

Pengelompokan Mahasiswa Berdasarkan Pencapaian Prestasi Belajar dari Mata Kuliah yang Ditempuh Berbasis *Web* Dengan *K-Means Clustering*

Student Grouping Based on Learning Achievement from Courses Taken Web-Based With K-Means Clustering

Didik Eko Wiyono^{1,*}

¹*Fakultas Teknologi Informasi, Universitas Islam Balitar
Jl. Majapahit No.2- 4, Sananwetan, Kec. Sananwetan, Kota Blitar, Jawa Timur 66137 Indonesia*

^{1,*}Penulis Korespondensi: ekowdidik@gmail.com

Received on 29-07-2020, accepted on 20-10-2020, published on 31-01-2020

Abstract

Penelitian bertujuan mengetahui cara merancang dan membangun aplikasi pengelompokan mahasiswa berdasarkan pencapaian prestasi belajar dari mata kuliah yang ditempuh berbasis *Web* dengan *K-Means Clustering*, mengetahui hasil pengujian akurasi terhadap pengelompokan mahasiswa berdasarkan pencapaian prestasi belajar dari mata kuliah yang ditempuh berbasis *Web* dengan *K-Means Clustering* menggunakan *Silhouette analysis*. Metode yang digunakan adalah Research and Development dengan *waterfall* sebagai model pengembangan perangkat lunak yang terdiri dari tahapan analisis kebutuhan, desain sistem, implementasi logika, pengujian program, implementasi program. Hasil dari penelitian ini diketahui cara merancang dan membangun aplikasi adalah dengan menyiapkan data set yang diperlukan untuk diolah. Diperlukan juga database untuk menampung seluruh data set yang ada. Diperlukan juga adanya level pengguna dengan kombinasi pemanfaatan fitur relasi tabel dalam database untuk membuat aplikasi yang dapat mengolah banyak data set yang berbeda. Hasil pengujian akurasi terhadap pengelompokan mahasiswa menggunakan *Silhouette analysis* dari data set yang dimiliki adalah 100% untuk mata kuliah Kecerdasan Buatan. Tingkat perbedaan capaian nilai untuk mata kuliah Kecerdasan Buatan oleh mahasiswa kontras. Sehingga, metode tertentu dapat diterapkan untuk memperbaiki keadaan mahasiswanya. Mata kuliah Sistem Basis Data adalah 48%. Sehingga, tingkat pemisahannya untuk mata kuliah Sistem Basis Data tidak begitu terpaut jauh, atau capaian nilai yang diperoleh mahasiswa lebih merata.

Kata kunci: aplikasi, *clustering*, mahasiswa, *silhouette*, web

Abstract

The research aims to find out how to design and build student grouping applications based on the learning achievement of the courses taken Web-based with K-Means Clustering, find out the results of accuracy testing on student grouping based on learning achievement of the courses taken Web-based with K-Means Clustering. using *Silhouette analysis*. The method used is Research and Development with *waterfall* as a software development model which consists of stages of requirements analysis, system design, logic implementation, program testing, program implementation. The results of this study are known how to design and build applications is to prepare the data set needed to be processed. A database is also required to accommodate all existing data sets. It is also necessary to have a user level with a combination of utilizing the table relation features in the database to create

applications that can process many different data sets. The results of testing the accuracy of student grouping using Silhouette analysis from the data set they have are 100% for the Artificial Intelligence course. The degree of difference in grades for Artificial Intelligence courses by contrast students. Thus, certain methods can be applied to improve the condition of students. Database Systems course is 48%. Thus, the level of separation for Database Systems courses is not that far apart, or the achievement of scores obtained by students is more evenly distributed.

Keywords: application, clustering, student, silhouette, web

I. PENDAHULUAN

Proses belajar mengajar adalah interaksi dari pencari ilmu dengan sumber atau pengajar yang dapat membagi pengetahuannya. Belajar merupakan suatu proses usaha yang dilakukan seseorang untuk memperoleh suatu perubahan tingkah laku yang baru secara keseluruhan, sebagai hasil pengalamannya sendiri dalam interaksi dengan lingkungannya [1]. Mengemukakan bahwa mengajar adalah segenap aktivitas kompleks yang dilakukan pengajar dalam mengorganisasi atau mengatur lingkungan sebaik-baiknya dan menghubungkannya dengan anak sehingga terjadi proses belajar [2].

Prestasi merupakan hasil yang telah dicapai seseorang dalam melakukan kegiatan. Prestasi belajar dibedakan menjadi lima aspek, yaitu : kemampuan intelektual, strategi kognitif, informasi verbal, sikap dan keterampilan [3]. Hasil belajar dibedakan menjadi tiga aspek yaitu kognitif, afektif dan psikomotorik. Prestasi merupakan kecakapan atau hasil kongkrit yang dapat dicapai pada saat atau periode tertentu. Berdasarkan pendapat tersebut, prestasi dalam penelitian ini adalah hasil yang telah dicapai siswa dalam proses pembelajaran [4].

Prestasi yang dicapai oleh mahasiswa di Universitas Islam Balitar Blitar belumlah merata. Begitulah sebabnya hingga dirasa perlu dilakukan penelitian. Salah satu langkah yang dapat ditempuh untuk menelusuri pencapaian prestasi yang tidak merata ini adalah dilakukannya pengelompokan mahasiswa berdasarkan kemiripan capaian prestasinya. Saat ini di Universitas Islam Balitar Blitar belum ada aplikasi yang dapat dipergunakan untuk membantu mengelompokkan mahasiswa berdasarkan capaian prestasi. Dosen masih menggunakan pemetaan secara manual tanpa metode yang terukur. Perlu adanya pengembangan aplikasi yang dapat membantu dosen mengelompokkan mahasiswanya.

Terdapat berbagai metode yang bisa dipergunakan untuk membantu proses pengelompokan mahasiswa. Salah satu dari metode tersebut adalah K-Means Clustering. Metode K-Means merupakan salah satu metode dalam fungsi pengelompokan. Clustering mengacu pada pengelompokan data, observasi atau kasus berdasar tingkat kemiripan objek yang diteliti. Sebuah cluster merupakan suatu kumpulan data yang mirip dengan lainnya atau ketidakmiripan data pada kelompok lain [5]. Penerapan metode K-Means Clustering dijalankan dengan kriteria pengelompokan berdasarkan nilai tugas, UTS dan UAS dari mahasiswa dalam kasus penelitian ini.

Penelitian yang akan dilakukan berjudul “Pengelompokan Mahasiswa Berdasarkan Pencapaian Prestasi Belajar dari Mata Kuliah yang Ditempuh Berbasis Web dengan K-Means Clustering”. Dengan adanya aplikasi dari penelitian ini diharapkan dapat digunakan oleh dosen untuk mengkomposisikan kelompok belajar yang seimbang sehingga membantu mahasiswa yang tertinggal untuk memperbaiki capaian prestasinya. Hasil akhir dari penelitian ini akan dikemas dalam aplikasi berbasis web yang dapat digunakan oleh dosen dalam membuat clustering kategori mahasiswanya.

II. KAJIAN PUSTAKA

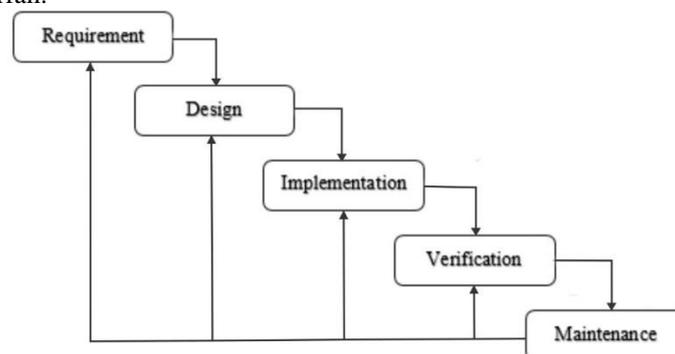
Penelitian Muningsih dan Kiswati pada tahun 2015 yang berjudul “Penerapan Metode K-Means Untuk Clustering Produk Online Shop Dalam Penentuan Stok Barang”. Penelitian tersebut bertujuan untuk mengelompokkan produk yang dijual pada online shop Ragam Jogja menjadi beberapa cluster untuk mengetahui produk mana yang paling diminati sehingga jumlah stok harus banyak, produk diminati untuk jumlah stok sedang dan produk kurang diminati untuk jumlah stok sedikit. Hasil dari penelitian tersebut adalah sebagai berikut. Pertama, berdasarkan pengolahan data yang dilakukan dengan atribut kode produk, jumlah transaksi, volume penjualan dan rata-rata penjualan, dihasilkan 3 kelompok produk yang paling diminati berjumlah 3 produk untuk jumlah stok banyak, 11 produk diminati untuk jumlah stok sedang dan 17 produk kurang diminati untuk jumlah stok sedikit. Kedua, penelitian ini bisa dimanfaatkan juga untuk pelaku bisnis lainnya bukan hanya online shop yang membutuhkan informasi penentuan stok

barang secara lebih cepat dan akurat. Ketiga, dengan adanya program aplikasi untuk penentuan stok barang diharapkan bisa memberikan kontribusi lebih baik lagi karena aplikasi yang dihasilkan bisa menampilkan kategori produk dalam jangka waktu tertentu sesuai yang diinginkan [5].

Penelitian Narwati pada tahun 2012 yang berjudul “Pengelompokan Mahasiswa Menggunakan Algoritma K-Means”. Penelitian tersebut bertujuan mengelompokkan mahasiswa berdasarkan data akademik menggunakan teknik clustering dan membuat aplikasinya kemudian menganalisis hasilnya sehingga diharapkan mampu memberikan informasi bagi yang berkepentingan. Hasil dari penelitian tersebut adalah sebagai berikut. Pola yang dihasilkan dari prestasi mahasiswa yang klusternya tetap, turun atau naik. pola mahasiswa tersebut dapat diketahui melalui pengamatan asal program studi, asal kota dan asal SMA. Diketahui juga hasil kluster untuk nilai tes prosentase dengan kluster terbesar adalah mahasiswa yang masuk dalam kluster 3, maka dapat diartikan bahwa sebagian mahasiswa yang masuk ke Universitas dapat di kategorikan sebagai mahasiswa yang mempunyai kemampuan yang tinggi berdasarkan hasil tes. Selain itu setelah mahasiswa menempuh kuliah 6 semester untuk program D3 dan 8 semester untuk program S1, mahasiswa yang berada di kluster 3 tetap memiliki jumlah prosentase terbanyak, tetapi jumlah ini tidak semua merupakan mahasiswa yang pada saat masuk berada pada kluster ini, karena terdapat mahasiswa dengan prestasinya yang menurun dan ada juga yang naik. Diketahui pula mahasiswa yang tetap pada klusternya, terutama kluster 3 jika ditinjau melalui program studinya, Komputerisasi Akuntansi (KA) berada di urutan pertama untuk mahasiswa dengan kemampuan tersebut tetap di kluster tiga, hal ini bisa dikatakan mereka mempunyai potensi dan kecerdasan yang tinggi. Untuk jenis SMA, prosentase tertinggi adalah mahasiswa yang berasal dari SMA Umum yang tetap bertahan di kluster 3. Sedangkan untuk asal SMA masing masing wilayah mempunyai persentase yang sama. Diketahui juga mahasiswa yang meningkat klusternya berasal dari kluster 1 ke kluster 3 mempunyai persentase kecil, penyumbang kenaikan jumlah persentase terbanyak adalah dari kluster 2 ke kluster 3. Maka dapat diartikan bahwa proses belajar mengajar bagi mahasiswa dengan kemampuannya berada pada kluster 1 harus lebih giat lagi. Diketahui juga mahasiswa yang turun kluster, paling tinggi persentasenya adalah dari kluster 3 ke kluster 2. Persentase tertinggi untuk prodi adalah TekKom yang mendapatkan penurunan persentase terbanyak sebesar 40% dari seluruh mahasiswa yang ada [6].

III. METODOLOGI PENELITIAN

Pada bab ini membahas mengenai proses perancangan aplikasi web yang dapat mengelompokkan tingkat pemahaman mahasiswa terhadap mata kuliah yang ditempuh dengan k-means clustering. metode pengembangan Research and Development (R&D). Peneliti mengembangkan produk berupa web aplikasi dengan menggunakan proses pengembangan yang mengacu pada tahapan model pengembangan perangkat lunak waterfall.



Gambar 1. Ilustrasi model waterfall

A. SIMULASI PENELITIAN

Proses penelitian menggunakan dataset yang didapat dari dosen pengampu mata kuliah Sistem Basis Data dan Kecerdasan Buatan. Data yang dibutuhkan adalah daftar nilai capaian prestasi mahasiswa UNISBA Blitar. Isi data yang dibutuhkan mencakup nilai rata-rata Tugas, nilai UTS dan nilai UAS yang diperoleh mahasiswa. Data set yang akan digunakan adalah tabel 1 dan tabel 2.

Tabel 1. Dataset kecerdasan buatan

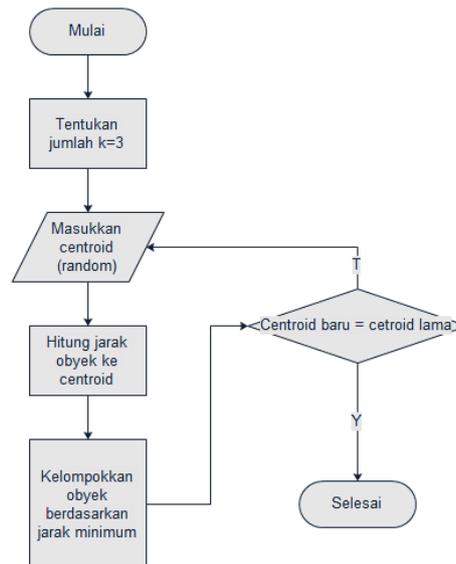
KODE	TUGAS	UTS	UAS
DB001	72,5	95,0	85,0
DB002	95,0	85,0	90,0
DB003	100,0	95,0	85,0
DB004	64,0	85,0	95,0
DB005	95,5	90,0	90,0
DB006	73,5	85,0	90,0
DB007	81,5	100,0	90,0
DB008	72,5	80,0	95,0
DB009	100,0	95,0	85,0
DB010	71,5	80,0	85,0
DB011	46,0	90,0	80,0
DB012	84,0	95,0	90,0
DB014	56,5	55,0	90,0
DB016	56,5	65,0	50,0
DB017	65,0	80,0	50,0
DB018	73,5	65,0	90,0
DB019	72,5	90,0	90,0
DB020	97,5	75,0	80,0
DB022	70,0	85,0	75,0
DB024	71,5	80,0	85,0
DB025	97,5	85,0	90,0
DB026	73,5	100,0	50,0
DB027	81,5	100,0	50,0
DB028	71,5	85,0	85,0
DB029	77,5	90,0	75,0
DB030	54,0	85,0	80,0
DB031	81,0	90,0	80,0
DB033	84,0	85,0	85,0
DB034	70,0	77,0	85,0
DB035	42,5	95,0	85,0
DB036	53,5	85,0	85,0
DB037	91,5	85,0	90,0
DB038	78,5	90,0	85,0
DB039	40,0	55,0	85,0
DB041	65,0	80,0	90,0
DB043	78,0	75,0	80,0
DB044	40,0	85,0	75,0
DB045	56,5	100,0	50,0
DB046	46,5	100,0	85,0
DB047	86,5	100,0	85,0
DB048	66,5	70,0	50,0
DB049	69,5	90,0	90,0
DB050	78,5	70,0	80,0
DB051	64,0	70,0	85,0
DB052	70,0	100,0	80,0
DB053	57,0	85,0	85,0
DB054	75,0	90,0	50,0
DB055	91,5	90,0	75,0
DB056	94,5	75,0	95,0
DB057	56,0	100,0	95,0
DB058	100,0	80,0	95,0
DB059	72,5	50,0	90,0
DB060	72,5	100,0	90,0
DB062	100,0	65,0	90,0
DB064	83,5	90,0	75,0
DB065	91,5	95,0	50,0
DB066	70,0	75,0	80,0
DB067	90,0	95,0	85,0
DB068	78,5	70,0	90,0

KODE	TUGAS	UTS	UAS
DB069	83,5	80,0	75,0

Tabel 2. Dataset sistem basis data

KODE	TUGAS	UTS	UAS
BK001	79,0	80,0	75,0
BK002	100,0	80,0	85,0
BK003	71,0	80,0	75,0
BK004	86,0	80,0	75,0
BK005	93,0	80,0	85,0
BK006	100,0	80,0	75,0
BK007	93,0	80,0	75,0
BK008	93,0	80,0	85,0
BK009	71,0	80,0	75,0
BK010	36,0	80,0	75,0
BK011	86,0	80,0	75,0
BK012	43,0	80,0	75,0
BK013	100,0	80,0	85,0
BK014	100,0	80,0	75,0
BK015	100,0	80,0	75,0
BK016	86,0	80,0	75,0
BK017	21,0	80,0	85,0
BK018	100,0	80,0	75,0
BK019	57,0	80,0	75,0
BK020	50,0	80,0	75,0
BK021	80,0	80,0	85,0
BK022	80,0	80,0	75,0
BK023	80,0	80,0	85,0
BK024	80,0	80,0	80,0
BK025	72,5	80,0	75,0
BK026	72,5	80,0	85,0
BK027	80,0	80,0	75,0
BK028	80,0	80,0	75,0
BK029	80,0	80,0	80,0
BK030	80,0	80,0	75,0
BK031	80,0	80,0	85,0
BK032	72,5	80,0	75,0
BK033	72,5	80,0	85,0
BK034	80,0	80,0	80,0
BK035	72,5	80,0	80,0
BK036	72,5	80,0	80,0
BK037	80,0	80,0	80,0
BK038	80,0	80,0	75,0
BK039	80,0	80,0	80,0
BK040	80,0	80,0	75,0
BK041	72,5	80,0	80,0
BK042	80,0	80,0	75,0

Pada penelitian ini data yang dimiliki selanjutnya dimasukkan dalam format xlsx untuk diunggah ke aplikasi web. Aplikasi ini telah dirancang untuk dapat melakukan clustering dengan metode k-means. Data yang diolah akan di kelompok kan menjadi 3 kelas diantaranya C1 (Advanced), C2 (Proficient) dan C3 (Basic). Dengan flowchart clustering k-means seperti gambar 3 berikut.



Gambar 3. Flowchart k-means

B. SKENARIO PENGUJIAN

Pada penelitian ini penelitian dilakukan dengan memasukkan dataset ke web aplikasi yang telah ditentukan k=3. Dari data yang telah dimiliki dilakukan pencarian centroid secara acak selanjutnya menghitung jarak obyek ke centroid. Obyek dikelompokkan berdasarkan jarak minimumnya. Centroid baru dihasilkan dari rata-rata jarak minimum, dilanjutkan dengan membandingkan centroid baru dengan centroid lama. Jika centroid lama dan centroid baru tidak sama maka kembali melakukan randomize, namun jika sama maka selesai.

C. PARAMETER ANALISIS DATA

Pada penelitian ini pengelompokan mahasiswa berdasarkan tingkat pencapaian prestasi belajar dengan k-means clustering terdapat 2 parameter yang dapat diamati, diantaranya:

1. Cluster

Cluster merupakan hasil dari pengolahan dataset yang sudah terkelompokkan dengan cara menentukan posisi centroid baru dengan cara menghitung nilai rata-rata dari data-data yang ada pada centroid yang sama.

$$C_i = \left(\frac{\sum x}{n} \right) \quad (1)$$

2. Jarak Cluster

Jarak cluster adalah nilai koefisien dari data dengan centroid oleh karena itu, yang diharapkan adalah koefisien menjadi sebesar mungkin dan mendekati 1 untuk memiliki kelompok yang baik atau berjarak kontras.

$$\frac{b^i - a^i}{\max(a^i, b^i)} \quad (2)$$

Dengan:

a = titik data dalam cluster yang sama

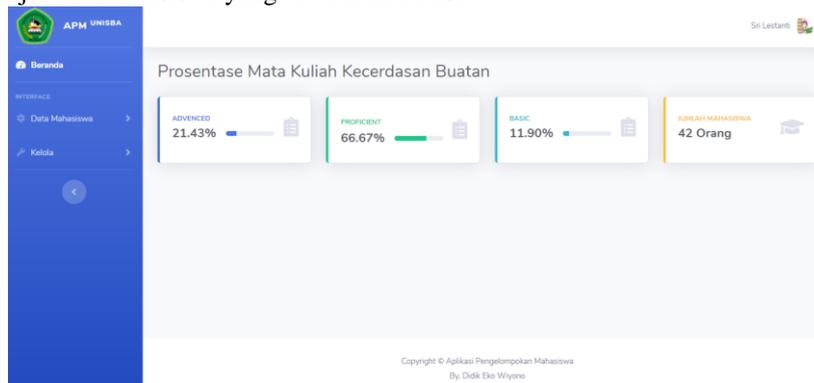
b = jarak rata-rata dari semua titik data di kluster terdekat

IV. HASIL DAN ANALISA

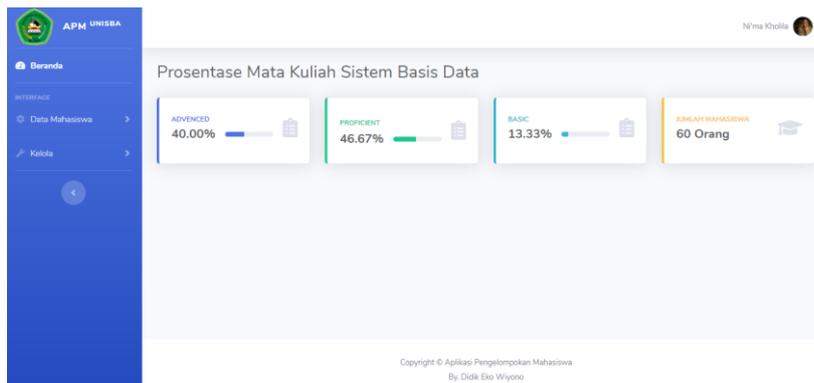
Penelitian ini bertujuan untuk mengetahui parameter cluster dan jarak cluster dari hasil pengelompokan mahasiswa berdasarkan tingkat pencapaian prestasi belajar dengan k-means. Pengujian dilakukan menggunakan 2 skenario dan 2 variasi dataset. Mengolah dataset menggunakan web aplikasi dan menggunakan metode silhouette analysis yang dijalankan pada program jupyter untuk menghitung nilai koefisien jarak clustering dari dataset yang digunakan dalam penelitian. Dataset mata kuliah Sistem Cerdas sebanyak 60 mahasiswa dan Sistem Basis Data sebanyak 42 mahasiswa.

A. Analisis Cluster

Berdasarkan penelitian yang dilakukan dengan 2 skenario dari 2 variasi dataset yang dimiliki menghasilkan nilai parameter salah satunya cluster. Hasil cluster dapat dilihat pada gambar 4 dan gambar 5. Dimana pada gambar tersebut didapatkan prosentase mahasiswa yang termasuk cluster C1, C2 dan C3 dari total jumlah mahasiswa yang ada dalam dataset.



Gambar 4. Cluster matakuliah kecerdasan buatan

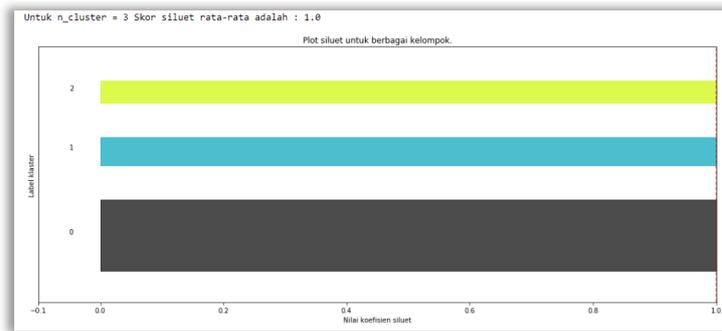


Gambar 5. Cluster mata kuliah sistem basis data

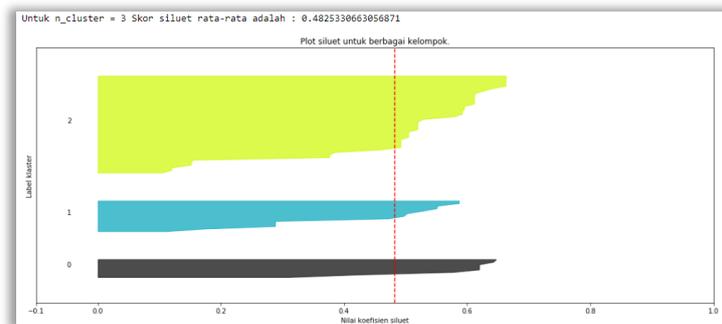
Pada kedua gambar diatas mempresentasikan hasil prosentase cluster dataset yang termasuk dalam kategori C1 (Advenced), C2 (Proficient) dan C3 (Basic) dari total mahasiswa yang ada.

B. Analisis Jarak Cluster

Berdasarkan penelitian yang dilakukan dengan 2 skenario dari 2 variasi dataset yang dimiliki menghasilkan nilai parameter salah satunya jarak cluster. Hasil penghitungan jarak luster dapat dilihat pada gambar 6 dan gambar 7. Dimana pada gambar tersebut nilai koefisien yang mempresentasikan kualitas dari pemisahan cluster yang terjadi dari dataset yang dimiliki.



Gambar 6. Shiluite dan nilai koefisien mata kuliah kecerdasan buatan



Gambar 7. Shiluite dan nilai koefisien mata kuliah sistem basis data

Gambar 6 mempresentasikan hasil pengujian menunjukkan skor silhouette sebesar 1 sehingga $1 : 1 \times 100\% = 100\%$. Sedangkan gambar 7 mempresentasikan hasil pengujian menunjukkan skor silhouette sebesar 0,48 sehingga $0,48 : 1 \times 100\% = 48\%$. Selanjutnya dalam tampilan diagram C1 ditandai dengan warna hitam, C2 ditandai dengan warna biru dan C3 ditandai dengan warna hijau.

V. KESIMPULAN

Berdasarkan penelitian pengelompokan tingkat pemahaman mahasiswa terhadap mata kuliah yang ditempuh dengan k-means clustering berbasis web dapat disimpulkan hasil sebagai berikut:

1. Untuk menjalankan aplikasi pengelompokan mahasiswa berdasarkan pencapaian prestasi belajar dari mata kuliah yang ditempuh berbasis Web dengan K-Means Clustering adalah dengan menyiapkan dataset yang diperlukan untuk diolah.
2. Hasil pengujian akurasi terhadap pengelompokan mahasiswa berdasarkan pencapaian prestasi belajar dari mata kuliah yang ditempuh berbasis Web dengan K-Means Clustering menggunakan Silhouette analysis dari dataset yang dimiliki adalah 100% untuk mata kuliah Kecerdasan Buatan. Dapat dikatakan tingkat perbedaan capaian nilai untuk untuk mata kuliah Kecerdasan Buatan oleh mahasiswa kontras, sehingga dapat diterapkan metode tertentu untuk memperbaiki keadaan mahasiswanya. Selanjutnya untuk mata kuliah Sistem Basis Data adalah 48%. Dapat dikatakan tingkat pemisahannya untuk mata kuliah Sistem Basis Data tidak begitu terpaut jauh atau juga dapat dikatakan capaian nilai yang diperoleh mahasiswa lebih merata.

REFERENSI

- [1] Slameto, *Belajar dan Faktor-faktor yang Mempengaruhinya*, Rev. Jakarta: PT Rineka Cipta, 2010.
- [2] S. Nasution, *Asas-asas Kurikulum*. Bandung: Jemmars, 1982.
- [3] R. M. Gagne, *The Condition of Learning Theory of Instruction*. New York: Rinehart, 1985.
- [4] S. Arikunto, *Dasar-dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara, 1990.
- [5] E. Muningsih and S. Kiswati, "Penerapan Metode K-Means untuk Clustering Produk Online Shop dalam Penentuan Stock Barang," *Bianglala Inform.*, vol. 3, no. 1, 2015.
- [6] N. Narwati, "Pengelompokan Mahasiswa Menggunakan Algoritma K-Means," *J. Din. Inform.*, vol. 2, no. 2, 2010.