam tampilan tersebut juga terlihat dua jalur pemrosesan data, masing-masing mewakili data training dan testing yang telah dipisahkan sebelumnya untuk kebutuhan pelatihan dan pengujian model secara paralel.

4.7. Input Data



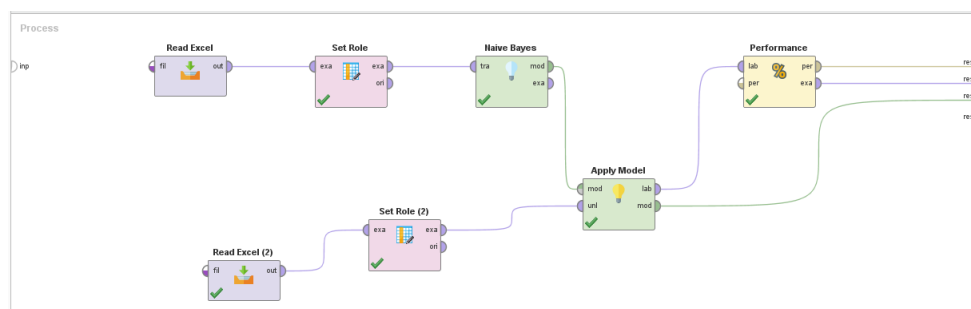
Gambar 4. 3. Input Data Penelitian

Gambar di atas merupakan tampilan konfigurasi input data menggunakan operator Read Excel pada aplikasi RapidMiner dalam konteks penelitian ini. Data yang dimasukkan berasal dari file Excel berjudul "*DATA TRAINING*", yang berisi hasil wawancara langsung dengan masyarakat Rantauprapat mengenai kriteria penerima bantuan BPJS. Dataset ini terdiri atas 100 entri dan memuat enam atribut utama, yaitu Nama, Usia, Pekerjaan, Pendapatan/Bulan, Jumlah Tanggungan, dan Status Rumah, serta satu atribut target yaitu Status Bantuan. Lima atribut pertama ditetapkan sebagai atribut independen (features) yang digunakan untuk memprediksi penerimaan bantuan, sementara atribut Status Bantuan berfungsi sebagai label klasifikasi dengan dua kelas: *Dapat* dan *Tidak Dapat*. Atribut Nama dikonfigurasi sebagai atribut tambahan (meta) yang tidak dilibatkan dalam pelatihan model. Pada tahap ini, peneliti memastikan bahwa setiap kolom telah diklasifikasikan sesuai perannya agar dataset siap untuk dianalisis dan digunakan dalam proses klasifikasi penerima bantuan dengan algoritma Naive Bayes di RapidMiner.

4.8. Perancangan Model Klasifikasi

Pada penelitian ini, perancangan model klasifikasi dilakukan dengan menyusun algoritma Naive Bayes yang bertujuan untuk mengklasifikasikan status penerimaan bantuan BPJS oleh masyarakat Rantauprapat berdasarkan beberapa atribut utama, yaitu Usia, Pekerjaan, Pendapatan/Bulan, Jumlah Tanggungan, dan Status Rumah. Model ini dirancang untuk belajar dari data training yang telah disiapkan sebelumnya, lalu diuji menggunakan data testing guna mengevaluasi akurasi dan efektivitas klasifikasinya dalam mengenali pola-pola yang relevan pada

data. Melalui proses pelatihan ini, algoritma Naive Bayes diharapkan mampu membangun pemahaman probabilistik terhadap hubungan antar atribut, sehingga dapat secara akurat mengelompokkan data masyarakat ke dalam kategori Dapat atau Tidak Dapat menerima bantuan berdasarkan kombinasi kondisi sosial ekonomi yang dimiliki.



Gambar 4. 4. Perancangan Model Klasifikasi

Gambar di atas menunjukkan alur perancangan model klasifikasi penerimaan bantuan BPJS menggunakan aplikasi RapidMiner. Proses diawali dengan komponen Read Excel yang digunakan untuk memuat dataset training berisi atribut seperti Usia, Pekerjaan, Pendapatan/Bulan, Jumlah Tanggungan, dan Status Rumah. Data tersebut kemudian diproses melalui komponen Set Role untuk menetapkan atribut Kelayakan (Dapat/Tidak Dapat) sebagai label atau target klasifikasi. Selanjutnya, algoritma Naive Bayes diterapkan melalui operator khusus yang berfungsi untuk membentuk model klasifikasi berdasarkan data training tersebut. Setelah model terbentuk, digunakan operator Apply Model untuk menerapkan model terhadap data testing, yang juga dimuat melalui komponen Read Excel (2) dan dikonfigurasi terlebih dahulu dengan Set Role agar label klasifikasi dikenali dengan benar. Tahapan terakhir dilakukan dengan operator Performance, yang digunakan untuk membandingkan hasil prediksi dengan label aktual dan

menghitung metrik evaluasi seperti akurasi, presisi, dan recall, guna menilai performa dari model klasifikasi yang telah dibangun.

4.9. Hasil Klasifikasi

Hasil klasifikasi pada penelitian ini ditampilkan dalam bentuk tabel yang memuat data testing dari opini masyarakat terkait layanan rehabilitasi sosial, lengkap dengan prediksi kategori kepuasan dari model Naive Bayes. Tabel ini mencakup kolom nama responden, atribut input seperti *Kualitas Pelayanan*, *Aksesibilitas Layanan*, dan *Transparansi Informasi*, serta label aktual *Kepuasan Masyarakat*. Di samping itu, tabel juga menampilkan hasil prediksi dari model klasifikasi yang telah dibangun menggunakan algoritma Naive Bayes, sehingga memudahkan peneliti dalam mengevaluasi ketepatan prediksi terhadap data yang belum pernah dilihat oleh model sebelumnya.

Tabel 4. 1. Hasil Klasifikasi

Row No.	Label	predictions	confidence	confidence	Nama	Usia	Pekerjaan	Pendapatan	Jumlah Tay.	Status Rumah
1	Dapat	Dapat	1.000	0.000	D1	36	Kuli Pasar	800000	3	Sewa
2	Tidak Dapat	Tidak Dapat	0.000	1.000	D2	35	PHS	2700000	1	Milik Sendiri
3	Dapat	Dapat	1.000	0.000	D3	30	Tidak Bekerja	400000	4	Menumpang
4	Dapat	Dapat	1.000	0.000	D4	28	Pedagang	1000000	2	Sewa
5	Dapat	Dapat	1.000	0.000	D5	56	Buruh	500000	4	Menumpang
6	Dapat	Dapat	1.000	0.000	D6	54	Tidak Bekerja	400000	4	Menumpang
7	Dapat	Dapat	1.000	0.000	D7	62	Tidak Bekerja	350000	3	Menumpang
8	Dapat	Dapat	1.000	0.000	D8	60	Tidak Bekerja	400000	5	Menumpang
9	Dapat	Dapat	1.000	0.000	D9	21	Buruh	800000	3	Sewa
10	Dapat	Dapat	1.000	0.000	D10	52	Buruh	500000	4	Menumpang
11	Tidak Dapat	Tidak Dapat	0.000	1.000	D11	37	PHS	2500000	1	Milik Sendiri
12	Dapat	Dapat	1.000	0.000	D12	66	Tidak Bekerja	500000	3	Sewa
13	Dapat	Dapat	1.000	0.000	D13	63	Tidak Bekerja	300000	4	Menumpang
14	Dapat	Dapat	1.000	0.000	D14	33	Buruh	850000	3	Menumpang
15	Tidak Dapat	Tidak Dapat	0.000	1.000	D15	32	Karyawan Sw.	2200000	0	Milik Sendiri
16	Dapat	Dapat	1	0	D16	70	Tidak Bekerja	250000	4	Menumpang
17	Dapat	Dapat	1.000	0.000	D17	63	Petani	700000	4	Sewa
18	Dapat	Dapat	0.739	0.261	D18	35	Pedagang	2000000	1	Sewa
19	Dapat	Dapat	1.000	0.000	D19	50	Tulang Sewa	800000	3	Sewa
20	Dapat	Dapat	1.000	0.000	D20	35	Tidak Bekerja	350000	3	Menumpang
21	Dapat	Dapat	1.000	0.000	D21	59	Tidak Bekerja	300000	5	Menumpang
22	Dapat	Dapat	1.000	0.000	D22	30	Buruh	650000	3	Menumpang
23	Tidak Dapat	Tidak Dapat	0.000	1.000	D23	38	PHS	2700000	1	Milik Sendiri
24	Dapat	Dapat	1.000	0.000	D24	62	Petani	800000	3	Sewa
25	Dapat	Dapat	1.000	0.000	D25	22	Buruh	800000	3	Sewa

Berdasarkan hasil klasifikasi terhadap 100 data responden masyarakat Rantauprapat, diperoleh bahwa mayoritas responden masuk ke dalam kategori

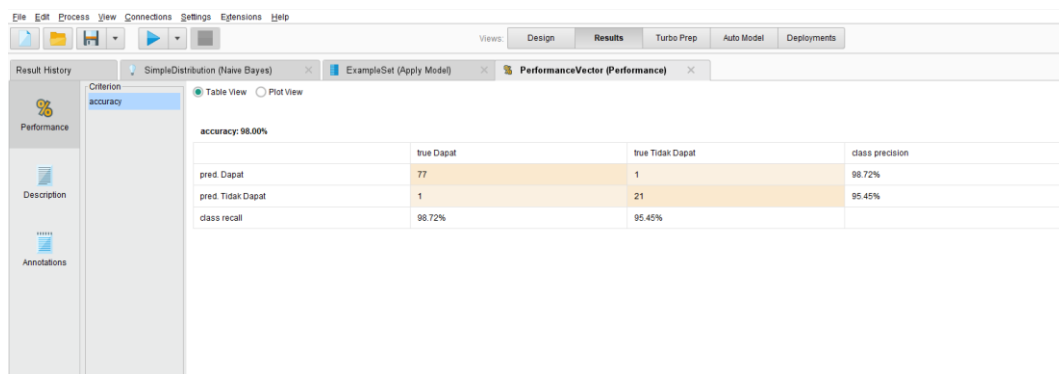
Layak menerima bantuan BPJS, yaitu sebanyak 78 orang, sedangkan 22 orang lainnya diklasifikasikan sebagai *Tidak Layak*. Kategori *Layak* menunjukkan bahwa responden memenuhi kriteria untuk mendapatkan bantuan berdasarkan faktor-faktor seperti penghasilan rendah, jumlah tanggungan keluarga yang tinggi, dan status tempat tinggal yang tidak dimiliki secara pribadi (seperti menumpang atau menyewa).

Distribusi ini menggambarkan bahwa sebagian besar responden dalam dataset berada pada kondisi sosial ekonomi yang kurang mampu. Mereka umumnya bekerja sebagai buruh, petani, atau bahkan tidak memiliki pekerjaan tetap, dengan rata-rata penghasilan bulanan antara Rp250.000 hingga Rp1.000.000. Selain itu, sebagian besar dari mereka memiliki lebih dari tiga tanggungan dan tinggal di rumah kontrakan, rumah keluarga, atau tempat tinggal sewa, yang merupakan indikator kerentanan ekonomi. Faktor-faktor inilah yang memengaruhi sistem klasifikasi Naive Bayes dalam mengelompokkan mereka ke dalam kategori *Layak* menerima bantuan.

Sebaliknya, sebanyak 22 responden yang termasuk dalam kategori *Tidak Layak* umumnya memiliki kondisi ekonomi yang relatif stabil. Mereka bekerja sebagai PNS, pegawai tetap, atau pelaku usaha dengan penghasilan di atas Rp2.000.000 per bulan, memiliki sedikit tanggungan, dan tinggal di rumah milik pribadi. Dengan demikian, kelompok ini dianggap tidak memenuhi kriteria prioritas penerima bantuan BPJS. Hasil klasifikasi ini menjadi dasar penting dalam mendukung kebijakan penyaluran bantuan secara objektif, agar tepat sasaran dan berbasis data.

4.10. Hasil Evaluasi

Hasil evaluasi model klasifikasi menunjukkan bahwa algoritma Naive Bayes mampu mengelompokkan data opini masyarakat dengan tingkat ketepatan yang cukup baik. Confusion matrix yang dihasilkan memperlihatkan bahwa sebagian besar prediksi sesuai dengan kategori aktual, meskipun masih terdapat beberapa kesalahan klasifikasi pada sejumlah data.



The screenshot shows the 'Performance Vector (Performance)' window in RapidMiner. It displays the following data:

	True Dapat	True Tidak Dapat	Class Precision
pred. Dapat	77	1	98.72%
pred. Tidak Dapat	1	21	95.45%
Class Recall	98.72%	95.45%	

Overall accuracy is 98.00%.

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:

$$\text{Accuracy} = \frac{77+21}{77+21+1+1} \times 100\% \quad \text{Then the Accuracy value} = 98\%$$

$$\text{Presisi Dapat} = \frac{77}{77+1} \times 100\% \quad \text{Then the Precision value} = 98,72\%$$

$$\text{Presisi Tidak Dapat} = \frac{21}{21+1} \times 100\% \quad \text{Then the Precision value} = 95,45\%$$

$$\text{Recall Dapat} = \frac{77}{77+1} \times 100\% \quad \text{Then the Recall value} = 98,72\%$$

$$\text{Recall Tidak Dapat} = \frac{21}{21+1} \times 100\% \quad \text{Then the Recall value} = 95,45\%$$

Berdasarkan hasil evaluasi model klasifikasi menggunakan algoritma Naive Bayes pada aplikasi RapidMiner, diperoleh tingkat akurasi sebesar 98,00%. Nilai ini menunjukkan bahwa model memiliki performa yang sangat baik dalam mengklasifikasikan data masyarakat Rantau prapat ke dalam dua kategori, yaitu

Layak dan *Tidak Layak* menerima bantuan BPJS. Tingginya akurasi tersebut mengindikasikan bahwa sebagian besar data testing dapat diprediksi secara tepat oleh model, baik untuk kategori positif (*Layak*) maupun negatif (*Tidak Layak*).

Berdasarkan hasil dari confusion matrix, diketahui bahwa jumlah True Positive (TP) adalah 77, yang berarti sebanyak 77 data dengan label aktual *Layak* berhasil diprediksi dengan benar oleh model. Sementara itu, terdapat True Negative (TN) sebanyak 21, yaitu data yang sebenarnya *Tidak Layak* dan berhasil diklasifikasikan dengan tepat. Untuk kesalahan prediksi, model mencatat hanya 1 data False Positive (FP), yakni data yang sebenarnya *Tidak Layak* namun diprediksi sebagai *Layak*, serta 1 data False Negative (FN), yaitu data yang seharusnya *Layak* tetapi diprediksi sebagai *Tidak Layak*.

Hasil evaluasi ini menunjukkan bahwa model Naive Bayes yang dibangun dalam penelitian ini mampu mengenali pola-pola karakteristik sosial ekonomi masyarakat secara efisien dan akurat. Dengan tingkat kesalahan klasifikasi yang sangat rendah, model ini dinilai cukup andal untuk digunakan sebagai alat bantu dalam proses seleksi calon penerima bantuan BPJS, sehingga kebijakan yang diambil oleh pihak terkait, seperti Dinas Sosial, dapat lebih tepat sasaran dan berbasis data.