

**ANALISIS PENGELOMPOKAN DATA PENJUALAN  
SMARTPHONE MENGGUNAKAN K-MEANS CLUSTERING**



**“SKRIPSI”**

**diajukan untuk memenuhi syarat memperoleh gelar Sarjana  
Komputer**

**QUINTA VERNANDYA IVANA SASTI**

**1803040124**

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS TEKNIK DAN SAINS  
UNIVERSITAS MUHAMMADIYAH PURWOKERTO  
2023**

## HALAMAN PERSETUJUAN

Skripsi yang diajukan oleh:

Nama : Quinta Vernandya Ivana Sasti

NIM : 1803040124

Program Studi : Teknik Informatika

Fakultas : Teknik dan Sains

Perguruan Tinggi : Universitas Muhammadiyah Purwokerto

Judul : Analisis Pengelompokan Data Penjualan *Smartphone*

Menggunakan *K-Means Clustering*



Telah diterima dan disetujui  
Purwokerto,

Elindra Ambar Pambudi, S.Kom., M.Kom

NIK.2160766

## HALAMAN PENGESAHAN

Skripsi yang diajukan oleh:

Nama : Quinta Vernandya Ivana Sasti  
NIM : 1803040124  
Program Studi : Teknik Informatika  
Fakultas : Teknik dan Sains  
Perguruan Tinggi : Universitas Muhammadiyah Purwokerto  
Judul : Analisis Pengelompokan Data Penjualan *Smartphone*  
Menggunakan *K-Means Clustering*

Telah berhasil dipertahankan dihadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Komputer (S.Kom) pada Program Studi Teknik Informatika, Fakultas Teknik dan Sains, Universitas Muhammadiyah Purwokerto.

### DEWAN PENGUJI

Penguji 1 (Pembimbing) : Elindra Ambar Pambudi, S.Kom., M.Kom.  
Penguji 2 : Maulida Ayu Fitriani, S.Kom., M.Cs.  
Penguji 3 : Dimara Kusuma Hakim, S.T., M.Cs.

Ditetapkan di : Purwokerto

Tanggal :

Mengetahui

Dekan Fakultas Teknik dan Sains

Ir. Teguh Marhehdi, S.T., M.T., ASEAN Eng., ACPE., IPM  
NIK. 2160172

## HALAMAN PERNYATAAN ORISINALITAS

Saya yang bertandatangan dibawah ini:

Nama : **Quinta Vernandya Ivana Sasti**

NIM : **1803040124**

Program Studi : **Teknik Informatika**

Fakultas : **Teknik dan Sains**

Perguruan Tinggi : **Universitas Muhammadiyah Purwokerto**

Menyatakan dengan sebenar-benarnya bahwa skripsi ini adalah hasil karya saya dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar, dan bukan hasil penjiplakan dari karya orang lain.

Demikian pernyataan ini saya buat dan apabila kelak di kemudian hari terbukti ada unsur penjiplakan, saya bersedia mempertanggungjawabkan sesuai dengan ketentuan yang berlaku.

Purwokerto,

nembuat pernyataan



Quinta Vernandya Ivana Sasti

## HALAMAN PERSEMBAHAN

Dengan segala kerendahan hati, serta rasa syukur terhadap Allah SWT yang telah memberikan rahmat dan nikmatnya, maka saya mempersembahkan tugas akhir ini kepada:

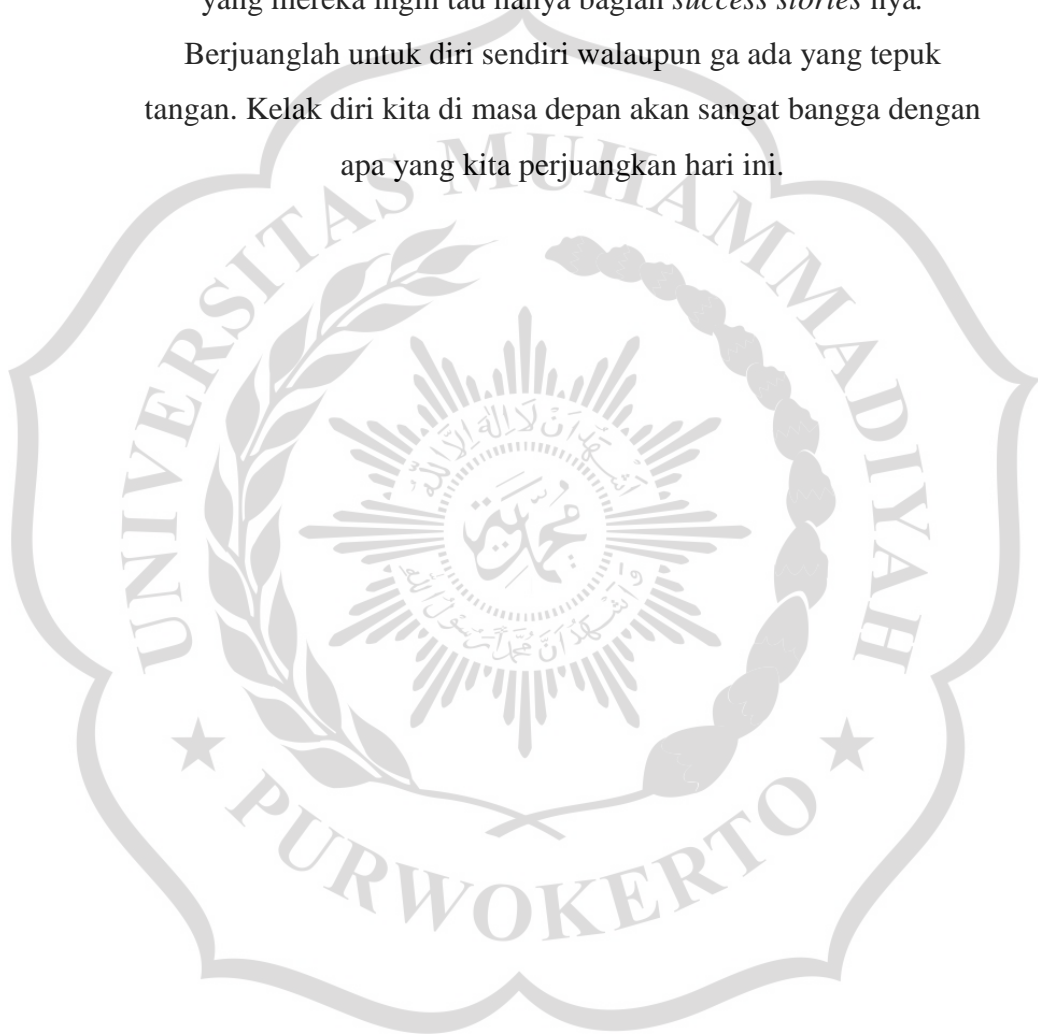
1. Kedua orang tua yang selama ini telah membantu dalam bentuk doa, kasih sayang, semangat, dan perhatian yang tidak henti-hentinya mengalir demi kelancaran dan kesuksesan dalam menyelesaikan tugas akhir ini. Semoga menjadi kebahagiaan bagi mereka.
2. Kepada bapak Elindra Ambar Pambudi, S.Kom., M.Kom. selaku dosen pembimbing yang telah memberikan bimbingan, arahan, dorongan sehingga tugas akhir ini dapat terselesaikan.
3. Kepada seluruh dosen program studi Teknik Informatika, Universitas Muhammadiyah Purwokerto yang telah memberikan banyak ilmu.
4. Kepada seluruh sahabat seperjuangan angkatan 2018 Teknik Informatika.
5. Kepada semua sivitas akademika Universitas Muhammadiyah Purwokerto yang telah membantu penyusunan tugas akhir ini baik secara langsung maupun tidak langsung.
6. Kepada NIM 1803040138 yang telah kebersamai penulis pada hari-hari yang tidak mudah selama proses pengerjaan Tugas Akhir. Terima kasih telah menjadi rumah yang tidak hanya berupa tanah dan bangunan. Tetap kebersamai dan tidak tunduk pada apa-apa. Tabah sampai akhir.

## MOTTO

*“Only you can change your life. Nobody else can do it for you”*

Orang lain ga akan bisa paham *struggle* dan masa sulitnya kita,  
yang mereka ingin tau hanya bagian *success stories* nya.

Berjuanglah untuk diri sendiri walaupun ga ada yang tepuk  
tangan. Kelak diri kita di masa depan akan sangat bangga dengan  
apa yang kita perjuangkan hari ini.



## KATA PENGANTAR

Segala puji bagi Allah SWT yang telah memberikan rahmat dan karunia-Nya kepada penulis, sehingga penulis dapat menyelesaikan tugas akhir ini dengan baik. Shalawat serta salam senantiasa tercurah kepada nabi Muhammad SAW yang mengantarkan manusia dari zaman kebodohan ke zaman terang seperti sekarang ini. Penyusunan tugas akhir ini dimaksudkan untuk memenuhi sebagian syarat-syarat guna mencapai gelar Sarjana Komputer di Universitas Muhammadiyah Purwokerto.

Selama proses penulisan tugas akhir ini, penulis menyadari adanya keterlibatan dari pihak lain. Oleh karena itu, penulis mengucapkan terima kasih kepada seluruh pihak yang telah memberi dukungan dalam proses penulisan tugas ini baik dalam bentuk moril dan materil. Semoga dengan terselesaikannya tugas akhir ini, menjadi sebuah penghargaan dan kebanggaan bagi semua pihak yang telah membantu secara langsung maupun tidak langsung.

Penulis menyadari masih banyak kekurangan dalam penulisan laporan maupun penelitian yang telah dilakukan. Oleh karena itu, penulis berharap kritik dan saran untuk perbaikan ataupun pengembangan di masa yang akan datang. Semoga tugas akhir ini bermanfaat bagi banyak pihak.

## HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademik Universitas Muhammadiyah Purwokerto dan demi pengembangan ilmu pengetahuan, saya yang bertanda tangan dibawah ini:

Nama : Quinta Vernandya Ivana Sasti  
NIM : 1803040124  
Program Studi : Teknik Informatika  
Fakultas : Teknik dan Sains  
Perguruan Tinggi : Universitas Muhammadiyah Purwokerto  
Jenis karya : Skripsi

Menyetujui untuk memberikan Hak Bebas Royalti Noneksklusif (*Non-exclusive Royalty-Free Right*) kepada Universitas Muhammadiyah Purwokerto atas karya ilmiah saya yang berjudul:

*Analisis Pengelompokan Data Penjualan Smartphone Menggunakan K-Means Clustering*

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Noneksklusif ini Universitas Muhammadiyah Purwokerto berhak menyimpan, mengalihmedia / mengalihformatkan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan skripsi saya dengan tetap mencantumkan nama saya sebagai penulis / pencipta dan sebagai pemilik Hak Cipta.

Pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Purwokerto  
Pada Tanggal :  
Yang menyatakan,



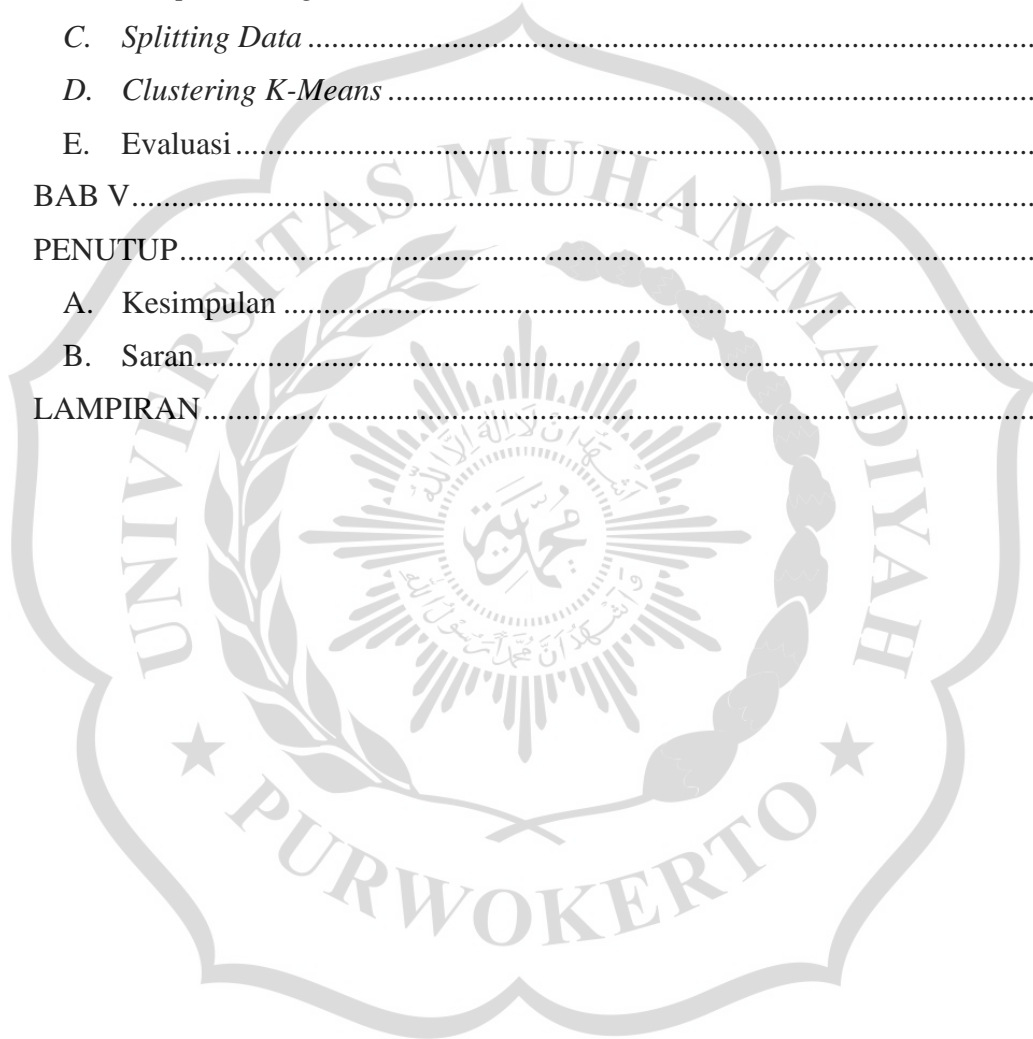
SPULUH RIBURUPIAH  
10000  
TEL. METERAI TEMPEL  
E9EAKX257783238

Quinta Vernandya Ivana Sasti

## DAFTAR ISI

HALAMAN PERSETUJUAN.....	i
HALAMAN PENGESAHAN.....	ii
HALAMAN PERNYATAAN ORISINALITAS.....	iii
HALAMAN PERSEMBAHAN .....	iv
MOTTO .....	v
KATA PENGANTAR .....	vi
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI.....	vii
DAFTAR ISI.....	viii
DAFTAR TABEL.....	x
DAFTAR GAMBAR .....	xi
DAFTAR LAMPIRAN.....	xii
ABSTRAK .....	xiii
<i>ABSTRACT</i> .....	xiv
BAB I.....	1
PENDAHULUAN .....	1
A. Latar Belakang .....	1
B. Perumusan Masalah .....	2
C. Batasan Masalah.....	3
D. Tujuan Penelitian .....	3
E. Manfaat Penelitian .....	3
BAB II.....	4
TINJAUAN PUSTAKA .....	4
A. Penelitian Terdahulu .....	4
B. Landasan Teori.....	5
BAB III .....	10
METODE PENELITIAN.....	10
A. Jenis Penelitian.....	10
B. Waktu dan Tempat Penelitian .....	10
C. Variabel yang diteliti.....	10
D. Pengumpulan Data .....	10

E. Alat Penelitian .....	10
F. Desain Alur Penelitian .....	11
BAB IV .....	18
HASIL DAN PEMBAHASAN .....	18
A. Data Penelitian .....	18
B. <i>Pre-processing Data</i> .....	19
C. <i>Splitting Data</i> .....	22
D. <i>Clustering K-Means</i> .....	24
E. Evaluasi .....	32
BAB V .....	39
PENUTUP .....	39
A. Kesimpulan .....	39
B. Saran .....	39
LAMPIRAN .....	43

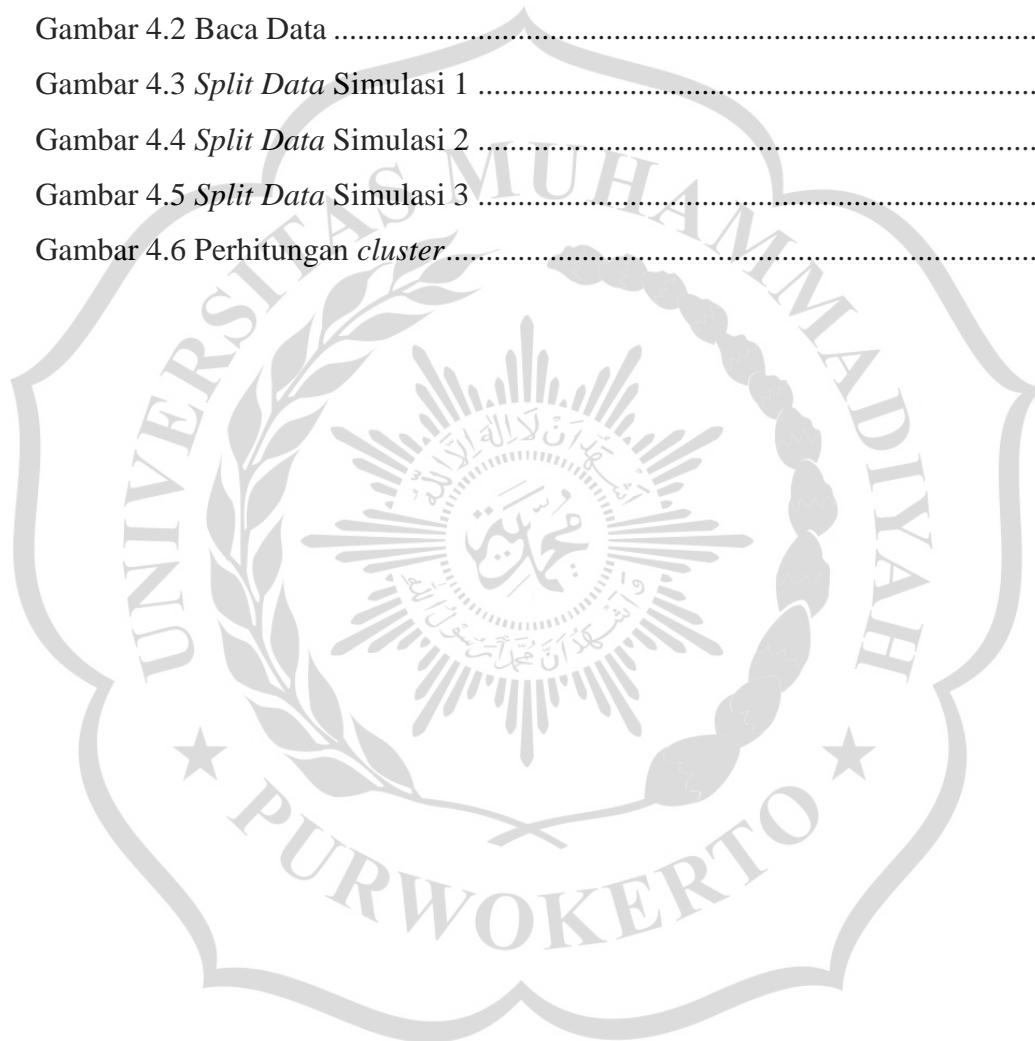


## DAFTAR TABEL

Tabel 3.1 Penafsiran nilai <i>Silhouette Index</i> .....	16
Tabel 4.1 Data sebelum normalisasi .....	22
Tabel 4.2 Data setelah normalisasi.....	22
Tabel 4. 3 <i>Centroid</i> 1 ( <i>Random</i> ) .....	24
Tabel 4. 4 Tabel Pengelompokkan Data Iterasi 1 .....	26
Tabel 4. 5 <i>Centroid</i> 2.....	27
Tabel 4. 6 Tabel Pengelompokkan Data Iterasi 2 .....	28
Tabel 4. 7 <i>Centroid</i> Akhir .....	29
Tabel 4. 8 Tabel Pengelompokkan Data Akhir.....	29
Tabel 4.9 Kelompok hasil <i>clustering</i> simulasi 1 pada data test.....	30
Tabel 4.10 Kelompok hasil <i>clustering</i> simulasi 1 pada data train .....	30
Tabel 4.11 Kelompok hasil <i>clustering</i> simulasi 2 pada data test.....	31
Tabel 4.12 Kelompok hasil <i>clustering</i> simulasi 2 pada data train .....	31
Tabel 4.13 Kelompok hasil <i>clustering</i> simulasi 3 pada data test.....	31
Tabel 4.14 Kelompok hasil <i>clustering</i> simulasi 3 pada data train .....	32
Tabel 4.15 Hasil SSB .....	32
Tabel 4. 16 Hasil Perhitungan Rasio.....	33
Tabel 4.17 Hasil perhitungan <i>Sillhouette Index</i> .....	35
Tabel 4.18 Hasil evaluasi simulasi 1 .....	36
Tabel 4.19 Hasil evaluasi simulasi 2.....	37
Tabel 4.20 Hasil evaluasi simulasi 3.....	37

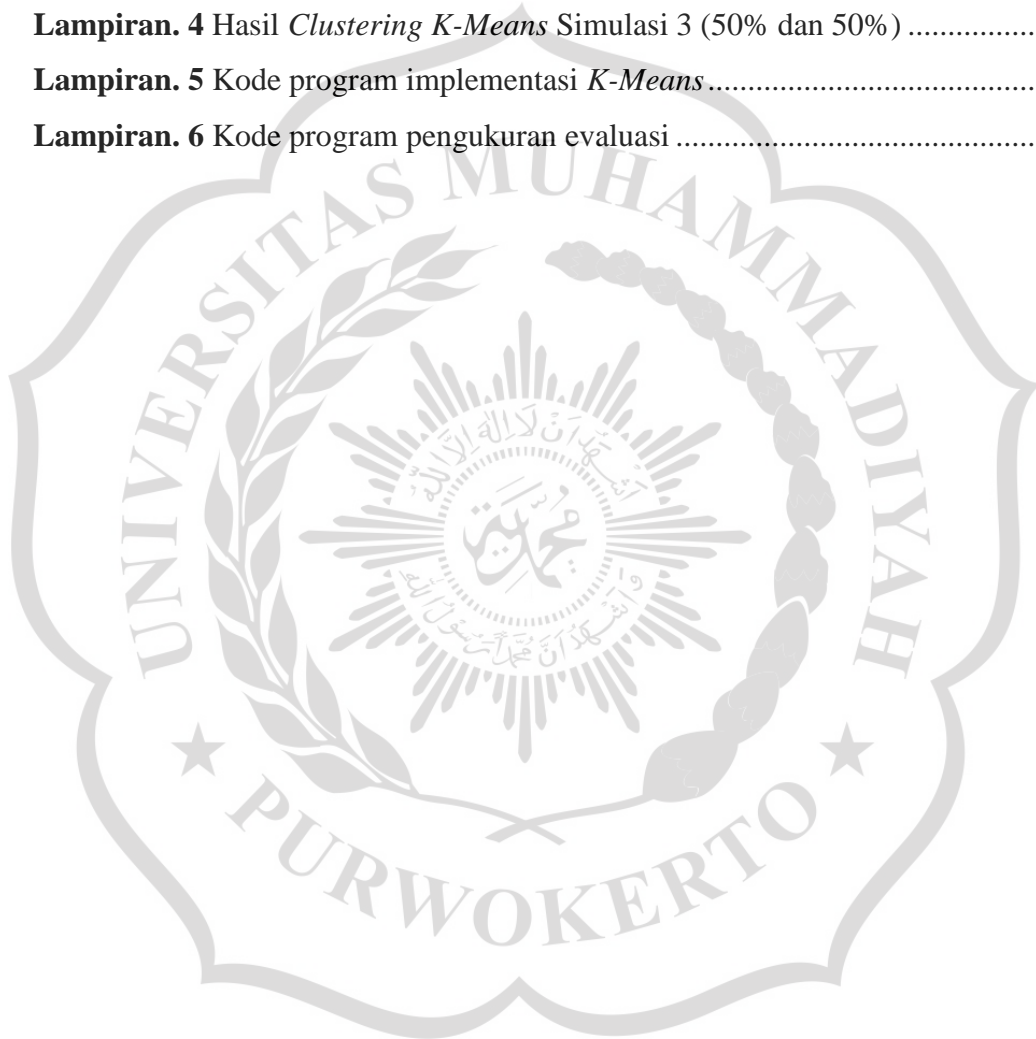
## DAFTAR GAMBAR

Gambar 2.1 <i>Flowchart</i> diagram algoritma K-Means.....	8
Gambar 3.1 Diagram alur penelitian.....	11
Gambar 4.1 <i>Import Library</i> .....	20
Gambar 4.2 Baca Data .....	21
Gambar 4.3 <i>Split Data</i> Simulasi 1 .....	23
Gambar 4.4 <i>Split Data</i> Simulasi 2 .....	23
Gambar 4.5 <i>Split Data</i> Simulasi 3 .....	23
Gambar 4.6 Perhitungan <i>cluster</i> .....	29



## DAFTAR LAMPIRAN

<b>Lampiran. 1</b> File data penjualan <i>smartphone</i> .....	43
<b>Lampiran. 2</b> Hasil <i>Clustering K-Means</i> Simulasi 1 (80% dan 20%) .....	46
<b>Lampiran. 3</b> Hasil <i>Clustering K-Means</i> Simulasi 2 (70% dan 30%) .....	49
<b>Lampiran. 4</b> Hasil <i>Clustering K-Means</i> Simulasi 3 (50% dan 50%) .....	52
<b>Lampiran. 5</b> Kode program implementasi <i>K-Means</i> .....	55
<b>Lampiran. 6</b> Kode program pengukuran evaluasi .....	56



## ABSTRAK

Persaingan di dunia teknologi semakin berkembang pesat, salah satunya *smartphone*. Tidak dipungkiri, segala urusan dari mulai pelajar sampai orang dewasa hampir semua melibatkan teknologi terutama pada *smartphone*. Banyaknya merk, tipe *smartphone*, serta jumlah penjualan yang semakin hari semakin melonjak, maka pencatatan penjualan dengan menggunakan teknologi diperlukan agar semua pasar terpenuhi. Pada penelitian ini dilakukan analisis pengelompokan data penjualan *smartphone* menggunakan algoritma *K-Means*. Data yang digunakan sebanyak 26015 data, serta jumlah cluster yang dipakai sebanyak 3, yang mana data penjualan akan dikelompokkan menjadi 3 kelompok / kelas. Pengolahan data akan dibagi menjadi 3 simulasi dengan simulasi data *train* dan data *test* 80% dan 20%, 70% dan 30%, 50% dan 50%. Data dievaluasi menggunakan evaluasi *Davies Bouldin Index*, *Silhouette Index*, dan *Calinski Harabasz Index*. Hasil dari penelitian menunjukkan bahwa hasil terbaik terdapat pada simulasi 1 ( 80% data train, dan 20% data test) karena terdapat 2 evaluasi dengan hasil yang optimal yaitu evaluasi *Davies Bouldin Index* dan *Silhouette Index*, dan nilai akhir *cluster* 0 sebanyak 4110 data, cluster 1 sebanyak 319 data, dan cluster 2 sebanyak 774 data.

**Kata Kunci :** *Smartphone, K-Means, Cluster*

## **ABSTRACT**

*Competition in the world of technology is growing rapidly, one of which is smartphones. It is undeniable that almost everything from students to adults involves technology, especially smartphones. There are many brands, types of smartphones, and the number of sales is increasing day by day, so recording sales using technology is necessary so that all markets are met. In this research, grouping analysis of smartphone sales data was carried out using the K-Means algorithm. The data used is 26015 data, and the number of clusters used is 3, where the sales data will be grouped into 3 groups/classes. Data processing will be divided into 3 simulations with simulated train data and test data 80% and 20%, 70% and 30%, 50% and 50%. Data were evaluated using the Davies Bouldin Index, Silhouette Index, and Calinski Harabasz Index evaluations. The results of the study show that the best results are in simulation 1 (80% data train, and 20% test data) because there are 2 evaluations with optimal results, namely the evaluation of the Davies Bouldin Index and the Silhouette Index, and the final value of cluster 0 is 4110 data, cluster 1 as many as 319 data, and cluster 2 as many as 774 data.*

**Keywords :** *Smartphone, K-Means, Cluster*