

**PENGELOMPOKAN PEMINATAN OUTLINE TUGAS AKHIR  
DENGAN MENGGUNAKAN ALGORITMA K-MEANS  
PADA AMIK MI BSI JAKARTA**



**TESIS**

**ACHMAD MAEZAR BAYU AJI**

**14001136**

**PROGRAM PASCASARJANA MAGISTER ILMU KOMPUTER  
SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER  
NUSA MANDIRI  
JAKARTA  
2015**

**PENGELOMPOKAN PEMINATAN OUTLINE TUGAS AKHIR  
DENGAN MENGGUNAKAN ALGORITMA K-MEANS  
PADA AMIK MI BSI JAKARTA**



**TESIS**

Diajukan sebagai salah satu syarat untuk memperoleh gelar Magister Ilmu Komputer  
(M.Kom)

**ACHMAD MAEZAR BAYU AJI**

**14001136**

**PROGRAM PASCASARJANA MAGISTER ILMU KOMPUTER  
SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER  
NUSA MANDIRI  
JAKARTA  
2015**

**SURAT PERNYATAAN ORISINALITAS**

Yang bertanda tangan di bawah ini, saya:

Nama : Achmad Maezar Bayu Aji  
NIM : 14000852  
Program Studi : Magister Ilmu Komputer  
Jenjang : Strata Dua (S2)

Konsentrasi : *Management Information System*

Dengan ini menyatakan bahwa Tesis yang telah saya buat dengan judul: **“PENGELOMPOKAN PEMINATAN OUTLINE TUGAS AKHIR DENGAN MENGGUNAKAN ALGORITMA K-MEANS PADA AMIK MI BSI JAKARTA”**, adalah asli (orsinil) atau tidak plagiat (menjiplak) dan belum pernah diterbitkan/dipublikasikan dimanapun dan dalam bentuk apapun.

Demikianlah surat pernyataan ini saya buat dengan sebenar-benarnya tanpa ada paksaan dari pihak manapun juga. Apabila dikemudian hari ternyata saya memberikan keterangan palsu dan atau ada pihak lain yang mengklaim bahwa Skripsi yang telah saya buat adalah hasil karya milik seseorang atau badan tertentu, saya bersedia diproses baik secara pidana maupun perdata dan kelulusan saya dari **Program Pasca Sarjana Magister Ilmu Komputer Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri** dicabut/dibatalkan.

Dibuat di : Jakarta  
Pada tanggal : 11 Desember 2015  
menyatakan,

*Materai 10.000*

**Achmad Maezar Bayu Aji**

## HALAMAN PENGESAHAN

Tesis ini diajukan oleh:

Nama : Achmad Maezar Bayu Aji  
NIM : 14001136  
Program Studi : Magister Ilmu Komputer  
Jenjang : Strata Dua (S2)  
Konsentrasi : *Management Information System*  
Judul Tugas Akhir : PENGELOMPOKAN PEMINATAN OUTLINE TUGAS  
AKHIR DENGAN MENGGUNAKAN ALGORITMA  
K- MEANS PADA AMIK MI BSI JAKARTA

Telah berhasil dipertahankan dihadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Magister Ilmu Komputer (M.Kom) pada Program Pascasarjana Magister Ilmu Komputer Sekolah Tinggi Manajemen Informatika Dan Komputer Nusa Mandiri .

Jakarta, 11 Desember 2015  
Pascasarjana Magister Ilmu Komputer  
STMIK Nusa Mandiri  
Direktur

**Prof. Dr. Ir. Kaman Nainggolan, MS**

Penguji I : Dr. Dwiza Riana, M.Kom, M.M, S.Si ( )

Penguji II : Drs. Bambang Setiarso, M.A ( )

Penguji III/  
Pembimbing : Dr. Sularso Budilaksono, M.Kom ( )

## KATA PENGANTAR

Puji syukur alhamdulillah, penulis panjatkan kehadiran Allah SWT yang telah melimpahkan rahmat dan karunia-Nya, sehingga pada akhirnya penulis dapat menyelesaikan tesis ini tepat pada waktunya. Tesis yang diberi judul “Pengelompokan Peminatan Outline Tugas Akhir Dengan Menggunakan Algoritma K-Means Pada AMIK MI BSI Jakarta” ini dibuat sebagai salah satu syarat untuk mendapatkan gelar Magister Ilmu Komputer (M.Kom) pada Program Pascasarjana Magister Ilmu Komputer Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri (STMIK Nusa Mandiri). Penulis menyadari bahwa tanpa bimbingan dan dukungan dari semua pihak dalam pembuatan tesis ini, maka penulis tidak dapat menyelesaikan tesis ini tepat pada waktunya. Untuk itu ijinkanlah penulis pada kesempatan ini untuk mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Bapak H. Mochamad Wahyudi, M.M, M.Kom, M.Pd, selaku Ketua STMIK Nusa Mandiri Jakarta.
2. Bapak Prof. Dr. Ir. Kaman Nainggolan, MS, selaku Direktur Pascasarjana Magister Ilmu Komputer STMIK Nusa Mandiri Jakarta.
3. Bapak Dr. Sularso Budilaksono, M.Kom, selaku pembimbing tesis yang telah menyediakan waktu, pikiran dan tenaga dalam membimbing penulis dalam menyelesaikan tesis ini.
4. Seluruh staff pengajar (dosen) Program Pascasarjana Magister Ilmu Komputer STMIK Nusa Mandiri yang telah memberikan pelajaran yang berarti bagi penulis selama menempuh studi.
5. Seluruh staff dan karyawan Program Pascasarjana Magister Ilmu Komputer STMIK Nusa Mandiri yang telah melayani penulis dengan baik selama kuliah.
6. Orang tua dan saudara yang selalu memberikan doa, semangat dan dukungan yang tak pernah henti.
7. Rekan-rekan seperjuangan selama kuliah di Program Pascasarjana Magister Ilmu Komputer STMIK Nusa Mandiri.
8. Semua pihak yang telah membantu penulis dalam menyelesaikan tesis yang tidak dapat disebutkan satu per satu.

Penulis menyadari bahwa penulisan tesis ini masih jauh dari sempurna, untuk itu penulis

mohon kritik dan saran dari semua pihak yang sangat diharapkan. Semoga tesis ini bermanfaat bagi semua pihak yang membutuhkan dan bagi penulis khususnya.

Jakarta, 11 Desember 2015

Achmad Maezar Bayu Aji

Penulis

## SURAT PENYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Yang bertanda tangan di bawah ini, saya:

Nama : Achmad Maezar Bayu Aji  
NIM : 14001136  
Program Studi : Magister Ilmu Komputer  
Jenjang : Strata Dua (S2)  
Konsentrasi : *Management Information System*  
Jenis Karya : Tesis

Demi perkembangan ilmu pengetahuan, dengan ini menyetujui untuk memberikan ijin kepada pihak Program Pasca Sarjana Magister Ilmu Komputer Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri (STMIK Nusa Mandiri) **Hak Bebas Royalti Non-Eksklusif** (*Non-Exclusive Royalti-Free Right*) atas karya ilmiah kami yang berjudul: “Penerapan Exponential Smoothing untuk Transformasi Data dalam Meningkatkan Akurasi Neural Network pada Prediksi Harga Emas”, beserta perangkat yang diperlukan (apabila ada). Dengan **Hak Bebas Royalti Non-Eksklusif** ini pihak STMIK Nusa Mandiri berhak menyimpan, mengalih media atau bentuk-kan, mengelolaknya dalam pangkalan data (*database*), mendistribusikannya dan menampilkan atau mempublikasikannya di internet atau media lain untuk kepentingan akademis tanpa perlu meminta ijin dari kami selama tetap mencantumkan nama kami sebagai penulis atau pencipta karya ilmiah tersebut. Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak STMIK Nusa Mandiri, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini. Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 11 Desember 2015  
Yang menyatakan

Achmad Maezar Bayu Aji

## ABSTRAK

Nama : Achmad Maezar Bayu Aji  
NIM : 14000831  
Program Studi : Magister Ilmu Komputer  
Jenjang : Strata Dua (S2)  
Konsentrasi : *Management Information System*  
Judul : Pengelompokan Peminatan Outline Tugas Akhir Dengan Menggunakan K- Means Pada AMIK MI BSI Jakarta

Sebelum membuat tugas akhir, mahasiswa memilih outline untuk menentukan minat dan tujuan dari pembuatan tugas akhir tersebut. Pada AMIK MI BSI Jakarta mempunyai banyak cabang dan banyak peminatan outline. Salah satu masalah dalam pemilihan outline tugas akhir di AMIK MI BSI Jakarta adalah terlalu banyaknya data peminatan outline dan cabang. Tujuan dari penelitian ini untuk mengetahui dan mengelompokkan outline pendaftaran tugas akhir yang paling banyak diminati dan yang paling sedikit diminati untuk nantinya akan diperbaharui lagi peminatan outline tugas akhir di AMIK MI BSI Jakarta. Algoritma yang digunakan untuk mengelompokkan data tersebut menggunakan algoritma k-means K-Means adalah suatu metode penganalisaan data. Metode k-means berusaha mengelompokkan data yang ada ke dalam beberapa kelompok, dimana data dalam satu kelompok mempunyai karakteristik yang sama satu sama lainnya dan mempunyai karakteristik yang berbeda dengan data yang ada di dalam kelompok yang lain. Atribut yang digunakan dalam pengelompokan peminatan outline tugas akhir adalah jenis outline pendaftaran tugas akhir, cabang AMIK MI BSI Jakarta, nilai tugas akhir mahasiswa. Cluster yang terbentuk ada 3 cluster, cluster pertama mempunyai 9 anggota data mahasiswa, cluster kedua mempunyai 8 anggota data mahasiswa, cluster ketiga mempunyai anggota data mahasiswa.

**Kata Kunci: Tugas Akhir, K-Means, Peminatan Outline**

## ABSTRACT

Nama : Achmad Maezar Bayu Aji  
NIM : 14000831  
Program Studi : Magister Ilmu Komputer  
Jenjang : Strata Dua (S2)  
Konsentrasi : *Management Information System*  
Judul : *Final Project Outline Specification Grouping Using K-Means Algorithm*  
On  
*AMIK MI BSI Jakarta*

*Before making the final project, students choose an outline to determine the interests and goals of making the final project. At AMIK MI BSI Jakarta, it has many branches and many outline interests. One of the problems in selecting the final project outline at AMIK MI BSI Jakarta is that there is too much data on outline and branch specialization. The purpose of this research is to identify and classify the outlines for the registration of the most popular and least desirable final assignments so that later the interest in the final project outlines at AMIK MI BSI Jakarta will be renewed. The algorithm used to group the data using the k-means algorithm K-Means is a data analysis method. The k-means method tries to group the existing data into several groups, where the data in one group has the same characteristics as each other and has different characteristics from the data in other groups. The attributes used in grouping the final project outline specialization are the type of final assignment registration outline, AMIK MI BSI Jakarta branch, student final project grades. There are 3 clusters formed, the first cluster has 9 student data members, the second cluster has 8 student data members, the third cluster has student data members.*

**Key Words:** *Final Project, K-Means, Outline Specialization*

## DAFTAR ISI

Halaman

HALAMAN SAMPUL .....	i
HALAMAN JUDUL .....	ii
HALAMAN PERNYATAAN ORISINALITAS.....	iii
HALAMAN PENGESAHAN.....	iv
KATA PENGANTAR .....	v
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS .....	vii
ABSTRAK .....	viii
ABSTRACT.....	ix
DAFTAR ISI.....	x
DAFTAR TABEL.....	xi
DAFTAR GAMBAR .....	xii
BAB 1. PENDAHULUAN	
1.1. Latar Belakang .....	1
1.2. Masalah Penelitian .....	2
1.3. Tujuan dan Manfaat Penelitian .....	3
1.4. Sistematika Penulisan.....	3
BAB 2. LANDASAN TEORI	
2.1. Tinjauan Pustaka .....	5
2.2. Tinjauan Studi .....	9
2.3. Kerangka Pemikiran.....	10
BAB 3. METODE PENELITIAN	
3.1. Perancangan Penelitian .....	13
3.2. Pengolahan Data Awal .....	14
3.3. Metode Yang Diusulkan .....	14
BAB 4. HASIL PENELITIAN DAN PEMBAHASAN .....	
4.1. Hasil .....	15
4.2. Pembahasan .....	16
BAB 5. PENUTUP	
5.1. Kesimpulan .....	26
5.2. Saran.....	27
DAFTAR PUSTAKA .....	28

## DAFTAR TABEL

Tabel 2.1 Tinjauan Studi Terdahulu .....	9
Tabel.4.1. Tabel Data Mahasiswa Bimbingan Periode 2014-1 AMIK MI BSI Jakarta .....	18
Tabel 4.2. Inisialisasi tabel nilai.....	19
Tabel 4.3. Titik Pusat Awal Cluster .....	20
Tabel 4.4 Contoh Hasil Perhitungan Data Mahasiswa Ke Setiap Titik Cluster ...	22
Tabel 4.5. Titik Pusat Cluster Baru .....	23
Tabel 4.6. Hasil Perhitungan Data Mahasiswa Dengan Titik Pusat Cluster Baru	25

## DAFTAR GAMBAR

Gambar 2.1 Konsep Clustering .....	6
Gambar 2.2 Kerangka Pemikiran .....	10

## **DAFTAR LAMPIRAN**



# **BAB I**

## **PENDAHULUAN**

### **1.1. Latar Belakang**

Dalam penentuan kelulusan suatu universitas atau akademik salah satu faktor yang menentukan adalah tugas akhir. Tugas akhir merupakan hasil tertulis dari pelaksanaan suatu penelitian, yang dibuat untuk pemecahan masalah tertentu dengan menggunakan kaidah- kaidah yang berlaku dalam bidang ilmu tersebut. Syarat untuk pendaftaran tugas akhir di AMIK MI BSI Jakarta, mahasiswa tersebut harus menyelesaikan 110 sks dan tidak ada mata kuliah unggulan yang di her, dan ipk minimal 2,75.

Kemudian, mahasiswa memilih outline untuk menentukan minat dan tujuan dari pembuatan tugas akhir tersebut. Di AMIK MI BSI Jakarta untuk outline tugas akhir diantaranya adalah : Jaringan Komputer, Pemrograman - Berorientasi Bisnis - Web Programming, Pemrograman - Berorientasi Science, Pemrograman - Animasi Interaktif, Pembuatan Alat, Pemrograman Berorientasi Bisnis Visual, Perancangan Sistem Berorientasi Objek, Pemrograman Berorientasi Objek, Aplikasi Program Akuntansi, Perancangan Sistem, Proyek Sistem Informasi Berorientasi Objek.

Dan untuk AMIK MI BSI Jakarta dibagi dalam beberapa cabang, yaitu cabang margonda, fatmawati, Kalimalang, Kramat- Salemba, Jatiwaringin, Ciledug, Wolter Monginsidi, Cengkareng, Pondok Aren, Dewi Sartika, Pemuda (Rawamangun).

## **1.2. Masalah Penelitian**

### **1.2.1. Identifikasi Masalah**

Sebelum membuat tugas akhir, mahasiswa memilih outline untuk menentukan minat dan tujuan dari pembuatan tugas akhir tersebut. Pada AMIK MI BSI Jakarta banyak cabang dan banyak peminatan outline dan dalam penentuan kelulusan suatu universitas atau akademik salah satu faktor yang menentukan adalah tugas akhir. Tugas akhir merupakan hasil tertulis dari pelaksanaan suatu penelitian, yang dibuat untuk pemecahan masalah tertentu dengan menggunakan kaidah-kaidah yang berlaku dalam bidang ilmu tersebut. Syarat untuk pendaftaran tugas akhir di AMIK MI BSI Jakarta, mahasiswa tersebut harus menyelesaikan 110 sks dan tidak ada mata kuliah unggulan yang di her, dan ipk minimal 2,75. Kemudian, mahasiswa memilih outline untuk menentukan minat dan tujuan dari pembuatan tugas akhir tersebut.

### **1.2.2. Batasan Masalah**

Penelitian ini akan melakukan pengelompokan data berdasarkan pada data mahasiswa AMIK MI BSI Jakarta dengan data bimbingan periode 2014-1 di seluruh cabang AMIK MI BSI Jakarta. Atribut yang digunakan dalam pengelompokan peminatan outline tugas akhir adalah jenis outline pendaftaran tugas akhir, cabang AMIK MI BSI Jakarta, nilai tugas akhir mahasiswa.

### **1.2.3. Rumusan Masalah**

Salah satu masalah dalam pemilihan outline tugas akhir di AMIK MI BSI Jakarta adalah terlalu banyaknya data peminatan outline dan cabang. Sehingga dalam penelitian ini focus pada metode data mining dengan kasus pengelompokan data (clustering). Menurut Chandra, dkk (2014) "Adanya data dalam besar

memungkinkan metode data mining dengan teknik clustering yang dapat mengelompokkan data ke dalam beberapa kelompok yang diinginkan. Teknik clustering yang digunakan yaitu K-Means”.

### **1.3. Tujuan dan Manfaat Penelitian**

#### **1.3.1. Tujuan**

Tujuan dari penelitian ini untuk mengetahui dan mengelompokkan outline pendaftaran tugas akhir yang paling banyak diminati dan yang paling sedikit diminati untuk nantinya akan diperbaharui lagi peminatan outline tugas akhir di AMIK MI BSI Jakarta.

#### **1.3.2. Manfaat**

Manfaat dari penelitian ini mahasiswa bisa lebih fokus dalam memilih outline tugas akhir karena sudah mengetahui outline yang paling banyak diminati dan mana yang sedikit, juga untuk institusi lebih diperbaharui dan direvisi kembali outline pendaftaran tugas akhir mahasiswa.

### **1.4. Sistematika Penulisan**

Sistematika penulisan tesis ini terdiri dari 5 (lima) bab, di mana tiap bab terdiri dari beberapa sub bab yaitu sebagai berikut :

#### **BAB I PENDAHULUAN**

Bab ini menguraikan tentang latar belakang, masalah penelitian, identifikasi masalah, batasan masalah, rumusan masalah, tujuan dan manfaat penelitian, serta sistematika penulisan.

## **BAB II LANDASAN TEORI DAN KERANGKA PEMIKIRAN**

Bab ini membahas tentang tinjauan pustaka, tinjauan studi terdahulu yang relevan, tinjauan obyek penelitian, kerangka konsep penelitian, hipotesis.

## **BAB III METODOLOGI PENELITIAN**

Bab ini membahas tentang metode penelitian, langkah-langkah penelitian, metode pemilihan sampel, metode pengumpulan data, instrumen penelitian, metode analisis deployment phase.

## **BAB IV HASIL DAN PEMBAHASAN**

Membahas mengenai hasil penelitian, pembahasan dan implikasi penelitian.

## **BAB V PENUTUP**

Memberikan kesimpulan dan saran berdasarkan bab-bab sebelumnya untuk kepentingan penelitian selanjutnya.

## **BAB II**

### **LANDASAN TEORI**

#### **2.1. Tinjauan Pustaka**

##### **2.1.1. Data**

Dalam *Webster's New World's Dictionary* tertulis bahwa datum: *something known or assumed*. Artinya, datum (bentuk tunggal data) merupakan suatu yang diketahui/dianggap. Dengan demikian, data dapat memberi gambaran tentang suatu keadaan atau persoalan. Sedangkan, data menurut kamus Oxford Dictionary adalah The Facts. Jadi, dapat disimpulkan bahwa data adalah sesuatu yang nyata diketahui atau dianggap yang dipakai untuk keperluan suatu analisa, diskusi, presentasi ilmiah atau tes statistik (Supranto, 2000)

##### **2.1.2. Data Mining**

Menurut Kusrini dan Lutfi (2009) "Data mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam database. Data mining adalah proses yang menggunakan Teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terakut dari berbagai database besar.

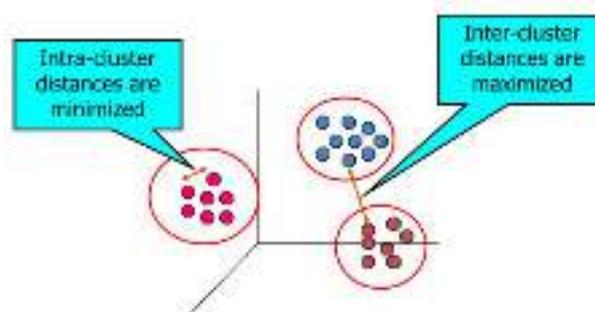
Data mining adalah suatu metode pengolahan data untuk menemukan pola yang tersembunyi dari data tersebut. Hasil dari pengolahan data dengan metode data mining ini dapat digunakan untuk mengambil keputusan di masa depan. Data mining ini juga dikenal dengan istilah pattern recognition (Santosa, 2007). Data mining merupakan metode pengolahan data berskala besar oleh karena itu data mining ini memiliki peranan penting dalam bidang industri, keuangan, cuaca, ilmu

dan teknologi. Secara umum kajian data mining membahas metode-metode seperti, clustering, klasifikasi, regresi, seleksi variable, dan market basket analisis (Santosa, 2007).

### 2.1.3. Clustering

Menurut Zaki (2014) “Clustering adalah suatu tugas yang membagi poin menjadi beberapa kelompok-kelompok alami yang disebut klaster, sehingga poin-poin dalam kelompok sangat mirip, sedangkan poin-poin yang diluar cluster sedikit berbeda. Tergantung pada data dan karakteristik klaster yang diinginkan, ada berbagai jenis pengelompokan paradigma seperti berbasis perwakilan, hirarkis, berdasarkan kepadatan, berdasarkan grafik dan spektral clustering.

Clustering melakukan pengelompokan data yang didasarkan pada kesamaan antar objek, oleh karena itu klasterisasi digolongkan sebagai metode *unsupervised learning*.



**Gambar 2.1 Konsep Clustering**

### 2.1.4. K-Means

Menurut Baginda Harahap K-Means Clustering adalah metode untuk mengkategorikan atau pengelompokan sekelompok objek sesuai dengan atribut yang sama atau karakteristik ke dalam sejumlah groups (jumlah bilangan bulat positif). Ini mendefinisikan sebuah cluster oleh massa yang yang mewakili mean

dari cluster. Dalam data mining analisis klaster populer adalah K- means.

Algoritma K-Means merupakan salah satu algoritma parsial karena K-Means didasarkan pada penentuan jumlah awal grup dengan menetapkan nilai centroid awal. Algoritma KMeans menggunakan proses iteratif untuk mendapatkan database cluster. Dibutuhkan jumlah cluster awal yang diinginkan sebagai input dan menghasilkan jumlah akhir cluster sebagai output. Jika suatu algoritma diperlukan untuk menghasilkan K cluster, akan ada K huruf pertama dan K huruf terakhir. Metode K-Means akan secara acak memilih sampel k sebagai titik awal pusat. Jumlah iterasi untuk mencapai cluster pusat akan dipengaruhi oleh calon cluster pusat asli secara acak, atau jika posisi pusat baru tidak berubah. Nilai K yang dipilih sebagai pusat awal akan dihitung menggunakan rumus jarak Euclidean, yang melibatkan pencarian jarak terdekat antara titik pusat dan data/objek. Data dengan jarak terpendek atau terdekat dengan centroid akan membentuk cluster.

Menurut Budi (2007) langkah-langkah melakukan clustering dengan metode K-Means adalah sebagai berikut[3]:

- a. Pilih jumlah cluster k.
- b. Inisialisasi k pusat cluster ini bisa dilakukan dengan berbagai cara. Namun yang paling sering dilakukan adalah dengan cara random. Pusat-pusat cluster diberi nilai awal dengan angka-angka random,
- c. Alokasikan semua data/ objek ke cluster terdekat. Kedekatan dua objek ditentukan berdasarkan jarak kedua objek tersebut. Demikian juga kedekatan suatu data ke cluster tertentu ditentukan jarak antara data dengan pusat cluster. Dalam tahap ini perlu dihitung jarak tiap data ke tiap pusat cluster. Jarak paling antara satu data dengan satu cluster tertentu akan menentukan suatu data masuk dalam cluster

mana. Untuk menghitung jarak semua data ke setiap titik pusat cluster dapat menggunakan teori jarak Euclidean yang dirumuskan sebagai berikut:  $D(i,j)$

Reference source not

found. Error! Reference source not found. Error! Reference source not found.

Dimana:

$D(i,j)$  = Jarak data ke  $i$  ke pusat cluster  $j$

$X_{ki}$  = Data ke  $i$  pada atribut data ke  $k$

$X_{kj}$  = Titik pusat ke  $j$  pada atribut ke  $k$  Error!

Reference source not found.

d. Hitung kembali pusat cluster dengan keanggotaan cluster yang sekarang.

Pusat cluster adalah rata-rata dari semua data/ objek dalam cluster tertentu. Jika dikehendaki bisa juga menggunakan median dari cluster tersebut. Jadi rata-rata (mean) bukan satu-satunya ukuran yang bisa dipakai.

Tugaskan lagi setiap objek memakai pusat cluster yang baru. Jika pusat cluster tidak berubah lagi maka proses clustering selesai. Atau, kembali ke langkah nomor 3 sampai pusat cluster tidak berubah lagi.

## 2.2. Tinjauan Studi

**Tabel 2.1 Tinjauan Studi Terdahulu**

No.	Peneliti, Judul	Tujuan	Metodologi	Hasil
1	Johan Oscar Ong, IMPLEMENTASI ALGORITMA K-MEANS CLUSTERING UNTUK MENENTUKAN STRATEGI MARKETING PRESIDENT UNIVERSITY	bagian marketing dapat melakukan pemasaran dengan strategi yang tepat untuk mendapatkan calon mahasiswa baru	Dengan metode Hard K-Means Clustering	Melakukan promosi dengan mengirim tim marketing yang sesuai dengan jurusan yang paling banyak diminati dan pada kota-kota di Indonesia yang didasarkan pada tingkat kemampuan akademik dari calon mahasiswa
2	Baginda Harahap, Penerapan Algoritma K-Means Untuk Menentukan Bahan Bangunan Laris (Studi Kasus Pada UD. Toko Bangunan YD Indarung)	mengumpulkan dan mengelompokkan bahan bangunan tersebut menjadi beberapa kategori yaitu sangat laris, laris dan kurang laris	algoritma K-Means	Hasilnya mendapatkan 3 kategori penjualan bahan bangunan, yaitu kurang laris, sangat laris dan laris
3	Ni Putu Eka Merliana, Ernawati, Alb. Joko Santoso, ANALISA PENENTUAN JUMLAH CLUSTER TERBAIK PADA METODE K-MEANS CLUSTERING	menghasilkan hasil cluster yang optimal dan dapat membantu menghasilkan informasi yang dibutuhkan	Elbow dan K-Means	jumlah cluster yang ideal adalah K=3 dan dijadikan default cluster untuk menentukan karakteristik dari data-data tersebut
4	Tutik Khotimah, PENGELOMPOKAN SURAT DALAM AL QUR'AN MENGGUNAKAN ALGORITMA K-MEANS	untuk melakukan pengelompokan surat-surat dalam Al Qur'an berdasarkan jumlah ayat, jumlah ruku' dan lama waktu membaca surat dengan menggunakan algoritma K-means	K-Means	Pengelompokan menghasilkan 4 cluster dengan anggota cluster_0 sebanyak 18 surat, cluster_1 sebanyak 26 surat, cluster_2 sebanyak 7 surat, cluster_3 sebanyak 63 surat. Pengelompokan ini telah mencapai nilai optimal pada iterasi ke-14

5	Slamet Handoko, Eko Sediono, Suhartono, Sistem Informasi Geografis Berbasis Web untuk Pemetaan Sebaran Alumni Menggunakan Metode K-Means	membangun sistem informasi geografis berbasis web untuk pemetaan sebaran alumni dengan menggunakan metode K-Means	SIG untuk simulasi dalam bentuk data spasial dan metode K-Means untuk menganalisis kluster sebaran alumni	berupa SIG berbasis web untuk pemetaan sebaran alumni Polines
---	--	---	---	---

### 2.3. Kerangka Pemikiran

Tahapan dalam poin ini yakni dimana kerangka kerja berfungsi buat mendeskripsikan proses yg sistematis pada penelitian ini, menurut termin menganalisa kebutuhan data sampai menyimpulkan hasil.



Gambar 2.2 Kerangka Pemikiran

#### 1. Identifikasi Masalah

Identifikasi ini dilakukan menjadi langkah awal penelitian. Jadi, secara ringkas, identifikasi merupakan mendefinisikan kasus penelitian. Selain itu, identifikasi kasus jua bisa diartikan menjadi proses & output sosialisasi kasus atau inventarisasi kasus. Makanya identifikasi ini sebagai langkah awal penelitian yg penting.

## 2. Menganalisis Masalah

Mendeskripsi masalah kasus secara naratif menggunakan memakai data / keterangan yg jeladan spesifik. Mengembangkan karena-karena yg mungkin berdasarkan masalah itu menggunakan memakai pengalaman & nalar berdasarkan diskripsi kasus tersebut. Menemukan karena yg sesungguhnya menggunakan menguji secara kritis buat menandakan data/keterangan yg ada. Terdapat beberapa langkah buat mengidentifikasi perseteruan, mulai berdasarkan merumuskan kasus hingga melakukan ferifikasi terhadap penyebab perseteruan tersebut, & ini akan diuraikan lebih lanjut.

## 3. Mempelajari Literatur

Studi Literatur merupakan cara buat menuntaskan dilema menggunakan menelusuri asal-asal goresan pena yg pernah dibentuk sebelumnya. Dengan istilah lain, kata Studi Literatur ini jua sangat familiar menggunakan sebutan studi pustaka. Sumber-asal yg diteliti pun nir boleh sembarangan. Sebab nir seluruh output penelitian sanggup dijadikan acuan. Selain itu seseorang peneliti bisa memperoleh warta mengenai penelitian-penelitian homogen atau yg terdapat kaitannya menggunakan penelitiannya. Dan penelitian-penelitian yg sudah dilakukan sebelumnya

## 4. Melakukan Preprocessing dan Transformasi Data

Preprocessing data adalah teknik awal data mining buat membarui data mentah atau biasa dikenal menggunakan raw data yg dikumpulkan menurut banyak sekali asal sebagai keterangan yg lebih higienis & sanggup dipakai buat pengolahan selanjutnya. Missing value adalah data yg nir seksama lantaran keterangan yg hilang mengakibatkan keterangan yg terdapat pada dalamnya nir relevan. Pencilan

& data keliru ini berisi keterangan yg nir berarti. Beberapa penyebab adanya data noise merupakan lantaran kesalahan insan berupa kesalahan anugerah label & perkara lain selama pengumpulan data.

#### 5. Menguji Hasil

Tata cara atau metode pembuktian tersebut harus dengan cara yang sesuai, artinya proses pembuktian tersebut ada tata cara atau metodenya. pembuktian adalah tiap-tiap bahan, proses, prosedur, kegiatan, sistem, perlengkapan atau mekanisme yang digunakan dalam produksi dan pengawasan mutu.

#### 6. Menyimpulkan Hasil

pernyataan yang diambil secara ringkas dari keseluruhan hasil pembahasan atau analisis. menjadi bagian terpenting dalam setiap penelitian, Namun tidak sedikit dari penulis mengalami kesulitan dalam menulis kesimpulan. Hal ini bisa saja karena penulis sulit memilih atau menemukan sesuatu yang menarik atau berguna untuk dijadikan kesimpulan

## BAB III

### METODOLOGI PENELITIAN

#### 3.1. Perancangan Penelitian

Dalam melakukan penelitian metode yang digunakan adalah metode K-Means. Menurut Ong (2013) dengan menggunakan metode ini, data data yang telah didapatkan dapat dikelompokkan kedalam beberapa cluster berdasarkan kemiripan dari data-data tersebut, sehingga data-data yang memiliki karakteristik yang sama dikelompokkan dalam satu cluster dan yang memiliki karakteristik yang berbeda dikelompokkan dalam cluster yang lain yang memiliki karakteristik yang sama. Tahapan – tahapan *clustering* dalam metode K-Means adalah sebagai berikut:

1. Pilih jumlah *cluster* K

Dalam menentukan jumlah cluster K bisa menggunakan beberapa metode dan ada juga yang *random*. Dalam penelitian ini menggunakan cara yang random

2. Inisialisasi nilai K

Inisialisasi adalah pemberian nilai awal pada titik-titik pusat *cluster* dan untuk pemberian nilai awal ini bisa juga dengan penggabungan beberapa algoritma dan yang paling sering digunakan adalah dengan menggunakan metode *random* data.

3. Alokasikan semua data/objek ke *cluster* terdekat

Menempatkan data-data ke dalam setiap cluster. Untuk menempatkan data-data ke dalam setiap *cluster* perlu dihitung jarak tiap data ke tiap pusat *cluster*. Jarak paling antara satu data dengan satu cluster tertentu akan menentukan suatu data masuk dalam *cluster* mana.

4. Hitung Kembali pusat *cluster* dengan keanggotaan *cluster* yang sekarang  
Titik pusat *cluster* yang baru adalah rata-rata dari semua anggota dalam setiap *cluster* tertentu
5. Tugaskan lagi setiap objek memakai pusat *cluster* yang baru.  
Tahapan ini merupakan tahapan terakhir dalam metode K-Means. Dalam tahapan ini akan dilakukan secara berulang hingga titik pusat *cluster* tidak berubah lagi.

### **3.2. Pengolahan Data Awal**

Penelitian ini akan melakukan pengelompokan data berdasarkan pada data mahasiswa AMIK MI BSI Jakarta dengan data bimbingan periode 2014-1. Data – data yang digunakan pada penelitian ini didapat dari bagian kemahasiswaan yang sedang melakukan bimbingan dalam pengerjaan tugas akhir. Data ini juga berisi tentang data mahasiswa bimbingan, dan dalam penelitian ini atribut yang digunakan hanya sebatas, nim, nama mahasiswa, cabang, outline tugas akhir dan nilai. Data yang sudah didapat harus diolah kembali dalam bentuk angka dikarenakan metode algoritma k-means hanya bisa diolah menggunakan angka. Dan data yang sudah diolah ini digunakan untuk menganalisa dan dapat digunakan untuk menyelesaikan permasalahan yang ada dalam pemilihan outline tugas akhir di AMIK MI BSI Jakarta.

### **3.3. Metode Yang Diusulkan**

Metode yang diusulkan dalam penelitian adalah algoritma *clustering* KMeans, atau lebih tepatnya algoritma hard KMeans *clustering*. Pada metode *hard* KMeans *clustering* ini, data yang diperoleh dikelompokkan sesuai dengan kesamaan fitur data masing-masing, sehingga dapat ditemukan informasi yang

tersembunyi dari data tersebut. Berdasarkan analisis dan pengolahan data, dapat ditarik kesimpulan dari penelitian yang dilakukan. Kesimpulan ini menjawab pertanyaan yang dinyatakan di awal

## **BAB IV**

### **HASIL DAN PEMBAHASAN**

#### **4.1. Hasil**

Dalam penelitian ini, titik pusat cluster tidak berubah lagi dan data-data tidak ada lagi yang berpindah ke masing-masing cluster pada iterasi ke-3, iterasi clustering yang terjadi sebanyak 3 kali.

Hasil dari penelitian ini dengan mengelompokkan ke dalam 3 cluster dengan menggunakan algoritma k-means adalah sebagai berikut

Untuk cluster pertama berisi 9 anggota dengan data mahasiswanya terdiri dari 2 dari Fatmawati, Jatiwaringin dan Dewi Sartika. Untuk cabang Kramat- Salemba, Ciledug, dan Warung Jati hanya berisi 1 data mahasiswa. Outline peminatannya didominasi oleh Pemrograman - Berorientasi Bisnis - Web Programming dengan nilai tugas akhir dominan A dengan hanya satu mahasiswa yang mempunyai nilai C.

Untuk cluster kedua berisi 8 anggota dengan data mahasiswanya terdiri dari 3 mahasiswa cabang margonda, 2 mahasiswa cabang Kramat- Salemba dan Jatiwaringin, dan hanya 1 mahasiswa dari cabang Ciledug, dengan outline peminatannya terdiri dari Pemrograman Berorientasi Bisnis Visual, Perancangan Sistem Berorientasi Objek, Pemrograman Berorientasi Objek dan Perancangan Sistem dengan nilai tugas akhir dominan A dan B.

Untuk cluster ketiga hanya berisi 3 anggota dengan data mahasiswanya 2 dari cabang Dewi Sartika dan dari cabang Warung Jati, dengan outline peminatannya terdiri dari Pemrograman Berorientasi Bisnis Visual, Perancangan Sistem dengan

nilai tugas akhir dominan A dan B.

#### 4.2. Pembahasan

Dalam penelitian ini, data diperoleh dari bagian akademik AMIK MI BSI Jakarta berupa data bimbingan periode 2014-1 dan data terkait hasil uji kelulusan. Data-data ini berisi data tentang pembelajaran siswa Dalam penelitian ini atribut yang digunakan dibatasi pada Nimes, nama siswa, cabang, garis besar tugas akhir dan nilai. Berikut ini contoh adalah data mahasiswa bimbingan periode 2014-1 yang ditampilkan dalam tabel 1.1.

**Table.4.1. Tabel Data Mahasiswa Bimbingan Periode 2014-1 AMIK MI  
BSI Jakarta**

No	Nim	Nama Mahasiswa	kd_ cab ang	Kd Outl ine	Nama Outline	NILAI TA
1	12112979	MIFTACHUL BASORI	1	98	Pemrograman Berorientasi Bisnis Visual ( KA & MI )	B
2	12112251	RESTI SEPRIYANI	1	108	Perancangan Sistem (Dekstop/Web)	B
3	18113933	PUSNANDAR	1	101	Pemrograman Berorientasi Objek (Desktop/Web)	B
4	12117758	MOCHAMAD RIFKI	2	42	Pemrograman - Berorientasi Bisnis -Web Programming (Untuk mahasiswa : AMIK)	A
5	12117702	ROBERTUS EKO PRASETYO	2	42	Pemrograman - Berorientasi Bisnis -Web Programming (Untuk mahasiswa : AMIK)	A
6	18112943	YEANNE SITI YANTO	7	98	Pemrograman Berorientasi Bisnis Visual ( KA & MI )	A
7	12112271	NUR HAMIDAH	7	42	Pemrograman - Berorientasi Bisnis - Web Programming (Untuk mahasiswa : AMIK)	C
8	18112718	WULAN DARI	7	98	Pemrograman Berorientasi Bisnis Visual ( KA & MI )	B
9	18113908	YULI NURMALITA	11	100	Perancangan Sistem Berorientasi Objek (Desktop/Web)	A

10	18112505	VIRGINIA RIA UNIKA URMAWATI	11	42	Pemrograman - Berorientasi Bisnis - Web Programming (Untuk mahasiswa : AMIK)	A
11	18111290	AHMAD SOFIYAN	11	42	Pemrograman - Berorientasi Bisnis - Web Programming (Untuk mahasiswa : AMIK)	A
12	18110781	YESSY EVIANTARI	11	101	Pemrograman Berorientasi Objek (Desktop/Web)	A
13	18112382	ASEP ZAKARIA	12	100	Pemrograman Berorientasi Objek (Desktop/Web)	A
14	12117839	IRFAN MAULANA	12	42	Pemrograman - Berorientasi Bisnis - Web Programming (Untuk mahasiswa : AMIK)	A
15	12112436	WAHYUNINGSIH	31	42	Pemrograman - Berorientasi Bisnis - Web Programming (Untuk mahasiswa : AMIK)	A
16	12112430	AFRILIA RAHMAN	31	42	Pemrograman - Berorientasi Bisnis - Web Programming (Untuk mahasiswa : AMIK)	A
17	12110067	ASEP SAEPUDIN	31	108	Perancangan Sistem (Dekstop/Web)	A
18	18111050	SOPIAH SITI NURAENI	31	98	Pemrograman Berorientasi Bisnis Visual ( KA & MI )	B
19	18113684	NANANG RIYADI	32	98	Pemrograman Berorientasi Bisnis Visual ( KA & MI )	A
20	12117817	AHMAD FAUZI	32	42	Pemrograman - Berorientasi Bisnis - Web Programming (Untuk mahasiswa : AMIK)	A

Pada saat melakukan pengolahan data pada algoritma k-means harus digunakan bentuk digital agar data yang menggunakan nilai nominal dapat diganti dengan angka. Dalam data ini, data nama asli, seperti nama siswa, nama cabang dan nama kerangka, dan nilai akan diganti. Nama siswa akan diganti dengan nim, nama cabang akan diganti dengan kode cabang, nama garis akan diganti dengan kode garis besar, dan nilai ta akan diinisialisasi pada tabel di bawah ini.

**Table.4.2. Inisialisasi tabel nilai**

NILAI TA	NILAI
A	1
B	2
C	3
D	4
E	5

Untuk mengelompokan data tersebut perlu dilakukan dalam beberapa tahap. Untuk tahapan-tahapan dalam algoritma k-means clustering adalah sebagai berikut:

Langkah pertama dalam menggunakan algoritma k-means adalah menentukan jumlah cluster yang dibutuhkan. Pada penelitian ini jumlah cluster yang digunakan untuk mengklasifikasikan data adalah 3 cluster.

Pemberian nilai pada titik-titik pusat cluster atau k pusat-pusat cluster untuk menetapkan nilai-nilai ke pusat-pusat cluster k dapat dilakukan dengan berbagai cara, Dalam penelitian ini ditentukan bahwa nilai-nilai awal dari pusat-pusat cluster memiliki bilangan acak. Berikut ini adalah pusat cluster pada Tabel 1.3.

**Table.4.3. Titik Pusat Awal Cluster**

No	Nim	kd_c abang	Kd Outline	NILAI TA
1	12117758	2	42	1
2	18110781	11	101	1
3	18111050	31	98	2

Menempatkan data-data ke dalam setiap cluster. Untuk menempatkan data-data ke dalam setiap cluster perlu dihitung jarak tiap data ke tiap pusat cluster. Jarak paling antara satu data dengan satu cluster tertentu akan menentukan suatu data masuk dalam cluster mana. dalam penelitian ini Untuk menghitung jarak semua data ke setiap titik pusat cluster menggunakan teori jarak Euclidean. Berikut ini adalah contoh perhitungan jarak dari data mahasiswa pertama ke setiap titik pusat cluster: Untuk perhitungan mahasiswa pertama di cluster pertama.

Titik mahasiswa pertama:(1,98,2) Untuk titik pusat cluster pertama : (2, 42, 1).

Perhitungan :

$$D_{(1,1)}=\sqrt{(1-2)^2+(98-42)^2+(2-1)^2}$$

Hasil : 56,01

Jadi, jarak dari data mahasiswa pertama dengan titik pusat cluster pertama adalah 56,01

Untuk perhitungan mahasiswa pertama di cluster kedua.

Titik mahasiswa pertama:(1, 98, 2)

Untuk titik pusat cluster kedua : (11, 101, 1).

Perhitungan :

$$D_{(1,1)}=\sqrt{(1-11)^2+(98-101)^2+(2-1)^2}$$

Hasil : 10,48

Jadi, jarak dari data mahasiswa pertama dengan titik pusat cluster kedua adalah 10,48

Untuk perhitungan mahasiswa pertama di cluster ketiga.

Titik mahasiswa pertama:(1, 98, 2)

Untuk titik pusat cluster ketiga : (31,98,2).

Perhitungan :

$$D_{(1,1)}=\sqrt{(1-31)^2+(98-98)^2+(2-2)^2}$$

Hasil : 30

Jadi, jarak dari data mahasiswa pertama dengan titik pusat cluster ketiga adalah 30

Berdasarkan hasil perhitungan dari ketiga titik cluster tersebut, dapat kita ketahui bahwa jarak data mahasiswa yang pertama dengan ketiga titik cluster tersebut yang paling dekat jaraknya adalah dengan titik pusat cluster yang kedua yaitu sebesar

10,48 jadi dapat kita simpulkan bahwa data mahasiswa pertama masuk kedalam cluster pertama. Berikut ini contoh untuk perhitungan data- data mahasiswa yang ditunjukkan pada tabel 1.4.

**Table.4.4. Contoh Hasil Perhitungan Data Mahasiswa Ke Setiap Titik Cluster**

No	Nim	kd_cabang	Kd Outline	NILAI TA	jarak ke cluster			jarak terdekat dengan cluster
					C1	C2	C3	
1	12112979	1	98	2	56,01785	10,48809	30	2
2	12112251	1	108	2	66,01515	12,24745	31,62278	2
3	18113933	1	101	2	59,01695	10,04988	30,14963	2
4	12117758	2	42	1	0	59,68249	63,07139	1
5	12117702	2	42	1	0	59,68249	63,07139	1
6	18112943	7	98	1	56,22277	5	24,02082	2
7	12112271	7	42	3	5,385165	59,16925	60,93439	1
8	18112718	7	98	2	56,23166	5,09902	24	2
9	18113908	11	100	1	58,69412	1	20,12461	2
10	18112505	11	42	1	9	59	59,47268	1
11	18111290	11	42	1	9	59	59,47268	1
12	18110781	11	101	1	59,68249	0	20,24846	2
13	18112382	12	100	1	58,85576	1,414214	19,13113	2
14	12117839	12	42	1	10	59,00847	59,14389	1
15	12112436	31	42	1	29	62,29767	56,00893	1
16	12112430	31	42	1	29	62,29767	56,00893	1
17	12110067	31	108	1	72,09022	21,18962	10,04988	3
18	18111050	31	98	2	63,07139	20,24846	0	3
19	18113684	32	98	1	63,52952	21,2132	1,414214	3
20	12117817	32	42	1	30	62,62587	56,01785	1

Menghitung kembali pusat cluster dengan keanggotaan cluster yang sudah terbentuk. titik pusat cluster yang baru adalah rata-rata dari semua anggota dalam setiap cluster tertentu. Berikut ini perhitungan untuk menentukan titik pusat cluster yang baru: Untuk titik cluster pertama dari contoh data mahasiswa sebagai berikut:

Total anggota cluster pertama : 9 anggota

Total anggota cluster kedua : 8 anggota

Total anggota cluster ketiga : 3 anggota

Rata-rata cluster pertama  $C1_{(1)} = (2+2+7+11+11+12+31+31+32)/9 = 15,44$

$$C1_{(2)}=(42+42+42+42+42+42+42+42+42)/9=42$$

$$C1_{(3)}=(1+1+1+3+1+1+1+1+1)/9=1,22$$

Jadi, titik pusat baru untuk cluster pertama adalah (15.44, 42, 1.22)

$$\text{Rata-rata cluster kedua } C2_{(1)}=(1+1+1+7+7+11+11+12)/8= 6,38$$

$$C2_{(2)}= (98+108+101+98+98+100+101+100)/8=100,5$$

$$C2_{(3)}=(2+2+2+1+2+1+1+1+1+1)/8=1,75$$

Jadi, titik pusat baru untuk cluster pertama adalah (6.38, 100.5, 1.75)

$$\text{Rata-rata cluster ketiga } C3_{(1)}=(31+31+32)/3=31,33 \quad C3_{(2)}= (108+108+98)/3=104,67$$

$$C3_{(3)}=(1+2+1)/3=1,3$$

Jadi, titik pusat baru untuk cluster pertama adalah (31.33, 104.67, 1.3)

**Tabel 4.5. Titik Pusat Cluster Baru**

Cluster	kd_cabang	Kd Outline	NILAI TA
1	15,44	42	1,22
2	6,38	100,5	1,75
3	31,33	104,67	1,3

Hitung kembali setiap anggota memakai titik pusat cluster yang baru. Jika titik pusat cluster tidak berubah lagi maka proses clustering selesai. Atau, kembali ke langkah nomor 3 sampai pusat cluster tidak berubah lagi.

Berikut ini adalah contoh perhitungan data mahasiswa pertama dengan titik pusat cluster yang baru

Titik mahasiswa pertama:(1,98,2)

Untuk titik pusat baru cluster pertama : (15.44, 42,1.22).

$$\text{Perhitungan : } D_{(1,1)}=\sqrt{(1-15,44)^2+(98-42)^2+(2-1,22)^2}$$

Hasil : 57,83

Jadi, jarak dari data mahasiswa pertama dengan titik pusat cluster pertama adalah 57,83

Untuk perhitungan mahasiswa pertama di cluster kedua.

Titik mahasiswa pertama:(1,98,2)

Untuk titik baru pusat cluster kedua : (6.38, 100.5,1.75).

Perhitungan :  $D_{(1,2)} = \sqrt{(1-6,38)^2 + (98-100,5)^2 + (2-1,75)^2}$

Hasil : 5,93772

Jadi, jarak dari data mahasiswa pertama dengan titik pusat cluster kedua adalah 5,93

Untuk perhitungan mahasiswa pertama di cluster ketiga.

Titik mahasiswa pertama:(1,98,2)

Untuk titik pusat baru cluster ketiga : (31.33, 104.67, 1.3).

Perhitungan :  $D_{(1,3)} = \sqrt{(1-31,33)^2 + (98-104,67)^2 + (2-1,3)^2}$

Hasil : 31,06

Jadi, jarak dari data mahasiswa pertama dengan titik pusat cluster ketiga adalah 31,06

Berdasarkan hasil perhitungan ketiga titik cluster terlihat bahwa jarak data siswa pertama dan ketiga titik cluster merupakan jarak terdekat dengan titik pusat cluster kedua yaitu 5,93, sehingga diperoleh kesimpulan: Pertama data siswa memasuki cluster. Berikut ini contoh untuk perhitungan data- data mahasiswa yang ditunjukkan pada tabel 1.6.

**Tabel 4.6. Hasil Perhitungan Data Mahasiswa Dengan Titik Pusat Cluster Baru**

No	Nim	kd_cabang	Kd Outline	NILAI TA	jarak ke cluster			jarak terdekat dengan cluster
					C1	C2	C3	
1	12117758	2	42	1	13,4418	58,6685 3	69,1944 2	1
2	12117702	2	42	1	13,4418	58,6685 3	69,1944 2	1
3	12112271	7	42	3	8,62565 9	58,5166 4	67,2485 5	1
4	18112505	11	42	1	4,44544 7	58,6869 4	65,88 2 57	1
5	18111290	11	42	1	4,44544 7	58,6869 4	65,8857 2	1
6	12117839	12	42	1	3,44702 8	58,7741 2	65,5840 5	1
7	12112436	31	42	1	15,5615 6	63,4740 6	62,6715 9	1
8	12112430	31	42	1	15,5615 6	63,4740 6	62,6715 9	1
9	12117817	32	42	1	16,5614 6	63,8685 9	62,6743	1
10	12112979	1	98	2	57,8370 3	5,93775 2	31,0626 4	2
11	12112251	1	108	2	67,5656 9	9,23346 6	30,5202 9	2
12	18113933	1	101	2	60,7463 7	5,40896 5	30,5592 5	2
13	18112943	7	98	1	56,6328 7	2,68270 4	25,2295	2
14	18112718	7	98	2	56,6378 1	2,58783 7	25,2374 3	2
15	18113908	11	100	1	58,1701 1	4,70711 2	20,8616 3	2
16	18110781	11	101	1	59,1672 4	4,70711 2	20,6607 8	2
17	18112382	12	100	1	58,1023 4	5,69182 7	19,8883 8	2
18	12110067	31	108	1	67,8097 5	25,7479 5	3,35973 2	3
19	18111050	31	98	2	58,1267 8	24,7478 7	6,71474 5	3
20	18113684	32	98	1	58,3976 2	25,7526 1	6,71027 6	3

## **BAB V**

### **PENUTUP**

#### **5.1. Kesimpulan**

Kesimpulan yang didapat dari penelitian ini adalah sebagai berikut

1. Metode Algoritma K-Means digunakan untuk metode pengelompokan data dengan menggunakan titik pusat K, dimana untuk menentukan nilai K pada saat memproses data pertama kali umumnya dengan random.
2. Dalam menentukan titik pusat cluster K baru adalah dengan menghitung nilai rata-rata dalam setiap anggotanya.
3. Pengelompokan dalam penelitian memakai 3 cluster. Untuk cluster pertama berisi 9 data mahasiswa, cluster kedua berisi 8 data mahasiswa, dan untuk cluster ketiga berisi 3 data mahasiswa. Iterasi dalam penelitian ini hanya mencapai 3 iterasi.
4. Data yang digunakan dalam penelitian ini adalah data mahasiswa bimbingan periode 2014-1 dan untuk pengolahan datanya hanya menggunakan 20 data mahasiswa

#### **5.1. Saran**

Dalam penelitian ini penulis memberikan saran-saran sebagai berikut

1. Outline pendaftaran tugas akhir yang paling dominan Pemrograman - Berorientasi Bisnis - Web Programming. Perlu ditingkatkan lagi untuk outline pendaftaran tugas akhir yang lain dengan melakukan sosialisasi atau membuat definisi yang jelas pada setiap masing-masing outline

pendaftaran tugas akhir.

2. Menambah lagi metode yang digunakan, tidak hanya sebatas algoritma k-means. Untuk mengoptimasi nilai pusat K cluster yang umumnya random.
3. Atribut yang digunakan dalam penelitian seharusnya bisa ditambah dengan IPK, untuk penelitian selanjutnya.
4. Untuk pengolahan data selanjutnya data yang digunakan jangan hanya sebatas satu periode tapi seharusnya bisa lebih dari satu periode

## DAFTAR PUSTAKA

- Andayani, Sri., 2007, Pembentukan Cluster dalam Knowledge Discovery in database dengan Algoritma K- Means, Semnas Matematika dan Pendidikan Matematika.
- Candra Nugraha, Deka Dwinavinta., dkk. 2014. Klasterisasi Judul Buku dengan Menggunakan Metode K-Means. *Jurnal SNATI*, Juni 2014
- Kusrini dan Lutfi, E.T. 2009. *Alogaritma Data Mining*. Yogyakarta: Andi Offset.
- Kusumadewi dan Purnomo. 2010. *Aplikasi Logika Fuzzy untuk mendukung keputusan*. Yogyakarta: Graha Ilmu.
- Khotimah, T. (2014). Pengelompokan Surat Dalam Al Qur'an Menggunakan Algoritma K-Means. *Simetris: Jurnal Teknik Mesin, Elektro Dan Ilmu Komputer*, 5(1), 83–88. <https://doi.org/10.24176/simet.v5i1.141>
- Handoko, S. H., Sediono, E., & Suhartono, S. (2014). Sistem Informasi Geografis Berbasis Web untuk Pemetaan Sebaran Alumni Menggunakan Metode K-Means. *Jurnal Sistem Informasi Bisnis*, 1(2), 81–86. <https://doi.org/10.21456/vol1iss2pp80-85>
- Harahap, B. (2019). *Penerapan Algoritma K-Means Untuk Menentukan Bahan Bangunan Laris (Studi Kasus Pada UD. Toko Bangunan YD Indarung)*. 394–403.
- Nugraha, D. D. C., Naimah, Z., Fahmi, M., & Setiani, N. (2014). Klasterisasi Judul Buku dengan Menggunakan Metode K-Means. *Seminar Nasional Aplikasi Teknologi Informasi (SNATI) Yogyakarta*, 21(1), 1907–5022.
- Ong, Johan Oscar, 2013. Implementasi Algoritma K-Means Clustering Untuk Menentukan Strategi Marketing President University. *JITI* Vol.12, No.1.
- Putu, N., Merliana, E., & Santoso, A. J. (n.d.). *Analisa Penentuan Jumlah Cluster Terbaik pada Metode K-Means*. 978–979.
- Santosa, Budi. 2007. *Data Mining Teknik Pemanfaatan Data untuk Keperluan Bisnis*. Yogyakarta: Graha Ilmu.
- Santoso, S. 2010. *Statistik Multivariat*. Jakarta: Elex Media Komputindo.
- Zaki, Mohammed J., and Wagner Meira Jr. *Data mining and analysis: fundamental concepts and algorithms*. Cambridge University Press, 2014.
- Supranto, J. M.A. 2000. *Statistik: Teori dan Aplikasi Edisi Keenam*. Jakarta: Erlangga.