

**PENERAPAN ALGORITMA *K-MEANS* UNTUK *CLUSTERING*
DATA PENDUDUK MISKIN PER KABUPATEN/KOTA DI
PULAU SUMATERA**

Diajukan Sebagai Syarat Untuk Menyelesaikan
Pendidikan Program Strata-1 Pada
Jurusan Teknik Informatika



Oleh:

Yulya Anita
NIM: 09021182025001

**Jurusan Teknik Informatika
FAKULTAS ILMU KOMPUTER UNIVERSITAS SRIWIJAYA
2024**

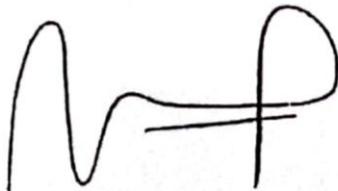
LEMBAR PENGESAHAN SKRIPSI

PENERAPAN ALGORITMA *K-MEANS* UNTUK *CLUSTERING* DATA
PENDUDUK MISKIN PER KABUPATEN/KOTA DI PULAU SUMATERA

Oleh:

Yulya Anita
NIM: 09021182025001

Pembimbing I



Al Farissi, M.Comp.Sc.
NIP. 198512152014041001

Indralaya, Februari 2024
Pembimbing II



Kanda Januar Miraswan, M.T.
NIP. 199001092019031012

Mengetahui,
Dekan Fakultas Ilmu Komputer



Dr. M. Fachrurrozi, S.Si., M.T.
NIP. 198005222008121002

TANDA LULUS UJIAN KOMPREHENSIF

Pada hari Jumat tanggal 02 Februari 2024 Telah dilaksanakan ujian komprehensif skripsi oleh Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.

Nama : Yulya Anita
NIM : 09021182025001
Judul : Penerapan Algoritma *K-Means* untuk *Clustering* Data Penduduk Miskin Per Kabupaten/kota di Pulau Sumatera.

dan dinyatakan LULUS.

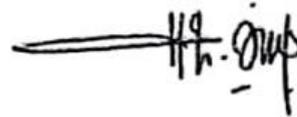
1. Ketua Penguji

Desty Rodiah, M.T.
NIP. 198912212020122011



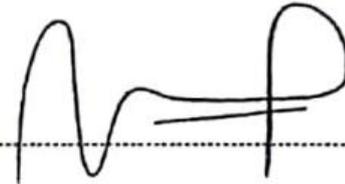
2. Penguji I

Annisa Darmawahyuni, M.Kom.
NIP. 199006302023212044



3. Pembimbing I

Al Farissi, M.Comp.Sc.
NIP. 198512152014041001



4. Pembimbing II

Kanda Januar Miraswan, M.T.
NIP. 199001092019031012



Mengetahui,
Ketua Jurusan Teknik Informatika



Dr. M. Fachrurrozi, S.Si., M.T.
NIP. 198005222008121002

HALAMAN PERNYATAAN

Yang bertanda tangan di bawah ini:

Nama : Yulya Anita
NIM : 09021182025001
Program Studi : Teknik Informatika
Judul Skripsi : Penerapan Algoritma *K-Means* untuk *Clustering* Data
Penduduk Miskin Per Kabupaten/kota di Pulau Sumatera.

Hasil pengecekan *Software iThenticate/Turnitin*: 7%

Menyatakan bahwa laporan proyek saya merupakan hasil karya sendiri dan bukan hasil penjiplakan/plagiat/ Apabila ditemukan unsur penjiplakan/plagiat dalam laporan ini, maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya sesuai dengan ketentuan yang berlaku.

Demikian pernyataan ini saya buat dengan sebenarnya dan tidak ada paksaan dari pihak mana pun.



Indralaya, Februari 2024



Yulya Anita
NIM. 09021182025001

MOTTO DAN PERSEMBAHAN

Tidak ada kesuksesan tanpa kerja keras, tidak ada keberhasilan tanpa kebersamaan, tidak ada kemudahan tanpa do'a, usaha tidak akan mengkhianati hasil, dan jika pada waktu kita khawatir, kita terkadang lebih percaya pada masalah kita dari pada janji Allah SWT. serta Do'a Orang Tua.

“Allah SWT. tidak akan membebani seorang hamba melainkan sesuai dengan kemampuannya.”

(QS. Al-Baqarah: 286)

Kupersembahkan karya tulis ini kepada:

- Allah SWT
- Kedua Orang Tua
- Diriku Sendiri
- Support System
- Keluarga Besar
- Sahabat dan Teman Seperjuangan
- Fakultas Ilmu Komputer
- Universitas Sriwijaya

ABSTRACT

APPLICATION OF K-MEANS ALGORITHM FOR CLUSTERING POOR POPULATION DATA PER DISTRICT / CITY ON THE ISLAND OF SUMATRA

By:

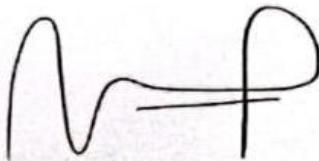
Yulya Anita (09021182025001)

ABSTRACT

Poverty is a problem that exists in every region in Indonesia. There are many factors that cause poverty that influence the level of welfare of the population. Therefore, it is necessary to group data on the poor population of districts/cities to assist in policy decision making. This research uses the K-Means algorithm which is a technique in data mining. Grouping large amounts of data with relatively fast and efficient computing time is the advantage of the K-Means algorithm. This research aims to determine the results of grouping poverty levels from low to highest levels. The results of the grouping will be evaluated using the Davies-Bouldin Index (DBI) value. After testing and evaluating 3 times with 10 groups, it was found that the 10th group was the best result. The comparison of results in 2020 obtained for the district/city category, the lowest was 4 group members and the highest 28 group members. In 2021, the lowest was 4 group members and the highest was 30 group members for the district/city category. In 2022, the lowest was 4 group members for the district/city category and the highest was 29 group members.

Keywords: Poverty, Clustering, K-Means Algorithm, Davies-Bouldin Index.

Supervisor I,



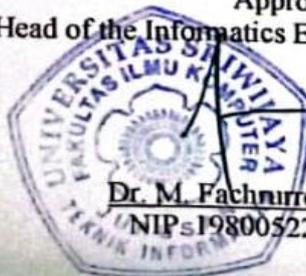
Al Farissi, M.Comp.Sc.
NIP. 198512152014041001

Indralaya, February 2024
Supervisor II,



Kanda Januar Miraswan, M.T.
NIP. 199001092019031012

Approved,
Head of the Informatics Engineering Department,



Dr. M. Fachrurozi, S.Si., M.T.
NIP. 198005222008121002

ABSTRAK

PENERAPAN ALGORITMA *K-MEANS* UNTUK *CLUSTERING* DATA PENDUDUK MISKIN PER KABUPATEN/KOTA DI PULAU SUMATERA

Oleh:

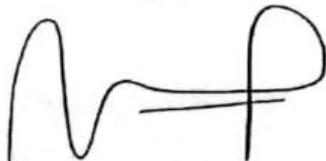
Yulya Anita (09021182025001)

ABSTRAK

Kemiskinan merupakan permasalahan yang ada di setiap wilayah di Indonesia. Banyaknya faktor-faktor penyebab kemiskinan yang mempengaruhi tingkat kesejahteraan penduduk. Oleh karena itu, pengelompokan data penduduk miskin kabupaten/kota perlu dilakukan untuk membantu dalam kebijakan pengambilan keputusan. Penelitian ini menggunakan algoritma *K-Means* yang merupakan salah satu teknik dalam *data mining*. Pengelompokan data dalam jumlah yang cukup besar dengan waktu komputasi yang relatif cepat dan efisien merupakan keunggulan dari algoritma *K-Means*. Penelitian ini bertujuan untuk mengetahui hasil pengelompokan tingkat kemiskinan dari tingkat yang rendah hingga tertinggi. Hasil dari pengelompokan akan dievaluasi menggunakan nilai *Davies-Bouldin Index* (DBI), setelah dilakukan pengujian dan evaluasi sebanyak 3 kali dengan 10 kelompok didapatkan kelompok ke-10 adalah hasil terbaik, perbandingan hasil pada tahun 2020 diperoleh untuk kategori kabupaten/kota yang terendah adalah 4 anggota kelompok dan tertinggi 28 anggota kelompok. Tahun 2021 diperoleh untuk kategori kabupaten/kota yang terendah 4 anggota kelompok dan tertinggi 30 anggota kelompok. Tahun 2022 diperoleh untuk kategori kabupaten/kota terendah 4 anggota kelompok dan tertinggi 29 anggota kelompok.

Kata Kunci: Kemiskinan, *Clustering*, Algoritma *K-Means*, *Davies-Bouldin Index*.

Pembimbing I,



Al Farissi, M.Comp.Sc.
NIP. 198512152014041001

Indralaya, Februari 2024
Pembimbing II,



Kanda Januar Miraswan, M.T.
NIP. 199001092019031012

Mengetahui,
Dekan Fakultas Ilmu Komputer



Dr. M. Fachrurrozi, S.Si., M.T.
NIP. 19800522008121002

KATA PENGANTAR



Puji syukur kepada Allah SWT karena atas berkat dan rahmat-Nya diberi kesehatan, kelancaran dan dikabulkan-Nya doa yang dipanjatkan, sehingga Penulis dapat menyelesaikan Tugas Akhir ini dengan baik. Tugas akhir ini disusun untuk memenuhi salah satu syarat guna menyelesaikan pendidikan program Strata-1 pada Fakultas Ilmu Komputer Program Studi Teknik Informatika di Universitas Sriwijaya.

Dalam menyelesaikan Tugas Akhir ini, banyak pihak yang telah memberikan bantuan dan dukungan baik secara langsung maupun tidak langsung. Pada kesempatan ini, penulis ingin menyampaikan ucapan terima kasih sebesar-besarnya kepada pihak-pihak yang telah membantu penulis dalam menyelesaikan Tugas Akhir ini, yaitu kepada:

1. Orang tua saya, **Bapak Nuril Absori** dan **Ibu Meliana**, Saudara/I, **Dwi Nurhaliza** dan **Iv Berohim**, serta semua keluarga besar yang memberikan cinta dan kasih sayang dan selalu mendoakan dan mendukung serta selalu ada untuk membantu dan memberikan semangat dalam menyelesaikan penelitian.
2. **Edo Fanjola, S.P.** sebagai Support System, terima kasih telah memberikan dukungan yang mengorbankan tenaga, pikiran serta material hingga bangkit kembali dan akhirnya sampai dititik menyelesaikan penelitian ini.

3. **Prof. DR. Erwin, S.Si., M.Si.**, selaku dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
4. **Dr. M. Fachrurrozi, S.Si., M.T.** selaku Ketua Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.
5. **Bapak Osvari. A., S.Kom., M.T.** selaku Sekretaris Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.
6. **Ibu Alvi Syahrini Utami, M.Kom.** selaku dosen pembimbing akademik, yang telah membimbing, mengarahkan, memberikan motivasi serta nasehat dan informasi dalam perkuliahan.
7. **Ibu Mastura Diana Marieska, M.T.** selaku dosen Pembimbing kerja praktek, yang telah banyak membimbing dan mengarahkan seputar perkuliahan.
8. **Bapak Al Farissi, M.Comp.Sc.** sebagai pembimbing Tugas Akhir I, yang telah membimbing, memberikan arahan pada pengerjaan Tugas Akhir.
9. **Bapak Kanda Januar Miraswan, M.T.** sebagai pembimbing Tugas Akhir II, yang memberi masukan dan bimbingannya pada pengerjaan Tugas Akhir.
10. **Ibu Desty Rodiah, M.T.** selaku Ketua Penguji yang telah memberi saran dan masukan dalam pengerjaan Tugas Akhir.
11. **Ibu Annisa Darmawahyuni, M.Kom.** selaku dosen penguji I yang telah memberi saran dan masukan dalam pengerjaan Tugas Akhir.
12. **Seluruh Dosen Program Studi Teknik Informatika dan seluruh Staff Fakultas Ilmu Komputer Universitas Sriwijaya** yang telah memberikan ilmunya. Serta membantu dalam kelancaran proses administrasi dan akademik selama masa perkuliahan.

13. Sahabat seperjuanganku **Dheanita Puspita Sari** dan **Adinda Septiana**, keluarga rantauku **Mba Siti Rodiah** dan teman-teman dari angkatan 2020, serta teman-teman lainnya yang telah berjuang selama menempuh Pendidikan di Universitas Sriwijaya.
14. Keluarga besar **UKM Beladiri** tekhhusus **Divisi PSHT, HIMA 4L UNSRI, LDF WIFI, dan LDK Nadwah.** yang telah memberikan semangat dan pengalaman dalam kehidupan maupun dunia perorganisasian.

Penulis menyadari dalam penyusunan Tugas Akhir ini, masih terdapat banyak kekurangan disebabkan keterbatasan pengetahuan dan pengalaman, oleh karena itu kritik dan saran yang membangun sangat diharapkan untuk kemajuan selanjutnya. Akhir kata semoga Tugas Akhir ini dapat berguna dan bermanfaat bagi kita semua.

Indralaya, Februari 2024
Penulis,

Yulya Anita
NIM. 09021182025001

DAFTAR ISI

	Halaman
HALAMAN JUDUL	1
LEMBAR PENGESAHAN SKRIPSI	ii
TANDA LULUS UJIAN KOMPREHENSIF	iii
HALAMAN PERNYATAAN	iv
MOTTO DAN PERSEMBAHAN	v
ABSTRACT	vi
ABSTRAK	vii
KATA PENGANTAR	viii
DAFTAR ISI	xi
DAFTAR TABEL	xv
DAFTAR GAMBAR	xvii
BAB I PENDAHULUAN	I-1
1.1 Pendahuluan	I-1
1.2 Latar Belakang	I-1
1.3 Rumusan Masalah	I-3
1.4 Tujuan Penelitian	I-4
1.5 Manfaat Penelitian	I-4
1.6 Batasan Masalah	I-5
1.7 Sistematika Penulisan	I-5
1.8 Kesimpulan	I-7
BAB II KAJIAN LITERATUR	II-1
2.1 Pendahuluan	II-1
2.2 Landasan Teori.....	II-1
2.2.1 <i>Data Mining</i>	II-1
2.2.2 <i>Fungsi Data Mining</i>	II-3
2.2.3 <i>Teknik Data Mining</i>	II-5
2.2.4 <i>Proses Data Mining</i>	II-6
2.2.5 <i>Clustering</i>	II-7

2.2.6 <i>K-Means</i>	II-9
2.2.6.1 Pengertian <i>K-Means</i>	II-9
2.2.6.2 <i>Davies-bouldin Index (DBI)</i>	II-11
2.2.6.3 Kekurangan dan Manfaat <i>K-Means</i>	II-13
2.2.7 <i>Rational Unified Process (RUP)</i>	II-14
2.3 Penelitian lain yang Relevan	II-16
2.3.1 Penerapan Metode <i>K-Means</i> Pemetaan Calon Penerima JAMKESDA II-16	
2.3.2 Pemetaan Kemiskinan Kabupaten/Kota di Provinsi Jawa Tengah Menggunakan Algoritma <i>K-Means</i>	II-17
2.3.3 Penerapan Metode <i>K-Means</i> Pada Data Penduduk miskin Per Kecamatan Kabupaten Blitar	II-17
2.4 Kesimpulan	II-18
BAB III METODELOGI PENELITIAN	III-1
3.1 Pendahuluan	III-1
3.2 Unit Penelitian.....	III-1
3.3 Pengumpulan Data	III-1
3.3.1 Jenis dan Sumber Data	III-1
3.3.2 Metode Pengumpulan Data	III-2
3.4 Tahapan Penelitian.....	III-3
3.4.1 Mengumpulkan Data	III-3
3.4.2 Kerangka Kerja Penelitian.....	III-4
3.4.3 Menentukan Kriteria Pengujian.....	III-5
3.4.4 Menentukan Format Data Pengujian	III-6
3.4.5 Alat Bantu Penelitian.....	III-7
3.4.6 Melakukan Analisis dan Mekanik Kesimpulan Penelitian.....	III-7
3.5 Metode Pengembangan Perangkat Lunak.....	III-8
3.5.1 Fase Insepsi	III-8
3.5.2 Fase Elaborasi.....	III-8
3.5.3 Fase Konstruksi	III-9
3.5.4 Fase Transisi.....	III-9
3.6 Manajemen Proyek Perangkat Lunak.....	III-9
3.7 Kesimpulan	III-12

BAB IV PENGEMBANGAN PERANGKAT LUNAK.....	IV-1
4.1 Pendahuluan.....	IV-1
4.2 Fase Insepsi.....	IV-1
4.2.1 Pemodelan Bisnis.....	IV-2
4.2.2 Kebutuhan Sistem.....	IV-2
4.2.3 Analisis Kebutuhan dan Desain.....	IV-4
4.2.3.1 Analisis Kebutuhan Perangkat Lunak.....	IV-4
4.2.3.2 Analisis Algoritma <i>K-Means Clustering</i>	IV-4
4.2.3.3 Analisis Evaluasi Hasil Pengelompokan.....	IV-5
4.2.3.4 Analisis Evaluasi Hasil Pengelompokkan.....	IV-12
4.2.3.5 Perancangan Perangkat Lunak.....	IV-14
4.3 Fase Elaborasi.....	IV-24
4.3.1 Pemodelan Bisnis.....	IV-24
4.3.2 Perancangan Data.....	IV-24
4.3.3 Diagram.....	IV-25
4.3.3.1 <i>Diagram Sequence</i>	IV-25
4.3.3.2 <i>Class Diagram</i>	IV-28
4.3.3.3 Perancangan Antarmuka.....	IV-29
4.4 Fase Konstruksi.....	IV-31
4.4.1 Kebutuhan Sistem.....	IV-31
4.4.2 Implementasi.....	IV-32
4.4.2.1 Implementasi Class.....	IV-32
4.4.2.2 Implementasi Antarmuka.....	IV-33
4.5 Fase Transisi.....	IV-35
4.5.1 Pemodelan Bisnis.....	IV-35
4.5.2 Rencana Pengujian.....	IV-35
4.5.2.1 Rencana Pengujian Memasukkan Berkas.....	IV-35
4.5.2.2 Rencana Pengujian <i>Clustering Kabupaten/Kota K-Means</i>	IV-36
4.5.2.3 Rencana Pengujian Evaluasi Nilai DBI.....	IV-37
4.5.3 Implementasi.....	IV-37
4.5.3.1 Pengujian Memasukkan Berkas.....	IV-38
4.5.3.2 Pengujian <i>Clustering Kabupaten/Kota K-Means</i>	IV-39

4.5.3.3 Pengujian Evaluasi Nilai DBI.....	IV-40
4.6 Kesimpulan	IV-41
BAB V	V-1
HASIL DAN ANALISIS PENELITIAN	V-1
5.1 Pendahuluan	V-1
5.2 Data Hasil Percobaan/Penelitian	V-1
5.2.1 Konfigurasi Percobaan	V-1
5.2.2 Data Hasil Konfigurasi	V-2
5.2.3 Hasil Pengujian.....	V-3
5.3 Analisis Hasil Penelitian	V-6
5.3.1 Hasil Pengujian <i>K-Means Clustering</i>	V-7
5.4 Kesimpulan	V-10
BAB VI KESIMPULAN DAN SARAN	VI-1
6.1 Kesimpulan	VI-1
6.2 Saran.....	VI-2
DAFTAR PUSTAKA.....	xlvii
DAFTAR LAMPIRAN.....	xlvii

DAFTAR TABEL

	Halaman
Tabel III-1. Format Data Pengujian.....	III-6
Tabel III-2. Format Hasil Pengujian <i>Clustering</i> dan penentuan nilai DBI.....	III-7
Tabel III-3. Tabel Work Breakdown Structure (WBS) Penelitian.....	III-10
Tabel IV-1. Kebutuhan Fungsional.....	IV-3
Tabel IV-2. Kebutuhan Non-Fungsional.....	IV-3
Tabel IV-3. Data Penduduk Miskin Kabupaten/kota.....	IV-6
Tabel IV-4. Nilai <i>Centroit</i> Awal.....	IV-7
Tabel IV-5. Jarak Data Terhadap <i>Centroit</i>	IV-8
Tabel IV-6. Daftar Kelompok <i>Cluster</i>	IV-9
Tabel IV-7. Nilai <i>Centroid</i> Baru.....	IV-9
Tabel IV-8. Daftar <i>Centroit</i> Akhir.....	IV-10
Tabel IV-9. Daftar Kelompok <i>Cluster</i> Akhir.....	IV-11
Tabel IV-10. Nilai SSW setiap <i>Cluster</i>	IV-12
Tabel IV-11. Nilai SSB.....	IV-13
Tabel IV-12. Hasil DBI.....	IV-13
Tabel IV-13. Definisi Aktor <i>Use Case</i>	IV-16
Tabel IV-14. Definisi <i>Use Case</i>	IV-16
Tabel IV-15. Skenario <i>Use Case</i> Memasukkan Berkas.....	IV-17
Tabel IV-16. Skenario <i>Use case Clustering</i> Data Algoritma <i>K-Means Clustering</i>	IV-19
Tabel IV-17. Skenario Use case Menghitung Nilai DBI.....	IV-20
Tabel IV-18. Implementasi <i>Diagram Class</i>	IV-32
Tabel IV-19. Rencana Pengujian Memasukkan Berkas.....	IV-36
Tabel IV-20. Rencana Pengujian <i>Clustering</i> Data Menggunakan <i>K-Means</i>	IV-36
Tabel IV-21. Rencana Pengujian Menentukan Nilai DBI.....	IV-37
Tabel IV-22. Pengujian Memasukkan Berkas.....	IV-38

Tabel IV-23. Pengujian <i>Use case</i> Melakukan Pengelompokan <i>K-Means Clustering</i>	IV-39
Tabel IV-24. Pengujian <i>Use case</i> Menentukan Nilai DBI.....	IV-40
Tabel V-1. Pengujian Nilai Evaluasi 1	V-3
Tabel V-2. Pengujian Nilai Evaluasi 2.....	V-4
Tabel V-3. Pengujian Nilai Evaluasi 3	V-5
Tabel V-4. Tabel Hasil <i>Clustering</i>	V-7
Tabel V-5. Hasil Nilai Rata-rata <i>Cluster</i> untuk <i>Cluster</i> 10, Tahun 2020.....	V-8
Tabel V-6. Hasil Nilai Rata-rata <i>Cluster</i> untuk <i>Cluster</i> 10, Tahun 2021.....	V-8
Tabel V-7. Hasil Nilai Rata-rata <i>Cluster</i> untuk <i>Cluster</i> 10, Tahun 2022.....	V-9
Tabel V-8. Tabel Hasil Jumlah <i>Clustering</i>	V-9

DAFTAR GAMBAR

	Halaman
Gambar II-1. Penemuan ilmu baru (Han, <i>et al.</i> , 2012).....	II-3
Gambar II-2. Fungsi <i>Data Mining</i> (Bagas Prayoga H, 2018).....	II-4
Gambar II-3. Identifikasi Kelompok (Wakhidah, 2010).....	II-8
Gambar II-4. Flowchart <i>K-Means Clustering</i> (Wakhida, 2010).....	II-13
Gambar II-5. Arsitektur RUP (IBM, 1998).....	II-15
Gambar III-1. Rincian Kegiatan Penelitian.....	III-3
Gambar III-2. Kerangka Kerja Penelitian.....	III-4
Gambar IV-1. <i>Use Case Diagram</i>	IV-15
Gambar IV-2. Activity Diagram Memasukkan Berkas.....	IV-21
Gambar IV-3. Activity Diagram Melakukan <i>Clustering</i>	IV-22
Gambar IV-4. Activity Diagram Melakukan Evaluasi Nilai DBI.....	IV-23
Gambar IV-5. <i>Sequence Diagram</i> Memasukkan Berkas.....	IV-25
Gambar IV-6. <i>Sequence Diagram K-Means Clustering</i>	IV-26
Gambar IV-7. <i>Sequence Diagram</i> Menghitung Nilai DBI.....	IV-27
Gambar IV-8. <i>Diagram Class</i> Perangkat Lunak.....	IV-28
Gambar IV-9. Rancangan Antarmuka Halaman Awal.....	IV-29
Gambar IV-10. Rancangan Antarmuka Hasil <i>Clustering</i> dan Evaluasi.....	IV-30
Gambar IV-11. Implementasi Antarmuka Halaman Awal.....	IV-33
Gambar IV-12. Implementasi Antarmuka Hasil <i>Clustering</i> dan Evaluasi.....	IV-34

BAB I

PENDAHULUAN

1.1 Pendahuluan

Konsep utama yang terdapat dalam bab ini berperan sebagai pedoman dalam menyusun tugas akhir atau tesis. Gagasan-gagasan utama tersebut mencakup latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, dan batasan masalah. Bab ini akan berisi penjelasan umum mengenai keseluruhan penelitian.

1.2 Latar Belakang

Kemiskinan merupakan permasalahan sosial serius yang beragam dalam tingkat keparahannya di berbagai wilayah di Indonesia. Menurut Badan Pusat Statistik (BPS) Menilai ketidakmampuan seseorang untuk memenuhi kebutuhan dasar, baik berupa makanan maupun kebutuhan non-makanan, melalui aspek ekonomi menjadi salah satu metode untuk mengukur tingkat kemiskinan. Menurut (Darsyah dan Wasono, 2013) Beberapa faktor yang ikut berperan dalam menyebabkan kemiskinan mencakup frekuensi bencana alam, ketidaksetaraan dalam perekonomian, kebijakan pemerintah yang kurang efektif, dan keterbatasan dalam akses terhadap peluang pekerjaan, layanan kesehatan, pendidikan, serta pembangunan infrastruktur.

Menurut (Kompas) Kepulauan Sumatera (bumi melayu) merupakan salah satu pulau terbesar di Indonesia dan pulau keenam terbesar di dunia yang memiliki luas wilayah sekitar 473.481 km², 10 provinsi yang ada yaitu Aceh, Sumatera Utara,

Sumatera Barat, Riau, Kepulauan Riau, Jambi, Bengkulu, Sumatera Selatan, Kepulauan Bangka Belitung dan Lampung. Memiliki 154 kabupaten/kota dengan penduduk sekitar 58.721.272 jiwa dari tahun 2020-2021 yang ada di Pulau Sumatera. Sehubungan dengan pengumpulan data statistik, terutama terkait penduduk miskin di Indonesia, khususnya di Pulau Sumatera, Badan Pusat Statistik (BPS) dianggap sebagai sumber informasi yang andal dan komprehensif, Selain itu lebih mudah dalam mengakses data pada Pulau Sumatera. Berdasarkan data yang dihimpun oleh Badan Pusat Statistik tahun 2020, proporsi penduduk miskin per provinsi di Pulau Sumatera, yang terdiri dari 10 provinsi, mencapai 17,5%.

Salah satu teknik *data mining* yang menggunakan sistem partisi untuk melakukan pemodelan tanpa supervise (*unsupervised*) dan pengelompokan data adalah *K-Means clustering* (Yunita, 2018). Dari sudut pandang ilmiah, *Clustering* adalah cara untuk mengelompokkan objek berdasarkan kriteria tertentu tergantung pada seberapa mirip sifat-sifat mereka. Barang dengan sifat yang hampir identik akan dikelompokkan menjadi satu kelompok. Sifat dari artefak bervariasi secara signifikan di antara kelompok-kelompok tersebut. Dengan kata lain, terdapat sedikit variasi di dalam sebuah kelompok dan banyak perbedaan antara kelompok-kelompok. Oleh karena itu, *K-Means*, sebuah teknik pengelompokan *non-hierarkis*, digunakan (Yustanti *et al.*, 2020).

Penelitian menunjukkan bahwa metode *K-Means clustering* unggul dibandingkan metode *K-medoids* (Harahap, 2021), Pendekatan *K-Means* menunjukkan kinerja yang lebih baik daripada pendekatan *K-medoids*, dengan menggunakan nilai DBI sebagai ukuran kinerja. Nilai DBI yang dihasilkan oleh

pendekatan *clustering K-Means* adalah 0,16, sedangkan *K-medoids* menghasilkan nilai 0,281. Hal ini menegaskan temuan dari penelitian lain (Batra, 2011) yang menyatakan bahwa teknik *K-Means* berkinerja baik pada dataset yang lebih kecil.

K-Means clustering merupakan algoritma pengelompokan yang terkenal. Berbeda dengan pengelompokan hirarkis, pendekatan ini lebih sederhana dipahami dan memerlukan daya komputasi yang lebih rendah (Dhanachandra *et al.*, 2015). Mengatasi tantangan tersebut, penelitian ini mencoba menggunakan teknik *Clustering* yang dipadukan dengan algoritma *K-Means* sebagai alat bantu dalam pemecahan masalah. Dengan demikian, judul yang diberikan oleh penulis untuk penelitian ini adalah "Penerapan Algoritma *K-Means* untuk *Clustering* Data Penduduk Miskin Per Kabupaten/Kota di Pulau Sumatera."

1.3 Rumusan Masalah

Berdasarkan permasalahan yang telah dijelaskan pada latar belakang di atas, maka rumusan masalah pada penelitian ini adalah:

1. Bagaimana menerapkan algoritma *K-Means* untuk mengelompokkan data penduduk miskin per kabupaten/kota di Pulau Sumatera?
2. Bagaimana hasil pengelompokan data penduduk miskin per kabupaten/kota di Pulau Sumatera menggunakan algoritma *K-Means*?
3. Bagaimana hasil implikasi *Clustering* dalam mengetahui tingkat kemiskinan di Pulau Sumatera?

1.4 Tujuan Penelitian

Tujuan penelitian ini adalah:

1. Untuk menerapkan algoritma *K-Means* dalam melakukan pengelompokan data penduduk miskin per kabupaten/kota di Pulau Sumatera.
2. Untuk menghasilkan kelompok-kelompok data penduduk miskin per kabupaten/kota di Pulau Sumatera yang memiliki karakteristik serupa.
3. Untuk memberikan rekomendasi kebijakan pengambilan keputusan dalam penanggulangan kemiskinan di Pulau Sumatera berdasarkan hasil *Clustering* yang diperoleh.

1.5 Manfaat Penelitian

Manfaat dari penelitian ini mencakup:

1. Mendapatkan pemahaman mengenai klaster-klaster yang terbentuk di Pulau Sumatera sebagai hasil dari keterbatasan data statistik penduduk yang dikelompokkan berdasarkan kabupaten/kota.
2. Mendapatkan evaluasi terhadap sejauh mana kinerja teknik *K-Means* dalam mengelompokkan data penduduk miskin di kabupaten/kota di Pulau Sumatera.
3. Memberikan informasi yang akurat dan terpercaya tentang karakteristik penduduk miskin per kabupaten/kota di Pulau Sumatera, sehingga dapat membantu pemerintah dan lembaga terkait dalam merencanakan dan melaksanakan program penanggulangan kemiskinan yang lebih efektif dan efisien.

1.6 Batasan Masalah

Adapun batasan masalah pada penelitian kali ini sebagai berikut:

1. Penelitian ini hanya memfokuskan pada data penduduk miskin per kabupaten/kota di Pulau Sumatera pada tahun 2020, 2021 dan 2022 yang diperoleh dari Badan Pusat Statistik (BPS).
2. Penelitian ini tidak membahas secara rinci tentang faktor-faktor penyebab terjadinya kemiskinan di masing-masing kabupaten/kota di Pulau Sumatera, tetapi hanya memfokuskan pada pengelompokan data penduduk miskin berdasarkan karakteristik serupa yang ditemukan dalam data tersebut.

1.7 Sistematika Penulisan

Berikut adalah gambaran umum mengenai struktur penulisan yang telah penulis susun untuk tugas akhir, yang akan mengikuti standar penulisan Fakultas Ilmu Komputer Universitas Sriwijaya.

BAB I. PENDAHULUAN

Pada Bab I penulis membahas tentang pokok dari landasan dasar yang meliputi hal-hal seperti landasan masalah, latar belakang, rumusan masalah, tujuan dan manfaat penelitian, batasan masalah, sistematika penelitian, dan metodologi dalam penelitian ini.

BAB II. KAJIAN LITERATUR

Pada Bab II penulis merangkum keseluruhan kerangka teoritis yang konsisten dengan penelitian studi kasus. Ini mencakup dasar-dasar teoritis dari konsep-konsep seperti data mining, *Clustering*, dan algoritma *K-Means*, serta beberapa literatur yang relevan.

BAB III. METODOLOGI PENELITIAN

Pada Bab III mekanisme dan prosedur yang akan digunakan selama penyelidikan dijelaskan secara mendalam, dengan setiap langkah dijelaskan dalam konteks kerangka kerja yang dikembangkan untuk penelitian ini.

BAB IV. PERANCANGAN PERANGKAT LUNAK

Pada Bab IV penulis mengulas rencana dan analisis pengembangan perangkat lunak. Menerapkan teknik *Rational Unified Process* (RUP) sebagai alat penelitian yang digunakan untuk merancang perangkat lunak, menganalisis kebutuhan, dan melakukan pengujian guna memastikan bahwa sistem yang dibuat sesuai dengan persyaratan dan rencana penelitian.

BAB V. HASIL DAN ANALISIS PENELITIAN

Pada Bab V penulis menyajikan hasil pengujian berdasarkan tahapan yang direncanakan. Analisis diberikan sebagai dasar pengambilan kesimpulan dalam penelitian ini.

BAB VI. KESIMPULAN DAN SARAN

Pada Bab VI penulis akan membahas kesimpulan dan saran berdasarkan penelitian yang telah dilakukan, seperti pengertian pada bab sebelumnya.

1.8 Kesimpulan

Dalam Bab ini dijelaskan latar belakang, rumusan masalah, tujuan, manfaat penelitian, batasan masalah, sistematika penelitian dan metodologi penelitian yang akan dijadikan pokok-pokok penelitian bagi peneliti. Dalam penelitian ini, x

DAFTAR PUSTAKA

- Alfiansyah dan Dede., (2022). Penerapan Metode K-Means pada Data Penduduk Miskin Per Kecamatan Kabupaten Blitar. *Jurnal Repositor*. 4. 49-58. 10.22219/repositor.v4i1.1416.
- Alifah, A. (2019). PEMETAAN KEMISKINAN KABUPATEN/KOTA DI PROVINSI JAWA TENGAH MENGGUNAKAN ALGORITMA *K-MEANS*. *EDUSAINTEK*, 3.
- Bagas, P. H., (2018). Penerapan Algoritma *K-Means* Untuk Memetakan Garis Kemiskinan Menurut Provinsi Di Indonesia, *Teknik Informatika Sekolah Tinggi Teknologi Pelita Bangsa Bekasi* <https://ecampus.pelitabangsa.ac.id/pb/AmbilLampiran?ref=22949&jurusan=&jenis=Item&usingId=false&download=false&clazz=ais.database.model.file.LampiranLain>
- Batra, A. (2011). Analysis and Approach: K-Means and K-Medoids Data Mining Algorithms. 5th IEEE International Conference on Advanced, 274, 274–279.
- Badan Pusat Statistik (BPS) Kemiskinan dan Penduduk Miskin <https://sumsel.bps.go.id/subject/23/kemiskinan.html#subjekViewTab1>
- Badan Pusat Statistik (BPS) Kemiskinan <https://bengkulu.bps.go.id/indicator/23/304/1/persentase-penduduk-miskin-menurut-provinsi-di-sumatera.html>

- Darsyah, M. Y., & Wasono, R. (2013). Pendugaan Tingkat Kemiskinan di Kabupaten Sumenep dengan pendekatan SAE. *Prosiding Seminar Nasional & Internasional*, 1(1), 1–10.
- Dhanachandra, N., Manglem, K., & Chanu, Y. J. (2015). Image Segmentation Using K-means Clustering Algorithm and Subtractive Clustering Algorithm. *Procedia Computer Science*, 54, 764–771.
- Hakim, Zaenal, and Robby Rizky. "Analisis Perancangan Sistem Informasi Pembuatan Paspor Di Kantor Imigrasi Bumi Serpong Damai Tangerang Banten Menggunakan Metode Rational Unified Process." vol 6 (2020): 103-112.
- Han, J., Kamber, M., Pei, J. 2012. *Data Mining Concepts and Techniques*. Waltham: Elsevier.
- Harahap, F. (2021). Perbandingan Algoritma K Means dan K Medoids Untuk *Clustering* Kelas Siswa Tunagrahita. *TIN: Terapan Informatika Nusantara*, 2(4), 191–197.
- IBM, 1998. *Rational Unified Process Best Practices for Software Development Teams*. United States: IBM Rational Software.
- Kompas.com <https://www.kompas.com/skola/read/2022/07/25/140000369/10-provinsi-di-pulau-sumatera>
- Merliana, N.P.E. & Santoso, A.J. 2015. Analisa Penentuan Jumlah Cluster Terbaik Pada Metode K-Means Clustering.
- Nanda dan Faisal., (2018) “Penerapan Metode *K-Means* Pemetaan Calon Penerima JAMKESDA” *Jurnal Ilmu Komputer*, 5(2), 190-200.

- Rahmawati, L., Widya Sihwi, S. & Suryani, E. 2016. *Analisa Clustering Menggunakan Metode K-Means dan Hierarchical Clustering (Studi Kasus : Dokumen Skripsi Jurusan Kimia, FMIPA, Universitas Sebelas Maret)*. *Jurnal Teknologi & Informasi ITSmart*, 3: 66.
- Stanton, J. 2013. *Inroduction to Data Science*. Syracuse: Syracuse University's School of Information Studies
- Suyanto. 2017. *Data Mining Untuk Klasifikasi dan Klasterisasi Data*. Bandung: Informatika
- Vulandari, R. T. 2017. *Data Mining Teori dan Aplikasi Rapidminer*. Yogyakarta: Gavamedia
- Wakhidah, N. (2010). *Clustering menggunakan K-Means algorithm*. *Jurnal Transformatika*, 8(1), 33-39.
- Yunita, F. (2018). *Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Pada Penerimaan Mahasiswa Baru*. *Sistemasi*, 7(3), 238.
- Yustanti, W., Rahmawati, N., & Yamasari, Y. (2020). *Klastering Wilayah Kota/Kabupaten Berdasarkan Data Persebaran Covid-19 Di Propinsi Jawa Timur dengan Metode K-Means*. *JIEET (Journal of Information Engineering and Educational Technology)*, 4(1), 1–9.