

Penerapan metode GA-Kmeans untuk pengelompokan pengguna pada Bapersip Provinsi Jawa Timur

Ferlyna K Wardhani, Erma Suryani, dan Ahmad Mukhlason

Sistem Informasi, Fakultas Teknologi Informasi, Institut Teknologi Sepuluh Nopember (ITS)

Jl. Arief Rahman Hakim, Surabaya 60111

E-mail: erma@is.its.ac.id

Abstrak— Badan Perpustakaan dan Kearsipan Provinsi Jawa Timur tidak hanya berperan sebagai institusi pengelola dan penyedia jasa informasi, namun juga berperan penting dalam pengembangan budaya dan minat baca masyarakat, sesuai dengan ditetapkannya program pemerintah yaitu “Jawa Timur Membaca dan Tertib Arsip 2014”. Berdasarkan data statistik peminjaman pada Bapersip serta data jumlah penduduk Surabaya dapat disimpulkan bahwa perbandingan jumlah buku yang dipinjam dengan jumlah masyarakat Surabaya hanya 1:76, yang mengindikasikan bahwa tingkat kesadaran minat baca masyarakat Surabaya masih rendah. Salah satu hal yang dapat dilakukan untuk meningkatkan minat baca masyarakat Surabaya adalah dengan meningkatkan intensitas kunjungan dan aktifitas peminjaman pada perpustakaan.

Peningkatan aktifitas peminjaman dapat dilakukan dengan meningkatkan kualitas layanan Bapersip yang lebih personal kepada pengguna. Oleh sebab itu, dilakukan klusterisasi dari demografis pengguna serta perilaku pengguna untuk lebih memahami kelompok-kelompok pengguna. Metode yang digunakan dalam klusterisasi ini adalah menggunakan GA-Kmeans. Metode ini dipilih karena dapat menghasilkan kinerja yang lebih optimal dibandingkan dengan menggunakan metode Kmeans sederhana, melalui penentuan initial centroid yang lebih optimal. Hasil pengelompokan yang diperoleh menunjukkan bahwa nilai *total within cluster variation* dari pengelompokan yang telah dioptimasi dengan algoritma genetika bernilai lebih kecil dibandingkan pengelompokan menggunakan algoritma K-means sederhana. Dengan kata lain pengelompokan menggunakan GA-Kmeans menghasilkan klaster yang lebih optimal dibandingkan metode K-means sederhana. Hasil pengelompokan menghasilkan empat klaster pengguna disertai rekomendasi kategori buku yang sesuai untuk setiap klaster pengguna Bapersip.

Kata Kunci—Perpustakaan, klusterisasi, GA-Kmeans

I. PENDAHULUAN

BADAN Perpustakaan dan Kearsipan Provinsi Jawa Timur merupakan lembaga yang dibentuk pemerintah untuk menjalankan tugas dan fungsinya yang secara spesifik dalam bidang Perpustakaan dan Kearsipan sesuai dengan Peraturan Pemerintah No. 41 tahun 2007 tentang Organisasi Perangkat Daerah [1]. Sesuai dengan visinya yaitu melaksanakan fungsi

dan tugasnya dalam mendukung program “Jawa Timur Membaca dan Tertib Arsip 2014” yang dicanangkan oleh Pemerintah Provinsi Jawa Timur, Bapersip tidak hanya menyediakan jasa informasi, namun juga berperan aktif dalam meningkatkan minat baca masyarakat.

Sebuah penelitian terhadap minat baca masyarakat yang diadakan oleh Bapersip dan sebuah lembaga *survey* pada tahun 2010 menyatakan bahwa hanya sebanyak 54% warga Surabaya membaca kurang dari tiga buah buku dalam sebulan [2]. Berdasarkan data yang terdapat pada Bapersip Provinsi Jawa Timur, pada Desember 2010 telah tercatat sebanyak lebih dari 40000 orang anggota perpustakaan, dengan rata-rata jumlah kunjungan mencapai 100-175 per hari [3]. Namun, jumlah aktifitas peminjaman dari pelayanan sirkulasi Bapersip dari hari senin-jumat hanya sekitar 165 buku/ hari [3]. Jika dikaitkan dengan jumlah penduduk Surabaya pada tahun 2011 yang mencapai 3.024.321 warga [4], dapat disimpulkan bahwa perbandingan jumlah penduduk dengan jumlah pengguna yang terdaftar pada Bapersip hanya sebesar 1:76. Jika dilihat dari aktifitas peminjaman buku, dapat dinyatakan bahwa dari 76 orang masyarakat Surabaya, hanya meminjam satu buah buku pada Bapersip. Rendahnya aktifitas masyarakat dalam meminjam buku ataupun melakukan kunjungan ke Bapersip tersebut, turut mengindikasikan masih rendahnya minat baca masyarakat Surabaya.

Salah satu upaya yang dapat dilakukan dalam meningkatkan minat baca masyarakat adalah dengan meningkatkan aktifitas kunjungan dan aktifitas peminjaman buku pada Bapersip. Aktifitas kunjungan dan peminjaman dapat ditingkatkan dengan memberikan fasilitas dan pelayanan prima kepada pengunjung. Sebuah pendekatan *customer centric* perlu diterapkan, sehingga dibutuhkan suatu teknik yang dapat memberikan informasi spesifik tentang anggota perpustakaan yang nantinya dapat digunakan untuk membantu mendesain dan menyediakan layanan yang dapat memenuhi kebutuhan anggota perpustakaan. Dalam hal ini, pengelompokan dari perilaku pengguna dalam aktifitas peminjaman buku pada Bapersip Jawa Timur perlu dilakukan, sehingga dapat menghasilkan analisis rekomendasi buku yang sesuai dengan karakteristik pengguna perpustakaan yang diperoleh berdasarkan data dari transaksi peminjaman pengguna, informasi personal pengguna, serta informasi katalog buku.

Dalam tugas akhir ini, dilakukan pengelompokan pengguna untuk menghasilkan rekomendasi buku dengan

menggunakan metode *GA-Kmeans clustering* pada Bapersip Provinsi Jawa Timur. Metode *GA-Kmeans* ini dipilih untuk melakukan *clustering* karena jika menggunakan metode *K-means* sederhana terdapat kekurangan pada pemilihan *initial seeds* [5]. Sebuah studi yang dilakukan oleh Kim & Ahn, 2008 yang diterapkan untuk segmentasi pasar perdagangan elektronik membuktikan bahwa *GA-Kmeans* dipandang dari perspektif *intraclass inertia* dapat menghasilkan segmentasi yang lebih baik daripada algoritma *clustering* tradisional seperti *simple K-Means*. Sehingga, metode *GA (Genetic Algorithm)* dalam tugas akhir ini dipakai sebagai optimization tool untuk pemilihan *initial seeds*. Dengan adanya analisis rekomendasi buku ini diharapkan dapat meningkatkan kualitas layanan pada Bapersip Jawa Timur yang berdampak kepada meningkatnya jumlah peminjaman buku, yang secara tidak langsung akan meningkatkan minat baca masyarakat Surabaya.

II. LANDASAN TEORI

Pada bab ini akan ditelaah mengenai dasar teori yang digunakan dalam pelaksanaan penelitian.

A. Pengelompokan pada perpustakaan

Pengelompokan pelanggan pada umumnya dilakukan pada organisasi profit. Namun, implementasinya pada organisasi *non profit*, seperti perpustakaan, tidak terlalu berhubungan dengan pencapaian keuntungan yang diperoleh organisasi. Pengelompokan pelanggan merupakan salah satu bagian dari pemasaran dari perpustakaan [6]. Pemasaran dalam perpustakaan menjadi bagian yang penting, contohnya mengetahui *audience target* diperlukan dalam memilih buku yang ingin dibaca pengguna, agar buku nantinya tidak berakhir tersimpan di rak.

Pengelompokan pada perpustakaan didefinisikan sebagai pembagian dari pasar atau informasi perpustakaan menjadi bagian yang lebih kecil, lebih dapat diatur, menjadi kelompok yang mempunyai karakteristik [7]. Acuan dalam memetakan pengguna perpustakaan dapat didasarkan pada metode yang umumnya banyak digunakan pada organisasi profit, [8].

- 1) Segmentasi geografis. Contoh: berdasarkan lokasi, kota, atau gedung yang berbeda.
- 2) Segmentasi demografis. Contoh: berdasarkan fungsi, disiplin ilmu, bahasa.
- 3) Segmentasi psikografis. Contoh: tingkatan manajemen.
- 4) Segmentasi tingkah laku. Contoh: penggunaan layanan, browser yang digunakan.

B. Analisis kluster

Analisis kluster membagi sejumlah data ke dalam kelompok -kelompok yang memberi arti dan berguna, berdasarkan informasi yang terdapat di dalamnya yang mendefinisikan objek-objek dan hubungannya [9]. Dalam satu kluster mengandung sekumpulan obyek-obyek yang memiliki kesamaan satu sama lain dan sebaliknya dengan obyek-obyek yang berbeda kluster. Obyek di dalam kluster yang berbeda mempunyai tingkat kesamaan yang lebih rendah daripada obyek-obyek yang berada di dalam satu kluster yang sama.

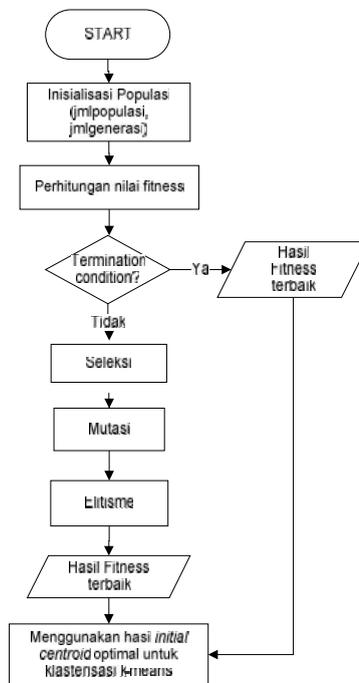
Terdapat manfaat dari dilakukannya analisis kluster pada

perpustakaan, yaitu dapat membantu meningkatkan kualitas pelayanan, menyediakan dukungan pembuatan keputusan yang lebih efektif, serta membuat perpustakaan memainkan peran yang lebih aktif dalam melayani pengguna [10].

C. Metode GA-Kmeans

K-means merupakan algoritma yang cukup cepat untuk melakukan pengklasteran. Namun, terdapat masalah pada tahap inialisasi dari initial seeds dari K-means yang dalam algoritma sederhananya ditentukan secara random, dimana pemilihan initial seeds yang berbeda akan menghasilkan kluster yang berbeda pula [11]. Dalam hal ini, algoritma genetika digunakan untuk mengoptimalkan *initial seeds* pada algoritma K-means dengan menggunakan *natural selection*.

Proses algoritma genetika dalam penentuan titik pusat kluster yang memiliki performa yang lebih cepat, yang disebut dengan algoritma genetika cepat (*Fast Genetic K-means Algorithm*) [12] dengan alur proses kerja dijelaskan pada Gambar 1.



Gambar. 1. Alur kerja algoritma GA-Kmeans [12]

III. METODOLOGI PENELITIAN

Metodologi proses pelaksanaan Tugas Akhir terdiri dari pengumpulan data, identifikasi permasalahan, studi literatur, praproses data, implementasi klasterisasi, uji coba dan validasi, analisis hasil klasterisasi, serta penyusunan buku Tugas Akhir.

A. Identifikasi Permasalahan

Pada tahap ini dilakukan identifikasi permasalahan melalui analisis singkat terhadap data yang telah dikumpulkan sebelumnya. Identifikasi ini memberikan dua hasil, pertama permasalahan yang diangkat yaitu mengenai pengelompokan pengguna untuk menghasilkan rekomendasi kategori buku, dan

permasalahan yang kedua yaitu algoritma yang tepat digunakan untuk analisis permasalahan tersebut.

B. Studi Literatur

Pada tahap ini dilakukan penggalian teori-teori yang relevan dengan permasalahan yang sedang dikaji. Studi literatur ini meliputi pemahaman metode GA-Kmeans untuk diterapkan dalam pengelompokan pengguna.

C. Praproses data

Praproses data dilakukan untuk memperbaiki kualitas data yang rendah, sehingga menghasilkan hasil pengklasteran yang berkualitas. Pada tahap ini dilakukan pengolahan data yang terdiri dari proses-proses yaitu:

- a. Pemilihan atribut (*feature selection*)
- b. Integrasi data
- c. Pembersihan data
- d. Transformasi data
- e. Reduksi data
- f. Mempersiapkan dataset

D. Implementasi Klasterisasi

Pada tahap ini dilakukan klasterisasi atau pengelompokan terhadap data Bapersip dengan menggunakan algoritma GA-Kmeans.

E. Uji coba dan validasi

Pada tahap ini dilakukan uji coba dan validasi untuk memeriksa apakah output/ keluaran serta metode yang digunakan telah valid. Uji validasi dilakukan dengan cara membandingkan tingkat performa algoritma GA-Kmeans dengan algoritma K-means sederhana. Dalam tahap ini juga dilakukan uji coba parameter algoritma genetika untuk mengetahui pengaruh nilai parameter algoritma genetika terhadap solusi yang dihasilkan.

F. Analisis Hasil Clustering

Analisis terhadap hasil klaster dilakukan untuk memperoleh relevansi informasi dengan permasalahan yang diajukan. Kemudian dilakukan pembuatan rekomendasi kategori buku yang sesuai dengan kelompok-kelompok pengguna yang didapatkan dari hasil pengklasteran.

IV. MODEL DAN IMPLEMENTASI

A. Masalah Pengelompokan Pengguna Bapersip

Dalam pengelompokan pengguna pada Bapersip ini dilakukan pengelompokan berdasarkan demografis dan tingkah laku (*behaviour*), sesuai dengan batasan data yang digunakan, yaitu data informasi (profil) pengguna, data transaksi peminjaman buku, serta data katalog (informasi) buku.

- 1. Data demografis meliputi data personal pengguna, seperti atribut gender, usia dan pekerjaan.
- 2. Data behavioural meliputi data yang berkaitan dengan transaksi peminjaman buku yang dilakukan pengguna, seperti jumlah transaksi dan jumlah buku yang

dipinjam per kategori.

B. Praproses data

Pada tahap ini akan dijelaskan proses transformasi dari data mentah menjadi data set yang siap untuk digunakan dalam proses klasterisasi. Proses-proses yang dilakukan dalam tahap praproses data ini meliputi integrasi data, pembersihan data dari *missing value dan outlier*, transformasi data, dan reduksi data.

C. Implementasi Algoritma Genetika

Tahap selanjutnya setelah melakukan praproses data adalah implementasi algoritma genetika untuk mendapatkan *initial centroid* yang optimal.

1) Representasi kromosom

Dalam kasus ini kromosom merupakan kumpulan dari

Tabel 1.

Atribut yang digunakan dalam klasterisasi		
No	Atribut	Keterangan atribut
1	Gender	Demografis
2	Usia	Demografis
2	Pekerjaan	Demografis
3	Jumlah transaksi peminjaman buku	Behavioural
4	Jumlah buku yang dipinjam kategori a	Behavioural
5	Jumlah buku yang dipinjam kategori b	Behavioural
6	Jumlah buku yang dipinjam kategori b, dst	behavioural

gen berupa bilangan real yang merepresentasikan pusat klaster. Panjang kromosom adalah sebanyak banyak jumlah atribut*jumlah klaster.

2) Inisialisasi populasi

Dalam proses pembangkitan populasi awal, pusat klaster K dikodekan pada setiap kromosom dan dibangkitkan secara random berdasarkan data yang telah dilakukan praproses data sebelumnya.

3) Perhitungan fungsi fitness

Nilai fitness dihitung berdasarkan TWCV (*Total within cluster variation*), atau dalam literatur disebut juga sebagai *square error*, dimana tujuan dari algoritma genetika adalah untuk meminimumkan TWCV. Alur dari proses perhitungan fungsi fitness meliputi:

- Menghitung jarak setiap data asli terhadap *centroid* yang direpresentasikan dalam kromosom.
- Mengambil nilai minimum dari jarak setiap data ke *centroid* ke i.
- Menjumlahkan jarak terpendek data terhadap *centroid* yang dimiliki masing-masing kromosom sebagai nilai fitness.

4) Proses seleksi

Proses seleksi digunakan untuk memilih individu-individu mana saja yang dipilih untuk dilakukan proses selanjutnya. Proses seleksi disini menggunakan seleksi roda *roulette*. Alur dari proses seleksi meliputi:

- Menghitung nilai frekuensi relatif dan frekuensi

- kumulatif dari fitness setiap kromosom.
- Membangkitkan suatu bilangan acak p.
- Kromosom yang memiliki nilai yang lebih kecil

Tabel 5.
Perubahan nilai fitness pada setiap generasi

Maksimum generasi	Nilai fitness
100	7196.7307
200	7572.1212
500	7744.1662
1000	7834.2859

dari nilai p akan dipilih sebagai kandidat induk yang melakukan mutasi.

- 5) Proses mutasi
Proses mutasi menggunakan mutasi geser, yaitu dengan menggeser nilai gen pada gen yang terpilih untuk dimutasikan [13]. Nilai geser dapat bernilai positif atau negatif. Proses mutasi geser dilakukan berdasarkan probabilitas mutasi yang ditentukan oleh *user*.
- 6) Proses elitisme
Proses elitisme dilakukan untuk melestarikan kromosom-kromosom terbaik ke populasi yang baru, sehingga tidak kehilangan solusi terbaik. Sistem elitisme yang dipakai adalah sistem elitisme *ranking* [13]. Proses ini dilakukan dengan meranking nilai fitness populasi induk dan anak, kemudian dipilih sejumlah kromosom sebagai populasi baru

D. Hasil Algoritma Genetika

Hasil algoritma genetika berupa *centroid* masing-masing variabel yang akan digunakan sebagai *initial centroid* dari proses klusterisasi K-means. Data *centroid* merupakan matriks berukuran jumlah kluster x jumlah atribut. Data *initial centroid* ini selanjutnya akan digunakan untuk klusterisasi K-means. Tabel 2 berisi data *initial centroid* sebanyak 4 klaster yang diperoleh dari implemmentasi algoritma genetika dengan parameter ukuran populasi sebanyak 100, maksimum generasi sebanyak 100, dan probabilitas mutasi sebanyak 0.3.

E. Proses Klusterisasi K-means

Proses klusterisasi ini dilakukan dengan menggunakan algoritma simple K-means seperti pada umumnya, hanya saja *initial centroid*nya ditentukan yaitu berdasarkan proses algoritma genetika yang telah dilakukan sebelumnya. Proses klusterisasi menggunakan algoritma K-means juga dilakukan menggunakan MATLAB dengan ukuran *proximity euclidean*.

V. UJI COBA DAN ANALISIS HASIL

Proses uji coba dilakukan dalam dua tahap, yaitu uji coba parameter algoritma genetika dan uji coba klusterisasi. Uji coba parameter algoritma genetika dilakukan untuk mengetahui pengaruh parameter algoritma genetika terhadap keluaran yang dihasilkan. Uji coba klusterisasi dilakukan untuk menganalisis kelompok pengguna yang dihasilkan serta karakteristiknya.

1) Data Uji Coba

Data uji coba yang digunakan adalah data hasil praproses

yang telah dilakukan pada tahap sebelumnya. Data uji coba ini berisi data demografis serta behavioural pengguna Bapersip.

Tabel 2.
Nilai *initial centroid* tiap klaster

Gen-der	usia	Peker-jaan	Jml transaksi	agama	Psiko-logi	Lite ratur	Kom pu ter
1	44	7	3	1	3	2	3
2	68	3	2	3	1	4	3
1	59	4	5	2	5	5	0
1	24	7	3	2	2	1	1

2) Uji Parameter Algoritma Genetika

Uji coba parameter dilakukan pada parameter ukuran populasi, probabilitas mutasi, serta maksimum generasi. Dalam menentukan ukuran populasi sebaiknya berkisar antara 30 hingga 1000 [14]. Nilai untuk ukuran populasi dibuat berbeda yaitu mulai dari 100, 200, dan 400 populasi. Sedangkan nilai probabilitas mutasi mulai dari 0.01, 0.05, 0.1, dan 0.3. Dari hasil uji coba pada Tabel 3 diketahui nilai fitness terbaik yaitu 6982.3814 adalah pada parameter dengan ukuran populasi 200 dan probabilitas mutasi sebesar 0.1.

Uji coba parameter maksimum generasi nilainya dibuat berbeda mulai dari 100 hingga 1000 generasi. Dari hasil uji coba parameter maksimum generasi pada Tabel 5 dapat dianalisis bahwa nilai maksimum generasi berbanding lurus dengan nilai fitness yang dihasilkan.

3) Uji Penentuan Jumlah Cluster

Penentuan jumlah klaster dapat dilakukan berdasarkan pertimbangan (*judgement*) dari peneliti sendiri atau dapat juga menggunakan metode yang umum telah digunakan. Rousseuw & Kauffman (1987) memperkenalkan suatu metode untuk menentukan jumlah klaster yang maksimal, yaitu berdasarkan *silhouette measure* [15]. Nilai *silhouette* dari setiap titik adalah ukuran seberapa besar kemiripan suatu titik terhadap titik dalam klasternya sendiri dibandingkan dengan titik pada klaster yang lain. Nilai ini

Tabel 3.
Nilai fitness pada uji coba parameter ukuran populasi dan probabilitas mutasi

Ukuran populasi	Probabilitas mutasi	Nilai fitness
100	0.01	6036.2033
100	0.05	6481.5271
100	0.1	6650.6408
100	0.3	6953.3401
200	0.01	5980.8426
200	0.05	6754.8998
200	0.1	6982.3814
200	0.3	6912.1992
400	0.01	6318.1748
400	0.05	6823.7922
400	0.1	6902.6118
400	0.3	6952.5661

bernilai antara -1 hingga +1. Hasil uji coba pada Tabel 4 menunjukkan bahwa jumlah klaster optimal adalah sebanyak 4 klaster dengan rata-rata nilai *silhouette* terbesar yaitu sebesar 0.5142.

4) Uji Perbandingan Klasterisasi GA-Kmeans dan simple K-means

Dalam uji coba klasterisasi ini, dilakukan dengan cara membandingkan *total within cluster variation* antara metode pengelompokan yang telah dioptimasi dengan menggunakan algoritma genetika dan metode pengelompokan menggunakan algoritma K-means sederhana. Pada uji coba ditentukan jumlah klaster sesuai dengan yang diinginkan yaitu sebanyak 5 klaster dengan ukuran *proximity euclidean*.

5) Analisis Hasil

Setelah melalui serangkaian tahap uji coba, maka diperoleh kelompok klaster dengan karakteristik pengguna seperti yang ditunjukkan pada Tabel 7.

VI. KESIMPULAN/RINGKASAN

Pada bab terakhir ini diuraikan mengenai kesimpulan yang dapat ditarik dari pengerjaan Tugas Akhir ini. Selain itu juga diuraikan mengenai saran-saran yang dapat diperhatikan untuk pengembangan selanjutnya.

Tabel 5.

Rata-rata nilai *silhouette* pada jumlah klaster yang berbeda

Jumlah klaster	Rata-rata nilai <i>silhouette</i>
2 klaster	0.5094
3 klaster	0.4505
4 klaster	0.5142
5 klaster	0.3562
6 klaster	0.3902
7 klaster	0.3887

- 1) Praproses data yang digunakan untuk proses pengelompokan pengguna Bapersip terdiri dari lima tahap, yaitu tahap penentuan atribut yang digunakan untuk klasterisasi, tahap integrasi data, tahap pembersihan data, tahap transformasi data, dan tahap reduksi data.
- 2) Algoritma GA-Kmeans mampu menghasilkan pengelompokan dengan tingkat variasi di dalam klaster yang lebih baik dibandingkan dengan algoritma K-means sederhana. Dengan nilai *total*

Tabel 6.

Perbandingan nilai *within cluster variation* antara GA-Kmeans dan simple K-means

	GA-Kmeans	Simple K-means
Klaster 1	1095.718	3893.636
Klaster 2	914.8125	1587.205
Klaster 3	1609.717	1425.363
Klaster 4	3519.519	1332.67
Klaster 5	413.6923	657.4286
Total within cluster variation	7553.459	8896.303

within cluster yang dihasilkan adalah sebesar

7553.459 dibandingkan dengan nilai yang diperoleh K-means sebesar 8896.303.

- 3) Dari hasil pengelompokan pengguna Bapersip menggunakan algoritma GA-Kmeans, terbentuk

Tabel 7.

Analisis tiap klaster pengguna dan karakteristiknya

Klaster 1	Perempuan, rata-rata usia 24 tahun, mahasiswa, jumlah peminjaman 23-33 kali Rekomendasi buku : literatur
Klaster 2	Laki-laki, rata-rata usia 44 tahun, pekerjaan umum, jumlah peminjaman 23-33 kali Rekomendasi buku : agama, psikologi, kedokteran dan kesehatan
Klaster 3	Laki-laki, rata-rata usia 27 tahun, mahasiswa, jumlah peminjaman 23-33 kali Rekomendasi buku : psikologi, sosial politik, ekonomi
Klaster 4	Laki-laki, rata-rata usia 20 tahun, mahasiswa, jumlah peminjaman 23-33 kali Rekomendasi buku : bahasa dan ilmu alam

empat klaster pengguna dengan persebaran data pada klaster 1 terdapat 156 pengguna, klaster 2 terdapat 107 pengguna, klaster 3 terdapat 151 pengguna, dan klaster 4 terdapat 57 pengguna.

- 4) Pengukuran kinerja algoritma GA-Kmeans dilakukan dengan mengubah parameter algoritma genetika. Dapat disimpulkan pengaruh parameter terhadap performa algoritma GA-Kmeans adalah:

- Ukuran populasi: semakin besar ukuran populasi tidak berarti nilai fitness yang dihasilkan semakin besar pula. Hal ini dikarenakan sifat algoritma genetika yang membangkitkan ukuran populasinya secara acak.
- Maksimum generasi: jumlah maksimum generasi berbanding lurus dengan nilai solusi yang dihasilkan. Semakin besar jumlah generasi, maka semakin baik solusi yang dihasilkan. Hal ini dikarenakan semakin besar jumlah generasi maka ruang observasi akan semakin besar pula.
- Probabilitas mutasi: probabilitas mutasi juga berbanding lurus dengan nilai solusi yang dihasilkan. Probabilitas mutasi yang tinggi akan memungkinkan gen yang terpilih untuk dimutasi lebih banyak.

VII. SARAN

Pengelompokan pengguna Bapersip dengan hanya mempertimbangkan atribut demografis dan behavioural saja belum dapat memberikan karakteristik yang mendalam terhadap pengguna. Atribut yang digunakan untuk pengelompokan sebaiknya ditambahkan, tidak hanya menggunakan atribut demografis dan behavioural saja, namun juga atribut psikografis. Pengelompokan berdasarkan atribut psikografis merupakan pendekatan pengelompokan yang memisahkan pelanggan berdasarkan kepribadian, gaya hidup,

maupun opini pengguna [16].

UCAPAN TERIMA KASIH

Dengan ini, penulis hendak menyampaikan penghormatan dan terimakasih kepada pihak Bapersip Provinsi Jawa Timur serta pihak-pihak lain yang telah memberikan bantuan dan dukungan secara langsung maupun tidak langsung terhadap pelaksanaan tugas akhir ini.

DAFTAR PUSTAKA

- [1] Bapersip. *Profil*. Dipetik Maret 22, 2012, dari Badan Perpustakaan dan Kearsipan Provinsi Jawa Timur: <http://www.bapersip.jatimprov.go.id> (2008).
- [2] Humas Pemkot Surabaya. (2011, September 15). *Release*. Dipetik Maret 2012, 27, dari Humas Surabaya, Informasi Cepat, Tepat, dan Obyektif: <http://humas.surabaya.go.id/index.php?option=press&det=39>.
- [3] Bapersip. (2010, December 23). *Statistik Perpustakaan*. Dipetik Maret 22, 2012, dari Badan Perpustakaan dan Kearsipan Propinsi Jawa Timur PNRI: <http://bpad-jawatimur.pnri.go.id/?q=node/67>
- [4] Dispendukcapil. *Jumlah Penduduk Surabaya Tahun 2011*. Dipetik Maret 28, 2012, dari Dispendukcapil Surabaya: <http://dispendukcapil.surabaya.go.id/index.php/media-a-publik/statistik-penduduk/43-pergerakan-penduduk/167-jumlah-penduduk-surabaya-2011> (2012, Februari 20).
- [5] Kim, K. J., & Ahn, H. A recommender system using GA-Kmeans klasterisasi in an online shopping market. *Expert Systems with Applications* , 1201-1209 (2008)..
- [6] Guilano, S. J. Library system Journal. We have books and computers: libraries and the importance of marketing (2009).
- [7] de Saez, E. Marketing concepts for libraries and information services (2nd ed.). London: Facet Publishing (2002).
- [8] Wicaksono, H. Pasarkan Layanan Perpustakaan Anda dengan Tepat (2008).
- [9] Tan, P. N., Steinbach, M., & Kumar, V. *Introduction to Data Mining*. United States: Pearson Education Inc (2006).
- [10] Li, J. W., & Chen, P. H. The application of cluster analysis in Library System. *IEEE* , 907-910 (2008).
- [11] Al-Shboul, B., & Sung, H. M. Initializing K-Means Using Genetic Algorithm. *World Academy of Science, Engineering and Technology* , 114-118 (2010).
- [12] Lu, Y., Fotouhi, F., Deng, Y., & Brown, S. J. FGKA: A Fast Genetic K-means Clustering Algorithm. *ACM* (2004).
- [13] Martiana, E. *Penerapan Algoritma Genetika Cepat dalam Perbaikan Penentuan Titik Pusat Awal Klasterisasi berbasis K-means*. Surabaya: ITS (2006).
- [14] Suyanto. *Algoritma Genetika dalam Matlab*. Yogyakarta: Penerbit Andi (2005).
- [15] Rousseeuw, J. P., & Kauffman. Silhouettes: a Graphical Aid to the Interpretation and Validation of Cluster Analysis. *Computational and Applied Mathematics* (1987).
- [16] Browne, M. Communicating value through strategic alignment. (2011).