

**PENERAPAN METODE *K-MEANS CLUSTERING* UNTUK
ANALISIS PROFIL LULUSAN TEKNIK INDUSTRI DI DUNIA
KERJA**

TUGAS AKHIR

**Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar Sarjana Strata-1
Program Studi Teknik Industri - Fakultas Teknologi Industri
Universitas Islam Indonesia**



Nama : Amalia Rahmatika
No. Mahasiswa : 20522260

**PROGRAM STUDI TEKNIK INDUSTRI PROGRAM SARJANA
FAKULTAS TEKNOLOGI INDUSTRI
UNIVERSITAS ISLAM INDONESIA
YOGYAKARTA
2024**

PERNYATAAN KEASLIAN

Saya mengakui bahwa tugas akhir ini adalah hasil karya saya sendiri kecuali kutipan dan ringkasan yang seluruhnya sudah saya jelaskan sumbernya. Jika dikemudian hari ternyata terbukti pengakuan saya ini tidak benar dan melanggar peraturan yang sah maka saya bersedia ijazah yang telah saya terima ditarik kembali oleh Universitas Islam Indonesia.

Yogyakarta, 10 November 2024



(Amalia Rahmatika)

20522260

SURAT BUKTI PENELITIAN



FAKULTAS
TEKNOLOGI INDUSTRI

Grinding KH. Mas Mansur
Kampus Terpadu Universitas Islam Indonesia
Jl. Kalirejo Km. 14,5 Yogyakarta 55584
T. (0274) 898444 ext. 4100, 4101
F. (0274) 895007
E. fti@iainid.ac.id
W. fti.iainid.ac.id

SURAT KETERANGAN PENELITIAN

Nomor: 016/Ka.Lab.Datmin/70/Lab.Datmin/XI/2024

Assalamu'alaikum Warahmatullahi Wabarakatuh

Kami yang bertanda tangan dibawah ini, menerangkan bahwa mahasiswa dengan keterangan sebagai berikut :

Nama : Amalia Rahmatika
No. Mhs : 20522260
Dosen Pembimbing : Annisa Uswatun Khasanah, S.T., M.Sc.

Telah selesai melaksanakan penelitian yang berjudul " Penerapan Metode K-Means Clustering untuk Analisis Profil Lulusan Teknik Industri di Dunia Kerja" di Laboratorium Data Mining, Program Studi Teknik Industri, Fakultas Teknologi Industri, Universitas Islam Indonesia tercatat mulai tanggal 05 Mei sampai dengan tanggal 21 Juni 2024.

Demikian surat keterangan kami keluarkan, agar dapat dipergunakan sebagaimana mestinya.

Wassalamu'alaikum Warahmatullahi Wabarakatuh

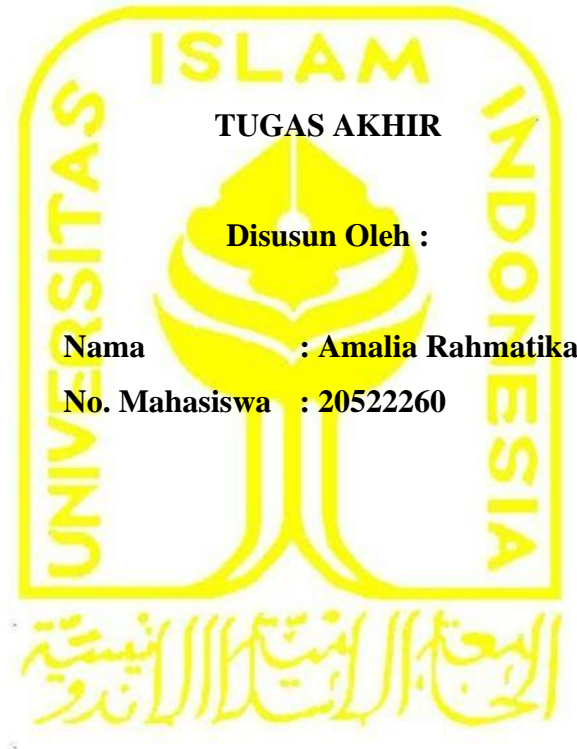
Yogyakarta, 12 November 2024

Kepala Laboratorium
Data Mining

Dr. Dwi Adi Purnama, S.T.

LEMBAR PENGESAHAN PEMBIMBING

**PENERAPAN METODE *K-MEANS CLUSTERING* UNTUK ANALISIS PROFIL
LULUSAN TEKNIK INDUSTRI DI DUNIA KERJA**



Yogyakarta, 10 November 2024

Dosen Pembimbing

(Annisa Uswatun Khasanah, S.T., M.Sc.)

LEMBAR PENGESAHAN DOSEN PENGUJI

PENERAPAN METODE *K-MEANS CLUSTERING* UNTUK ANALISIS PROFIL LULUSAN TEKNIK INDUSTRI DI DUNIA KERJA

TUGAS AKHIR

Disusun Oleh :

Nama : Amalia Rahmatika

No. Mahasiswa : 20522260

Telah dipertahankan di depan sidang penguji sebagai salah satu syarat untuk memperoleh gelar Sarjana Strata-1 Teknik Industri Fakultas Teknologi Industri Universitas Islam Indonesia

Yogyakarta, 21 November 2024

Tim Penguji

Annisa Uswatun Khasanah, S.T., M.Sc.

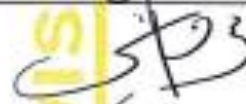
Ketua

Dr. Harwati, S.T., M.T.

Anggota I

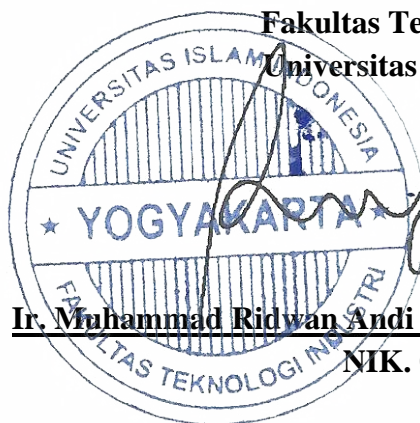
Elanjati Worldailmi, S.T., M.Sc.

Anggota II


Mengetahui,

Ketua Program Studi Teknik Industri Program Sarjana
Fakultas Teknologi Industri
Universitas Islam Indonesia



Ir. Muhammad Ridwan Andi Purnomo, S.T., M.Sc., Ph.D., IPM.

NIK. 015220101

HALAMAN PERSEMBAHAN

Bismillahirrahmanirrahim

Dengan mengucapkan syukur, Alhamdulillahirabbil'alamin, segala puji bagi Allah SWT yang telah memberikan nikmat dan karunia-Nya sehingga tugas akhir ini dapat diselesaikan.

Shalawat serta salam senantiasa tercurah kepada Nabi Muhammad SAW, beserta keluarga serta sahabatnya.

Tugas akhir ini saya persembahkan untuk kedua orang tua saya yang telah memberikan saya dukungan dan doa yang tiada henti. Terima kasih atas pengorbanan yang dilakukan sejak saya lahir hingga dapat menyelesaikan studi S1 di Teknik Industri FTI UII.

MOTTO

“Cukuplah Allah bagiku. Tidak ada Tuhan selain Dia. Hanya kepada-Nya aku bertawakal dan Dia adalah Tuhan pemilik ‘Arasy (singgasana) yang agung”

(Q.S. At-Taubah: 129)

“Percayalah bahwa kamu bisa, dan kamu sudah setengah jalan menuju kesuksesan”

(Theodore Roosevelt)

KATA PENGANTAR

Assalamu'alaikum Warahmatullahi Wabarakatuh.

Alhamdulillah *rabbi'l'alam*, dengan menyebut nama Allah SWT yang Maha Pengasih lagi Maha Penyayang, segala puji dan syukur kami panjatkan kepada Allah SWT atas anugerah dan nikmat, serta ridho-Nya sehingga kami dapat menyelesaikan tugas akhir yang berjudul “Penerapan Metode *K-Means Clustering* untuk Analisis Profil Lulusan Teknik Industri di Dunia Kerja”. Sholawat dan salam pun senantiasa tidak lupa tucurahkan kepada junjungan kita Nabi Muhammad SAW yang telah membimbing umat manusia dari zaman kegelapan menuju zaman yang terang benderang seperti saat ini dengan penuh ilmu dan anugerah.

Pada kesempatan kali ini, penulis ingin mengucapkan terima kasih yang tulus kepada semua pihak yang telah ikut serta membantu dan membimbing penulis dalam penulisan laporan tugas akhir ini, baik yang terlibat secara langsung maupun tidak langsung. Oleh karena itu, penulis ingin berterima kasih kepada:

1. Bapak Prof. Dr. Ir. Hari Purnomo, M.T., IPU, ASEAN.Eng. selaku Dekan Fakultas Teknologi Industri, Universitas Islam Indonesia.
2. Bapak Dr. Drs. Imam Djati Widodo, M.Eng.Sc., selaku Ketua Jurusan Program Studi Teknik Industri, Fakultas Teknologi Industri, Universitas Islam Indonesia.
3. Bapak Ir. Muhammad Ridwan Andi Purnomo, S.T., M.Sc., Ph.D., IPM., selaku Ketua Program Studi Teknik Industri Program Sarjana Fakultas Teknologi Industri, Universitas Islam Indonesia.
4. Ibu Annisa Uswatun Khasanah, S.T., M.Sc., selaku dosen pembimbing tugas akhir yang telah memberikan bimbingan, ilmu, serta arahan yang telah beliau berikan sehingga penulis dapat menyelesaikan laporan tugas akhir ini.
5. Kedua Orang Tua penulis, Bapak Supama dan Ibu Ellyani, yang selalu memberikan doa, motivasi, serta dukungan moral maupun materil kepada penulis.
6. Teman-teman Teknik Industri UII angkatan 2020 yang telah memberikan semangat dan motivasi dalam pelaksanaan tugas akhir
7. Semua pihak yang telah membantu penulis dalam penyelesaian tugas akhir yang tidak dapat disebutkan satu persatu.

Penulis menyadari bahwa penulisan laporan tugas akhir ini masih banyak kekurangan dan jauh dari kata sempurna. Maka dari itu, penulis mengharapkan kritik dan saran yang bersifat membangun dengan tujuan untuk menyempurnakan tugas akhir. Akhir kata, semoga laporan tugas akhir ini dapat bermanfaat bagi penulis dan pembaca.

Yogyakarta, 10 November 2024



Amalia Rahmatika

NIM 20522260

ABSTRAK

Profil lulusan merupakan salah satu acuan yang diperlukan untuk merancang kurikulum, metode pembelajaran, evaluasi, serta informasi yang dapat membentuk pemahaman dan keterampilan mahasiswa. Dalam dunia kerja, profil lulusan diperlukan untuk melakukan perekrutan karyawan dimana calon karyawan yang merupakan lulusan institusi pendidikan diharapkan memiliki berbagai macam keterampilan dan kompetensi yang sesuai dengan program studi yang dijalani. Namun, kenyataannya, masih banyak lulusan institusi pendidikan yang kurang dalam menguasai keterampilan teknis atau *hard skill* dan keterampilan soft *skill*. Dampak negatif yang dapat ditimbulkan apabila hal tersebut tidak diatasi adalah seperti kesalahan dalam perekrutan dimana perusahaan merekrut karyawan yang tidak memiliki kompetensi serta keterampilan yang sesuai untuk mengisi posisi yang ada. Program Studi Teknik Industri Universitas Islam Indonesia, yang melakukan perancangan kurikulum berdasarkan perkembangan teknologi yang ada, penting untuk mengetahui profil lulusan seperti apa yang diharapkan di dunia kerja. Maka dari itu, diperlukan analisis untuk mengetahui profil lulusan teknik industri yang seperti apa yang paling diharapkan di dunia kerja menggunakan metode *K-Means Clustering* sehingga nantinya dapat diketahui kriteria profil lulusan teknik industri seperti apa yang paling dibutuhkan di dunia kerja. Tujuan dari dilakukannya penelitian ini adalah untuk melakukan analisis clustering profil lulusan teknik industri menggunakan metode *K-Means Clustering* dan mencari karakteristik di tiap clusternya dan memberikan strategi terkait pengembangan profil lulusan berdasarkan hasil analisis cluster yang didapatkan. Hasil dari penelitian ini adalah didapatkan 3 *cluster* yang terbentuk beserta saran di setiap *clusternya*.

Kata Kunci: Profil Lulusan, Teknik Industri, *Data Mining*, *Clustering*, *K-Means Clustering*

DAFTAR ISI

PERNYATAAN KEASLIAN	ii
SURAT BUKTI PENELITIAN	iii
LEMBAR PENGESAHAN PEMBIMBING	iv
LEMBAR PENGESAHAN DOSEN PENGUJI	v
HALAMAN PERSEMBAHAN	vi
MOTTO	vii
KATA PENGANTAR	viii
ABSTRAK	ix
DAFTAR ISI	x
DAFTAR TABEL	xii
DAFTAR GAMBAR	xiii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan Penelitian	3
1.4 Manfaat Penelitian	3
1.5 Batasan Penelitian	3
BAB II TINJAUAN PUSTAKA	4
2.1 Kajian Literatur	4
2.1.1 Analisis <i>Clustering</i>	4
2.1.2 Analisis <i>Clustering</i> Menggunakan <i>K-Means Clustering</i>	6
2.2 Landasan Teori	11
2.2.1 Profil Lulusan	11
2.2.2 <i>Knowledge Discovery in Database (KDD)</i>	12
2.2.3 <i>Data Mining</i>	14
2.2.4 <i>Clustering</i>	15
2.2.5 <i>K-Means Clustering</i>	17
2.2.6 <i>Elbow Method</i>	18
BAB III METODE PENELITIAN	20
3.1 Jenis Penelitian	20
3.2 Objek dan Subjek Penelitian	20
3.3 Metode Pengumpulan Data	20
3.4 Alur Penelitian	20
BAB IV PENGUMPULAN DAN PENGOLAHAN DATA	28
4.1 <i>Data Selection</i>	28
4.2 <i>Pre-Processing Data</i>	30
4.3 <i>Data Transformation</i>	32
4.4 Pengolahan Data	34
4.4.1 Penentuan Nilai optimal <i>Cluster</i>	34
4.4.2 Analisis <i>K-Means Clustering</i>	36
BAB V PEMBAHASAN	46
5.1 Nilai Optimal <i>Cluster</i>	46
5.2 Proporsi <i>Cluster</i>	46
5.3 <i>Profiling Cluster</i>	47
5.3.1 <i>Cluster 1</i>	47
5.3.2 <i>Cluster 2</i>	49
5.3.3 <i>Cluster 3</i>	50

5.4	Strategi Pengembangan Profil Lulusan	52
5.4.1	<i>Cluster 1</i>	54
5.4.2	<i>Cluster 2</i>	55
5.4.3	<i>Cluster 3</i>	55
BAB VI	PENUTUP	56
6.1	Kesimpulan.....	56
6.2	Saran.....	57
DAFTAR PUSTAKA	58
LAMPIRAN	A-1

DAFTAR TABEL

Table 2. 1 Kajian Literatur	10
Tabel 3. 1 Deskripsi Variabel.....	23
Tabel 4. 1 Sumber Data.....	28
Tabel 4. 2 Jumlah Perusahaan dan Lowongan Berdasarkan Jenisnya	28
Tabel 4. 3 <i>Cleaning</i> Data.....	30
Tabel 4. 4 Aturan Transformasi Data.....	32
Tabel 4. 5 Tabel Transformasi Data.....	33
Tabel 4. 6 Perhitungan SSE.....	35
Tabel 4. 7 Tabel Deskriptif Statistik	37
Tabel 4. 8 Tabel <i>Initial Cluster Centers</i>	38
Tabel 4. 9 Tabel Iterasi.....	38
Tabel 4. 10 Tabel <i>Final Cluster Centers</i>	39
Tabel 4. 11 <i>Number of Cases in each Cluster</i>	41
Tabel 5. 1 Strategi Pengembangan Profil Lulusan.....	52

DAFTAR GAMBAR

Gambar 4. 1 Grafik <i>Elbow Method</i>	36
Gambar 4. 2 Visualisasi	45
Gambar 5. 1 Jumlah Anggota <i>Cluster</i>	46

BAB I PENDAHULUAN

1.1 Latar Belakang

Pada suatu institusi pendidikan, terdapat beberapa acuan yang diperlukan untuk merancang kurikulum, metode pembelajaran, evaluasi, serta informasi yang dapat membentuk pemahaman dan keterampilan mahasiswa, salah satunya adalah profil lulusan. Pada hakikatnya, profil menurut Mulyani dalam Lisdarti (2021) merupakan pandangan sisi, garis besar, atau biografi dari diri seseorang atau kelompok yang memiliki usia yang sama. Profil lulusan merupakan gambaran umum mengenai kompetensi, seperti *soft skill*, *hard skill*, sikap dan etika, kemampuan beradaptasi, serta kemampuan umum yang diharapkan ada pada setiap lulusan setelah menyelesaikan suatu program pendidikan di program studi tertentu (Hidayat & Zainuddin, 2017). Salah satu faktor yang dapat menentukan keberhasilan mahasiswa adalah profil lulusan dimana dengan adanya profil lulusan diharapkan dapat meningkatkan motivasi belajar mahasiswa serta dapat membantu mahasiswa untuk mempersiapkan diri dalam menghadapi dunia kerja.

Di dunia kerja sendiri, profil lulusan diperlukan untuk melakukan perekrutan karyawan. Terdapat berbagai macam keterampilan dan kompetensi yang diharapkan ada pada seorang lulusan institusi pendidikan, contohnya seperti kompetensi teknis atau *hard skill* yang mendalam serta relevan dengan bidang studi yang ditempuh, *soft skill* seperti kemampuan komunikasi, kepemimpinan, dan kerja sama tim yang baik, sikap dan etika profesional yang tinggi, kemampuan beradaptasi yang baik dalam menghadapi perubahan, serta memiliki pengetahuan umum yang luas. Namun, pada kenyataannya terdapat kesenjangan antara profil lulusan yang diharapkan dengan kenyataan yang ada, dimana pada kenyataannya masih banyak lulusan institusi pendidikan yang kurang dalam menguasai keterampilan teknis atau *hard skill*, keterampilan *soft skill* yang kurang baik akibat kurang mendapat perhatian pada kurikulum formal, kurangnya kesadaran akan etika dan sikap profesional di dunia kerja, kurang siap dalam menghadapi perubahan akibat sistem pendidikan yang cenderung konvensional, serta pengetahuan umum dan wawasan sosial yang kurang terasah akibat terlalu fokus pada aspek akademis dan teknis (Maryana, et al., 2022)

Profil lulusan juga merupakan hal penting yang perlu diketahui di dunia kerja, contohnya di suatu perusahaan. Apabila perusahaan tidak mengetahui profil lulusan yang diharapkan ketika melakukan perekrutan karyawan, hal tersebut dapat menimbulkan beberapa dampak negatif bagi perusahaan itu sendiri. Dampak negatif yang dapat ditimbulkan adalah seperti kesalahan dalam perekrutan dimana perusahaan merekrut karyawan yang tidak memiliki kompetensi serta

keterampilan yang sesuai untuk mengisi posisi yang ada. Selain itu, perusahaan juga harus mengeluarkan biaya tambahan untuk pengembangan karyawan apabila kompetensi karyawan tidak memadai untuk memenuhi ekspektasi kerja, kualitas produk dan layanan yang menurun akibat karyawan yang tidak kompeten juga dapat mengakibatkan produktivitas perusahaan juga ikut menurun dan akhirnya perusahaan berisiko tertinggal dalam persaingan industri.

Berdasarkan penjelasan di atas, dapat diketahui bahwa profil lulusan merupakan suatu hal sangat penting, baik itu bagi institusi pendidikan, mahasiswa, ataupun perusahaan. Pada Program Studi Teknik Industri Universitas Islam Indonesia, yang melakukan perancangan kurikulum berdasarkan perkembangan teknologi yang ada sehingga lulusan teknik industri nantinya dapat bersaing di dunia kerja, penting untuk mengetahui profil lulusan seperti apa yang diperlukan di dunia kerja. Selain itu, banyaknya jumlah lulusan dari perguruan tinggi tidak sebanding dengan jumlah lapangan pekerjaan yang ada sehingga menyebabkan banyak lulusan yang bekerja tidak sesuai jurusannya (Haryanto, et al., 2024). Maka dari itu, diperlukan analisis untuk mengetahui profil lulusan teknik industri yang seperti apa yang paling diharapkan di dunia kerja.

Clustering merupakan salah satu teknik pada *data mining* yang digunakan untuk mengelompokkan data berdasarkan kemiripan atau kedekatan antar data. Pada penelitian ini, *clustering* digunakan untuk mengetahui keterampilan atau kompetensi apa saja yang paling dibutuhkan di dunia kerja. Terdapat beberapa metode *clustering* yang dapat digunakan, seperti *K-Means Clustering*, *Hierarchical Clustering*, *Density-Based Spatial Clustering of Applications with Noise*, dan *Self Organizing Maps (SOM)* (Syaripudin, et al., 2013). Pada penelitian tentang analisis profil lulusan teknik industri di dunia kerja ini, metode yang digunakan adalah metode *K-Means Clustering* dimana metode ini merupakan salah satu metode yang paling populer dan memiliki keunggulan seperti kemudahan dan kemampuannya untuk meng*cluster* data yang berukuran besar dan data *outlier* dengan sangat cepat (Metisen & Sari, 2015). Selain itu, jumlah *cluster* (k) pada *K-Means Clustering* ditentukan di awal sehingga proses clustering menjadi lebih terarah. Metode *K-Means* merupakan metode iteratif yang melakukan pembagian dataset menjadi k *cluster* dimana setiap kelompok *cluster* memiliki *centroid* yang merupakan representasi dari kelompok *cluster* tersebut (Syaripudin, et al., 2013).

Hasil dari penelitian ini nantinya adalah kriteria profil lulusan teknik industri seperti apa yang paling dibutuhkan di dunia kerja, dimana hasil tersebut dapat dijadikan sebagai acuan untuk melakukan perancangan kurikulum, metode pembelajaran, evaluasi, serta informasi yang dapat membentuk pemahaman dan keterampilan yang dapat berguna bagi mahasiswa sebagai persiapan untuk menghadapi dunia kerja.

1.2 Rumusan Masalah

Berdasarkan latar belakang, didapatkan rumusan masalah sebagai berikut:

1. Bagaimana hasil dari tiap *cluster* yang terbentuk berdasarkan analisis *clustering* dengan menggunakan metode *K-Means Clustering* dan bagaimana karakteristik di tiap *clusternya*?
2. Apa saja strategi yang dapat diberikan untuk pengembangan profil lulusan berdasarkan hasil analisis *cluster*?

1.3 Tujuan Penelitian

Berikut ini merupakan tujuan penelitian yang menjawab rumusan masalah di atas:

1. Melakukan analisis *clustering* profil lulusan teknik industri menggunakan metode *K-Means Clustering* dan mencari karakteristik di tiap *clusternya*.
2. Memberikan strategi terkait pengembangan profil lulusan berdasarkan hasil analisis *cluster* yang didapatkan.

1.4 Manfaat Penelitian

Berikut ini merupakan manfaat penelitian yang dapat dirasakan oleh berbagai pihak:

1. Dapat membantu mahasiswa untuk meningkatkan motivasi belajar mahasiswa serta membantu mahasiswa untuk mempersiapkan diri dalam menghadapi dunia kerja.
2. Dapat membantu dalam Program Studi Teknik Industri dalam merancang strategi pengembangan profil lulusan.
3. Dapat membantu perusahaan dalam melakukan perekrutan karyawan yang tepat dan sesuai untuk mengisi posisi di perusahaan.

1.5 Batasan Penelitian

Berikut merupakan batasan-batasan dalam dilakukannya penelitian ini:

1. Data penelitian yang digunakan merupakan data lowongan kerja yang ditujukan untuk lulusan teknik industri.
2. Penelitian ini berfokus pada strategi pengembangan profil lulusan teknik industri berdasarkan hasil analisis *cluster*.

BAB II

TINJAUAN PUSTAKA

2.1 Kajian Literatur

Kajian literatur merupakan desain penelitian yang digunakan dalam pengumpulan sumber data yang berkaitan dengan suatu topik penelitian (El Hafiz & Himawan, 2021). Kajian literatur pada penelitian ini merujuk pada 16 penelitian terdahulu yang memiliki topik yang sama dengan penelitian ini.

2.1.1 Analisis *Clustering*

Pada penelitian yang berjudul “*Research on Clustering of Material Inspection Data for Highway Constuction Projects Based on Agglomerative Hierarchical Clustering*” (Zhang, et al., 2024), menjelaskan bahwa pada pelaksanaan proyek pembangunan jalan raya skala besar, akan dilakukan pengujian secara rinci terhadap berbagai indikator material yang digunakan dalam proyek. Hasil dari penelitian dengan menggunakan metode *Hierarchical Clustering* ini adalah didapatkan jumlah klasifikasi terbaik dilakukan dilakukan jika jumlah *cluster* adalah 3. Namun, jumlah *cluster* 4 juga masuk akal jika mempertimbangkan kebutuhan sebenarnya dari aplikasi. Jadi, dapat disimpulkan bahwa kita dapat secara fleksibel memilih 3 atau 4 kelas sebagai jumlah *cluster* terbaik tergantung pada kebutuhan aplikasi yang berbeda.

Pada penelitian yang berjudul “Pengukuran Tingkat Kesesuaian Pekerjaan Alumni Berdasarkan Profil Lulusan Dengan Metode *Hierarchical Clustering*” (Zebua, et al., 2022), menjelaskan bahwa pada Universitas Methodist Indonesia, khususnya pada Fakultas Ilmu Komputer, belum menyediakan sistem yang dapat mempermudah pengelompokan data alumni di dunia kerja. Pengelompokan alumni dibutuhkan berdasarkan kesesuaian pekerjaan dengan bidang keahlian tiap lulusan. Selain itu, masa tunggu lulusan dari mereka wisuda sampai mendapatkan pekerjaan pertama serta lokasi tempat mereka bekerja juga merupakan hal yang penting untuk diketahui oleh bagian studi yang bersangkutan. Hasil dari penelitian dengan menggunakan metode *Hierarchical Clustering*, tepatnya dengan teknik *single linkage clustering* ini diketahui bahwa dari 163 alumni, sebanyak 96 (59%) lulusan memiliki pekerjaan yang sesuai dengan profil lulusan mereka, sedangkan sebanyak 67 (41%) memiliki pekerjaan yang tidak sesuai dengan profil lulusan mereka.

Pada penelitian yang berjudul “*Comparative Study of Hydrochemical Classification Based on Different Hierarchical Cluster Analysis Methods*” (Bu, et al., 2020), menjelaskan klasifikasi hidrokimia merupakan salah satu langkah penting untuk memahami variasi kimia air di dalam lingkungan dalam hidrogeologi. Untuk mengelompokkan data hidrokimia, digunakan satu metode umum dalam analisis *cluster*, yaitu metode *Hierarchical Cluster Analysis* (HCA). Namun, meskipun menggunakan dataset yang sama, penggunaan berbagai metode HCA yang ada dapat menghasilkan jumlah cluster yang berbeda. Maka dari itu, dilakukanlah evaluasi komparatif dengan membandingkan beberapa metode HCA dalam mengklasifikasikan data hidrokimia untuk menentukan metode yang paling sesuai. Dengan membandingkan metode *ward's method*, *average linkage*, dan *complete linkage*, dan *single linkage* didapatkan hasil *ward's method* dan *average linkage* merupakan metode yang lebih konsisten dalam menghasilkan cluster yang relevan dan stabil. Khususnya pada *ward's method*, metode ini menunjukkan performa terbaik dalam menjaga integritas struktur data asli.

Pada penelitian yang berjudul “*Impact of Complete Lockdown on Total Infection and Death Rates: A Hierarchical Cluster Analysis*” (Ghosal, et al., 2020), menjelaskan bahwa lockdown merupakan Upaya yang dilakukan berbagai negara untuk membatasi penyebaran virus pada pandemi Covid-19. Terdapat beberapa faktor yang dapat mempengaruhi efektivitas *lockdown* dalam mengurangi tingkat infeksi virus dan kematian di berbagai negara, contohnya adalah kepatuhan masyarakat, kesiapan sistem kesehatan, serta durasi *lockdown*. Untuk mengevaluasi bagaimana *lockdown* total dapat mempengaruhi penyebaran virus, tingkat kematian, serta untuk mengidentifikasi pola-pola yang muncul di berbagai negara, maka dilakukan penelitian ini dengan menggunakan *Hierarchical Cluster Analysis* (HCA). Hasil dari penelitian ini adalah didapatkan tiga *cluster* yang masing-masing beranggotakan negara-negara yang memiliki karakteristik yang sama dalam hal tingkat infeksi dan kematiannya. *Cluster 1* yaitu *cluster* dengan dampak tinggi berisi negara yang memiliki tingkat infeksi dan kematian yang tinggi meskipun menerapkan *lockdown* total. *Cluster 2* yaitu *cluster* dengan dampak sedang berisi negara yang berhasil menurunkan laju infeksi dan kematian, namun belum sepenuhnya berhasil mengendalikan pandemi. Sedangkan *Cluster 3* yaitu *cluster* dengan dampak rendah berisi negara yang berhasil mengendalikan pandemi dan memiliki tingkat infeksi serta kematian yang relatif rendah.

Pada penelitian yang berjudul “*A Hierarchical Clustering Method of Hydrogen Bond Networks in Liquid Water Undergoing Shear Flow*” (Gao, et al., 2021), menjelaskan bahwa ikatan *hydrogen* merupakan interaksi penting dalam struktur dan dinamika molekul air. Dalam disiplin ilmu seperti kimia, fisika, biologi, dan ilmu material, pemahaman yang mendalam tentang bagaimana jaringan ikatan hidrogen berperilaku di bawah kondisi aliran geser memiliki implikasi signifikan dalam

berbagai disiplin ilmu tersebut. Untuk mengidentifikasi dan mengelompokkan pola-pola dalam jaringan ikatan hidrogen yang muncul saat air cair mengalami aliran geser, dilakukanlah penelitian dengan menggunakan metode *Hierarchical Clustering* ini. Hasil dari penelitian ini berupa teridentifikasinya pola-pola unik dalam jaringan ikatan hidrogen yang berkembang saat air mengalami aliran geser. Diketahui beberapa cluster menunjukkan stabilitas yang tinggi, sementara yang lain mengalami perubahan yang dinamis yang menunjukkan bahwa ikatan hidrogen di dalam air tidak hanya dipengaruhi oleh tegangan geser, tetapi juga berinteraksi dengan ikatan lain secara kompleks.

2.1.2 Analisis Clustering Menggunakan K-Means Clustering

Pada penelitian yang berjudul “*K-Means Clustering* dalam Pengelompokan Relevansi Pekerjaan S1 Informatika (Studi Kasus Jurusan Teknik Informatika UMM Malang)” (Haryanto, et al., 2024), menjelaskan bahwa evaluasi mengenai tingkat keberhasilan capaian pembelajaran lulusan yang ada di perguruan tinggi sangat diperlukan dikarenakan banyaknya jumlah lulusan dari perguruan tinggi tidak sebanding dengan jumlah lapangan pekerjaan yang ada sehingga menyebabkan banyak lulusan yang bekerja tidak sesuai jurusannya. Sehingga, untuk menganalisa data relevansi pekerjaan para lulusan S1 Teknik Informatika dengan apa yang sudah mereka pelajari, digunakanlah metode *K-Means Clustering*. Berdasarkan analisis *clustering*, hasil dari penelitian ini didapatkan jumlah *cluster* sebanyak 3 dimana sebanyak 29.92% lulusan Teknik Informatika UMM dapat memenuhi capaian pembelajaran lulusan dan memiliki pekerjaan yang relevan, sebanyak 49.63% lulusan kurang menguasai keterampilan khusus namun memiliki pekerjaan yang relevan, dan sebanyak 20.45% lulusan memiliki pekerjaan yang kurang relevan dengan bidang Teknik Informatika.

Pada penelitian yang berjudul “*Implementation of K-Means Clustering Method to Lectures Based on Publications of National Journals and Accredited Sinta*”) (Dahnial, 2023), menjelaskan bahwa penilaian kinerja merupakan aspek yang sangat penting dalam rangka meningkatkan kualitas sumber daya manusia (SDM), hal ini termasuk kinerja seorang dosen dalam mempublikasi hasil karya ilmiahnya baik itu dalam bentuk jurnal nasional maupun jurnal internasional. Maka dari itu, dilakukanlah penerapan metode *K-Means Clustering* untuk melihat kinerja dosen dengan mengacu pada hasil publikasi penelitian dosen berbasis jurnal nasional dan jurnal terakreditasi sinta. Hasil dari penelitian mengenai implementasi metode *K-Means Clustering* ini adalah terdapat empat kelompok yang terbentuk yaitu kelompok dosen yang mempunyai bobot jurnal nasional tinggi dan bobot jurnal akreditasi sinta tinggi sebanyak

0, kelompok dosen yang memiliki bobot jurnal nasional tinggi dan bobot jurnal akreditasi rendah sebanyak 5, kelompok dosen yang memiliki bobot jurnal nasional rendah dan bobot jurnal akreditasi tinggi sebanyak 9, serta kelompok dosen yang memiliki bobot jurnal nasional rendah dan bobot jurnal akreditasi rendah sebanyak 20. Saran yang diberikan peneliti yaitu diharapkan dapat menambahkan lebih banyak variabel masukan yang detail untuk dijadikan data yang diimplementasikan ke dalam sistem.

Pada penelitian yang berjudul “Klasterisasi Data Hasil Studi Pelacakan Tentang Karir dan Pekerjaan Lulusan Perguruan Tinggi Menggunakan Algoritma *K-Means*” (Sutrisno, et al., 2023), menjelaskan bahwa kesesuaian bidang pekerjaan dengan program pendidikan yang ditempuh dan masa tunggu sebelum mendapatkan pekerjaan merupakan salah satu indikator yang menentukan kualitas lulusan. Maka dari itu, dilakukanlah studi pelacakan (*tracking study*) bagi para lulusan. Pada penelitian ini, metode *K-Means Clustering* digunakan untuk menganalisis data hasil studi pelacakan lulusan dimana hasil yang didapatkan adalah didapatkan jumlah *cluster* terbaik yaitu dua menurut *Davies-Bouldin Index* (DBI) dimana *cluster* utama terdiri dari lulusan yang ketika mendapatkan pekerjaan pertamanya memiliki masa tunggu yang relatif singkat, yaitu dalam rentang waktu nol sampai dengan enam bulan. Sedangkan, sisanya memiliki masa tunggu yang relatif lebih lama ketika mendapatkan pekerjaan pertamanya, yaitu dalam rentang waktu enam bulan sampai dengan satu tahun.

Pada penelitian yang berjudul “Klasterisasi Lulusan Pendidikan Tinggi Vokasi Berbasis Hasil Studi Pelacakan Menggunakan Algoritma *K-Means*” (Utari & Sofiani, 2021), menjelaskan bahwa pendidikan vokasi kesekretarisan memiliki kompetensi kerja dan profil lulusan yang sangat spesifik dikarenakan program pendidikan ini bersifat khas dan unik yang mana belum tersedia di jenjang pendidikan strata satu. Namun, pada kenyataannya, pendidikan vokasi kesekretarisan memiliki lulusan yang bekerja di berbagai bidang dikarenakan posisi sekretaris sendiri juga bisa diisi oleh lulusan pendidikan non-sekretari. Maka dari itu, hal tersebut menjadi tantangan bagi pengelola dan penyelenggara program studi untuk mengetahui tentang profil lulusan program studi vokasi kesekretarisan. Hasil dari penelitian menggunakan metode *K-means Clustering* ini adalah lulusan dikelompokkan ke dalam 6 *cluster* yang memiliki karakteristik yang berbeda seperti tingkat pendapatan, waktu tunggu kerja, dan relevansi pekerjaan dengan pendidikan. Hasil dari penelitian ini nantinya dapat digunakan oleh pemangku kepentingan untuk mengembangkan kebijakan dan strategi yang lebih baik.

Pada penelitian yang berjudul “Penerapan *Data Mining K-Means Clustering* untuk Analisa Penjualan pada Toko *Fashion Hijab Banten*” (Normah, et al., 2021), menjelaskan bahwa industri *fashion* hijab di Indonesia, termasuk di daerah banten, mengalami persaingan yang sangat ketat akibat pertumbuhan yang pesat. Maka dari itu, untuk membantu toko memahami pola penjualan yang lebih mendalam, mengidentifikasi segmen pelanggan, serta merancang strategi yang efektif untuk pemasaran, dilakukanlah analisis *cluster* dengan menggunakan metode *K-Means Clustering*. Hasil dari penelitian ini adalah didapatkan tiga *cluster* dimana setiap *cluster* berisi sekelompok pelanggan yang memiliki kesamaan dalam pola pembelian mereka, seperti kesamaan preferensi terhadap jenis produk, kesamaan waktu pembelian, dan kesamaan respons terhadap promosi.

Pada penelitian yang berjudul “*The Application of K-Means Clustering for Province Clustering in Indonesia of The Risk of The Covid-19 Pandemic Based on Covid-19 Data*” (Abdullah, et al., 2022), menjelaskan bahwa tingkat penyebaran Covid-19 beserta dampaknya bervariasi di berbagai provinsi di Indonesia saat pandemi Covid-19 tengah mempengaruhi dunia. Sehingga untuk membantu pemerintah dalam menentukan prioritas serta strategi penanganan yang efektif pada setiap wilayah, maka penting untuk dilakukan pengklasifikasian provinsi-provinsi berdasarkan risiko pandemi. Pengelompokan dilakukan berdasarkan jumlah kasus positif, tingkat kesembuhan, angka kematian, serta jumlah tes yang telah dilakukan. Dengan menggunakan algoritma *K-means Clustering*, hasil penelitian didapatkan sebanyak tiga *cluster* dimana *cluster* tinggi risiko memiliki jumlah kasus positif yang tinggi, tingkat kematian yang tinggi, dan kapasitas tes yang rendah. *Cluster* menengah risiko memiliki jumlah kasus positif yang sedang, tingkat kematian yang sedang, dan kapasitas tes yang lebih baik dibandingkan *cluster* tinggi risiko. Sedangkan, *cluster* rendah risiko memiliki jumlah kasus positif yang rendah, tingkat kematian yang rendah, dan kapasitas tes yang baik. Sehingga provinsi yang terdapat pada *cluster* tinggi risiko akan diprioritaskan untuk dilakukan strategi penanganan.

Pada penelitian yang berjudul “*Implementation of Clustering Using K-Means Method to Determine Nutritional Status*” (Nagari & Inayati, 2020), menjelaskan bahwa salah satu indikator penting dalam kesehatan masyarakat adalah status gizi dimana status gizi dapat mencerminkan kecukupan asupan nutrisi serta risiko terkait dengan malnutrisi ataupun obesitas. Penentuan status gizi suatu individu perlu dilakukan sehingga nantinya dapat dilakukan intervensi kesehatan apabila diperlukan, Untuk melakukan intervensi kesehatan yang lebih

tepat sasaran, maka dilakukanlah identifikasi pola atau kelompok yang berbeda dalam hal status gizi dengan menggunakan metode *K-means Clustering*. Hasil dari penelitian ini adalah didapatkan 3 cluster dimana cluster 1 dikategorikan sebagai *cluster* dengan gizi baik yang berisi individu dengan BMI yang berada dalam rentang normal, asupan kalori seimbang, dan Tingkat aktivitas fisik yang cukup. *Cluster 2* dikategorikan sebagai *cluster* dengan gizi kurang yang berisi individu dengan BMI yang rendah, asupan kalori yang kurang, dan tanda-tanda malnutrisi. *Cluster 3* dikategorikan sebagai *cluster* obesitas yang berisi individu dengan BMI yang tinggi, asupan kalori yang berlebihan, dan sering kali disertai dengan risiko kesehatan terkait obesitas, seperti hipertensi.

Pada penelitian yang berjudul “*Factors Influencing Carbon Emissions from China’s Electricity Industry: Analysis Using the Combination of LMDI and K-Means Clustering*” (He, et al., 2022), menjelaskan bahwa China merupakan salah satu negara dengan kontribusi emisi karbon terbesar di dunia. Emisi karbon tersebut berasal dari industri listrik yang ada di China. Terdapat beberapa faktor yang dapat memengaruhi emisi karbon dari industri listrik di China, diantaranya adalah struktur energi, efisiensi energi, dan permintaan listrik. Untuk mengidentifikasi pola emisi karbon yang signifikan dan menganalisis faktor-faktor utama yang mempengaruhi emisi karbon, maka dilakukan penelitian dengan menggunakan metode *K-Means Clustering* ini. Hasil dari penelitian ini adalah didapatkan beberapa *cluster* yang menunjukkan perbedaan pola dalam emisi karbon dimana *cluster* pertama menunjukkan intensitas karbon yang tinggi terkait penggunaan batu bara yang dominan, sementara *cluster* lainnya menunjukkan intensitas yang lebih rendah terkait peningkatan efisiensi energi dan penggunaan energi terbarukan.

Pada penelitian yang berjudul “*The Water Supply Association Analysis Method in Shenzhen Bases on K-Means Clustering Discretization and Apriori Algorithm*” (Liu, et al., 2021), menjelaskan bahwa seiring dengan pertumbuhan populasi dan industrialisasi, permintaan akan air terus meningkat. Maka dari itu, diperlukan pengelolaan sumber daya air yang efisien di kota-kota besar seperti di Shenzhen, Tiongkok. Untuk menemukan pola tersembunyi dalam data konsumsi air, digunakan kombinasi algoritma *K-means Clustering* dan *Apriori* sehingga dapat digunakan untuk meningkatkan efisiensi distribusi air dan mengoptimalkan pengelolaan sumber daya. Dari penelitian ini didapatkan hasil pengelompokan data konsumsi air berdasarkan kesamaan pola konsumsinya, *cluster* pertama yaitu *cluster* rumah tangga, memiliki tingkat konsumsi air yang tinggi. Sedangkan *cluster* kedua, yaitu *cluster* industri memiliki

tingkat konsumsi air yang lebih besar namun lebih stabil. Kemudian, dengan algoritma *Apriori* ditemukan bahwa terdapat hubungan yang kuat antara penggunaan air di area tertentu dalam sehari yang dapat memberikan wawasan tentang puncak permintaan air dan potensi efisiensi dalam distribusi air.

Pada penelitian yang berjudul “*Suspended Sediment Modeling Using Heuristic Regression Method Hybridized with K-Means Clustering*” (Adnan, et al., 2021), menjelaskan bahwa sedimen tersuspensi merupakan partikel padat yang terbawa oleh aliran air di Sungai atau badan air lainnya. Untuk pengelolaan sumber daya air, pengendalian erosi, dan mitigasi bencana alam, pemodelan sedimen tersuspensi penting untuk dilakukan. Dalam pemodelan sedimen tersuspensi ini, digunakan algoritma *K-Means Clustering* dan metode regresi *heuristic* sehingga dapat meningkatkan akurasi serta efisiensi prediksi. Hasil dari penelitian ini adalah pola yang lebih rinci dalam data sedimen tersuspensi dimana pola dibedakan berdasarkan yang terjadi selama periode curah hujan tinggi dan periode aliran rendah. Kemudian, data yang sudah dikelompokkan tersebut dapat digunakan untuk meningkatkan akurasi model regresi.

Pada penelitian yang berjudul “*Road Network Partitioning Method Based on Canopy-K-Means Clustering Algorithm*” (Lin & Xu, 2020), menjelaskan bahwa partisi jalan merupakan proses membagi jaringan jalan yang luas menjadi bagian-bagian yang lebih kecil dan homogen untuk memudahkan analisis, pengelolaan, dan perencanaan lalu lintas. Untuk manajemen lalu lintas, sistem navigasi, serta analisis data geospasial, partisi jaringan penting untuk dilakukan. Penelitian ini mengintegrasikan algoritma *K-Means* dan *Canopy* untuk meningkatkan efisiensi dan akurasi dalam partisi jaringan yang lebih kompleks. Hasil dari penelitian ini adalah didapatkan pembagian jaringan wilayah jalan menjadi empat subarea. Terdapat perbedaan kecil antara subarea 1, 2, dan 4. Namun, terdapat perbedaan besar di subarea 3.

Table 2. 1 Kajian Literatur

No.	Penulis	Research Focus		
		<i>K-Means Clustering</i>	Profil Lulusan	<i>Elbow Method</i>
1	(Zhang, et al., 2024)			
2	(Zebua, et al., 2022)		√	
3	(Bu, et al., 2020)			
4	(Ghosal, et al., 2020)			
5	(Gao, et al., 2021)			

No.	Penulis	Research Focus		
		<i>K-Means Clustering</i>	Profil Lulusan	<i>Elbow Method</i>
6	(Haryanto, et al., 2024)	√	√	√
7	(Dahnial, 2023)	√		
8	(Sutrisno, et al., 2023)	√	√	
9	(Utari & Sofiani, 2021)	√	√	
10	(Normah, et al., 2021)	√		
11	(Abdullah, et al., 2022)	√		√
12	(Nagari & Inayati, 2020)	√		
13	(He, et al., 2022)	√		
14	(Liu, et al., 2021)	√		
15	(Adnan, et al., 2021)	√		
16	(Lin & Xu, 2020)	√		
17	(Rahmatika, 2024)	√	√	√

Pada tabel 2.1, menunjukkan penelitian terdahulu yang memiliki fokus pada *K-Means Clustering*, *Elbow Method*, dan profil lulusan. Pada penelitian ini, peneliti akan melakukan analisis menggunakan *Elbow Method* dan *K-Means Clustering* untuk melakukan *clustering* profil lulusan Teknik Industri serta memberikan strategi pengembangan profil lulusan berdasarkan hasil *clustering*.

2.2 Landasan Teori

2.2.1 Profil Lulusan

Profil lulusan merupakan gambaran umum mengenai kompetensi, seperti *soft skill*, *hard skill*, sikap dan etika, kemampuan beradaptasi, serta kemampuan umum yang diharapkan ada pada setiap lulusan setelah menyelesaikan suatu program pendidikan di program studi tertentu. Pada hakikatnya, profil menurut Mulyani dalam Lisdarti (2021) merupakan pandangan sisi, garis besar, atau biografi dari diri seseorang atau kelompok yang memiliki usia yang sama. Profil lulusan merupakan gambaran umum mengenai kompetensi, seperti *soft skill*, *hard skill*, sikap dan etika, kemampuan beradaptasi, serta kemampuan umum yang diharapkan ada pada setiap lulusan setelah menyelesaikan suatu program pendidikan di program studi tertentu (Hidayat & Zainuddin, 2017). Profil lulusan ini biasanya disusun oleh institusi pendidikan, baik itu sekolah, perguruan tinggi,

atau universitas, untuk memberikan panduan kepada para mahasiswa mengenai tujuan akhir dari proses pembelajaran yang mereka jalani. Beberapa aspek yang biasanya tercakup dalam profil lulusan meliputi:

- *Hard Skill: Skill* atau kemampuan spesifik yang terkait dengan bidang studi atau profesi yang dipelajari, seperti keterampilan teknis, keterampilan menggunakan software, dan kemampuan bahasa (Maryana, et al., 2022).
- *Soft Skill: Skill* atau kemampuan non-teknis yang penting dalam dunia kerja dan kehidupan sehari-hari, seperti komunikasi, kerja sama tim, kepemimpinan, pemecahan masalah, dan manajemen waktu (Maryana, et al., 2022).

Profil lulusan dapat digunakan sebagai acuan bagi institusi pendidikan dalam perancangan kurikulum, metode pengajaran, dan evaluasi, serta sebagai informasi bagi para calon mahasiswa dan industri mengenai kualitas dan kompetensi yang dimiliki oleh lulusan dari suatu program studi.

2.2.2 *Knowledge Discovery in Database (KDD)*

Knowledge Discovery in Database (KDD) merupakan sebuah proses yang dilakukan pada suatu data besar atau basis data untuk menemukan pengetahuan yang berguna, seperti mengidentifikasi pola dalam data, mengidentifikasi hubungan, serta informasi yang sebelumnya tidak diketahui sehingga hal-hal tersebut dapat digunakan dalam membantu pengambilan keputusan (Widaningsih, 2019). Proses KDD sendiri terdapat beberapa tahapan yang perlu dilakukan, yang mana masing-masing tahapan memainkan peranan penting untuk menghasilkan pengetahuan yang berguna dan valid. Berikut ini merupakan tahapan-tahapan utama dari proses KDD menurut Tomar & Agarwal (2013):

1. *Data Selection*

Pada tahapan ini akan dilakukan pemilihan subset data yang relevan dari basis data yang lebih besar. Pemilihan subset data ini didasari oleh tujuan atau masalah yang ingin diselesaikan.

2. *Pre-processing*

Pada tahapan ini akan dilakukan pembersihan data seperti penghapusan data yang tidak konsisten, tidak relevan, ataupun hilang sehingga nantinya dapat meningkatkan akurasi hasil yang diperoleh serta meningkatkan kualitas data.

3. *Data Transformation*

Pada tahapan ini subset yang sudah dipilih dan melalui data cleaning sebelumnya kemudian akan diubah menjadi format yang sesuai untuk analisis dimana pada tahapan ini dapat melibatkan normalisasi, agregasi, atau pengkodean data.

4. *Data Mining*

Tahapan ini merupakan tahapan inti dalam proses KDD dimana akan diterapkan teknik-teknik analitik untuk menemukan pola atau model dalam data. Metode-metode yang sering digunakan dalam data mining antara lain adalah klusterisasi, klasifikasi, asosiasi, dan analisis regresi.

5. *Evaluation and Interpretation*

Pada tahapan ini akan dilakukan evaluasi pada pola atau model yang dihasilkan dari proses data mining berdasarkan kriteria tertentu seperti validitas, relevansi, dan kegunaannya sehingga hasilnya dapat diinterpretasikan untuk memastikan bahwa pengetahuan yang dihasilkan bermakna serta dapat digunakan dalam konteks praktis.

Knowledge Discovery in Database (KDD) ini dapat diterapkan di berbagai bidang untuk tujuan yang berbeda, misalnya dalam bidang bisnis dan pemasaran akan diterapkan untuk menganalisis perilaku pelanggan, menemukan pola pembelian, segmentasi pasar, dan analisis loyalitas pelanggan (Nurdiawan & Salim, 2018). Dalam bidang kesehatan akan diterapkan untuk membantu dalam menganalisis data pasien seperti menemukan pola yang berkaitan dengan diagnosis, perawatan, prediksi, ataupun klasifikasi penyakit (Maulana, et al., 2023). Selain itu, *Knowledge Discovery in Database (KDD)* juga dapat diterapkan untuk analisa tingkat kecelakaan lalu lintas dimana setelah melakukan pengumpulan data, pendekatan KDD akan digunakan untuk memperoleh informasi yang bisa mewakili masalah, dimana dapat diketahui ruas jalan yang mana yang tingkat frekuensi kecelakaannya paling sering terjadi serta penyebab terjadinya kecelakaan tersebut (Utomo, et al., 2023).

2.2.3 Data Mining

Data Mining merupakan serangkaian proses yang dilakukan untuk mengetahui pola, hubungan, serta informasi baru dari sejumlah besar data yang ada pada basis data atau format penyimpanan lainnya (Zebua, et al., 2022). *Data Mining* juga merupakan salah satu dari proses *Knowledge Discovery in Database* (KDD) yang memiliki fokus untuk mengidentifikasi pola menarik pada data yang tidak diketahui sebelumnya, bermanfaat, serta nantinya dapat dipahami oleh manusia (Meilani & Nurdiawan, 2023). Menurut (Metisen & Sari, 2015), proses *Data Mining* terdiri dari beberapa tahap, diantaranya adalah:

1. *Data Collection*

Pada tahapan ini akan dilakukan pengumpulan data data berbagai sumber seperti website, *database*, dan lainnya.

2. *Pre-processing*

Pada tahapan ini akan dilakukan pembersihan data, transformasi data, serta seleksi data sehingga nantinya data yang digunakan relevan dan berkualitas tinggi.

3. *Data Selection*

Pada tahapan ini akan dilakukan penentuan subset dari data yang sudah melalui tahap *pre-processing* sebelumnya untuk dilakukan analisis lebih lanjut.

4. *Transformation*

Pada tahapan ini akan dilakukan perubahan data menjadi format yang sesuai untuk proses *data mining*.

5. *Data Mining*

Pada tahapan ini akan dilakukan pengaplikasian algoritma untuk mengekstrak pola dalam data.

6. Evaluasi dan Interpretasi

Pada tahapan ini akan dilakukan analisis dan interpretasi hasil dari proses data mining sehingga informasi yang diperoleh dapat dipahami dan digunakan.

7. *Presentation*

Pada tahapan ini hasil dari data mining akan disajikan dalam format yang lebih mudah dipahami oleh *user* seperti grafik, tabel, ataupun laporan.

Dalam melakukan proses data mining, terdapat beberapa metode utama yang umum digunakan, metode data mining tersebut meliputi (Syaripudin, et al., 2013):

- *Classification*

Metode ini merupakan metode yang digunakan untuk mengklasifikasikan data ke dalam beberapa kategori dimana kategori tersebut telah ditentukan berdasarkan atributnya. Contoh algoritma yang digunakan untuk *classification* adalah *Decision Tree*, *Naïve Bayes*, dan *Support Vector Machine* (SVM).

- *Clustering*

Metode ini merupakan metode yang digunakan untuk mengelompokkan data berdasarkan kedekatan atau kesamaan antara data-data tersebut. Contoh algoritma yang populer untuk *clustering* adalah *K-Means*, *Hierarchical Clustering*, dan *DBSCAN*.

- *Association Rule*

Metode ini merupakan metode yang digunakan untuk menemukan hubungan antar item dalam suatu basis data. Contoh algoritma yang sering digunakan untuk *association rule* adalah *Apriori* dan *FP-Growth*.

Data mining ini dapat diterapkan di berbagai bidang untuk tujuan yang berbeda, misalnya dalam bidang pemasaran akan diterapkan untuk segmentasi pelanggan, analisis keranjang belanja, dan prediksi perilaku konsumen (Nurdiawan & Salim, 2018). Dalam bidang kesehatan akan diterapkan untuk memprediksi penyakit, analisis genom, dan peningkatan kualitas pelayanan medis (Meilani & Nurdiawan, 2023). Dalam bidang telekomunikasi akan diterapkan untuk analisis *churn*, optimasi jaringan, dan deteksi penipuan. Dalam bidang *e-commerce* akan diterapkan untuk mengidentifikasi tren dan preferensi pelanggan, meningkatkan *targeting* promosi, dan mengoptimalkan strategi pemasaran (Apriana & Yuliansyah, 2024).

2.2.4 *Clustering*

Clustering merupakan salah satu metode dalam *data mining* dimana pada metode ini dilakukan pengelompokan pada sekumpulan objek ke dalam beberapa *cluster* sedemikian rupa sehingga objek-objek yang ada di dalam satu *cluster* memiliki kemiripan yang tinggi satu sama lain serta memiliki perbedaan yang signifikan dengan objek-objek yang ada di dalam *cluster* lain (Wakhidah, 2010). *Clustering* memiliki beberapa tujuan dan manfaat, antara lain adalah untuk mengidentifikasi struktur data dimana dengan adanya *clustering* lebih memungkinkan untuk pengguna mengidentifikasi kelompok ataupun pola yang tidak terlihat secara langsung serta membantu dalam menemukan struktur yang terdapat pada data tidak berlabel. Manfaat lain dari *clustering* adalah pengurangan dimensi dimana dengan adanya *clustering* kompleksitas data

dapat dikurangi juga dapat membagi data ke dalam beberapa *cluster* sehingga memudahkan analisis lebih lanjut. Selain itu, *clustering* juga bermanfaat untuk segmentasi pasar dimana dengan adanya *clustering* pelanggan dapat dikelompokkan berdasarkan perilaku ataupun preferensi mereka sehingga memungkinkan strategi pemasaran yang lebih terarah dan efektif (Wakhidah, 2010). Dalam melakukan *clustering*, terdapat beberapa teknik yang dapat digunakan menurut Wakhidah (2010), antara lain adalah:

- *Partitioning Clustering*

Teknik *clustering* ini bekerja dengan cara membagi data ke dalam beberapa cluster yang sebelumnya telah ditentukan. *K-Means Clustering* merupakan metode yang paling terkenal pada kategori ini dimana metode *K-Means* akan mengelompokkan data dengan cara meminimalkan varians dalam setiap *cluster*.

- *Hierarchical Clustering*

Teknik *clustering* ini bekerja dengan cara membangun hierarki dari *cluster* dengan dua pendekatan utama, yaitu *agglomerative* yang merupakan pendekatan *bottom-up* dan *divisive* yang merupakan pendekatan *top-down*. Hasil dari *Hierarchical Clustering* sering kali digambarkan dalam bentuk dendogram.

- *Density-Based Clustering*

Teknik *clustering* ini bekerja dengan cara mengelompokkan data berdasarkan kepadatan data di sekitarnya. DBSCAN (*Density-Based Spatial Clustering of Applications with Noise*) merupakan metode yang paling terkenal pada kategori ini dimana metode DBSCAN dapat mengidentifikasi *cluster* dalam suatu data dengan bentuk yang arbitrer serta dapat menangani *outlier* dengan baik.

- *Model-Based Clustering*

Teknik *clustering* ini bekerja dengan cara mengasumsikan bahwa data dihasilkan oleh sekumpulan model probabilistik, contohnya yaitu pada model *Gaussian Mixture*. Pada *Model-Based Clustering*, setiap *cluster* direpresentasikan oleh model probabilistik tertentu.

- *Fuzzy Clustering*

Teknik *clustering* ini bekerja dengan cara yang berbeda dari metode *clustering* tradisional dimana *Fuzzy Clustering* mengelompokkan setiap objek hanya ke satu *cluster*. Salah satu contoh *Fuzzy Clustering* adalah *Fuzzy C-Means* yang memungkinkan

sebuah objek menjadi anggota lebih dari satu *cluster* dengan derajat keanggotaan tertentu.

2.2.5 *K-Means Clustering*

K-Means Clustering merupakan salah satu metode dalam *clustering* yang paling sederhana dan paling populer yang digunakan dalam analisis data dimana metode ini merupakan salah satu metode yang paling populer dan memiliki keunggulan seperti kemudahan dan kemampuannya untuk meng*cluster* data yang berukuran besar dan data *outlier* dengan sangat cepat (Metisen & Sari, 2015). Metode *K-Means* merupakan metode iteratif yang melakukan pembagian dataset menjadi *k cluster* dimana setiap kelompok *cluster* memiliki *centroid* yang merupakan representasi dari kelompok *cluster* tersebut (Haryanto, et al., 2024). Terdapat beberapa langkah yang dilakukan dalam proses *clustering* dengan menggunakan *K-Means Clustering*. Berikut ini merupakan langkah-langkah dalam *K-Means Clustering* menurut (Maori & Evanita, 2023).

1. Menentukan nilai *k* sebagai jumlah *cluster* yang akan dibuat.
2. Menentukan titik pusat *cluster* (*k centroid*) awal secara acak.
3. Menghitung jarak setiap data ke masing-masing *centroid* dengan menggunakan rumus korelasi antar dua objek, seperti yang dilihat pada rumus *Euclidian Distance*.

$$D_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + \dots + (n_i - n_j)^2} \quad (2. 1)$$

Dimana:

D_{ij} = jarak antara *i* dan *j*

x_i = koordinat *x* objek

x_j = koordinat *x* pusat

y_i = koordinat *y* objek

y_j = koordinat *y* pusat

4. Menggabungkan setiap data yang dilihat berdasarkan jarak terdekat antara data dan *cluster*.
5. Menentukan *centroid* baru dengan menghitung nilai rata-rata berdasarkan data pada *cluster* yang serupa. Berikut ini merupakan rumus untuk perhitungannya.

$$C_i = \frac{x_1 + x_2 + x_3 + \dots + x_n}{\sum x} \quad (2. 2)$$

Dimana:

x_1 = nilai *record* data ke-1

x_2 = nilai *record* data ke-2

$\sum x$ = jumlah *record* data

K-Means Clustering dapat diterapkan di berbagai bidang untuk tujuan yang berbeda, misalnya dalam bidang segmentasi pelanggan akan diterapkan untuk mengelompokkan pelanggan berdasarkan demografi dan perilaku pelanggan ketika berbelanja sehingga memungkinkan untuk membuat strategi yang lebih terarah. Selain itu, *K-Means Clustering* juga dapat digunakan dalam bidang pengelolaan informasi yaitu untuk mengelompokkan dokumen dimana pengelompokan dilakukan berdasarkan kemiripan konten sehingga dapat berguna dalam pencarian informasi dan pengelompokan topik. Pada analisis citra, *K-Means Clustering* digunakan untuk segmentasi gambar dimana gambar akan dikelompokkan berdasarkan warna atau intensitas untuk mengidentifikasi objek ataupun area tertentu pada gambar. *K-Means Clustering* juga dapat digunakan untuk mendeteksi anomali dimana akan dilakukan identifikasi mengenai data yang tidak cocok dengan *cluster* utama (Prianto & Bunyamin, 2020).

2.2.6 *Elbow Method*

Elbow Method merupakan salah satu teknik yang dapat digunakan untuk menentukan jumlah *cluster* yang tepat untuk digunakan pada algoritma seperti *K-Means* (Rahman, et al., 2024). Metode ini membantu menemukan keseimbangan yang pas antara jumlah *cluster* yang tidak terlalu banyak dan tingkat kesalahan yang rendah sehingga dapat memudahkan memilih jumlah *cluster* yang optimal. Cara menentukan jumlah *cluster* dengan *Elbow Method* adalah dengan menghitung dan menggambarkan *Within-Cluster Sum of Squares* (WCSS) atau *Sum of Squared Errors* (SSE) untuk setiap jumlah *cluster* yang dicoba. Tujuan utamanya adalah untuk menemukan titik di grafik yang mana menambahkan *cluster* tidak lagi memberikan pengurangan yang signifikan pada WCSS atau SSE. Titik tersebut nantinya akan terlihat seperti siku pada grafik (Maori & Evanita, 2023). Berikut ini merupakan rumus perhitungan SSE:

$$SSE = \sum_{k=1}^k \sum_{x_i \in C_k} (X_i - \varphi_k)^2 \quad (2.3)$$

Dimana:

C_k = k *cluster* yang terbentuk

k = banyak *cluster*

C_k = data x pada fitur ke- i

φ_k = rata-rata k cluster pada nilai k ($k=1, 2, 3, \dots, k$)

Dalam melakukan *clustering* dengan *K-Means*, setiap data ditempatkan ke dalam *cluster* yang *centroid*-nya paling dekat dengan tujuan untuk meminimalkan WCSS, yaitu jumlah total jarak kuadrat antara setiap data dan pusat *cluster*-nya. Terdapat beberapa langkah yang dilakukan dalam proses penentuan jumlah *cluster* optimal dengan menggunakan *Elbow Method*. Berikut ini merupakan langkah-langkah dalam *Elbow Method* menurut (Prianto & Bunyamin, 2020).

1. Menginisialisasikan nilai awal k .
2. Menaikkan nilai k .
3. Menghitung nilai hasil *sum of square error* dari setiap nilai k .
4. Menetapkan nilai k yang membentuk siku.

BAB III

METODE PENELITIAN

3.1 Jenis Penelitian

Penelitian ini bertujuan untuk menganalisis tentang profil lulusan seperti apa yang diharapkan dari lulusan teknik industri dengan menggunakan metode *K-Means Clustering* yang mana metode ini termasuk ke dalam metode kuantitatif. Pendekatan kuantitatif sendiri dipilih karena dapat membantu untuk analisis profil lulusan yang lebih mendalam dimana profil lulusan akan dikelompokkan ke dalam beberapa *cluster*.

3.2 Objek dan Subjek Penelitian

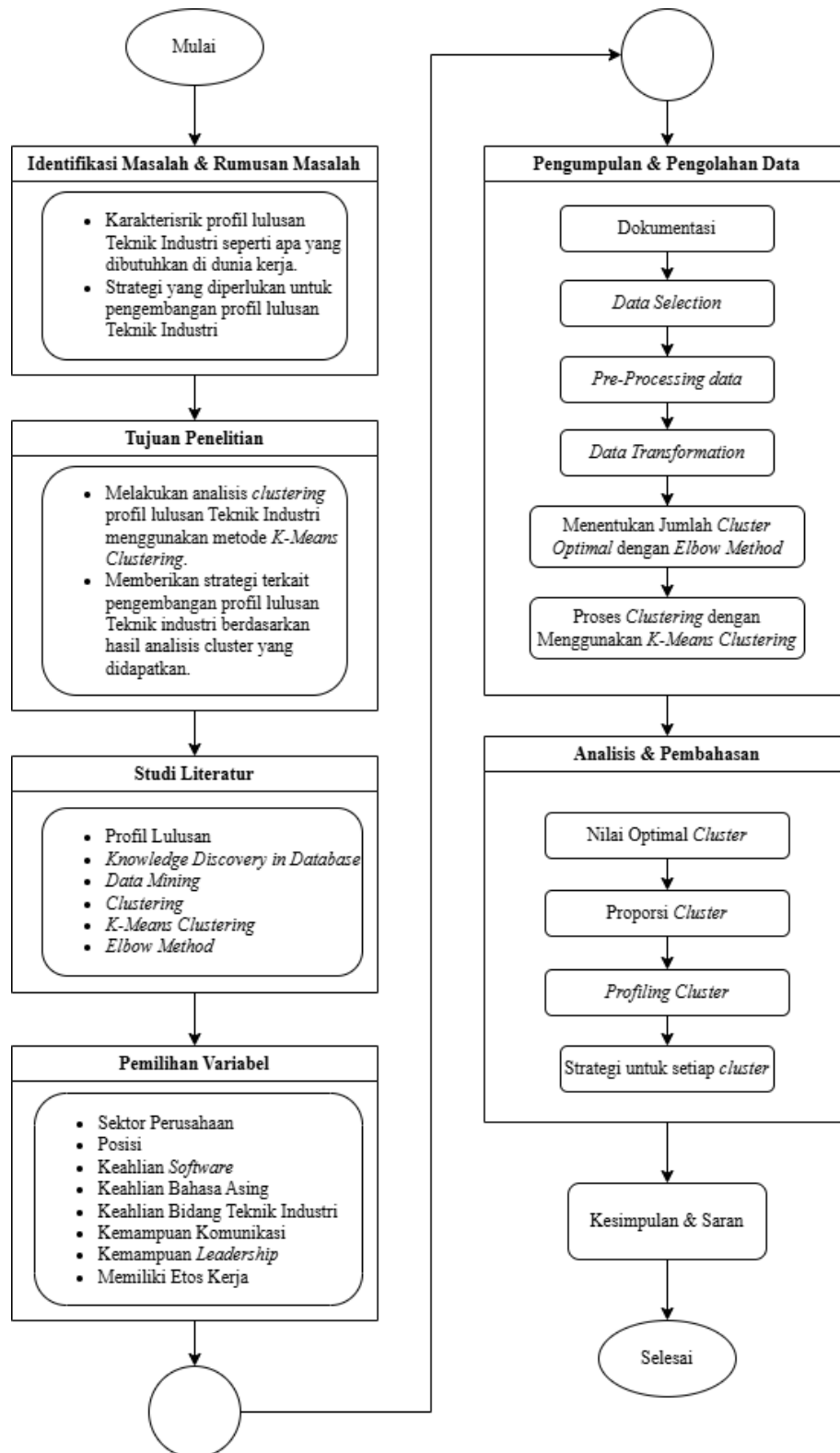
Objek yang diteliti pada penelitian ini adalah profil lulusan teknik industri yang diharapkan ada pada setiap lulusan teknik industri di dunia kerja. Penelitian ini berfokus pada identifikasi kriteria lulusan teknik industri seperti apa yang dicari oleh perusahaan untuk mengisi kekosongan posisi di perusahaan tersebut sehingga dari hasil identifikasi yang didapatkan akan dibuat strategi untuk pengembangan profil lulusan teknik industri. Adapun subjek pada penelitian ini merupakan perusahaan-perusahaan yang menawarkan lowongan pekerjaan pada situs pencari kerja dimana kriteria yang dibutuhkan oleh perusahaan untuk setiap posisi yang ditawarkan tersebut yang nantinya akan dijadikan data untuk melakukan penelitian.

3.3 Metode Pengumpulan Data

Metode pengumpulan data pada penelitian ini dilakukan dengan menggunakan metode dokumentasi. Metode dokumentasi merupakan salah satu metode pengumpulan data sekunder dimana metode ini digunakan untuk mendapatkan data yang valid dan sesuai dengan informasi yang dibutuhkan oleh peneliti. Pada penelitian ini, data yang digunakan adalah data sekunder dimana data diperoleh dari situs pencari kerja seperti Jobstreet, Indeed, Loker.id, Glints, Kupu.id, dan LinkedIn. Data yang diambil berupa perusahaan-perusahaan yang membuka dan *posting* lowongan kerja untuk lulusan teknik industri pada situs pencari kerja, sektor perusahaan, posisi yang ditawarkan, serta *skill* yang dibutuhkan sebagai syarat untuk mengisi posisi tersebut.

3.4 Alur Penelitian

Berikut ini merupakan diagram alir yang menjelaskan tentang alur penelitian:



Gambar 3.1 Alur penelitian

Berikut ini merupakan penjelasan mengenai diagram alur penelitian di atas:

1. Mulai

Pada tahap ini, peneliti akan mempersiapkan dan memulai penelitian.

2. Identifikasi Masalah dan Perumusan Masalah

Pada tahap ini, akan dilakukan identifikasi mengenai masalah yang akan diteliti, dalam hal ini yaitu terkait kesenjangan antara profil lulusan teknik industri yang diharapkan di dunia kerja dengan kenyataannya yang ada serta dampak yang bisa terjadi apabila masalah tersebut tidak diatasi. Selanjutnya, akan dilakukan perumusan masalah yang ingin dipecahkan berdasarkan identifikasi masalah yang telah dilakukan sebelumnya.

3. Tujuan Penelitian

Pada tahap ini akan dilakukan penetapan tujuan penelitian yang ingin dicapai peneliti berdasarkan rumusan masalah yang telah dibuat sehingga penelitian yang dilakukan akan lebih fokus dan terarah. Tujuan yang ditetapkan pada penelitian ini yaitu melakukan analisis *clustering* profil lulusan teknik industri menggunakan metode *K-Means Clustering* dan mengidentifikasi karakteristik di tiap *clusternya* serta memberikan strategi terkait pengembangan profil lulusan berdasarkan hasil analisis *cluster* yang telah didapatkan.

4. Studi Literatur

Pada tahap ini akan dilakukan pengkajian mengenai literatur terdahulu yang memiliki kesamaan topik dengan penelitian ini. Literatur terdahulu dapat diperoleh dari berbagai sumber, seperti internet, buku, serta jurnal. Studi literatur dibutuhkan sebagai bahan pendukung untuk memecahkan permasalahan penelitian dimana dengan studi literatur peneliti dapat memperoleh pemahaman yang mendalam tentang topik penelitian. Pada penelitian ini, digunakan beberapa landasan teori, seperti Profil Lulusan, *Knowledge Discovery in Database (KDD)*, *Data Mining*, *Clustering*, *K-Means Clustering*, dan *Elbow Method*. Selain itu, studi literatur juga digunakan untuk menetapkan variabel apa yang dibutuhkan dalam penelitian ini.

5. Pemilihan Variabel Penelitian

Pada tahapan ini akan dipilih variabel-variabel yang diperlukan untuk pengolahan data. Variabel-variabel inilah yang akan menjadi fokus dalam penelitian yang sesuai dengan identifikasi dan rumusan masalah. Metode studi literatur digunakan untuk membantu dalam penetapan variabel penelitian ini. Variabel yang dibutuhkan pada penelitian ini antara lain adalah:

Tabel 3. 1 Deskripsi Variabel

Variabel	Deskripsi	Sumber
Sektor	Variabel sektor dapat menunjukkan keterkaitan antara profil lulusan teknik industri dengan tempat mereka bekerja dimana dengan variabel sektor nantinya dapat dilihat sektor perusahaan apa yang paling banyak membutuhkan lulusan teknik industri.	Keperluan Prodi Teknik Industri
Posisi	Variabel ini menggambarkan lowongan atau jabatan yang ditawarkan perusahaan untuk lulusan teknik industri dimana dengan variabel ini nantinya dapat dilihat posisi apa saja yang paling banyak ditawarkan untuk lulusan teknik industri.	(Zebua, et al., 2022)
Keahlian <i>Software</i>	Variabel ini merupakan salah satu kualifikasi yang disyaratkan oleh perusahaan ketika membuka lowongan kerja dimana variabel ini berfokus pada <i>software</i> apa saja yang perlu dikuasai untuk bisa diterima di posisi yang ditawarkan.	(Mulyanan, et al., 2023)
Keahlian Bahasa Asing	Variabel ini merupakan salah satu kualifikasi yang disyaratkan oleh perusahaan, terutama perusahaan multinasional, ketika membuka lowongan kerja dimana variabel ini berfokus pada Bahasa Inggris, Bahasa Mandarin, dan lainnya yang relevan dengan dunia kerja.	(Othman, et al., 2018)
Keahlian Bidang Teknik Industri	Variabel ini merupakan salah satu kualifikasi yang disyaratkan oleh perusahaan ketika membuka lowongan kerja dimana variabel berfokus pada bidang ergonomi, sistem produksi, manajemen industri dan <i>operational research</i> .	Keperluan Prodi Teknik Industri
Kemampuan Komunikasi	Variabel ini merupakan salah satu kualifikasi yang disyaratkan oleh perusahaan ketika membuka lowongan kerja dimana variabel ini mencakup kemampuan komunikasi secara verbal maupun nonverbal.	(Othman, et al., 2018)
Kemampuan <i>Leadership</i>	Variabel ini merupakan salah satu kualifikasi yang disyaratkan oleh perusahaan ketika membuka lowongan kerja dimana variabel ini mencakup kemampuan kepemimpinan, <i>critical thinking</i> , memiliki kreativitas yang tinggi, dan <i>decision making</i> .	(Othman, et al., 2018)
Memiliki Etos Kerja	Variabel ini merupakan salah satu kualifikasi yang disyaratkan oleh perusahaan ketika membuka lowongan kerja dimana variabel	(Mulyanan, et al., 2023)

Variabel	Deskripsi	Sumber
	berfokus pada sikap kerja, kedisiplinan, serta inisiatif untuk belajar dan berkembang.	

- Sektor Perusahaan: Pada variabel ini, sektor perusahaan dikategorikan menjadi 4 sektor, yaitu sektor manufaktur, jasa, dagang, dan konstruksi.
- Posisi: Pada variabel ini, posisi lowongan dikategorikan menjadi 4, yaitu posisi di bagian operasional produksi, *marketing*, IT, dan *Managerial*.
- Kualifikasi: Pada variabel kualifikasi ini terbagi menjadi 2, yaitu kualifikasi *hard skill* dan *soft skill* dimana kualifikasi *hard skill* terdiri dari keahlian *software*, keahlian bahasa asing, dan keahlian bidang Teknik industri. Sedangkan kualifikasi *soft skill* terdiri dari kemampuan komunikasi, kemampuan *leadership*, dan memiliki etos kerja.

6. Pengumpulan Data

Pada tahapan ini akan dilakukan pengumpulan data yang dibutuhkan dalam penelitian, dimana data dikumpulkan dengan mengambil informasi lowongan kerja yang dibuka untuk lulusan teknik industri dan *diposting* oleh perusahaan pada situs pencari kerja. Pengumpulan data dilakukan dengan metode dokumentasi yang merupakan metode pengumpulan data sekunder.

7. Pengolahan Data

Pada tahapan ini akan dilakukan pengolahan data dengan menggunakan data yang sudah dikumpulkan dari situs pencari kerja. Pengolahan data akan dilakukan untuk menjawab tujuan penelitian yang telah ditentukan. Berikut ini merupakan uraian tahapan pengolahan data untuk nantinya dilakukan identifikasi karakteristik di tiap *clusternya*:

Pre-Processing Data

Pada tahapan ini, data yang sudah terkumpul sebelumnya akan melalui proses *cleaning data* atau pembersihan data yang tidak sesuai untuk penelitian dan dianggap mengganggu. Tahap *cleaning data* ini perlu dan penting untuk dilakukan untuk mendapatkan data yang lebih berkualitas sehingga hasil pengolahan data pun akan lebih optimal.

Data Transformation

Pada tahapan ini, data yang sudah melalui tahap *data cleaning* akan ditransformasikan ke dalam format yang lebih seragam dan sesuai sehingga dapat lebih mudah digunakan pada tahap pengolahan data dimana variabel data posisi serta kualifikasi *hard skill* dan *soft skill* akan dibagi menjadi beberapa kelompok untuk membuat data menjadi lebih seragam dan memudahkan dalam pengolahan data. Pada variabel posisi, akan dikelompokkan menjadi 4 posisi yaitu posisi dalam bidang operasional produksi, *marketing*, IT, dan *managerial*. Kemudian untuk variabel kualifikasi *hard skill* dan *soft skill*, akan dikelompokkan menjadi 6 kualifikasi yaitu keahlian *software*, keahlian bahasa asing, keahlian bidang teknik industri, kemampuan komunikasi, kemampuan *leadership*, dan memiliki etos kerja. Selain itu transformasi akan dilakukan untuk mengubah format data *string* menjadi data *numeric* sehingga memudahkan untuk proses *data mining*.

Menemukan *Cluster* Optimal Menggunakan *Elbow Method*

Pada tahapan ini, *Elbow Method* digunakan untuk menemukan jumlah cluster optimal dimana nilai *cluster* optimal dapat dilihat pada grafik *Elbow Method* dimana grafik ini terdiri dari beberapa titik yang merupakan nilai *Sum of Square Error* (SSE) yang akan membentuk lengkungan seperti siku (Maori & Evanita, 2023). Berikut ini merupakan rumus perhitungan SSE:

$$SSE = \sum_{k=1}^k \sum_{X_i \in C_k} (X_i - \varphi_k)^2 \quad (3.1)$$

Dimana:

$C_k = k$ cluster yang terbentuk

k = banyak cluster

C_k = data x pada fitur ke-i

φ_k = rata-rata k cluster pada nilai k ($k=1, 2, 3, \dots, k$)

Berikut ini merupakan langkah-langkah perhitungan SSE menurut Prianto & Bunyamin (2020).

5. Menginialisasikan nilai awal k .
6. Menaikkan nilai k .
7. Menghitung nilai hasil *sum of square error* dari setiap nilai k .
8. Menetapkan nilai k yang membentuk siku.

K-Means Clustering

Setelah menetapkan jumlah *cluster optimal* menggunakan *Elbow Method*, maka selanjutnya akan dilakukan clustering dengan menggunakan metode *K-Means Clustering*. Pengolahan data ini akan dilakukan dengan bantuan *software SPSS*. Berikut ini merupakan tahapan dalam melakukan *K-Means Clustering* menurut Maori (2023).

1. Menentukan nilai k sebagai jumlah *cluster* yang akan dibuat.
2. Menentukan titik pusat *cluster* (k *centroid*) awal secara acak.
3. Menghitung jarak setiap data ke masing-masing *centroid* dengan menggunakan rumus korelasi antar dua objek, seperti yang dilihat pada rumus *Eucledian Distance*.

$$D_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + \dots + (n_i - n_j)^2} \quad (3. 2)$$

Dimana:

D_{ij} = jarak antara i dan j

x_i = koordinat x objek

x_j = koordinat x pusat

y_i = koordinat y objek

y_j = koordinat y pusat

4. Menggabungkan setiap data yang dilihat berdasarkan jarak terdekat antara data dan *cluster*.
5. Menentukan *centroid* baru dengan menghitung nilai rata-rata berdasarkan data pada *cluster* yang serupa. Berikut ini merupakan rumus untuk perhitungannya.

$$C_i = \frac{x_1 + x_2 + x_3 + \dots + x_n}{\sum x} \quad (3. 3)$$

Dimana:

x_1 = nilai *record* data ke-1

x_2 = nilai *record* data ke-2

$\sum x$ = jumlah *record* data

Output

Output yang terbentuk dari pengolahan data dengan *K-Means Clustering* adalah pengelompokan perusahaan-perusahaan ke dalam setiap *cluster* yang ditentukan serta karakteristik di setiap *clusternya* dimana dari *output* ini nantinya akan dirancang strategi pengembangan profil lulusan teknik industri.

8. Analisis & Pembahasan

Pada tahapan ini akan dilakukan analisis dan pembahasan mengenai hasil dari pengumpulan serta pengolahan data yang telah dilakukan sebelumnya yang sesuai dengan kedua tujuan penelitian. Pada tahapan ini pula akan diberikan strategi yang tepat berdasarkan hasil analisis *cluster*.

9. Kesimpulan & Saran

Pada tahapan ini akan diambil kesimpulan dari analisis dan pembahasan yang telah dilakukan pada tahapan sebelumnya. Kesimpulan ini berisi poin-poin yang menjawab tujuan penelitian serta akan diberikan saran sebagai perbaikan dan masukan untuk penelitian selanjutnya.

10. Selesai

Tahapan ini merupakan tanda selesainya penelitian yang dilakukan.

BAB IV

PENGUMPULAN DAN PENGOLAHAN DATA

4.1 Data Selection

Sebelum melakukan pengolahan data, langkah pertama yang perlu dilakukan adalah *data selection* atau pemilihan data dimana tidak semua data yang diperoleh dapat digunakan dalam proses pengolahan data. Pengumpulan data dilakukan dengan metode dokumentasi dimana data yang diambil merupakan data lowongan kerja yang diposting oleh perusahaan di situs pencari kerja seperti Jobstreet, Indeed, Loker.id, Glints, Kupu.id, dan LinkedIn. Berikut ini merupakan situs pencari kerja yang menjadi sumber data beserta jumlah data per situsnya.

Tabel 4. 1 Sumber Data

Sumber	Jumlah Data
Jobstreet	155
Indeed	94
Loker.id	51
Kupu.id	11
Glints	6
LinkedIn	87

Data lowongan kerja yang diambil hanya berkisar pada lowongan kerja yang dibuka untuk lulusan teknik industri dengan total data yang terkumpul untuk penelitian ini adalah sebanyak 404 data lowongan yang berasal dari berbagai jenis perusahaan. Berikut ini merupakan tabel jenis perusahaan beserta jumlah perusahaan dan jumlah lowongan pekerjaannya.

Tabel 4. 2 Jumlah Perusahaan dan Lowongan Berdasarkan Jenisnya

Jenis Perusahaan	Jumlah Perusahaan	Jumlah Lowongan
Multinasional	37	92
BUMN	3	9
Swasta Nasional	250	278
Swasta Asing	9	11
UMKM	11	13
Koperasi	1	1
Jumlah	301	404

Dari tabel di atas, dapat dilihat terdapat enam jenis perusahaan, yaitu perusahaan multinasional yang merupakan perusahaan besar yang beroperasi di banyak negara karena jangkauannya yang sangat luas, Badan Usaha Milik Negara (BUMN) yang sesuai dengan namanya merupakan perusahaan yang dimiliki oleh negara atau pemerintah baik sepenuhnya, sebagian besar, ataupun sebagian kecil sehingga pemerintah dapat mengontrol perusahaan tersebut, swasta nasional yang merupakan perusahaan yang bukan milik pemerintah dan modalnya dikuasai sepenuhnya oleh masyarakat dalam negeri, swasta asing yang merupakan perusahaan yang beroperasi di Indonesia namun modalnya berasal dari luar negeri, Usaha Mikro Kecil Menengah (UMKM) yang merupakan usaha yang dilakukan oleh individu ataupun sekelompok orang dimana usaha produktif ini memiliki peranan penting dalam perekonomian Indonesia, dan koperasi yang merupakan badan usaha yang didirikan oleh badan hukum koperasi untuk menjalankan usaha bersama demi kepentingan bersama. Untuk perusahaan multinasional beberapa contoh perusahaannya adalah PT Unilever Indonesia Tbk, PT Coca-Cola Amatil Indonesia, dan PT Chevron Pacific Indonesia. Untuk BUMN, contoh perusahaannya adalah PT Bank Mandiri Tbk, PT Aerotrans Services Indonesia (Garuda Indonesia Group), dan PT Krakatau Steel Persero Tbk. Untuk swasta nasional beberapa contoh perusahaannya adalah PT Sri Indah Labetama, PT Ungaran Wana Karya, dan PT Furindo Sakti Sejahtera. Untuk swasta asing contohnya adalah PT Newtrend Nutrition Ingridient, PT Changhong Electric Indonesia, dan Toppan Plasindo Lestari. Untuk UMKM, contohnya adalah DMC Interior, Chocolate Monggo, dan Factory Print. Sedangkan untuk koperasi yaitu adalah Koperasi Jasa Amanah LPPOM MUI.

Untuk menganalisis *clustering* profil lulusan teknik industri seperti apa yang diharapkan di dunia kerja dengan menggunakan metode *K-Means Clustering*, terdapat beberapa variabel yang digunakan, yaitu sektor perusahaan, posisi lowongan, dan *skill* yang dibutuhkan untuk mengisi posisi tersebut. *Skill* dapat berupa *hard skill* seperti keahlian *software*, keahlian bahasa asing, dan keahlian bidang teknik industri dan *soft skill* seperti kemampuan komunikasi, kemampuan *leadership*, dan memiliki etos kerja. Dalam pemrosesan data, variabel gaji tidak dipertimbangkan dalam analisis sehingga data tidak digunakan dalam proses pengolahan data. Proses *data selection* ini dilakukan secara manual menggunakan *software excel*.

4.2 Pre-Processing Data

Setelah selesai melakukan proses *data selection*, selanjutnya data akan masuk ke proses *pre-processing data* dimana pada proses data mentah dipersiapkan terlebih dahulu dan akan melalui proses seperti *data cleaning* atau pembersihan data dengan tujuan agar tidak ada data yang tidak konsisten, tidak lengkap, mengurangi *noise*, ataupun mengilangkan data duplikasi sehingga akan meningkatkan kualitas data dan data menjadi lebih siap untuk digunakan. Pada proses pengumpulan data, terdapat sebanyak 404 data lowongan kerja yang terkumpul dari berbagai situs pencari kerja dalam kurun waktu 2 minggu. Proses pengumpulan data ini dilakukan secara daring dengan menggunakan metode dokumentasi. Namun, dari hasil pengumpulan data tersebut terdapat masalah seperti adanya data yang terduplikasi akibat sumber data yang berasal dari berbagai situs pencari kerja dan informasi data lowongan kerja yang tidak lengkap karena tidak mencantumkan persyaratan *skill* apa yang dibutuhkan sehingga terdapat sebanyak 4 data yang tidak bisa digunakan. Oleh karena itu, data yang akan digunakan dalam penelitian ini adalah sebanyak 400 data. Berikut ini merupakan tabel *data cleaning*.

Tabel 4. 3 *Cleaning Data*

Nama Perusahaan	Sektor	Posisi	Kualifikasi Hard Skill	Kualifikasi Soft Skill
PT. Maxindo Karya Anugerah Tbk	Manufaktur	Operasional Produksi	Memahami ISO, <i>Food Safety Management</i> , GMP, HACCP, <i>Quality Control & Management System</i> .	Memiliki kemampuan komunikasi yang baik, dapat bekerja sama dalam tim, berorientasi pada target, memiliki kepribadian. & <i>attitude</i> yang baik.
Chocolate Monggo	Manufaktur	Operasional Produksi	Menguasai Bahasa Inggris, memahami <i>Food Safety Management</i> , HACCP, 5R, K3, perencanaan produksi, alur produksi, manajemen SDM, sertifikat jaminan halal.	Memiliki kemampuan analisis, kemampuan komunikasi yang baik, kemampuan <i>decision making</i> .

Nama Perusahaan	Sektor	Posisi	Kualifikasi Hard Skill	Kualifikasi Soft Skill
PT. Semanggi Tiga	Jasa	Operasional Produksi	Memiliki keahlian <i>Microsoft Office</i> , memahami manajemen operasional produksi, dan <i>customer relation management</i> .	Memiliki kemampuan analisis, kemampuan komunikasi yang baik.
PT. Rajafa Group	Dagang	Operasional Produksi	Memahami <i>supply chain management, inventory control & management</i> , memiliki keahlian <i>management</i> , memahami <i>management SDM</i> , strategi <i>management penjualan</i> serta pemasaran, dan <i>customer relation management</i> .	
Chocolate Monggo	Manufaktur	Operasional Produksi	Menguasai Bahasa Inggris, memahami <i>Food Safety Management, HACCP, 5R, K3</i> , perencanaan produksi, alur produksi, manajemen SDM, sertifikat jaminan halal.	Memiliki kemampuan analisis, kemampuan komunikasi yang baik, kemampuan <i>decision making</i> .
PT. Tidung Jaya Mandiri Indonesia	Manufaktur	Operasional Produksi	Memahami tentang <i>warehouse management</i> ,	Memiliki kemampuan komunikasi yang baik, kemampuan

Nama Perusahaan	Sektor	Posisi	Kualifikasi Hard Skill	Kualifikasi Soft Skill
			memiliki keahlian <i>data analytics</i> .	kepemimpinan yang baik, dan mampu bekerja dalam tim.
PT. Kalibaru	Manufaktur	PPIC Staff	-	-

Pada tabel di atas, dapat dilihat terdapat duplikasi data, yaitu pada data lowongan *Chocolate Monggo*, hal ini terjadi akibat sumber data berasal dari berbagai situs yang berbeda. Kemudian pada PT. Kalibaru tidak mencantumkan kualifikasi yang lengkap sehingga terjadi *missing data* dan data tidak bisa digunakan.

4.3 Data Transformation

Setelah melalui proses *data cleaning*, maka selanjutnya data akan melalui proses transformasi dimana pada proses ini format data akan diubah sehingga dapat lebih mudah digunakan untuk pengolahan data. Pada penelitian ini transformasi pada variabel data posisi serta kualifikasi *hard skill* dan *soft skill* untuk membuat data menjadi lebih seragam dan memudahkan dalam pengolahan data. Pada variabel posisi, akan dikelompokkan menjadi 4 posisi yaitu posisi dalam bidang operasional produksi, *marketing*, IT, dan *managerial*. Aturan transformasi pada variabel posisi dapat dilihat pada lampiran 1. Kemudian untuk variabel kualifikasi *hard skill* dan *soft skill*, akan dikelompokkan menjadi 6 kualifikasi yaitu keahlian *software*, keahlian bahasa asing, keahlian bidang teknik industri, kemampuan komunikasi, kemampuan *leadership*, dan memiliki etos kerja. Aturan transformasi pada variabel posisi dapat dilihat pada lampiran 2. Setelah itu data *string* akan akan ditransformasikan menjadi data *numeric* sehingga dapat memudahkan dalam pengolahan data. Proses *data transformation* ini dilakukan secara manual menggunakan *software excel*. Berikut ini merupakan aturan transformasi pada penelitian ini.

Tabel 4. 4 Aturan Transformasi Data

Aturan Transformasi		
Sektor	Manufaktur	1
	Jasa	2
	Dagang	3
	Konstruksi	4
Posisi	Operasional Produksi	1
	<i>Marketing</i>	2
	<i>IT</i>	3

Aturan Transformasi				
Kualifikasi	Hard Skill	Keahlian Software	<i>Managerial</i>	4
			Ya	1
		Tidak	0	
		Keahlian Bahasa Asing	Ya	1
			Tidak	0
		Keahlian Bidang Teknik Industri	Ya	1
	Tidak		0	
	Soft Skill	Kemampuan Komunikasi	Ya	1
			Tidak	0
		Kemampuan Leadership	Ya	1
			Tidak	0
		Memiliki Etos kerja	Ya	1
Tidak			0	

Berikut ini merupakan tabel data yang telah ditransformasikan.

Tabel 4. 5 Tabel Transformasi Data

No.	Sektor	Posisi	Keahlian Software	Keahlian Bahasa Asing	Keahlian Bidang Teknik Industri	Kemampuan Komunikasi	Kemampuan Leadership	Memiliki Etos Kerja
1	1	1	1	0	1	0	1	1
2	1	1	0	0	1	1	1	1
3	1	1	1	0	1	0	1	0
4	1	1	0	0	1	1	1	0
5	1	1	0	0	1	1	1	0
6	1	1	0	0	1	0	0	0
7	2	1	1	0	0	1	0	0
8	2	2	0	0	1	1	1	0
9	1	1	1	0	1	1	0	0
10	1	1	1	0	1	1	1	0
...								
391	1	4	1	1	1	1	1	1
392	1	2	1	1	1	1	1	1
393	1	4	1	0	0	1	1	1
394	1	1	1	1	1	1	1	0
395	1	1	1	0	0	1	1	1
396	1	1	1	1	1	1	1	0
397	1	1	1	0	0	1	1	1
398	1	2	0	1	1	1	1	1

No.	Sektor	Posisi	Keahlian Software	Keahlian Bahasa Asing	Keahlian Bidang Teknik Industri	Kemampuan Komunikasi	Kemampuan Leadership	Memiliki Etos Kerja
399	1	2	1	1	1	1	0	1
400	1	1	1	0	1	1	1	1

Berdasarkan tabel di atas, dapat dilihat bahwa transformasi dilakukan pada setiap variabel. Pada variabel sektor, sektor manufaktur akan ditransformasikan menjadi 1, sektor jasa akan ditransformasikan menjadi 2, sektor dagang akan ditransformasikan menjadi 3, dan sektor konstruksi akan ditransformasikan menjadi 4. Pada variabel posisi, posisi di bidang operasional produksi akan ditransformasikan menjadi 1, posisi di bagian *marketing* akan ditransformasikan menjadi 2, posisi di bagian IT akan ditransformasikan menjadi 3, dan posisi di bagian *managerial* akan ditransformasikan menjadi 4. Sedangkan untuk variabel kualifikasi *hard skill* dan *soft skill*, masing-masing kualifikasi akan dibagi menjadi ya dan tidak atau 1 dan 0.

4.4 Pengolahan Data

4.4.1 Penentuan Nilai optimal *Cluster*

Pada tahapan ini, nilai optimal *cluster* akan ditentukan dengan menggunakan *Elbow Method*. Nilai *cluster* optimal dapat dilihat pada grafik *Elbow Method* dimana grafik ini terdiri dari beberapa titik yang merupakan nilai *Sum of Square Error* (SSE) yang akan membentuk lengkungan seperti siku. Dalam mencari nilai *cluster* optimal, dapat dilakukan dengan membandingkan nilai *Sum of Square Error* (SSE) yang ada dalam grafik. Berikut ini merupakan langkah-langkah melakukan perhitungan dengan *Elbow Method*.

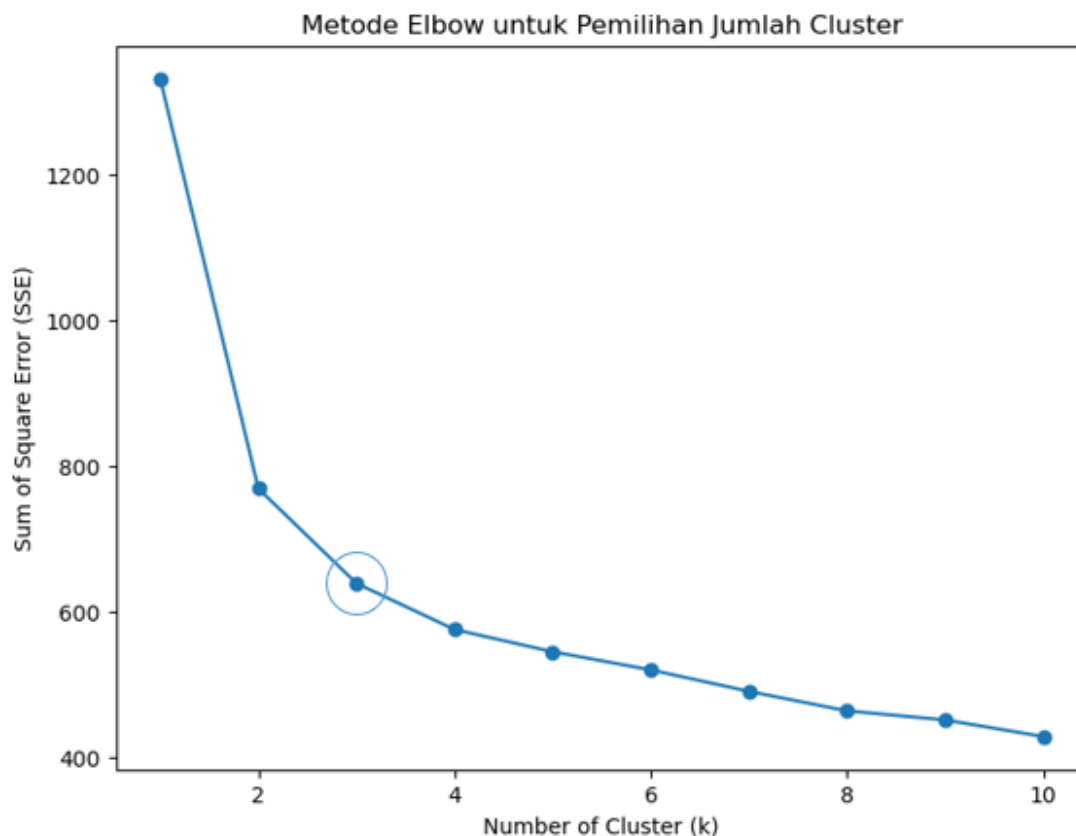
1. Inisialisasi awal nilai k
2. Naikkan nilai k
3. Hitung hasil *Sum of Square Error* (SSE) dari setiap nilai k
4. Lihat hasil *Sum of Square Error* (SSE) dari nilai k yang turun drastis
5. Tetapkan nilai k

Berikut ini merupakan hasil perhitungan *Elbow Method* dengan menggunakan *Python*.

Tabel 4. 6 Perhitungan SSE

<i>Cluster</i>	SSE	Selisih SSE
2	769,305	0
3	639,340	129,965
4	576,001	63,339
5	545,528	30,473
6	520,656	24,872
7	491,234	29,422
8	464,443	26,791
9	451,982	12,461
10	429,224	22,758

Pada tabel di atas, dapat dilihat perhitungan nilai SSE di setiap cluster dan selisihnya. Pada *cluster* 3, nilai SSE mengalami penurunan yang paling besar, yaitu sebesar 129,965 sedangkan setelah *cluster* 3, penurunan SSE pada *cluster-cluster* tersebut mulai melambat dan mulai stabil sehingga ditetapkan bahwa nilai *cluster* optimalnya adalah 3. Nilai optimal *cluster* juga dapat dilihat langsung pada grafik *Elbow Method*, titik awal akan mengalami penurunan yang tajam, kemudian selanjutnya akan menjadi lebih landai atau tidak signifikan. Titik sebelum grafik melandai itulah yang akan menjadi nilai *cluster* optimal. Berikut ini merupakan grafik *Elbow Method* yang telah dilakukan dengan *Python*.



Gambar 4. 1 Grafik *Elbow Method*

Berdasarkan gambar grafik di atas, dapat dilihat bahwa pada *cluster* 1 sampai 3 mengalami penurunan yang tajam, sedangkan pada *cluster* 4 dan seterusnya penurunan menjadi lebih landai dan tidak terlalu signifikan sehingga dapat disimpulkan bahwa nilai *cluster* optimalnya adalah 3. Hasil penentuan nilai *cluster* optimal menggunakan *Elbow Method* ini akan menjadi input pada proses *clustering* dengan menggunakan *software* SPSS.

4.4.2 Analisis *K-Means Clustering*

Setelah mendapatkan nilai *cluster* optimal, maka selanjutnya akan dilakukan pengolahan data menggunakan SPSS sebagai berikut.

4.4.2.1 Analisis Deskriptif

Berikut ini merupakan *output* deskriptif statistik yang didapatkan dari pengolahan data menggunakan SPSS.

Tabel 4. 7 Tabel Deskriptif Statistik

<i>Descriptive Statistics</i>					
	<i>N</i>	<i>Minimum</i>	<i>Maximum</i>	<i>Mean</i>	<i>Std. Deviation</i>
Sektor	400	1	4	1,295	0,65118
Posisi	400	1	4	1,9175	1,24482
Keahlian_ <i>Software</i>	400	0	1	0,57	0,4957
Keahlian_ <i>Bahasa_Asing</i>	400	0	1	0,24	0,42762
Keahlian_ <i>Bidang_Teknik_Industri</i>	400	0	1	0,665	0,47258
Kemampuan_ <i>Komunikasi</i>	400	0	1	0,685	0,4651
Kemampuan_ <i>Leadership</i>	400	0	1	0,555	0,49759
Memiliki_ <i>Etos_Kerja</i>	400	0	1	0,4475	0,49786
<i>Valid N (listwise)</i>	400				

Berdasarkan tabel deskriptif statistik di atas, terdapat delapan variabel yang meliputi sektor, posisi, keahlian *software*, keahlian bahasa asing, keahlian bidang teknik industri, kemampuan komunikasi, kemampuan *leadership*, dan memiliki etos kerja. Pada tabel deskriptif statistik, variabel-variabel yang menjadi kriteria dalam melakukan *clustering* tersebut dianalisis terlebih dahulu berdasarkan jumlah data, nilai minimum, nilai maksimum, *mean*, dan standar deviasinya.

Nilai N pada tabel menunjukkan jumlah data yang telah melalui tahap *pre-processing* sehingga nilai N pada setiap variabel adalah 400. Untuk nilai minimum merupakan nilai terkecil pada setiap variabel. Pada variabel sektor nilai minimumnya adalah 1 yang mana angka 1 pada variabel sektor merupakan transformasi dari sektor “manufaktur”. Pada variabel posisi nilai minimumnya adalah 1 yang mana angka 1 pada variabel posisi merupakan transformasi dari posisi di bagian “operasional produksi” dan begitu seterusnya. Untuk nilai maksimum merupakan nilai terbesar pada setiap variabel. Pada variabel sektor nilai maksimumnya adalah 4 yang mana angka 4 pada variabel sektor merupakan transformasi dari sektor “konstruksi”. Pada variabel posisi nilai maksimumnya adalah 4 yang mana angka 4 pada variabel posisi merupakan transformasi dari posisi di bagian “*managerial*” dan begitu pula seterusnya. Kemudian untuk nilai *mean* merupakan nilai rata-rata pada setiap variabel dimana nilai mean pada variabel sektor adalah sebesar 1,295, nilai *mean* pada variabel posisi adalah sebesar 1,9175, dan begitu seterusnya. Untuk nilai standar deviasi digunakan untuk mengukur persebaran data dari nilai rata-rata dimana nilai standar deviasi dari variabel sektor adalah sebesar 0,65118 dan untuk variabel posisi sebesar 1,24482, dan begitu seterusnya.

4.4.2.2 K-Means Clustering

Berikut ini merupakan *output Initial Cluster Centers* dari hasil pengolahan data dengan menggunakan SPSS.

Tabel 4. 8 Tabel *Initial Cluster Centers*

	<i>Initial Cluster Centers</i>		
	<i>Cluster</i>		
	1	2	3
Zscore(Sektor)	-0,45302	1,08265	4,15400
Zscore(Posisi)	0,06627	1,67294	-0,73706
Zscore(Keahlian_ <i>Software</i>)	-1,14990	0,86747	-1,14990
Zscore(Keahlian_ <i>Bahasa_Asing</i>)	1,77729	-0,56125	-0,56125
Zscore(Keahlian_ <i>Bidang_Teknik_Industri</i>)	0,70887	-1,40716	0,70887
Zscore(Kemampuan_ <i>Komunikasi</i>)	0,67728	-1,47281	0,67728
Zscore(Kemampuan_ <i>Leadership</i>)	0,89431	-1,11538	0,89431
Zscore(Memiliki_ <i>Etos_Kerja</i>)	1,10975	-0,89885	-0,89885

Sebelum melakukan iterasi, perlu untuk mencari nilai *Initial Cluster Centers* yang merupakan titik awal yang digunakan pada proses *clustering* dan menunjukkan dimana posisi awalnya dari pusat *cluster* sebelum nantinya akan dibagi ke dalam kelompok-kelompok yang lebih terstruktur. Tabel di atas menunjukkan bagaimana gambaran awal distribusi data ke dalam kelompok dimana tabel menampilkan hasil pengelompokkan data di tahap awal setelah menetapkan 3 sebagai jumlah *cluster*. Berdasarkan tabel, dapat dilihat hasil perhitungan *Initial Cluster Centers* pada *Zscore* sektor *cluster* 1 menunjukkan nilai -0,45302 yang mana jika hasilnya negatif, maka artinya *cluster* ada di bawah rata-rata total. Pada *Zscore* sektor *cluster* 2 menunjukkan nilai 1,08265 yang mana hasilnya positif, maka artinya *cluster* ada di atas rata-rata total. Sedangkan pada *Zscore* sektor *cluster* 3 menunjukkan nilai 4,15400 yang mana menunjukkan jika hasilnya juga positif, maka artinya *cluster* ada di atas rata-rata total dan begitu pula dengan variabel lainnya.

Tabel 4. 9 Tabel Iterasi

<i>Iteration History^a</i>			
<i>Change in Cluster Centers</i>			
<i>Iteration</i>	1	2	3
1	2,112	2,394	1,964

<i>Iteration History^a</i>			
<i>Iteration</i>	<i>Change in Cluster Centers</i>		
	1	2	3
2	0,110	0,178	0,430
3	0,115	0,151	0,175
4	0,056	0,085	0,160
5	0,086	0,109	0,097
6	0,100	0,129	0,130
7	0,034	0,044	0,000
8	0,000	0,000	0,000

a. Convergence achieved due to no or small change in cluster centers. The maximum absolute coordinate change for any center is ,000. The current iteration is 8. The minimum distance between initial centers is 5,601.

Berdasarkan tabel di atas, untuk melakukan pengelompokan, proses iterasi yang dilakukan adalah sebanyak 8 kali untuk menemukan *cluster* yang tepat yang mana pada iterasi ke-8 pusat *cluster* telah menunjukkan angka 0,000 sehingga proses *clustering* telah mencapai titik konvergensi yang artinya perubahan pusat *cluster* sudah sangat kecil atau tidak ada sama sekali dan setiap pusat *cluster* telah berada di posisi yang stabil. Dari tabel di atas, dijelaskan bahwa *the minimum distance between initial centers*nya adalah 5,601 dimana diantara semua pasangan pusat *cluster* awal yang dipilih secara acak, terdapat pasangan pusat *cluster* yang paling dekat dan kedua pusat *cluster* ini berjarak 5,601 satu sama lain.

Tabel 4. 10 Tabel *Final Cluster Centers*

	<i>Final Cluster Centers</i>		
	<i>Cluster</i>		
	1	2	3
Zscore(Sektor)	-0,35168	-0,28350	2,30492
Zscore(Posisi)	-0,16208	0,20190	0,01709
Zscore(Keahlian_ <i>Software</i>)	-0,05417	0,25178	-0,57351
Zscore(Keahlian_ <i>Bahasa_Asing</i>)	0,09164	-0,06013	-0,17945

	<i>Final Cluster Centers</i>		
	<i>Cluster</i>		
	1	2	3
Zscore(Keahlian_Bidang_Teknik_Industri)	0,21477	-0,21174	-0,19800
Zscore(Kemampuan_Komunikasi)	0,37168	-0,52342	0,15073
Zscore(Kemampuan_Leadership)	0,61887	-0,78913	-0,00800
Zscore(Memiliki_Etos_Kerja)	0,46741	-0,49452	-0,32496

Berdasarkan tabel di atas, dapat dilihat hasil analisis *final cluster centers* yang menunjukkan masing-masing kriteria dan *cluster* yang terbentuk. Berbeda dengan *initial cluster centers* yang merupakan titik awal yang digunakan pada proses *clustering* dan menunjukkan dimana posisi awalnya dari pusat *cluster* sebelum dibagi ke dalam kelompok-kelompok yang lebih terstruktur, *final cluster centers* merupakan tahap terakhir dari proses *clustering* dimana pada *final cluster centers*, pusat *cluster* sudah berada di posisi terakhir setelah proses iterasi selesai dilakukan.

Sama seperti pada *initial cluster centers*, pada *final cluster centers* apabila hasil perhitungan *cluster* negatif, maka *cluster* tersebut dominan di bawah rata-rata total. Sedangkan apabila hasil perhitungan *cluster* positif, maka *cluster* tersebut dominan di atas rata-rata total. Berikut ini merupakan penjelasan tabel *final cluster centers* yang didapatkan dari pengolahan data menggunakan SPSS di atas:

1. *Cluster 1*

Pada *cluster 1*, hasil perhitungan menunjukkan bahwa pada variabel sektor dominan di bawah rata-rata total, variabel posisi dominan di bawah rata-rata total, variabel keahlian *software* berada di bawah rata-rata total, variabel keahlian bahasa asing dominan di atas rata-rata total, variabel keahlian bidang teknik industri dominan di atas rata-rata total, variabel kemampuan komunikasi dominan di atas rata-rata total, variabel kemampuan *leadership* dominan di atas rata-rata total, dan variabel memiliki etos kerja dominan di atas rata-rata total.

2. *Cluster 2*

Pada *cluster 2*, hasil perhitungan menunjukkan bahwa pada variabel sektor dominan di bawah rata-rata total, variabel posisi dominan di atas rata-rata total, variabel keahlian *software* berada di atas rata-rata total, variabel keahlian bahasa asing dominan di bawah

rata-rata total, variabel keahlian bidang teknik industri dominan di bawah rata-rata total, variabel kemampuan komunikasi dominan di bawah rata-rata total, variabel kemampuan *leadership* dominan di bawah rata-rata total, dan variabel memiliki etos kerja dominan di bawah rata-rata total.

3. *Cluster 3*

Pada *cluster 3*, hasil perhitungan menunjukkan bahwa pada variabel sektor dominan di atas rata-rata total, variabel posisi dominan di atas rata-rata total, variabel keahlian *software* berada di bawah rata-rata total, variabel keahlian bahasa asing dominan di bawah rata-rata total, variabel keahlian bidang teknik industri dominan di bawah rata-rata total, variabel kemampuan komunikasi dominan di atas rata-rata total, variabel kemampuan *leadership* dominan di bawah rata-rata total, dan variabel memiliki etos kerja dominan di bawah rata-rata total.

Berikut ini merupakan hasil pengelompokan 400 data ke dalam 3 *cluster*:

Tabel 4. 11 *Number of Cases in each Cluster*

<i>Number of Cases in each Cluster</i>	
<i>Cluster</i>	
1	197,000
2	154,000
3	49,000
<i>Valid</i>	400,000
<i>Missing</i>	0,000

Berdasarkan tabel hasil pengolahan data menggunakan SPSS di atas, dapat dilihat jumlah anggota pada setiap *cluster*. Total 400 data yang digunakan pada proses *clustering* terbagi ke dalam 3 *cluster*, yaitu *cluster 1*, *cluster 2*, dan *cluster 3* dimana jumlah anggota pada *cluster 1* yaitu sebanyak 197 anggota, pada *cluster 2* sebanyak 154 anggota, dan pada *cluster 3* sebanyak 49 anggota.

4.4.2.3 Profilisasi *Cluster*

Tahap profilisasi *cluster* ini dilakukan dengan cara menghitung nilai persentase rata-rata setiap nilai variabel pada masing-masing *cluster*. Perhitungan dilakukan secara manual dengan menggunakan bantuan *software excel*. Berikut ini merupakan rumus yang digunakan untuk perhitungan profilisasi *cluster*.

Persentase rata – rata

$$= \frac{\text{Jumlah total nilai variabel dalam cluster}}{\text{Jumlah data dalam cluster}} \times 100\% \quad (4.1)$$

Berikut ini contoh merupakan langkah perhitungannya.

$$\text{Persentase rata – rata} = \frac{\text{Jumlah total nilai variabel dalam cluster}}{\text{Jumlah data dalam cluster}} \times 100\%$$

$$\text{Persentase rata – rata} = \frac{184}{197} \times 100\%$$

$$\text{Persentase rata – rata} = 0,93 \times 100\%$$

$$\text{Persentase rata – rata} = 93\%$$

Contoh perhitungan di atas merupakan contoh perhitungan untuk menghitung persentase rata-rata nilai variabel sektor manufaktur pada *cluster* 1 dimana nilai 184 merupakan jumlah data lowongan pada *cluster* 1 yang sektornya adalah sektor manufaktur. Sedangkan nilai 197 merupakan jumlah seluruh data yang terdapat pada *cluster* 1. Sehingga didapatkan persentase rata-rata nilai variabel sektor manufaktur pada *cluster* 1 adalah 1,93 atau 93%. Berdasarkan rumus perhitungan tersebut, berikut ini merupakan persentase rata-rata nilai pada variabel sektor dan posisi pada setiap *cluster* yang terbentuk.

Gambat 4. 11 Persentase Sektor dan Posisi di Setiap *Cluster*

Cluster 1			
Sektor		Posisi	
Manufaktur	93%	Operasional Produksi	66%
Jasa	7%	Marketing	13%
		IT	3%
		Managerial	18%
Cluster 2			
Sektor		Posisi	
Manufaktur	89%	Operasional Produksi	54%
Jasa	11%	Marketing	8%
		IT	5%
		Managerial	33%
Cluster 3			
Sektor		Posisi	
Jasa	29%	Operasional Produksi	39%
Dagang	63%	Marketing	45%
Konstruksi	8%	Managerial	16%

Berdasarkan tabel di atas, dapat dilihat pada *cluster* 1 terdapat dua sektor perusahaan yaitu sektor manufaktur dan sektor jasa yang masing-masing sebanyak 93% dan 7%. Sedangkan untuk posisi lowongan pada *cluster* 1 terdapat empat bidang, bidang operasional produksi sebanyak 66%, bidang *marketing* sebanyak 13%, bidang IT sebanyak 3%, dan bidang *managerial* sebanyak 18%. Pada *cluster* 2 juga terdapat dua sektor perusahaan yaitu sektor manufaktur dan sektor jasa yang masing-masing sebanyak 89% dan 11%. Sedangkan untuk posisi lowongan yang mengisi *cluster* 2 terdapat empat, yaitu operasional produksi sebanyak 54%, *marketing* sebanyak 8%, IT sebanyak 5%, dan *managerial* sebanyak 33%. Selanjutnya pada *cluster* 3 berisi tiga sektor perusahaan yaitu sektor jasa sebanyak 29%, sektor dagang sebanyak 63%, dan sektor konstruksi sebanyak 8%. Sedangkan untuk posisi yang mengisi *cluster* 3 terdapat tiga bidang, yaitu bidang operasional produksi sebanyak 39%, bidang *marketing* sebanyak 45%, dan bidang *managerial* sebanyak 16%.

Berikutnya, untuk menghitung persentase rata-rata pada setiap variabel kualifikasi akan diterapkan rumus yang sama. Berikut ini merupakan contoh langkah perhitungannya.

$$\text{Persentase rata - rata} = \frac{\text{Jumlah total nilai variabel dalam cluster}}{\text{Jumlah data dalam cluster}} \times 100\%$$

$$\text{Persentase rata - rata} = \frac{107}{197} \times 100\%$$

$$\text{Persentase rata - rata} = 0,54 \times 100\%$$

$$\text{Persentase rata - rata} = 54\%$$

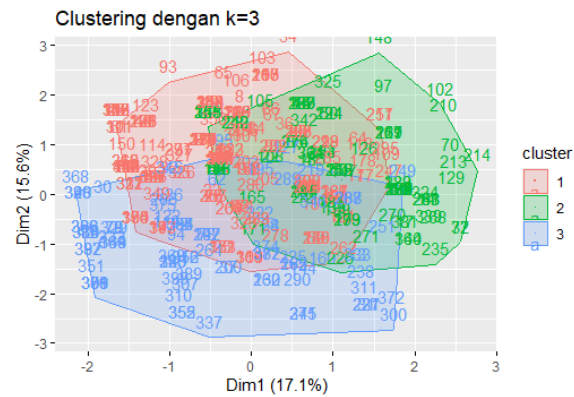
Contoh perhitungan di atas merupakan contoh perhitungan untuk menghitung persentase rata-rata nilai variabel keahlian *software* pada *cluster* 1 dimana nilai 107 merupakan jumlah data lowongan pada *cluster* 1 yang memberikan kualifikasi keahlian *software* sebagai persyaratannya. Sedangkan nilai 197 merupakan jumlah seluruh data yang terdapat pada *cluster* 1. Sehingga didapatkan persentase rata-rata nilai variabel keahlian *software* pada *cluster* 1 adalah 0,54 atau 54%. Berdasarkan rumus perhitungan tersebut, berikut ini merupakan persentase rata-rata nilai setiap variabel kualifikasi pada ketiga *cluster* yang terbentuk.

Gambat 4. 12 Persentase Kualifikasi di Setiap *Cluster*

Kualifikasi	Cluster		
	Cluster 1	Cluster 2	Cluster 3
Keahlian <i>Software</i>	54%	69%	29%

Kualifikasi	Cluster		
	Cluster 1	Cluster 2	Cluster 3
Keahlian Bahasa Asing	28%	21%	16%
Keahlian Bidang Teknik Industri	77%	56%	57%
Kemampuan Komunikasi	86%	44%	76%
Kemampuan <i>Leadership</i>	86%	16%	55%
Memiliki Etos Kerja	68%	20%	29%

Berdasarkan tabel di atas, dapat dilihat data lowongan kerja pada *cluster 1* membutuhkan kualifikasi keahlian *software* sebanyak 54%, keahlian bahasa asing sebanyak 28%, keahlian bidang teknik industri sebanyak 77%, kemampuan komunikasi sebanyak 86%, kemampuan *leadership* sebanyak 86%, dan memiliki etos kerja sebanyak 68%. Pada *cluster 2* membutuhkan kualifikasi keahlian *software* sebanyak 69%, keahlian bahasa asing sebanyak 21%, keahlian bidang teknik industri sebanyak 56%, kemampuan komunikasi sebanyak 44%, kemampuan *leadership* sebanyak 16%, dan memiliki etos kerja sebanyak 20%. *cluster 3* membutuhkan kualifikasi keahlian *software* sebanyak 29%, keahlian bahasa asing sebanyak 16%, keahlian bidang teknik industri sebanyak 57%, kemampuan komunikasi sebanyak 76%, kemampuan *leadership* sebanyak 55%, dan memiliki etos kerja sebanyak 29%.



Gambar 4. 2 Visualisasi

Berdasarkan visualisasi *cluster* di atas, dapat dilihat *cluster* 1 ditandai dengan area dan titik yang berwarna merah, *cluster* 2 ditandai dengan area dan titik yang berwarna hijau, dan *cluster* 3 ditandai dengan area dan titik berwarna biru. Dari gambar 4.2 terlihat bahwa terjadi tumpang tindih (*overlapping*) antar *cluster* sehingga membuat pembagian kelompok menjadi tidak merata.

BAB V

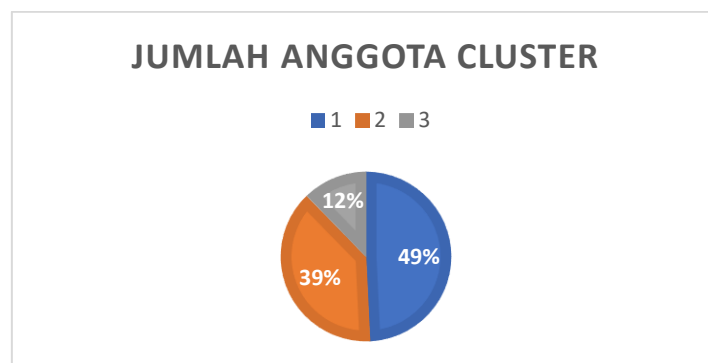
PEMBAHASAN

5.1 Nilai Optimal Cluster

Berdasarkan hasil penentuan nilai optimal *cluster* dengan menggunakan *Elbow Method*, grafik *elbow* pada gambar 4.1 menunjukkan 3 dimana titik awal akan mengalami penurunan yang tajam, kemudian selanjutnya akan menjadi lebih landai atau tidak signifikan. Titik sebelum grafik melandai itulah yang akan menjadi nilai *cluster* optimal. Selain dari grafik, penentuan nilai *cluster* optimal juga dapat dilihat dari selisih nilai *Sum of Square Error* (SSE) pada setiap *clusternya* untuk memastikan kembali jumlah *cluster* optimalnya dimana jika dilihat dari tabel nilai SSE, pada *cluster* 3, nilai SSE mengalami penurunan yang paling besar, yaitu sebesar 129,965 sedangkan setelah *cluster* 3, penurunan SSE pada *cluster-cluster* tersebut mulai melambat sehingga ditetapkan bahwa nilai *cluster* optimalnya adalah 3. Nilai optimal *cluster* 3 dipilih karena selisih nilai SSE setelah *cluster* 3 mengalami penurunan yang stabil yang artinya manfaat dari penambahan *cluster* mulai menurun drastis dimana penambahan *cluster* dapat menurunkan nilai SSE. Namun, apabila jumlah *cluster* terlalu banyak dapat menyebabkan hasilnya terlalu kompleks.

5.2 Proporsi Cluster

Proporsi *cluster* dapat digunakan untuk melihat proporsi anggota pada setiap *cluster*. Berikut ini merupakan began yang menunjukkan proporsi anggota pada *cluster* 1, 2, dan 3 yang dinyatakan dalam persen.



Gambar 5. 1 Jumlah Anggota Cluster

Berdasarkan bagan di atas, dapat dilihat persentase anggota pada *cluster* 1 yaitu sebesar 49% yang mana jumlah anggota pada *cluster* 1 ini adalah sebanyak 197 data dari 400 data yang diolah. Persentase anggota pada *cluster* 2 yaitu sebesar 39% yang mana jumlah anggota pada *cluster* 2 ini adalah sebanyak 154 data dari 400 data yang diolah. Dan yang terakhir persentase anggota pada *cluster* 3 yaitu sebesar 12% yang mana jumlah anggota pada *cluster* 3 ini adalah sebanyak 49 data dari 400 data yang diolah. *Cluster* 1 menjadi *cluster* yang memiliki anggota terbanyak, disusul dengan *cluster* 2, dan *cluster* 3 memiliki anggota yang paling sedikit. Selain itu, berdasarkan visualisasi *cluster* pada gambar 4.2, diketahui bahwa terjadi *overlapping* atau tumpang tindih terutama pada bagian tengah grafik sehingga membuat pembagian kelompok menjadi tidak merata. *Overlapping* sendiri yaitu kondisi dimana terdapat data yang memiliki keanggotaan lebih dari satu yang mana hal tersebut bisa terjadi karena data yang kurang jelas, variabel yang kurang sesuai sehingga pengelompokan menjadi kurang optimal, ataupun penggunaan metode *clustering* yang kurang tepat (Adwitya, et al., 2015).

5.3 Profiling Cluster

Berdasarkan hasil pengolahan data pada bab sebelumnya, jumlah *cluster* optimal yang didapatkan dengan *elbow method* adalah 3 yang kemudian nilai ini akan menjadi *input* pada proses *clustering* menggunakan *software* SPSS. Kemudian dilakukan perhitungan manual untuk mencari rata-rata setiap variabel di setiap *cluster* dengan bantuan *software excel* sehingga dapat digunakan untuk membantu memahami karakteristik di setiap *clusternya*. Berikut ini merupakan penjelasan mengenai karakteristik pada setiap *cluster* yang terbentuk.

5.3.1 Cluster 1

Berdasarkan tabel karakteristik *cluster* yang sudah dijelaskan pada bab sebelumnya, diketahui bahwa pada *cluster* 1 terdiri dari 197 data lowongan dimana data lowongan di *cluster* 1 terdiri dari dua sektor, yaitu sektor manufaktur yang mana perusahaan dengan sektor ini merupakan perusahaan yang melakukan kegiatan ekonominya dengan melakukan kegiatan mengubah suatu barang dasar atau bahan baku menjadi barang jadi atau setengah jadi yang sifatnya lebih dekat ke pemakaian akhir dengan secara mekanis, kimia, ataupun dengan tangan dan sektor jasa yang mana perusahaan dengan sektor ini merupakan perusahaan yang melakukan kegiatan ekonominya dengan menyediakan kegiatan atau menjalankan suatu tugas dengan komersial dan perusahaan dengan sektor ini cenderung lebih banyak mencari sumber daya manusia. Untuk sektor manufaktur berjumlah sebanyak 93% atau 184 data dari total 197 data pada *cluster* 1 merupakan perusahaan dengan sektor manufaktur. Kemudian untuk sektor jasa berjumlah

sebanyak 7% atau 13 data dari total 197 pada cluster 1 merupakan perusahaan dengan sektor jasa. Pada variabel posisi di *cluster* 1 ini juga terdapat empat bagian, yaitu lowongan dengan posisi di bagian operasional produksi sebanyak 66% dimana posisi pada bagian operasional produksi ini mencakup posisi sebagai *PPIC staff*, *supervisor PPIC*, *manager PPIC*, *quality assurance staff*, *quality control staff*, *staff warehouse*, *PIC project interior*, *staff logistic executive*, dan *maintenance section head*. Lowongan dengan posisi di bagian *marketing* sebanyak 13% dimana posisi pada bagian *marketing* ini mencakup posisi sebagai *sales B2B*, *marketing & sales*, *sales machinery manager*, *marketing communication*, *IT sales*, *account manager*, dan *staff marketing logistic*. Lowongan dengan posisi di bagian IT sebanyak 3% dimana posisi pada bagian IT ini mencakup posisi sebagai *business analyst*, *business process analyst*, *CNC programmer*, dan *collection database analyst*. Lowongan dengan posisi di bagian *managerial* sebanyak 18% dimana posisi pada bagian *managerial* ini mencakup posisi sebagai *staff HRD*, *purchasing staff*, *junior auditor*, *MEP document controller*, ahli K3 kimia, dan *general affair*.

Selanjutnya, pada *cluster* ini, data lowongan kerja yang memberikan kualifikasi keahlian *software* adalah sebanyak 107 data dari total 197 data lowongan atau sebanyak 54% dimana keahlian *software* ini mencakup *software-software* seperti *Microsoft Office (Excel, Word, Powerpoint, dan Ms. Project)*, *Autocad*, *Solidworks*, dan lainnya. Kualifikasi keahlian bahasa asing yang mencakup Bahasa Inggris dan Bahasa Mandarin sebanyak 55 data dari total 197 data lowongan atau sebanyak 28%. Kualifikasi keahlian bidang teknik industri yang mencakup bidang ergonomi, sistem produksi, dan bidang manajemen dan *operational research* sebanyak 151 data dari total 197 data lowongan atau sebanyak 77%. Kualifikasi kemampuan komunikasi yang mencakup kemampuan negosiasi, bekerja sama dalam tim, dan memiliki kemampuan presentasi yang baik sebanyak 169 data dari total 197 data lowongan atau sebanyak 86%. Kualifikasi kemampuan *leadership* yang mencakup kemampuan *decision making*, tegas dan berani mengatur bawahan, kemampuan *problem solving*, kemampuan analisis dan memiliki kreativitas yang tinggi sebanyak 170 data dari total 197 data lowongan atau sebanyak 86%. Kemudian kualifikasi memiliki etos kerja yang mencakup memiliki semangat yang tinggi, berorientasi pada target, mampu bekerja di bawah tekanan, *multitasking* dan bisa bekerja secara efisien, serta memiliki manajemen waktu yang baik sebanyak 134 data dari total 197 data lowongan atau sebanyak 68%.

Berdasarkan persentase setiap variabel di atas, pada *cluster 1* ini data lowongan kerja didominasi oleh perusahaan dengan sektor manufaktur dan posisi di bagian operasional produksi dan untuk mengisi posisi pada bagian operasional produksi di perusahaan sektor manufaktur tersebut, banyak perusahaan yang memberi kualifikasi kemampuan komunikasi, kemampuan *leadership*, dan keahlian bidang teknik industri. Selain itu, kualifikasi keahlian *software* dan memiliki etos kerja juga cukup banyak diberikan oleh perusahaan.

5.3.2 Cluster 2

Berdasarkan tabel karakteristik *cluster* yang sudah dijelaskan pada bab sebelumnya, diketahui bahwa pada *cluster 2* terdiri dari 154 data lowongan dimana data lowongan di *cluster 2* terdiri dari dua sektor, yaitu sektor manufaktur yang mana perusahaan dengan sektor ini merupakan perusahaan yang melakukan kegiatan ekonominya dengan melakukan kegiatan mengubah suatu barang dasar atau bahan baku menjadi barang jadi atau setengah jadi yang sifatnya lebih dekat ke pemakaian akhir dengan secara mekanis, kimia, ataupun dengan tangan dan sektor jasa yang mana perusahaan dengan sektor ini merupakan perusahaan yang melakukan kegiatan ekonominya dengan menyediakan kegiatan atau menjalankan suatu tugas dengan komersial dan perusahaan dengan sektor ini cenderung lebih banyak mencari sumber daya manusia. Untuk sektor manufaktur berjumlah sebanyak 89% atau 137 data dari total 154 data pada *cluster 2* merupakan perusahaan dengan sektor manufaktur. Kemudian untuk sektor jasa berjumlah sebanyak 11% atau 17 data dari total 154 pada *cluster 2* merupakan perusahaan dengan sektor jasa. Pada variabel posisi di *cluster 2* ini juga terdapat empat bagian, yaitu lowongan dengan posisi di bagian operasional produksi sebanyak 54% dimana posisi pada bagian operasional produksi ini mencakup posisi sebagai *PPIC staff*, *supervisor PPIC*, *manager PPIC*, *quality assurance staff*, *quality control staff*, *staff warehouse*, *PIC project interior*, *staff logistic executive*, dan *maintenance section head*. Lowongan dengan posisi di bagian *marketing* sebanyak 8% dimana posisi pada bagian *marketing* ini mencakup posisi sebagai *sales B2B*, *marketing & sales*, *sales machinery manager*, *marketing communication*, *IT sales*, *account manager*, dan *staff marketing logistic*. Lowongan dengan posisi di bagian IT sebanyak 5% dimana posisi pada bagian IT ini mencakup posisi sebagai *business analyst*, *business process analyst*, *CNC programmer*, dan *collection database analyst*. Lowongan dengan posisi di bagian *managerial* sebanyak 33% dimana posisi pada bagian *managerial* ini mencakup posisi sebagai *staff HRD*, *purchasing staff*, *junior auditor*, *MEP document controller*, ahli K3 kimia, dan *general affair*.

Selanjutnya, pada *cluster* ini, data lowongan kerja yang memberikan kualifikasi keahlian *software* adalah sebanyak 107 data dari total 154 data lowongan atau sebanyak 69% dimana keahlian *software* ini mencakup *software-software* seperti *Microsoft Office (Excel, Word, Powerpoint, dan Ms. Project)*, *Autocad, Solidworks*, dan lainnya. Kualifikasi keahlian bahasa asing yang mencakup Bahasa Inggris dan Bahasa Mandarin sebanyak 33 data dari total 154 data lowongan atau sebanyak 21%. Kualifikasi keahlian bidang teknik industri yang mencakup bidang ergonomi, sistem produksi, dan bidang manajemen dan *operational research* sebanyak 87 data dari total 154 data lowongan atau sebanyak 56%. Kualifikasi kemampuan komunikasi yang mencakup kemampuan negoisasi, bekerja sama dalam tim, dan memiliki kemampuan presentasi yang baik sebanyak 68 data dari total 154 data lowongan atau sebanyak 44%. Kualifikasi kemampuan *leadership* yang mencakup kemampuan *decision making*, tegas dan berani mengatur bawahan, kemampuan *problem solving*, kemampuan analisis dan memiliki kreativitas yang tinggi sebanyak 25 data dari total 154 data lowongan atau sebanyak 16%. Kemudian kualifikasi memiliki etos kerja yang mencakup memiliki semangat yang tinggi, berorientasi pada target, mampu bekerja di bawah tekanan, *multitasking* dan bisa bekerja secara efisien, serta memiliki manajemen waktu yang baik sebanyak 31 data dari total 154 data lowongan atau sebanyak 20%.

Berdasarkan persentase setiap variabel di atas, pada *cluster 2* ini data lowongan kerja didominasi oleh perusahaan dengan sektor manufaktur serta posisi di bagian operasional produksi dan *managerial*. Untuk mengisi posisi pada bagian operasional produksi dan *managerial* di perusahaan sektor manufaktur tersebut, banyak perusahaan yang memberi kualifikasi keahlian *software* dan keahlian bidang teknik industri serta cukup banyak perusahaan yang memberikan kualifikasi kemampuan komunikasi.

5.3.3 *Cluster 3*

Berdasarkan tabel karakteristik *cluster* yang sudah dijelaskan pada bab sebelumnya, diketahui bahwa pada *cluster 3* terdiri dari 49 data lowongan dimana data lowongan di *cluster 3* terdiri dari tiga sektor, yaitu sektor jasa yang mana perusahaan dengan sektor ini merupakan perusahaan yang melakukan kegiatan ekonominya dengan menyediakan kegiatan atau menjalankan suatu tugas dengan komersial dan perusahaan dengan sektor ini cenderung lebih banyak mencari sumber daya manusia, sektor dagang yang mana perusahaan dengan sektor ini merupakan perusahaan yang melakukan kegiatan utamanya adalah membeli dan menjual barang di atas harga pokok sehingga mendapatkan keuntungan, dan sektor konstruksi yang berfokus pada

perancangan, pembangunan, renovasi, pemeliharaan infrastruktur dimana sektor ini memainkan peran penting dalam pembangunan ekonomi suatu negara. Untuk sektor jasa berjumlah sebanyak 29% atau 14 data dari total 49 data pada *cluster 3* merupakan perusahaan dengan sektor jasa. Untuk sektor dagang berjumlah sebanyak 63% atau 31 data dari total 49 data pada *cluster 3* merupakan perusahaan dengan sektor dagang. Kemudian untuk sektor konstruksi berjumlah sebanyak 8% atau 4 data dari total 49 pada *cluster 3* merupakan perusahaan dengan sektor konstruksi. Pada variabel posisi di *cluster 3* ini juga terdapat tiga bagian, yaitu lowongan dengan posisi di bagian operasional produksi sebanyak 39% dimana posisi pada bagian operasional produksi ini mencakup posisi sebagai *PPIC staff, supervisor PPIC, manager PPIC, quality assurance staff, quality control staff, staff warehouse, PIC project interior, staff logistic executive*, dan *maintenance section head*. Lowongan dengan posisi di bagian *marketing* sebanyak 45% dimana posisi pada bagian *marketing* ini mencakup posisi sebagai *sales B2B, marketing & sales, sales machinery manager, marketing communication, IT sales, account manager*, dan *staff marketing logistic*. Lowongan dengan posisi di bagian *managerial* sebanyak 16% dimana posisi pada bagian *managerial* ini mencakup posisi sebagai *staff HRD, purchasing staff, junior auditor, MEP document controller*, ahli K3 kimia, dan *general affair*.

Selanjutnya, pada *cluster* ini, data lowongan kerja yang memberikan kualifikasi keahlian *software* adalah sebanyak 14 data dari total 49 data lowongan atau sebanyak 29% dimana keahlian *software* ini mencakup *software-software* seperti *Microsoft Office (Excel, Word, Powerpoint, dan Ms. Project), Autocad, Solidworks*, dan lainnya. Kualifikasi keahlian bahasa asing yang mencakup Bahasa Inggris dan Bahasa Mandarin sebanyak 8 data dari total 49 data lowongan atau sebanyak 16%. Kualifikasi keahlian bidang teknik industri yang mencakup bidang ergonomi, sistem produksi, dan bidang manajemen dan *operational research* sebanyak 28 data dari total 197 data lowongan atau sebanyak 57%. Kualifikasi kemampuan komunikasi yang mencakup kemampuan negoisasi, bekerja sama dalam tim, dan memiliki kemampuan presentasi yang baik sebanyak 37 data dari total 49 data lowongan atau sebanyak 76%. Kualifikasi kemampuan *leadership* yang mencakup kemampuan *decision making*, tegas dan berani mengatur bawahan, kemampuan *problem solving*, kemampuan analisis dan memiliki kreativitas yang tinggi sebanyak 27 data dari total 49 data lowongan atau sebanyak 55%. Kemudian kualifikasi memiliki etos kerja yang mencakup memiliki semangat yang tinggi, berorientasi pada target, mampu bekerja di bawah tekanan, *multitasking* dan bisa bekerja secara

efisien, serta memiliki manajemen waktu yang baik sebanyak 14 data dari total 49 data lowongan atau sebanyak 29%.

Berdasarkan persentase setiap variabel di atas, pada *cluster* 3 ini data lowongan kerja didominasi oleh perusahaan dengan sektor dagang diikuti dengan perusahaan dengan sektor jasa serta posisi di bagian *marketing* dan operasional produksi. Untuk mengisi posisi pada perusahaan dengan kriteria tersebut, banyak perusahaan yang memberi kualifikasi kemampuan komunikasi, keahlian bidang teknik industri, dan kemampuan *leadership*.

5.4 Strategi Pengembangan Profil Lulusan

Berdasarkan hasil *profiling cluster* sebelumnya, diketahui bagaimana karakteristik di setiap *clusternya*. Selanjutnya, dari profil *cluster* tersebut akan dirancang strategi-strategi untuk setiap *cluster* yang terbentuk. Berikut ini merupakan strategi pengembangan profil lulusan yang disarankan untuk setiap *cluster*.

Tabel 5. 1 Strategi Pengembangan Profil Lulusan

Referensi	Strategi	Cluster 1	Cluster 2	Cluster 3
(Al Qadri, et al., 2024)	Pelaksanaan <i>Training Microsoft Office</i>	√	√	
(Chairani & Nurhazana, 2020)	Penerapan Mata Kuliah Pengembangan Kepribadian & Etika Profesi	√		
(Nurhayati & Yunitasari, 2020)	Pelaksanaan Sertifikasi Profesional	√	√	√
(Sholeha, et al., 2024)	Melakukan Kunjungan Industri	√	√	√
(Astuti, 2021)	Penerapan Mata kuliah Kepemimpinan dan Pengambilan Keputusan	√	√	√

Berdasarkan tabel di atas, terdapat lima strategi yang disarankan, strategi yang pertama yaitu pelaksanaan *training Microsoft Office* bagi mahasiswa teknik industri. Perancangan strategi ini berdasar akan kebutuhan perusahaan terhadap lulusan teknik industri yang memiliki keahlian

software, terutama *Microsoft Office*. Pada *cluster 2*, keahlian *software* merupakan kualifikasi yang paling banyak dibutuhkan yaitu sebesar 69%, selain itu pada *cluster 1* juga sebanyak 54% data lowongan memberikan kualifikasi keahlian *software*. Pada prodi Teknik Industri Universitas Islam Indonesia sendiri, selama beberapa tahun ke belakang, belum diadakan *training Microsoft Office* sehingga dirancanglah strategi untuk melakukan *training Microsoft Office* bagi mahasiswa Teknik Industri UII dimana dengan adanya pelatihan ini mahasiswa dapat lebih mahir dalam menggunakan *Microsoft Office (Word, Excel, Powerpoint)*. Selain itu mahasiswa juga akan mendapatkan sertifikat yang dapat dicantumkan pada *curriculum vitae*.

Strategi yang kedua adalah penerapan mata kuliah Pengembangan Kepribadian dan Etika Profesi. Perancangan strategi ini didasari oleh banyaknya perusahaan yang memberikan kualifikasi memiliki etos kerja pada *cluster 1* yaitu sebanyak 68%. Selain itu perancangan strategi ini dilakukan karena pada prodi Teknik Industri UII sendiri belum ada mata kuliah yang berfokus pada pengembangan etika mahasiswa. Penting bagi mahasiswa untuk memiliki etika kerja yang baik sebelum menghadapi dunia kerja nantinya. Maka dari itu strategi ini disarankan agar nantinya terbentuk individu yang produktif, berintegritas, dan profesional dalam dunia kerja.

Strategi yang ketiga adalah pelaksanaan sertifikasi profesional seperti sertifikasi *project management, lean six sigma*, ataupun manajemen rantai pasok, dan lainnya yang dapat meningkatkan keahlian bidang teknik industri mahasiswa. Perancangan strategi ini didasari oleh banyaknya perusahaan yang memberikan kualifikasi keahlian teknik industri pada *cluster 1, cluster 2, dan cluster 3* dimana pada *cluster 1* sebanyak 77%, *cluster 2* sebanyak 56%, dan *cluster 3* sebanyak 57%. Pada prodi Teknik Industri UII sendiri sudah beberapa kali mengadakan sertifikasi yang dapat meningkatkan keahlian bidang teknik industri mahasiswa walaupun belum rutin diadakan sehingga strategi ini disarankan dengan harapan agar prodi Teknik Industri UII dapat lebih sering mengadakan sertifikasi profesional agar keahlian dan keterampilan mahasiswa dalam bidang teknik industri semakin meningkat.

Strategi yang keempat adalah melakukan kunjungan industri terutama kunjungan industri ke perusahaan-perusahaan besar dimana dalam pelaksanaan kunjungan industri ini mahasiswa dapat diberikan studi kasus dari perusahaan sehingga dapat membantu mahasiswa dalam memahami kondisi nyata dalam dunia kerja seperti wawasan mengenai bagaimana perusahaan-perusahaan tersebut mengelola proses produksi dan *supply chain* dalam skala besar.

Perancangan strategi ini didasari oleh banyaknya perusahaan yang memberikan kualifikasi keahlian teknik industri pada *cluster 1*, *cluster 2*, dan *cluster 3* dimana pada cluster 1 sebanyak 77%, cluster 2 sebanyak 56%, dan cluster 3 sebanyak 57%. Pada prodi Teknik Industri UII sendiri sudah beberapa kali melaksanakan kunjungan industri pada beberapa mata kuliah. Namun kunjungan industri belum rutin dilakukan sehingga strategi ini disarankan dengan harapan agar prodi Teknik Industri UII dapat lebih sering mengadakan kunjungan industri agar keahlian dan keterampilan mahasiswa dalam bidang teknik industri semakin meningkat.

Strategi yang kelima adalah penerapan mata kuliah Kepemimpinan dan Pengambilan Keputusan dimana mata kuliah ini dapat membantu mahasiswa dalam mengembangkan kemampuan kepemimpinan dan pengambilan keputusan dengan cara membekali mahasiswa dengan konsep kepemimpinan, pengambilan keputusan, manajemen tim, dan studi kasus kepemimpinan. Perancangan strategi ini didasari oleh banyaknya perusahaan yang memberikan kualifikasi kemampuan *leadership* terutama pada *cluster 1* yaitu sebanyak 86%, kemudian pada *cluster 3* yaitu sebanyak 55%. Pada prodi Teknik Industri UII sendiri terdapat mata kuliah *Student Soft Skill & Development*. Namun mata kuliah ini berfokus pada pengembangan soft skill secara umum dan tidak berfokus pada kemampuan kepemimpinan sehingga strategi ini disarankan dengan harapan agar prodi Teknik Industri UII dapat mengadakan mata kuliah Kepemimpinan dan Pengambilan Keputusan agar dapat meningkatkan kemampuan kepemimpinan mahasiswa.

5.4.1 *Cluster 1*

Cluster 1 terdiri dari data lowongan kerja yang didominasi oleh perusahaan dengan sektor manufaktur dan posisi di bagian operasional produksi dan untuk mengisi posisi pada bagian operasional produksi di perusahaan sektor manufaktur tersebut, banyak perusahaan yang memberi kualifikasi kemampuan komunikasi, kemampuan *leadership*, dan keahlian bidang teknik industri. Selain itu, kualifikasi keahlian *software* dan memiliki etos kerja juga cukup banyak diberikan oleh perusahaan. Sehingga strategi yang disarankan adalah pelaksanaan *training Microsoft Office*, penerapan mata kuliah Pengembangan Kepribadian dan Etika Profesi, pelaksanaan sertifikasi profesional, melakukan kunjungan industri, dan penerapan mata kuliah Kepemimpinan dan Pengambilan Keputusan.

5.4.2 Cluster 2

Cluster 2 terdiri dari data lowongan kerja yang didominasi oleh perusahaan dengan sektor manufaktur serta posisi di bagian operasional produksi dan *managerial*. Untuk mengisi posisi pada bagian operasional produksi dan *managerial* di perusahaan sektor manufaktur tersebut, banyak perusahaan yang memberi kualifikasi keahlian *software* dan keahlian bidang teknik industri serta cukup banyak perusahaan yang memberikan kualifikasi kemampuan komunikasi. Sehingga strategi yang disarankan adalah pelaksanaan *training Microsoft Office*, pelaksanaan sertifikasi profesional, melakukan kunjungan industri, dan penerapan mata kuliah Kepemimpinan dan Pengambilan Keputusan.

5.4.3 Cluster 3

Cluster 3 terdiri dari data lowongan kerja yang didominasi oleh perusahaan dengan sektor dagang diikuti dengan perusahaan dengan sektor jasa serta posisi di bagian *marketing* dan operasional produksi. Untuk mengisi posisi pada perusahaan dengan kriteria tersebut, banyak perusahaan yang memberi kualifikasi kemampuan komunikasi, keahlian bidang teknik industri, dan kemampuan *leadership*. Sehingga strategi yang disarankan adalah penerapan mata kuliah Pengembangan Kepribadian dan Etika Profesi, pelaksanaan sertifikasi profesional, melakukan kunjungan industri, dan penerapan mata kuliah Kepemimpinan dan Pengambilan Keputusan.

BAB VI

PENUTUP

6.1 Kesimpulan

Berdasarkan hasil penelitian yang didapatkan, berikut ini merupakan kesimpulan dari penelitian:

1. Karakteristik yang terbentuk dari proses *clustering* menggunakan *K-Means Clustering* menunjukkan bahwa *cluster 1* terdiri dari data lowongan kerja yang didominasi oleh perusahaan dengan sektor manufaktur dan posisi di bagian operasional produksi dan untuk mengisi posisi pada bagian operasional produksi di perusahaan sektor manufaktur tersebut, banyak perusahaan yang memberi kualifikasi kemampuan komunikasi, kemampuan *leadership*, dan keahlian bidang teknik industri. Selain itu, kualifikasi keahlian *software* dan memiliki etos kerja juga cukup banyak diberikan oleh perusahaan. *Cluster 2* terdiri dari data lowongan kerja yang didominasi oleh perusahaan dengan sektor manufaktur serta posisi di bagian operasional produksi dan *managerial*. Untuk mengisi posisi pada bagian operasional produksi dan *managerial* di perusahaan sektor manufaktur tersebut, banyak perusahaan yang memberi kualifikasi keahlian *software* dan keahlian bidang teknik industri serta cukup banyak perusahaan yang memberikan kualifikasi kemampuan komunikasi. Kemudian yang terakhir, *cluster 3* terdiri dari data lowongan kerja yang didominasi oleh perusahaan dengan sektor dagang diikuti dengan perusahaan dengan sektor jasa serta posisi di bagian *marketing* dan operasional produksi. Untuk mengisi posisi pada perusahaan dengan kriteria tersebut, banyak perusahaan yang memberi kualifikasi kemampuan komunikasi, keahlian bidang teknik industri, dan kemampuan *leadership*.
2. Strategi yang disarankan untuk pengembangan profil lulusan teknik industri UII adalah yang pertama pada *cluster 1* strategi yang disarankan adalah pelaksanaan *training Microsoft Office*, penerapan mata kuliah Pengembangan Kepribadian dan Etika Profesi, pelaksanaan sertifikasi profesional, melakukan kunjungan industri, dan penerapan mata kuliah Kepemimpinan dan Pengambilan Keputusan. Pada *cluster 2* strategi yang disarankan adalah pelaksanaan *training Microsoft Office*, pelaksanaan sertifikasi profesional, melakukan kunjungan industri, dan penerapan mata kuliah Kepemimpinan dan Pengambilan Keputusan. Pada *cluster 3* strategi yang disarankan adalah penerapan mata

kuliah Pengembangan Kepribadian dan Etika Profesi, pelaksanaan sertifikasi profesional, melakukan kunjungan industri, dan penerapan mata kuliah Kepemimpinan dan Pengambilan Keputusan.

6.2 Saran

Berikut ini merupakan saran yang dapat diberikan berdasarkan penelitian yang telah dilakukan.

1. Untuk penelitian selanjutnya dapat menambahkan variabel penelitian seperti membagi variabel keahlian bidang teknik industri menjadi lebih spesifik sehingga memungkinkan untuk mengurangi tumpang tindih atau *overlapping* antar *cluster* saat dilakukan proses *clustering* serta dapat berguna untuk mendapatkan hasil *clustering* yang lebih akurat.
2. Untuk penelitian selanjutnya juga dapat menggunakan metode lain untuk proses *clustering* seperti *Fuzzy C-Means* dan *Gaussian Mixture Models* yang memungkinkan data untuk berada pada lebih dari satu *cluster* dengan tingkat probabilitas yang berbeda untuk mendapatkan hasil yang lebih akurat.
3. Untuk Program Studi Teknik Industri Universitas Islam Indonesia dapat menggunakan strategi yang disarankan untuk pengembangan profil lulusan teknik industri seperti strategi penerapan mata kuliah pengembangan kepribadian dan etika profesi, mata kuliah kepemimpinan dan pengambilan keputusan, dan pelaksanaan *training Microsoft Office*. Untuk Strategi seperti kunjungan industri dan pelaksanaan sertifikasi profesional yang sudah ada dapat dilakukan dengan lebih konsisten sehingga membantu mahasiswa untuk lebih siap dalam menghadapi dunia kerja.
4. Untuk Program Studi Teknik Industri Universitas Islam Indonesia dapat memiliki database alumni yang lebih lengkap sehingga dapat memungkinkan dilakukannya perbandingan antara profil lulusan yang diharapkan di dunia kerja dengan kenyataannya.

DAFTAR PUSTAKA

- Abdullah, D. et al., 2022. The Application of K-Means Clustering for Province Clustering in Indonesia of The Risk of The Covid-19 Pandemic Based on Covid-19 Data. *Quality & Quantity*, 56(3), pp. 1283-1291.
- Adnan, R. M. et al., 2021. Suspended Sediment Modeling Using Heuristic Regression Method Hybridized with K-Means Clustering. *Sustainability*, 13(9), p. 4648.
- Adwitya, G. N. D., Shaufiah, S. & Effendy, V., 2015. Analisis dan Implementasi Algoritma Overlapping Cover Coefficient-Based Clustering Method (OC3M) pada Dokumen Teks Berbahasa Indonesia. *eProceeding of Engineering*, 2(2), pp. 6315-6322.
- Al Qadri, M., Zakaria, M. R. & Bukhari, I., 2024. Pendampingan dan Pelatihan Microsoft Office Untuk Meningkatkan Keterampilan Mahasiswa STAI-JM dan Guru YASPEND Ar-Ridha Tanjung Pura. *Center of Knowledge: Jurnal Pendidikan Dan Pengabdian Masyarakat*, 4(1), pp. 36-44.
- Apriana, D. & Yuliansyah, C., 2024. Optimalisasi Penjualan Online Melalui Teknik Data Mining (Studi Kasus E-Commerce). *AL-MIKRAJ Jurnal Studi Islam dan Humaniora*, 4(02), pp. 514-527.
- Astuti, E. D., 2021. Berpikir Kritis Mahasiswa dalam Mata Kuliah Teknik Pengambilan Keputusan. *Prosiding Seminar Nasional Ilmu Sosial dan Teknologi*, Volume 1, pp. 40-44.
- Bu, J., Liu, W., Pan, Z. & Ling, K., 2020. Comparative Study of Hydrochemical Classification Based on Different Hierarchical Cluster Analysis Methods. *International Journal of Environmental Research and Public Health*, 17(24), p. 9515.
- Chairani, S. & Nurhazana, N., 2020. Peran Mata Kuliah Etika Profesi Terhadap Perkembangan Perilaku Etis Mahasiswa. *Jurnal IAKP: Jurnal Inovasi Akuntansi Keuangan & Perpajakan*, 1(2), pp. 114-127.
- Dahnial, D., 2023. Implementation of K-Means Clustering Method to Lectures Based on Publications of National Journals and Accredited Sinta. *JEECS (Journal of Electrical Engineering and Computer Sciences)*, 8(1), pp. 27-40.
- El Hafiz, S. & Himawan, K. K., 2021. Tantangan Melakukan Kajian Literatur Psikologi di Indonesia: Masalah Mendasar dan Solusinya. *Jurnal Psikologi Ulayat*, 8(1), pp. 6-17.
- Gao, Y., Fang, H. & Ni, K., 2021. A hierarchical clustering method of hydrogen bond networks in liquid water undergoing shear flow. *Scientific Reports*, 11(1), p. 9542.
- Ghosal, S., Bhattacharyya, R. & Majumder, M., 2020. Impact of Complete Lockdown on Total Infection and Death Rates: A Hierarchical Cluster Analysis. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, 14(4), pp. 707-711.
- Haryanto, D. C., Harini, S. & Chamidy, T., 2024. K-Means Clustering dalam Pengelompokan Relevansi Pekerjaan S1 Informatika (Studi Kasus Jurusan Teknik Informatika UMM Malang). *JUPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, 9(2), pp. 792-797.

- He, Y. et al., 2022. Factors Influencing Carbon Emissions from China's Electricity Industry: Analysis Using the Combination of LMDI and K-Means Clustering. *Environmental Impact Assessment Review*, Volume 93, p. 106724.
- Hidayat, R. & Zainuddin, M., 2017. Profil Lulusan Program Studi Administrasi Bisnis Terapan Tahun 2015 dan Relevansinya dengan Penyerapan Dunia Kerja. *Jurnal Akuntansi, Ekonomi, dan Manajemen Bisnis*, 5(1), pp. 91-104.
- Lin, X. & Xu, J., 2020. Road Network Partitioning Method Based on Canopy-K-Means Clustering Algorithm. *Archives of Transport*, 54(2), pp. 95-106.
- Lisdarti, L., 2021. Analisa dan Perancangan Web Profil Perusahaan (Studi Kasus: Butik Chaniago Jambi). *FORTECH (Journal of Information Technology)*, 5(1), pp. 41-46.
- Liu, X. et al., 2021. The Water Supply Association Analysis Method in Shenzgen Bases on K-Means Clustering Discretization and Apriori Algorithm. *PloS one*, 16(8), pp. 1-21.
- Maori, N. A. & Evanita, E., 2023. Metode Elbow dalam Optimasi Jumlah Cluster pada K-Means Clustering. *Simetris: Jurnal Teknik Mesin, Elektro dan Ilmu Komputer*, 14(2), pp. 277-288.
- Maryana, et al., 2022. *Menyiapkan Soft Skills bagi Lulusan Mahasiswa Kesehatan*. Bantul: Samudra Biru.
- Maulana, R. et al., 2023. Komparasi Algoritma Decision Tree dan Naive Bayes dalam Klasifikasi Penyakit Diabetes. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(6), pp. 3865-3870.
- Meilani, N. & Nurdiawan, O., 2023. Data Mining untuk Klasifikasi Penderita Kanker Payudara Menggunakan Algoritma K-Nearest Neighbor. *Jurnal Wahana Informatika*, 2(1), pp. 177-187.
- Metisen, B. M. & Sari, H. L., 2015. Analisis Clustering Menggunakan Metode K-Means dalam Pengelompokan Penjualan Produk pada Swalayan Fadhila. *Jurnal Media Infotama*, 11(2), pp. 110-118.
- Mulyanan, I. N., Setyawan, M. Y. H. & Rahayu, W. I., 2023. Penerapan Metode Naive Bayes untuk Merekomendasikan Pekerjaan yang Sesuai Terhadap Fresh Graduate. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(5), pp. 3453-3460.
- Nagari, S. S. & Inayati, L., 2020. Implementation of Clustering Using K-Means Method to Determine Nutritional Status. *Jurnal Biometrika dan Kependudukan*, 9(1), pp. 62-68.
- Normah, N., Nurajizah, S. & Salbinda, A., 2021. Penerapan Data Mining Metode K-Means Clustering Untuk Analisa Penjualan Pada Toko Fashion Hijab Banten. *Jurnal Khatulistiwa Informatika*, 7(2), pp. 158-163.
- Nurdiawan, O. & Salim, N., 2018. Penerapan Data Mining pada Penjualan Barang Menggunakan Metode Naine Bayes Classifier untuk Optimasi Strategi Pemasaran. *Jurnal Teknologi Informasi dan Komunikasi*, 13(1), pp. 84-95.
- Nurhayati, E. & Yunitasari, E. W., 2020. Strategi Peningkatan Kualitas Lulusan Mahasiswa Program Studi Teknik Industri Universitas XYZ Yogyakarta. *Jurnal Optimasi Sistem Industri*, 13(1), pp. 70-75.

- Othman, Z., Shan, S. W., Yusoff, I. & Kee, C. P., 2018. Classification Techniques for Predicting Graduate Employability. *International Journal on Advanced Science Engineering Information Technology*, 8(4-2), pp. 1712-1720.
- Prianto, C. & Bunyamin, S., 2020. *Panduan Pembuatan Aplikasi Clustering Gangguan Jaringan Menggunakan Metode Kmeans Clustering*. Bandung: Kreatif Industri Nusantara.
- Rahman, F. D., Zulfa, M. I. & Taryana, A., 2024. Clustering dan Klasifikasi Data Cuaca Kota Cilacap Menggunakan K-Means dan Random Forest. *Jurnal SINTA: Sistem Informasi dan Teknologi Komputasi*, 1(2), pp. 90-97.
- Sahara, S., Amri, S., Ningrum, A. F. & Purwanto, D., 2024. Dashboard Lingkungan Hidup untuk Analis Diare Menggunakan Metode K-Means Clustering. *Journal of Data Insights*, 2(1), pp. 18-25.
- Sholeha, D. et al., 2024. Peningkatan Wawasan Mahasiswa Dengan Kunjungan Industri Ke PLTA PT. Indonesia Asahan Aluminium (INALUM). *Medani: Jurnal Pengabdian Masyarakat*, 3(1), pp. 16-21.
- Sutrisno, J., Wibowo, A. & Pratama, B. S., 2023. Klasterisasi Data Hasil Studi Pelacakan Tentang Karir dan Pekerjaan Lulusan Perguruan Tinggi Menggunakan Algoritma K-Means. *J-Icon : Jurnal Komputer dan Informatika*, 11(2), pp. 157-164.
- Syaripudin, U. et al., 2013. Studi Komparatif Penerapan Metode Hierarchicalm K-Means, dan Self Organizing Maps (SOM) Clustering pada Basis Data. *Jurnal ISTEK*, 7(1), pp. 132-149.
- Tomar, D. & Agarwal, S., 2013. A survey on Data Mining approaches for Healthcare. *International Journal of Bio-Science and Bio -Technology*, 5(5), pp. 241-266.
- Utari, D. R. & Sofiani, F., 2021. Klasterisasi Lulusan Pendidikan Tinggi Vokasi Berbasis Hasil Studi Pelacakan Menggunakan Algoritma K-Means. *Jurnal Sekretari & Administrasi (Serasi)*, 19(2), pp. 75-86.
- Utomo, Y. B., Kurniasari, I. & Yanuartanti, I., 2023. Penerapan Knowledge Discovery in Database untuk Analisa Tingkat Kecelakaan Lalu Lintas. *Jurnal Teknik Informatika Kaputama (JTIK)*, 7(1), pp. 171-180.
- Wakhidah, N., 2010. Clustering Menggunakan K-Means Algorithm. *Jurnal Transformatika*, 8(1), pp. 33-39.
- Widaningsih, S., 2019. Perbandingan Metode Data Mining untuk Prediksi Nilai dan Waktu Kelulusan Mahasiswa Prodi Teknik Informatika dengan Algoritma C4.5, Naive Bayes, KNN, dan SVM. *Jurnal Tekno Insentif*, 13(1), pp. 16-25.
- Zebua, R. T., Saragih, N. F. & Larosa, F. G. N., 2022. Pengukuran Tingkat Kesesuaian Pekerjaan Alumni Berdasarkan Profil Lulusan Dengan Metode Hierarchical Clustering. *METHOSISFO: Jurnal Ilmiah Sistem Informasi*, 2(2), pp. 1-11.
- Zhang, Q., Fang, S. & Ye, H., 2024. Research on Clustering of Material Inspection Data for Highway Construction Projects Based on Agglomerative Hierarchical Clustering. *Journal of Physics: Conference Series*, 2706(1), pp. 1-7.

LAMPIRAN
A-Tabel Uraian Variabel Posisi

Posisi			
Operasional Produksi	Marketing	IT	Managerial
AGRICULTURE SPECIALIST	Account Executive (IT Sales)	Business Analyst	Ahli K3 Kimia
Area Manager Grosir	Account Manager	Business Process Analyst	Asst. Manager General Affair
Asisten Logistik	AREA PART DEVELOPMENT	CNC Programmer	Business Process Officer
Cost Control Production	Business Development & Account Executive	COLLECTION DATABASED ANALYST	Business Process Staff
Design Engineering	Kaizen Staff	staff SAP	COMMUNITY COORDINATOR STAFF
Drafter CAD	Manager Bisnis	SAP Abaper	Compliance Staff
Engineering Production Coordinator	Marketing & Sales	CS Network Data Analyst	Costing Staff
Factory Manager	Marketing Communication	OBSdigital designer (OS)	Engineering Analysis
FOLLOW UP DYEING & FINISHING	Marketing Project	business analytics supervisor	General Affair
Foodcourt Manager	Product Marketing Officer		General Affair Administrator
FOREMAN BAHAN BAKU & OPERATOR PRODUKSI	RETAIL BUSINESS OFFICER		HR Training & Staff Personalia
KAIZEN TEAM	Sales B2B		Internal Audit Officer
Kasie Produksi	Sales Engineer		Internal Audit Staff
Kasie Produksi Packaging	Sales Executive		Junior Auditor
Kepala Produksi	Sales Machinery Manager		Logistic Excellence
Maintenance Section Head	Staff Marketing Logistic		Management Trainee (MT)
Manager Operasional	Store Manager		Mechanical Costing Staff
Manager PPIC	Technical Marketing Staff		MEP Document Controller

Manager SCM (Supply Chain Management)	supervisor retail engagement		Organization Development Commercial Officer
Manajer Restoran Sederhana	User Growth Associate		Organization Development Staff
Mechanical Engineer	Profit revenue growth manager		Personalia & Payroll Staff
PIC & Koordinator Lapangan	HNS Representative		Planner Logistik (CS2)
PIC Project Interior	brand inovation manager		Policy & Process Development
PPIC Admin			PROCUREMENT STAFF
PPIC Material Planner			Purchasing Staff
PPIC Officer			QUALITY MANAGEMENT SYSTEM STAFF
PPIC Staff			Revenue Staff
PPIC Supervisor			Safety Officer
Primary Factory QA Engineer			Staff Continous Improvement
Production Assistant Manager			Staff HRD
PRODUCTION CONTROL			Staff ISO
Production Engineering Staff			Staff Kalibrasi
Production Manager			Staff Purchasing Jewelry
Production Section Head			Staff Research and Development
Project Manager			System and Procedure Officer
Quality Assurance Engineer			HSE staff
Quality Assurance Staff			commercial process analyst
Quality Control Inspector			product specialist
Quality Control staff			employee and labor relations lead

Quality Management Representative (QMR)			staff ER & Personnel Adm
Staff Engineering			staff compliance
Staff Logistik Executive			People services lead
Staff Perizinan Halal			management system
Staff Produksi			human resources & planning system team member
Staff Warehouse			HR Benefit specialist
Supply Demand Supervisor			HSE specialist
Technical Product Development			Manager HR & Sustainability
Teknisi dan Maintenance			Talent Acquisition Officer
Quality Engineer			safety & utility part leader
regional distribution executive			R&D staff & Interpreter
E2E Supply Chain			P2P supporting & reporting
staff industrial engineering			sc customer solution manager
resources & recycling engineer			
Bulk warehouse			
specialist heavy equipment			
Reability Analyst			

B-Tabel Uraian Variabel Kualifikasi

Kualifikasi					
Hard Skill			Soft Skill		
Keahlian Software	Keahlian Bahasa	Keahlian Bidang Teknik Industri	Kemampuan Komunikasi yang Baik	Kemampuan Kepemimpinan yang Baik	Memiliki Etos Kerja yang Tinggi
Keahlian Microsoft Office (excel, word, power point, microsoft project)	Menguasai Mandarin	Memahami lean manufacturing	Memiliki kemampuan komunikasi yang baik	Memiliki kemampuan kepemimpinan	Teliti dan perhatian terhadap detail
Menguasai software Autocad	Menguasai Bahasa Inggris	Memahami Supply Chain Management	Memiliki kemampuan negosiasi	Memiliki kemampuan decision making yang baik	Berorientasi pada target
Menguasai software Solidworks	Menguasai Bahasa Jepang	Memahami mesin produksi	Mampu bekerja sama dalam tim	Tegas dan berani mengatur bawahan	Mampu bekerja di bawah tekanan
Menguasai aplikasi Wiring Diagram		Memahami inventory control & management	Memiliki kemampuan presentasi yang baik	Memiliki wawasan yang luas	Memiliki kemampuan adaptasi yang baik
Menguasai mastecam		Memahami warehouse management	Memiliki koneksi	Memiliki kreativitas yang tinggi	Multitasking dan bisa bekerja secara efisien
Menguasai aplikasi Design (photoshop, corel draw, dll)		Memahami tentang distribusi		Memiliki kemampuan analisis yang kuat	Memiliki manajemen waktu yang baik
Menguasai aplikasi inventory stock		Memahami tentang alur proses produksi		Memiliki kemampuan problem solving	Memiliki kepribadian & attitude yang baik
Menguasai SAP ERP		Memahami tentang PPIC dan tugasnya			Memiliki tanggung jawab

Menguasai Google Sketch Up		Mampu membaca & membuat gambar teknik			Memiliki semangat yang tinggi
Menguasai sistem accurate		Memiliki kemampuan teknis & lapangan			Menyukai tantangan
Mampu mengoperasikan komputer		Menguasai control CNC			Menyukai pekerjaan di lapangan
		Memahami tentang demand management			Dapat belajar dengan cepat
		Memahami tentang perencanaan produksi			
		Memahami VSM, Time Study atau Motion Study			
		Menguasai penjadwalan			
		Memahami tentang purchasing			
		Menguasai alat ukur (kaliper, micrometer, high gauge, dll)			
		Mahir dalam matematika dan perhitungan			
		Memahami MRP			
		Memahami standar ISO			
		Memahami standar SNI			

		Memahami Food Safety Management System			
		Memahami SHE System			
		Memahami standar GMP			
		Memahami HACCP			
		Memahami 5R			
		Memiliki keahlian K3			
		Memahami tentang CSMS			
		Memahami tentang Emergency Response Plan			
		Memiliki keahlian manajemen risiko			
		Memiliki keahlian manajemen proyek			
		Menguasai data analytics			
		Memahami tentang manajemen SDM			
		Memahami tentang manajemen data			
		Memahami tentang quality control & management system			

		Memiliki keahlian management			
		Memahami strategi dan manajemen penjualan serta pemasaran			
		Memahami tentang customer relation management			
		Mampu membuat RAB			
		Memahami tentang produk yang dijual			
		Memiliki keahlian welding inspector			
		Memahami tentang General Affair dan tugasnya			
		Memahami alur forecasting			
		Memahami tentang proses bisnis			
		Memiliki kemampuan filing yang jelas dan akurat			
		Memahami kegiatan administrasi			
		Mampu membuat program kerja			

		Mampu membuat konsep & analisa marketing			
		Memahami Profit & Lost			
		Memahami sertifikat jaminan halal			
		Menguasai tentang manajemen operasional produksi			
		Memahami human resource			

C-Anggota Cluster 1

Cluster 1		Cluster 1		Cluster 1		Cluster 1	
No.	Nama Perusahaan	No.	Nama Perusahaan	No.	Nama Perusahaan	No.	Nama Perusahaan
1	PT. Cilia Prisma Utama Ma	51	PT. Tidung Jaya Mandiri I	101	DHARMA GROUP	151	PT Yamaha Music Manufactu
2	CV. Multilami Plasindo	52	PT. Rancang Resto Indones	102	PT Industrial Robotic Aut	152	PT Yamaha Music Manufactu
3	PT. Wahana Kasih Mulia	53	PT. Tectona Sinergi Jaya	103	PT Gema Mulia Semesta	153	Danone Indonesia
4	Zeppelin Indonesia	54	PT. Mustika Land	104	PT. Tri-Saudara Sentosa I	154	Danone Indonesia
5	PT. Furindo Sakti Sejahte	55	PT. Dukung Aksi Rakyat Ev	105	Sun Paper Source	155	Danone Indonesia
6	PT. Sebastian Jaya Metal	56	PT. Jet Transport Service	106	PT. NAMI SELERA GLOBAL	156	Danone Indonesia
7	PT. Formosa Ingredient Fa	57	PT. Muliaoffset Packindo	107	PT IGP Internasional	157	Danone Indonesia
8	PT. Dwikar Robinson	58	PT. Hasanayu Putri Jelita	108	PT. Anugerah Mekanika Suk	158	Danone Indonesia
9	PT. Newtrend nutrition In	59	PT. Mempeng Pilar Kreasi	109	PT Sariguna Primatirta Tb	159	Danone Indonesia
10	PT. Cilia Prisma Utama Ma	60	PT. Sumiden Sintered Comp	110	PT. Dwida Jaya Tama	160	Danone Indonesia
11	PT. Sinar Sakti Jaya	61	PT Stampindo Lancar Jaya	111	PT. LOTUS LINGGA PRATAMA	161	PT Riau Andalan Pulp and
12	Anaria Souvenir	62	PT. International Leather	112	PT Tirta Alam Segar	162	Asia Pulp & Paper
13	PT. Cilia Prisma Utama Ma	63	Factory Print	113	PT. Pou Yuen Indonesia	163	Asia Pulp & Paper
14	Greebel	64	PT. Hokinda Citra Lestari	114	PT Bogajaya Manna Lestari	164	Asia Pulp & Paper
15	PT. Inter Mitra Trasindo	65	PT. Sorin Maharasa	115	PT Multi Indojoya Makmur	165	Asia Pulp & Paper
16	PT. Karya Artha Perdana	66	Bisnis Indonesia Group	116	PT Yamaha Indonesia Motor	166	Asia Pulp & Paper
17	PT. Pelita Wijaya Karya	67	Telasih Makmur Offset Pri	117	Technoskin Global Kosmeti	167	Asia Pulp & Paper
18	PT. Agung Manufaktur Desa	68	PT. Padama Bahtera Labeli	118	CV. KHARIS INDO PASTI	168	PT Kalbe Farma Tbk
19	PT. Ferrum Design	69	PT. Sunson Textile Manufa	119	PT OMEGA STEEL PERKASA	169	PT Nestle Indonesia
20	PT. Berkah Multitech	70	PT. Out Of Asia	120	PT Triyuda Perkasa	170	PT Nestle Indonesia
21	CV. Cahaya Prima Mandiri	71	Union Plastik	121	Cimory Group	171	PT Nestle Indonesia
22	PT. Sinergi Abadi Sentosa	72	PT. Enerel Kosmetika Biot	122	PT Karet Ngagel Surabaya	172	PT Nestle Indonesia
23	PT. Jabaku Karaba Technol	73	PT. Bumindo Sakti	123	PT Sri Rejeki Isman Tbk	173	PT Nestle Indonesia

Cluster 1		Cluster 1		Cluster 1		Cluster 1	
No.	Nama Perusahaan	No.	Nama Perusahaan	No.	Nama Perusahaan	No.	Nama Perusahaan
24	PT. Indoseiki Metal Utama	74	PT. Hokinda Citralestari	124	PT Sri Rejeki Isman Tbk	174	PT Nestle Indonesia
25	PT. Star Abadi Ratu Indon	75	PT. Sinar Continental	125	PT Indorama Synthetics Tb	175	PT Krakatau Steel Persero
26	PT. Megacon Bangun Perkas	76	PT. Sinergi Inti Plastind	126	PT Samsung Electronics In	176	PT Krakatau Steel Persero
27	PT. Arisa Mandiri Pratama	77	Kompas Gramedia	127	PT Samsung Electronics In	177	PT Krakatau Steel Persero
28	PT. Furindo Sakti Sejahte	78	Astra Honda Motor	128	PT LG Electronics Indones	178	PT Krakatau Steel Persero
29	PT. Pinako Rotari Permai	79	PT. Soraya Interindo	129	PT LG Electronics Indones	179	PT Philips Industries
30	PT. Ungaran Wana Karya	80	PT. Arkanindoplast Utama	130	PT LG Electronics Indones	180	Adidas Group
31	PT. Ungaran Wana Karya	81	PT. Indonesia Quality Foc	131	PT LG Electronics Indones	181	PT Pan Brothers Tbk
32	Konveksi Toko Abi	82	PT. Multi Sari Sedap	132	PT Greenfields Indonesia	182	Johnson & Johnson
33	PT. Rukun Sejahtera Tekni	83	PT. Petra Sejahtera Abadi	133	PT Chevron Pacific Indone	183	Johnson & Johnson
34	PT. Benza Teknik Mandiri	84	Massindo Group	134	PT Chevron Pacific Indone	184	Johnson & Johnson
35	PT. Suly Bersama Jaya Ste	85	Papandayan Cargo	135	PT Chevron Pacific Indone	185	PT L'Oreal Indonesia
36	PT. Newtrend nutrition In	86	PT. Erlimpex	136	PT Chevron Pacific Indone	186	Boeing
37	PT. Madusari Nusaperdana	87	PT. CJ Foodville Bakery a	137	PT Indocement Tunggal Pra	187	PT Coca-Cola Amatil Indon
38	PT. Inspiran Ideal Indota	88	PT Leading Garment Indust	138	PT Indocement Tunggal Pra	188	PT HM Sampoerna Tbk
39	grob.media	89	Oelek	139	PT Industri Jamu dan Farm	189	PT HM Sampoerna Tbk
40	grob.media	90	PT. Furindo Sakti Sejahte	140	PT ABC President Indonesi	190	PT HM Sampoerna Tbk
41	PT. Arpu Selaras Cemerlan	91	CV. Cahaya Prima Mandiri	141	Gojek Indonesia	191	PT HM Sampoerna Tbk
42	PT Hasanayu Putri Jelita	92	PT. Hema Medhajaya	142	GoTo Group	192	PT HM Sampoerna Tbk
43	PT. Adiprima Suraprinta	93	PT. Sungai Budi	143	Gojek Indonesia	193	PT HM Sampoerna Tbk
44	PT Solomon Indo Global	94	PT. Tandi Tirta Mas	144	PT GoTo Gojek Tokopedia	194	PT Unilever Indonesia Tbk
45	PT. Annjoy Indonesia	95	PT. Rivindi Artha Mandiri	145	PT Pabrik Kertas Tjiwi Ki	195	PT Unilever Indonesia Tbk
46	PT. Maxindo Karya Anugera	96	PT. Cahaya Bintang Olympi	146	PT Pabrik Kertas Tjiwi Ki	196	PT Unilever Indonesia Tbk
47	Chocolate Monggo	97	PT. Mega Niaga Sejahtera	147	PT Pupuk Indonesia	197	PT Unilever Indonesia Tbk
48	CV. Perajutan Sahabat	98	Samora Group	148	PT Vale Indonesia Tbk		
49	PT. Etta Indotama Persada	99	PT Kosena Lestari Makmur	149	PT Unicharm Indonesia Tbk		
50	PT. Semanggi Tiga	100	PT Pipamas Primasejati	150	PT Samator Indo Gas Tbk		

D-Anggota Cluster 2

Cluster 2		Cluster 2		Cluster 2		Cluster 2	
No.	Nama Perusahaan	No.	Nama Perusahaan	No.	Nama Perusahaan	No.	Nama Perusahaan
1	PT. Ridho Agung Mitra Aba	41	PT. Tiga Rasa Indonesia S	81	PT Supernova Flexible Pac	121	PT Milan Ecowood Indonesi
2	Purwana Group	42	Indofood CBP	82	PT Pabrik Kertas Indonesi	122	PT Hokkan Deltapack Indus
3	PT. Imecon Teknindo	43	Dekoruma	83	PT Propan Raya Industrial	123	PT Duta Nichirindo Pratam
4	DMC Interior	44	PT. Kairos Utama Indonesi	84	PT Rusli Vinilon Sakti	124	PT Arwana Citramulia Tbk
5	PT. Delta Jaya Mas	45	Indofood CBP	85	PT Arwana Citramulia Tbk	125	PT Propan Raya Industrial
6	PT. Sebastian Jaya Metal	46	PT. Family Indo Pratama I	86	PT Sinar Metrindo Perkasa	126	PT Rototama Berlianplast
7	PT. Madusari Nusaperdana	47	PT. Nippon Indosari Corpi	87	Tunggal Idaman Abdi	127	PT Gunanusa Eramandiri
8	PT. Gebang Surya Harapan	48	PT. Mega Central Finance	88	PT Baskara Prima Sarana	128	PT Tins Jaya Abadi
9	CV. Weston Textile	49	PT. Distritama Perkasa	89	PT Sari Kresna Kimia	129	PT Mizan Pustaka
10	PT. Rajawali Karya Perkas	50	PT. Hema Medhajaya	90	PT Margacipta Wirasantosa	130	PT Hokinda Citralestari
11	PT. Lautan Energi Nusanta	51	PT. Rollent Indonesia	91	Orang Tua Group	131	PT Bukit Muria Jaya
12	PT. Formosa Ingredient Fa	52	PT. Bhakti Bersama Roarth	92	PT Nojorono Tobacco Inter	132	PT LG Electronics Indones
13	PT. Formosa Ingredient Fa	53	PT. Tera Data Indonusa, T	93	PT Yamaha Indonesia Motor	133	PT LG Electronics Indones
14	PT. Fico Indo Rekayasa	54	PT. Margacipta Wirasantos	94	PT Rodamas	134	PT Indofood CBP Sukses ma
15	PT. Maxon Prime Technolog	55	PT. Cahaya Bintang Plasti	95	SINAR ANTJOL GROUP (B29)	135	PT ABC President Indonesi
16	PT. Sari Kresna Kimia	56	PT. Duniatex Group	96	PT Penguin Indonesia	136	PT ABC President Indonesi
17	PT. Insan daya sentosa	57	PT. Faber Castell Interna	97	PT Rekatama Inti Anugerah	137	PT ABC President Indonesi
18	PT. Tabitha Express	58	SATEX Group	98	PT Epiterma Mas Indonesia	138	PT Vale Indonesia Tbk
19	PT. Suryacipta Interindo	59	PT. Prokemas Adhikari Kre	99	PT Multi Indomandiri	139	PT Vale Indonesia Tbk
20	Salsa Beauty Centre	60	PT. Santos Jaya Abadi	100	PT Richeese Kuliner Indon	140	PT Unicharm Indonesia Tbk
21	PT. Federal Mardhika Citr	61	PT. Ibara Lioho Indonesia	101	PT Epiterma Mas Indonesia	141	PT Samator Indo Gas Tbk
22	PT. Cosmopack Plastindo U	62	PT. Bajaindo Eraprima	102	PT Agro Mitra Alimentare	142	PT Samator Indo Gas Tbk
23	PT. Maju Makmur Utomo	63	PT. Sanggar Sarana Baja	103	PT Cendrawasih Garmindo I	143	PT Samator Indo Gas Tbk

Cluster 2		Cluster 2		Cluster 2		Cluster 2	
No.	Nama Perusahaan	No.	Nama Perusahaan	No.	Nama Perusahaan	No.	Nama Perusahaan
24	PT. Changhong Electric In	64	PT. Adil Makmur Fajar	104	PT SADHANA	144	PT Samator Indo Gas Tbk
25	PT. Dharma Karya Perdana	65	Cimory Group	105	PT Onna Prima Utama	145	PT Samator Indo Gas Tbk
26	PT. Aisa Food Industry	66	PT. ZEF Energi	106	PT Yamaha Indonesia Motor	146	Asia Pulp & Paper
27	PT. Reckat Composites Del	67	PT. Bintang Toedjoe	107	PT Mitra Agro Sembada	147	PT Krakatau Steel Persero
28	PT. Lovary Corpora Indone	68	PT. Mustika Citra Rasa	108	PT MAS ARYA INDONESIA	148	PT Krakatau Steel Persero
29	PT. Kharisma Potensia Ind	69	PT. Shukaku Indonesia	109	PT Cut Make Trim	149	PT Krakatau Steel Persero
30	Bali Interio	70	PT. Romi Violeta	110	PT Yolita Jaya Indonesia	150	PT Philips Industries
31	PT. Sri Indah Labetama	71	Toppan Plasindo Lestari	111	PT Ethos Holding Company	151	PT Pan Brothers Tbk
32	PT Ameya Livingstyle Indo	72	PT. Indah Putih Cemerlang	112	PT Mandala Multifinance,	152	PT Pan Brothers Tbk
33	PT. Panah Perak Megasaran	73	PT. Sahabat Mitra Intrabu	113	PT Mecosuprin Grafia	153	PT Pan Brothers Tbk
34	CV. Tiga Bintang Rezeki	74	PT Indo Non Woven	114	Simatelex Manufactory Co	154	PT Pan Brothers Tbk
35	CV. Megah Sejahtera	75	PT Sariguna Primatirta Tb	115	PT CATUR PUTRA MAPAN		
36	PT U&I Holidays	76	CV Green Aluminium (Jakar	116	PT Panasonic Industrial C		
37	PT. Dewata Telematika	77	PT Murino Bersatu Indones	117	PT Lami Packaging Indones		
38	CV. Megah Sejahtera	78	PT Nojorono Tobacco Inter	118	PT Tera Data Indonusa, Tb		
39	PT. Aneka Bina Citra	79	PT Kompindo Wiratama	119	PT Kirana Megatara, Tbk (
40	Speicher Concepts	80	PT Maxxis International I	120	PT Tenera Inti Sawit		

E-Anggota Cluster 3

Cluster 3		Cluster 3	
No.	Nama Perusahaan	No.	Nama Perusahaan
1	PT. Integrasia Utama	26	PT. Umran Rubi Perkasa
2	PT. Haswara Anjaya Akeh	27	PT. Artha Demo Engineerin
3	PT. Adhiguna Samasta Hars	28	PT. Agung Automall
4	PT. Grasindo Multi Sentos	29	PT. Far East seating
5	PT. Hydro Industrial Auto	30	Siloam Hospitals Group (T
6	PT Semeru Teknik	31	PT EMS Technology
7	PT. Sarana Komunikasi Dat	32	PT Medion Farma Jaya
8	Arka Group	33	PT Ega Mekinka Grup
9	PT. Karya Delitama	34	PT. ENAM GUNUNG EMAS
10	CV. Suwarna Indosemesta	35	P.T. Airindo Sakti
11	Home Detailing	36	PT Green Plus Indonesia
12	PT. Pesona Cipta	37	PT Ikimura Indotools Cent
13	Beautyverse Group	38	Koperasi Jasa Amanah LPPO
14	Nakumi Nakabi	39	PT. Megah Estetika Perkas
15	PT. Korsa Persada Asia	40	PT Global sarana Persada
16	PT. Anugerah Mida Prestas	41	PT Star Cosmos
17	PT. Rajafa Group	42	PT Cimory Dairy Shop
18	Andaf Corp	43	PT Comextra Majora
19	PT. Warna Warni Investama	44	PT Budijaya Makmursentosa
20	PT. Akutahu Media Positif	45	PT Evoluzione Tyres
21	PT. Panca Makmur Baru	46	PT BCA Finance
22	Amplitude Skatepark	47	PT. Asia Inti Prima
23	PT. Sinar Jaya Komunika	48	PT AeroTrans Services Ind
24	The Clinic Beautylosophy	49	PT Bank Mandiri Tbk
25	PT. Anak Panah Perkasa		

F-Script Coding Elbow Method

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.cluster import KMeans
from sklearn.datasets import make_blobs
import seaborn as sns

# Memasukkan data
dataset = 'K-Means_Skripsi.csv'
data = pd.read_csv(dataset)
X,_=make_blobs(n_samples=400, centers=4, cluster_std=2.0, random_state=0)
inertia = []
for k in range(1, 11):
    kmeans = KMeans(n_clusters=k, random_state=42)
    kmeans.fit(X) # Indentasi juga berlaku di sini
    inertia.append(kmeans.inertia_)

# Menampilkan nilai sum square of error
print("Nilai SSE untuk setiap jumlah cluster:")
for i in range(len(sse)):
    print("Jumlah Cluster =", i+1, "| SSE =", sse[i])

# Menampilkan grafik elbow
plt.figure(figsize=(8, 6))
plt.plot(range(1, 11), sse, marker='o')
plt.xlabel('Number of Cluster (k)')
plt.ylabel('Sum of Square Error (SSE)')
plt.title('Metode Elbow untuk Pemilihan Jumlah Cluster')
plt.show()
```

G-Script R Coding Cluster Plot

```
library(tidyverse)
library(factoextra)
library(ggplot2)
library(cluster)
# Memasukkan data
data<-read.csv(file.choose())
# Menampilkan data
data
# Memilih variabel
d<-data.frame(data[2:9])
# Summary
summary(d)
# Menampilkan grafil elbow
fviz_nbclust(d,kmeans,method="wss")
# Standarisasi data
datafix<-scale(d)
View(datafix)
# Clustering
Clustering<-kmeans(datafix,3,nstart=25)
Clustering
# Visualisasi
fviz_cluster(Clustering, geom = "text", data = datafix) + ggtitle("Clustering dengan k=3")
```