

OPTIMASI NILAI KLASTER PADA ALGORITMA K-MEANS  
MENGUNAKAN ALGORITMA *FIREFLY*

*Diajukan sebagai Syarat untuk Menyelesaikan  
Pendidikan Program Strata-1 pada  
Jurusan Teknik Informatika*



Oleh:

Febiyanti W  
NIM: 09021282025094

**Jurusan Teknik Informatika**

**FAKULTAS ILMU KOMPUTER UNIVERSITAS SRIWIJAYA**

**2024**

## LEMBAR PENGESAHAN SKRIPSI

### OPTIMASI NILAI KLASTER PADA ALGORITMA K-MEANS MENGUNAKAN ALGORITMA *FIREFLY*

Oleh:

Febiyanti W  
NIM : 09021282025094


Palembang, 1 Juli 2024

Pembimbing I

Mengetahui,  
Ketua Jurusan Teknik Informatika



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



Dian Palupi Rini, M.Kom., Ph.D.  
NIP.197802232006042002

## TANDA LULUS UJIAN SIDANG SKRIPSI

Pada hari Senin tanggal 1 Juli 2024 telah dilaksanakan ujian sidang skripsi oleh Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.

Nama : Febiyanti W  
NIM : 09021282025094  
Judul : Optimasi Nilai Klaster Pada Algoritma K-Means Menggunakan Algoritma *Firefly*.

Dan dinyatakan LULUS.

1. Ketua Penguji

Alvi Syahrini Utami, S.Si., M.Kom.  
NIP. 197812222006042003



2. Penguji

Mastura Diana Marieska, M.T.  
NIP. 198603212018032001



3. Pembimbing I

Dian Palupi Rini, M.Kom., Ph.D.  
NIP. 197802232006042002



Mengetahui,  
Ketua Jurusan Teknik Informatika



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

## HALAMAN PERNYATAAN

Yang bertanda tangan dibawah ini:

Nama : Febiyanti W

NIM : 09021282025094

Program Studi : Teknik Informatika

Judul : Optimasi Nilai Klaster Pada Algoritma K-Means Menggunakan Algoritma *Firefly*.

Hasil Pengecekan *Software iThenticate/Turnitin*: 17%

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 Universitas Sriwijaya sesuai dengan ketentuan yang berlaku.

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



Palembang, 1 Juli 2024



METERA  
TEMPEL  
ED 79AK0762568998

Febiyanti W  
NIM 09021282025094

## **MOTTO DAN PERSEMBAHAN**

**“Make your own journey as you go,  
And let your dreams lead the way.”**

**Kupersembahkan Karya Tulis Ini Kepada:**

- **Keluarga**
- **Teman-teman Seperjuangan**
- **Fakultas Ilmu Komputer**
- **Universitas Sriwijaya**

# OPTIMIZATION OF CLUSTER VALUES IN THE K-MEANS ALGORITHM USING THE FIREFLY ALGORITHM

By

Febiyanti W  
09021282025094

## ABSTRACT

K-Means is one of the most commonly used clustering algorithms for grouping data with similarities within a cluster. However, there are limitations in K-Means clustering, such as clustering results being initialized based on random centroid points and the number of clusters used. To improve the performance of the K-Means algorithm, the Firefly Algorithm is used for clustering optimization. The Firefly Algorithm offers flexibility in parameter determination and can deliver excellent performance. In the testing phase, the best optimization values with the Firefly Algorithm were obtained with the number of iterations = 60,  $\alpha = 0.1$ ,  $\beta_0 = 0.1$ , and  $\gamma = 0.01$ . This study shows that optimizing the Firefly Algorithm for K-Means clustering can improve clustering results, using the Silhouette Score as a benchmark. The closer the Silhouette Score is to one, the better the clustering result. The Silhouette Score for K-Means was 0.381, while the result for K-Means clustering optimized with the Firefly Algorithm was 0.431.

Keywords: Clustering, Optimization, K-Means, Firefly Algorithm



Palembang, 1 July 2024  
Supervisor I

Dian Palupi Rini, M.Kom., Ph.D.  
NIP.197802232006042002



# OPTIMASI NILAI KLASSTER PADA ALGORITMA K-MEANS MENGUNAKAN ALGORITMA *FIREFLY*

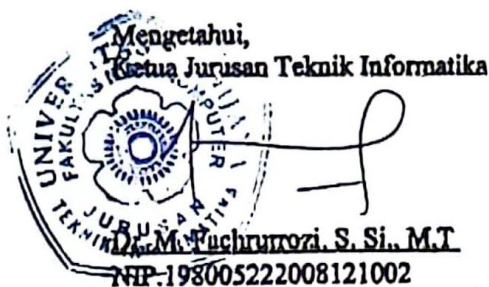
Oleh

**Febiyanti W**  
09021282025094

## ABSTRAK

K-Means adalah salah satu algoritma *clustering* yang umum digunakan untuk mengelompokkan data-data yang memiliki kesamaan dalam suatu *cluster*. Namun, terdapat keterbatasan dalam *clustering* K-Means, seperti hasil *clustering* yang diinisialisasi berdasarkan titik *centroid* acak dan jumlah *cluster* yang digunakan. Untuk itu, demi meningkatkan kinerja algoritma K-Means, digunakanlah algoritma *Firefly* sebagai optimasi *clustering*. Algoritma *Firefly* memiliki fleksibilitas dalam penentuan parameter dan dapat memberikan performa yang bagus. Pada pengujiannya nilai optimasi terbaik dengan Algoritma *Firefly* didapatkan dengan jumlah iterasi = 60,  $\alpha = 0,1$ ,  $\beta_0 = 0,1$ , dan  $\gamma = 0,01$ . Penelitian ini menunjukkan bahwa hasil dari optimasi algoritma *Firefly* terhadap *clustering* K-Means dapat meningkatkan hasil *clustering* dengan melihat *Silhouette Score* sebagai patokannya. Semakin nilainya mendekati satu maka, semakin baik hasil *clustering* tersebut. Didapatkan hasil *Silhouette Score* untuk K-Means = 0,381 dan hasil dari *Clustering* K-Means dengan optimasi Algoritma *Firefly* = 0,431.

Kata kunci: *Clustering*, Optimasi, K-Means, Algoritma *Firefly*



Palembang, 1 Juli 2024

Pembimbing I



Dian Palupi Kani, M.Kom., Ph.D.  
NIP.197802232006042002

## KATA PENGANTAR

Puji Syukur penulis ucapkan ke hadirat Tuhan Yang Maha Esa atas rahmat dan karunianya sehingga penulis dapat menyelesaikan skripsi ini dengan baik. Skripsi yang berjudul “Optimasi Nilai Klaster Pada Algoritma K-Means menggunakan Algoritma *Firefly*” berfungsi sebagai salah satu syarat untuk menyelesaikan program Strata-1 Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.

Penulis ingin memberikan ucapan terima kasih secara spesial terhadap dukungan dan bantuan secara langsung dan tidak langsung kepada pihak-pihak berikut ini:

1. Keluarga tercinta yang telah memberikan dukungan serta doa kepada penulis sehingga penulis dapat menyelesaikan skripsi ini.
2. Bapak Prof. Erwin, S.Si., M.Si selaku Dekan Fakultas Ilmu Komputer.
3. Bapak Dr. M. Fachrurrozi, S. Si., M.T selaku Ketua Jurusan Teknik Informatika.
4. Ibu Dian Palupi Rini, M.Kom., Ph.D. selaku Dosen Pembimbing yang telah memberikan arahan, masukan, kritik dan saran terhadap tugas akhir penulis ini.
5. Ibu Mastura Diana Marieska, M.T. selaku dosen penguji yang telah memberikan masukan dan pengetahuan kepada penulis.
6. Seluruh Bapak dan Ibu Dosen serta Staff Teknik Informatika Fakultas Ilmu



Komputer Universitas Sriwijaya yang telah memberikan ilmu dan dukungan kepada penulis selama masa perkuliahan.

7. Kezia Kurniati, Helen Vania dan Hanny Putri Gayatri selaku teman dekat yang telah memberikan waktu dan dukungan kepada penulis selama perkuliahan hingga skripsi ini terselesaikan.
8. Serta semua pihak yang telah mendukung dalam penyelesaian tugas akhir ini terima kasih banyak atas semua doa dan dukungannya.

Penulis menyadari bahwa dalam penyusunan tugas akhir ini masih terdapat kekurangan yang dikarenakan keterbatasan pengetahuan serta pengalaman penulis. Oleh karena itu, penulis mengharapkan kritik dan saran yang membangun untuk dapat menyempurnakan skripsi ini. Akhir kata, semoga skripsi ini dapat bermanfaat bagi kita semua.

Palembang, 1 Juli 2024



Febiyanti W

## DAFTAR ISI

Halaman

<b>HALAMAN JUDUL .....</b>	<b>i</b>
<b>HALAMAN PENGESAHAN.....</b>	<b>ii</b>
<b>HALAMAN PERSETUJUAN KOMISI PENGUJI.....</b>	<b>iii</b>
<b>HALAMAN PERNYATAAN.....</b>	<b>iv</b>
<b>HALAMAN MOTTO DAN PERSEMBAHAN.....</b>	<b>v</b>
<b>ABSTRACT .....</b>	<b>vi</b>
<b>ABSTRAK .....</b>	<b>vii</b>
<b>KATA PENGANTAR.....</b>	<b>viii</b>
<b>DAFTAR ISI.....</b>	<b>x</b>
<b>DAFTAR TABEL .....</b>	<b>xii</b>
<b>DAFTAR GAMBAR.....</b>	<b>xiii</b>
<b>DAFTAR PUSTAKA.....</b>	<b>xiv</b>
<b>DAFTAR LAMPIRAN .....</b>	<b>xv</b>
<b>BAB I PENDAHULUAN.....</b>	<b>I-1</b>
1.1 Pendahuluan.....	I-1
1.2 Latar Belakang Masalah.....	I-1
1.3 Rumusan Masalah .....	I-3
1.4 Tujuan Penelitian .....	I-3
1.5 Manfaat Penelitian .....	I-3
1.6 Batasan Masalah .....	I-4
1.7 Sistematika Penulisan .....	I-4
1.8 Kesimpulan .....	I-5
<b>BAB II KAJIAN LITERATUR.....</b>	<b>II-1</b>
2.1 Pendahuluan.....	II-1

2.2 Landasan Teori .....	II-1
2.2.1 <i>Data Mining</i> .....	II-1
2.2.2 <i>Clustering</i> .....	II-3
2.2.3 Jenis Metode <i>Clustering</i> .....	II-4
2.2.4 K-Means .....	II-5
2.2.5 <i>Silhouette Coefficient</i> .....	II-9
2.2.6 Algoritma <i>Firefly</i> .....	II-11
2.3 Penelitian yang Relevan .....	II-15
2.4 Kesimpulan.....	II-15
<b>BAB III METODOLOGI PENELITIAN .....</b>	<b>III-1</b>
3.1 Pendahuluan .....	III-1
3.2 Pengumpulan Data .....	III-1
3.3 Jenis dan Sumber Data .....	III-1
3.3.1 Metode Pengumpulan Data .....	III-2
3.4 Tahapan Penelitian .....	III-2
3.4.1 Kerangka Kerja .....	III-4
3.4.2 Kriteria Pengujian .....	III-8
3.4.3 Format Data Pengujian.....	III-9
3.4.4 Alat yang Digunakan dalam Pelaksanaan Penelitian .....	III-11
3.4.5 Pengujian Penelitian.....	III-12
3.4.6 Analisis Hasil Pengujian dan Membuat Kesimpulan .....	III-12
3.5 Metode Pengembangan Perangkat Lunak .....	III-12
3.5.1 Fase Insepsi .....	III-12
3.5.2 Fase Elaborasi .....	III-13
3.5.3 Fase Kontruksi.....	III-13
3.5.4 Fase Transisi.....	III-14
3.6 Manajemen Proyek Perangkat Lunak .....	III-15
3.7 Kesimpulan.....	III-18
<b>BAB IV PENGEMBANGAN PERANGKAT LUNAK .....</b>	<b>IV-1</b>
4.1 Pendahuluan .....	IV-1

4.2 Fase Insepsi .....	IV-1
4.2.1 Pemodelan Bisnis .....	IV-1
4.2.2 <i>Requirement</i> .....	IV-2
4.2.3 Analisis Kebutuhan dan Desain .....	IV-2
4.2.3.1 Analisis <i>Clustering</i> K-Means dengan Algoritma <i>Firefly</i> ...	IV-3
4.2.3.2 Desain Perangkat Lunak .....	IV-8
4.3 Fase Elaborasi .....	IV-12
4.3.1 <i>Activity</i> Diagram .....	IV-12
4.3.2 <i>Sequence</i> Diagram.....	IV-14
4.4 Fase Konstruksi .....	IV-16
4.4.1 Kebutuhan Sistem .....	IV-16
4.4.2 Perancangan Antarmuka .....	IV-16
4.4.3 Class Diagram .....	IV-17
4.4.4 Implementasi Antarmuka .....	IV-18
4.5 Fase Transisi.....	IV-21
4.5.1 Rencana Pengujian .....	IV-21
4.5.1.1 Rencana Pengujian <i>Upload File</i> .....	IV-21
4.5.1.2 Rencana Pengujian Perhitungan <i>Clustering</i> K-Means dan Algoritma <i>Firefly</i> .....	IV-22
4.5.2 Implementasi .....	IV-22
4.5.2.1 Pengujian <i>Upload File</i> .....	IV-22
4.5.2.2 Pengujian Perhitungan <i>Clustering</i> K-Means dengan Algoritma <i>Firefly</i> .....	IV-23
4.6 Kesimpulan.....	IV-24
<b>BAB V HASIL DAN PEMBAHASAN .....</b>	<b>V-1</b>
5.1 Pendahuluan .....	V-1
5.2 Data Hasil Percobaan .....	V-1
5.2.1 Konfigurasi Percobaan .....	V-1
5.2.2 Data Hasil Percobaan .....	V-2
5.2.2.1 Hasil Pengujian <i>Clustering</i> K-Means .....	V-2
5.2.2.2 Hasil Pengujian <i>Clustering</i> K-Means dan Algoritma <i>Firefly</i> .....	V-3

5.2.2.1 Hasil Pengujian Iterasi Maksimum.....	V-3
5.2.2.2 Hasil Pengujian Parameter Pengacakan.....	V-3
5.2.2.3 Hasil Pengujian Koefisien Daya Tarik .....	V-4
5.2.2.4 Hasil Pengujian Koefisien Penyerapan Cahaya.....	V-5
5.3 Analisis Hasil Penelitian .....	V-5
5.4 Kesimpulan.....	V-6
<b>BAB VI KESIMPULAN DAN SARAN.....</b>	<b>VI-1</b>
6.1 Kesimpulan.....	VI-1
6.2 Saran.....	VI-1
<b>DAFTAR PUSTAKA .....</b>	<b>xvii</b>

## DAFTAR TABEL

	Halaman
<b>Tabel II-1.</b> Interpretasi Nilai <i>Silhouette Coefficient</i> .....	II-10
<b>Tabel III-1.</b> Pengujian Jumlah Iterasi.....	III-9
<b>Tabel III-2.</b> Pengujian Parameter Pengacakan .....	III-10
<b>Tabel III-4.</b> Pengujian Koefisien Daya Tarik .....	III-10
<b>Tabel III-5.</b> Pengujian Koefisien Penyerapan Cahaya .....	III-11
<b>Tabel III-6.</b> Tabel <i>Work Breakdown Structure</i> (WBS) Penelitian .....	III-15
<b>Tabel IV-1.</b> Kebutuhan Fungsional .....	IV-2
<b>Tabel IV-2.</b> Kebutuhan Non-Fungsional .....	IV-2
<b>Tabel IV-3.</b> Nilai Input <i>Dataset</i> .....	IV-3
<b>Tabel IV-4.</b> <i>Preprocessing</i> Nilai Input <i>Dataset</i> .....	IV-3
<b>Tabel IV-5.</b> Inisialisasi Populasi 1 .....	IV-4
<b>Tabel IV-6.</b> Inisialisasi Populasi 2 .....	IV-4
<b>Tabel IV-7.</b> Titik <i>Centroid Clustering</i> K-Means .....	IV-5
<b>Tabel IV-8.</b> Jarak Populasi dengan Titik <i>Centroid</i> .....	IV-5
<b>Tabel IV-9.</b> Label Populasi 1 .....	IV-5
<b>Tabel IV-10.</b> Label Populasi 2 .....	IV-6
<b>Tabel IV-11.</b> Titik <i>Centroid</i> Populasi 1.....	IV-6
<b>Tabel IV-12.</b> Titik <i>Centroid</i> Populasi 2 .....	IV-6
<b>Tabel IV-13.</b> <i>Silhouette Score</i> Populasi 1 .....	IV-6
<b>Tabel IV-14.</b> <i>Silhouette Score</i> Populasi 2 .....	IV-6
<b>Tabel IV-15.</b> Intensitas Cahaya.....	IV-7
<b>Tabel IV-16.</b> Perbandingan Intensitas Cahaya .....	IV-7
<b>Tabel IV-17.</b> Definisi Aktor .....	IV-9
<b>Tabel IV-18.</b> Definisi <i>Use Case</i> .....	IV-9
<b>Tabel IV-19.</b> Skenario <i>Use Case</i> Melakukan <i>Upload Dataset</i> .....	IV-10
<b>Tabel IV-20.</b> Skenario <i>Use Case</i> Melakukan Perhitungan Clustering K-Means dan Algoritma Firefly .....	IV-11

<b>Tabel IV-21.</b> Rencana Pengujian Mengunggah File.....	IV-22
<b>Tabel IV-22.</b> Rencana Pengujian Perhitungan <i>Clustering</i> K-Means dan Algoritma <i>Firefly</i> .....	IV-22
<b>Tabel IV-23.</b> Pengujian <i>Upload File</i> .....	IV-23
<b>Tabel IV-24.</b> Pengujian <i>Upload File</i> .....	IV-23
<b>Tabel IV-25.</b> Pengujian Perhitungan <i>Clustering</i> K-Means dan Algoritma <i>Firefly</i> .....	IV-23
<b>Tabel V-1.</b> Hasil Pengujian Iterasi <i>Clustering</i> K-Means .....	V-2
<b>Tabel V-2.</b> Hasil Pengujian Iterasi Maksimum.....	V-3
<b>Tabel V-3.</b> Hasil Pengujian Parameter Pengacakan.....	V-3
<b>Tabel V-4.</b> Hasil Pengujian Koefisien Daya Tarik.....	V-4
<b>Tabel V-5.</b> Hasil Pengujian Koefisien Penyerapan Cahaya.....	V-5
<b>Tabel V-6.</b> Perbandingan Nilai <i>Silhouette Score</i> .....	V-5



## DAFTAR GAMBAR

	Halaman
<b>Gambar II-1.</b> Tahapan Proses Algoritma K-Means .....	II-6
<b>Gambar II-2.</b> Alur Algoritma <i>Firefly</i> (Arianti et al., 2020) .....	II-12
<b>Gambar III-1.</b> Diagram Tahapan Penelitian .....	III-2
<b>Gambar III-2.</b> Diagram Alir Clustering K-Means.....	III-5
<b>Gambar III-3.</b> Diagram Alir Optimasi K-Means dengan Algoritma <i>Firefly</i> .....	III-6
<b>Gambar IV-1.</b> Diagram <i>Use Case</i> .....	IV-8
<b>Gambar IV-2.</b> Diagram Aktivitas <i>Upload Dataset</i> .....	IV-12
<b>Gambar IV-3.</b> Diagram Aktivitas Perhitungan <i>Clustering</i> K-Means dan Algoritma <i>Firefly</i> .....	IV-13
<b>Gambar IV-4.</b> Diagram <i>Sequence Upload Dataset</i> .....	IV-14
<b>Gambar IV-5.</b> Diagram <i>Sequence</i> Perhitungan <i>Clustering</i> K-Means dan <i>Firefly</i> .....	IV-15
<b>Gambar IV-6.</b> Rancangan Desain Antarmuka <i>Upload File</i> .....	IV-17
<b>Gambar IV-7.</b> Rancangan Desain Antarmuka Perhitungan Clustering K- Means .....	IV-17
<b>Gambar IV-8.</b> Diagram Kelas .....	IV-18
<b>Gambar IV-9.</b> Implementasi Antarmuka <i>Upload Dataset</i> .....	IV-18
<b>Gambar IV-10.</b> Implementasi Antarmuka Hasil <i>Upload Dataset</i> .....	IV-19
<b>Gambar IV-11.</b> Implementasi Antarmuka Hasil Perhitungan <i>Clustering</i> K- Means .....	IV-19
<b>Gambar IV-12.</b> Implementasi Antarmuka Hasil Perhitungan Clustering K- Means dalam Grafik.....	IV-20
<b>Gambar IV-13.</b> Implementasi Antarmuka Hasil Perhitungan <i>Clustering</i> K- Means dan Algoritma <i>Firefly</i> .....	IV-20
<b>Gambar IV-14.</b> Implementasi Antarmuka Hasil Perhitungan <i>Clustering</i> K- Means dan Algoritma <i>Firefly</i> dalam Grafik.....	IV-21

# **BAB I**

## **PENDAHULUAN**

### **1.1 Pendahuluan**

Pada bab ini akan membahas tentang gagasan-gagasan utama mengenai penelitian yang meliputi latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, dan sistematika penulisan.

### **1.2 Latar Belakang Masalah**

*Clustering* adalah proses memisahkan kumpulan data besar ke dalam banyak unit yang serupa tanpa pengetahuan tentang label. Di dalam unit yang sama, titik data yang identik satu sama lain akan digabungkan menjadi satu grup dan yang berbeda akan dipisahkan ke grup yang lain (Waboke W. R., 2023). Metode ini berfungsi sebagai analisis data statistik yang dapat digunakan di berbagai bidang, seperti *machine learning*, *data mining*, *image analysis*, pengenalan pola dan *bioinformatics* (Mandhulatha T. S., 2012).

*Clustering* dibedakan menjadi dua, yakni *Hierarchical Clustering* dan *Partitional Clustering*. Pada *Hierarchical Clustering* didapatkan baik dengan *Agglomerative* atau *Divisive Techniques*. Sedangkan, *Partition* bisa dengan *Soft Technique (Fuzzy Clustering)* atau *Hard Technique (K-Means)*. Algoritma K-Means telah banyak digunakan karena kesederhanaannya dan mudah diimplementasikan (Waboke W. R., 2023).

Algoritma K-Means merepresentasikan proses iterasi dari pencarian *centroid*, yaitu pusat *cluster* (Tayfur et al., 2018). Metode ini sering digunakan untuk pengelompokan berbagai data dengan tanpa label atau pengelompokan berdasarkan kesamaan data. Namun, K-means memiliki kekurangan terhadap pembangkitan titik pusat awal yang *random* dan hasil pengelompokan yang selalu berubah-ubah. Oleh karena itu, untuk mengatasi permasalahan tersebut digunakanlah algoritma optimasi (Windarto A. P. et al., 2020).

Pengoptimasian *modern* dilakukan dengan paradigma *metaheuristic*. Diketahui bahwa, paradigma ini lebih bisa menyelesaikan permasalahan *NP-Hard Optimization* (Xin She Yang, 2010). Algoritma dari *metaheuristic* seperti *Particle Swarm Optimization* telah diteliti sebelumnya di dalam jurnal pengelompokan gambar satelit untuk mencari solusi dari titik pusat yang optimal di K-Means (Kumar et al., 2016). Selain itu, terdapat jurnal yang mengulas tentang *Genetic Algorithm* dapat digunakan untuk masalah optimasi pusat *cluster* di K-Means yang meningkatkan kualitas dari optimasi itu sendiri (Zeebaree et al., 2017). Ditambah penelitian tentang K-Means dan *Ant Colony Algorithm* dimana keduanya dapat meningkatkan akurasi dan kecepatan sebelumnya (Jue Lu & Rongqiang Hu, 2013).

Berdasarkan metode optimasi di atas terdapat banyak lagi algoritma yang digunakan untuk optimasi. Seiring berjalannya waktu terdapat algoritma lain yang muncul dan disempurnakan, salah satunya adalah algoritma *Firefly*. Algoritma ini telah berkembang pesat dan banyak digunakan dalam berbagai bidang optimasi dan masalah teknik (Ali N. et al., 2014). *Firefly* banyak digunakan karena tingkat performa dan fleksibilitas yang lebih tinggi dari algoritma optimasi lainnya.

Hal ini dibuktikan berdasarkan penelitian yang telah dilakukan. Algoritma *Firefly* memberikan fleksibilitas lebih dalam penyesuaian parameter seperti parameter pengacakan, koefisien daya tarik dan koefisien penyerapan cahaya. Selain itu, memiliki kemampuan untuk menangani *noise* tinggi dibandingkan *Particle Swarm Optimization* (Michael F. Lohrer, 2013). Ditambah penelitian lainnya yang membuktikan bahwa Algoritma *Firefly* membuat performa yang lebih baik dari pada *Genetic Algorithm* dan *Particle Swarm Optimization* dalam pengujian merancang desain antena (Mohammed et al., 2016)

Berdasarkan masalah yang disebutkan diatas. Maka dari itu, penelitian ini akan menganalisis pengelompokkan berdasarkan algoritma *clustering* K-means dengan algoritma *Metaheuristic*, yaitu algoritma *Firefly* untuk menghasilkan kualitas *clustering* yang lebih optimal.

### **1.3 Rumusan Masalah**

Fokus utama dalam penelitian ini adalah untuk mengetahui apakah Algoritma *Firefly* mampu mengoptimalkan nilai *cluster* dalam metode *clustering* K-Means. Oleh karena itu, dirumuskanlah beberapa pertanyaan sebagai berikut:

1. Bagaimana algoritma *Firefly* dalam mengoptimasi proses *clustering* K-Means?
2. Bagaimana kinerja algoritma *clustering* K-Means dalam pengelompokkan dengan menggunakan algoritma *Firefly*?

#### **1.4 Tujuan Penelitian**

Terdapat beberapa tujuan melakukan penelitian ini yang dijabarkan sebagai berikut:

1. Mengoptimalkan proses *clustering* pada K-Means dengan menerapkan algoritma *Firefly*
2. Meningkatkan hasil akurasi *clustering* K-Means menggunakan algoritma *Firefly*
3. Membandingkan kinerja dari K-Means dengan algoritma *Firefly* dan K-Means tanpa algoritma *Firefly*.

#### **1.5 Manfaat Penelitian**

Berikut adalah beberapa manfaat yang bisa diperoleh dari penelitian ini sebagai berikut:

1. Meningkatkan kualitas pengelompokan data dengan menerapkan *clustering* K-Means dan algoritma *Firefly*.
2. Sebagai kontribusi di bidang *data mining* dengan menggabungkan dua teknik, yaitu *clustering* K-Means dan algoritma *Firefly*.

## **1.6 Batasan Masalah**

Data yang digunakan adalah data publik yang diambil dari *website* Kaggle dengan judul *Wine Dataset*

## **1.7 Sistematika Penulisan**

Sistematika penulisan pada penelitian ini yaitu:

### **BAB I. PENDAHULUAN**

Pada bab satu akan menjabarkan tentang latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah serta sistematika penulisan dalam penelitian yang akan dibuat.

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

Pada kajian literatur berisi beberapa dasar teori yang digunakan dalam penelitian, seperti *Clustering*, metode K-Means, *Silhouette Coefficient* dan Algoritma *Firefly* serta terdapat beberapa rujukan dari penelitian lain yang relevan.

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

Pada bab metodologi penelitian akan berisi tahapan yang akan dilaksanakan pada penelitian ini. Tiap-tiap tahapan pada penelitian yang didasarkan pada kerangka kerja III-1 akan dijabarkan secara *detail*. Lalu pada subbab akhir akan menjelaskan mengenai perancangan manajemen proyek pada pelaksanaan penelitian.

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

Bab ini akan membahas mengenai langkah-langkah pengembangan aplikasi perangkat lunak untuk optimasi nilai *Cluster* Perhitungan K-Means dengan menggunakan Algoritma *Firefly*.

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

Pada bab ini akan menjelaskan hasil penelitian yang diperoleh dari perancangan perangkat lunak yang telah dikembangkan pada bab sebelumnya.

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

Bab ini akan menguraikan hasil kesimpulan yang didapat dan membuat saran terhadap penelitian yang akan dilakukan kedepannya.

### **1.8 Kesimpulan**

Berdasarkan masalah yang telah dijelaskan pada subbab-subbab sebelumnya maka akan dilakukan penelitian mengenai *clustering* K-Means yang menggunakan metode optimasi algoritma *Firefly* dalam pengelompokkan data *Wine*. Optimasi dilakukan agar mendapatkan nilai *cluster* yang akurat, konsisten dan juga cepat.



## DAFTAR PUSTAKA

- Ali, M. (2012). Kontrol Kecepatan Motor DC Menggunakan PID Kontroler yang *Ditunning* dengan *Firefly* Algorithm. *Jurnal Intake*, 3(2).
- Arianti, R. W., Via, Y. V., & Purbasari, I. Y. (2020). Implementasi Algoritma *Firefly* dalam Menyelesaikan Pengoptimalan Produksi Sepatu (Studi Kasus: *Home industry “PAK KICU Shoes”* Sidoarjo). *Jurnal Informatika Dan Sistem Informasi (JIFoSI)*, 1(2).
- Hermawati, F. A. (2013). *Data Mining*. Yogyakarta: ANDI.
- Jue Lu & Rongqiang Hu. (2013). *A New Hybrid Clustering Algorithm Based on K-Means and Ant Colony Algorithm*. *2<sup>nd</sup> International Conference on Computer Science and Electronics Engineering (ICCSEE 2013)*.
- Khotimah, T., Teknik, D. F., Studi, P., Informatika, T., & Kudus, U. M. (2014). Pengelompokan Surat Dalam AL QUR 'AN Menggunakan Algoritma K-Means, 5(1), 83–88.
- Lohrer, Michael F. (2013). A Comparison Between the Firefly Algorithm and Particle Swarm Optimization. <https://our.oakland.edu/server/api/core/bitstreams/6f31ccda-9f90-490e-b5d9-27cd71faf6d4/content>
- Mandhulatha T. S. (2012). *An OvervieEw On Clustering Methods*. *IOSR Journal of Engineering*, 2(4), 719-725.
- Mohammed, H. J., F. Abdulsalam, A.S. Abdulla, R.S. Ali, R.A. Abd-Alhameed, J.M. Noras, Y.I. Abdulraheem, A. Ali, J. Rodriguez, & Abdelgader M. Abdalla. (2016). *Evaluation of genetic algorithms, particle swarm optimisation, and Firefly algorithms in antenna design*. *13th International Conference on Synthesis, Modeling, Analysis and Simulation Methods and Applications to Circuit Design (SMACD)*.
- Nasari, F., Darma, S., & Informasi, S. (2015). Penerapan K-Means *Clustering* Pada Data Penerimaan Mahasiswa Baru, 6–8.
- Nasution, I., Windarto, A. P., & Fauzan, M. (2020). Penerapan Algoritma K-Means Dalam Pengelompokan Data Penduduk Miskin Menurut Provinsi. *Building of Informatics, Technologu and Science (BITS)*, 2(2), 76-83.
- Nofriansyah, D. & Nurcahyo, G. W. (2015). Algoritma *Data Mining* dan Pengujian. Yogyakarta: *Deepublish*.

- Primartha, R. (2018). *Belajar Machine Learning Teori dan Praktik*. Bandung: Informatika Bandung.
- Putra, I M. S. (2018). *Algoritma DBSCAN (Density-Based Spatial Clustering of Applications with Noise) dan Contoh Perhitungannya*. <http://erepo.unud.ac.id/id/eprint/20097/1/b8f162572f7b5c80a51c1816883153d1.pdf>
- Salih, R. N., & Al-Jawaherry, M. A. (2020). *Finding Minimum and Maximum Values of Variables in Mathematical Equations by Applying Firefly and PSO Algorithm*. *Tikrit Journal of Pure Science*, 25(5). <https://doi.org/10.25130/tjps.25.2020.096>
- Suyanto. (2017). *Data Mining Untuk Klasifikasi dan Klasterisasi Data*. Bandung: Informatika Bandung.
- Suyanto. (2017). *Swarm Intelligence Komputasi Modern untuk Optimasi dan Big Data Mining*. Bandung: Informatika Bandung.
- Tayfur, S., Alver, N., Abdi S., Saatci, S., & Ghiami, A. (2018). *Characterization of concrete matrix/steel fiber de-bonding in an SFRC beam: Principal component analysis and k-means algorithm for clustering AE data*. *Engineering Fracture Mechanics*, 194, 73-85. [www.elsevier.com/locate/engfractmech](http://www.elsevier.com/locate/engfractmech)
- Waboke W. R., Bagiwa M. A., Obiniyi A. A., Obasa A. I. (2023). *Centroid Initialization In K-Means Clustering Using GATCAM*. *Science World Jurnal*, 18(1). <https://www.scienceworldjournal.org/article/view/23559>
- Wicaksono A. E. (2016). Implementasi Data Mining dalam Pengelompokan Data Peserta Didik di Sekolah Untuk Memprediksi Calon Penerima Beasiswa Dengan Menggunakan Algoritma K-Means (Studi Kasus SMAN 16 Bekasi). <https://ejournal.gunadarma.ac.id/index.php/tekno/article/view/1599>
- Xin She Yang. (2010). *Firefly Algorithm, Stochastic Test Functions and Design Optimisation*. *Int. J. Bio-Inspired Computation*, 2(2), 78-84.
- Zeebaree, D. Q., Haron, H., Abdulazeez, A. M., Zeebaree, S. R. M. (2017). *Combination of K-Means Clustering with Genetic Algorithm: A review*. *International Journal of Applied Engineering Research*, 12(24), 14238-14245. <http://www.ripublication.com>