

BAB III

METODE PENELITIAN

3.1 Identifikasi Masalah

Pada saat ini dalam menentukan konsentrasi prodi hanya berdasarkan pada keinginan mahasiswa saja tanpa adanya suatu sistem yang bisa memberikan bahan pertimbangan atau acuan dalam menentukan konsentrasi pada prodi. Akibatnya banyak mahasiswa yang mengalami kesulitan saat mengambil konsentrasi yang akan mereka pilih. Dengan adanya salah satu teknik data mining *clustering k-means* yang akan memberikan pengelompokan data yang bisa dijadikan bahan pertimbangan mahasiswa untuk menentukan konsentrasi berdasarkan nilai akademik yang berhubungan dengan konsentrasi yang ada.

3.2 Penentuan Kebutuhan dan Data

Collecting Data dan Seleksi Data. Berikut ini beberapa tahapan metode penelitian yang dilakukan:

1. *Collecting data* Ditahapan ini dicari data yang tersedia, memperoleh data tambahan yang dibutuhkan, mengintegrasikan semua data kedalam dataset.
2. *Seleksi data* Pada tahap ini setelah data didapat, data akan diseleksi untuk membuat sebuah target data, fokus dalam bagian dari variabel atau sampel data yang akan dilakukan pada penelitian.
3. *Metode* yang diusulkan Pada tahap ini data akan dianalisis kemudian dikumpulkan untuk menjadi data yang saling berhubungan satu sama lainnya (Purwati et al., 2020)

Setelah menganalisa identifikasi masalah tahap dalam menentukan data yang akan dibutuhkan dalam proses *clustering* yaitu data-data nilai yang akan digunakan dalam proses penelitian ini diperoleh dari bagian *information communication technology* (ICT) dan bagian *biro administrasi akademik kemahasiswaan* (BAAK) yaitu data nilai mahasiswa dari semester 6 sampai 7.

3.2 Penentuan sampel masing-masing matakuliah

Nilai sampel dalam penelitian ini ditentukan dengan menggunakan persamaan Slovin sebagai berikut :

$$n = \frac{N}{1 + Ne^2}$$

$$n = \frac{2742}{1 + 2742(0.05)^2}$$

$$n = \frac{2742}{1 + (2742 \times 0.0025)}$$

$$n = \frac{2742}{1 + 6.855}$$

$$n = \frac{2742}{7.855} = 349 \text{ record}$$

Jadi sampel yang diambil adalah sebanyak 349 record. Kemudian dilakukan perhitungan data sampling dengan *Proportionate Stratified Random* pada atribut matakuliah (Riduwan, 2010). Metode pengambilan sampel ini dari anggota populasi yang diperoleh secara acak, maka dilakukan perhitungan sebagai berikut :

Jurusan TI

- IPK (2,00 – 2,75) = (35/ 2742) x 349 = 4 sampel
- IPK (2,76 – 3,50) = (334/2742) x 349 = 43 sampel
- IPK (3,51 – 4,00) = (67/2742) x 349 = 9 sampel

Jurusan SI

- IPK (2,00 – 2,75) = (48/ 2742) x 349 = 6 sampel
- IPK (2,76 – 3,50) = (411/2742) x 349 = 52 sampel
- IPK (3,51 – 4,00) = (46/2742) x 349 = 6 sampel

Jurusan SK

- IPK (2,00 – 2,75) = (5/2742) x 349 = 1 sampel
- IPK (2,76 – 3,50) = (70/2742) x 349 = 9 sampel

- IPK (3,51 – 4,00) = (12/2742) x 349 = 1 sampel

Jurusan MA

- IPK (2,00 – 2,75) = (21/2742) x 349 = 3 sampel
- IPK (2,76 – 3,50) = (709/2742) x 349 = 90 sampel
- IPK (3,51 – 4,00) = (308/2742) x 349 = 39 sampel

Jurusan AK

- IPK (2,00 – 2,75) = (16/2742) x 349 = 2 sampel
- IPK (2,76 – 3,50) = (454/742) x 349 = 58 sampel
- IPK (3,51 – 4,00) = (206/2742) x 349 = 26 sampel

3.3 Perancangan Aplikasi Data Mining

Dalam proses penelitian perlu adanya perancangan aplikasi data mining dengan metode clustering kmeans agar memudahkan penulis. Beberapa tahapan-tahapan yang diperlukan dalam prosesnya, yaitu:

Pengubahan Range Nilai dalam format Numerik (angka) Data nilai mata kuliah yang awalnya berupa huruf akan diolah untuk clustering dengan algoritma k-means harus diubah ke dalam bentuk angka agar dapat diproses.

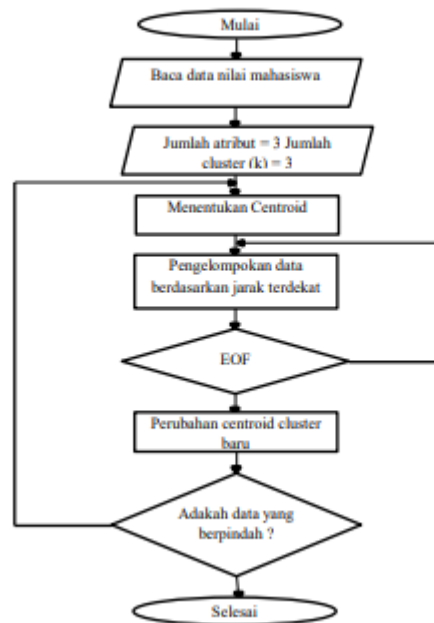
Tabel. Perubahan Parameter Nilai (Surat Edaran Rektor IIB Darmajaya)

Range	Nilai	Bobot
80 – 100	A	4
75 – 79,5	A-	3,75
70 – 74,5	B+	3,5
65 – 69,5	B	3
55 – 64,5	C	2
30 – 54,5	D	1
< 30	E	0

Komponen Penilaian :	
UTS	= 30%
UAS	= 30%
Tugas Mandiri	= 30%
Presensi Kehadiran	= 10%
	<u>100%</u>

3.4 Menentukan Metode

Dalam menentukan sebuah metode untuk penelitian penulis menggunakan metode *clustering k-means*. Beberapa kelebihan dari k-means yaitu dalam prosesnya lebih cepat dan dinilai cukup efisien dalam proses pengelompokan pada jumlah data yang banyak. Selain itu, dalam iterasinya, algoritma ini akan terhenti dalam sebuah kondisi optimum local (Dwirohayati et al., n.d.). Secara umum pada urutan proses clustering dengan algoritma k-means ditunjukkan pada gambar 3



Gambar 3 Flowchart Proses K-Means(Dwirohayati et al., n.d.)