

Clustering Tingkat Pemahaman Dasar Mahasiswa Pada Pra-Perkuliah Probabilitas Statistika Dengan Metode *K-Means*

Rezqiwati Ishak
Program Studi Teknik Informatika
Universitas Ichsan Gorontalo
Gorontalo, Indonesia
rezqi.uig@gmail.com

Amiruddin
Program Studi Teknik Informatika
Universitas Ichsan Gorontalo
Gorontalo, Indonesia
amier.76@gmail.com

Diterima : November 2021
Disetujui : Desember 2021
Dipublikasi : Januari 2022

Abstrak—Probabilitas Statistik merupakan mata kuliah logika perhitungan dimana sebagian Mahasiswa merasa kesulitan dalam penerimaan materi sehingga bagi Dosen pengampu mata kuliah hal ini merupakan suatu permasalahan dalam pemberian materi disebabkan tingkat pemahaman Mahasiswa yang tidak merata, hal ini tentunya akan berdampak pada sasaran dan tujuan pembelajaran yang tidak bisa tercapai sesuai Rencana Pembelajaran Semester. Tujuan penelitian ini adalah untuk melakukan pengelompokan atau *clustering* tingkat pemahaman dasar mahasiswa berdasarkan nilai mata kuliah prasyarat yaitu Kalkulus dan Matematika Diskrit sebelum perkuliahan Probabilitas Statistik dimulai. Metode yang digunakan yaitu *K-Means* untuk *Clustering* dan *Elbow* untuk menentukan jumlah *cluster* yang optimum. Berdasarkan hasil penelitian didapatkan hasil sebanyak 3 *cluster*, dimana *cluster* 1 adalah kelompok Mahasiswa yang tingkat pemahamannya digolongkan kategori Baik, *Cluster* 2 adalah kelompok kategori Sangat Baik dan *Cluster* 3 adalah kategori Kurang. Dari 110 Mahasiswa yang dijadikan sebagai dataset, masing-masing diperoleh untuk *cluster* 1 sebanyak 30 Mahasiswa, *cluster* 2 sebanyak 58 Mahasiswa dan *cluster* 3 sebanyak 22 Mahasiswa. Mahasiswa yang masuk pada *cluster* 3 perlu dilakukan tambahan perkuliahan sebelum mereka mengikuti mata kuliah Probabilitas Statistik agar capaian pembelajaran dapat dicapai, dengan demikian metode *K-Means* dan metode *Elbow* sangat cocok digunakan untuk melakukan *clustering* karena metode ini bisa didapatkan hasil *clustering* yang optimal.

Kata Kunci— *Clustering*; Probabilitas Statistik; *K-Means*; *Elbow*; *Python*

Abstract — *Statistical Probability is a calculation logic course where some students find it difficult to receive material so that for lecturers who master the course this is a problem in the provision of materials due to the uneven level of student understanding, this will certainly have an impact on the goals and objectives of learning that cannot be achieved in accordance with the Semester Learning Plan. The purpose of this research is to group or cluster the level of basic understanding of students based on the value of prerequisite courses namely Calculus and Discrete Mathematics before the Statistical Probability lecture begins. The method used*

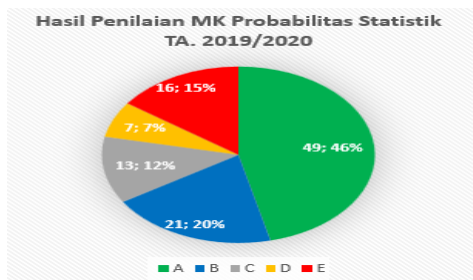
is K-Means for Clustering and Elbow to determine the optimum number of clusters. Based on the results of the study obtained results as many as 3 clusters, where cluster 1 is a group of Students whose level of understanding is classified as Good category, Cluster 2 is a category of Very Good and Cluster 3 is a category of Less. Of the 110 students who served as datasets, each was obtained for cluster 1 as many as 30 students, cluster 2 as many as 58 students and cluster 3 as many as 22 students. Students who enter cluster 3 need to do additional lectures before they take the Statistical Probability course so that learning achievement can be achieved, thus the K-Means method and elbow method are very suitable for clustering because this method can get optimal clustering results.

Keywords— *Clustering*; *Statistical probability*; *K-Means*; *Elbow*; *Python*

I. PENDAHULUAN

Statistika merupakan ilmu pengetahuan yang berkaitan dengan susunan, teknik ataupun cara mengumpulkan, mengolah dan menganalisis serta menginterpretasikan data untuk disajikan secara lengkap dalam bentuk yang mudah dipahami, sementara Probabilitas adalah derajat/tingkat kepastian dari munculnya hasil percobaan suatu statistika [1]. Statistika Probabilitas disajikan sebagai Mata Kuliah pada Jurusan Teknik Informatika khususnya semester 4 (Genap), dimana mata kuliah ini sebagai pendukung bagi mahasiswa dalam pengolahan data untuk penyusunan tugas akhir.

Pengambilan Mata kuliah Probabilitas Statistika memiliki beberapa mata kuliah prasyarat kelulusan diantaranya Kalkulus dan Matematika Diskrit sebagaimana yang tertuang dalam RPS (Rencana Pembelajaran Semester) Program Studi Teknik Informatika, Fakultas Ilmu Komputer Universitas Ichsan Gorontalo[2]. Berdasarkan hasil penilaian mata kuliah Probabilitas Statistik pada semester sebelumnya yaitu Tahun Akademik 2019/2020 sebanyak 106 Mahasiswa dapat dilihat persentasenya pada gambar 1 berikut ini[3] :



Gambar 1. Persentase Penilaian Statistik

Berdasarkan gambar 1 di atas, dapat dilihat bahwa terdapat 36,97% Mahasiswa yang mendapatkan nilai dibawah standar, terdiri dari 13,17% Mahasiswa mendapatkan nilai C, 7,7% Mahasiswa mendapatkan nilai D dan 16,15% Mahasiswa mendapatkan nilai E (tidak lulus). Hal ini menunjukkan bahwa perlu dilakukan analisis terkait pemahaman dasar Mahasiswa khususnya masalah logika dan matematika dasar sebelum perkuliahan Probabilitas Statistik dimulai agar hal ini tidak terulang di semester berikutnya. Salah satu cara yang bisa dilakukan adalah dengan melakukan pemetaan atau pengelompokkan tingkat pemahaman dasar Mahasiswa berdasarkan nilai mata kuliah prasyarat yang didapat di semester sebelumnya yaitu mata kuliah Kalkulus dan Matematika Diskrit.

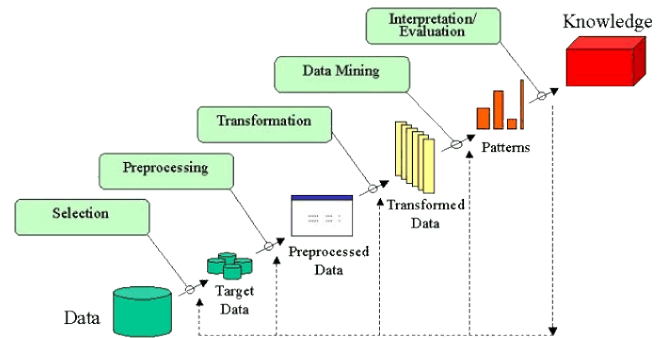
Pemetaan atau pengelompokkan ini tentunya bertujuan agar sebelum memasuki awal perkuliahan/pembelajaran Probabilitas Statistik, Dosen sudah bisa mengetahui kelompok tingkat pemahaman mahasiswa yang ada dalam kelasnya sehingga Mahasiswa yang masuk pada kelompok pemahamannya kurang, maka perlu dilakukan pra-perkuliahan 2-3 kali pertemuan penambahan materi terkait pemahaman awal yang harus dipahami sebelum perkuliahan Probabilitas Statistik dimulai, agar mereka bisa mengikuti materi perkuliahan dengan baik serta mereka juga tetap diberikan perhatian lebih sehingga mereka bisa memahami materi perkuliahan dengan baik dan lulus dengan nilai terbaik

Manfaat dari penelitian ini diharapkan bisa didapatkan hasil pemetaan atau pengelompokkan tingkat pemahaman mahasiswa dalam proses pra-perkuliahan mata kuliah Statistika, untuk itu perlu dilakukan penelitian lebih lanjut dengan menggunakan metode atau algoritma *Machine Learning* yaitu *K-Means*. Beberapa penelitian sebelumnya yang telah menggunakan metode *K-Means* dan *Elbow* diantaranya penelitian yang dilakukan oleh Adnan dengan judul "Penerapan Metode *K-Means* Dengan Metode *Elbow* Untuk Segmentasi Pelanggan Menggunakan Model RFM"[4] dan penelitian yang dilakukan oleh Andre dengan judul "Optimasi *Cluster K-Means* Menggunakan Metode *Elbow* Pada Data Pengguna Narkoba Dengan Pemrograman Python"[5], kedua penelitian tersebut menekankan bahwa penggunaan metode *Elbow* sangat diperlukan dalam penentuan jumlah *cluster* yang optimum pada metode *K-Means*, karena dengan menggunakan metode tersebut bisa langsung ditentukan jumlah *cluster* yang tepat digunakan pada kasus yang diteliti. Proses Analisa hasil *clustering* yang didapatkan dari metode *K-Means* sangat ditentukan dari jumlah *cluster* yang terbentuk, karena jika jumlah *cluster* terlalu kecil maka nilai errornya sangat tinggi namun analisisnya mudah dilakukan sedangkan jika jumlah *cluster* terlalu tinggi maka analisisnya sangat rumit namun tingkat errornya sangat rendah, untuk itu sangat penting penentuan

jumlah *cluster* yang optimum agar proses analisa bisa dilakukan dengan tepat sehingga bisa menghasilkan suatu informasi yang sangat berguna.

II. METODE

Pada penelitian ini akan menggunakan metodologi KDD (*Knowledge Discovery in Databases*) [6] dengan tahapan penelitian pada gambar 2 berikut :



Gambar 2. Tahapan Penelitian [7]

1. Tahap *Selection* : pada tahapan ini akan dilakukan seleksi data terhadap dataset yang dikumpulkan dari nilai Mahasiswa pada mata kuliah Kalkulus dan Matematika Diskrit untuk dijadikan sebagai target data (dataset)
2. Tahap *Preprocessing* : pada tahap ini dilakukan serangkaian proses yaitu menghapus data yang duplikat, pengecekan data yang tidak konsisten dan melengkapi data yang kosong.
3. Tahap *Transformation* : pada tahap ini dilakukan perubahan atau konversi data yaitu nilai mata kuliah dalam bentuk huruf dikonversi ke bentuk angka (A=4, B=3, C=2, D=1) agar dapat digunakan dalam pemodelan *K-Means*
4. Tahap *Data Mining* : pada tahap ini dilakukan pengembangan model *K-Means* untuk mendapatkan hasil *cluster* yang akan terbentuk
5. Tahap *Interpretation/Evaluation* : pada tahap ini dilakukan penentuan berapa jumlah *cluster* yang optimal dengan menggunakan metode *Elbow*. Hasil nilai K optimum dari metode *Elbow* diterapkan kedalam metode *K-Means* kembali. Hasil dari *Clustering* tersebut divisualisasikan jumlah kluster beserta anggota klusternya, sehingga mudah dipahami arti dari setiap kelompok *cluster* tersebut yang selanjutnya dijadikan sebagai pengetahuan atau keputusan untuk *clustering* tingkat pemahaman Mahasiswa pada perkuliahan Probabilitas Statistika.

A. Algoritma *Machine Learning*

Algoritma *Machine Learning* didefinisikan sebagai algoritma yang digunakan untuk melatih model, dalam *Machine Learning* dibagi menjadi tiga jenis yang berbeda, yaitu (a). *Supervised Learning* (dalam dataset ini diberi label dan teknik Regresi dan Klasifikasi digunakan), (b). *Unsupervised Learning* (dalam dataset ini tidak diberi label dan teknik yang digunakan seperti *Dimensionality reduction* dan *Clustering*) dan (c). *Reinforcement Learning* (algoritma di mana model belajar dari setiap tindakannya)[8]

B. Clustering

Cluster adalah sebuah grup yang memiliki kemiripan tertentu. Clustering adalah salah satu metode *machine learning unsupervised* untuk mengelompokkan objek-objek menjadi beberapa group atau *cluster* berdasarkan kemiripan dengan obyek yang lainnya [9]. Karena termasuk kategori *unsupervised*, maka dataset yang digunakan model *clustering* tidak memiliki label. Metode *Clustering* yang digunakan dalam *machine learning unsupervised* yaitu metode *K-Means, K-Medoids, Fuzzy C-means, Density methods, Grid-based methods, Hierarchical bases method* dan lain-lain [10].

C. K-Means Clustering

Pengklasteran *K-Means* adalah sebuah metode yang dikembangkan oleh Stuart Lloyd dari Bell Labs pada tahun 1957 [11]. Algoritma *K-Means* didefinisikan sebagai metode *Unsupervised Learning* yang memiliki proses berulang di mana dataset dikelompokkan ke dalam k jumlah *cluster* atau subkelompok yang tidak tumpang tindih yang telah ditentukan, membuat titik-titik dalam *cluster* sedekat mungkin ketika mencoba untuk menjaga *cluster* pada ruang yang berbeda itu mengalokasikan titik data ke *cluster* sehingga jumlah jarak kuadrat antara *cluster centroid* dan titik data berada pada titik data. minimal, pada posisi ini *centroid* dari *cluster* adalah nilai rata-rata dari titik data yang ada dalam *cluster* [12]. Secara umum tahapan algoritma *K-Means* sebagai berikut[9] :

1. Tentukan k buah *cluster*
2. Pilih sejumlah k buah objek secara acak yang akan dijadikan sebagai titik *centroid cluster*
3. Tentukan k buah *centroid* (titik tengah)
4. Kelompokkan obyek ke *centroid cluster* terdekat berdasarkan *Euclidean distance* :

$$d_{ik} = \sqrt{\sum_{i=1}^n (C_{ij} - C_{kj})^2} \quad (1)$$

Dimana:

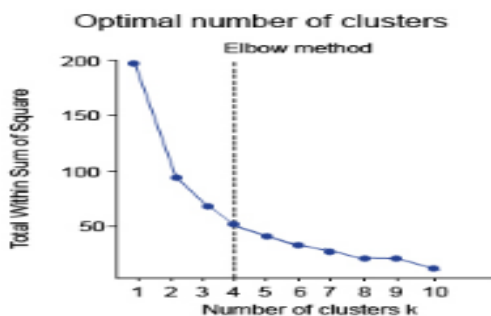
C_{ij} = pusat *cluster*

C_{kj} = data

5. Hitung Kembali semua titik *centroid*
6. Ulangi Langkah 3-5 sehingga nilai titik *centroid* tidak lagi berubah.

D. Metode Elbow

Metode *Elbow* adalah salah satu metode atau teknik yang bisa diterapkan pada algoritma *K-Means* untuk menentukan jumlah *K* yang optimal. [5] [4]. Untuk menentukan jumlah *K* yang optimal dengan melihat titik pada *graph* dimana penurunan inersia tidak lagi signifikan seperti pada gambar 3 di bawah ini, jumlah *K* yang optimal adalah *K=4*.



Gambar 3. Contoh Visualisasi Metode *Elbow* [10]

III. HASIL DAN PEMBAHASAN

Proses *Clustering* dilakukan sesuai dengan tahapan metode *KDD* yang diuraikan di atas, mulai tahapan seleksi sampai dengan tahapan evaluasi menggunakan *tools Jupyter Notebook* dengan bahasa pemrograman *Python* dengan menggunakan beberapa *library* seperti *Numpy, Pandas, Matplotlib dan Scikit Learn* [13][8]. Dataset yang digunakan sebanyak 110 data sesuai dengan jumlah mahasiswa yang mengambil mata kuliah Kalkulus dan Matematika Diskrit pada Tahun Akademik 2020/2021 di Program Studi Teknik Informatika Universitas Ichsan Gorontalo

A. Tahap Selection

Pada tahapan ini dilakukan seleksi dataset yang akan digunakan yaitu nilai mahasiswa pada mata kuliah Kalkulus pada tabel 1 dan Matematika Diskrit pada tabel 2 [14].

TABEL 1. SAMPLE DATASET NILAI KALKULUS

No	NIM	Nilai
1	T3119101	D
2	T3120002	A
3	T3120003	B
4	T3120004	A
5	T3120005	A
6	T3120006	B
7	T3120007	A
8	T3120008	A
9	T3120009	A
10	T3120010	A

TABEL 2. DATASET NILAI MATEMATIKA DISKRIT

No	NIM	Nilai
1	T3114097	B
2	T3116126	D
3	T3116342	D
4	T3117006	B
5	T3117057	B
6	T3117069	D
7	T3117149	B
8	T3117168	C
9	T3118003	D
10	T3120010	A

B. Tahap Preprocessing

Pada tahap ini dilakukan serangkaian proses yaitu menghapus data yang duplikat, melengkapi data yang kosong dan melakukan penggabungan data nilai Kalkulus dan nilai Matematika Diskrit, proses penggabungan data ini langsung diproses dalam pemrograman *Python*, dengan hasil akhir penggabungan data seperti pada tabel 3 berikut ini :

TABEL 3. SAMPLE HASIL PENGGABUNGAN DATASET

No	NIM	Kalkulus	M_Diskrit
1	T3120002	A	A
2	T3120003	B	C
3	T3120004	A	D
4	T3120005	A	A
5	T3120006	B	D
6	T3120007	A	A
7	T3120008	A	B
8	T3120009	A	B
9	T3120010	A	A
10	T3120011	B	A

C. Tahap Transformation

Pada tahap ini dilakukan konversi data nilai yaitu nilai mata kuliah dalam bentuk huruf dikonversi ke bentuk angka yaitu nilai A=4, B=3, C=2, D=1, hal ini dilakukan agar dapat digunakan dalam pemodelan metode *K-Means*. Tahap konversi ini juga langsung dilakukan didalam Python dengan hasil akhir seperti pada tabel 4 berikut ini :

TABEL 4. SAMPLE HASIL KONVERSI NILAI HURUF KE ANGKA

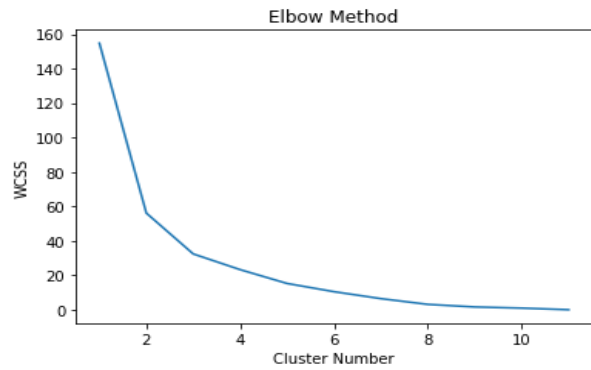
No	NIM	Kalkulus	M_Diskrit
1	T3120002	4	4
2	T3120003	3	2
3	T3120004	4	1
4	T3120005	4	4
5	T3120006	3	1
6	T3120007	4	4
7	T3120008	4	3
8	T3120009	4	3
9	T3120010	4	4
10	T3120131	3	4

D. Tahap Data Mining :

Pada tahap ini dilakukan pengembangan model *K-Means* untuk melakukan *clustering*. Perintah untuk proses pemodelan metode *K-Means* ini menggunakan *library* dari *Scikit Learn*.

E. Tahap Interpretation/Evaluation :

Pada tahap ini dilakukan evaluasi terhadap jumlah *cluster* yang optimum dengan menggunakan metode *elbow*, kemudian hasil metode tersebut dibuatkan dalam bentuk visualisasi dengan menggunakan *library Matplotlib*, setelah dijalankan perintanya, maka didapatkan hasil visualisasi pada gambar 4 berikut ini :

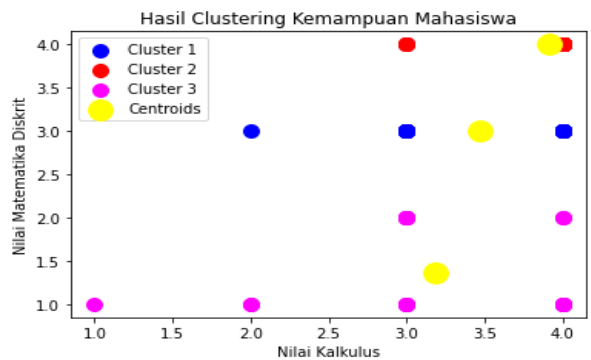


Gambar 4. Visualisasi Hasil Metode Elbow

Pada gambar 4, di atas dapat dilihat bahwa pada penelitian ini jumlah *cluster* yang optimum adalah sebanyak 3 *cluster*. Dengan menggunakan metode *Elbow* ini, jumlah *cluster* yang optimum bisa langsung diketahui tanpa harus menguji jumlah *cluster* secara berulang. Setelah didapatkan jumlah *cluster* yang optimum, maka jumlah *cluster* tersebut dimasukkan kedalam fungsi *K-Means* yang tersedia pada *library Scikit Learn* dengan tahapan sebagai berikut :

1. Baca dataset yang sudah diproses:
`X = dfData.iloc[:, [1,2]].values`
2. Panggil Fungsi K-Means :
`kmeans = KMeans(n_clusters = 3, init = 'k-means++')`
3. Terapkan hasil pemodelan terhadap dataset :
`y_kmeans = kmeans.fit_predict(X)`
4. Lakukan visualisasi hasil *clustering* atau pemodelan

Berikut Visualisasi hasil *clustering* berdasarkan pemodelan *K-Means* yang sudah dilakukan :



Gambar 5. Visualisasi Hasil Cluster

Berdasarkan gambar 5 di atas, dapat ditentukan titik *Centroids* masing-masing *cluster* dengan melihat posisi koordinat pusat *cluster* yang diberikan warna kuning atau dapat juga diketahui secara langsung dengan menggunakan perintah “`(kmeans.cluster_centers_)`” sehingga didapatkan hasil sebagai berikut :

```
[[3.46666667 3.         ]
 [3.9137931  4.         ]
 [3.18181818 1.36363636]]
```

Gambar 6. Informasi Pusat Cluster

Jumlah anggota masing-masing *cluster* dapat diketahui dengan menggunakan perintah `dfData['Kluster'].value_counts()`. Berdasarkan kedua perintah tersebut di atas, dapat dimodel hasilnya dalam bentuk tabel seperti pada tabel 5 berikut ini :

TABEL 5. CENTROIDS DAN JUMLAH DATA

Cluster	Centroids		Jumlah Mahasiswa
	Kalkulus	Matematika Diskrit	
1	3,46	3	30
2	3,91	4	58
3	3,18	1,36	22

Berdasarkan tabel 5 di atas, dengan melihat titik *Centroids* setiap *cluster*, maka dapat dianalisa atau dikategorikan bahwa *cluster* 1 adalah kategori tingkat pemahaman Mahasiswa kategori baik, *cluster* 2 adalah kategori sangat baik dan *cluster* 3 adalah kategori kurang, dengan demikian dapat dinyatakan bahwa *cluster* 3 adalah merupakan kelompok Mahasiswa yang perlu diberikan materi tambahan sebelum mereka mengikuti perkuliahan Mata Kuliah Probabilitas Statistik sebanyak 22 Mahasiswa, agar proses perkuliahan dapat berjalan lancar dan apa yang menjadi tujuan capaian pembelajaran dapat tercapai. Hasil pelabelan *clustering* untuk masing-masing Mahasiswa berdasarkan file Excel yang sudah di *download* dari hasil pemodelan di aplikasi Python dapat dilihat pada tabel 6 di bawah ini :

TABEL 6. SAMPEL HASIL CLUSTERING

NIM	Kalkulus	M_Diskrit	Kluster
T3120002	4	4	2
T3120003	3	2	3
T3120004	4	1	3
T3120005	4	4	2
T3120006	3	1	3
T3120007	4	4	2
T3120008	4	3	1
T3120009	4	3	1
T3120010	4	4	2
T3120011	3	4	1

Berdasarkan informasi tabel 6 di atas, maka dengan mudah dapat diketahui setiap Mahasiswa masuk ke *cluster* berapa. Sehingga jika ingin dicari atau diketahui Mahasiswa yang masuk *cluster* 3, maka cukup dilakukan filter untuk

cluster tersebut, dengan demikian informasi ini sangat membantu dalam mengambil keputusan lebih lanjut.

IV. KESIMPULAN

Berdasarkan hasil penelitian yang sudah diuraikan di atas, maka jumlah *cluster* optimum yang dihasilkan adalah sebanyak 3 *cluster*, dimana *cluster* 1 adalah kelompok Mahasiswa yang tingkat pemahamannya kategori Baik, *cluster* 2 adalah kelompok kategori Sangat Baik dan *cluster* 3 adalah kategori Kurang. Dari 110 Mahasiswa yang dijadikan sebagai dataset, masing-masing diperoleh untuk *cluster* 1 sebanyak 30 Mahasiswa, *cluster* 2 sebanyak 58 Mahasiswa dan *cluster* 3 sebanyak 22 Mahasiswa. Berdasarkan hasil *cluster* tersebut, maka Mahasiswa yang masuk pada *cluster* 3 perlu dilakukan tambahan perkuliahan sebelum mereka mengikuti mata kuliah Probabilitas Statistik, dengan demikian metode *K-Means* dan metode *Elbow* sangat cocok digunakan untuk melakukan *clustering* karena dapat menghasilkan *clustering* yang optimal.

UCAPAN TERIMA KASIH

Penulis menyampaikan terima kasih yang sebesar-besarnya kepada Lembaga Penelitian Universitas Ichsan Gorontalo yang telah memberikan dukungan materil sehingga penelitian ini dapat kami laksanakan. Ucapan terima kasih juga kami sampaikan kepada Kepala Biro Administrasi dan Akademik Universitas Ichsan Gorontalo.

REFERENSI

- [1] Sudaryono, *Statistika Probabilitas (Teori & Aplikasi)*. Yogyakarta: Andi, 2012.
- [2] FIKOM-UNISAN, "Rencana Pembelajaran Semester (RPS) Mata Kuliah Probabilitas Statistik," Gorontalo, 2021.
- [3] BAAK, "Daftar Nilai Mata Kuliah Probabilitas Statistik Tahun Akademik 2019/2020," Gorontalo, 2020.
- [4] A. B. H. Kiat, Y. Azhar, and V. Rahmayanti, "Penerapan Metode K-Means Dengan Metode Elbow Untuk Segmentasi Pelanggan Menggunakan Model RFM (Recency, Frequency & Monetary)," *Repositor*, vol. 2, no. 7, pp. 945–952, 2020.
- [5] A. Winarta and W. J. Kurniawan, "Optimasi Cluster K-Means Menggunakan Metode Elbow Pada Data Pengguna Narkoba Dengan Pemrograman Python," *J. Tek. Inform. Kaputama*, vol. 5, no. 1, pp. 113–119, 2021.
- [6] F. A. Hermawati, *Data Mining*. Yogyakarta: Andi, 2013.
- [7] O. Maimon and L. Rokach, *Data Mining and Knowledge Discovery Handbook*, Second. London: Springer Science+Business Media, 2010.
- [8] Y. Heryadi and T. Wahyono, *Machine Learning Konsep dan Implementasi*. Yogyakarta: Gava Media, 2020.
- [9] R. Primartha, *Algoritma Machine Learning*. Bandung: Informatika, 2021.
- [10] P. Pedamkar, "Clustering in Machine Learning," *EDUCBA. All Rights Reserved*, 2020. <https://www.educba.com/clustering-in-machine-learning/> (accessed Sep. 29, 2021).
- [11] Suyanto, *Machine Learning Tingkat Dasar dan Lanjut*. Bandung: Informatika, 2018.
- [12] P. Pedamkar, "K-Means Clustering Algorithm," *EDUCBA. All Rights Reserved*, 2020. <https://www.educba.com/k-means-clustering-algorithm/> (accessed Sep. 29, 2021).
- [13] T. Wahyono, *Fundamental of Python for Machine Learning*. Yogyakarta: Gava Media, 2018.
- [14] BAAK UNISAN, "Daftar Nilai Mahasiswa Kalkulus dan Matematika Diskrit," Gorontalo, 2021. [Online]. Available: <https://siakun.unisan.ac.id/>.