

Implementasi Metode K-Means Clustering Pada Analisis Motivasi Verbal Terhadap Semangat Belajar Mahasiswa

Welshy*, Indra Samsie, Santi

Teknik Informatika, Universitas Dipa Makassar, Makassar Sulawesi Selatan
e-mail: ^{*1}Welshypersum28@gmail.com, ²indra.samsie@undipa.ac.id, ³santi@undipa.ac.id

Abstrak

Tujuan dari penelitian ini adalah untuk mengetahui seberapa efektif metode k-means clustering dalam mengukur respon mahasiswa berdasarkan tingkat semangat belajar mereka dan untuk mengetahui berapa jumlah kulster optimal yang dapat menggambarkan variasi tingkat semangat belajar mahasiswa berdasarkan respon mereka terhadap kata-kata motivasi. Latar belakang penelitian ini adalah bagaimana cara menerapkan metode k-means clustering untuk mengukur respon mahasiswa terhadap kata-kata motivasi dan mengelompokkan mahasiswa berdasarkan tingkat semangat belajar mereka dan berapa jumlah cluster optimal yang dapat menggambarkan variasi tingkat semangat belajar mahasiswa berdasarkan respon mereka terhadap kata-kata motivasi. Metode yang digunakan untuk mengumpulkan data dalam penelitian ini adalah quisioner. Ruang lingkup penelitian ini melibatkan mahasiswa aktif dari semua prodi yang ada di Universitas Dipa Makassar. Berdasarkan hasil kuesioner yang telah diisi dari semua prodi yang ada pada Universitas Dipa Makassar, peneliti menyimpulkan bahwa menggunakan metode k-means bisa sangat efektif dalam mengelompokkan mahasiswa berdasarkan tingkat semangat belajar mereka dan menentukan jumlah kluster optimal yang dapat menggambarkan variasi tingkat semangat belajar mahasiswa berdasarkan respon mereka terhadap kata-kata motivasi.

Kata kunci—K-MEANS, Motivasi Verbal, Semangat Belajar, Analisis Clustering, ANOVA.

Abstract

The purpose of this study is to determine how effective the k-means clustering method is in measuring students' responses based on their level of learning enthusiasm and to determine the optimal number of clusters that can describe the variation in students' level of learning enthusiasm based on their responses to motivational words. The background of this study is how to apply the k-means clustering method to measure students' responses to motivational words and group students based on their level of learning enthusiasm and how many optimal clusters can describe the variation in students' level of learning enthusiasm based on their responses to motivational words. The method used to collect data in this study is a questionnaire. The scope of this study involves active students from all study programs at Dipa University Makassar. Based on the results of the questionnaires that have been completed from all study programs at Dipa University Makassar, the researcher concluded that using the k-means method can be very effective in grouping students based on their level of learning the enthusiasm and determining the optimal number of clusters that can describe the variation in students' level of learning enthusiasm based on their responses to motivational words.

Keywords—K-MEANS, Verbal Motivation, Learning Enthusiasm, Clustering Analysis, ANOVA.

1. PENDAHULUAN

Motivasi belajar merupakan faktor kunci yang mempengaruhi prestasi akademik mahasiswa. Salah satu faktor yang dapat memengaruhi semangat belajar dalam kata-kata motivasi dapat memberikan dorongan positif kepada mahasiswa, mendorong mereka untuk tetap fokus dan gigih dalam menghadapi tantangan akademik. Kata-kata motivasi akan berdampak pada hasil belajar yang lebih baik. Mahasiswa yang termotivasi cenderung lebih bertanggung jawab dengan pekerjaan akademis mereka. Selain itu, mereka sering menunjukkan ketekunan yang lebih besar dalam mengatasi tantangan belajar, selain itu motivasi sangat diperlukan untuk mendorong mahasiswa meraih pencapaian yang diinginkan. Motivasi belajar juga dapat diartikan sebagai penggerak yang diaktifkan [1]. Namun masih sedikit penelitian yang secara khusus meneliti tentang pengaruh kata-kata motivasi terhadap semangat belajar mahasiswa, terutama menggunakan metode k-means.

Di Universitas Dipa Makassar ada beberapa mahasiswa yang termotivasi untuk menyelesaikan tugas akademik mereka dan ada juga mahasiswa yang belum termotivasi untuk menyelesaikan tugas akademik mereka dikarenakan berbagai faktor, baik dari segi pribadi masing-masing atau dari pengaruh lingkungan mereka. Sehingga diperlukan kata-kata motivasi yang tepat untuk mendorong semangat mahasiswa dalam menyelesaikan tugas akademik mereka di Universitas Dipa Makassar. Dalam penelitian ini, penulis menggunakan algoritma K-Means Clustering karena algoritma K-Means Clustering merupakan algoritma yang sederhana dan prosesnya cepat, sehingga algoritma ini menjadi populer dan banyak digunakan untuk penelitian. Dengan mengintegrasikan beberapa metode, yaitu Knowledge Discovery in Database (KDD), Data Mining, K-Means, Clustering, Euclidean Distance, Manhattan Distance, RapidMiner, Davis Bouldin Index (DBI), dan Dewey Decimal Classification (DDC). Metode ini dipilih karena keunggulannya dalam memberikan hasil yang lebih relevan dan akurat dibandingkan metode lain yang tersedia. Dengan demikian, berdasarkan penjabaran di atas maka penulis tertarik melakukan penelitian dengan judul “Implementasi Metode k-means Pada Analisis Pengaruh Motivasi Verbal Terhadap Semangat Belajar Mahasiswa”. Hasil dari penelitian ini yaitu dapat mengukur respon mahasiswa terhadap kata-kata motivasi dan mengelompokkan mahasiswa berdasarkan tingkat semangat belajar mereka.

Penelitian mengenai motivasi belajar mahasiswa telah banyak dilakukan dengan berbagai pendekatan. Misalnya, Anugrahwati & Silitonga (2020) meneliti pengaruh pembelajaran jarak jauh terhadap motivasi belajar mahasiswa keperawatan selama masa pandemi, dan menemukan bahwa faktor lingkungan belajar memiliki peran signifikan dalam menjaga semangat belajar. Apriana (2020) juga mengungkapkan bahwa motivasi intrinsik dan ekstrinsik menjadi faktor utama yang memengaruhi prestasi akademik mahasiswa kedokteran. Namun, kedua penelitian tersebut lebih menyoroti aspek motivasi umum tanpa menjelaskan secara spesifik pengaruh motivasi verbal yakni dorongan berupa kata-kata penyemangat atau afirmasi positif terhadap semangat belajar mahasiswa. Selain itu, penelitian terdahulu juga belum mengeksplorasi perbedaan efektivitas berbagai jenis motivasi verbal, seperti ungkapan ketekunan, dorongan semangat, kesuksesan, pengingat, dan inspirasi, dalam memengaruhi semangat belajar mahasiswa. Padahal, setiap jenis motivasi verbal dapat menimbulkan respons emosional yang berbeda pada individu. Keterbatasan ini menunjukkan adanya kesenjangan penelitian dalam mengidentifikasi bentuk motivasi verbal yang paling efektif bagi mahasiswa.

Penelitian dengan konteks lokal di Universitas Dipa Makassar atau wilayah Sulawesi Selatan juga masih jarang ditemukan, sehingga data empiris mengenai pengaruh motivasi verbal terhadap semangat belajar mahasiswa di lingkungan tersebut belum tersedia. Dengan demikian, penelitian ini hadir untuk mengisi kesenjangan penelitian dengan mengimplementasikan metode K-Means Clustering dalam menganalisis pengaruh motivasi verbal terhadap semangat belajar mahasiswa, serta mengelompokkan mahasiswa berdasarkan tingkat motivasi yang mereka miliki. Pendekatan ini diharapkan dapat menghasilkan pola pengelompokan yang lebih akurat dan memberikan kontribusi baru dalam penerapan data mining di bidang pendidikan, khususnya pada

analisis faktor psikologis dan perilaku belajar mahasiswa.

Metode K-Means merupakan teknik clustering yang didasarkan pada jarak antar data, di mana data dikelompokkan ke dalam beberapa klaster tertentu. Algoritma ini hanya dapat digunakan pada data yang memiliki atribut numerik [2]. Sementara itu, semangat belajar menggambarkan tingkat antusiasme dan dorongan internal mahasiswa dalam menjalani proses pembelajaran. Motivasi sendiri terbagi menjadi dua jenis, yaitu motivasi intrinsik dan motivasi ekstrinsik [3].

Motivasi berperan sebagai proses yang mengarahkan mahasiswa untuk memperoleh berbagai pengalaman yang mendukung mereka dalam kegiatan belajar. Melalui adanya motivasi, seseorang akan memiliki ketekunan dan konsistensi dalam menjalani proses pembelajaran [4]. Sementara itu, analisis merupakan suatu proses yang mencakup berbagai kegiatan seperti membedakan, memecah, serta mengelompokkan informasi berdasarkan tujuan tertentu agar dapat dipahami keterkaitannya dan dimaknai secara lebih mendalam. Dengan kata lain, analisis adalah upaya untuk menguraikan suatu kesatuan menjadi bagian-bagian yang lebih kecil [5].

Secara umum, implementasi dalam Kamus Besar Bahasa Indonesia (KBBI) diartikan sebagai pelaksanaan atau penerapan. Istilah implementasi sering dihubungkan dengan suatu kegiatan yang dilakukan untuk mencapai tujuan tertentu. Implementasi juga dapat dipahami sebagai proses penerapan suatu gagasan, konsep, kebijakan, atau inovasi ke dalam tindakan nyata sehingga menimbulkan dampak, baik dalam bentuk perubahan pengetahuan, keterampilan, maupun sikap dan nilai [6].

Data mining berlandaskan pada bidang keilmuan seperti kecerdasan buatan, machine learning, statistik, dan sistem basis data. Secara umum, data mining merupakan proses untuk menemukan pola atau informasi tersembunyi dari suatu kumpulan data. Berdasarkan buku Konsep Data Mining Vs Sistem Pendukung Keputusan [7], penentuan jumlah klaster yang optimal menjadi tahap penting dalam proses clustering. Salah satu teknik yang sering digunakan adalah metode Elbow, yang berasumsi bahwa penambahan jumlah klaster akan menurunkan variasi di dalam klaster (intra-cluster variance) [8].

Rumus Slovin digunakan untuk menentukan jumlah sampel ketika ukuran populasi serta tingkat kesalahan (margin of error) telah diketahui [9]. Sementara itu, SPSS merupakan perangkat lunak yang memiliki kemampuan tinggi dalam melakukan analisis statistik serta pengelolaan data. Aplikasi ini bekerja dalam lingkungan grafis yang intuitif dengan menu-menu deskriptif dan kotak dialog yang sederhana, sehingga memudahkan pengguna dalam mengoperasikannya. Berbagai aktivitas analisis dapat dilakukan dengan mudah hanya dengan menggunakan fungsi point and click pada mouse [10].

Metode Skala Likert merupakan pendekatan yang digunakan untuk mengukur tingkat kepuasan atau persepsi responden terhadap suatu objek penelitian dengan menggunakan skala penilaian tertentu. Skala ini pertama kali dikembangkan oleh Rensis Likert pada tahun 1932. Dalam penerapannya, skala Likert terdiri dari beberapa pernyataan yang dirancang untuk memperoleh tanggapan responden, kemudian hasilnya dikombinasikan menjadi satu nilai atau skor yang merepresentasikan karakteristik individu, seperti pengetahuan, sikap, dan perilaku. Skala ini juga dikenal sebagai skala psikometrik yang banyak digunakan dalam kuesioner dan merupakan salah satu metode pengukuran paling umum dalam penelitian [11].

Python merupakan salah satu bahasa pemrograman yang paling banyak digunakan pada saat ini. Dalam penerapannya pada bidang data science, Python menawarkan kemudahan serta fleksibilitas yang tinggi bagi para penggunanya dalam menyelesaikan berbagai permasalahan dan tugas analisis data. Saat ini, sebagian besar ilmuwan data (data scientists) memanfaatkan kemampuan Python sebagai alat utama dalam kegiatan profesional mereka [12].

Validitas merupakan suatu ukuran yang menunjukkan sejauh mana sebuah instrumen benar-benar mampu mengukur apa yang seharusnya diukur. Semakin tinggi tingkat validitas suatu instrumen, maka semakin akurat pula alat ukur tersebut dalam memperoleh data yang sesuai dengan tujuan penelitian. Pengujian validitas sangat penting dilakukan agar butir pertanyaan yang digunakan tidak menyimpang dari variabel yang ingin diteliti. Secara teoritis, uji validitas dapat

dihitung menggunakan korelasi Product Moment atau korelasi Pearson [13]. Sementara itu, salah satu perangkat lunak penting yang sering digunakan dalam pengolahan data numerik adalah Microsoft Excel [14]. Aplikasi ini tidak hanya berfungsi untuk mengelola data dalam bentuk tabel, tetapi juga menyediakan berbagai fitur yang memungkinkan pengguna membuat visualisasi data yang informatif dan menarik [15].

2. METODE PENELITIAN

Instrumen yang digunakan untuk mengumpulkan data berdasarkan teori yang digunakan oleh peneliti dalam penelitian ini adalah kuesioner sebagai instrumen utama. Kuesioner adalah serangkaian pertanyaan atau pernyataan tertulis yang diajukan kepada responden untuk dijawab, yang bertujuan mengumpulkan data dan informasi mengenai suatu topik atau masalah penelitian. Peneliti menggunakan metode analisis data kuantitatif melalui empat tahap utama, yaitu: pengumpulan data, praproses data, K-Means, dan interpretasi hasil.

3. HASIL DAN PEMBAHASAN

3.1 Pembahasan

3.1.1 Praproses Data

a. Pemeriksaan kelengkapan data

Data hasil kuesioner yang dikumpulkan melalui Google Form diperiksa satu per satu untuk memastikan seluruh pertanyaan telah dijawab oleh responden. Responden yang tidak mengisi seluruh butir pertanyaan atau memberikan jawaban di luar skala Likert (1–5) dikeluarkan dari dataset.

Tabel 1. Data Penelitian

NAMA LENGKAP	PRODI	P1	P2
Indah hardiansah	Teknik Informatika	Tidak Memotivasi	Sangat Memotivasi
Rahman	Teknik Informatika	Memotivasi	Memotivasi
Saleo	Teknik Informatika	Memotivasi	Memotivasi
Mayanti Sepselia	Teknik Informatika	Netral	Sangat Memotivasi
Muh Ikhwan Imran	Teknik Informatika	Memotivasi	Memotivasi
Yusril	Teknik Informatika	Memotivasi	Sangat Memotivasi
Geostro	Teknik Informatika	Sangat Memotivasi	Memotivasi
Muhammad Sadik	Teknik Informatika	Sangat Memotivasi	Memotivasi
Sutan Azwar Zeus	Teknik Informatika	Sangat Memotivasi	Memotivasi
BENYAMIN BUNGA'	Teknik Informatika	Memotivasi	Memotivasi
Yati sampe bamba	Sistem Informasi	Sangat Memotivasi	Sangat Memotivasi
Ricard Salubongga	Teknik Informatika	Sangat Memotivasi	Sangat Memotivasi
Nurul Qalby Paembonan	Teknik Informatika	Memotivasi	Sangat Tidak Memotivasi
ELIASER kondorura	Teknik Informatika	Memotivasi	Memotivasi
Victor Tandi	Teknik Informatika	Sangat Memotivasi	Memotivasi
Trigyan Paschal	Teknik Informatika	Sangat Memotivasi	Sangat Memotivasi
Ananda Dhykha Aura Putri	Sistem Informasi	Memotivasi	Sangat Memotivasi
Rafael melano	Teknik Informatika	Memotivasi	Netral
GLEN	Sistem Informasi	Sangat Memotivasi	Sangat Memotivasi
Andika Musawir R	Sistem Informasi	Netral	Sangat Memotivasi

NAMA LENGKAP	PRODI	P1	P2
Muhammad Azril	Sistem Informasi	Netral	Netral
Juan Alfino	Sistem Informasi	Memotivasi	Netral
Ines Theresia	Teknik Informatika	Memotivasi	Memotivasi
Gabriel William	Teknik Informatika	Sangat Memotivasi	Memotivasi
Muh. Asrul	Teknik Informatika	Netral	Netral
Amanda	Teknik Informatika	Tidak Memotivasi	Netral
Ardi Patoding	Teknik Informatika	Memotivasi	Sangat Memotivasi
Qyron Surya Malaihollo	Sistem Informasi	Memotivasi	Sangat Memotivasi
Muhammad Kareem	Teknik Informatika	Tidak Memotivasi	Tidak Memotivasi
Welly Imanuel Laduk	Sistem Informasi	Netral	Memotivasi
Yasser Anshar	Teknik Informatika	Memotivasi	Netral
Jhozua C Rerung Allo	Sistem Informasi	Netral	Netral
Arjuna	Teknik Informatika	Memotivasi	Sangat Memotivasi
Herlina	Teknik Informatika	Sangat Memotivasi	Memotivasi
Raditya palebangan	Teknik Informatika	Memotivasi	Sangat Memotivasi
Miftahul Rahma	Teknik Informatika	Sangat Memotivasi	Sangat Memotivasi
Salimah	Teknik Informatika	Sangat Memotivasi	Sangat Memotivasi
Qalby	Teknik Informatika	Memotivasi	Memotivasi
Willhelmus	Teknik Informatika	Memotivasi	Sangat Memotivasi
Bahri	Teknik Informatika	Memotivasi	Sangat Memotivasi
Dwi	Teknik Informatika	Tidak Memotivasi	Sangat Tidak Memotivasi
Yoel	Teknik Informatika	Sangat Tidak Memotivasi	Tidak Memotivasi
David	Teknik Informatika	Tidak Memotivasi	Sangat Tidak Memotivasi
Elizabeth Ratte Datu	Teknik Informatika	Tidak Memotivasi	Sangat Tidak Memotivasi
Rosie	Teknik Informatika	Sangat Tidak Memotivasi	Tidak Memotivasi
Septiana Lisu Allo	Teknik Informatika	Sangat Tidak Memotivasi	Tidak Memotivasi
Muhammad Azah	Teknik Informatika	Sangat Tidak Memotivasi	Tidak Memotivasi
Ahmad Efendy	Teknik Informatika	Sangat Tidak Memotivasi	Tidak Memotivasi
Alif	Teknik Informatika	Tidak Memotivasi	Sangat Tidak Memotivasi

P1=Pernyataan 1 dan seterusnya

- b. Penghapusan data tidak valid

Data yang terdeteksi memiliki pola jawaban tidak logis atau tidak konsisten (misalnya memberikan nilai yang sama untuk seluruh pertanyaan secara ekstrem tanpa pertimbangan) juga dihapus dari dataset untuk menghindari bias pada hasil clusterisasi

- c. Konversi data ke format seragam

Seluruh jawaban responden dikonversi ke dalam format numerik skala Likert 1–5, di mana angka 1 menunjukkan “sangat tidak setuju” dan angka 5 menunjukkan “sangat setuju”. Proses ini dilakukan untuk memastikan semua variabel berada pada skala yang sama.

- d. Penyimpanan data siap olah

Data yang sudah bersih kemudian disimpan dalam format Microsoft Excel (.xlsx) untuk memudahkan pengolahan selanjutnya di SPSS dan RapidMiner. Dengan dilakukannya praproses ini, data yang digunakan pada tahap analisis.

K-Means dipastikan bebas dari nilai kosong (missing values) dan kesalahan format, sehingga hasil clusterisasi yang diperoleh dapat lebih akurat dan representatif.

3.1.2 ANOVA (*Analisis Of Variance*)

Uji ANOVA dilakukan untuk mengidentifikasi variabel mana yang paling membedakan antar cluster. Berdasarkan Tabel ANOVA, seluruh variabel (P01– P10) memiliki nilai signifikansi (Sig.) sebesar $< 0,001$. Hal ini menunjukkan bahwa rata-rata skor setiap variabel berbeda secara signifikan antar cluster, sehingga semua variabel berperan penting dalam proses pembentukan cluster.

Tabel 2. ANOVA

ANOVA						
		Sum of squares	df	Mean square	F	Sig.
P01	Between groups	152.147	30	5.072	8.827	<.001
	Within Groups	39.643	69	.575		
	Total	191.790	99			
P02	Between groups	152.927	30	5.098	10.703	<.001
	Within Groups	32.863	69	.476		
	Total	185.790	99			
P03	Between groups	159.989	30	5.333	8.992	<.001
	Within Groups	40.921	69	.593		
	Total	200.910	99			
P04	Between groups	139.126	30	4.638	9.472	<.001
	Within Groups	33.784	69	.490		
	Total	172.910	99			
P05	Between groups	136.223	30	4.541	7.855	<.001
	Within Groups	39.887	69	.578		
	Total	176.110	99			
P06	Between groups	143.817	30	4.794	7.401	<.001
	Within Groups	44.693	69	.648		
	Total	188.510	99			
P07	Between groups	110.809	30	3.694	7.091	<.001
	Within Groups	35.941	69	.521		
	Total	146.750	99			
P08	Between groups	134.645	30	4.488	13.198	<.001
	Within Groups	23.465	69	.340		
	Total	158.110	99			
P09	Between groups	110.585	30	3.686	8.432	<.001
	Within Groups	30.165	69	.437		
	Total	140.750	99			
P10	Between groups	144.523	30	4.817	8.327	<.001
	Within Groups	39.917	69	.579		
	Total	184.440	99			

Nilai F pada ANOVA merepresentasikan tingkat kontribusi variabel terhadap perbedaan antar cluster. Semakin tinggi nilai F, semakin besar peran variabel tersebut dalam membedakan kelompok responden. Dari hasil analisis, variabel dengan nilai F tertinggi adalah:

a. P08 ($F = 13.198$)

Variabel P08 atau pertanyaan ke-8 memiliki nilai F tertinggi, artinya indikator ini paling membedakan antar cluster mahasiswa. P08 adalah item pernyataan yang terkait dengan kata-kata inspiratif yang mampu memberikan dorongan emosional yang kuat. Hasil ini menunjukkan bahwa mahasiswa sangat dipengaruhi oleh motivasi yang bersifat inspiratif dalam meningkatkan semangat belajar.

b. P02 ($F = 10.703$)

Variabel P02 menempati urutan kedua. hasil ini mengindikasikan bahwa sebagian mahasiswa benar-benar terinspirasi dengan pesan pentingnya memanfaatkan waktu untuk belajar,

sementara sebagian lainnya menanggapi dengan tingkat motivasi yang berbeda. Namun secara umum, kalimat ini terbukti efektif meningkatkan semangat belajar, karena hasil uji statistik menunjukkan pengaruh yang nyata.

c. P04 ($F = 9.472$)

Nilai F Variabel P04 sebesar 9.472 mengindikasikan bahwa pernyataan ini cukup kuat dalam membedakan tingkat semangat belajar mahasiswa. Kalimat motivasi ini menekankan pentingnya pendidikan sebagai pembentukan karakter dan keterampilan hidup, bukan hanya pencapaian formal. Pesan ini mampu memberi dorongan psikologis yang berbeda bagi mahasiswa: sebagian merasa sangat termotivasi untuk belajar demi pengembangan diri, sementara sebagian lain mungkin masih melihat pendidikan lebih sebatas gelar.

Variabel-variabel ini merupakan indikator paling kuat dalam membedakan responden antar cluster. Artinya, jawaban responden terhadap butir-butir kuesioner ini sangat menentukan apakah mereka masuk ke dalam kelompok motivasi tinggi, sedang, atau rendah.

Dengan demikian, hasil ANOVA ini menegaskan bahwa semua item kuesioner valid untuk digunakan dalam analisis cluster dan mendukung hasil pengelompokan yang telah dilakukan. Variabel-variabel dengan nilai F tertinggi dapat menjadi fokus utama dalam perancangan strategi peningkatan motivasi verbal mahasiswa karena memiliki pengaruh yang dominan dalam membedakan karakteristik kelompok.

3.1.3 Distribusi anggota cluster

Hasil Number of Cases in Each Cluster menunjukkan distribusi responden sebagai berikut:

- Cluster 1 : 46 Responden (46%)
- Cluster 2 : 22 Responden (22%)
- Cluster 3 : 32 Responden (32%)

Distribusi ini menggambarkan bahwa sebagian besar responden berada pada Cluster 3, yang nantinya dapat diidentifikasi sebagai kelompok dengan tingkat motivasi lebih tinggi. Cluster 2 merupakan kelompok dengan jumlah responden paling sedikit, yang kemungkinan memiliki karakteristik motivasi paling rendah berdasarkan nilai rata-rata yang akan dianalisis pada Final Cluster Centers.

3.1.4 Profil Cluster

Tabel Final Cluster Centers menunjukkan nilai rata-rata untuk setiap variabel pertanyaan (P01–P10) pada masing-masing cluster. Nilai ini merepresentasikan tingkat persetujuan responden terhadap pernyataan motivasi verbal dan semangat belajar pada skala Likert 1–5.

Tabel 3. Final Cluster Centers

	1	2	3
P1	4	3	1
P2	5	3	2
P3	4	3	1
P4	4	3	2
P5	4	3	2
P6	4	3	2
P7	4	4	2
P8	4	4	2
P9	4	4	2
P10	4	4	2

3.2 Interpretasi hasil

Berdasarkan hasil pengolahan data, diperoleh interpretasi sebagai berikut:

- Cluster 1- Motivasi tinggi (46 responden/46%). Mahasiswa dalam cluster ini memiliki nilai rata-rata jawaban pada setiap item P1-P10 berada pada rentang 4–5. Bahkan pada item P2, nilai rata-rata mencapai angka 5 yang berarti responden dalam kelompok ini sangat setuju terhadap pernyataan motivasi terkait semangat belajar. Hal ini menunjukkan bahwa mahasiswa dalam cluster ini memberikan respon positif dan konsisten terhadap kata-kata motivasi verbal, sehingga dapat disimpulkan bahwa mereka termasuk kelompok dengan tingkat motivasi belajar yang tinggi.

- b. Cluster 2 – Motivasi Sedang (22 responden / 22%). Cluster ini Memiliki Nilai rata-rata pada item pernyataan berada di kisaran 3–4. Kondisi ini memperlihatkan bahwa mahasiswa dalam cluster ini cenderung cukup setuju terhadap pernyataan motivasi, namun efek motivasi tidak sepenuhnya kuat. Dengan demikian, semangat belajar mahasiswa dalam cluster ini berada pada kategori sedang. Mereka dapat terdorong oleh kata-kata motivasi, tetapi pengaruhnya tidak sebesar kelompok pertama sehingga tingkat semangat belajarnya tidak sepenuhnya stabil.
- c. Cluster 3 – Motivasi Rendah (32 responden / 32%). Cluster ini Memiliki Rata-rata nilai item P1–P10 dalam cluster ini relatif rendah, berada di kisaran 1–2. Bahkan pada item P1 dan P3 rata-rata hanya mencapai angka 1. Hal ini menunjukkan bahwa mahasiswa dalam kelompok ini kurang merespons kata-kata motivasi verbal, sehingga semangat belajar mereka dapat dikategorikan rendah. Mahasiswa pada cluster ini perlu mendapat perhatian lebih dari dosen maupun pihak kampus karena kelompok ini memiliki kecenderungan untuk tidak mudah terpengaruh oleh motivasi verbal, sehingga membutuhkan strategi lain untuk meningkatkan semangat belajar mahasiswa.

Pada penelitian ini, dilalui beberapa tahap untuk mendapatkan kluster optimal menggunakan python yaitu:

- a. Menghitung SSE

Potongan kode pada gambar adalah implementasi fungsi `hitung_sse` yang digunakan untuk menghitung Sum of Squared Error (SSE) pada algoritma K-Means. Nilai SSE ini menjadi dasar dalam menentukan jumlah cluster yang paling optimal.

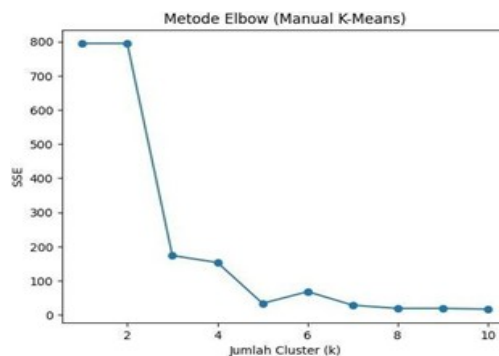
```
# 2. Fungsi untuk menghitung SSE
def hitung_sse(data, k):
    # Inisialisasi centroid secara acak
    centroid = random.sample(list(data), k)
    # Looping: Assign data ke cluster lalu hitung centroid baru
    cluster = [[] for _ in range(k)]
    for d in data:
        jarak_ke_centroid = [abs(d - c) for c in centroid]
        idx = np.argmin(jarak_ke_centroid)
        cluster[idx].append(d)
    # Update centroid
    centroid_baru = [np.mean(c) if len(c) > 0 else 0 for c in cluster]
    # Hitung SSE (total jarak kuadrat)
    sse = 0
    for i, c in enumerate(cluster):
        sse += sum((np.array(c) - centroid_baru[i])**2)
    return sse
```

- b. Menghitung Perulangan Nilai SSE

```
# 3. Hitung SSE untuk k = 1 sampai 10
sse_list = []
for k in range(1, 11):
    sse = hitung_sse(data, k)
    sse_list.append(sse)
    print(f"k={k}, SSE={sse:.2f}")
```

Pada gambar di atas ditunjukkan potongan kode Python yang digunakan untuk menghitung perulangan Sum of Squared Error (SSE) dari algoritma K-Means untuk berbagai nilai k (jumlah cluster). Proses ini merupakan tahap penting dalam metode Elbow untuk menentukan jumlah cluster optimal

Dalam penelitian ini, fungsi `hitung_sse` digunakan untuk menghitung nilai SSE pada berbagai percobaan jumlah cluster (k = 1 s.d. 10). Nilai SSE kemudian diplot dalam bentuk grafik metode Elbow.



Gambar 1. Grafik Elbow Method

Semakin besar jumlah cluster (k), SSE akan semakin kecil, karena data terbagi lebih detail. Namun, setelah titik tertentu, penurunan SSE menjadi tidak signifikan. Titik inilah yang disebut “elbow” (siku). Pada penelitian ini, titik elbow diperoleh pada $k = 3$, yang berarti jumlah cluster optimal adalah 3 cluster.

Hasil ini konsisten dengan pengelompokan mahasiswa menjadi tiga kategori motivasi, yaitu motivasi tinggi, sedang, dan rendah. Dengan demikian, implementasi fungsi `hitung_sse` menjadi dasar penting dalam memastikan jumlah cluster yang tepat dan sesuai dengan tujuan penelitian.

4. KESIMPULAN

Berdasarkan hasil penelitian mengenai pengelompokan mahasiswa berdasarkan motivasi verbal terhadap semangat belajar menggunakan metode K-Means Clustering, dapat disimpulkan:

- Menggunakan metode k-means sangat efektif untuk mengelompokkan mahasiswa terhadap Tingkat semangat belajar mereka.
- Pengelompokan responden berhasil dilakukan. K-Means Clustering menghasilkan 3 cluster dengan kategori motivasi tinggi, sedang, dan rendah. Mayoritas responden berada pada cluster motivasi tinggi.

5. SARAN

Penelitian selanjutnya agar menggunakan jumlah responden yang lebih besar dan beragam untuk mendapatkan hasil yang lebih general, menambahkan variabel lain seperti faktor lingkungan belajar atau dukungan keluarga untuk memperkaya analisis, dan mencoba metode cluster lain untuk perbandingan hasil.

DAFTAR PUSTAKA

- [1] R. Anugrahwati and J. Silitonga, “Pendidikan Jarak Jauh (PJJ) Saat Wabah Covid-19 Memengaruhi Motivasi Belajar Mahasiswa Akademi Keperawatan Hermina Manggala Husada,” *Jurnal Ilmiah Keperawatan Altruistik*, vol. 3, no. 2, pp. 10–16, 2020.
- [2] K. S. H. Kusuma Al Atros, A. R. Padri, O. Nurdyawan, A. Faqih, and S. Anwar, “Model Klasifikasi Analisis Kepuasan Pengguna Perpustakaan Online Menggunakan K-Means dan Decision Tree,” *JURIKOM (Jurnal Riset Komputer)*, vol. 8, no. 6, p. 323, 2021.
- [3] R. Apriana, “Faktor-Faktor yang Berpengaruh Terhadap Motivasi Belajar Mahasiswa Kedokteran,” *Jurnal Medika Utama*, vol. 2, no. 1, pp. 382–389, 2020.
- [4] M. R. Nanggo, “Motivasi Mahasiswa dalam Meningkatkan Prestasi Belajar Pasca Bencana Alam (Studi Kasus Mahasiswa IAIN Palu yang Tinggal di Asrama HPMIG Cabang Palu),” *Doctoral dissertation*, IAIN Palu, 2019.

- [5] Y. Septiani, E. Aribbe, and R. Diansyah, "Analisis Kualitas Layanan Sistem Informasi Akademik Universitas Abdurrah Terhadap Kepuasan Pengguna Menggunakan Metode Servqual," *Jurnal Teknologi dan Open Source*, vol. 3, no. 1, pp. 131–143, 2020.
- [6] B. T. Haji, "Pengertian Implementasi," *Laporan Akhir*, p. 31, 2020.
- [7] F. Handayani, "Aplikasi Data Mining Menggunakan Algoritma K-Means Clustering untuk Mengelompokkan Mahasiswa Berdasarkan Gaya Belajar," *Jurnal Teknologi dan Informasi*, 2022.
- [8] N. Rohman and A. Wibowo, "Perbandingan Metode K-Medoids dan Metode K-Means dalam Analisis Segmentasi Pelanggan Mall," *SINTECH Journal*, vol. 7, no. 1, 2024.
- [9] A. Santoso, "Rumus Slovin: Panacea Masalah Ukuran Sampel?," *Suksma: Jurnal Psikologi Universitas Sanata Dharma*, vol. 4, no. 2, pp. 24–43, 2023.
- [10] W. T. Bhirawa, "Proses Pengolahan Data dari Model Persamaan Regresi dengan Menggunakan Statistical Product and Service Solution (SPSS)," *Jurnal Mitra Manajemen*, vol. 7, no. 1, 2020.
- [11] R. A. Setyawan and W. F. Atapukan, "Pengukuran Usability Website E-Commerce Sambal Nyoss Menggunakan Metode Skala Likert," *Jurnal Compiler*, vol. 7, no. 1, pp. 54–61, 2018.
- [12] S. Saabith, V. Thangarajah, A. S. Saabith, T. Vinothraj, and M. Fareez, "A Review on Python Libraries and IDEs for Data Science," *International Journal of Research in Engineering*, 2021.
- [13] L. Amanda, F. Yanuar, and D. Devianto, "Uji Validitas dan Reliabilitas Tingkat Partisipasi Politik Masyarakat Kota Padang," *Jurnal Matematika UNAND*, vol. 8, no. 1, pp. 179–188, 2019.
- [14] G. Germecca, N. A. Wardhani, and M. M. Dewi, "Implementasi Sistem Informasi Antrian Berbasis Website dengan Metodologi Scrum," *Journal of Information System Management (JOISM)*, vol. 5, no. 2, 2024.
- [15] W. Wisnumurti, T. Faulina, and S. Novari, "Pelatihan Optimalisasi Microsoft Office untuk Meningkatkan Kegiatan Mahasiswa pada Pengabdian Masyarakat di Universitas Mahakarya Asia Baturaja," *BERNAS: Jurnal Pengabdian kepada Masyarakat*, vol. 4, no. 1, 2023.