

# Sistem Informasi Geografis Pemetaan Persebaran Alumni dengan Analisa *Clustering*

I Kadek Dwi Gandika Supartha<sup>1</sup>, Made Sudarma<sup>2</sup>, Dewa Made Wiharta<sup>3</sup>

**Abstract**— Clustering is carried out on alumni data with occupational attributes, Commulative Achievement Index (GPA), length of study and length of work on the final project with the aim of knowing the characteristics of alumni data and their distribution. Using data from the STIKI Indonesia alumni in 2012-2014, the results are expected to be used to assist in policy making at STIKI-Indonesia. The clustering method used is Fuzzy C-Means (FCM) and cluster validity measurement using Modified Partition Coefficient (MPC) and Entropy (CE) Classification. The results of testing the black box method on the Geographic Information System of Alumni Distribution Mapping with Clustering Analysis shows that all modules in the system are functioning properly and based on the validity index of MPC and CE the most optimal number of clusters is 7 clusters and the best characteristics are cluster 1 with 49 members (9.3155% of the total alumni). This number is still very small compared to the total number of alumni.

**Keyword:** Fuzzy C Means, Clustering, Alumni, SIG

**Intisari**—Clustering dilakukan pada data alumni dengan atribut bidang pekerjaan, Indeks Prestasi Kumulatif (IPK), lama study dan lama pengerjaan tugas akhir dengan tujuan untuk mengetahui karakteristik dari data alumni dan persebarannya di dunia kerja. Menggunakan data alumni STIKI Indonesia tahun 2012-2014, hasilnya diharapkan dapat digunakan untuk membantu dalam pengambilan kebijakan di STIKI-Indonesia. Metode clustering yang digunakan yaitu Fuzzy C-Means (FCM) dan pengukuran validitas cluster menggunakan Modified Partition Coefficient (MPC) dan Classification Entropy (CE). Hasil pengujian metode black box pada Sistem Informasi Geografis Pemetaan Persebaran Alumni dengan Analisa Clustering didapatkan hasil bahwa semua modul dalam sistem telah berfungsi dengan baik dan berdasarkan indeks validitas MPC dan CE jumlah cluster yang paling optimal adalah 7 cluster dan yang memiliki karakteristik terbaik adalah cluster ke 1 dengan jumlah anggota 49 (9,3155% dari jumlah keseluruhan alumni). Jumlah ini masih sangat kecil dibandingkan dengan jumlah keseluruhan alumni.

**Kata Kunci**— Fuzzy C Means, Clustering, Alumni, SIG

## I. PENDAHULUAN

STMIK STIKOM Indonesia (STIKI Indonesia) adalah perguruan tinggi yang terletak di Denpasar Bali dan sampai saat ini memiliki alumni kurang lebih 700 orang. Data alumni tersebut kurang dimaksimalkan dan tidak diolah menjadi sebuah informasi yang mungkin berguna. STIKI Indonesia juga kurang mengetahui informasi persebaran alumni secara

spesifik di dunia kerja karena belum adanya sistem yang bisa melakukan hal tersebut. Untuk mengatasi permasalahan tersebut diatas dapat memanfaatkan teknologi Sistem Informasi Geografis (SIG) dan data mining. SIG dapat menyajikan informasi dalam bentuk grafis dan menggunakan peta untuk antar mukanya [1], sehingga persebaran alumni dapat diketahui secara spesifik. Data mining adalah proses untuk mengekstraksi atau mendeteksi pola-pola yang tersembunyi dari sebuah database besar dan salah satu metode dalam data mining adalah clustering[2].

Clustering akan dilakukan pada data alumni dari tahun 2012-2014 dan hasilnya ditampilkan dalam bentuk SIG berbasis web. Sebelumnya penelitian tentang clustering data alumni sudah pernah dilakukan, diantaranya yaitu penelitian tentang clustering data alumni Politeknik Negeri Semarang (IKA POLINES) dengan metode K-Means dan hasilnya disajikan dalam bentuk peta digital berbasis web. Mekanisme clustering didasarkan pada empat atribut yaitu: jenis perusahaan, klasifikasi jabatan, bidang kerja dan kompetensi prodi. Pengujian menunjukkan 51 alumni bekerja sesuai dengan kompetensinya, sedangkan 23 alumni dengan pekerjaan kurang sesuai dan 26 alumni tidak sesuai dengan kompetensinya. Hasil dari penelitian ini digunakan untuk membantu pengambilan keputusan apakah perlu mengadakan perubahan kurikulum atau tidak [3]. Penelitian selanjutnya menggunakan Fuzzy C-Means untuk melakukan clustering lulusan jurusan Matematika FMIPA Universitas Tangjungpura (UNTAN) dengan menggunakan variabel IPK dan lama studi. Berdasarkan clustering yang dilakukan diperoleh 4 cluster. Dari hasil keempat cluster tersebut diketahui bahwa pada cluster 4 memiliki anggota lulusan yang paling banyak dengan 33 lulusan. Cluster 4 terdiri dari lulusan dengan kisaran lama studi 5,91 tahun. Hal ini menunjukkan bahwa masih banyak mahasiswa jurusan Matematika di Fakultas MIPA UNTAN Pontianak yang menempuh lama studi lebih dari 10 semester atau 5 tahun. Hasil ini diharapkan dapat dijadikan sebagai bahan pertimbangan jurusan dalam meningkatkan IPK mahasiswa untuk menyelesaikan masa studinya dengan cepat[4]. Hasil cluster dari dua penelitian yang sudah dilakukan sebelumnya tidak mengalami proses validasi sehingga cluster yang dihasilkan belum tentu merupakan cluster yang paling optimal [5]. Pada penelitian ini akan dilakukan clustering dengan atribut: bidang pekerjaan, Indeks Prestasi Kumulatif (IPK), lama study dan lama pengerjaan tugas akhir. Metode yang akan digunakan dalam melakukan cluster data alumni adalah metode fuzzy clustering, yaitu dengan algoritma Fuzzy C-Means Clustering (FCM).

FCM dipilih karena logika fuzzy terus berkembang dan pada umumnya data tidak bisa dipisahkan secara tegas ke dalam kelompok tetapi memiliki kecenderungan yang dinyatakan dengan derajat keanggotaan yang bernilai antara 0 dan 1 terhadap pengelompokannya [6]. FCM memiliki tingkat p-ISSN:1693 – 2951; e-ISSN: 2503-2372

<sup>1</sup>Mahasiswa, Magister Teknik Elektro Universitas Udayana, Gedung Pascasarjana Jl. PB Sudirman Denpasar-Bali 80232 (telp/fax: 0361-239599; e-mail: [gandika23@gmail.com](mailto:gandika23@gmail.com))

<sup>2, 3</sup> Dosen, Magister Teknik Elektro Universitas Udayana, Gedung Pascasarjana Jl. PB Sudirman Denpasar-Bali 80232 (telp/fax: 0361-239599; e-mail: [sudarma@unud.ac.id](mailto:sudarma@unud.ac.id), [wiharta@unud.ac.id](mailto:wiharta@unud.ac.id))

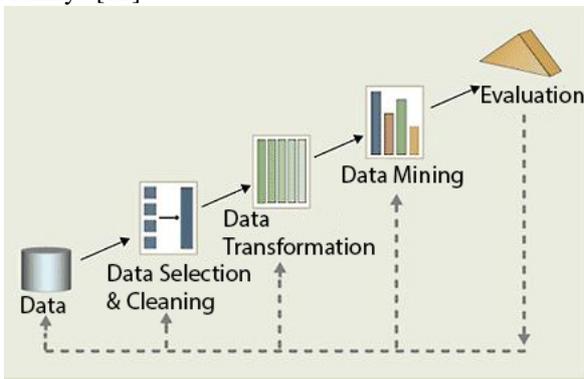


akurasi yang tinggi dan waktu komputasi yang cepat [7], [8], [9], juga karena dapat memberikan hasil yang halus dan cukup efektif untuk meningkatkan *homogenitas* tiap *cluster* yang dihasilkan [10], [11]. Untuk memvalidasi apakah *cluster* yang terbentuk merupakan jumlah *cluster* yang paling optimal, digunakan indeks pengukuran validitas *cluster*[5]. Metode yang akan digunakan untuk menguji validitas *cluster* pada penelitian ini adalah indeks *Modified Partition Coefficient* (MPC) (untuk mengukur *cluster* yang mengalami *overlap*) dan *Classification Entropy* (CE) (untuk mengukur tingkat keaburan/*fuzziness* dari partisi *cluster*). Hasil penelitian ini diharapkan dapat digunakan dalam membantu proses pengambilan kebijakan di STIKI Indonesia.

II. LANDASAN TEORI

A. Data Mining

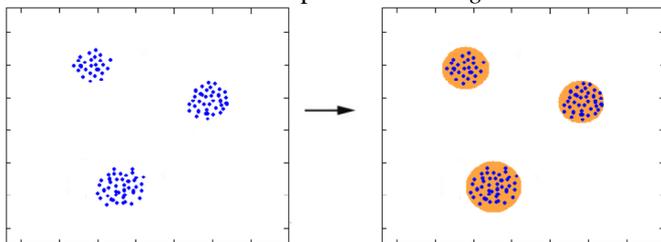
Data mining merupakan proses pencarian pola-pola yang menarik dan tersembunyi (*hidden pattern*) dari suatu kumpulan data yang berukuran besar yang tersimpan dalam suatu basis data, data *warehouse*, atau tempat penyimpanan data lainnya [12].



Gambar 1: Proses Data Mining

B. Metode Fuzzy C-Means Clustering

Metode *Fuzzy C-Means Clustering* pertama kali dikenalkan oleh Jim Bezdek pada tahun 1981 [13]. *Fuzzy C-Means* adalah salah satu teknik pengelompokan data yang mana keberadaan tiap titik data dalam suatu kelompok (*cluster*) ditentukan oleh derajat keanggotan. Metode *Fuzzy C-Means* termasuk metode *supervised clustering* dimana jumlah pusat *cluster* ditentukan di dalam proses *clustering*.

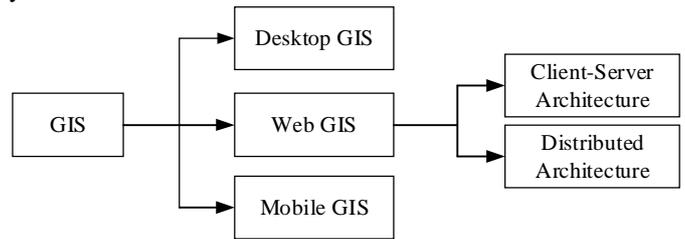


Gambar 2: Contoh Proses Clustering

C. Sistem Informasi Geografis

Sistem Informasi Geospasial atau juga dikenal sebagai Sistem Informasi Geografis (SIG) mulai dikenal pada awal tahun 1980-an. SIG adalah suatu sistem untuk memperoleh, menyimpan, menganalisis dan mengelola data spasial beserta

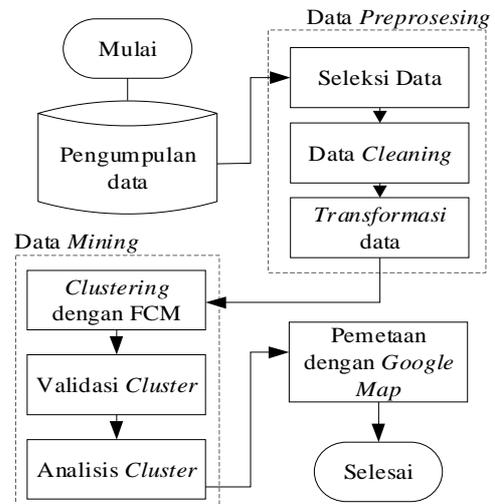
data atribut terkait yang secara keruangan direferensikan pada bumi[14]. Berdasarkan teknologi dan implementasinya, sistem informasi geografis dapat dikategorikan dalam tiga aplikasi yaitu



Gambar 3: Kategori Sistem Informasi Geografis

III. PERANCANGAN SISTEM

Gambaran umum penelitian dapat dilihat pada Gambar berikut ini



Gambar 4: Gambaran Umum Penelitian

A. Pengumpulan Data

Pada tahapan ini yang dilakukan yaitu mengumpulkan data yang akan digunakan dalam penelitian ini. Data tersebut berjenis data sekunder yaitu data alumni STIKI Indonesia. Atribut data alumni yang akan digunakan dalam penelitian ini selengkapny terdiri dari: Nomor Induk Mahasiswa (NIM), nama mahasiswa, jenis kelamin, alamat, no telepon, program studi, angkatan, tahun lulus, Indeks Prestasi Kumulatif (IPK), lama studi (dalam semester), lama pengerjaan tugas akhir (dalam semester), tempat bekerja, alamat tempat bekerja, kesesuaian bidang kerja.

B. Data Preprocessing

Dalam tahapan ini terdapat tiga langkah yaitu seleksi data, data *cleaning* dan *transformasi* data. Tujuan dari tahapan ini adalah untuk mengubah data mentah menjadi data yang berkualitas, seperti memperbaiki data yang *incomplete* (tidak lengkap), *noisy* (data berisi kesalahan nilai) dan data tidak konsisten.

1) *Seleksi Data*: Pada tahapan seleksi data yang akan dilakukan yaitu melakukan seleksi atribut yang akan digunakan. Tentunya tidak semua atribut dimasukkan dalam set data yang digunakan dalam proses data mining karena

hanya yang berperan sebagai referensi identifikasi yang akan dipilih. Dari semua atribut yang disebutkan diatas maka yang akan digunakan dalam proses data mining adalah atribut NIM, Nama mahasiswa, kesesuaian bidang kerja, Indeks Prestasi Kumulatif (IPK), lama masa studi, lama pengerjaan Tugas Akhir (TA).

2) *Data Cleaning*: Sebelum proses data mining dapat dilaksanakan, perlu dilakukan proses data cleaning. Tujuan dari proses ini adalah untuk memastikan kualitas data yang telah dipilih pada tahapan seleksi data. Misalkan penulisan jenis kelamin masih belum inkonsisten masih ada yang menggunakan p, l, laki-laki dan perempuan.

3) *Transformasi Data*: Pada tahapan ini data yang akan dilakukan yaitu data akan ditransformasi sehingga cocok digunakan dalam proses data mining. Seperti atribut bidang kerja berjenis data kategorikal sedangkan IPK, Lama Studi, Pengerjaan TA berjenis data numerik. Untuk dapat dilakukan proses data mining dalam hal ini dengan clustering metode Fuzzy C-Means maka keempat atribut tersebut harus diskalakan terlebih dahulu. Nilai skala yang akan digunakan yaitu mulai dengan skala terkecil dengan nilai 1 sampai dengan skala terbesar dengan nilai 10. Secara lengkap prosesnya dapat dilihat di bawah ini.

TABEL I  
PROSES PENSKALAAN

| Atribut       |                         |                              |                        | Nilai Skala |
|---------------|-------------------------|------------------------------|------------------------|-------------|
| Bidang Kerja  | Lama Studi              | IPK                          | Pengerjaan TA          |             |
| Sesuai        | Studi $\leq 8$          | IPK $> 3,5$                  | TA $\leq 1$            | 10          |
| Kurang Sesuai | $8 < \text{Studi} < 12$ | $2,75 < \text{IPK} \leq 3,5$ | $1 < \text{TA} \leq 2$ | 5           |
| Tidak Sesuai  | Studi $\geq 12$         | IPK $\leq 2,75$              | TA $> 2$               | 1           |

Pada Tabel I merupakan tabel penskalaan yang akan digunakan dalam penskalaan data. Tabel II memperlihatkan contoh atribut sebelum penskalaan dan Tabel III merupakan contoh atribut setelah penskalaan

TABEL III  
ATRIBUT SEBELUM PENSKALAAN

| NIM      | Nama          | Bidang Kerja | IPK | Lama Studi | Lama TA |
|----------|---------------|--------------|-----|------------|---------|
| 12101461 | I Made Andika | Sesuai       | 3,2 | 8          | 1       |
| 11101050 | Kadek Dimas   | Kurang       | 2,6 | 8          | 2       |
| 11101118 | Kadek Ari Y   | Sesuai       | 3,8 | 10         | 1       |
| 12101059 | Wina Yuana    | Tidak        | 3,3 | 12         | 2       |
| 12101377 | Ni Putu Rista | Sesuai       | 2,8 | 9          | 4       |

TABEL IIIII  
ATRIBUT SETELAH PENSKALAAN

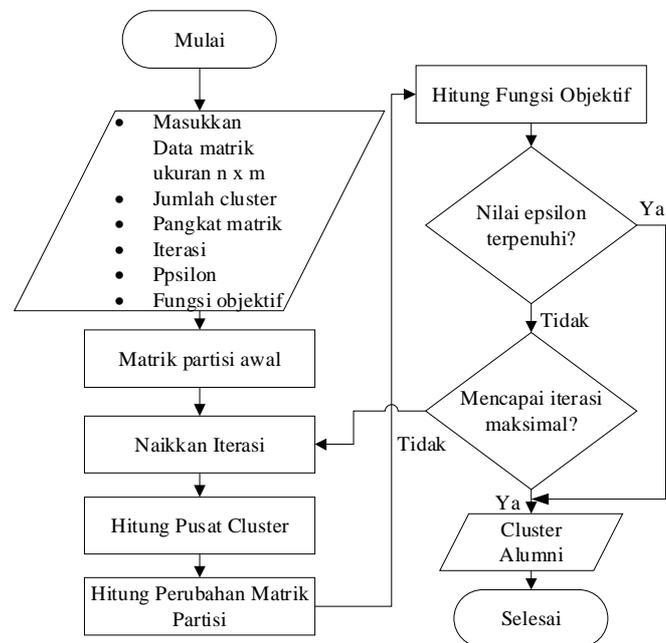
| NIM      | Nama          | Bidang Kerja | IPK | Lama Studi | Lama TA |
|----------|---------------|--------------|-----|------------|---------|
| 12101461 | I Made Andika | 10           | 5   | 10         | 10      |
| 11101050 | Kadek Dimas   | 5            | 1   | 10         | 5       |
| 11101118 | Kadek Ari Y   | 10           | 10  | 5          | 10      |
| 12101059 | Wina Yuana    | 1            | 5   | 1          | 5       |
| 12101377 | Ni Putu Rista | 10           | 5   | 5          | 1       |

I Kadek Dwi Gandika Supartha: SIG.Clustering.Data.Alumni

### C. Data Mining

Pada tahapan ini terdapat tiga proses yaitu melakukan *clustering* dengan *Fuzzy C-Means*, dan *Validasi Cluster Analisis Cluster*.

1) *Clustering dengan Fuzzy C-Means*: Setelah data melewati *preprocessing* maka selanjutnya yaitu melakukan *clustering* dengan *Fuzzy C-Mean*, proses algoritmanya dapat dilihat pada Gambar 5



Gambar 5: Clustering Fuzzy C Means

2) *Validasi Cluster*: Proses *clustering* diuji coba dengan berbagai nilai parameter. Dalam penelitian ini menggunakan validasi *cluster* indeks *Modified Partition Coefficient* (MPC) (untuk mengukur *cluster* yang mengalami *overlap*) dan *Classification Entropy* (CE) (untuk mengukur tingkat kekaburan/*fuzziness* dari partisi *cluster*). Validasi *cluster* memiliki tujuan untuk mengetahui apakah jumlah *cluster* yang terbentuk merupakan *cluster* yang paling bagus.

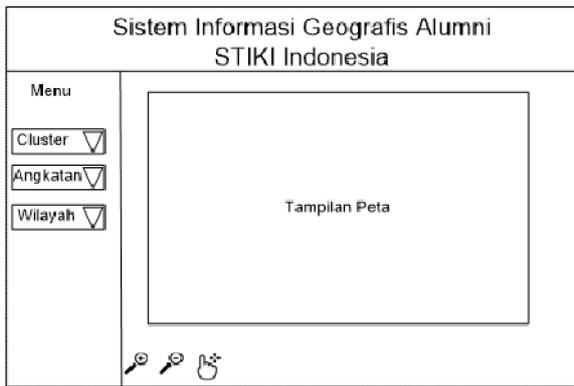
3) *Analisis Cluster*: Setelah melewati proses data *mining* dan sudah divalidasi maka tahapan selanjutnya yaitu analisis *cluster-cluster* yang telah terbentuk, hasilnya berupa karakteristik dari *profile* alumni STIKI Indonesia. Misalkan pada *cluster* pertama cenderung ditempati oleh alumni yang berkerja sesuai bidang studi dengan IPK diatas 3,6 memiliki lama studi kurang dari 8 semester dan lama pengerjaan tugas akhir selama satu semester.

### D. Pemetaan dengan Google Maps

Pada tahapan ini data hasil *clustering* akan diaplikasikan dalam Sistem Informasi Geografis dengan memanfaatkan peta online dari *Google Maps*, *database MySql* dan Bahasa pemrograman menggunakan PHP. Atribut yang akan dimunculkan dalam peta yaitu NIM, Nama mahasiswa, Jenis kelamin, program studi, Angkatan, Pekerjaan, Alamat tempat bekerja, *Cluster*. Dari aplikasi ini akan dapat diketahui



persebaran alumni STIKI Indonesia lengkap dengan *profile* alumni tersebut.



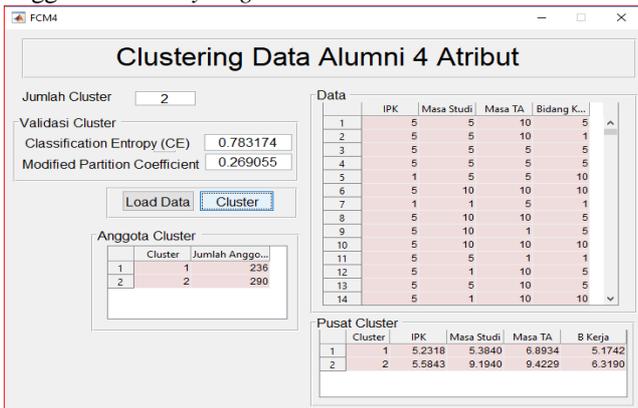
Gambar 6: Rancangan Sistem Informasi Geografis Alumni

IV. HASIL DAN PEMBAHASAN

Untuk proses *clustering* implementasi sistemnya menggunakan MATLAB versi 2017a dan untuk sistem informasi geografisnya menggunakan bahasa PHP, *database* MySql dan *Google Maps*

A. Implementasi *Clustering* dengan *Fuzzy C-Means*

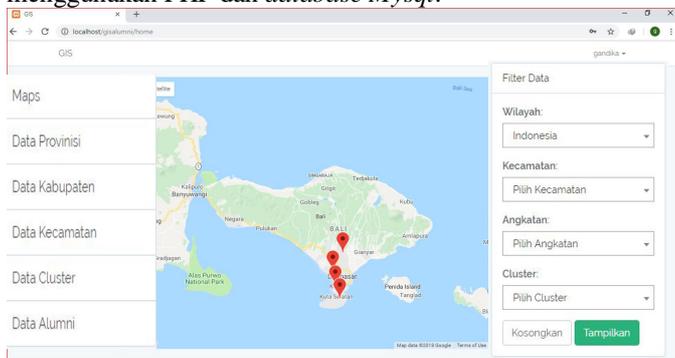
Implementasi *clustering* dengan algoritma *Fuzzy C-Means* dibuat berbasis GUI menggunakan MATLAB versi R2017a menggunakan *Fuzzy Logic Toolbox*.



Gambar 7: Form *Clustering* Alumni

B. Implementasi Sistem Informasi Geografis Alumni

Sistem Informasi Geografis Pemetaan berbasis *web* dengan analisa *clustering*. Sistem ini dibuat dengan menggunakan *google maps* sebagai peta dasarnya, *google map api* untuk dapat mengakses fitur dari *google maps*, bahasa pemrograman menggunakan PHP dan *database* *MySql*.



Gambar 8: Tampilan Sistem Informasi Geografis Alumni

Proses *clustering* data alumni menggunakan algoritma *Fuzzy C-Means*, parameter yang akan diubah yaitu jumlah *cluster*. Untuk proses *clustering* parameter awal akan ditentukan nilainya seperti berikut:

1. Jumlah *cluster* = ditentukan *user*
2. Pangkat = 2
3. Maksimum iterasi = 100
4. Error terkecil yang diharapkan = 10-5
5. Fungsi objektif awal = 0
6. Iterasi awal = 1

C. Pengujian *Clustering* Data Alumni

Proses pengujian *cluster* data alumni dilakukan dengan mengubah jumlah *cluster* dan mengamati nilai CE dan MPC untuk menentukan *cluster* terbaik.

1) *Jumlah Cluster* 2: Hasil proses *clustering* data alumni dengan empat atribut dengan jumlah *cluster* = 2 menggunakan algoritma FCM dengan nilai fungsi objektif 5381.491298, proses *clustering* berhenti pada iterasi ke 26. Dalam Tabel IV dapat dilihat no *cluster*, jumlah anggota tiap *cluster* dan koordinat titik pusat setiap *cluster* yang terbentuk.

TABEL IVV  
KOORDINAT TITIK PUSAT CLUSTER(JUMLAH CLUSTER = 2)

| No Cluster | Jumlah anggota | IPK    | Masa Studi | Masa TA | B Kerja |
|------------|----------------|--------|------------|---------|---------|
| 1          | 266            | 5,5655 | 9,3574     | 9,6011  | 4,1113  |
| 2          | 260            | 5,1845 | 5,1038     | 6,4459  | 4,1991  |

2) *Jumlah Cluster* 3: Hasil proses *clustering* data alumni dengan empat atribut dengan jumlah *cluster* = 3 menggunakan algoritma FCM dengan nilai fungsi objektif 3204.693481, proses *clustering* berhenti pada iterasi ke 26. Dalam Tabel V dapat dilihat no *cluster*, jumlah anggota tiap *cluster* dan koordinat titik pusat setiap *cluster* yang terbentuk.

TABEL V  
KOORDINAT TITIK PUSAT CLUSTER(JUMLAH CLUSTER = 3)

| No Cluster | Jumlah anggota | IPK    | Masa Studi | Masa TA | B Kerja |
|------------|----------------|--------|------------|---------|---------|
| 1          | 237            | 5,4129 | 9,8658     | 9,7747  | 4,3743  |
| 2          | 133            | 5,1864 | 5,1481     | 9,6407  | 4,1365  |
| 3          | 156            | 5,1672 | 5,3038     | 4,8026  | 4,4761  |

3) *Jumlah Cluster* 4: Hasil proses *clustering* data alumni dengan empat atribut dengan jumlah *cluster* = 4 menggunakan algoritma FCM dengan nilai fungsi objektif 2199.810226, proses *clustering* berhenti pada iterasi ke 23. Dalam Tabel VI dapat dilihat no *cluster*, jumlah anggota tiap *cluster* dan koordinat titik pusat setiap *cluster* yang terbentuk.

TABEL V  
KOORDINAT TITIK PUSAT CLUSTER(JUMLAH CLUSTER = 4)

| No Cluster | Jumlah anggota | IPK    | Masa Studi | Masa TA | B Kerja |
|------------|----------------|--------|------------|---------|---------|
| 1          | 127            | 5,1060 | 5,0755     | 4,7446  | 4,6670  |
| 2          | 133            | 5,0918 | 4,9150     | 9,7374  | 4,6327  |
| 3          | 75             | 5,4843 | 9,5557     | 9,5413  | 1,3411  |
| 4          | 191            | 5,2257 | 9,8954     | 9,8422  | 4,9623  |

4) *Jumlah Cluster* 5: Hasil proses *clustering* data alumni dengan empat atribut dengan jumlah *cluster* = 5 menggunakan

algoritma FCM dengan nilai fungsi objektif 1649.683051, proses *clustering* berhenti pada iterasi ke 22. Dalam Tabel VII dapat dilihat no *cluster*, jumlah anggota tiap *cluster* dan koordinat titik pusat setiap *cluster* yang terbentuk.

TABEL VII  
KOORDINAT TITIK PUSAT CLUSTER (JUMLAH CLUSTER = 5)

| No Cluster | Jumlah anggota | IPK    | Masa Studi | Masa TA | B Kerja |
|------------|----------------|--------|------------|---------|---------|
| 1          | 191            | 5,1699 | 9,9504     | 9,8816  | 4,9933  |
| 2          | 127            | 5,0901 | 5,0907     | 4,7179  | 4,8634  |
| 3          | 44             | 5,1486 | 4,9925     | 9,4946  | 1,2937  |
| 4          | 75             | 5,3490 | 9,8814     | 9,6760  | 1,2501  |
| 5          | 89             | 5,0804 | 4,9909     | 9,7989  | 5,0177  |

5) *Jumlah Cluster 6*: Hasil proses *clustering* data alumni dengan empat atribut dengan jumlah *cluster* = 6 menggunakan algoritma FCM dengan nilai fungsi objektif 1649.683051, proses *clustering* berhenti pada iterasi ke 22. Dalam Tabel VIII dapat dilihat no *cluster*, jumlah anggota tiap *cluster* dan koordinat titik pusat setiap *cluster* yang terbentuk.

TABEL VIII  
KOORDINAT TITIK PUSAT CLUSTER (JUMLAH CLUSTER = 6)

| No Cluster | Jumlah anggota | IPK    | Masa Studi | Masa TA | B Kerja |
|------------|----------------|--------|------------|---------|---------|
| 1          | 156            | 5,0131 | 9,9589     | 9,9062  | 4,9990  |
| 2          | 63             | 5,0685 | 9,9037     | 9,7491  | 1,0898  |
| 3          | 49             | 9,8287 | 9,8466     | 9,7245  | 4,7768  |
| 4          | 127            | 5,0280 | 5,0211     | 4,7223  | 4,8707  |
| 5          | 44             | 5,0211 | 4,8819     | 9,5939  | 1,1822  |
| 6          | 87             | 5,0015 | 4,9245     | 9,8353  | 5,0152  |

6) *Jumlah Cluster 7*: Hasil proses *clustering* data alumni dengan empat atribut dengan jumlah *cluster* = 7 menggunakan algoritma FCM dengan nilai fungsi objektif 905.862815, proses *clustering* berhenti pada iterasi ke 19. Dalam Tabel IX dapat dilihat no *cluster*, jumlah anggota tiap *cluster* dan koordinat titik pusat setiap *cluster* yang terbentuk.

TABEL IX  
KOORDINAT TITIK PUSAT CLUSTER (JUMLAH CLUSTER = 7)

| No Cluster | Jumlah anggota | IPK    | Masa Studi | Masa TA | B Kerja |
|------------|----------------|--------|------------|---------|---------|
| 1          | 49             | 9,8912 | 9,9013     | 9,8237  | 4,8259  |
| 2          | 43             | 5,0167 | 5,1824     | 4,7035  | 1,4452  |
| 3          | 96             | 5,0271 | 5,0200     | 4,8511  | 5,0608  |
| 4          | 44             | 5,0223 | 4,9093     | 9,8680  | 1,1317  |
| 5          | 51             | 5,0601 | 9,9529     | 9,8635  | 1,0649  |
| 6          | 156            | 5,0112 | 9,9743     | 9,9396  | 5,0057  |
| 7          | 87             | 5,0023 | 4,9409     | 9,9097  | 5,0344  |

Dari proses *clustering* yang telah dilakukan terhadap data alumni dengan 4 buah atribut (IPK, masa studi, masa TA dan bidang kerja) untuk mengetahui jumlah *cluster* yang paling bagus/optimal dapat dilihat dari nilai *indeks* validitas MPC

dan CE. Tabel X menunjukkan nilai nilai indeks MPC dan CE untuk setiap jumlah *cluster* yang diujikan pada algoritma *Fuzzy C-Means* (FCM). Jumlah *cluster* yang paling optimal adalah saat nilai indeks MPC semakin besar dan nilai indeks CE yang semakin kecil.

TABEL X  
NILAI INDEK MPC DAN CE UNTUK SETIAP CLUSTER

| Jumlah Cluster | Indeks CE | Indeks MPC |
|----------------|-----------|------------|
| 2              | 0.66186   | 0.39636    |
| 3              | 0.83373   | 0.5145     |
| 4              | 0.84786   | 0.59532    |
| 5              | 0.82182   | 0.65799    |
| 6              | 0.727483  | 0.724591   |
| 7              | 0.650175  | 0.762504   |

Dari Tabel X dapat dilihat bahwa nilai *indeks* CE yang terkecil muncul pada saat jumlah *cluster* = 7. Untuk nilai *indeks* MPC pada Tabel X nilai terbesar muncul pada saat jumlah *cluster* 7. Sehingga berdasarkan indeks CE dan *indeks* MPC, jumlah *cluster* yang paling bagus/optimal untuk data alumni dengan 4 buah atribut (IPK, masa sudi, masa TA dan bidang kerja) adalah 7 *cluster*.

Untuk dapat mengetahui karakteristik masing-masing *cluster* yang terbentuk (7 *cluster*), untuk koordinat titik pusat *cluster* dapat dibandingkan dengan nilai skala pada Tabel I proses penskalaan, misalkan nilai pusat *cluster* untuk bidang kerja pada sebuah *cluster* = 10 maka dapat diartikan bahwa rata-rata anggota pada *cluster* tersebut berkerja sesuai dengan bidang studinya. Berdasarkan Tabel IX dapat disimpulkan bahwa *cluster* yang memiliki karakteristik terbaik adalah *cluster* ke 1 yang jumlah anggotanya 49 (9,3155%) orang memiliki rata-rata IPK diatas 3,5, masa studi rata-rata kurang dari 8 semester, masa pengerjaan TA kurang dari 1 semester dan berkerja kurang sesuai dengan bidang studi. Dilihat dari jumlahnya alumni yang termasuk dalam *cluster* 1 jumlahnya masih sedikit dibandingkan dengan keseluruhan alumni.

#### D. Pengujian *Black Box Testing*

Pengujian *black box* pada Sistem Informasi Geografis Pemetaan Persebaran Alumni dengan Analisa *Clustering* memiliki fungsi untuk mengetahui adanya kesalahan pada sistem yang telah dibuat. Cara yang digunakan yaitu dengan menginputkan data tertentu dan mengamati output data yang didapatkan. Pengujian *black box*, yang diuji adalah masukan serta keluarannya. Input data tertentu akan dilakukan oleh user ke dalam aplikasi yang sedang berjalan dan akan dilakukan pengamatan pada outputnya dengan tujuan untuk mengetahui kesesuaian fungsi dari aplikasi yang telah dibuat. Apakah telah sesuai dengan perancangan yang telah ditentukan sebelumnya.

Berikut ini merupakan hasil pengujian *black box* testing dengan berbagai skenario pada semua halaman *web*. Hasil pengujian disajikan dalam Tabel XI



TABEL XI  
HASIL PENGUJIAN BLACK BOX

| No | Komponen                     | Skenario  | Input                                  | Output diharapkan   | Output hasil eksekusi   |
|----|------------------------------|---|--|---|---|
| 1  | Authentifikasi administrator | Seorang user memiliki username admin dan password admin | username = admin,<br>password = admin  | Masuk ke tampilan utama SIG   | Berhasil Masuk ke tampilan utama SIG  |
| 2  | Authentifikasi administrator | Seorang user memiliki username admin dan password admin | username = admin,<br>password = adminn | Gagal Masuk ke tampilan utama SIG dan user diharuskan mengisi ulang username dan password | Gagal Masuk ke tampilan utama SIG dan user diharuskan mengisi ulang username dan password |
| 3  | Pengelolaan data provinsi    | Seorang user menginputkan data provinsi                 | Data provinsi                          | Berhasil ditambahkan  | Berhasil ditambahkan  |
| 4  | Pengelolaan data provinsi    | Seorang user menginputkan data provinsi                 | Data kosong                            | Data gagal ditambahkan  | Berhasil ditambahkan  |
| 5  | Pengelolaan data kabupaten   | Seorang user menginputkan data kabupaten                | Data kosong                            | Data gagal ditambahkan  | Data gagal ditambahkan  |
| 6  | Pengelolaan data kabupaten   | Seorang user menginputkan data kabupaten                | Data kabupaten                         | Berhasil ditambahkan  | Berhasil ditambahkan  |
| 7  | Pengelolaan data Kecamatan   | Seorang user menginputkan data Kecamatan                | Data Kecamatan                         | Berhasil ditambahkan  | Berhasil ditambahkan  |
| 8  | Pengelolaan data Kecamatan   | Seorang user menginputkan data Kecamatan                | Data kosong                            | Data gagal ditambahkan  | Data gagal ditambahkan  |
| 9  | Pengelolaan data Cluster     | Seorang user menginputkan data Cluster                  | Data Cluster                           | Berhasil ditambahkan  | Berhasil ditambahkan  |
| 10 | Pengelolaan data Cluster     | Seorang user menginputkan data Cluster                  | Data kosong                            | Data gagal ditambahkan  | Data gagal ditambahkan  |
| 11 | Pengelolaan data Alumni      | Seorang user menginputkan data Alumni                   | Data Alumni                            | Berhasil ditambahkan  | Berhasil ditambahkan  |
| 12 | Pengelolaan data Alumni      | Seorang user menginputkan data Alumni                   | Data kosong                            | Data gagal ditambahkan  | Data gagal ditambahkan  |

## V. KESIMPULAN

Berdasarkan hasil pengujian yang telah dilakukan dapat disimpulkan beberapa hal sebagai berikut:

- 1) Untuk mengukur tingkat validitas FCM dalam melakukan *clustering* terhadap data alumni dapat menggunakan *indeks* MPC dan CE. Pada *clustering* alumni dengan empat atribut, jumlah *cluster* yang paling optimal adalah 7 *cluster* dan *cluster* yang memiliki karakteristik terbaik adalah *cluster* ke 1 yang jumlah anggotanya 49 (9,3155% dari jumlah keseluruhan alumni). Dari hasil *clustering* data alumni dapat dilihat bahwa jumlah alumni STIKI Indonesia yang memiliki karakteristik ideal (IPK bagus, masa studi kurang dari 8 semester, pengerjaan TA kurang dari 1 semester dan berkerja sesuai dengan bidang studi) jumlahnya masih

sangat sedikit dibanding dengan total keseluruhan alumni.

- 2) Berdasarkan pengujian menggunakan metode *black box testing* pada Sistem Informasi Geografis Pemetaan Persebaran Alumni dengan Analisa *Clustering* didapatkan hasil bahwa semua modul dalam sistem telah berfungsi dengan baik.

## REFERENSI

- [1] Riyanto, *Membuat Sendiri Aplikasi Mobile GIS Platform Java ME, Blackberry dan Android*, Jogjakarta: Andi, 2010.
- [2] M. J. A. Berry & G. S. Linoff, *Data mining techniques second edition-for marketing, sales, and customer relationship management*, Wiley, 2004.
- [3] S. Handoko, "Sistem Informasi Geografis Berbasis Web Untuk pemetaan Alumni Menggunakan Metode K-Means", tesis, Semarang, Program Pascasarjana, Universitas Diponegoro, 2012.
- [4] L. C. Simbolon, N. Kusumastuti, B. Irawan, "Clustering Lulusan mahasiswa Matematika FMIPA Untan Pontianak Menggunakan

- Algoritma Fuzzy C-Means”, *Buletin Ilmiah Mat. Stat. dan Terapannya (Bimaster)*, Volume 02, No.1 ,2013.
- [5] K. Wu, M. Yang, “A Cluster Validity Index for Fuzzy Clustering”, *Pattern Recognition Letter*, vol.26, pp. 1275-1291, Juli 2005.
- [6] F. Hoppner dan F. Klowon , “Learning Fuzzy System - An Objective Function-Approach”, *Mathware & Soft Computing*, vol.11, pp. 143-162, 2004.
- [7] K. Hammouda, F. Karaay, *A Comparative Study of Data Clustering Techniques*. University of Waterloo, Ontario, Canada, 2000.
- [8] R. Hadi, I. K. G. D. Putra, I. N. S. Kumara, “Penentuan Kompetensi Mahasiswa dengan Algoritma Genetik dan Metode Fuzzy C-Means”, *Majalah Ilmiah Teknologi Elektro*, vol. 15, n. 2, p. 101-106, Desember 2016.
- [9] A. A. G. B. Ariana, I. K. G. D. Putra, Linawati,” Perbandingan Metode SOM/Kohonen dengan ART 2 pada Data Mining Perusahaan Retail”, *Majalah Ilmiah Teknologi Elektro*, vol. 16, n. 2, p. 55-59, Agustus 2017.
- [10] A.I. Shihab, “Fuzzy Clustering Algorithm and Their Application to Medical Image Analysis”. Dissertation, University of London, London, 2000.
- [11] I. M. B. Adnyana, I. K. D. Putra, I. P. A. Bayupati, “Segmentasi Citra Berbasis Clustering Menggunakan Algoritma Fuzzy C-Means”, *Majalah Ilmiah Teknologi Elektro*, vol. 14, n. 1, Juni 2015.
- [12] P. N. Tan, M. Steinbach, & V. Kumar, *Introduction to Data Mining*. Pearson Education, Inc, 2016.
- [13] A. K. Jain, M. N. Murthy & P. J. Flynn, “Data Clustering: A Review”, *ACM Computing Surveys*, vol. 31, No. 3, 1999.
- [14] E. Prahasta, *Sistem Informasi Geografis: Aplikasi Pemrograman MapInfo*, CV. Informatika: Bandung, 2005.



{ Halaman ini sengaja dikosongkan }