

BAB 4

HASIL DAN PEMBAHASAN

Pada bab ini penulisi akan membahas mengenai implementasi data yang digunakan. Data tersebut akan diimplementasikan menggunakan rapidminer menggunakan K-means dan Kmedoids.

4.1 Persiapan Data

Data yang digunakan dalam penelitian ini adalah dataset dari sekolah SMK N1 Katibung dengan jumlah sebanyak 601 data yang menggunakan tools rapid miner. Metode yang digunakan dalam penelitian ini K-means dan Kmedoids. Data pendaftar beasiswa terdapat pada gambar 4.1 dibawah ini:

	A	B	C	D	E	F	G	H
1	NAMA	DIKIPAN KEPALA RUMAH TANGGA	PEKERJAAN ORANG TUA	PENGHASILAN ORANG TUA	JUMLAH TANGGUNGAN ORANG TUA	JUMLAH/STATUS ORANG TUA	KEADAAN FISIK RUMAH	DAYA LISTRIK
2	Anisa Luzai M. S	1	3	2	3	1	10	2
3	ARDA IKBAL SANDIKA	1	2	1	2	1	8	1
4	ARIF SETIAWAN YUSUP	1	1	1	3	1	8	2
5	Arif Kurniawan putra	1	1	1	2	1	8	1
6	Bagas abrii prakoso	1	1	1	2	1	8	1
7	BAYU ANGGARA	1	2	1	2	1	8	2
8	Eka putri ardianis	1	2	1	2	1	8	2
9	Emilia Aprilianti	1	1	1	2	1	8	2
10	FAHMI RAHMAN	1	1	1	2	1	8	2
11	Indra Julianto	1	1	1	2	1	8	2
12	MALIK FAJAR TRI KURNIA	1	2	1	2	1	8	2
13	Mario Yuliyadi	1	2	1	1	1	8	2
14	Mufid Budi Sofyan	1	2	1	1	1	8	2
15	Muhammad Akmal	1	1	1	1	1	8	2
16	Muhamad Ridwan	1	2	1	1	1	8	2
17	NIA ARYUNIKA	1	1	1	1	1	8	2
18	PUPUT AYU RISWANI	1	1	1	2	1	8	2
19	PUTRI MELIANA WULANDA	1	3	2	3	1	10	2
20	RENDI SAPUTRA	1	2	1	1	1	8	1
21	Reni	1	1	1	1	1	8	1
22	RICKY FEBRIANES FERNAND	1	1	1	1	1	8	2
23	RIZAL MUHAMMAD AL GHIF	1	2	1	1	1	8	1
24	Sarwani	1	2	1	1	1	9	1

Gambar 4. 1 Dataset Pendaftar Beasiswa

Gambar 4.1 menunjukkan karakter data pada penelitian ini, Data yang digunakan dalam penelitian ini adalah dataset dari sekolah SMK N 1 Katibung dengan jumlah sebanyak 601 data . Atribut yang digunakan seperti Pendidikan Kepala Rumah Tangga, Pekerjaan Orang Tua, Penghasilan Orang Tua, Jumlah Tanggungan Orang Tua, Jumlah/Status Orang Tua, Keadaan Fisik Rumah, Daya Listrik, Status Kepemilikan Tempat Tinggal, Perabot Rumah Dan Alat Komunikasi. Atribut-atribut ini memiliki hubungan atau pengaruh terhadap masalah yang sedang diteliti. Dalam penelitian ini,

atribut-atribut ini kemudian digunakan sebagai dasar untuk analisis menggunakan metode K-means dan K-medoids. Metode tersebut digunakan untuk mengelompokkan data berdasarkan pola-pola yang terdapat dalam atribut-atribut ini, sehingga dapat membantu dalam pemahaman dan penyelesaian masalah yang sedang diteliti.

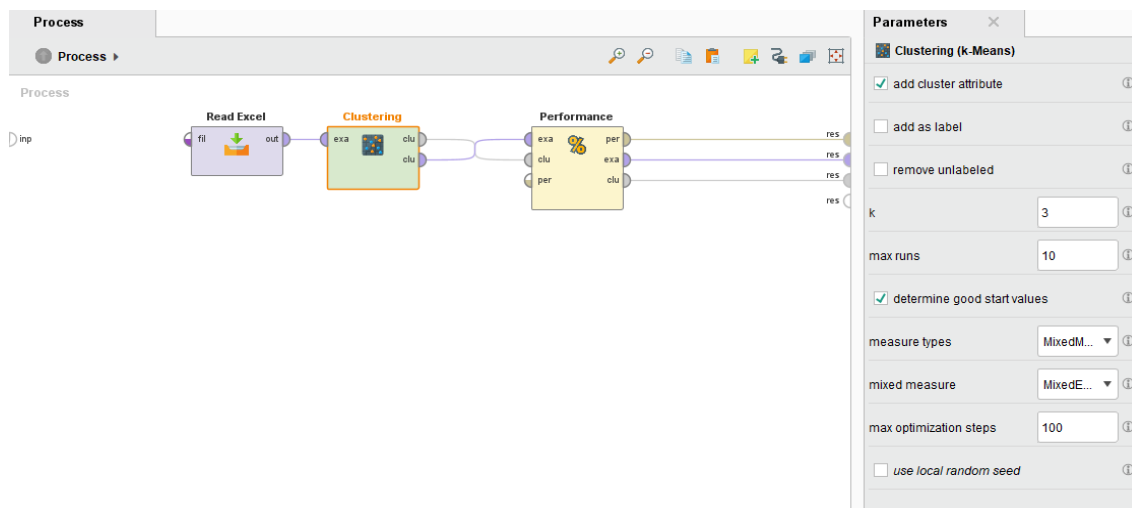
4.2 Pemodelan

Pemilihan dan penerapan teknik pemodelan yang sesuai sangat penting dalam tahapan analisis data. Dalam konteks penelitian ini, teknik pemodelan yang dipilih adalah teknik clusterisasi. Clusterisasi adalah metode analisis data yang bertujuan untuk mengelompokkan objek data ke dalam kelompok-kelompok yang serupa berdasarkan kesamaan fitur atau karakteristik tertentu. Pemilihan teknik clusterisasi dalam penelitian ini sangatlah relevan karena penelitian ini bertujuan untuk mengelompokkan data siswa sekolah menengah ke dalam kelompok-kelompok berdasarkan atribut-atribut yang terkait dengan kondisi sosial-ekonomi dan lingkungan rumah tangga mereka. Dengan menggunakan teknik clusterisasi, peneliti dapat mengidentifikasi pola-pola yang mungkin tersembunyi di dalam data, serta memahami lebih baik struktur dan karakteristik dari populasi siswa yang diteliti.

Teknik clusterisasi memberikan fleksibilitas dalam mengelompokkan data tanpa memerlukan informasi target yang telah ditetapkan sebelumnya. Hal ini cocok untuk penelitian di mana peneliti ingin mendapatkan pemahaman yang lebih holistik tentang data tanpa memiliki asumsi tertentu tentang klasifikasi yang mungkin ada. Selain itu, teknik clusterisasi juga memiliki keunggulan dalam menangani data yang besar dan kompleks. Dengan memanfaatkan algoritma clustering yang efisien, peneliti dapat mengelompokkan data dengan cepat dan efektif, sehingga memungkinkan untuk analisis yang lebih mendalam dan temuan yang lebih berharga. Oleh karena itu, pemilihan teknik clusterisasi sebagai metode pemodelan dalam penelitian ini merupakan langkah yang tepat dan relevan dengan tujuan penelitian. Dengan menggunakan teknik ini, diharapkan peneliti dapat menghasilkan pemahaman yang lebih mendalam tentang karakteristik siswa SMK dan mengidentifikasi potensi pola atau tren yang signifikan dalam data.

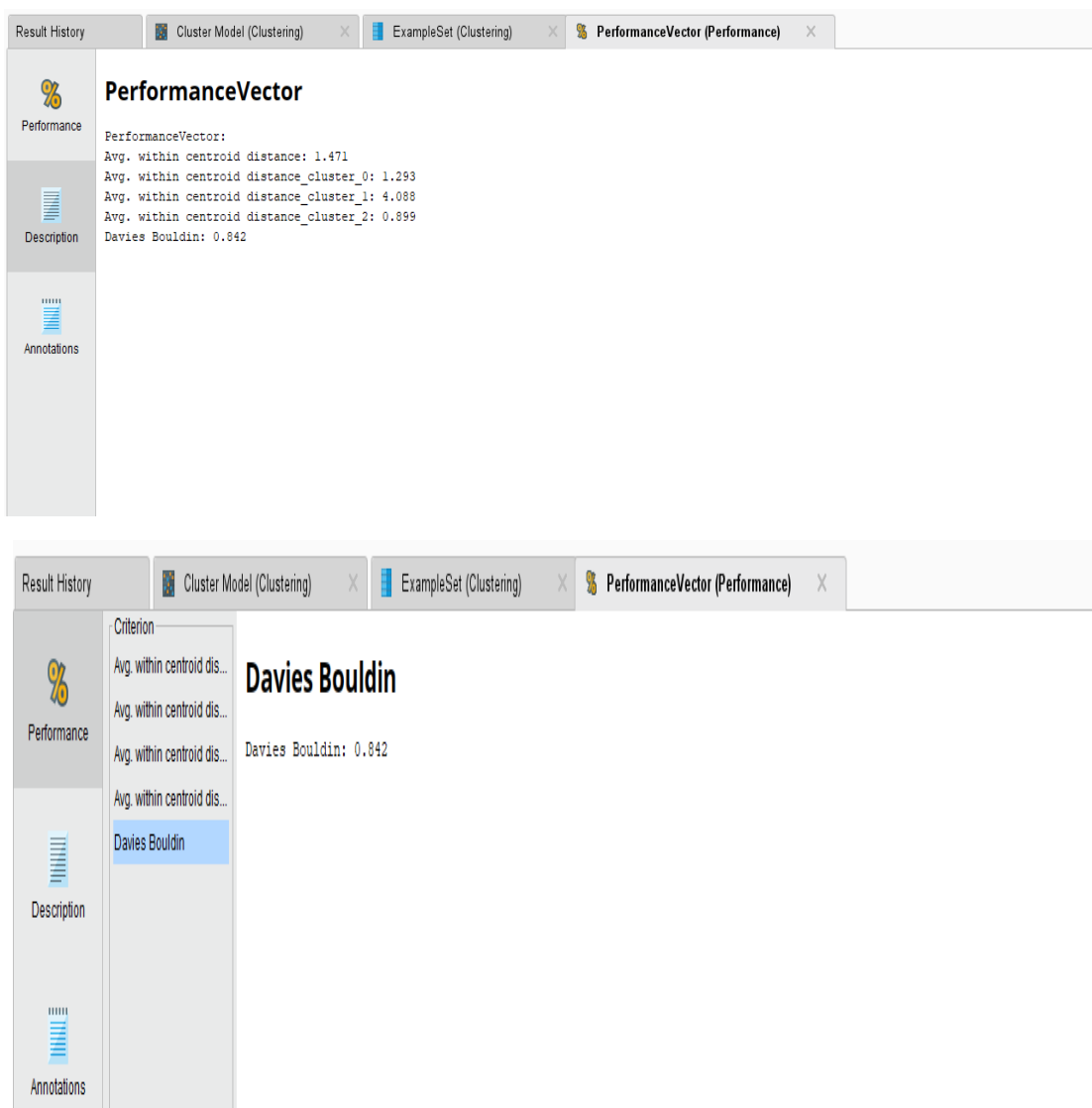
4.3 Penelitian Menggunakan Algoritma K-Means

Penerapan data pada Rapidminer untuk penentuan beasiswa menggunakan algoritma K-means ditunjukkan pada gambar dibawah ini:



Gambar 4. 2 Penerapan Data Pendaftar Beasiswa Menggunakan Algoritma K-Means pada Rapidminer

Gambar 4.2 menunjukkan proses penerapan data yang telah disiapkan ke dalam aplikasi RapidMiner. Dalam konteks ini, data yang telah disiapkan digunakan untuk melakukan eksperimen menggunakan teknik clusterisasi K-means. Proses eksperimen ini bertujuan untuk mengelompokkan data ke dalam beberapa kluster berdasarkan kemiripan atau pola yang terdapat dalam atribut-atribut yang telah ditentukan sebelumnya. Setelah data diterapkan pada aplikasi RapidMiner dan eksperimen menggunakan algoritma K-means dilakukan, hasil dari eksperimen tersebut ditampilkan pada Gambar 4.3. Gambar ini menyajikan hasil klasterisasi yang dihasilkan oleh algoritma K-means, di mana data telah dikelompokkan ke dalam beberapa kluster berdasarkan karakteristik yang dimiliki.



Gambar 4. 3 Hasil Eksperimen Menggunakan Algoritma

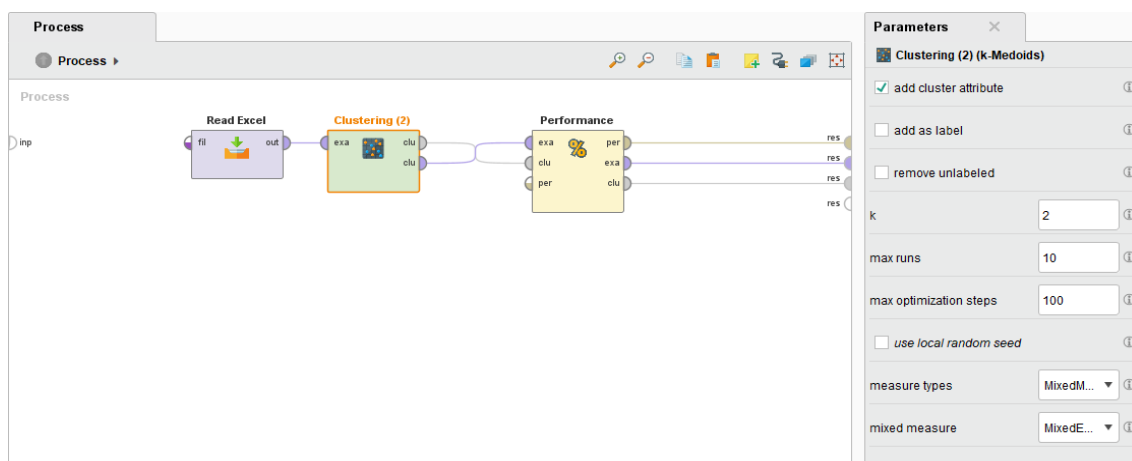
Gambar 4.3 menunjukkan hasil dari proses clustering, yang mungkin dilakukan menggunakan algoritma tertentu seperti K-means atau Hierarchical Clustering, di mana data telah dikelompokkan ke dalam beberapa cluster berdasarkan kesamaan fitur atau karakteristik tertentu. Hasil dari clusterisasi yang menunjukkan eksperimen didalam Davies Bouldin kita dapat melihat hasil Davies Bouldin yaitu sebesar 0.842.

Davies Bouldin Index (DBI) adalah sebuah metrik evaluasi yang digunakan untuk mengukur kualitas dari sebuah clustering. Semakin rendah nilai DBI, semakin baik hasil

clusteringnya. Nilai DBI dapat diinterpretasikan sebagai rata-rata dari kedekatan antara setiap cluster, diukur dalam konteks varian dalam setiap cluster dan jarak antar cluster. Nilai yang lebih rendah menunjukkan bahwa antar-cluster lebih berbeda dan intra-cluster lebih kompak. Dalam kasus ini, hasil Davies Bouldin sebesar 0.842 menunjukkan bahwa clustering memiliki tingkat kualitas yang baik.

4.4 Penelitian Menggunakan K-Medoids Pada Rapidminer

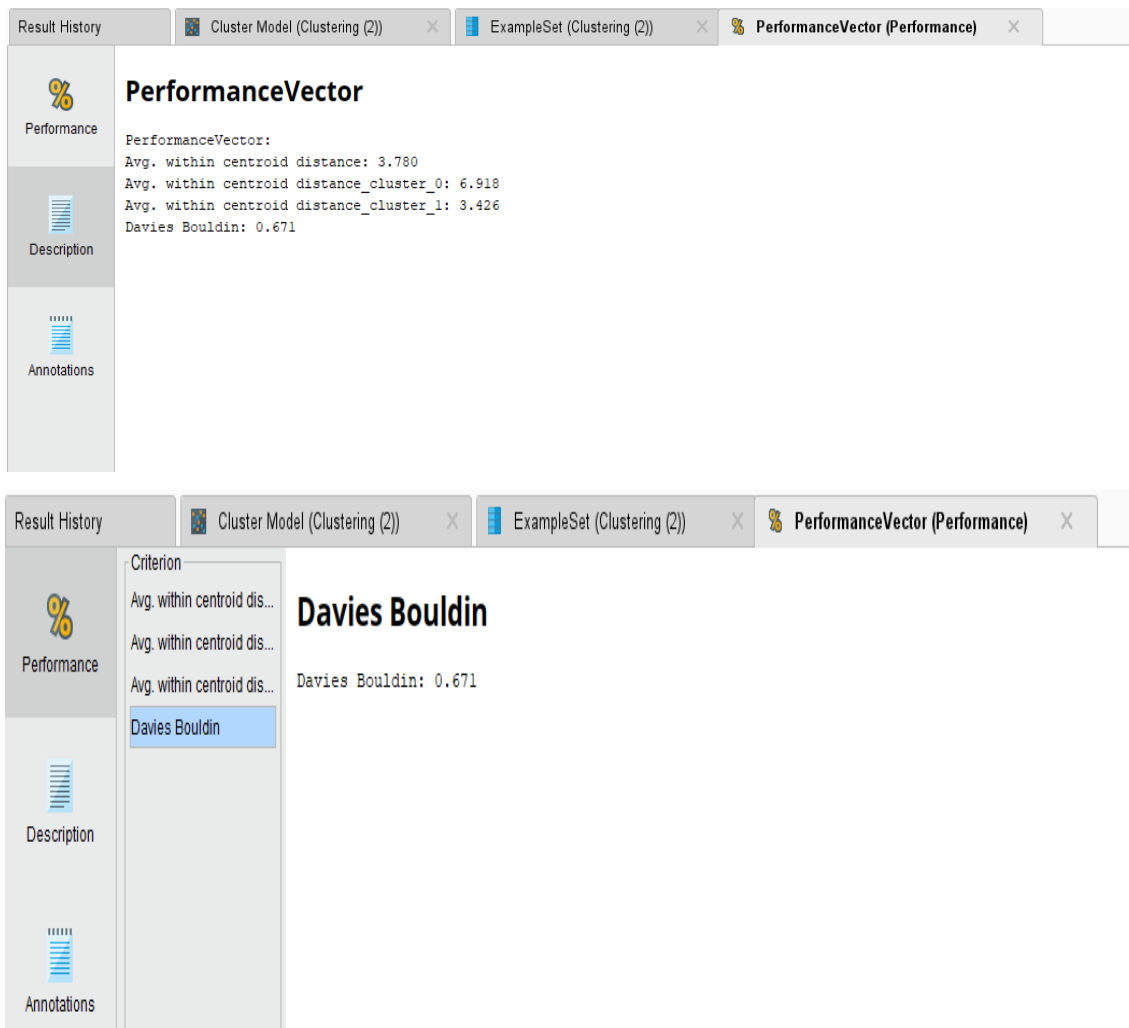
Penerapan data pada Rapidminer untuk penentuan beasiswa menggunakan algoritma K-medoids ditunjukkan pada gambar dibawah ini:



Gambar 4. 4 Penerapan Data Pendaftar Beasiswa Menggunakan Algoritma K-Medoids pada Rapidminer

Gambar 4.4 menggambarkan proses penerapan data yang telah dipersiapkan ke dalam aplikasi RapidMiner. Dalam tahap ini, data yang telah dipersiapkan tersebut digunakan untuk melakukan eksperimen menggunakan teknik clusterisasi K-Medoids. Tujuan dari eksperimen ini adalah untuk mengelompokkan data ke dalam beberapa klaster berdasarkan kemiripan atau pola yang terdapat dalam atribut-atribut yang telah ditentukan sebelumnya. Setelah data diterapkan pada aplikasi RapidMiner dan eksperimen menggunakan algoritma K-Medoids dilakukan, hasil dari eksperimen

tersebut ditampilkan pada Gambar 4.5. Gambar ini menunjukkan hasil klasterisasi yang dihasilkan oleh algoritma K-Medoids, di mana data telah dikelompokkan ke dalam beberapa klaster berdasarkan karakteristik yang dimiliki.



Gambar 4. 5 Hasil Eksperimen Menggunakan Algoritma

Gambar 4.5 menampilkan hasil dari proses clustering yang dilakukan menggunakan algoritma K-Medoids. K-Medoids adalah varian dari algoritma K-Means yang menggunakan medoids (representative data points) sebagai pusat cluster, yang membuatnya lebih tahan terhadap outliers. Dalam proses clustering ini, data telah dikelompokkan ke dalam beberapa cluster berdasarkan kesamaan fitur atau karakteristik

tertentu. Hasil dari clusterisasi yang menunjukkan eksperimen didalam Davies Bouldin kita dapat melihat hasil Davies Bouldin yaitu sebesar 0.671.

Davies Bouldin Index (DBI) digunakan sebagai metrik evaluasi untuk mengukur kualitas dari clustering yang dihasilkan. Semakin rendah nilai DBI, semakin baik hasil clusteringnya. Nilai DBI adalah rata-rata dari kedekatan antara setiap cluster, diukur dalam konteks varian dalam setiap cluster dan jarak antar cluster. Nilai yang lebih rendah menunjukkan bahwa antar-cluster lebih berbeda dan intra-cluster lebih kompak.

Perbandingan hasil eksperimen yang tidak menggunakan metode K-means dan eksperimen yang menggunakan K-Medoids ditunjukkan pada tabel dibawah ini:

Tabel 4. 1 Perbandingan Hasil Akurasi Penggunaan K-Means dan K-Medoids

Penelitian	Davies Bouldin Index
K-Means	0.842
K-Medoids	0.671

Dalam konteks ini, perbandingan antara hasil Davies Bouldin Index (DBI) dari clustering menggunakan algoritma K-Medoids dengan eksperimen sebelumnya menggunakan K-Means memungkinkan untuk mengevaluasi dan memahami kualitas dari kedua metode clustering tersebut. Davies Bouldin Index (DBI) adalah salah satu metrik evaluasi yang umum digunakan untuk mengukur kualitas dari clustering yang dihasilkan. Nilai DBI mencerminkan seberapa baik kluster yang dihasilkan dalam memisahkan data, dengan nilai yang lebih rendah menandakan kualitas clustering yang lebih baik. Hasil Davies Bouldin sebesar 0.671 untuk clustering dengan menggunakan algoritma K-Medoids menunjukkan bahwa partisi data yang dihasilkan memiliki tingkat kualitas yang lebih baik dibandingkan dengan eksperimen sebelumnya yang menggunakan K-Means, di mana nilai DBI sebesar 0.842. Ini menandakan bahwa algoritma K-Medoids menghasilkan kluster yang lebih baik dalam memisahkan data menjadi kelompok-kelompok yang bermakna.

4.5 Evaluasi

Pemodelan yang telah dilakukan dengan melakukan eksperimen menggunakan algoritma k-means kurang menghasilkan yang baik. Dalam konteks ini, hasil Davies Bouldin K-medoids sebesar 0.671 menunjukkan bahwa clustering memiliki tingkat kualitas yang lebih baik daripada eksperimen sebelumnya menggunakan K-means (dengan nilai DBI sebesar 0.842). Hasil Davies Bouldin Index (DBI) sebesar 0.671 untuk clustering menggunakan algoritma K-Medoids menunjukkan peningkatan kualitas yang signifikan dibandingkan dengan eksperimen sebelumnya yang menggunakan K-Means, yang memiliki nilai DBI sebesar 0.842. Peningkatan dalam kualitas clustering hasil penelitian menunjukkan peningkatan sekitar 20.33% dalam kualitas clustering ketika menggunakan algoritma K-Medoids dibandingkan dengan K-Means. Ini merupakan peningkatan yang signifikan dan menunjukkan bahwa penggunaan K-Medoids dapat menghasilkan hasil clustering yang lebih baik dalam konteks penelitian tersebut.