

## **BAB II**

### **LANDASAN TEORI**

#### **2.1 Data Mining**

Tan (2006) mendefinisikan *data mining* sebagai proses untuk mendapatkan informasi yang berguna dari gudang basis data yang besar. *Data mining* juga dapat diartikan sebagai pengekstrakan informasi baru yang diambil dari bongkahan data besar yang membantu dalam pengambilan keputusan. Istilah *data mining* kadang disebut juga *knowledge discovery*. Selanjutnya perbedaan *data mining* dan *data warehouse* adalah *data mining* merupakan bidang yang sepenuhnya menggunakan apa yang sepenuhnya digunakan oleh *data warehouse*, bersama dengan bidang yang menangani masalah pelaporan dan manajemen data. Sementara *data warehouse* sendiri bertugas untuk menarik/meng-*query* data dari basis data mentah untuk memberikan hasil data yang nantinya digunakan oleh bidang yang menangani manajemen, pelaporan, dan *data mining*. Dengan *data mining* inilah, penggalian informasi baru dapat dilakukan dengan bekal data mentah yang diberikan oleh *data warehouse*. Hasil yang diberikan oleh ketiga bidang tersebut berguna untuk mendukung aktivitas bisnis cerdas (*business intellegent*).

Pekerjaan yang berkaitan dengan *data mining* dibagi menjadi 4 (empat) kelompok, yaitu (Prasetyo, E, 2012) :

1. *Prediction Modelling* (model prediksi), berkaitan dengan sebuah model yang dapat melakukan pemetaan dari setiap himpunan variabel ke setiap targetnya, kemudian menggunakan model tersebut untuk memberikan nilai target pada himpunan baru yang didapat.
2. *Cluster Analysis* (analisis kelompok), pengelompokan data-data kedalam sejumlah kelompok (*cluster*) berdasarkan kesamaan karakteristik masing-masing data pada kelompok-kelompok yang ada.

3. *Association Analysis* (analisis asosiasi), digunakan untuk menemukan pola yang menggambarkan kekuatan fitur dalam data dengan tujuan untuk menemukan pola yang menarik dengan cara yang efisien.
4. *Anomaly Detection* (deteksi anomali), berkaitan dengan pengamatan sebuah data dari sejumlah data yang secara signifikan mempunyai karakteristik yang berbeda dari sisa data yang lain.

## 2.2 PHP (Personal Home Page)

PHP adalah akronim dari *Hypertext Preprocessor*, yaitu suatu bahasa pemrograman berbasis kode – kode ( *script* ) yang digunakan untuk mengolah suatu data dan mengirimkannya kembali ke *web browser* menjadi kode HTML. PHP diciptakan pertama kali oleh Ramus Lerdorf pada tahun 1994. Awalnya, PHP digunakan untuk mencatat jumlah serta untuk mengetahui siapa saja pengunjung pada *homepage*-nya. Rasmus Lerdorf adalah salah seorang pendukung *Open Source*.

Tahun 2004 bulan juli dirilis PHP 5 dengan inti *Zend Engine 2.0*. PHP 5 adalah versi PHP terbaru yang mendukung penuh *object-oriented programming* (OOP), integrasi XML, mendukung semua eksistensi terbaru MySQL, pengembangan *web service* dengan SOAP dan REST, serta ratusan peningkatan lainnya dibandingkan dengan versi sebelumnya PHP 4.0. PHP dapat digunakan pada semua sistem operasi, antara lain Linux, Unix, Microsoft Windows, Mac OS X, RISC OS. PHP juga mendukung banyak *Web Server*.(Harahap, L, 2009)

## 2.3 Apache

Server HTTP Apache atau Server Web/WWW Apache adalah server web yang dapat dijalankan di banyak sistem operasi (Unix, BSD, Linux, Microsoft Windows dan Novell Netware serta platform lainnya) yang berguna untuk melayani dan memfungsikan situs web. Protokol yang digunakan untuk melayani fasilitas web/www ini menggunakan HTTP. Apache memiliki fitur-fitur canggih seperti pesan kesalahan yang dapat

dikonfigur, autentikasi berbasis basis data dan lain-lain. Apache juga didukung oleh sejumlah antarmuka pengguna berbasis grafik (GUI) yang memungkinkan penanganan server menjadi mudah.

#### **2.4 SQLyog Comunity**

SQLyog adalah salah satu tool administrasi untuk database MySQL. Jika kita biasanya menggunakan PhpMyAdmin yang include di dalam aplikasi Xampp untuk melakukan administrasi database, SQLyog adalah aplikasi alternatif untuk melakukan proses administrasi database MySQL. Banyak fitur yang disediakan oleh SQLyog yang tidak disediakan oleh PhpMyAdmin maupun tool administrasi database lainnya seperti MySQLQueryBrowser. Dengan SQLyog kita dapat membuat Store Procedure, Function maupun Trigger dengan mudah.

#### **2.5 Dreamweaver CS3**

Dreamweaver CS3 adalah sebuah HTML editor profesional untuk mendesign web secara visual dan mengelolah situs atau halaman web. Dreamweaver CS3 ini terdapat beberapa kemampuan bukan hanya untuk mendesign web tapi juga untuk menyunting kode serta membuat aplikasi web dengan menggunakan berbagai bahasa pemrograman Web, antara lain : JPS, PHP, ASP dan ColdFusion. Fasilitas dalam dreamweaver CS3 memiliki peningkatan mengenai toolbar dimana dreamweaver CS3 dapat digunakan untuk memodifikasi tampilan-tampilan pada toolbar atau menambahkan fungsi yang baru. Ruang kerja Dreamweaver CS3 memiliki atribut-atribut yang memberikan fasilitas dan ruang untuk membuat para user berkreasi.

#### **2.6 Rekayasa Perangkat Lunak**

Rekayasa Perangkat Lunak (RPL) telah dikenal sebagai bidang ilmu sejak tahun 1960-an. RPL yaitu ilmu yang mempelajari tentang bagaimana memanfaatkan sebuah perangkat keras (*hardware*) dan perangkat lunak

(*software*) untuk dijadikan suatu project yang bisa digunakan dengan baik serta ilmu yang mempelajari tentang tahapan dalam pembuatan suatu project. Rekayasa Perangkat Lunak juga pada dasarnya terdiri dari komponen-komponen perangkat lunak yang bertugas membuat suatu desain aplikasi yang ada dilingkungan tugas atau pekerjaan serta membuat deadline berupa pendahuluan, perumusan masalah, analisis, desain, dan kesimpulan (Ladjamuddin B, Al-Bahra, 2006).

Rekayasa yang dilakukan terhadap suatu perangkat lunak memiliki beberapa tujuan diantaranya adalah (Ladjamuddin B, Al-Bahra, 2006) :

1. Biaya produksi rendah.
2. Kinerja program yang tinggi.
3. Portabilitas yang tinggi
4. Biaya perawatan yang rendah
5. Keandalan sistem dan penyerahan tepat waktu

## 2.7 Sistem Cerdas

Menurut Lofti A.Zadeh sistem cerdas adalah sistem yang mampu melakukan penalaran dalam ketidakpastian (*uncertainty*), ketidaktepatan (*imprecision*), kebenaran sebagian (*partial truth*), pendekatan (*approximation*). Sistem cerdas sendiri merupakan suatu sistem yang terinspirasi dari kecerdasan manusia. Kecerdasan yang diciptakan tersebut diinputkan ke dalam sebuah mesin dengan tujuan pekerjaan-pekerjaan tertentu seperti pekerjaan yang dilakukan oleh seorang manusia. Sistem cerdas sendiri mempunyai tiga konsep, yaitu (Latifah, 2013) :

### 1. *Artificial Neural Network*

Konsep tersebut mengenai kerja dari sel saraf dari otak manusia. Untuk mendapatkan mesin dengan cara kerja sel saraf yang sama dengan manusia, diperlukan perumusan persamaan model matematik. Model persamaan tersebut dituliskan pada program-program seperti Visual Basic, Delphi, dan Matlab.

## 2. *Artificial Intelligence*

Konsep ini membahas mengenai sebuah mesin yang dapat mengambil keputusan tanpa diperintah oleh seorang manusia. Sehingga diperlukan logika berpikir mesin yang sama dengan manusia.

## 3. *Fuzzy Logic*

merupakan logika yang mempunyai perasaan dan toleransi seperti manusia. Pada logika konvensional sebelumnya, hanya dikenal option “YA” atau “TIDAK”. Sedangkan logika samar dimaksudkan bahwa logika yang bernilai samar yang benar terlihat tidak benar dan salah bernilai tidak samar. Sehingga dalam kerjanya, logika tersebut menggunakan perasaan. Contohnya, sebuah robot berlogika samar dapat mengerti arti kata sedikit atau banyak dengan menggunakan logika yang ia miliki.

## 2.8 Clustering

Clustering merupakan sebuah teknik penggalian data yang bersifat *unsupervised*, karena tidak ada satu atributpun yang digunakan untuk memandu proses pembelajaran, jadi seluruh atribut input diperlakukan sama. Pengklasteran merupakan satu dari sekian banyak fungsi proses data mining untuk menemukan kelompok atau identifikasi kelompok obyek yang hampir sama. Pengelompokan data ini didasarkan pada kesamaan karakter atau kriteria dari data-data yang dianalisis. Data-data yang ada dalam cluster yang sama memiliki karakter atau kriteria yang sama, sementara data-data yang berada dalam cluster yang berbeda juga memiliki karakter atau kriteria yang berbeda (Agusta, Y, 2007).

Tujuan pengelompokan (*clustering*) data dapat dibedakan menjadi dua, yaitu pengelompokan untuk pemahaman dan pengelompokan untuk penggunaan dan mencari prototipe kelompok yang paling representative terhadap data, memberikan abstraksi dari setiap objek data dalam kelompok dimana sebuah data terletak di dalamnya (Prasetyo, E, 2012). Menurut Tan (2006) Metode klasterisasi secara umum dapat dibagi menjadi dua yaitu :

### 1. *Hierarchical clustering*

Pengelompokkan data melalui suatu bagan yang berupa hirarki, dimana terdapat penggabungan dua grup yang terdekat disetiap iterasinya ataupun pembagian dari seluruh set data kedalam klaster-klaster.

### 2. *Partitional clustering*

Pengelompokkan ke dalam sejumlah klaster tanpa adanya struktur hirarki antara satu dengan yang lainnya. Pada metode ini setiap klaster memiliki titik pusat klaster (*centroid*) dan secara umum metode ini memiliki fungsi tujuan yaitu meminimumkan jarak (*dissimilarity*) dari seluruh data ke pusat klaster masing-masing.

## **2.9 Buku Panduan Akademik UMG 2013/2014**

Buku panduan akademik merupakan buku yang menjadi petunjuk bagi setiap mahasiswa selama melaksanakan perkuliahan. Buku panduan akademik bisa jadi berbeda untuk setiap universitas. Buku panduan akademik yang akan digunakan dalam penelitian ini merupakan buku panduan akademik Universitas Muhammadiyah Gresik tahun 2013/2014. Buku ini berisi informasi umum, peraturan akademik, prosedur-prosedur kerja dan kurikulum untuk setiap program studi di Universitas Muhammadiyah Gresik. Selain itu buku panduan akademik juga memberikan informasi tentang penjurusan suatu mahasiswa terhadap fakultas yang dipilihnya.

## **2.10 Logika Fuzzy**

*Logika Fuzzy* merupakan suatu komponen pembentuk *softcomputing*. Logika fuzzy pertama kali diperkenalkan oleh Lotfi A. Zadeh pada tahun 1965. Dasar logika fuzzy adalah teori himpunan fuzzy. Pada teori himpunan fuzzy, peranan derajat keanggotaan sebagai penentu keberadaan elemen dalam suatu himpunan sangatlah penting. Nilai keanggotaan atau derajat keanggotaan atau *membership function* menjadi ciri utama dari penalaran dengan logika fuzzy tersebut (Kusumadewi, S ; Purnomo, H, 2010).

Fungsi keanggotaan (*membership function*) adalah suatu kurva yang menunjukkan pemetaan titik-titik *input* data ke dalam nilai keanggotaan (sering juga disebut derajat keanggotaan) yang memiliki interval antara 0 sampai 1 (Kusumadewi, S ; Purnomo, H, 2010).

Menurut Cox (1994), ada beberapa alasan mengapa orang menggunakan logika fuzzy, antara lain (Kusumadewi, S ; Purnomo, H, 2010):

1. Konsep logika fuzzy mudah dimengerti. Karena logika fuzzy menggunakan dasar teori himpunan, maka konsep matematis yang mendasari penalaran fuzzy tersebut cukup mudah untuk dimengerti.
2. Logika fuzzy sangat fleksibel, artinya mampu beradaptasi dengan perubahan-perubahan, dan ketidakpastian yang menyertai permasalahan.
3. Logika fuzzy memiliki toleransi terhadap data yang tidak tepat. Jika diberikan sekelompok data yang cukup homogen, dan kemudian ada beberapa data yang “*ekskusif*”, maka logika fuzzy memiliki kemampuan untuk menangani data eksklusif tersebut.
4. Logika fuzzy mampu memodelkan fungsi-fungsi nonlinier yang sangat kompleks.
5. Logika fuzzy dapat membangun dan mengaplikasikan pengalaman-pengalaman para pakar secara langsung tanpa harus melalui proses pelatihan.
6. Logika fuzzy dapat bekerjasama dengan teknik-teknik kendali secara konvensional. Hal ini umumnya terjadi pada aplikasi di bidang teknik mesin atau teknik elektro.
7. Logika fuzzy didasarkan pada bahasa alami. Logika fuzzy menggunakan bahasa sehari-hari sehingga mudah dimengerti.

#### ***A. Himpunan Fuzzy***

Pada himpunan tegas (*crisp*), nilai keanggotaan suatu item  $x$  dalam suatu himpunan  $A$ , yang sering ditulis dengan  $\mu_A(x)$ , memiliki dua kemungkinan, yaitu (Kusumadewi, S ; Purnomo, H, 2010) :

1. Satu (1), yang berarti bahwa suatu item menjadi anggota dalam suatu himpunan, atau
2. Nol (0), yang berarti bahwa suatu item tidak menjadi anggota dalam suatu himpunan.

**Contoh :**

Jika diketahui :

$S = \{1, 2, 3, 4, 5, 6\}$  adalah semesta pembicaraan

$A = \{1, 2, 3\}$

$B = \{3, 4, 5\}$

Bisa dikatakan bahwa :

- Nilai keanggotaan 2 pada himpunan A,  $\mu_A(2)=1$ , karena  $2 \in A$ .
- Nilai keanggotaan 3 pada himpunan A,  $\mu_A(3)=1$ , karena  $3 \in A$ .
- Nilai keanggotaan 4 pada himpunan A,  $\mu_A(4)=0$ , karena  $4 \notin A$ .
- Nilai keanggotaan 2 pada himpunan B,  $\mu_B(2)=0$ , karena  $2 \notin B$ .
- Nilai keanggotaan 3 pada himpunan B,  $\mu_B(3)=1$ , karena  $3 \in B$ .

### 2.11 Fuzzy C-Means (FCM)

Pengelompokan dengan metode *Fuzzy C-Means* (FCM) didasarkan pada teori logika fuzzy. Teori ini pertama kali diperkenalkan oleh Lotfi Zadeh (1965) dengan nama himpunan fuzzy (*fuzzy set*). Dalam teori fuzzy, keanggotaan sebuah data tidak diberi nilai secara tegas dengan nilai 1 (menjadi anggota) dan 0 (tidak menjadi anggota), melainkan dengan suatu nilai derajat keanggotaan yang dijangkau nilai 0 sampai 1. Nilai keanggotaan suatu data dalam sebuah himpunan menjadi 0 ketika data sama sekali bukan anggota, dan 1 ketika data menjadi anggota secara penuh dalam suatu himpunan. Umumnya nilai keanggotaannya antara 0 dan 1. Semakin tinggi nilai keanggotaannya, semakin tinggi derajat keanggotaannya, dan semakin kecil nilai keanggotaannya, semakin rendah derajat keanggotaannya. Dikaitkan dengan *K-Means*, sebenarnya FCM merupakan versi fuzzy dari *K-Means*

dengan beberapa modifikasi yang membedakannya dengan *K-Mean* (Prasetyo, E, 2012).

Konsep dasar FCM, pertama kali adalah menentukan pusat *cluster*, yang akan menandai lokasi rata-rata untuk tiap-tiap *cluster*. Pada kondisi awal, pusat *cluster* ini masih belum akurat. Tiap-tiap titik data memiliki derajat keanggotaan untuk tiap-tiap *cluster*. Dengan cara memperbaiki pusat *cluster* dan derajat keanggotaan tiap-tiap titik data secara berulang, maka akan dapat dilihat bahwa pusat *cluster* akan bergerak menuju lokasi yang tepat. Perulangan ini didasarkan pada minimisasi fungsi objektif yang menggambarkan jarak dari titik data yang diberikan pada pusat *cluster* yang terbobot oleh derajat keanggotaan titik data tersebut. *Output* dari FCM bukan merupakan *fuzzy inference system*, namun merupakan deretan pusat *cluster* dan beberapa derajat keanggotaan untuk tiap-tiap titik data. Informasi ini dapat digunakan untuk membangun *fuzzy inference system* (Kusumadewi, S ; Purnomo, H, 2010).

Asumsikan ada sejumlah data dalam data set ( $X$ ) yang berisi  $n$  data :  $x_1, x_2, \dots, x_n$  dinotasikan  $X = \{ x_1, x_2, \dots, x_n \}$ , dimana setiap data mempunyai fitur  $r$  dimensi :  $x_{i1}, x_{i2}, \dots, x_{ir}$ , dinotasikan  $x_i = \{ x_{i1}, x_{i2}, \dots, x_{ir} \}$ . Ada sejumlah cluster  $C$  dengan centroid :  $c_1, c_2, \dots, c_k$ , dimana  $k$  adalah jumlah cluster. Setiap data mempunyai derajat keanggotaan pada setiap cluster, dinyatakan dengan  $u_{ij}$ , dengan nilai diantara 0 dan 1,  $i$  menyatakan data  $x_i$  dan  $j$  menyatakan cluster  $c_j$ . Jumlah nilai derajat keanggotaan setiap data  $x_i$  selalu sama dengan 1, diformulasikan pada persamaan (2.1) (Prasetyo, E, 2012).

$$\sum_{j=1}^k u_{ij} = 1 \quad \dots\dots\dots (2.1)$$

Untuk setiap cluster  $c_j$ , berisi paling sedikit satu data dengan nilai keanggotaan tidak nol, tapi tidak berisi derajat satu pada semua data, diformulasikan pada persamaan (2.2).

$$0 < \sum_{i=1}^n U_{ij} < n \quad \dots\dots\dots (2.2)$$

Sepeti halnya teori himpunan fuzzy yang menyatakan bahwa suatu data bisa menjadi anggota dibeberapa himpunan yang dinyatakan dengan nilai derajat keanggotaan pada setiap himpunan dalam FCM, setiap data juga menjadi anggota pada setiap cluster dengan derajat keanggotaan  $u_{ij}$  (Prasetyo,E, 2012). Nilai derajat keanggotaan data  $x_i$  pada cluster  $c_j$ , diformulasikan pada persamaan (2.3).

$$U_{ij} = \frac{D(x_i, c_j)^{-2/w-1}}{\sum_{j=1}^k D(x_i, c_j)^{-2/w-1}} \quad \text{dengan } D=d(x_i - c_j) = \sum_{l=1}^k D(x_{il} - c_{jl})^{1/2} \quad \dots\dots\dots(2.3)$$

Parameter  $c_j$  adalah centroid cluster ke- $j$ ,  $D()$  adalah jarak antara data dengan centroid, untuk fungsi jarak yang digunakan adalah jarak Euclidean.  $w$  adalah parameter bobot pangkat (*Weighting Exponent*) yang diperkenalkan dalam FCM. Tidak ada nilai ketetapan, biasanya nilai  $w > 1$ , dan umumnya diberi nilai 2.

Nilai keanggotaan tersebut disimpan dalam matrik *fuzzy pseudo-partition* berukuran  $N \times k$  dimana baris merupakan data, sedangkan kolom adalah nilai keanggotaan pada setiap cluster. Bentuknya seperti persamaan (2.4).

$$U = \begin{pmatrix} U_{11}(X_1) & U_{12}(X_1) & \dots & U_{1k}(X_1) \\ U_{21}(X_2) & U_{22}(X_2) & \dots & U_{2k}(X_2) \\ \dots & \dots & \dots & \dots \\ U_{n1}(X_n) & U_{n2}(X_n) & \dots & U_{nk}(X_n) \end{pmatrix} \quad \dots\dots\dots (2.4)$$

Untuk menghitung centroid pada cluster  $c_i$  pada fitur  $j$ , digunakan persamaan (2.5).

$$c_{ij} = \frac{\sum_{i=1}^N (u_{il})^w x_{ij}}{\sum_{i=1}^N (u_{il})^w} \quad \dots\dots\dots (2.5)$$

Parameter  $N$  adalah jumlah data,  $w$  adalah bobot pangkat, dan  $u_{ij}$  adalah derajat keanggotaan data  $x_i$  ke cluster  $c_j$ . Sementara fungsi objektif yang digunakan persamaan (2.6).

$$J = \sum_{i=1}^N \sum_{l=1}^k (u_{ij})^w D(x_i, c_j)^2 \quad \dots\dots\dots (2.6)$$

### A. Algoritma FCM

Secara prinsip, algoritma FCM banyak kesamaan dengan K-Means. Adapun algoritma FCM adalah sebagai berikut (Agusta, 2007) :

1. Inisialisasi : tentukan jumlah cluster ( $k \geq 2$ ), tentukan bobot pangkat ( $w > 1$ ), tentukan jumlah maksimal iterasi, tentukan threshold perubahan nilai fungsi objektif, jika perlu juga perubahan nilai centroid.
2. Berikan nilai awal pada matrik fuzzy pseudo-partition dengan syarat seperti persamaan (2.1).
3. Lakukan langkah 4 sampai 5 selama syarat masih terpenuhi : (1) apabila perubahan pada nilai fungsi objektif masih di atas nilai threshold yang ditentukan; atau (2) perubahan pada nilai centroid masih di atas nilai threshold yang ditentukan; atau (3) iterasi maksimal belum tercapai.
4. Hitung nilai centroid dari masing-masing cluster menggunakan persamaan (2.5).
5. Hitung kembali matrik fuzzy pseudo-partition (derajat keanggotaan setiap data pada setiap cluster) menggunakan persamaan (2.3).

### 2.12 Penelitian Sebelumnya

Fuzzy C-Means (FCM) adalah suatu teknik pengclusteran data yang mana keberadaan tiap-tiap titik data dalam suatu cluster ditentukan oleh derajat keanggotaan. Teknik ini pertama kali diperkenalkan oleh Jim Bezdek

pada tahun 1981. Konsep dari *Fuzzy C-Means* pertama kali adalah menentukan pusat *cluster*, yang akan menandai lokasi rata-rata untuk tiap-tiap *cluster*. Pada kondisi awal, pusat *cluster* ini masih belum akurat. Tiap-tiap titik data memiliki derajat keanggotaan untuk tiap-tiap *cluster*. Dengan cara memperbaiki pusat *cluster* dan derajat keanggotaan tiap-tiap titik data secara berulang, maka akan dapat dilihat bahwa pusat *cluster* akan bergerak menuju lokasi yang tepat. Perulangan ini didasarkan pada minimisasi fungsi objektif yang menggambarkan jarak dari titik data yang diberikan kepusat *cluster* yang terbobot oleh derajat keanggotaan titik data tersebut.

Penelitian sebelumnya dilakukan oleh seorang mahasiswi Universitas Muhammadiyah Gresik yang bernama Syarifatul Mudawamah untuk skripsi yang berjudul “**Sistem Pencarian Potensi Bidang Minat Rekayasa Perangkat Lunak dan Sistem Cerdas Menggunakan *K-Harmonic Means***”. Pengujian konsistensi sistem diketahui sistem memiliki konsistensi terbaik saat nilai  $K=2$  yaitu sebesar 77.50% untuk bidang minat rekayasa perangkat lunak sedangkan untuk bidang minat sistem cerdas yaitu sebesar 79.17%. Dari hasil pengujian diketahui bahwa semakin besar nilai  $K$  maka presentase konsistensi sistem semakin kecil. Nilai validitas terbaik diperoleh ketika nilai mendekati 1. Hasil terbaik didapatkan ketika nilai  $K$  sebanyak dua yaitu sebesar 0.8996 dan nilai validitas terbesar untuk tiap cluster dari semua nilai  $K$  yang telah dimasukkan adalah pada cluster 2 dengan nilai  $K$  dua yaitu sebesar 0.960. Sedangkan pengujian validitas sistem pencarian kecenderungan potensi mahasiswa di bidang sistem cerdas terbaik didapatkan ketika nilai  $K$  sebanyak dua yaitu sebesar 0.928 dan nilai validitas terbesar untuk tiap cluster dari semua nilai  $K$  yang telah dimasukkan adalah pada cluster 2 dengan nilai  $K$  dua yaitu sebesar 0.960.

Hasil pencarian kecenderungan potensi yang dilakukan oleh Kaprodi, Dosen Wali angkatan 2010, dan dosen yang ahli di bidang rekayasa perangkat lunak dan sistem cerdas dapat disimpulkan bahwa akurasi sistem terbaik diperoleh pada saat input  $K$  bernilai 3 yaitu sebesar 68,33% jika dibandingkan dengan hasil pencarian kecenderungan potensi oleh Kaprodi dan Dosen Wali.

Sedangkan nilai akurasi sistem sebesar 84% jika dibandingkan dengan Dosen saat input K sebesar 2.

Selain itu metode FCM juga digunakan untuk penelitian seseorang yang bernama *Bahar* dalam menyelesaikan tesis Magister Komputer (M.Kom) di Universitas Dian Nuswantoro Semarang dengan judul tesis “**Penentuan Jurusan Sekolah Menengah Atas Dengan Algoritma Fuzzy C-Means**”. Dalam hal ini, peneliti melakukan pembuatan aplikasi dengan metode FCM untuk menentukan jurusan ditingkat SMA yang didapatkan terdapat 3 cluster yang dilakukan, yaitu jurusan IPA, IPS, dan Bahasa. Dari data yang dilatih, diperoleh tiga kelompok berdasarkan nilai rata-rata mata pelajaran peminatan, yaitu:

- Kelompok pertama, terdiri atas siswa yang memiliki nilai rata-rata mata pelajaran peminatan IPA sekitar 72,0635; nilai rata-rata mata pelajaran peminatan IPS sekitar 76,3067; dan nilai rata-rata mata pelajaran peminatan Bahasa sekitar 71,5032.
- Kelompok kedua, terdiri atas siswa yang memiliki nilai rata-rata mata pelajaran peminatan IPA sekitar 73,5371; nilai rata-rata mata pelajaran peminatan IPS sekitar 74,7951; dan nilai rata-rata mata pelajaran peminatan Bahasa sekitar 79,7301.
- Kelompok ketiga, terdiri atas siswa yang memiliki nilai rata-rata mata pelajaran peminatan IPA sekitar 80,0742; nilai rata-rata mata pelajaran peminatan IPS sekitar 75,0224; dan nilai rata-rata mata pelajaran peminatan Bahasa sekitar 74,4123.