PENGELOMPOKAN PRESTASI SISWA GUNA KUALIFIKASI BEASISWA BERDASARKAN DATA NILAI MENGGUNAKAN ALGORITMA *K-MEANS*

Syafina Haviyola, Susilawati, Mohamad Jajuli
Program Studi Informatika S1, Fakultas Ilmu Komputer, Universitas Singaperbangsa Karawang
Jl. HS. Ronggo Waluyo, Telukjambe Timur, Karawang, Indonesia
1910631170049@student.unsika.ac.id

ABSTRAK

Penelitian ini merupakan sebuah studi kasus yang dilakukan di SMKN 3 Karawang dengan tujuan untuk mengelompokkan prestasi siswa dengan memanfaatkan data nilai guna kualifikasi beasiswa, dan metode yang digunakan adalah algoritma *K-Means*. Implementasi algoritma *K-Means* menghasilkan tiga kelompok prestasi yang optimal (C=3). Hasil evaluasi dengan menggunakan metode *Davies Bouldien Index* menunjukkan bahwa model *cluster* dengan titik C=3 sudah optimal dengan nilai DBI 0.86. Dalam penelitian ini, terdapat 526 data siswa yang dianalisis, dimana 430 siswa tergolong dalam kelompok berprestasi rendah (*cluster* 0), 84 siswa berada dalam kelompok berprestasi menengah (*cluster* 1), dan 12 siswa berprestasi unggul (*cluster* 2). Penelitian ini memberikan kontribusi penting dalam penentuan kualifikasi beasiswa berdasarkan prestasi siswa di SMKN 3 Karawang, dan menggunakan metode *clustering* algoritma K-Means yang terbukti optimal dan efektif dalam analisis data nilai.

Kata kunci: *Algoritma K-Means, KDD, Data Mining, Davies Bouldien Index*

1. PENDAHULUAN

Pendidikan Indonesia selalu berkembang seiring dengan semakin tingginya kebutuhan untuk melanjutkan ke jenjang berikutnya. Kualitas sumber daya manusia harus ditingkatkan agar investasi ini menjadi salah satu langkah awal untuk dapat mewujudkan suatu keinginan. Pendidikan merupakan faktor penting dalam pencapaian tujuan yang dapat dicapai, sehingga menjadi penentu nasib masa depan suatu bangsa [1].

Sekolah membutuhkan biaya oprasional setiap bulannya, dana tersebut didapatkan dengan adanya Bantuan Operasional Sekolah yang diberikan oleh pemerintah tiap beberapa bulan tertentu dengan persyaratan yang perlu sekolah penuhi untuk mendapatkannya. Namun, pada pelaksanaannya dana bos sering kali mengalami keterlambatan pencairan sehingga pihak sekolah kesulitan dalam memenuhi kebutuhan operasional tiap bulannya sehingga sekolah menerapkan sistem Sumbangan Pembinaan Pendidikan (SPP) yang diberatkan kepada siswa setiap bulannya dengan nominal tertentu.

Hal ini tentunya memberatkan siswa yang memiliki kendala ekonomi contohnya saat harus melunasi SPP ketika hendak mengikuti ujian sekolah. Dengan demikian perlu diadakan bantuan pembayaran SPP khususnya untuk siswa yang memiliki kendala ekonomi dengan memberikan beasiswa kepada siswa yang berprestasi dan memiliki kendala perihal ekonomi. Beasiswa merupakan bantuan yang diberikan kepada individu dalam bentuk keuangan yang betujuan untuk digunakan demi kelangsungan pendidikan yang sedang dijalankan oleh penerima beasiswa pembagian beasiswa dilakukan oleh beberapa lembaga untuk membantu seseorang yang membutuhkan ataupun orang yang berprestasi selama menempuh studinya [2].

Perilaku dan prestasi peserta didik merupakan hal yang dinilai oleh suatu lembaga pendidikan seperti sekolah. Lembaga pendidikan menerapkan berbagai sistem untuk memotivasi dan mengembangkan potensi yang dimiliki para siswa, salah satunya adalah dengan pemilihan siswa berprestasi. Hal tersebut juga dilakukan oleh Sekolah Menengah Kejurusan (SMK) Negeri 3 Karawang.

Di sekolah ini menerapkan program untuk siswa kelas XI yang berprestasi dimana penilaian dilakukan berdasarkan nilai akademik, nilai kehadiran di kelas X serta beberapa aspek penilaian lainnya. Penghargaan yang diberikan berupa beasiswa dengan di bebaskanya kewajiban dalam pembayaran Sumbangan Pembinaan Pendidikan (SPP). Namun Saat ini pembagian beasiswa berupa pembebasan pembayaran SPP dirasa tidak merata karena parameter utamanya yang dapat mengajukan atau menerima yaitu yang memiliki kartu KIS/KIP/SKTM sedangkan tidak semua siswa yang kurang perihal ekonomi memliki kartu-kartu tersebut.

Berdasarkan latar belakang tersebut maka penulis melakukan penelitian dengan judul pengelompokan prestasi siswa guna kualifikasi beasiswa berdasarkan data nilai menggunakan algoritma k-means (studi kasus: SMKN 3 Karawang)". Penelitian ini menggunakan metode *Davies Bouldin Index* (DBI) untuk mengukur kebaikan atau kualitas hasil klastering yang telah dibentuk. Pada semua atribut hasil dapat di lihat dari perhitungan *K-Means* dalam mengelompokan dan menentukan siswa yang berprestasi mana yang layak diberikan beasiswa berprestasi.

2. TINJAUAN PUSTAKA

2.1. Sekolah Menengah Kejuruan

Sesuai dengan penjelasan pada *website* resmi *campus.quipper.com,* Sekolah Menengah Kejuruan

(SMK) adalah suatu bentuk satuan pendidikan formal dengan menyelenggarakan pendidikan menengah kejuruan dan mempersiapkan peserta didik untuk bekerja dalam bidang tertentu. Siswa dapat melanjutkan pendidikan kejuruan mereka setelah menyelesaikan SMP atau sederajat.

2.2. Prestasi Belajar

Prestasi belajar merupakan penguasaan pengetahuan atau keterampilan yang dikembangkan melalui mata pelajaran, biasanya ditunjukkan dengan nilai tes atau nilai yang diberikan oleh guru. Tingkat prestasi belajar tergantung pada usaha masing-masing peserta didik [3].

2.3. *Data Mining*

Data mining merupakan proses atau kegiatan mengumpulkan data dalam jumlah besar dan kemudian menggali suatu kumpulan data menjadi sebuah Informasi yang nantinya dapat digunakan [4]. *Data mining* dikenal juga sebagai *Knowledge Discovery in Database* (KDD) dalam proses pengumpulan informasi dari kumpulan data. Proses KDD interaktif dan berulang, termasuk beberapa Langkah yang melibatkan pengguna dalam pembuatan keputusan dan dapat dilakukan pengulangan dalam dua langkah [5].

2.4. *Clustering*

Clustering adalah sebuah proses pengelompokan barang sejenis ke dalam kelompok yang berbeda atau membagi *dataset* menjadi *subset*, sehingga data di setiap *subset* memiliki satu arti yang bermanfaat. *Clustering* sendiri disebut juga dengan unsupervised *classification* karena bertujuan untuk mempelajari dan memperhatikan [6].

Pengklasteran (*clustering*) merupakan pengelompokan data yang memiliki kemiripan nilai serta bentuk data yang dapat dikelompokkan dalam pengklasteran adalah hasil pengamatan, record data, atau kelas-kelas dan objek-objek yang memiliki kemiripan. Tujuan dari pengklasteran adalah mengelompokan beberapa data atau objek ke dalam klaster sehingga klaster tersebut berisi data yang mirip [7].

2.5. *K-Means*

Pengelompokan data merupakan metode eksplorasi dalam analisis data yang mengelompokkan objek dengan karakteristik serupa ke dalam satu kelompok, sehingga memungkinkan pengklasifikasian yang lebih terperinci dalam proses pengolahan data selanjutnya [8].

Algoritma *K-Means* merupakan metode analisis tanpa tingkatan yang mengelompokan data berdasarkan karakteristiknya ke dalam satu atau lebih *cluster*. Dan data yang memiliki memiliki kesamaan dikelompokan dalam satu *cluster* yang sama dan sebaliknya jika data tidak sama dikelompokan kedalam cluster yang tidak sama juga [8].

2.6. *Knowledge Discovery in Database*

Knowledge Discovery in Database adalah proses meneliti dan menganalisis banyak informasi dan mengekstraksinya hasil pengetahuan yang bermanfaat. Kemudian Hasil informasi yang didapat dapat digunakan sebagai database untuk tujuan pengambilan keputusan [9].

2.7. Elbow

Pengelompokan *K-Means* meminimalkan jumlah kesalahan kuadrat antara setiap titik massa *cluster* dan titik sampel dalam *cluster* sebagai tingkat distorsi. jika distorsi dalam sebuah *cluster* lebih rendah, koneksi antar anggota internalnya lebih dekat. Sebaliknya semakin tinggi distorsi semakin longgar struktur internalnya. Derajat distorsi menurun, ketika jumlah cluster meningkat, dan kecepatan Penurunannya lebih lambat, poin ini sering dipertimbangkan sesuai dengan kinerja kelompok baik dan karena desainnya menyerupai siku, maka disebut metode siku [10].

2.8. *Silhouette Coefficient*

Dalam proses validasi klaster, terdapat metrik internal yang dikenal sebagai koefisien siluet, yang mempertimbangkan jarak di dalam klaster dan antar klaster. Kualitas pengelompokan dapat dinilai melalui rata-rata siluet, dimana peningkatan nilai indeks ini dapat dijadikan panduan untuk mencari jumlah klaster yang optimal [10]. Jarak antara data dihitung dengan menggunakan metode jarak *Euclidean*. Untuk memberikan pemahaman mengenai kualitas pengelompokan pada proses clustering, kita dapat menghitung nilai siluet dari setiap klaster dan juga nilai siluet untuk keseluruhan klaster yang dihasilkan oleh algoritma *clustering*.

2.9. *Sum of Square Error*

Sum of Square Error (SSE) adalah rumus yang diterapkan untuk mengukur perbedaan antara data yang didapat dengan model perkiraan yang telah dilakukan sebelumnya [11]. dinyatakan baik jika cluster memiliki nilai SSE yang paling kecil dibanding dengan model lainnya.

2.10. *Davies Bouldin Index*

DBI adalah salah satu metode evaluasi internal yang evaluasinya diukur pada suatu metode pengelompokan berdasarkan nilai-nilai kohesi dan separasi. didalam suatu pengelompokan, kohesi didefinisikan sebagai jumlah kedekatan data terhadap *centroid cluster* yang diikuti Meskipun separasi didasarkan pada jarak antar *centroid cluster*. Jika jarak dalam *cluster* minimal signifikan Setiap objek dalam *cluster* memiliki level kemiripan fitur yang tinggi. Semakin rendah nilainya DBI diperoleh (non-negatif >= 0), semakin baik *cluster* diperoleh dengan cara clustering algoritma pengelompokan [10].

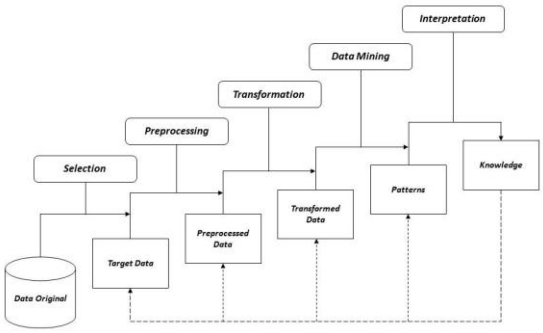
2.11. *Google Collabolatory*

Google colab, juga dikenal sebagai *google collaboratory*, adalah alat atau *editor* kode yang memanfaatkan sistem penyimpanan awan secara gratis untuk keperluan riset. *Platform google colab* dirancang dengan mengambil elemen dari Jupyter, dan hampir semua pustaka masukan tersedia untuk digunakan dalam kegiatan riset, termasuk eksplorasi data menggunakan *data mining*. Dalam hal kesamaan fungsi dan tujuan, *Google colab* memiliki kemiripan dengan *Jupyter Notebook*, karena keduanya memiliki tujuan dan fungsi yang serupa. Namun, perbedaannya terletak pada batasan sistem, dengan *Google colab* dapat diakses secara gratis dan dalam mode online [12].

3. METODE PENELITIAN

Metodologi penelitian merupakan proses memeriksa dan menganalisis data dan mengekstraksi informasi yang berguna. hasil informasi yang di peroleh dapat digunakan sebagai dasar pengambilan pengambilan keputusan [9].

Metodologi yang diterapkan pada penelitian kali ini yaitu *Knowledge Discovery in Databases* (KDD) yang yang telah terstruktur dalam 5 tahapan yaitu *data selection* (penyeleksian data), *preprocessing* atau *cleaning* (pembersihan data), *transformation* (penyesuaian data), *data mining* (pencarian pola atau informasi), dan *interpretation* atau *evaluation* (penafsiran data).

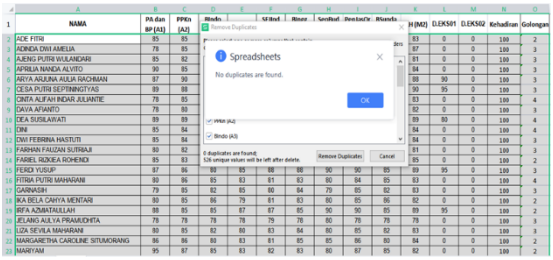


Gambar 1. Rancangan penelitian

4. HASIL DAN PEMBAHASAN

Pada tahap *data selection* dimulai dengan memilih informasi yang paling penting dari tabel yang diberikan, memilih atribut-atribut yang benar-benar berpengaruh dan berguna untuk menjawab pertanyaan atau tujuan analisis. Data yang didapat diberikan dalam bentuk tabel dengan jumlah tupel 527 dengan 20 atribut. Pada tahap awal akan dilakukan seleksi data untuk memilih atribut yang akan berpengaruh terhadap penentuan label data dari data nilai siswa untuk beasiswa. Setelah melakukan *interview* terkait data penerimaan beasiswa yang didapat, ada beberapa atribut yang dijadikan acuan dan perhitungan sebagai penilaian kelolosan pengajuan beasiswa. Sebanyak 13 atribut sebagai perhitungan kelolosan dan 7 atribut sebagai kebutuhan database. Sedangkan atribut nama lengkap dan jurusan tetap dipertahankan untuk keperluan tahap *preprocessing*.

Selanjutnya pada tahap *preprocessing* dilakukan pengecekan *missing value*. Pada *dataset* hasil *data selection* masih terdapat *fields* dengan data yang kosong dan *strip* (-) pada atribut prestasi yang berarti "tidak ada atau nol". Data-data tersebut perlu diatasi untuk meningkatkaan kualitas data. Untuk mengatasi masalah tersebut dengan mengganti niai yang kosong dan tanda *strip* dengan angka 0. Setelah pengecekan *missing value* dilakukan pula pengecekan duplikasi data menggunakan *excel* dan *google collab*.



Gambar 2. Hasil pengecekan *excel*

Gambar 3. Hasil pengecekan *google collab*

Dalam analisis data yang dilakukan, tidak ditemukan adanya entri ganda atau data yang memiliki kesamaan dengan entri lainnya.

Kemudian data akan melewati tahap transformasi. Pada tahap transformasi data dilakukan perubahan pada data kompleks menjadi lebih mudah di olah seperti melakukan normalisasi data. Proses ini akan memastikan bahwa rentang nilai dari setiap variabel menjadi seragam, sehingga perbedaan skala tidak akan mempengaruhi hasil analisis. Transformasi data membantu dalam merubah informasi menjadi bentuk yang lebih berguna atau lebih cocok untuk analisis atau presentasi. Setelah melakukan analisis terhadap data yang tersedia, tidak ada data yang memerlukan transformasi tambahan. Data yang telah diberikan telah diverifikasi dan terbukti akurat serta sesuai dengan kebutuhan analisis. Maka dari itu dapat dilanjutkan kepada tahap berikutnya dalam proses evaluasi tanpa perlu melakukan perubahan lebih lanjut pada data yang telah diberikan.

Setelah dilakukan transformasi data menggunakan penskalaan. Tahapan selanjutnya yaitu proses *data mining* untuk melihat pola yang terdapat pada data. Teknik yang diterapkan pada proses ini adalah teknik *clustering* yaitu melakukan pengelompokan data berdasarkan tingkat kemiripan dalam *cluster*. Algoritma yang digunakan yaitu algoritma *K-Means* dengan optimalisasi penentuan jumlah *cluster* menggunakan metode *elbow* dan *silhouette* yang digunakan sebagai pertimbangan

penentujuan jumlah *cluster* dan didasarkan pada hasil perhitungan *Sum of Square.*

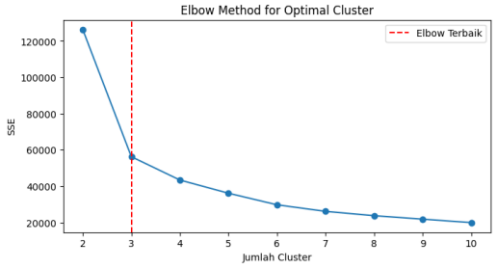
Tabel 1. Hasil SSE

Cluster	SSE
1	1650.0
2	1025.0
3	649.5543293718167

Nilai SSE untuk Cluster 1 adalah sebesar 1650.0. Ini mengindikasikan bahwa penyebaran data di dalam *cluster* 1 relatif lebih besar, yang dapat menunjukkan tingkat variasi yang signifikan antara anggota kluster ini. *Cluster* 2 memiliki nilai SSE sebesar 1025.0. Nilai SSE yang lebih rendah ini menandakan bahwa data di dalam *cluster* 2 memiliki penyebaran yang lebih padat dan homogen dibandingkan dengan *cluster* 1. Nilai SSE *cluster* 3 adalah 649.5543293718167. *Cluster* ini memiliki nilai SSE terendah di antara ketiga kluster, yang menunjukkan bahwa data di dalamnya sangat padat dan homogen, dengan anggota kluster yang cenderung mendekati pusat kluster. Nilai tersebut kemudian divisualisasikan menjadi grafik *elbow.*

```
# Plot Elbow Method (SSE)
plt.figure(figsize=(8, 4))
plt.plot(jumlah_cluster, sse_scores, marker='o')
plt.xlabel('Jumlah Cluster')
plt.ylabel('SSE')
plt.title('Elbow Method for Optimal Cluster')

# Menemukan indeks elbow terbaik
elbow_index = 0
for i in range(1, len(sse_scores) - 1):
    elbow_diff = sse_scores[i] - sse_scores[i + 1]
    if elbow_diff < 0.1 * sse_scores[0]:  # Atur ambang sesuai kebutuhan
        elbow_index = i
        break
```

Gambar 4. Proses *elbow*

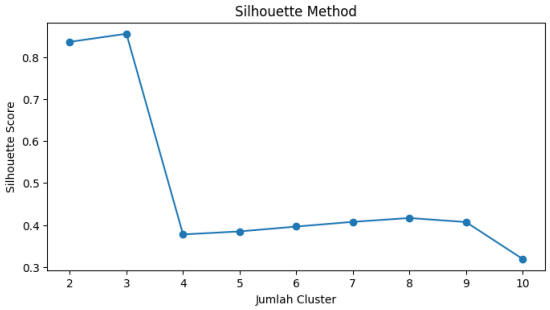


Gambar 5. Grafik *elbow*

Berdasarkan perhitungan nilai SSE menunjukan bahwa titik yang memiliki nilai paling rendah serta merupakan titik yang sudah stabil dan tidak mengalami banyak perubahan pada *cluster* atau titik berikutnya terdapat pada C=3. Sedangkan pada grafik *elbow* juga ditujukan bahwa titik yang membentuk siku sebelum menjadi garis lurus dan menjadi stabil terdapat pada C=3. Maka dari itu berdasarkan metode *elbow* baik pada data dengan penskalaan memiliki cluster optimal sejumlah 3 cluster. Sedangkan pada grafik *silhouette coefficient* yaitu:

Tabel 2. Hasil SSE

Cluster	***Indeks Silhouette***
2	0.8361193108696747
3	0.8553663402794811
4	0.377565433678523
5	0.38469542845931826



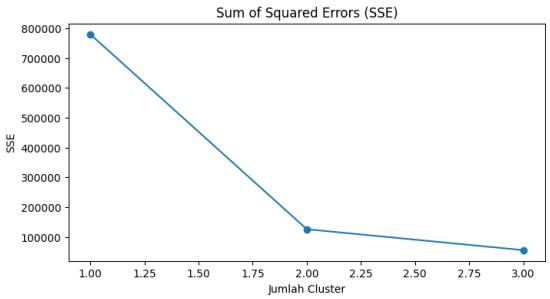
Gambar 5. Grafik *silhouette coefficient*

Koefisien *Silhouette* memiliki skala nilai dari -1 hingga 1. Cluster dianggap berkualitas dengan struktur yang kuat ketika mendekati nilai 1. Oleh karena itu, berdasarkan grafik koefisien Silhouette pada Gambar 4.6 menunjukan titik optimal pada C=3 dengan *silhouette score* sebesar 0.8553663402794811 dan masuk ke dalam kategori struktur kuat.

Karena berdasarkan grafik *elbow* dan *silhouette coefficient* berada pada titik optimal C=3 maka pada tahapan selanjutnya yaitu *data mining* akan dibagi ke dalam 3 *cluster* dengan grafik SSE.

```
# Melakukan clustering dengan jumlah cluster yang berbeda
for k in jumlah_cluster:
    kmeans = KMeans(n_clusters=k)
    kmeans.fit(data)

    # Mendapatkan nilai SSE untuk jumlah cluster saat ini
    sse_current = kmeans.inertia_

    sse_scores.append(sse_current)

    # Cetak nilai SSE untuk setiap jumlah cluster
    print(f"Nilai SSE untuk {k} cluster: {sse_current}")

# Jumlah cluster dengan SSE terendah
best_cluster_sse = jumlah_cluster[sse_scores.index(min(sse_scores))]
print(f"Jumlah cluster terbaik berdasarkan SSE: {best_cluster_sse}")
```

Gambar 6. Proses mendapatkan nilai SSE

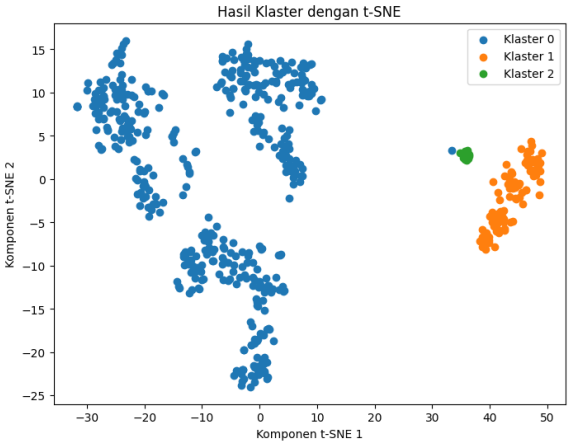


Gambar 7. Grafik SSE

```
Nilai SSE untuk 1 cluster: 1650.0
Nilai SSE untuk 2 cluster: 1025.0
Nilai SSE untuk 3 cluster: 649.5543293718167
Jumlah cluster terbaik berdasarkan SSE: 3
```

Gambar 8. Nilai SSE

Hasil *clustering* kemudian divisualisasikan menggunakan *scatter plot* dan dilakukan analisis karakteristik *cluster* sebagai acuan penentuan label. Terdapat tiga klaster yang telah teridentifikasi, masing-masing diberi label dengan kategori prestasi. Klaster 0 memiliki jumlah 430 data dan dikategorikan sebagai kelompok yang berprestasi rendah. Klaster 1 terdiri dari 84 data dan dikategorikan sebagai kelompok berprestasi menengah. Sementara itu, klaster 2 hanya memiliki 12 data dan diidentifikasi sebagai kelompok berprestasi unggul.



Gambar 9. *Scatter plot* prestasi siswa

Tabel 3. Label *cluster*

Cluster	Label
0	Berprestasi Rendah
1	Berprestasi Menengah
2	Berprestasi Unggul

Sesuai dengan yang dihasilkan oleh *scatter plot* untuk siswa yang berprestasi rendah ditandai oleh warna biru sebanyak 430 siswa, siswa yang berprestasi menengah ditandai oleh warna orange sebanyak 84 siswa, siswa yang berprestasi unggul ditandai oleh warna hijau sebanyak 12 siswa.

```
1
# Menampilkan total isi cluster
cluster_counts = data['Cluster'].value_counts()
print("Total isi setiap cluster:")
print(cluster_counts)
```

```
Total isi setiap cluster:
0    430
1     84
2     12
Name: Cluster, dtype: int64
```

Gambar 10. Hasil karakteristik *google collab*

Setelah proses *data mining*, langkah selanjutnya yaitu proses evaluasi menggunakan *Davies Bouldin Index* (DBI) untuk mengukur hasil atau kualitas dari cluster dan model yang dibentuk.

Tabel 4. Hasil *index clustering*

Jumlah *Cluster*	DBI
2	1.1786205411170696
3	0.8624322253305081
4	0.8497457050956073
5	0.853208529619683

Pada *Davies Bouldin Index* (DBI), pengelompokan data dinyatakan optimal atau semakin baik jika memiliki *index* yang lebih rendah atau mendekati angka 0. Berdasarkan tabel, *index* yang lebih rendah ada pada C=3. Maka titik optimal *cluster* berdasarkan nilai DBI berada di C=3

5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian dapat diperoleh kesimpulan bahwa penerapan *clustering* menggunakan algoritma *k-means* pada data nilai siswa tahun 2021 dengan *tools Google Colab* menghasilkan jumlah *cluster* optimal yaitu C=3. Pada 526 data siswa yang di dapatkan sebanyak 430 siswa berprestasi rendah pada *cluster* 0, sebanyak 84 siswa berprestasi menengah pada *cluster* 1 dan sebanyak 12 siswa berprestasi unggul pada *cluster* 2. Hasil evaluasi *cluster* dengan metode *silhouette* dengan nilai 0.85, *elbow* dan SSE 649.554 dan *Davies Bouldien Index* (DBI) memiliki nilai DBI 0.86, adapun saran untuk penngembangan penelitian yaitu diharapkan menggunakan *dataset* yang lebih besar dan algoritma serta metode evaluasi yang berbeda.

DAFTAR PUSTAKA

[1] Amadea, K. dan Ayuningtyas, M. 2020. Perbandingan Efektivitas Pembelajaran Sinkronus dan Asinkronus pada Materi Program Linear. Jurnal PRIMATIKA. 9 (2): 111-120.

[2] Sembiring, S. S., Nasyuha, A. H., & Mariami, I. (2020). Pemanfaatan Algoritma K-means Dalam Pengelompokan Data Siswa Berdasarkan Prestasi Akademik Sebagai Penentuan Penerima Beasiswa di SMK Negeri 1 Lubuk Pakam. Jurnal Cyber Tech, 3(9).

[3] Darmuki, A., & Hariyadi, A. (2019). Peningkatan Keterampilan Berbicara Menggunakan Metode Kooperatif Tipe Jigsaw Pada Mahasiswa PBSI Tingkat I-B IKIP PGRI Bojonegoro Tahun Akademik 2018/2019. Jurnal Kredo. 2(2): halaman 256-267.

[4] Saputra, E. A., & Nataliani, Y. (2021). Analisis Pengelompokan Data Nilai Siswa untuk Menentukan Siswa Berprestasi Menggunakan Metode Clustering K-Means. Journal of Information Systems and Informatics, 3(3), 424-439.

[5] Yunita, D., Rosyani, P., & Amalia, R. (2018). Analisa Prestasi Siswa Berdasarkan Kedisiplinan, Nilai Hasil Belajar, Sosial Ekonomi dan Aktivitas Organisasi Menggunakan Algoritma Naïve Bayes. J. Inform. Univ. Pamulang, 3(4), 209..

[6] Zyen, A. K., Wibowo, G. W. N., Widiastuti, N. A., & Wahono, B. B. (2023). Klasterisasi Penerima Bantuan Fasilitas Sekolah di MI Datuk Singaraja Menggunakan Metode K-Means. JTINFO: Jurnal Teknik Informatika, 2(1), 25-35.

[7] Kurniawan, S., Siregar, A. M. and Novita, H. Y. (2023) 'Penerapan Algoritma K-Means dan Fuzzy C-Means Dalam Mengelompokan Prestasi Siswa Berdasarkan Nilai Akademik', Scientific Student Journal for Information, Technology and Science, IV(1), pp. 73–81.

[8] Primanda, R. P., Alwi, A., & Mustikasari, D. (2021). Data Mining Seleksi Siswa Berprestasi Untuk Menentukan Kelas Unggulan Menggunakan Metode K-Means Clustering (Studi Kasus di MTS Darul Fikri). Komputek, 5(1), 88-100

[9] Qisthiano, M. R., Kurniawan, T. B., Negara, E. S., & Akbar, M. (2021). Pengembangan Model Untuk Prediksi Tingkat Kelulusan Mahasiswa Tepat Waktu dengan Metode Naïve Bayes. JURNAL MEDIA INFORMATIKA BUDIDARMA, 5(3), 987-994.

[10] Nuraeni, F., Kurniadi, D., & Dermawan, G. F. (2023). Pemetaan Karakteristik Mahasiswa Penerima Kartu Indonesia Pintar Kuliah (KIP-K) menggunakan Algoritma K-Means++. Jurnal Sisfokom (Sistem Informasi Dan Komputer), 11(3), 437-443.

[11] Winarta, A., & Kurniawan, W. J. (2021). Optimasi Cluster K-Means Menggunakan Metode Elbow Pada Data Pengguna Narkoba Dengan Pemrograman Python. Jurnal Teknik Informatika Kaputama (JTIK), 5, 113–119.

[12] M. Khandava Mulyadien, A. Kurniawan, D. R. Utami, and U. Enri, "Rancang Bangun Sistem Penjualan Sate Taichan Berbasis Web," J. Sains Komput. Inform. (J-SAKTI, vol. 6, no. 1, pp. 101–114, 2022.

.