

PENERAPAN ALGORITMA K-MEANS UNTUK MENENTUKAN PENJURUSAN MAHASISWA PROGRAM STUDI TEKNIK INFORMATIKA PADA STMIK MUSIRAWAS

Ali Amran

Departement of Information System, Institute of Technology and Business

Correspondence Author: aliamran37705@gmail.com

| | |
|--|--|
| Info Artikel : | ABSTRACT (in English) |
| Sejarah Artikel : Menerima : 26 Nopember 2023 Revisi : Diterima : 30 Desember 2023 Online : 31 Desember 2023 | <i>K-Means is a data clustering algorithm over and over again, by determining cluster values (K) randomly, for a while these values become pisat of clusters or commonly referred to as centroid, mean or "means". High School of Management and Computer Science Musirawas has three study programs, namely Informatics Engineering (IT), Information Systems (SI), and Computer Systems (SK). Informatics engineering study programs have a concentration of expertise such as Web Engineering and Database Engineering in accordance with the informatics engineering study program curriculum at STMIK Musirawas , the process of determining the current majors can be directly determined based on the interest of students who have chosen the skill concentration course on the Study Plan Card page located on the website of the High School of Management and Computer Science. . The process of determining concentration majors like this is not effective because it is not necessarily that students who choose the concentration of Web Engineering have a good value of Web Engineering concentration support courses and vice versa students who choose Software Engineering concentration do not necessarily have a good Software Engineering concentration support course.</i> |
| Keyword : K-Means Algorithm, Concentration Department | |
| | INTISARI (in Indonesia) |
| Kata Kunci : Algoritma K- Means, Penjurusan Konsentrasi | <i>K-Means merupakan algoritma clustering data berulang-ulang, dengan menetapkan nilai-nilai cluster (K) secara random, untuk sementara nilai tersebut menjadi pisat dari cluster atau biasa disebut dengan centroid, mean atau "means". Sekolah Tinggi Manajemen dan Ilmu Komputer Musirawas memiliki tiga program studi yaitu Teknik Informatika(TI), Sistem Informasi(SI), dan Sistem Komputer(SK). Program studi teknik informatika mempunyai konsentrasi keahlian seperti Web Engineering dan Database Engineering yang sesuai dengan kurikulum program studi teknik informatika di STMIK Musirawas , proses penentuan Penjurusan saat ini secara langsung dapat ditentukan berdasarkan minat mahasiswa yang telah memilih mata kuliah konsentersasi keahlian pada halaman Kartu Rencana Studi yang terdapat di website Sekolah Tinggi Manajemen dan Ilmu Komputer Musirawas . . Proses penentuan Penjurusan kosentrasi seperti ini tidak efektif karena belum tentu mahasiswa yang memilih konsentrasi Web Engineering mempunyai nilai mata kuliah pendukung konsentrasi Web Engineering yang baik dan juga sebaliknya mahasiswa yang memilih konsentrasi Software Engineering belum tentu mempunyai nilai mata kuliah pendukung konsentrasi Software Engineering yang baik juga.</i> |

1. PENDAHULUAN

Berbagai institusi, universitas, perusahaan atau organisasi pasti memiliki data yang tersimpan diserver masing-masing dan jumlahnya berlimpah. Berlimpahnya data ini merupakan akumulasi data transaksi yang terekam bertahun-tahun. Dengan data tersebut institusi, universitas, perusahaan atau organisasi bisa digunakan untuk mendapatkan informasi baru. Salah satu cara efektif dapat digunakan untuk mendapatkan informasi baru dengan menggunakan klasifikasi data, teknik yang dilakukan untuk klasifikasi data secara umum yaitu memanfaatkan proses data mining dengan menggunakan Algoritma K-Means. K-Means merupakan algoritma clustering data berulang-ulang, dengan menetapkan nilai-nilai cluster (K) secara random, untuk sementara nilai tersebut menjadi pisat dari cluster atau biasa disebut dengan centroid, mean atau “means” (Vulandari, 2017).

Sekolah Tinggi Manajemen dan Ilmu Komputer Musirawas didirikan pada tanggal 15 juni 2004 berdasarkan surat keputusan menteri pendidikan dan kebudayaan Republik Indonesia No.81/D/0/2004, merupakan lembaga pendidikan yang didirikan oleh yayasan dwi tunggal, berdasarkan surat keputusan Kemendikbud Republik Indonesia saat itu, STMIK Musirawas memiliki tiga program studi yaitu Teknik Informatika (TI), Sistem Informasi (SI), dan Sistem Komputer (SK).

Program studi teknik informatika mempunyai konsentrasi keahlian seperti Web Engineering dan Database Engineering yang sesuai dengan kurikulum program studi teknik informatika di STMIK Musirawas , proses penentuan Penjurusan saat ini belum dilakukan dengan alat bantu khusus yang dapat melakukan pemilihan penjurusan konsentrasi pada prodi teknik informatika, Sistem penjurusan secara langsung dapat ditentukan berdasarkan minat mahasiswa yang telah memilih mata kuliah konsentersasi keahlian pada halaman Kartu Rencana Studi yang terdapat di website Sekolah Tinggi Manajemen dan Ilmu Komputer Musirawas . Hasil dari keseluruhan Kartu Rencana Studi yang telah dipilih setiap mahasiswa untuk menentukan penjurusan konsentersasi, akan dapat diketahui Penjurusan yang paling banyak diminati oleh mahasiswa. Proses penentuan Penjurusan kosentersasi seperti ini tidak efektif karena belum tentu mahasiswa yang memilih konsentrasi Web Engineering mempunyai nilai mata kuliah pendukung konsentrasi Web Engineering yang baik dan juga sebaliknya mahasiswa yang memilih konsentrasi Software Engineering belum tentu mempunyai nilai mata kuliah pendukung konsentrasi Software Engineering yang baik juga.

Berdasarkan permasalahan yang ada, maka penulis akan menerapkan algoritma k-means untuk dapat membantu mengelompokan pemilihan Penjurusan berdasarkan sampel pada nilai Kartu Hasil Studi mahasiswa semester empat dan lima pada angkatan 2015 dan minat setiap mahasiswa. Beberapa kelebihan dari k-means yang akan dapat memaksimalkan populasi penentuan Penjurusan yaitu dalam prosesnya cepat dan nilai cukup efisien untuk pengelompokan dalam jumlah data mahasiswa yang banyak.

2. TINJAUAN PUSTAKA

2.1 Data Mining

Data Mining merupakan proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidenfikasi informasi yang bermanfaat dan pengetahuan yang terakit dari berbagai database besar/Data Warehouse (Vulandari, 2017).

Data Mining mendefinisikan data mining adalah serangkaian proses untuk menggali nilai tambah berupa informasi yang selama ini tidak diketahui secara manual dari suatu basisdata. Informasi yang dihasilkan diperoleh dengan cara mengekstrasi dan mengenali pola yang menarik dari suatu data (Nugraha, 2014).

2.2 Algoritma K-Means

K-Means merupakan algoritma clustering data berulang-ulang, dengan menetapkan nilai-nilai cluster (K) secara random, untuk sementara nilai tersebut menjadi pisat dari cluster atau biasa disebut dengan *centroid*, *mean* atau “*means*” (Vulandari, 2017).

Berikut langkah-langkah dari penerapan algoritma *K-Means* Untuk dapat menentukan Penjurusan Mahasiswa prodi Teknik Informatika pada STMIK Musirawas :

- a. Menentukan banyak *k-cluster* yang ingin dibentuk.

Dalam menentukan jumlah *cluster* berasal dari banyak data jurusan pada Prodi Teknik Informatika yang ada akan dibuat menjadi 2 *cluster* yaitu *Cluster Software Engineering* dan *Cluster Web Engineering* yang akan diidentifikasi sebagai *Cluster A*, dan *Cluster B*, agar lebih mudah dalam proses perhitungan.

- b. Membangkitkan nilai random untuk pusat *cluster* awal (*centroid*) sebanyak *k-cluster*.

Dalam menghitung nilai *centroid* awal akan digunakan nilai *meandari* seluruh *Cluster*. Nilai mata kuliah pada masing-masing kolom konsentrasi diambil sampel 10 *dataset*.

- c. Menghitung *dataset* konsentrasi ke 1

Menghitung jarak setiap data *input* terhadap masing-masing *centroid* konsentrasi menggunakan rumus jarak (*Euclidian Distance*) hingga akan ditemukan jarak yang paling dekat dari setiap data dengan *centroid* konsentrasi. Berikut persamaan *Euclidian Distance*:

$$d(x_i, \mu_i) = \sqrt{(x_i - \mu_i)^2}$$

- d. Mengklasifikasikan setiap data konsentrasi berdasarkan kedekatannya dengan *centroid* konsentrasi diidentifikasi sebagai jarak terkecil.

- e. Melakukan update nilai *centroid* konsentrasi. Nilai *centroid* konsentrasi baru diperoleh dari rata-rata *cluster* konsentrasi dengan menggunakan rumus:

$$C_k = \frac{1}{n_k} \sum d_i$$

Keterangan :

n_k = jumlah data dalam setiap *cluster* konsentrasi

d_i = jumlah nilai jarak masuk dalam setiap *cluster* konsentrasi

2.3 Penjurusan

Penjurusan merupakan upaya untuk membantu siswa dalam memilih jenis program pengajaran atau program jurusan studi yang akan diikuti oleh siswa dalam pendidikan lanjut (Wikarta, 2013).

Definisi lain, Penjurusan adalah satu seri materi pendidikan yang sudah ditentukan secara sistematis sesuai dengan bidangnya (Nugraha, 2014).

2.4 Konsentrasi

Menurut Kamus Besar Bahasa Indonesia (2014), konsentrasi adalah pemusatan perhatian atau pikiran pada suatu hal, sedangkan Kelas Konsentrasi adalah suatu penjurusan yang menitik beratkan pada satu pembelajaran di bidang tertentu.

3. METODE PENELITIAN

3.1. Metode Pengumpulan Data

Metode penelitian dibutuhkan untuk mengetahui kinerja sistem sehingga sistem perangkat lunak menjadi lebih baik dan memiliki data yang akurat. Metode-metode penelitian tersebut adalah sebagai berikut :

- a. Observasi

Peneliti melakukan Observasi pengamatan langsung ke STMIK Musirawas pada jurusan Teknik Informatika untuk memperoleh informasi yang diperlukan Menentukan Penjurusan konsentrasi Mahasiswa. Hasil dari observasi ini peneliti mendapatkan data mahasiswa, data mata kuliah, dan data Kartu Rencana Studi (KRS) yang akan mengganti konsentrasi jurusan setelah melewati semester lima.

b. Wawancara

Peneliti melakukan wawancara, merupakan teknik pengumpulan data dengan cara mengadakan tanya jawab kepada Kaprodi Teknik Informatika STMIK Musi Rawas .

c. Literatur

Peneliti mencari dan mengumpulkan materi-materi yang berkaitan dengan penyusunan skripsi, seperti buku-buku, jurnal, dan skripsi yang telah selesai dikerjakan. Materi-materi yang tersedia dapat digunakan sebagai bahan pembelajaran dan referensi bagi penulisan penelitian ini.

3.2. Analisa Data

Analisis sistem pada penelitian ini berperan sebagai penguraian dari sistem informasi utuh kedalam bagian-bagian komponennya dengan maksud untuk mengidentifikasi dan mengevaluasi permasalahan-permasalahan, kesempatan-kesempatan, hambatan-hambatan yang terjadi dan kebutuhan yang diharapkan dapat terpenuhi.

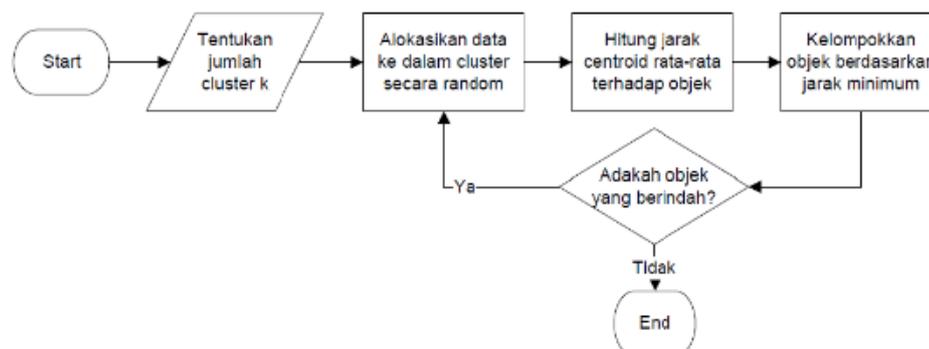
Dalam analisis ini, penulis mengumpulkan data-data yang diperlukan dalam penyusunan skripsi ini, seperti observasi ke STMIK Musirawas pada jurusan Teknik Informatika, melakukan wawancara dengan Kaprodi Teknik Informatika pada STMIK Musirawas , serta studi pustaka terhadap judul penulis.

Analisis Permasalahan

- Menentukan Variabel, disini penulis menggunakan nilai rata-rata Mata Kuliah Web Enggining dan Software Engineering dan IPK Semester V pada mahasiswa untuk dijadikan variabel.
- Menentukan Sampel, berdasarkan jumlah mahasiswa reguler jurusan Teknik Informatika pada angkatan 2015.
- Inisialisasi Data Nilai, data yang sudah diubah dikelompokkan berdasarkan mata kuliah yang dominan kemasing-masing konsentrasi dan dicari rata-rata nilainya pada setiap mata kuliah.

3.3. Metode Clustering K-Means

Menentukan metode untuk penelitian penulis menggunakan metode *clustering k-means*. Beberapa kelebihan dari *k-means* yaitu dalam prosesnya cepat dan dinilai cukup efisien untuk proses pengelompokan dalam jumlah data yang banyak. Selain itu, dalam iterasinya, algoritma ini akan berhenti dalam kondisi *optimum local* (Vulandari, 2017). Secara umum urutan proses *clustering* dengan algoritma *k-means* ditunjukkan pada gambar 1



Gambar 1. Flowchart Proses K-Means

Beberapa metode yang digunakan untuk menghitung selisih jarak adalah *Euclidean distance*. Adapun persamaanya adalah sebagai berikut:

$$d(x, y) = ||x - y||^2$$

$$\sqrt{\sum_{i=1}^n (xi - yi)^2}$$

Keterangan :

$d(x,y)$ = Jarak objek antara X_i dan Y_i

n = Dimensi data

X_i = Koordinat dari objek X_i pada dimensi i

Y_i = Koordinat dari objek Y_i pada dimensi i

3.4. Rancangan Basis Data

Tahap ini akan diterapkan metode *K-Means* untuk mengelompokkan data. Hasil pengelompokkan ini kemudian akan digunakan untuk pertimbangan Menentukan Penjurusan konsentrasi Mahasiswa Program Studi Teknik Informatika. Adapun algoritma *K-Means Clustering* pada Penjurusan konsentrasi Mahasiswa Program Studi Teknik Informatika adalah sebagai berikut:

1. Jumlah *cluster* yang dibentuk sebagai nilai k adalah dua ($k = 2$), yaitu Web Engineering dan Software Engineering.
2. Membangkitkan nilai *random* untuk pusat *cluster* awal (*centroid*) sebanyak 2 dari data sample. *Centroid* kriteria 1 adalah rata-rata nilai mata kuliah yang dominan ke penjurusan konsentrasi *Web Engineering* dan *centroid* kriteria 2 adalah rata-rata nilai mata kuliah yang dominan ke penjurusan konsentrasi *Software Engineering* dan *centroid* kriteria 3 adalah IPK Semester V.
3. Dalam menentukan penjurusan konsentrasi disini penulis menentukan paramater diantaranya :
 - a. Software Engineering rata- rata = 3,00 – 4,00
 - b. Web Engineering rata - rata = 2,00 – 2,90

4. HASIL DAN ANALISA

4.1 Data Pengujian

Tabel 1. Data Mata Kuliah yang dominan berdasarkan konsentrasi

| | |
|--------------------|---|
| WEB ENGINEERING | c. Kriptografi (X1) d. Praktek Kriptografi (X2) e. Praktek Pemrograman Lanjut II Web Dinamis (X3) f. Praktek Pemrograman Web Mobile (X4) |
| | g. IT Interpreneur (X5) h. Keamanan Komputer (X6) i. Praktek Design Perangkat Lunak (X7) j. Praktek Rekayasa Perangkat Lunak (X8) |

| | |
|-------------------------|--|
| SOFTWARE ENGINEERING | k. Praktek Uji Kualitas Perangkat Lunak (X9) l. Rekayasa Perangkat Lunak (X10) m. Uji Kualitas Perangkat Lunak (X11) |
|-------------------------|--|

Tabel 2. Data IPK mahasiswa dan Rata-rata Nilai Mata Kuliah yang dominan berdasarkan konsentrasi

| NIM | NAMA MAHASISWA | IPK | | | | | Web Engineering | | | | | | Software Engineering | | | | | |
|----------|----------------------------|------|----|----|----|----|-----------------|----|----|----|----|----|----------------------|-----|-----------|--|--|--|
| | | V | X1 | X2 | X3 | X4 | Rata Rata | X5 | X6 | X7 | X8 | X9 | X10 | X11 | Rata Rata | | | |
| 15010002 | FERI SAHRIA | 3,22 | 3 | 3 | 3 | 3 | 3 | 3 | 4 | 3 | 3 | 4 | 3 | 3 | 3,29 | | | |
| 15010003 | BUSTAN ARIFIN | 3,87 | 3 | 4 | 4 | 4 | 3,75 | 3 | 4 | 3 | 3 | 4 | 3 | 4 | 3,43 | | | |
| 15010004 | RINI RAHMAN AIFA RABY | 3,42 | 3 | 3 | 3 | 3 | 3 | 4 | 4 | 3 | 3 | 4 | 3 | 4 | 3,57 | | | |
| 15010005 | ARIANTO | 2,98 | 1 | 0 | 1 | 1 | 0,75 | 3 | 0 | 0 | 3 | 3 | 3 | 1 | 1,86 | | | |
| 15010006 | MARTIN SYAHPUTRA | 2,86 | 3 | 2 | 3 | 3 | 2,75 | 0 | 4 | 0 | 3 | 4 | 2 | 3 | 2,29 | | | |
| 15010007 | ANGGA SAPUTRA | 3,27 | 3 | 3 | 3 | 3 | 3 | 4 | 4 | 3 | 3 | 4 | 3 | 3 | 3,43 | | | |
| 15010009 | TOMI IRRAWAN | 3,07 | 3 | 0 | 3 | 3 | 2,25 | 4 | 3 | 3 | 3 | 4 | 2 | 4 | 3,29 | | | |
| 15010010 | WENDI SAPUTRA | 3,33 | 3 | 3 | 3 | 3 | 3 | 4 | 3 | 3 | 3 | 3 | 3 | 3 | 3,14 | | | |
| 15010014 | UVA OKTAVIA | 3,8 | 3 | 3 | 4 | 3 | 3,25 | 4 | 4 | 2 | 3 | 4 | 3 | 4 | 3,43 | | | |
| 15010016 | ZULHAKIM | 3 | 3 | 0 | 3 | 3 | 2,25 | 3 | 0 | 3 | 3 | 4 | 3 | 3 | 2,71 | | | |
| 15010017 | M. TAUFIK JAMIL AL WASHOLI | 3,41 | 3 | 3 | 3 | 3 | 3 | 4 | 4 | 3 | 3 | 4 | 3 | 4 | 3,57 | | | |
| 15010018 | MUHAMMAD DIMAS PRAYOGA | 2,9 | 3 | 0 | 3 | 2 | 2 | 3 | 2 | 3 | 2 | 1 | 2 | 2 | 2,14 | | | |
| 15010020 | BAYU SUBRATA | 3,35 | 3 | 3 | 3 | 3 | 3 | 4 | 3 | 3 | 3 | 4 | 3 | 4 | 3,43 | | | |
| 15010021 | YUNI LESTARI | 3,32 | 3 | 3 | 3 | 3 | 3 | 4 | 3 | 2 | 3 | 3 | 3 | 3 | 3,00 | | | |
| 15010022 | SUNGATI | 3,27 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 4 | 3 | 4 | 3,29 | | | |
| 15010023 | ELAM SARI REDO ERATAMA | 3,26 | 3 | 3 | 3 | 3 | 3 | 4 | 3 | 3 | 3 | 4 | 3 | 4 | 3,43 | | | |
| 15010024 | NOVYAN MARYANSYAH | 3,24 | 3 | 3 | 3 | 3 | 3 | 4 | 3 | 3 | 3 | 4 | 3 | 4 | 3,43 | | | |
| 15010025 | SEPTI MARYANI | 3,26 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 4 | 3 | 4 | 3,29 | | | |
| 15010028 | EKA YUMANTI | 3,34 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 4 | 3 | 4 | 3,29 | | | |
| 15010030 | LIA NATALIA | 3,35 | 3 | 2 | 3 | 3 | 2,75 | 4 | 3 | 2 | 3 | 3 | 3 | 3 | 3,00 | | | |
| 15010031 | EKA UTIDIA NINGSIH | 3,06 | 3 | 1 | 3 | 3 | 2,5 | 0 | 3 | 3 | 3 | 3 | 3 | 3 | 2,57 | | | |
| 15010032 | ANITA MIRRA | 3,2 | 3 | 3 | 3 | 3 | 3 | 4 | 2 | 3 | 3 | 3 | 3 | 3 | 3,00 | | | |
| 15010033 | DEPI KURNIAWAN | 3,27 | 3 | 3 | 3 | 3 | 3 | 4 | 3 | 3 | 3 | 3 | 3 | 3 | 3,14 | | | |
| 15010034 | SUPRIYADI | 3,48 | 3 | 3 | 3 | 3 | 3 | 4 | 3 | 3 | 3 | 4 | 3 | 3 | 3,29 | | | |
| 15010037 | DWIKO ARIF PUTRA | 3,4 | 3 | 3 | 3 | 3 | 3 | 4 | 3 | 3 | 3 | 3 | 3 | 3 | 3,14 | | | |
| 15010038 | MURMA BAITI | 3,21 | 4 | 2 | 3 | 3 | 3 | 3 | 4 | 3 | 3 | 3 | 3 | 3 | 3,14 | | | |
| 15010039 | DENI NOVERTA | 3,15 | 3 | 1 | 3 | 3 | 2,5 | 4 | 3 | 2 | 3 | 3 | 3 | 1 | 2,71 | | | |
| 15010040 | CANDERA DHINATA | 3,15 | 3 | 2 | 3 | 3 | 2,75 | 4 | 3 | 2 | 3 | 3 | 3 | 3 | 3,00 | | | |
| 15010041 | RAHMAT WUDODO | 3,58 | 4 | 4 | 3 | 4 | 3,75 | 4 | 4 | 4 | 3 | 4 | 3 | 3 | 3,57 | | | |
| 15010042 | TRY RIZKI | 2,33 | 3 | 0 | 3 | 2 | 2 | 0 | 1 | 0 | 2 | 3 | 2 | 3 | 1,57 | | | |
| 15010044 | EKO PRIOGA | 3,61 | 4 | 3 | 3 | 4 | 3,5 | 4 | 3 | 3 | 3 | 4 | 3 | 4 | 3,43 | | | |
| 15010046 | WAHYU SARI ARIANINGSIH | 3,33 | 3 | 3 | 3 | 3 | 3 | 4 | 3 | 2 | 3 | 3 | 3 | 3 | 3,00 | | | |
| 15010049 | ADI IRRAWAN | 3,27 | 3 | 2 | 3 | 3 | 2,75 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3,00 | | | |
| 15010050 | REFO CATUR PRAYOGA | 3,33 | 3 | 2 | 3 | 3 | 2,75 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3,00 | | | |
| 15010051 | DELA OLYMPIA PRATIWI | 3 | 3 | 2 | 3 | 3 | 2,75 | 4 | 3 | 3 | 3 | 3 | 3 | 3 | 3,14 | | | |
| 15010052 | ITA KURNIA | 3,14 | 3 | 2 | 3 | 3 | 2,75 | 4 | 3 | 3 | 3 | 3 | 3 | 3 | 3,14 | | | |
| 15010054 | ZERLAN NOVIANTO | 3,28 | 3 | 4 | 3 | 3 | 3,25 | 4 | 3 | 3 | 2 | 3 | 2 | 3 | 2,86 | | | |
| 15010055 | MURUL AL ANDY | 3,15 | 2 | 2 | 3 | 2 | 2,25 | 3 | 3 | 2 | 3 | 3 | 3 | 3 | 2,86 | | | |
| 15010058 | EKO MUZAMEL | 3,5 | 4 | 3 | 3 | 4 | 3,5 | 4 | 3 | 2 | 3 | 3 | 3 | 3 | 3,00 | | | |
| 15010059 | SANDI HARTANTO | 2,85 | 0 | 0 | 3 | 3 | 1,5 | 3 | 0 | 2 | 3 | 3 | 3 | 1 | 2,14 | | | |

4.2 Hasil Pengujian

Selanjutnya akan digunakan algoritma klasifikasi K-Means untuk mengelompokkan data yang ada. Data yang ada akan dikelompokkan menjadi 3 kelompok. Adapun langkah dari pengelompokkan data adalah sebagai berikut :

1. Tentukan pusat cluster secara acak, misalkan kita tentukan $c1 = (3.00, 2.50, 2.80)$
 $c2 = (3.50, 2.80, 3.00)$

Tabel 3. Nilai Cluster Awal

| | C1 | C2 |
|--|------|------|
| IPK | 3,00 | 3,50 |
| Rata-rata nilai dominan Web Engineering | 2,50 | 2,80 |
| Rata-rata nilai dominan Software Engineering | 2,80 | 3,00 |

2. Menghitung jarak setiap data yang ada terhadap setiap pusat cluster menggunakan rumus jarak *Euclidian Distance* hingga ditemukan jarak paling dekat dari setiap data dengan *centroid*. Berikut adalah persamaan *Euclidian Distance*:

$$d(x_i, \mu_j) = \sqrt{(x_{ia} - \mu_{ja})^2 + (x_{ib} - \mu_{jb})^2}$$

Keterangan :

x_i : data mahasiswa

μ_j : *centroid* pada *cluster* ke-j

x_{ia} : rata-rata nilai penjurusan *Web Engineering*

x_{ib} : rata-rata nilai penjurusan *Software Engineering*

μ_{ja} : nilai kriteria 1 dari *centroid cluster* ke-j adalah data rata-rata nilai penjurusan *Web Engineering*

μ_{jb} : nilai kriteria 2 dari *centroid cluster* ke-j adalah nilai random data rata-rata nilai *Software Engineering*

jarak data mahasiswa pertama dengan pusat cluster pertama adalah :

$$D1 = \sqrt{(3,22 - 3,00)^2 + (3,00 - 2,50)^2 + (3,29 - 2,80)^2} = 0,73$$

jarak data mahasiswa pertama dengan pusat cluster kedua adalah :

$$D2 = \sqrt{(3,22 - 3,50)^2 + (3,00 - 2,80)^2 + (3,29 - 3,00)^2} = 0,45$$

Hasil perhitungan selengkap nya dapat dilihat pada tabel dibawah ini

Tabel 4. Hasil Perhitungan Jarak Setiap Data

| NIM | NAMA | IPK | RATA RATA NILAI WEB ENGINEERING | RATA RATA NILAI SOFTWARE ENGINEERING | C1 | C2 |
|----------|--------------------------|------|---------------------------------------|---|------|------|
| 15010002 | FERI SAHRIA | 3,22 | 3,00 | 3,29 | 0,73 | 0,45 |
| 15010003 | BUSTAN ARIFIN | 3,87 | 3,75 | 3,43 | 1,65 | 1,11 |
| 15010004 | RIKI RAHMAN ALFA RABY | 3,42 | 3,00 | 3,57 | 1,01 | 0,61 |
| 15010005 | ARIANTO | 2,98 | 0,75 | 1,86 | 1,99 | 2,40 |
| 15010006 | MARTIN SYAHPUTRA | 2,86 | 2,75 | 2,29 | 0,59 | 0,96 |
| 15010007 | ANGGA SAPUTRA | 3,27 | 3,00 | 3,43 | 0,85 | 0,53 |
| 15010009 | TOMI IRAWAN | 3,07 | 2,25 | 3,29 | 0,55 | 0,75 |

| | | | | | | |
|----------|-------------------------------|------|------|------|------|------|
| 15010010 | WENDI SAPUTRA | 3,33 | 3,00 | 3,14 | 0,69 | 0,30 |
| 15010014 | UVA OKTAVIA | 3,8 | 3,25 | 3,43 | 1,26 | 0,69 |
| 15010016 | ZULHAKIM | 3 | 2,25 | 2,71 | 0,26 | 0,80 |
| 15010017 | M. TAUFIK JAMIL AL WASHOLI | 3,41 | 3,00 | 3,57 | 1,01 | 0,61 |
| 15010018 | MUHAMMAD DIMAS PRAYOGA | 2,9 | 2,00 | 2,14 | 0,83 | 1,32 |
| 15010020 | BAYU SUBRATA | 3,35 | 3,00 | 3,43 | 0,88 | 0,50 |
| 15010021 | YUNI LESTARI | 3,32 | 3,00 | 3,00 | 0,63 | 0,27 |
| 15010022 | SUMIATI | 3,27 | 3,00 | 3,29 | 0,75 | 0,42 |
| 15010023 | IMAM SARI REDO PRATAMA | 3,26 | 3,00 | 3,43 | 0,84 | 0,53 |
| 15010024 | NOVYAN MARVIANSYAH | 3,24 | 3,00 | 3,43 | 0,84 | 0,54 |
| 15010025 | SEFTI MARYANI | 3,26 | 3,00 | 3,29 | 0,74 | 0,42 |
| 15010028 | EKA YUMANTI | 3,34 | 3,00 | 3,29 | 0,78 | 0,38 |
| 15010030 | LIA NATALIA | 3,35 | 2,75 | 3,00 | 0,47 | 0,16 |
| 15010031 | IKA WIDIA NINGSIH | 3,06 | 2,50 | 2,57 | 0,24 | 0,68 |
| 15010032 | ANITA MIRA | 3,2 | 3,00 | 3,00 | 0,57 | 0,36 |
| 15010033 | DEPI KURNIAWAN | 3,27 | 3,00 | 3,14 | 0,66 | 0,34 |
| 15010034 | SUPRIYADI | 3,48 | 3,00 | 3,29 | 0,85 | 0,35 |
| 15010037 | DWIKO ARIF PUTRA | 3,4 | 3,00 | 3,14 | 0,73 | 0,27 |
| 15010038 | NURMA BAITI | 3,21 | 3,00 | 3,14 | 0,64 | 0,38 |
| 15010039 | DENI NOVERTA | 3,15 | 2,50 | 2,71 | 0,17 | 0,54 |
| 15010040 | CANDERA DHINATA | 3,15 | 2,75 | 3,00 | 0,35 | 0,35 |
| 15010041 | RAHMAT WIDODO | 3,58 | 3,75 | 3,57 | 1,58 | 1,11 |
| 15010042 | TRY RIZKI | 2,33 | 2,00 | 1,57 | 1,49 | 2,01 |
| 15010044 | EKO PRIOGA | 3,61 | 3,50 | 3,43 | 1,33 | 0,83 |
| 15010046 | WAHYU SARI APRIANINGSIH | 3,33 | 3,00 | 3,00 | 0,63 | 0,26 |
| 15010049 | ADI IRAWAN | 3,27 | 2,75 | 3,00 | 0,42 | 0,24 |
| 15010050 | REFO CATUR PRAYOGA | 3,33 | 2,75 | 3,00 | 0,46 | 0,18 |
| 15010051 | DELA OLYMPIA PRATIWI | 3 | 2,75 | 3,14 | 0,42 | 0,52 |
| 15010052 | ITA KURNIA | 3,14 | 2,75 | 3,14 | 0,45 | 0,39 |
| 15010054 | ZERIAN NOVIANTO | 3,28 | 3,25 | 2,86 | 0,80 | 0,52 |
| 15010055 | NURUL AL AMIN | 3,15 | 2,25 | 2,86 | 0,30 | 0,67 |
| 15010058 | EKO MUZAMIL | 3,5 | 3,50 | 3,00 | 1,14 | 0,70 |
| 15010059 | SANDI HARYANTO | 2,85 | 1,50 | 2,14 | 1,21 | 1,69 |

3. Suatu data akan menjadi anggota dari suatu cluster yang memiliki jarak terkecil dari pusat clusternya. Misalkan untuk data pertama, jarak terkecil diperoleh pada cluster kedua, sehingga data pertama akan menjadi anggota dari cluster kedua. Demikian juga untuk data kedua, jarak terkecil ada pada cluster kedua, maka data tersebut akan masuk pada cluster kedua. Posisi cluster selengkapya dapat dilihat pada Tabel 5

Tabel 5. Posisi cluster s

| NIM | NAMA | IPK | RATA RATA NILAI WEB ENGINEERING | RATA RATA NILAI SOFTWARE ENGINEERING | C1 | C2 |
|-----|--------------------------|------|---------------------------------------|---|----|----|
| 1 | FERI SAHRIA | 3,22 | 3,00 | 3,29 | | 1 |
| 2 | BUSTAN ARIFIN | 3,87 | 3,75 | 3,43 | | 1 |
| 3 | RIKI RAHMAN ALFA RABY | 3,42 | 3,00 | 3,57 | | 1 |
| 4 | ARIANTO | 2,98 | 0,75 | 1,86 | 1 | |
| 5 | MARTIN SYAHPUTRA | 2,86 | 2,75 | 2,29 | 1 | |
| 6 | ANGGA SAPUTRA | 3,27 | 3,00 | 3,43 | | 1 |

| | | | | | | |
|----|-------------------------------|------|------|------|---|---|
| 7 | TOMI IRAWAN | 3,07 | 2,25 | 3,29 | 1 | |
| 8 | WENDI SAPUTRA | 3,33 | 3,00 | 3,14 | | 1 |
| 9 | UVA OKTAVIA | 3,8 | 3,25 | 3,43 | | 1 |
| 10 | ZULHAKIM | 3 | 2,25 | 2,71 | 1 | |
| 11 | M. TAUFIK JAMIL AL WASHOLI | 3,41 | 3,00 | 3,57 | | 1 |
| 12 | MUHAMMAD DIMAS PRAYOGA | 2,9 | 2,00 | 2,14 | 1 | |
| 13 | BAYU SUBRATA | 3,35 | 3,00 | 3,43 | | 1 |
| 14 | YUNI LESTARI | 3,32 | 3,00 | 3,00 | | 1 |
| 15 | SUMIATI | 3,27 | 3,00 | 3,29 | | 1 |
| 16 | IMAM SARI REDO PRATAMA | 3,26 | 3,00 | 3,43 | | 1 |
| 17 | NOVYAN MARVIANSYAH | 3,24 | 3,00 | 3,43 | | 1 |
| 18 | SEFTI MARYANI | 3,26 | 3,00 | 3,29 | | 1 |
| 19 | EKA YUMANTI | 3,34 | 3,00 | 3,29 | | 1 |
| 20 | LIA NATALIA | 3,35 | 2,75 | 3,00 | | 1 |
| 21 | IKA WIDIA NINGSIH | 3,06 | 2,50 | 2,57 | 1 | |
| 22 | ANITA MIRA | 3,2 | 3,00 | 3,00 | | 1 |
| 23 | DEPI KURNIAWAN | 3,27 | 3,00 | 3,14 | | 1 |
| 24 | SUPRIYADI | 3,48 | 3,00 | 3,29 | | 1 |
| 25 | DWIKO ARIF PUTRA | 3,4 | 3,00 | 3,14 | | 1 |
| 26 | NURMA BAITI | 3,21 | 3,00 | 3,14 | | 1 |
| 27 | DENI NOVERTA | 3,15 | 2,50 | 2,71 | 1 | |
| 28 | CANDERA DHINATA | 3,15 | 2,75 | 3,00 | | |
| 29 | RAHMAT WIDODO | 3,58 | 3,75 | 3,57 | | 1 |
| 30 | TRY RIZKI | 2,33 | 2,00 | 1,57 | 1 | |
| 31 | EKO PRIOGA | 3,61 | 3,50 | 3,43 | | 1 |
| 32 | WAHYU SARI APRIANINGSIH | 3,33 | 3,00 | 3,00 | | 1 |
| 33 | ADI IRAWAN | 3,27 | 2,75 | 3,00 | | 1 |
| 34 | REFO CATUR PRAYOGA | 3,33 | 2,75 | 3,00 | | 1 |
| 35 | DELA OLYMPIA PRATIWI | 3 | 2,75 | 3,14 | 1 | |
| 36 | ITA KURNIA | 3,14 | 2,75 | 3,14 | | 1 |
| 37 | ZERIAN NOVIANTO | 3,28 | 3,25 | 2,86 | | 1 |
| 38 | NURUL AL AMIN | 3,15 | 2,25 | 2,86 | 1 | |
| 39 | EKO MUZAMIL | 3,5 | 3,50 | 3,00 | | 1 |
| 40 | SANDI HARYANTO | 2,85 | 1,50 | 2,14 | 1 | |

4. Hitung pusat cluster baru. Untuk cluster pertama, ada 11 data yaitu data ke-4, 5, 7, 10, 12, 21, 27, 30, 35, 38 dan data ke-40, sehingga:

$$C11 = (2,98 + 2,86 + 3,07 + 3,00 + 2,9 + 3,06 + 3,15 + 2,33 + 3,00 + 3,15 + 2,85) / 11 = 2,94$$

$$C12 = (0,75 + 2,75 + 2,25 + 2,25 + 2,00 + 2,50 + 2,50 + 2,00 + 2,75 + 2,25 + 1,50) / 11 = 2,14$$

$$C13 = (1,86 + 2,29 + 3,29 + 2,71 + 2,14 + 2,57 + 2,71 + 1,57 + 3,14 + 2,86 + 2,14) / 11 = 2,48$$

Untuk cluster kedua, ada 28 data yaitu data ke- 1, 2, 3, 6, 8, 9, 11, 13, 14, 15, 16, 17, 18, 19, 20, 22, 23, 24, 25, 26, 29, 31, 32, 33, 34, 36, 37 dan data ke -39, sehingga :

$$C21 = (3,22 + 3,87 + 3,42 + 3,27 + 3,33 + 3,8 + 3,41 + 3,35 + 3,32 + 3,27 + 3,26 + 3,24 + 3,26 + 3,34 + 3,34 + 3,35 + 3,2 + 3,27 + 3,48 + 3,4 + 3,21 + 3,58 + 3,61 + 3,33 + 3,27 + 3,33 + 3,14 + 3,28 + 3,5) / 28 = 3,37$$

$$C22 = (3,00 + 3,75 + 3,00 + 3,00 + 3,00 + 3,25 + 3,00 + 3,00 + 3,00 + 3,00 + 3,00 + 3,00 + 2,75 + 3,00 + 3,00 + 3,00 + 3,00 + 3,75 + 3,50 + 3,00 + 2,75 + 2,75 + 2,75 + 3,25 + 3,50) / 28 = 3,07$$

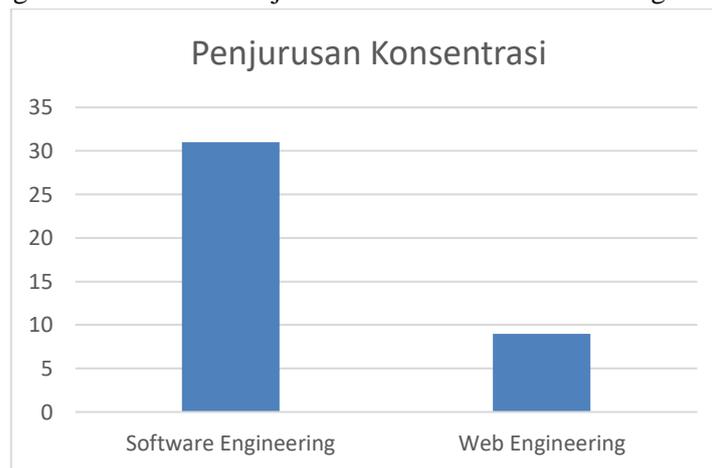
$$C23 = (3,29 + 3,43 + 3,57 + 3,43 + 3,14 + 3,43 + 3,57 + 3,43 + 3,00 + 3,29 + 3,43 + 3,43 + 3,29 + 3,29 + 3,00 + 3,00 + 3,14 + 3,29 + 3,14 + 3,14 + 3,57 + 3,43 + 3,00 + 3,00 + 3,00 + 3,14 + 2,86 + 3,00) / 28 = 3,24$$

5. Klasifikasi Hasil *Clustering*

Setelah proses *clustering*, tahap selanjutnya adalah proses klasifikasi. Di sini akan ditentukan *cluster* Penjurusan Mahasiswa Program Studi Teknik Informatika apakah *Web Engineering* atau *Software Engineering*, dan hasil akhir yang diperoleh adalah 2 cluster :

- a. Cluster pertama memiliki pusat (2,92 ; 2,06 ; 2,32) yang dapat diartikan bahwa mahasiswa tersebut dapat memilih konsentrasi web engineering.
- b. Cluster kedua memiliki pusat (3,34 ; 3,02 ; 3,23) yang dapat diartikan bahwa mahasiswa tersebut dapat memilih konsentrasi software engineering

Adapun grafik Penentuan Penjurusan Konsentrasi adalah sebagai berikut :



Gambar 2. Grafik Penjurusan Konsentrasi

5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian ini, dapat disimpulkan beberapa hal mengenai penerapan algoritma k-means untuk menentukan penjurusan konsentrasi mahasiswa program studi teknik informatika pada stmik musirawas .

1. Dengan menggunakan penerapan metode algoritma k-means dapat mengetahui minat dan kemampuan mahasiswa berdasarkan rata-rata nilai yang dominan terhadap masing-masing konsentrasi dan ipk nya dalam pemilihan konsentrasi.
2. Setelah dilakukannya penerapan metode algoritma k-means ini dapat membantu mahasiswa dan program studi teknik informatika dalam penentuan penjurusan konsentrasi setiap mahasiswa.
3. Dari data yang diolah dapat dilihat dari 40 mahasiswa 31 orang layak di Software Engineering dan 9 orang di Web Engineering.

Berdasarkan kesimpulan penelitian, maka penulis merekomendasikan berupa saran-saran sebagai berikut :

1. Pada penelitian ini dapat dikembangkan menjadi aplikasi maupun dibuat menggunakan bahasa pemrograman lainnya.
2. Untuk pengembangan penelitian ini dapat dikembangkan menggunakan perbandingan algoritma lainnya.

DAFTAR PUSTAKA

- R. T. Vulandari, *Data Mining Teori dan Aplikasi Rapidminer*. Yogyakarta: Gava Media, 2017.
- Y. P. Ario, “Implementasi Peraturan Presiden Nomor 26 Tahun 2009 Tentang Penerapan E-KTP Berbasis Nomor Induk Kependudukan Di Kecamatan Samarinda Utara Kota Samarinda.” 2014.
- Kbbi.web.id, “Definisi Penerapan,” 2017. [Online]. Available: <https://kbbi.web.id/terap-2>.
- W. R. Danang Aditya Nugraha, “Sistem Pendukung Keputusan Penjurusan Di Sma Menggunakan Metode Neural Network Backpropagation,” pp. 1–6, 2014.
- M. N. Desta Putu Wikarta, “Penerapan Konseling Trait And Factor Pada Siswa Yang Mengalami Kesulitan Memilih Program Penjurusan Bahasa,” pp. 1–10, 2013.
- S. Arikunto, *Prosedur Penelitian Suatu Pendekatan Praktik*. Jakarta: Rineka Cipta, 2010.
- Sugiyono, *Statistika untuk Penelitian*. Bandung: Alfabeta, 2013.