

**OPTIMASI METODE K-MEANS DENGAN ALGORITMA PSO  
PADA PENGLASTERAN DATA BERDIMENSI TINGGI**

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



Oleh:

Ardi Wasila Chandra  
09021181520014

**JURUSAN TEKNIK INFORMATIKA  
FAKULTAS ILMU KOMPUTER UNIVERSITAS SRIWIJAYA  
2019**

**LEMBAR PENGESAHAN TUGAS AKHIR**

**OPTIMASI METODE K-MEANS DENGAN ALGORITMA PSO  
PADA PENGLASTERAN DATA BERDIMENSI TINGGI**

Oleh :

**ARDI WASILA CHANDRA**

**NIM : 09021181520014**

Mengetahui,  
Ketua Jurusan Teknik Informatika,



Rifkie Primartha, M.T  
NIP. 198505102015041002

Indralaya, Oktober 2019  
Pembimbing,

Rifkie Primartha, M.T  
NIP. 198505102015041002

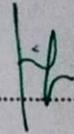
### TANDA LULUS SIDANG TUGAS AKHIR

Pada hari Jumat, 4 Oktober 2019 telah dilaksanakan ujian sidang tugas akhir oleh Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.

Nama : Ardi Wasila Chandra  
NIM : 09021181520014  
Judul : Optimasi Metode K-means dengan Algoritma PSO pada Pengklasteran Data Berdimensi Tinggi

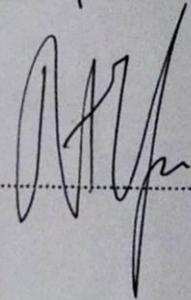
1. Pembimbing I

Rifkie Primartha, M.T  
NIP. 197706012009121004



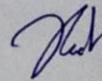
2. Penguji I

Novi Yusliani, M.T  
NIP. 198211082012122001



3. Penguji II

Kanda Januar Miraswan, M.T  
NIP. 199001092019031012



Mengetahui,  
Ketua Jurusan Teknik Informatika



Rifkie Primartha, M.T  
NIP. 197706012009121004

## HALAMAN PERNYATAAN

Yang bertanda tangan di bawah ini :

Nama : Ardi Wasila Chandra  
NIM : 09021181520014  
Program Studi : Teknik Informatika  
Judul Skripsi : Optimasi Metode K-means dengan Algoritma PSO  
Pada Pengklasteran Data Berdimensi Tinggi  
Hasil Pengecekan Software *iThenticate/Turnitin* : 7%

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

Demikian, pernyataan ini saya buat dengan sebenarnya dan tidak ada paksaan oleh siapapun.



Indralaya, 5 November 2019



Ardi Wasila Chandra  
NIM. 09021181520014

*Motto:*

- *“Berdoalah kepada ku pastilah aku kabulkan untukmu”*. Setiap kali memiliki hajat atau menginginkan sesuatu hendaknya mengusahakan dengan sungguh sungguh dan meminta pada Allah untuk mengabulkan hajat anda. Allah senang pada hamba Nya yang senantiasa berdoa, karena doa menghubungkan langsung antara seorang hamba dengan Allah. - **QS Al Mukmin : 60**

*Kupersembahkan karya tulis ini kepada:*

- *Orang tuaku tersayang*
- *Keluarga besarku*
- *Sahabat dan teman seperjuanganku*
- *Fakultas Ilmu Komputer*
- *Universitas Sriwijaya*

OPTIMIZATION K-MEANS METHOD USING  
PARTICLE SWARM OPTIMIZATION ALGORITHM  
ON CLUSTERING OF HIGH-DIMENSIONAL DATA

by:

Ardi Wasila Chandra  
09021181520014

**ABSTRACT**

K-means is one of the clustering methods whose clustering results depend on the position of the initial centroid. The initial centroid commonly used in the k-means method is the initial centroid randomly generated, initial centroids that are randomly generated usually often causes k-means to be trapped in the optimum local solution, therefore in this study we will examine the effect of particle swarm optimization which are wrong one optimization algorithm that can do a global search in determining the initial centroid of k-means. Clustering k-means with random initial centroids and initial centroids from particle swarm optimization calculations are each tested on data dimensional reduction and without dimensional reduction. Based on the results of evaluation of the results of k-means clustering with the initial centroid of particle swarm optimization able to improve cluster quality, both if tested on reduction and without reduction data, namely with the percentage change value of 43.8% in data without dimensional reduction and 53.4% in the data with dimensional reduction. Although it can increase overall computational time but, the initial centroid obtained from particle swarm optimization makes the complexity of k-means work simpler.

**Keywords:** *Clustering, K-Means, Particle Swarm Optimization, Principal Component Analysis, Davies Bouldin Index.*

Mengetahui,  
Ketua Jurusan Teknik Informatika,



Rifkie Primartha, M.T  
NIP. 198505102015041002

Indralaya, Oktober 2019  
Pembimbing,

Rifkie Primartha, M.T  
NIP. 198505102015041002

OPTIMASI METODE K-MEANS DENGAN  
ALGORITMA PSO PADA PENGLASTERAN DATA  
BERDIMENSI TINGGI

Oleh:

Ardi Wasila Chandra  
09021181520014

**ABSTRAK**

*K-means* merupakan salah satu metode *clustering* yang hasil pengklasterannya sangat bergantung pada posisi *centroid* awal. *Centroid* awal yang umum digunakan pada metode *k-means* merupakan *centroid* awal yang dibangkitkan secara acak, *centroid* awal yang dibangkitkan secara acak sering kali menyebabkan *k-means* terjebak pada solusi lokal optimum, oleh karena itu pada penelitian ini akan menguji pengaruh *particle swarm optimization* yang merupakan salah satu algoritma optimasi yang dapat melakukan pencarian global dalam menentukan *centroid* awal *k-means*. *Clustering k-means* dengan *centroid* awal acak maupun dengan *centroid* awal dari perhitungan *particle swarm optimization* masing-masing diuji pada data reduksi dan tanpa reduksi. Berdasarkan hasil evaluasi hasil *clustering k-means* dengan *centroid* awal dari *particle swarm optimization* mampu meningkatkan kualitas *cluster*, baik jika diuji pada data reduksi maupun tanpa reduksi, yaitu dengan nilai persentase perubahan sebesar 43,8% pada data tanpa reduksi dimensi dan 53,4% pada data yang dilakukan reduksi dimensi. Meskipun dapat menambah waktu komputasi secara keseluruhan namun, *centroid* awal yang di peroleh dari *particle swarm optimization* membuat kompleksitas kerja *k-means* menjadi sederhana.

**Keywords:** *Clustering, K-Means, Particle Swarm Optimization, Principal Component Analysis, Davies Bouldin Index.*

Mengetahui,  
Ketua Jurusan Teknik Informatika,



Rifkie Primartha, M.T  
NIP. 198505102015041002

Indralaya, Oktober 2019  
Pembimbing,

Rifkie Primartha, M.T  
NIP. 198505102015041002

## KATA PENGANTAR



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

Dalam menyelesaikan Tugas Akhir ini banyak pihak yang telah memberikan bantuan dan dukungan baik secara langsung maupun secara tidak langsung. Penulis ingin menyampaikan rasa terima kasih kepada:

1. Allah Subhanahu wa Ta'ala yang telah melimpahkan kepada hamba keimanan, kesehatan, kecerdasan, kemudahan, dan kelancaran sehingga hamba dapat menyelesaikan tugas-tugas hamba sebagai seorang mahasiswa.
2. Nabi Muhammad Shallallahu 'Alaihi wa Sallam yang telah memberikan jalan yang terang dan memberikan ilmu yang bermanfaat kepada umatnya atas seizin Allah.
3. Bapak Jaidan Jauhari, S.Pd., M.T selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
4. Bapak Rifkie Primartha, M.T selaku Ketua Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya dan sekaligus pembimbing Tugas Akhir saya yang telah mengarahkan dan memberikan motivasi penulis dalam proses perkuliahan dan pengerjaan Tugas Akhir serta Ibu Hardini Novianti, M.T. selaku Sekretaris Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.
5. Ibu Novi Yusliani, M.T selaku dosen penguji I dan Bapak Kanda Januar Miraswan, M.T selaku dosen penguji II yang telah memberikan masukan dan dorongan dalam proses pengerjaan Tugas Akhir.
6. Seluruh dosen Program Studi Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.
7. Orang tua saya, Zauhari, M.Pd dan Nurmadiyah, S.Pd yang telah memberikan dorongan dan dukungan baik secara moril maupun materil, serta kepada ayuk saya Liza Risdiyanti, S.Pd beserta suaminya kak Senapati, adik saya Syahrizal Dian Nugraha, dan ponakan saya yang lagi lucu-lucunya dedek Ulya Nurhasanah.
8. Kak Ricy dan Mbak Winda serta seluruh staf tata usaha yang telah membantu dalam kelancaran proses administrasi dan akademik selama masa perkuliahan.
9. Devi Permata Hati, Wak Tirta, Bang Udin, M. Agung Pribadi, Kerel Khalif Afif, Albar, Genda Ananta, Gian Agnar, Rizky Blend, Bang Rasyid Hilmi, Rayyan, Arfah, Dian Ayu, Zakia, Destrilia, Rani dan teman-teman IF REG B 2015. Serta seluruh teman-teman Teknik Informatika yang telah berbagi cerita kehidupan bersama pada masa perkuliahan ini.
10. Website kursus gratis online pemrograman yang telah membantu penulis dalam menimba ilmu dalam pemrograman.

Penulis menyadari dalam penyusunan Tugas Akhir ini masih terdapat banyak kekurangan disebabkan keterbatasan pengetahuan dan pengalaman, oleh karena itu kritik dan saran yang membangun sangat diharapkan untuk kemajuan penelitian selanjutnya.

Akhir kata semoga Tugas Akhir ini dapat berguna dan bermanfaat bagi kita semua.

Indralaya, 5 November 2019

Ardi Wasila Chandra

## DAFTAR ISI

HALAMAN JUDUL .....	i
LEMBAR PENGESAHAN .....	ii
TANDA LULUS UJIAN .....	iii
HALAMAN PERNYATAAN .....	iv
MOTTO DAN PERSEMBAHAN .....	v
ABSTRACT.....	vi
ABSTRAK.....	vii
KATA PENGANTAR .....	viii
DAFTAR ISI.....	x
DAFTAR TABEL.....	xiii
DAFTAR GAMBAR .....	xvi
DAFTAR LAMPIRAN.....	xiv
BAB I.....	I-1
PENDAHULUAN .....	I-1
1.1    Pendahuluan.....	I-1
1.2    Latar Belakang.....	I-1
1.3    Rumusan Masalah.....	I-3
1.4    Tujuan Penelitian .....	I-3
1.5    Manfaat Penelitian .....	I-4
1.6    Batasan Masalah .....	I-4
1.7    Sistematika Penulisan .....	I-5
1.8    Kesimpulan .....	I-6
BAB II.....	<b>II-Error! Bookmark not defined.</b>
KAJIAN TEORITIS .....	<b>II-Error! Bookmark not defined.</b>
2.1    Pendahuluan.....	<b>II-Error! Bookmark not defined.</b>
2.2    Penelitian Terkait.....	<b>II-Error! Bookmark not defined.</b>
2.3    K-Means .....	<b>II-Error! Bookmark not defined.</b>
2.4. <i>Particle Swarm Optimizati</i> on (PSO) .....	<b>II-Error! Bookmark not defined.</b>
2.5.    Principal Component Analysis (PCA).....	<b>II-Error! Bookmark not defined.</b>
2.6.    Evaluasi Menggunakan <i>Davies Bouldin Index</i> (DBI) .....	<b>II-Error! Bookmark not defined.</b>
2.7    Kesimpulan .....	<b>II-Error! Bookmark not defined.</b>
BAB III .....	<b>III-Error! Bookmark not defined.</b>
METODOLOGI PENELITIAN.....	<b>III-Error! Bookmark not defined.</b>
3.1    Pendahuluan.....	<b>III-Error! Bookmark not defined.</b>
3.2    Pengumpulan Data.....	<b>III-Error! Bookmark not defined.</b>
3.3    Tahapan Penelitian.....	<b>III-Error! Bookmark not defined.</b>
3.3.1    Menentukan Ruang Lingkup dan Unit Penelitian .....	<b>III-Error! Bookmark not defined.</b>
3.3.2    Menemukan Dasar Teori yang Berkaitan dengan Penelitian .....	<b>III-Error! Bookmark not defined.</b>

3.3.3	Menetapkan Kerangka Kerja / Framework	III-Error! Bookmark not defined.
3.3.4	Membuat Kriteria Pengujian.....	III-Error! Bookmark not defined.
3.3.5	Menentukan Alat yang Digunakan dalam Pelaksanaan Penelitian .....	III-Error! Bookmark not defined.
3.3.6	Melakukan Pengujian Penelitian	III-Error! Bookmark not defined.
3.3.7	Melakukan Analisa Hasil Pengujian dan Membuat Kesimpulan Penelitian .....	III-Error! Bookmark not defined.
3.4	Metode Pengembangan Perangkat Lunak .	III-Error! Bookmark not defined.
3.4.1	Rational Unified Process (RUP).	III-Error! Bookmark not defined.
3.4.1.1	Fase Insepsi.....	III-Error! Bookmark not defined.
3.4.1.2	Fase Elaborasi .....	III-Error! Bookmark not defined.
3.4.1.3	Fase Konstruksi .....	III-Error! Bookmark not defined.
3.4.1.4	Fase Transisi .....	III-Error! Bookmark not defined.
3.5	Penjadwalan Penelitian .....	III-Error! Bookmark not defined.
3.7	Kesimpulan .....	III-Error! Bookmark not defined.
BAB IV		IV-Error! Bookmark not defined.
PENGEMBANGAN PERANGKAT LUNAK		IV-Error! Bookmark not defined.
4.1	Pendahuluan.....	IV-Error! Bookmark not defined.
4.2	Fase Insepsi .....	IV-Error! Bookmark not defined.
4.2.1	Pemodelan Bisnis.....	IV-Error! Bookmark not defined.
4.2.3	Analisis dan Desain.....	IV-Error! Bookmark not defined.
4.3	Fase Elaborasi .....	IV-Error! Bookmark not defined.
4.3.1	Pemodelan Bisnis.....	IV-Error! Bookmark not defined.
4.3.2	Kebutuhan Sistem .....	IV-Error! Bookmark not defined.
4.3.3	Analisis dan Desain .....	IV-Error! Bookmark not defined.
4.4	Fase Konstruksi .....	IV-Error! Bookmark not defined.
4.4.1	Pemodelan Bisnis.....	IV-Error! Bookmark not defined.
4.4.2	Kebutuhan Sistem .....	IV-Error! Bookmark not defined.
4.5.	Fase Transisi .....	IV-Error! Bookmark not defined.
4.5.1.	Pemodelan Bisnis.....	IV-Error! Bookmark not defined.
4.5.2.	Kebutuhan Sistem .....	IV-Error! Bookmark not defined.
4.5.3.	Pengujian .....	IV-Error! Bookmark not defined.
4.5.3.1.	Rencana Pengujian.....	IV-Error! Bookmark not defined.
4.5.3.2.	Implementasi Pengujian.....	IV-Error! Bookmark not defined.
4.6.	Kesimpulan .....	IV-Error! Bookmark not defined.
BAB V		V-Error! Bookmark not defined.
HASIL DAN ANALISIS PENELITIAN.....		V-Error! Bookmark not defined.
5.1	Pendahuluan.....	V-Error! Bookmark not defined.
5.2.	Konfigurasi Percobaan.....	V-Error! Bookmark not defined.
5.3.	Hasil Pengujian .....	V-Error! Bookmark not defined.
5.3.1.	Hasil Pengujian $k$ Optimum.....	V-Error! Bookmark not defined.
5.3.2.	Hasil Pengujian <i>Clustering K-Means</i> dengan <i>Centroid</i> Awal Acak.....	V-Error! Bookmark not defined.

5.3.3.	Hasil Pengujian Parameter <i>Particle Swarm Optimization</i> .....	V-Error! Bookmark not defined.
5.3.3.1	Hasil Pengujian Jumlah Iterasi .....	V-Error! Bookmark not defined.
5.3.3.2	Hasil Pengujian Jumlah Partikel .....	V-Error! Bookmark not defined.
5.3.3.3	Hasil Pengujian Nilai $c_1$ .....	V-Error! Bookmark not defined.
5.3.3.4	Hasil Pengujian Nilai $c_2$ .....	V-Error! Bookmark not defined.
5.3.4.	Hasil Pengujian <i>Clustering K-Means</i> dengan <i>Centroid</i> Awal dari Perhitungan <i>Particle Swarm Optimization</i> .	V-Error! Bookmark not defined.
5.4.	Olah Data .....	V-Error! Bookmark not defined.
5.4.1.	Olah Data Hasil Pengujian $k$ Optimum .....	V-Error! Bookmark not defined.
5.4.2.	Olah Data Hasil Perbandingan Nilai DBI...	V-Error! Bookmark not defined.
5.4.3.	Olah Data Hasil Perbandingan Waktu Komputasi .....	V-Error! Bookmark not defined.
5.4.4.	Olah Data Hasil Perbandingan Jumlah Iterasi ...	V-Error! Bookmark not defined.
5.5.	Analisis Hasil Pengujian .....	V-Error! Bookmark not defined.
5.5.1.	Analisis Hasil Pengujian Nilai $k$ optimum..	V-Error! Bookmark not defined.
5.5.2.	Analisis Hasil Pengujian Parameter <i>Particle Swarm Optimization</i> .....	V-Error! Bookmark not defined.
5.5.3.	Analisis Hasil Pengujian <i>Clustering K-Means</i> dengan <i>Centroid</i> Awal Acak pada Data Hasil Reduksi dan Tanpa Reduksi .....	V-Error! Bookmark not defined.
5.5.4.	Analisis Hasil Pengujian <i>Clustering K-Means</i> dengan <i>Centroid</i> Awal dari Perhitungan <i>Particle Swarm Optimization</i> pada Data Hasil Reduksi dan Tanpa Reduksi .....	V-Error! Bookmark not defined.
5.5.5.	Analisis Hasil Pengujian <i>Clustering K-Means</i> dengan <i>Centroid</i> Awal Acak dan <i>Centroid</i> Awal dari Perhitungan <i>Particle Swarm Optimization</i> pada Data Reduksi Maupun Tanpa Reduksi Dimensi .....	V-Error! Bookmark not defined.
5.6.	Kesimpulan .....	V-Error! Bookmark not defined.
BAB VI	.....	VI-Error! Bookmark not defined.
KESIMPULAN DAN SARAN	.....	VI-Error! Bookmark not defined.
6.1.	Pendahuluan .....	VI-Error! Bookmark not defined.
6.2.	Kesimpulan .....	VI-Error! Bookmark not defined.
6.3.	Saran .....	VI-Error! Bookmark not defined.
DAFTAR PUSTAKA	.....	xii

## DAFTAR TABEL

- Tabel III-1. Rancangan Tabel Rancangan Tabel Hasil Pengujian Penentuan Nilai  $k$  Optimum ..... **Error! Bookmark not defined.**
- Tabel III-2. Rancangan Tabel Hasil Pengujian Jumlah Iterasi pada Data yang Telah dilakukan Reduksi dan Tanpa Reduksi Dimensi ..... **Error! Bookmark not defined.**
- Tabel III-3. Rancangan Tabel Hasil Pengujian *learning rates* pada Data yang Telah dilakukan Reduksi dan Tanpa Reduksi Dimensi ..... **Error! Bookmark not defined.**
- Tabel III-4. Rancangan Tabel Hasil Perbandingan Algoritma *K-Means* dengan *Centroid* yang ditetapkan dengan *Particle Swarm Optimization* dan *K-Means* dengan *Centroid random* yang Telah di reduksi dan Tanpa Reduksi Dimensi. .... **Error! Bookmark not defined.**
- Tabel III-5. Penjadwalan Penelitian dalam Bentuk *Work Breakdown Structure* (WBS) ..... **Error! Bookmark not defined.**
- Tabel IV-1. Definisi Aktor ..... **Error! Bookmark not defined.**
- Tabel IV-2. Definisi *Use Case* ..... **Error! Bookmark not defined.**
- Tabel IV-3. Kebutuhan Fungsional ..... **Error! Bookmark not defined.**
- Tabel IV-4. Kebutuhan Non-Fungsional ..... **Error! Bookmark not defined.**
- Tabel IV-5. Skenario *Use Case* Melakukan Konversi Data Teks ke Numerik ..... **Error! Bookmark not defined.**
- Tabel IV-6. Skenario *Use Case* Melakukan Clustering K-means dengan Centroid Awal Acak ..... **Error! Bookmark not defined.**
- Tabel IV-7. Skenario *Use Case* Melakukan Clustering K-means dengan Centroid Awal dari *Particle Swarm Optimization* ..... **Error! Bookmark not defined.**
- Tabel IV-8. Skenario *Use Case* Melakukan Clustering K-Means dengan Centroid Awal Acak Menggunakan Data Hasil Reduksi. **Error! Bookmark not defined.**
- Tabel IV-9. Skenario *Use Case* Melakukan Clustering K-means dengan Centroid Awal dari *Particle Swarm Optimization* Menggunakan Data Hasil Reduksi ..... **Error! Bookmark not defined.**
- Tabel IV-10. Implementasi Kelas ..... **Error! Bookmark not defined.**
- Tabel IV-11. Rencana Pengujian *Use Case* Melakukan Konversi Data ..... **Error! Bookmark not defined.**
- Tabel IV-12. Rencana Pengujian *Use Case* Melakukan Clustering K-means dengan Centroid Awal Acak ..... **Error! Bookmark not defined.**

- Tabel IV-13. Rencana Pengujian *Use Case* Melakukan *Clustering K-means* dengan *Centroid* Awal dari Perhitungan *Particle Swarm Optimization* .....**Error! Bookmark not defined.**
- Tabel IV-14. Rencana Pengujian *Use Case* Melakukan *Clustering K-means* dengan *Centroid* Awal Acak Menggunakan Data Hasil Reduksi**Error! Bookmark not defined.**
- Tabel IV-15. Rencana Pengujian *Use Case* Melakukan *Clustering K-means* dengan *Centroid* Awal dari *Particle Swarm Optimization* Menggunakan Data Hasil Reduksi.....**Error! Bookmark not defined.**
- Tabel IV – 16. Pengujian Fitur Konversi Data .....**Error! Bookmark not defined.**
- Tabel IV-17. Pengujian Fitur *Clustering K-Means* dengan *Centroid* Awal Acak .....**Error! Bookmark not defined.**
- Tabel IV-18. Pengujian Fitur *Clustering k-means* dengan *Centroid* Awal dari Perhitungan *Particle Swarm Optimization*..... **Error! Bookmark not defined.**
- Tabel IV-19. Pengujian Fitur *Clustering k-means* dengan *Centroid* Awal Acak Menggunakan Data Hasil Reduksi Dimensi ..... **Error! Bookmark not defined.**
- Tabel IV-20. Pengujian Fitur *Clustering K-means* dengan *Centroid* Awal dari Perhitungan *Particle Swarm Optimization* Menggunakan Data Hasil Reduksi Dimensi .....**Error! Bookmark not defined.**
- Tabel V-2. Hasil Pengujian *Clustering K-Means* dengan *Centroid* Awal Acak .....**Error! Bookmark not defined.**
- Tabel V-3. Hasil Pengujian Jumlah Iterasi .....**Error! Bookmark not defined.**
- Tabel V-4. Hasil Pengujian Jumlah *Partikel* .....**Error! Bookmark not defined.**
- Tabel V-5. Hasil Pengujian Nilai  $c_1$  .....**Error! Bookmark not defined.**
- Tabel V-6. Hasil Pengujian Nilai  $c_2$  .....**Error! Bookmark not defined.**
- Tabel V-7. Hasil Pengujian *Clustering K-Means* dengan *Centroid* Awal dari Perhitungan *Particle Swarm Optimization*..... **Error! Bookmark not defined.**
- Tabel V-8. Deskripsi Nilai DBI .....**Error! Bookmark not defined.**
- Tabel V-9. Hasil Uji Normalitas Nilai DBI .....**Error! Bookmark not defined.**
- Tabel V-10. Hasil Uji Kruskal-Wallis Perbandingan Nilai DBI **Error! Bookmark not defined.**
- Tabel V – 11. Hasil Uji Mann-Whitney Perbandingan Nilai DBI pada Data Hasil Pengklasteran K-Means dan PSO+K-Means **Error! Bookmark not defined.**

- Tabel V – 12. Hasil Uji Mann-Whitney Perbandingan Nilai DBI pada Data Hasil Pengklasteran K-Means dan PCA+K-Means **Error! Bookmark not defined.**
- Tabel V – 13. Hasil Uji Mann-Whitney Perbandingan Nilai DBI pada Data Hasil Pengklasteran K-Means dan PCA+PSO+K-Means ..... **Error! Bookmark not defined.**
- Tabel V – 14. Hasil Uji Mann-Whitney Perbandingan Nilai DBI pada Data Hasil Pengklasteran PSO+K-Means dan PCA+K-Means ..... **Error! Bookmark not defined.**
- Tabel V – 15. Hasil Uji Mann-Whitney Perbandingan Nilai DBI pada Data Hasil Pengklasteran PSO+K-Means dan PCA+PSO+K-Means..... **Error! Bookmark not defined.**
- Tabel V – 16. Hasil Uji Mann-Whitney Perbandingan Nilai DBI pada Data Hasil Pengklasteran PCA+K-Means dan PCA+PSO+K-Means ..... **Error! Bookmark not defined.**
- Tabel V-17. Deskripsi Waktu Komputasi.....**Error! Bookmark not defined.**
- Tabel V-18. Hasil Uji Normalitas Waktu Komputasi..... **Error! Bookmark not defined.**
- Tabel V – 20. Hasil Uji Mann-Whitney Perbandingan Waktu Komputasi pada Data Hasil Pengklasteran K-Means dan PSO+K-Means ..... **Error! Bookmark not defined.**
- Tabel V – 21. Hasil Uji Mann-Whitney Perbandingan Waktu Komputasi pada Data Hasil Pengklasteran K-Means dan PCA+K-Means..... **Error! Bookmark not defined.**
- Tabel V – 22. Hasil Uji Mann-Whitney Perbandingan Waktu Komputasi pada Data Hasil Pengklasteran K-Means dan PCA+PSO+K-Means .....**Error! Bookmark not defined.**
- Tabel V – 23. Hasil Uji Mann-Whitney Perbandingan Waktu Komputasi pada Data Hasil Pengklasteran PSO+K-Means dan PCA+K-Means .....**Error! Bookmark not defined.**
- Tabel V – 24. Hasil Uji Mann-Whitney Perbandingan Waktu Komputasi pada Data Hasil Pengklasteran PSO+K-Means dan PCA+PSO+K-Means .....**Error! Bookmark not defined.**
- Tabel V – 25. Tabel Hasil Uji Mann-Whitney Perbandingan Waktu Komputasi pada Data Hasil Pengklasteran PCA+K-Means dan PCA+PSO+K-Means .....**Error! Bookmark not defined.**
- Tabel V-26. Deskripsi Jumlah Iterasi .....**Error! Bookmark not defined.**

- Tabel V-27. Hasil Uji Normalitas Jumlah Iterasi ..... **Error! Bookmark not defined.**
- Tabel V-28. Hasil Uji Kruskal-Wallis Perbandingan Jumlah Iterasi ..... **Error! Bookmark not defined.**
- Tabel V – 29. Hasil Uji Mann-Whitney Perbandingan Jumlah Iterasi pada Data Hasil Pengklasteran K-Means dan PSO+K-Means ..... **Error! Bookmark not defined.**
- Tabel V – 30. Hasil Uji Mann-Whitney Perbandingan Jumlah Iterasi pada Data Hasil Pengklasteran K-Means dan PCA+K-Means ..... **Error! Bookmark not defined.**
- Tabel V – 33. Tabel Hasil Uji Mann-Whitney Perbandingan Jumlah Iterasi pada Data Hasil Pengklasteran PSO+K-Means dan PCA+PSO+K-Means ..... **Error! Bookmark not defined.**
- Tabel V – 34. Tabel Hasil Uji Mann-Whitney Perbandingan Jumlah Iterasi pada Data Hasil Pengklasteran PCA+K-Means dan PCA+PSO+K-Means ..... **Error! Bookmark not defined.**

## DAFTAR GAMBAR

- Gambar III-1. Diagram Tahapan Pengujian Penelitian..... **Error! Bookmark not defined.**
- Gambar III-3. Arsitektur Rational Unified Process (RUP).. **Error! Bookmark not defined.**
- Gambar IV-1. Diagram *Use Case* .....**Error! Bookmark not defined.**
- Gambar IV-2. Diagram *Sequence* Melakukan Konversi Data....**Error! Bookmark not defined.**
- Gambar IV-3. Diagram *Sequence* Melakukan Clustering K-Means dengan Centroid Acak.....**Error! Bookmark not defined.**
- Gambar IV-4. Diagram *Sequence* Melakukan Clustering K-Means dengan Centroid Awal dari PSO.....**Error! Bookmark not defined.**
- Gambar IV-5. Diagram *Sequence* Melakukan *Clustering K-Means* dengan *Centroid* Acak Menggunakan Data yang Telah direduksi Menggunakan *Particle Swarm Optimization* **Error! Bookmark not defined.**
- Gambar IV-6. Diagram *Sequence* Melakukan *Clustering K-means* dengan *Centroid* Awal dari Perhitungan *Particle Swarm Optimization* dan Menggunakan Data yang Telah direduksi Menggunakan PCA .....**Error! Bookmark not defined.**
- Gambar IV-7. Rancangan Antar Muka Perangkat Lunak.... **Error! Bookmark not defined.**
- Gambar IV-8. Rancangan Fitur Muat Dokumen ...**Error! Bookmark not defined.**
- Gambar IV-9. Rancangan Fitur Input Jumlah Cluster ..... **Error! Bookmark not defined.**
- Gambar IV-10. Rancangan Input Parameter *Particle Swarm Optimization*.. **Error! Bookmark not defined.**
- Gambar IV-11. Rancangan Fitur Metode *Clustering* ..... **Error! Bookmark not defined.**
- Gambar IV-12. Rancangan Fitur Tombol Mulai Proses ..... **Error! Bookmark not defined.**
- Gambar IV-13. Rancangan Fitur Menampilkan Waktu Komputasi dan nilai DBI .....**Error! Bookmark not defined.**
- Gambar IV-14. Rancangan Fitur Menampilkan *Centroid* Akhir **Error! Bookmark not defined.**

- Gambar IV-15. Rancangan Fitur Menampilkan Hasil *Clustering* Dokumen **Error! Bookmark not defined.**
- Gambar IV-16. Diagram Aktivitas Konversi Data **Error! Bookmark not defined.**
- Gambar IV-17. Diagram Aktivitas Melakukan *Clustering K-means* dengan *Centroid* Awal Acak.....**Error! Bookmark not defined.**
- Gambar IV-18. Diagram Aktivitas Melakukan *Clustering K-means* dengan *Centroid* Awal dari Perhitungan *Particle Swarm Optimization* .....**Error! Bookmark not defined.**
- Gambar IV-19. Diagram Aktivitas Melakukan *Clustering K-means* dengan *Centroid* Awal Acak Menggunakan Data Hasil Reduksi *Principal Component Analysis* .....**Error! Bookmark not defined.**
- Gambar IV-20. Diagram Aktivitas Melakukan *Clustering K-means* dengan *Centroid* Awal dari Perhitungan *Particle Swarm Optimization* Menggunakan Data Hasil Reduksi *Principal Component Analysis* .....**Error! Bookmark not defined.**
- Gambar IV-21. Diagram Kelas .....**Error! Bookmark not defined.**
- Gambar V-1. Kurva Hasil Penentuan  $k$  Optimum .**Error! Bookmark not defined.**

# BAB I

## PENDAHULUAN

### 1.1 Pendahuluan

Pada bagian pendahuluan ini menjelaskan secara keseluruhan penelitian yang dilakukan antara lain mengenai latar belakang masalah, rumusan masalah, tujuan dan manfaat penelitian serta batasan masalah.

Permasalahan dimulai dengan penjelasan mengenai *data mining*, serta penelitian yang berkaitan dengan metode *clustering* yaitu *k-means* dan dioptimasi menggunakan *particle swarm optimization* dalam menentukan initial centroid terhadap hasil pengklasteran data berdimensi tinggi.

### 1.2 Latar Belakang

*Data mining* menggambarkan suatu proses untuk menemukan suatu pola pada data yang menghasilkan informasi baru (Kurniasih, Kumaladewi, & Katjong, 2001). Salah satu metode yang ada pada *data mining* ialah pengelompokan yang bersifat *unsupervised learning*. Struktur sederhana dan laju konvergensi cepat menjadi salah satu kelebihan yang dimiliki oleh algoritma *k-means* yang mana sering diterapkan dibidang *data mining* (Lin et al., 2012). Pada algoritma pengelompokan, *k-means* terkenal dalam hal efisiensi perhitungan tetapi lemah dalam hal menginisialisasi *centroid* (Majhi & Biswal, 2018). Untuk meningkatkan akurasi dari algoritma *k-means*, *particle swarm optimization (PSO)* dapat menjadi alternatif dalam hal optimasi. Karena pada penggambarannya *particle swarm*

*optimization (PSO)* menghasilkan partikel-partikel acak yang bertemu secara keseluruhan untuk memperbarui posisi dan kecepatan secara bersama (Gupta, Science, Pattanaik, Singh, & Science, 2017).

Berdasarkan penelitian Lin et al. (2014) mengatakan bahwa penggunaan algoritma *k-means* dengan *particle swarm optimization (PSO)* membuat hasil *cluster analysis* lebih baik dan lebih stabil apabila dataset bersifat kompleks karena dapat mengintegrasikan dengan baik nilai *k-centroid*. Penelitian yang lain juga dilakukan oleh Science Gupta et al. (2017) yang mengatakan bahwa tingkat akurasi *k-means* dengan *particle swarm optimization* meningkat, ini dibuktikan dengan adanya empat data uji yang masing-masing tingkat akurasinya sebesar 72.3404 menjadi 74.4681, 11.94 menjadi 14.94, 20.35 menjadi 25.2, dan 50.6667 menjadi 96. Algoritma *k-means* sering kali terjebak dalam *curse of dimentionality* pada data yang bersifat dimensi tinggi. Untuk mengatasi hal tersebut maka harus dilakukan reduksi dimensi. Pada penelitian yang dilakukan (Jambak, Mohammed, Hidayati, Efendi, & Primartha, 2017) dengan melakukan reduksi dimensi metode-metode pada clustering, dapat memperbaiki hasil *clustering k-means*.

Berdasarkan penelitian-penelitian diatas, maka penelitian ini akan menguji algoritma *k-means* yang dioptimasi menggunakan algoritma *particle swarm optimization* yang tujuannya untuk menguji pengaruh penentuan *initial centroid* terhadap pengklasteran data berdimensi tinggi.

### 1.3 Rumusan Masalah

Rumusan masalah pada penelitian ini adalah pengaruh dari *particle swarm optimization* dalam menentukan *initial centroid* terhadap hasil pengklasteran data berdimensi tinggi.

Untuk menjawab rumusan masalah tersebut, diuraikan beberapa pertanyaan penelitian sebagai berikut :

1. Bagaimana mekanisme *particle swarm optimization* dalam menentukan nilai *centroid* pada *k-means*?
2. Bagaimana pengaruh *particle swarm optimization* dalam menentukan *initial* terhadap *clustering* data berdimensi tinggi yang direduksi dan tanpa direduksi?
3. Bagaimana perbandingan hasil *clustering* pada *k-means* dengan *initial centroid* yang ditentukan menggunakan *particle swarm optimization* dan *k-means* dengan *random centroid* terhadap *clustering* dokumen berdimensi tinggi yang direduksi dan tanpa direduksi?

### 1.4 Tujuan Penelitian

Tujuan dilakukannya penelitian adalah sebagai berikut:

1. Mengetahui mekanisme *particle swarm optimization* dalam menentukan nilai *centroid* pada *k-means*.
2. Mengetahui pengaruh *particle swarm optimization* dalam menentukan *initial* terhadap *clustering* data berdimensi tinggi yang direduksi dan tanpa direduksi.

3. Mengetahui perbandingan hasil *clustering* pada *k-means* dengan *initial centroid* yang ditentukan menggunakan *particle swarm optimization* dan *k-means* dengan *random centroid* terhadap *clustering* dokumen berdimensi tinggi yang direduksi dan tanpa direduksi.

### **1.5 Manfaat Penelitian**

Manfaat yang dapat diperoleh dari penelitian ini adalah sebagai berikut :

1. Memahami mekanisme *particle swarm optimization* dalam menentukan nilai *centroid* pada *k-means*;
2. Memahami pengaruh *particle swarm optimization* dalam menentukan *initial* terhadap *clustering* data berdimensi tinggi yang direduksi dan tanpa direduksi;
3. Memahami perbandingan hasil *clustering* pada *k-means* dengan *initial centroid* yang ditentukan menggunakan *particle swarm optimization* dan *k-means* dengan *random centroid* terhadap *clustering* dokumen berdimensi tinggi yang direduksi dan tanpa direduksi.

### **1.6 Batasan Masalah**

Batasan masalah pada penelitian ini adalah sebagai berikut :

1. Data yang digunakan adalah dokumen jurnal berbahasa Indonesia yang diperoleh dari situs *garuda.ristekdikti.go.id*.
2. Untuk proses reduksi dimensi metode yang digunakan adalah *Principal Component Analysis* (PCA).

3. Hasil *clustering* akan diukur menggunakan metode *Davies Bouldin index (DBI)*.

## **1.7 Sistematika Penulisan**

Penyusunan skripsi ini disusun dengan sistematika penulisan sebagai berikut :

### **BAB I. PENDAHULUAN**

Pada bab ini diuraikan mengenai latar belakang, perumusan masalah, tujuan dan manfaat penelitian, batasan masalah atau ruang lingkup serta sistematika penulisan.

### **BAB II. KAJIAN LITERATUR**

Pada bab ini membahas dasar-dasar teori yang digunakan didalam penelitian, seperti definisi dan algoritma optimasi, teknik pengklasteran dan perhitungan evaluasinya serta beberapa kajian literatur mengenai penelitian lain yang relevan pada penelitian ini.

### **BAB III. METODOLOGI PENELITIAN**

Pada bab ini membahas mengenai tahapan yang akan dilaksanakan pada penelitian ini. Masing-masing rencana tahapan penelitian dideskripsikan dengan rinci mengacu pada suatu kerangka kerja. Diakhir bab, berisi perancangan manajemen proyek pelaksanaan penelitian.

### **BAB IV. PENGEMBANGAN PERANGKAT LUNAK**

Pada bab ini membahas mengenai analisis dan perancangan perangkat lunak yang akan digunakan sebagai alat penelitian nantinya. Diawali dari pengumpulan dan analisa kebutuhan, rancangan dan konstruksi perangkat lunak serta pengujian untuk memastikan semua kebutuhan perangkat lunak sesuai dengan kebutuhan. Penyusunan pada bab ini memiliki kerangka penulisan dengan fase-fase dan elemen-elemen pengembangan perangkat lunak bersifat berorientasi objek.

## **BAB V. HASIL DAN ANALISA PENELITIAN**

Pada bab ini diuraikan hasil pengujian berdasarkan langkah-langkah yang telah direncanakan. Tabel hasil pengujian serta analisisnya disajikan sebagai basis dari kesimpulan yang akan diambil dalam penelitian ini.

## **BAB VI. KESIMPULAN DAN SARAN**

Pada bab ini berisi kesimpulan dari semua uraian-uraian pada bab-bab sebelumnya dan juga saran-saran yang diharapkan berguna dalam melakukan pengembangan perangkat lunak ini selanjutnya.

### **1.8 Kesimpulan**

Dari penjelasan diatas, telah diuraikan latar belakang dan rumusan masalah, tujuan dan manfaat penelitian, batasan masalah serta sistematika penulisan. Adapun algoritma pengelompokkan yang digunakan yaitu *k-means* yang dioptimasi dengan

menggunakan algoritma *particle swarm optimization* (PSO) yang mana algoritma *particle swarm optimization* (PSO) untuk mengatasi tingkat akurasi dalam menentukan *centroid k-means* yang dapat melakukan pencarian secara menyeluruh.

## DAFTAR PUSTAKA

- Gupta, A., Science, C., Pattanaik, V., Singh, M., & Science, C. (2017). using PSO Algorithm, 228–233.
- Jambak, M. I., Mohammed, F., Hidayati, N., Efendi, R., & Primartha, R. (2017). *The Impacts of Singular Value Decomposition Algorithm Toward Indonesian Language Text Documents Clustering* (Vol. 1). Springer International Publishing. <https://doi.org/10.1007/978-3-319-99007-1>
- Kaushik, A., & Ghosh, S. (2014). A Survey on Optimization Approaches to K-Means Clustering using Simulated Annealing, *847*(3), 845–847.
- Khotimah, B. K., Irhamni, F., & Sundarwati, T. (2016). A GENETIC ALGORITHM FOR OPTIMIZED INITIAL CENTERS K-MEANS CLUSTERING IN SMEs. *Journal of Theoretical and Applied Information Technology*, *1590*(1). Retrieved from [www.jatit.org](http://www.jatit.org)
- Kurniasih, F., Kumaladewi, N., & Katjong, dan B. La. (2001). Analisa Dan Perancangan Data Mining Dengan Metode Market Basket Analysis Untuk Analisa Pola Belanja Konsumen pada Tendencies Store, *5*(1), 1–10. Retrieved from <http://web.stanford.edu/group/4D/%5Chttp://web.stanford.edu/group/4D/issues/wishlist.shtml>
- Lin, Y., Tong, N., Shi, M., Fan, K., Yuan, D., Qu, L., & Fu, Q. (2012). K-means optimization clustering algorithm based on particle swarm optimization and multiclass merging. *Advances in Intelligent and Soft Computing*, *168 AISC(VOL. 1)*, 569–578. [https://doi.org/10.1007/978-3-642-30126-1\\_90](https://doi.org/10.1007/978-3-642-30126-1_90)
- Majhi, S. K., & Biswal, S. (2018). Optimal cluster analysis using hybrid K-Means and Ant Lion Optimizer. *Karbala International Journal of Modern Science*. <https://doi.org/10.1016/j.kijoms.2018.09.001>

- Raval, U. R., & Jani, C. (2016). Implementing & Improvisation of K-means Clustering Algorithm. *International Journal of Computer Science and Mobile Computing*, 55(5), 191–203. Retrieved from <http://www.ijcsmc.com/docs/papers/May2016/V5I5201647.pdf>
- Rendón, E., Abundez, I., Arizmendi, A., & Quiroz, E. M. (2011). Internal versus External cluster validation indexes, 5(1).
- Roiha, N. U., Suprpto, Y. K., & Wibawa, A. D. (2016a). The optimization of the weblog central cluster using the genetic K-means algorithm. In *2016 International Seminar on Application for Technology of Information and Communication (ISEmantic)* (pp. 278–284). IEEE. <https://doi.org/10.1109/ISEMANTIC.2016.7873851>
- Roiha, N. U., Suprpto, Y. K., & Wibawa, A. D. (2016b). The optimization of the weblog central cluster using the genetic K-means algorithm. In *2016 International Seminar on Application for Technology of Information and Communication (ISEmantic)* (pp. 278–284). IEEE. <https://doi.org/10.1109/ISEMANTIC.2016.7873851>
- Santoso, B., Cholissodin, I., & Setiawan, B. D. (2017). Optimasi K-Means untuk Clustering Kinerja Akademik Dosen Menggunakan Algoritme Genetika, *I*(12), 1652–1659.
- Taufik, M., Darwiyanto, E. and P, S. Y. (2015) ‘Analisis dan Implementasi Perancangan Metode Rational Unified Process pada layanan SDB dan Metode Pengujian Product Metric pada Bank Mandiri Cabang Palu Sam Ratulangi Analysis and Implementation Rational Unified Process Method In SDB Service and Product Metric Testing Method In Bank Mandiri Branch Palu Sam Ratulangi’, 2(3), pp. 7693–7704.
- Younus, Z. S., Mohamad, D., Saba, T., Alkawaz, M. H., Rehman, A., Al-Rodhaan, M., & Al-Dhelaan, A. (2014). Content-based image retrieval using PSO and k-means clustering algorithm. *Arabian Journal of Geosciences*, (Salamah 2010). <https://doi.org/10.1007/s12517-014-1584-7>