

**DATA *MINING* UNTUK PENGELOMPOKAN  
JENIS USAHA DI RUMAH BUMN BATAM  
MENGUNAKAN METODE *CLUSTERING***

**SKRIPSI**



**Oleh :  
Arif Hernawan  
180210071**

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS TEKNIK DAN KOMPUTER  
UNIVERSITAS PUTERA BATAM  
TAHUN 2023**

**DATA *MINING* UNTUK PENGELOMPOKAN  
JENIS USAHA DI RUMAH BUMN BATAM  
MENGUNAKAN METODE *CLUSTERING***

**SKRIPSI**

**Untuk memenuhi salah satu syarat  
memperoleh gelar Sarjana**



**Oleh :  
Arif Hernawan  
180210071**

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS TEKNIK DAN KOMPUTER  
UNIVERSITAS PUTERA BATAM  
TAHUN 2023**

## SURAT PERNYATAAN ORISINALITAS

Yang bertanda tangan di bawah ini saya:

Nama : Arif Hernawan  
NPM : 180210071  
Fakultas : Teknik dan Komputer  
Program Studi : Teknik Informatika

Menyatakan bahwa “Skripsi” yang saya buat dengan judul:

**“DATA MINING UNTUK PENGELOMPOKAN JENIS USAHA DI RUMAH BUMH BATAM MENGGUNAKAN METODE *CLUSTERING*”**

Adalah hasil karya sendiri dan bukan “duplikasi” dari karya orang lain. Sepengetahuan saya, di dalam naskah Skripsi ini tidak terdapat karya ilmiah atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis dikutip didalam naskah ini dan disebutkan dalam sumber kutipan dan daftar pustaka.

Apabila ternyata di dalam naskah Skripsi ini dapat dibuktikan terdapat unsur-unsur PLAGIASI, saya bersedia naskah Skripsi ini digugurkan dan gelar akademik yang saya peroleh dibatalkan, serta diproses sesuai dengan peraturan perundang-undangan yang berlaku.

Demikian pernyataan ini saya buat dengan sebenarnya tanpa ada paksaan dari siapapun.

Batam, 27 Januari 2023



**Arif Hernawan**

NPM. 180210071

**DATA *MINING* UNTUK PENGELOMPOKAN  
JENIS USAHA DI RUMAH BUMN BATAM  
MENGUNAKAN METODE *CLUSTERING***

**SKRIPSI**

**Untuk memenuhi salah satu syarat  
memperoleh gelar Sarjana**

**Oleh:  
Arif Hernawan  
180210071**

**Telah disetujui oleh Pembimbing pada tanggal  
seperti tertera dibawah ini**

**Batam, 27 Januari 2023**



**Rahmat Fauzi, S.Kom., M.Kom.  
Pembimbing**

## ABSTRAK

Pelaku usaha di Indonesia pada umumnya dikategorikan menjadi usaha besar dan usaha kecil menengah atau yang sering dikenal dengan UKM. Ada salah satu Badan Usaha Milik Negara yang memiliki peran menghimpun dan mendorong para pelaku UKM untuk naik kelas agar lebih sejahtera, yaitu Rumah BUMN Batam. Masih banyak pelaku UKM yang perlu dibantu untuk melakukan promosi agar mendapat jumlah pemesanan dengan frekuensi yang tinggi. Tujuan penelitian yaitu untuk mengelompokkan jenis-jenis UKM berdasarkan frekuensi penjualan, sehingga nantinya perusahaan dapat melakukan promosi lebih lanjut terhadap UKM yang mendapat jumlah pemesanan dengan frekuensi rendah. Dalam Penelitian ini digunakan metode *Clustering* dengan algoritma *K-means clustering*. Algoritma *k-means clustering* dapat dimanfaatkan oleh Rumah BUMN Batam untuk mempermudah pengelompokan jenis usaha dan frekuensi pemesanan UKM pertahun. Penulis menggunakan proses *Knowledge Discovery in Database (KDD)* yang terdiri dari pembersihan data, integrasi data, seleksi data, transformasi data, penambangan data, evaluasi pola, dan presentasi pengetahuan. Pada tahap awal KDD akan dilakukan pembersihan data dan integrasi data secara manual melalui aplikasi *WEKA 3.9.6*. Selanjutnya dilakukan seleksi data, transformasi data, dan penambangan data dengan menggunakan perangkat lunak yang dibuat. Tahap selanjutnya yaitu evaluasi pola dan presentasi pengetahuan. Hasil penelitian ini memiliki 4 *cluster* yang paling optimal sesuai dengan kebutuhan. Penelitian ini diharapkan dapat membantu mempermudah Rumah BUMN Batam untuk mengelompokkan jenis-jenis usaha kedepannya.

**Kata Kunci:** *Algoritma K-Means, Clustering, Data Mining*

## **ABSTRACT**

*Business actors in Indonesia are generally categorized into large businesses and small and medium enterprises or often known as UKM. There is one State-Owned Enterprise that has the role of gathering and encouraging UKM players to upgrade their classes to be more prosperous, namely Rumah BUMN Batam. There are still many UKM players who need help to carry out promotions in order to get a high number of orders. The research objective is to classify the types of UKM based on the frequency of sales, so that later the company can carry out further promotions for UKM that get a low number of orders. In this study, the clustering method was used with the K-means clustering algorithm. The k-means clustering algorithm can be used by Rumah BUMN Batam to facilitate the grouping of types of business and the frequency of orders for UKM per year. The author uses the Knowledge Discovery in Database (KDD) process which consists of data cleaning, data integration, data selection, data transformation, data mining, pattern evaluation, and knowledge presentation. In the early stages of KDD, data cleaning and data integration will be carried out manually through the WEKA 3.9.6 application. Furthermore, data selection, data transformation, and data mining were carried out using the software that was created. The next stage is pattern evaluation and knowledge presentation. The results of this study have 4 clusters that are the most optimal according to the needs. This research is expected to help make it easier for Rumah BUMN Batam to classify types of businesses in the future.*

**Keywords:** *Algoritma K-Means, Clustering, Data Mining*

## KATA PENGANTAR

Segala Puji dan Syukur penulis panjatkan atas kehadiran Allah S.W.T atas ridha-Nya sehingga penulis bisa menyusun dan menyelesaikan skripsi yang berjudul “Data *Mining* untuk Pengelompokan Jenis Usaha di Rumah BUMN Batam Menggunakan Metode *Clustering*”

Penulisan skripsi ini disusun guna memenuhi persyaratan dalam menyelesaikan program studi Strata Satu (S1) pada program Studi Teknik Informatika Universitas Putera Batam.

Penulis sadar bahwa skripsi ini tidak dapat terselesaikan dengan baik tanpa dukungan, bimbingan serta bantuan dari berbagai pihak. Untuk itu pada kesempatan ini dengan segala kerendahan hati, penulis mengucapkan terima kasih yang sebesar-besarnya kepada yang terhormat:

1. Ibu Dr. Nur Elfi Husda, S.Kom., M.SI. Selaku Rektor Universitas Putera Batam.
2. Dekan Fakultas Teknik dan Komputer, Bapak Welly Sugianto, S.T., M.M.
3. Ketua Program Studi Teknik Informatika, Bapak Andi Maslan, S.T., M.SI.
4. Kepada Bapak Rahmat Fauzi, S.Kom., M.Kom. Selaku pembimbing Skripsi pada Program Studi Teknik Informatika Universitas Putera Batam.
5. Kepada Ibu Alfannisa Annurrullah Fajrin, S.Kom., M.Kom. Selaku pembimbing Akademik pada Program Studi Teknik Informatika Universitas Putera Batam.
6. Para Dosen serta Staff di Universitas Putera Batam.
7. Kepada kedua orang tua dan seluruh keluarga besar penulis yang selalu

memberikan dukungan baik dari segi materil maupun moril kepada penulis.

8. Kepada istri tercinta Chania Aprilia Yonanta yang selalu mendukung, menemani serta memotivasi dalam suka dan duka.
9. Rumah BUMN Batam yang telah memberikan izin kepada penulis untuk pengumpulan data penelitian.
10. Bapak Firgiawan Adi Nugroho, S.T. yang telah bersedia membantu dalam penelitian ini.
11. Para teman-teman seperjuangan yang dengan setia mendampingi dan memberikan dukungan dan juga semangat.

Penulis juga menyadari keterbatasan pengalaman dan juga pengetahuan yang dimiliki penulis, sehingga dalam penulisan tugas akhir ini masih sangat jauh dari kata sempurna. Oleh sebab itu, penulis sangat mengharapkan saran dan kritik yang konstruktif dari berbagai pihak. Akhir kata semoga tugas akhir ini dapat diterima dan bermanfaat bagi para pembaca.

Batam, 27 Januari 2023

Penulis,



Arif Hernawan



## DAFTAR ISI

<b>HALAMAN COVER</b> .....	i
<b>HALAMAN JUDUL</b> .....	ii
<b>SURAT PERNYATAAN ORISINALITAS</b> .....	iii
<b>HALAMAN PENGESAHAN</b> .....	iv
<b>ABSTRAK</b> .....	v
<b>ABSTRACT</b> .....	vi
<b>KATA PENGANTAR</b> .....	vii
<b>DAFTAR ISI</b> .....	ix
<b>DAFTAR GAMBAR</b> .....	xi
<b>DAFTAR TABEL</b> .....	xii
<b>DAFTAR RUMUS</b> .....	xiii
<b>BAB I PENDAHULUAN</b> .....	1
1.1 Latar Belakang Masalah .....	1
1.2 Identifikasi Masalah .....	3
1.3 Batasan Masalah .....	3
1.4 Rumusan Masalah .....	4
1.5 Tujuan Penelitian .....	4
1.6 Manfaat Penelitian .....	5
1.6.1 Manfaat Praktis .....	5
1.6.2 Manfaat Teoritis .....	5
<b>BAB II TINJAUAN PUSTAKA</b> .....	6
2.1 <i>Knowledge Discovery of Database (KDD)</i> .....	6
2.2 <i>Data Mining</i> .....	8
2.3 <i>Metode Data Mining (Clustering)</i> .....	9
2.4 <i>Algoritma K-Means</i> .....	21
2.5 <i>Software Pendukung</i> .....	23
2.5.1 <i>Microsoft Excel 2016</i> .....	23
2.5.2 <i>WEKA 3.9.6</i> .....	24
2.6 <i>Penelitian Terdahulu</i> .....	27
2.7 <i>Kerangka Pemikiran</i> .....	31
<b>BAB III METODE PENELITIAN</b> .....	33
3.1 <i>Desain Penelitian</i> .....	33
3.2 <i>Teknik Pengumpulan Data</i> .....	35
3.3 <i>Operasional Variabel</i> .....	36
3.4 <i>Lokasi dan Jadwal Penelitian</i> .....	37
3.4.1 <i>Lokasi Penelitian</i> .....	37
3.4.2 <i>Jadwal Penelitian</i> .....	38
<b>BAB IV HASIL DAN PEMBAHASAN</b> .....	39
4.1 <i>Analisis Data</i> .....	39
4.1.1 <i>Pengumpulan Data</i> .....	39
4.1.2 <i>Seleksi Data</i> .....	40
4.1.3 <i>Pembersihan Data</i> .....	41
4.1.4 <i>Transformasi Data</i> .....	44
4.1.5 <i>Implementasi</i> .....	46
4.1.6 <i>Perhitungan Algoritma K-Means Clustering</i> .....	47

4.2 Hasil Pengujian .....	62
<b>BAB V PENUTUP</b> .....	69
5.1 Kesimpulan.....	69
5.2 Saran.....	69

**DAFTAR PUSTAKA**

**LAMPIRAN**

Lampiran 1 : Wawancara
Lampiran 2 : Foto Dokumentasi
Lampiran 3 : Foto Beberapa Produk UKM
Lampiran 4 : Lanjutan Perhitungan Euclidean Ke-2 sampai Ke-8
Lampiran 5 : Lanjutan Tabel Iterasi Ke-2 sampai Ke-8
Lampiran 6 : Daftar Riwayat Hidup
Lampiran 7 : Hasil Turnitin Skripsi
Lampiran 8 : Hasil Turnitin Jurnal
Lampiran 9 : Surat Izin Penelitian
Lampiran 10 : Balasan Surat Izin Penelitian

## DAFTAR GAMBAR

<b>Gambar 2. 1</b> Tahapan <i>Knowledge Discovery Database (KDD)</i> .....	6
<b>Gambar 2. 2</b> Visualisasi <i>Cluster</i> .....	10
<b>Gambar 2. 3</b> Ruang Fitur pada Iterasi Awal .....	17
<b>Gambar 2. 4</b> Ruang Fitur pada Iterasi Ke 2 .....	20
<b>Gambar 2. 5</b> <i>Software Microsoft Excel 2016</i> .....	23
<b>Gambar 2. 6</b> Tampilan Utama <i>WEKA</i> .....	24
<b>Gambar 2. 7</b> Kerangka Pemikiran .....	32
<b>Gambar 3. 1</b> Diagram Alur Penelitian .....	33
<b>Gambar 3. 2</b> Lokasi Penelitian .....	37
<b>Gambar 3. 3</b> Peta Tempat Penelitian .....	38
<b>Gambar 4. 1</b> Data Mentah UKM Tahun 2020 .....	39
<b>Gambar 4. 2</b> Data Mentah UKM Tahun 2021 .....	40
<b>Gambar 4. 3</b> Tampilan <i>Menu Explorer Software WEKA 3.9.6</i> .....	63
<b>Gambar 4. 4</b> <i>Open File</i> pada <i>WEKA 3.9.6</i> .....	64
<b>Gambar 4. 5</b> Tampilan <i>Preprocess</i> .....	64
<b>Gambar 4. 6</b> Pemilihan Algoritma .....	65
<b>Gambar 4. 7</b> Penentuan Jumlah <i>Cluster</i> .....	65
<b>Gambar 4. 8</b> Hasil Jumlah <i>Attribut</i> .....	66
<b>Gambar 4. 9</b> Hasil Pemilihan <i>Centroid Random</i> .....	66
<b>Gambar 4. 10</b> Hasil <i>Cluster</i> Akhir .....	67
<b>Gambar 4. 11</b> Visualisasi <i>Cluster</i> .....	67

## DAFTAR TABEL

<b>Tabel 2. 1</b> Contoh Kasus <i>Clustering</i> .....	16
<b>Tabel 2. 2</b> Hasil Iterasi Terakhir.....	21
<b>Tabel 3. 1</b> Jadwal Penelitian .....	38
<b>Tabel 4. 1</b> Data Setelah Proses Seleksi dan <i>Cleaning</i> .....	41
<b>Tabel 4. 2</b> Data Setelah Proses Transformasi.....	44
<b>Tabel 4. 3</b> Pusat <i>Centroid</i> Baru.....	47
<b>Tabel 4. 4</b> Pusat <i>Centroid</i> Baru Hasil Iterasi 1.....	51
<b>Tabel 4. 5</b> Pusat <i>Centroid</i> Baru Hasil Iterasi 2.....	52
<b>Tabel 4. 6</b> Pusat <i>Centroid</i> Baru Hasil Iterasi 3.....	53
<b>Tabel 4. 7</b> Pusat <i>Centroid</i> Baru Hasil Iterasi 4.....	55
<b>Tabel 4. 8</b> Pusat <i>Centroid</i> Baru Hasil Iterasi 5.....	56
<b>Tabel 4. 9</b> Pusat <i>Centroid</i> Baru Hasil Iterasi 6.....	57
<b>Tabel 4. 10</b> Pusat <i>Centroid</i> Baru Hasil Iterasi 7.....	59
<b>Tabel 4. 11</b> Pusat <i>Centroid</i> Baru Hasil Iterasi 8.....	60
<b>Tabel 4. 12</b> Hasil Iterasi Ke-1 .....	61

## DAFTAR RUMUS

<b>Rumus 2. 1</b> Rumus <i>SSE (Sum of Squared Errors)</i> .....	15
<b>Rumus 2. 2</b> Rumus <i>Euclidean</i> .....	22