

APPLICATION OF K-MEANS CLUSTERING METHOD TO CLUSTER STUDENTS' ENGLISH SKILL JASON ENGLISH COURSE

Khairunnisa^{1*}, Juna Eska², Mustika Fitri Larasati³

^{1,2,3}Program Studi Sistem Informasi, Sekolah Tinggi Manajemen Informatika dan Komputer (STMIK) Royal Kisaran, Indonesia

Email: ¹kn378979@gmail.com, ²dosen.junaeska@gmail.com, ³bukmus.inaction@gmail.com

(Naskah masuk: 14 Februari 2022, Revisi: 26 Februari 2022, diterbitkan: 28 Juni 2022)

Abstract

The role of students with high English skills and higher quality standards in an English course institution is one aspect that has a significant impact on the course institution's quality. To generate positive competition among students and to produce the best quality of students, student English skills should be clustered. The determination of students' English skills at the Jason English Course is currently done with the results of decisions that are off target, impenetrable, and inefficient because the calculations are still not computerized and depend on the tutor's personal feelings, causing it unachievable the initial goal of clustering these skills. The goal of this research is to make it easier for tutors to group students' English skills with targeted, transparent, and efficient decision results so that students can recognize their abilities and tutors can create appropriate learning methods so that students do not encounter learning obstacles. As a result, the course institution requires the appropriate technology to assist tutors in grouping students' abilities. The method used in this research was k-means clustering, which is a data mining technique. This research data is calculated and tested using Microsoft Excel, RapidMiner, Design Systems with PHP programming languages, and databases MySQL. After calculating and testing the data, the accuracy obtained on the data sample reached 80%. As a result, using K-means clustering, Student English Skill Jason English Course can be clustered quickly, clearly, efficiently, and accurately.

Keywords: Algorithm, K-Means Clustering, Skill, PHP.

PENGGUNAAN K-MEANS CLUSTERING UNTUK MENGELOMPOKKAN KEMAMPUAN BAHASA INGGRIS SISWA LEMBAGA KURSUS JASON ENGLISH COURSE

Abstrak

Peranan siswa yang memiliki kemampuan bahasa inggris yang baik dan standar kualitas lebih disebuah lembaga kursus pendidikan bahasa inggris merupakan salah satu bagian yang sangat berpengaruh pada kualitas lembaga kursus tersebut. Untuk menumbuhkan daya saing yang positif antar siswa dan menghasilkan kualitas siswa yang terbaik maka dilakukan *clustering student english skill* atau pengelompokkan kemampuan bahasa inggris siswa. Penentuan kemampuan Bahasa Inggris siswa pada Lembaga Kursus *Jason English Course* saat ini dilakukan dengan hasil keputusan yang tidak tepat sasaran, tidak transparan, dan kurang efisien karena perhitungan masih belum terkomputerisasi dan menggunakan perasaan pribadi tutor sehingga tidak dapat mencapai tujuan awal dari pengelompokan kemampuan ini. Tujuan dari penelitian ini adalah untuk memudahkan tutor dalam mengelompokkan kemampuan Bahasa Inggris Siswa dengan hasil keputusan yang tepat sasaran, transparan, dan efisien sehingga siswa dapat mengenali kemampuan yang dimilikinya dan tutor mampu membuat metode belajar yang sesuai agar siswa tidak mengalami hambatan dalam belajar. Maka dari itu pihak lembaga kursus membutuhkan teknologi yang tepat untuk membantu tutor dalam mengelompokkan kemampuan yang dimiliki siswa. Pada penelitian ini metode yang diterapkan menggunakan teknik *data mining* yaitu *k-means clustering*. Data penelitian ini dihitung dan diuji dengan *Microsoft Excel*, *RapidMiner*, Sistem Rancangan dengan Bahasa Pemrograman *PHP* dan basis data dengan *MySQL*. Setelah melakukan perhitungan dan pengujian data, akurasi yang diperoleh mencapai 80% pada sampel data. Sehingga penggunaan *k-means clustering* dapat mengelompokkan *Student English Skill Jason English Course* dengan cepat, jelas, efisien, dan akurat.

Kata kunci: Algoritma, Clusterisasi K-Means, Kemampuan, PHP.

1. PENDAHULUAN

Teknologi yang sangat berkembang pesat dari waktu ke waktu mampu membantu pekerjaan yang dilakukan manusia dapat diselesaikan dengan cepat. Derasnya perkembangan globalisasi dari tahun ke tahun dalam segala bidang terutama dalam bidang teknologi dan pendidikan, menyebabkan timbulnya persaingan dalam penyediaan sumber daya manusia yang unggul dan berkualitas [1]. Pendidikan merupakan sarana yang penting dalam menciptakan Sumber Daya Manusia (SDM) yang berkualitas [2]. Permasalahan yang dihadapi para penerus bangsa Indonesia ialah masih banyaknya yang menganggap pelajaran itu sebuah mimpi buruk, terkhusus untuk pelajaran Bahasa Inggris, tetapi anggapan tersebut tidaklah benar. Jika pembelajaran Bahasa Inggris menggunakan metode belajar yang menarik dan cara penyampaian yang baik, maka Bahasa Inggris adalah pelajaran yang sangat mudah untuk dipahami [3]. Komunikasi dan interaksi global yang pesat menjadikan Bahasa Inggris sebagai salah satu cara berkomunikasi yang sangat diperlukan [4]. Tanpa keterampilan Bahasa Inggris yang memadai sumber daya manusia akan menghadapi masalah dalam membangun interaksi global karena Bahasa Inggris merupakan bahasa Internasional [5].

Kendala yang dihadapi banyak siswa yaitu ketidaktahuan skill yang dimiliki siswa sendiri [6]. Maka dari itu, pihak lembaga kursus menggunakan *clustering Students' English Skill* nya guna menerapkan metode belajar yang cocok untuk siswa-siswa nya. Permasalahan yang terjadi selama ini penentuan *Students' English Skill* dilakukan dengan cara sederhana dengan hasil keputusan yang didapatkan tidak tepat sasaran, tidak bersifat transparan dan kurang efisien karena masih dihitung secara manual yang mana tidak dapat mencapai tujuan awal diadakannya pengelompokan tersebut. Dari permasalahan yang melatarbelakangi tersebut maka didapat cara yaitu pengelompokan *Students' English Skill* dengan menerapkan teknik *data mining* dengan metode algoritma *k-means clustering* sebagai solusi dari permasalahan yang terdapat pada lembaga kursus tersebut. Dengan data yang sudah dikelompokkan menggunakan algoritma *k-means clustering* dapat mempermudah *Jason English Course* dalam mengelompokkan kemampuan tiap siswa-siswi nya.

Data Mining ialah suatu metode yang bermanfaat dalam mengolah data yang bertujuan untuk memilih suatu pola tersembunyi dari data yang diolah [7]. Berdasarkan hasil data tersebut, kemudian didapatkan suatu pengetahuan baru yang bersumber dari data yang lama dan dari pengolahan data tersebut dapat dipergunakan untuk menentukan sebuah keputusan di masa depan[8]. Pada proses *data mining* dilakukan beberapa tahapan diantaranya diawali dari seleksi data dari data sumber ke data target, tahap *pre-processing* untuk memperbaiki kualitas data, transformasi, *data mining* serta tahap

interpretasi dan evaluasi yang menghasilkan *output* berupa pengetahuan baru yang diharapkan memberikan kontribusi yang lebih baik[9][10].

K-means adalah salah satu algoritma non-hierarki di mana proses *clustering* berdasarkan jarak terdekat dengan titik pusat yang ditentukan [11]. Secara umum, metode *k-means clustering* hanya bisa mengolah data dalam bentuk angka, maka untuk data yang berbentuk nominal seperti wilayah, jenis dan lain-lain harus diinisialisasikan terlebih dahulu dalam bentuk angka [12]. Tujuannya untuk meminimalkan fungsi objektif yang diatur dalam proses pengelompokan, yang pada umumnya berusaha meminimalkan variasi di dalam suatu kelompok dan memaksimalkan variasi antar kelompok[11].

2. METODE PENELITIAN

Dalam penelitian ini metode yang digunakan adalah pendekatan kuantitatif. Penelitian kuantitatif adalah jenis penelitian yang sistematis, terstruktur, tersusun dari awal hingga akhir sehingga riset ini memiliki kecenderungan untuk menggunakan teknik analisis angka-angka secara statistik [13]. Tahapan penelitian yang dilakukan sebagai berikut:

2.1. Pengumpulan Data

Pengumpulan data pada penelitian ini dilakukan dengan metode wawancara ke pihak terkait di *Jason English Course*. Data yang didapatkan berupa data-data siswa yang berkaitan dengan kriteria *Clustering English Skill Jason English Course* sejak tahun 2019 – 2021.

2.2. Analisis Data

Data yang didapatkan dari Lembaga Kursus *Jason English Course* selanjutnya akan diolah yang mana akan digunakan untuk *Clustering Student English Skill Jason English Course* dengan menerapkan metode algoritma *k-means clustering*.

2.3. Perancangan Sistem

Perancangan sistem merupakan tahapan awal untuk merancang sebuah sistem. Hal-hal yang perlu dirancang untuk membangun sebuah sistem ialah merancang *Flowmap*, ERD (*Entity Relationships Diagram*), UML (*Unified Modeling Language*) dan *Flowchart*.

2.4. Pembangunan Sistem

Pada tahapan ini sistem harus dibangun dengan rancangan yang sudah dibuat. Sistem yang dibangun akan digunakan oleh pihak Lembaga Kursus *Jason English Course* yang mana sebelumnya sistem akan melewati tahap uji coba.

Sistem ini nantinya dibangun dengan bahasa pemrograman *PHP* dan memiliki fitur yang membantu pihak Lembaga Kursus dalam

mengelompokkan kemampuan siswanya dengan menggunakan teknik *data mining* dengan menerapkan metode *k-means clustering*.

2.5. Uji Coba Sistem

Tahapan ini dilakukan untuk menilai kelayakan sistem yang akan digunakan agar bisa digunakan oleh Lembaga Kursus *Jason English Course*.

2.6. Implementasi Sistem

Implementasi sistem merupakan tahapan akhir yang mana sistem yang sudah melewati tahapan uji coba kelayakan sistem akan dilakukan implementasian kepada Lembaga Kursus *Jason English Course*.

3. HASIL DAN PEMBAHASAN

3.1. Langkah- Langkah Algoritma *K-means clustering*

Langkah-langkah dalam algoritma *k-means clustering* [13]:

- 1) Menentukan jumlah cluster.
- 2) Menentukan nilai centroid.
 Dalam menentukan nilai centroid untuk awal iterasi, nilai awal centroid dilakukan secara acak. Sedangkan jika menentukan nilai centroid yang merupakan tahap dari iterasi, maka digunakan rumus sebagai berikut (1):

$$\bar{v}_{ij} = \frac{1}{N_i} \sum_{k=0}^{N_i} x_{kj} \tag{1}$$

- 3) Menghitung jarak antara titik centroid dengan titik tiap objek dengan menggunakan persamaan *Euclidean Distance Space* (2).

$$D_e = \sqrt{(x_i - s_i)^2 + (y_i - t_i)^2} \tag{2}$$

- 4) Pengelompokan object untuk menentukan anggota cluster adalah dengan memperhitungkan jarak minimum objek.
- 5) Kembali ke tahap 2, lakukan perulangan hingga nilai centroid yang dihasilkan tetap dan anggota cluster tidak berpindah ke cluster lain.

Metode *K-means clustering* hanya bisa mengolah data dalam bentuk angka, maka untuk data yang berbentuk nominal harus di inialisasikan terlebih dahulu dalam bentuk angka. Langkahnya adalah:

- Urutkan data berdasarkan frekuensi kemunculannya
- Inialisasikan data tersebut mulai dari data tertinggi dengan nilai 1, kemudian data selanjutnya 2, 3 dan seterusnya [15].

3.2. Perhitungan Algoritma *K-Means*

Tabel 1 di bawah ini merupakan tabel yang berisikan data nilai siswa Lembaga Kursus *Jason English Course*.

Table 1 Data Nilai Siswa Jason English Course

NISN	Nama Siswa	Grammar	Speaking & Reading	Listening
1	Ratu	89	90	75
2	Feby	90	71	95
3	Elisah	70	75	80
4	Mutiara	45	65	59
5	Alga	65	75	52
6	Algi	80	70	75
7	Keysa	90	85	81
8	Puja	70	70	73
9	Andrian	95	92	85
10	Farhan	60	55	48
11	Syifa	45	60	58
12	Jauhara	60	70	72
13	Novita	85	90	88
14	Nazwa	52	68	55
15	Wali	40	60	70

Data pada Tabel 1 merupakan data yang diolah untuk penelitian ini merupakan sampel data yang peserta didik Lembaga Kursus *Jason English Course* tahun ajaran 2019/2020. Atribut yang dijadikan variabel dalam algoritma *k-means* terdiri dari nisn, nama, nilai *grammar*, *speaking & reading* serta *listening*. Data sampel diuji coba pada 15 peserta didik, kemudian mengelompokkan data di atas menjadi 3 kelompok. Dalam penelitian ini jumlah cluster yang digunakan sebanyak 3 cluster dan sudah teruji pembuktiannya dengan menggunakan *RapidMiner*[16][17]. Berikut langkah-langkah *clustering* data ini dengan menggunakan algoritma *k-means*:

- 1) Penentuan jumlah cluster dan dalam penelitian ini jumlah cluster yang akan digunakan sebanyak 3 cluster.
- 2) Penentuan centroid awal secara acak. Dalam penelitian ini centroid yang didapat adalah:

Table 2 Centroid Awal yang Diambil Secara Acak

Cluster	Grammar	Speaking & Reading	Listening
C1	95	92	85
C2	65	75	52
C3	40	60	70

- 3) Perhitungan jarak setiap data terhadap setiap pusat cluster dengan menggunakan persamaan *Euclidean Distance Space* berikut;
 - a. Tahap pertama yaitu perhitungan jarak antara data siswa pertama dengan pusat cluster pertama dengan menggunakan persamaan (1).

$$C1 \sqrt{SQRT(((89 - 95)^2) + (90 - 92)^2) + (75 - 85)^2} \tag{1}$$

- b. Tahap kedua yaitu perhitungan jarak antara data siswa pertama dengan pusat cluster kedua dengan menggunakan persamaan (2).

$$C1 \sqrt{SQRT(((89 - 65)^2) + (90 - 75)^2) + (75 - 52)^2)} \quad (2)$$

c. Selanjutnya, menghitung jarak antara data siswa pertama dengan pusat *cluster* ketiga (3).

$$C1 \sqrt{SQRT(((89 - 40)^2) + (90 - 60)^2) + (75 - 70)^2)} \quad (3)$$

- 4) Setelah menghitung jarak, data pada masing-masing centroid secara manual dengan *Microsoft Excel* maka diperoleh hasil perhitungan seperti pada Tabel 3.
- 5) Lalu kelompokan data menjadi 3 kelompok yaitu kelompok 1, 2 dan 3. Hasil pada pengelompokan data dapat dilihat pada Tabel 3.

Table 3 Hasil Perhitungan Jarak Data Nilai Siswa pada Masing-Masing Centroid Pada Iterasi 1

Grammar	Speaking & Reading	Listening	C1	C2	C3	C1	C2	C3
89	90	75	12	36	58	1	2	3
90	71	95	24	50	57	1	2	3
70	75	80	31	28	35	1	2	3
45	65	59	62	23	13	1	2	3
65	75	52	48	0	34	1	2	3
80	70	75	28	28	42	1	2	3
90	85	81	9	40	57	1	2	3
70	70	73	35	22	32	1	2	3
95	92	85	0	48	65	1	2	3
60	55	48	63	21	30	1	2	3
45	60	58	65	26	13	1	2	3
60	70	72	43	21	22	1	2	3
85	90	88	11	44	57	1	2	3
52	68	55	58	15	21	1	2	3
40	60	70	65	34	0	1	2	3

Table 4 Centroid Baru Iterasi 2

Cluster	Grammar	Speaking & Reading	Listening
C1	90	65	43
C2	86	69	62
C3	85	65	62

- 6) Setelah semua data dikelompokkan. Kemudian hitung kembali pusat *cluster* yang baru dengan nilai siswa, sesuai dengan anggota kelompok yang ada pada *cluster* tersebut. Pada Tabel 4 diatas adalah tabel dari *centorid* baru, hasil dari perhitungan pada cluster awal di iterasi kedua yang dihitung dengan *Microsoft Excel*.
- 7) Setelah didapat *centroid* baru dari tiap *cluster*, hitung kembali nilai siswa dengan pusat *cluster*

yang baru sesuai anggota yang sudah dikelompokkan sebelumnya. Lakukan perulangan sampai didapat pola terakhir yang sudah tidak berpindah. Berikut hasil perhitungan pada iterasi ke-2 yang dilakukan dengan *Microsoft Excel* dapat dilihat pada Tabel 5.

Table 5 Hasil Perhitungan Jarak Data Nilai Siswa pada Masing-Masing Centroid Pada Iterasi 2

Grammar	Speaking & Reading	Listening	C1	C2	C3	JARAK	C1	C2	C3
89	90	75	117	1117	3069	117	1	2	3
90	71	95	325	1529	3379	325	1	2	3
70	75	80	546	286	1222	286	1	2	3
45	65	59	3142	452	22	22	1	2	3
65	75	52	1835	205	753	205	1	2	3
80	70	75	456	326	1602	326	1	2	3
90	85	81	17	1137	3099	17	1	2	3
70	70	73	800	90	914	90	1	2	3
95	92	85	61	1829	4133	61	1	2	3
60	55	48	3230	510	534	510	1	2	3
45	60	58	3430	530	24	24	1	2	3
60	70	72	1325	75	453	75	1	2	3
85	90	88	50	1370	3224	50	1	2	3
52	68	55	2668	270	166	166	1	2	3
40	60	70	3401	731	77	77	1	2	3

- 8) Karena belum didapat pola yang sama pada perhitungan iterasi ke-1 dan ke-2 maka dilakukan perhitungan iterasi yang ke-3, lakukan sampai perhitungan mendapatkan pola yang sama. Untuk melakukan perhitungan

iterasi ke-3 harus menggunakan centroid baru. Maka nilai centroid baru pada iterasi ke-3 didapat seperti pada Tabel 6 dibawah ini:

Table 6 Centroid Baru Iterasi 3

Cluster	Grammar	Speaking & Reading	Listening
C1	90	69	48
C2	86	72	62
C3	85	70	58

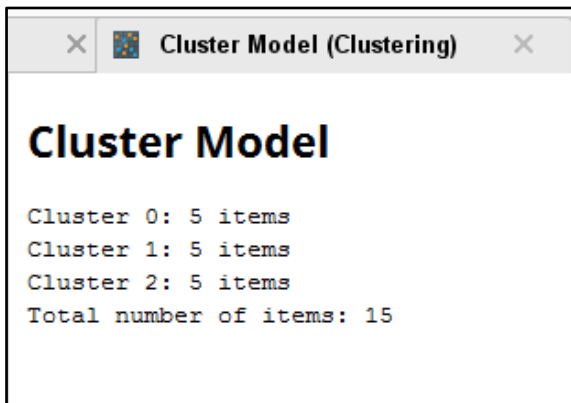
9) Selanjutnya lakukan perhitungan nilai siswa dengan centroid baru diatas sesuai anggota yang sudah dikelompokkan sebelumnya. Maka didapat hasil perhitungan sampai iterasi ke-3 seperti pada tabel dibawah ini:

Table 7 Hasil Perhitungan Jarak Data Nilai Siswa pada Masing-Masing Centroid Pada Iterasi 3 yang Tidak Ada Lagi Perpindahan Cluster

	Grammar	Speaking & Reading	Listening	C1	C2	C3	JARAK	C1	C2	C3
	89	90	75	117	749	2754	117	1	2	3
	90	71	95	325	1067	3214	325	1	2	3
	70	75	80	546	110	1137	110	1	2	3
	45	65	59	3142	746	19	19	1	2	3
	65	75	52	1835	349	494	349	1	2	3
	80	70	75	456	150	1377	150	1	2	3
	90	85	81	17	731	2822	17	1	2	3
	70	70	73	800	14	773	14	1	2	3
	95	92	85	61	1301	3838	61	1	2	3
	60	55	48	3230	854	293	293	1	2	3
	45	60	58	3430	864	13	13	1	2	3
	60	70	72	1325	89	404	89	1	2	3
	85	90	88	50	904	3053	50	1	2	3
	52	68	55	2668	530	61	61	1	2	3
	40	60	70	3401	985	212	212	1	2	3

Berdasarkan hasil perhitungan yang dilakukan dengan *Microsoft Excel* maka hasil yang didapat yaitu masing-masing cluster terdiri dari 5 data.

Hasil pengujian dengan menggunakan *RapidMiner* pada Gambar 1 sama dengan perhitungan manual dengan menggunakan *Microsoft Excel*. Untuk melihat kecenderungan kemampuan siswa pada suatu kelompok dapat dilihat pada Gambar 2.



Gambar 1 Cluster Model

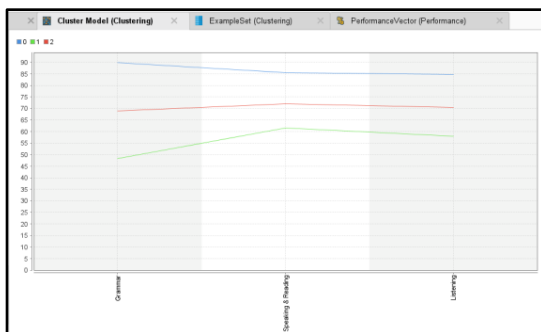
Row No.	Id	cluster	Nama	Grammar	Speaking & Reading	Listening
1	1	cluster_0	Ratu Mumtaz	89	90	75
2	2	cluster_0	Fely Rahma	90	71	95
3	3	cluster_2	Elsah	70	75	80
4	4	cluster_1	Mubara Bulan	45	65	59
5	5	cluster_2	Alga	65	75	52
6	6	cluster_2	Algi	80	70	75
7	7	cluster_0	Kaysa Ivanca	90	85	81
8	8	cluster_2	Puja Kesuma	70	70	73
9	9	cluster_0	Ardian	95	92	85
10	10	cluster_1	Fathan Ram	60	55	48
11	11	cluster_1	Silla Aulia	45	60	58
12	12	cluster_2	Jauhara Tsani	60	70	72
13	13	cluster_0	Novita Sari	85	90	88
14	14	cluster_1	Nazka	52	68	55
15	15	cluster_1	Mutawally	40	60	70

Gambar 3 Nama-Nama Siswa yang Sudah Dikelompokkan Berdasarkan Kemampuan

Setelah hasil dari analisis data dihitung. Maka selanjutnya, data yang sudah di hitung akan di uji menggunakan *Rapid Miner*. Pengujian ini dilakukan untuk melihat kesesuaian hasil perhitungan data dengan menggunakan *Microsoft Excel*. Cluster model yang diperoleh dari hari pengujian sistem terhadap data akan terlihat pada Gambar 1 di atas.

Pada Gambar 3 diatas menampilkan nama-nama siswa yang sudah dikelompokkan dengan *RapidMiner*.

Dilihat dari hasil perhitungan dan pengujian terhadap sampel data maka dapat disimpulkan *k-means clustering* mampu mengelompokkan *Students' English Skill Jason English Course* dengan akurasi mencapai 80% dari sampel data yang digunakan.



Gambar 2 Grafik Kecenderungan Kemampuan Siswa

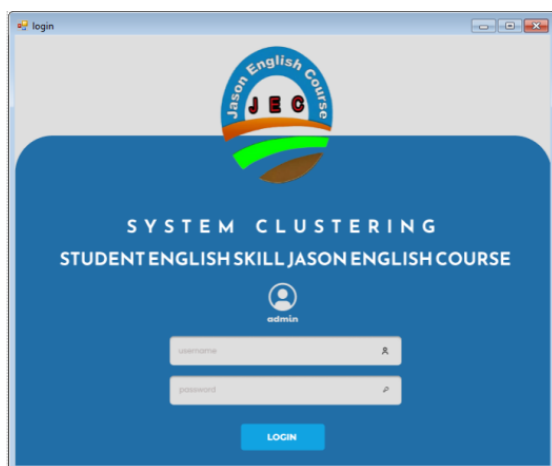
3.3. Implementasi Sistem

Implementasi sistem adalah tahapan yang mana sistem yang sudah melewati tahapan uji coba kelayakan sistem akan dilakukan implementasian kepada Lembaga Kursus *Jason English Course*. Implementasi dari pengelompokkan *Students' English Skill* ini menerapkan bahasa pemograman *PHP* dan untuk perhitungan data dilakukan dengan hitungan persamaan *Euclidean Distance Space*

menggunakan *Microsoft Excel* dan diuji dengan *RapidMiner*.

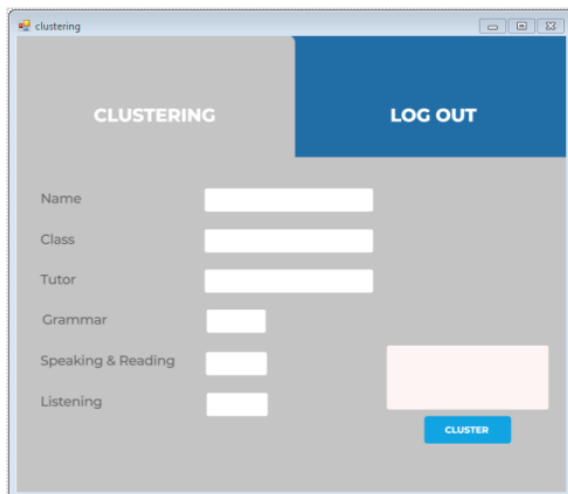
Tampilan sistem yang akan diterapkan pada LKP *Jason English Course* dapat dilihat pada Gambar 4 dan 5.

- 1) Tampilan login sistem menunjukkan admin harus memasukkan username dan password agar bisa masuk ke dalam sistem. Tampilan login sistem dapat dilihat pada Gambar 5.



Gambar 4. Tampilan Login Sistem

Tampilan clustering merupakan tampilan menu clustering yang digunakan untuk mengkluster siswa agar dikelompokkan sesuai skill yang dimilikinya. Tampilan menu clustering dapat dilihat pada Gambar 6.



Gambar 5 Tampilan Clustering

4. KESIMPULAN

Kesimpulan yang diperoleh yaitu penggunaan *k-means clustering* mampu mengkluster *Student English Skill* atau kemampuan bahasa Inggris siswa *Jason English Course* dengan lebih cepat, jelas dan akurat. Pengelompokan *Student English Skill* dengan menggunakan metode algoritma *K-means* dilakukan perhitungan sampai iterasi ke-3. Keakuratan yang terjadi mencapai 80% dapat mengkluster data pada

sampel yang dikluster setelah dilakukan perhitungan secara manual dengan *Microsoft Excel* dan dilakukan pengujian dengan *RapidMiner*.

DAFTAR PUSTAKA

- [1] S. Aisyah, "Dampak Teknologi Informasi dan Komunikasi Terhadap Pembentukan Akhlak Peserta Didik Perspektif Pendidikan Islam di MTs. Olang Kecamatan Ponrang Selatan ...," *J. Konsepsi*, vol. 8, no. 2, pp. 45–54, 2019, [Online]. Available: <https://p3i.my.id/index.php/konsepsi/article/view/18%0Ahttps://p3i.my.id/index.php/konsepsi/article/download/18/80>.
- [2] D. Society, D. A. N. Revolusi, and I. Di, "Generasi Milenial Yang Siap Menghadapi Era Revolusi," no. November, pp. 1–6, 2021.
- [3] D. S. Adzkiya and M. Suryaman, "Penggunaan Media Pembelajaran Google Site dalam Pembelajaran Bahasa Inggris Kelas V SD," *J. Teknol. Pendidik.*, vol. 6, no. 2, pp. 20–32, 2021, doi: 10.32832/educate.v6i2.4891.
- [4] I. Sari, "Motivasi Belajar Mahasiswa Program Studi Manajemen dalam Penguasaan Keterampilan Berbicara (Speaking) Bahasa Inggris," *Manaj. Tools*, vol. 9, no. 1, pp. 41–52, 2018, [Online]. Available: <http://jurnal.pancabudi.ac.id/index.php/JUMANT/article/view/191>.
- [5] M. Maliki and S. A. Hasyiyati, "Pelatihan Bahasa Inggris Pekerja Migran Indonesia di Brunei Darussalam," no. 1996, pp. 39–48, 2022.
- [6] Z. Aini, H. Nirwana, and M. Marjohan, "Kontribusi Penguatan Guru Mata Pelajaran Dan Kepercayaan Diri Siswa Terhadap Keaktifan Siswa Dalam Belajar," *Biblio Couns J. Kaji. Konseling dan Pendidik.*, vol. 1, no. 1, pp. 1–11, 2018, doi: 10.30596/bibliocouns.v1i1.1941.
- [7] E. Srikanti, R. F. Yansi, Norvahina, I. Permana, and F. N. Salisah, "Penerapan Data Mining Untuk Menganalisis Penjualan Barang dengan Menggunakan Metode Apriori pada Supermarket Sejahtera Lhoksumawe," *J. Ilm. Rekayasa dan Manaj. Sist. Inf.*, vol. 4, pp. 77–80, 2018.
- [8] R. Nofitri and J. Eska, "Implementasi Data Mining Klasifikasi C4.5 Dalam Menentukan Kelayakan Pengambilan Kredit," *Semin. Nas. R.*, vol. 9986, no. September, pp. 307–310, 2018.
- [9] E. Ongko, A. Teknologi, and I. Immanuel, "Perbaikan Performance K-Means Melalui Perbaikan," vol. 9, no. 2, pp. 37–40, 2021.
- [10] N. Purba, P. Poningsih, and H. S. Tambunan, "Penerapan Algoritma K-Means Clustering Pada Penyebaran Penyakit Infeksi Saluran

- Pernapasan Akut (ISPA) di Provinsi Riau,” *J. Inf. Syst. Res.*, vol. 2, no. 3, pp. 220–226, 2021, [Online]. Available: <http://ejurnal.seminar-id.com/index.php/josh/article/view/736>.
- [11] L. Maulida, “Penerapan Datamining Dalam Mengelompokkan Kunjungan Wisatawan Ke Objek Wisata Unggulan Di Prov. Dki Jakarta Dengan K-Means,” *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 2, no. 3, p. 167, 2018, doi: 10.14421/jiska.2018.23-06.
- [12] R. Nainggolan, “OPTIMASI PERFORMA CLUSTER K-MEANS MENGGUNAKAN SUM OF Squared Error,” *J. Gortap Lumbantoruan*, vol. 2, no. 2, pp. 103–108, 2018.
- [13] R. Yulianti, “Identifikasi Gaya Belajar Statistika Mahasiswa Pendidikan Matematika di Papua (Survei: Kota Jayapura),” *J. Pendidik.*, vol. 10, no. 1, pp. 104–113, 2022.
- [14] E. Ramadanti, “PENERAPAN DATA MINING ALGORITMA K-MEANS CLUSTERING PADA POPULASI AYAM PETELUR DI INDONESIA,” *J. Teknol. dan Sist. Inf. Univrab*, vol. 7, no. 1, pp. 1–7, 2022.
- [15] M. R. Muttaqin *et al.*, “PENERAPAN K-MEANS CLUSTERING DAN CROSS-INDUSTRY STANDARD PROCESS FOR DATA MINING (CRISP-DM) UNTUK,” vol. 19, no. 1, pp. 38–53, 2022.
- [16] B. G. Sudarsono, M. I. Leo, A. Santoso, and F. Hendrawan, “Analisis Data Mining Data Netflix Menggunakan Aplikasi Rapid Miner,” *JBASE - J. Bus. Audit Inf. Syst.*, vol. 4, no. 1, pp. 13–21, 2021, doi: 10.30813/jbase.v4i1.2729.
- [17] L. P. Refialy, H. Maitimu, and M. S. Pesulima, “Perbaikan Kinerja Clustering K-Means pada Data Ekonomi Nelayan dengan Perhitungan Sum of Square Error (SSE) dan Optimasi nilai K cluster,” *Techno.Com*, vol. 20, no. 2, pp. 321–329, 2021, doi: 10.33633/tc.v20i2.4572.