

**PENERAPAN *DATA MINING* MENGGUNAKAN ALGORITMA  
*K-MEDOIDS* DALAM PENGELOMPOKAN NASABAH  
PENERIMA *REWARD* PADA PT DOTRI  
GADAI JAYA**

**SKRIPSI**

**OLEH:**

**MENIATI ZEBUA**

**208160007**



**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS TEKNIK  
UNIVERSITAS MEDAN AREA  
MEDAN  
2024**

**UNIVERSITAS MEDAN AREA**

© Hak Cipta Di Lindungi Undang-Undang

1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Document Accepted 10/9/24

Access From (repository.uma.ac.id)10/9/24

PENERAPAN *DATA MINING* MENGGUNAKAN ALGORITMA  
*K-MEDOIDS* DALAM PENGELOMPOKAN NASABAH  
PENERIMA *REWARD* PADA PT DOTRI  
GADAI JAYA

SKRIPSI

Diajukan Sebagai Salah Satu Syarat untuk Memperoleh  
Gelar Sarjana (S1) di Fakultas Teknik  
Universitas Medan Area

OLEH:

MENIATI ZEBUA

208160007

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS TEKNIK

UNIVERSITAS MEDAN AREA

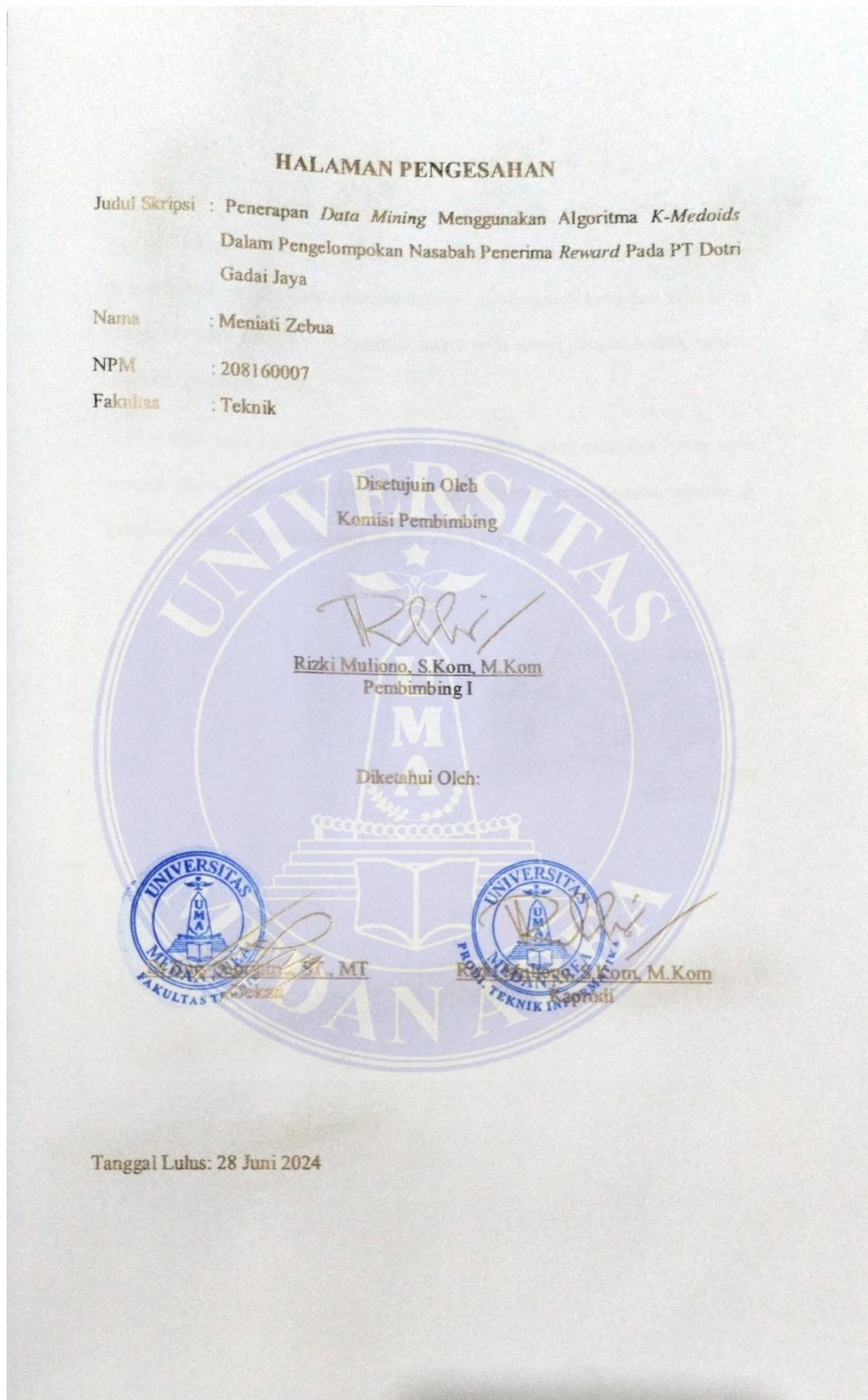
MEDAN

2024

UNIVERSITAS MEDAN AREA

© Hak Cipta Di Lindungi Undang-Undang

1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area



### HALAMAN PERNYATAAN

Saya menyatakan bahwa skripsi yang saya susun, sebagai syarat memperoleh gelar serjana merupakan hasil karya tulis saya sendiri. Adapun bagian-bagian tertentu dalam penulisan skripsi ini yang saya kutip dari hasil karya orang lain telah dituliskan sumbernya secara jelas sesuai dengan norma, kaidah, dan etika penulisan ilmiah.

Saya bersedia menerima sanksi pencabutan gelar akademik yang saya peroleh dan sanksi- sanksi lainnya dengan peraturan yang berlaku, apabila di kemudian hari ditemukan adanya plagiat dalam skripsi ini.

Medan, Agustus 2024



Meniati Zebua  
208160007

## HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR/SKRIPSI/TESIS UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademik Universitas Medan Area, saya yang bertanda tangan di bawah ini:

Nama : Meniati Zebua  
NPM : 208160007  
Program Studi : Teknik Informatika  
Fakultas : Teknik  
Jenis Karya : Skripsi

demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Medan Area Hak Bebas Royalti Noneksklusif (Non-exclusive Royalty-Free Right) atas karya ilmiah saya yang berjudul:

**Penerapan *Data Mining* Menggunakan Algoritma *K-Medoids* Dalam Pengelompokan Nasabah Penerima *Reward* Pada PT Dotri Gadai Jaya**

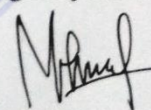
berserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Noneksklusif ini Universitas Medan Area berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (database), merawat, dan memublikasikan tugas akhir/skripsi/tesis saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di: Medan

Pada tanggal: 23 Agustus 2024

Yang Menyatakan



(Meniati Zebua)

## ABSTRAK

Pertumbuhan pesat industri keuangan menjadikan perusahaan-perusahaan mengalami persaingan yang ketat. PT Dotri Gadai Jaya (PT DGJ) adalah sebuah perusahaan pergadaian swasta yang juga merasakan dampaknya dan berupaya dalam mempertahankan dan meningkatkan loyalitas nasabah di tengah persaingan tersebut, salah satu strategi yang diterapkan oleh PT DGJ adalah memberikan *reward* kepada nasabah berdasarkan jumlah transaksi pinjaman gadai. Namun, dalam pelaksanaannya, perusahaan menghadapi kesulitan dalam pengelompokan nasabah yang pantas menerima *reward* secara efisien dan akurat, selama ini pengelompokan nasabah penerima *reward* dilakukan secara manual yang dapat memakan waktu yang cukup lama, adanya potensi kesalahan dan kurangnya transparan. Penelitian ini bertujuan untuk menerapkan *data mining* menggunakan algoritma *K-Medoids* dalam pengelompokan nasabah penerima *reward* di PT DGJ, mengetahui hasil pengelompokan untuk membantu PT DGJ dalam menentukan nasabah penerima *reward*, dan mengevaluasi hasilnya menggunakan *Davies Bouldin Index* (DBI). Untuk mengatasi masalah dan mencapai tujuan tersebut, maka diperlukannya penerapan *data mining* menggunakan algoritma *K-Medoids*. Hasil pengelompokan dari 1.085 nasabah menunjukkan bahwa jumlah nasabah yang mendapatkan *reward* sebesar 30% sebanyak 314 nasabah, 20% sebanyak 540 nasabah dan 10% sebanyak 231 nasabah. Evaluasi kualitas *cluster* menggunakan DBI menghasilkan nilai sebesar 0.368812, hal ini membuktikan bahwa nilai tersebut mendekati angka 0 atau terhitung cukup kecil. Sehingga dapat dikatakan bahwa *cluster* yang dihasilkan cukup baik. Penelitian ini memberikan kontribusi penting bagi PT DGJ dalam mengoptimalkan strategi pemberian *reward* kepada nasabah guna meningkatkan efisiensi dan keakuratan dalam pengelompokan nasabah.

**Kata kunci:** *Data Mining, Clustering, K-Medoids, Nasabah, Reward*

### ABSTRACT

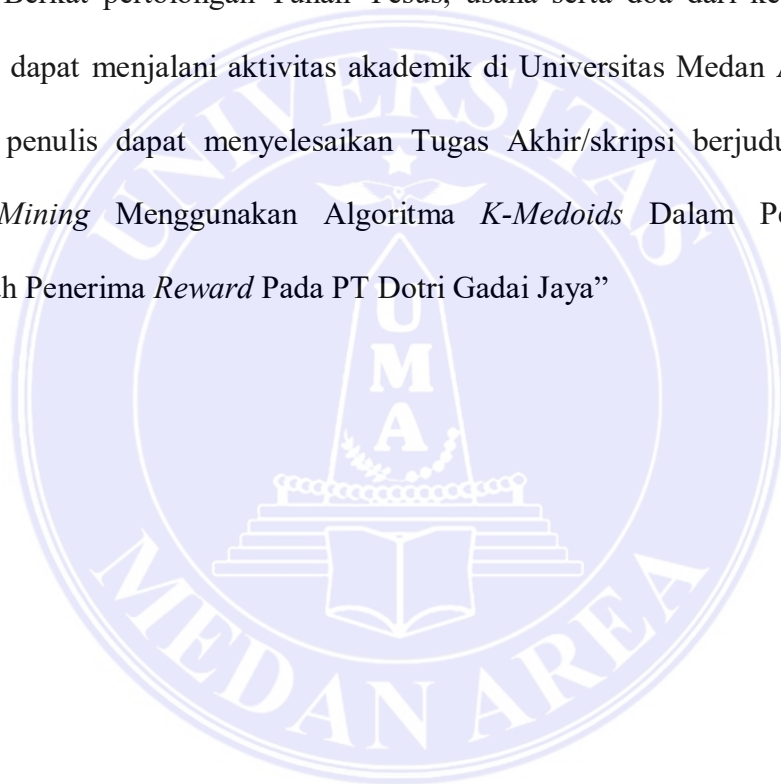
The rapid growth of the financial industry has led companies to face intense competition. PT Dotri Gadai Jaya (PT DGJ), a private pawnshop company, is also experiencing the impact and is striving to maintain and enhance customer loyalty amid this competition. One of the strategies implemented by PT DGJ is to provide rewards to customers based on the number of pawn loan transactions. However, in practice, the company faces difficulties in efficiently and accurately grouping customers eligible for rewards. Previously, customer grouping for reward distribution was done manually, which could be time-consuming, prone to errors, and lacking in transparency. This study aimed to apply data mining using the K-Medoids algorithm to group reward-eligible customers at PT DGJ, determine the grouping results to assist PT DGJ in identifying reward recipients, and evaluate the results using the Davies-Bouldin Index (DBI). To address the challenges and achieve these objectives, the application of data mining using the K-Medoids algorithm was necessary. The clustering results from 1,085 customers showed that 30% of the customers, totaling 314, were eligible for rewards, 20%, totaling 540 customers, and 10%, totaling 231 customers. The evaluation of cluster quality using the DBI produced a value of 0.368812, indicating that this value was close to 0, or relatively small. Therefore, it was concluded that the clusters generated were of good quality. This study provides significant contributions to PT DGJ in optimizing its reward distribution strategy to improve efficiency and accuracy in customer grouping.

**Keywords:** Data Mining, Clustering, K-Medoids, Customers, Reward

## RIWAYAT HIDUP

Penulis bernama lengkap MENIATI ZEBUA, lahir di Sungai Rumbia 1, Rokan Hilir, Riau pada 26 Oktober 2001 anak pertama dari 4 bersudara dari ayah MARTINUS ZEBUA dan Ibu ADELIA NDURU. Tahun 2020 penulis lulus dari SMA Tunas Bangsa Rokan Hilir, Riau. Dan pada 2020 terdaftar sebagai mahasiswa Program Studi Teknik Informatika Universitas Medan Area.

Berkat pertolongan Tuhan Yesus, usaha serta doa dari kedua orang tua penulis dapat menjalani aktivitas akademik di Universitas Medan Area, dan puji Tuhan penulis dapat menyelesaikan Tugas Akhir/skripsi berjudul “Penerapan *Data Mining* Menggunakan Algoritma *K-Medoids* Dalam Pengelompokan Nasabah Penerima *Reward* Pada PT Dotri Gadai Jaya”





## KATA PENGANTAR

Puji syukur kepada Tuhan Yang Maha Esa atas berkat dan rahmat-Nya sehingga penulis dapat menyelesaikan skripsi dengan judul “**Penerapan Data Mining Menggunakan Algoritma K-Medoids Dalam Pengelompokan Nasabah Penerima Reward pada PT Dotri Gadai Jaya**” ini dengan baik.

Skripsi ini disusun untuk memenuhi salah satu syarat menyelesaikan pendidikan Program Strata-1 pada Fakultas Teknik Program Studi Teknik Informatika di Universitas Medan Area. Dalam proses menyelesaikan skripsi ini, penulis menyadari bahwa tulisan ini masih jauh dari kata sempurna dan juga terdapat banyak kekurangan. Penulis juga mengharapkan kritik dan saran yang sifatnya membangun dari para pembaca. Kemudian penulis ingin mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. Dadan Ramdan, M.Eng., M.Sc., selaku Rektor Universitas Medan Area.
2. Bapak Dr.Eng. Supriatno, S.T., M.T. selaku Dekan Fakultas Teknik Universitas Medan Area.
3. Bapak Rizki Muliono, S.Kom., M.Kom., selaku Kepala Program Studi Teknik Informatika yang juga selaku dosen pembimbing yang telah meluangkan waktu untuk membimbing dan memberikan arahan dengan penuh kesabaran kepada penulis dalam proses menyelesaikan skripsi ini.
4. Bapak Dr. Dian Noviandri, ST, M.Kom sebagai ketua, bapak Andre Hasudungan Lubis, S.Ti, M.Sc sebagai sekretaris dan bapak Dr. Sayuti Rahman, ST, M.Kom sebagai pembanding II yang telah memberikan saran

dan masukan dalam skripsi penulis dan meluangkan waktunya menjadi panitia dalam seminar proposal sampai penulis sidang.

5. Seluruh Dosen Teknik Informatika Universitas Medan Area yang selama ini telah membekali penulis dengan ilmu yang sangat bermanfaat.
6. IT Support Teknik Informatika bang Robby Kurniawan Sari Damanik, ST yang telah banyak membantu penulis dalam menyiapkan pemberkasan yang diperlukan dari awal sampai selesai.
7. Teristimewa kepada kedua orang tua penulis, Bapak Martinus Zebua dan Ibu Adelia Nduru yang dengan penuh kasih sayang telah memberikan semangat dan selalu mendoakan penulis dapat menyelesaikan pendidikan hingga perguruan tinggi.
8. Adik-adik tersayang, Antorius Zebua, Seniawati Zebua dan Arman Zebua yang selalu mendukung dan membantu serta mendoakan penulis hingga berhasil menyelesaikan pendidikan saat ini.
9. Masriyanto Tafanao, S.H. yang selalu mendukung, mendoakan serta sangat banyak membantu penulis hingga berhasil menyelesaikan skripsi ini.
10. Teman satu kost yaitu kak Rena Gulo dan Kak Yeprien Zalukhu yang selalu memberi semangat dan mendoakan penulis hingga berhasil menyelesaikan skripsi ini.
11. Muhammad Qori Ramadhan Nasution, S.Kom dan Teman-teman Teknik Informatika 2020, terima kasih atas persahabatan dan persaudaraannya selama ini. Semoga Tuhan memudahkan untuk menyelesaikan study S-1 ini dan semoga kita bisa sukses bersama.

12. Dan yang paling terakhir dan paling penting saya mengucapkan terima kasih kepada diri saya sendiri yaitu Meniati Zebua yang telah berusaha kuat sejauh ini walaupun banyak takut dan mengeluhnya tetapi tidak berhenti berusaha dan berjuang. Semoga kedepannya menjadi seorang yang sukses dan menggapai cita-cita, Amin.

Medan, 2024

Meniati Zebua

208160009



## DAFTAR ISI

ABSTRAK.....	vi
ABSTRAC .....	<b>Error! Bookmark not defined.</b>
KATA PENGANTAR .....	ix
DAFTAR ISI.....	xii
DAFTAR GAMBAR.....	xiii
DAFTAR TABEL .....	xiv
BAB I	
PENDAHULUAN .....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	3
1.3 Tujuan Penelitian .....	4
1.4 Manfaat Penelitian .....	4
1.5 Batasan Masalah .....	4
1.6 Sistematika Penulisan.....	5
BAB II.....	
LANDASAN TEORI.....	6
2.1 Data Mining.....	6
2.2 Clustering.....	9
2.3 Algoritma K-Medoids .....	9
2.4 Nasabah .....	11
2.5 Reward .....	11
2.6 Flowchart.....	12
2.7 Use case Diagram .....	13
2.9 Davies-Bouldin Index (DBI) .....	14
2.10 Penelitian Terdahulu .....	16
BAB III .....	
METODOLOGI PENELITIAN.....	18
3.1 TAHAPAN PENELITIAN.....	18
3.1.1 Identifikasi Masalah.....	18
3.1.2 Pengumpulan Data.....	19
3.1.3 Data Selection.....	20
3.1.4 Pre-processing .....	20
3.1.5 Perhitungan K-Medoids .....	23
3.1.6 Evaluasi Cluster dengan DBI .....	26
3.2 Pseudocode.....	28
3.3 Perancangan Sistem .....	29
3.3.1 Perancangan UML .....	29
3.2.2 Perancangan User Interface .....	31
BAB IV .....	
HASIL DAN PEMBAHASAN.....	36
4.2 Pembahasan .....	42
BAB V .....	
KESIMPULAN DAN SARAN.....	44
5.1 Kesimpulan .....	44
5.2 Saran .....	44
DAFTAR PUSTAKA .....	45

## DAFTAR GAMBAR

Gambar 2.1 Tahap proses Data Mining .....	7
Gambar 3. 1 Tahapan Penelitian.....	18
Gambar 3. 2 Use Case Diagram .....	29
Gambar 3. 3 Activity Diagram .....	30
Gambar 3. 4 Rancangan Form Login.....	31
Gambar 3. 5 Rancangan Form Home .....	31
Gambar 3. 6 Rancangan Form Upload Data .....	32
Gambar 3. 7 Rancangan Form Hasil Proses Data Awal .....	33
Gambar 3. 8 Rancangan Form Hasil proses K-Medoids .....	34
Gambar 3. 9 Rancangan Form Hasil Cluster dan Evaluasi .....	35
Gambar 4. 1 Form Login.....	36
Gambar 4. 2 Form Utama.....	37
Gambar 4. 3 Form Upload Data .....	37
Gambar 4. 4 Form Data Awal .....	38
Gambar 4. 5 Form Hasil Proses K-Medoids .....	39
Gambar 4. 6 Form Hasil Cluster.....	40
Gambar 4. 7 Laporan Hasil Pengelompokan.....	40

## DAFTAR TABEL

Tabel 2. 1 Flowchart .....	12
Tabel 2. 2 Simbol-simbol Usecase Diagram .....	13
Tabel 2. 3 Simbol-simbol Activity Diagram .....	14
Tabel 2. 4 Penelitian Terdahulu.....	16
Tabel 3. 1 Contoh Data Transaksi Pinjaman Nasabah.....	19
Tabel 3. 2 Transformasi Data "Bunga Pinjaman" .....	21
Tabel 3. 3 Transformasi Data "Status Pinjaman" .....	21
Tabel 3. 4 Nilai Max dan Nilai Min.....	22
Tabel 3. 5 Hasil Normalisasi Data .....	22
Tabel 3. 6 Medoids Awal Sebagai Iterasi 1 .....	23
Tabel 3. 7 Jarak Data Terhadap Cluster Pada Iterasi Ke-1 .....	24
Tabel 3. 8 Medoid Baru Pada Iterasi 2 .....	24
Tabel 3. 9 Perhitungan Iterasi 2.....	24
Tabel 3. 10 Hasil Perhitungan K-Medoids.....	25

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Meningkatnya pertumbuhan industri keuangan menjadikan perusahaan menghadapi persaingan yang ketat. Dalam menghadapi persaingan tersebut perusahaan tidak hanya diharapkan untuk bertahan, tetapi juga dituntut untuk mengoptimalkan layanan agar sesuai dengan kebutuhan nasabah. Meningkatnya persaingan mendorong perusahaan harus mengetahui apa yang di butuhkan oleh nasabah serta berusaha untuk memenuhinya dengan cara memberikan pelayanan terbaik dan memuaskan, sehingga dapat menguasai pasar (Sabella dkk., 2022).

Pemberian *reward* kepada nasabah adalah salah satu strategi yang umum digunakan untuk meningkatkan loyalitas nasabah atau pelanggan. Memberikan *reward* yang relevan dan sesuai dengan kebutuhan nasabah dapat menjadi poin dalam meningkatkan kepuasan nasabah. Program *reward* yang efektif dapat menjadi pendorong kuat dalam mempertahankan pelanggan, meningkatkan loyalitas, dan pada akhirnya, menciptakan hubungan jangka panjang yang bermanfaat.

PT Dotri Gadai Jaya atau biasa disebut PT DGJ adalah perusahaan pergadaian swasta yang bergerak di bidang jasa pemberian pinjaman dengan dengan jaminan berupa barang. PT DGJ salah satu perusahaan swasta yang memiliki nasabah yang cukup banyak, dalam mempertahankan dan meningkatkan loyalitas nasabah salah satu strategi yang dilakukan PT DGJ adalah memberikan *reward* kepada nasabah berdasarkan jumlah transaksi pinjaman gadai. *Reward* yang diberikan berupa penambahan pinjaman sebesar 30%, 20%, 10% dari total

pinjaman awal. Program ini diadakan setiap ulang tahun PT DGJ. Namun permasalahan yang muncul saat ini, perusahaan mengalami kesulitan dalam mengelompokkan nasabah yang pantas menerima *reward* tersebut. Selama ini pengelompokan nasabah penerima *reward* masih dilakukan secara manual yang dapat memakan waktu yang cukup lama, adanya potensi kesalahan dan kurangnya transparan.

Untuk membantu perusahaan dalam mengelompokkan nasabah penerima *reward*, *Data Mining* dapat menjadi solusi dalam menyelesaikannya. *Data mining* adalah proses pengambilan data dalam jumlah yang besar untuk mendapatkan informasi dari sistem yang dibutuhkan pengguna (Shinta Sibatuara dkk., 2022). Salah satu teknik *data mining* yang sering digunakan adalah *clustering*. *Clustering* merupakan proses pengelompokan sekumpulan objek ke dalam kelas yang serupa (Bahauddin dkk., 2021). Pada *clustering* terdapat beberapa algoritma yaitu *K-Means*, *Fuzzy C-Means*, *DBSCAN* dan *K-Medoids*. Pada penelitian ini algoritma yang digunakan adalah *K-Medoids*. *K-medoids* merupakan salah satu algoritma yang mampu mengatasi kelemahan dari *K-Means* yang *sensitive* terhadap outlier (Andini & Arifin, 2020). *K-medoids* deterministik jika inisialisasi medoidnya tetap sama, artinya akan memberikan hasil yang konsisten setiap kali dijalankan dengan kondisi yang sama sedangkan *Fuzzy C-Means* menggunakan derajat keanggotaan fuzzy yang bersifat probabilistik, sehingga hasil akhirnya mungkin tidak selalu konsisten dan dapat bervariasi tergantung pada inisialisasi dan parameter *fuzziness*. *K-medoids* hanya memerlukan satu parameter utama, yaitu jumlah kluster ( $k$ ), yang membuatnya lebih mudah untuk diatur dalam beberapa kasus, *DBSCAN* memerlukan dua parameter utama, epsilon (jarak maksimum



untuk dianggap tetangga) dan minPts (jumlah minimum titik untuk membentuk kluster). Pemilihan parameter yang tidak tepat dapat sangat mempengaruhi hasil klusterisasi.

Penelitian mengenai algoritma *K-Medoids* pernah dilakukan oleh Aulanda, dkk. (2021) penelitian tersebut tentang pengelompokan pembiayaan nasabah klaim asuransi pengguna kendaraan bermotor menggunakan *K-Medoids* berhasil menampilkan hasil *cluster* dengan presentase keakuratan sebesar 100%. Kemudian, Mirantika, dkk. (2023) penelitian ini mengimplementasikan algoritma *K-Medoids* dalam menentukan segmentasi pelanggan untuk membantu perusahaan membuat strategi bisnis dalam mempertahankan hubungan dengan pelanggan.

Berdasarkan latar belakang yang telah diuraikan maka penelitian ini mengambil judul yaitu “Penerapan *Data Mining* menggunakan Algoritma *K-Medoids* dalam Pengelompokan Nasabah Penerima *Reward* pada PT Dotri Gadai Jaya”.

## 1.2 Rumusan Masalah

Adapun yang menjadi rumusan masalah dalam penelitian ini adalah bagaimana menerapkan *data mining* menggunakan algoritma *k-medoids* untuk mengelompokan nasabah penerima *reward* pada PT DGJ?

### 1.3 Tujuan Penelitian

Salah satu tujuan penelitian ini adalah sebagai berikut:

1. Menerapkan algoritma *K-Medoids* dalam pengelompokan nasabah penerima *reward* di PT DGJ.
2. Mengetahui hasil pengelompokan nasabah penerima *reward* pada PT DGJ.
3. Mengetahui hasil evaluasi pengelompokan dengan *Davies-Bouldin Index* (DBI).

### 1.4 Manfaat Penelitian

Adapun manfaat penelitian ini adalah:

1. Penelitian ini dapat menjadi rujukan pada PT DGJ untuk menentukan kelompok nasabah yang mendapatkan *reward*.
2. Penelitian ini dapat menjadi rujukan dalam mengelompokkan nasabah menggunakan algoritma *K-Medoids*.

### 1.5 Batasan Masalah

Adapun batasan masalah dalam penelitian ini, yaitu:

1. Data yang digunakan adalah data primer yang diperoleh dari data transaksi nasabah PT DGJ tahun 2023 dengan jumlah data sebanyak 1.085.
2. Penerapan *cluster* yang akan digunakan dalam penelitian ini adalah *K-Medoids*.
3. Variabel pengelompokan nasabah yang digunakan yaitu jumlah pinjaman, bunga pinjaman dan status transaksi.
4. Penelitian ini menggunakan *Davies Bouldin Index* (DBI) untuk mengevaluasi hasil pengelompokan.

## 1.6 Sistematika Penulisan

Dalam penulisan skripsi ini, penulis memberikan gambaran sistematika penulisan dalam mengelompokan nasabah penerima *reward* pada PT DGJ terbagi mejadi 5 (lima) bagian yaitu:

### **BAB I PENDAHULUAN**

Bagian ini berisi tentang latar belakang penelitian, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan msalah dan sistematika penulisan.

### **BAB II TINJAUAN PUSATAKA**

Bab ini membahas tentang teori yang diambil dari beberapa sumber yang sesuai dengan permasalahan dalam penelitian ini.

### **BAB III METODOLOGI PENELITIAN**

Bab ini membahas mengenai metodologi penelitian yang akan dilakukan serta langkah-langkah yang digunakan pada penelitian yang akan dilakukan

### **BAB IV HASIL DAN PEMBAHASAN**

Bab ini berisi pengujian dan analisis dari hasil pengujian dari sistem yang telah dibangun berdasarkan hasil perancangan pada bab 3 sebelumnya.

### **BAB V PENUTUP**

Bab ini berisi kesimpulan dan saran seluruh penelitian yang dilakukan.

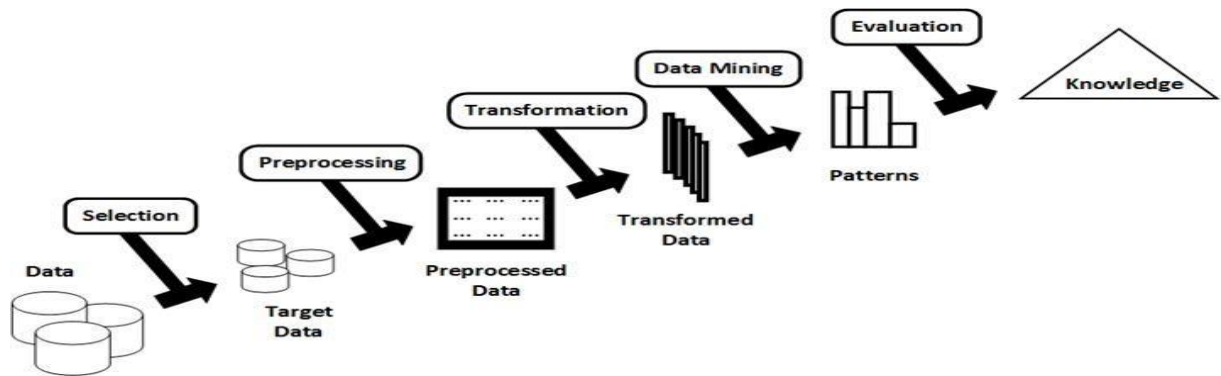
## BAB II

### LANDASAN TEORI

#### 2.1 *Data Mining*

*Data Mining* adalah proses pengambilan data yang besar menjadi suatu informasi penting yang dapat digunakan (Rahmah dkk., 2022). Informasi penting ini didapat dari sebuah proses yang rumit seperti menggunakan ilmu matematika, *artificial intelligence* (AI), teknik statistik, dan *machine learning* (Sudarsono dkk., 2021). Pengertian lain *data mining* ialah proses penggalian data yang besar yang sangat penting digunakan untuk keputusan suatu bisnis (Ahmad dkk., 2022). Tujuan utama dari *Data Mining* yaitu untuk mendapatkan pengetahuan dari sejumlah data yang besar (Z. Nabila dkk., 2021).

*Data Mining* merupakan disiplin ilmu yang mempelajari metode untuk mengekstrak informasi atau menemukan pola dari data yang besar (Wibowo & Jananto, 2020). *Data Mining* sangat penting dilakukan dalam mengolah data yang besar untuk mempermudah pencatatan sebuah transaksi dan mengolah data agar memberikan informasi yang tepat pada penggunaannya (Situmorang dkk., 2023). Istilah *data mining* sering disebut sebagai *Knowledge Discovery in Database* (KDD). KDD adalah proses yang meliputi pengumpulan, penggunaan data, historis untuk menemukan pola atau hubungan set data berukuran besar. Sebagai rangkaian proses, *data mining* dibagi menjadi beberapa tahap yaitu:



Gambar 2.1 Tahap proses *Data Mining*

(Sumber: Nasir, 2021)

Berdasarkan Gambar 2.1 tahapan *data mining* terbagi atas:

1. *Data Selection*

Sekumpulan data harus melalui tahap KDD. Data yang digunakan harus melalui proses seleksi data dari sekumpulan data, kemudian data disimpan dalam database secara terpisah dan disimpan dalam suatu file/berkas.

2. *Pre-processing*

a. Pembersihan Data

Proses mempersiapkan dan membersihkan data mentah sebelum dianalisis atau dimodelkan.

b. Transformasi Data

Proses mengolah atau mengubah data dalam bentuk tertentu agar lebih sesuai untuk analisis atau pemodelan.

c. Normalisasi Data

Proses mengubah nilai-nilai dalam suatu data set sehingga data tersebut memenuhi kriteria tertentu atau berada dalam rentang yang spesifik. Untuk

data yang bersifat numerik dilakukan normalisasi untuk menstandarisasi perbedaan skala yang dapat memberikan efek pada hasil yang diperoleh. Berikut rumus normalisasi menggunakan metode Min-Max data yang umum digunakan:

$$X_{baru} = \frac{X - \min A}{\max A - \min A} \quad (2.1)$$

Dimana:

$X_{baru}$  = Nilai hasil normalisasi

$X$  = Nilai Asli

$\min A$  = Nilai minimum dari data

$\max A$  = Nilai maximum dari data

### 3. *Data Mining*

*Data Mining* merupakan teknik dalam pembentukan pola sehingga menghasilkan informasi yang dibutuhkan. Pemilihan metode yang dibutuhkan bergantung pada tujuan dan proses KDD secara keseluruhan.

### 4. *Evaluation*

Pola yang dihasilkan maka didapatkanlah sebuah informasi yang mudah dipahami oleh pengguna dan menjadi sumber pengetahuan dalam pengambilan keputusan.

## 2.2 Clustering

*Clustering* salah satu teknik yang cukup dikenal dan banyak digunakan dalam *data mining*. *Clustering* sangat penting dalam beberapa permasalahan antarlain analisa pola, pembuatan keputusan, *machine learning*, *data mining* dan sebagainya (Paembonan & Abduh, 2021). *Clustering* adalah proses membagi sekumpulan objek data dari satu set menjadi beberapa kelas (Simanjuntak dkk., 2023). Menurut Aulia (2021), *clustering* ialah sekumpulan objek data yang memiliki kemiripan dalam cluster yang sama dan tidak cocok terhadap objek yang berbeda *cluster*. Objek akan dikelompokan ke dalam satu atau lebih *cluster* sehingga objek yang berada dalam satu *cluster* mempunyai kesamaan yang tinggi antara satu dengan lainnya. Terdapat banyak algoritma *clustering* yang telah digunakan oleh peneliti sebelumnya seperti *k-means*, *improvet k-means*, *Fuzzy C-Means*, *DBSCAN*, *K-Medoids (PAM)*, *CLARANS* dan *fuzzy subtractive*.

## 2.3 Algoritma K-Medoids

Algoritma *K-Medoids* atau PAM (*algoritma Partitioning Around Medoids*) adalah teknik *data mining* untuk membagi data kedalam beberapa kelompok berdasarkan jarak kriteria kondisi atau karakteristik (Situmorang dkk., 2023). *K-Medoids* adalah algoritma *clustering* yang mirip dengan *K-Means*. *K-Medoids* lebih kuat daripada *K-Means*, karena algoritma *K-medoids* menggunakan objek sebagai pusat *cluster* untuk setiap *cluster* sedangkan *K-Means* menggunakan nilai rata-rata (*mean*) sebagai pusat *cluster* (Herviany dkk., 2021). Menurut Zeilani (2022) kelebihan dari algoritma *K-Medoids* yaitu hasil proses *clustering* tidak bergantung pada urutan masuk data set. Beberapa langkah-langkah perhitungan algoritma *K-Medoids* yaitu (Ningrum dkk., 2021):

1. Inisialisasi pusat *cluster* sebanyak  $k$  (jumlah *cluster*)
2. Hitung setiap objek ke *cluster* terdekat menggunakan persamaan ukuran jarak *Euclidian Distance* dengan persamaan (2.2):

$$d(i, j) = \sqrt{(x_{1i} - x_{1j})^2 + (x_{2i} - x_{2j})^2 + \dots + (x_{ki} - x_{kj})^2} \quad (2.3)$$

Dimana:

$d(I, j)$  = jarak data  $I$  ke pusat cluster  $j$

$x_{ki}$  = data ke  $i$  pada atribut data ke  $k$

$x_{kj}$  = titik pusat ke  $j$  pada atribut ke  $k$

3. Selanjutnya tentukan objek pada masing-masing *cluster* sebagai *medoid* baru secara acak.
4. Hitung masing-masing jarak terhadap *medoids* dengan menggunakan rumus *Euclidian Distance* diatas.
5. Selanjutnya hitung simpangan ( $S$ ) dengan cara menghitung selisih nilai dari total *distance* baru dan total *distance* lama. Jika  $S < 0$ , maka tukar objek dengan data *cluster* untuk membentuk sekumpulan  $k$  objek baru sebagai *medoids*.
6. Kemudian ulangi langkah ke-3 sampai ke-5 hingga tidak terjadi perubahan *medoids*, sehingga didapatkan *cluster* beserta anggota *cluster* masing-masing.



## 2.4 Nasabah

Nasabah adalah orang yang memiliki hubungan atau menjadi pelanggan tetap dan sering berinteraksi dalam waktu yang cukup lama (Andriyani & Ardianto, 2020). Menurut OJK Otoritas Jasa Keuangan, pengertian nasabah adalah individu atau kelompok yang memanfaatkan atau menggunakan fasilitas perbankan, baik dalam bentuk jasa maupun produk. Pengertian lain menurut Berutu (2021) nasabah merupakan pihak yang menggunakan dan secara sengaja berlangganan dengan bank yang dipercayainya. Dengan kata lain nasabah adalah orang atau badan hukum yang mempunyai rekening atau pinjaman pada pihak bank (Nurdin dkk., 2020).

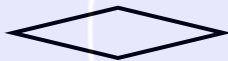




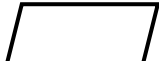
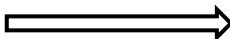


## 2.5 Reward

*Reward* merupakan salah satu cara untuk memotivasi seseorang agar meningkatkan prestasinya (Kiu dkk., 2021). *Reward* adalah bentuk penghargaan atau apresiasi kepada suatu pencapaian tertentu yang diberikan oleh perorangan atau suatu organisasi yang biasanya diberikan dalam bentuk materi atau ucapan (Zai & Aripin, 2021). Menurut Saleh & Mardiana (2021) *reward* merupakan hadiah, imbalan atau penghargaan yang bertujuan agar seseorang menjadi giat lagi untuk memperbaiki atau meningkatkan kinerja yang telah dicapai. Pemberian *reward* dalam dunia bisnis merupakan cara untuk membangun hubungan bisnis yang bertahan lama.

## 2.6 Flowchart

Flowchart merupakan representasi visual dari langkah-langkah penyelesaian masalah atau proses yang dituliskan dalam bentuk diagram (Khesya, 2021). Flowchart bertujuan untuk menggambarkan langkah penyelesaian masalah dengan sederhana dan jelas menggunakan simbol-simbol (Furqani & Muliono, 2021). Menyajikan tahapan penyelesaian masalah secara sederhana, teratur dan rapi dengan menggunakan simbol-simbol yang standar yang dapat di mengerti oleh orang yang terlibat dalam proses pengembangan atau analisis adalah tujuan dari penggunaan flowchart.

Tabel 2. 1 Flowchart  
(Sumber: Khesya, 2021)





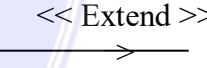
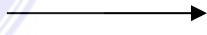
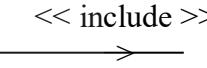
No	Simbol	Fungsi
1		Pemilihan proses berdasarkan kondisi yang ada.
2		Pelaksanaan suatu bagian (sub program).
3		Penghubung proses pada halaman yang sama.
4		Permulaan atau akhir dari suatu program.
5		Proses pengolahan data yang dilakukan komputer.
6		Proses <i>input</i> atau <i>output</i> data/informasi.
7		Penghubung antara simbol satu dengan simbol yang lain.
8		Penghubung alur yang berada pada halaman berbeda.
9		Proses pemberian nilai awal suatu variabel.

## 2.7 Use case Diagram

*Use case* diagram digunakan untuk menggambarkan hubungan antara aktor dan sistem (S. Nabila dkk., 2021). *Use case* mendeskripsikan interaksi antara satu atau lebih aktor dengan sistem informasi yang akan dibuat (Hardiyanti, 2021).

Tabel 2. 2 Simbol-simbol *Use case* Diagram

(Sumber: Hardiyanti, 2021)







Nama	Keterangan	Gambar
<i>Use case</i>	Menerangkan apa yang dikerjakan system.	
<i>Actor</i>	Menggambarkan orang, sistem atau eksternal entitas/stakeholder yang menyediakan atau menerima informasi dari sistem.	
<i>Sistem Boundary</i>	Menggambarkan Jangkauan sistem.	
<i>Association</i>	Menggambarkan bagaimana aktor terlihat dalam use case.	
<i>Extend</i>	Perluasan dari use case lain jika kondisi atau syarat terpenuhi.	
<i>Generalization</i>	Dibuat ketika ada sebuah keadaan yang lain/perlakuan khusus.	
<i>Include</i>	Menjelaskan bahwa <i>use case</i> termasuk didalam <i>use case</i> lain.	

## 2.8 Activity Diagram

*Activity* diagram menggambarkan aliran kerja, kondisi, kendala atau aktivitas dari sebuah sistem pada perangkat lunak (Pakaya dkk., 2020). Menurut Widyastuti (2022) *activity* diagram merupakan gambaran proses alir dari aktivitas di dalam sistem yang berjalan.

Tabel 2. 3 Simbol-simbol *Activity Diagram*

(Sumber: Widyastuti, 2022)

Nama	Simbol	Deskripsi
Status awal		Status awal aktivitas sistem, sebuah diagram aktivitas memiliki sebuah status awal
Aktivitas		Aktivitas yang dilakukan sistem, aktivitas biasanya diawali dengan kata kerja
Percabangan/ <i>decision</i>		Asosiasi percabangan dimana jika ada pilihan aktivitas lebih dari satu
penggabungan/ <i>join</i>		Asosiasi penggabungan dimana lebih dari satu aktivitas digabungkan menjadi satu
Status akhir		Status akhir yang dilakukan sistem, sebuah diagram aktivitas memiliki sebuah status akhir
Swimlane		Memisahkan organisasi bisnis yang bertanggung jawab terhadap aktivitas yang terjadi

## 2.9 *Davies-Bouldin Index (DBI)*

Metode yang diperkenalkan oleh D.L. Davies dan D.W. Bouldin dan nama metode ini menggunakan kedua nama yaitu Davies-Bouldin (DBI). DBI adalah ukuran untuk mengevaluasi kinerja pengelompokan (Sopyan dkk., 2022). Evaluasi kualitas *cluster* mempunyai peran penting dalam memahami efektivitas algoritma klasterisasi adalah salah satu tujuan DBI yang mengukur dissimilaritas rata-rata antara *clusternya* (Umagapi dkk., 2023). Berikut adalah langkah-langkah perhitungan DBI:

1. *Sum Of Square Within-Cluster (SSW)*

Menghitung nilai dari SSW. Dengan rumus sebagai berikut:

$$SSW_i = \frac{1}{m_i} \sum_{j=i}^{m_i} d(X_j, C_j) \quad (2.4)$$

Dimana:

$m_i$  = jumlah data dalam *cluster* ke- $i$

$c_i$  = titik pusat *cluster* ke- $i$

$d(X_j, C_j)$  = jarak setiap data ke titik pusat  $i$  yang dihitung menggunakan jarak *euclidiance*

2. *Sum Of Square Between-Cluster (SSB)*

Perhitungan SSB bertujuan untuk mengetahui jarak antar *cluster*. Dengan rumus perhitungan sebagai berikut:

$$SSB_{ij} = d(X_i, X_j) \quad (2.5)$$

Dimana :

$d(X_i, X_j)$  = jarak antara data ke  $i$  dengan data ke  $j$  di *cluster* lain.

3. *Ratio (Rasio)*

Perhitungan rasio untuk mengetahui nilai perbandingan antara cluster ke- $i$  dan cluster ke- $j$ . Untuk menentukan nilai rasio dengan rumus sebagai berikut:

$$R_{ij, \dots, n} = \frac{SSW_i + SSW_j + \dots + SSW_n}{SSB_{i,j} + \dots + SSB_{n,i,nj}} \quad (2.6)$$

Dimana :

$SSW_i$  = *Sum Of Square Within-Cluster* pada titik pusat  $i$

$SSB_{i,j}$  = *Sum of Square Between Cluster* data ke  $i$  dengan  $j$  pada *cluster* yang berbeda

#### 4. Davies Bouldin Index (DBI)

Nilai rasio yang diperoleh dari persamaan 2.5 digunakan untuk mencari nilai DBI dengan menggunakan perhitungan sebagai berikut :

$$DBI = \frac{1}{k} \sum_{i=1}^k \max_{i \neq j} (R_i, j, \dots k) \quad (2.7)$$

Dari perhitungan DBI dapat disimpulkan bahwa jika semakin kecil nilai DBI yang diperoleh (non negatif  $\geq 0$ ) maka *cluster* tersebut semakin baik.

### 2.10 Penelitian Terdahulu

Dalam menyusun penelitian ini, terdapat beberapa penelitian terdahulu yang dijadikan sebagai referensi. Adapun penelitian terdahulu yang berkaitan dengan pengelompokan menggunakan algoritma *K-Medoids* yaitu:

Tabel 2. 4 Penelitian Terdahulu

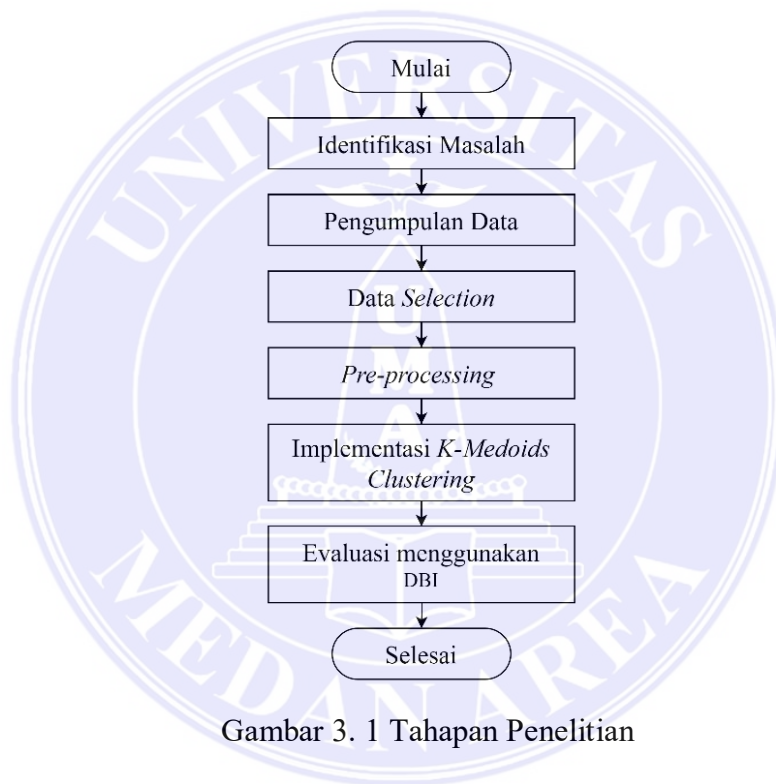
No	Peneliti (Tahun)	Judul Penelitian	Hasil Penelitian
1	Silaban, (2023)	Analisis Pengelompokan Tingkat Minat Baca Siswa Menggunakan Algoritma <i>K-Medoids</i>	Pada penelitian ini, menggunakan 4 variabel yaitu daya tarik siswa membaca, data peminjam buku, jangka waktu dan frekuensi membaca di perpustakaan. Hasil penelitian menunjukkan bahwa algoritma <i>K-Medoids</i> dapat membantu pengelompokan Tingkat minat baca siswa.
2	Gulo, (2023)	Penerapan Metode	Dengan menggunakan <i>K-medoids</i>

		K-Medoids Untuk Pengelompokan Penduduk Kurang Mampu di kecamatan Afulu	dalam pengelompokan penduduk kurang mampu dapat membantu mempercepat proses pengelompokan pada kecamatan Afulu.
3	Puspita dkk. (2023)	Pengelompokan Puskesmas Berdasarkan Kasus Balita Stunting di Kabupaten Paser Menggunakan Metode K-medoids	Hasil penelitian yang diperoleh dengan menggunakan 2 <i>cluster</i> yaitu nilai DBI sebesar 0,977
4	Susilawaty dkk. (2021)	Penerapan Algoritma K-Medoids untuk Menentukan Segmentasi Pelanggan	Hasil penelitian yang telah dilakukan, segmentasi pelanggan dengan menerapkan algoritma K-Medoids menghasilkan jumlah <i>cluster</i> optimal adalah 3 yaitu yaitu Lost Customer, Core Customer, dan New Customer dengan uji validitas <i>cluster</i> dengan nilai DBI 1,030.

## BAB III METODOLOGI PENELITIAN

### 3.1 Tahapan Penelitian

Berikut tahapan penelitian yang akan menjadi pedoman untuk mengelompokkan nasabah penerima *reward*. Tahapan yang digunakan pada penelitian ini dijelaskan pada gambar dibawah ini:



Gambar 3. 1 Tahapan Penelitian

#### 3.1.1 Identifikasi Masalah

Tujuan dari penelitian ini adalah mengidentifikasi masalah mengenai pengelompokan nasabah penerima *reward* di PT DGJ. Untuk mengetahui hasil pengelompokan nasabah penerima *reward* pada PT DGJ diperlukannya penerapan *data mining* yaitu menggunakan algoritma *k-medoids* menggunakan bahasa pemrograman *Python*.



### 3.1.2 Pengumpulan Data

Pengumpulan data bertujuan untuk memperoleh informasi yang dibutuhkan dalam mencapai tujuan penelitian. Teknik pengumpulan data dalam penelitian ini dilakukan dengan cara wawancara dan observasi langsung di PT Dotri Gadai Jaya.

#### a. Wawancara

Wawancara dilakukan untuk mencari informasi mengenai proses pengelompokan nasabah penerima *reward* pada PT Dotri Gadai Jaya

#### b. Observasi

Observasi penelitian melakukan pengamatan secara langsung kegiatan pengolahan data nasabah penerima *reward*

#### c. Dokumentasi

Pengambilan data yang diperlukan untuk mengelompokan nasabah penerima *reward*, yakni data transaksi nasabah.

#### a. Data

Data yang digunakan sebanyak 1.085 yang berisi Nama Nasabah, Bunga, Status Pinjaman.

Tabel 3. 1 Contoh Data Transaksi Pinjaman Nasabah

No	Nama Nasabah	Jumlah Pinjaman	Bunga	Status Pinjaman
1	N1	Rp70,000,000	5 %	Kurang Lancar
2	N2	Rp 50,100,000	5 %	Lancar
3	N3	Rp 55,000,000	5 %	Lancar
4	N4	Rp 50,000,000	10 %	Tidak Lancar
5	N5	Rp 61,100,000	5 %	Lancar
6	N6	Rp 57,000,000	5 %	Lancar
7	N7	Rp 65,600,000	5 %	Tidak Lancar
8	N8	Rp 56,900,000	5 %	Tidak Lancar
9	N9	Rp92,500,000	5 %	Lancar
10	N10	Rp 51,500,000	10%	Kurang Lancar

## **b. Alat yang Digunakan**

Dalam menyelesaikan penelitian ini, beberapa alat atau perangkat lunak yang mungkin digunakan untuk mendukung penelitian tersebut adalah sebagai berikut:

1. Spesifikasi Kebutuhan *Hardware*
  - a. *Processor: Intel Core i3-1115G4*
  - b. *Ram: 8 GB*
  - c. *SSD: 256 GB*
  - d. *System tipe: 64-bit Operating System*
2. Spesifikasi Kebutuhan *Software*
  - a. *Sistem Operasi: Microsoft windows 11*
  - b. *Bahasa Pemrograman: Python*
  - c. *Google Colab*

### **3.1.3 Data Selection**

*Data Selection* adalah proses seleksi penggunaan variabel atau atribut yang paling relevan dari data set yang tersedia untuk mencapai tujuan penelitian.

Variabel yang digunakan dalam penelitian ini yaitu:

1. Jumlah Pinjaman
2. Bunga Pinjaman
3. Status Pinjaman

### **3.1.4 Pre-processing**

#### **a. Pembersihan Data**

Proses mempersiapkan dan membersihkan data mentah sebelum dianalisis atau dimodelkan. Tujuan dari pembersihan data adalah untuk memastikan data yang digunakan kualitas yang baik dan dapat memberikan hasil yang lebih akurat.

b. Transformasi Data

Proses mengolah atau mengubah data dalam bentuk numerik agar sesuai untuk analisis atau pemodelan. Tujuannya adalah untuk menyamakan skala atribut data dalam range yang lebih spesifik, sehingga data dapat diolah dengan mudah menggunakan metode *K-Medoids Clustering*, tahap ini dilakukan dengan mengubah data bunga pinjaman dan status pinjaman.

Berikut adalah bobot transformasi data terhadap kriteria atau variabel Bunga Pinjaman:

Tabel 3. 2 Transformasi Data "Bunga Pinjaman"

Kriteria	Pengelompokan	Transformasi
Bunga Pinjaman	5%	1
	10%	2

Berikut adalah bobot transformasi data terhadap kriteria atau variabel Status Pinjaman:

Tabel 3. 3 Transformasi Data "Status Pinjaman"

Kriteria	Pengelompokan	Transformasi
Status Pinjaman	Tidak Lancar	1
	Kurang Lancar	2
	Lancar	3

c. Normalisasi Data

Sebelum melakukan perhitungan manual algoritma *k-medoids*, normalisasikan data untuk mengubah nilai-nilai dalam suatu data set agar memiliki skala yang seragam. Tahap ini dilakukan untuk mengubah data jumlah pinjaman. Sebelum melakukan normalisasi, setiap data harus ditentukan nilai maximum dan nilai minimum.

Tabel 3. 4 Nilai Max dan Nilai Min

No	Nama	X1	X2	X3
1	N1	Rp 70.000.000	1	2
2	N2	Rp 50.100.000	1	3
3	N3	Rp 55.000.000	1	3
4	N4	Rp 50.000.000	2	2
5	N5	Rp 61.100.000	1	3
6	N6	Rp 57.000.000	1	3
7	N7	Rp 65.600.000	1	2
8	N8	Rp 56.900.000	1	2
9	N9	Rp 92.500.000	1	3
10	N10	Rp 51.500.000	2	2
<b>MIN</b>		<b>50.000.000</b>	<b>1</b>	<b>1</b>
<b>MAX</b>		<b>92.500.000</b>	<b>2</b>	<b>3</b>

Normalisasi dapat dilakukan dengan menerapkan rumus persamaan 2.1.

$$X_{baru} = \frac{70.000.000 - 50.000.000}{92.500.000 - 50.000.000} = 0,47$$

$$X_{baru} = \frac{1 - 1}{2 - 1} = \frac{0}{1} = 0$$

$$X_{baru} = \frac{2 - 1}{3 - 1} = \frac{1}{2} = 0,5$$

Lakukanlah perhitungan pada setiap data, sehingga mendapatkan hasil seperti pada Tabel 3.5.

Tabel 3. 5 Hasil Normalisasi Data

No	Nama Nasabah	X1	X2	X3
1	N1	0.333	0	0.5
2	N2	0.002	0	1
3	N3	0.083	0	1
4	N4	0.000	1	0.5

5	N5	0.185	0	1
6	N6	0.117	0	1
7	N7	0.260	0	0.5
8	N8	0.115		0.5
9	N9	0.708	0	1
10	N10	0.025	1	0.5

### 3.1.5 Perhitungan K-Medoids

Adapun perhitungan manual yang dilakukan untuk pengelompokan nasabah penerima *reward*:

1. Inisialisasi pusat kluster sebanyak k (jumlah *cluster*). Kluster sebanyak 3 (k=3), yaitu *reward* 30%, 20%, 10%.
2. Menentukan medoid awal dari setiap kluster secara acak.

Tabel 3. 6 Medoids Awal Sebagai Iterasi 1

Data Ke-	Pusat Cluster	X1	X2	X3
2	C1	0.002	0	1
5	C2	0.185	0	1
10	C3	0.025	1	0.5

3. Menghitung jarak *Euclidian Distance* data yang telah dinormalisasi dengan pusat *cluster* dengan persamaan (2.2):

$$d(1,1) = \sqrt{(0.333 - 0.002)^2 + (0 - 0)^2 + (0,5 - 1)^2} = 0,600002$$

$$d(1,2) = \sqrt{(0.333 - 0.185)^2 + (0 - 0)^2 + (0,5 - 1)^2} = 0,5215$$

$$d(1,3) = \sqrt{(0.333 - 0.025)^2 + (0 - 1)^2 + (0,5 - 0.5)^2} = 1,0464$$

Dan selanjutnya dilakukan menghitung jarak untuk data ke-2 sampai data ke-10. Berikut adalah hasil perhitungan jarak setiap data terhadap pusat *cluster* baru sebagai berikut:

Tabel 3. 7 Jarak Data Terhadap Cluster Pada Iterasi Ke-1

No	C1	C2	C3	JARAK TERPENDEK	CLUSTER
1	0,600002	0,521539	1,046456	0,521538855	2
2	0	0,183333	1,118277	0	1
3	0,081667	0,101667	1,119555	0,081666667	1
4	1,118035	1,133237	0,025	0,025	3
5	0,183333	0	1,129425	0	2
6	0,115	0,068333	1,121786	0,068333333	2
7	0,562793	0,505594	1,027241	0,50559371	2
8	0,512684	0,504876	1,004042	0,504876222	2
9	0,706667	0,523333	1,310322	0,523333333	2
10	1,118277	1,129425	0	0	3
<b>TOTAL KEDEKATAN</b>				<b>2.230342122</b>	

- Pilih medoid baru secara acak objek pada masing-masing *cluster* untuk melakukan iterasi ke-2.

Tabel 3. 8 Medoid Baru Pada Iterasi 2

Data Ke-	Pusat Cluster	X1	X2	X3
1	C1	0.333	0	0.5
3	C2	0.083	0	1
6	C3	0.117	0	1

- Gunakan rumus Euclidian Distance untuk menghitung jarak dari setiap objek pada masing-masing *cluster* dengan medoids baru.

Tabel 3. 9 Perhitungan Iterasi 2

No	C1	C2	C3	JARAK TERPENDEK	CLUSTER
1	0	0,559017	0,544926	0	1
2	0,600002	0,081667	0,115	0,081666667	2

3	0,559017	0	0,033333	0	2
4	1,054093	1,121135	1,124105	1,054092553	1
5	0,521539	0,101667	0,068333	0,068333333	3
6	0,544926	0,033333	0	0	3
7	0,073333	0,530293	0,520139	0,073333333	1
8	0,218333	0,501002	0,500003	0,218333333	1
9	0,625	0,625	0,591667	0,591666667	3
10	1,046456	1,119555	1,121786	1,046455658	1
<b>TOTAL KEDEKATAN</b>				<b>3.133881545</b>	

6. Hitung total simpangan (S) dengan menghitung nilai total kedekaan baru dikurang total kedekaan lama. Jika nilai  $S > 0$  maka iterasi berhenti.

$$S = 3.133881545 - 2.230342122 = 0,903539423$$

Nilai  $S > 0$  maka iterasi berhenti dan hasil yang diperoleh adalah perhitungan sebelumnya seperti pada Tabel 3.9. Berikut hasil perhitungan *K-Medoids*:

Tabel 3. 10 Hasil Perhitungan *K-Medoids*

No	Nama Nasabah	CLUSTER	KETERANGAN CLUSTER
1	N1	1	Reward 30%
2	N2	2	Reward 20%
3	N3	2	Reward 20%
4	N4	1	Reward 30%
5	N5	3	Reward 10%
6	N6	3	Reward 10%
7	N7	1	Reward 30%
8	N8	1	Reward 30%
9	N9	3	Reward 10%
10	N10	1	Reward 30%

### 3.1.6 Evaluasi *Cluster* dengan DBI

Setelah didapat hasil pengklasteran, kemudian dilakukan evaluasi *cluster* dari algoritma *K-Medoids* menggunakan DBI. Untuk mendapatkan nilai DBI, terlebih dahulu hitung nilai SWW, SBB dan Ratio. Untuk melakukan perhitungan DBI data yang digunakan adalah data yang sudah terbentuk pada clustering akhir dan titik pusat cluster yang terakhir seperti pada Tabel 3.8.

1. Menghitung SSW dengan persamaan (2.3)

$$SSW_1 = \frac{0 + 1.054 + 0.0733 + \dots + 1.0464}{5} = 0,478$$

$$SSW_2 = \frac{0,0816 + 0}{2} = 0,04$$

$$SSW_3 = \frac{0.0683 + 0 + 0,5916}{3} = 0,22$$

2. Menghitung SSB dengan persamaan (2.4)

Setelah mengetahui nilai SSW maka selanjutnya dilakukan perhitungan *Sum of square between-cluster (SSB)*.

$$SSB_{1,2} = \sqrt{(0.333 - 0.083)^2 + (0 - 0)^2 + (0.5 - 1)^2} = 0,559$$

$$SSB_{1,3} = \sqrt{(0.333 - 0.117)^2 + (0 - 0)^2 + (0.5 - 1)^2} = 0,544$$

$$SSB_{2,3} = \sqrt{(0.083 - 0.117)^2 + (0 - 0)^2 + (1 - 1)^2} = 0,0333$$



### 3. Perhitungan Rasio dengan persamaan (2.5)

Setelah nilai SSW dan nilai SSB telah dihitung dan mendapatkan hasilnya, maka selanjutnya adalah mencari nilai rasio antar cluster dengan perhitungan sebagai berikut:

$$R_1 = \frac{0,478}{0,559 + 0,544 + 0,0333} = \frac{0,478}{1,137} = 0,4206$$

$$R_2 = \frac{0,04}{0,559 + 0,544 + 0,0333} = \frac{0,04}{1,137} = 0,0359$$

$$R_3 = \frac{0,22}{0,559 + 0,544 + 0,0333} = \frac{0,22}{1,137} = 0,1934$$

### 4. Menghitung nilai *Davies Boulding Index* (DBI) dengan persamaan (2.6)

Setelah mengetahui rasio antar *cluster* kemudian hitunglah nilai DBI sebagai berikut:

$$DBI = \frac{R_1 + R_2 + R_3}{K} = \frac{0,4206 + 0,0359 + 0,1934}{3} = 0,216$$

### 3.2 Pseudocode

1. Mulai
2. Baca data.
3. Lakukan prapemrosesan data.
  1. Ubah nilai pada kolom "Bunga" dari persentase menjadi nilai numerik.
  2. Enkode nilai di kolom "Bunga" menggunakan LabelEncoder.
  3. Memetakan nilai pada kolom "Status Pinjaman" ke nilai numerik.
  4. Normalisasi data menggunakan MinMaxScaler.
4. Lakukan pengelompokan K-Medoids:
  1. Tentukan jumlah cluster ( $n\_clusters$ ).
  2. Inisialisasi objek KMedoids dengan jumlah cluster dan parameter lainnya.
  3. Lakukan dua iterasi: Pilih medoid secara acak.
  4. Hitung jarak terpendek antara data dan medoid menggunakan  $pairwise\_distances\_argmin\_min$ .
  5. Hitung total deviasi (inersia) untuk setiap iterasi.
  6. Hitung pusat cluster dan atribut lainnya.
5. Hitung SSW (Jumlah Kuadrat Jarak Dalam Cluster):
  1. Hitung jarak antara setiap titik data dan pusat cluster.
6. Hitung SSB (Jumlah Kuadrat Jarak Antar Cluster):
  1. Hitung total pusat semua data.
  2. Hitung jarak antara pusat cluster dan pusat total.
  3. Kalikan dengan jumlah titik data dalam cluster.
  4. Jumlahkan nilai-nilai ini.
7. Hitung DBI (Davies-Bouldin Indeks):
  1. Hitung perbandingan SSW dan SSB.
  2. Hitung jarak antara pusat cluster dengan pusat cluster lainnya.
  3. Hitung perbandingan antara SSW dan SSB dikalikan jarak antar pusat cluster.
  4. Ambil nilai maksimum setiap Rij.
  5. Jumlahkan nilai maksimumnya.
  6. Bagilah jumlah nilai maksimum dengan jumlah cluster untuk mendapatkan DBI.
8. Selesai

### 3.3 Perancangan Sistem

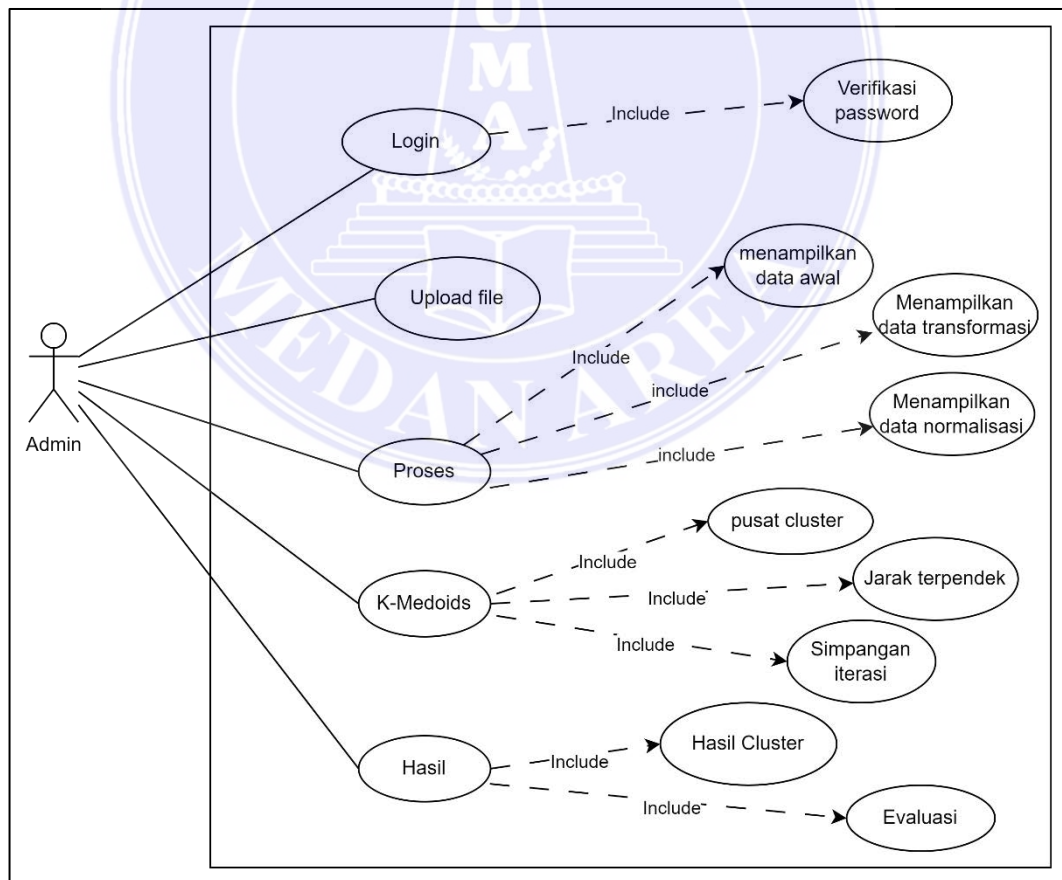
Perancangan sistem adalah proses merancang untuk menciptakan sebuah sistem yang memenuhi kebutuhan atau tujuan tertentu.

#### 3.3.1 Perancangan UML

Perancangan UML adalah proses merancang sebuah sistem menggunakan UML. Adapun perancangan UML yang digunakan dalam penelitian ini adalah *use case diagram* dan *activity diagram*.

##### a. Use Case Diagram

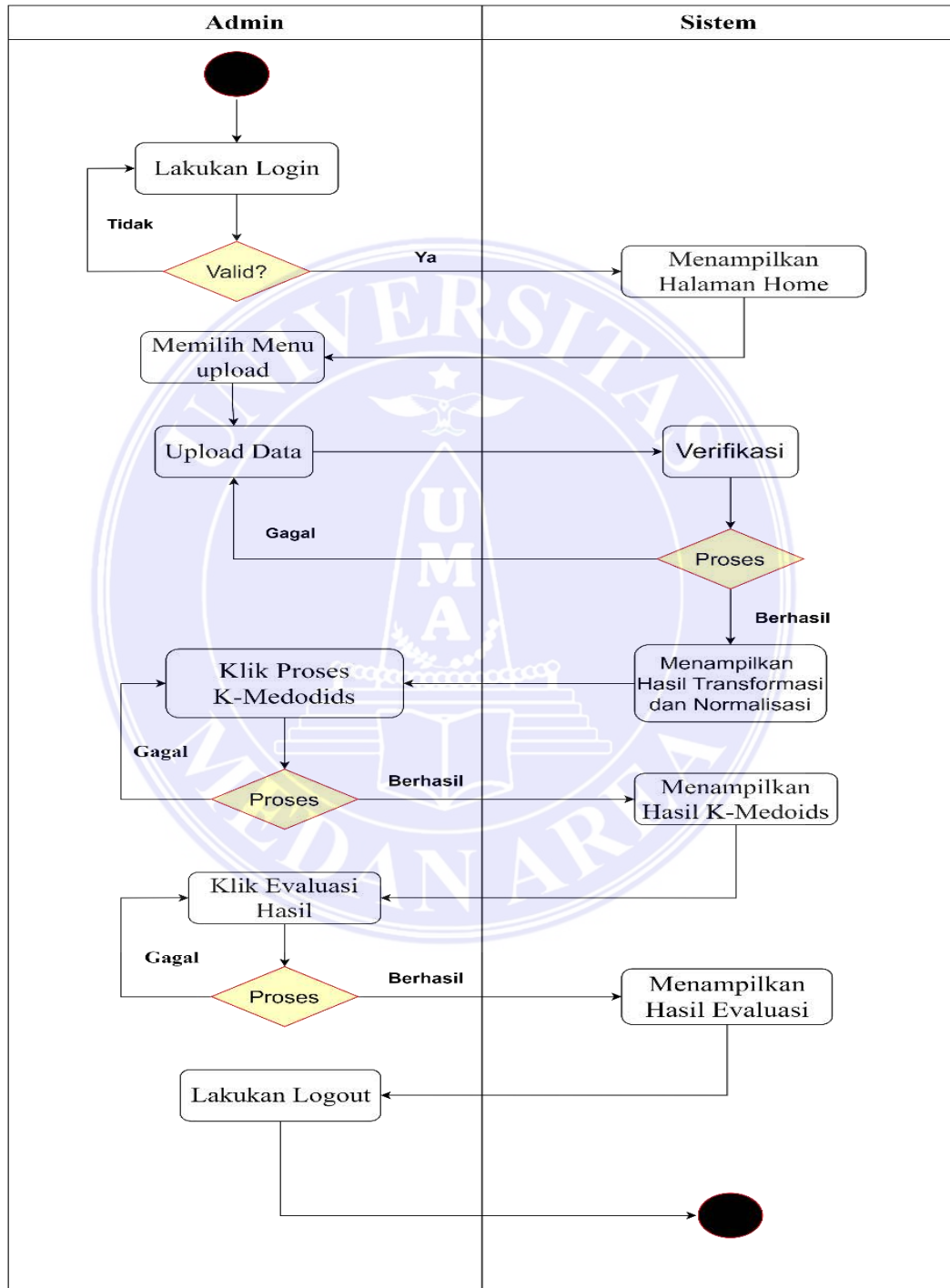
Berikut adalah pemodelan *use case diagram* yang digunakan untuk perancangan aplikasi penerapan *data mining* dalam pengelompokan nasabah penerima *reward*.



Gambar 3. 2 Use Case Diagram

b. *Activity Diagram*

Berikut adalah pemodelan *Activity Diagram K-Medoids* untuk pengelompokan penerima *reward* pada PT DGJ.



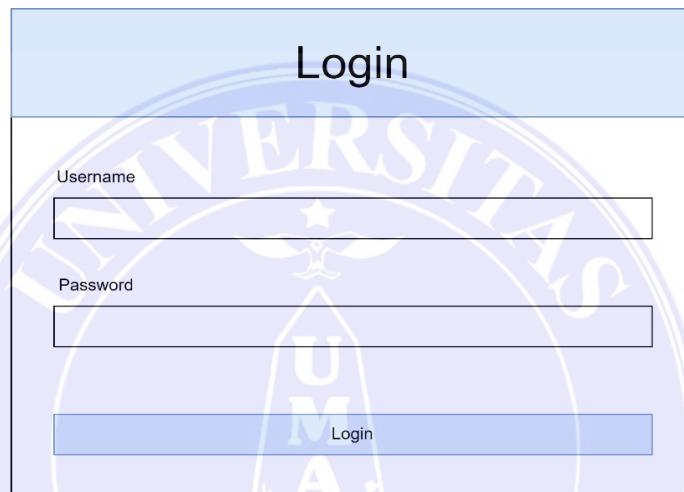
Gambar 3. 3 *Activity Diagram*

### 3.2.2 Perancangan *User Interface*

Perancangan User Interface (UI) adalah proses merancang antarmuka yang digunakan oleh pengguna untuk berinteraksi dengan sebuah sistem.

#### 1. Rancangan *Form Login*

*Form login* sebagai autentikasi pengguna sebelum mereka dapat mengakses sistem.



The image shows a login form with a light blue header containing the word "Login". Below the header, there are three input fields: "Username", "Password", and a "Login" button. The form is overlaid on a large, faint watermark of the Universitas Medan Area logo.

Gambar 3. 4 Rancangan Form Login

#### 2. Rancangan *Form Home*

*Form Home* adalah halaman pertama yang dilihat pengguna saat mengakses sistem dan menampilkan menu-menu pada sistem.



The image shows a home page layout. On the left is a sidebar menu with the following items: "K-Medoids", "Home", "Upload file", and "Logout". The main content area features a search bar at the top, a logo for "dotri. GADAI" in the center, and a text box below it that reads: "Pengelompokan Nasabah Penerima Reward pada PT Dotri Gadaai Jaya menggunakan Algoritma K-medoids".

Gambar 3. 5 Rancangan Form Home

### 3. Rancangan *Form Upload Data*

Form yang digunakan untuk memasukan atau mengupload data nasabah yang terdiri dari nomor urut, nama nasabah, pinjaman, bunga dan status pinjaman. Format data yang digunakan adalah .xlsx. Setelah data kita upload, selanjutnya sistem akan otomatis memproses data, proses data dimulai dari transformasi data, normalisasi data, proses *k-medoids* dan proses evaluasi.

The image shows a web application interface for K-Medoids. On the left is a vertical navigation menu with the following items: 'K-Medoids', 'Home', 'Upload file', and 'Logout'. The 'K-Medoids' item is currently selected, indicated by a small square icon. The main content area displays a form titled 'Upload dataset'. This form contains a file selection input field with two options: 'Chosee file' and 'No Chosee file'. Below the input field is a black button labeled 'Upload data'. The entire interface is overlaid on a large, faint watermark of the Universitas Medan Area logo.

Gambar 3. 6 Rancangan Form Upload Data

#### 4. Rancangan *Form* Hasil Proses Data Awal

*Form* ini menampilkan data awal yang kita upload, hasil transformasi data dan hasil normalisasi data.

The figure shows three sequential screenshots of a web application interface for a K-Medoids algorithm. Each screenshot has a sidebar on the left with the following menu items: 'K-Medoids', 'Home', 'Data', 'Proses K-medoids', and 'Evaluasi dan hasil'. The main content area of each screenshot displays a table of data.

**Screenshot 1: Tampilan data awal**

Nama nasabah	jumlah pinjaman	bunga	status pinjaman
X	X	X	X
X	X	X	X
X	X	X	X
X	X	X	X
X	X	X	X
X	X	X	X
X	X	X	X
X	X	X	X

**Screenshot 2: Tampilan transform**

Nama nasabah	jumlah pinjaman	bunga	status pinjaman
X	X	X	X
X	X	X	X
X	X	X	X
X	X	X	X
X	X	X	X
X	X	X	X
X	X	X	X
X	X	X	X

**Screenshot 3: Tampilan Normalisasi**

Nama nasabah	jumlah pinjaman	bunga	status pinjaman
X	X	X	X
X	X	X	X
X	X	X	X
X	X	X	X
X	X	X	X
X	X	X	X
X	X	X	X
X	X	X	X

Gambar 3. 7 Rancangan Form Hasil Proses Data Awal

5. Rancangan *Form* Hasil proses *K-Medoids*

*Form* yang menampilkan hasil proses pengelompokan yang dimulai dari penentuan pusat cluster, perhitungan jarak data dengan pusat cluster, penentuan pusat cluster baru.

The form is titled 'K-Medoids' and has a sidebar menu with 'Home', 'Data', 'Proses K-medoids', and 'Evaluasi dan hasil'. The main content area is titled 'Data dengan informasi jarak terpendek dan clustering:'.

**Iterasi 1**  
Menentukan Medoid 1 secara acak objek pada masing-masing cluster untuk melakukan iterasi ke-1:

No	jumlah pinjaman	bunga	status pinjaman	jarak terpendek	cluster
X	X	X	X	X	X
X	X	X	X	X	X
X	X	X	X	X	X
X	X	X	X	X	X
X	X	X	X	X	X
X	X	X	X	X	X
X	X	X	X	X	X
X	X	X	X	X	X

Total Jarak Terpendek dari Masing-masing Iterasi:  
 • :XXXX  
 • :XXXXX  
 Selisih Total Jarak Terpendek dari Kedua Iterasi: xxxx

**Iterasi 1**      **Iterasi 2**

Menentukan Medoid 1 secara acak objek pada masing-masing cluster untuk melakukan iterasi ke-1:      Menentukan Medoid 1 secara acak objek pada masing-masing cluster untuk melakukan iterasi ke-2:

data	x1	x2	x3	cluster
X	X	X	X	X
X	X	X	X	X
X	X	X	X	X

data	x1	x2	x3	cluster
X	X	X	X	X
X	X	X	X	X
X	X	X	X	X

Gambar 3. 8 Rancangan Form Hasil proses *K-Medoids*



6. Rancangan Form Hasil *Cluster* dan Evaluasi

*Form* yang menampilkan hasil pengelompokan dan hasil evaluasi pengelompokan. Pada form ini juga disediakan untuk download hasil pengelompokannya dalam format pdf.

No	Nama Nasabah	Clustering	Keterangan
X	X	X	X
X	X	X	X
X	X	X	X
X	X	X	X
X	X	X	X
X	X	X	X
X	X	X	X

Gambar 3. 9 Rancangan Form Hasil Cluster dan Evaluasi

## BAB V

### KESIMPULAN DAN SARAN

#### 5.1 Kesimpulan

Berdasarkan hasil penelitian pada skripsi ini, maka diperoleh kesimpulan sebagai berikut:

1. *K-medoids clustering* dapat digunakan untuk mengelompokan nasabah penerima *reward* pada PT DGJ.
2. Berdasarkan hasil pengelompokan, jumlah nasabah yang mendapatkan *reward* 30% sebanyak 314 nasabah, *reward* 20% sebanyak 540 nasabah dan *reward* 10% sebanyak 231 nasabah, dari total 1.085 nasabah.
3. Penelitian ini menggunakan DBI sebagai metode untuk mengevaluasi hasil *cluster*. Adapun nilai DBI yang diperoleh adalah 0.368812. Hal ini membuktikan bahwa nilai tersebut mendekati angka 0 atau terhitung cukup kecil. Sehingga dapat dikatakan bahwa *cluster* yang dihasilkan cukup baik.

#### 5.2 Saran

Pada penelitian ini, yang digunakan adalah algoritma *k-medoids*. Evaluasi kualitas klaster dilakukan menggunakan DBI, dan hasilnya menunjukkan nilai DBI sebesar 0.368812 untuk data sebanyak 1.085 dengan 3 variabel. Penulis menyarankan agar penelitian selanjutnya menggunakan jumlah data dan variabel yang lebih banyak agar pengelompokan menjadi lebih optimal dan mendapatkan hasil yang lebih baik dari penelitian ini.

## DAFTAR PUSTAKA

- Ahmad, I., Samsugi, S., & Irawan, Y. (2022). Implementasi Data Mining Sebagai Pengolahan Data. *Jurnal Teknoinfo*, 16(1), 46. <http://portaldata.org/index.php/portaldata/article/view/107>
- Andini, A. D., & Arifin, T. (2020). Implementasi Algoritma K-Medoids Untuk Klasterisasi Data Penyakit Pasien Di Rsud Kota Bandung. *Jurnal Responsif: Riset Sains dan Informatika*, 2(2), 128–138. <https://doi.org/10.51977/jti.v2i2.247>
- Andriyani, M., & Ardianto, R. (2020). Pengaruh Kualitas Layanan dan Kualitas Produk Terhadap Kepuasan Nasabah Bank. *EKOMABIS: Jurnal Ekonomi Manajemen Bisnis*, 1(02), 133–140. <https://doi.org/10.37366/ekomabis.v1i02.73>
- Aulanda, L., Windarto, A. P., & Okprana, H. (2021). Pengelompokan Pembiayaan Nasabah Klaim Asuransi Pengguna Kendaraan Bermotor dengan Metode K-Medoids. *TIN: Terapan Informatika Nusantara*, 2(4), 263–270.
- Aulia, S. (2021). Klasterisasi Pola Penjualan Pestisida Menggunakan Metode K-Means Clustering (Studi Kasus Di Toko Juanda Tani Kecamatan Hutabayu Raja). *Djtechno: Jurnal Teknologi Informasi*, 1(1), 1–5. <https://doi.org/10.46576/djtechno.v1i1.964>
- Bahauddin, A., Fatmawati, A., & Permata Sari, F. (2021). Analisis Clustering Provinsi Di Indonesia Berdasarkan Tingkat Kemiskinan Menggunakan Algoritma K-Means. *Jurnal Manajemen Informatika dan Sistem Informasi*, 4(1), 1–8. <https://doi.org/10.36595/misi.v4i1.216>
- Berutu, I. (2021). RESOLUSI: Rekayasa Teknik Informatika dan Informasi Penerapan Metode C5.0 Untuk Pengelompokan Potensi Nasabah PT.Pegadaian Berdasarkan Pola Pembayaran Angsuran. *Media Online*, 1(4), 232–240. <https://djournal.com/resolusi>
- Furqani, N. El, & Muliono, R. (2021). Web-Based Library Information System Design at SDN 056004 Basilam. *Journal of Research Computer Science*, 1(1), 14–26. <http://journal.station-it.org/index.php/jrcs>
- Hardiyanti, D. (2021). Sistem Informasi Sekolah Berbasis Web pada Sekolah Dasar Negeri (SDN) Seriti. *Indonesian Journal Of Education And Humanity*, 1(3), 156–168. <http://ijoejm.rcipublisher.org/index.php/ijoejm/article/view/28>
- Herviany, M., Putri Delima, S., Nurhidayah, T., & Kasini, K. (2021). Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokan

- Daerah Rawan Tanah Longsor Pada Provinsi Jawa Barat. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 1(1), 34–40. <https://doi.org/10.57152/malcom.v1i1.60>
- Khesya, N. (2021). Mengenal Flowchart dan Pseudocode Dalam Algoritma dan Pemrograman. *Preprints*, 1, 1–15. <https://osf.io/dq45ef>
- Kiu, V., Seng, K., & Harman, R. (2021). Penerapan Algoritma C4.5 Untuk Menentukan Reward Karyawan Pada Pt Indoland Batam. *Jurnal Comasie*, 04(04).
- Mirantika, N., Syamfithriani, T. S., & Trisudarmo, R. (2023). Implementasi Algoritma K-Medoids Clustering Untuk Menentukan Segmentasi Pelanggan. *Jurnal Nuansa Informatika*, 17, 2614–5405. <https://journal.uniku.ac.id/index.php/ilkom>
- Nabila, S., Putri, A. R., Hafizhah, A., Rahmah, F. H., & Muslikhah, R. (2021). Pemodelan Diagram UML Pada Perancangan Sistem Aplikasi Konsultasi Hewan Peliharaan Berbasis Android (Studi Kasus: Alopel). *Jurnal Ilmu Komputer dan Bisnis*, 12(2), 130–139. <https://doi.org/10.47927/jikb.v12i2.150>
- Nabila, Z., Rahman Isnain, A., & Abidin, Z. (2021). Analisis Data Mining Untuk Clustering Kasus Covid-19 Di Provinsi Lampung Dengan Algoritma K-Means. *Jurnal Teknologi dan Sistem Informasi (JTSI)*, 2(2), 100. <http://jim.teknokrat.ac.id/index.php/JTSI>
- Ningrum, H., Irawan, E., & Lubis, M. R. (2021). Implementasi Metode K-Medoids Clustering Dalam Pengelompokan Data Penyakit Alergi Pada Anak. *Jurasik (Jurnal Riset Sistem Informasi dan Teknik Informatika)*, 6(1), 130. <https://doi.org/10.30645/jurasik.v6i1.277>
- Nurdin, N., Musyawarah, I., Nurfitriani, N., & Jalil, A. (2020). Pengaruh Pelayanan Mobile Banking Terhadap Kepuasan Nasabah (Studi Pada Mahasiswa Perbankan Syariah IAIN Palu). *Jurnal Ilmu Perbankan dan Keuangan Syariah*, 2(1), 87–104. <https://doi.org/10.24239/jipsya.v2i1.24.87-104>
- Paembonan, S., & Abduh, H. (2021). Penerapan Metode Silhouette Coefficient untuk Evaluasi Clustering Obat. *PENA TEKNIK: Jurnal Ilmiah Ilmu-Ilmu Teknik*, 6(2), 48. [https://doi.org/10.51557/pt\\_jiit.v6i2.659](https://doi.org/10.51557/pt_jiit.v6i2.659)
- Pakaya, R., Tapate, A. R., & Suleman, S. (2020). Perancangan Aplikasi Penjualan Hewan Ternak Untuk Qurban Dan Aqiqah Dengan Metode Unified Modeling Language (Uml). *Jurnal Technopreneur (JTech)*, 8(1), 31–40. <https://doi.org/10.30869/jtech.v8i1.531>

- Puspita, I., Hayati, M. N., & Nohe, D. A. (2023). Pengelompokan Puskesmas Berdasarkan Kasus Balita Stunting di Kabupaten Paser Menggunakan Metode K-Medoids. *Eksponensial*, 14(1), 1. <https://doi.org/10.30872/eksponensial.v14i1.1089>
- Rahmah, E., Haerani, E., Nazir, A., & Ramadhani, S. (2022). Penerapan Algoritma K-Medoids Clustering Untuk Menentukan Strategi Promosi Pada Data Mahasiswa (Studi Kasus: Stikes Perintis Padang). *Jurnal Nasional Komputasi dan Teknologi Informasi (JNKTI)*, 5(3), 556–564. <https://doi.org/10.32672/jnkti.v5i3.4355>
- Sabella, R. A. N., Muftifiandi, & Lemiyana. (2022). Analisis Kompetensi Komunikasi Customer Service Dalam Upaya Meningkatkan Pelayanan Prima Pada Bank Syariah Indonesia Kcp Tulang Bawang Barat. *Jurnal Manajemen, Akuntansi dan Ekonomi*, 1(3), 9–26. <https://jurnal.erapublikasi.id/index.php/JMAE/index>
- Saleh, A., & Mardiana, A. (2021). Pemberian Reward Terhadap Peningkatan Motivasi Kerja Karyawan Dalam Perspektif Islam. *MUTAWAZIN (Jurnal Ekonomi Syariah)*, 2(1), 1–14. <https://doi.org/10.54045/mutawazin.v2i1.233>
- Shinta Sibatuara, J., Khairina, N., Sembiring, Z., & Muliono, R. (2022). Clasterization of Student's Level of Understanding in Online Learning System with K-Means Clustering. *Jrcs*, 2(1), 2770–1800. <http://journal.station-it.org/index.php/jrcs>
- Simanjuntak, D. S. M., Gunawan, I., Sumarno, S., Poningsih, P., & Sari, I. P. (2023). Penerapan Algoritma K-Medoids Untuk Pengelompokan Pengangguran Umur 25 tahun Keatas Di Sumatera Utara. *Jurnal Krisnadana*, 2(2). <https://doi.org/10.58982/krisnadana.v2i2.264>
- Situmorang, S., Hondro, R. K., & Sumiati, S. (2023). Penerapan Algoritma K-Medoids Pada Pengelompokan Penerima Dana Bantuan Sosial Kelurahan Medan Sinembah. *KOMIK (Konferensi Nasional Teknologi Informasi dan Komputer)*, 6(1), 738–747. <https://doi.org/10.30865/komik.v6i1.5721>
- Sopyan, Y., Lesmana, A. D., & Juliane, C. (2022). Analisis Algoritma K-Means dan Davies Bouldin Index dalam Mencari Cluster Terbaik Kasus Perceraian di Kabupaten Kuningan. *Building of Informatics, Technology and Science (BITS)*, 4(3), 1464–1470. <https://doi.org/10.47065/bits.v4i3.2697>
- Sudarsono, B. G., Leo, M. I., Santoso, A., & Hendrawan, F. (2021). Analisis Data Mining Data Netflix Menggunakan Aplikasi Rapid Miner. *JBASE - Journal of Business and Audit Information Systems*, 4(1), 13–21. <https://doi.org/10.30813/jbase.v4i1.2729>

- Umagapi, I. T., Umaternate, B., Hazriani, H., & ... (2023). Uji Kinerja K-Means Clustering Menggunakan Davies-Bouldin Index Pada Pengelompokan Data Prestasi Siswa. *Prosiding ...*, 303–308. <http://www.seminar.iaii.or.id/index.php/SISFOTEK/article/view/411>
- Wibowo, A. R., & Jananto, A. (2020). Implementasi Data Mining Metode Asosiasi Algoritma FP-Growth Pada Perusahaan Ritel. *Inspiration: Jurnal Teknologi Informasi dan Komunikasi*, 10(2), 200. <https://doi.org/10.35585/inspir.v10i2.2585>
- Widyastuti, R. (2022). Penerapan Sistem Informasi Akademik Di Smk Yaspem Jakarta. *PROSISKO: Jurnal Pengembangan Riset dan Observasi Sistem Komputer*, 9(2), 9–24. <https://doi.org/10.30656/prosisko.v9i2.4938>
- Zai, I. W., & Aripin, S. (2021). Penerapan Metode Profile Matching Untuk Menentukan Anggota Polri Yang Mendapatkan Reward Pada Polda Sumut. *Seminar Nasional Sains dan Teknologi ...*, 338–344. <http://prosiding.seminar-id.com/index.php/sensasi/article/view/672>
- Zeilani, M. S. (2022). Implementasi Algoritma K-Medoids Dalam Mengklasifikasi Barang Layak Lelang. *BEES: Bulletin of Electrical and Electronics Engineering*, 2(3), 106–111. <https://doi.org/10.47065/bees.v2i3.1131>

## LAMPIRAN

### 1. Source Code

```
!pip install scikit-learn-extra

import pandas as pd
from sklearn.preprocessing import LabelEncoder
from sklearn.preprocessing import MinMaxScaler
from sklearn_extra.cluster import KMedoids
from sklearn.metrics import pairwise_distances_argmin_min
from scipy.spatial.distance import cdist
import numpy as np
from sklearn.metrics import davies_bouldin_score
from sklearn.metrics import pairwise_distances
from sklearn.metrics import davies_bouldin_score
import numpy as np
from google.colab import drive
drive.mount('/content/drive/')

df = pd.read_excel('/content/drive/MyDrive/SKRIPSI MENIATI
ZEBUA/DATA NASABAH/K-MEDOIDSs.xlsx')
df = df.drop('No',axis=1)
df

bunga_encoder = LabelEncoder()

df['Bunga'] = df['Bunga'].str.replace('%', '').astype(float)

df['Bunga'] = bunga_encoder.fit_transform(df['Bunga']) + 1

mapping = {'Tidak Lancar': 1, 'Kurang Lancar': 2, 'Lancar': 3}

df['Status Pinjaman'] = df['Status Pinjaman'].map(mapping)
df

nama_nasabah = df['Nama Nasabah']
df = df.drop(columns=['Nama Nasabah'])

scaler = MinMaxScaler()
```

```

data = pd.DataFrame(scaler.fit_transform(df), columns=df.columns)

df_normalized = pd.concat([nama_nasabah, data], axis=1)
df_normalized

n_clusters = 3
kmedoids = KMedoids(n_clusters=n_clusters,
random_state=42,max_iter=2)
kmedoid = KMedoids(n_clusters=n_clusters, random_state=42,
max_iter=2)
total_simpangan_iterasi_list = []
medoids_list = []
clusterings = []
distances_list = []
centers_list = []
total_distance_difference = 0
medoid_indices_list = []

for _ in range(2):
    medoids_indices = np.random.choice(len(data), n_clusters,
replace=False)
    medoids = data.iloc[medoids_indices]
    cluster_labels, distances = pairwise_distances_argmin_min(data,
medoids)

    total_simpangan_iterasi = 0
    for i, medoid_index in enumerate(medoids_indices):
        total_simpangan_iterasi +=
np.sum(np.linalg.norm(data[cluster_labels == i] -
data.iloc[medoid_index], axis=1))

    total_simpangan_iterasi_list.append(total_simpangan_iterasi)
    medoids_list.append(medoids)
    clusterings.append(cluster_labels)
    distances_list.append(distances)
    medoid_indices_list.append(medoids_indices)

cluster_centers = []
for medoid_index in range(len(medoids)):
    cluster_data = data[cluster_labels == medoid_index]
    cluster_center = cluster_data.mean()

```



```

        cluster_center['Jumlah Data'] = cluster_data.shape[0]
        cluster_centers.append(cluster_center)
    centers_list.append(cluster_centers)

for i, (medoids, cluster_labels, distances, centers, medoid_indices) in
enumerate(zip(medoids_list, clusterings, distances_list, centers_list,
medoid_indices_list), start=1):
    print(f'Iterasi {i}:')
    centers_df = pd.DataFrame(centers, columns=['Jumlah Pinjaman',
'Bunga', 'Status Pinjaman'])
    centers_df.columns = ['X1', 'X2', 'X3']
    centers_df['Cluster'] = [c + str(i+1) for i in range(n_clusters)]
    centers_df.insert(0, 'Data Ke', medoid_indices + 1)
    centers_df = centers_df.set_index('Cluster')

    print(f'Menentukan Medoid {i} secara acak objek pada masing masing
cluster untuk melakukan iterasi ke- {i}:')
    print(centers_df)
    print()

    data_with_info = data.copy()
    data_with_info['Jarak Terpendek'] = distances
    data_with_info['Clustering'] = [c + str(label+1) for label in
cluster_labels]
    print("Data dengan informasi jarak terpendek dan clustering:")
    print(data_with_info)
    print()

total_distance_difference = total_simpangan_iterasi_list[-1] -
total_simpangan_iterasi_list[0]

print("Total Jarak Terpendek dari Masing-masing Iterasi:")
for i, total in enumerate(total_simpangan_iterasi_list, start=1):
    print(f'Iterasi {i}: {total}')

print(f'Selisih Total Jarak Terpendek dari Kedua Iterasi:
{total_distance_difference}')

data_iterasi_pertama = data.copy()
data_iterasi_pertama['Jarak Terpendek'] = distances_list[0]

```

```
data_iterasi_pertama['Clustering'] = ['c' + str(label+1) for label in
clusterings[0]]
data_iterasi_pertama['Clustering'] =
data_iterasi_pertama['Clustering'].replace({'c1': '1', 'c2': '2', 'c3': '3'})
```

```
keterangan_persentase = {'1': 'Reward 30%', '2': 'Reward 20%', '3':
'Reward 10%'}
```

```
data_iterasi_pertama['Keterangan'] =
data_iterasi_pertama['Clustering'].map(keterangan_persentase)
hasil = pd.concat([nama_nasabah, data_iterasi_pertama], axis=1)
```

```
hasil = hasil[['Nama Nasabah', 'Clustering', 'Keterangan']]
hasil
```

```
# kmedoid.fit(data)
```

```
def calculate_SSW(data, labels, centers):
SSW = 0
for i, center in enumerate(centers):
cluster_data = data[labels == i]
cluster_distance = np.linalg.norm(cluster_data - center, axis=1)
SSW += np.sum(cluster_distance)
return SSW
```

```
def calculate_SSB(data, labels, centers):
SSB = 0
total_center = np.mean(data, axis=0)
for i, center in enumerate(centers):
cluster_data = data[labels == i]
num_points = len(cluster_data)
SSB += num_points * np.linalg.norm(center - total_center)
return SSB
```

```
def calculate_DBI(data, labels, centers):
k = len(np.unique(labels))
SSW = calculate_SSW(data, labels, centers)
SSB = calculate_SSB(data, labels, centers)
```

```
DBI = 0
for i in range(k):
    max_Rij = 0
    for j in range(k):
        if i != j:
            Rij = (SSW / SSB) * np.linalg.norm(centers[i] - centers[j])
            if Rij > max_Rij:
                max_Rij = Rij
    DBI += max_Rij
DBI /= k
return DBI
```

```
medoids_first_iter = medoids_list[0]
cluster_labels_first_iter = clusterings[0]

medoids_first_iter = medoids_first_iter.values

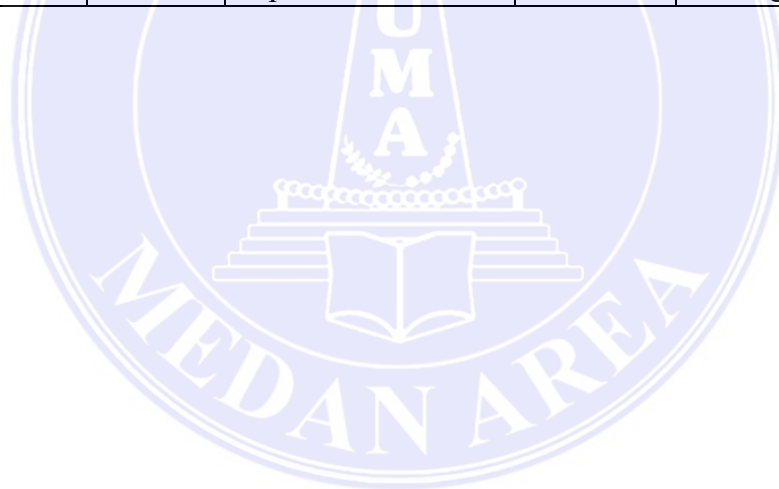
# SSB = calculate_SSB(centroids)

SSB = calculate_SSB(data, cluster_labels_first_iter, medoids_first_iter)
SSW = calculate_SSW(data, cluster_labels_first_iter, medoids_first_iter)
DBI = calculate_DBI(data, cluster_labels_first_iter, medoids_first_iter)
# from sklearn.metrics import davies_bouldin_score

# DBI = davies_bouldin_score(data, kmedoid.labels_)
print("SSW:", SSW)
print("SSB:", SSB)
print("DBI:", DBI)
```

## 2. Sampel Data

No	Nama Nasabah	Jumlah Pinjaman	Bunga	Status Pinjaman
1	N1	Rp 70.000.000	5 %	Kurang Lancar
2	N2	Rp 50.100.000	5 %	Lancar
3	N3	Rp 55.000.000	5 %	Lancar
4	N4	Rp 50.000.000	10 %	Lancar
5	N5	Rp 61.100.000	5 %	Lancar
6	N6	Rp 57.000.000	5 %	Lancar
7	N7	Rp 65.600.000	5 %	Lancar
8	N8	Rp 56.900.000	5 %	Lancar
9	N9	Rp 92.500.000	5 %	Lancar
10	N10	Rp 51.500.000	10%	Kurang Lancar
11	N11	Rp 54.700.000	5 %	Lancar
12	N12	Rp 70.000.000	5 %	Tidak Lancar
13	N13	Rp 56.000.000	5 %	Lancar
14	N14	Rp 61.000.000	5 %	Kurang Lancar





# UNIVERSITAS MEDAN AREA FAKULTAS TEKNIK

Kampus I : Jalan Kolam Nomor 1 Medan Estate/Jalan PBSI Nomor 132 (061) 7366878, 7360168, 7364348, 7366781, Fax (061) 7366998 Medan 20223  
Kampus II : Jalan Setiabudi Nomor 79 / Jalan Sei Serayu Nomor 70 A, (061) 8225602, Fax (061) 8226331 Medan 20122  
Website [www.teknik.uma.ac.id](http://www.teknik.uma.ac.id) E-mail [univ\\_medanarea@uma.ac.id](mailto:univ_medanarea@uma.ac.id)

Nomor : FT.6/01.10/III/2024

18 Maret 2024

Lamp : -  
Hal :

**: Penelitian Dan Pengambilan Data Tugas Akhir**

Yth. Pimpinan PT Dotri Gadai Jaya  
Jln. William Iskandar No. 91 A  
Di  
Medan

Dengan hormat,

Kami mohon kesediaan Bapak/Ibu berkenan untuk memberikan izin dan kesempatan kepada mahasiswa kami tersebut dibawah ini :

NO	NAMA	NPM	PRODI
1	Meniati Zebua	208160007	Teknik Informatika

Untuk melaksanakan Penelitian dan Pengambilan Data Tugas Akhir pada perusahaan/Instansi yang Bapak/Ibu Pimpin.

Perlu kami jelaskan bahwa Pengambilan Data tersebut adalah semata-mata untuk tujuan ilmiah dan Skripsi yang merupakan salah satu syarat bagi mahasiswa tersebut untuk mengikuti ujian sarjana pada Fakultas Teknik Universitas Medan Area dan tidak untuk dipublikasikan, dengan judul penelitian :

**Penerapan Data Mining menggunakan Algoritma K-Medoids dalam Pengelompokan Nasabah Penerima Reward pada PT. Dotri Gadai Jaya**

Atas perhatian dan kerja sama yang baik diucapkan terima kasih.

Dekan,  
  
Dr. Eng Supriatno, ST, MT

**Tembusan :**  
1. Ka. BPMPP  
2. Mahasiswa  
3. File



## PT. DOTRI GADAI JAYA

Jl. Willem Iskandar No. 91 A  
Hp. 0895612013404

Medan, 22 April 2024

Nomor : 53/DG/IV/2024  
Lampiran : -  
Perihal : **Telah Melakukan Penelitian dan Pengambilan Data**

Kepada Yth.  
Dekan Fakultas Teknik  
Universitas Medan Area  
di-

Medan, Sumatera Utara

Sehubungan dengan surat saudara Nomor: 162/FT.6/01.10/III/2024 perihal Permohonan Penelitian dan Pengambilan Data di PT.Dotri Gadai Jaya, maka dengan ini menerangkan nama mahasiswa/i dibawah ini sebagai berikut:

Nama : Meniati Zebua  
Npm : 208160007  
Fakultas : Teknik  
Prodi : Teknik Informatika

Bahwa Mahasiswa/i tersebut benar telah selesai melaksanakan penelitian dan pengambilan data di PT.Dotri Gadai Jaya untuk mendukung Tugas Akhir (Skripsi) dengan judul "Penerapan *Data Mining* Menggunakan Algoritma *K-Medoids* dalam Pengelompokan Nasabah Penerima *Reward* pada PT Dotri Gadai Jaya".

Demikian disampaikan, untuk dapat dipergunakan sebagaimana mestinya.

HRD PT.Dotri Gadai Jaya

(RAHMADINA)

