

## Optimasi Jumlah *Cluster* pada Algoritme K-Means untuk Evaluasi Kinerja Dosen

Vega Purwayoga

Teknik Informatika, Universitas Muhammadiyah Cirebon, Jl. Fatahillah, Watubelah, Sumber, Cirebon, Indonesia, 45611

e-mail: vega.purwayoga@umc.ac.id

Submitted Date: February 06<sup>th</sup>, 2020  
Revised Date: June 04<sup>th</sup>, 2021

Reviewed Date: June 02<sup>nd</sup>, 2020  
Accepted Date: June 15<sup>th</sup>, 2021

### Abstract

*Lecturers are one of the main actors in universities. Lecturer performance can affect the quality of a college. Because the quality of higher education is strongly influenced by the lecturers, the performance of the lecturers needs to be assessed. Lecturer performance can be assessed by evaluating the lecturer's performance in teaching. Lecturer performance can be evaluated by classifying student assessments of lecturers. Lecturers are assessed based on how the lecturer mastered the material, the lecturer's discipline in teaching, and the presentation of the material. The process of grouping lecturer scores can be done using the K-Means algorithm. K-Means is a popular clustering algorithm that performs well. K-Means requires a parameter that is K or the number of clusters. The importance of the number of clusters so that there is a need for optimization in determining the number of K. In this study, optimization was carried out using the Elbow method so as to produce the ideal number of groups of 4 groups. The results of the clustering evaluation calculated using the SSE were 54.4% which showed that the results were not optimal. The results of the cluster evaluation are not optimal due to the lack of data for the K-Means application.*

*Keywords: Elbow Method; K-Means; Lecturer performance evaluation; Optimization*

### Abstrak

Dosen merupakan salah satu aktor utama yang berada pada perguruan tinggi. Kinerja dosen dapat mempengaruhi kualitas suatu perguruan tinggi. Dikarenakan kualitas perguruan tinggi sangat dipengaruhi oleh dosen maka kinerja dosen perlu dinilai. Kinerja dosen dapat dinilai dengan cara mengevaluasi kinerja dosen dalam mengajar. Kinerja dosen dapat dievaluasi dengan mengklasifikasikan penilaian mahasiswa terhadap dosen. Dosen dinilai berdasarkan bagaimana dosen menguasai materi, kedisiplinan dosen dalam mengajar, dan penyajian materi. Proses pengelompokan nilai dosen dapat dilakukan dengan menggunakan algoritma K-Means. K-Means merupakan algoritme clustering populer yang kinerjanya baik. K-Means membutuhkan parameter yaitu K atau jumlah cluster. Pentingnya jumlah cluster sehingga perlu adanya optimasi dalam menentukan jumlah K. Pada penelitian ini dilakukan optimasi dengan menggunakan metode Elbow sehingga menghasilkan jumlah kelompok yang ideal sebanyak 4 kelompok. Hasil evaluasi clustering yang dihitung menggunakan SSE adalah 54,4% yang menunjukkan hasil yang belum optimal. Hasil evaluasi klaster belum optimal karena kurangnya data untuk aplikasi K-Means.

Kata kunci: Evaluasi kinerja dosen; K-Means; Metode Elbow; Optimasi cluster

### 1. Introduction

Evaluasi kinerja sumber daya manusia merupakan salah satu bagian penting dalam suatu organisasi atau perusahaan. Salah satu SDM pada suatu instansi akademik atau kampus yang harus dievaluasi adalah dosen. Dosen merupakan salah

satu faktor utama dalam perguruan tinggi (Purwayoga, 2020). Dosen menjadi faktor utama dikarenakan dosen merupakan objek yang berkomunikasi langsung dengan mahasiswa. Dosen dikatakan baik ketika dosen tersebut dapat mengajarkan dan memotivasi mahasiswa dengan

baik juga (Widaningrum, 2012). Dosen yang baik merupakan salah satu faktor utama untuk mendukung perguruan tinggi yang berkualitas (Hartiti, 2019). Dosen yang salah satu tugas utamanya sebagai aktor proses pembelajaran, dituntut untuk memiliki kualitas pembelajaran yang baik. Salah satu bentuk usaha untuk menilai cara dosen belajar itu baik atau tidak yaitu dengan mengevaluasi kinerja dosen (Hartiti, 2019).

Evaluasi dosen sendiri dapat dilakukan oleh mahasiswa, karena mahasiswa adalah aktor yang berkomunikasi langsung dengan seorang dosen. Mahasiswa dapat menilai dosen dari cara dia mengajar, menguasai materi dan mencontohkan kedisiplinan mahasiswa (Widaningrum, 2012).

Hasil evaluasi dosen dapat dilihat dari penilaian mahasiswa yang menilai baik atau menilai buruk (Seniwati, 2011). Penilaian tersebut dapat dilakukan dengan cara mengelompokkan mahasiswa yang menilai baik atau buruk (Safira, 2016). Pengelompokan Mahasiswa dapat dilihat dari nilai evaluasi mahasiswa terhadap seorang dosen pada setiap akhir semester (Safira, 2016).

Pengelompokan mahasiswa untuk evaluasi kinerja dosen dapat dilakukan dengan teknik data mining dengan memanfaatkan data yang telah tersedia (Purwayoga, 2020). Metode yang sering

digunakan dan performanya baik dalam data mining salah satunya adalah K-Means clustering (Andri, 2015).

Penelitian terkait pengelompokan mahasiswa telah dilakukan oleh (Purwayoga, 2020), namun masih terdapat kekurangan. Kekurangan dalam penelitian sebelumnya yaitu belum melakukan optimasi jumlah kelompok yang dibentuk oleh metode clustering. Optimasi perlu dilakukan untuk menentukan jumlah cluster yang terbaik sehingga menghasilkan kualitas cluster yang baik juga (Muningsih, 2018). Penelitian ini dilakukan untuk menerapkan dan mengoptimasi algoritme K-Means untuk evaluasi penilaian dosen agar mendapat cluster yang optimal

## 2. Metode Penelitian

### 2.1 Pengumpulan Data

Data yang digunakan penelitian ini didapat dari penelitian (Seniwati, 2011). Terdapat 5 faktor/atribut yang digunakan untuk menilai kinerja dosen dalam proses belajar mengajar (Seniwati, 2011). Masing-masing faktor memiliki nilai yang sama dari mulai nilai = kurang sampai dengan nilai = sangat baik sebagaimana yang telah disajikan pada penelitian.

Table 1. Hasil Transformasi

Kode	Atribut	Keterangan	Nilai
A	Penguasaan materi	Penguasaan dosen terkait materi yang disampaikan kepada mahasiswa	1. Kurang 2. Cukup 3. Baik 4. Sangat Baik
B	Penyajian materi	Penyajian materi dilakukan secara berurutan	1. Kurang 2. Cukup 3. Baik 4. Sangat baik
C	Kedisiplinan dosen	Kepatuhan dosen terhadap waktu yang telah disepakati	1. Kurang 2. Cukup 3. Baik 4. Sangat baik
D	Meningkatkan minat mahasiswa	Pengaruh gaya dosen dalam mengajar terhadap minat belajar mahasiswa	1. Kurang 2. Cukup 3. Baik 4. Sangat baik
E	Penguasaan kelas	Mengendalikan mahasiswa untuk tetap fokus kepada materi atau dosen	1. Kurang 2. Cukup 3. Baik 4. Sangat baik

### 2.2 Transformasi Data

Transformasi data dalam penelitian ini menghasilkan nilai hasil konversi dari nilai

variabel A, B, C, D dan E menjadi numerik. Perubahan nilai variabel menjadi numerik bertujuan untuk mempermudah proses atau

tahapan selanjutnya khususnya dalam pengukuran jarak antar sebuah objek (Agarwal, 2016). Hasil transformasi yang dihasilkan sama dengan hasil yang digunakan oleh penelitian (Purwayoga, 2020).

### 2.3 Penerapan K-Means

Algoritme K-Means memiliki tahapan sebagaimana berikut:

- Menentukan jumlah *cluster* (K)  
 Penentuan *cluster* dilakukan dengan menggunakan metode *elbow*. Dimana *elbow* menghasilkan nilai K yang terbaik. Perhitungan metode *elbow* didapatkan dari nilai *Sum of Squares Error (SSE)* yang mengalami penurunan secara signifikan (Rahman, Dewi, & Andini, 2017).
- Memilih *centroid*  
 Pemilihan *centroid* atau titik pusat *cluster* dapat dilakukan secara acak / random.
- Menghitung jarak antara data dengan titik pusat cluster (titik pusat diambil secara acak) Proses pengukuran jarak dilakukan dengan menggunakan *euclidean distance*. *Euclidean distance* merupakan metode pengukuran jarak yang populer (Nishom, 2019).

$$dist(i, j) = \sqrt{(x_{i1} - x_{j1})^2 + \dots + (x_{in} - x_{jn})^2}$$

Dimana,  $x =$  objek,  $i = (x_{i1}, x_{i2}, \dots, x_{in})$ ,  $j = (x_{j1}, x_{j2}, \dots, x_{jn})$  adalah data objek dua dimensi.

- Memasukkan data pada *cluster* berdasarkan jarak minimum antara data dengan *centroid*.
- Lakukan langkah ke dua hingga *centroid* tidak berubah (Singh, Yadav, & Rana., 2013)(Putu & Merliana, 2015).

### 2.4 Evaluasi Hasil Pengelompokan Kinerja Dosen

Evaluasi *cluster* bertujuan untuk mengetahui kualitas *cluster* yang telah terbentuk. Metode yang digunakan untuk evaluasi cluster dalam penelitian ini adalah *Sum of Squares Error (SSE)* (Rahman, Dewi, & Andini, 2017).

$$SSE = \sum_{K=1}^K \sum_{x_i \in S_K} \|X_i - C_K\|_2^2$$

Di mana K = jumlah cluster, C = titik tengah cluster, X = data yang berada pada setiap cluster (Rahman et al., 2017).

### 2.5 Pengkategorian Kinerja Dosen

Pengkategorian dilakukan untuk membentuk suatu label apakah penilaian mahasiswa terhadap dosen itu kurang, cukup, baik atau sangat baik (Purwayoga, 2020). Kategori didapat dari nilai rata-rata setiap *cluster*.

## 3. Hasil dan Pembahasan

### 3.1 Pengumpulan Data

Pengumpulan data menghasilkan nilai dari setiap faktor penilaian kinerja dosen seperti penguasaan materi, penyajian materi, kedisiplinan dosen, apakah dosen meningkat minat belajar siswa dan penguasaan kelas.

### 3.2 Transformasi Data

Transformasi data dalam penelitian ini menghasilkan nilai hasil perubahan dari nilai variabel A, B, C, D dan E menjadi numerik untuk proses pengukuran jarak antar sebuah objek. Hasil transformasi yang dihasilkan sama dengan hasil yang digunakan oleh penelitian (Purwayoga, 2020). Kurang = 1, Cukup = 2, Baik = 3, Sangat Baik = 4. Hasil Transformasi telah disajikan pada Tabel 1.

Table 2. Hasil Transformasi

No Kasus	Variabel				
	A	B	C	D	E
1	1	2	3	3	1
2	3	3	1	2	4
3	3	3	2	1	1
4	2	3	3	3	3
5	3	1	1	1	3
6	4	4	3	2	3
7	4	4	3	1	1
8	1	2	4	4	4
9	3	3	3	3	2
10	1	3	3	2	2

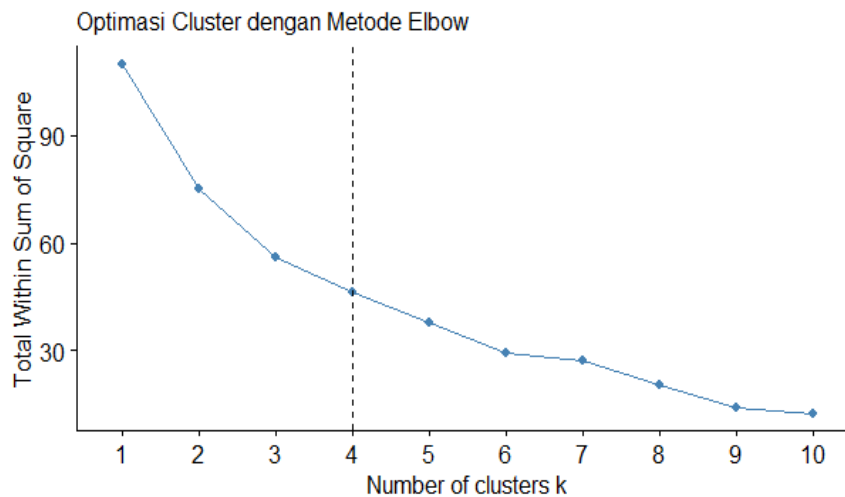
11	1	4	4	4	3
12	1	2	4	4	4
13	1	1	1	4	3
14	4	1	4	4	3
15	3	2	3	2	3
16	2	3	3	3	3
17	3	1	1	1	2
18	2	1	2	2	3
19	3	4	2	2	2
20	4	4	3	2	3

### 3.3 Penerapan Algoritme K-Means

Penerapan Algoritme K-Means untuk mengelompokkan data penilaian evaluasi dosen adalah sebagai berikut:

- Tahapan pertama yaitu menentukan jumlah *cluster* (K). Jumlah *cluster* dilakukan dengan menggunakan Metode *Elbow* dengan formula

yang telah dijelaskan pada penelitian (Muningsih, 2018). Jumlah *cluster* yang direkomendasikan oleh Metode *Elbow* adalah 4, sesuai dengan penelitian yang dilakukan oleh (Purwayoga & Cirebon, 2020). Visualisasi metode *Elbow* telah disajikan pada Gambar 1.



Gambar 1. Example of figure caption

- Setelah jumlah *cluster* ditentukan maka proses selanjutnya adalah memilih *centroid* masing-masing *cluster*. *Centroid* pada awal mulanya dilakukan secara random, lalu dievaluasi

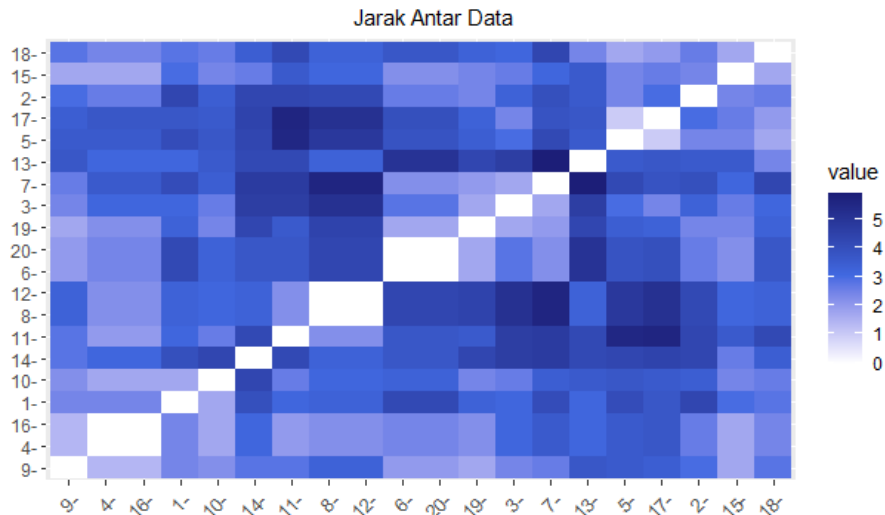
sampai dengan nilai *centroid* tidak berbeda signifikan dengan nilai *centroid* sebelumnya. Nilai *centroid* dapat dilihat pada Tabel 2.

Table 2. Nilai centroid

<i>Centroid</i>	A	B	C	D
1	2.25	1	1.25	2
2	3.33	3.33	2.50	2.17
3	1.63	2.50	3.50	3.38
4	3.50	3.50	2.50	1

- Centroid* yang telah ditentukan digunakan untuk menghitung jarak antara data dengan *centroid* tersebut. Visualisasi proses

pengukuran jarak dapat dilihat pada Gambar 2.



Gambar 2. Example of figure caption

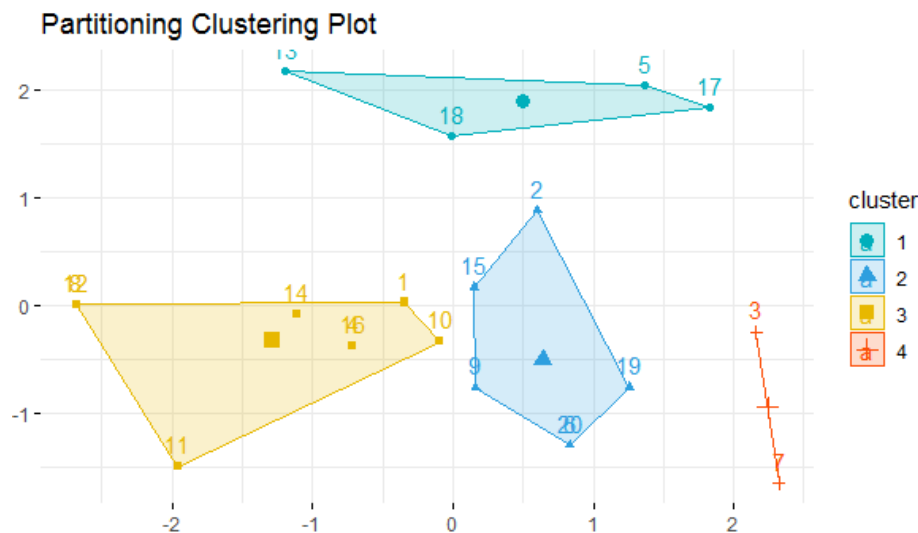
- Hasil pengukuran jarak menentukan suatu objek masuk ke dalam suatu *cluster*. Ketika objek dan *centroid cluster* memiliki jarak minimum, maka objek tersebut masuk ke dalam *cluster* tersebut. Hasil dari proses *clustering* telah disajikan pada Tabel 3. Penelitian (Purwayoga, 2020) menentukan jumlah *cluster* yang sama yaitu 4, namun anggota *cluster* yang terbentuk berbeda. Perbedaan anggota *cluster* dapat disebabkan oleh data yang digunakan (Rahmansyah et al., 2018). Pada penelitian (Purwayoga, 2020) data yang digunakan untuk proses *clustering* adalah jumlah total bobot pada seluruh atribut, sedangkan pada penelitian ini tidak

- melakukan proses penjumlahan terlebih dahulu. Penelitian ini diharapkan dapat menunjukkan karakteristik data yang terdapat pada suatu *cluster*. Karakteristik yang dimaksud adalah nilai atribut dari setiap data.
- Visualisasi hasil pengelompokan dapat dilihat pada Gambar 3. Visualisasi menggunakan library *factoextra* yang terdapat pada bahasa pemrograman R. Hasil visualisasi merangkum objek atau data yang berada pada *cluster*. Anggota *cluster* pada *cluster* 1 sampai 4 berturut-turut berjumlah 4, 6, 8 dan 2. Simbol lingkaran untuk *cluster* 1, segitiga untuk *cluster* 2, kotak untuk *cluster* 3 dan + untuk *cluster* 4. Hasil pengelompokan dapat dilihat pada Tabel 3.

Table 3. Hasil pengelompokan

No Kasus	Variabel					Cluster
	A	B	C	D	E	
1	1	2	3	3	1	3
2	3	3	1	2	4	2
3	3	3	2	1	1	4
4	2	3	3	3	3	3
5	3	1	1	1	3	1
6	4	4	3	2	3	2
7	4	4	3	1	1	4
8	1	2	4	4	4	3
9	3	3	3	3	2	2
10	1	3	3	2	2	3
11	1	4	4	4	3	3
12	1	2	4	4	4	3
13	1	1	1	4	3	1
14	4	1	4	4	3	3
15	3	2	3	2	3	2
16	2	3	3	3	3	3
17	3	1	1	1	2	1

18	2	1	2	2	3	1
19	3	4	2	2	2	2
20	4	4	3	2	3	2



Gambar 3. Example of figure caption

### 3.3 Evaluasi Hasil Pengelompokan

Evaluasi hasil pengelompokan dilakukan untuk mengetahui kualitas hasil *cluster* yang terbentuk dalam pengelompokan penilaian dosen. Evaluasi pengelompokan dilakukan dengan menggunakan SSE. Nilai SSE yang didapat pada penelitian ini yaitu sebesar 54.4 %. Perbedaan nilai SSE yang signifikan antara penelitian ini dengan penelitian (Purwayoga, 2020) dikarenakan pada penelitian tersebut tidak menggunakan 5 atribut, tetapi hanya menggunakan satu atribut yaitu jumlah bobot seluruh atribut. Perbedaan atribut yang digunakan dapat mempengaruhi hasil pengelompokan. Dapat dibandingkan bahwa semakin banyak atribut, maka semakin kompleks proses pengukuran jarak antara objek dengan *centroid* (Widystuti & Darmawan, 2018).

### 3.3 Evaluasi Hasil Pengelompokan

Pengkategorian ditentukan menjadi kelompok atau kategori dimana terdapat kategori kurang, cukup, baik dan sangat baik. Kategori ditentukan berdasarkan nilai rata-rata dari setiap *centroid* dengan aturan sebagaimana berikut

- Jika suatu rata-rata nilai *centroid* bernilai 1 atau mendekati satu, maka seluruh anggota yang berada pada suatu *cluster* dengan rata-rata *centroid* tersebut berkategori kurang.

- Jika suatu rata-rata nilai *centroid* bernilai 2 atau mendekati dua, maka seluruh anggota yang berada pada suatu *cluster* dengan rata-rata *centroid* tersebut berkategori cukup.
- Jika suatu rata-rata nilai *centroid* bernilai 3 atau mendekati tiga, maka seluruh anggota yang berada pada suatu *cluster* dengan rata-rata *centroid* tersebut berkategori baik.
- Jika suatu rata-rata nilai *centroid* bernilai 4 atau mendekati empat, maka seluruh anggota yang berada pada suatu *cluster* dengan rata-rata *centroid* tersebut berkategori sangat baik.

Penentuan kategori dapat dilihat pada Tabel 4. *Centroid* 1 dan *centroid* 4 yang merupakan titik pusat pada *cluster* 1 dan 4 masuk ke dalam kategori 2 yaitu, kategori cukup. Anggota yang masuk dalam *cluster* 2 dan *cluster* 3 masuk pada kategori tiga sebagaimana yang telah disajikan pada Tabel 4.

Hasil dari pengkategorian dapat dilihat pada Tabel 5. Jumlah yang masuk ke dalam kategori cukup berjumlah 6 kasus dan kasus yang masuk ke dalam kategori baik berjumlah 14 kasus. Tidak adanya data atau kasus yang masuk kategori kurang dan sangat baik dikarenakan tidak ada nilai rata-rata *centroid* yang mendekati satu dan mendekati nilai 4.

Table 4. Hasil pengelompokan

<i>Centroid</i>	A	B	C	D	Rataan	Pembula
-----------------	---	---	---	---	--------	---------

						<b>tan</b>
1	2.25	1	1.25	2	1.85	2
2	3.33	3.33	2.50	2.17	2.83	3
3	1.63	2.50	3.50	3.38	2.775	3
4	3.50	3.50	2.50	1	2.3	2

Table 5. Hasil penkategorian

No Kasus	Variabel					Cluster	Kategori
	A	B	C	D	E		
1	1	2	3	3	1	3	Baik
2	3	3	1	2	4	2	Baik
3	3	3	2	1	1	4	Cukup
4	2	3	3	3	3	3	Baik
5	3	1	1	1	3	1	Cukup
6	4	4	3	2	3	2	Baik
7	4	4	3	1	1	4	Cukup
8	1	2	4	4	4	3	Baik
9	3	3	3	3	2	2	Baik
10	1	3	3	2	2	3	Baik
11	1	4	4	4	3	3	Baik
12	1	2	4	4	4	3	Baik
13	1	1	1	4	3	1	Cukup
14	4	1	4	4	3	3	Baik
15	3	2	3	2	3	2	Baik
16	2	3	3	3	3	3	Baik
17	3	1	1	1	2	1	Cukup
18	2	1	2	2	3	1	Cukup
19	3	4	2	2	2	2	Baik
20	4	4	3	2	3	2	Baik

#### 4. Kesimpulan

Penelitian ini berhasil menerapkan algoritme K-Means yang telah dioptimasi untuk mengelompokkan evaluasi penilaian dosen dari sejumlah kasus. Algoritme K-Means dioptimasi dengan menggunakan metode *elbow* yang menghasilkan jumlah *cluster* sebanyak 4. Walaupun jumlah *cluster* telah ditentukan dengan menggunakan metode *elbow*, belum tentu kualitas hasil clusternya baik. Dibuktikan bahwa nilai evaluasi pengelompokan yang dihitung dengan menggunakan SSE menghasilkan nilai 54.4 % yang berarti menunjukkan hasil yang kurang baik. Hasil yang kurang baik dapat dipengaruhi oleh data yang banyak memiliki kemiripan, sehingga K-Means tidak mudah dalam proses pengelompokan (Widystuti, 2018). Mayoritas mahasiswa menilai bahwa dosen yang dinilai berdasarkan data tersebut kinerjanya sudah baik. Penelitian selanjutnya diharapkan dapat menambahkan data penelitian untuk melihat pengaruh jumlah data pada proses *clustering*. Penelitian ini diharapkan dapat dikembangkan atau diterapkan pada suatu sistem informasi sebagai bahan evaluasi untuk penilaian kinerja dosen.

#### References

- Agarwal, V. (2016). Research on Data Preprocessing and Categorization Technique for Smartphone Review Analysis. November. *International Journal of Computer Applications*. 131(4):30-36. doi : <https://doi.org/10.5120/ijca2015907309>
- Andri & Marlindawati. (2015). Pengelompokan Minat Belajar Mahasiswa Menggunakan Teknik. *Jurnal Ilmiah Matrik*. 3(1):67-76
- Muningsih, E., & Kiswati, S. (2018). Sistem Aplikasi Berbasis Optimasi Metode Elbow Untuk. *JOUTICA*. 3(1):1-8.
- Nishom, M. (2019). Perbandingan Akurasi Euclidean Distance, Minkowski Distance, Dan Manhattan Distance Pada Algoritma K-Means Clustering Berbasis Chi-Square. *Jurnal Informatika: Jurnal Pengembangan IT*. 4(1):20-24. doi : <https://doi.org/10.30591/Jpit.V4i1.1253>
- Purwayoga, V. (2020). Implementasi Algoritme K-Means Untuk Evaluasi Kinerja Dosen Pada Rancangan Sistem Informasi Manajemen ( SIM ) Kampus. *INTI TALIFA : Jurnal Teknik Informatika*.12(1):1-6.

- doi:<https://doi.org/10.32534/Int.V12i01.1265>
- Purwayoga, V., & Sitanggang, I. S. (2020). Clustering Potential Area Of Fusarium Oxysporum As A Disease Of Garlic. *International Conference on Environment and Forest Conservation*. 528:012040
- Putu, N., Merliana, E. (2015). *Analisa Penentuan Jumlah Cluster Terbaik Pada Metode K-Means*. 978–979.
- Rahman, A. T., Wiranto, & Rini, A. (2017). Coal Trade Data Clustering Using K-Means (Case Study Pt. Global Bangkit Utama). *ITSMART: Jurnal Teknologi Dan Informasi*, 6(1):24–31. doi:<https://doi.org/10.20961/ITS.V6I1.11296>
- Rahmansyah, A., Dewi, O., Andini, P., Hastuti, T., Ningrum, P., & Suryana, M. E. (2018). Membandingkan Pengaruh Feature Selection Terhadap Algoritma Naïve Bayes Dan Support Vector Machine. 1–7.
- Retnowati, T. H. (2019). Model Evaluasi Kinerja Dosen : Pengembangan Instrumen untuk Mengevaluasi Kinerja. *Jurnal Penelitian dan Evaluasi Pendidikan*. 21(2):206-214.
- Safira, W. (2016). Pengelompokan Minat Baca Mahasiswa Menggunakan Metode K-Means. *ILKOM Jurnal Ilmiah*. 8(2):89–94.
- Seniwati, E. (2011). Perhitungan Penilaian Mahasiswa Terhadap Mengajar Dosen. *Jurnal DASI*. 12(2):1–5.
- Singh, A., Yadav, A., & Rana, A. (2013). K-Means With Three Different Distance Metrics. *International Journal Of Computer Applications*. 67(10):13–17. doi:<https://doi.org/10.5120/11430-6785>
- .Widaningrum, I. D. A. (2012). Sistem Informasi Evaluasi Kinerja Dosen Aspe Kompetensi dan Tridarma di Lingkungan Universitas Muammadiyah Ponorogo. 6(1), 49–60.
- Widystuti, W., & Darmawan, J. B. B. (2018). Pengaruh Jumlah Data Set Terhadap Akurasi Pengenalan Dalam Deep Convolutional Network. 8–13.