

Clustering Profil Pengunjung Perpustakaan (Studi Kasus Perpustakaan BP Batam)

Fauziah Mahmuda^{1*}, Maya Armys Roma Sitorus^{2*}, Hilda Widyastuti^{3*}, Dwi Ely Kurniawan^{4*}

* Teknik Informatika, Politeknik Negeri Batam
fauziahmahmuda.usi@gmail.com¹

Article Info

Article history:

Received 07-03-2017
Revised 09-04-2017
Accepted 06-05-2017

Keyword:

Data Mining, Clustering,
K-Means Algorithm,
Libraries.

ABSTRACT

Business Entity library Batam (BP Batam) is a public library located in Batam city with thw number of visitors. Every visitor who comes to do the charging guest book manually by writing system. It causes a buildup of data which are not organized. Data mining is one of the analytical tools that can be used to address the backlog of data. The method of Clustering with the K-Means Algorithm used in analyzing the data library visitors BP Batam. Library visitors using the data processing method of Elbow to get the best number of clusters K i.e., $K = 3$, and by using the center point (centroid) initial i.e, $P1 = (4,1)$, $P2 = (2,4)$, $P3 = (4,2)$. The purpose of this research is to apply the algorithm for K-Means clustering in the data library visitors (case study library BP Batam). K-Means clustering results obtained from 1556 dataset data library visitors are grouped into three clusters, Clusters 1 is dominated by a college student and visitor located at Batam Center, Cluster 2 is dominated by a college student and visitor located at Bengkong, Cluster 3 is dominated by public and visitor status in Batam Center.

Copyright © 2017 Journal of Applied Informatics and Computing.
All rights reserved.

I. PENDAHULUAN

Pendidikan merupakan sesuatu yang sangat penting hingga saat ini, dan sekarang telah menjadi kebutuhan pokok yang harus dimiliki setiap orang agar bisa menghadapi dan menjawab tantangan kehidupan. Pendidikan harus diberikan kepada setiap orang sejak dini. Waktu yang produktif untuk pemberian pendidikan yaitu kepada pelajar yang sedang duduk di bangku Sekolah Dasar (SD), Sekolah Menengah Pertama (SMP) dan Sekolah Menengah Atas (SMA). Banyak cara yang dapat digunakan untuk mencapai pendidikan, salah satunya adalah melalui perpustakaan. Karena dari perpustakaan berbagai sumber dapat kita peroleh dan juga perpustakaan merupakan upaya peningkatan efisiensi dan efektifitas proses belajar-mengajar.

Perpustakaan sendiri memiliki beberapa jenis, salah satunya adalah perpustakaan daerah yang berkedudukan di tiap provinsi di Indonesia. Batam juga memiliki perpustakaan daerah yaitu, Kantor Perpustakaan Umum dan Arsip Kota Batam dan Perpustakaan Badan Pengusahaan Batam. Tujuan perpustakaan adalah untuk menyediakan

fasilitas dan sumber informasi dan menjadi pusat pembelajaran [14].

Banyaknya pengunjung perpustakaan yang tercatat pada daftar buku pengunjung perpustakaan dapat diolah untuk mengetahui informasi yang tersembunyi dari data tersebut. Informasi yang terdapat pada data pengunjung perpustakaan yaitu berupa informasi profil pengunjung perpustakaan meliputi status dan alamat dari data pengunjung perpustakaan. Kalangan apa yang mendominasi mengunjungi perpustakaan dan berasal dari daerah mana berdasarkan masing-masing kelompok sejumlah kelompok yang telah ditentukan. Berdasarkan uraian tersebut, solusi yang dapat diberikan yaitu dengan mengolah data pengunjung perpustakaan yang menumpuk dan menuangkan dalam penelitian dengan judul "Clustering Profil Pengunjung Perpustakaan (Studi Kasus Perpustakaan BP Batam)". *Clustering* digunakan untuk mengelompokkan data pengunjung perpustakaan tahun 2015 untuk mengetahui profil yang terdapat pada setiap pengelompokan dari data pengunjung perpustakaan, untuk mengetahui informasi profil yang didapat dari pengolahan data pengunjung perpustakaan menggunakan metode *Clustering* dengan algoritma K-Means yang akan diterapkan ke dalam sistem

berbasis desktop yang dibuat menggunakan bahasa pemrograman *Java Netbeans*.

A. Tinjauan Pustaka

Perbandingan antara penelitian analisis data yang dilakukan dengan penelitian analisis data sebelumnya terdapat pada Tabel 1.

TABEL I
PERBANDINGAN PENELITIAN

Penulis (Tahun)	Fina Nasari dan Surya Darma (2015)	Fauziah Mahmuda (2017)
Objek yang diteliti	Penerapan K-Means Clustering pada Data Penerimaan Mahasiswa Baru	Clustering Profil Pengunjung Perpustakaan (Studi Kasus Perpustakaan BP Batam).
Studi Kasus	Universitas Potensi Utama	Perpustakaan BP Batam
Algoritma	K-Means	K-Means
Seleksi Atribut	Nilai UAN, Asal Sekolah, Prodi.	Data Pengunjung: Status dan Alamat.
Tujuan	Mengelompokkan data mahasiswa Universitas Potensi Utama tahun ajaran 2014/2015.	Mengelompokkan profil data pengunjung perpustakaan meliputi status dan alamat. Hasil pengelompokkan dengan melihat data profil berupa status dan alamat yang mendominasi dari masing-masing cluster.
Software Pengembang	-	Netbeans (JAVA)

B. Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut:

- Membuat *clustering* dengan algoritma K-Means untuk mengelompokkan profil data pengunjung perpustakaan meliputi status dan alamat yang mendominasi pada setiap cluster.
- Membangun aplikasi yang dapat melakukan *clustering* untuk mengelompokkan profil data pengunjung perpustakaan meliputi status dan alamat pengunjung perpustakaan dengan menggunakan algoritma K-Means.

C. Manfaat

Manfaat dari penelitian yang dilakukan adalah bagaimana memanfaatkan basis data perpustakaan yang menumpuk (besar), sehingga menjadi data mining yang dapat memberikan informasi profil pengunjung perpustakaan berupa status dan alamat yang mendominasi berdasarkan pengelompokkannya.

II. LANDASAN TEORI

Perpustakaan adalah mencakup suatu ruangan, bagian dari gedung/bangunan atau gedung tersendiri yang berisi buku-buku koleksi, yang diatur dan disusun demikian rupa, sehingga mudah untuk dicari dan dipergunakan apabila sewaktu-waktu diperlukan oleh pembaca [14].

Perpustakaan adalah kumpulan atau bangunan fisik sebagai tempat buku dikumpulkan dan disusun menurut sistem tertentu atau keperluan pemakai [8]. Dari beberapa pengertian diatas, dapat disimpulkan bahwa perpustakaan

adalah tempat untuk mengembangkan informasi dan pengetahuan yang dikelola oleh suatu lembaga pendidikan, sekaligus sebagai sarana edukatif untuk membantu memperlancar cakrawala pendidik dan peserta didik dalam kegiatan belajar mengajar. Dan juga perpustakaan dapat dirumuskan sebagai suatu unit kerja dari sebuah lembaga pendidikan yang berupa tempat penyimpanan koleksi buku-buku pustaka untuk menunjang proses pendidikan [11].

K-Means merupakan salah satu metode data *clustering* non hirarki yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih *cluster*/kelompok. Metode ini mempartisi data ke dalam *cluster*/kelompok sehingga data yang memiliki karakteristik sama dikelompokkan ke dalam satu *cluster* yang sama [1]. Algoritma K-Means merupakan algoritma yang relatif sederhana untuk mengklasifikasikan atau mengelompokkan sejumlah besar obyek dengan atribut tertentu ke dalam kelompok-kelompok (*cluster*) sebanyak K. Pada algoritma K-Means, jumlah cluster K sudah ditentukan lebih dahulu. Setiap cluster memiliki titik pusat dan anggota-anggota dari satu cluster dipilih berdasarkan jarak dari titik pusat cluster terdekat. Penentuan keanggotaan dan titik pusat cluster kemudian menjadi tidak mudah, karena penambahan satu anggota pada lokasi yang signifikan akan merubah lokasi titik pusat cluster, dan status keanggotaan harus ditinjau kembali, perubahan keanggotaan kemudian akan kembali merubah lokasi titik pusat dan seterusnya keanggotaan mungkin akan berubah lagi. Karena itu, proses penentuan titik pusat dan keanggotaan cluster harus dilakukan dalam iterasi (perulangan) hingga posisi titik pusat dan anggota-anggota cluster benar-benar stabil [15].

Algoritma K-Means adalah sebagai berikut:

- Tentukan nilai k sebagai jumlah cluster yang ingin dibentuk;
- Tentukan k *centroid* (titik pusat cluster) awal secara random;
- Hitung jarak setiap data ke masing-masing centroid menggunakan rumus korelasi antar dua objek seperti rumus Euclidean Distance;

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \dots\dots\dots(1)$$

Keterangan:

- dij = jarak antara i dan j
- x_i = koordinat x objek
- x_j = koordinat x pusat
- y_i = koordinat y objek
- y_j = koordinat y pusat

- Kelompokkan setiap data berdasarkan jarak terdekat antara data dengan centroidnya;
- Tentukan posisi centroid baru dengan cara menghitung nilai rata-rata dari data-data yang ada pada cluster yang sama.

$$C_i = \frac{x_1+x_2+x_3+\dots+x_n}{x} \dots\dots\dots(2)$$

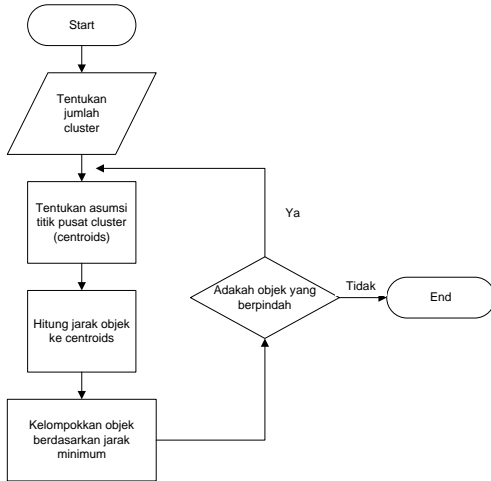
Keterangan:

x1 = nilai data *record* ke-1

x2 = nilai data *record* ke-2

$\sum x$ = jumlah data *record*

- Kembali ke langkah 3 jika posisi centroid baru dengan centroid lama tidak sama. Secara sederhana algoritma K-Means dapat digambarkan dalam diagram alur sebagai berikut:



Gambar 2. Diagram Alur Algoritma K-Means Cluster

Metode *Elbow* merupakan suatu metode yang digunakan untuk menghasilkan informasi dalam menentukan jumlah cluster terbaik dengan cara melihat persentase hasil perbandingan antara jumlah cluster yang akan membentuk siku pada suatu titik [9]. Hasil persentase yang berbeda dari setiap nilai cluster dapat ditunjukkan dengan menggunakan grafik sebagai sumber informasinya. Jika nilai cluster pertama dengan nilai cluster kedua memberikan sudut dalam grafik atau nilainya mengalami penurunan paling besar maka nilai cluster tersebut yang terbaik [3]. Untuk mendapatkan perbandingannya adalah dengan menghitung SSE (*Sum of Square Error*) dari masing-masing nilai cluster.

Rumus SSE pada K-Means:

$$SSE = \sum_{i=1}^k \sum_{p \in C_i} d(p, m_i)^2 \dots\dots\dots(3)$$

Keterangan :

p ∈ Ci = tiap data poin pada cluster i

mi = centroid dari cluster i

d = jarak/ distances/ variance terdekat pada masing-masing cluster i

Setelah dilihat akan ada beberapa nilai K yang mengalami penurunan paling besar dan selanjutnya hasil dari nilai K akan turun secara perlahan-lahan sampai hasil dari nilai K tersebut stabil. Misalnya nilai cluster K=2 ke K=3, kemudian dari K=3 ke K=4, terlihat penurunan drastis membentuk siku pada titik K=3 maka nilai cluster k yang ideal adalah K=3 [7].

Algoritma Metode Elbow dalam menentukan nilai K pada K-Means:

1. Mulai
2. Inisialisasi awal nilai K
3. Naikkan nilai K
4. Hitung hasil sum of square error dari tiap nilai K
5. Melihat hasil sum of square error dari nilai K yang turun secara drastis
6. Tetapkan nilai K yang berbentuk siku
7. Selesai

III. ANALISIS DAN PERANCANGAN

A. Simulasi Perhitungan Algoritma K-Means

Simulasi perhitungan terhadap data pengunjung perpustakaan dengan menggunakan 23 dataset sebagai data sampel dari jumlah data utuh sejumlah 1556 dataset. Atribut yang digunakan yaitu status dan alamat, bertujuan untuk mengelompokkan profil pengunjung berupa status dan alamat yang mendominasi mengunjungi perpustakaan dari masing-masing cluster.

Berikut data sampel pengunjung perpustakaan yang akan diperhitungkan dengan algoritma K-Means:

TABEL II
DATA PENGUNJUNG PERPUSTAKAAN

No	Status	Alamat
A1	4	6
A2	4	1
A3	3	1
A4	3	1
A5	2	4
A6	4	2
A7	1	2
A8	2	2
A9	2	2
A10	4	2
A11	2	1
A12	2	1
A13	3	3
A14	1	1
A15	2	1
A16	2	6
A17	2	4
A18	2	3
A19	2	3
A20	2	4
A21	2	3
A22	2	1
A23	2	1

Dari data di atas akan dipisah menjadi 3 kelompok (K=3), yang tujuannya adalah mengelompokkan profil pengunjung perpustakaan dengan melihat data status dan alamat yang mendominasi dari masing-masing cluster.

K-Means Clustering

Membagi menjadi K=3, pertama tentukan secara acak titik pusat sebagai pusat dari masing-masing cluster.

Selanjutnya, sebut saja sebagai P1=A2=(4,1), P2=A5=(2,4), P3=A10=(4,2).

Untuk mengetahui jarak dapat dihitung dengan menggunakan rumus Euclidian Distance:

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \dots\dots\dots(4)$$

Keterangan:

- dij = jarak antara i dan j
- xi = koordinat x objek
- xj = koordinat x pusat
- yi = koordinat y objek
- yj = koordinat y pusat

Iterasi 1

TABEL III
MENGHITUNG JARAK OBJEK DENGAN PUSAT PADA ITERASI 1

	P1 (4,1)	P2 (2,4)	P3 (4,2)
A1 (4,6)	$\sqrt{(4-4)^2 + (6-1)^2}$ =sqrt(0+25) =5	$\sqrt{(4-2)^2 + (6-4)^2}$ =sqrt(4+4) =2.828	$\sqrt{(4-4)^2 + (6-2)^2}$ =sqrt(0+16) =4
A2 (4,1)	0	3.605	1
A3 (3,1)	1	3.162	1.414
A4 (3,1)	1	3.162	1.414
A5 (2,4)	3.605	0	2.828
A6 (4,2)	1	2.828	0
A7 (1,2)	3.162	2.236	3
A8 (2,2)	2.236	2	2
A9 (2,2)	2.236	2	2
A10 (4,2)	1	2.828	0
A11 (2,1)	2	3	2.236
A12 (2,1)	2	3	2.236
A13 (3,3)	2.236	1.414	1.414
A14 (1,1)	3	3.162	3.162
A15 (2,1)	2	3	2.236
A16 (2,6)	5.385	2	4.472
A17 (2,4)	3.605	0	2.828
A18 (2,3)	2.828	1	2.236
A19 (2,3)	2.828	1	2.236
A20 (2,4)	3.605	0	2.828
A21 (2,3)	2.828	1	2.236
A22 (2,1)	2	3	2.236
A23 (2,1)	2	3	2.236

Sehingga cluster yang diperoleh pada iterasi 1 adalah:
 P1 : {A2, A3, A4, A11, A12, A14, A15, A22, A23}
 P2 : {A1, A5, A7, A8, A9, A13, A16, A17, A18, A19, A20, A21}
 P3 : {A6, A10}

Langkah selanjutnya adalah menghitung pusat cluster baru dengan rumus:

$$C_i = \frac{x_1+x_2+x_3+\dots+x_n}{x}$$

Keterangan:

- x1 = nilai data record ke-1
- x2 = nilai data record ke-2
- Σx = jumlah data record

P1 = ((4+3+3+2+2+1+2+2+2)/9, (1+1+1+1+1+1+1+1)/9) = (2.333, 1)
 P2 = ((4+2+1+2+2+3+2+2+2+2+2)/12, (6+4+2+2+2+3+6+4+3+3+4+3)/12) = (2.166, 3.5)
 P3 = ((4+4)/2, (2+2)/2) = (4, 2)

Iterasi 2

TABEL IV
MENGHITUNG JARAK OBJEK DENGAN PUSAT PADA ITERASI 2

	P1 (2.333, 1)	P2 (2.166, 3.5)	P3 (4, 2)
A1 (4,6)	$\sqrt{(4-2.333)^2 + (6-1)^2}$ =sqrt(2.778+25) =5.270	$\sqrt{(4-2.166)^2 + (6-3.5)^2}$ =sqrt(3.363+6.25) =3.100	$\sqrt{(4-4)^2 + (6-2)^2}$ =sqrt(0+16) =4
A2 (4,1)	1.666	3.100	1
A3 (3,1)	0.666	2.635	1.414
A4 (3,1)	0.666	2.635	1.414
A5 (2,4)	3.018	0.527	2.828
A6 (4,2)	1.943	2.368	0
A7 (1,2)	1.666	1.900	3
A8 (2,2)	1.054	1.509	2
A9 (2,2)	1.054	1.509	2
A10 (4,2)	1.943	2.368	0
A11 (2,1)	0.333	2.505	2.236
A12 (2,1)	0.333	2.505	2.236
A13 (3,3)	2.108	0.971	1.414
A14 (1,1)	1.333	2.758	3.162
A15 (2,1)	0.333	2.505	2.236
A16 (2,6)	5.011	2.505	4.472
A17 (2,4)	3.018	0.527	2.828
A18 (2,3)	2.027	0.527	2.236
A19 (2,3)	2.027	0.527	2.236
A20 (2,4)	3.018	0.527	2.828
A21 (2,3)	2.027	0.527	2.236
A22 (2,1)	0.333	2.505	2.236
A23 (2,1)	0.333	2.505	2.236

Sehingga cluster yang diperoleh pada iterasi 2 adalah:
 P1 : {A3, A4, A7, A8, A9, A11, A12, A14, A15, A22, A23}
 P2 : {A1, A5, A13, A16, A17, A18, A19, A20, A21}
 P3 : {A2, A6, A10}

Langkah selanjutnya adalah menghitung pusat cluster baru, yaitu:

P1= ((3+3+1+2+2+2+2+1+2+2+2)/11, (1+1+2+2+2+1+1+1+1+1+1)/11) = (2, 1.272)
 P2 = ((4+2+3+2+2+2+2+2+2)/9, (6+4+3+6+4+3+3+4+3)/9) = (2.333, 4)
 P3 = ((4+4+4)/3, (1+2+2)/3) = (4, 1.666)

Iterasi 3

TABEL V
MENGHITUNG JARAK OBJEK DENGAN PUSAT PADA ITERASI 3

	P1 (2, 1.272)	P2 (2.333, 4)	P3 (4, 1.666)
A1 (4,6)	$\sqrt{(4-2)^2 + (6-1.272)^2}$ =sqrt(4+22.353) =5.132	$\sqrt{(4-2.333)^2 + (6-4)^2}$ =sqrt(2.778+4) =2.603	$\sqrt{(4-4)^2 + (6-1.666)^2}$ =sqrt(0+18.783) =4.333

A2 (4,1)	2.018	3.431	0.666
A3 (3,1)	1.036	3.073	1.201
A4 (3,1)	1.036	3.073	1.201
A5 (2,4)	2.727	0.333	3.073
A6 (4,2)	2.128	2.603	0.333
A7 (1,2)	1.236	2.403	3.018
A8 (2,2)	0.727	2.027	2.027
A9 (2,2)	0.727	2.027	2.027
A10 (4,2)	2.128	2.603	0.333
A11 (2,1)	0.272	3.018	2.108
A12 (2,1)	0.272	3.018	2.108
A13 (3,3)	1.995	1.201	1.666
A14 (1,1)	1.036	3.282	3.073
A15 (2,1)	0.272	3.018	2.108
A16 (2,6)	4.727	2.027	4.772
A17 (2,4)	2.727	0.333	3.073
A18 (2,3)	1.727	1.054	2.403
A19 (2,3)	1.727	1.054	2.403
A20 (2,4)	2.727	0.333	3.073
A21 (2,3)	1.727	1.054	2.403
A22 (2,1)	0.272	3.018	2.108
A23 (2,1)	0.272	3.018	2.108

Sehingga cluster yang diperoleh pada iterasi 3 adalah:
 P1 : {A3, A4, A7, A8, A9, A11, A12, A14, A15, A22, A23}
 P2 : {A1, A5, A13, A16, A17, A18, A19, A20, A21}
 P3 : {A2, A6, A10}

Karena cluster yang terbentuk pada iterasi 3 sama dengan yang terbentuk pada cluster sebelumnya (cluster 2), maka iterasi selesai atau berhenti. Sehingga didapatkan hasil clustering pada Tabel 6 berikut.

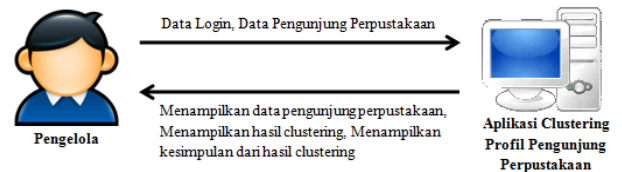
TABEL VI
HASIL CLUSTERING

No	Status	Alamat
Kelompok 1 berjumlah 11 orang		
A3	Pelajar	Batam Center
A4	Pelajar	Batam Center
A7	Karyawan BP	Sekupang
A8	Mahasiswa	Sekupang
A9	Mahasiswa	Sekupang
A11	Mahasiswa	Batam Center
A12	Mahasiswa	Batam Center
A14	Karyawan BP	Batam Center
A15	Mahasiswa	Batam Center
A22	Mahasiswa	Batam Center
A23	Mahasiswa	Batam Center
Kelompok 2 berjumlah 9 orang		
A1	Umum	Nongsa
A5	Mahasiswa	Bengkong

A13	Pelajar	Batu Aji
A16	Mahasiswa	Nongsa
A17	Mahasiswa	Bengkong
A18	Mahasiswa	Batu Aji
A19	Mahasiswa	Batu Aji
A20	Mahasiswa	Bengkong
A21	Mahasiswa	Batu Aji
Kelompok 3 berjumlah 3 orang		
A2	Umum	Batam Center
A6	Umum	Sekupang
A10	Umum	Sekupang

Data dari masing-masing kelompok dapat diambil kesimpulan informasi profil berupa status dan alamat yang mendominasi mengunjungi perpustakaan BP Batam dari masing-masing kelompok. Kelompok 1 didominasi oleh pengunjung Mahasiswa dan beralamat di Batam Center. Kelompok 2 didominasi oleh pengunjung Mahasiswa dan beralamat di Batu Aji. Kelompok 3 didominasi oleh pengunjung Umum dan beralamat di Sekupang.

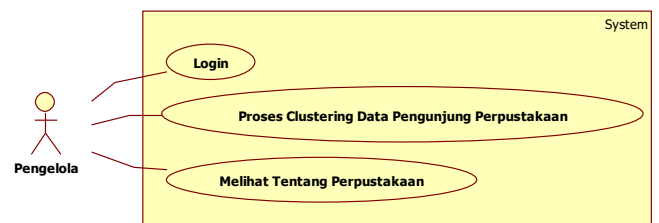
B. Deskripsi Umum Sistem



Gambar 4. Deskripsi Umum Sistem

Gambar 4 menjelaskan bahwa sistem hanya memiliki satu user, yaitu pengelola perpustakaan itu sendiri. Pengelola memasukkan data login berupa username dan password, lalu pengelola dapat memasukkan data pengunjung perpustakaan berekstensi.csv. Setelah memasukkan data pengunjung yang akan dicluster didalam aplikasi, aplikasi akan menampilkan informasi isi data yang diinputkan setelah itu data akan diproses menggunakan metode clustering dengan algoritma K-Means, data yang diolah hanya berupa atribut status dan alamat, dan aplikasi akan menampilkan hasil clustering. Setelah itu aplikasi akan menarik kesimpulan dari hasil clustering tersebut dan menampilkannya.

C. Use Case Diagram

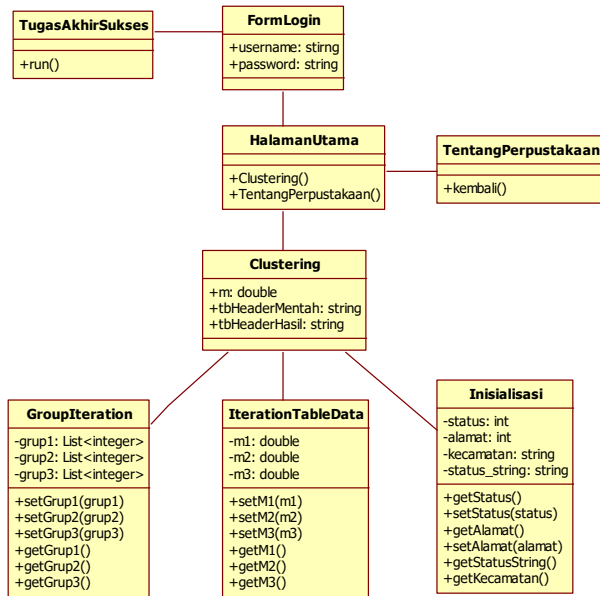


Gambar 5. Use Case Diagram

Gambar 5 menjelaskan bahwa pengelola memiliki akses login. Pada proses clustering pengelola dapat

memasukkan data pengunjung perpustakaan yang berekstensi.csv. Pengelola dapat melihat informasi Tentang Perpustakaan BP Batam.

D. Class Diagram



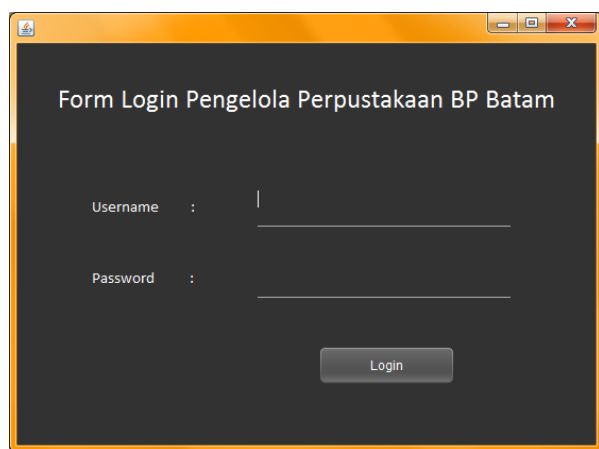
Gambar 6. Class Diagram

Berdasarkan gambar 6 dapat dijelaskan bahwa dari class FormLogin terhubung dengan class HalamanUtama. Class HalamanUtama terdapat class TentangPerpustakaan dan class Clustering, Class Clustering terdiri dari class GroupIteration, class IterationTableData dan class Inisialisasi.

IV. IMPLEMENTASI

A. Implementasi Antarmuka Aplikasi

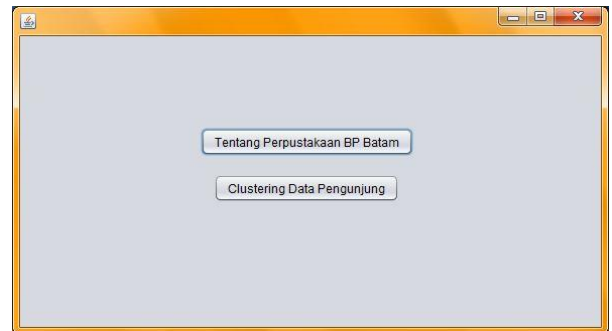
1) Implementasi Antarmuka Halaman Login:



Gambar 7. Implementasi Antarmuka Halaman Login

Gambar 7 diatas adalah gambar antarmuka halaman login pengelola, pengelola mengisi data login berupa username dan password pada halaman login. Login berguna untuk menjaga aplikasi dari segi keamanan. Apabila username dan password terisi dengan benar maka pengelola dapat mengakses aplikasi.

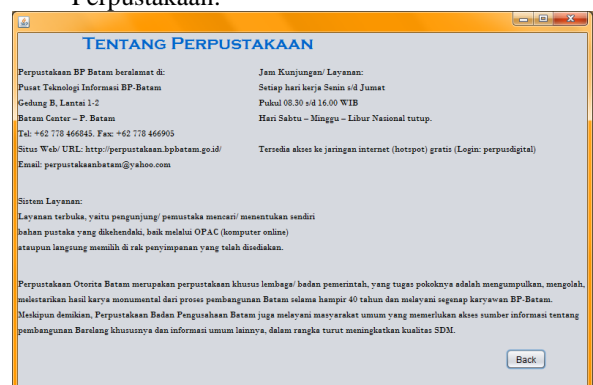
2) Implementasi Antarmuka Halaman Utama:



Gambar 8. Implementasi Antarmuka Halaman Utama

Gambar 8 diatas adalah gambar antarmuka halaman utama. Pengelola dapat mengakses halaman tersebut apabila telah berhasil login. Terdapat dua pilihan menu, menu tentang perpustakaan BP Batam dan menu Clustering Data Pengunjung.

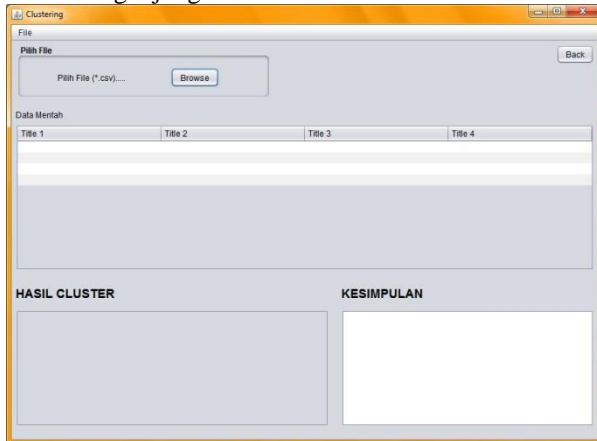
3) Implementasi Antarmuka Halaman Tentang Perpustakaan:



Gambar 9. Implementasi Antarmuka Halaman Tentang Perpustakaan

Gambar 9 diatas adalah gambar antarmuka halaman tentang perpustakaan yang terdapat pada menu “Tentang Perpustakaan BP Batam”. Berisi informasi mengenai perpustakaan.

4) Implementasi Antarmuka Halaman *Clustering* Data Pengunjung:



Gambar 10. Implementasi Antarmuka Halaman Clustering Data Pengunjung

Gambar 10 diatas ini adalah antarmuka halaman *clustering* data pengunjung. Pada halaman ini pengelola dapat memasukkan *file* berekstensi *csv*. *File* yang diinput tampil berdasarkan isi data sesuai format yang telah ditentukan. Data yang diinput diolah menggunakan algoritma K-Means di dalam aplikasi, hasil data yang telah diolah akan ditampilkan pada kolom **Hasil Cluster**. Data dari hasil cluster disimpulkan pada kolom **Kesimpulan**.

5) Implementasi Antarmuka Halaman Data Mentah:

No.	Tanggal	Nama	Status	Alamat
1	2/1/2015	Anggarda Bhayangkara	Umum	Kabil
2	2/1/2015	Fani	Umum	Citra Batam
3	2/1/2015	Aidir	Pelajar	Legenda
4	2/1/2015	Ian	Pelajar	Legenda
5	2/1/2015	Abdul Karim Nasution	Mahasiswa	Bengkong
6	4/1/2015	Jamil	Umum	Sekupang
7	5/1/2015	Martini	Karyawan BP	Sekupang
8	5/1/2015	Tika	Mahasiswa	Sekupang
9	5/1/2015	Arif F	Mahasiswa	Sekupang
10	5/1/2015	Salita N	Umum	Sekupang
11	5/1/2015	Halimatussadiyah	Mahasiswa	Baloi Harapan 2

Gambar 11. Implementasi Antarmuka Halaman Data Mentah

Gambar 11 merupakan tampilan isi data dari data yang diinputkan pada aplikasi, dengan menampilkan semua atribut dan isi data yang ada.

6) Implementasi Antarmuka Halaman Data Hasil Cluster :

No.	Status	Alamat	Cluster
464	Mahasiswa	Anggrek Sari	Cluster 1
93	Mahasiswa	Anggrek Sari	Cluster 1
94	Mahasiswa	Anggrek Sari	Cluster 1
1128	Karyawan BP	Baloi	Cluster 1
1168	Mahasiswa	Baloi	Cluster 1
1187	Mahasiswa	Baloi	Cluster 1
121	Karyawan BP	Baloi	Cluster 1
1255	Mahasiswa	Baloi	Cluster 1
1276	Mahasiswa	Baloi	Cluster 1

No.	Status	Alamat	Cluster
757	Karyawan BP	Anggrek Permai	Cluster 2
1283	Umum	Bali	Cluster 2
1176	Mahasiswa	Baloi Center	Cluster 2
1201	Mahasiswa	Baloi Center	Cluster 2
1404	Mahasiswa	Baloi Center	Cluster 2
176	Mahasiswa	Baloi Center	Cluster 2
177	Mahasiswa	Baloi Center	Cluster 2
187	Mahasiswa	Baloi Center	Cluster 2
32	Mahasiswa	Baloi Center	Cluster 2

No.	Status	Alamat	Cluster
137	Umum	Baloi Ditpam	Cluster 3
491	Umum	Baloi Ditpam	Cluster 3
1116	Pelajar	Baloi Harapan	Cluster 3
101	Umum	Batam Center	Cluster 3
102	Pelajar	Batam Center	Cluster 3
1086	Umum	Batam Center	Cluster 3
1089	Umum	Batam Center	Cluster 3
1090	Umum	Batam Center	Cluster 3
1092	Umum	Batam Center	Cluster 3

Gambar 12. Implementasi Antarmuka Halaman Data Hasil Cluster

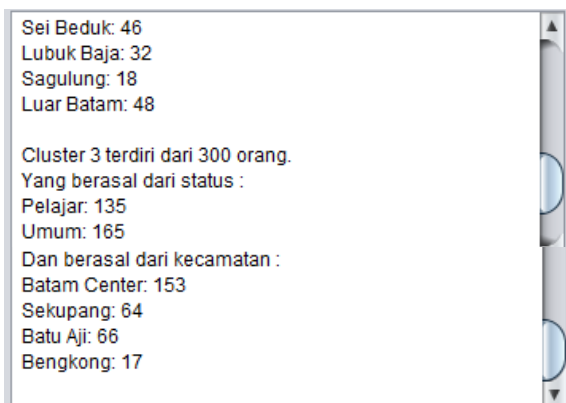
Gambar 12 merupakan implementasi antarmuka hasil *clustering* yang menampilkan hasil cluster dari data masukkan yang diolah dengan algoritma K-Means pada aplikasi berdasarkan status dan alamat. Pada aplikasi ditetapkan pengelompokkan menjadi 3 cluster, dan menggunakan titik pusat (*centroid*) awal $P1=(4,1)$, $P2=(2,4)$, $P3=(4,2)$.

7) Implementasi Antarmuka Halaman Kesimpulan Hasil Cluster:

KESIMPULAN

Cluster 1 terdiri dari 814 orang.
 Yang berasal dari status :
 Karyawan BP: 308
 Mahasiswa: 506
 Dan berasal dari kecamatan :
 Batam Center: 475
 Sekupang: 184
 Batu Aji: 155

Cluster 2 terdiri dari 442 orang.
 Yang berasal dari status :
 Karyawan BP: 49
 Mahasiswa: 237
 Pelajar: 63
 Umum: 93
 Dan berasal dari kecamatan :
 Bengkong: 168
 Batu Ampar: 70
 Nongsa: 60



Gambar 13. Implementasi Antarmuka Halaman Kesimpulan Hasil Cluster

Gambar 13 merupakan implementasi antarmuka kesimpulan hasil *clustering* yang menampilkan kesimpulan dari masing-masing data yang terdapat pada setiap cluster berdasarkan jumlah data setiap cluster dan rincian jumlah status dan alamat pada setiap cluster. Hasil cluster data pengunjung didapatkan dari proses K-Means dalam aplikasi, dengan jumlah data 1556 dataset. Setiap data pada suatu cluster memiliki tingkat kemiripan yang tinggi dan sangat berbeda dengan data yang ada pada cluster lainnya.

V. KESIMPULAN

Berdasarkan tahapan analisis, perancangan, dan implementasi yang dilakukan pada analisis data pengunjung perpustakaan Badan Pengusahaan Batam, maka kesimpulan yang dapat diambil:

- *Clustering* dibuat dengan menggunakan algoritma K-Means dan untuk mengetahui jumlah K cluster data pengunjung perpustakaan dibantu dengan menggunakan metode Elbow. Pada studi kasus perpustakaan BP Batam dapat disimpulkan pengelompokan profil pengunjung perpustakaan yang mendominasi pada setiap cluster yang meliputi 1556 dataset, yaitu: Cluster 1 yang didominasi oleh pengunjung berstatus Mahasiswa dan beralamat di Batam Center. Cluster 2 didominasi oleh pengunjung berstatus Mahasiswa dan beralamat di Bengkong. Cluster 3 didominasi oleh pengunjung berstatus Umum dan beralamat di Batam Center.
- Aplikasi yang dibangun berhasil melakukan *clustering* profil data pengunjung perpustakaan meliputi status dan alamat pengunjung perpustakaan dengan menggunakan algoritma K-Means. Dengan menetapkan jumlah K cluster yang digunakan yaitu $K=3$, dan titik pusat (*centroid*) awal yang digunakan yaitu $P1=(4,1)$, $P2=(2,4)$, $P3=(4,2)$. Hasil cluster dipengaruhi dari nilai *centroid* awal yang digunakan, perbedaan pengambilan *centroid* awal yang digunakan akan mempengaruhi hasil *centroid* akhirnya.

DAFTAR PUSTAKA

- [1] Agusta, Y., 2007, K-means - Penerapan, Permasalahan dan Metode Terkait, Jurnal Sistem dan Informatika Vol. 3 (Februari 2007): 47-60.
- [2] Bertalya., 2009, Data Mining & Knowledge Discovery in Database, Universitas Gunadarma.
- [3] Bholowalia, Purnima., Kumar, Arvind., 2014, EBK-Means: A Clustering Techniques based on Elbow Method and K-Means in WSN, International Journal of Computer Application (0975-8887), IX(105), pp. 17-24.
- [4] Cios, Krzysztof J., Lukasz A. Kurgan., 2002, Trends in Data Mining and Knowledge Discovery.
- [5] Han, Jiawei., Kamber, Micheline., 2006, Data Mining: Concepts and Techniques, Morgan Kaufman: San Fransisco.
- [6] Hilda, Widyastuti., 2012, Modul Pembelajaran Data Mining, Perpustakaan Politeknik Negeri Batam, Batam.
- [7] Kodinariya, T. M., Makwana, P. R., 2013, Review on determining number of cluster in K-Means Clustering, International Journal of Advance Research in Computer Science and Management Studies, I(6), pp. 90-95.
- [8] Lasa HS., 2007, Manajemen Perpustakaan Sekolah, Yogyakarta: Pinus Book Publisher.
- [9] Madhulatha, T. S., 2012, An Overview On Clustering Methods, IOSR Journal of Engineering, II(4), pp.719-725.
- [10] Michael Berry., Gordon S. Linoff., 2004, Data Mining Techniques: For Marketing, Sales, and Customer Relationship Management, John Wiley & Sons, Inc.
- [11] Muchlisin, R., 2012, Pengertian, Jenis dan Tujuan Perpustakaan.
- [12] Santoso, S., 2010, Statistik Multivariat, Jakarta: Elex Media Komputindo.
- [13] Supranto, J., 2004, Analisis Multivariat: Arti Dan Interpretasi, Jakarta: Rineka Cipta.
- [14] Sutarno NS., 2006, Perpustakaan dan Masyarakat, Jakarta: Sagung Seto.
- [15] Widya, S.A., Dedy, A., 2016, Pengelompokan Minat Baca Mahasiswa menggunakan Metode K-Means, Teknik Informatika, Jurnal Ilmiah ILKOM Vol.8 No.2, Universitas Muslim Indonesia.
- [16] Yudho, GS., 2003, Penerapan Data Mining, Artikel Populer IlmuKomputer.com.