

BAB IV

HASIL DAN PEMBAHASAN

4.1. Pengumpulan Data

Informasi yang digunakan dalam penelitian ini diperoleh melalui kombinasi wawancara, observasi, dan dokumentasi transaksi penjualan Coffeeshop Bobskuy. Wawancara dilakukan secara langsung dengan pemilik/pengelola untuk menggali penjelasan operasional mengenai alur penjualan harian, cara pencatatan transaksi. Data utama (primer) penelitian ini berupa data historis transaksi penjualan yang diberikan langsung oleh pemilik Coffeeshop Bobskuy dalam periode September 2025 hingga November 2025. Dataset dikumpulkan dari dokumen internal seperti laporan penjualan dan catatan transaksi. Data memuat informasi inti yang relevan untuk implementasi algoritma Naïve Bayes.

4.2. Transformasi Data

Tahap transformasi data dilakukan untuk mengubah data penjualan yang semula terpisah berdasarkan periode pencatatan menjadi satu dataset terintegrasi (data merging) agar siap diproses pada tahap data mining. Secara teknis, data transaksi penjualan Coffeeshop Bobskuy yang tersimpan di Microsoft Excel dikonsolidasikan dengan cara menggabungkan rekap dari beberapa periode waktu ke dalam satu tabel ringkas, sehingga setiap produk memiliki atribut yang konsisten untuk kebutuhan pelatihan model klasifikasi. Langkah ini sejalan dengan kebutuhan penelitian, karena setelah proses awal data, sumber data perlu disatukan menjadi

satu dalam bentuk data merging. Setelah data digabungkan, transformasi juga mencakup penyeragaman format agar kompatibel dengan proses pemodelan, misalnya memastikan nama kolom seragam, tidak ada nilai kosong yang mengganggu, serta tipe data sesuai (numerik/kategorikal) untuk proses klasifikasi. Orange dipilih karena mendukung integrasi dataset umum seperti CSV/Excel dan menyediakan alur kerja visual dari input data, preprocessing/penyesuaian atribut, hingga pemodelan Naïve Bayes.

Tabel 4. 1 Data Merging

Nama Produk	Kategori	Harga	Produk Terjual	Pendapatan	Promo	Klasifikasi
American o	Minuman	Rp10.000,00	169	Rp1.690.000,00	Tidak	Laris
Aren Latte	Minuman	Rp10.000,00	141	Rp1.410.000,00	Tidak	Tidak Laris
Bakso Bakar	Makanan	Rp5.000,00	303	Rp1.515.000,00	Tidak	Laris
Boba Brown Sugar	Minuman	Rp10.000,00	391	Rp3.910.000,00	Ya	Laris
Brownies	Makanan	Rp10.000,00	161	Rp1.610.000,00	Ya	Tidak Laris
Burger	Makanan	Rp10.000,00	212	Rp2.120.000,00	Tidak	Tidak Laris
Cappuccino	Minuman	Rp15.000,00	268	Rp4.020.000,00	Ya	Tidak Laris
Caramel Latte	Minuman	Rp15.000,00	187	Rp2.805.000,00	Tidak	Tidak Laris
Choco Latte	Minuman	Rp15.000,00	273	Rp4.095.000,00	Tidak	Tidak Laris

Coffee Latte	Minuman	Rp10.000,00	140	Rp1.400.000,00	Tidak	Tidak Laris
Dimsum	Makanan	Rp10.000,00	366	Rp3.660.000,00	Tidak	Laris
Espresso	Minuman	Rp10.000,00	447	Rp4.470.000,00	Ya	Laris
Hazelnut Coffee	Minuman	Rp15.000,00	178	Rp2.670.000,00	Ya	Tidak Laris
Kebab	Makanan	Rp10.000,00	260	Rp2.600.000,00	Tidak	Tidak Laris
Kentang Goreng	Makanan	Rp10.000,00	188	Rp1.880.000,00	Tidak	Tidak Laris
Kopi Tarik	Minuman	Rp10.000,00	377	Rp3.770.000,00	Ya	Laris
Kopi Tubruk	Minuman	Rp10.000,00	408	Rp4.080.000,00	Ya	Laris
Matcha	Minuman	Rp15.000,00	354	Rp5.310.000,00	Ya	Laris
Roti Bakar	Makanan	Rp5.000,00	102	Rp510.000,00	Tidak	Tidak Laris
Sanger	Minuman	Rp15.000,00	456	Rp6.840.000,00	Ya	Laris
Robusta	Minuman	Rp15.000,00	372	Rp5.580.000,00	Ya	Laris
Kopi Gula Aren	Minuman	Rp10.000,00	401	Rp4.010.000,00	Tidak	Laris
Tiramisu Latte	Minuman	Rp10.000,00	326	Rp3.260.000,00	Tidak	Laris
Red Velvet Latte	Minuman	Rp15.000,00	382	Rp5.730.000,00	Ya	Laris
Kopi Hitam	Minuman	Rp10.000,00	349	Rp3.490.000,00	Ya	Laris
Thai Tea	Minuman	Rp10.000,00	301	Rp3.010.000,00	Tidak	Laris
Air Mineral	Minuman	Rp.5.000,00	377	Rp1.885.000,00	Tidak	Laris

Pisang Cokelat	Makanan	Rp10.000,00	211	Rp2.110.000,00	Ya	Tidak Laris
Vanila Latte	Minuman	Rp10.000,00	244	Rp2.440.000,00	Tidak	Tidak Laris
Teh Hijau	Minuman	Rp10.000,00	300	Rp3.000.000,00	Ya	Laris

4.3. Pembersihan Data

Pada tahapan pembersihan data, dilakukan proses seleksi untuk memastikan bahwa hanya data penjualan yang layak, relevan, dan konsisten yang digunakan dalam analisis klasifikasi produk Laris dan Tidak Laris di Coffeeshop Bobskuy. Tahap ini diposisikan sebagai bagian penting dari *data preprocessing* karena data yang akan dilatih ke dalam algoritma harus berada dalam kondisi bersih dan terstruktur; pembersihan data mencakup penghapusan *noise*, penanganan nilai hilang (*missing values*), serta koreksi data yang tidak konsisten. Pembersihan dilakukan dengan memeriksa kelengkapan atribut inti yang digunakan pada dataset.

Pembersihan juga diarahkan untuk menyiapkan dataset agar kompatibel dengan alur kerja di Orange Data Mining, yang pada *workflow* praprosesnya mencakup pembersihan nilai kosong, penghapusan duplikasi, dan normalisasi sebelum pemodelan.

4.4. Implementasi Sistem

Implementasi sistem klasifikasi produk terlaris pada Coffeeshop Bobskuy dilakukan dengan menggunakan algoritma Naïve Bayes untuk mengelompokkan produk ke dalam kelas laris dan tidak laris berdasarkan pola pada data historis transaksi penjualan. Proses implementasi mencakup tahapan: pengumpulan data, prapemrosesan, pelatihan model, pengujian model, dan evaluasi hasil klasifikasi.

Data yang telah siap selanjutnya dibagi menjadi dua bagian, yaitu data training dan data testing. Data training digunakan untuk membangun model probabilistik Naïve Bayes, sedangkan data testing digunakan untuk menguji kemampuan prediksi model pada data yang belum dilihat saat pelatihan.

4.5. Pembagian Data (Split Data)

Data yang diambil dibagi menjadi dua bagian, yaitu data pelatihan (training data) dan data pengujian (testing data). Operator Split Data digunakan di sini untuk memisahkan dataset berdasarkan persentase tertentu, misalnya 80% untuk pelatihan dan 20% untuk pengujian. Data ini akan digunakan untuk membangun dan menguji model prediksi.

Tabel 4. 2 Data Training

Nama Produk	Kategori	Harga	Produk Terjual	Pendapatan	Promo	Klasifikasi
Americano	Minuman	Rp10.000,00	169	Rp1.690.000,00	Tidak	Laris
Aren Latte	Minuman	Rp10.000,00	141	Rp1.410.000,00	Tidak	Tidak Laris
Bakso Bakar	Makanan	Rp5.000,00	303	Rp1.515.000,00	Tidak	Laris
Boba Brown Sugar	Minuman	Rp10.000,00	391	Rp3.910.000,00	Ya	Laris
Brownies	Makanan	Rp10.000,00	161	Rp1.610.000,00	Ya	Tidak Laris
Burger	Makanan	Rp10.000,00	212	Rp2.120.000,00	Tidak	Tidak Laris
Cappuccino	Minuman	Rp15.000,00	268	Rp4.020.000,00	Ya	Tidak Laris
Caramel Latte	Minuman	Rp15.000,00	187	Rp2.805.000,00	Tidak	Tidak Laris
Choco Latte	Minuman	Rp15.000,00	273	Rp4.095.000,00	Tidak	Tidak Laris
Coffee Latte	Minuman	Rp10.000,00	140	Rp1.400.000,00	Tidak	Tidak Laris
Dimsum	Makanan	Rp10.000,00	366	Rp3.660.000,00	Tidak	Laris
Espresso	Minuman	Rp10.000,00	447	Rp4.470.000,00	Ya	Laris
Hazelnut Coffee	Minuman	Rp15.000,00	178	Rp2.670.000,00	Ya	Tidak Laris

Kebab	Makanan	Rp10.000,00	260	Rp2.600.000,00	Tidak	Tidak Laris
Kentang Goreng	Makanan	Rp10.000,00	188	Rp1.880.000,00	Tidak	Tidak Laris
Kopi Tarik	Minuman	Rp10.000,00	377	Rp3.770.000,00	Ya	Laris
Kopi Tubruk	Minuman	Rp10.000,00	408	Rp4.080.000,00	Ya	Laris
Matcha	Minuman	Rp15.000,00	354	Rp5.310.000,00	Ya	Laris
Roti Bakar	Makanan	Rp5.000,00	102	Rp510.000,00	Tidak	Tidak Laris
Sanger	Minuman	Rp15.000,00	456	Rp6.840.000,00	Ya	Laris
Robusta	Minuman	Rp15.000,00	372	Rp5.580.000,00	Ya	Laris
Kopi Gula Aren	Minuman	Rp10.000,00	401	Rp4.010.000,00	Tidak	Laris
Tiramisu Latte	Minuman	Rp10.000,00	326	Rp3.260.000,00	Tidak	Laris
Red Velvet Latte	Minuman	Rp15.000,00	382	Rp5.730.000,00	Ya	Laris

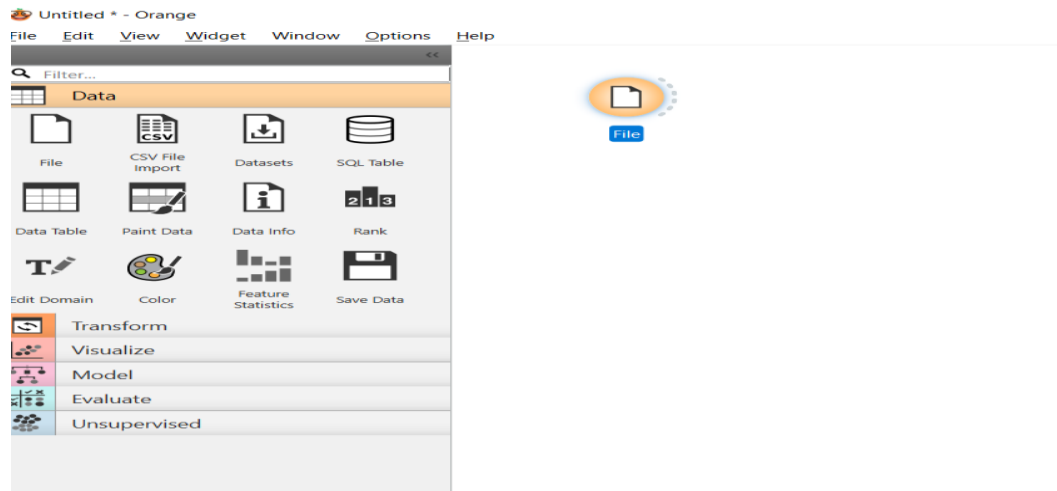
Tabel 4. 3 Data Testing

Nama Produk	Kategori	Harga	Produk Terjual	Pendapatan	Promo	Klasifikasi
Kopi Hitam	Minuman	Rp10.000,00	349	Rp3.490.000,00	Ya	Laris
Thai Tea	Minuman	Rp10.000,00	301	Rp3.010.000,00	Tidak	Tidak Laris
Air Mineral	Minuman	Rp5.000,00	377	Rp1.885.000,00	Tidak	Laris
Pisang Cokelat	Makanan	Rp10.000,00	211	Rp2.110.000,00	Ya	Tidak Laris
Vanila Latte	Minuman	Rp10.000,00	244	Rp2.440.000,00	Ya	Tidak Laris
Teh Hijau	Minuman	Rp10.000,00	300	Rp3.000.000,00	Tidak	Laris

4.6. Pengujian Menggunakan Software Orange

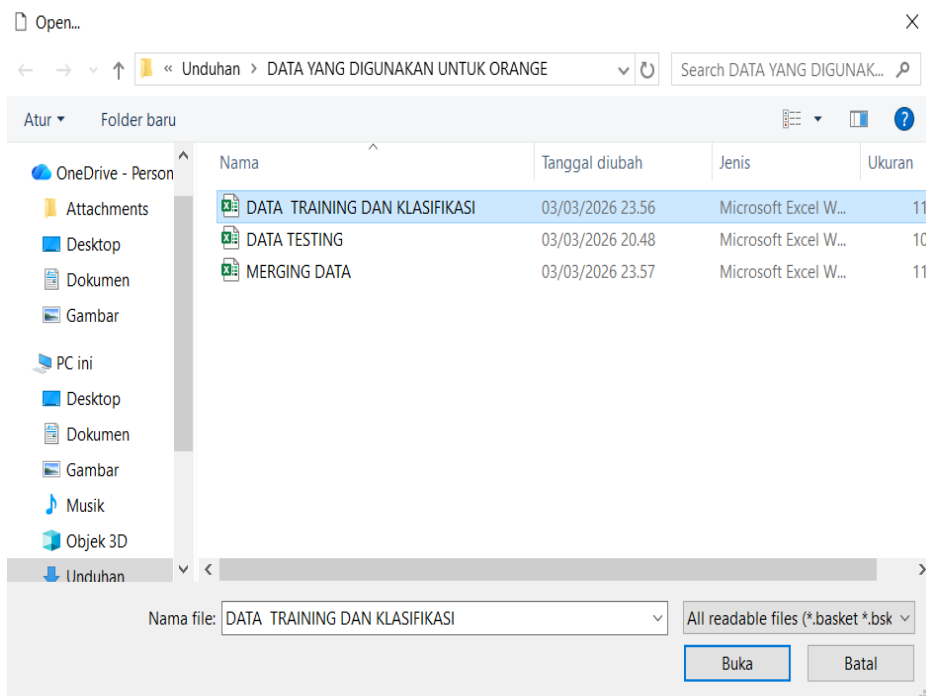
Pada tahap pengujian menggunakan aplikasi Orange, penulis membagi dataset menjadi dua bagian, data training untuk melatih model, dan data testing untuk menguji kinerja serta akurasi hasil klasifikasi.

Gambar 4.1 Menambahkan Widget File



Langkah pertama pada pengujian di Orange adalah menambahkan widget File pada canvas untuk memuat dataset yang akan diproses dan diuji.

Gambar 4.2 Import Data Training



Tahap selanjutnya adalah memasukkan data training ke dalam widget File agar dataset dapat terbaca di Orange dan digunakan sebagai input pada model Naïve Bayes, sehingga algoritma memperoleh informasi dari data yang diberikan untuk membentuk pola klasifikasi.

Gambar 4. 3 Tampilan Data Training

24 instances
6 features (no missing values)
Data has no target variable.
1 meta attribute

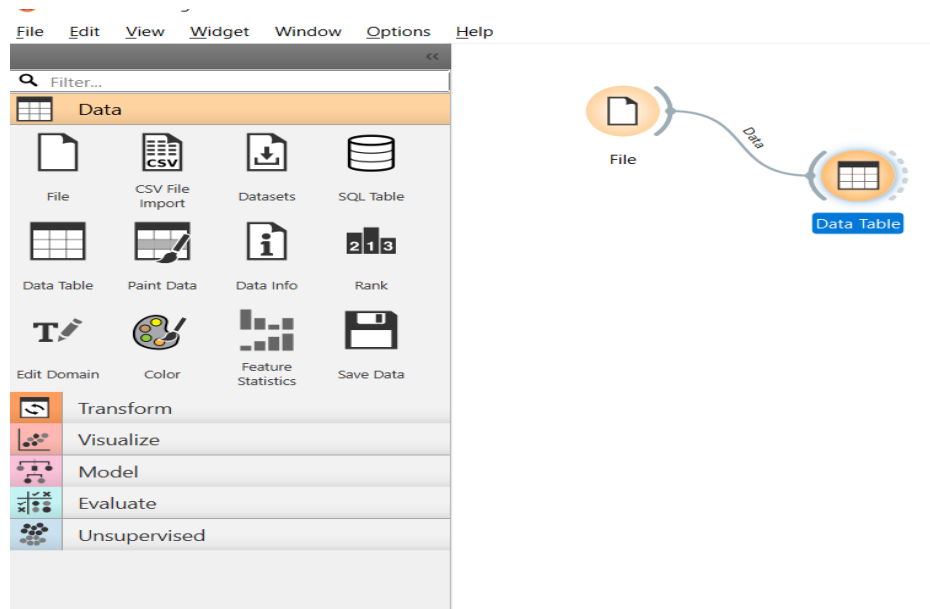
Columns (Double click to edit)

	Name	Type	Role	Values
1	Kategori	C categorical	feature	Makanan, Minuman
2	Harga	N numeric	feature	
3	Produk Terjual	N numeric	feature	
4	Pendapatan	N numeric	feature	
5	Promo	C categorical	feature	Tidak, Ya
6	Klasifikasi	C categorical	target	Laris, Tidak Laris
7	Nama Produk	S text	meta	

Setelah data dimasukkan, akan muncul tampilan seperti pada gambar 4.3.

Pada tahap ini, pengguna cukup menekan tombol Apply agar data training yang telah diinput dapat diterapkan dan masuk ke dalam alur yang sudah dibuat.

Gambar 4. 4 Menambahkan Widget Data Table



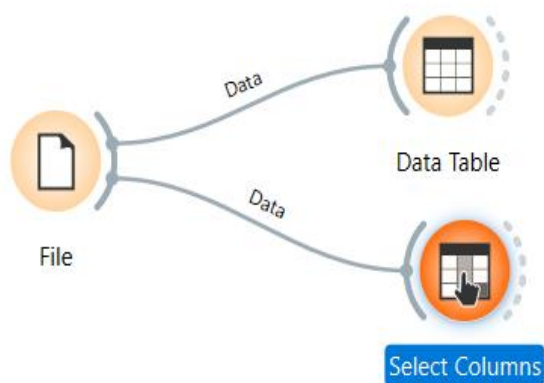
Tahap berikutnya adalah menambahkan Widget Data Table untuk menampilkan dan memeriksa dataset yang telah diinput. Melalui Widget Data Table ini, peneliti dapat melihat seluruh isi data yang berasal dari widget File, sehingga memastikan data telah terbaca dengan benar sebelum dilanjutkan ke proses pemodelan.

Gambar 4. 5 Seluruh Data Yang Ada Pada Data Training

	Klasifikasi	Nama Produk	Kategori	Harga	Produk Terjual	Pendapatan	Promo
1	Laris	Americano	Minuman	10000	169	1690000	Tidak
2	Tidak Laris	Aren Latte	Minuman	10000	141	1410000	Tidak
3	Laris	Bakso Bakar	Makanan	5000	303	1515000	Tidak
4	Laris	Boba Brown Su...	Minuman	10000	391	3910000	Ya
5	Tidak Laris	Brownies	Makanan	10000	161	1610000	Ya
6	Tidak Laris	Burger	Makanan	10000	212	2120000	Tidak
7	Tidak Laris	Cappuccino	Minuman	15000	268	4020000	Ya
8	Tidak Laris	Caramel Latte	Minuman	15000	187	2805000	Tidak
9	Tidak Laris	Choco Latte	Minuman	15000	273	4095000	Tidak
10	Tidak Laris	Coffee Latte	Minuman	10000	140	1400000	Tidak
11	Laris	Dimsum	Makanan	10000	366	3660000	Tidak
12	Laris	Espresso	Minuman	10000	447	4470000	Ya
13	Tidak Laris	Hazelnut Coffee	Minuman	15000	178	2670000	Ya
14	Tidak Laris	Kebab	Makanan	10000	260	2600000	Tidak
15	Tidak Laris	Kentang Goreng	Makanan	10000	188	1880000	Tidak
16	Laris	Kopi Tarik	Minuman	10000	377	3770000	Ya
17	Laris	Kopi Tubruk	Minuman	10000	408	4080000	Ya
18	Laris	Matcha	Minuman	15000	354	5310000	Ya
19	Tidak Laris	Roti Bakar	Makanan	5000	102	510000	Tidak
20	Laris	Sanger	Minuman	15000	456	6840000	Ya
21	Laris	Robusta	Minuman	15000	372	5580000	Ya
22	Laris	Kopi Gula Aren	Minuman	10000	401	4010000	Tidak
23	Laris	Tiramisu Latte	Minuman	10000	326	3260000	Tidak
24	Laris	Red Velvet Latte	Minuman	15000	382	5730000	Ya

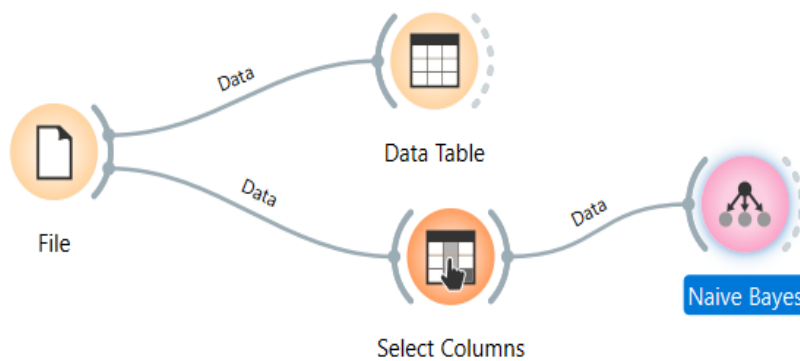
Setelah menambahkan widget Data Table, seluruh data training yang sebelumnya dimuat melalui widget File akan ditampilkan pada Data Table.

Gambar 4. 6 Menambahkan Widget Select Columns



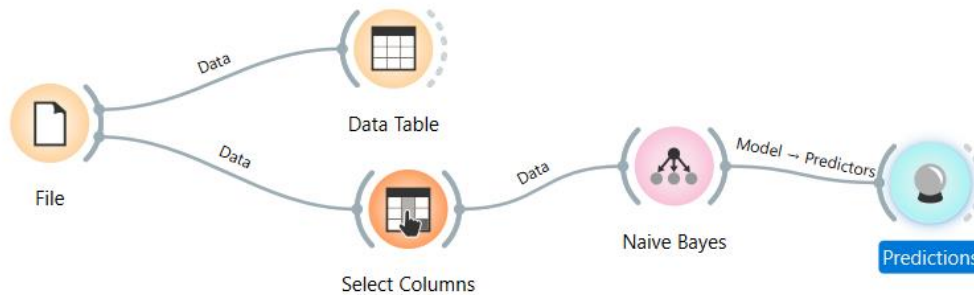
Tahap berikutnya adalah menambahkan widget Select Columns yang berfungsi untuk menentukan atribut mana yang akan digunakan sebagai variabel input dan mana yang dijadikan target/kelas, sehingga fitur yang dianalisis dapat ditetapkan terlebih dahulu sebelum model Naïve Bayes dijalankan.

Gambar 4. 7 Menambahkan Widget Naïve Bayes



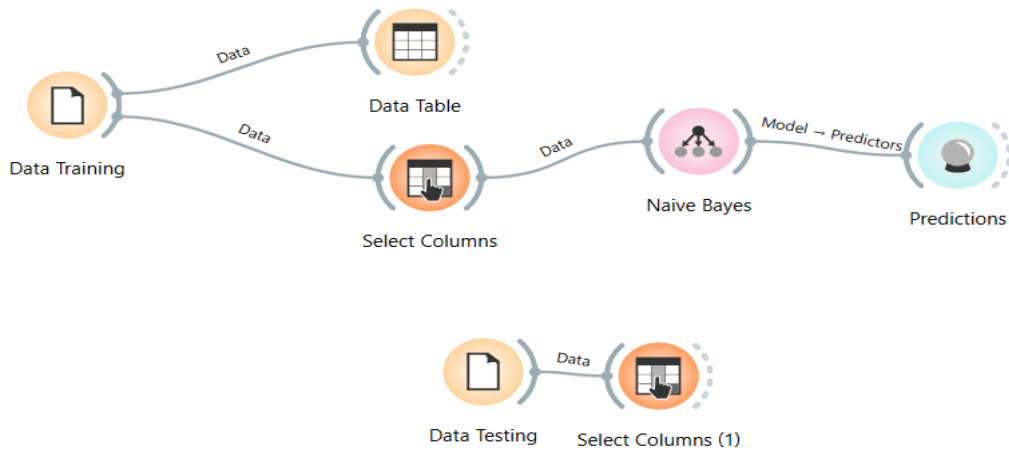
Setelah widget Select Columns ditambahkan, tahap berikutnya adalah memasukkan widget Naïve Bayes sebagai model yang digunakan untuk menghasilkan output klasifikasi. Namun, pada tahap ini proses dan hasil kerja Naïve Bayes belum dapat diamati secara lengkap karena dataset testing belum dihubungkan, sehingga evaluasi kinerja model belum dapat dilakukan.

Gambar 4. 8 Menambahkan Widget Predictions



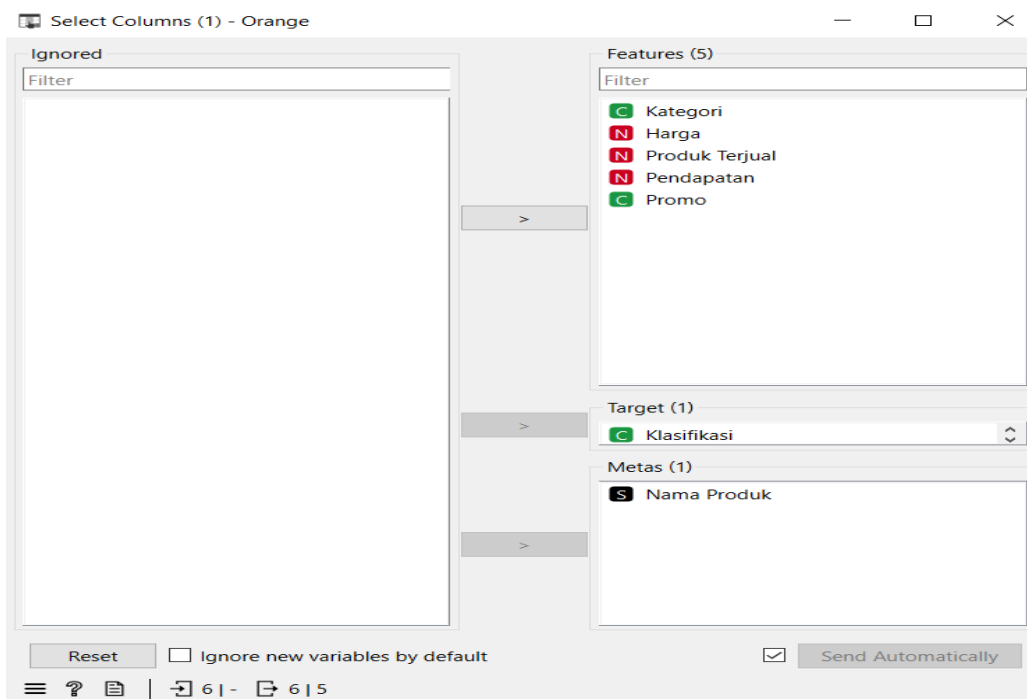
Tahap berikutnya adalah menambahkan widget Predictions yang berfungsi untuk menghasilkan prediksi berdasarkan model dan data yang telah dimasukkan. Namun, pada tahap ini hasil prediksi belum dapat ditampilkan karena data testing belum dihubungkan ke dalam rangkaian proses, sehingga keluaran prediksi belum terbentuk secara utuh.

Gambar 4. 9 Menambahkan Data Testing



Data testing ini kemudian dihubungkan ke rangkaian widget Naïve Bayes dan widget Predictions agar proses pengujian dapat berjalan, sehingga model dapat melakukan perhitungan Naïve Bayes sekaligus menghasilkan prediksi berdasarkan data uji yang diberikan.

Gambar 4. 10 Menjadikan klasifikasi Sebagai Target



Pada tahap ini, peneliti menetapkan variabel yang menjadi fokus analisis, yaitu variabel Klasifikasi sebagai target/kelas. Penetapan target ini bertujuan untuk membandingkan kesesuaian label klasifikasi pada data dengan hasil perhitungan yang dihasilkan oleh aplikasi. Selanjutnya, widget Select Columns dihubungkan dengan widget Predictions agar proses prediksi dapat dijalankan dan keluaran perhitungan berbasis Naïve Bayes dapat diperoleh secara sistematis.

Gambar 4. 11 Hasil Predictions

	Naive Bayes	error	Klasifikasi	Nama Produk	Kategori	Harga	Produk Terjual	Pendapatan	Promo
1	0.98 : 0.02 → Laris	0.017	Laris	Kopi Hitam	Minuman	10000	349	3490000	Ya
2	0.65 : 0.35 → Laris	0.649	Tidak Laris	Thai Tea	Minuman	10000	301	3010000	Tidak
3	0.59 : 0.41 → Laris	0.410	Laris	Air Mineral	Minuman	5000	377	1885000	Tidak
4	0.04 : 0.96 → Tidak Laris	0.036	Tidak Laris	Pisang Cokelat	Makanan	10000	211	2110000	Ya
5	0.11 : 0.89 → Tidak Laris	0.113	Tidak Laris	Vanilla Latte	Minuman	10000	244	2440000	Ya
6	0.65 : 0.35 → Laris	0.351	Laris	Teh Hijau	Minuman	10000	300	3000000	Tidak

Show performance scores							Target class: (Average over classes)	
Model	AUC	CA	F1	Prec	Recall	MCC		
Naive Bayes	0.833	0.833	0.829	0.875	0.833	0.707		

Setelah seluruh widget terhubung, widget Predictions akan menampilkan hasil prediksi secara otomatis. Pada tampilan tersebut, label Naive Bayes (error) tidak menunjukkan bahwa model Naïve Bayes mengalami kegagalan (*error*) sistem, melainkan merupakan nama kolom keluaran yang merepresentasikan tingkat kesalahan (ketidakakuratan) prediksi model untuk setiap baris data. Dengan

demikian, kolom ini digunakan sebagai indikator seberapa jauh hasil prediksi menyimpang dari label aktual pada data uji.

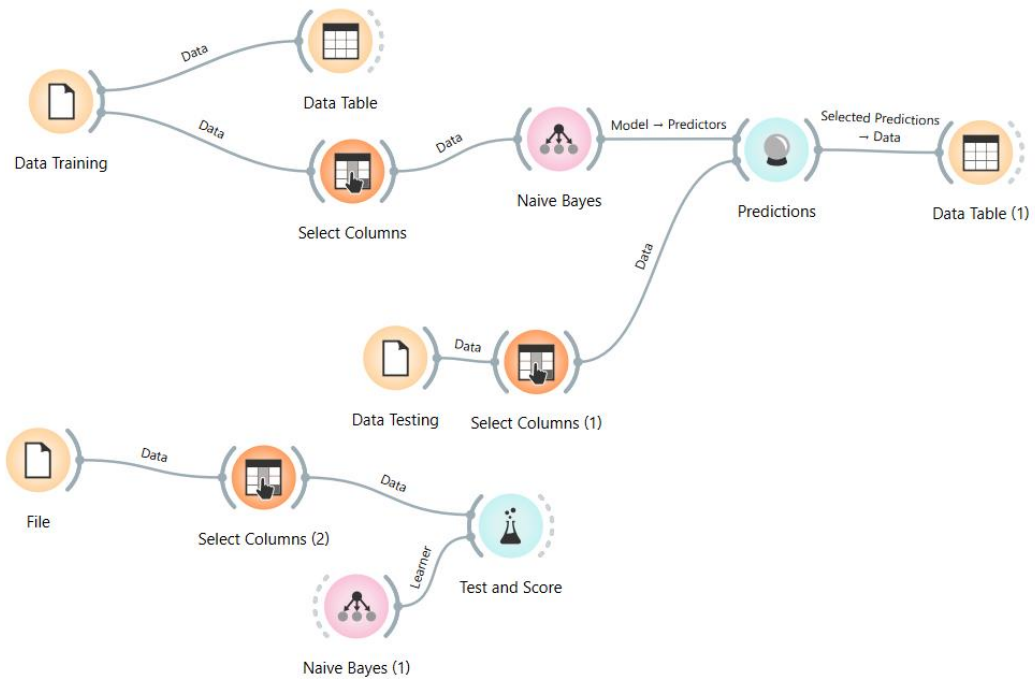
Gambar 4. 12 Analisis Hasil Perhitungan

	Klasifikasi	Nama Produk	Naive Bayes	Naive Bayes (Laris)	Naive Bayes (Tidak Laris)	Naive Bayes (error)	Kategori	Harga	Produk Terjual	Pendapatan	Promo
1	Laris	Kopi Hitam	Laris	0.982522	0.0174783	0.0174783	Minuman	10000	349	3490000	Ya
6	Laris	Teh Hijau	Laris	0.649204	0.350796	0.350796	Minuman	10000	300	3000000	Tidak
3	Laris	Air Mineral	Laris	0.590063	0.409937	0.409937	Minuman	5000	377	1885000	Tidak
2	Tidak Laris	Thai Tea	Laris	0.649204	0.350796	0.649204	Minuman	10000	301	3010000	Tidak
5	Tidak Laris	Vanilla Latte	Tidak Laris	0.113057	0.886943	0.113057	Minuman	10000	244	2440000	Ya
4	Tidak Laris	Pisang Cokelat	Tidak Laris	0.0358457	0.964154	0.0358457	Makanan	10000	211	2110000	Ya

Berdasarkan tampilan hasil pada widget Data Table/Predictions, proses perhitungan Naive Bayes telah menghasilkan kolom prediksi sehingga keluaran model dapat diamati. Pada bagian ini terlihat bahwa nilai pada variabel Klasifikasi (label aktual) tidak selalu sama dengan hasil pada kolom Naive Bayes (label prediksi). Perbedaan tersebut menunjukkan adanya ketidaksesuaian antara label yang dijadikan acuan pada data dengan keputusan klasifikasi yang dihasilkan model berdasarkan pola statistik pada dataset.

Ketidaksamaan antara Klasifikasi dan Naive Bayes dapat terjadi karena label Klasifikasi dalam data penelitian ditetapkan berdasarkan hasil wawancara/penilaian pemilik atau pihak terkait, sedangkan hasil Naive Bayes merupakan keluaran perhitungan algoritma yang dibentuk dari distribusi data dan hubungan antar-atribut yang dipelajari selama proses pelatihan. Dengan kata lain, model tidak mengikuti hasil wawancara, tetapi melakukan prediksi berdasarkan probabilitas yang dihitung dari data yang dimasukkan ke aplikasi.

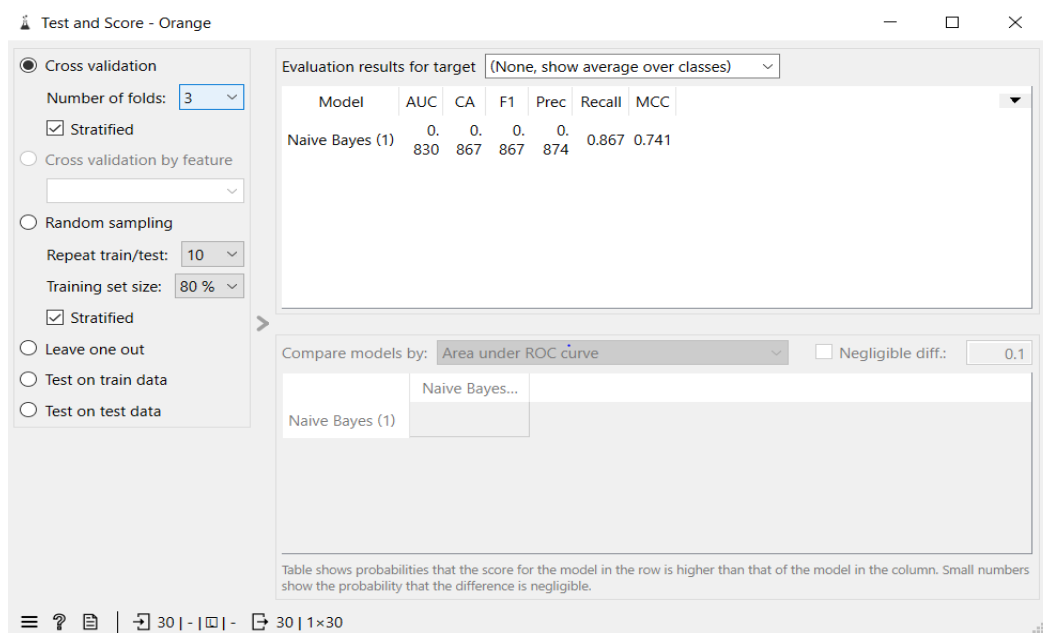
Gambar 4. 13 Menambahkan widget file, select columns, test and score dan juga naive bayes



Selanjutnya ditambahkan rangkaian widget baru, diawali dengan file. Pada widget ini dimasukkan dataset yang akan digunakan sebagai dasar perhitungan Naïve Bayes, sehingga data tersebut dapat diproses dalam tahap klasifikasi. Setelah dataset berhasil dimuat, proses dilanjutkan dengan menambahkan widget Select Columns serta Test & Score. Kedua widget ini berfungsi untuk menetapkan atribut yang digunakan dalam pemodelan sekaligus menghitung metrik evaluasi, meliputi precision, recall, F1-score, AUC (Area Under the Curve), CA (Classification Accuracy), dan MCC (Matthews Correlation Coefficient). Agar widget Test & Score dapat menampilkan hasil evaluasi, selanjutnya perlu ditambahkan widget Naïve Bayes dan dihubungkan ke rangkaian tersebut. Widget Naïve Bayes berperan sebagai model klasifikasi yang dijalankan dalam proses pengujian pada Test & Score. Apabila widget Naïve Bayes tidak disertakan, maka perhitungan pada Test

& Score tidak dapat dieksekusi karena tidak tersedia algoritma yang membentuk model klasifikasinya. Oleh karena itu, penambahan widget Naïve Bayes bersifat wajib agar proses perhitungan dan evaluasi dapat berjalan sebagaimana mestinya.

Gambar 4. 14 Hasil Perhitungan Naïve Bayes



Hasil evaluasi model klasifikasi menggunakan algoritma Naïve Bayes menunjukkan performa yang tergolong baik. Berdasarkan pengujian cross-validation dengan 3-fold stratified, model memperoleh nilai AUC (Area Under the Curve) sebesar 0.830. Nilai ini mengindikasikan bahwa model memiliki kemampuan yang memadai dalam membedakan kelas, sehingga secara umum dapat dikatakan bahwa kapasitas diskriminatif model berada pada kategori baik, walaupun belum menunjukkan pemisahan kelas yang sangat kuat.

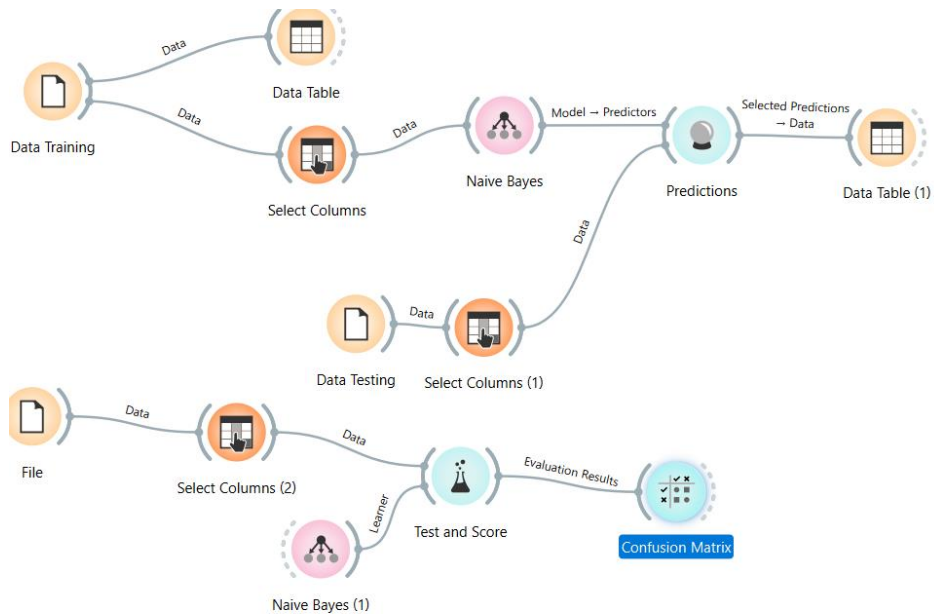
Dari sisi ketepatan klasifikasi, nilai CA (Classification Accuracy) sebesar 0.867 menunjukkan bahwa model mampu mengklasifikasikan 86,7% data dengan

benar. Nilai ini menandakan bahwa secara keseluruhan model telah bekerja dengan cukup efektif dalam mengenali pola data. Temuan ini juga didukung oleh nilai F1-score sebesar 0.867, yang menunjukkan adanya keseimbangan yang baik antara precision dan recall. Dengan demikian, performa model tidak hanya baik dalam menghasilkan prediksi yang tepat, tetapi juga cukup konsisten dalam menangkap data dari kelas yang relevan.

Nilai precision sebesar 0.874 mengindikasikan bahwa sebagian besar hasil prediksi positif yang dihasilkan model adalah benar. Dalam konteks klasifikasi, hal ini menunjukkan tingkat ketepatan prediksi yang baik ketika model menetapkan suatu data ke kelas tertentu. Sementara itu, nilai recall sebesar 0.867 menunjukkan bahwa model mampu mengidentifikasi sebagian besar data yang memang termasuk ke dalam kelas yang seharusnya terdeteksi. Kedua metrik ini memperlihatkan bahwa model memiliki performa yang relatif seimbang antara ketepatan prediksi dan kemampuan menangkap data yang relevan.

Nilai MCC (Matthews Correlation Coefficient) sebesar 0.741 memperkuat bahwa hasil klasifikasi model memiliki kualitas yang cukup kuat. Nilai MCC yang berada jauh di atas nol menunjukkan adanya korelasi positif yang baik antara hasil prediksi model dan label aktual. Hal ini menandakan bahwa model Naïve Bayes tidak hanya menunjukkan akurasi yang baik secara umum, tetapi juga memiliki kemampuan yang cukup memadai dalam menghasilkan klasifikasi

Gambar 4. 15 Menambahkan Widget Confusion Matrix



Widget Confusion Matrix sebagai rangkaian proses baru yang berfungsi untuk memperoleh nilai true positive, true negative, false positive, dan false negative. Selain itu, penambahan widget ini bertujuan untuk mengetahui jumlah klasifikasi produk yang termasuk terlaris dan tidak laris, sehingga data tersebut dapat dihitung secara rinci berdasarkan kategori yang benar maupun keliru. Dengan demikian, melalui Confusion Matrix dapat ditentukan secara jelas perolehan true positive, true negative, false positive, serta false negative sebagai dasar evaluasi kinerja model.

Gambar 4. 16 Probabilitas Prior dan Likelihood

	Bagian	Variabel	Nilai	Kelas	Jumlah	Pembagi	Probabilitas
1	PRIOR	-	-	Laris	16	30	0.5333
2	PRIOR	-	-	Tidak Laris	14	30	0.4667
3	LIKELIHOOD	Kategori	Makanan	Laris	2	16	0.125
4	LIKELIHOOD	Kategori	Makanan	Tidak Laris	6	14	0.4286
5	LIKELIHOOD	Kategori	Minuman	Laris	14	16	0.875
6	LIKELIHOOD	Kategori	Minuman	Tidak Laris	8	14	0.5714
7	LIKELIHOOD	Harga	Menengah	Laris	10	16	0.625
8	LIKELIHOOD	Harga	Menengah	Tidak Laris	9	14	0.6429
9	LIKELIHOOD	Harga	Rendah	Laris	2	16	0.125
10	LIKELIHOOD	Harga	Rendah	Tidak Laris	1	14	0.0714
11	LIKELIHOOD	Harga	Tinggi	Laris	4	16	0.25
12	LIKELIHOOD	Harga	Tinggi	Tidak Laris	4	14	0.2857
13	LIKELIHOOD	Produk Terjual	Menengah	Laris	5	16	0.3125
14	LIKELIHOOD	Produk Terjual	Menengah	Tidak Laris	5	14	0.3571
15	LIKELIHOOD	Produk Terjual	Rendah	Laris	1	16	0.0625
16	LIKELIHOOD	Produk Terjual	Rendah	Tidak Laris	9	14	0.6429
17	LIKELIHOOD	Produk Terjual	Tinggi	Laris	10	16	0.625
18	LIKELIHOOD	Produk Terjual	Tinggi	Tidak Laris	0	14	0
19	LIKELIHOOD	Pendapatan	Menengah	Laris	5	16	0.3125
20	LIKELIHOOD	Pendapatan	Menengah	Tidak Laris	5	14	0.3571
21	LIKELIHOOD	Pendapatan	Rendah	Laris	3	16	0.1875
22	LIKELIHOOD	Pendapatan	Rendah	Tidak Laris	7	14	0.5
23	LIKELIHOOD	Pendapatan	Tinggi	Laris	8	16	0.5
24	LIKELIHOOD	Pendapatan	Tinggi	Tidak Laris	2	14	0.1429
25	LIKELIHOOD	Promo	Tidak	Laris	6	16	0.375
26	LIKELIHOOD	Promo	Tidak	Tidak Laris	10	14	0.7143
27	LIKELIHOOD	Promo	Ya	Laris	10	16	0.625
28	LIKELIHOOD	Promo	Ya	Tidak Laris	4	14	0.2857

Berdasarkan output probabilitas pada Orange, probabilitas awal atau prior menunjukkan bahwa kelas Laris memiliki nilai yang sedikit lebih besar dibandingkan kelas Tidak Laris. Hal ini terlihat dari jumlah data kelas Laris sebanyak 16 dengan probabilitas 0,5333, sedangkan kelas Tidak Laris berjumlah 14 dengan probabilitas 0,4667. Kondisi ini menunjukkan bahwa secara umum data pada penelitian ini sedikit lebih banyak didominasi oleh produk yang termasuk kategori laris. Nilai prior ini penting karena menjadi dasar awal perhitungan probabilitas posterior pada algoritma Naïve Bayes sebelum mempertimbangkan atribut lainnya.

Bagian likelihood pada output Orange menunjukkan hubungan masing-masing atribut dengan kelas Laris dan Tidak Laris. Pada atribut Kategori, nilai probabilitas menunjukkan bahwa produk Minuman memiliki kecenderungan lebih tinggi berada pada kelas Laris dibandingkan Makanan. Pada kelas Laris, kategori Minuman memiliki probabilitas 0,875, sedangkan Makanan hanya 0,125. Hal ini menunjukkan bahwa dalam dataset yang digunakan, produk minuman lebih sering muncul sebagai produk laris. Sebaliknya, pada kelas Tidak Laris, kategori Minuman juga masih cukup dominan, namun proporsinya lebih rendah dibandingkan pada kelas Laris. Dengan demikian, atribut kategori memberikan gambaran bahwa produk minuman memiliki peluang yang lebih besar untuk tergolong laris dibandingkan produk makanan.

Pada atribut Harga, hasil likelihood menunjukkan bahwa kategori Menengah/Sedang memiliki probabilitas yang paling besar baik pada kelas Laris maupun Tidak Laris. Pada kelas Laris, harga menengah memiliki probabilitas 0,625, sedangkan pada kelas Tidak Laris sebesar 0,6429. Temuan ini menunjukkan bahwa harga menengah merupakan kategori harga yang paling sering muncul dalam dataset, tetapi belum menjadi pembeda yang sangat kuat antara produk laris dan tidak laris. Dengan kata lain, harga tetap berperan dalam klasifikasi, namun pengaruhnya tidak sekuat beberapa atribut lain.

Pada atribut Produk Terjual, terlihat pola yang jauh lebih tegas. Produk dengan tingkat penjualan Tinggi memiliki probabilitas 0,625 pada kelas Laris dan 0 pada kelas Tidak Laris. Sebaliknya, produk dengan tingkat penjualan Rendah

hanya memiliki probabilitas 0,0625 pada kelas Laris, tetapi mencapai 0,6429 pada kelas Tidak Laris. Hasil ini menunjukkan bahwa jumlah produk terjual merupakan salah satu atribut yang paling kuat dalam membedakan dua kelas. Semakin tinggi tingkat penjualan suatu produk, maka semakin besar peluang produk tersebut diklasifikasikan sebagai laris.

Pada atribut Pendapatan, pola yang muncul juga cukup jelas. Kategori pendapatan Tinggi memiliki probabilitas 0,5 pada kelas Laris, sedangkan pada kelas Tidak Laris hanya 0,1429. Sebaliknya, pendapatan Rendah lebih banyak muncul pada kelas Tidak Laris, yaitu sebesar 0,5, dibandingkan pada kelas Laris yang hanya 0,1875. Hal ini menunjukkan bahwa semakin tinggi pendapatan yang dihasilkan suatu produk, semakin besar pula kecenderungannya untuk termasuk ke dalam kategori laris.

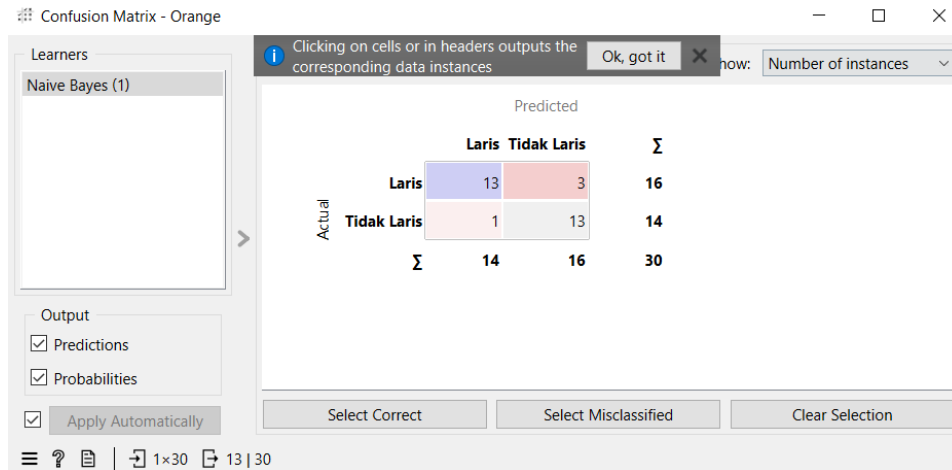
Sementara itu, pada atribut Promo, hasil likelihood menunjukkan bahwa produk dengan status promo Ya memiliki probabilitas 0,625 pada kelas Laris, sedangkan pada kelas Tidak Laris hanya 0,2857. Sebaliknya, status promo Tidak lebih dominan pada kelas Tidak Laris dengan probabilitas 0,7143. Temuan ini memperlihatkan bahwa promosi memiliki kontribusi yang cukup penting dalam mendorong produk masuk ke kategori laris. Dengan demikian, promo dapat dipandang sebagai salah satu faktor pendukung dalam meningkatkan penjualan produk pada Coffeeshop Bobskuy.

Secara keseluruhan, output probabilitas pada Orange menunjukkan bahwa atribut yang paling menonjol dalam membedakan kelas Laris dan Tidak Laris

adalah Produk Terjual, Pendapatan, dan Promo. Sementara itu, atribut Harga cenderung memiliki distribusi yang lebih seimbang antar kelas sehingga kekuatannya sebagai pembeda relatif lebih lemah. Temuan ini memperlihatkan bahwa model Naïve Bayes tidak hanya memberikan hasil klasifikasi, tetapi juga dapat membantu menjelaskan pola faktor-faktor yang memengaruhi kelarisan produk secara probabilistik.

4.7. Confusion Matrix

Gambar 4. 17 Hasil Confusion Matrix



Gambar 4.17 memperlihatkan *confusion matrix* hasil klasifikasi menggunakan algoritma Naïve Bayes pada dua kelas, yaitu Laris dan Tidak Laris. Berdasarkan matriks tersebut, pada kelas Laris terdapat 16 data aktual, di mana 13 data berhasil diklasifikasikan dengan benar sebagai Laris, sedangkan 3 data lainnya salah diklasifikasikan sebagai Tidak Laris. Pada kelas Tidak Laris terdapat 14 data aktual, dengan 13 data berhasil diprediksi secara benar sebagai Tidak Laris dan 1 data salah diprediksi sebagai Laris.

Secara total, dari 30 data yang dievaluasi, model menghasilkan 26 prediksi benar dan 4 prediksi salah. Hasil ini menunjukkan bahwa model memiliki kinerja klasifikasi yang tergolong baik dalam membedakan kedua kelas, dengan nilai:

TP (True Positive) = 13, yaitu produk yang benar-benar “Laris” dan diprediksi “Laris”.

FP (False Positive) = 1, yaitu produk yang diprediksi “Laris”, tetapi sebenarnya “Tidak Laris”.

FN (False Negative) = 3, yaitu produk yang diprediksi “Tidak Laris”, tetapi sebenarnya “Laris”.

TN (True Negative) = 13, yaitu produk yang benar-benar “Tidak Laris” dan diprediksi “Tidak Laris”

4.8. Analisis Faktor-Faktor yang Mempengaruhi Klasifikasi Produk

Berdasarkan hasil pengolahan data dan pengujian menggunakan algoritma Naïve Bayes pada software Orange, dapat diketahui bahwa klasifikasi produk ke dalam kategori Laris dan Tidak Laris dipengaruhi oleh beberapa atribut yang terdapat dalam dataset, yaitu kategori produk, jumlah produk terjual, pendapatan, dan promo. Dari atribut-atribut tersebut, produk terjual menjadi faktor yang paling kuat dalam membedakan produk yang masuk kategori laris dan tidak laris, karena jumlah penjualan secara langsung menunjukkan tingkat minat konsumen terhadap suatu produk.

Selain itu, atribut pendapatan juga memberikan kontribusi dalam proses klasifikasi. Produk dengan jumlah penjualan yang tinggi cenderung menghasilkan pendapatan yang lebih besar, sehingga peluang produk tersebut untuk diklasifikasikan sebagai Laris menjadi lebih tinggi. Sebaliknya, produk dengan jumlah penjualan rendah dan pendapatan yang lebih kecil cenderung masuk ke dalam kategori Tidak Laris. Hal ini menunjukkan bahwa pola klasifikasi yang

dihasilkan model Naïve Bayes terbentuk dari hubungan probabilistik antaratribut dalam data penjualan.

Atribut promo juga turut memengaruhi hasil klasifikasi, meskipun pengaruhnya tidak sebesar jumlah produk terjual. Pada beberapa data, produk yang mendapatkan promo memiliki kecenderungan penjualan yang lebih baik dibandingkan produk tanpa promo. Namun demikian, promo bukan satu-satunya penentu, karena hasil klasifikasi tetap dipengaruhi oleh kombinasi seluruh atribut yang digunakan dalam pemodelan. Dengan kata lain, keberadaan promo dapat memperkuat peluang suatu produk menjadi Laris, tetapi tidak secara otomatis menjamin produk tersebut akan masuk ke kategori tersebut.

Sementara itu, atribut kategori menunjukkan bahwa jenis produk juga ikut berperan dalam proses klasifikasi. Produk yang termasuk kategori minuman dan makanan memiliki pola penjualan yang berbeda-beda, sehingga model memanfaatkan informasi kategori untuk membantu membedakan karakteristik masing-masing produk. Akan tetapi, keputusan akhir klasifikasi tetap ditentukan oleh keterkaitan seluruh atribut secara bersama-sama, bukan hanya berdasarkan satu kategori tertentu saja.

Berdasarkan hasil pengolahan data, model lebih banyak memanfaatkan atribut yang memiliki kontribusi informasi lebih jelas, yaitu jumlah produk terjual, pendapatan, promo, dan kategori produk. Kondisi ini menunjukkan bahwa pembentukan kelas Laris dan Tidak Laris pada penelitian ini lebih dipengaruhi oleh variasi penjualan dan karakteristik produk daripada oleh atribut harga.

4.9. Pemanfaatan Hasil Analisis untuk Strategi Penjualan

Berdasarkan hasil analisis menggunakan algoritma Naïve Bayes, diketahui bahwa atribut produk terjual, pendapatan, promo, dan kategori produk memiliki peran dalam menentukan klasifikasi produk ke dalam kategori Laris dan Tidak Laris. Hasil analisis ini dapat dimanfaatkan oleh pihak Coffeeshop Bobskuy sebagai dasar dalam menyusun strategi penjualan yang lebih efektif dan tepat sasaran. Dengan mengetahui karakteristik produk yang cenderung laris maupun tidak laris, pihak usaha dapat mengambil keputusan yang lebih akurat dalam pengelolaan produk.

Produk yang termasuk dalam kategori Laris dapat dipertahankan ketersediaannya agar mampu memenuhi permintaan konsumen secara berkelanjutan. Selain itu, produk-produk tersebut juga dapat dijadikan sebagai produk unggulan dalam strategi promosi, karena telah terbukti memiliki tingkat penjualan yang baik. Dengan demikian, hasil klasifikasi tidak hanya berfungsi sebagai informasi data, tetapi juga sebagai bahan pertimbangan dalam menentukan prioritas penjualan.

Sementara itu, produk yang masuk dalam kategori Tidak Laris dapat menjadi bahan evaluasi bagi pihak Coffeeshop Bobskuy. Evaluasi tersebut dapat dilakukan melalui beberapa langkah, seperti memperbaiki strategi promosi, mengubah cara penyajian produk, memberikan paket bundling, atau meninjau kembali daya tarik produk tersebut di mata konsumen. Dengan memanfaatkan hasil

analisis klasifikasi, pihak usaha dapat lebih mudah mengidentifikasi produk yang membutuhkan perhatian khusus agar penjualannya dapat ditingkatkan.

Selain itu, atribut promo yang ikut memengaruhi hasil klasifikasi menunjukkan bahwa program promosi dapat dimanfaatkan sebagai salah satu strategi untuk meningkatkan minat beli konsumen. Produk yang sebelumnya tergolong Tidak Laris dapat didorong penjualannya melalui pemberian promo tertentu, sehingga peluang produk tersebut untuk lebih dikenal dan dibeli konsumen menjadi lebih besar. Dengan demikian, hasil analisis ini dapat membantu dalam menentukan produk mana yang layak dipromosikan secara lebih intensif.

Dari sisi kategori produk, hasil analisis juga dapat digunakan untuk melihat kecenderungan minat konsumen terhadap jenis produk tertentu, baik makanan maupun minuman. Informasi ini bermanfaat dalam perencanaan stok, penyusunan menu, serta pengembangan produk baru yang sesuai dengan pola pembelian konsumen. Dengan memahami kategori produk yang lebih diminati, pihak Coffeeshop Bobskuy dapat menyusun strategi penjualan yang lebih terarah dan efisien.