

Penerapan Algoritma *K-Means* dan *Fuzzy C-Means* Dalam Mengelompokan Prestasi Siswa Berdasarkan Nilai Akademik

Salim Kurniawan
Universitas Buana Perjuangan
Karawang, Indonesia

if18.salimkurniawan@mhs.ubpkarawang.ac.id

Amril Mutoi Siregar
Universitas Buana Perjuangan
Karawang, Indonesia

amrilmutoi@ubpkarawang.ac.id

Hilda Yulia Novita
Universitas Buana Perjuangan
Karawang, Indonesia

hilda.yulia@ubpkarawang.ac.id

Abstract— SMK PGRI 2 Karawang setiap tahun akan memberikan beasiswa pendidikan dan merekomendasikan siswa-siswinya kepada perusahaan berdasarkan prestasi dalam capaian Ujian Nasional namun mengalami kesulitan dalam pengelompokan siswa yang berprestasi setelah Ujian Nasional ditiadakan dari tahun 2020 karena covid-19 untuk itu tujuan penelitian ini yaitu mengelompokan prestasi siswa berdasarkan nilai akademik menggunakan data mining dengan menerapkan algoritma *K-Means* dan *Fuzzy C-Means* serta membandingkan hasilnya untuk mengetahui algoritma terbaik, penelitian ini bertempat di SMK PGRI 2 Karawang dan data yang digunakan merupakan nilai raport siswa dari semester 1 sampai 6. Hasil perhitungan manual dan program *python* algoritma *K-Means* sebanyak 52 siswa berprestasi, 25 siswa dengan prestasi sedang dan 28 siswa tidak berprestasi sedangkan pada algoritma *Fuzzy C-Means* sebanyak 48 siswa berprestasi, 29 siswa dengan prestasi sedang, dan 28 siswa tidak berprestasi. Pada perhitungan *rapidminer studio* hanya bisa menggunakan algoritma *K-Means* saja karena *Fuzzy C-Means* tidak tersedia dan berhasil mendapatkan 49 siswa berprestasi, 27 siswa dengan prestasi sedang dan 29 siswa tidak berprestasi serta berdasarkan nilai evaluasi dengan metode *Davies Bouldien Index (DBI)* dari setiap jenis perhitungan menunjukkan algoritma *K-Means* merupakan algoritma terbaik karena memiliki nilai *DBI* yang mendekati 0 dan tidak negatif.

Kata kunci — *data mining, davies bouldien index, fuzzy c-means, k-means, pengelompokan, python*

I. PENDAHULUAN

Proses belajar mengajar adalah sebuah proses interaksi antara siswa dengan guru atau siswa dengan lingkungannya, yang mengarah pada perubahan perilaku, sehingga memberikan pengalaman baik pengetahuan, sikap dan keterampilan [1]. Pendidikan menurut [2] dalam pasal 1 ayat 1 merupakan upaya menciptakan suasana belajar dan proses pembelajaran yang memungkinkan peserta didik secara aktif mengembangkan potensi dirinya dan secara sadar menumbuhkembangkan kekuatan spiritual keagamaan, kepribadian, akhlak mulia, kecerdasan, pengendalian diri serta keterampilan yang diperlukan dirinya, masyarakat, bangsa dan negara. Penilaian Hasil Belajar seperti yang termaktub dalam [3] adalah proses pengumpulan dan pengolahan informasi untuk menggambarkan hasil belajar siswa yang salah satu tujuannya untuk mengetahui tingkat capaian hasil belajar. Melalui proses tersebut dapat diukur capaian kemampuan, pengetahuan dan pemahaman terhadap materi pembelajaran di sekolah yang disebut dengan prestasi belajar [1].

Sekolah Menengah Kejuruan (SMK) PGRI 2 Karawang yang berpusat di Jalan Pangkal Perjuangan/By Pass Paracis Kelurahan Tanjung Pura Kecamatan Karawang Barat Kabupaten Karawang merupakan lembaga pendidikan formal di bawah naungan Yayasan Pembina Lembaga Pendidikan (YPLP) PGRI. Sebagai salah satu lembaga penyedia pendidikan tentunya SMK PGRI 2 Karawang setiap tahunnya akan meluluskan siswa-siswinya sesuai dengan peraturan perundang-undangan. Sebelum meluluskan siswa-siswinya, SMK PGRI 2 Karawang akan mengelompokan prestasi siswanya berdasarkan nilai dalam proses belajar. Hal itu dilakukan untuk menjadi bahan acuan dalam memberikan penghargaan berupa beasiswa serta merekomendasikan siswa-siswi tersebut dalam melanjutkan pendidikan ke jenjang lebih tinggi dan atau merekomendasikannya ke perusahaan agar menjadi bahan pertimbangan dalam proses perekrutan. Saat ini Wakil Kepala Sekolah Bidang Kesiswaan belum dapat melakukan hal tersebut dikarenakan biasanya pemberian beasiswa pendidikan dan rekomendasi rekrutmen kerja diambil dari nilai mata pelajaran yang di Ujian Nasional akan tetapi pada tahun 2020 Ujian Nasional ditiadakan oleh Pemerintah Republik Indonesia melalui [4] dikarenakan situasi pandemi covid-19 yang terus meningkat.

Berdasarkan permasalahan tersebut, maka dalam penelitian ini akan menggunakan teknik data mining dalam mengelompokan prestasi siswa dengan metode *clustering*. Penelitian mengenai metode *clustering* telah ada sebelumnya yakni terdapat dalam jurnal penelitian [5] dengan tingkat keberhasilan sebesar 88% dan pada penelitian [6] yang mengkomparasi dua algoritma yakni *K-Means* dan *Fuzzy C-Means* menyebutkan bahwa algoritma *Fuzzy C-Means* memiliki performa yang lebih baik dibandingkan algoritma *K-Means* serta pada penelitian [7] menghasilkan algoritma terbaik yaitu algoritma *Fuzzy C-Means*.

II. DATA DAN METODE

A. Prestasi Belajar

Prestasi belajar dapat dikatakan sebagai seluruh keterampilan atau perubahan perilaku yang diperoleh melalui proses belajar berdasarkan tes yang diselenggarakan serta hasilnya dinyatakan dalam bentuk nilai ulangan harian, ujian, atau rapor [8]. Tingkat prestasi belajar tergantung pada usaha masing-masing peserta didik [8]. Faktor-faktor yang mempengaruhi kegiatan belajar dan prestasi belajar dalam prestasi dan motivasi belajar dengan konsep diri adalah sebagai berikut :

- a. Faktor internal peserta didik yakni kematangan atau pertumbuhan, kecerdasan, kepribadian dan motivasi.
- b. Faktor eksternal atau faktor sosial yakni faktor keluarga dengan keadaan rumah tangga, guru dengan cara mengajarnya, lingkungan, perangkat pembelajaran serta motivasi sosial

B. Data Mining

Data mining atau kadang juga dikenal dengan *Knowledge Discovery in Database* (KDD) adalah suatu kegiatan yang berkaitan dengan pengumpulan data, penggunaan data historis untuk menemukan pengetahuan, informasi, keteraturan, pola atau hubungan dalam data yang besar. Hasil dalam data mining dapat digunakan sebagai alternatif dalam pengambilan keputusan atau untuk meningkatkan pengambilan keputusan di masa yang akan datang. Terdapat enam fungsi dalam data mining, yaitu sebagai berikut deskripsi (*description*), estimasi (*estimation*), prediksi (*prediction*), klasifikasi (*classification*), pengelompokan (*clustering*) dan asosiasi (*association*).

C. Algoritma K-Means

Algoritma *K-Means Clustering* merupakan algoritma yang sering digunakan karena penggunaannya yang sederhana dan mudah untuk diimplementasikan [9]. Pada algoritma *K-Means Clustering*, data dibagi menjadi beberapa kelompok, dimana data pada satu kelompok memiliki karakteristik yang sama, dan data pada kelompok lainnya memiliki karakteristik yang berbeda. Langkah-langkah dalam algoritma *K-Means Clustering* [10] sebagai berikut :

1. Memasukan data
2. Tentukan jumlah *cluster*
3. Ambil sembarang data sebanyak *cluster* secara acak sebagai pusat *cluster* (*centroid*)
4. Hitung jarak antara data dengan *centroid*, dengan menggunakan persamaan *Euclidean distance* :

$$D_{ij} = \sqrt{(x_{1j} - v_{1j})^2 + \dots + (x_{ij} - v_{kj})^2} \quad (1)$$

Dimana:

D_{ij} = jarak data i ke pusat cluster ke k

x_{ij} = data ke i pada atribut ke j

v_{kj} = titik pusat ke k pada atribut ke j

5. Hitung kembali pusat cluster (*centroid*) dengan keanggotaan *cluster* yang baru, berikut rumus untuk menghitung *centroid* :

$$C_1 = \frac{1}{M} \sum_{j=1}^M x_j \quad (2)$$

6. Jika pusat cluster tidak berubah maka proses cluster telah selesai, jika belum maka ulangi langkah ke 4 sampai pusat cluster tidak berubah lagi

D. Algoritma Fuzzy C-Means

Algoritma *Fuzzy C-Means* (FCM) merupakan algoritma pengelompokan data dimana ditentukan oleh derajat keanggotaan [10]. Algoritma FCM menyangkut konsep kesamaan fungsional objek yang berdekatan dan mendefinisikan titik pusat *cluster* sebagai prototipe. Untuk beberapa objek data, tidak ada batasan untuk satu kelas, tetapi data dapat dikelompokkan berdasarkan tingkat keanggotaan, yang berkisar dari 0 hingga 1, yang menunjukkan keanggotaan sebagian data [9]. Langkah-langkah dalam algoritma FCM [10] sebagai berikut :

1. Memasukan data yang akan dikelompokkan X, berupa matriks berukuran $n \times m$
2. Tentukan :
 - a. Jumlah *cluster* = c
 - b. Pangkat = w
 - c. Maksimum iterasi = Max Iter
 - d. Error terkecil yang diharapkan = ϵ
 - e. Fungsi objektif awal: $P_0 = 0$

f. Iterasi awal: $t = 1$;

3. Bangkitkan bilangan acak (μ_{ik}), dengan $i = 1, 2, \dots, n$; $k = 1, 2, \dots, n$; dan c sebagai elemen-elemen matriks partisi awal U.

Hitung jumlah setiap kolom :

$$Q_i = \sum_k^c \mu_{ik} \quad \text{dengan } j = 1, 2, \dots, n \quad (3)$$

Hitung :

$$\mu_{ik} = \frac{\mu_{ik}}{Q_i} \quad (4)$$

4. Hitung pusat *cluster* ke-k : V_{kj}

$$V_{kj} = \frac{\sum_{i=1}^n (\mu_{ik})^w \times x_{ij}}{\sum_{i=1}^n (\mu_{ik})^w} \quad (5)$$

5. Hitung fungsi objektif pada iterasi ke-t

$$P_t = \sum_{t=1}^n \sum_{k=1}^c ([\sum_{j=1}^m (X_{ij} - V_{kj})^2] (\mu_{ik})^w) \quad (6)$$

6. Hitung perubahan matriks partisi

$$\mu_{ik} = \frac{[\sum_{j=1}^m (X_{ij} - V_{kj})^2]^{-\frac{1}{w-1}}}{\sum_{k=1}^c [\sum_{j=1}^m (X_{ij} - V_{kj})^2]^{-\frac{1}{w-1}}} \quad (7)$$

7. Cek kondisi berhenti :

- Jika $(|P_t - P_{t-1}| \leq \text{MaxIter})$ maka berhenti.
- Jika tidak $: t = t + 1$, ulangi langkah ke 4. Output yang dihasilkan dari FCM merupakan deretan pusat *cluster* dan beberapa derajat keanggotaan untuk tiap-tiap titik data.

E. Evaluasi Cluster

Evaluasi *cluster* ditemukan oleh para peneliti untuk menentukan jumlah *cluster* yang terbaik atau optimal [11]. *Davies Bouldin Index* (DBI) merupakan salah satu teknik evaluasi *cluster* yang diperlukan untuk menyempurnakan kinerja sebuah metode atau algoritma optimasi demi mendapatkan jumlah *cluster* yang terbaik [12].

Sum of Square Within Cluster (SSW) merupakan persamaan yang digunakan untuk mengetahui matrik kohesi dalam sebuah *cluster* ke-i yang dirumuskan sebagai berikut (Zulfa Nabila, et al., 2021) :

$$SSW_i = \frac{1}{m_i} \sum_{j=1}^{m_i} d(x_i, c_i) \quad (8)$$

Keterangan :

m = jumlah data dalam *cluster* ke-i

c = *centroid cluster* ke-i

d = jarak

Untuk menghitung separasi antar *cluster* dapat menggunakan *Sum of Square Between Cluster* (SSB) dengan penjabaran sebagai berikut (Zulfa Nabila, et al., 2021) :

$$SSB_{ij} = d(c_i, c_j) \quad (9)$$

Keterangan :

d = jarak

c = *centroid cluster* ke-i

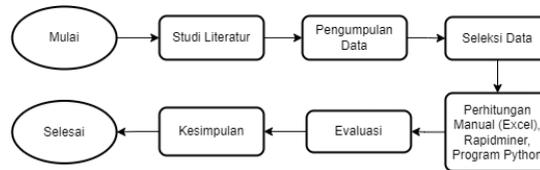
Setelah memperoleh nilai kohesi dan separasi langkah selanjutnya melakukan pengukuran rasio (R_{ij}) untuk mengetahui nilai perbandingan antar *cluster* ke-i dan ke-j. *Cluster* yang memiliki nilai kohesi terkecil dan separasi yang terbesar adalah *cluster* yang baik. Berikut persamaan dalam menghitung nilai rasio (Zulfa Nabila, et al., 2021) :

$$DBI = \frac{1}{k} \sum_{i=1}^k \max_{i \neq j} (R_{ij}) \tag{10}$$

k merupakan jumlah *cluster* yang digunakan. Semakin kecil nilai DBI yang diperoleh (non-negatif ≥ 0), maka semakin baik *cluster* yang diperoleh

F. Prosedur Penelitian

Prosedur penelitian dimulai dengan studi literatur dengan rincian penelitian sebagai berikut:



Gambar 1. Prosedur Penelitian

III. HASIL DAN PEMBAHASAN

A. Pengumpulan Data

Data yang diperoleh dari SMK PGRI 2 Karawang merupakan nilai rapor siswa dari semester 1 sampai dengan semester 6 jurusan Teknik Komputer dan Jaringan tahun pelajaran 2020-2021 berjumlah 105 siswa yang dapat dilihat pada tabel 1.

Tabel 1. Nilai Rapor Siswa

No	NAMA	NIS	KELAS	Semester I			...	Semester VI		
				1. PAIBP			...	10. PSPJ		
				P (40 %)	K (60 %)	NA	...	P (40 %)	K (60 %)	NA
1	Adtya Agustin	181910148	XII TKJ 1	80	83	82	...	80	80	80
2	Agung F. Sukandar	181910149	XII TKJ 1	80	85	83	...	80	80	80
3	Akmal Muflih Hakim	181910150	XII TKJ 1	83	88	86	...	84	84	84
4	Almuriah	181910151	XII TKJ 1	82	87	85	...	83	83	83
5	Andreas Kayis Akmal	181910152	XII TKJ 1	85	92	89	...	84	84	84
6	Asri Kurniyasih	181910153	XII TKJ 1	85	87	86	...	83	83	83
7	Astri Berliana	181910154	XII TKJ 1	82	87	85	...	83	83	83
8	Cahyana Adinata	181910155	XII TKJ 1	82	87	85	...	80	80	80
9	Cantika	181910156	XII TKJ 1	80	85	83	...	83	83	83
10	Cucu Kuraesin	181910157	XII TKJ 1	82	87	85	...	83	83	83
...
...
104	Sutrisna	181910260	XII TKJ 3	77	77	77	...	80	80	80
105	Sutrisno Purnama	181910261	XII TKJ 3	77	77	77	...	80	80	80

B. Seleksi Data

Tahap ini adalah hasil dari proses penyeleksian data dari atribut yang tidak diperlukan. Pada tabel 1 dapat dilihat bahwa setiap mata pelajaran mempunyai tiga komponen yaitu nilai Pengetahuan (P) sebesar 40%, nilai Keterampilan (K) sebesar 60% dan Nilai Akhir (NA) yang merupakan gabungan dari bobot nilai Pengetahuan dan Keterampilan. Namun pada penelitian ini hanya akan menggunakan Nilai Akhir (NA) pada setiap mata pelajaran karena merupakan gabungan dari kedua komponen nilai.

Setelah didapatkan Nilai Akhir dari setiap mata pelajaran data nilai rapor masih harus diseleksi kembali karena data yang didapat merupakan nilai rapor dari semester 1 sampai semester 6 yang mana didalam setiap semester akan ada mata pelajaran yang sama sehingga harus diseleksi terlebih dahulu dengan mencari nilai rata-ratanya sehingga didapatkan sebuah data yang akan dipakai untuk dataset pada penelitian ini dengan jumlah data 105 dan 22 atribut. Adapun contoh dataset yang akan dipakai dapat dilihat pada tabel 2 berikut.

Tabel 2. Dataset

PAIBP	PPKn	B. INDO	MTK	SEJ. INDO	B. ING	SENBUD	PJOK	SIKOMDIG	FISIKA	KIMIA	KJD	PEMDAS	DDG	SISKOM	TJBL	B. SUNDA	PSPJ
81	83	82	82	83	83	81	82	82	81	82	86	83	84	80	80	83	80
82	83	83	82	82	82	83	83	82	80	81	86	85	84	80	82	83	80
85	85	83	85	87	85	82	86	88	84	84	89	85	88	86	83	87	84
83	84	85	83	87	83	84	84	86	82	83	89	85	85	86	83	86	83
86	84	84	84	87	83	81	83	87	84	83	90	84	87	88	82	87	84
83	84	85	84	85	83	83	82	84	83	84	88	84	85	88	82	86	83
84	84	85	85	86	83	85	83	85	82	84	90	84	86	86	83	88	83
84	83	84	83	87	83	86	83	86	82	83	89	85	86	86	82	86	80
83	83	85	82	83	82	81	84	82	81	81	86	84	85	80	80	83	83
85	86	85	85	85	83	87	83	87	83	84	89	84	87	86	83	87	83
...
...
...
81	78	79	80	82	82	85	81	77	78	81	82	81	83	78	85	84	81
79	78	79	79	78	82	81	82	76	77	80	80	78	76	76	85	88	80
81	81	81	80	78	83	81	80	78	80	81	78	76	76	76	79	85	80

C. Perhitungan

1. Perhitungan Manual

a. Algoritma *K-Means*

Perhitungan manual algoritma *K-Means* menggunakan Microsoft Excel 2019 menghasilkan perhitungan sebanyak 10 iterasi seperti dengan langkah-langkah yang telah disebutkan sebelumnya sehingga hasil iterasi terakhir dapat dilihat pada tabel 3 dimana angka yang bercetak tebal menunjukkan data tersebut masuk ke cluster 1, 2 atau 3.

Tabel 3. Hasil Perhitungan Manual *K-Means*

No	Nama	C1	C2	C3
1	ADTYA AGUSTIN	13,14999	7,772229	14,92493

No	Nama	C1	C2	C3
2	AGUNG FATKHURROHMAN SUKANDAR	12,50117	7,210025	14,9584
3	AKMAL MUFLIH HAKIM	8,026344	18,8063	28,06552
4	ALMURIAH	3,493982	14,19046	23,6777
5	ANDREAN KAYIS AKMAL	6,922586	17,59108	27,29164
6	ASRI KURNIYASIH	5,21612	12,3718	22,73485
7	ASTRI BERLIANA	3,644082	15,09151	24,52904
8	CAHYANA ADINATA	4,526988	14,1891	24,2395
9	CANTIKA	11,56012	7,551353	16,00605
10	CUCU KURAESIN	4,832263	16,08582	26,09509
...
...
...
103	SURYA AJIALPIAN	19,6903	11,10543	9,661967
104	SUTRISNA	26,79007	17,8019	8,43763
105	SUTRISNO PURNAMA	25,44449	14,97382	9,376225

Rincian jumlah anggota *cluster* 1 sebanyak 28 Siswa, *Cluster* 2 sebanyak 52 Siswa dan *Cluster* 3 sebanyak 25 Siswa.

b. Algoritma *Fuzzy C-Means*

Perhitungan manual algoritma *Fuzzy C-Means* menggunakan Microsoft Excel 2019 menghasilkan perhitungan sebanyak 26 iterasi seperti dengan langkah-langkah yang telah disebutkan sebelumnya sehingga hasil iterasi terakhir dapat dilihat pada tabel 4 dimana angka yang bercetak tebal menunjukkan data tersebut masuk ke cluster 1, 2 atau 3.

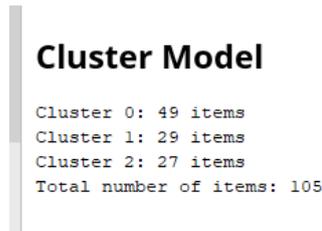
Tabel 4. Hasil Perhitungan Manual *Fuzzy C-Means*

No	Nama	C1	C2	C3
1	ADTYA AGUSTIN	0,22590	0,54338	0,23072
2	AGUNG FATKHURROHMAN SUKANDAR	0,20910	0,55129	0,23961
3	AKMAL MUFLIH HAKIM	0,08423	0,14999	0,76578
4	ALMURIAH	0,02533	0,05244	0,92223
5	ANDREAN KAYIS AKMAL	0,07109	0,13426	0,79465
6	ASRI KURNIYASIH	0,04614	0,11241	0,84145
7	ASTRI BERLIANA	0,03194	0,06447	0,90359
8	CAHYANA ADINATA	0,03834	0,08232	0,87935
9	CANTIKA	0,18629	0,52101	0,29270
10	CUCU KURAESIN	0,04422	0,08896	0,86682
...
...
...
103	SURYA AJIALPIAN	0,53788	0,34717	0,11495
104	SUTRISNA	0,65743	0,23791	0,10466
105	SUTRISNO PURNAMA	0,60326	0,29571	0,10103

Rincian jumlah anggota *cluster* 1 sebanyak 28 Siswa, *Cluster* 2 sebanyak 48 Siswa dan *Cluster* 3 sebanyak 29 Siswa

2. Perhitungan RapidMiner Studio

Perhitungan rapidminer studio versi 9.10 hanya bisa menggunakan algoritma *K-Means* dikarenakan algoritma *Fuzzy C-Means* tidak tersedia. Hasil algoritma *K-Means* pada rapidminer studio 9.10 dapat dilihat pada gambar 2 berikut.

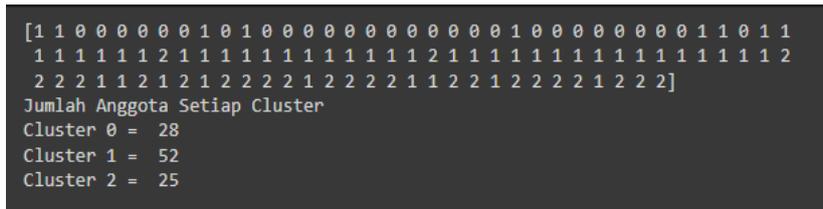


Gambar 2. Hasil Perhitungan Rapidminer Studio

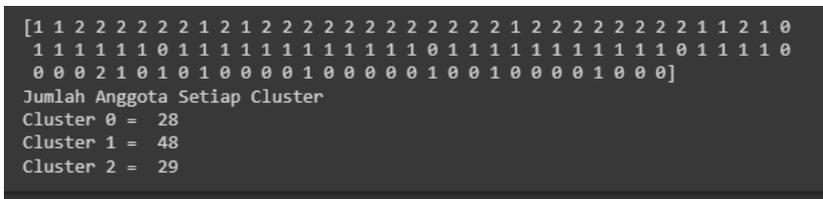
3. Program Python

Perhitungan dengan program python dimaksudkan untuk memperkuat perhitungan sebelumnya yakni perhitungan manual dan menggunakan *RapidMiner*. Dalam membuat program python, peneliti menggunakan *library Google Colab*.

Hasil perhitungan dapat dilihat pada gambar 3 dan 4.



Gambar 3. Hasil Perhitungan Program Python Algoritma K-Means



Gambar 4. Hasil Perhitungan Program Python Algoritma Fuzzy C-Means

D. Evaluasi

Evaluasi *cluster* menggunakan metode *Davies Bouldien Index* (DBI) dengan persamaan 8, 9 dan 10 maka menghasilkan nilai DBI sebagai berikut :

1. Algoritma *K-Means*

a. Manual

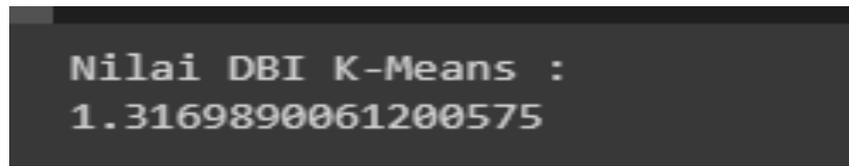
Hasil evaluasi DBI secara manual untuk algoritma *K-Means* menghasilkan nilai sebesar 1,3207

b. RapidMiner Studio



Gambar 5. Nilai DBI *K-Means* pada RapidMiner

c. Program Python



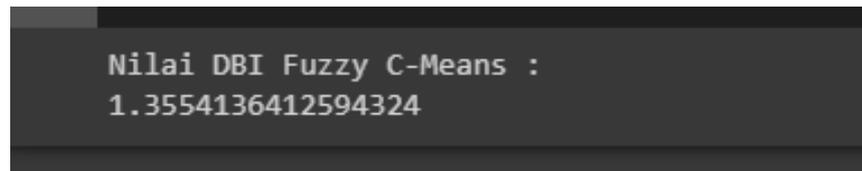
Gambar 6. Nilai DBI *K-Means* pada Program Python

2. Algoritma *Fuzzy C-Means*

a. Manual

Hasil evaluasi DBI secara manual untuk algoritma *Fuzzy C-Means* menghasilkan nilai sebesar 1,7965

b. Program Python

Gambar 7. Nilai DBI *Fuzzy C-Means* pada Program Python

E. Hasil Penelitian

Setelah hasil dari semua jenis perhitungan dan nilai DBI telah diketahui selanjutnya menentukan cluster siswa berprestasi, prestasi sedang dan siswa tidak berprestasi dengan menjumlahkan nilai perhitungan jarak setiap anggota *cluster* pada algoritma *K-Means* dan menjumlahkan nilai matriks partisi setiap *cluster* pada *algoritma Fuzzy C-Means* kemudian mengurutkannya dari nilai terkecil, sedang dan tertinggi. Sehingga hasil akhir penelitian ini dapat dilihat pada tabel 5 dan 6.

Tabel 5. Hasil Perhitungan Manual dan Program Python Dengan Kategori

ALGORITMA	CLUSTER	JUMLAH	KATEGORI
<i>K-MEANS</i>	0	28	TIDAK BERPRESTASI
	1	52	BERPRESTASI
	2	25	PRESTASI SEDANG
<i>FUZZY C-MEANS</i>	0	28	TIDAK BERPRESTASI
	1	48	BERPRESTASI
	2	29	PRESTASI SEDANG

Berdasarkan hasil perhitungan secara manual menggunakan *Microsoft Excel 2019* dan Program *Python*. *Cluster* yang dihasilkan sama yaitu baik menggunakan algoritma *K-Means* maupun algoritma *Fuzzy C-Means* seperti yang telah disajikan pada tabel 5 diatas. Akan tetapi pada perhitungan menggunakan *tools Rapidminer* menghasilkan *cluster* yang berbeda. Hasil perhitungan *tools Rapidminer* dengan kategori dapat dilihat pada tabel 6.

Tabel 6. Hasil Perhitungan Tools *Rapidminer* Dengan Kategori

ALGORITMA	CLUSTER	JUMLAH	KATEGORI
<i>K-MEANS</i>	0	49	BERPRESTASI
	1	29	TIDAK BERPRESTASI
	2	27	PRESTASI SEDANG

IV. KESIMPULAN

Kesimpulan hasil penelitian dari pengelompokan prestasi siswa berdasarkan nilai akademik menggunakan algoritma *K-Means* dan *Fuzzy C-Means* adalah perhitungan manual dan program python pada algoritma *K-Means* menghasilkan 52 siswa berprestasi, 25 siswa dengan prestasi sedang serta 28 siswa tidak berprestasi, adapun pada algoritma *Fuzzy C-Means* menunjukkan 48 siswa berprestasi, 29 siswa dengan prestasi sedang dan 28 siswa tidak berprestasi. Perhitungan menggunakan *rapidminer studio* dengan algoritma *K-Means* menghasilkan 49 siswa berprestasi, 27 siswa dengan prestasi sedang serta 29 siswa tidak berprestasi, adapun algoritma *Fuzzy C-Means* tidak dapat diproses pada *rapidminer studio* karena tidak adanya algoritma tersebut dalam *rapidminer studio*. Hasil evaluasi *cluster* dalam mengelompokan prestasi siswa berdasarkan nilai akademik menggunakan algoritma *K-Means* dan *Fuzzy C-Means* dengan

metode *Davies Bouldien Index* (DBI) dari setiap jenis perhitungan menunjukkan algoritma *K-Means* merupakan algoritma terbaik karena memiliki nilai DBI yang mendekati 0 dan tidak negatif.

PENGAKUAN

Naskah ilmiah ini adalah sebagian dari penelitian Tugas Akhir milik Salim Kurniawan, dengan judul Pengelompokan Presrtasi Siswa Berdasarkan Nilai Akademik Menggunakan Algoritma *K-Means* dan Algoritma *Fuzzy C-Means* yang dibimbing oleh Amril Mutoi Siregar dan Hilda Yulia Novita.

DAFTAR PUSTAKA

- [1] Rosyid, M.Z., Mansyur, M., IP, S. and Abdullah, A.R., 2019. *Prestasi belajar*. Literasi Nusantara.
- [2] Undang-undang Republik Indonesia Nomor 20 Tahun 2003. "Sistem pendidikan nasional". Jakarta: Direktorat Pendidikan Menengah Umum.
- [3] Peraturan Menteri Pendidikan dan Kebudayaan Nomor 34 Tahun 2018. "*Standar nasional pendidikan sekolah Menengah Kejuruan (SMK) / Madrasah Aliyah Kejuruan (MAK)*".
- [4] Surat Edaran Menteri Pendidikan dan Kebudayaan Nomor 4 Tahun 2020. "*Pelaksanaan Kebijakan Pendidikan dalam Masa Darurat Penyebaran Coronavirus Disease (COVID-19)*"
- [5] Prayoga, Y., Mahmudi, A. and Zahro, H.Z., 2021. Penerapan Metode K Means Pada Sistem Informasi Akademik Untuk Pengelompokan Siswa Berprestasi Di Upt Sma Negeri 3 Kota Pasuruan Berbasis Web. *Jati (Jurnal Mahasiswa Teknik Informatika)*, 5(2), pp.822-828.
- [6] Wiharto, W. and Suryani, E., 2020. The comparison of clustering algorithms k-means and fuzzy c-means for segmentation retinal blood vessels. *Acta Informatica Medica*, 28(1), p.42
- [7] Luka, S.S.P., Candiasa, I.M. and Aryanto, K.Y.E., 2019. Analisis Pembentukan Kelompok Diskusi Panel Siswa Menggunakan Algoritma Fuzzy C-Means dan K-Means. *Jurnal Pendidikan Teknologi dan Kejuruan*, 16(2), pp.267-277.
- [8] Hariyadi, A. and Darmuki, A., 2019. Prestasi Dan Motivasi Belajar Dengan Konsep Diri. In *Prosiding Seminar Nasional Penguatan Muatan Lokal Bahasa Daerah Sebagai Pondasi Pendidikan Karakter Generasi Milenial* (Pp. 280-286).
- [9] Agustina, N. and Prihandoko, P., 2018. Perbandingan Algoritma K-Means dengan Fuzzy C-Means Untuk Clustering Tingkat Kedisiplinan Kinerja Karyawan. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 2(3), pp.621-626.
- [10] Pramitasari, A.E. and Nataliani, Y., 2021. Perbandingan Clustering Karyawan Berdasarkan Nilai Kinerja Dengan Algoritma K-Means Dan Fuzzy C-Means. *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, 8(3), pp.1119-1132.
- [11] Jollyta, D., Siddik, M., Mawengkang, H. and Efendi, S., 2021. *Teknik Evaluasi Cluster Solusi Menggunakan Python Dan Rapidminer*. Deepublish.
- [12] Gie, W. and Jollyta, D., 2020, July. Perbandingan Euclidean dan Manhattan Untuk Optimasi Cluster Menggunakan Davies Bouldin Index: Status Covid-19 Wilayah Riau. In *Prosiding Seminar Nasional Riset Information Science (SENARIS)* (Vol. 2, pp. 187-191).