

Penerapan Algoritma K-Means Clustering untuk Mengetahui Pola Penerima Beasiswa Bank Indonesia (BI)

Qurrata A'yuni*, Alwis Nazir, Lestari Handayani, Iis Afrianty

Sains dan Teknologi, Teknik Informatika, Universitas Islam Negeri Sultan Syarif Kasim Riau, Pekanbaru, Indonesia

Email: ¹*11950121728@students.uin-suska.ac.id, ²alwis.nazir@uin-suska.ac.id, ³lestari.handayani@uin-suska.ac.id,

⁴iis.afrianty@uin-suska.ac.id

Email Penulis Korespondensi: 11950121728@students.uin-suska.ac.id

Abstrak—Beasiswa Bank Indonesia merupakan jenis beasiswa yang bersumber dari Bank Indonesia untuk mahasiswa dari Perguruan Tinggi Negeri, Perguruan Tinggi Swasta, serta Politeknik yang terpilih. Dari data penerima beasiswa yang sudah lulus seleksi dari tahun 2020, 2021, 2022 perguruan tinggi di Riau, perlu dilakukan pencarian pola karakteristik penerima beasiswa, karena pihak Bank Indonesia belum mempunyai pola. Untuk mengetahui pola penerima beasiswa, peneliti menerapkan metode data mining yakni dengan algoritma K-Means Clustering. Parameter yang digunakan berjumlah 4 yaitu program studi, semester, IPK, dan jenjang. Hasil dari penelitian dengan menggunakan RapidMiner diperoleh cluster 0 didominasi oleh mahasiswa dari program studi Ketatalaksanaan Pelayaran Niaga, berada di semester 5, dan jenjang D3. Cluster 1 didominasi oleh mahasiswa dari program studi Akuntansi dan Manajemen, berada di semester 7, dengan IPK besar dari atau sama dengan 3.51, dan jenjang S1. Cluster 2 didominasi oleh mahasiswa dari program studi Keperawatan, berada di semester 5, dengan IPK besar dari atau sama dengan 3.51, dan jenjang D3. Cluster 3 didominasi oleh mahasiswa oleh program studi Hubungan Internasional, berada di semester 7, dengan IPK besar dari atau sama dengan 3.51, dan jenjang S1. Cluster 4 didominasi oleh mahasiswa dari program studi Teknik Informatika, berada di semester 5, dengan IPK besar dari atau sama dengan 3.51, dan jenjang S1. Dapat dilihat bahwa penerima beasiswa Bank Indonesia didominasi oleh mahasiswa dengan nilai IPK tinggi atau sama dengan 3.51. Selain itu juga didominasi oleh mahasiswa yang berada di jenjang S1. Pengujian dilakukan menggunakan DBI dengan k=5 menghasilkan nilai validitas sebesar 0.121.

Kata Kunci: Beasiswa; Data Mining; K-Means; Pola

Abstract—Bank Indonesia Scholarships are a type of scholarship sourced from Bank Indonesia for students from selected State Universities, Private Universities, and Polytechnics. From the data on scholarship recipients who have passed the selection from 2020, 2021, 2022 universities in Riau, it is necessary to look for the behavior patterns of scholarship recipient because Bank Indonesia does not yet have a pattern. To find the pattern from scholarship recipients using the method of data mining with K-Means Clustering algorithm. The parameters used are 4, namely study program, semester, GPA, and level. The results of the study using RapidMiner showed that cluster 0 was dominated by students from the Commerce Shipping Management study program, who were in semester 5 and D3 level. Cluster 1 is dominated by students from the Accounting and Management study program, in semester 7, with GPA greater than or equal to 3.51, and S1 level. Cluster 2 is dominated by students from the Nursing study program, in semester 5, with GPA greater than or equal to 3.51, and D3 level. Cluster 3 is dominated by students from the International Relations study program, in semester 7, with GPA greater than or equal to 3.51, and S1 level. Cluster 4 is dominated by students from the Informatics Engineering study program, in semester 5, with GPA greater than or equal to 3.51, and S1 level. It shows that the recipients of Bank Indonesia scholarships are dominated by students with high GPA scores or equal to 3.51. In addition, it is also dominated by students who are at the S1 level. Tests were carried out using DBI with k=5 resulting in a validity value of 0.121.

Keywords: Scholarship; Data Mining; K-Means; Pattern

1. PENDAHULUAN

Bangsa Indonesia memiliki banyak anak-anak yang mempunyai kemampuan belajar bagus dan keinginan tinggi untuk mencapai cita-citanya [1],[2]. Namun, hal ini terhambat oleh masalah keuangan dan tekanan ekonomi, terutama yang dialami oleh keluarga dan mahasiswa. Saat ini pemerintah melakukan berbagai upaya untuk menyelesaikan permasalahan ekonomi yang dihadapi oleh mahasiswa. Upaya yang dilakukan pemerintah untuk membantu mahasiswa yaitu dengan dibentuknya program dana bantuan berupa beasiswa [3].

Beasiswa adalah pemberian berupa bantuan baik itu berupa barang maupun uang yang diberikan untuk perindividu bertujuan membantu pendidikan yang tengah dijalani [4],[5]. Beasiswa juga disebut dengan pendanaan perindividu yang tidak disediakan oleh dana sendiri atau orang tua untuk dipakai sebagai kelanjutan pendidikan [6],[7]. Akan tetapi, hal itu disediakan oleh pemerintah, lembaga pendidikan, universitas, dan perusahaan swasta. Tujuan dari program beasiswa adalah untuk meringankan tekanan keuangan yang dihadapi mahasiswa dalam menempuh masa studi kuliah [3]. Mahasiswa yang dipilih untuk beasiswa berdasarkan jenis beasiswa yang ditawarkan. Salah satu program beasiswa yang ada saat ini ialah beasiswa yang dipersembahkan Bank Indonesia (BI).

Mahasiswa pada Perguruan Tinggi Negeri, Perguruan Tinggi Swasta, dan Politeknik tertentu di Indonesia dapat mengajukan dan menerima Beasiswa Bank Indonesia [6]. Beasiswa dari Bank Indonesia ini akan membantu hal-hal seperti biaya kuliah, buku, dan biaya hidup [8]. Agar memenuhi syarat untuk mendapatkan Beasiswa Bank Indonesia, seorang siswa harus berasal dari latar belakang berpenghasilan rendah, telah menyelesaikan 40 atau berada di semester ketiga perguruan tinggi, tidak lebih dari 23 tahun pada saat mereka dipilih, bukan penerima beasiswa lainnya, dan bersedia berkontribusi aktif dalam pengelolaan dan pengembangan Generasi Baru Indonesia

(GenBI) [9]. Persyaratan tersebut harus dilengkapi oleh mahasiswa untuk melakukan pendaftaran beasiswa Bank Indonesia. Setelah memenuhi persyaratan, maka akan dilakukan proses seleksi.

Dari data penerima beasiswa yang sudah lulus seleksi dari tahun 2020, 2021, 2022 perguruan tinggi di Riau, perlu dilakukan pencarian pola karakteristik penerima beasiswa Bank Indonesia (BI). Menurut hasil wawancara dengan pembina Beasiswa Bank Indonesia, hal ini perlu dilakukan karena pihak Bank Indonesia (BI) belum mempunyai pola penerima beasiswa tersebut. Untuk pencarian pola tersebut, akan dilakukan pengelompokan dari data penerima beasiswa. Metode yang dipakai untuk pengelompokan dalam penelitian ini yaitu algoritma K-Means Clustering [10],[11].

Algoritma K-Means clustering merupakan teknik data mining yang dapat mengelompokkan objek berdasarkan kesamaan karakteristiknya [12]. Algoritma K-Means Clustering bisa meng-cluster banyak data dalam waktu relatif singkat dan efisiensi yang tinggi [13],[14]. Penggunaan algoritma ini pilihan yang tepat dalam melakukan pengelompokan sebuah data [1]. Teknik K-Means Clustering berguna untuk mengatur dan menyederhanakan berbagai macam data terkait cluster karena mudah digunakan dan diimplementasikan.

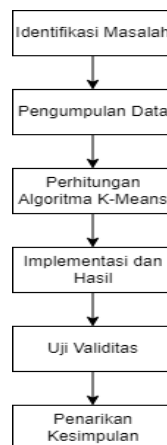
Beberapa penelitian terkait yang menggunakan K-Means Clustering adalah pengelompokan pemeroleh beasiswa dari Yayasan untuk mahasiswa dengan menggunakan algoritma K-Means Clustering yang diimplementasikan menggunakan RapidMiner. Hasil penelitian menunjukkan 3 kategori beasiswa Yayasan, yaitu 50%, 60%, dan 100% [15]. Dalam penelitian lain terkait rekomendasi calon mahasiswa yang mendapatkan beasiswa Yayasan dengan algoritma K-Means Clustering. Hasil penelitian menunjukkan 23% calon mahasiswa ditolak, 61% dipertimbangkan, dan 16% diterima [16]. Penelitian selanjutnya oleh [17] terkait penentuan calon beasiswa UKT dengan K-Means Clustering yang diimplementasikan menggunakan RapidMiner. Hasilnya menunjukkan 215 anggota layak menerima beasiswa UKT, 104 anggota tidak layak, dan 202 kurang layak.

Para peneliti juga membandingkan K-Means dengan berbagai metode pengelompokan lain dalam studi yang diteliti. Penelitian [18] menunjukkan bahwa dalam hal pengelompokan produk asuransi yang ditawarkan oleh perusahaan nasional, teknik K-Means mencapai nilai Davies Boldin Index (DBI) yang lebih baik daripada metode K-Medoids. Selain itu, dikembangkan aplikasi clustering armada kendaraan berbasis web dengan metode studi K-Means karena relevansinya tinggi berdasarkan nilai DBI [19].

Tujuan dilakukannya penelitian ini ialah untuk memberi informasi yang memiliki manfaat untuk pihak-pihak yang terkait dalam pengidentifikasian pola khas penerima beasiswa Bank Indonesia (BI) di Provinsi Riau dengan menggunakan teknik K-Means Clustering.

2. METODOLOGI PENELITIAN

Tahapan dari penelitian yang direncanakan digambarkan dalam Gambar 1 sebagai berikut.



Gambar 1. Tahapan Penelitian

2.1 Identifikasi Masalah

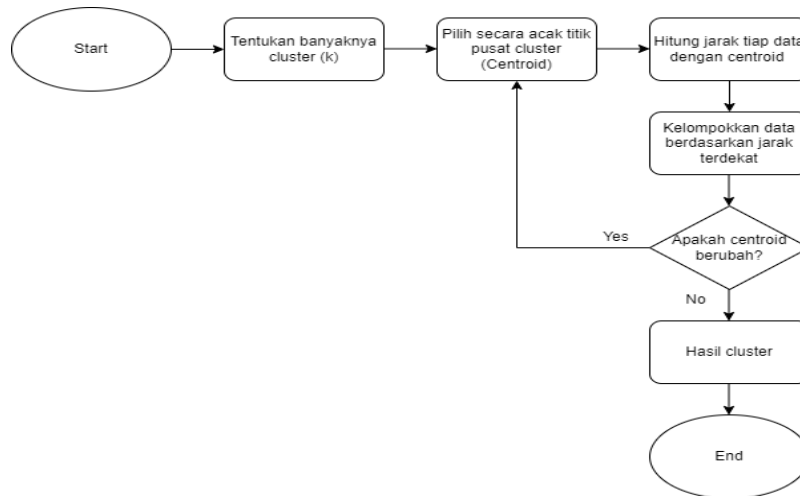
Analisa masalah dilakukan pada tahap ini, dimulai dengan deskripsi masalah hingga mengetahui permasalahan yang ada untuk mengidentifikasi solusi yang paling efektif.

2.2 Pengumpulan Data

Data dalam penelitian ini didapatkan dari pihak Humas Bank Indonesia yang sekaligus menjadi pembina penerima beasiswa Bank Indonesia. Data yang diperoleh yakni data penerima beasiswa tahun 2020, 2021, dan 2022 berjumlah 1077 data yang terdiri dari 6 perguruan tinggi di Riau yaitu Universitas Islam Negeri Sultan Syarif Kasim Riau (USR), Universitas Lancang Kuning (UNILAK), Politeknik Bengkalis (POLBENG), Sekolah Tinggi Teknologi Dumai (STTD), Universitas Riau (UR), dan Universitas Muhammadiyah Riau (UMRI). Parameter yang diperoleh terdiri dari nama, program studi, indeks prestasi kumulatif (IPK), semester, dan jenjang.

2.3 Perhitungan Algoritma K-Means

Data yang didapatkan pada tahap pengumpulan data, akan diproses dan diolah mengimplementasikan algoritma K-Means. Flowchart untuk algoritma K-Means dapat dilihat pada Gambar 2.



Gambar 2. Flowchart Algoritma K-Means

Tahapan untuk perhitungan algoritma K-Means dijelaskan di bawah ini:

- Tetapkan banyaknya jumlah cluster (k)
- Pilih random titik pusat untuk cluster (centroid)
- Pakai rumus Euclidean Distance untuk mendapatkan jarak tiap data terhadap centroid, rumusnya yakni:

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1)$$

Keterangan:

$d(x, y)$: jarak data x dan y
 x : titik data objek
 y : titik data centroid
 i : banyaknya objek

- Kelompokkan data yang sudah dikalkulasikan menurut jarak terkecil (minimum) antara data tersebut dengan pusat cluster atau data centroid dan mendapatkan cluster baru.
- Lakukan perhitungan kembali berdasarkan data yang mengikuti cluster masing-masing untuk pusat cluster (centroid) baru. Nilai centroid baru didapatkan dari hasil perhitungan rata-rata data terhadap setiap cluster. Rumus sebagai Berikut:

$$CI = \frac{x_1 + x_2 + x_3 + \dots + x_n}{\sum x} \quad (2)$$

Keterangan:

CI : centroid baru
 x_1 : nilai cluster ke-1
 x_n : nilai cluster ke-n
 x : jumlah data

- Setelah dapat centroid baru, maka lakukan iterasi selanjutnya atau ulangi langkah c sampai e sampai tidak ditemukan data yang berpindah-pindah dari cluster.

2.4 Implementasi dan Hasil

Implementasi dalam penelitian ini menggunakan bantuan tools rapidminer untuk mendapatkan hasil yang diinginkan.

2.5 Uji Validitas

Tahap ini dilakukan guna menemukan jumlah cluster terbaik dari pemodelan cluster. Pengujian pada penelitian ini menggunakan Davies Bouldien Index (DBI). DBI merupakan suatu teknik yang digunakan untuk menilai kualitas sebuah cluster dengan melihat kedekatan antar hasil cluster [20]. Hasil cluster nilai DBI akan semakin baik jika dekat dengan 0 yang diperoleh menggunakan metode clustering.

2.6 Penarikan Kesimpulan

Hasil kajian yang diperoleh digunakan untuk membuat kesimpulan dengan menggunakan algoritma K-Means.

3. HASIL DAN PEMBAHASAN

Proses penelitian menggunakan data dari penerima Beasiswa Bank Indonesia di Provinsi Riau tahun 2020, 2021, 2022. Parameter yang digunakan pada penelitian adalah nama, program studi, indeks prestasi kumulatif (IPK), semester, dan jenjang. Data penerima beasiswa bisa dilihat pada Tabel 1.

Tabel 1. Penerima Beasiswa Bank Indonesia

No	Nama	Program Studi	Semester	IPK	Jenjang
1	Fauzi Nur Rizki	Pendidikan Matematika	5	3,40	S1
2	Nadya Tafari	Pendidikan Matematika	7	3,54	S1
3	Dhiya U Syamsi Jamali	Sistem Informasi	7	3,65	S1
4	Dinda Andani Putri	Agroteknologi	7	3,46	S1
5	Fiya Fhadilah Ihsani	Agroteknologi	7	3,82	S1
6	Fahrul Rozi	Ilmu Komunikasi	7	3,56	S1
7	Iismuliati	Peternakan	7	3,53	S1
8	Firman Syahputra	Peternakan	5	3,53	S1
9	Haridsyah	Peternakan	5	3,67	S1
10	Sherin Rizkina	Agroteknologi	5	3,89	S1
...	...	...	...	...	...
...	...	...	...	...	...
1076	Rizma Tia Yuniar	Keperawatan	5	3,84	D3
1077	Yurdi Halawa	Keperawatan	5	3,89	D3

3.1 Pemodelan Data

Data kriteria program studi, semester, IPK, dan jenjang digunakan dalam proses clustering. Untuk mempermudah penelitian agar bisa melihat pola karakteristik data program studi, IPK, dan jenjang akan dikategorikan menjadi beberapa bagian. Proses pengategorian pada Tabel 2 untuk parameter program studi sebagai berikut.

Tabel 2. Kategori Program Studi

Program Studi	Ket
Pendidikan Matematika	1
Sistem Informasi	2
Agroteknologi	3

Untuk parameter IPK dilakukan pemodelan data dengan diubah ke dalam bentuk numerical. Proses pengategorian pada Tabel 3 untuk parameter IPK sebagai berikut.

Tabel 3. Kategori IPK

IPK	Ket
$\leq 3,50$	1
$> 3,51$	2

Parameter Jenjang dilakukan pemodelan data diubah ke dalam bentuk numerical. Proses pengategorian pada Tabel 4 untuk parameter jenjang sebagai berikut.

Tabel 4. Kategori Jenjang

Jenjang	Ket
D3	1
S1	2

3.2 Pengolahan Data

Data pada Tabel 1 kemudian diproses dengan perhitungan K-Means Clustering. Berikut proses pengolahan datanya.

- Tentukan banyaknya cluster(k). Jumlah cluster ditentukan berdasarkan kebutuhan yang ingin dihasilkan. Jumlah cluster yang akan digunakan yaitu 5 cluster.
- Tentukan nilai titik cluster (centroid) awal sebanyak cluster (k). Nilai centroid diperoleh secara acak. Berikut titik cluster awal pada Tabel 5.

Tabel 5. Titik Cluster Awal (Iterasi-1)

Centroid	Program Studi	Semester	IPK	Jenjang
C0	2	7	2	2
C1	3	7	2	2

Centroid	Program Studi	Semester	IPK	Jenjang
C2	4	5	2	2
C3	4	5	2	2
C4	5	5	2	2

- c. Menghitung jarak terdekat data terhadap setiap nilai centroid yakni dengan Persamaan (1).

1. Perhitungan data ke-1 terhadap centroid cluster 0

$$d(1,0) = \sqrt{(1-2)^2 + (5-7)^2 + (1-2)^2 + (2-2)^2} = 6$$

2. Perhitungan data ke-1 terhadap centroid cluster 1

$$d(1,1) = \sqrt{(1-3)^2 + (5-7)^2 + (1-2)^2 + (2-2)^2} = 7$$

3. Perhitungan data ke-1 terhadap centroid cluster 2

$$d(1,2) = \sqrt{(1-4)^2 + (5-5)^2 + (1-2)^2 + (2-2)^2} = 4$$

4. Perhitungan data ke-1 terhadap centroid cluster 3

$$d(1,3) = \sqrt{(1-4)^2 + (5-7)^2 + (1-2)^2 + (2-2)^2} = 8$$

5. Perhitungan data ke-1 terhadap centroid cluster 4

$$d(1,4) = \sqrt{(1-5)^2 + (5-5)^2 + (1-2)^2 + (2-2)^2} = 5$$

Berdasarkan hasil perhitungan dengan rumus Euclidean Distance pada data ke-1, maka didapatkan hasil bahwa data ke-1 paling dekat dengan cluster 2 berdasarkan nilai terkecil diantara nilai perhitungan cluster lainnya. Sehingga data ke-1 termasuk kedalam kelompok cluster 2. Selanjutnya, perhitungan dilakukan sampai data ke-1077 hingga masing-masing data menemukan hasil jarak terdekat dengan cluster.

- d. Setelah dilakukan perhitungan jarak data terhadap centroid, selanjutnya adalah pengelompokan data berdasarkan nilai terdekat (terkecil). Hasil pengelompokan dapat dilihat dalam Tabel 6.

Tabel 6. Hasil Pengelompokan Data

Objek ke-	C0	C1	C2	C3	C4	Jarak Terkecil	Cluster
1	6	7	4	8	5	4	2
2	1	2	7	3	8	1	0
3	0	1	6	2	7	0	0
4	1	0	5	1	6	0	1
5	1	0	5	1	6	0	1
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
1076	53	52	47	51	46	46	4
1077	53	52	47	51	46	46	4

- e. Hitung nilai rata-rata setiap atribut data dalam cluster dengan memperbarui nilai centroid baru untuk iterasi berikutnya. Berikut perhitungan nilai centroid baru yakni dengan Persamaan (2).

1. Rata-rata data pada cluster 0

$$CI(\text{prodi}) = \frac{51}{28} = 1.82$$

$$CI(\text{semester}) = \frac{196}{28} = 7$$

$$CI(\text{ipk}) = \frac{50}{28} = 1.78$$

$$CI(\text{strata}) = \frac{56}{28} = 2$$

2. Rata-rata data pada cluster 1

$$CI(\text{prodi}) = \frac{45}{15} = 3$$

$$CI(\text{semester}) = \frac{105}{15} = 7$$

$$CI(\text{ipk}) = \frac{29}{15} = 1.93$$

$$CI(\text{strata}) = \frac{30}{15} = 2$$

3. Rata-rata data pada cluster 2

$$CI(\text{prodi}) = \frac{193}{61} = 3.16$$

$$CI(\text{semester}) = \frac{305}{61} = 5$$

$$CI(ipk) = \frac{119}{61} = 1.95$$

$$CI(strata) = \frac{122}{61} = 2$$

4. Rata-rata data pada cluster 3

$$CI(prodi) = \frac{7341}{432} = 16.99$$

$$CI(semester) = \frac{3026}{432} = 7.00$$

$$CI(ipk) = \frac{770}{432} = 1.78$$

$$CI(strata) = \frac{855}{432} = 1.97$$

5. Rata-rata data pada cluster 4

$$CI(prodi) = \frac{11277}{541} = 20.84$$

$$CI(semester) = \frac{2536}{541} = 4.68$$

$$CI(ipk) = \frac{932}{541} = 1.72$$

$$CI(strata) = \frac{931}{541} = 1.72$$

Hasil dari kalkulasi untuk nilai centroid baru di iterasi-2 bisa dilihat dalam Tabel 7 berikut.

Tabel 7. Centroid Baru (Iterasi-2)

Centroid	Program Studi	Semester	IPK	Jenjang
C0	1.82	7	1.78	2
C1	3	7	1.93	2
C2	3.16	5	1.95	2
C3	16.99	7.00	1.78	1.97
C4	20.84	4.68	1.72	1.72

- f. Lakukan perhitungan kembali jarak data dengan nilai centroid baru. Proses c hingga e akan diulang kembali, jika tidak ada data yang berpindah cluster atau iterasi terakhir memiliki data yang sama dengan iterasi sebelumnya.

3.3 Implementasi dan Hasil

Setelah data diolah, selanjutnya dilakukan pengimplementasian data kedalam sebuah tools untuk menghasilkan sebuah model dan hasil. RapidMiner yang digunakan sebagai Tools. Pada Gambar 3 menggambarkan hasil cluster model dengan menggunakan RapidMiner dapat dilihat sebagai berikut.

Cluster Model

```
Cluster 0: 145 items
Cluster 1: 463 items
Cluster 2: 82 items
Cluster 3: 147 items
Cluster 4: 240 items
Total number of items: 1077
```

Gambar 3. Cluster Model

Hasil cluster pada Gambar 3 menunjukkan cluster 0 mempunyai 145 items, cluster 1 mempunyai 463 items, cluster 2 mempunyai 82 items, cluster 3 mempunyai 147 items, dan cluster 4 mempunyai 240 items. Berikut data hasil cluster 0 pada Tabel 8.

Tabel 8. Hasil Cluster 0

No	Nama	Program Studi	Semester	IPK	Jenjang	Cluster
1	Anwar Bukhori	Pertanian	5	1	2	0
2	Rubiati	Ketatalaksanaan Pelayaran Niaga	3	1	1	0
3	Fitri Handayani	Ketatalaksanaan Pelayaran Niaga	3	1	1	0
4	Sindi Claudia	Ketatalaksanaan Pelayaran Niaga	3	2	1	0
5	Dumadi	Ketatalaksanaan Pelayaran Niaga	3	2	1	0
6	Muhammad Padil	Ketatalaksanaan Pelayaran Niaga	3	1	1	0

No	Nama	Program Studi	Semester	IPK	Jenjang	Cluster
7	Akbar Maulana	Ketatalaksanaan Pelayaran Niaga	3	2	1	0
8	M. Rizki	Ketatalaksanaan Pelayaran Niaga	5	2	1	0
9	Muhamad Syahir	Ketatalaksanaan Pelayaran Niaga	5	2	1	0
10	Andre Syahreza	Ketatalaksanaan Pelayaran Niaga	5	2	1	0
11	Budi Rahmadani Putra	Ketatalaksanaan Pelayaran Niaga	3	1	1	0
12	Suwandi	Ketatalaksanaan Pelayaran Niaga	3	1	1	0
13	Murniati	Ketatalaksanaan Pelayaran Niaga	3	1	1	0
14	Alif Mukminin	Ketatalaksanaan Pelayaran Niaga	3	1	1	0
...	...	...	...	...	...	...
...	...	...	...	...	...	...
144	Mutia Fatma	Administrasi Bisnis	5	2	2	0
145	Sinta Mutiara Br	Administrasi Bisnis	5	2	2	0

Hasil cluster 0 menunjukkan penerima beasiswa dari program studi ketatalaksanaan pelayaran niaga berjumlah 38 orang, administrasi bisnis 25 orang, teknik sipil 24 orang, akuntansi keuangan publik 24 orang, nautika 16 orang, teknik perkapalan 6 orang, pertanian 2 orang, bahasa inggris 3 orang, teknik elektro 3 orang, teknik mesin 2 orang, dan rekayasa perangkat lunak 2 orang. Parameter semester menunjukkan mahasiswa yang berada di semester tiga berjumlah 34 orang, semester lima 91 orang, semester tujuh 19 orang, dan semester sembilan 1 orang. Parameter IPK menunjukkan mahasiswa yang mempunyai nilai IPK kecil atau sama dengan 3.50 berjumlah 73 orang dan besar atau sama dengan 3.51 berjumlah 72 orang. Sedangkan untuk parameter jenjang mahasiswa yang berada di D3 berjumlah 89 orang dan S1 56 orang. Dapat dilihat bahwa hasil cluster 0 didominasi oleh mahasiswa dari program studi ketatalaksanaan pelayaran niaga, berada di semester 5, dan jenjang D3. Berikut data hasil cluster 1 pada Tabel 9.

Tabel 9. Hasil Cluster 1

No	Nama	Program Studi	Semester	IPK	Jenjang	Cluster
1	Nadia Tifany	Manajemen	5	2	2	1
2	Al Fiona Sinta	Manajemen	5	2	2	1
3	Nina Sriwahyuni	Manajemen	7	2	2	1
4	Galuh Diah Pitaloka	Manajemen	7	2	2	1
5	Muftia Fitri Zahra	Manajemen	7	2	2	1
6	Eriska Yuliarti	Akuntansi	7	2	2	1
7	Wulan	Akuntansi	5	2	2	1
...	...	...	...	...	...	...
...	...	...	...	...	...	...
457	Ridho Dwi Darma	Akuntansi	7	2	2	1
458	Shandy Ramadhan	Akuntansi	7	2	2	1
459	Sopia Ananda	Akuntansi	5	2	2	1
460	Tio Irfandi	Ilmu Komunikasi	5	2	2	1
461	Tri Rahayu Ramadani	Akuntansi	7	2	2	1
462	Wahyu Miftahani	Akuntansi	5	2	2	1
463	Yudi Winaldi	Akuntansi	7	2	2	1

Hasil cluster 1 menunjukkan penerima beasiswa dari program studi akuntansi berjumlah 110 orang, manajemen 108 orang, ilmu komunikasi 68 orang, ilmu hukum 65 orang, sistem informasi 40 orang, ekonomi syariah 31 orang, agroteknologi 23 orang, peternakan 10 orang, dan pendidikan matematika 8 orang. Parameter semester menunjukkan mahasiswa yang berada di semester tiga berjumlah 6 orang, semester lima 195 orang, semester tujuh 260 orang, dan semester sembilan 2 orang. Parameter IPK menunjukkan mahasiswa yang memiliki nilai IPK kecil atau sama dengan 3.50 berjumlah 59 orang dan besar atau sama dengan 3.51 berjumlah 404 orang. Sedangkan untuk parameter jenjang mahasiswa yang berada di D3 berjumlah 9 orang dan S1 454 orang. Dapat dilihat bahwa hasil cluster 1 didominasi oleh mahasiswa dari program studi akuntansi dan manajemen, berada di semester 7, dengan IPK tinggi atau sama dengan 3.51, dan jenjang S1. Berikut data hasil cluster 2 pada Tabel 10.

Tabel 10. Hasil Cluster 2

No	Nama	Program Studi	Semester	IPK	Jenjang	Cluster
1	Chyntya Dwi Rahmadani	Keperawatan	5	2	1	2
2	Delvi Delvira	Keperawatan	3	2	1	2
3	Deya Oktarysa	Keperawatan	5	2	1	2
4	Dinda Puspita Dewi	Keperawatan	3	2	1	2
5	Evitasari	Keuangan dan Perbankan	3	2	1	2
6	Friska Nabila Putri	Keperawatan	3	2	1	2

No	Nama	Program Studi	Semester	IPK	Jenjang	Cluster
7	Hasambua	Keperawatan	3	2	1	2
8	Lala Narwiyana	Keperawatan	3	2	1	2
9	Razky Ihsanul Sably	Keperawatan	3	2	1	2
10	Retni Rindayu Putri	Keperawatan	3	2	1	2
11	Rizky Nur Handayani	Keuangan dan Perbankan	5	2	1	2
12	Roma Dela Br Manik	Keperawatan	5	2	1	2
...	...	...	...	...	...	...
...	...	...	...	...	...	...
81	Rizma Tia Yuniar	Keperawatan	5	2	1	2
82	Yurdi Halawa	Keperawatan	5	2	1	2

Hasil cluster 2 menunjukkan penerima beasiswa dari program studi keperawatan berjumlah 37 orang, keuangan dan perbankan 9 orang, pendidikan ipa 7 orang, usaha perjalanan wisata 7 orang, pendidikan bahasa inggris 5 orang, perbankan syariah 5 orang, kehutanan 3 orang, hukum tata negara 2 orang, psikologi islam 2 orang, perpajakan 2 orang, perdata bisnis 1 orang, ilmu komputer 1 orang, dan hukum pidana 1 orang. Parameter semester menunjukkan mahasiswa yang berada di semester tiga berjumlah 10 orang, semester lima 44 orang, dan semester tujuh 28 orang. Parameter IPK menunjukkan mahasiswa yang memiliki nilai IPK kecil atau sama dengan 3.50 berjumlah 3 orang dan besar atau sama dengan 3.51 79 orang. Sedangkan untuk parameter jenjang mahasiswa yang berada di D3 berjumlah 49 orang dan S1 33 orang. Dapat dilihat bahwa hasil cluster 2 didominasi oleh mahasiswa dari program studi Keperawatan, berada di semester 5, dengan IPK tinggi atau sama dengan 3.51, dan jenjang D3. Berikut hasil cluster 3 pada Tabel 11.

Tabel 11. Hasil Cluster 3

No	Nama	Program Studi	Semester	IPK	Jenjang	Cluster
1	Bob Marta	Hubungan Internasional	5	2	2	3
2	Rud Calara Cintia	Hubungan Internasional	7	1	2	3
3	Irfandi Hernaldi	Pendidikan Ekonomi	7	2	2	3
4	Elfira	Teknologi Hasil Pertanian	5	1	2	3
5	Aisyah Sri Utari	Ekonomi Pembangunan	7	2	2	3
6	Jeantama Akmadani	Ekonomi Pembangunan	7	1	2	3
7	Eko Sumantri	Teknologi Hasil Pertanian	5	1	2	3
8	Anisa Rahmadani	Pendidikan Ekonomi	7	2	2	3
9	Jodi Misyendi	Hubungan Internasional	7	2	2	3
10	Laposa Mirta Dea Roja	Hubungan Internasional	5	2	2	3
11	Dymas Ilham D	Hubungan Internasional	7	2	2	3
...	...	...	...	...	...	...
...	...	...	...	...	...	...
144	Tazkia Surryihda K	Hubungan Internasional	5	2	2	3
145	Virginia Asoka Puri	Ekonomi Pembangunan	7	2	2	3
146	Winda Safitri	Ekonomi Pembangunan	5	2	2	3
147	Yessy Anggina Lubis	Ekonomi Pembangunan	5	2	2	3

Hasil cluster 3 menunjukkan penerima beasiswa dari program studi hubungan internasional berjumlah 34 orang, pendidikan ekonomi 26 orang, ekonomi pembangunan 19 orang, ilmu pemerintahan 16 orang, administrasi bisnis internasional 16 orang, teknik perancangan jalan dan jembatan 8 orang, teknik pertanian 7 orang, sosiologi 7 orang, teknologi hasil pertanian 5 orang, bisnis internasional 3 orang, teknologi industri pertanian 2 orang, teknik listrik 2 orang, agrobisnis perikanan 1 orang, dan TM produksi dan perawatan 1 orang. Parameter semester menunjukkan mahasiswa yang berada di semester tiga berjumlah 5 orang, semester lima 68 orang, dan semester tujuh 74 orang. Parameter IPK menunjukkan mahasiswa yang memiliki nilai IPK kecil atau sama dengan 3.50 berjumlah 39 orang dan besar atau sama dengan 3.51 108 orang. Sedangkan untuk parameter jenjang mahasiswa berada di S1 berjumlah 33 orang. Dapat dilihat bahwa hasil cluster 3 didominasi oleh mahasiswa oleh program studi Hubungan Internasional, berada di semester 7, dengan IPK tinggi atau sama dengan 3.51, dan jenjang S1. Berikut hasil cluster 4 pada Tabel 12.

Tabel 12. Hasil Cluster 4

No	Nama	Program Studi	Semester	IPK	Jenjang	Cluster
1	Chonie Merintiona	Hukum Ekonomi Syariah	5	2	2	4
2	Nabila Frizka Putri F	Hukum Ekonomi Syariah	5	2	2	4
3	Silvia Nadrah N	Ekonomi Islam	5	2	2	4
4	Erika Lumban Gaol	Agribisnis	7	1	2	4
5	Rafesa Berlian Naldo	Teknik Informatika	7	2	2	4

No	Nama	Program Studi	Semester	IPK	Jenjang	Cluster
6	Juni Yulfianti Hutabarat	Ilmu Administrasi Negara	5	2	2	4
7	Talenta Yohana K. Panjaitan	Ilmu Administrasi Negara	5	1	2	4
8	Diana Sheren Hutagalung	Ilmu Administrasi Publik	7	1	2	4
9	Lisna Christiany S	Ilmu Administrasi Negara	5	2	2	4
10	Feryandi Hutapea	Ilmu Administrasi Negara	5	2	2	4
11	Ilfan Sukri	Ilmu Administrasi Negara	5	1	2	4
12	Delfy Agustina	Ilmu Administrasi Negara	5	1	2	4
13	Elfira Nurul Ilma	Teknik Informatika	3	1	1	4
14	Fitrah Desmalini	Teknik Informatika	3	1	1	4
15	Muhammad Hafiz	Teknik Informatika	5	1	1	4
16	Zulfikar Fitrah	Teknik Informatika	5	1	1	4
17	Rabiatul Rahmadanis	Teknik Informatika	5	1	1	4
18	Fera Fitrianti	Teknik Informatika	5	2	1	4
19	Abu Dzar Musaiyid	Teknik Informatika	3	2	2	4
20	Ahmad Nasa'i	Teknik Informatika	3	2	2	4
21	Alfian Nur Ardhi	Teknik Informatika	3	2	2	4
22	Alfina Sukma	Teknik Informatika	3	2	2	4
...	...	...	...	...	...	...
...	...	...	...	...	...	...
237	Muhammad Abdullah	Teknik Informatika	5	2	2	4
238	Muhammad Albani	Teknik Informatika	7	2	2	4
239	Muhammad Leonza	Teknik Informatika	5	2	2	4
240	Putri Altika Sari	Teknik Informatika	5	2	2	4

Hasil cluster 4 menunjukkan penerima beasiswa dari program studi teknik informatika berjumlah 130 orang, ilmu administrasi negara 44 orang, teknik industri 29 orang, agribisnis 16 orang, matematika 8 orang, ilmu administrasi publik 7 orang, hukum ekonomi syariah 4 orang, ekonomi islam 1 orang, dan gizi 1 orang. Parameter semester menunjukkan mahasiswa yang berada di semester tiga berjumlah 29 orang, semester lima 120 orang, dan semester tujuh 91 orang. Parameter IPK menunjukkan mahasiswa yang memiliki nilai IPK kecil atau sama dengan 3.50 berjumlah 79 orang dan besar atau sama dengan 3.51 161 orang. Sedangkan untuk parameter jenjang mahasiswa yang berada di D3 berjumlah 11 orang dan S1 229 orang. Dapat dilihat bahwa hasil cluster 4 didominasi oleh mahasiswa dari program studi Teknik Informatika, berada di semester 5, dengan IPK tinggi atau sama dengan 3.51, dan jenjang S1.

3.4 Pengujian Data

Tahapan pengujian data dilakukan untuk memperoleh informasi terkait total cluster yang terbaik. Dalam penelitian ini, metode DBI dipakai untuk pengujian data. Pengujian dilakukan untuk melihat nilai jarak antar cluster dengan melakukan percobaan mulai dari 3 sampai 7 cluster. Metode DBI menyatakan bahwa 5 cluster adalah cluster yang terbaik diantara cluster lainnya. Pada metode DBI menghasilkan validitas sebesar 0.121. Berikut pada Gambar 4 merupakan hasil dari uji performance dengan implementasi DBI.

PerformanceVector

```
PerformanceVector:
Avg. within centroid distance: 1.839
Avg. within centroid distance_cluster_0: 3.068
Avg. within centroid distance_cluster_1: 1.471
Avg. within centroid distance_cluster_2: 1.991
Avg. within centroid distance_cluster_3: 2.697
Avg. within centroid distance_cluster_4: 1.230
Davies Bouldin: 0.121
```

Gambar 4. Hasil Uji Performance

4. KESIMPULAN

Setelah dilakukan penelitian, diperoleh kesimpulan yaitu parameter yang digunakan berjumlah 4 yaitu program studi, semester, IPK, dan jenjang. Total k yang digunakan yakni sebanyak 5 (lima). Hasil pengujian kelayakan menggunakan RapidMiner untuk 1077 data penerima beasiswa Bank Indonesia tahun 2020, 2021, 2022 diperoleh cluster 0 didominasi oleh mahasiswa dari program studi Ketatalaksanaan Pelayaran Niaga, berada di semester 5, dan jenjang D3. Cluster 1 didominasi oleh mahasiswa dari program studi Akuntansi dan Manajemen, berada di semester 7, dengan IPK besar dari atau sama dengan 3.51, dan jenjang S1. Cluster 2 didominasi oleh mahasiswa dari program studi Keperawatan, berada di semester 5, dengan IPK besar dari atau sama dengan 3.51, dan jenjang

D3. Cluster 3 didominasi oleh mahasiswa oleh program studi Hubungan Internasional, berada di semester 7, dengan IPK besar dari atau sama dengan 3.51, dan jenjang S1. Cluster 4 didominasi oleh mahasiswa dari program studi Teknik Informatika, berada di semester 5, dengan IPK besar dari atau sama dengan 3.51, dan jenjang S1. Dapat dilihat bahwa penerima beasiswa Bank Indonesia didominasi oleh mahasiswa yang memiliki IPK tinggi atau sama dengan 3.51 dan juga berada di jenjang S1. Hasil pengujian yang dilakukan menggunakan metode DBI dengan nilai validitas sebesar 0.121 diperoleh cluster terbaik pada $k=5$.

REFERENCES

- [1] B. G. Sudarsono and S. P. Lestari, "Clustering Penerima Beasiswa Yayasan Untuk Mahasiswa Menggunakan Metode K-Means," *J. Media Inform. Budidarma*, vol. 5, no. 1, p. 258, 2021, doi: 10.30865/mib.v5i1.2670.
- [2] H. S. Tambunan, I. Gunawan, and S. Sumarno, "Prediksi Jumlah Pendapatan Beasiswa PPA dan BBP Menggunakan Jaringan Syaraf Tiruan Backpropagation," *J. Media Inform. Budidarma*, vol. 3, no. 4, p. 346, 2019, doi: 10.30865/mib.v3i4.1327.
- [3] M. Muhandi, "Penentuan Penerima Beasiswa Dengan Algoritma Fuzzy C-Means," *J. Inf. dan Komput.*, vol. 7, no. 2, pp. 37–46, 2019, doi: 10.35959/jik.v7i2.148.
- [4] N. A. Manihuruk, M. Zarlis, E. Irawan, and H. S. Tambunan, "Penerapan Data Mining Dalam Mengelompokkan Calon Penerima Beasiswa Dengan Menggunakan Algoritma K-Means," *KOMIK (Konferensi Nas. Teknol. Inf. dan Komputer)*, vol. 4, no. 1, pp. 29–34, 2020, doi: 10.30865/komik.v4i1.2575.
- [5] A. E. Rahayu, K. Hikmah, N. Yustia, and A. C. Fauzan, "Penerapan K-Means Clustering Untuk Penentuan Klasterisasi Beasiswa Bidikmisi Mahasiswa," *Ilk. J. Comput. Sci. Appl. Informatics*, vol. 1, no. 2, pp. 82–86, 2019, doi: 10.28926/ilkomnika.v1i2.23.
- [6] P. Putriani and A. Mardiana, "Sistem Pendukung Keputusan Penerimaan Beasiswa Bank Indonesia Menggunakan Metode Logika Fuzzy Dan Saw (Studi Kasus Universitas Majalengka)," *Infotech J.*, vol. 8, no. 1, pp. 13–21, 2022.
- [7] F. S. Sagala, M. Mugiarto, and W. Priatna, "Algoritma K-Means Clustering Untuk Menentukan Rekomendasi Pemberian Beasiswa Bagi Siswa Berprestasi," *J. Students' Res. Comput. Sci.*, vol. 2, no. 2, pp. 111–120, 2021, doi: 10.31599/jsrsc.v2i2.840.
- [8] *Generasibaruindonesia.com*, "Program Beasiswa Bank Indonesia." <https://www.generasibaruindonesia.com/beasiswa> (accessed Nov. 04, 2022).
- [9] R. Widayanto, "Seleksi Beasiswa Bank Indonesia 2022." <http://baakk.unnes.ac.id/berita-1122-seleksi-beasiswa-bank-indonesia-2022.html> (accessed Nov. 05, 2022).
- [10] D. J. Lubis and M. B. Tamam, "Penerapan K-Means Untuk Pengelompokkan Beasiswa Santri di Pondok Pesantren Miftahul Huda Bogor," *Teknois J. Ilm. Teknol. Inf. dan Sains*, vol. 12, no. 1, pp. 7–20, 2022, doi: 10.36350/jbs.v12i1.125.
- [11] U. Ma'rifat, "Implementasi Algoritma K-Means untuk Pengelompokan Penyakit Pasien Pada Puskesmas Warujayeng," *Semin. Nas. Inov. Teknol.*, no. 2549–7952, pp. 285–291, 2020.
- [12] S. Aulia, "Klasterisasi Pola Penjualan Pestisida Menggunakan Metode K-Means Clustering (Studi Kasus Di Toko Juanda Tani Kecamatan Hutabayu Raja)," *Djtechno J. Teknol. Inf.*, vol. 1, no. 1, pp. 1–5, 2021, doi: 10.46576/djtechno.v1i1.964.
- [13] A. Salam, D. Adiatma, and J. Zeniarja, "Implementasi Algoritma K-Means Dalam Pengklasteran untuk Rekomendasi Penerima Beasiswa PPA di UDINUS," *JOINS (Journal Inf. Syst.)*, vol. 5, no. 1, pp. 62–68, 2020, doi: 10.33633/joins.v5i1.3350.
- [14] C. Kamila, M. A. A. S, G. R. Namang, and R. R. F. Syah, "Systematic Literature Review: Penggunaan Algoritma K-Means Untuk Clustering di Indonesia dalam Bidang Pendidikan," *J. Intech*, vol. 2, no. 1, pp. 19–24, 2021.
- [15] B. Harahap and A. Rambe, "Implementasi K-Means Clustering Terhadap Mahasiswa yang Menerima Beasiswa Yayasan Pendidikan Battuta di Universitas Battuta Tahun 2020/2021 Studi Kasus Prodi Informatika," *Informatika*, vol. 9, no. 3, pp. 90–97, 2021, doi: 10.36987/informatika.v9i3.2185.
- [16] S. A. Rahmah and J. Antares, "Klasterisasi Seleksi Mahasiswa Calon Penerima Beasiswa Yayasan Menggunakan K-Means Clustering," *INFORMATIKA*, vol. 13, no. 2, p. 25, 2022, doi: 10.36723/juri.v13i2.282.
- [17] D. N. Cahya, M. Hidayat, and M. F. Asnawi, "Implementasi Algoritma K-Means Clustering untuk Menentukan Calon Penerima Beasiswa UKT di Universitas Sains Al- Qur ' an," vol. 1, no. 1, pp. 26–32, 2022.
- [18] F. Farahdinna, I. Nurdiansyah, A. Suryani, and A. Wibowo, "Perbandingan Algoritma K-Means Dan K-Medoids Dalam Klasterisasi Produk Asuransi Perusahaan Nasional," *J. Ilm. FIFO*, vol. 11, no. 2, p. 208, 2019, doi: 10.22441/fifo.2019.v11i2.010.
- [19] A. Supriyadi, A. Triyudi, and I. D. Sholihati, "Perbandingan Algoritma K-Means Dengan K-Medoids Pada Pengelompokan Armada Kendaraan Truk Berdasarkan Produktivitas," *JIPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.)*, vol. 6, no. 2, pp. 229–240, 2021, doi: 10.29100/jipi.v6i2.2008.
- [20] S. B. F. Ginting, S. Sawaluddin, and M. Zarlis, "Kombinasi Pembobotan Symmetrical Uncertainty Pada K-Means Clustering Dalam Peningkatan Kinerja Pengelompokan Data," *J. Media Inform. Budidarma*, vol. 6, no. 1, p. 484, 2022, doi: 10.30865/mib.v6i1.3366.