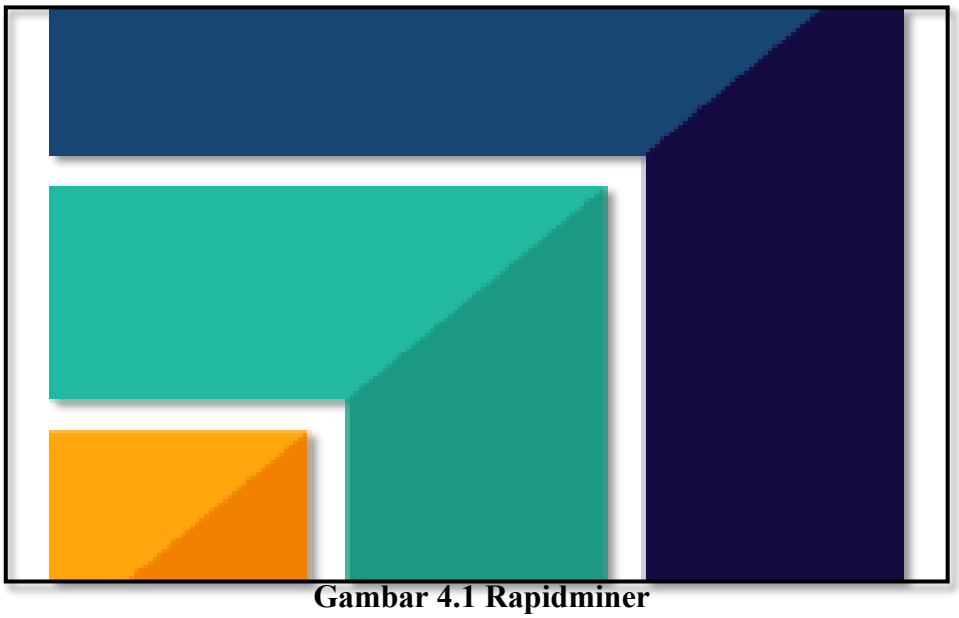
BAB IV

IMPLEMENTASI DAN PEMBAHASAN

4.1 Implementasi Sistem

Pada bab ini, dilakukan implementasi dari data tersebut dengan menggunakan aplikasi Rapid Miner. Rapid Miner merupakan perangkat lunak yang digunakan untuk mengolah data secara sistematis dengan menerapkan prinsip-prinsip serta algoritma yang berasal dari teknik *data mining*. Aplikasi ini memungkinkan pengguna untuk melakukan berbagai analisis data, seperti klasifikasi, klastering, asosiasi, dan prediksi, melalui antarmuka visual yang memudahkan dalam menyusun alur kerja analisis tanpa perlu menulis kode secara langsung.



Gambar 4.1 Rapidminer

Langkah Awal pada penelitian ini terdapat proses seleksi data. Proses seleksi data ini merupakan proses pemngambilan data testing. Pemilihan data ini diambil secara acak untuk dan diwakili 25% dari keseluruhan data yang diperoleh.

Selain itu data kuisioner juga sudah di kelompokkan berdasarkan atribut Kasih Sayang dan Proteksi. Dari penjelesan di atas ditunjukkan dengan gambar di bawah ini:

1	No	Nama / Inisial	Status	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
2	1	R1	TB	3	1	3	2	3	4	3	1	2	2	3	3	3	2	4	3	3	3	3	3	4	3	4	2	1
3	2	R2	TB	3	1	3	2	3	4	2	3	4	3	2	3	3	3	2	3	4	4	4	3	4	4	4	2	3
4	3	R3	Psh	4	1	4	1	4	4	3	1	3	2	3	4	4	2	4	4	4	3	3	3	1	3	4	3	2
5	4	R4	Psh	4	1	2	1	3	4	2	3	3	2	3	4	3	1	3	3	4	4	4	1	2	2	4	3	4
6	5	R5	TB	4	1	1	1	4	4	2	1	4	3	4	4	4	1	3	4	4	3	3	3	4	3	4	2	1
7	6	R6	Psh	3	1	1	1	3	4	3	1	4	3	4	4	4	2	1	3	4	4	4	4	3	2	4	4	4
8	7	R7	TB	4	1	3	1	4	4	2	1	2	1	4	4	4	1	2	4	4	4	2	2	3	2	4	4	3
9	8	R8	Psh	3	3	4	2	3	3	4	1	1	1	2	2	1	3	4	2	3	2	4	4	1	1	4	2	1
10	9	R9	TB	3	2	4	1	4	4	4	1	2	2	3	3	1	2	4	4	4	4	4	4	1	3	4	3	1
11	10	R10	TB	3	2	4	2	3	4	4	1	1	1	2	2	1	3	4	2	2	3	4	4	1	1	3	3	1
12	11	R11	Psh	3	3	4	3	2	3	4	1	1	2	3	2	1	2	3	2	2	2	4	4	2	1	3	2	1
13	12	R12	Psh	4	2	4	1	3	4	4	1	2	2	3	3	1	2	4	4	4	2	4	4	2	1	4	3	1
14	13	R13	TB	4	3	4	1	4	4	4	1	3	2	2	3	1	3	4	3	4	2	4	4	2	2	4	3	1
15	14	R14	Psh	2	2	4	2	2	4	4	1	1	1	2	2	1	1	4	2	4	2	4	4	2	1	4	2	1
16	15	R15	TB	4	3	4	1	3	4	4	1	3	3	3	3	2	3	3	4	3	4	3	3	3	4	4	3	2
17	16	R16	Psh	3	2	4	2	3	4	3	2	2	3	3	3	2	1	4	3	3	2	3	4	2	2	4	3	2
18	17	R17	TB	3	2	2	2	2	4	3	3	4	3	2	2	1	3	4	3	4	3	4	3	4	4	4	3	3
19	18	R18	Psh	3	3	4	1	3	4	4	2	3	3	2	2	3	3	3	2	2	3	4	4	2	1	4	2	1
20	19	R19	TB	4	1	4	2	4	4	4	2	3	2	4	4	3	1	3	4	4	4	4	3	4	4	4	3	4
21	20	R20	Psh	3	2	4	2	3	3	4	1	2	2	2	2	1	4	3	2	2	2	4	4	2	1	2	1	1
22	21	R21	TB	3	2	4	2	3	4	4	1	3	2	3	3	2	3	4	2	3	3	4	4	2	3	4	2	2
23	22	R22	TB	4	1	4	1	3	4	2	3	2	1	2	2	3	3	2	3	3	2	4	4	2	2	3	2	1
24	23	R23	TB	4	3	4	2	2	4	4	1	1	2	3	3	1	1	4	2	4	3	4	4	3	3	4	3	2
25	24	R24	TB	4	1	4	2	4	4	4	1	2	2	3	3	2	2	4	4	4	3	4	4	3	3	4	3	3
26	25	R25	TB	3	1	4	2	3	3	4	1	2	2	3	2	2	2	3	3	3	3	4	4	3	3	4	2	2
27	26	R26	TB	3	2	4	2	4	4	4	1	3	3	3	2	2	1	4	3	3	3	4	4	3	2	4	2	3
28	27	R27	TB	4	2	4	1	4	4	3	2	3	1	3	2	2	2	2	2	3	2	3	3	4	2	3	3	3
29	28	R28	TB	4	2	4	1	3	4	3	1	2	2	3	3	2	2	3	3	3	3	4	3	3	4	4	4	4
30	29	R29	Psh	3	3	4	1	3	4	4	1	2	2	3	2	2	4	4	3	4	2	4	4	1	1	3	2	1
31	30	R30	TB	4	1	4	1	4	4	2	1	3	2	4	4	2	1	4	4	4	4	3	2	4	4	4	4	4
32	31	R31	TB	3	2	3	2	2	4	4	1	3	2	3	3	2	3	4	3	2	3	3	4	3	4	4	2	4
33	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
34	101	R101	TB	4	1	4	1	4	4	4	1	3	1	4	4	1	1	4	4	4	4	4	4	1	1	4	3	1

Gambar 4.2 Data Sebelum Seleksi

Data diperoleh dari hasil penyebaran kuesioner yang diisi oleh sebanyak 101 responden melalui platform Google Form. Selanjutnya, data yang terkumpul

dianalisis berdasarkan dua kelompok atribut, yaitu Kasih Sayang dan Proteksi. Untuk menjaga kerahasiaan dan kenyamanan para responden, identitas mereka diubah menjadi kode anonim, seperti R1, R2, dan seterusnya. Penggunaan kode ini bertujuan agar tidak mendiskreditkan individu tertentu, khususnya terkait dengan apakah pola asuh dalam keluarganya berpengaruh terhadap kesejahteraan emosional yang dimilikinya atau tidak.

Tabel 4.1 Data Setelah Diseleksi

Kasih Sayang	Proteksi	Tingkat Emosional
30	38	Normal
33	43	Pengaruh
34	40	Pengaruh
32	38	Pengaruh
33	39	Pengaruh
32	43	Pengaruh
31	39	Pengaruh
29	32	Normal
33	39	Pengaruh
29	32	Normal
31	29	Normal
33	36	Normal
35	37	Pengaruh
27	32	Normal
36	41	Pengaruh
34	35	Normal
32	43	Pengaruh
34	34	Normal
38	45	Pengaruh
30	29	Normal

34	38	Pengaruh
29	34	Normal
33	38	Pengaruh
34	43	Pengaruh
30	38	Normal

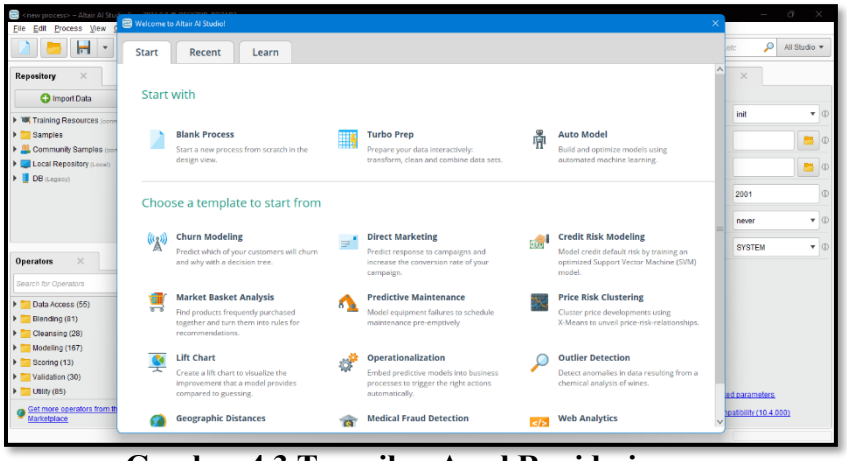
Pengaruh tingkat emosional remaja terhadap pola asuh orangtua dikategorikan berdasarkan total nilai kuisioner berjumlah 70. Dengan total nilai diatas 70 maka data dikategorikan menjadi tingkat emosional yang berpengaruh.

4.2 Teknik Pengujian

proses pengolahan data dalam penelitian ini terdiri dari beberapa tahapan penting yang harus dilakukan secara sistematis agar hasil yang diperoleh akurat dan dapat dipertanggungjawabkan. Salah satu tahapan awal yang sangat krusial dalam proses tersebut adalah *preprocessing* data, di mana data mentah yang telah dikumpulkan akan dipersiapkan dan dibersihkan sebelum masuk ke tahap analisis utama. Pada tahapan ini, peneliti menggunakan bantuan perangkat lunak RapidMiner sebagai alat bantu dalam pengolahan data.

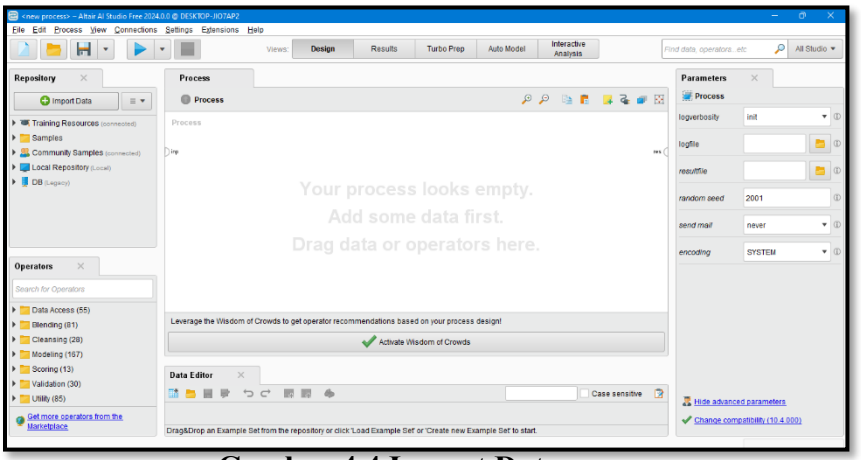
Untuk memulai proses pengolahan data menggunakan RapidMiner, langkah pertama yang dilakukan adalah menjalankan aplikasi dan memilih opsi Blank Process pada tampilan pop-up awal yang muncul. Opsi ini digunakan untuk membuat proses baru secara manual, sehingga pengguna dapat menyusun dan mengatur sendiri tahapan alur analisis yang akan dilakukan sesuai dengan kebutuhan dan metode yang digunakan dalam penelitian. Tampilan awal ini memberikan fleksibilitas kepada pengguna dalam merancang alur kerja mulai dari

proses *retrieving data*, *splitting data*, penerapan model, hingga evaluasi kinerja model yang dibangun. Proses awal ini dapat dilihat secara visual pada gambar berikut, yang menunjukkan bagaimana aplikasi RapidMiner disiapkan sebelum memulai proses analisis yang lebih lanjut.



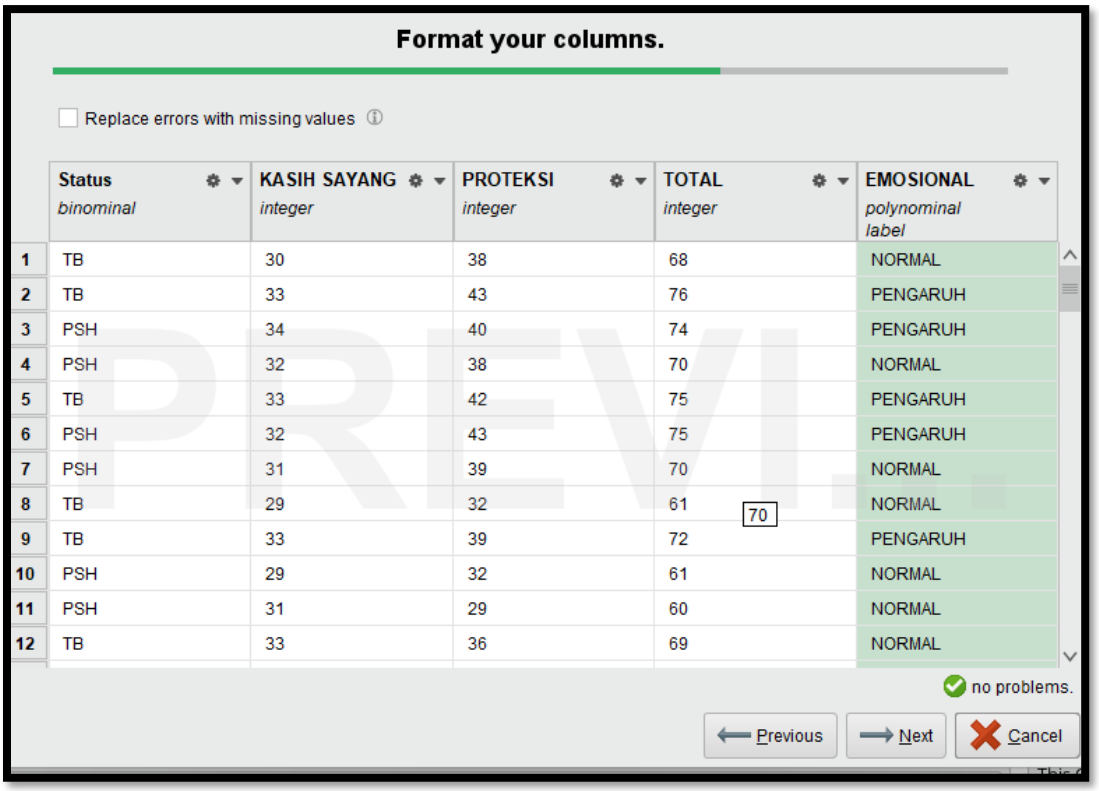
Gambar 4.3 Tampilan Awal Rapidminer

Setelah itu untuk pengolahan data, data di impor kedalam rapidminer dengan cara memilih menu import data kemudian memilih data berdasarkan lokasi tempat data yang sudah di siapkan diletakkan. Data yang sudah di siapkan dapat dimasukkan langsung kedalam rapidminner.



Gambar 4.4 Import Data

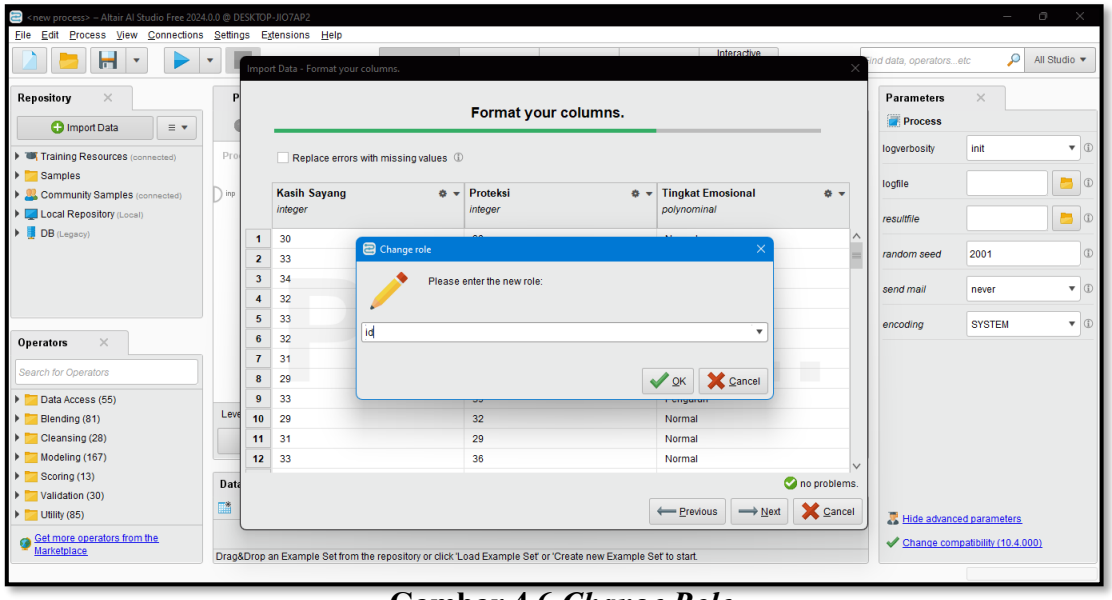
Setelah data diimpor, tahap berikutnya adalah seleksi cell, yaitu memilih kolom-kolom yang relevan yang akan digunakan dalam analisis. Misalnya, hanya kolom 'Kasih Sayang', 'Proteksi', dan 'Total' yang dipilih, sementara kolom seperti 'Nama' atau 'Jenis Kelamin' yang tidak diperlukan dalam proses clustering diabaikan. Seleksi ini bertujuan untuk fokus pada atribut yang berpengaruh terhadap pembentukan cluster. Pada tahap *select the cell to import*, tahap ini merupakan tahap untuk memastikan pada bagian mana *cell* yang ingin digunakan. Jika ada data yang tidak ingin digunakan dapat diubah pada *cell range*. Apabila data sudah tepat, maka dapat mengklik *next* untuk ketahapan berikutnya.



Gambar 4.5 *Select Cell*

Change Role digunakan untuk mengubah peran atribut dalam dataset. Misalnya, kolom 'Total' dapat diubah menjadi 'label' bila akan digunakan dalam

supervised learning. Namun dalam konteks clustering (unsupervised), biasanya semua atribut tetap bertindak sebagai regular attribute. Ini penting supaya algoritma dapat mengelompokkan data berdasarkan kemiripan tanpa adanya target label.Tahapan selanjutnya adalah pengubahan role pada bagian tingkat emosional. Hal ini dilakukan agar data yang diproses hanya data yang bersifat angka atau dalam bentuk integer. Sedangkan untuk tingkat emosional daa berbentuk *polynominal*. Untuk pengubahan *role* dapat dilakukan dengan cara menekan tombol panah kearah bawah pada kolom tingkat emosional dan memilih *change role* lalu mengubahnya menjadi id. Untuk penamaan role dapat diubah sesuai dengan keinginan. Hal ini hanya untuk memudahkan pengolahan data dalam mengidentifikasi data.

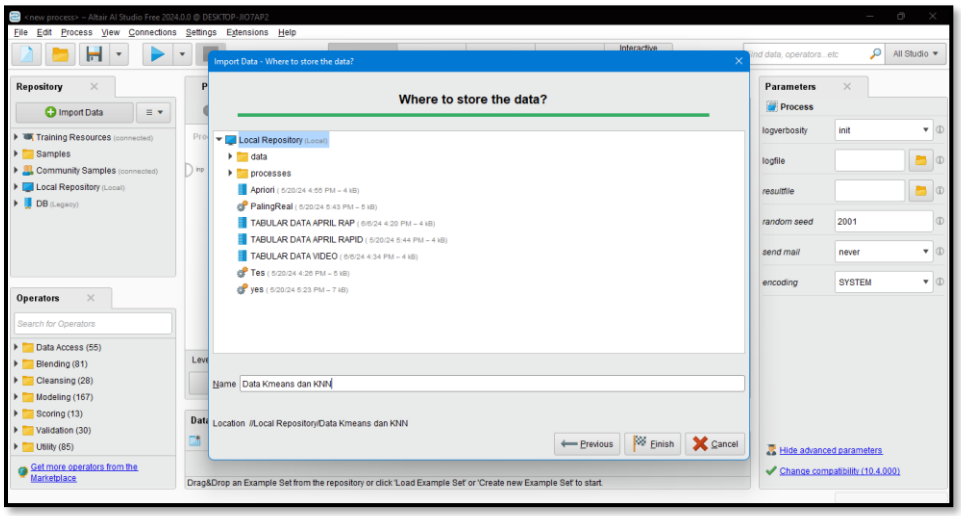


Gambar 4.6 *Change Role*

Dalam proses clustering menggunakan operator K-Means, ditentukan jumlah cluster yang diinginkan (k). Misalnya, peneliti dapat menetapkan k = 2 untuk membagi data ke dalam dua kelompok kesejahteraan emosional: normal dan bermasalah. Algoritma kemudian akan berusaha meminimalkan jarak antar anggota

dalam cluster yang sama dan memaksimalkan jarak antar cluster berbeda. Pada tahap ini, ditentukan lokasi penyimpanan data sementara atau hasil olahan dalam RapidMiner. Ini memudahkan tracing atau audit data saat ingin melakukan analisis lebih lanjut atau membandingkan hasil eksperimen di masa depan. Tahap terakhir dalam proses ini adalah menentukan lokasi penyimpanan data yang akan digunakan dalam pengolahan. Pemilihan lokasi penyimpanan bersifat fleksibel, artinya pengguna bebas memilih direktori atau folder mana saja, selama lokasi tersebut mudah diakses dan dapat menunjang kelancaran proses analisis data. Hal ini penting untuk memastikan efisiensi kerja, terutama ketika pengguna perlu mengakses kembali data yang sama di lain waktu.

Setelah lokasi penyimpanan dipilih sesuai kebutuhan, langkah selanjutnya adalah menekan tombol *Finish* pada antarmuka RapidMiner. Setelah tombol tersebut diklik, sistem akan secara otomatis membuka lembar kerja (*workspace*) yang menjadi tempat utama dalam melakukan proses analisis. Pada tahap ini, pengguna sudah dapat mulai menyusun alur kerja dan mengaplikasikan berbagai operator analisis sesuai dengan metode yang akan digunakan.



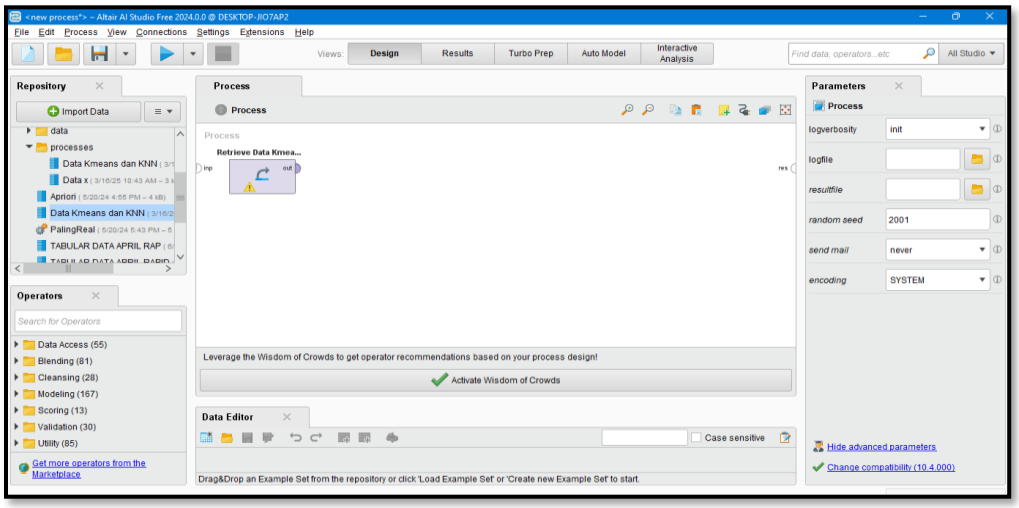
Gambar 4.7 *Select Location*

4.2.1 Pengujian *K-Means Clustring*

Pada tahap awal, data hasil kuesioner yang telah dikumpulkan diimpor ke Rapid Miner menggunakan fitur "Import Data". Data ini biasanya dalam format CSV atau Excel. Pada proses ini, perlu diperhatikan kesesuaian tipe data untuk setiap atribut, misalnya apakah bertipe numerik atau kategorikal. Ini penting untuk menghindari kesalahan saat proses clustering nantinya. Pada tahap pengujian pertama, data dianalisis menggunakan metode *K-Means Clustering*, yang merupakan salah satu teknik *unsupervised learning* dalam data mining. Metode ini digunakan untuk mengelompokkan data ke dalam beberapa kelompok (*cluster*) berdasarkan kemiripan atau kedekatan nilai dari atribut-atribut yang dimiliki. Dalam konteks penelitian ini, *K-Means* digunakan untuk melihat pola-pola tertentu dari data kuesioner yang telah dikumpulkan, khususnya dalam kaitannya dengan atribut *Kasih Sayang* dan *Proteksi*, guna mengetahui kemungkinan adanya pengelompokan responden berdasarkan pola asuh yang mereka terima.

Langkah pertama dalam pengujian ini dilakukan dengan masuk ke dalam menu *Design* pada aplikasi RapidMiner. Pada menu ini, pengguna dapat mengatur dan menyusun alur proses analisis data secara visual. Selanjutnya, data yang akan dianalisis dimasukkan ke dalam lembar kerja (*workspace*) dengan cara menarik dan meletakkan (*drag and drop*) file data yang telah disiapkan sebelumnya. Proses ini bertujuan untuk mempersiapkan data agar dapat diolah lebih lanjut menggunakan operator-operator analisis yang tersedia di RapidMiner. Dengan menggunakan

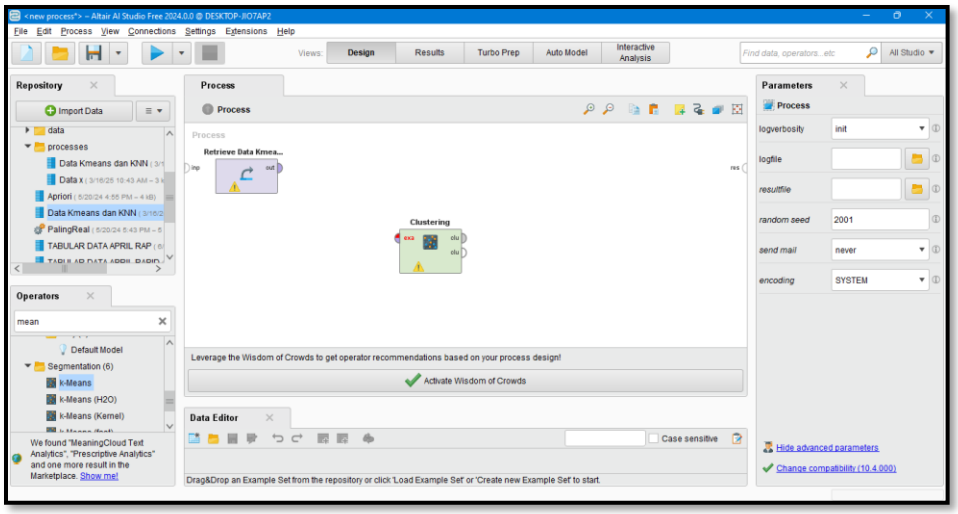
antarmuka yang intuitif, pengguna dapat dengan mudah mengatur jalannya proses analisis tanpa harus menuliskan kode secara manual.



Gambar 4.8 Olah Data

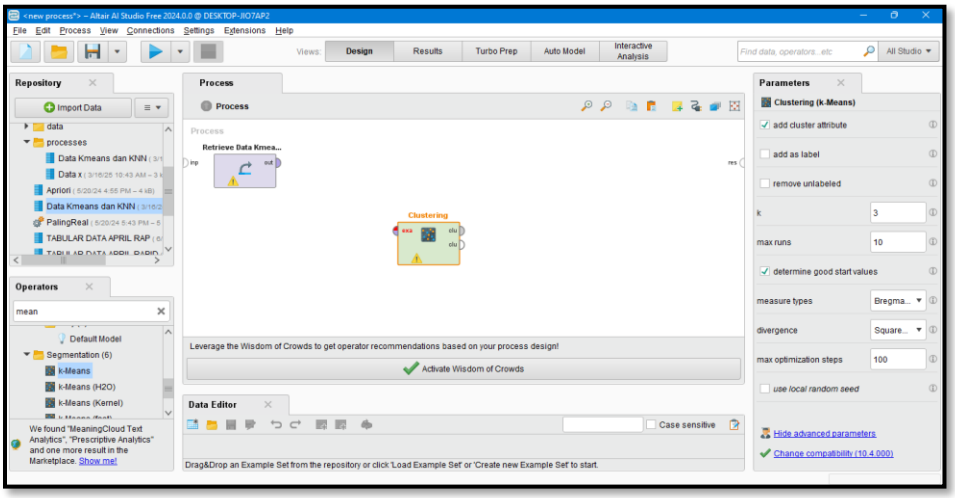
Dalam proses clustering menggunakan operator K-Means, ditentukan jumlah cluster yang diinginkan (k). Misalnya, peneliti dapat menetapkan k = 2 untuk membagi data ke dalam dua kelompok kesejahteraan emosional: normal dan bermasalah. Algoritma kemudian akan berusaha meminimalkan jarak antar anggota dalam cluster yang sama dan memaksimalkan jarak antar cluster berbeda. Tahapan selanjutnya dalam proses pengujian *K-Means Clustering* adalah memasukkan operator pengolahan data ke dalam lembar kerja RapidMiner. Operator ini berfungsi sebagai alat untuk menjalankan proses clustering sesuai metode yang diinginkan. Pada tahap ini, pengguna dapat mencari operator dengan mengetikkan kata kunci *"K-Means"* pada kolom pencarian yang tersedia di panel *Operators*. Setelah ditemukan, operator K-Means tersebut kemudian dimasukkan ke dalam *workspace* dengan cara menyeret dan meletakkannya (*drag and drop*) seperti halnya saat memasukkan data sebelumnya.

Langkah ini merupakan bagian penting dari proses desain alur analisis, karena operator *K-Means* inilah yang akan mengatur mekanisme pengelompokan data berdasarkan parameter-parameter yang akan ditentukan selanjutnya, seperti jumlah cluster dan atribut yang digunakan. Dengan menyusun data dan operator dalam satu lembar kerja, RapidMiner akan dapat menjalankan proses pengolahan data secara terstruktur dan efisien.



Gambar 4.9 *Clustering*

Parameter penting seperti jumlah cluster, metode inisialisasi centroid, dan kriteria penghentian iterasi diatur pada tahap ini. Misalnya, bisa diatur iterasi maksimal 100 kali atau toleransi konvergensi di angka 0.0001, agar proses clustering berhenti saat sudah cukup stabil.Berikutnya adalah mengatur clustering dengan menekan menu clustering kemudian akan muncul pilihan pada tampilan sebelah kanan lembar kerja. Untuk pilihan k merupakan penentuan banyaknya kluster yang kita butuhkan dan max runs merupakan benyaknya percobaan yang dilakukan. Semakin banyak percobaan yang dilakukan maka semakin tinggi tingkat kebenaran data. Dalam tahapan kali ini kluster data dibagi menjadi 2 bagian.

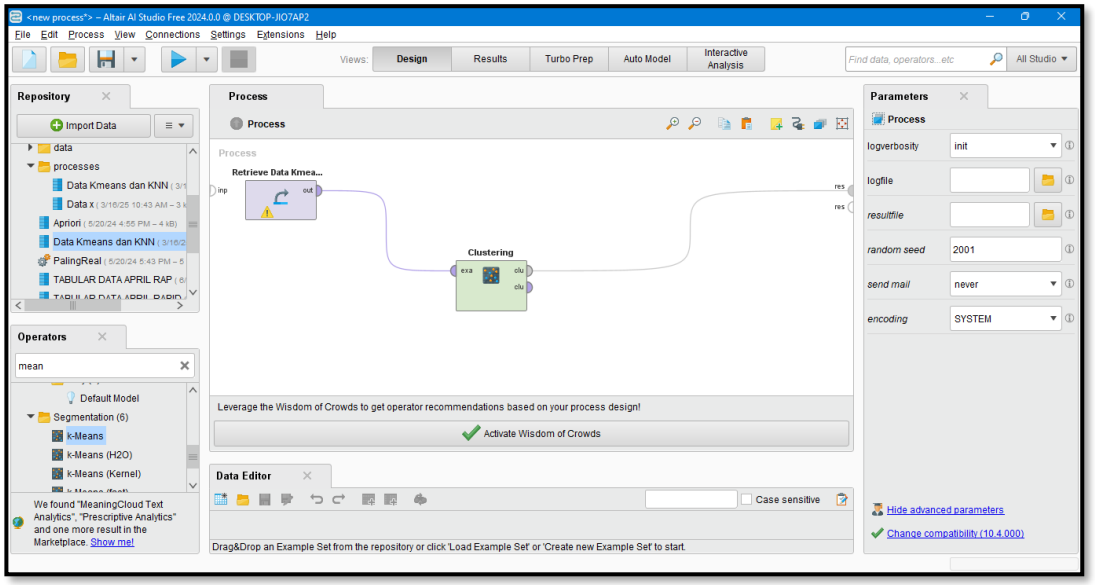


Gambar 4.10 *Parameters Clustering*

Setelah seluruh parameter yang dibutuhkan telah selesai ditentukan dan dipersiapkan dengan baik, tahapan selanjutnya adalah menjalankan proses *clustering*. Dalam konteks ini, peneliti menggunakan algoritma K-Means Clustering yang telah disediakan dalam aplikasi RapidMiner. Pada saat proses ini dijalankan, sistem akan secara otomatis menghitung jarak antara masing-masing data dengan centroid (pusat cluster), kemudian melakukan pembaruan posisi centroid berdasarkan perhitungan rata-rata dari anggota dalam masing-masing cluster yang terbentuk. Setelah proses ini selesai, hasil clustering dapat divisualisasikan oleh aplikasi RapidMiner dalam berbagai format, baik dalam bentuk grafik scatter plot, tabel klasifikasi anggota cluster, ataupun diagram evaluasi performa model, tergantung pada kebutuhan analisis dan preferensi pengguna.

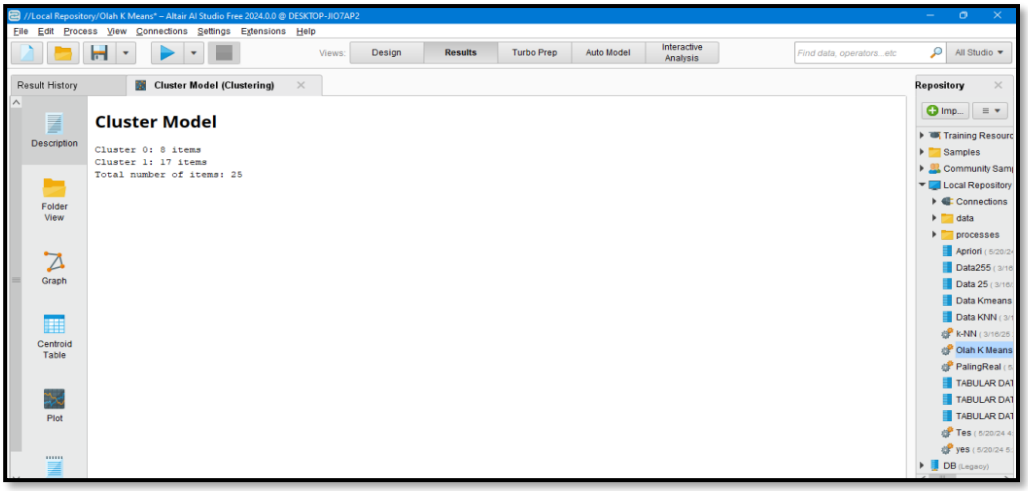
Selanjutnya, tahapan penting lainnya adalah menghubungkan *dataset* yang digunakan dengan operator *clustering process* (dalam hal ini K-Means) yang telah

dimasukkan ke dalam area kerja RapidMiner sebelumnya. Proses ini dilakukan dengan cara menghubungkan komponen data (misalnya operator “Retrieve”) ke input operator K-Means dengan menarik garis koneksi antar node yang tersedia. Setelah koneksi berhasil dibuat secara lengkap, dan semua alur telah tersusun dengan baik, maka data tersebut siap untuk dieksekusi. Pengguna dapat langsung menekan tombol *Run* pada toolbar untuk memulai proses, dan hasil dari clustering akan ditampilkan pada tab *Results*.



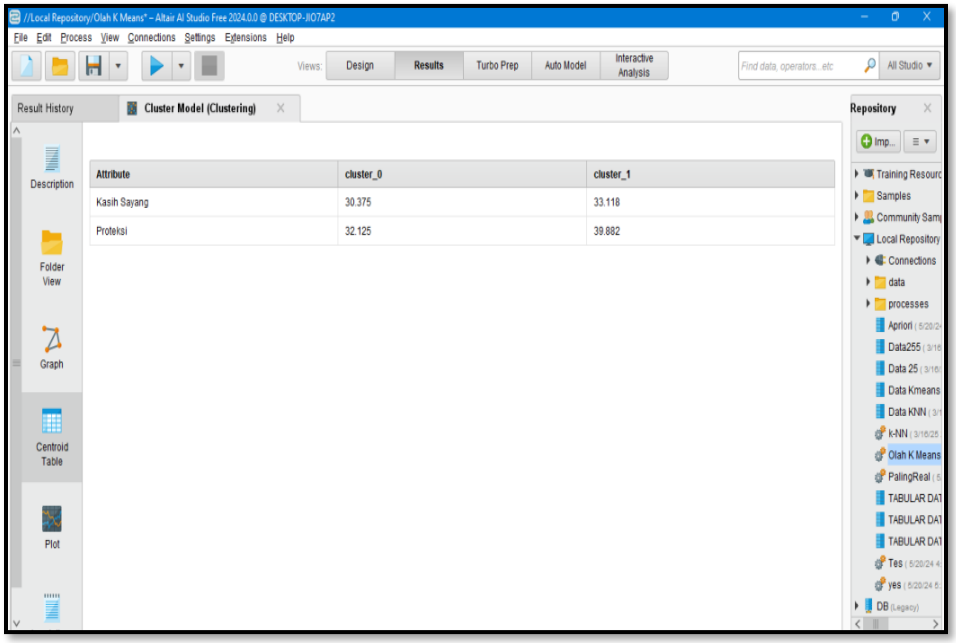
Gambar 4.11 Jalankan Pengolahan

Centroid Table menunjukkan nilai rata-rata atribut pada masing-masing cluster. Dari tabel ini, peneliti bisa menginterpretasikan karakteristik umum tiap cluster, misalnya: cluster 1 memiliki rata-rata kasih sayang lebih tinggi dibanding cluster 2. Setelah berhasil, data akan tampil seperti tampilan dibawah ini. Dari 2 kluster yang dibuat terdapat 8 item pada kluster pertama (0) dan 17 item pada kluster kedua (1).



Gambar 4.12 Hasil Clustering

Pada *centroid table*, data klastering yang ditampilkan merupakan nilai rata-rata dari setiap atribut dalam masing-masing klaster. Nilai-nilai ini dihitung berdasarkan hasil pengelompokan data pada setiap fase dalam proses *clustering*. Tabel centroid berfungsi untuk menunjukkan posisi pusat dari masing-masing klaster, yang menjadi acuan dalam menentukan kedekatan data pada proses selanjutnya. Untuk informasi yang lebih rinci, dapat dilihat pada gambar berikut:

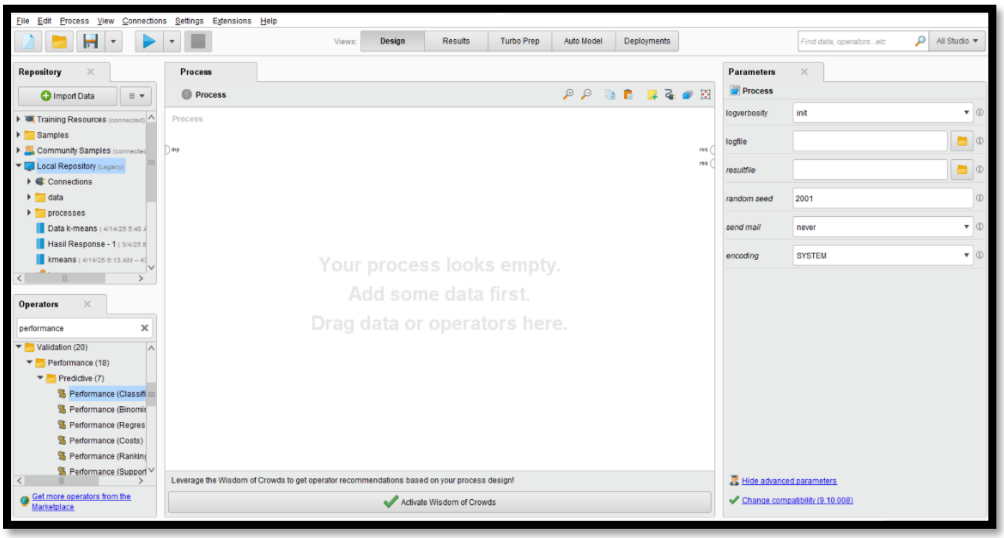


Gambar 4.13 Hasil *Centroid Table*

4.2.2 Pengujian *K-Nearest Neighbor*

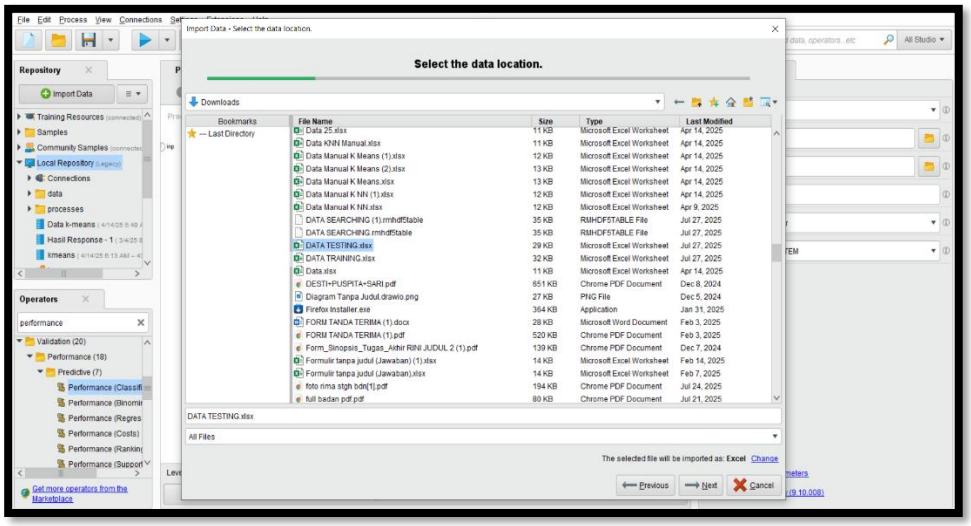
Sama seperti pada proses clustering, tahap awal dalam pengujian KNN adalah mengimpor data ke RapidMiner. File yang berisi data kuesioner yang sudah diproses sebelumnya diunggah. Pada tahap ini, sangat penting untuk memastikan bahwa semua atribut sudah bertipe numerik atau sesuai kebutuhan model, terutama atribut-atribut yang akan digunakan untuk prediksi. Pada tahapan ini, pengujian dilakukan dengan metode KNN. Metode ini digunakan untuk mengklasifikasikan data kuisioner menjadi klasifikasi baru yaitu apakah pengaruh atau tidaknya pola asuh orangtua terhadap kesejahteraan emosional seorang anak remaja.

Untuk tahapan pertama yang dilakukan adalah mengentri beberapa data didalamnya seperti tahapan sebelumnya. Yang pertama ialah memasukkan data yang sama untuk di olah selanjutnya. Kemudian memasukkan beberapa operator seperti Retrive, Split Data, K-NN, Apply Model, dan *Performance* . Untuk lebih jelasnya gambarkan dan hubungkan seperti gambar berikut ini :



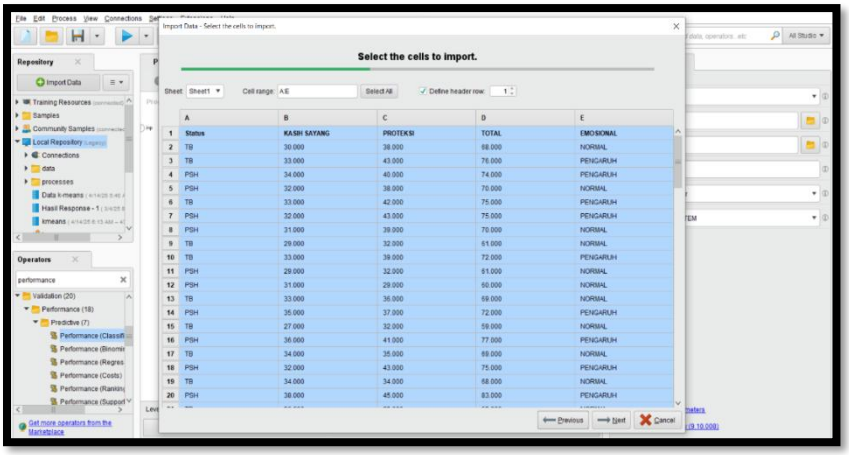
Gambar 4.14 Tampilan awal *workspace* Rapidminer

Pada halaman awal workspace RapidMiner pada tampilan tab *Design*. Halaman ini masih kosong dan belum memiliki alur proses. Langkah selanjutnya yang perlu dilakukan adalah menambahkan data atau operator ke dalam area kerja (*Process*) dengan cara drag & drop dari panel sebelah kiri. Operator seperti *Retrieve*, *Split Data*, *K-NN*, dan *Performance* akan digunakan dalam proses klasifikasi data.



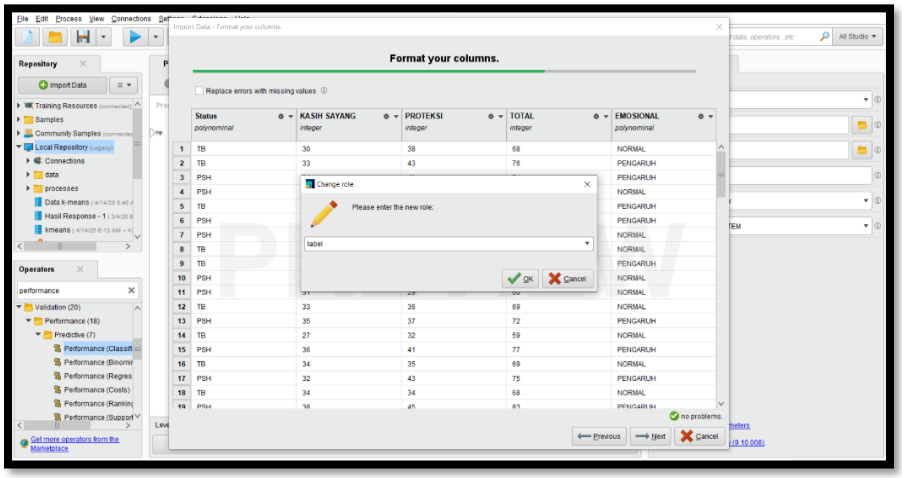
Gambar 4.15 *Implementasi Sistem* atau *Langkah-Langkah Pengujian*

Pada tahap ini, dilakukan proses pemilihan file data yang akan diimpor ke dalam RapidMiner**.** Pengguna memilih file dengan nama DATA TESTING.xlsx dari direktori penyimpanan lokal. File ini berisi data kuesioner yang akan digunakan dalam proses klasifikasi atau pengujian model. Setelah file dipilih, jenis file secara otomatis dikenali sebagai format Excel Worksheet dan akan diimpor ke RapidMiner untuk tahap pemrosesan selanjutnya, seperti pengaturan atribut, pelabelan, dan penerapan algoritma klasifikasi. Pengguna kemudian dapat melanjutkan ke tahap berikutnya dengan menekan tombol Next.



Gambar 4.16 Tahap Import Data

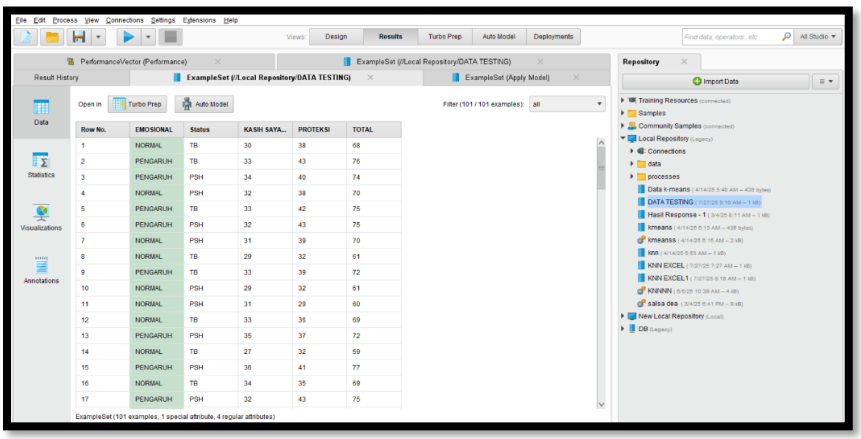
Setelah file data dipilih, tahap selanjutnya adalah menentukan rentang sel yang akan diimpor ke dalam RapidMiner**.** Pada tampilan ini, seluruh kolom dari A hingga E yang mencakup atribut Status**,** KASIH SAYANG**,** PROTEKSI**,** TOTAL, dan EMOSIONAL telah dipilih secara otomatis. memastikan bahwa baris pertama merupakan baris judul kolom (*header)* dengan mencentang opsi *Define header row*. Data ini merupakan hasil kuesioner yang akan digunakan dalam proses klasifikasi untuk mengukur pengaruh pola asuh terhadap kesejahteraan emosional. Setelah data terverifikasi, proses dilanjutkan dengan menekan tombol *Nex*t untuk masuk ke tahap pengaturan atribut.



Gambar 4.17 Proses Klasifikasi

Pada tahap ini, dilakukan proses pengaturan peran atau *role* dari setiap atribut yang terdapat dalam dataset menggunakan aplikasi RapidMiner. Salah satu langkah penting yang harus diperhatikan adalah menentukan atribut mana yang akan dijadikan sebagai *label*, yaitu variabel target yang akan diprediksi dalam proses klasifikasi. Dalam konteks penelitian ini, atribut yang dipilih sebagai label adalah atribut EMOSIONAL, yang berisi nilai-nilai hasil klasifikasi seperti *NORMAL* dan *PENGARUH*. Pemilihan atribut ini sebagai label merupakan langkah yang sangat krusial, karena RapidMiner akan menjadikan atribut tersebut sebagai acuan utama dalam melakukan pembelajaran dan prediksi. Pengaturan role dilakukan melalui operator "*Change Role*", di mana pengguna dapat memilih atribut yang ingin dijadikan label dengan cara mengatur peran kolom tersebut ke opsi *label*.

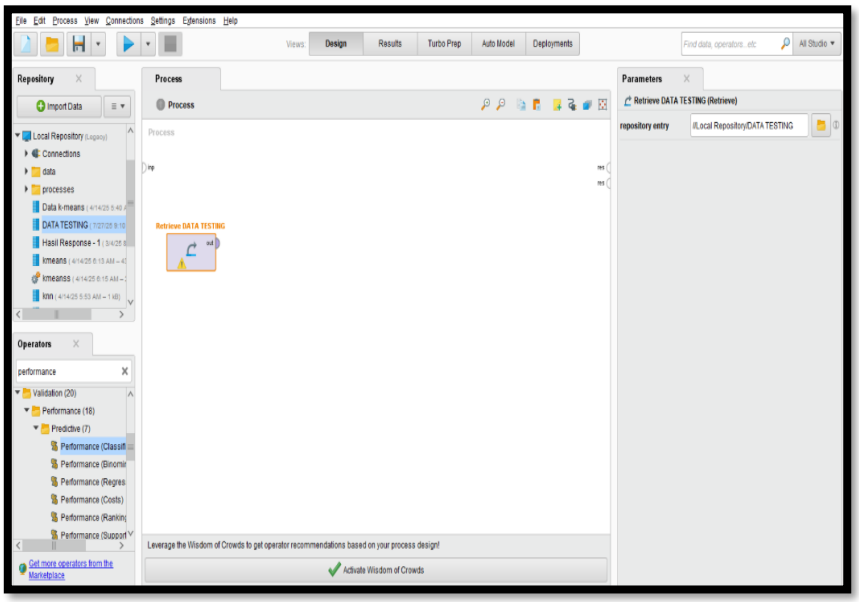
Langkah berikutnya adalah melanjutkan proses dengan menekan tombol *Next* pada tampilan antarmuka, yang akan membawa pengguna ke tahapan selanjutnya dalam desain proses klasifikasi. Tahapan ini memastikan bahwa sistem dapat berjalan sesuai dengan logika pengolahan data yang diinginkan, serta menghasilkan output klasifikasi yang akurat dan relevan dengan tujuan analisis.



Gambar 4.18 Hasil Pengujian Sistem Atau Klasifikasi

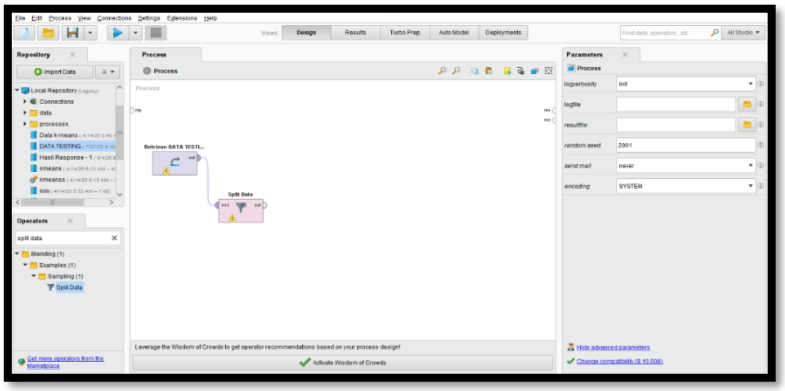
Hasil akhir setelah proses klasifikasi selesai dijalankan di RapidMiner. Tab *Results* menampilkan dataset DATA TESTING yang telah berhasil diproses, dengan informasi lengkap pada setiap baris meliputi atribut EMOSIONAL**,** Status**,** KASIH SAYANG**,** PROTEKSI**,** dan TOTAL. Kolom *EMOSIONAL* yang merupakan hasil klasifikasi berisi dua kategori, yaitu *NORMAL* dan *PENGARUH*, sesuai dengan target klasifikasi yang telah ditentukan sebelumnya. Terdapat 101 baris data yang berhasil dianalisis, ditunjukkan oleh keterangan di bagian bawah. Tampilan ini memungkinkan untuk memverifikasi hasil klasifikasi berdasarkan input yang telah diberikan dan memastikan bahwa data telah diproses dengan benar.

Tampilan area kerja (*workspace*) awal pada tab *Design* di RapidMiner yang masih kosong. Pada tahap ini telah berhasil mengimpor file data DATA TESTING.xlsx ke dalam *Local Repository*, yang dapat dilihat di panel sebelah kiri. Langkah selanjutnya adalah menambahkan operator-operator penting ke area kerja seperti *Retrieve*, *Split Data*, *K-NN*, *Apply Model*, dan *Performance*, dengan cara drag & drop dari panel operator. Setelah semua operator tersusun dan terhubung dengan benar, proses klasifikasi dapat dijalankan untuk memprediksi hasil berdasarkan data yang telah diimpor.



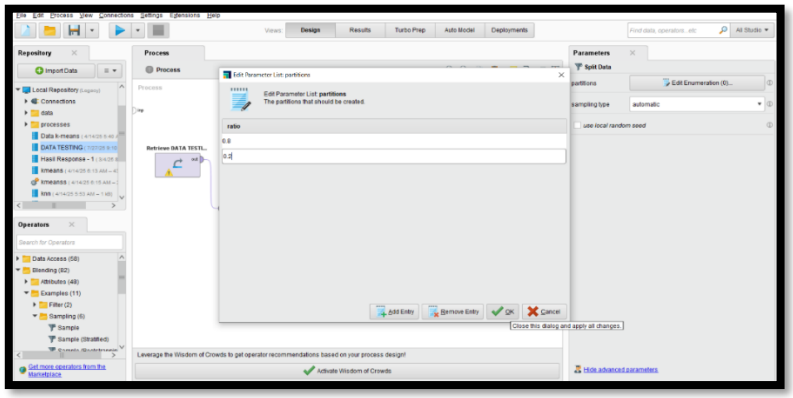
Gambar 4.20 Pemanggilan Data Dalam Sistem Klasifikasi

Retrieve telah ditambahkan ke area kerja (*Process*) di tab *Design* RapidMiner. Operator ini berfungsi untuk mengambil data yang telah diimpor sebelumnya, dalam hal ini file DATA TESTING.xlsx yang tersimpan di *Local Repository*. Pengaturan pada panel kanan menunjukkan bahwa *repository entry* telah diarahkan ke lokasi file tersebut. Munculnya simbol peringatan (ikon segitiga kuning) menandakan bahwa operator ini masih belum terhubung dengan operator lain, sehingga proses belum dapat dijalankan. Langkah berikutnya adalah menambahkan operator lain seperti *Split Data*, *K-NN*, dan *Performance* untuk membentuk alur klasifikasi yang lengkap.



Gambar 4.21 Split data

Selanjutnya, data tersebut diproses menggunakan operator *Split Data* untuk membagi data menjadi beberapa subset, biasanya untuk keperluan pelatihan dan pengujian model. Operator *Split Data* belum dikonfigurasi sepenuhnya, sehingga perlu ditentukan proporsi pembagiannya agar proses dapat dijalankan dengan benar.

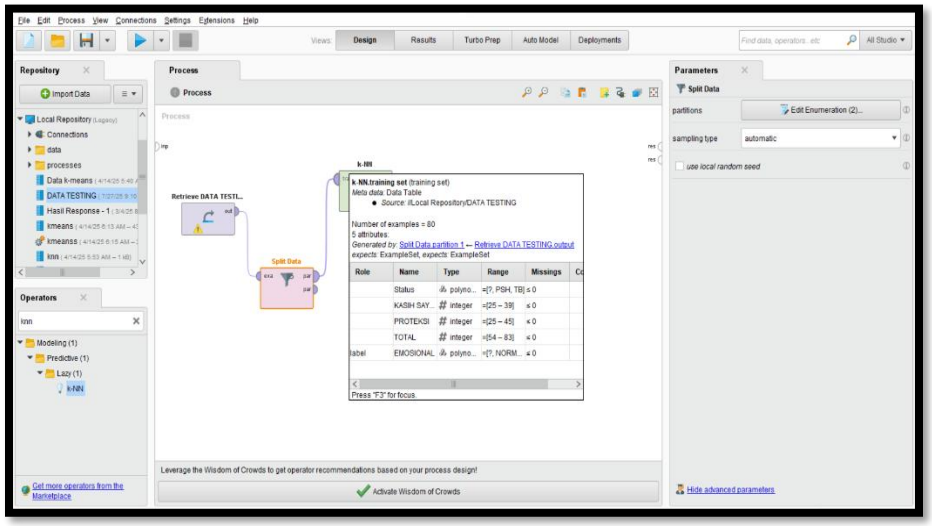


Gambar 4.22 Rasio

Proses selanjutnya adalah membagi data tersebut menggunakan operator *Split Data*. Parameter rasio telah diatur menjadi 0.8 dan 0.2, yang berarti 80% dari data akan digunakan untuk pelatihan model, sedangkan 20% sisanya akan digunakan untuk pengujian model. Pembagian ini penting untuk memastikan bahwa

model dapat dilatih dan dievaluasi secara adil dengan data yang belum pernah dilihat sebelumnya

Pada tahapan Split Data, memilih proporsi 80:20 bertujuan untuk memberi model cukup data dalam proses belajar (training) sambil tetap menyediakan sampel data testing yang representatif. Rasio ini membantu menghindari *overfitting*, yaitu kondisi saat model terlalu cocok dengan data training.

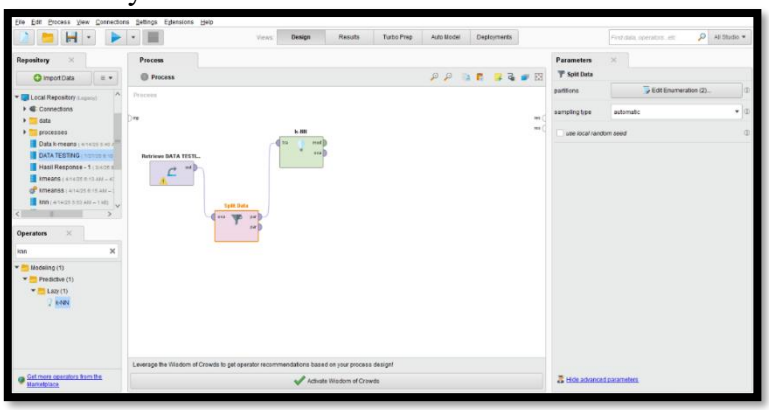


Gambar 4.23 K-NN

Pada langkah pertama dalam alur proses ini, data yang telah disimpan dalam repository lokal dengan nama *DATA TESTING* diambil menggunakan operator *Retrieve*. Data yang diambil kemudian dibagi menjadi dua bagian menggunakan operator *Split Data*. Pembagian ini dilakukan dengan rasio 80% untuk data pelatihan (*training set*) dan 20% untuk data pengujian (*test set*). Bagian yang digunakan untuk pelatihan, yang berjumlah 80 contoh data, selanjutnya diproses menggunakan algoritma klasifikasi *k-Nearest Neighbors (k-NN)*. Algoritma ini dipilih untuk memprediksi nilai label berdasarkan kedekatan data dengan tetangga terdekatnya dalam ruang fitur.

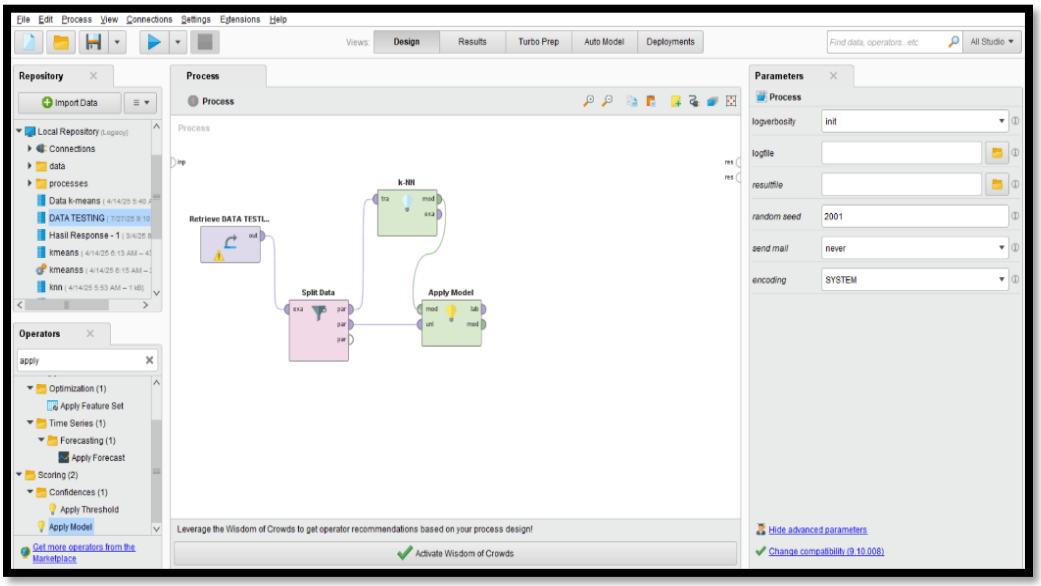
Data yang digunakan untuk pelatihan ini terdiri dari lima atribut yang relevan, yaitu *Status*, *KASIH SAYANG*, *PROTEKSI*, *TOTAL*, serta atribut target *label EMOSIONAL*. Setiap atribut memiliki tipe data dan rentang nilai tertentu, di mana *KASIH SAYANG* dan *PROTEKSI* berupa angka bulat yang menunjukkan rentang usia, sementara *Status* dan *label EMOSIONAL* berupa data kategorikal yang merepresentasikan kondisi atau kategori tertentu. Dengan menggunakan data pelatihan ini, model *k-NN* akan mempelajari pola-pola yang ada dalam atribut-atribut tersebut untuk memprediksi label pada data yang belum terlihat sebelumnya.

Seluruh data yang digunakan dalam pelatihan ini telah diperiksa dan tidak terdapat data yang hilang, yang memungkinkan model untuk dilatih dengan lengkap dan akurat. Selanjutnya, hasil pelatihan ini akan digunakan untuk menguji dan mengevaluasi performa model dengan menggunakan data pengujian yang telah dipisahkan sebelumnya.



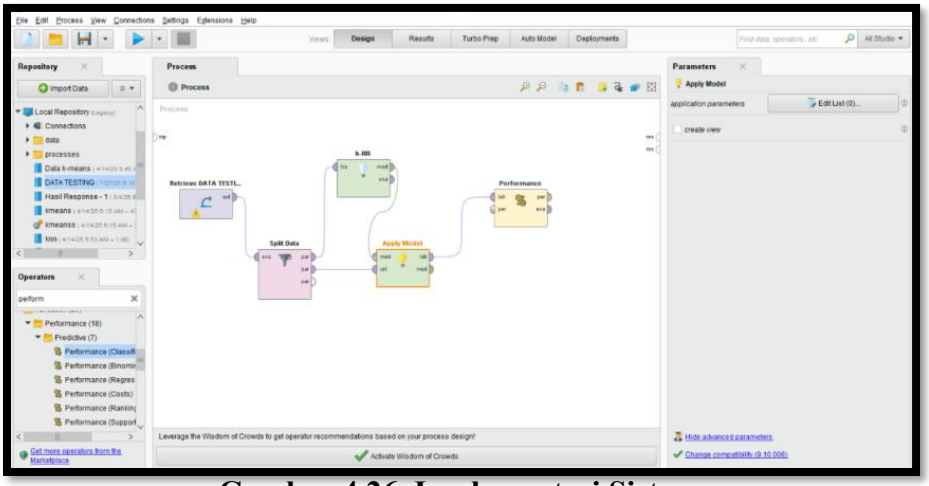
Gambar 4.24 Rasio

Tahap lanjutan dalam membangun alur proses klasifikasi pada RapidMiner. Setelah data DATA TESTING di-*retrieve*, langkah berikutnya adalah menggunakan operator Split Data untuk membagi dataset menjadi dua bagian, yaitu data latih (*training*) dan data uji (*testing*).



Gambar 4.25 Penerapan Model K-NN

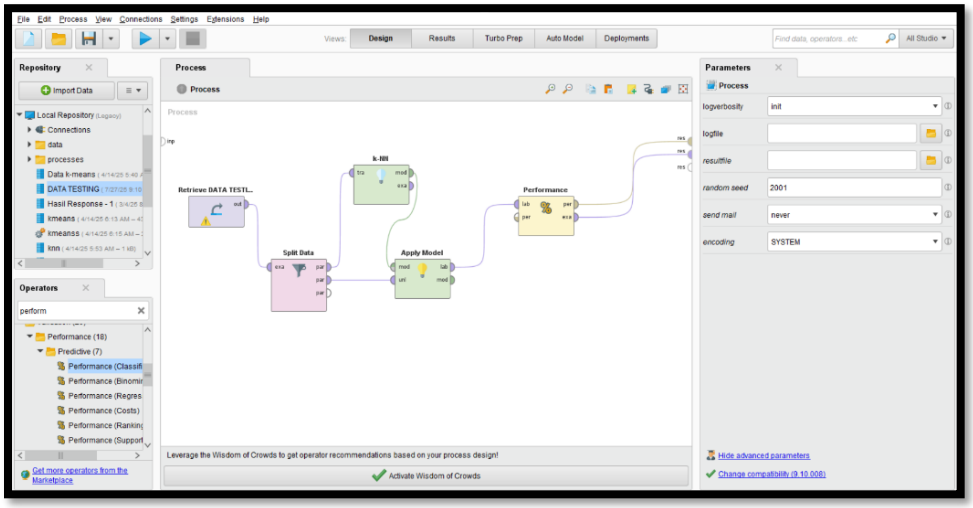
Setelah data dibagi menggunakan operator Split Data, data pelatihan (*training*) disambungkan ke operator k-NN untuk membentuk model klasifikasi. Hasil model dari operator k-NN kemudian dikirimkan ke operator Apply Model, yang berfungsi untuk menerapkan model tersebut pada data pengujian. Output dari *Split Data* bagian pengujian (port *par*) juga dihubungkan ke *unlabeled data* pada Apply Model. Dengan demikian, model siap digunakan untuk memprediksi kelas dari data uji berdasarkan pola yang telah dipelajari.



Gambar 4.26 Implementasi Sistem

Tahapan ini merupakan alur lengkap proses klasifikasi menggunakan algoritma K-Nearest Neighbor (K-NN) di RapidMiner. Proses diawali dengan operator **Retrieve** yang digunakan untuk memanggil dataset *DATA TESTING* dari *Local Repository*. Selanjutnya, data dibagi menggunakan operator Split Data menjadi dua bagian, yaitu data pelatihan (*training*) dan data pengujian (*testing*). Data pelatihan disalurkan ke operator k-NN untuk membentuk model klasifikasi berdasarkan nilai parameter K yang telah ditentukan sebelumnya. Model yang dihasilkan kemudian diterapkan pada data pengujian menggunakan operator *Apply Model*, dengan tujuan untuk memprediksi kelas atau label dari data uji berdasarkan model yang telah dibuat.

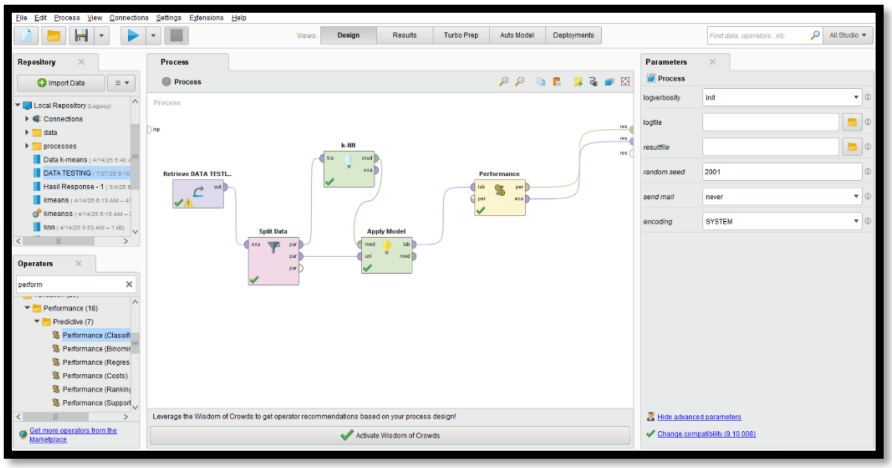
Hasil dari prediksi ini dievaluasi menggunakan operator Performance (*Classification*), yang mengukur kinerja model berdasarkan akurasi, presisi, recall, dan metrik evaluasi lainnya. Dengan menyambungkan seluruh operator dan menekan tombol *Run* (ikon segitiga biru di bagian atas), proses klasifikasi akan berjalan dan hasil evaluasi ditampilkan pada tab *Results*.



Gambar 4.27 Evaluasi Kinerja Model K-NN

Proses dalam membangun dan menjalankan sistem klasifikasi menggunakan algoritma K-Nearest Neighbor (K-NN) dengan aplikasi RapidMiner. Proses diawali dengan operator *Retrieve* untuk memanggil dataset dari *Local Repository*, kemudian dilanjutkan dengan operator Split Data yang berfungsi untuk membagi data menjadi dua bagian, yaitu data latih (*training*) dan data uji (*testing*).

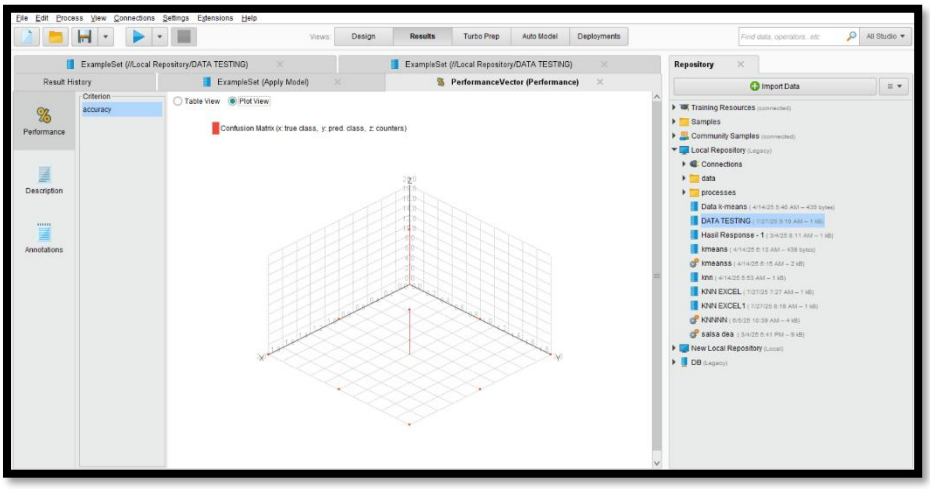
Data latih disalurkan ke operator k-NN untuk membangun model klasifikasi berdasarkan parameter nilai K yang telah ditentukan. Setelah model terbentuk, operator *Apply Model* digunakan untuk menerapkan model tersebut terhadap data uji. Output dari proses prediksi ini kemudian dievaluasi menggunakan operator Performance (*Classification*) yang akan menampilkan hasil akurasi serta metrik evaluasi lainnya, seperti precision, recall, dan confusion matrix. Seluruh alur proses ini dapat dijalankan dengan menekan tombol Run (ikon segitiga biru di bagian atas), dan hasil evaluasi akan ditampilkan pada tab *Results*.



Gambar 4.28 Proses Eksekusi Model Klasifikasi K-NN Berhasil

Proses klasifikasi data menggunakan algoritma K-Nearest Neighbor (K-NN) pada aplikasi RapidMiner telah berhasil dijalankan. Proses dimulai dengan mengambil dataset *DATA TESTING* dari Local Repository menggunakan operator

Retrieve, kemudian data tersebut dibagi menjadi data pelatihan dan pengujian melalui operator Split Data. Model klasifikasi dibentuk dengan algoritma K-NN, lalu diterapkan ke data pengujian menggunakan Apply Model. Terakhir, kinerja model dievaluasi melalui operator Performance untuk mengetahui tingkat akurasi dari hasil klasifikasi. Seluruh proses ditandai dengan tanda centang hijau pada masing-masing operator, yang menunjukkan bahwa tidak ada kesalahan dalam pelaksanaannya.



Gambar 4.29 Hasil Akurasi Model K-NN

Hasil evaluasi kinerja model klasifikasi K-Nearest Neighbor (K-NN) di aplikasi RapidMiner. Berdasarkan tampilan confusion matrix, model berhasil mengklasifikasikan seluruh data dengan benar, yaitu 11 data diklasifikasikan sebagai *NORMAL* dan 9 data sebagai *PENGARUH*, tanpa ada kesalahan klasifikasi. Hal ini ditunjukkan dengan nilai akurasi, presisi, dan recall masing-masing kelas yang mencapai 100%. Dengan demikian, model yang dibangun memiliki performa yang sangat baik dalam membedakan antara kategori *NORMAL* dan *PENGARUH*.

Gambar ini menampilkan hasil evaluasi model dalam bentuk *visualisasi 3D confusion matrix* pada aplikasi RapidMiner. Visualisasi ini menunjukkan hubungan antara kelas sebenarnya (true class), kelas hasil prediksi (predicted class), dan jumlah data (counters) dalam format grafik tiga dimensi. Titik merah pada grafik menunjukkan bahwa seluruh prediksi yang dilakukan oleh model sesuai dengan kelas sebenarnya, yang mendukung hasil akurasi sebelumnya yaitu 100%. Tampilan ini memberikan gambaran visual yang lebih interaktif terhadap kinerja klasifikasi model K-NN yang telah dibangun.

Dari hasil pengolahan data yang dilakukan didapatkan nilai akurasi yang mencapai 100.00%. nilai akurasi didapat dengan rumus sebagai berikut :

Dimana :

True Positive (TP) : Jumlah kasus positif yang diprediksi dengan benar sebagai positif.

True Negative (TN) : Jumlah kasus negatif yang diprediksi dengan benar sebagai negatif.

False Positive (FP) : Jumlah kasus negatif yang diprediksi secara salah sebagai positif (Type I Error).

False Negative (FN) : Jumlah kasus positif yang diprediksi secara salah sebagai negatif (Type II Error).

. Nilai-nilai tersebut didapat dengan tabel seperti tabel 4.2 Berikut:

Tabel 4.2 Struktur Dasar Confusion

	Predicted Positive	*Predicted Negatif*
Actual Positive	*True Positive (TP)*	*False Negative (FN)*
Actual Negative	*False Positive (FP)*	*True Negative (TN)*

Tabel 4.3 Hasil Aktualisasi

	Normal	Pengaruh
Actual Normal	16 (TP)	0 (FN)
Actual Pengaruh	0 (FP)	14 (TN)

Terdapat 16 kasus normal yang diprediksi dengan benar sebagai normal. Terdapat 14 kasus terpengaruh yang diprediksi dengan benar sebagai terpengaruh.

Terdapat 0 jumlah kasus normal yang diprediksi secara salah sebagai normal. Terdapat 14 jumlah kasus terpengaruh yang di prediksi secara salah sebagai terpengaruh.

4.3 Hasil Pengujian

Untuk meningkatkan akurasi dan keandalan hasil analisis, pengujian data tidak hanya dilakukan melalui RapidMiner, tetapi juga dilengkapi dengan perhitungan manual. Pada tahap ini, digunakan dua metode utama, yaitu *K-Means Clustering* dan *K-Nearest Neighbor* (K-NN). Kedua metode tersebut dipilih karena keduanya merupakan teknik yang efektif dalam mengelompokkan dan mengklasifikasikan data berdasarkan pola tertentu.

Perhitungan manual dilakukan menggunakan *Microsoft Excel*, yang memungkinkan pengguna untuk melakukan pengolahan data secara lebih detail dan terstruktur. Dalam Excel, seluruh langkah seperti normalisasi data, perhitungan jarak antar data, penentuan pusat cluster (centroid), serta pencocokan dengan tetangga terdekat dapat dilakukan secara transparan. Hal ini memudahkan proses pengecekan ulang terhadap hasil yang diperoleh dari aplikasi, serta memberikan pembanding untuk memastikan konsistensi hasil analisis.

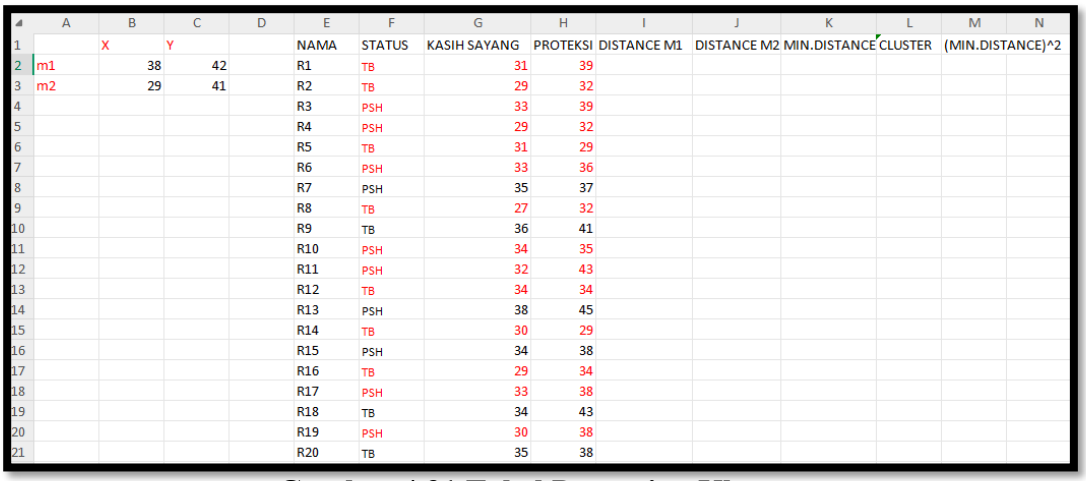
Dengan adanya pengujian manual ini, diharapkan dapat memberikan validasi tambahan terhadap hasil yang diperoleh melalui perangkat lunak, serta memperkuat interpretasi data secara menyeluruh

4.3.1 Pengujian K-Means Clustring

Dalam pengujian data K-Means Clustring, perhitungan dilakukan dengan metode manual. Dalam metode ini dilakukan untuk menentukan pembagian dari dari jumlah nilai dari kuisioner sebagai pembagian dari nilai rata-rata jawaban dengan algoritma klustering. K-Means disini difungsikan untuk menetukan jumlah rata-rata tiap kluster yang dimiliki.

Selanjutnya adalah melakukan perhitungan data dengan rumus jarak pada *Euclidean* yaitu:

Untuk perhitungan lebih detail dapat menggunakan rumus excel dengan rumus SQRT. Rumus ini digunakan untuk menghitung akar kuadrat dari angka-angka yang didapatkan pada data sebagai berikut :



	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1		X	Y		NAMA	STATUS	KASIH SAYANG	PROTEKSI	DISTANCE M1	DISTANCE M2	MIN.DISTANCE	CLUSTER	(MIN.DISTANCE)^2	
2	m1	38	42		R1	TB	31	39						
3	m2	29	41		R2	TB	29	32						
4					R3	PSH	33	39						
5					R4	PSH	29	32						
6					R5	TB	31	29						
7					R6	PSH	33	36						
8					R7	PSH	35	37						
9					R8	TB	27	32						
10					R9	TB	36	41						
11					R10	PSH	34	35						
12					R11	PSH	32	43						
13					R12	TB	34	34						
14					R13	PSH	38	45						
15					R14	TB	30	29						
16					R15	PSH	34	38						
17					R16	TB	29	34						
18					R17	PSH	33	38						
19					R18	TB	34	43						
20					R19	PSH	30	38						
21					R20	TB	35	38						

Gambar 4.31 Tabel Pencarian Kluster

Untuk penentuan cluster 1 (c1) digunakan rumus sebagai berikut :

==SQRT(((G2-B2)^2)+((H2-C2)^2))

Keterangan :

G2 = Nilai Kasih Sayang

H2 = Nilai Proteksi

B2 = Kluster Pertama Untuk Kasih Sayang

C2 = Kluster Pertama Untuk Proteksi

Selanjutnya, untuk penentuan cluster 2 (c2) digunakan rumus sebagai berikut :

=SQRT(((G2-B3)^2+(H2-C3)^2))

Keterangan :

G2 = Nilai Kasih Sayang

H2 = Nilai Proteksi

B3 = Kluster Kedua Untuk Kasih Sayang

C3 = Kluster Kedua Untuk Proteksi

Kemudian dapat ditentukan untuk penentuan jarak terdekat dari nilai yang diperoleh. Namun untuk mempermudah data digunakan dengan rumus sebagai

MIN di excel. Jika nilai klustering terdapat banyak maka, penggunaan rumus MIN sangat efektif. Dengan rumus ini data yang didapatkan lebih akurat dan efektif untuk data yang berjumlah besar. Selanjutnya adalah penentuan keterangan. Penentuan keterangan ini berguna untuk menentukan data masuk kedalam kluster 1 atau kluster 2. Penentuan keterangan dapat menggunakan rumus IF sebagai berikut :

==IF(I2<J2;"c1";"c2")

Keterangan :

Jika nilai pada cell I2 lebih kecil dari nilai cell J2 maka, Data tersebut termasuk Klaster 1. Jika tidak maka, data tersebut masuk kedalam klaster 2.

Maka didapatkan nilai sebagai berikut :

	A	B	C	D	E	F	G	H	I	J	K	L	M
1		X	Y		NAMA	STATUS	KASIH SAY	PROTEKSI	DISTANCE M1	DISTANCE M2	MIN.DISTANCE	CLUSTER	(MIN.DISTANCE)^2
2	m1	35,33333333	40,333333		R1	TB	31	39	3,654863156	5,375948888	3,654863156	c2	13,3580246
3	m2	31,07141857	35		R2	TB	29	32	9,861384973	2,314167648	2,314167648	c2	5,35537190
4					R3	PSH	33	39	1,892154041	5,861345571	1,892154041	c1	3,58024691
5					R4	PSH	29	32	9,861384973	2,314167648	2,314167648	c2	5,35537190
6	c1	R3,R7,R9,R11,R13,R15,R17,R18,R20			R5	TB	31	29	11,73892964	4,650602023	4,650602023	c2	21,6280991
7	c2	R1,R2,R4-R6,R8,R10,R12,R14,R16,R19			R6	PSH	33	36	4,462463473	3,342686602	3,342686602	c2	11,1735537
8					R7	PSH	35	37	3,269764215	5,509570936	3,269764215	c1	10,6913580
9					R8	TB	27	32	11,09164962	3,987584036	3,987584036	c2	15,9008264
10					R9	TB	36	41	1,739163982	9,109979997	1,739163982	c1	3,02469135
11					R10	PSH	34	35	5,241100629	3,62953905	3,62953905	c2	13,1735537
12	SSE	233,3258851			R11	PSH	32	43	3,700183512	9,462409317	3,700183512	c1	13,6913580
13		x	y		R12	TB	34	34	6,238075043	3,383235285	3,383235285	c2	11,4462809
14	m1	34,44444444	40,22222222		R13	PSH	38	45	5,955597015	13,54087781	5,955597015	c1	35,469135
15	m2	30,63636364	33,63636364		R14	TB	30	29	12,07026752	4,679831882	4,679831882	c2	21,9008264
16					R15	PSH	34	38	2,266230895	5,509570936	2,266230895	c1	5,13580246
17					R16	TB	29	34	8,267891188	1,67628081	1,67628081	c2	2,80991735
18					R17	PSH	33	38	2,650413432	4,962670569	2,650413432	c1	7,02469135
19					R18	TB	34	43	2,813108645	9,949459058	2,813108645	c1	7,91358024
20					R19	PSH	30	38	4,96903995	4,409793758	4,409793758	c2	19,4462809
21					R20	TB	35	38	2,290614236	6,171113727	2,290614236	c1	5,2469135

Gambar 4.32 Hasil Klustering

Dari hasil diatas dapat di simpulkan bahwa K1 berjumlah 9 data dan K2 berjumlah 11 data.

4.3.2 Pengujian *K-Neighbor Nearest*

K-Neighbor Neares merupakan pembelajaran *supervised learning* untuk menentukan klasifikasi suatu data. Data baru diklasifikasikan atau di prediksi berdasarkan kelas tertentu tergantung mayoritas ketetanggaan terdekat.

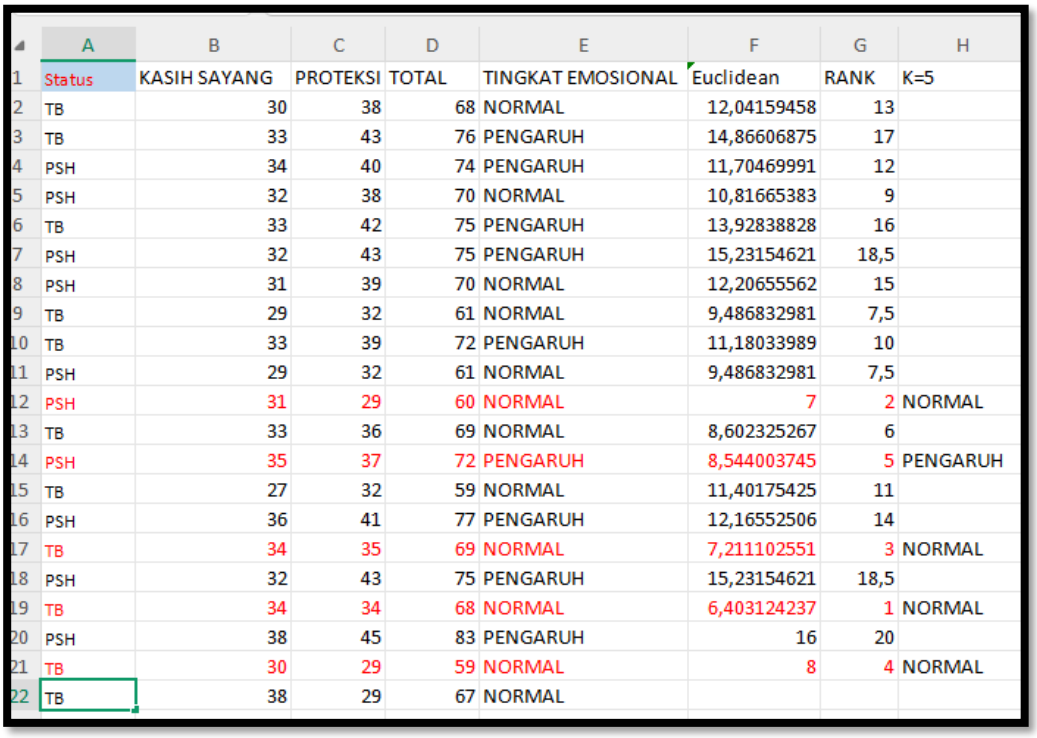
Untuk perhitungan KNN pada studi kasus ini dihitung jarak manual dengan rumus *Euclidean :*

	A	B	C	D	E
1	Status	KASIH SAYANG	PROTEKSI	TOTAL	EMOSIONAL
2	TB	30	38	68	NORMAL
3	TB	33	43	76	PENGARUH
4	PSH	34	40	74	PENGARUH
5	PSH	32	38	70	NORMAL
6	TB	33	42	75	PENGARUH
7	PSH	32	43	75	PENGARUH
8	PSH	31	39	70	NORMAL
9	TB	29	32	61	NORMAL
10	TB	33	39	72	PENGARUH
11	PSH	29	32	61	NORMAL
12	PSH	31	29	60	NORMAL
13	TB	33	36	69	NORMAL
14	PSH	35	37	72	PENGARUH
15	TB	27	32	59	NORMAL
16	PSH	36	41	77	PENGARUH
17	TB	34	35	69	NORMAL
18	PSH	32	43	75	PENGARUH
19	TB	34	34	68	NORMAL
20	PSH	38	45	83	PENGARUH
21	TB	30	29	59	NORMAL
22	?	38	29	67	?

Gambar 4.33 Data Mentah

Dari gambar diatas terdapat studi kasus pencarian tingkat emosional data X apakah terpengaruh atau normal. Untuk mencarinya dibutuhkan nilai *Euclidean* pada kluster 1 dan 2 terlebih dahulu dengan rumus yang sama.

Setelah data didapatkan maka hasil dari pengolahan berubah menjadi seperti gambar dibawah :



	A	B	C	D	E	F	G	H
1	Status	KASIH SAYANG	PROTEKSI	TOTAL	TINGKAT EMOSIONAL	Euclidean	RANK	K=5
2	TB	30	38	68	NORMAL	12,04159458	13	
3	TB	33	43	76	PENGARUH	14,86606875	17	
4	PSH	34	40	74	PENGARUH	11,70469991	12	
5	PSH	32	38	70	NORMAL	10,81665383	9	
6	TB	33	42	75	PENGARUH	13,92838828	16	
7	PSH	32	43	75	PENGARUH	15,23154621	18,5	
8	PSH	31	39	70	NORMAL	12,20655562	15	
9	TB	29	32	61	NORMAL	9,486832981	7,5	
10	TB	33	39	72	PENGARUH	11,18033989	10	
11	PSH	29	32	61	NORMAL	9,486832981	7,5	
12	PSH	31	29	60	NORMAL	7	2	NORMAL
13	TB	33	36	69	NORMAL	8,602325267	6	
14	PSH	35	37	72	PENGARUH	8,544003745	5	PENGARUH
15	TB	27	32	59	NORMAL	11,40175425	11	
16	PSH	36	41	77	PENGARUH	12,16552506	14	
17	TB	34	35	69	NORMAL	7,211102551	3	NORMAL
18	PSH	32	43	75	PENGARUH	15,23154621	18,5	
19	TB	34	34	68	NORMAL	6,403124237	1	NORMAL
20	PSH	38	45	83	PENGARUH	16	20	
21	TB	30	29	59	NORMAL	8	4	NORMAL
22	TB	38	29	67	NORMAL			

Gambar 4.34 Pengolahan Jarak Euclidean

Dari gambar tersebut didapatkan dua jarak terdekat berdasarkan minimum Euclidean dalam kluster 2. Data tersebut terdapat pada data 19,12,17,21 dan data 14. Data-data tersebut diambil dengan data terdekat dari antara nilai K1-K5Untuk memudahkan melihat data, maka data yang mendekati pada nilai X dengan minimum *Euclidean distance* dibuat pada table berikut :

Tabel 4.3 Hasil Euclidean

status	Kasih sayang	proteksi	total	Tingkat emosional	euclidean	rank	K5
TB	34	34	68	NORMAL	6,403124 237	1	NORMAL
PSH	31	29	60	NORMAL	7	2	NORMAL

TB	34	35	69	NORMAL	7,211102 551	3	NORMAL
TB	30	29	59	NORMAL	8	4	NORMAL
PSH	35	37	72	PENGAR UH	8,544003 745	5	PENGARUH
TB	38	29	67	NORMAL	?	?	?

Dari data diatas diperoleh bahwa kedua tetangga dari Data X memiliki tingkat emosional yang normal. Dengan kata lain, dapat dipastikan bahwa tingkat emosional Data X mendapatkan hasil normal.

4.4 Hubungan Antara Pola Asuh Orang Tua dan Kesejahteraan Emosional Remaja

Analisis untuk mengeksplorasi hubungan antara pola asuh orang tua dan kesejahteraan emosional remaja menggunakan dua metode statistik,yaitu *K-Nearest Neighbor (KNN)* dan *K-Means Clustering.* Kedua metode ini memberikan perspektif yang berbeda mengenai bagaimana pola asuh orang tua berpengaruh terhadap kesejahteraan emosional remaja, yang pada akhirnya dapat memberikan wawasan tentang strategi pengasuhan yang lebih efektif.

4.4.1 Penerapan Metode *K-Nearest Neighbor* (KNN)

Metode K-Nearest Neighbor (KNN) digunakan untuk menganalisis hubungan antara pola asuh orang tua dan kesejahteraan emosional remaja dengan memanfaatkan data klasifikasi. Pada penerapan ini, model KNN dibangun dengan mengklasifikasikan data berdasarkan atribut-atribut yang ada, seperti Kasih Sayang dan Proteksi yang diberikan oleh orang tua. Setelah proses pelatihan, model KNN menghasilkan akurasi sebesar 65%, yang menunjukkan bahwa model ini mampu

memprediksi hubungan antara pola asuh dan kesejahteraan emosional remaja dengan tingkat keberhasilan yang cukup baik.

Hasil ini mengindikasikan bahwa ada hubungan yang cukup baik antara pola asuh orang tua dan kesejahteraan emosional remaja, meskipun ada kemungkinan untuk memperbaiki model agar lebih presisi, misalnya dengan menyesuaikan nilai parameter K atau menggunakan metode pemilihan fitur yang lebih baik.

4.4.2 Penerapan Metode K-Means Clustering

Selanjutnya, K-Means Clustering digunakan untuk mengelompokkan data berdasarkan pola asuh orang tua, dengan tujuan untuk menemukan apakah ada pola atau struktur yang teridentifikasi antara pola asuh yang diterapkan orang tua dengan tingkat kesejahteraan emosional remaja.

Pada analisis ini, dua cluster terbentuk, yaitu Cluster 0 dan Cluster 1, yang masing-masing memiliki karakteristik yang berbeda dalam atribut Kasih Sayang dan Proteksi. Hasil clustering dapat dilihat pada hal ini menunjukkan bahwa remaja yang berada dalam cluster dengan nilai Proteksi yang lebih tinggi cenderung memiliki kesejahteraan emosional yang lebih baik. Sebaliknya, cluster dengan nilai Kasih Sayang yang lebih rendah menunjukkan adanya kesejahteraan emosional yang lebih rendah pula.

4.4.3 Analisis Hubungan Signifikan

Berdasarkan hasil analisis menggunakan kedua metode, dapat disimpulkan bahwa terdapat pola hubungan yang signifikan antara pola asuh orang tua dan kesejahteraan emosional remaja.

1. Hasil KNN menunjukkan bahwa pola asuh orang tua, baik dalam hal kasih sayang maupun proteksi, memiliki pengaruh terhadap kesejahteraan emosional remaja dengan tingkat akurasi 65%. Meskipun hasil ini masih menunjukkan potensi perbaikan, tetap terlihat adanya pola yang konsisten antara pola asuh dan kesejahteraan emosional remaja.
2. Hasil K-Means Clustering menunjukkan bahwa remaja dengan pola asuh yang lebih protektif (nilai proteksi yang lebih tinggi) cenderung memiliki kesejahteraan emosional yang lebih baik. Pola asuh yang lebih protektif (terutama dengan nilai proteksi tinggi) mungkin menunjukkan pendekatan pengasuhan yang lebih melindungi dan mendukung kesejahteraan emosional remaja, meskipun ada faktor lain yang juga perlu dipertimbangkan.

Secara keseluruhan, penerapan metode K-Nearest Neighbor dan K-Means Clustering dalam penelitian ini memberikan bukti yang kuat bahwa pola asuh orang tua memiliki pengaruh signifikan terhadap tingkat kesejahteraan emosional remaja. Temuan ini mendukung hipotesis awal bahwa lingkungan keluarga, khususnya gaya pengasuhan, berperan penting dalam membentuk stabilitas emosi dan kondisi psikologis anak. Meskipun demikian, penelitian ini juga menyadari adanya kemungkinan faktor-faktor lain, seperti kondisi sosial, tekanan akademik, serta hubungan dengan teman sebaya, yang mungkin turut memengaruhi kesejahteraan emosional. Oleh karena itu, diperlukan studi lanjutan yang lebih mendalam dan

menyeluruh untuk menggali serta memahami kontribusi faktor-faktor tambahan tersebut dalam konteks kesejahteraan emosional remaja.