

**PROGRAM STUDI TEKNIK INFORMATIKA**

**IMPLEMENTASI ALGORITMA K-MEANS CLUSTERING DALAM  
PENENTUAN PENGELOMPOKAN DATA PAMSIMAS**

**EKA PERMATA SARI**

**15.142.0144**

**Skripsi ini diajukan sebagai syarat memperoleh gelar Sarjana Komputer**



**FAKULTAS ILMU KOMPUTER**

**UNIVERSITAS BINADARMA**

**2020**



**IMPLEMENTASI ALGORITMA K-MEANS CLUSTERING DALAM  
PENENTUAN PENGELOMPOKAN DATA PAMSIMAS**

**EKA PERMATA SARI**

**15.142.0144**

**Skripsi ini diajukan sebagai syarat memperoleh gelar Sarjana Komputer**

**PROGRAM STUDI TEKNIK INFORMATIKA**

**FAKULTAS ILMU KOMPUTER**

**UNIVERSITAS BINA DARMA**

**2020**

**HALAMAN PENGESAHAN**

**IMPLEMENTASI ALGORITMA K-MEANS CLUSTERING DALAM  
PENENTUAN PENGELOMPOKAN DATA PAMSIMAS**

**OLEH :**

**EKA PERMATA SARI**

**151420144**

**Telah diterima sebagai salah satu syarat untuk memperoleh gelar Sarjana  
Komputer pada Program Studi Teknik Informatika**

**Palembang, Maret 2020**

**Fakultas Ilmu Komputer**

**Universitas Bina Darma**

**Dosen Pembimbing,**



**Nia Oktaviani, M.Kom**

**Dekan,**




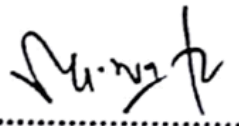
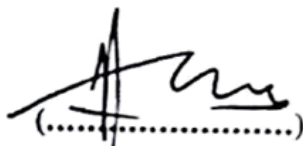
Universitas **Bina Darma**  
Fakultas Ilmu Komputer

**Dedy Syamsuar, S.kom., M.I.T., Ph.D.**

HALAMAN PERSETUJUAN

Skripsi Berjudul "IMPLEMENTASI ALGORITMA K-MEANS  
CLUSTERING DALAM PENENTUAN PENGELOMPOKAN DATA  
PAMSIMAS" Oleh "EKA PERMATA SARI" telah dipertahankan didepan  
komisi penguji pada hari SENIN tanggal 02 MAERT 2020.

Komisi Penguji

1. Ketua : Nia Oktaviani, M.Kom   
(.....)
2. Anggota : Susan Dian Purnama Sari, M.Kom.   
(.....)
3. Anggota : Hadi Syaputra, M.Kom.   
(.....)

Mengetahui,  
Program Studi Teknik Informatika  
Fakultas Ilmu Komputer  
Universitas Bina Darma

Ketua,

Universitas Bina Darma  
Fakultas Ilmu Komputer

Dr. Widya Cholil, S.Kom., M.I.T.

## SURAT PERNYATAAN

Saya yang bertanda tangan di bawah ini :

Nama : EKA PERMATA SARI

NIM : 15.142.0144

Dengan ini menyatakan :

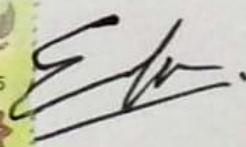
1. Skripsi ini adalah asli dan belum pernah diajukan untuk mendapatkan gelar sarjana di Universitas Bina Darma atau perguruan tinggi lain;
2. Skripsi ini murni gagasan, rumusan dan penelitian saya sendiri dengan arahan tim pembimbing;
3. Di dalam Skripsi ini tidak terdapat karya atau pendapat yang telah di tulis atau dipublikasikan orang lain, kecuali secara tertulis dengan jelas dikutip dengan mencantumkan nama pengarang dan memasukkan ke dalam rujukan;
4. Saya bersedia Skripsi yang saya hasilkan dicek keasliannya menggunakan *turnitin* serta di unggah ke internet, sehingga dapat di akses publik secara daring;
5. Surat pernyataan ini saya tulis dengan sungguh-sungguh dan apabila terbukti melakukan penyimpangan atau ketidakbenaran dalam pernyataan ini, saya bersedia menerima sanksi sesuai dengan peraturan dan perundang-undangan yang berlaku.

Demikian Surat Pernyataan ini saya buat agar dapat dipergunakan sebagaimana mestinya.

Palembang, Februari 2020

Yang membuat pernyataan,



  
**EKA PERMATA SARI**  
NIM : 151420144

## MOTTO DAN PERSEMBAHAN

### MOTTO

- *“Dan bahwasannya seorang manusia tiada memperoleh selain apa yang telah diusahakannya” (An Najm : 39).*
- *“Barangsiapa yang mempelajari ilmu pengetahuan yang seharusnya yang ditujukan untuk mencari ridho Allah bahkan hanya untuk mendapatkan kedudukan/kekayaan duniawi maka ia tidak akan mendapatkan baunya surga nanti pada hari kiamat” (riwayat Abu Hurairah radhiallahu anhu).*

### PERSEMBAHAN

**Alhamdulillah, Puji dan Syukur Kehadirat Allah SWT ku persembahkan kepada :**

- ❖ **Orang Tuaku tercinta, mama dan abah yang selalu support dan mendoakan disetiap aktivitasku.**
- ❖ **Saudara kandungku, ayuk dan kakak yang juga selalu membantu dan menasehatiku.**
- ❖ **Dosen pembimbing, ibu Nia Oktaviani, M.Kom**
- ❖ **Teman-teman seperjuangan prodi teknik informatika.**

## KATA PENGANTAR



Puji syukur kehadirat Allah SWT karena berkat rahmat dan karunia- Nya jualah, skripsi ini dapat diselesaikan guna memenuhi salah satu persyaratan untuk menyelesaikan program studi Informatika di Universitas Bina Darma Palembang. Dalam penulisan skripsi ini, tentunya masih jauh dari sempurna. Hal ini dikarenakan keterbatasnya pengetahuan yang dimiliki. Oleh karena itu dalam rangka melengkapi kesempurnaan dari penulisan skripsi ini diharapkan adanya saran dan kritik yang diberikan bersifat membangun. Pada kesempatan yang baik ini, tak lupa penulis menghaturkan terima kasih kepada semua pihak yang telah memberikan bimbingan, pengarahan, nasehat dan pemikiran dalam penulisan skripsi ini, terutama kepada :

- 1) Ibu Dr. Sunda Ariana., M.Pd., M.M. selaku Rektor Universitas Bina Darma Palembang.
- 2) Bapak Dedy Syamsuar, Ph.D. selaku Dekan Fakultas Ilmu Komputer
- 3) Ibu Dr. Widya Cholil, S.Kom., M.I.T. selaku Ketua Program Studi Informatika.
- 4) Ibu Nia Oktaviani, M.Kom. selaku Pembimbing yang telah memberikan bimbingan dalam penulisan skripsi ini.
- 5) Orang Tua, saudara-saudaraku, seluruh teman dan sahabat-sahabatku yang selalu memberikan dorongan dan masukan serta bantuan baik moril maupun materil yang tak ternilai harganya.

- 6) Seluruh Dosen Fakultas Ilmu Komputer Universitas Bina Darma yang telah memberikan ilmu dan bimbingannya kepada kami.
- 7) Seluruh teman-teman seperjuangan di Universitas Bina Darma Palembang

Palembang, Februari 2020

Peneliti



## ABSTRAK

Kemajuan luar biasa yang terus berlanjut dalam bidang data mining didorong oleh beberapa faktor salah satunya adalah pertumbuhan yang cepat dalam kumpulan data dan perkembangan yang hebat dalam kemampuan komputasi dan pengembangan kapasitas media penyimpanan. Dalam penentuan pengelompokan data PAMSIMAS banyak metode yang dapat digunakan untuk menggali nilai emas dari kumpulan data yang besar salah satunya menggunakan clustering. Teknik clustering dapat dijalankan dengan menggunakan salah satu algoritma yang mudah di implementasikan yaitu k-means. Algoritma ini dapat dijalankan dengan menggunakan tool weka dan hasilnya dapat di evaluasi sehingga dapat memperoleh metode pengumpulan data yang tepat.

**Keyword** : Pengelompokan, *Data mining, Algoritma K-means, KDD, Weka.*

## ABSTRACT

The extraordinary progress that continues in the field of data mining is driven by several factors one of which is the rapid growth in data collection and the great development in computing capability and storage media capacity development. In determining the grouping of PAMSIMAS data, many methods can be used to extract gold values from large data sets, one of which uses clustering. Clustering techniques can be run using one algorithm that is easy to implement, namely k-means. This algorithm can be run using the Weka tool and the results can be evaluated so that it can obtain the right data collection method.

**Keyword:** *Clustering, Data mining, K-means Algorithm, KDD, Weka*

## DAFTAR ISI

<b>HALAMAN COVER..</b> .....	<b>i</b>
<b>HALAMAN JUDUL..</b> .....	<b>ii</b>
<b>HALAMAN PENGESAHAN..</b> .....	<b>iii</b>
<b>HALAMAN PERSETUJUAN..</b> .....	<b>iv</b>
<b>HALAMAN PERNYATAAN..</b> .....	<b>v</b>
<b>MOTTO DAN PERSEMBAHAN..</b> .....	<b>vi</b>
<b>KATA PENGANTAR..</b> .....	<b>vii</b>
<b>ABSTRAK .....</b>	<b>viii</b>
<b>DAFTAR ISI .....</b>	<b>ix</b>
<b>DAFTAR TABLE .....</b>	<b>x</b>
<b>DAFTAR GAMBAR .....</b>	<b>xii</b>

### **BAB I PENDAHULUAN**

1.1 Latar Belakang .....	1
1.2 Perumusan Masalah .....	3
1.3 Batasan Masalah .....	3
1.4 Tujuan dan Manfaat Penelitian .....	4
1.4.1 Tujuan Penelitian .....	4
1.4.2 Manfaat Penelitian .....	4
1.5 Metodologi Penelitian .....	5
1.5.1 Waktu dan Tempat Penelitian .....	5
1.5.2 Metode Penelitian .....	5
1.5.3 Metode Pengumpulan Data .....	5
1.6 Sistematika Penulisan .....	7

### **BAB II TUJUAN PUSTAKA**

2.1 Tinjauan Umum .....	8
2.1.1 Sejarah PAMSIMAS .....	8
2.1.2 Struktur Organisasi .....	11
2.2 Landasan Teori .....	11
2.2.1 Data Mining .....	11
2.2.2 Clustering .....	15
2.2.3 <i>K-Means</i> .....	16
2.2.4 WEKA .....	20
2.3 Penelitian Sebelumnya .....	23
2.4 Kerangka Berpikir .....	25

### **BAB III ANALISIS DATA MINING**

3.1 Metode Analisis .....	28
3.1.1 <i>Data Selection</i> .....	28
3.1.2 <i>Data Cleaning</i> .....	29
3.1.3 <i>Data Transformation</i> .....	31

<b>BAB IV HASIL DAN PEMBAHASAN</b>	
4.1 Data Mining .....	45
4.1.1 Perhitungan K-Means Secara Manual .....	45
4.1.1 Proses Data Mining Menggunakan WEKA .....	56
4.2 Evaluation .....	66
<b>BAB V KESIMPULAN DAN SARAN</b>	
5.1 Kesimpulan .....	68
5.2 Saran .....	69
<b>DAFTAR PUSTAKA</b>	
<b>LAMPIRAN</b>	

## DAFTAR TABEL

Tabel 3.1	Tranformasi Data Pada Atribut Kabupaten.....	32
Table 3.2	Transformasi Data Pada Atribut Kabupaten Lampung Barat.....	33
Tabel 3.3	Tranformasi Data Pada Atribut Kabupaten Tanggamus .....	34
Tabel 3.4	Tranformasi Data Pada Atribut Kabupaten Lampung Selatan .....	35
Tabel 3.5	Tranformasi Data Pada Kabupaten Lampung Timur .....	36
Tabel 3.6	Tranformasi Data Atribut Kabupaten Lampung Tengah .....	37
Tabel 3.7	Transformasi Data Pada Kabupaten Lampung Utara.....	38
Tabel 3.8	Transformasi Data Pada Kabupaten Way Kanan .....	39
Tabel 3.9	Transformasi Data Pada Kabupaten Tulang Bawang .....	40
Tabel 3.10	Transformasi Data Pada Kabupaten Pesawaran .....	41
Tabel 3.11	Transformasi Data Pada Kabupaten Pringsewu .....	42
Tabel 3.12	Transformasi Data Pada Kabupaten Mesuji .....	42
Tabel 3.13	Transformasi Data Pada Kabupaten Tulang Bawang Barat .....	43
Tabel 3.14	Transformasi Data Pada Kabupaten Pesisir Barat .....	44
Tabel 4.1	Sample Data Pamsimas Yang Sudah Di Transform.....	46
Tabel 4.2	Centroid Awal.....	47
Tabel 4.3	Iterasi 1 .....	49
Tabel 4.4	Contoh Pengelompokan Data Cluster Terdekat .....	50
Tabel 4.5	Centroid Baru.....	52
Tabel 4.6	Iterasi 4 .....	53
Tabel 4.7	Contoh Pengelompokan Kelompok Pada Iterasi 4 .....	54

## DAFTAR GAMBAR

Gambar 2.2	Struktur Organisasi Pamsimas.....	11
Gambar 2.2	Proses Kdd .....	13
Gambar 2.3	Flowcart Algoritma K-Means.....	19
Gambar 3.1	Data Awal .....	28
Gambar 3.2	Contoh Hasil Seleksi Data .....	29
Gambar 3.3	Contoh Data Kosong .....	30
Gambar 3.4	Hasil Pembersihan Data Kosong .....	31
Gambar 4.1	Tampilan Awal Aplikasi Weka .....	56
Gambar 4.2	Menu Explorer .....	57
Gambar 4.3	Tampilan Awal Proses .....	57
Gambar 4.4	Data Yang Siap Proses .....	58
Gambar 4.5	Diagram Pada Atribut Tahun 2017.....	59
Gambar 4.6	Diagram Pada Atribut Tahun 2018.....	60
Gambar 4.7	Diagram Pada Atribut Tahun 2019.....	60
Gambar 4.8	Pemilihan K-Means Clustering.....	60
Gambar 4.9	Hasil Pengelompoakn Cluster 2017 .....	62
Gambar 4.10	Hasil Pengelompoakn Cluster 2018 .....	64
Gambar 4.11	Hasil Pengelompoakn Cluster 2019 .....	65
Gambar 4.12	Cluster Visualize Tahun 2017 .....	66
Gambar 4.13	Cluster Visualize Tahun 2018 .....	66
Gambar 4.14	Cluster Visualize Tahun 2019 .....	67