

**OPTIMASI *K-MEANS CLUSTERING* DENGAN *PARTICLE SWARM OPTIMIZATION* UNTUK PENGELOMPOKAN
DEPRESI BERDASARKAN DASS-42**



SKRIPSI

Diajukan untuk memenuhi salah satu syarat kelulusan Program Sarjana

MOHAMAD ABDUL YUSUF

NIM: 15200056

Program Studi Informatika Kampus Margonda

Fakultas Teknik dan Informatika

Universitas Bina Informatika

2024

PERSEMBAHAN

Maka, sesungguhnya beserta kesulitan ada kemudahan. Sesungguhnya beserta kesulitan ada kemudahan.
(QS Al-Insyirah: 5-6)

Alhamdulillah segala puja, puji, sembah, sujud, syukur hanya untuk Allah SWT. Sholawat serta salam kepada Nabi Muhammad SAW. Karenanya dengan nikmat islam dan iman ini penulis dapat menjalani hidup sejauh ini sehingga dapat menyelesaikan skripsi ini.

Halaman persembahan ini juga ditunjukkan sebagai ungkapan terimakasih kepada:

1. Bapak Rachmat Adi Purnama, M.Kom yang telah membimbing penulis agar menyelesaikan skripsi ini sehingga jauh lebih baik.
2. Kedua orang tua penulis yaitu Naeng dan Timah karena telah penuh kasih dan sayang membesarkan dan mengajari penulis segala yang mereka tau.
3. Kakak penulis yaitu Desy Yanti Umami yang sangat mensupport penulis dalam segala hal.
4. Elena Monika alias Lebar yang selalu memberi support dan masukan kepada penulis.
5. Seluruh teman penulis yang tidak dapat penulis sebutkan namanya satu-persatu namun tidak mengurangi rasa terima kasih penulis barang satu titik saja dari Daftest, Omend, After-Server, BLK JKM PPKPI, dan Informatika UBSI hingga teman-teman baru penulis dari Grup Main, sungguh penulis sangat menyayangi seluruhnya.
6. Secara khusus juga kepada Leo Agustian Winarno, Gilang Apdillah, Dzul Arsi Majid Anwar, Rhomadona Bambang Pamungkas, Rafly Saputra, Gatot Primaulida Saputro, Muhammad Riesky Pratama, Mochamad Ardiansyah, Dimas Rizki Kurniawan, Arya Dimas Wardani, Andri Atma Wijaya, Muhammad Rizqi, Fazar Setiawan, Vivi Andriani, dan Abudzar Al Gifari yang memberikan bantuan dan support kepada penulis dalam bentuk yang berbeda-beda.

Terima kasih banyak.

Sekali lagi,
Terima kasih banyak.

SURAT PERNYATAAN KEASLIAN SKRIPSI PADA PROGRAM SARJANA

SURAT PERNYATAAN KEASLIAN SKRIPSI PADA PROGRAM SARJANA

Yang bertanda tangan dibawah ini:

Nama : Mohamad Abdul Yusuf
NIM : 15200056
Jenjang : Sarjana (S1)
Program Studi : Informatika
Fakultas : Teknik dan Informatika
Perguruan Tinggi : Universitas Bina Sarana Informatika

Dengan ini menyatakan bahwa Tugas Akhir yang telah saya buat dengan judul: "**Optimasi K-Means Clustering Dengan Particle Swarm Optimization Untuk Pengelompokan Depresi Berdasarkan DASS-42**", adalah asli (orsinil) atau tidak plagiat (menjiplak) dan belum pernah diterbitkan/dipublikasikan dimanapun dan dalam bentuk apapun.

Demikianlah surat pernyataan ini saya buat dengan sebenar-benarnya tanpa ada paksaan dari pihak manapun juga. Apabila di kemudian hari ternyata saya memberikan keterangan palsu dan atau ada pihak lain yang mengklaim bahwa Skripsi yang telah saya buat adalah hasil karya milik seseorang atau badan tertentu, saya bersedia diproses baik secara pidana maupun perdata dan kelulusan saya dari **Universitas Bina Sarana Informatika** dicabut/dibatalkan.

Dibuat di : Jakarta
Pada tanggal: 28 Juni 2024
Yang Menyatakan,



Mohamad Abdul Yusuf

SURAT PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

SURAT PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Yang bertanda tangan di bawah ini, saya:

Nama : Mohamad Abdul Yusuf
NIM : 15200056
Jenjang Studi : Sarjana (S1)
Program Studi : Informatika
Fakultas : Teknik dan Informatika
Perguruan Tinggi : Universitas Bina Sarana Informatika

Dengan ini menyatakan bahwa seluruh data, informasi, interpretasi serta pernyataan yang terdapat dalam karya ilmiah Penulis dengan judul "**Optimasi K-Means Clustering Dengan Particle Swarm Optimization Untuk Pengelompokan Depresi Berdasarkan DASS-42**" ini, kecuali yang disebutkan sumbernya adalah hasil pengamatan, penelitian, pengelolaan, serta pemikiran saya.

Penulis menyetujui untuk memberikan ijin kepada pihak Universitas Bina Sarana Informatika untuk mendokumentasikan karya ilmiah saya tersebut secara internal dan terbatas, serta tidak untuk mengunggah karya ilmiah Penulis pada repository Universitas Bina Sarana Informatika.

Penulis bersedia untuk bertanggung jawab secara pribadi, tanpa melibatkan pihak **Universitas Bina Sarana Informatika**, atas materi/isi karya ilmiah tersebut, termasuk bertanggung jawab atas dampak atau kerugian yang timbul dalam bentuk akibat tindakan yang berkaitan dengan data, informasi, interpretasi serta pernyataan yang terdapat pada karya ilmiah saya ini.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta

Pada tanggal : 28 Juni 2024

Yang menyatakan,



Mohamad Abdul Yusuf

PERSETUJUAN DAN PENGESAHAN SKRIPSI

PERSETUJUAN DAN PENGESAHAN SKRIPSI

Skripsi ini diajukan oleh:

Nama : Mohamad Abdul Yusuf
NIM : 15200056
Jenjang : Sarjana (S1)
Program Studi : Informatika
Fakultas : Teknik dan Informatika
Perguruan Tinggi : Universitas Bina Sarana Informatika
Judul Skripsi : Optimasi K-Means Clustering Dengan Particle Swarm Optimization Untuk Pengelompokan Depresi Berdasarkan DASS-42

Telah dipertahankan pada periode 2024-1 dihadapan penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh Sarjana Komputer (S.Kom) pada Program Sarjana (S1) Program Studi Informatika di Universitas Bina Sarana Informatika.

Jakarta, 19 Agustus 2024

PEMBIMBING SKRIPSI

Pembimbing I : Rachmat Adi Purnama, M.Kom.

DEWAN PENGUJI

Penguji I : Budi Sudrajat, M.Kom.

Penguji II : Rian Septian Anwar, M.Kom.

PEDOMAN PENGGUNAAN HAK CIPTA

Skripsi sarjana yang berjudul “Optimasi *K-Means Clustering* Dengan *Particle Swarm Optimization* Untuk Pengelompokan Depresi Berdasarkan Dass-42” adalah hasil karya tulis asli Rio Septian dan bukan hasil terbitan sehingga peredaran karya tulis hanya berlaku dilingkungan akademik saja, serta memiliki hak cipta. Oleh karena itu, dilarang keras untuk menggandakan baik sebagian maupun seluruhnya karya tulis ini, tanpa seizin penulis.

Referensi kepustakaan diperkenankan untuk dicatat tetapi pengutipan atau peringkasan isi tulisan hanya dapat dilakukan dengan seizin penulis dan disertai ketentuan pengutipan secara ilmiah dengan menyebutkan sumbernya.

Untuk keperluan perizinan pada pemilik dapat menghubungi informasi yang tertera di bawah ini:

Nama : Mohamad Abdul Yusuf
Alamat : Jl. Praji / Rt. 05/001 Jakarta Timur
No. Telp : 085780499281
E-mail : yusufppkpi@gmail.com



UNIVERSITAS

KATA PENGANTAR

Alhamdulillah, penulis sangat bahagia bisa menyusun penulisan skripsi ini dengan baik. Puji serta syukur penulis limpahkan hanya untuk Allah SWT, karena telah menggerakkan diri penulis selama merampungkan skripsi ini sebagaimana skripsi ini dibutuhkan sebagai syarat kelulusan penulis. Juga, sholawat dan salam tak lupa penulis haturkan kepada Nabi Besar Muhammad SAW dan keluarganya karena telah membawa perubahan yang paling baik kepada dunia ini. Adapun judul skripsi yang penulis selesaikan ialah “Optimasi K-Means Clustering Dengan Particle Swarm Optimization Untuk Pengelompokan Depresi Berdasarkan Dass-42”.

Seperti yang sudah penulis singgung sebelumnya, maksud dari penulisan Skripsi ini utamanya adalah sebagai syarat kelulusan penulis dalam menyelesaikan Program Sarjana (S1) di Universitas Bina Sarana Informatika. Sedangkan lebih dalam skripsi ini penulis gunakan sebagai media penyaluran penelitian yang ingin penulis lakukan secara pribadi. Dan mengenai sumber penelitian, penulisan dalam skripsi ini disusun dari banyak sumber seperti hasil pengumpulan data melalui Kuesioner, dan beberapa studi ilmiah yang selaras dengan penelitian penulis. Penulis juga mengetahui tanpa bantuan dan arahan dari berbagai pihak penulisan tugas akhir ini tidak akan pernah selesai, maka penulis ingin mengucapkan terima kasih kepada:

1. Rektor Universitas Bina Sarana Informatika.
2. Dekan Fakultas Teknik Dan Informatika.
3. Ketua Program Studi Informatika di Universitas Bina Sarana Informatika.
4. Bapak Rachmat Adi Purnama M.Kom selaku dosen pembimbing.
5. Orang Tua serta Kakak penulis yang selalu melayangkankan support moril dan materil tanpa batas.

6. Seluruh teman-teman penulis yang selalu menemani penulis kala suka dan duka.
7. Staff atau karyawan Universitas Bina Sarana Informatika.

Sampai pada waktu penyusunan penulis menyadari bahwa dalam penyusunan skripsi ini masih terdapat banyak sekali kesalahan, maka sebab itu peneliti meminta agar diberikan masukan juga saran yang mengapikan untuk kesempuenaan penulis di masa yang akan datang. Penulis juga memohon maaf karena penelitian ini masih jauh dari kata sempurna karena terbatasnya ilmu dan pengalaman penulis. Namun, penulis masih berharap jika hasil penelitian ini dapat membawa manfaat dalam meningkatkan keilmuan, Masyarakat dan Bangsa. Semoga Allah SAW meridhoi penelitian penulis ini. Aamiin.

Jakarta, 28 Juni 2024
Penulis



Mohamad Abdul Yusuf

ABSTRAKSI

Bagi manusia depresi adalah hal yang sangat lekat dengan kehidupan. Walaupun depresi membawa banyak dampak negatif bagi manusia, pada kenyataannya masih banyak di antara kita yang menyepelekan depresi ini. Dampak dari depresi juga diketahui dapat mengganggu kinerja seseorang, mengganggu kemampuan bersosialisasi, mempengaruhi kesehatan fisik, hingga menjadi perantara bunuh diri. Penelitian ini diharapkan dapat meningkatkan kesadaran diri terhadap sekitar mengenai masalah depresi di lingkungan masyarakat atau paling tidak di lingkungan orang-orang terdekat kita. Data dari penelitian ini berupa data kualitatif yang didapat dari hasil observasi dengan pengamatan secara langsung dan dengan cara pengumpulan data menggunakan angket kuis berupa *questionnaire Google Form* yang berupa DASS-42. DASS-42 sendiri adalah seperangkat skala subyektif yang dibentuk untuk mengukur status emosional negatif dari depresi, kecemasan, dan stres. Namun, didalam penelitian ini nilai difokuskan hanya kepada nilai depresinya saja. Selanjutnya untuk melihat hasil berupa tingkat depresi, data dimasukkan ke dalam model algoritma *K-Means Clustering* yang dioptimasi dengan *Particle Swarm Optimization* untuk membuktikan sedikit apa sebenarnya kita dengan depresi ini. Penelitian ini mendapatkan hasil bahwa *K-Means* yang dioptimasi menggunakan *Particle Swarm Optimization* mampu meningkatkan kinerja *clustering*.

Kata Kunci: *K-Means, Particle Swarm Optimization, Machine Learning.*

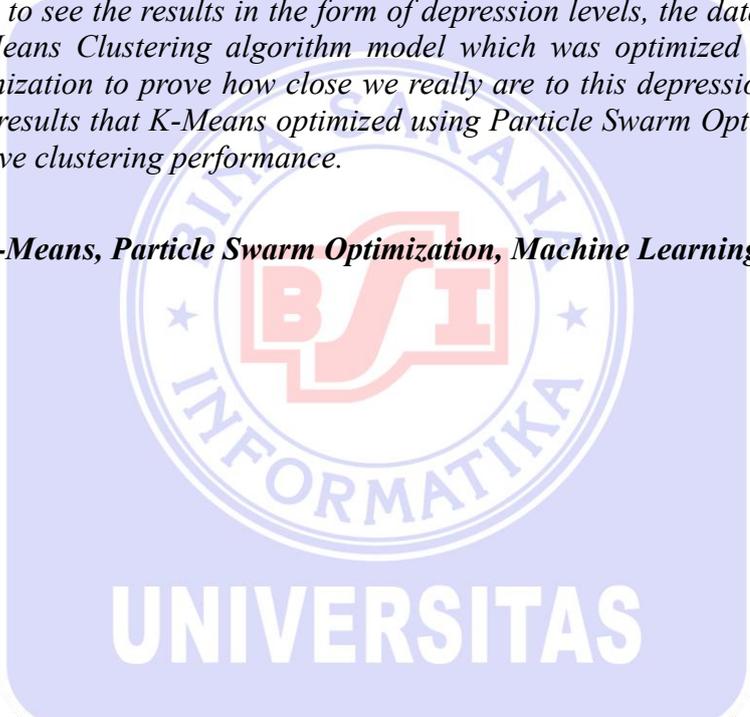


UNIVERSITAS

ABSTRACT

For humans, depression is something that is very closely related to life. Although depression has many negative impacts on humans, in fact there are still many of us who underestimate this depression. The impact of depression is also known to interfere with a person's performance, interfere with social skills, affect physical health, and become a suicide broker. This research is expected to increase self-awareness of the surroundings regarding depression problems in the community or at least in the environment of the people closest to us. The data from this study is in the form of qualitative data obtained from observation results with direct observation and by collecting data using a quiz questionnaire in the form of a Google Form questionnaire in the form of DASS-42. The DASS-42 itself is a set of subjective scales formed to measure the negative emotional status of depression, anxiety, and stress. However, in this study, the value is focused only on the value of depression. Furthermore, to see the results in the form of depression levels, the data was entered into the K-Means Clustering algorithm model which was optimized with Particle Swarm Optimization to prove how close we really are to this depression. This study obtained the results that K-Means optimized using Particle Swarm Optimization was able to improve clustering performance.

Keywords: K-Means, Particle Swarm Optimization, Machine Learning.

The logo of Universitas BSI Informatika is a circular emblem. It features a red stylized 'BSI' monogram in the center, surrounded by a white ring containing the text 'UNIVERSITAS BSI INFORMATIKA' in blue capital letters. Two small blue stars are positioned on the left and right sides of the ring. Below the emblem, the word 'UNIVERSITAS' is written in large, bold, white capital letters on a light blue rectangular background.

UNIVERSITAS

DAFTAR ISI

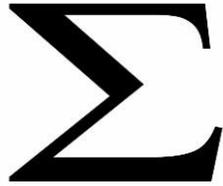
PERSEMBAHAN.....	ii
SURAT PERNYATAAN KEASLIAN SKRIPSI PADA PROGRAM SARJANA	iii
SURAT PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS.....	iv
PERSETUJUAN DAN PENGESAHAN SKRIPSI.....	v
PEDOMAN PENGGUNAAN HAK CIPTA	vi
KATA PENGANTAR.....	vii
ABSTRAKSI	ix
<i>ABSTRACT</i>	x
DAFTAR ISI.....	xi
DAFTAR SIMBOL	xiv
DAFTAR GAMBAR	xv
DAFTAR TABEL.....	xvi
DAFTAR LAMPIRAN	xvii
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang Masalah.....	1
1.2. Rumusan Masalah	2
1.3. Tujuan dan Manfaat.....	2
1.4. Hipotesis	4
1.5. Batasan Masalah.....	5
BAB II LANDASAN TEORI.....	6
2.1. Tinjauan Pustaka	6
2.1.1. Depresi.....	6
2.1.2. DASS-42.....	6
2.1.3. Algoritma.....	7
2.1.4. K-Means Clustering.....	7
2.1.5. <i>Particle Swarm Optimization</i>	7
2.1.6. <i>Silhouette Score</i>	8
2.1.7. <i>Calinski-Harabasz Score</i>	8
2.1.8. <i>Davies-Bouldin Score</i>	8
2.2. Penelitian Terkait	8
BAB III METODOLOGI PENELITIAN.....	13

3.1. Proses dan Langkah Penelitian.....	13
3.1.1. Pengumpulan Data.....	14
3.1.2. Pembersihan Dataset Secara Manual.....	14
3.1.3. Pengkodean Data Secara Manual	15
3.1.4. Memuat Dataset	15
3.1.5. Membuat Data Training.....	16
3.1.6. Penskalaan Fitur.....	16
3.1.7. <i>Clustering</i>	16
3.1.8. <i>Clustering</i> Menggunakan K-Terbaik	17
3.1.9. Membuat <i>Dataset</i> Baru.....	17
3.1.10. Optimasi Dengan <i>Particle Swarm Optimization</i>	18
3.2. Metode Pengolahan dan Analisis Data.....	18
3.2.1. Metode Pengolahan Data.....	18
3.2.2. Metode Analisis Data	19
BAB IV HASIL DAN PEMBAHASAN.....	22
4.1. Hasil Penelitian.....	22
4.1.1. Pengumpulan Data.....	22
4.1.2. Pembersihan Dataset Secara Manual.....	22
4.1.3. Pengkodean Data Secara Manual	23
4.1.4. Memuat Dataset	23
4.1.5. Membuat <i>Data Training</i>	24
4.1.6. Penskalaan Fitur.....	25
4.1.7. <i>Clustering</i>	26
4.1.8. Mencari K Terbaik menggunakan <i>Elbow Method</i>	29
4.1.9. Membuat <i>Dataset</i> Baru Dengan Label HSL2.....	30
4.1.10. Optimasi Dengan PSO	31
4.2. Hasil Pengujian.....	32
4.2.1. Dengan Jumlah <i>Cluster</i> Acak	32
4.2.2. Penggunaan K Terbaik.....	36
4.2.3. Optimasi K-Means <i>Clustering</i> Menggunakan <i>Particle Swarm Optimization</i>	39
4.2.4. Perbandingan Nilai-Nilai Skor Validasi <i>Clustering</i>	44
BAB V PENUTUP.....	47
5.1. Kesimpulan.....	47

5.2. Saran.....	49
DAFTAR PUSTAKA.....	50
DAFTAR RIWAYAT HIDUP.....	52
LEMBAR KONSULTAN.....	53
SURAT PERNYATAN KEBENARAN/KEABSAHAN DATA HASIL RISET UNTUK KARYA ILMIAH.....	54
LAMPIRAN-LAMPIRAN.....	55



DAFTAR SIMBOL



SIGMA

Simbol ini menunjukkan notasi serangkaian perhitungan

PROCESSING SYMBOL



Simbol ini digunakan untuk mendefinisikan Alur kerja

FLOWLINE



Simbol ini menunjukkan hubungan proses satu dengan yang lainnya



UNIVERSITAS

DAFTAR GAMBAR

	Halaman
Gambar III. 1 Tahapan Penelitian	13
Gambar III. 2 Persamaan K-Means.....	20
Gambar IV. 1 Visualisasi Data	24
Gambar IV. 2 Kode Untuk Membuat Data Training	24
Gambar IV. 3 Sample Data Training	25
Gambar IV. 4 Code Untuk Feature Scaling Dengan MinMaxScaler	25
Gambar IV. 5 Sample Data Training Setelah Feature Scaling	26
Gambar IV. 6 Code Untuk Menjalankan K-Means	26
Gambar IV. 7 K-Means Dengan 2 Titik 2 Centroid	27
Gambar IV. 8 Crosstab Antara KODE JK dan KODE HSL	28
Gambar IV. 9 Visualisasi Centroid 2 Cluster	28
Gambar IV. 10 Grafik Elbow Method.....	29
Gambar IV. 11 K-Means Dengan 4 Titik Centroid	29
Gambar IV. 12 Visualisasi K-Means 4 Cluster	30
Gambar IV. 13 Crosstab Antara KODE JK dan KODE HSL2 pada 4 Cluster	30
Gambar IV. 14 Crosstab Antara KODE JK dan KODE HSL3 pada 5 Cluster	32
Gambar IV. 15 Koordinat 2 Titik Centroid	33
Gambar IV. 16 Nilai Inertia Pada Tiap Cluster	33
Gambar IV. 17 Skor Validasi 2 Cluster	34
Gambar IV. 18 Crosstab KODE HSL dan KODE JK	35
Gambar IV. 19 Crosstab Antara KODE HSL2 dan KODE JK	37
Gambar IV. 20 Skor Validasi Cluster Dengan K Terbaik.....	38
Gambar IV. 21 Visualisasi Elbow Method	40
Gambar IV. 22 Koordinat 5 Titik Centroid	41
Gambar IV. 23 Cluster Data dan 5 Centroid	42
Gambar IV. 24 Crosstab Antara KODE JK dan KODE HSL3	43

DAFTAR TABEL

	Halaman
Tabel IV. 1 Sampel Data.....	22
Tabel IV. 2 Sampel Data Yang telah Dibersihkan	22
Tabel IV. 3 Sample Data Setelah Dikodekan	23



DAFTAR LAMPIRAN

Lampiran A. 1 Kuesioner Depression Anxiety Stress Scales (DASS 42).....	55
Lampiran A. 2 Kuesioner Depression Anxiety Stress Scales (DASS 42).....	56
Lampiran A. 3 Kuesioner Depression Anxiety Stress Scales (DASS 42).....	57
Lampiran B. 1 Data Tingkat Depresi	58
Lampiran B. 2 Data Tingkat Depresi	59
Lampiran B. 3 Data Tingkat Depresi	60
Lampiran C. 1 Kode Program K-Means dan Particle Swarm Optimization.....	61
Lampiran C. 2 Kode Program K-Means dan Particle Swarm Optimization.....	62
Lampiran C. 3 Kode Program K-Means dan Particle Swarm Optimization.....	63
Lampiran C. 4 Kode Program K-Means dan Particle Swarm Optimization.....	64
Lampiran C. 5 Kode Program K-Means dan Particle Swarm Optimization.....	65
Lampiran C. 6 Kode Program K-Means dan Particle Swarm Optimization.....	66
Lampiran C. 7 Kode Program K-Means dan Particle Swarm Optimization.....	67
Lampiran C. 8 Kode Program K-Means dan Particle Swarm Optimization.....	68
Lampiran C. 9 Kode Program K-Means dan Particle Swarm Optimization.....	69
Lampiran C. 10 Kode Program K-Means dan Particle Swarm Optimization.....	70
Lampiran C. 11 Kode Program K-Means dan Particle Swarm Optimization.....	71
Lampiran C. 12 Kode Program K-Means dan Particle Swarm Optimization.....	72
Lampiran C. 13 Kode Program K-Means dan Particle Swarm Optimization.....	73
Lampiran C. 14 Kode Program K-Means dan Particle Swarm Optimization.....	74
Lampiran C. 15 Kode Program K-Means dan Particle Swarm Optimization.....	75
Lampiran D. 1 Cluster Hasil Tingkat Depresi.....	76

INFORMATIKA
UNIVERSITAS

BAB I

PENDAHULUAN

1.1. Latar Belakang Masalah

Pesatnya perkembangan teknologi di era ini membuat kita juga harus cepat belajar dan menyesuaikan diri dalam mengikuti perkembangan sebagai pengimbang agar tidak tertinggal oleh teknologi belakangan ini. Karena, setiap perkembangannya biasanya selalu diakhiri dengan digitalisasi. Digitalisasi teknologi belakangan ini memang berkembang sangat cepat. Bahkan, pada masa sekarang ini penerapan digitalisasi teknologi yang merujuk pada kecerdasan buatan (AI) juga *Internet of Things* (IoT) sangat mudah kita temui di sekitar kita. Padahal jika kita sedikit melihat 30 tahun ke belakang untuk berkomunikasi saja kita masih menggunakan telepon koin dan untuk berkomunikasi jarak jauh.

Dalam setiap dinamika perubahan pasti akan membawa dampak baik dan juga dampak buruk bagi kehidupan manusia. Terlebih lagi, pada arus dinamika kemajuan yang sangat pesat di bidang ilmu Teknologi dan Informasi yang seakan tidak ada habisnya juga tidak ada batasnya. Contoh kasus menarik dari perkembangan teknologi ialah aktivitas sosial yang terbentuk dari munculnya media sosial menjadi tidak terbendung, terlebih didalamnya dampak baik dan buruk terjadi secara bersamaan. Yang tentu saja pasti menimbulkan masalah yang besar bagi kehidupan sosial manusia. “Seperti depresi yang merebak belakangan ini juga diyakini merupakan sebuah masalah baru buah dari perkembangan teknologi. Hal buruk yang difasilitasi teknologi dapat menimbulkan berbagai dampak buruk pula salah satunya ialah depresi”(Aqsa Mu’azzaroh & Cahyanti, 2023).

Depresi menurut *Division* adalah merupakan kondisi emosional yang biasanya dapat ditandai dengan adanya kesedihan atau perasaan yang tidak berarti dan memiliki rasa bersalah, menjauhkan diri dari lingkungan, kehilangan selera makan, tidak dapat tidur dengan tenang ataupun gelisah. Dan menurut Aries Dirgayunita Depresi merupakan kondisi emosional yang biasanya ditandai dengan kesedihan yang amat sangat, perasaan tidak berarti dan bersalah seperti menarik diri tidak dapat tidur, kehilangan selera, minat dalam aktivitas sehari-hari (Adi Zayd Bintang et al., 2021).

Sebagai manusia modern sebagai bentuk penyeimbang kemajuan teknologi

yang seakan tidak ada habisnya, sudah seharusnya bagi kita untuk meningkatkan juga kepedulian kita terhadap sekitar, terlebih lagi kepada sesama manusia.

1.2. Rumusan Masalah

Dalam melakukan penelitian ini, penulis membuat perumusan masalah sebagai berikut:

1. Bagaimana metode *K-Means Clustering* berkerja terhadap data depresi yang ada berdasarkan Instrumen DASS-42?
2. Apakah optimasi *K-Means* dengan *Particle Swarm Optimization* lebih baik akurasiya ketika model dijalankan?
3. Sedekat apa sebenarnya kehidupan manusia dengan depresi?
4. Berapa banyak orang yang memasuki kategori depresi berat?

1.3. Tujuan dan Manfaat

Tujuan dari penenelitian ini ialah:

1. Mengevaluasi performa *K-Means Clustering* yang dioptimasi menggunakan *Particle Swarm Optimization* (PSO) dalam menganalisis data terkait depresi berat.
2. Mengidentifikasi sejauh mana depresi berat terkait dengan kehidupan sehari-hari.

3. Menganalisis kemungkinan-kemungkinan buruk yang dapat terjadi pada pengidap depresi berat, sehingga dapat dilakukan tindakan pencegahan atau intervensi yang lebih baik

Selanjutnya manfaat sebagai berikut:

1. Manfaat untuk penulis:

- a. Penelitian ini dapat dimanfaatkan oleh penulis sebagai salah satu persyaratan untuk menyelesaikan program pendidikan Sarjana. (S1) yang dibebankan kepada penulis pada Program Studi Informatika Kampus Margonda Kota Depok Universitas Bina Sarana Informatika.
- b. Menjadikan penelitian ini sebagai jawaban dari banyak hal yang ingin penulis lakukan dan ketahui.

2. Manfaat untuk objek penelitian:

- a. Para responden yang memang sedang merasa kesehatan mentalnya sedang tidak baik dapat lebih mengetahui keadaan diri mereka dengan lebih spesifik jika mereka menginginkan hasil dari kuis yang mereka isi untuk mereka ketahui hasilnya.
- b. Para responden dapat lebih kritis terhadap diri mereka sendiri setelah mengetahui bahwa mereka memang perlu penanganan professional.
- c. Beberapa responden yang terbuka dan sudah mengetahui hasil, penulis berikan dukungan berupa semangat dan imbauan untuk lebih mencintai dan memperdulikan diri mereka sendiri.

3. Manfaat untuk pembaca:

- a. Pembaca diharapkan mendapat pengetahuan baru mengenai depresi dan dapat mengecek diri mereka masing-masing menggunakan instrument sederhana DASS-42.
- b. Pembaca dapat meningkatkan kesadaran mereka mengenai kesehatan mental khususnya depresi.
- c. Pembaca dapat mengikuti alur penelitian ini sebagai bahan rujukan studi pustaka untuk membuat penelitian sendiri.

1.4. Hipotesis

Hipotesis yang peneliti dapatkan dari penelitian *K-Means* dan *Particle Swaem Optimization* (PSO) adalah:

H0: *Clustering* dengan algoritma *K-Means* yang dioptimasi dengan *Particle Swarm Optimization* tidak akan menghasilkan akurasi yang lebih tinggi dibandingkan dengan *clustering* murni hanya algoritma *K-Means*.

H1: *Clustering* menggunakan algoritma *K-Means* yang dioptimasi dengan *Particle Swarm Optimization* akan menghasilkan nilai hasil yang lebih presisi setelah dikomparasi dengan *clustering* yang hanya memanfaatkan murni algoritma *K-Means* saja.

1.5. Batasan Masalah

Data yang penulis kumpulkan penulis dapat dari menggunakan kuesioner *Instrumen DASS-42*. Meskipun begitu, penulis hanya berfokus *Variable* depresinya saja. Dalam menjalankan model algoritma hasil pengelompokan, penulis juga hanya berfokus pada Algoritma *Clustering* menggunakan *K-Means* yang kemudian penulis optimasi menggunakan *Particle Swarm Optimazation*.

BAB II

LANDASAN TEORI

2.1. Tinjauan Pustaka

Beberapa dukungan fundamental penulis dalam melaksanakan dan penyusunan penelitian ini, penulis tentu saja memerlukan sumber dari studi lain yang memiliki koneksi yang sama dengan penelitian penulis yang berjudul “Penerapan Algoritma K-Means *Clustering* Sebagai Media Pengelompokan Depresi Menggunakan Instrumen Dass-42” ini sebagai dasar dari sumber keilmuan yang terkait yaitu, sebagai berikut.

2.1.1. Depresi

Menurut Hoban pada tahun 2009, dan Reavley, Jorm pada tahun 2010 depresi adalah gangguan mental yang ditandai dengan munculnya perasaan sedih dan cemas. Mahasiswa rentan mengalami depresi, berdasarkan hasil riset sebelumnya menunjukkan bahwa prevalensi terjadinya depresi pada mahasiswa cukup tinggi yakni sekitar 33%, selain itu pada penelitian lainnya dilaporkan bahwa 43% mahasiswa mengalami gejala depresi yang mengganggu proses pembelajaran, sehingga hal tersebut berdampak negatif pada kinerja akademis mahasiswa (Senjaya et al., 2022).

2.1.2. DASS-42

DASS-42 adalah sekumpulan form pertanyaan skala pengukuran depresi, kecemasan, dan stress. DASS-42 adalah suatu instrumen pengukuran psikologis yang didesain untuk mengevaluasi tiga kondisi emosional negatif, yakni depresi, kecemasan, dan stres. “Skala Depresi, Kecemasan, dan Stres-42 (DASS-42) merupakan skala yang menguji tiga subskala depresi, kecemasan, dan stres - baik secara dimensional maupun kategoris, dan membedakan masing-masing dalam kuesioner yang sama” (Karadoğan & Altay, 2024).

2.1.3. Algoritma

Konsep algoritma telah ada sejak abad ke-9, ditemukan oleh Muhammad Ibn Musa al-Khawarizmi. Algoritma merupakan serangkaian langkah atau prosedur yang sistematis, digunakan untuk memecahkan suatu masalah secara teratur. Dalam pemrograman komputer, algoritma berkaitan erat dengan logika dalam menentukan program yang akan ditulis atau dibuat. “Algoritma dapat diartikan sebagai sebuah perhitungan atau operasi untuk memecahkan masalah, terutama proses atau tahapan yang harus diikuti oleh computer” (Tharisa amalia, 2020). Sementara itu, dari perspektif lain algoritma dapat diartikan sebagai “sebuah upaya melalui urutan operasi yang disusun secara logis dan sistematis, dengan tujuan untuk menyelesaikan suatu masalah tertentu. Algoritma dirancang sedemikian rupa sehingga dapat menghasilkan suatu hasil atau output yang diinginkan dalam waktu yang relatif singkat” (Hasibuan & Yahfizham, 2023).

2.1.4. K-Means Clustering

“*Clustering* merupakan salah satu alat bantu dalam data *mining*. *Clustering* bertujuan untuk melakukan pengelompokkan objek-objek ke dalam cluster-cluster” (Virgo et al., 2020). *Cluster* adalah kumpulan record yang memiliki kemiripan satu dengan yang lainnya dan memiliki ketidakmiripan dengan *record* lain dalam *cluster*. Salah satu algoritma *machine learning* yang tidak membutuhkan data berlabel adalah *K-Means*. *Unsupervised* sendiri dapat diartikan menjadi tidak berlabel atau tidak diawasi, yang berarti dataset yang kita gunakan tidak diberi label.

2.1.5. Particle Swarm Optimization

“*Particle Swarm Optimization* adalah metode penelitian populasi yang menggunakan populasi (*swarm*) individu (partikel) yang diperbarui dari setiap iterasi yang dilakukan. Dalam membuat partikel mencapai solusi optimalnya, setiap partikel

bergerak ke arah posisi terbaik sebelumnya (pbest) dan posisi global terbaik (gbest)”(Hayatin et al., 2020).

2.1.6. Silhouette Score

“Pengujian silhouette score adalah proses evaluasi yang digunakan dengan tujuan mengevaluasi seberapa efektif suatu algoritma clustering mampu memisahkan data menjadi kelompok-kelompok yang berbeda. Metode ini melibatkan perhitungan silhouette score untuk setiap titik data, yang mencerminkan seberapa baik titik tersebut cocok dengan kelompoknya sendiri dibandingkan dengan kelompok lainnya” (Hendrastuty, 2024).

2.1.7. Calinski-Harabasz Score

“Indeks Calinski-Harabasz adalah perbandingan antara jumlah varians antar kelompok dan jumlah varians dalam kelompok untuk semua kelompok, di mana varians dihitung sebagai jumlah kuadrat jarak” (Sikana & Wijayanto, 2021).

2.1.8. Davies-Bouldin Score

“Davies-Bouldin Index (DBI) adalah salah satu metode yang digunakan untuk mengukur validitas cluster dalam metode clustering. Pengukuran ini dengan Indeks DaviesBouldin memaksimalkan jarak antar-cluster dan pada saat yang sama mencoba meminimalkan jarak antar titik dalam sebuah cluster” (Tarigan, 2023)

2.2. Penelitian Terkait

Dalam melakukan penelitian ini, penulis tentu saja melihat bagaimana metode yang penulis gunakan pada penelitian sebelumnya sebagai acuan. Berikut adalah beberapa penelitian yang terkait dengan penelitian penulis.

Pertama, menurut (Pasaribu, 2021) dalam jurnal dengan judul “*Memfaatkan Algoritma K-Means Dalam Memetakan Potensi Hasil Produksi Kelapa Sawit PTPN IV Marihat*” peneliti menggunakan algoritma K-Means untuk memetakan potensi hasil dari kelapa sawit pada PT. Perkebunan Nusantara (PTPN). Peneliti menyimpulkan Metode yang diterapkan dalam perhitungan skripsi ini adalah penerapan teknik data mining dengan menggunakan metode Clustering K-Means. Hasil penelitian menggunakan metode K-Means mengidentifikasi 2 kelompok (cluster), yaitu cluster dengan hasil produksi tinggi dan cluster dengan hasil produksi rendah. Algoritma K-Means dapat diimplementasikan untuk melakukan pemetaan terhadap hasil produksi buah kelapa sawit dengan menggunakan perangkat lunak RapidMiner. Dari hasil penelitian, didapatkan cluster tertinggi (C1) terdiri dari 14 data Blok Panen, sedangkan cluster terendah (C2) terdiri dari 11 data Blok Panen.

Kedua, menurut (Fauziah & Purnamasari, 2023) dalam penelitiannya yang berjudul “*Implementasi Algoritma K-Means pada Kasus Kekerasan Anak dan Perempuan Berdasarkan Usia*” ia menggunakan algoritma K-Means untuk mengcluster data kekerasan pada anak dan mendapatkan hasil sebagai pengelompokan kasus kekerasan terhadap anak dan perempuan dengan menggunakan algoritma K-Means menghasilkan beberapa cluster atau kelompok, data tersebut terbagi menjadi 6 cluster berdasarkan kesamaan karakteristik usia korban. Cluster 0 (C0) memiliki 55 anggota, didominasi oleh korban berusia 13-17 tahun. Cluster 1 (C1) memiliki 193 anggota. Cluster 2 (C2) memiliki 184 anggota, didominasi oleh korban berusia 45-59 tahun. Cluster 3 (C3) memiliki 197 anggota, didominasi oleh korban berusia 45-59 tahun. Cluster 4 memiliki 4 anggota, korban berusia 25-44 tahun. Cluster 5 (C5) memiliki 177 anggota, didominasi oleh korban berusia 0-5 tahun. Dari 810 dataset yang

digunakan, hasil terbaik diperoleh pada cluster 6 dengan nilai Davies-Bouldin Index (DBI) 0.211, yang mendekati 0. Oleh karena itu, untuk penelitian selanjutnya disarankan untuk menggunakan metode atau algoritma lain dalam mengelompokkan kasus kekerasan anak dan perempuan, serta melakukan eksperimen dengan parameter yang berbeda untuk mendapatkan evaluasi yang lebih optimal.

Ketiga, menurut (Homepage et al., 2021) dalam jurnalnya yang berjudul "*Comparison of K-Means and K-Medoids Algorithms for Grouping Landslide Prone Areas in West Java Province Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokan Daerah Rawan Tanah Longsor di Provinsi Jawa Barat*" penulis melakukan pengelompokan data kejadian tanah longsor di Jawa Barat pada tahun 2019 menggunakan algoritma K-Means menghasilkan pengelompokan yang lebih optimal dibandingkan dengan algoritma K-Medoids. Jumlah kelompok optimal yang diperoleh adalah 6. Hasil pengelompokan menunjukkan bahwa kelompok atau kluster 2 memiliki jumlah daerah paling banyak, yang berarti kluster tersebut merupakan kluster dominan. Sementara itu, jumlah kejadian tanah longsor terbanyak terdapat pada kluster 5, yaitu sebanyak 106 kejadian di 4 daerah. Jadi, dapat disimpulkan bahwa metode K-Means menghasilkan pengelompokan yang lebih optimal dibandingkan K-Medoids pada data tanah longsor di Jawa Barat 2019, dengan jumlah kluster optimal sebanyak 6 kluster, di mana kluster 2 kluster tersebut merupakan kluster yang paling dominan, dan kluster 5 memiliki jumlah kejadian tanah longsor terbanyak.

Keempat, menurut (Lashiyanti et al., 2023) dalam jurnalnya yang berjudul "*Optimisasi Klasterisasi Nilai Ujian Nasional dengan Pendekatan Algoritma K-Means, Elbow, dan Silhouette*" peneliti mendapatkan hasil berupa Berdasarkan analisis yang telah dilakukan, dapat disimpulkan beberapa hal terkait data nilai ujian

nasional tingkat SMK. Untuk kelompok nilai UN dengan $K=3$, diperoleh rata-rata nilai UN Bahasa Indonesia sebesar 72,79906, rata-rata nilai UN Bahasa Inggris sebesar 45,941, rata-rata nilai UN Matematika sebesar 41,324, dan rata-rata nilai UN Kompetensi sebesar 48,1947. Sementara untuk kelompok nilai UN dengan $K=2$, diperoleh rata-rata nilai UN Bahasa Indonesia sebesar 76,95, rata-rata nilai UN Bahasa Inggris sebesar 33,425, rata-rata nilai UN Matematika sebesar 45,65, dan rata-rata nilai UN Kompetensi sebesar 52,54. Dalam pengelompokan data nilai ujian nasional menggunakan metode cluster k-means dengan optimasi elbow, diperoleh 3 kluster dengan rincian: Kluster 1 memiliki 707 anggota, Kluster 2 memiliki 152 anggota, dan Kluster 3 memiliki 675 anggota. Cluster sum of squares masing-masing kluster adalah 717,4831 untuk Kluster 1, 516,4798 untuk Kluster 2, dan 776,3202 untuk Kluster 3, dengan nilai *between_ss* sebesar 67,2%. Sementara itu, penggunaan metode cluster k-means dengan optimasi silhouette menghasilkan 2 kluster, dengan Kluster 1 memiliki 1044 anggota dan Kluster 2 memiliki 490 anggota. Sum of squares Kluster 1 adalah 1416,049 dan Kluster 2 adalah 1675,411, dengan total *ss* sebesar 49,6%. Secara keseluruhan, optimasi cluster elbow dan silhouette yang digunakan pada metode clustering k-means memberikan hasil yang cukup baik dalam mengelompokkan permasalahan yang ada.

Kelima, menurut (Citra Pengenalan Aksara Bugis Kurniati & Reza Wardana, 2020) pada jurnalnya yang berjudul "*Penerapan Algoritma Particle Swarm Optimization pada Segmentasi Citra Pengenalan Aksara Bugis*" peneliti membahas penerapan algoritma Particle Swarm Optimization (PSO) pada segmentasi citra aksara Bugis. Salah satu temuan utama adalah bahwa spesifikasi perangkat yang digunakan berpengaruh terhadap waktu proses pengenalan huruf Bugis. Sistem yang didesain dalam penelitian

ini mengasumsikan bahwa citra input memiliki noise yang sekecil mungkin. Hal ini penting karena noise dapat menurunkan akurasi dalam proses pengenalan, meskipun bagi manusia karakter tersebut masih dapat terbaca. Sistem yang dikembangkan sangat bergantung pada kualitas pemisahan aksara yang baik untuk mencapai keberhasilan dalam mengenali suatu karakter. Jadi, penelitian ini berfokus pada penerapan algoritma PSO untuk segmentasi citra aksara Bugis, dengan perhatian khusus pada dampak spesifikasi perangkat terhadap waktu proses pengenalan, serta pentingnya kualitas pemisahan aksara dalam meningkatkan akurasi pengenalan huruf Bugis oleh sistem

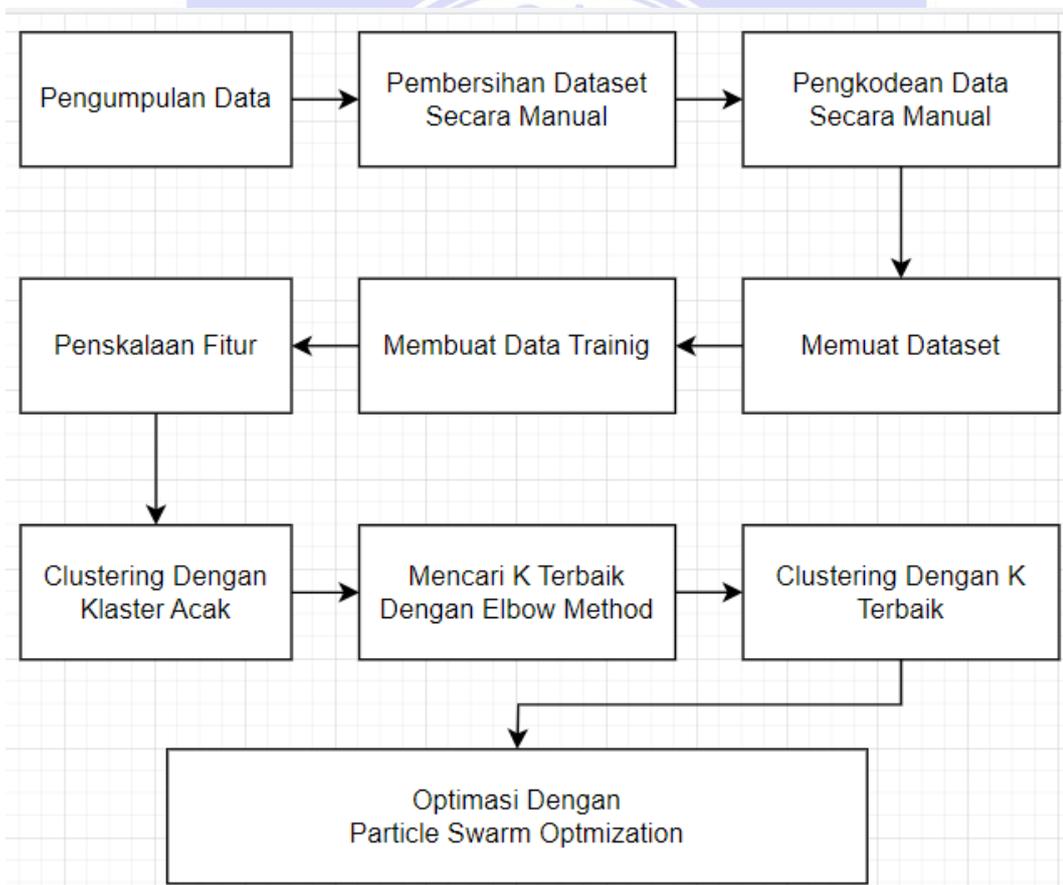


BAB III

METODOLOGI PENELITIAN

3.1. Proses dan Langkah Penelitian

Dalam menyusun penelitian “Optimasi *K-Means Clustering* Dengan *Particle Swarm Optimization* Untuk Pengelompokan Depresi Berdasarkan DASS-42” ini, terdapat proses penelitian yang berisi: kerangka penelitian, sumber data, teknik pengambilan data. Berikut adalah tahapan penelitian penulis:



Sumber: (Peneliti, 2024)

Gambar III. 1
Tahapan Penelitian

3.1.1. Pengumpulan Data

Dalam penelitian ini, peneliti menggunakan kuesioner *Google Forms* untuk mengumpulkan data terkait tingkat depresi responden. Kuesioner ini disebar secara daring dan berhasil memperoleh 106 data berbentuk *Comma Separated Values* tanggapan dari peserta. Melalui kuesioner ini, peneliti dapat mengukur dan menganalisis tingkat depresi yang dialami oleh para responden. Data yang terkumpul akan diolah dan dianalisis lebih lanjut untuk memahami profil dan karakteristik depresi di populasi yang diteliti. Hasil penelitian ini diharapkan dapat memberikan wawasan berharga mengenai isu kesehatan mental dan membantu dalam pengembangan intervensi yang tepat sasaran.

3.1.2. Pembersihan Dataset Secara Manual

Penulis mengumpulkan 106 data melalui kuesioner *Google Forms*, peneliti kemudian membersihkan data tersebut secara manual di *Google Sheets*. Terdapat beberapa kuesioner yang tidak terisi lengkap dan juga data *double*, sehingga peneliti memutuskan untuk mengeliminasi data-data yang kosong tersebut. Setelah proses pembersihan data, jumlah data yang valid dan dapat digunakan dalam analisis selanjutnya menjadi 94 data. Dengan demikian, dari total 106 tanggapan yang terkumpul, hanya 94 data yang dianggap lengkap dan memenuhi syarat untuk diproses lebih lanjut dalam penelitian ini. Langkah ini dilakukan untuk memastikan akurasi dan reliabilitas analisis yang akan dilakukan pada data depresi yang terkumpul.

3.1.3. Pengkodean Data Secara Manual

Setelah membersihkan data dan menghasilkan 94 data yang valid, peneliti selanjutnya melakukan pengkodean pada variabel jenis kelamin. Jenis kelamin responden dikodekan menjadi "KODE JK". Selain itu, peneliti juga melakukan pengkodean terhadap tingkat depresi yang dialami oleh responden. Variabel ini diberi kode "KODE TD" agar dapat diintegrasikan dengan mudah dalam analisis data. Dimana perempuan diberi kode 0 dan laki-laki diberi kode 1. Tujuan dari pengkodean ini adalah untuk memudahkan pengelolaan dan analisis data lebih lanjut.

Dengan mengubah data jenis kelamin menjadi format kode numerik, peneliti dapat dengan mudah memasukkan dan memproses variabel ini ke dalam analisis statistik. Pengkodean data jenis kelamin menjadi KODE JK, dengan 0 untuk perempuan dan 1 untuk laki-laki. Selain itu, peneliti juga melakukan pengkodean terhadap tingkat depresi yang dialami oleh responden. Variabel ini diberi kode "KODE TD". Pengkodean data jenis kelamin menjadi KODE JK, dengan 0 untuk perempuan dan 1 untuk laki-laki, serta pengkodean tingkat depresi menjadi KODE TD dengan skala 1 hingga 5, merupakan langkah penting dalam mempersiapkan data untuk analisis yang lebih komprehensif nantinya di *Google Colab*.

3.1.4. Memuat Dataset

Dalam menjalankan penelitian ini, penulis mengumpulkan dan mempersiapkan data untuk dijalankan di model *clustering*. Data yang penulis muat atau upload adalah data yang telah penulis kodekan pada proses sebelumnya.

3.1.5. Membuat Data Training

Dalam menjalankan *K-Means clustering*, data training harus dibuat terlebih dahulu. Membuat data training untuk keperluan *clustering* dibutuhkan sebagai awal agar dapat memahami struktur data. Data training digunakan sebagai model yang melatih atau mengajarkan algoritma K-Means bagaimana mengelompokkan data berdasarkan kesamaan fitur. Algoritma akan mempelajari pola dan karakteristik data sehingga dapat menentukan centroid (pusat *cluster*) yang optimal.

3.1.6. Penskalaan Fitur

Penskalaan fitur (*feature scaling*) dilakukan untuk mengubah skala nilai pada fitur atau array. Peneliti menggunakan penskalaan fitur *MinMaxScaler* dari sklearn agar array berada dalam rentang yang seragam yaitu antara 0 dan 1. Untuk menghitung parameter skala (minimum dan maksimum) berdasarkan data training x_{train} , lalu mentransformasikan x_{train} .

3.1.7. Clustering

Berdasarkan dampak yang diharapkan, *machine learning* dibagi menjadi 3 yaitu *Supervised Learning*, *Unsupervised Learning* dan *Reinforcement Learning*. Peneliti menggunakan *Clustering* yang masuk ke dalam kategori *Unsupervised Learning* dengan algoritma *K-Means*. Peneliti akan mengkategorikan data ke dalam kluster-kluster berdasarkan kemiripan karakteristik data ke tingkat depresi yang dialami oleh responden.

Setelah melakukan *clustering*, penulis merubah nilai numerik yang ada pada cluster menjadi label tekstual guna memudahkan dalam memahami data. Penulis awalnya mengisi nilai-nilai dari $y_{cluster}$ dengan “KODE HSL” dalam datagrame df .

Lalu, penulis mengganti nilai-nilai yang ada pada kolom “KODE HSL” yang sebelumnya adalah 0 dan 1 menjadi “SEHAT” dan “DEPRESI”.

3.1.8. *Clustering* Menggunakan K-Terbaik

Dalam menjalankan algoritma *K-Means* terdapat metode *Elbow-Method* untuk mencari K Terbaik. K terbaik disini dimaksudkan untuk jumlah *cluster* yang paling optimal, sehingga hasil dari *clustering* selanjutnya akan menggunakan nilai K jumlah *cluster* yang ditunjukkan dari hasil *Elbow Method*. Metode *Elbow* menentukan jumlah *cluster* yang paling tepat berdasarkan perubahan nilai *Sum of Squared Errors* (SSE) atau inerti yang ditunjukkan dengan grafik paling patah atau siku (*elbow*).

Peneliti menjalankan kembali proses *clustering* dengan menerapkan kembali algoritma *K-Means* dengan jumlah *cluster* yang optimal, dan menyimpan hasilnya dalam variabel *y_cluster*.

3.1.9. Membuat *Dataset* Baru

Setelah menentukan jumlah *cluster* yang optimal menggunakan metode *Elbow*, peneliti akan membuat dataset baru yang mencakup label *cluster* untuk setiap data. Pembuatan dataset baru ini ditujukan untuk memudahkan memahami perbandingan dari hasil *clustering* yang sebelumnya. Dalam membuat dataset baru, penulis menambahkan sebuah kolom baru ke dalam dataframe *df* yang disebut 'KODE HSL2', dan mengisi kolom tersebut dengan nilai-nilai dari variabel *y_cluster*.

Dengan menambahkan kolom 'KODE HSL 2' ke dalam dataset awal, peneliti telah membuat *dataset* baru yang memiliki informasi tambahan mengenai *cluster* yang lebih mudah dipahami dengan jumlah *cluster* didasarkan dari hasil K terbaik menggunakan *elbow method*.

3.1.10. Optimasi Dengan *Particle Swarm Optimization*

Particle swarm optimization (PSO) adalah salah satu metode optimasi populasi yang terinspirasi oleh perilaku kawanan hewan seperti burung atau ikan. Disebut demikian karena *Particle Swarm Optimization* adalah algoritma yang *metaheuristic* yang berjalan dengan teknik swarm atau gerombolan.

Dengan mengombinasikan *K-Means* dan *Particle Swarm Optimization*, peneliti dapat memperoleh hasil clustering yang lebih baik dan lebih akurat dibandingkan hanya menggunakan *K-Means* biasa.

3.2. Metode Pengolahan dan Analisis Data

Dalam proses penyusunan tugas akhir ini terdapat beberapa metode pengolahan dan analisis data yang penulis gunakan ialah sebagai berikut.

3.2.1. Metode Pengolahan Data

Dalam proses pengolahan data, penulis menggunakan metode kuesioner dari instrumen *form* DASS-42. Pengumpulan data penulis menggunakan *form* DASS-42 ini penulis gunakan untuk penyebaran kuis kepada objek penelitian penulis. Instrumen DASS-42 dibuat untuk memproses lebih lanjut pemahaman, pemahaman, dan pengukuran keadaan emosi yang umum, termasuk stres dan keadaan emosi tradisional. Berikut adalah detail instrumen DASS-42 yang penulis gunakan

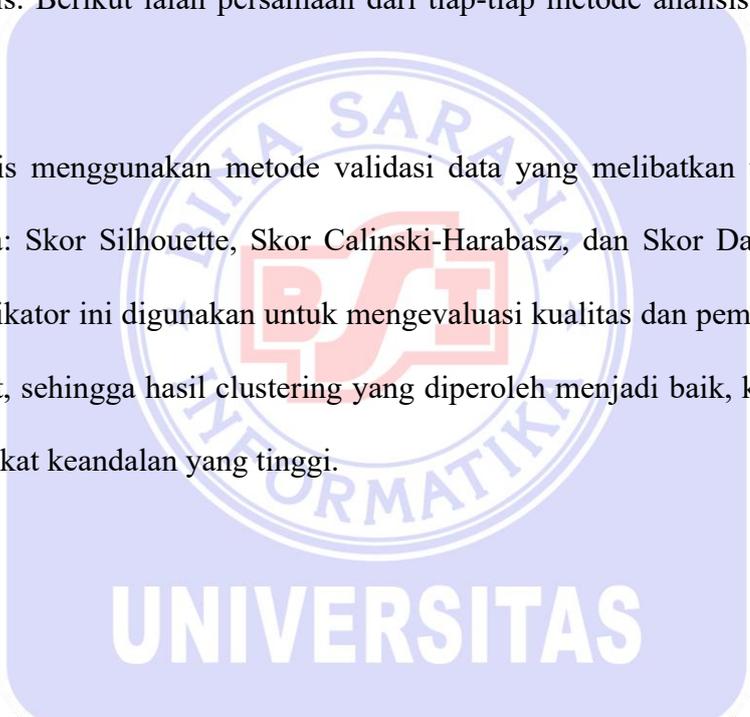
Penulis juga memanfaatkan alat-alat digital seperti *Microsoft Excel* dan *Google Colab* dalam mengelola dan mengolah data. *Excel*, yang merupakan aplikasi lembar sebar populer, digunakan oleh penulis sebagai platform untuk melakukan berbagai operasi dan analisis data. Sementara itu, *Google Colab*, sebuah lingkungan pemrograman berbasis web, juga dimanfaatkan oleh penulis sebagai sarana untuk

mengolah dan memproses data secara interaktif. dengan menggunakan kombinasi dari kedua alat tersebut, penulis dapat melakukan tugas-tugas terkait pengolahan data dengan lebih efisien dan efektif.

3.2.2. Metode Analisis Data

Dalam melakukan analisis data, penulis menerapkan pendekatan berbasis clustering dengan menggunakan algoritma k-means. Selain itu, penulis juga memanfaatkan metode *Particle Swarm Optimization* (PSO) sebagai alat bantu dalam proses analisis. Berikut ialah persamaan dari tiap-tiap metode analisis yang penulis gunakan.

Penulis menggunakan metode validasi data yang melibatkan tiga indikator yang berbeda: Skor Silhouette, Skor Calinski-Harabasz, dan Skor Davies-Bouldin. Indikator-indikator ini digunakan untuk mengevaluasi kualitas dan pemisahan kluster dalam dataset, sehingga hasil clustering yang diperoleh menjadi baik, konsisten, dan memiliki tingkat keandalan yang tinggi.



UNIVERSITAS

1. *K-Means*

Berikut adalah persamaan *K-Means* dari <https://scikit-learn.org>:

$$\sum_{i=0}^n \min_{\mu_j \in C} (\|x_i - \mu_j\|^2)$$

Sumber: <https://scikit-learn.org/stable/modules/clustering.html#k-means>

Gambar III. 2
Persamaan K-Means

Keterangan gambar III. 2:

x_j : Data point ke- j .

C_i : Cluster ke- i .

c_i : Centroid cluster ke- i .

$1_{\{x_j \in C_i\}}$: Fungsi indicator yang bernilai 1 jika x_j tergolong ke dalam cluster C_i dan bernilai 0 jika tidak.

$\|x_j - c_i\|^2$: Jarak antara data point x_j dengan centroid cluster C_i

2. Optimasi dengan *Particle Swarm Optimization*

Particle Swarm Optimization adalah algoritma iteratif berbasis populasi dengan banyak partikel. Inisialisasi populasi dilakukan secara acak untuk menyelesaikan masalah optimasi. Optimasi dengan *Particle Swarm Optimization* dilakukan dengan melakukan fungsi optimasi dengan urutan tahapan sebagai berikut:

- a). Optimasi dengan Particle Swarm Optimization.
- b). Cari jumlah cluster optimal dengan inertia.

c). Klasterisasi final dengan K-Means.

3. Validasi Hasil

Dalam penelitian ini, penulis menerapkan beberapa metode validasi data untuk mengevaluasi hasil pengelompokan. Salah satunya adalah *Silhouette Score*, yang mengukur seberapa baik setiap objek ditempatkan dalam klasternya. Nilai *Silhouette Score* berkisar antara -1 dan 1, di mana nilai yang lebih tinggi menunjukkan pengelompokan yang lebih baik. Selain itu, penulis juga menggunakan *Calinski-Harabasz Score*, yang mengukur rasio antara *varians* antar-klaster dan *varians* dalam-klaster. Semakin tinggi nilai *Calinski-Harabasz Score*, semakin baik pula hasil pengelompokan data. Selanjutnya, penulis menerapkan *Davies-Bouldin Score*, yang mengevaluasi kemiripan antar-klaster. Nilai *Davies-Bouldin Score* yang lebih rendah menandakan pengelompokan yang lebih baik. Dengan mengombinasikan ketiga metode validasi ini, penulis dapat menilai kualitas dan kesesuaian hasil pengelompokan data yang diperoleh dalam penelitian.



UNIVERSITAS

BAB IV

HASIL DAN PEMBAHASAN

4.1. Hasil Penelitian

Dalam menyusun tugas akhir ini, penulis mendapatkan hasil dalam beberapa proses. Berikut rinciannya.

4.1.1. Pengumpulan Data

Pengumpulan data yang penulis lakukan mendapatkan hasil yang relevan dengan yang penulis harapkan. Data yang penulis dapatkan berisi data kualitatif dan kuantitatif yang seluruhnya berjumlah 106 data. Berikut adalah sample data yang penulis dapatkan.

Tabel IV. 1
Sampel Data

31/03/2024 15:14:43	Gilang apdillah	Laki- Laki	20- 30	Belajar + Berkerja	0	2	1	2	0	1	0	1	1	0	0	1	1	1	11	RINGAN	
31/03/2024 15:18:34	Dzul Arsi Majid	Laki- Laki	20- 30	Belajar	1	1	1	0	0	0	0	0	0	0	1	0	0	0	0	4	NORMAL

Sumber: (Peneliti, 2024)

4.1.2. Pembersihan Dataset Secara Manual

Pembersihan data secara manual yang penulis lakukan menghasilkan data bersih berjumlah 94 data. Berikut sample data yang telah penulis bersihkan.

Tabel IV. 2
Sampel Data Yang telah Dibersihkan

Laki-Laki	20-30	Belajar + Berkerja	0	2	1	2	0	1	0	1	1	0	0	1	1	1	11	RINGAN	
Laki-Laki	20-30	Belajar	1	1	1	0	0	0	0	0	0	0	1	0	0	0	0	4	NORMAL

Sumber: (Peneliti, 2024)

4.1.3. Pengkodean Data Secara Manual

Penulis melakukan pengkodean dengan jenis kelamin responden dikodekan menjadi "KODE JK". tingkat depresi diberi kode "KODE TD". Jenis kelamin Perempuan diwakili dengan kode 0, sedangkan Laki-laki diwakili dengan kode 1. Pengkodean terhadap tingkat depresi yang dialami oleh responden diberi kode "KODE TD".

Tabel IV. 3
Sample Data Setelah Dikodekan

Laki-Laki	1	20-30	Belajar + Berkerja	0	2	1	2	0	1	0	1	1	0	0	1	1	1	11	RINGAN	2
Laki-Laki	1	20-30	Belajar	1	1	1	0	0	0	0	0	0	1	0	0	0	0	4	NORMAL	1

Sumber: (Peneliti, 2024)

KODE JK

KODE TD

Keterangan:

1 = Normal

2 = Ringan

3 = Sedang

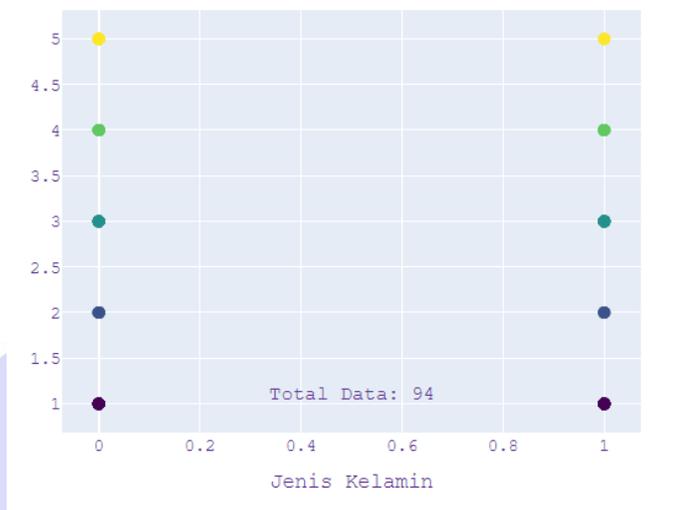
4 = Parah

5 = Sangat Parah

4.1.4. Memuat Dataset

Memuat *dataset* sederhananya ialah mengupload *dataset*. Dalam proses pemuatan *dataset*, penulis menggunakan *google colab* sebagai media untuk

menjalankan *syntax Phyton* sebagai *platform* untuk menjalankan model *K-Means* penulis.



Sumber: (Peneliti, 2024)

Gambar IV. 1
Visualisasi Data

4.1.5. Membuat *Data Training*

Setelah memuat dataset yang telah dibersihkan dan dikodekan ke dalam *Google Colab*, langkah selanjutnya adalah membuat data *training* untuk keperluan clustering. Peneliti membuat data training dengan menggunakan perintah:

```
x_train = df[['KODE TD','KODE TD']].values
x_train
```

Sumber: (Peneliti, 2024)

Gambar IV. 2
Kode Untuk Membuat *Data Training*

Dalam perintah tersebut, `x_train` merupakan variabel yang berisi data training. Data training diambil dari kolom 'KODE TD' dan 'KODE TD' pada dataset `df`. Penggunaan `df[['KODE TD','KODE TD']].values` akan mengekstrak nilai-nilai dari kedua kolom tersebut dan menyimpannya dalam variabel `x_train`. Proses ini mempersiapkan data training yang akan digunakan untuk *clustering*. Dengan memiliki

data training yang telah siap, peneliti dapat melanjutkan ke tahap selanjutnya dalam proses *clustering* data depresi.

Proses ini mempersiapkan data training yang akan digunakan untuk clustering. Dengan memiliki data training yang telah siap, peneliti dapat melanjutkan ke tahap selanjutnya dalam proses clustering data depresi. Berikut adalah sample array dari data training `x_train`:

```
array([[3, 3],
       [1, 1],
       [5, 5],
       [2, 2],
       [1, 1],
       [1, 1],
       [3, 3],
       [3, 3],
       [3, 3],
       [4, 4],
       [4, 4],
       [3, 3],
```

Sumber: (Peneliti, 2024)

Gambar IV. 3
Sample Data Training

4.1.6. Penskalaan Fitur

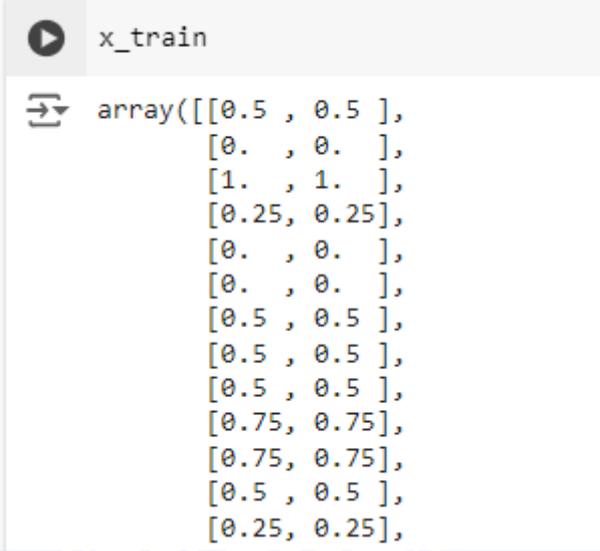
Dalam proses penskalaan fitur atau *Feature Scaling* peneliti menggunakan metode *MinMaxScaler* dari library *sklearn.preprocessing* untuk melakukan penskalaan fitur. Yaitu, dengan *code*:

```
[7] from sklearn.preprocessing import MinMaxScaler
scaler = MinMaxScaler()
x_train=scaler.fit_transform(x_train)
```

Sumber: (Peneliti, 2024)

Gambar IV. 4
Code Untuk Feature Scaling Dengan *MinMaxScaler*

Dalam kode di atas, pertama-tama peneliti mengimpor `MinMaxScaler` dari `sklearn.preprocessing`. Kemudian, peneliti membuat objek scaler dari kelas `MinMaxScaler()`. Selanjutnya, `scaler.fit_transform(x_train)` digunakan untuk menghitung parameter skala (minimum dan maksimum) berdasarkan data training `x_train`, lalu mentransformasikan `x_train` agar berada dalam rentang 0 hingga 1. Hasil dari *code* diatas adalah array seperti sample berikut:



```
x_train
array([[0.5 , 0.5 ],
       [0. , 0. ],
       [1. , 1. ],
       [0.25, 0.25],
       [0. , 0. ],
       [0. , 0. ],
       [0.5 , 0.5 ],
       [0.5 , 0.5 ],
       [0.5 , 0.5 ],
       [0.75, 0.75],
       [0.75, 0.75],
       [0.5 , 0.5 ],
       [0.25, 0.25],
```

Sumber: (Peneliti, 2024)

Gambar IV. 5
Sample Data Training Setelah Feature Scaling

4.1.7. Clustering

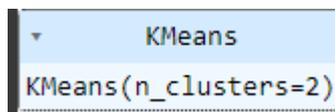
Dalam *clustering* menggunakan algoritma *K-Means*, peneliti akan mengelompokkan data berdasarkan tingkat depresi yang dialami oleh responden dengan menggunakan *code*:

```
[9] from sklearn.cluster import KMeans
     Kmeans = KMeans(n_clusters=2)
     Kmeans
```

Sumber: (Peneliti, 2024)

Gambar IV. 6
Code Untuk Menjalankan K-Means

Dalam kode di atas, pertama-tama peneliti mengimpor *K-Means* dari `sklearn.cluster`. Kemudian, peneliti membuat objek *Kmeans* dari kelas `KMeans()` dan menentukan jumlah *cluster* yang akan dibuat, yaitu 2 *cluster*.



Sumber: (Peneliti, 2024)

Gambar IV. 7
K-Means Dengan 2 Titik 2 Centroid

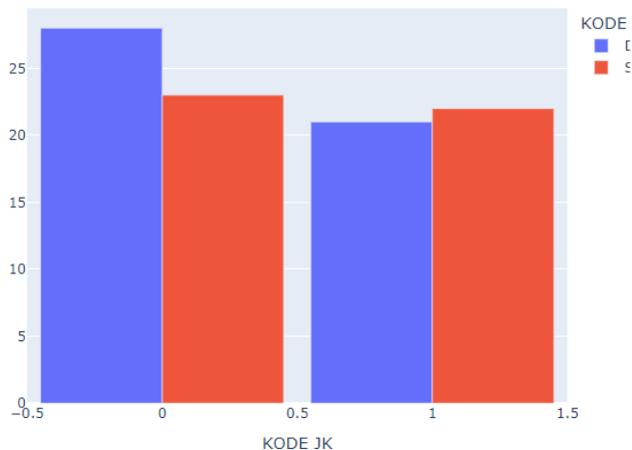
Kode pada gambar IV.6 dapat dijelaskan `Kmeans.fit_predict(x_train)` digunakan untuk melakukan proses clustering pada data *training* `x_train`. Hasil dari proses *clustering* ini disimpan dalam variabel `y_cluster`, yang berisi *label cluster* untuk setiap data dalam `x_train`. Contoh *sample* `y_cluster`: [1, 0].

Selanjutnya, penulis membuat kolom baru yang bernama “KODE HSL” yang diisi dengan data dari `y_cluster` dan penulis simpan di dalam *dataframe*. Kolom baru ini dimaksudkan sebagai data yang berisi hasil pertama dari clustering. Berikut adalah *sample* data dari KODE HSL: [1, 0]. Setelah KODE HSL terbentuk penulis mengganti nilai dari KODE HSL. Nilai 0 diganti menjadi “SEHAT” sedangkan nilai 1 diganti menjadi “DEPRESI” menggunakan *code*:

```
df['KODE HSL'] = y_cluster = df['KODE HSL'].replace([0,1],['SEHAT','DEPRESI'])
```

Hasil dari penggantian nilai-nilai pada KODE HSL menjadikan data lebih mudah dimengerti dan dipahami, berikut contoh *sample* dari KODE HSL setelah nilai-nilai di dalamnya diganti agar lebih mudah dipahami. Contoh *sample*: [DEPRESI, SEHAT].

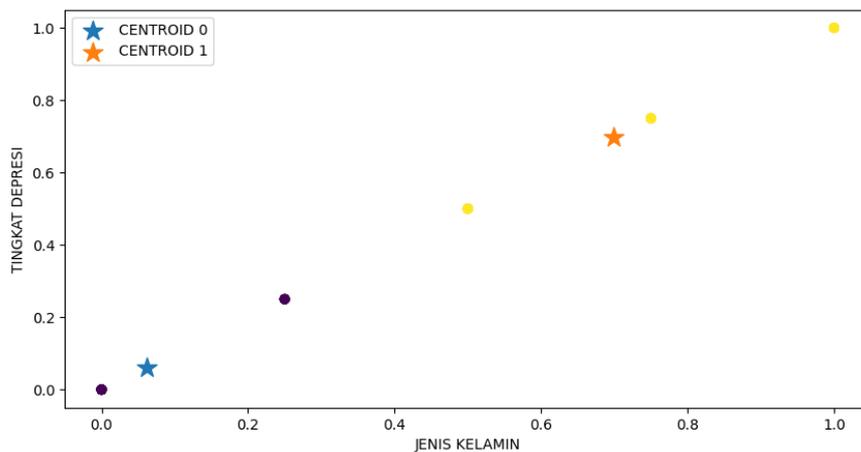
Crosstab of KODE JK and KODE HSL



Sumber: (Peneliti, 2024)

Gambar IV. 8
Crosstab Antara KODE JK dan KODE HSL

Proses clustering ini bertujuan untuk mengelompokkan responden berdasarkan tingkat depresi yang mereka alami seperti pada gambar diatas. Berikut adalah gambar centroid.



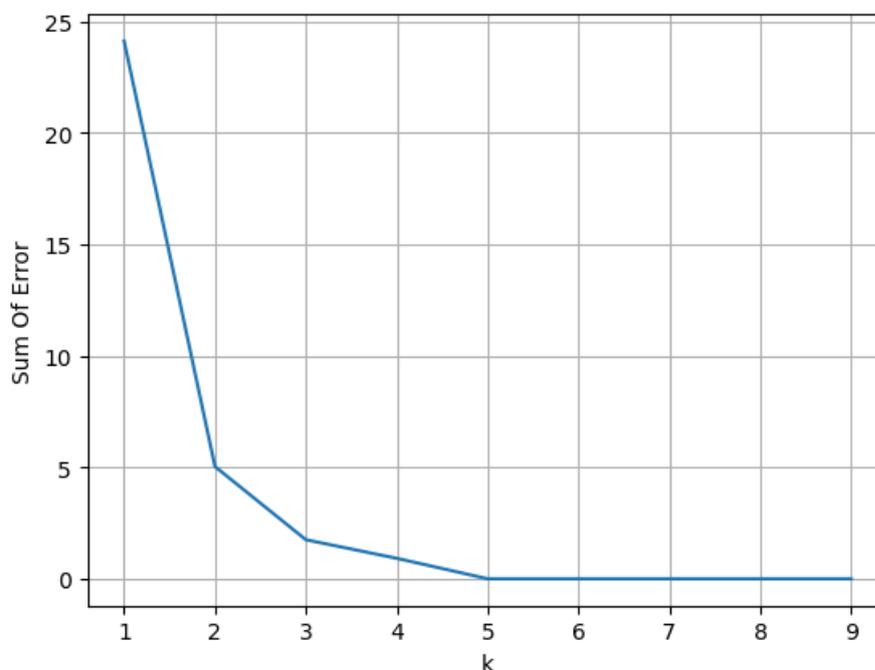
Sumber: (Peneliti, 2024)

Gambar IV. 9
Visualisasi Centroid 2 Cluster

4.1.8. Mencari K Terbaik menggunakan *Elbow Method*

Peneliti menggunakan *code* seperti dibawah ini untuk melakukan *Elbow Method*:

Method:



Sumber: (Peneliti, 2024)

Gambar IV. 10
Grafik *Elbow Method*

Pada gambar IV. 10, grafik yang dihasilkan akan menunjukkan perubahan nilai inerti berdasarkan jumlah cluster yang berbeda. Titik 4 di mana grafik menunjukkan perubahan tajam atau "siku" (elbow) menunjukkan jumlah cluster yang paling optimal yaitu 4 cluster. 4 cluster ini selanjutnya digunakan untuk clustering K-Means sehingga menghasilkan data cluster seperti sample berikut: 2, 1, 0 dan 3.

```

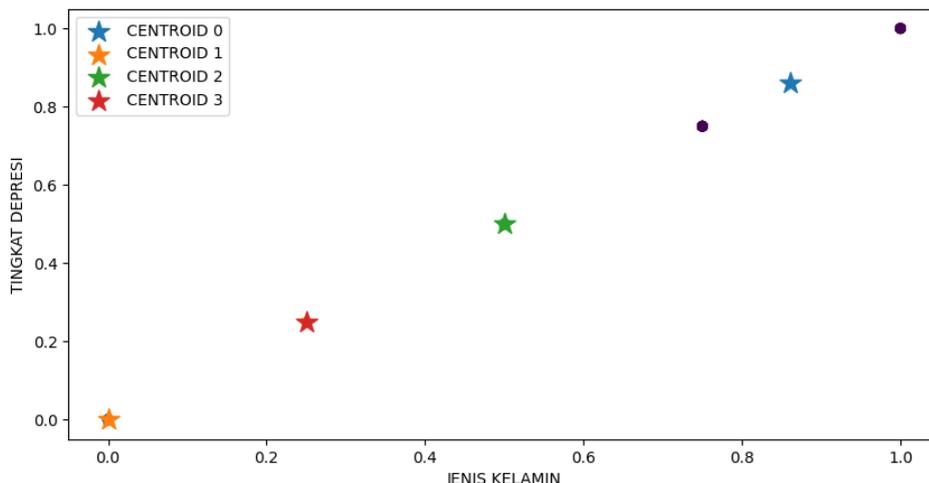
KMeans
KMeans(n_clusters=4)

```

Sumber: (Peneliti, 2024)

Gambar IV. 11
K-Means Dengan 4 Titik *Centroid*

Centroid 4 cluster:

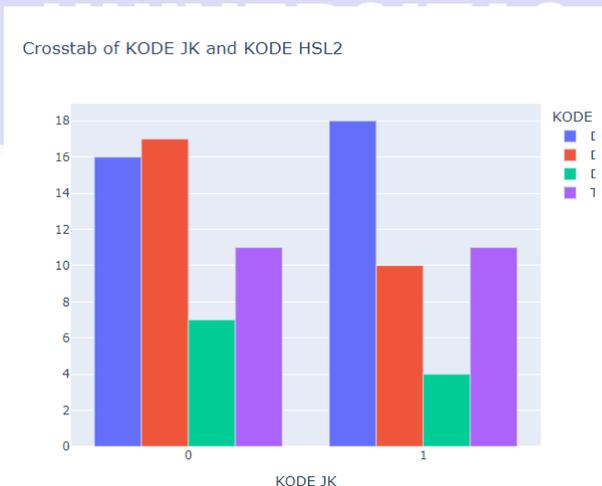


Sumber: (Peneliti, 2024)

Gambar IV. 12
Visualisasi K-Means 4 Cluster

4.1.9. Membuat *Dataset Baru Dengan Label HSL2*

Setelah menentukan jumlah *cluster* yang optimal menggunakan metode *Elbow*, peneliti akan membuat *dataset* baru yang mencakup label *cluster* untuk setiap data. Penulis juga mengubah kode Cluster menjadi ‘KODE HSL 2’. Contoh hasil *sample* data ialah seperti berikut: TIDAK DEPRESI, DEPRESI PARAH, DEPRESI RINGAN, DEPRESI SEDANG.



Sumber: (Peneliti, 2024)

Gambar IV. 13
Crosstab Antara KODE JK dan KODE HSL2 pada 4 Cluster

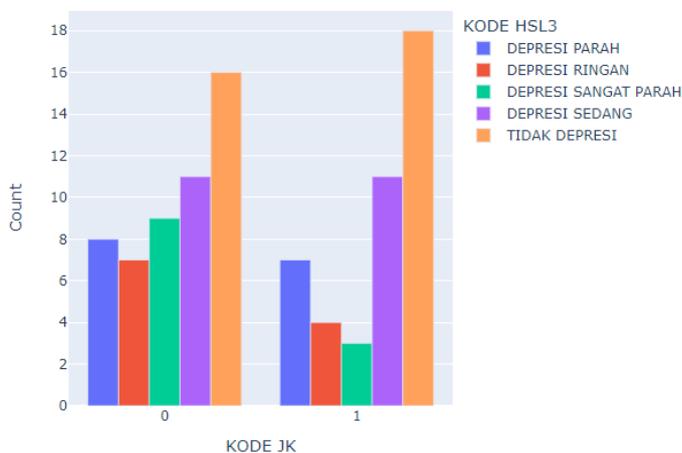
4.1.10. Optimasi Dengan PSO

Dalam konteks *K-Means*, *Particle swarm optimization* dapat digunakan untuk mencari nilai *centroid* yang optimal, sehingga proses *clustering* dapat menghasilkan kelompok-kelompok yang lebih terdefinisi dengan baik. Secara umum, proses optimasi *K-Means* menggunakan *Particle swarm optimization* dapat dengan langkah-langkah berikut:

- a). Inisialisasi populasi awal berupa posisi dan kecepatan partikel (calon *centroid*) secara acak.
- b). Hitung fungsi *fitness* untuk setiap partikel, misalnya dengan meminimalkan nilai *sum of squared errors* (SSE) atau *Silhouette score*.
- c). Perbarui posisi dan kecepatan partikel berdasarkan persamaan update pada *particle swarm optimization*.
- d). Lakukan iterasi hingga kriteria penghentian terpenuhi, misal jumlah iterasi maksimum atau konvergensi tercapai.
- e). Setelah proses optimasi, *centroid* akhir yang diperoleh dapat digunakan untuk proses *clustering K-Means* yang lebih optimal.

Dengan mengombinasikan *K-Means* dan *particle swarm optimization*, peneliti dapat memperoleh hasil *clustering* yang optimal dan lebih akurat dibandingkan hanya menggunakan *K-Means* biasa. Dalam berkerja, *particle swarm optimization* membagi cluster menjadi 5 sehingga didapat sample data seperti: 3, 2, 4, 0 dan 2. Selanjutnya penulis melakukan normalisasi data kembali agar data menjadi lebih mudah dipahami seperti sebelumnya. *Sample* sebelumnya berubah nilainya menjadi: DEPRESI SEDANG, TIDAK DEPRESI, DEPRESI SANGAT PARAH, DEPRESI RINGAN, dan TIDAK DEPRESI.

Crosstab of KODE JK and KODE HSL3



Sumber: (Peneliti, 2024)

Gambar IV. 14
Crosstab Antara KODE JK dan KODE HSL3 pada 5 Cluster

4.2. Hasil Pengujian

Dalam Menyusun penelitian ini, penulis berhasil melakukan pengujian algoritma *K-Means* dan menunjukkan bahwa data telah berhasil dikelompokkan ke dalam kluster-kluster yang memiliki karakteristik serupa. Penulis mendapatkan perbedaan jumlah kluster dalam melakukan clustering dengan rincian sebagai berikut. Jumlah kluster didapat menggunakan titik *centroid*. *Centroid* sendiri adalah titik pusat atau titik paling tengah dari kumpulan data.

4.2.1. Dengan Jumlah *Cluster* Acak

Pada saat menggunakan kluster dengan jumlah yang acak, penulis menggunakan 2 kluster. Saat menggunakan 2 kluster dalam menjalankan model *clustrering* ini, penulis mendapatkan nilai *centroid* sebagai berikut: dapat dilihat pada gambar IV. 15.

```
[ ] Kmeans.cluster_centers_
array([[0.06111111, 0.06111111],
       [0.69897959, 0.69897959]])
```

Sumber: (Peneliti, 2024)

Gambar IV. 15
Koordinat 2 Titik *Centroid*

Centroid pertama memiliki koordinat (0.06111111, 0.06111111).

Centroid kedua memiliki koordinat (0.69897959, 0.69897959).

1. *Inertia*

Nilai *inertia* berkerja dengan melihat nilai yang paling kecil, yang mana semakin kecil nilai menunjukkan jumlah klaster yang lebih optimal. Ketika menggunakan 2 klater ini, penulis mendapatkan nilai *inertia* sebagai berikut.

```
inertias
[24.122340425531927,
 5.033786848072566,
 1.7500000000000004,
 0.9166666666666667,
 2.2071157162677723e-30,
 2.2071157162677723e-30,
 3.0945842346414106e-30,
 2.2071157162677723e-30,
 2.2071157162677723e-30]
```

Sumber: (Peneliti, 2024)

Gambar IV. 16
Nilai *Inertia* Pada Tiap *Cluster*

Pada gambar IV. 16 menunjukkan, penulis mendapatkan nilai *inertia* untuk setiap *cluster* sebagai berikut:

- a). 24.122340425531927 - *Inertia* untuk *cluster* pertama.
- b). 5.033786848072566 - *Inertia* untuk *cluster* kedua.

- c). 1.7500000000000004 - *Inertia* untuk *cluster* ketiga.
- d). 0.9166666666666667 - *Inertia* untuk *cluster* keempat.
- e). 2.2071157162677723e-30 - *Inertia* untuk *cluster* kelima.
- f). 2.2071157162677723e-30 - *Inertia* untuk *cluster* keenam.
- g). 3.0945842346414106e-30 - *Inertia* untuk *cluster* ketujuh.
- h). 2.2071157162677723e-30 - *Inertia* untuk *cluster* kedelapan.
- i). 2.2071157162677723e-30 - *Inertia* untuk *cluster* kesembilan.

Dari hasil nilai diatas penulis mendapati bahwa nilai inertia yang sangat kecil seperti pada cluster 5, 6, 8, dan 9 mengindikasikan bahwa *cluster-cluster* tersebut adalah yang paling optimal atau terbaik di antara semua cluster yang terbentuk.

2. Kualitas Model Algoritma *K-Means* dengan *Cluster* acak Terhadap Data

Dalam mengukur sejauh mana model algoritma *K-Means* ini berjalan terhadap data yang penulis gunakan, penulis menggunakan pengukuran menggunakan *Silhouette Score*, *Calinski-Harabasz Score*, dan *Davies-Bouldin Score*.

Silhouette Score: 0.73
 Calinski-Harabasz Score: 348.87
 Davies-Bouldin Score: 0.42

Sumber: (Peneliti, 2024)

Gambar IV. 17
 Skor Validasi 2 *Cluster*

Berdasarkan gambar IV. 17 diketahui,

a). *Silhouette Score*, Nilai 0.73 untuk *Silhouette Score* menunjukkan bahwa secara keseluruhan, *clustering* yang dilakukan memiliki kualitas yang baik. Ini berarti bahwa *cluster-cluster* yang terbentuk cukup terpisah dan kompak.

b). *Calinski-Harabasz Score*, Nilai 348.87 untuk *Calinski-Harabasz Score* mengindikasikan bahwa *cluster-cluster* yang terbentuk memiliki kohesi internal yang tinggi dan pemisahan antar cluster yang baik. Ini merupakan indikasi *clustering* yang efektif.

c). *Davies-Bouldin Score*, Nilai 0.42 untuk *Davies-Bouldin Score* menunjukkan bahwa *cluster-cluster* yang terbentuk memiliki kehomogenan internal yang baik dan pemisahan antar cluster yang jelas. Ini juga merupakan indikasi *clustering* yang efektif.

Secara keseluruhan, ketiga skor evaluasi ini mengindikasikan bahwa proses *clustering* yang dilakukan menghasilkan *cluster-cluster* yang kompak dan terpisah dengan baik. Ini menunjukkan kualitas *clustering* yang cukup tinggi.

3. Hasil

Dalam menjalankan penelitian ini. Di titik penggunaan 2 klaster, penulis dapat merumuskan hasil *Clustering* ini hasil pada Gambar IV. 18.

KODE HSL	DEPRESI	SEHAT
KODE JK		
0	28	23
1	21	22

Sumber: (Peneliti, 2024)

Gambar IV. 18
Crosstab KODE HSL dan KODE JK

Tabel pada Gambar IV. 18 ini menunjukkan data frekuensi dari dua kondisi kesehatan mental (DEPRESI dan SEHAT) berdasarkan jenis kelamin (JK).

Keterangan:

KODE JK: 0 untuk Perempuan, 1 untuk Laki-laki

KODE HSL: Kondisi kesehatan mental, dengan 'DEPRESI' dan 'SEHAT'

Dari tabel tersebut, dapat dilihat bahwa, data frekuensi dari dua kondisi kesehatan mental, yaitu DEPRESI dan SEHAT, berdasarkan jenis kelamin (JK). Dari tabel tersebut, dapat dilihat bahwa untuk Perempuan (KODE JK = 0), terdapat 28 orang yang mengalami DEPRESI dan 23 orang dalam kondisi SEHAT. Sementara untuk Laki-laki (KODE JK = 1), terdapat 21 orang yang mengalami DEPRESI dan 22 orang dalam kondisi SEHAT.

Secara keseluruhan, data ini mengindikasikan bahwa jumlah perempuan yang mengalami DEPRESI (28) lebih banyak daripada laki-laki (21). Di sisi lain, jumlah Perempuan yang dalam kondisi SEHAT (23) sedikit lebih banyak daripada Laki-laki (22).

4.2.2. Penggunaan K Terbaik

Sebagaimana seharusnya *K-Means* berjalan. Penulis tentu saja menggunakan cara yang umum untuk menjalankan *K-Means*, yaitu menggunakan *Elbow Method* untuk menentukan nilai K terbaik, berikut ialah hasil dari *Clustering* menggunakan nilai K terbaik (4 klaster).

1. Hasil

KODE HSL2 KODE JK	DEPRESI PARAH	DEPRESI RINGAN	DEPRESI SEDANG	TIDAK DEPRESI
0	16	17	7	11
1	18	10	4	11

Sumber: (Peneliti, 2024)

Gambar IV. 19
Crosstab Antara KODE HSL2 dan KODE JK

Pada Gambar IV. 19 menunjukkan data frekuensi dari empat kondisi kesehatan mental, yaitu DEPRESI PARAH, DEPRESI RINGAN, DEPRESI SEDANG, dan TIDAK DEPRESI, berdasarkan jenis kelamin (JK).

Keterangan:

KODE JK: 0 untuk Perempuan, 1 untuk Laki-laki

KODE HSL2: Kondisi kesehatan mental, dengan 'TIDAK DEPRESI', 'DEPRESI RINGAN', 'DEPRESI SEDANG', dan 'DEPRESI PARAH'

Untuk Perempuan (KODE JK = 0), data menunjukkan bahwa 16 orang mengalami DEPRESI PARAH, 17 orang mengalami DEPRESI RINGAN, 7 orang mengalami DEPRESI SEDANG, dan 11 orang tidak mengalami DEPRESI. Sementara untuk Laki-laki (KODE JK = 1), 18 orang mengalami DEPRESI PARAH, 10 orang mengalami DEPRESI RINGAN, 4 orang mengalami DEPRESI SEDANG, dan 11 orang tidak mengalami DEPRESI.

Dari data ini, dapat diamati bahwa jumlah Laki-laki yang mengalami DEPRESI PARAH (18) lebih banyak daripada Perempuan (16). Namun, untuk DEPRESI RINGAN, jumlah Perempuan (17) lebih banyak daripada Laki-laki (10). Untuk kategori DEPRESI SEDANG, Perempuan (7) juga memiliki kuantitas yang lebih tinggi daripada Laki-laki (4). Untuk kondisi TIDAK DEPRESI, jumlah Perempuan dan Laki-laki sama, yaitu 11 orang.

2. Kualitas Model Algoritma *K-Means* dengan K Terbaik Terhadap Data

Sebagai pembandingan dengan hasil dari klaster sebelumnya, peneliti juga mengukur sejauh mana model Algoritma K-Means dengan K-terbaik ini berjalan.

Silhouette Score: 0.89
Calinski-Harabasz Score: 838.40
Davies-Bouldin Score: 0.26

Sumber: (Peneliti, 2024)

Gambar IV. 20
Skor Validasi Cluster Dengan K Terbaik

Berdasarkan gambar IV. 20,

- a). *Silhouette Score*, Nilai 0.89 untuk *Silhouette Score* Skor 0.89 mengungkapkan bahwa klasterisasi yang telah dilakukan memiliki kualitas yang luar biasa baik.
- b). *Calinski-Harabasz Score*, Nilai 838,40 untuk *Calinski-Harabasz* Skor 838.40 mengindikasikan bahwa klasterisasi yang dilakukan memiliki kualitas yang sangat baik.
- c). *Davies-Bouldin Score*, Nilai 0.26 untuk *Davies-Bouldin* Skor 0.26 mengindikasikan bahwa proses klasterisasi yang dilaksanakan memiliki tingkat kualitas yang sangat tinggi.

Secara keseluruhan, ketiga skor yang ditampilkan pada hasil dari proses klasterisasi yang diterapkan menunjukkan kualitas yang sangat baik. Hal ini berarti data telah dikelompokkan dengan sangat efektif berdasarkan metrik-metrik evaluasi klasterisasi K-Means yang penulis gunakan.

4.2.3. Optimasi K-Means Clustering Menggunakan Particle Swarm Optimization

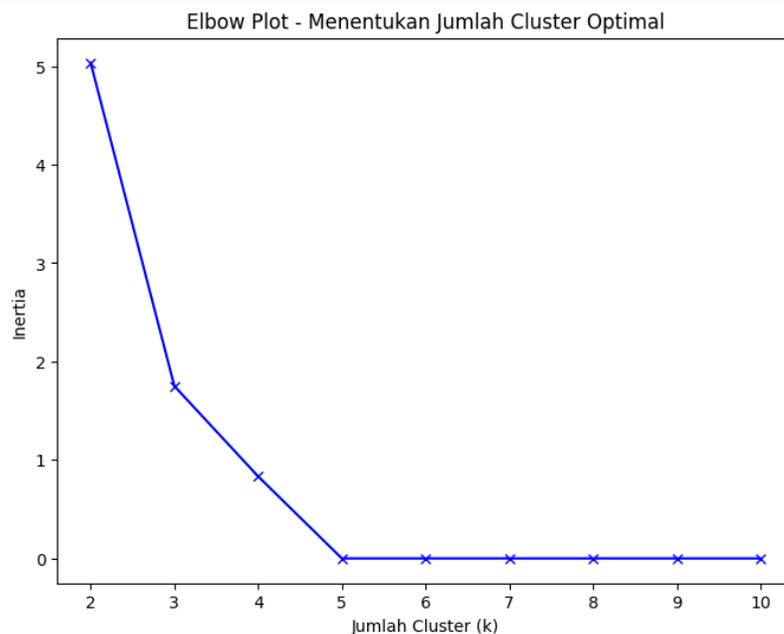
Optimasi *K-Means Clustering* menggunakan *Particle Swarm Optimization* memiliki fungsi untuk mencari jumlah cluster optimal. Fungsi ini menerima input berupa data X , jumlah *cluster* $n_clusters$, jumlah partikel $n_particles$, dan jumlah iterasi $n_iterations$. Berikut yang terjadi

- a). Inisialisasi posisi dan kecepatan partikel secara acak.
- b). Hitung *fitness* (*error K-Means*) untuk setiap partikel dan *update*
- c). *best_positions*, *best_fitness*, *global_best*, dan *global_best_fitness*.
- d). Perbarui posisi dan kecepatan partikel menggunakan persamaan update PSO.
- e). Kembalikan *global_best* sebagai pusat cluster optimal.

Selanjutnya, cari jumlah *cluster* optimal: *loop* dilakukan untuk mencoba berbagai jumlah *cluster* (2 hingga 10). Untuk setiap jumlah *cluster*, fungsi *pso_kmeans* dipanggil untuk mencari pusat *cluster* optimal, kemudian algoritma *K-Means* dijalankan dengan pusat cluster tersebut. Nilai *inertia* (*sum of squared distances*) disimpan dalam *inertias*.

1. Elbow Method

Visualisasi *Elbow Plot*: Grafik *Elbow Plot* dibuat untuk memvisualisasikan nilai *inertia* terhadap jumlah *cluster*. Titik "siku" pada grafik menunjukkan jumlah *cluster* optimal.



Sumber: (Peneliti, 2024)

Gambar IV. 21
Visualisasi *Elbow Method*

Grafik ini menampilkan *Elbow Plot*, yang digunakan untuk menentukan jumlah kluster optimal dalam analisis klusterisasi. Pada grafik ini, sumbu x menunjukkan jumlah kluster (k), sementara sumbu y menunjukkan metrik Inertia. Semakin kecil nilai *Inertia*, semakin homogen data di dalam kluster tersebut. Berdasarkan grafik, dapat dilihat bahwa pada saat jumlah kluster (k) mencapai 5, terjadi penurunan yang cukup tajam pada nilai *Inertia*. Setelah itu, penurunan Inertia menjadi lebih landai. Titik ini sering disebut sebagai "siku" (*elbow*) pada grafik, yang menunjukkan jumlah kluster optimal.

Oleh karena itu, berdasarkan interpretasi *Elbow Plot* ini, jumlah kluster optimal untuk mengelompokkan data adalah 5 kluster. Ini berarti data dapat dibagi menjadi 5 kelompok dengan tingkat homogenitas yang baik di dalam masing-masing kluster.

2. Klusterisasi Final dengan K-Means

Jumlah *cluster* optimal ditentukan berdasarkan nilai indeks yang memberikan nilai *inertia* minimum. Algoritma *K-Means* kemudian dijalankan kembali dengan jumlah *cluster* optimal, dan hasil klasterisasi disimpan dalam kolom '*Cluster*' pada dataframe *df*. Berikut adalah nilai centroid yang penulis dapatkan.

```
Centroids:
[[2.50000000e-01 2.50000000e-01]
 [7.50000000e-01 7.50000000e-01]
 [1.66533454e-16 1.66533454e-16]
 [5.00000000e-01 5.00000000e-01]
 [1.00000000e+00 1.00000000e+00]]
```

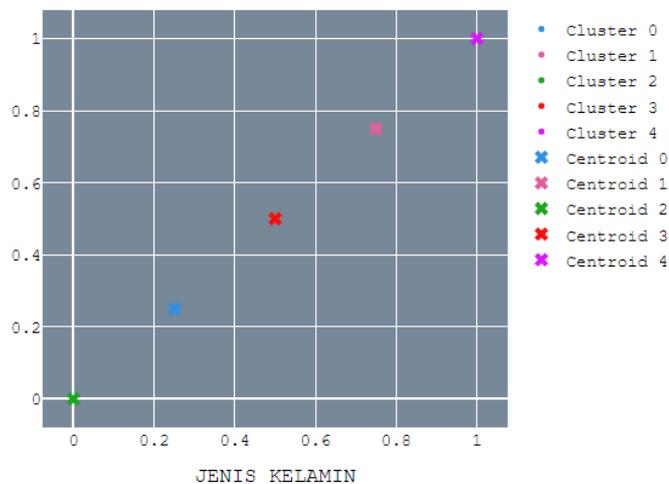
Sumber: (Peneliti, 2024)

Gambar IV. 22
Koordinat 5 Titik *Centroid*

Berdasarkan gambar terdapat 5 *centroid*:

- a). [2.50000000e-01, 2.50000000e-01]
- b). [7.50000000e-01, 7.50000000e-01]
- c). [1.66533454e-16, 1.66533454e-16]
- d). [5.00000000e-01, 5.00000000e-01]
- e). [1.00000000e+00, 1.00000000e+00]

Visualisasi Cluster Data dan Centroid



Sumber: (Peneliti, 2024)

Gambar IV. 23
Cluster Data dan 5 Centroid

Pada gambar IV. 23, kita dapat melihat bahwa seluruh data terpusat tepat berada di titik centroid pada klasternya masing-masing. Ini menunjukkan, *algoritma K-Means clustering* yang dioptimasi memakai *Particle Swarm Optimization* telah berhasil mengidentifikasi 5 kelompok data yang diberikan dengan sangat baik. Ini menunjukkan bahwa data dalam setiap *cluster* sangat kompak dan homogen, dengan jarak antar anggota *cluster* yang sangat dekat dengan titik pusatnya. Secara keseluruhan, jika seluruh data berada tepat di titik *centroid* pada masing-masing klasternya, maka dapat disimpulkan bahwa algoritma *clustering* telah berhasil memisahkan data dengan baik dan membentuk kelompok-kelompok yang sangat terdefinisi.

Situasi seperti ini terjadi karena penulis telah melabeli data dengan baik dan struktur model Algoritma K-Means dan PSO memiliki pola distribusi yang jelas. Karena hal ini dapat memudahkan proses pengelompokan data dan memungkinkan analisis lanjutan yang lebih akurat.

3. Hasil

KODE HSL3	DEPRESI PARAH	DEPRESI RINGAN	DEPRESI SANGAT PARAH
0	8	7	9
1	7	4	3

KODE HSL3	DEPRESI SEDANG	TIDAK DEPRESI
0	11	16
1	11	18

Sumber: (Peneliti, 2024)

Gambar IV. 24
Crosstab Antara KODE JK dan KODE HSL3

Pada gambar IV. 24 tersebut, terdapat dua kolom utama, yaitu KODE JK (Kode Jenis Kelamin) dan KODE HSL3 (Kode Hasil). Kode JK 0 mewakili Perempuan, sedangkan KODE JK 1 mewakili Laki-laki. Untuk KODE HSL3, nilainya menunjukkan jumlah individu yang termasuk ke dalam kategori depresi tertentu, baik depresi parah, depresi ringan, depresi sangat parah, depresi sedang, maupun tidak depresi.

Pada Perempuan (KODE JK 0): Terdapat 8 individu yang mengalami depresi parah, 7 individu yang mengalami depresi ringan, 9 individu yang mengalami depresi sangat parah, 11 individu yang mengalami depresi sedang, dan 16 individu yang tidak mengalami depresi.

Pada Laki-laki (KODE JK 1): Terdapat 7 individu yang mengalami depresi parah, 4 individu yang mengalami depresi ringan, 3 individu yang mengalami depresi sangat parah, 11 individu yang mengalami depresi sedang, dan 18 individu yang tidak mengalami depresi.

Secara keseluruhan berdasarkan analisis, kesimpulannya adalah bahwa, pada kedua kelompok Perempuan dan Laki-laki, populasi orang yang bebas dari depresi

memiliki angka yang lebih tinggi dibandingkan dengan populasi yang menderita depresi dalam berbagai tingkat. Selain itu, jumlah individu Perempuan yang mengalami depresi lebih banyak daripada Laki-laki, terutama untuk kategori depresi ringan, depresi parah, dan depresi sangat parah.

4. Kualitas *K-Means* menggunakan Optimasi *Particle Swarm Optimization*

- a). *Silhouette Score*, Nilai Skor 1.00 menunjukkan bahwa *clustering* yang dilakukan sangat baik, dengan data-data dalam *cluster* sangat kompak dan terpisah jelas antar *cluster*. Hal ini mengindikasikan bahwa algoritma *clustering* telah berhasil memisahkan data menjadi kelompok-kelompok yang sangat homogen dan *distinct*.
- b). *Calinski-Harabasz Score*, Nilai Skor 1.00 menunjukkan bahwa *cluster-cluster* yang terbentuk sangat terpisah dengan jelas, dengan varians antar *cluster* jauh lebih besar dibandingkan varians dalam *cluster*. Hal ini mengkonfirmasi bahwa *clustering* telah dilakukan dengan sangat baik, menghasilkan kelompok-kelompok data yang sangat terdefinisi.
- c). *Davies-Bouldin Score* 0.00 menyatakan bahwa tidak ada satupun kelompok yang memiliki kesamaan dengan kelompok lainnya. Hal ini menunjukkan bahwa kelompok-kelompok yang terbentuk benar-benar terpisah satu sama lain, tanpa adanya *overlap* atau persamaan antar kumpulan data.

4.2.4. Perbandingan Nilai-Nilai Skor Validasi *Clustering*

1. Dengan Jumlah Klaster Acak:

- a). *Silhouette Score*: 0.73

b). *Calinski-Harabasz Score*: 348.87

c). *Davies-Bouldin Score*: 0.42

Nilai-nilai ini menunjukkan bahwa *clustering* dengan jumlah kluster acak memiliki kualitas yang cukup baik, namun belum optimal.

2. Dengan K-Terbaik:

a). *Silhouette Score*: 0.89

b). *Calinski-Harabasz Score*: 838.40

c). *Davies-Bouldin Score*: 0.26

Nilai-nilai ini menunjukkan bahwa *clustering* dengan jumlah kluster terbaik kualitasnya unggul dibandingkan dengan jumlah kluster yang dibentuk secara acak. *Clustering* ini sudah mendekati optimal.

3. Setelah Optimasi *Particle Swarm Optimization*:

a). *Silhouette Score*: 1.00

b). *Calinski-Harabasz Score*: 1.00

c). *Davies-Bouldin Score*: 0.00

Nilai-nilai ini menunjukkan bahwa *clustering* setelah optimasi menggunakan PSO memiliki kualitas yang sangat baik. *Clustering* ini sudah optimal, dengan data-data terkelompok secara kompak dan terpisah dengan jelas.

Dapat disimpulkan bahwa optimasi K-Means menggunakan PSO berjalan sangat baik terhadap data, ini dibuktikan dengan skor-skor validasi *clustering* setelah optimasi dari PSO (*Silhouette* 1.00, *Calinski-Harabasz* 1.00, *Davies-Bouldin* 0.00)

menunjukkan bahwa clustering yang dilakukan memiliki kualitas yang sangat baik. Data-data telah dikelompokkan secara optimal, dengan cluster-cluster yang sangat kompak, terpisah jelas satu sama lain, dan membentuk kelompok-kelompok yang sangat terdefinisi.



BAB V

PENUTUP

5.1. Kesimpulan

1. Bagaimana metode *K-Means Clustering* berkerja terhadap data depresi yang ada berdasarkan Instrumen DASS-42?

K-Means Clustering berkerja dengan sangat baik, namun jumlah *cluster* belum masuk ke tingkat yang paling optimal, ini dibuktikan dengan skor validasi *clustering* dengan jumlah klaster acak: *Silhouette Score*: 0.73, *Calinski-Harabasz Score*: 348.87, dan *Davies-Bouldin Score*: 0.42. Sedangkan, *Clustering* dengan jumlah klaster terbaik: *Silhouette Score*: 0.89, *Calinski-Harabasz Score*: 838.40, dan *Davies-Bouldin Score*: 0.26. Kualitas lebih baik dibandingkan dengan jumlah klaster acak, mendekati optimal.

2. Apakah optimasi K-Means dengan *Particle Swarm Optimization* lebih baik akurasiya ketika model dijalankan?

Dapat disimpulkan bahwa optimasi *K-Means* menggunakan *Particle Swarm Optimization* (PSO) menghasilkan *clustering* yang lebih optimal dibandingkan dengan *clustering* menggunakan jumlah klaster acak maupun jumlah klaster terbaik. Berikut perbandingan kualitas *clustering* dari ketiga pendekatan tersebut:

Clustering Setelah Optimasi PSO: *Silhouette Score*: 1.00, *Calinski-Harabasz Score*: 1.00, dan *Davies-Bouldin Score*: 0.00. Kualitas sangat baik, *clustering* sudah optimal dengan data-data terkelompok secara kompak dan terpisah dengan jelas.

Berdasarkan hasil penelitian ini, dapat disimpulkan bahwa optimasi *K-Means* menggunakan *Particle Swarm Optimization* (PSO) menghasilkan akurasi model yang lebih baik dibandingkan dengan *clustering* menggunakan jumlah kluster acak maupun jumlah kluster terbaik.

Hal ini dibuktikan dengan nilai-nilai evaluasi clustering yang lebih optimal, seperti *Silhouette Score*, *Calinski-Harabasz Score*, dan *Davies-Bouldin Score* yang menunjukkan hasil terbaik pada *clustering* setelah optimasi PSO. Oleh karena itu, dapat disimpulkan bahwa optimasi *K-Means* oleh *Particle Swarm Optimization* (PSO) memang menghasilkan Algoritma yang lebih akurat dan optimal dibandingkan dengan pendekatan lainnya.

3. Berapa banyak orang yang memasuki kategori depresi berat?

Dari hasil penelitian penulis, penulis mendapati jumlah individu yang masuk ke dalam kategori depresi berat yaitu yaitu, pada perempuan terdapat 8 individu yang mengalami depresi parah, dan 9 individu yang mengalami depresi sangat parah, Sementara itu, pada laki-laki, terdapat 7 individu yang mengalami depresi parah, dan 3 individu yang mengalami depresi sangat parah, Jika dijumlahkan, total perempuan yang mengalami depresi parah adalah 17 individu, sedangkan total laki-laki yang mengalami depresi adalah 10 individu. Dengan demikian, jumlah keseluruhan individu yang mengalami depresi adalah 27 individu.

4. Sedekat apa sebenarnya kehidupan manusia dengan depresi?

Depresi memang tampaknya semakin dekat dengan kehidupan manusia modern. Hal ini dapat dilihat dari beberapa faktor. Pertama, prevalensi depresi yang cukup tinggi. Studi ini menunjukkan sekitar 52% dari 94 narasumber mengalami

depresi dengan tingkat keparahan yang signifikan. Angka ini mengindikasikan bahwa depresi bukanlah kondisi yang langka, melainkan cukup umum terjadi di masyarakat. Selain itu, tingkat keparahan depresi juga beragam, tidak hanya depresi sedang tetapi juga depresi parah ditemukan pada sejumlah orang. Hal ini menunjukkan bahwa gangguan mental ini dapat menyerang individu dengan berbagai tingkat keparahan, mulai dari yang ringan hingga yang sangat parah.

5.2. Saran

Berdasarkan penelitian penulis yang berjudul "*Optimasi K-Means Clustering Dengan Particle Swarm Optimization Untuk Pengelompokan Depresi Berdasarkan DASS-42*", penulis menyarankan untuk penelitian selanjutnya agar menggunakan metode K-Means yang dioptimalisasi dengan algoritma lain serta menerapkannya pada dataset yang lebih beragam dan terperinci. Hal ini diharapkan dapat meningkatkan efektivitas dan akurasi dari proses pengelompokan data menggunakan algoritma K-Means.

DAFTAR PUSTAKA

- Adi Zayd Bintang, A. M. M., Jember, K., Bintang, A. Z., Mandagi, A. M., Keselamatan, D., Kerja, K., & Masyarakat, K. (2021). Kejadian Depresi Pada Remaja Menurut Dukungan Sosial Di Kabupaten Jember. *JOURNAL OF COMMUNITY MENTAL HEALTH AND PUBLIC POLICY*, Vol. 3 No. 2(Mental Health), 1–10.
<http://cmhp.lenterakaji.org/index.php/cmhp>
- Aqsa Mu'azzaroh, A., & Cahyanti, I. Y. (2023). PENGARUH KEKERASASN SEKSUAL YANG DIFASILITASI TEKNOLOGI TERHADAP DEPRESI PADA WANITA DEWASA AWAL PENGGUNA MEDIA SOSIAL. *JURNAL FUSION*, 3(08).
<https://doi.org/10.54543/fusion.v3i05.332>
- Citra Pengenalan Aksara Bugis Kurniati, S., & Reza Wardana, R. (2020). Penerapan Algoritma Particle Swarm Optimization pada Segmentasi Citra Pengenalan Aksara Bugis. In *Jurnal Pengembangan Sistem Informasi dan Informatika* (Vol. 1, Issue 3). <http://journal.jis-institute.org/index.php/jpsii/index>
- Fauziah, R., & Purnamasari, A. I. (2023). Implementasi Algoritma K-Means pada Kasus Kekerasan Anak dan Perempuan Berdasarkan Usia. *Hello World Jurnal Ilmu Komputer*, 2(1), 34–41. <https://doi.org/10.56211/helloworld.v2i1.232>
- Hasibuan, C. K., & Yahfizham, Y. (2023). Analisis Pembelajaran Algoritma Pemrograman. *Bahasa Dan Matematika*, 1(5), 274–285. <https://doi.org/10.61132/arjuna.v1i5.337>
- Hayatin, N., Marthasari, G. I., & Nuraini, L. (2020). Optimization of Sentiment Analysis for Indonesian Presidential Election using Naïve Bayes and Particle Swarm Optimization. *Jurnal Online Informatika*, 5(1), 81–88. <https://doi.org/10.15575/join.v5i1.558>
- Hendrastuty, N. (2024). Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Dalam Evaluasi Hasil Pembelajaran Siswa. *JURNAL ILMIAH INFORMATIKA DAN ILMU KOMPUTER(JIMA-ILKOM)*. <https://doi.org/10.58602/jima-ilkom.v3i1.26>
- Homepage, J., Herviany, M., Putri Delima, S., Nurhidayah, T., Studi Sistem Informasi, P., & Sains dan Teknologi UIN Sultan Syarif Kasim Riau Jl Soebrantas, F. H. (2021). Comparison of K-Means and K-Medoids Algorithms for Grouping Landslide Prone Areas in West Java Province Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokan Daerah Rawan Tanah Longsor di Provinsi Jawa Barat. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 1, 34–40.
- Karadoğan, S. K., & Altay, B. U. (2024). Dermatology Quality of Life and Depression, Anxiety, and Stress Scale-42 in Scabies Patients. *Dermatology Practical and Conceptual*, 14(2). <https://doi.org/10.5826/dpc.1402a112>
- Lashiyanti, A. R., Rasyid Munthe, I., Nasution, F. A., & Korespondensi, E. P. (2023). Optimisasi Klasterisasi Nilai Ujian Nasional dengan Pendekatan Algoritma K-Means, Elbow, dan Silhouette. *Jurnal Ilmu Komputer Dan Sistem Informasi (JIKOMSI)*, 6(1), 14–20.

- Pasaribu, D. F., Damanik, I. S., Irawan, E., Suhada, & Tambunan, H. S. (2021). Memanfaatkan Algoritma K-Means Dalam Memetakan Potensi Hasil Produksi Kelapa Sawit PTPN IV Marihat. *BIOS : Jurnal Teknologi Informasi Dan Rekayasa Komputer*, 2(1), 11–20. <https://doi.org/10.37148/bios.v2i1.17>
- Senjaya, S., Rahmawati, S., Aisyah, S., Andriani, N., Heryani, H., Nur Apni, S., Ilham Mahardika Jati, B., Sintia, Y., Nurhafifah, S., Keperawatan Jiwa, D., Keperawatan, F., Padjadjaran, U., Raya Bandung-Sumedang, J. K., Barat, J., Keperawatan Anak, D., & Sarjana Keperawatan, P. (2022). GAMBARAN TINGKAT DEPRESI PADA MAHASISWA DI MASA PANDEMI: NARRATIVE REVIEW. <https://journal.ppnijateng.org/index.php/jikj>
- Sikana, A. M., & Wijayanto, A. W. (2021). Analisis Perbandingan Pengelompokan Indeks Pembangunan Manusia Indonesia Tahun 2019 dengan Metode Partitioning dan Hierarchical Clustering. *Jurnal Ilmu Komputer*, 14(2), 66. <https://doi.org/10.24843/jik.2021.v14.i02.p01>
- Tarigan, D. A. (2023). Optimization of the K-Means Clustering Algorithm Using Davies Bouldin Index in Iris Data Classification. *Media Online*, 4(1), 545–552. <https://doi.org/10.30865/klik.v4i1.964>
- Virgo, I., Defit, S., & Yuhandri, Y. (2020). Klasterisasi Tingkat Kehadiran Dosen Menggunakan Algoritma K-Means Clustering. *Jurnal Sestim Informasi Dan Teknologi*, 23–28. <https://doi.org/10.37034/jsisfotek.v2i1.17>



UNIVERSITAS

DAFTAR RIWAYAT HIDUP

DAFTAR RIWAYAT HIDUP

I. Biodata Mahasiswa

NIM : 15200056
Nama Lengkap : Mohamad Abdul Yusuf
Tempat/Tanggal Lahir : Jakarta, 15200056
Alamat lengkap : Jl. Praji Rt.05/001 Kelapa Dua Wetan,
Ciracas, Jakarta Timur

II. Pendidikan

a. Formal

1. SDN Kelapa Dua Wetan 05 (2013)
2. SMPN 147 Jakarta (2016)
3. MAN 14 Jakarta (2019)
4. Universitas Bina Sarama Informatika (2020-Saat Ini)

b. Non Formal

1. PPKPI Pasar Rebo (2024-2024)

III. Riwayat Pengalaman Berorganisasi / Pekerjaan

1. Anggota Dewan Pramuka MAN 14 Jakarta (2018-2022)
2. Anggota Saka Wira Kartika (2018-2019)
3. PT. Adipurna Mranata Jaya (2023-2023)



Jakarta, 28 Juni 2024

Mohamad Abdul Yusuf

LEMBAR KONSULTAN

	LEMBAR KONSULTAN
	UNIVERSITAS BINA SARANA INFORMATIKA

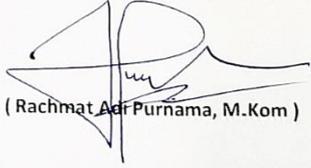
NIM : 15200056
Nama Lengkap : Mohamad Abdul Yusuf
Dosen Pembimbing I : Rachmat Adi Purnama, M.Kom
Judul Skripsi : Optimasi K-Means Clustering Dengan Particle Swarm Optimization Untuk Pengelompokan Depresi Berdasarkan Dass-42

No	Tanggal Bimbingan	Pokok Bahasa	Paraf Dosen Pembimbing I
1.	1 April 2024	Bimbingan Perdana dan Pengajuan Judul	
2.	19 April 2024	Penyerahan Bab 1	
3.	29 April 2024	Revisi Bab 1 dan Penyerahan Bab 2	
4.	17 Mei 2024	Penyerahan Bab 3	
5.	03 Juni 2024	Revisi Bab 3 dan Penyerahan Bab 4	
6.	13 Juni 2024	Revisi Bab 4 dan Penyerahan Bab 5	
7.	19 Juni 2024	Revisi Bab 5	
8.	27 Juni 2024	Penyerahan Bab dan Tugas Akhir Revisi Skripsi	

Catatan Untuk Dosen Pembimbing Bimbingan Skripsi

Dimulai pada tanggal : 1 April 2024
Diakhiri pada tanggal : 27 Juni
Jumlah pertemuan bimbingan: 8 (Delapan)

Disetujui oleh,
Dosen Pembimbing I


(Rachmat Adi Purnama, M.Kom)

SURAT PERNYATAAN KEBENARAN/KEABSAHAN DATA

HASIL RISET UNTUK KARYA ILMIAH

SURAT PERNYATAAN KEBENARAN/KEABSAHAN DATA HASIL RISET UNTUK KARYA ILMIAH

Yang bertanda tangan di bawah ini, saya:

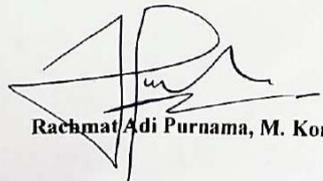
Nama : Mohamad Abdul Yusuf
NIM : 15200056
Jenjang : Sarjana (S1)
Program Studi : Informatika
Fakultas : Teknik dan Informatika
Perguruan Tinggi : Universitas Bina Sarana Informatika

Dengan ini menyatakan bahwa data dan atau informasi yang saya gunakan dalam penulisan karya ilmiah Penulis dengan judul "**Optimasi K-Means Clustering Dengan Particle Swarm Optimization Untuk Pengelompokan Depresi Berdasarkan DASS-42**" merupakan data dan atau informasi yang saya peroleh melalui hasil penelitian sendiri dan tidak didasarkan pada data atau informasi hasil riset dari perusahaan/instansi/lembaga manapun.

Saya bersedia untuk bertanggung jawab secara pribadi, tanpa melibatkan pihak **Universitas Bina Sarana Informatika**, atas materi/isi karya ilmiah tersebut, termasuk bertanggung jawab atas dampak atau kerugian yang timbul dalam bentuk akibat tindakan yang berkaitan dengan data dan atau informasi yang terdapat pada karya ilmiah saya ini.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dosen Pembimbing



Rachmat Adi Purnama, M. Kom

Dibuat di : Jakarta

Yang Menyatakan,



Mohamad Abdul Yusuf

LAMPIRAN-LAMPIRAN

Lampiran A. 1 Kuesioner *Depression Anxiety Stress Scales* (DASS 42)

Kuesioner *Depression Anxiety Stress Scales* (DASS 42)

Keterangan:

- 0: Tidak ada atau tidak pernah
- 1: Sesuai dengan yang dialami sampai tingkat tertentu, atau kadang-kadang
- 2: Sering
- 3: Sangat sesuai dengan yang dialami, atau hampir setiap saat.

No.	Aspek Penilaian	0	1	2	3
1.	Menjadi marah karena hal-hal kecil/sepele				
2.	Mulut terasa kering				
3.	Tidak dapat melihat hal yang positif dari suatu kejadian				
4.	Merasakan gangguan dalam bernapas (napas cepat, sulit bernapas)				
5.	Merasa sepertinya tidak kuat lagi untuk melakukan suatu kegiatan				
6.	Cenderung bereaksi berlebihan pada situasi				
7.	Kelemahan pada anggota tubuh				
8.	Kesulitan untuk relaksasi/bersantai				
9.	Cemas yang berlebihan dalam suatu situasi namun bisa lega jika hal/situasi itu berakhir				
10.	Pesimis				
11.	Mudah merasa kesal				
12.	Merasa banyak menghabiskan energi karena cemas				

Lampiran A. 2 Kuesioner *Depression Anxiety Stress Scales* (DASS 42)

13.	Merasa sedih dan depresi				
14.	Tidak sabaran				
15.	Kelelahan				
16.	Kehilangan minat pada banyak hal (misal: makan, ambulasi, sosialisasi)				

	Aspek Penilaian	0	1	2	3
17.	Merasa diri tidak layak				
18.	Mudah tersinggung				
19.	Berkeringat (misal: tangan berkeringat) tanpa stimulasi oleh cuaca maupun latihan fisik				
20.	Ketakutan tanpa alasan yang jelas				
21.	Merasa hidup tidak berharga				
22.	Sulit untuk beristirahat				
23.	Kesulitan dalam menelan				
24.	Tidak dapat menikmati hal-hal yang saya lakukan				
25.	Perubahan kegiatan jantung dan denyut nadi tanpa stimulasi oleh latihan fisik				
26.	Merasa hilang harapan dan putus asa				
27.	Mudah marah				
28.	Mudah panik				
29.	Kesulitan untuk tenang setelah sesuatu yang mengganggu				
30.	Takut diri terhambat oleh tugas-tugas yang tidak biasa dilakukan				
31.	Sulit untuk antusias pada banyak hal				

Lampiran A. 3 Kuesioner *Depression Anxiety Stress Scales* (DASS 42)

32.	Sulit mentoleransi gangguan-gangguan terhadap hal yang sedang dilakukan				
33.	Berada pada keadaan tegang				
34.	Merasa tidak berharga				
35.	Tidak dapat memaklumi hal apapun yang menghalangi anda untuk menyelesaikan hal yang sedang Anda lakukan				
36.	Ketakutan				
37.	Tidak ada harapan untuk masa depan				
38.	Merasa hidup tidak berarti				
No.	Aspek Penilaian	0	1	2	3
39.	Mudah gelisah				
40.	Khawatir dengan situasi saat diri Anda mungkin menjadi panik dan mempermalukan diri sendiri				
41.	Gemetar				
42.	Sulit untuk meningkatkan inisiatif dalam melakukan sesuatu				

-Skala depresi : 3, 5, 10, 13, 16, 17, 21, 24, 26, 31,34, 37, 38, 42.

-Skala kecemasan : 2, 4, 7, 9, 15, 19, 20, 23, 25, 28, 30,36, 40, 41.

-Skala stress : 1, 6, 8, 11, 12, 14, 18, 22, 27, 29, 32, 33, 35, 39.

Indikator Penilaian

Tingkat	Depresi	Kecemasan	Stress
Normal	0-9	0-7	0-14
Ringan	10-13	8-9	15-18
Sedang	14-20	10-14	19-25
Parah	21-27	15-19	26-33
Sangat Parah	>28	>20	>34

Lampiran B. 1 Data Tingkat Depresi

Jenis Kelamin	Usia	Kesibukan	Tidak dapat melihat hal yang positif dari suatu kejadian	Merasa seperti yang tidak kuat lagi untuk melakukan kegiatan	Pesimis	Merasa sedih dan depresi	Saya kehilangan minat pada banyak hal (misal: makan, sosialisasi, hobi)	Saya merasa diri saya tidak layak untuk orang lain	Saya merasa hidup saya tidak berharga	Tidak dapat menikmati hal-hal yang saya lakukan	Merasa hilang harapan dan putus asa	Sulit untuk antusias pada banyak hal	Saya merasa diri saya tidak berharga	Saya merasa tidak ada harapan untuk masa depan saya	Saya merasa hidup saya tidak berarti	Sulit untuk meningkatkan inisiatif dalam melakukan sesuatu	Column 1
Laki-Laki	20-30	Belajar + Berkerja	0	3	3	1	1	1	0	2	0	3	1	1	1	3	20
Laki-Laki	20-30	Belajar + Berkerja	0	0	2	0	0	0	0	0	0	0	0	0	0	0	2
Perempuan	20-30	Sedang Tidak Memiliki Kesibukan	1	3	3	3	3	3	3	3	3	3	3	3	3	3	40
Laki-Laki	20-30	Belajar + Berkerja	0	2	1	2	0	1	0	1	1	0	0	1	1	1	11
Laki-Laki	20-30	Belajar	1	1	1	0	0	0	0	0	0	1	0	0	0	0	4
Laki-Laki	20-30	Belajar	2	1	1	0	0	1	0	0	0	0	1	1	1	1	9
Perempuan	20-30	Berkerja	1	0	1	1	2	1	2	1	1	3	1	2	1	3	20
Laki-Laki	20-30	Sedang Tidak Memiliki Kesibukan	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Laki-Laki	20-30	Belajar + Berkerja	1	2	2	2	1	0	1	1	1	1	0	0	1	1	14
Laki-Laki	20-30	Belajar	2	1	1	1	1	1	1	1	1	1	1	1	1	1	15
Laki-Laki	20-30	Berkerja	1	2	1	2	2	3	3	1	2	1	2	1	1	2	24
Perempuan	20-30	Belajar	1	1	1	2	3	2	3	1	2	2	3	1	2	1	25
Laki-Laki	20-30	Berkerja	1	3	1	2	2	1	0	1	2	2	0	1	1	2	19
Laki-Laki	20-30	Belajar + Berkerja	2	1	0	1	1	0	0	0	0	1	0	0	0	1	7
Laki-Laki	20-30	Belajar	1	1	1	0	0	1	1	1	1	1	1	1	1	1	12
Perempuan	20-30	Belajar + Berkerja	1	2	2	3	2	2	1	1	2	1	2	1	2	1	23
Laki-Laki	20-30	Sedang Tidak Memiliki Kesibukan	2	1	3	3	1	2	2	1	3	2	2	2	1	2	27
Perempuan	30-40	Sedang Tidak Memiliki Kesibukan	0	0	3	3	3	3	3	3	3	2	3	3	3	3	35
Perempuan	20-30	Berkerja	2	3	3	3	2	2	2	1	3	3	1	2	3	3	33
Perempuan	20-30	Belajar	0	1	1	1	1	1	0	0	0	0	0	0	0	1	6
Perempuan	20-30	Belajar + Berkerja	1	1	0	2	1	1	1	1	1	0	0	0	0	1	11
Perempuan	20-30	Belajar + Berkerja	1	1	0	0	0	0	0	0	1	0	0	0	0	0	3
Perempuan	20-30	Berkerja	3	1	2	3	2	2	0	2	0	1	0	1	0	2	19
Perempuan	15-20	Berkerja	1	0	0	2	0	0	0	0	1	0	0	0	0	0	4
Perempuan	15-20	Berkerja	1	3	3	3	0	3	3	3	3	1	3	3	3	3	35
Laki-Laki	20-30	Belajar + Berkerja	0	0	1	1	0	3	0	0	0	1	0	0	0	0	6
Laki-Laki	20-30	Berkerja	1	0	2	2	1	1	0	0	0	0	0	1	0	1	9
Perempuan	15-20	Belajar + Berkerja	1	1	2	2	2	3	2	1	0	2	0	1	1	2	20
Perempuan	20-30	Belajar	0	1	1	2	3	3	1	3	1	3	2	1	1	2	24
Laki-Laki	20-30	Belajar	1	0	0	0	0	0	0	0	0	1	0	0	0	1	3
Laki-Laki	20-30	Belajar + Berkerja	2	1	0	1	1	0	0	0	1	0	0	0	0	1	7
Laki-Laki	20-30	Belajar + Berkerja	3	0	0	0	0	3	0	0	0	0	0	0	0	0	6
Laki-Laki	20-30	Belajar + Berkerja	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Laki-Laki	20-30	Belajar	0	1	0	0	1	2	1	1	0	0	0	0	0	2	8

Lampiran B. 2 Data Tingkat Depresi

Laki-Laki	15-20	Belajar	0	2	2	1	1	1	0	0	0	0	0	1	0	1	9
Perempuan	20-30	Belajar	2	2	3	2	2	2	1	2	2	2	2	2	2	2	28
Laki-Laki	20-30	Sedang Tidak Memiliki Kesibukan	1	2	2	3	0	1	1	3	3	3	3	1	2	0	25
Perempuan	15-20	Belajar	1	3	3	3	1	3	3	1	1	1	2	0	2	1	25
Perempuan	20-30	Sedang Tidak Memiliki Kesibukan	0	1	2	1	1	2	1	0	0	1	0	0	0	0	9
Laki-Laki	20-30	Belajar	2	3	2	2	1	1	2	2	2	1	0	0	0	3	21
Laki-Laki	15-20	Sedang Tidak Memiliki Kesibukan	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Perempuan	20-30	Belajar + Berkerja	0	0	1	1	0	0	1	0	1	1	1	0	0	0	6
Perempuan	20-30	Belajar	1	3	1	3	1	1	0	0	2	0	1	0	0	0	13
Laki-Laki	15-20	Berkerja	1	1	1	2	0	1	0	1	1	0	0	0	0	0	8
Laki-Laki	20-30	Sedang Tidak Memiliki Kesibukan	2	0	0	3	1	3	3	0	3	3	3	0	3	2	26
Perempuan	15-20	Belajar	1	2	1	2	3	3	3	3	3	1	2	3	3	1	31
Perempuan	20-30	Berkerja	0	0	1	0	0	2	0	1	0	3	0	1	0	0	8
Perempuan	20-30	Berkerja	0	1	1	0	2	1	1	1	0	1	1	0	1	0	10
Perempuan	20-30	Berkerja	3	2	2	1	1	2	0	0	1	1	1	1	1	1	17
Laki-Laki	15-20	Berkerja	2	0	1	1	0	0	1	1	1	0	0	0	0	0	7
Perempuan	20-30	Belajar + Berkerja	0	1	1	1	1	2	0	0	0	2	0	0	0	1	9
Laki-Laki	15-20	Belajar	2	1	2	1	1	1	0	1	1	1	0	1	1	1	14
Perempuan	15-20	Belajar	0	2	1	2	0	0	1	1	1	1	1	1	1	0	12
Perempuan	20-30	Berkerja	1	0	1	2	0	1	0	1	0	1	1	0	1	1	10
Perempuan	20-30	Berkerja	0	1	1	1	0	0	0	0	0	0	0	0	0	0	3
Laki-Laki	20-30	Belajar	0	1	1	1	1	1	1	1	0	1	1	1	1	1	12
Laki-Laki	20-30	Belajar + Berkerja	1	1	3	2	3	1	2	2	2	2	1	1	2	1	24
Perempuan	20-30	Belajar	0	0	2	1	0	0	0	0	1	0	0	0	0	1	5
Perempuan	20-30	Belajar	1	0	1	1	1	0	0	0	0	1	0	0	0	1	6
Perempuan	20-30	Berkerja	1	1	1	1	0	0	0	1	1	0	0	1	0	1	8
Laki-Laki	20-30	Sedang Tidak Memiliki Kesibukan	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Laki-Laki	20-30	Berkerja	3	0	0	0	0	0	3	3	0	3	3	0	0	3	18
Perempuan	20-30	Belajar + Berkerja	1	2	2	2	1	2	0	0	2	1	0	1	0	2	16
Perempuan	20-30	Belajar	0	1	2	1	0	1	1	0	1	0	1	1	1	2	12
Laki-Laki	20-30	Belajar	1	2	3	3	3	3	3	3	3	3	3	3	3	2	38
Perempuan	20-30	Belajar + Berkerja	1	2	2	2	2	2	1	1	2	2	2	0	0	1	20
Perempuan	20-30	Sedang Tidak Memiliki Kesibukan	1	2	3	3	3	3	3	1	1	1	1	2	2	2	28
Laki-Laki	20-30	Berkerja	0	1	0	1	2	0	0	1	0	1	3	1	1	1	12
Laki-Laki	20-30	Berkerja	1	1	1	2	1	2	1	0	0	2	1	1	1	2	16

Lampiran B. 3 Data Tingkat Depresi

Perempuan	20-30	Belajar	1	1	1	1	0	0	1	1	3	3	1	1	1	1	16
Laki-Laki	15-20	Belajar	3	0	0	0	0	0	0	0	0	0	0	0	0	0	3
Perempuan	20-30	Belajar + Berkerja	3	2	2	2	3	2	2	3	3	3	2	3	3	3	36
Perempuan	20-30	Belajar	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Perempuan	20-30	Belajar + Berkerja	0	1	0	2	2	0	0	1	0	0	0	0	0	1	7
Perempuan	15-20	Sedang Tidak Memiliki Kesibukan	2	2	2	3	3	3	3	3	1	3	3	0	3	2	33
Laki-Laki	20-30	Berkerja	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Laki-Laki	20-30	Belajar	0	0	1	1	0	1	0	0	0	0	0	0	0	1	4
Perempuan	20-30	Berkerja	0	1	2	1	2	0	0	0	0	2	0	0	0	0	8
Perempuan	20-30	Berkerja	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Perempuan	15-20	Belajar + Berkerja	0	1	1	1	1	1	1	0	1	0	1	1	0	1	10
Laki-Laki	20-30	Berkerja	1	1	1	1	1	1	1	1	1	1	1	1	1	1	14
Laki-Laki	20-30	Berkerja	3	0	0	0	2	0	0	0	0	1	0	0	0	0	6
Laki-Laki	20-30	Berkerja	1	1	2	3	2	3	2	1	2	3	1	0	0	2	23
Laki-Laki	15-20	Sedang Tidak Memiliki Kesibukan	3	2	2	3	1	3	3	3	2	1	1	2	3	3	32
Laki-Laki	20-30	Sedang Tidak Memiliki Kesibukan	3	3	3	3	2	2	3	3	3	3	3	3	3	3	40
Perempuan	20-30	Belajar + Berkerja	2	3	1	1	3	0	1	1	1	2	0	0	0	3	18
Perempuan	20-30	Belajar + Berkerja	0	1	1	1	0	0	0	0	0	3	0	0	0	2	8
Laki-Laki	20-30	Berkerja	2	0	1	0	0	0	0	1	0	1	0	0	0	1	6
Perempuan	20-30	Sedang Tidak Memiliki Kesibukan	1	1	2	1	2	2	2	1	1	2	2	2	2	1	22
Laki-Laki	20-30	Berkerja	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Laki-Laki	15-20	Belajar	2	1	0	3	3	3	1	1	3	0	1	0	1	0	19
Perempuan	20-30	Belajar + Berkerja	0	0	0	0	0	3	0	0	3	0	0	0	3	0	9
Perempuan	20-30	Belajar + Berkerja	1	1	2	2	1	3	2	1	2	1	1	1	1	1	20
Laki-Laki	20-30	Belajar	0	0	1	2	0	3	3	2	1	1	2	2	2	1	20
Laki-Laki	20-30	Belajar + Berkerja	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1
Perempuan	30-40	Sedang Tidak Memiliki Kesibukan	0	1	1	3	1	0	0	1	0	0	0	0	0	1	8
Perempuan	30-40	Sedang Tidak Memiliki Kesibukan	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
Perempuan	20-30	Sedang Tidak Memiliki Kesibukan	0	0	1	2	0	0	0	0	0	0	0	0	0	0	3
Laki-Laki	20-30	Belajar	2	2	1	1	0	0	0	1	1	1	0	0	0	0	9
Perempuan	20-30	Berkerja	1	3	2	2	1	1	0	2	1	0	0	0	0	1	14
Perempuan	15-20	Belajar	2	2	2	2	1	3	2	0	2	2	2	3	2	2	27
Laki-Laki	20-30	Belajar + Berkerja	0	2	2	1	2	2	1	1	3	1	2	1	2	0	20
Laki-Laki	30-40	Belajar + Berkerja	2	0	0	0	0	0	0	0	0	0	0	0	0	0	2
Perempuan	20-30	Berkerja	0	1	1	1	0	1	0	0	0	1	0	0	0	1	6
Laki-Laki	20-30	Belajar + Berkerja	1	0	1	0	0	1	0	0	0	1	1	1	1	1	8

Lampiran C. 1 Kode Program *K-Means* dan *Particle Swarm Optimization*

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
df = pd.read_csv('/content/DATA-DEPRESI-CLEAN-2.csv',
index_col=0, sep=',')
df.head()
df.info()
df.describe()
import pandas as pd
import plotly.graph_objects as go

# Load the data
df = pd.read_csv('/content/DATA-DEPRESI-CLEAN-2.csv',
index_col=0, sep=',')

# Create the interactive plot
fig = go.Figure()
fig.add_trace(go.Scatter(
    x=df['KODE JK'],
    y=df['KODE TD'],
    mode='markers',
    marker=dict(
        size=10,
        color=df['KODE TD'],
        colorscale='Viridis',
        opacity=0.7
    ),
    text=df.index,
    hovertemplate='<b>ID: %{text}</b><br>Jenis Kelamin:
%{x}<br>Tingkat Depresi: %{y}'
))
```

Lampiran C. 2 Kode Program *K-Means* dan *Particle Swarm Optimization*

```
# Add text annotation to show the total number of data points
total_data = len(df)
fig.add_annotation(
    x=0.5,
    y=1.05,
    text=f"Total Data: {total_data}",
    showarrow=False,
    font=dict(
        family="Courier New, monospace",
        size=16,
        color="RebeccaPurple"
    )
)

# Update layout
fig.update_layout(
    title='Perbandingan Jenis Kelamin vs Tingkat Depresi',
    xaxis_title='Jenis Kelamin',
    yaxis_title='Tingkat Depresi',
    font=dict(
        family="Courier New, monospace",
        size=14,
        color="RebeccaPurple"
    )
)

# Display the plot
fig.show()
x_train = df[['KODE TD', 'KODE TD']].values
x_train
from sklearn.preprocessing import MinMaxScaler
scaler = MinMaxScaler()
x_train=scaler.fit_transform(x_train)
x_train
from sklearn.cluster import KMeans
Kmeans = KMeans(n_clusters=2)
Kmeans
y_cluster = Kmeans.fit_predict(x_train)
y_cluster
df['KODE HSL'] = y_cluster
df
df['KODE HSL'] = y_cluster = df['KODE
HSL'].replace([0,1], ['SEHAT', 'DEPRESI'])
df['KODE HSL'] = y_cluster
df
```

Lampiran C. 3 Kode Program *K-Means* dan *Particle Swarm Optimization*

```
import pandas as pd

# Baca data dari file CSV
df = pd.read_csv('/content/DATA-DEPRESI-CLEAN-2.csv',
index_col=0, sep=',')
df['KODE HSL'] = y_cluster
# Buat crosstab antara KODE JK dan KODE HSL
pivot_table = pd.crosstab(df['KODE JK'], df['KODE HSL'])

# Tampilkan hasil
print(pivot_table)
import pandas as pd
import plotly.graph_objects as go

# Baca data dari file CSV
df = pd.read_csv('/content/DATA-DEPRESI-CLEAN-2.csv',
index_col=0, sep=',')
df['KODE HSL'] = y_cluster

# Buat crosstab antara KODE JK dan KODE HSL
pivot_table = pd.crosstab(df['KODE JK'], df['KODE HSL'])

# Visualisasi menggunakan Plotly
fig = go.Figure()

for col in pivot_table.columns:
    fig.add_trace(go.Bar(
        x=pivot_table.index,
        y=pivot_table[col],
        name=col
    ))

fig.update_layout(
    title='Crosstab of KODE JK and KODE HSL',
    xaxis_title='KODE JK',
    yaxis_title='Count',
    legend_title='KODE HSL',
    bargap=0.1
)

fig.show()
Kmeans.cluster_centers_
```

Lampiran C. 4 Kode Program *K-Means* dan *Particle Swarm Optimization*

```
plt.figure(figsize=(10,5))
plt.scatter(x_train[:,0],x_train[:,1],c=Kmeans.labels_)
plt.scatter(Kmeans.cluster_centers_[0,0],Kmeans.cluster_centers_[
0,1],cmap='rainbow', marker='*', s=200, label='CENTROID 0')
plt.scatter(Kmeans.cluster_centers_[1,0],Kmeans.cluster_centers_[
1,1],cmap='rainbow', marker='*', s=200, label='CENTROID 1')
plt.xlabel('JENIS KELAMIN')
plt.ylabel('TINGKAT DEPRESI')
plt.legend()
inertias = []
k_range = range(1,10)
for k in k_range:
    km = KMeans(n_clusters=k).fit(x_train)
    inertias.append(km.inertia_)
inertias
from sklearn.metrics import silhouette_score,
calinski_harabasz_score, davies_bouldin_score

# Hitung silhouette score
silhouette = silhouette_score(x_train, y_cluster)
print(f"Silhouette Score: {silhouette:.2f}")

# Hitung Calinski-Harabasz score
calinski_harabasz = calinski_harabasz_score(x_train, y_cluster)
print(f"Calinski-Harabasz Score: {calinski_harabasz:.2f}")

# Hitung Davies-Bouldin score
davies_bouldin = davies_bouldin_score(x_train, y_cluster)
print(f"Davies-Bouldin Score: {davies_bouldin:.2f}")
plt.xlabel("k")
plt.ylabel("Sum Of Error")
plt.plot(k_range,inertias)
plt.grid()
from sklearn.cluster import KMeans
Kmeans = KMeans(n_clusters=4)
Kmeans
y_cluster = Kmeans.fit_predict(x_train)
y_cluster
df['KODE HSL2'] = y_cluster
df
```

Lampiran C. 5 Kode Program *K-Means* dan *Particle Swarm Optimization*

```
df['KODE HSL2'] = y_cluster = df['KODE
HSL2'].replace([0,1,2,3], ['DEPRESI RINGAN', 'DEPRESI PARAH',
'TIDAK DEPRESI', 'DEPRESI SEDANG'])
df['KODE HSL2'] = y_cluster
df
import pandas as pd

# Baca data dari file CSV
df = pd.read_csv('/content/DATA-DEPRESI-CLEAN-2.csv',
index_col=0, sep=',')
df['KODE HSL2'] = y_cluster
# Buat crosstab antara KODE JK dan KODE HSL
pivot_table = pd.crosstab(df['KODE JK'], df['KODE HSL2'])

# Tampilkan hasil
print(pivot_table)
Kmeans.cluster_centers_
plt.figure(figsize=(10,5))
plt.scatter(x_train[:,0],x_train[:,1],c=Kmeans.labels_)
plt.scatter(Kmeans.cluster_centers_[0,0],Kmeans.cluster_centers_[
0,1],cmap='rainbow', marker='*', s=200, label='CENTROID 0')
plt.scatter(Kmeans.cluster_centers_[1,0],Kmeans.cluster_centers_[
1,1],cmap='rainbow', marker='*', s=200, label='CENTROID 1')
plt.scatter(Kmeans.cluster_centers_[2,0],Kmeans.cluster_centers_[
2,1],cmap='rainbow', marker='*', s=200, label='CENTROID 2')
plt.scatter(Kmeans.cluster_centers_[3,0],Kmeans.cluster_centers_[
3,1],cmap='rainbow', marker='*', s=200, label='CENTROID 3')
plt.xlabel('JENIS KELAMIN')
plt.ylabel('TINGKAT DEPRESI')
plt.legend()
```

Lampiran C. 6 Kode Program *K-Means* dan *Particle Swarm Optimization*

```
import pandas as pd
import plotly.graph_objects as go

# Baca data dari file CSV
df = pd.read_csv('/content/DATA-DEPRESI-CLEAN-2.csv',
index_col=0, sep=',')
df['KODE HSL2'] = y_cluster

# Buat crosstab antara KODE JK dan KODE HSL2
pivot_table = pd.crosstab(df['KODE JK'], df['KODE HSL2'])

# Visualisasi menggunakan Plotly
fig = go.Figure()

for col in pivot_table.columns:
    fig.add_trace(go.Bar(
        x=pivot_table.index,
        y=pivot_table[col],
        name=col
    ))

fig.update_layout(
    title='Crosstab of KODE JK and KODE HSL2',
    xaxis_title='KODE JK',
    yaxis_title='Count',
    legend_title='KODE HSL2',
    bargap=0.2
)

fig.show()
inertias = []
k_range = range(1,10)
for k in k_range:
    km = KMeans(n_clusters=k).fit(x_train)
    inertias.append(km.inertia_)
inertias
```

Lampiran C. 7 Kode Program *K-Means* dan *Particle Swarm Optimization*

```
from sklearn.metrics import silhouette_score,
calinski_harabasz_score, davies_bouldin_score

# Hitung silhouette score
silhouette = silhouette_score(x_train, y_cluster)
print(f"Silhouette Score: {silhouette:.2f}")

# Hitung Calinski-Harabasz score
calinski_harabasz = calinski_harabasz_score(x_train, y_cluster)
print(f"Calinski-Harabasz Score: {calinski_harabasz:.2f}")

# Hitung Davies-Bouldin score
davies_bouldin = davies_bouldin_score(x_train, y_cluster)
print(f"Davies-Bouldin Score: {davies_bouldin:.2f}")
Xcell=pd.DataFrame(df)
Xcell.to_excel(excel_writer=r'/content/final.xlsx')
import pandas as pd
import numpy as np
import plotly.graph_objects as go
import plotly.colors as pc
from sklearn.cluster import KMeans
from scipy.spatial.distance import cdist
import random

# Baca data
df = pd.read_csv('/content/DATA-DEPRESI-CLEAN-2.csv',
index_col=0, sep=',')

# Lakukan normalisasi fitur
X = df[['KODE TD', 'KODE TD']].values
scaler = MinMaxScaler()
X_norm = scaler.fit_transform(X)

# Fungsi optimisasi PSO
def pso_kmeans(X, n_clusters, n_particles, n_iterations):
    # Inisialisasi posisi dan kecepatan partikel
    positions = np.random.rand(n_particles, n_clusters,
X.shape[1])
    velocities = np.zeros((n_particles, n_clusters, X.shape[1]))
```

Lampiran C. 8 Kode Program *K-Means* dan *Particle Swarm Optimization*

```
# Inisialisasi best_positions dan best_fitness
best_positions = positions.copy()
best_fitness = np.inf * np.ones(n_particles)

# Inisialisasi global best
global_best = positions[0].copy()
global_best_fitness = np.inf

# Iterasi PSO
for _ in range(n_iterations):
    for i in range(n_particles):
        # Hitung fitness (error K-Means)
        kmeans = KMeans(n_clusters=n_clusters,
init=positions[i], n_init=1, max_iter=100)
        kmeans.fit(X)
        fitness = kmeans.inertia_

        # Update best_positions dan best_fitness
        if fitness < best_fitness[i]:
            best_positions[i] = positions[i].copy()
            best_fitness[i] = fitness

        # Update global_best dan global_best_fitness
        if fitness < global_best_fitness:
            global_best = positions[i].copy()
            global_best_fitness = fitness

    # Update posisi dan kecepatan partikel
    r1 = random.random()
    r2 = random.random()
    for j in range(n_particles):
        velocities[j] = 0.5 * velocities[j] + 2 * r1 *
(best_positions[j] - positions[j]) + 2 * r2 * (global_best -
positions[j])
        positions[j] = positions[j] + velocities[j]
```

Lampiran C. 9 Kode Program *K-Means* dan *Particle Swarm Optimization*

```
    return global_best

# Cari jumlah cluster optimal menggunakan PSO
n_particles = 20
n_iterations = 100
inertias = []
for n_clusters in range(2, 11):
    centroids = pso_kmeans(X_norm, n_clusters, n_particles,
n_iterations)
    kmeans = KMeans(n_clusters=n_clusters, init=centroids,
n_init=1)
    kmeans.fit(X_norm)
    inertias.append(kmeans.inertia_)

# Visualisasi jumlah cluster optimal
plt.figure(figsize=(8, 6))
plt.plot(range(2, 11), inertias, 'bx-')
plt.xlabel('Jumlah Cluster (k)')
plt.ylabel('Inertia')
plt.title('Elbow Plot - Menentukan Jumlah Cluster Optimal')
plt.show()

# Lakukan klusterisasi final dengan K-Means
n_clusters = np.argmin(inertias) + 2
kmeans = KMeans(n_clusters=n_clusters, random_state=42)
y_cluster = kmeans.fit_predict(X_norm)
df['Cluster'] = y_cluster

# Buat visualisasi interaktif
cluster_sizes = df.groupby('Cluster').size()
palette = pc.qualitative.Dark24

fig = go.Figure()
fig.add_trace(go.Bar(
    x=cluster_sizes.index,
    y=cluster_sizes,
    marker_color=[palette[i] for i in cluster_sizes.index]
)))
```

Lampiran C. 10 Kode Program *K-Means* dan *Particle Swarm Optimization*

```
fig.update_layout(  
    title='Cluster Setelah PSO',  
    xaxis_title='Cluster',  
    yaxis_title='Jumlah Data',  
    plot_bgcolor='lightslategray',  
    font=dict(  
        family="Courier New, monospace",  
        size=14,  
        color="Black"  
    )  
)  
  
fig.show()  
  
# Lakukan klasterisasi final dengan K-Means  
n_clusters = np.argmin(inertias) + 2  
kmeans = KMeans(n_clusters=n_clusters, random_state=42)  
y_cluster = kmeans.fit_predict(X_norm)  
df['Cluster'] = y_cluster  
  
# Tampilkan centroid cluster terakhir  
centroids = kmeans.cluster_centers_  
print("Centroids:")  
print(centroids)  
import pandas as pd  
import numpy as np  
import plotly.graph_objects as go  
import plotly.colors as pc  
from sklearn.cluster import KMeans  
from scipy.spatial.distance import cdist  
import random  
  
# Baca data  
df = pd.read_csv('/content/DATA-DEPRESI-CLEAN-2.csv',  
    index_col=0, sep=',')  
  
# Lakukan normalisasi fitur  
X = df[['KODE TD', 'KODE TD']].values  
scaler = MinMaxScaler()  
X_norm = scaler.fit_transform(X)
```

Lampiran C. 11 Kode Program *K-Means* dan *Particle Swarm Optimization*

```
# Fungsi optimisasi PSO
def pso_kmeans(X, n_clusters, n_particles, n_iterations):
    # Inisialisasi posisi dan kecepatan partikel
    positions = np.random.rand(n_particles, n_clusters,
X.shape[1])
    velocities = np.zeros((n_particles, n_clusters, X.shape[1]))

    # Inisialisasi best_positions dan best_fitness
    best_positions = positions.copy()
    best_fitness = np.inf * np.ones(n_particles)

    # Inisialisasi global best
    global_best = positions[0].copy()
    global_best_fitness = np.inf

    # Iterasi PSO
    for _ in range(n_iterations):
        for i in range(n_particles):
            # Hitung fitness (error K-Means)
            kmeans = KMeans(n_clusters=n_clusters,
init=positions[i], n_init=1, max_iter=100)
            kmeans.fit(X)
            fitness = kmeans.inertia_

            # Update best_positions dan best_fitness
            if fitness < best_fitness[i]:
                best_positions[i] = positions[i].copy()
                best_fitness[i] = fitness

            # Update global_best dan global_best_fitness
            if fitness < global_best_fitness:
                global_best = positions[i].copy()
                global_best_fitness = fitness

        # Update posisi dan kecepatan partikel
        r1 = random.random()
        r2 = random.random()
        for j in range(n_particles):
            velocities[j] = 0.5 * velocities[j] + 2 * r1 *
(best_positions[j] - positions[j]) + 2 * r2 * (global_best -
positions[j])
            positions[j] = positions[j] + velocities[j]
    return global_best
```

Lampiran C. 12 Kode Program *K-Means* dan *Particle Swarm Optimization*

```
# Cari jumlah cluster optimal menggunakan PSO
n_particles = 20
n_iterations = 100
inertias = []
for n_clusters in range(2, 11):
    centroids = pso_kmeans(X_norm, n_clusters, n_particles,
n_iterations)
    kmeans = KMeans(n_clusters=n_clusters, init=centroids,
n_init=1)
    kmeans.fit(X_norm)
    inertias.append(kmeans.inertia_)

# Lakukan klusterisasi final dengan K-Means
n_clusters = np.argmin(inertias) + 2
kmeans = KMeans(n_clusters=n_clusters, random_state=42)
y_cluster = kmeans.fit_predict(X_norm)
df['Cluster'] = y_cluster

# Visualisasi kluster dan centroid
fig = go.Figure()

# Plot data
for i in range(n_clusters):
    cluster_data = X_norm[y_cluster == i]
    fig.add_trace(go.Scatter(
        x=cluster_data[:, 0], y=cluster_data[:, 1],
        mode='markers',
        name=f'Cluster {i}',
        marker=dict(
            color=palette[i],
            size=5
        )
    ))
))
```

Lampiran C. 13 Kode Program *K-Means* dan *Particle Swarm Optimization*

```
fig.update_layout(  
    title='Visualisasi Cluster Data dan Centroid',  
    xaxis_title='JENIS KELAMIN',  
    yaxis_title='TINGKAT DEPRRSi',  
    plot_bgcolor='lightslategray',  
    font=dict(  
        family="Courier New, monospace",  
        size=14,  
        color="Black"  
    )  
)  
  
fig.show()  
  
from sklearn.metrics import silhouette_score,  
calinski_harabasz_score, davies_bouldin_score  
  
# Hitung silhouette score  
silhouette = silhouette_score(X_norm, y_cluster)  
print(f"Silhouette Score: {silhouette:.2f}")  
  
# Hitung Calinski-Harabasz score  
calinski_harabasz = calinski_harabasz_score(X_norm, y_cluster)  
print(f"Calinski-Harabasz Score: {calinski_harabasz:.2f}")  
  
# Hitung Davies-Bouldin score  
davies_bouldin = davies_bouldin_score(X_norm, y_cluster)  
print(f"Davies-Bouldin Score: {davies_bouldin:.2f}")  
df['KODE HSL3'] = y_cluster  
df  
df['KODE HSL3'] = y_cluster = df['KODE  
HSL3'].replace([0,1,2,3,4], ['DEPRESI RINGAN', 'DEPRESI PARAH',  
'TIDAK DEPRESI', 'DEPRESI SEDANG', 'DEPRESI SANGAT PARAH'])  
df['KODE HSL3'] = y_cluster  
df
```

Lampiran C. 14 Kode Program *K-Means* dan *Particle Swarm Optimization*

```
import pandas as pd
import plotly.graph_objects as go

# Baca data dari file CSV
df = pd.read_csv('/content/DATA-DEPRESI-CLEAN-2.csv',
index_col=0, sep=',')
df['KODE HSL3'] = y_cluster

# Buat crosstab antara KODE JK dan KODE HSL2
pivot_table = pd.crosstab(df['KODE JK'], df['KODE HSL3'])

# Visualisasi menggunakan Plotly
fig = go.Figure()

for col in pivot_table.columns:
    fig.add_trace(go.Bar(
        x=pivot_table.index,
        y=pivot_table[col],
        name=col
    ))

fig.update_layout(
    title='Crosstab of KODE JK and KODE HSL3',
    xaxis_title='KODE JK',
    yaxis_title='Count',
    legend_title='KODE HSL3',
    bargap=0.2
)

fig.show()
import pandas as pd

# Baca data dari file CSV
df = pd.read_csv('/content/DATA-DEPRESI-CLEAN-2.csv',
index_col=0, sep=',')
df['KODE HSL3'] = y_cluster
# Buat crosstab antara KODE JK dan KODE HSL
pivot_table = pd.crosstab(df['KODE JK'], df['KODE HSL3'])

# Tampilkan hasil
print(pivot_table)
```

Lampiran C. 15 Kode Program *K-Means* dan *Particle Swarm Optimization*

```
df['KODE HSL3'] = y_cluster  
y_cluster  
Xcell=pd.DataFrame(df['KODE HSL3'])  
Xcell.to_excel(excel_writer=r'/content/hasilcls.xlsx')
```



Lampiran D. 1 *Cluster* Hasil Tingkat Depresi

NO	KODE HSL3
1	DEPRESI SEDANG
2	TIDAK DEPRESI
3	DEPRESI SANGAT PA
4	DEPRESI RINGAN
5	TIDAK DEPRESI
6	TIDAK DEPRESI
7	DEPRESI SEDANG
8	DEPRESI SEDANG
9	DEPRESI SEDANG
10	DEPRESI PARAH
11	DEPRESI PARAH
12	DEPRESI SEDANG
13	DEPRESI RINGAN
14	DEPRESI PARAH
15	DEPRESI PARAH
16	DEPRESI SANGAT PA
17	DEPRESI SANGAT PA
18	TIDAK DEPRESI
19	DEPRESI RINGAN
20	TIDAK DEPRESI
21	DEPRESI SEDANG
22	TIDAK DEPRESI
23	DEPRESI SANGAT PA
24	TIDAK DEPRESI
25	TIDAK DEPRESI
26	DEPRESI SEDANG
27	DEPRESI PARAH
28	TIDAK DEPRESI
29	TIDAK DEPRESI
30	TIDAK DEPRESI
31	TIDAK DEPRESI
32	TIDAK DEPRESI
33	DEPRESI SANGAT PA
34	DEPRESI PARAH
35	DEPRESI PARAH
36	TIDAK DEPRESI
37	DEPRESI PARAH
38	TIDAK DEPRESI
39	DEPRESI RINGAN
40	TIDAK DEPRESI
41	DEPRESI PARAH
42	DEPRESI SANGAT PA
43	TIDAK DEPRESI
44	DEPRESI RINGAN
45	DEPRESI SEDANG
46	TIDAK DEPRESI
47	TIDAK DEPRESI
48	DEPRESI SEDANG

48	DEPRESI SEDANG
49	DEPRESI RINGAN
50	DEPRESI RINGAN
51	TIDAK DEPRESI
52	DEPRESI RINGAN
53	DEPRESI PARAH
54	TIDAK DEPRESI
55	TIDAK DEPRESI
56	TIDAK DEPRESI
57	DEPRESI SEDANG
58	DEPRESI SEDANG
59	DEPRESI RINGAN
60	DEPRESI SANGAT PA
61	DEPRESI SEDANG
62	DEPRESI PARAH
63	DEPRESI RINGAN
64	DEPRESI SEDANG
65	DEPRESI SEDANG
66	TIDAK DEPRESI
67	DEPRESI SANGAT PA
68	TIDAK DEPRESI
69	DEPRESI SANGAT PA
70	TIDAK DEPRESI
71	TIDAK DEPRESI
72	DEPRESI RINGAN
73	DEPRESI SEDANG
74	TIDAK DEPRESI
75	DEPRESI PARAH
76	DEPRESI SANGAT PA
77	DEPRESI SANGAT PA
78	DEPRESI SEDANG
79	TIDAK DEPRESI
80	TIDAK DEPRESI
81	DEPRESI PARAH
82	DEPRESI SEDANG
83	TIDAK DEPRESI
84	DEPRESI SEDANG
85	DEPRESI SEDANG
86	TIDAK DEPRESI
87	TIDAK DEPRESI
88	TIDAK DEPRESI
89	DEPRESI SEDANG
90	DEPRESI PARAH
91	DEPRESI SEDANG
92	DEPRESI SEDANG
93	DEPRESI PARAH
94	DEPRESI SANGAT PA