



PENERAPAN K-MEANS *CLUSTERING* DALAM MENENTUKAN BIDANG MAGANG MAHASISWA JURUSAN TEKNIK INFORMATIKA UNIVERSITAS PAPUA

Sri Putri Aulia Syam^{*}, Christian Dwi Suhendra, Lion Ferdinand Marini

Program Studi Teknik Informatika, Fakultas Teknik, Universitas Papua, Indonesia

Abstrak: Kampus merdeka merupakan kebijakan Menteri Pendidikan dan Kebudayaan dengan beberapa program yang diterapkan di perguruan tinggi salah satunya adalah magang dan studi independent bersertifikat (MSIB). Program magang memberikan kebebasan kepada mahasiswa untuk memilih bidang magang serta instansi penyedia magang. Namun dalam penerapannya bidang magang yang dipilih belum sesuai dengan kemampuan yang dimiliki mahasiswa. Kajian ini bertujuan untuk menerapkan K-Means Clustering dalam menentukan bidang magang mahasiswa. Studi kasus yang diambil yaitu mahasiswa Jurusan Teknik Informatik Universitas Papua. Kemampuan mahasiswa akan diukur berdasarkan nilai mata kuliah serta menerapkan metode K-Means *Clustering* untuk mengelompokkan bidang magang mahasiswa. Setelah melakukan pengujian diperoleh nilai DBI terendah yaitu 0.203 dari nilai $k=2$. *Cluster* yang terbentuk dalam penelitian ini yaitu *cluster_0* dengan 57 data nilai mahasiswa dan *cluster_1* dengan 134 data nilai mahasiswa. Dengan metode K-Means *Clustering* berhasil menentukan bidang magang mahasiswa berdasarkan nilai mata kuliah.

Kata kunci: *Clustering*, K-Means, Mahasiswa, Nilai Mata Kuliah, Program Magang

I. PENDAHULUAN

Kampus Merdeka merupakan kebijakan Menteri Pendidikan dan Kebudayaan yang ditetapkan dengan tujuan mendorong mahasiswa agar dapat menguasai berbagai keilmuan untuk diimplementasikan dalam dunia kerja nantinya. Berbagai program telah dilaksanakan guna mencapai tujuan tersebut salah satunya adalah program Magang dan Studi Independen Bersertifikat (MSIB). Salah satu program yang termasuk dalam MSIB tersebut adalah program magang. Program magang merupakan program yang bertujuan

memperkenalkan dan menumbuhkan kemampuan mahasiswa dalam dunia kerja secara nyata (Tohir, 2020). Program magang ini memberikan kebebasan kepada mahasiswa untuk memilih bidang serta instansi magang.

Kebebasan untuk memilih bidang dan instansi magang yang diberikan ternyata dinilai belum efektif dalam menjalankan program tersebut. Mahasiswa yang cenderung memilih bidang magang tidak sesuai dengan kemampuan yang dimiliki mengakibatkan program tidak berjalan sebagaimana mestinya dan hasil yang diperoleh tidak seperti yang diharapkan. Dengan demikian dibutuhkan metode untuk mengklaster bidang magang sesuai dengan kemampuan mahasiswa yang diukur menggunakan nilai mata kuliah. Kajian ini

^{*} sriputriauliasyam823@gmail.com

Diterima: 27 April 2023

Direvisi: 3 Mei 2023

Disetujui: 30 Mei 2023

DOI: 10.23969/infomatek.v25i1.7585

bertujuan untuk menerapkan metode K-Means Clustering dalam menentukan bidang magang untuk mahasiswa Jurusan Teknik Informatika Universitas Papua.

Clustering adalah proses pengelompokan data data menjadi beberapa kelompok sehingga objek di dalam satu kelompok memiliki banyak kemiripan serta memiliki banyak perbedaan dengan objek dikelompok lain (Alam Jusia et al., 2019). Pengklasteran dapat dilakukan dengan memanfaatkan teknologi data mining menggunakan pengklasteran K-Means. Metode K-Means akan mengelompokkan data yang ada ke dalam beberapa *cluster*, cluster yang akan dikelompokkan memiliki kemiripan data satu sama lain serta data yang memiliki karakteristik yang berbeda dengan cluster lainnya (Indah Werdiningsih et al., 2020). Hasil dari pengelompokkan data itu nantinya akan digunakan sebagai acuan untuk memilih bidang magang sesuai dengan kemampuan masing-masing mahasiswa.

II. METODOLOGI

Data yang digunakan pada penelitian ini merupakan data sekunder berupa rekap nilai mata kuliah mahasiswa Program Studi S1 Teknik Informatika Angkatan 2019 sampai Angkatan 2020 dan mahasiswa D3 Teknik Komputer Angkatan 2020 sampai 2021 Universitas Papua. Rekap nilai diperoleh dari masing-masing Ketua Program Studi Jurusan Teknik Informatika. Data tersebut kemudian dianalisis dengan tahapan seleksi data, pembersihan data, interpretasi/evaluasi, dan kesimpulan. Setelah data dianalisis maka akan diolah menggunakan Algoritma K-Means *clustering* melalui beberapa tahapan berikut ini (Ekasetya & Jananto, 2020).

a. Tentukan nilai K sebagai jumlah *cluster* yang diinginkan dan metrik ketidakmiripan (jarak) yang diinginkan. Jika perlu, tetapkan

ambang batas perubahan fungsi objektif dan ambang batas perubahan *centroid*.

b. Pilih K data baru set data X sebagai *centroid*.

Penentuan *centroid* dilakukan secara acak, sedangkan pada tahap iterasi digunakan persamaan berikut ini:

$$V_{ij} = \frac{1}{N_i} = \sum_{k=0}^{N_i} X_{kj} \dots\dots\dots(1)$$

Keterangan:

V_{ij} = *centroid* rata-rata pada *cluster* ke – i untuk variabel ke – j

N_i = jumlah data *cluster* ke – i

i, k = indeks dari *cluster*

j = indeks variabel

X_{kj} = nilai data ke – k variabel ke – j untuk *cluster* tersebut

c. Alokasikan semua data ke *centroid* terdekat dengan metrik jarak yang sudah ditetapkan (memperbaharui ID setiap data).

d. Hitung kembali *centroid* C berdasarkan data yang mengikuti *cluster* masing masing.

e. Ulangi langkah tiga dan empat hingga kondisi konvergen tercapai, yaitu (a) perubahan fungsi objektif sudah dibawah ambang batas yang diinginkan; atau (b) tidak ada data yang berpindah *cluster*; atau (c) perubahan posisi *centroid* sudah dibawah ambang batas yang ditetapkan.

$$d_{ij} = \sqrt{\sum_{k=1}^p \{X_{ik} - X_{jk}\}^2} \dots\dots\dots(2)$$

dimana:

d_{ij} = jarak objek antara objek i dan j

P = dimensi data

X_{ik} = koordinat dari obyek i pada dimensi k

X_{jk} = koordinat dari obyek j pada dimensi

Data nilai yang digunakan pada penelitian ini adalah nilai yang sesuai dengan ketentuan pelaksanaan program magang yakni nilai mahasiswa yang sedang menempuh minimal semester 4 pada program D3 dan minimal semester 5 pada program S1. Dengan demikian tidak semua nilai mata kuliah menjadi tolak ukur dalam penelitian ini. Nilai matakuliah yang digunakan dikelompokkan dalam 5 kelompok matakuliah yaitu MK1, MK2, MK3, MK4, MK5. Anggota dari setiap kelompok mata kuliah dapat dilihat pada tabel 1.

Tabel 1. Anggota Kelompok Matakuliah

Dasar Komputer Dan Algoritma (MK1)	Algoritma Dan Pemrograman
	Logika Informatika
	Dasar Teknik Komputer
	Struktur Data
	Aplikasi Perkantoran
Algoritma Pengembangan PL (MK2)	Konsep Teknologi Informasi
	Pemrograman Web I
	Pemrograman Web II
	Pemrograman Berbasis Komponen
	Basis Data
Komputer Arsitektur (MK3)	Pemrograman Berorientasi Objek
	Arsitektur Komputer
	Pengantar Ilmu Komputer
	Sistem Digital
Sistem Terdistribusi (MK4)	Sistem Operasi
	Jaringan Komputer I
	Jaringan Komputer II
	Teknologi Wireless
	Administrasi Sistem

Matematika Dan Statistika (MK5)	Jaringan Komputer Lanjut
	Metode Komunikasi
	Matematika Dasar I
	Matematika Dasar II
	Matematika Komputasi
	Probabilitas Dan Statistika

Variabel data data dilihat pada tabel 2.

Tabel 2. Variabel Data

No	Kode_Mhs	MK1	MK2	MK3	MK4	MK5
1	IT1	3,835	3,89	3,582	3,33	3,582
2	IT2	3	4	3,5	3,37	3,25
-	-	-	-	-	-	-
-	-	-	-	-	-	-
212	IT212	0,75	0	0	0	2,67
213	IT213	1,707	2	0,333	2,5	3,33

Setelah diperoleh 213 data nilai mahasiswa selanjutnya akan dibersihkan untuk mencegah adanya inkonsisten pada data. Data yang dibersihkan adalah data nilai mahasiswa yang bernilai 0 (nol). Dari hasil pembersihan data yang awalnya sebanyak 213 data menjadi 191 data, hal ini berarti terdapat 22 data nilai mahasiswa yang memiliki nilai 0 (nol). Setelah data dibersihkan selanjutnya akan di transformasi agar data siap diminingkan. Pada tahap transformasi ini karena data yang diperoleh sudah berbentuk angka (numerik) sehingga tidak terjadi transformasi data.

III. HASIL DAN PEMBAHASAN

Data yang sudah siap diminingkan kemudian dimasukkan pada aplikasi RapidMiner untuk diolah. Data akan diuji untuk mendapatkan nilai Davies Bouldin Index (DBI) terendah sebagai nilai *cluster* terbaik pada setiap nilai k (Irhamni et al., 2014) . Hasil pengujian nilai DBI dapat dilihat pada tabel 3 berikut ini.

Tabel 3. Hasil Pengujian Nilai DBI

Percobaan	Nilai K	Nilai DBI
1	2	0.203
2	3	0.259
3	4	0.264
4	5	0.267

Berdasarkan hasil pengujian diperoleh nilai DBI terendah pada 4 percobaan nilai k yaitu pada k=2 dengan nilai 0.203. Sehingga *cluster* terbaik pada penelitian ini adalah 2 *cluster*. Hasil perhitungan nilai centroid oleh RapidMner pada *cluster* terbaik dapat dilihat pada gambar 1 berikut ini.

Attribute	cluster_0	cluster_1
MK1	2.157	3.104
MK2	1.817	3.389
MK3	2.190	3.309
MK4	1.743	3.304
MK5	2.371	3.222

Gambar 1. Nilai Centroid Setiap *Cluster*

Selanjutnya *cluster model* yang terbentuk pada penelitian ini dapat dilihat pada gambar 2 berikut ini.

Cluster Model

```
Cluster 0: 57 items
Cluster 1: 134 items
Total number of items: 191
```

Gambar 2. *Cluster Model* Pada Penelitian

Berdasarkan *cluster model* yang terbentuk diperoleh 57 data nilai mahasiswa termasuk dalam *cluster 0* dan 134 data nilai mahasiswa termasuk dalam *cluster 1*. Anggota *cluster* untuk setiap *cluster* yang terbentuk dapat dilihat pada tabel 4 berikut:

Tabel 4. Hasil *Clustering*

No	Kode_						Cluster
	Mhs	MK1	MK2	MK3	MK4	MK5	
1	IT1	3,83	3,89	3,58	3,33	3,58	cluster_1
2	IT2	3	4	3,5	3,37	3,25	cluster_1
3	IT3	2,75	2,33	3	1,5	0,91	cluster_0
4	IT4	2,91	2,17	2,5	3,12	0,75	cluster_0
5	IT5	3,08	3,67	3,58	3,20	3,16	cluster_1
6	IT6	2,5	2,94	3,41	3,12	1,16	cluster_1
-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-
188	IT188	1,75	3,16	0,67	1,25	2	cluster_0
189	IT189	0,95	0,25	1	0,75	2	cluster_0
190	IT190	2,62	3,25	1,77	1,75	2	cluster_0
191	IT191	1,70	2	0,33	2,75	3,33	cluster_0

Berdasarkan hasil *cluster* selanjutnya bidang magang akan dipilih berdasarkan nilai mahasiswa setiap *clusternya*. Untuk setiap *cluster* yang terbentuk nilai akan dianalisis berdasarkan nilai rata-rata yang diperoleh pada setiap MK. Nilai rata-rata setiap MK pada *cluster* dapat dilihat pada tabel 5.

Tabel 5. Nilai Rata-Rata MK

No	Cluster	MK1	MK2	MK3	MK4	MK5
1	Cluster_0	2,16	1,81	2,19	1,74	2,37
2.	Cluster_1	3,10	3,39	3,31	3,30	3,22

Berdasarkan nilai rata-rata pada setiap kelompok MK dapat disimpulkan bahwa pada *cluster_0* kemampuan mahasiswa baik dalam Matematika dan Statistika, Komputer Arsitektur, serta Dasar Komputer dan Algoritma dan masih kurang dalam Algoritma Pengembangan Perangkat Lunak dan Sistem Distribusi sehingga dapat dikategorikan sebagai mahasiswa yang memiliki kemampuan Dasar Informatika. Pada

cluster_1 kemampuan mahasiswa sangat baik terutama pada Algoritma Pengembangan dan Perangkat Lunak yang merupakan kemampuan dengan nilai tertinggi serta kemampuan Sistem Distribusi sehingga dapat dikategorikan sebagai mahasiswa yang memiliki kemampuan Pemrograman dan Sistem Distribusi.

IV. KESIMPULAN

Berdasarkan uraian hasil dan pembahasan yang telah dikemukakan sebelumnya, maka penulis menyimpulkan hasil penelitian sebagai berikut:

1. Penerapan metode K-Means *Clustering* dalam menentukan bidang magang mahasiswa Jurusan Teknik Informatika membentuk 2 *cluster* yaitu *cluster_0* dan *cluster_1*. Data yang termasuk dalam *cluster_0* terdiri dari 57 data nilai mahasiswa dengan kategori nilai rendah. Data yang dalam *cluster_1* terdiri dari 134 data nilai mahasiswa dengan kategori nilai tinggi.
2. Mahasiswa yang termasuk dalam *cluster_0* adalah mahasiswa dengan kemampuan Dasar Informatika dan dapat memilih bidang magang diantaranya *Web Developer*, *Starter Front End Web Engineer Internship Program*, *Starter Backend Web Engineer Internship Program*, *Data Support Intern*, *IT support*, dan *Business Development*. dan mahasiswa yang termasuk dalam *cluster_1* adalah mahasiswa dengan kemampuan Pemrograman dan Jaringan dan dapat memilih bidang magang *UI/UX*, *System Analyst*, *Data Scientist*, *System Engineer*, *Software Development*, *Software Programmer*, *Frontend Engineer*, *Backend Engineer*, *AI Engineer*, dan *Mobile Apps Developer*.

DAFTAR PUSTAKA

- Alam Jusia, P., Muhammad Irfan, F., & Dinamika Bangsa Jambi Jl Jend Sudirman Thehok Jambi, S. (2019). *Clustering Data Untuk Rekomendasi Penentuan Jurusan Perguruan Tinggi Menggunakan Metode K-Means*. *Jurnal IKRA-ITH Informatika*, 3(3), 75.
- Aprilla Dennis. (2013). Belajar Data Mining dengan RapidMiner. *Innovation and Knowledge Management in Business Globalization: Theory & Practice, Vols 1 and 2*, 5(4), 1–5. http://esjournals.org/journaloftechnology/archive/vol1no6/vol1no6_6.pdf%5Cnhttp://www.airccse.org/journal/nsa/5413nsa02.pdf
- Hermawan, B. (2018). *DASAR TERHADAP MOTIVASI DAN KEMAMPUAN PEMROGRAMAN Bayu Hermawan Adi Pratama Bambang Sujatmiko*. 03.
- Irhamni, F., Damayanti, F., Khusnul K, B., & A, M. (2014). Optimalisasi Pengelompokan Kecamatan Berdasarkan Indikator Pendidikan Menggunakan Metode *Clustering* dan Davies Bouldin Index. *Seminar Nasional Dan Teknologi UMJ*, 11, 1–5.
- Kampus Merdeka. (2020). Kementerian Pendidikan Dan Kebudayaan RI. <https://kampusmerdeka.kemdikbud.go.id/>
- Lase, Y., & Panggabean, E. (2019). Implementasi Metode K-Means *Clustering* Dalam Sistem Pemilihan Jurusan Di SMK Swasta Harapan Baru. *Jurnal Teknologi Dan Ilmu Komputer Prima (JUTIKOMP)*, 2(2), 43. <https://doi.org/10.34012/jutikomp.v2i2.723>
- Ndruru, E., & Limbong, R. (2018).

- Implementasi Data Mining Dalam Pengelompokan Jurusan yang Diminati Siswa SMK Negeri 1 Lolowa'u menggunakan Metode *Clustering* | Ndruru | MEANS (Media Informasi Analisa dan Sistem). *MEANS (Media Informasi Analisa Dan Sistem)*, 3(2), 107–113.
http://ejournal.ust.ac.id/index.php/Jurnal_Means/article/view/273/pdfdsdx11
- Nur Khormarudin, A. (2016). Teknik Data Mining: Algoritma K-Means *Clustering*. *Jurnal Ilmu Komputer*, 1–12.
<https://ilmukomputer.org/category/datamining/>
- Rini, O., & Kunang, S. O. (2021). Implementasi Data Mining Menggunakan Metode Naive Bayes Untuk Penentuan Penerima Bantuan Program Indonesia Pintar (Pip) (Studi Kasus : Sd Negeri 9 Air Kumbang). *Bina Darma Conference on Computer Science*, 3(4), 714–722.
<https://conference.binadarma.ac.id/index.php/BDCCS/article/view/2450>
- Sibuea, M. L., & Safta, A. (2017). Pemetaan Siswa Berprestasi Menggunakan Metode K-Means Clustering. *Jurteks*, 4(1), 85–92.
<https://doi.org/10.33330/jurteks.v4i1.28>
- Sopyan, Y., Lesmana, A. D., & Juliane, C. (2022). Analisis Algoritma K-Means dan Davies Bouldin Index dalam Mencari *Cluster* Terbaik Kasus Perceraian di Kabupaten Kuningan. *Building of Informatics, Technology and Science (BITS)*, 4(3), 1464–1470.
<https://doi.org/10.47065/bits.v4i3.2697>
- Sulistiyawati, A., & Supriyanto, E. (2021). Implementasi Algoritma K-means Clustering dalam Penentuan Siswa Kelas Unggulan. *Jurnal Tekno Kompak*, 15(2), 25.
<https://doi.org/10.33365/jtk.v15i2.1162>
- Tohir, M. (2020). *Buku Panduan Merdeka Belajar - Kampus Merdeka*.
<https://doi.org/10.31219/osf.io/ujmte>
- Zaefuan, W. (2018). *Penerapan Clustering Pada Sistem Pembagian Rombongan Belajar Siswa Menggunakan Algoritma K-Means*.
<http://eprintslib.ummgl.ac.id/3103/>