

## Penerapan Metode *Fuzzy C-Means* Pada Aplikasi Simulasi TOEFL (*Test Of English As a Foreign Language*) Berbasis Web (Studi Kasus: Fakultas MIPA Universitas Mulawarman)

Rizky Fitriani Kemala<sup>1)</sup>, Indah Fitri Astuti<sup>2)</sup>, Septya Maharani<sup>3)</sup>

Program Studi Ilmu Komputer, Fakultas Ilmu Komputer dan Teknologi Informasi, Universitas Mulawarman  
Jalan Panajam Kampus Gn. Kelua Universitas Mulawarman Samarinda, Kalimantan Timur  
email : : rifikadr@gmail.com<sup>1)</sup>, indah@fkti.unmul.ac.id<sup>2)</sup>, septyamaharani@yahoo.com<sup>3)</sup>

### ABSTRAK

Fakultas Matematika dan Ilmu Pengetahuan Alam (FMIPA) di Universitas Mulawarman provinsi Kalimantan Timur merupakan salah satu fakultas yang berbasis pada keilmuan murni. Mahasiswa FMIPA dituntut untuk mampu mengikuti perkembangan ilmu dan teknologi serta meningkatkan penguasaan Bahasa Inggris untuk keperluan akademik. FMIPA mensyaratkan kepada mahasiswa calon lulusan untuk menguasai kemampuan bahasa Inggris sebagai bahasa internasional yang dibuktikan dengan sertifikat *Test Of English as a Foreign Language* (TOEFL) dengan skor minimal yang telah ditetapkan oleh pihak kampus. Mahasiswa satu dengan yang lain pasti mempunyai kemampuan yang berbeda dalam mengerjakan soal TOEFL, sehingga perlu pengelompokan data skor TOEFL berdasarkan jenis soal yang dikerjakan oleh mahasiswa. Algoritma yang digunakan untuk pengelompokan adalah *Fuzzy C-Means*. *Fuzzy C-Means* adalah suatu teknik pengelompokan data yang mana keberadaan tiap-tiap titik data dalam suatu *cluster* ditentukan oleh derajat keanggotaan. Sistem aplikasi simulasi TOEFL menggunakan metode *Fuzzy C-Means* ini dapat mengelompokkan data menjadi beberapa *cluster* berdasarkan jenis soal dengan skor tertinggi yaitu *Cluster Listening Comprehension*, *Cluster Structure and written expression*, dan *Cluster Reading Comprehension*. Penelitian ini memberikan solusi kepada pengguna agar mempunyai persiapan yang lebih kuat menghadapi ujian TOEFL dan sebagai media belajar bagi siapapun yang ingin meningkatkan kemampuan berbahasa Inggris, juga memberikan saran kepada pengguna yang mempunyai nilai terendah pada jenis soal TOEFL tertentu.

**Kata Kunci** : TOEFL, *Fuzzy C-Means*, bahasa Inggris

### 1. PENDAHULUAN

Internet merupakan faktor utama yang mendorong terbukanya akses ke berbagai jenis informasi dan sumber ilmu. Sekarang bukan hal yang sulit untuk belajar Bahasa Inggris atau melakukan persiapan TOEFL (*Test Of English as a Foreign Language*) dimana saja dan kapan saja. Adalah perlu bagi peserta yang mengukur kemampuan Bahasa Inggris untuk mengetahui kelemahan dirinya dalam pengerjaan soal TOEFL. Belum banyak penyelenggara tes TOEFL mencantumkan atau memberikan evaluasi kepada peserta tes, seperti lembaga ETS (*Educational Testing Service*) yang memberikan saran dan evaluasi terhadap hasil skor TOEFL peserta yang mengikuti tes. TOEFL adalah suatu tes yang digunakan untuk mengukur kemampuan berbahasa Inggris bagi orang yang menggunakan bahasa Inggris sebagai bahasa keduanya (*nonnative*). Kemampuan yang diukur yaitu *listening*, *structure*, dan *reading comprehension* (Widodo, 2013).

Fakultas Matematika dan Ilmu Pengetahuan Alam (FMIPA) Universitas Mulawarman merupakan salah satu fakultas yang berbasis keilmuan murni dan mahasiswa FMIPA dituntut untuk dapat mampu mengikuti perkembangan global serta meningkatkan penguasaan Bahasa Inggris dalam dunia akademik. Untuk lulus, FMIPA

mensyaratkan kepada mahasiswa calon lulusan untuk mempunyai sertifikat TOEFL dengan skor yang telah ditetapkan oleh Universitas Mulawarman.

Berdasarkan latar belakang tersebut maka perlu dibangun aplikasi simulasi TOEFL sebagai salah satu solusi yang memberikan alternatif pembelajaran dan membantu penggunaannya mengenal jenis-jenis soal TOEFL serta memberikan simulasi seperti ujian sebenarnya dengan media teknologi yaitu *website*. Aplikasi ini juga akan mengelompokkan skor peserta TOEFL berdasarkan jenis-jenis soal dalam TOEFL.

### 2. TINJAUAN PUSTAKA

#### A. *Test Of English as a Foreign Language* (TOEFL)

TOEFL dimaksudkan sebagai alat ukur atau evaluasi atas kemampuan berbahasa Inggris seseorang. TOEFL sendiri dikembangkan dan dikeluarkan oleh ETS (*Educational Testing Service*) yang berpusat di New Jersey, USA.

Terdapat beberapa bagian tes yang menjadi dasar pada TOEFL, yaitu:

#### 1. *Listening Comprehension*

Bagian ini terdiri dari 50 pertanyaan untuk jangka waktu selama 30-40 menit, peserta diminta untuk mendengarkan soal yang diucapkan hanya satu kali

melalui media audio dimana soal tidak tidak ditulis dilembar naskah soal, melainkan hanya pilihan jawaban.

## 2. Structure and Written Expression

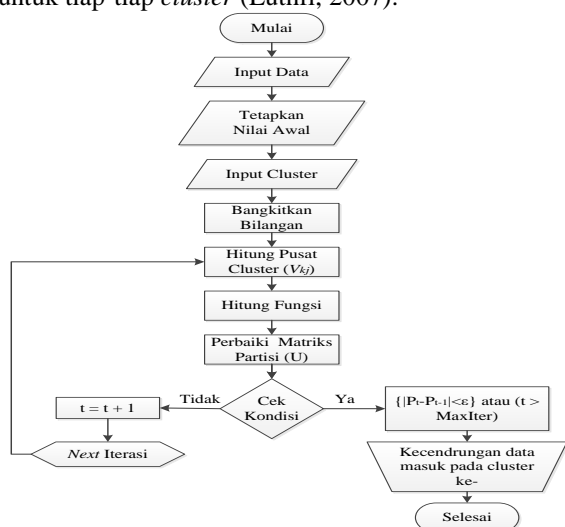
Bagian ini terdiri dari 40 pertanyaan mengenai bahasa Inggris formal atau bahasa-bahasa yang menggunakan kaidah-kaidah tata bahasa. Waktu yang disediakan dalam menjawab pertanyaan dibagian ini adalah 25 menit untuk melengkapi kalimat dan menentukan kesalahan yang ada.

## 3. Reading Comprehension

Bagian ini terdiri dari 50 pertanyaan dengan jangka waktu yang disediakan untuk mengerjakan soal adalah 45 menit. Disisi ini terdapat beberapa bacaan singkat yang masing-masing diikuti oleh pertanyaan, dimana setiap bacaan tidak sama panjangnya dan diikuti oleh pertanyaan yang tidak sama jumlahnya (Putra, Suryono, & Darmini, 2009).

## B. Metode Fuzzy C-Means

*Fuzzy C-Means* adalah teknik pengelompokan data dimana keberadaan tiap-tiap titik data dalam suatu *cluster* ditentukan oleh derajat keanggotaan. *Fuzzy C-Means* adalah algoritma pengelompokan yang terawasi, karena pada algoritma ini jumlah *cluster* yang akan dibentuk perlu diketahui terlebih dahulu. konsep dasarnya adalah menentukan pusat kelompok yang akan menandai lokasi rata-rata untuk tiap-tiap *cluster* (Luthfi, 2007).



Gambar 1. Flowchart Algoritma Fuzzy C-Means

Pusat *cluster* ini masih belum akurat pada kondisi awal. Tiap-tiap titik data memiliki derajat keanggotaan pada tiap-tiap *cluster*. dengan cara memperbaiki pusat cluster dan derajat keanggotaan tiap-tiap titik data secara berulang-ulang, maka didapat lokasi pusat cluster optimal. Perulangan ini didasarkan pada minimasi fungsi obyektif yang menggambarkan jarak dari titik data yang diberikan ke pusat *cluster* yang terbobot oleh derajat

keanggotaan titik data tersebut (Sediyono, Widiasari, & Diponegoro, 2006).

Penjelasan algoritma *Fuzzy C-Means* sebagai berikut:

1. Input data yang akan di kelompokkan, yaitu  $X$  berupa matriks berukuran  $n \times m$  ( $n$  = jumlah sampel data dan  $m$  = atribut setiap data).  $X_{ij}$  = data sampel ke- $i$  ( $i = 1, 2, \dots, n$ ), atribut ke- $j$  ( $j = 1, 2, \dots, m$ ).
2. Tetapkan jumlah *cluster* ( $c$ ), pangkat untuk matriks partisi ( $w$ ), maksimum iterasi (*MaxIter*), error terkecil yang diharapkan ( $\xi$ ), fungsi objektif awal ( $P_0 = 0$ ), iterasi awal ( $t = 1$ ).
3. Bangkitkan bilangan random  $\mu_{ik}$ , dimana  $i = 1, 2, \dots, n$  dan  $k = 1, 2, \dots, c$ ; sebagai elemen-elemen matriks partisi awal  $U$ . Kemudian menghitung jumlah setiap kolom bilangan random yang terbentuk. Hitung jumlah setiap kolom (atribut) :

$$Q_i = \sum_{k=1}^c \mu_{ik} \quad (1)$$

Setelah menjumlahkan setiap kolom dari bilangan random yang terbentuk, maka kemudian menghitung matriks partisi awal  $U$ . Matriks partisi awal  $U$  dapat dihitung dengan persamaan:

$$\mu_{ik} = \frac{\mu_{ik}}{Q_i} \quad (2)$$

4. Hitung pusat cluster ke- $k$ :  $V_{kj}$ , dengan  $k = 1, 2, \dots, c$ ; dan  $j = 1, 2, \dots, m$ ; menggunakan persamaan:

$$V_{kj} = \frac{\sum_{i=1}^n ((\mu_{ik})^w \cdot X_{ij})}{\sum_{i=1}^n (\mu_{ik})^w} \quad (3)$$

dengan keterangan,  $V_{kj}$  adalah pusat *cluster* ke- $k$  untuk atribut ke- $j$ ,  $\mu_{ik}$  adalah derajat keanggotaan untuk data sampel ke- $i$  pada *cluster* ke- $k$ , dan  $X_{ij}$  adalah data ke- $i$ , atribut ke- $j$ .

5. Hitung fungsi objektif pada iterasi ke- $t$  menggunakan persamaan:

$$P_t = \sum_{i=1}^n \sum_{k=1}^c \left( \left[ \sum_{j=1}^m (X_{ij} - V_{kj})^2 \right] (\mu_{ik})^w \right) \quad (4)$$

dengan keterangan  $V_{kj}$  adalah pusat *cluster* ke- $k$  untuk atribut ke- $j$ ,  $\mu_{ik}$  adalah derajat keanggotaan untuk data sampel ke- $i$  pada *cluster* ke- $k$ ,  $X_{ij}$  adalah data ke- $i$ , atribut ke- $j$  dan,  $P_t$  adalah fungsi objektif pada iterasi ke- $t$ .

6. Hitung perubahan matriks partisi menggunakan persamaan:

$$\mu_{ik} = \frac{\left[ \sum_{j=1}^m (X_{ij} - V_{kj})^2 \right]^{\frac{1}{w-1}}}{\sum_{k=1}^c \left[ \sum_{j=1}^m (X_{ij} - V_{kj})^2 \right]^{\frac{1}{w-1}}} \quad (5)$$

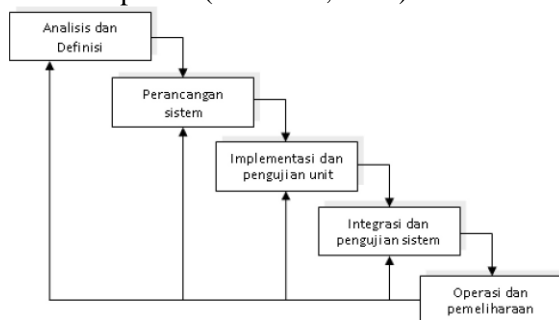
dengan keterangan  $i$  bernilai  $1, 2, \dots, n$  dan  $k$  bernilai  $1, 2, \dots, c$ . Dimana  $V_{kj}$  adalah pusat *cluster* ke- $k$  untuk atribut ke- $j$ ,  $\mu_{ik}$  adalah derajat

keanggotaan untuk data sampel ke-*i* pada *cluster* ke-*k*, dan  $X_{ij}$  adalah data ke-*i*, atribut ke-*j*.

7. Cek kondisi berhenti jika  $t > \text{MaxIter}$  maka berhenti atau jika tidak  $t = t + 1$  dan ulangi langkah ke-4 (Wijaya, 2014).

### 3. METODE PENELITIAN

Perancangan sistem pada penelitian ini menggunakan Model Air Terjun (*waterfall*). *Waterfall model* merupakan salah satu model yang mengambil kegiatan proses dasar seperti spesifikasi, pengembangan, validasi dan evolusi dengan memresentasikannya sebagai fase-fase proses (Pressman, 2012).



Gambar 2. Metode *Waterfall*

Dalam pengembangan dan pembangunan sistem, terdapat 7 tahap yang dilakukan:

1. Tahap Pengumpulan Data  
Melakukan pengumpulan data yang berhubungan dengan data borang akreditasi program studi, basis data menggunakan MySQL, desain antarmuka dan desain sistem. Pengumpulan data diperoleh melalui wawancara dengan pihak Penjaminan Mutu Fakultas dan Universitas Mulawarman, studi literatur melalui buku-buku maupun artikel serta jurnal yang bisa didapatkan di internet.
2. Tahap Perencanaan  
Memahami permasalahan, mengidentifikasi dan mendefinisikan secara rinci setiap masalah beserta kendalanya dan menyusun langkah-langkah dalam mencari penyelesaian penelitian.
3. Tahap Analisis  
Masalah dan langkah penyelesaian yang telah disusun kemudian dianalisis untuk mencari solusi yang paling tepat untuk diterapkan pada sistem. Tahapan ini dilakukan analisis kebutuhan pengguna hingga sistem yang akan dibangun.
4. Tahap Perancangan  
Membuat struktur basis data, dan *interface* yang sesuai dan mudah digunakan oleh pengguna sistem informasi.

5. Tahap Penerapan  
Penulisan kode menggunakan bahasa pemrograman dan framework yang telah dipilih dan diimplementasikan menjadi sebuah program
6. Tahap Evaluasi (Penguujian)  
Penguujian Sistem Menggunakan metode Black-Box dan user acceptance pada akhir pembangunan sistem.
7. Tahap Pemeliharaan  
Memperbaiki kesalahan dan kemampuan kinerja dari sistem informasi yang telah dibuat.

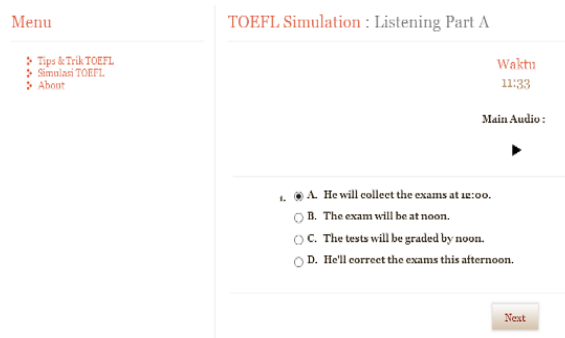
### 4. HASIL DAN PEMBAHASAN

Aplikasi Simulasi TOEFL menyimpan data skor mahasiswa FMIPA yang telah melakukan simulasi dan menampilkan datanya serta menghasilkan *clustering* apakah mahasiswa tersebut memiliki skor tertinggi *Listening*, *Structure* atau *Reading*.



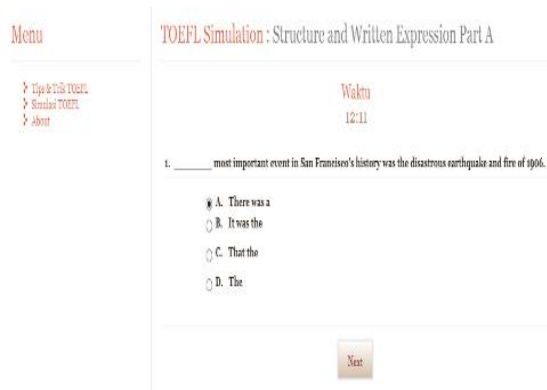
Gambar 3. Halaman Utama Aplikasi Simulasi TOEFL

Gambar 3 adalah halaman utama Simulasi TOEFL. Tampilan awal dari aplikasi simulasi TOEFL yang dilihat dari sisi *user*. Halaman ini terdapat menu-menu yang terdiri dari menu Tips dan Trik TOEFL, menu Simulasi TOEFL, dan menu *About*.



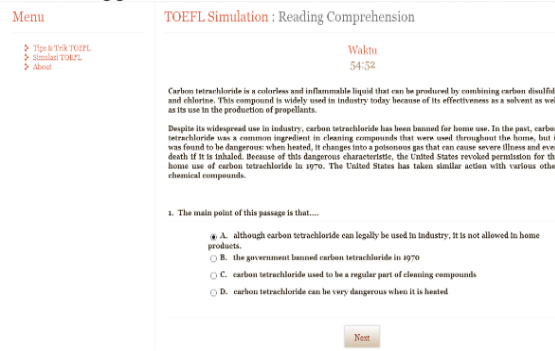
Gambar 4. Simulasi Soal TOEFL *Listening*

Gambar 4 adalah halaman simulasi soal TOEFL *Listening Comprehension*. *Audio* soal pada halaman ini hanya dapat diputar satu kali dan tidak dapat dijeda. Tombol *Next* di bawah soal untuk ke halaman soal berikutnya. Sebelum masuk halaman simulasi TOEFL *user* mengisi data diri peserta



Gambar 5. Simulasi Soal TOEFL Structure

Gambar 5 adalah halaman simulasi soal TOEFL structure and written expression yang berkaitan dengan tata bahasa dan kaidah dalam bahasa Inggris.



Gambar 6. Simulasi Soal TOEFL Reading

Gambar 6 adalah halaman simulasi soal TOEFL Reading Comprehension. Halaman ini terdapat beberapa bacaan singkat yang masing-masing diikuti oleh pertanyaan, dan setiap bacaan tidak sama panjangnya serta diikuti oleh pertanyaan yang juga tidak sama jumlahnya.



Gambar 7. Hasil Simulasi

Gambar 7 adalah tampilan hasil simulasi yang berisi informasi skor tiap jenis soal, skor TOEFL, dan saran untuk setiap jenis soal dengan skor rendah.



No	Nama	NIM	Prodi	skor listening	skor structure	skor reading	Cluster	Nilai Cluster
1	EMI HARMIANI	1207015026	PRODI STATISTIKA	46	40	39	Listening Comprehension	0.33342420
2	AL GHAZALI	1207032100	PRODI STATISTIKA	42	38	32	Listening Comprehension	0.33343441
3	SISWANTO DANAR EDSHA	1207015061	PRODI STATISTIKA	45	40	41	Listening Comprehension	0.33342365
4	ADE NIA SARI	1207015069	PRODI STATISTIKA	41	35	32	Listening Comprehension	0.33343819
5	HARDIANTY	1207015072	PRODI STATISTIKA	35	25	31	Listening Comprehension	0.33345748
6	LETIZIA RIZKA PUTRI	1207025009	PRODI BIOLOGI	49	35	36	Listening Comprehension	0.33342663
7	APRIANI SARASWATI	1207035020	PRODI KEMIA	42	43	31	Structure and Written Expression	0.33325566
8	FITRIANI	1207015010	PRODI STATISTIKA	41	33	45	Reading Comprehension	0.33331508

Gambar 8. Halaman Clustering Data Skor

Gambar 8 menunjukkan hasil clustering data skor berdasarkan jenis soal dengan skor tertinggi menggunakan metode *Fuzzy C-Means*. Terdapat 3 cluster pada hasil di halaman ini. *Cluster Listening Comprehension, Cluster Structure and Written Expression, dan Cluster Reading Comprehension*.

## 5. KESIMPULAN

Metode *fuzzy c-means* dapat diterapkan untuk membantu mengelompokkan data skor TOEFL berdasarkan jenis soal dengan skor tertinggi dengan masukan berupa nilai skor *Listening Comprehension, skor Structure and written expression, dan skor Reading Comprehension*.

## 6. DAFTAR PUSTAKA

- Luthfi, E. T. (2007). *Fuzzy c-means untuk Clustering Data (Studi Kasus Data Performance Mengajar Dosen)*. In *Seminar Nasional Teknologi*. Yogyakarta.
- Pressman, R. S. (2012). *Rekayasa Perangkat Lunak jilid I*. Yogyakarta: Penerbit Andi.
- Putra, A. D., Suryono, R. R., & Darmi. (2009). Rancang Bangun Media Pembelajaran Toefl Berbasis Web. In *Seminar Nasional Aplikasi Teknologi Informasi (SNATI)* (pp. 73–78).
- Sediyono, E., Widiyari, I. R., & Diponegoro, J. (2006). Penentuan Lokasi Fasilitas Gudang Menggunakan Fuzzy C – Means ( FCM ). *Jurnal Informatika*, 2(2), 155–166.
- Widodo, G. S. (2013). Perancangan Latihan TOEFL Menggunakan Framework Code Igniter untuk English Course Centre Banjarnegara. *Jurnal Ilmiah Universitas Dian Nuswantoro*, 1–10.
- Wijaya, A. K. (2014). Implementasi Data Mining dengan Algoritma Fuzzy C - Means Studi Kasus Penjualan di UD Subur Baru. *Jurnal Ilmiah Universitas Dian Nuswantoro*, 1–8.