IMPLEMENTASI ALGORITMA *K-MEANS* PADA PENERIMAAN CALON KARYAWAN MENGGUNAKAN COMPUTER BASED TEST (CBT)

TUGAS AKHIR

Chika Anggun Nabila
41517120063

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2021

http://digilib.mercubuana.ac.id/

IMPLEMENTASI ALGORITMA *K-MEANS* PADA PENERIMAAN CALON KARYAWAN MENGGUNAKAN COMPUTER BASED TEST (CBT)

Tugas Akhir

Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer

Oleh:
Chika Anggun Nabila
41517120063

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2021

http://digilib.mercubuana.ac.id/

LEMBAR PERNYATAAN ORISINALITAS

Yang bertanda tangan dibawah ini:

NIM	: 41517120063
Nama	: Chika Anggun Nabila
Judul Tugas Akhir	: Implementasi Algoritma *K-Means* pada Penerimaan Calon Karyawan Menggunakan Computer Based Test (CBT)

Menyatakan bahwa Laporan Tugas Akhir saya adalah hasil karya sendiri dan bukan plagiat. Apabila ternyata ditemukan didalam laporan Tugas Akhir saya terdapat unsur plagiat, maka saya siap untuk mendapatkan sanksi akademik yang terkait dengan hal tersebut.

Jakarta, 24 Februari 2022

METERAI TEMPEL

Chika Anggun Nabila

UNIVERSITAS
MERCU BUANA

http://digilib.mercubuana.ac.id/

SURAT PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR

SURAT PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR

Sebagai mahasiswa Universitas Mercu Buana, saya yang bertanda tangan di bawah ini :

Nama Mahasiswa	:	41517120063
NIM	:	Chika Anggun Nabila
Judul Tugas Akhir	:	Implementasi Algoritma *K-Means* pada Penerimaan Calon Karyawan Menggunakan Computer Based Test (CBT)

Dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Noneksklusif** (*None-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul diatas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti/Noneksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat dan mempublikasikan tugas akhir saya.

Selain itu, demi pengembangan ilmu pengetahuan di lingkungan Universitas Mercu Buana, saya memberikan izin kepada Peneliti di Lab Riset Fakultas Ilmu Komputer, Universitas Mercu Buana untuk menggunakan dan mengembangkan hasil riset yang ada dalam tugas akhir untuk kepentingan riset dan publikasi selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 24 Februari 2022

METERAI TEMPEL

3DAJX743614073

Chika Anggun Nabila

UNIVERSITAS MERCU BUANA

http://digilib.mercubuana.ac.id/

SURAT PERNYATAAN LUARAN TUGAS AKHIR

SURAT PERNYATAAN LUARAN TUGAS AKHIR

Sebagai mahasiswa Universitas Mercu Buana, saya yang bertanda tangan di bawah ini :

Nama Mahasiswa	:	Chika Anggun Nabila
NIM	:	41517120063
Judul Tugas Akhir	:	Implementasi Algoritma *K-Means* pada Penerimaan Calon Karyawan Menggunakan Computer Based Test (CBT)

Menyatakan bahwa :

1. Luaran Tugas Akhir saya adalah sebagai berikut :

No	Luaran	Jenis		Status	
1	Publikasi Ilmiah	Jurnal Nasional Tidak Terakreditasi		Diajukan	√
		Jurnal Nasional Terakreditasi	√		
		Jurnal International Tidak Bereputasi		Diterima	
		Jurnal International Bereputasi			
	Disubmit/dipublikasikan di :	Nama Jurnal	: JURNAL MATRIK		
		ISSN	: E-ISSN 2476-9843		
		Link Jurnal	: https://journal.universitasbumigora.ac.id/index.php/matrik/		
		Link File Jurnal Jika Sudah di Publish	:		

2. Bersedia untuk menyelesaikan seluruh proses publikasi artikel mulai dari submit, revisi artikel sampai dengan dinyatakan dapat diterbitkan pada jurnal yang dituju.
3. Diminta untuk melampirkan scan KTP dan Surat Pernyataan (Lihat Lampiran Dokumen HKI), untuk kepentingan pendaftaran HKI apabila diperlukan

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 24 Februari 2022

Mengetahui
Dosen Pembimbing TA

Wawan Gunawan, S.Kom., MT.

METERAI TEMPEL
FFD0AJX743614093

Chika Anggun Nabila

UNIVERSITAS MERCU BUANA

http://digilib.mercubuana.ac.id/

LEMBAR PERSETUJUAN PENGUJI

NIM	:	41517120063
Nama	:	Chika Anggun Nabila
Judul Tugas Akhir	:	Implementasi Algoritma *K-Means* pada Penerimaan Calon Karyawan Menggunakan Computer Based Test (CBT)

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 24 Februari 2022

(Sabar Rudiarto, S.Kom., M.Kom.)

UNIVERSITAS
MERCU BUANA

http://digilib.mercubuana.ac.id/

LEMBAR PERSETUJUAN PENGUJI

NIM	:	41517120063
Nama	:	Chika Anggun Nabila
Judul Tugas Akhir	:	Implementasi Algoritma *K-Means* pada Penerimaan Calon Karyawan Menggunakan Computer Based Test (CBT)

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 24 Februari 2022

(Anis Cherid, SE., MTI.)

UNIVERSITAS
MERCU BUANA

http://digilib.mercubuana.ac.id/

LEMBAR PERSETUJUAN PENGUJI

NIM	:	41517120063
Nama	:	Chika Anggun Nabila
Judul Tugas Akhir	:	Implementasi Algoritma *K-Means* pada Penerimaan Calon Karyawan Menggunakan Computer Based Test (CBT)

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 24 Februari 2022

(Ir. Emil R. Kaburuan, Ph.D., IPM.)

UNIVERSITAS
MERCU BUANA

http://digilib.mercubuana.ac.id/

LEMBAR PENGESAHAN

LEMBAR PENGESAHAN

NIM	:	41517120063
Nama	:	Chika Anggun Nabila
Judul Tugas Akhir	:	Implementasi Algoritma *K-Means* pada Penerimaan Calon Karyawan Menggunakan Computer Based Test (CBT)

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 24 Februari 2022

Menyetujui,

(Wawan Gunawan, S.Kom., MT.)
Dosen Pembimbing

Mengetahui,

(Wawan Gunawan, S.Kom., MT.)
Koord. Tugas Akhir Teknik Informatika

(Ir. Ernis K. Kaburuan, Ph.D., IPM.)
Ka. Prodi Teknik Informatika

UNIVERSITAS
MERCU BUANA

http://digilib.mercubuana.ac.id/

KATA PENGANTAR

Puji syukur kepada Allah SWT yang telah melimpahkan segala rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan laporan tugas akhir yang merupakan salah satu persyaratan untuk menyelesaikan program studi strata satu (S1) pada jurusan Informatika Universitas Mercu Buana.

Penulis menyadari pula bahwa laporan tugas akhir ini tidak akan terwujud tanpa bantuan, bimbingan, dan dorongan dari berbagai pihak. Oleh karena itu, penulis mengucapkan terima kasih kepada:

1. Bapak Wawan Gunawan,S.Kom.,M.T, selaku Pembimbing Tugas Akhir yang telah membimbing penulis dengan semua nasihat, semangat, dan ilmunya dalam menyusun laporan tugas akhir ini.
2. Bapak Ir. Emil R. Kaburuan, Ph.D., IPM. selaku Kepala Program Studi Informatika Universitas Mercu Buana.
3. Ibu Umniy Salamah S.Kom,MMSI selaku dosen pembimbing akademik yang telah membimbing penulis sejak awal semester yang selalu memberi dukungan dan motivasi agar lulus tepat waktu.
4. Bapak dan ibu dosen yang telah memberikan ilmu dan pengetahuan selama penulis menempuh pendidikan di Universitas Mercu Buana.
5. Kedua orang tua yang telah mendukung penulis baik secara moral maupun materi.
6. Seluruh teman-teman yang memberikan dukungan dan membantu saya dalam menyusun laporan Tugas Akhir ini.

Akhir kata, penulis berharap kritik dan saran yang bersifat membangun demi kesempurnaan penulisan di masa akan datang. Semoga laporan ini dapat berguna bagi para pembaca

Jakarta, November 2021

Chika Anggun Nabila

UNIVERSITAS MERCU BUANA

http://digilib.mercubuana.ac.id/

DAFTAR ISI

HALAMAN SAMPUL i
HALAMAN JUDUL i
LEMBAR PERNYATAAN ORISINALITAS ii
SURAT PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR ... iii
SURAT PERNYATAAN LUARAN TUGAS AKHIR iv
LEMBAR PERSETUJUAN PENGUJI v
LEMBAR PENGESAHAN viii
ABSTRAK ix
ABSTRACT x
KATA PENGANTAR xi
DAFTAR ISI xii
NASKAH JURNAL 1
KERTAS KERJA 11
BAB 1. LITERATUR REVIEW 14
BAB 2. ANALISIS DAN PERANCANGAN 20
BAB 3. SOURCE CODE 38
BAB 4. DATASET 88
BAB 5. TAHAPAN EKSPERIMEN 89
BAB 6. HASIL SEMUA EKSPERIMEN 92
DAFTAR PUSTAKA 99
LAMPIRAN DOKUMEN HAKI 102
LAMPIRAN KORESPONDENSI 104

UNIVERSITAS MERCU BUANA

http://digilib.mercubuana.ac.id/

NASKAH JURNAL

Implementasi Algoritma K-Means pada Penerimaan Calon Karyawan menggunakan Computer Based Test (CBT)

Implementation K-Means Algorithm on the Acceptance of Prospective Employees using a Computer Based Test (CBT)

Chika Anggun Nabila[1], Wawan Gunawan[2]

[1,2]Universitas Mercu Buana, Indonesia

Article Info

Genesis Artikel:

Diterima, Tanggal Bulan Tahun
Direvisi, Tanggal Bulan Tahun
Disetujui, Tanggal Bulan Tahun

Kata Kunci:

Computer Based Test
Clustering
K-Means
Penerimaan Karyawan

ABSTRAK

Terdapat banyak perusahaan saat ini dalam melaksanakan seleksi calon karyawan masih memberlakukan serangkaian tes secara manual atau tertulis. Computer Based Test atau biasa disebut dengan CBT merupakan sebuah sistem aplikasi sebagai alternative untuk melaksanakan test yang akan memberikan kemudahan dalam menentukan kompetensi calon karyawan bagi perusahaan dengan lebih cepat serta efisien. Hal tersebut menjadikan sebuah harapan besar perusahaan dalam meningkatkan efektivitas pelayanan dalam sistem penerimaan calon karyawan. Dengan adanya sistem tersebut proses penentuan akhir dapat segera di eksekusi dengan mengelompokkan data hasil score akhir calon karyawan. Maka dari itu dibutuhkan sebuah sistem pengelompokkan data untuk mengetahui penentuan kelayakan diterimanya calon karyawan. Dalam penelitian ini, penulis mencoba mengaplikasikan clustering k-means dengan menentukan 2 buah cluster untuk mengelompokkan kelayakan penerimaan calon karyawan dengan output diterima dan tidak lolos.

Keywords:

Computer Based Test
Clustering
K-Means
Employee Recruitment

ABSTRACT

There are many companies currently in carrying out the selection of prospective employees still apply a series of tests manually or in writing. Computer Based Test or commonly referred to as CBT is an application system as an alternative to carrying out tests that will provide ease in determining the competence of prospective employees for the company more quickly and efficiently. This makes it a great expectation of the company in improving the effectiveness of services in the admission system of prospective employees. With the system, the final determination process can be immediately executed by grouping the final score data of prospective employees. Therefore, a data grouping system is needed to determine the eligibility of prospective employees. In this study, the authors tried to apply k-means clustering by determining 2 clusters to group the eligibility of prospective employees with the output received and not reveived.

This is an open access article under the CC BY-SAlicense.

Penulis Korespondensi:

Chika Anggun Nabila,
Program Studi Teknik Informatika,
Universitas Mercu Buana,
Email: 41517120063@ student.mercubuana.ac.id

http://digilib.mercubuana.ac.id/

1. PENDAHULUAN

Kemajuan teknologi saat ini terus berkembang hingga menjadi bagian dalam kehidupan manusia. Dengan adanya kemajuan yang semakin meningkat, teknologi menjadi peran penting untuk saling bertukar informasi maupun berkomunikasi [1][2]. Hal tersebut tentunya menuntut manusia untuk lebih melek terhadap teknologi sehingga dapat dimanfaatkan untuk mempermudah kehidupan mereka secara umum karena tidak ada batasan terhadap penggunaan teknologi [3]. Informasi terhadap teknologi yang didapatkan contohnya seperti dalam mencari sebuah pekerjaan, tentunya informasi penting terkait persyaratan serta ujian test yang akan dihadapi.

Permasalahan yang terjadi saat ini, perusahaan dalam melaksanakan seleksi calon karyawan masih memberlakukan serangkaian test secara manual atau tertulis. Tentunya dengan sistem tersebut akan sangat membutuhkan waktu yang cukup lama ketika melakukan koreksi terhadap hasil akhir dalam menghitung score test calon karyawan [4]. Pada era modern yang serba cepat ini, waktu sangatlah penting dan menghargai waktu adalah prinsip setiap orang [5]. Hal ini menjadikan sebuah harapan besar perusahaan dalam meningkatkan efektivitas pelayanan dalam sistem penerimaan calon karyawan.

Pada penelitian sebelumnya menjelaskan tentang sistem operasional dalam transaksi layanan pasien hanya mencatat laporan terkait data. Sehingga menjadi kurang optimal karena data yang menumpuk dalam database server tidak memberikan sebuah informasi. Dengan memanfaatkan metode *clustering k-means* dalam pengelompokkan data informasi rekam medis pasien, sebanyak 253 buah data penularan penyakit berhasil dikelompokkan dalam waktu 0,06 detik [6].

Penelitian kedua, dalam menentukan klasifikasi individu untuk dikelompokkan dalam kategori yang sama maupun berbeda berdasarkan perbandingan kuantitatif variable diperlukan menerapkan pengelompokkan proses. Penerapan yang dilakukan pada penelitian tersebut adalah proses *clustering* dataset UN SMA 2019 bidang IPA, IPS, dan program bahasa di Kabupaten Banyumas. Hasil dari *clustering* tersebut menunjukkan bahwa sekolah di *cluster* 2 memiliki pendidikan yang berkualitas dibanding dengan cluster 1 dan 3 [7].

Penelitan ketiga, untuk mendapatkan sebuah informasi penting terhadap data mahasiswa perlu melakukan pengolahan data terdahulu. Dalam mencari data yang terpilih perlu diterapkan sebuah metode dan teknik tertentu. Contoh penerapan yang dilakukan peneliti menggunakan *clustering k-means* dengan hasil mampu mengelompokkan data-data, baik yang sama maupun berbeda. *Cluster* terbentuk mengacu pada calon mahasiswa dengan program studi atau jurusan yang paling banyak diminati. Hasil yang diterapkan pada penelitian ini menunjukkan bahwa pemanfaatan perhitungan *k-means* dalam menyusun informasi data mahasiswa yang sebelumnya tidak terstruktur menjadi terorganisir [1].

Dalam penelitian ini akan dibahas mengenai sistem yang akan dibangun menggunakan website sebagai alternative untuk melaksanakan test secara sistem yang akan memberikan kemudahan perusahaan dalam melakukan penilaian hasil test calon karyawan dengan cepat dalam memberi keputusan. Metode *clustering* adalah metode yang cocok digunakan dalam sistem pengelompokkan, dan *k-means* adalah *clustering* dengan kecepatan dan tingkat akurasi yang tinggi [8]. Penerapan yang dilakukan adalah dengan menghitung score masing-masing kategori test yang sudah dilakukan. Setelah melalui proses tersebut, ditentukan 2 buah *cluster* untuk mengelompokkan kelayakan penerimaann calon karyawan dengan output diterima dan tidak lolos.

2. METODE PENELITIAN

Metode penelitian yang digunakan dalam pengerjaan penelitian ini adalah *Clustering K-Means. Clustering* merupakan teknik untuk menerapkan pengelompokan proses [9], salah satunya adalah algoritma *k-means*. *K-means* termasuk dalam kategori kelompok *unsupervised learning* dengan prinsip utama mengklasifikasikan objek berdasarkan pengukuran jarak *Euclidean (Euclidean Distance)* [7][10]. *Euclidean distance* merupakan salah satu teknik untuk melakukan perhitungan dalam mengukur tingkat kemiripan suatu data [11].

Pada metode *K-Means* data dikelompokkan berdasarkan nilai *cluster* [12], dimana setiap *cluster* memiliki karakteristik kamiripan satu dengan lainnya dan memiliki berbagai karakteristik berbeda dalam pengelompokkan [13]. *K-Means* merupakan teknik yang terkenal dikarenakan kecepatannya dalam melakukan pengklasteran data

http://digilib.mercubuana.ac.id/

[14]. Dalam menentukan pusat *cluster*, nilai akan dipilih secara acak dari *history score test* calon karyawan yang di peroleh dari sistem CBT.

Dalam penelitian ini proses perhitungan K-means dilakukan. Berikut adalah tahapan perhitungan yang dilakukan :

1. Menentukan jumlah *cluster*
2. Pilih titik cluster secara acak untuk dijadikan pusat (centroid)
3. Hitung jarak dan alokasi setiap data ke *centroid* atau rata-rata terdekat. Untuk menentukan jarak pada proses clustering menggunakan teori *Euclidean Distance* yang dapat dilihat pada persamaan (1) :

$$D(x,y) = \sqrt{(\mathrm{Xn} - \mathrm{Yn})^2 + (\mathrm{Xn} - \mathrm{Yn})^2 + \cdots} \quad (1)$$

Ket :

D = Jarak data dan centroid

X_n dan Y_n = Objek data X dan Y dari atribut ke-n

4. Tentukan centroid baru atau rata-rata dari setiap cluster.
5. Kembali pada step 3, jika data tidak ada perubahan maka proses clustering berhenti.

2.1 Sumber Data

Sumber data dalam penelitian yang digunakan merupakan data history hasil score test calon karyawan yang telah melaksanakan test pada aplikasi CBT (*Computer Based Test*). Data yang digunakan untuk dilakukan menghitung clustering k-means yaitu berupa 3 data nilai dari hasil ujian (Test Potensi Akademik, Test Psikologi, dan Test Kompetensi Keahlian). Tabel 1 berikut merupakan dataset yang digunakan.

Tabel 1. Dataset

ID Pendaftar	TPA	PSI	TKK
LDP001	72	72	74
LDP002	69	90	88
LDP003	90	95	93
LDP004	47	57	44
LDP005	76	78	75
LDP006	71	83	78
LDP007	88	95	92
LDP008	40	43	39
LDP009	64	64	67
.....	.....	.....	.....
LDP289	88	75	76

3. HASIL DAN ANALISIS

3.1. Simulasi Algoritma K-Means

Simulasi ini dilakukan dengan menghitung secara manual menggunakan rumus *eucledian distance* guna melakukan eksperimen algoritma *clustering k-means*. Hasil dari proses perhitungan yang diperoleh secara manual digunakan sebagai komparasi pada aplikasi yang dibuat sehingga dapat diketahui keakuratan data yang diperoleh [15]. Berikut adalah tahapan eksperimen yang dilakukan.

1. Data latih yang merupakan dataset hasil *score test* calon karyawan pada Table 1.

http://digilib.mercubuana.ac.id/

2. Menentukan Jumlah Cluster dengan menetapkan 2 buah kluster untuk dapat menentukan hasil akhir diterima atau tidak lolos.
3. Menentukan titik cluster untuk dijadikan pusat atau Centroid, data yang diambil adalah pada ID Pendaftar LDP175 dan LDP233. Tabel 2 merupakan titik cluster awal.

Tabel 2. Titik Cluster

Cluster	TPA	PSI	TKK
C1	43	51	54
C2	67	86	83

4. Melakukan perhitungan jarak untuk menghitung rata-rata terdekat dengan alokasi data ke *centroid* dapat dihitung dengan rumus *Euclidean Distance*, yang ditujukan pada persamaan(1):

Data ke-1 dengan pusat cluster 1 (C1)

$$D(x1, y1) = \sqrt{(X1.1 - Y1)^2 + (X1.2 - Y1)^2 + (X1.3 - Y1)^2}$$
$$= \sqrt{(72 - 43)^2 + (72 - 51)^2 + (74 - 54)^2}$$
$$= 41.01219330882$$

Data ke-1 dengan pusat cluster 2 (C2)

$$D(x1, y2) = \sqrt{(X1.1 - Y2)^2 + (X1.2 - Y2)^2 + (X1.3 - Y2)^2}$$
$$= \sqrt{(72 - 67)^2 + (72 - 86)^2 + (74 - 83)^2}$$
$$= 17.378147196983$$

Dimana :

X = data pada Tabel 1 berdasarkan baris dan kolom ke-n

Y = titik cluster pada Tabel 2 ke-n

Lakukan berulang sampai data ke-n. Tabel 3 merupakan hasil perhitungan jarak terdekat pada iterasi pertama.

Tabel 3. Menentukan Cluster Jarak Terdekat

ID Pendaftar	C1	C2	Cluster
LDP001	41.01219330882	17.37814719698	2
LDP002	57.90509476721	6.70820393250	2
LDP003	75.27283706624	26.64582518895	2
LDP004	12.32882800594	52.55473337388	1
LDP005	47.52893855326	14.45683229480	2
LDP006	48.82622246293	7.07106781187	2
LDP007	73.51870510285	24.55605831562	2
LDP008	17.26267650163	67.18630812896	1
LDP009	27.91057147391	27.36786436681	2
….	….	….	….
LDP289	55.54277630800	24.71841418862	2

Selanjutnya menentukan cluster terbaru berdasarkan rata-rata dengan rumus :

Cluster Baru = nilai hasil cluster / banyak hasil cluster.

http://digilib.mercubuana.ac.id/

Cluster 1 Baru

C1 (tpa) = (47+40+38+58+40+....data ke-n) / 113
= 5831 / 113
= 51.601769911504

C1 (psi) = (57+43+60+54+52+…data ke-n) / 113
= 5988 / 113
= 52.991150442478

C1 (tkk) = (44+39+50+52+43+…date ke-n) / 113
= 5807 / 113
= 51.389380530973

Cluster 2 Baru

C2 (tpa) = (72+69+90+76+71+…data ke-n) / 176
= 13093 / 176
= 74.392045454545

C2 (psi) = (72+90+95+78+83+...date ke-n) / 176
= 13759 / 176
= 78.176136363636

C2 (tkk) = (74+88+93+75+78+…data ke-n) / 176
= 13585 / 176
= 77.1875

Tabel 4. Menentukan Titik Cluster Baru

ID Pendaftar	C1			C2		
	TPA	PSI	TKK	TPA	PSI	TKK
LDP001	0	0	0	72	72	74
LDP002	0	0	0	69	90	88
LDP003	0	0	0	90	95	93
LDP004	47	57	44	0	0	0
LDP005	0	0	0	76	78	75
LDP006	0	0	0	71	83	78
LDP007	0	0	0	88	95	92
LDP008	40	43	39	0	0	0
LDP009	0	0	0	64	64	67
…..	…..	…..	…..	…..	…..	…..
Tot nilai Cluster	5831	5988	5807	13093	13759	13585
Banyak Cluster	113	113	113	176	176	176
Cluster Baru	51.60176991150	52.99115044248	51.38938053097	74.39204545455	78.17613636364	77.1875

5. Setelah berhasil mendapatkan hasil clustering baru, hitung kembali jarak antara data dengan centroid sampai hasil clustering tetap atau iterasi berhenti.
6. Dalam melakukan simulasi algoritma k-means dilakukan, proses perhitungan berhenti pada iterasi ke-10. Dengan hasil perhitungan clustering sebagai berikut :

Tabel 5. Cluster Jarak Terdekat pada iterasi ke-10

ID Pendaftar	C1	C2	Cluster
LDP001	38.57778376216	5.96197228429	2
LDP002	57.50500326059	17.88632927390	2
LDP003	73.18705759900	29.38127633771	2
LDP004	8.61077232308	46.47398011581	1
LDP005	44.84668772607	2.84430783560	2
LDP006	47.21806222199	6.50438454280	2
LDP007	71.52318085768	27.73314125916	2
LDP008	17.02543391518	60.76548843232	1
LDP009	25.44062499232	18.81924974960	2
….	….	….	….
LDP289	51.72741439508	14.60974279096	2

Tabel 6. Titik Cluster Baru pada iterasi ke-10

ID Pendaftar	C1			C2		
	TPA	PSI	TKK	TPA	PSI	TKK
LDP001	0	0	0	72	72	74
LDP002	0	0	0	69	90	88
LDP003	0	0	0	90	95	93
LDP004	47	57	44	0	0	0
LDP005	0	0	0	76	78	75
LDP006	0	0	0	71	83	78
LDP007	0	0	0	88	95	92

http://digilib.mercubuana.ac.id/

ID Pendaftar	C1			C2		
	TPA	PSI	TKK	TPA	PSI	TKK
LDP008	40	43	39	0	0	0
LDP009	0	0	0	64	64	67
…..	….	….	….	**…..**	**…..**	**…..**
Tot nilai Cluster	5019	5142	4973	13905	14605	14419
Banyak Cluster	100	100	100	189	189	189
Cluster Baru	50.19	51.42	49.73	73.57142857143	77.27513227513	76.29100529101

3.2. Analisis dan Perancangan

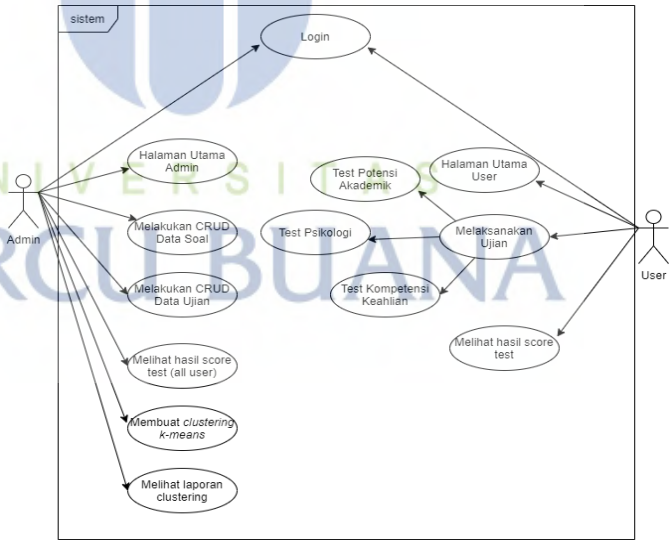
- **Analisis Kebutuhan**

  Identifikasi kebutuhan diperlukan untuk menentukan kebutuhan aplikasi yang akan di bangun. Berikut identifikasi kebutuhan pada sistem aplikasi CBT.

Tabel 7. Analisis Kebutuhan

Menu	Fungsi
Login	Menu login yaitu menu yang digunakan sebagai syarat untuk dapat mengakses aplikasi CBT. Hanya pengguna yang memiliki akun saja yang dapat login.
User	Menu user merupakan menu yang berfungsi untuk mengelola user aplikasi CBT
Data Soal	Menu data soal merupakan menu yang berfungsi untuk mengelola soal yang akan di ujikan pada sistem CBT
Data Ujian	Menu data ujian merupakan menu untuk mengelola jadwal ujian atau tes yang akan di ujikan.
Proses Clustering	Menu untuk mengelompokkan nilai hasil ujian pelamar untuk menentukan kelayakan calon karyawan dalam membantu mengambil keputusan menggunakan clustering k-means
Laporan (Output)	Memberikan informasi dari detail proses clustering.

- **Perancangan Diagram UML**



Gambar 1. Diagram UML Aplikasi

Berdasarkan Gambar 1 pada diagram UML aplikasi dapat diuraikan sebagai berikut :

1. Terdapat 2 aktor yang dapat menjalankan aplikasi yaitu Admin dan User.
2. Aktor yang bertugas sebagai admin memiliki hak akses penuh terhadap aplikasi. Diantaranya login, halaman utama admin, melakukan CRUD data soal, melakukan CRUD data ujian, melihat hasil score test (all user), membuat clustering k-means, melihat laporan clustering.

http://digilib.mercubuana.ac.id/

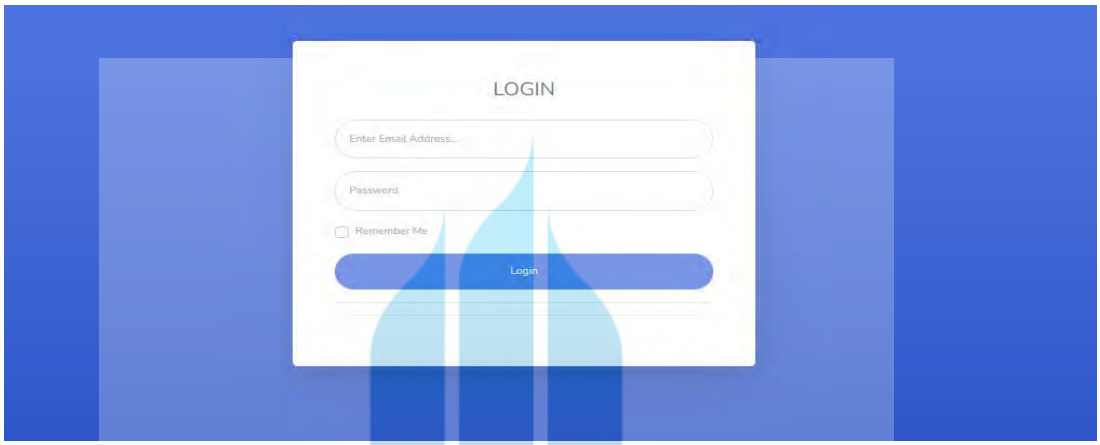
3. Aktor user memiliki hak akses yang terbatas dimana pengguna hanya dapat mengakses menu yang berkaitan dengan pelaksanaan ujian. Diantaranya login, halaman utama user, dan halaman ujian untuk melaksanakan ujian CBT.

3.3. Implementasi

Hasil implementasi atau eksperimen merupakan tahapan dalam pembuatan perangkat lunak dari hasil perancangan yang kemudian diolah menjadi perangkat lunak yang sebenarnya. Aplikasi yang dibuat berbasis web, berikut tampilan hasil implementasi :

1. Halaman Login

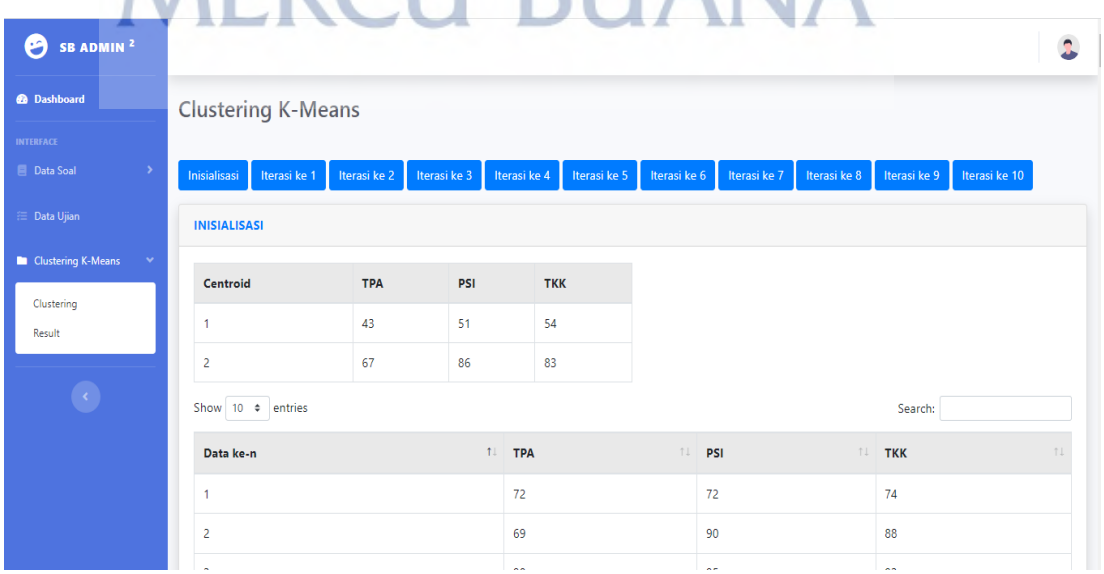
   Halaman Login akan digunakan oleh admin dan user untuk dapat masuk ke dalam aplikasi.



Gambar 2. Halaman Login

2. Halaman *Clustering K-Means*

   Halaman *Clustering K-Means* merupakan halaman yang digunakan untuk membuat *clustering* pada penerimaan calon karyawan berdasarkan data *history score test* pada aplikasi CBT.



UNIVERSITAS MERCU BUANA

http://digilib.mercubuana.ac.id/

Gambar 3. Menentukan Titik Pusat Clustering K-Means

Pada Gambar 3 diatas adalah inisialisasi atau menentukan nilai titik pusat *clustering*. Dimana dalam menentukan nilai titik pusat dipilih secara acak oleh sistem dengan jumlah 2 buah kluster.

Iterasi ke 1

Centroid	TPA	PSI	TKK
1	43	51	54
2	67	86	83

Show 10 entries Search:

Data ke i	Nama	TPA	PSI	TKK	Jarak ke centroid 1	Jarak ke centroid 2	Jarak terdekat	Cluster
1	LDP001	72	72	74	41.01219330882	17.378147196983	17.378147196983	2
2	LDP002	69	90	88	57.905094767214	6.7082039324994	6.7082039324994	2
3	LDP003	90	95	93	75.272837066235	26.645825188948	26.645825188948	2
4	LDP004	47	57	44	12.328828005938	52.554733373884	12.328828005938	1
5	LDP005	76	78	75	47.528938553265	14.456832294801	14.456832294801	2
6	LDP006	71	83	78	48.826222462935	7.0710678118655	7.0710678118655	2
7	LDP007	88	95	92	73.518705102851	24.556058315617	24.556058315617	2

Gambar 4. Iterasi pertama Perhitungan K-Means

Pada Gambar 4 diatas menunjukkan proses clustering data pada iterasi pertama. Dalam menentukan pengelompokkan data untuk masuk dalam *cluster* 1 dan *cluster* 2 diambil dari persamaan rumus *Euclidean Distance* dengan mengambil jarak yang terdekat.

Iterasi ke 10

Centroid	TPA	PSI	TKK
1	50.19	51.42	49.73
2	73.571428571429	77.275132275132	76.291005291005

Show 10 entries Search:

Data ke i	Nama	TPA	PSI	TKK	Jarak ke centroid 1	Jarak ke centroid 2	Jarak terdekat	Cluster
1	LDP001	72	72	74	38.57778376216	5.9619722842914	5.9619722842914	2
2	LDP002	69	90	88	57.505003260586	17.886329273899	17.886329273899	2
3	LDP003	90	95	93	73.187057599004	29.381276337711	29.381276337711	2
4	LDP004	47	57	44	8.6107723230846	46.473980115806	8.6107723230846	1
5	LDP005	76	78	75	44.846687726074	2.8443078355959	2.8443078355959	2
6	LDP006	71	83	78	47.218062221993	6.504384542796	6.504384542796	2
7	LDP007	88	95	92	71.523180857677	27.733141259163	27.733141259163	2
8	LDP008	40	43	39	17.025433915175	60.765488432318	17.025433915175	1

Gambar 5. Iterasi terakhir Perhitungan K-Means

http://digilib.mercubuana.ac.id/

Pada Gambar 5 diatas merupakan iterasi ke-10 atau iterasi terakhir perhitungan K-Means, dimana pada iterasi sebelumnya memiliki titik cluster yang sama atau tidak berubah, sehingga proses iterasi selesai.

4. KESIMPULAN

Berdasarkan serangkaian teori, analisa, perancangan, dan implementasi dapat disimpulkan bahwa penentuan *centroid* atau titik pusat akan berpengaruh dalam hasil *clustering*. Pada hasil eksperimen menggunakan dataset sebanyak 289 data, dimana melibatkan 3 kriteria yang digunakan yaitu score hasil test potensi akademik, psikologi, dan kompetensi keahlian, dengan menentukan 2 buah *cluster*, *output* yang dihasilkan yaitu 100 data masuk dalam *cluster* 1 (C1) dan sebanyak 189 data yang masuk *cluster* 2 (C2). Dalam menentukan hasil akhir dengan *output* diterima atau tidak nya calon karyawan dapat di lihat berdasarkan karakteristik nilai dari *centroid* atau titik pusat. Pada penelitian, *Cluster* 2 (C2) dapat dikelompokkan dalam *cluster* yang direkomendasikan untuk diterima karena memiliki karakteristik nilai yang lebih tinggi dengan rata-rata nilai 78,67 berdasarkan nilai pada *score* test potensi akademik = 67, psikologi = 86, kompetensi keahlian = 83. Dibandingkan dengan *Cluster* 1 (C1) yang memiliki rata-rata nilai 49,33 dengan *score* test potensi akademik = 43, psikologi = 51, kompetensi keahlian = 54.

REFERENSI

[1] F. Yunita, "Penerapan Data Mining Menggunkan Algoritma K-Means Clustring Pada Penerimaan Mahasiswa Baru," *Sistemasi*, vol. 7, no. 3, p. 238, 2018, doi: 10.32520/stmsi.v7i3.388.

[2] S. Rustam, "Analisa Clustering Phising Dengan K-Means Dalam," *Ilk. J. Ilm.*, vol. 10, no. 2, pp. 175–181, 2018.

[3] D. Triyansyah and D. Fitrianah, "Analisis Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Marketing," *J. Telekomun. dan Komput.*, vol. 8, no. 3, p. 163, 2018, doi: 10.22441/incomtech.v8i3.4174.

[4] N. A. Sinaga and T. Firinta, "Indonesian Business Polytechnic New Student Admissions Test Using a Computer Based Test (CBT)," *J. Mantik*, vol. 4, no. 3, pp. 2237–2244, 2020.

[5] L. D. Herliandry, N. Nurhasanah, M. E. Suban, and H. Kuswanto, "Pembelajaran Pada Masa Pandemi Covid-19," *JTP - J. Teknol. Pendidik.*, vol. 22, no. 1, pp. 65–70, 2020, doi: 10.21009/jtp.v22i1.15286.

[6] A. Ali, "Klasterisasi Data Rekam Medis Pasien Menggunakan Metode K-Means Clustering di Rumah Sakit Anwar Medika Balong Bendo Sidoarjo," *MATRIK J. Manajemen, Tek. Inform. dan Rekayasa Komput.*, vol. 19, no. 1, pp. 186–195, 2019, doi: 10.30812/matrik.v19i1.529.

[7] R. Ananda, "Silhouette Density Canopy K-Means for Mapping the Quality of Education Based on the Results of the 2019 National Exam in Banyumas Regency," *Khazanah Inform. J. Ilmu Komput. dan Inform.*, vol. 5, no. 2, pp. 158–168, 2019, doi: 10.23917/khif.v5i2.8375.

[8] S. Rustam, H. A. Santoso, and C. Supriyanto, "Optimasi K-Means Clustering Untuk Identifikasi Daerah Endemik Penyakit Menular Dengan Algoritma Particle Swarm Optimization Di Kota Semarang," *Ilk. J. Ilm.*, vol. 10, no. 3, pp. 251–259, 2018, doi: 10.33096/ilkom.v10i3.342.251-259.

[9] S. Rudiarto, "Implementasi Algoritma K-Means Clustering Pada Aplikasi Pencari Pelanggan Potensial Pada Restoran Xyz," *J. Ilmu Tek. dan Komput.*, vol. 2, no. 1, pp. 1–8, 2018.

[10] M. R. Muttaqin and M. Defriani, "Algoritma K-Means untuk Pengelompokan Topik Skripsi Mahasiswa," *Ilk. J. Ilm.*, vol. 12, no. 2, pp. 121–129, 2020, doi: 10.33096/ilkom.v12i2.542.121-129.

[11] M. Nishom, "Perbandingan Akurasi Euclidean Distance, Minkowski Distance, dan Manhattan Distance pada Algoritma K-Means Clustering berbasis Chi-Square," *J. Inform.*

http://digilib.mercubuana.ac.id/

J. Pengemb. IT, vol. 4, no. 1, pp. 20–24, 2019, doi: 10.30591/jpit.v4i1.1253.

[12] N. Agustina and P. Prihandoko, "Perbandingan Algoritma K-Means dengan Fuzzy C-Means Untuk Clustering Tingkat Kedisiplinan Kinerja Karyawan," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 2, no. 3, pp. 621–626, 2018, doi: 10.29207/resti.v2i3.492.

[13] S. S. Nagari and L. Inayati, "Implementation of Clustering Using K-Means Method To Determine Nutritional Status," *J. Biometrika dan Kependud.*, vol. 9, no. 1, p. 62, 2020, doi: 10.20473/jbk.v9i1.2020.62-68.

[14] I. Vhallah, S. Sumijan, and J. Santony, "Pengelompokan Mahasiswa Potensial Drop Out Menggunakan Metode Clustering K-Means," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 2, no. 2, pp. 572–577, 2018, doi: 10.29207/resti.v2i2.308.

[15] S. Handoko, F. Fauziah, and E. T. E. Handayani, "Implementasi Data Mining Untuk Menentukan Tingkat Penjualan Paket Data Telkomsel Menggunakan Metode K-Means Clustering," *J. Ilm. Teknol. dan Rekayasa*, vol. 25, no. 1, pp. 76–88, 2020, doi: 10.35760/tr.2020.v25i1.2677.

[16] N. Rendy and G. Leonard, "Penerapan Data Mining menggunakan Algoritma K-Means Clustering untuk Menganalisa Bisnis Perusahaan Asuransi," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 6, no. 1, pp. 85–95, 2019, doi: 10.35957/jatisi.v6i1.150.

[17] B. Biantoro and Hernadewita, "K-means clustering on quality of radial run out tires," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 1034, no. 1, p. 012122, 2021, doi: 10.1088/1757-899x/1034/1/012122.

[18] S. Setiawan, "Analisis Cluster Menggunakan Algoritma K-Means Untuk Mengetahui Kemampuan Pegawai Dibidang It Pada Cv . Roxed Ltd," *J. Pelita Inform.*, vol. 18, pp. 80–86, 2019, [Online]. Available: https://ejurnal.stmik-budidarma.ac.id/index.php/pelita/article/view/1142.

[19] S. Suprihatin, Y. R. W. Utami, and D. Nugroho, "K-Means Clustering Untuk Pemetaan Daerah Rawan Demam Berdarah," *J. Teknol. Inf. dan Komun.*, vol. 7, no. 1, 2019, doi: 10.30646/tikomsin.v7i1.408.

[20] I. Maryani, A. Ishaq, and D. S. Mulyadi, "Sistem Informasi Pemesanan Minuman Berbasis Client Server Pada Kampung Dahar Purwokerto," *Evolusi J. Sains dan Manaj.*, vol. 6, no. 2, pp. 84–90, 2018, doi: 10.31294/evolusi.v6i2.4455.

[21] I. H. G. Manurung, "Sistem Informasi Lembaga Kursus Dan Pelatihan (LKP) City Com Berbasis Web Menggunakan Php Dan Mysql," *J. Mahajana Inf.*, vol. 4, no. 1, pp. 42–50, 2019.

[22] M. F. Santoso, "Teknik Responsive Web Design Bootstrap 4 Serta Penerapannya Dalam Rancang Bangun Layout Web," *J. Pilar Nusa Mandiri*, vol. 15, no. 1, pp. 61–68, 2019, doi: 10.33480/pilar.v15i1.101.

UNIVERSITAS
MERCU BUANA

http://digilib.mercubuana.ac.id/

KERTAS KERJA

Ringkasan

Kertas kerja ini merupakan material kelengkapan artikel jurnal dengan judul Implementasi Algoritma *K-Means* pada Penerimaan Calon Karyawan Menggunakan Computer Based Test (CBT) yang berisi semua material hasil penelitan Tugas Akhir yang tidak dimuat atau disertakan di artikel jurnal. Dalam kertas kerja ini akan dijelaskan mengenai literature review, dataset yang digunakan, serta langkah-langkah perancangan, tahapan implementasi dan hasil pengujian penelitian.

Pendahuluan

Kemajuan teknologi saat ini terus berkembang hingga menjadi bagian dalam kehidupan manusia. Dengan adanya kemajuan yang semakin meningkat, teknologi menjadi peran penting untuk saling bertukar informasi maupun berkomunikasi [1]. Hal tersebut tentunya menuntut manusia untuk lebih melek terhadap teknologi sehingga dapat dimanfaatkan untuk mempermudah kehidupan mereka secara umum karena tidak ada batasan terhadap penggunaan teknologi. Informasi terhadap teknologi yang didapatkan contohnya seperti dalam mencari sebuah pekerjaan, tentunya informasi penting terkait persyaratan serta ujian test yang akan dihadapi.

Permasalahan yang terjadi saat ini, perusahaan dalam melaksanakan seleksi calon karyawan masih memberlakukan serangkaian test secara manual atau tertulis. Tentunya dengan sistem tersebut akan sangat membutuhkan waktu yang cukup lama ketika melakukan koreksi terhadap hasil akhir dalam menghitung score test calon karyawan [4]. Pada era yang serba cepat dan modern ini waktu sangat dijadikan prioritas, dimana menghargai waktu adalah prinsip setiap orang [5]. Hal ini menjadikan sebuah harapan besar perusahaan dalam meningkatkan efektivitas pelayanan dalam sistem penerimaan calon karyawan.

Pada penelitian sebelumnya menjelaskan tentang sistem operasional dalam transaksi layanan pasien hanya mencatat laporan terkait data. Sehingga menjadi

http://digilib.mercubuana.ac.id/

kurang optimal karena data yang menumpuk dalam database server tidak memberikan sebuah informasi. Penggunaan teknik data mining metode k-means clustering untuk mengelompokkan data rekam medis pada pasien telah berhasil melakukan pengelompokkan data penyebaran penyakit sebanyak 253 data dapat diselesaikan dalam waktu 0.06 detik [6].

Penelitian kedua, dalam menentukan klasifikasi individu untuk di kelompokkan dalam kategori yang sama maupun berbeda berdasarkan perbandingan kuantitatif variable diperlukan menerapkan pengelompokkan proses. Penerapan yang dilakukan pada penelitian tersebut adalah proses clustering dataset UN SMA 2019 bidang IPA, IPS, dan program bahasa di Kabupaten Banyumas. Hasil dari *clustering* tersebut menunjukkan bahwa sekolah di *cluster* 2 memiliki pendidikan yang berkualitas dibanding dengan *cluster* 1 dan 3 [7].

Penelitan ketiga, untuk mendapatkan sebuah informasi penting terhadap data mahasiswa perlu melakukan pengolahan data terdahulu. Dalam mencari data yang terpilih perlu diterapkan sebuah metode dan teknik tertentu. Contoh penerapan yang dilakukan peneliti menggunakan metode *clustering k-means* dengan hasil mampu mengelompokkan data-data, baik yang sama maupun berbeda. *Cluster* yang terbentuk mengacu pada program studi yang paling banyak diminati calon mahasiswa. Hasil yang diterapkan pada penelitian tersebut memperlihatkan bahwa penerapan algoritma *k-means* dalam mengelompokkan data mahasiswa yang sebelumnya tidak terstruktur menjadi terstruktur [1].

Dalam penelitian ini akan dibahas mengenai sistem yang akan dibangun menggunakan website sebagai alternative untuk melaksanakan test secara sistem yang akan memberikan kemudahan perusahaan dalam melakukan penilaian hasil test calon karyawan dengan cepat dalam memberi keputusan. Metode *clustering* merupakan metode yang cocok digunakan dalam sistem pengelompokkan, dan *k-means* adalah *clustering* yang memiliki kecepatan dan tingkat akurasi yang tinggi [8]. Penerapan yang dilakukan adalah dengan menghitung *score* masing-masing kategori test yang sudah dilakukan. Setelah melalui proses tersebut, ditentukan 2 buah *cluster* untuk mengelompokkan kelayakan penerimaann calon karyawan dengan output diterima dan tidak lolos.

http://digilib.mercubuana.ac.id/

Rumusan Masalah

Berdasarkan latar belakang yang sudah disampaikan maka dapat ditentukan rumusan masalah sebagai berikut :

1. Bagaimana membuat sistem untuk menentukan kelayakan calon karyawan dalam membantu mengambil keputusan menggunakan *clustering k-means*?
2. Bagaimana cara mengolah data hasil perhitungan untuk dapat ditampilkan dalam bentuk dashboard pada aplikasi web?

Tujuan dan Manfaat

Tujuan

1. Mengetahui penentuan kelayakan diterimanya calon pegawai pada sistem keputusan menggunakan *clustering k-means.*
2. Mengolah data hasil dari pemrosesan untuk dapat ditampilkan dalam bentuk dashboard pada aplikasi berbasis web.

Manfaat

1. Memberikan kemudahan pada pihak keputusan untuk menentukan kelayakan diterimanya calon karyawan.
2. Mengetahui *history* nilai data hasil pemrosesan calon karyawan pada sistem yang sudah dibangun.

UNIVERSITAS MERCU BUANA

Batasan Masalah

Dari rumusan masalah yang ada maka dapat diberi batasan-batasan sehingga pembahasan penelitian akan lebih terarah. Adapun batasan masalah pada penelitian adalah :

1. Data yang digunakan merupakan data dummy *score test* penerimaan calon karyawan.
2. Data yang diujikan adalah *history* hasil *score test* calon karyawan yang telah melaksanakan test pada aplikasi CBT.
3. Metode yang digunakan adalah *clustering k-means*.
4. Hasil dari pemrosesan *clustering* akan ditampilkan dalam bentuk aplikasi berbasis web.

http://digilib.mercubuana.ac.id/