

BAB I

PENDAHULUAN

1.1 Latar Belakang Penelitian

Knowledge Discovery in Database (KDD) didefinisikan sebagai ekstraksi informasi potensial, implisit dan tidak dikenal dari sekumpulan data. Proses *Knowledge Discovery* melibatkan hasil dari proses *Data Mining* (proses mengekstrak kecenderungan pola suatu data), kemudian mengubah hasilnya secara akurat menjadi informasi yang mudah dipahami (Fitri Larasati Sibuea dan Andy Saputra, 2017).

Data Mining adalah bagian dari proses penemuan pengetahuan dalam *Knowledge Discovery in Database* (KDD) (Singh dan Sharma, 2017). Menurut Indhu dan Porkodi (2018), istilah KDD atau disebut penemuan pengetahuan data karena tujuan utama *Data Mining* adalah untuk memanfaatkan data dalam basis data dengan mengolahnya sehingga menghasilkan informasi baru yang berguna.

Algoritma *K-means* merupakan metode analisis kelompok yang mengarah pada partisian N objek pengamatan ke dalam K kelompok, dimana setiap objek pengamatan sebuah kelompok data dengan *mean* (rata-rata) terdekat (Kumar dan Arulanandam, 2018). Algoritma ini merupakan salah satu *centroid* terkenal yang berbasis *algoritma clustering* (Ravali, et al. 2017). Didalam *Data Mining* proses pengelompokan merupakan salah satu metodologi yang paling penting (Madhumitha dan Kathiresan, 2018).

STMIK Indonesia Padang merupakan salah satu perguruan tinggi swasta dibawah naungan Lembaga Layanan DIKTI (LLDIKTI) Wilayah X, menghasilkan lulusan yang kompeten dalam bidang *system analyst* dan *database administrator*. Saat ini jumlah mahasiswa aktif di STMIK Indonesia Padang sebanyak 1.500 mahasiswa. Dengan rata-rata jumlah mahasiswa yang mengambil tugas akhir per semester ganjil dan genap sebanyak 200 mahasiswa.

Tabel 1.1 Perbandingan Bidang Pemilihan Topik Tugas Akhir

No.	Jumlah Mahasiswa	Periode	Bidang	
			<i>System Analyst</i>	<i>Database Administrator</i>
1.	125	Desember 2017	45 Mahasiswa	80 Mahasiswa
2.	108	Juni 2018	35 Mahasiswa	73 Mahasiswa
3.	126	September 2018	34 Mahasiswa	92 Mahasiswa

Sumber: Data lulusan STMIK Indonesia Padang, 2018

Dari tabel dapat dilihat bahwa terjadi ketimpangan jumlah bidang keahlian dalam setiap periode wisuda. Dalam hal ini peneliti melakukan wawancara kepada Wakil Ketua 1 Bidang Akademik. Menurut Stephanie (2019), hal ini disebabkan mahasiswa tingkat akhir belum mampu mengarahkan potensi yang mereka miliki, dalam penentuan topik dan tugas akhirnya.

Banyak mahasiswa yang sebenarnya mampu dibidang *system analyst* tetapi cenderung mengambil topik *database administrator* begitupun sebaliknya. Sehingga mahasiswa kesulitan dalam penyusunan tugas akhir karena kurang memahami topik tugas akhir yang dibuatnya. Pada penelitian ini penulis mencoba mencari solusi yang berguna bagi mahasiswa tingkat akhir dalam penyusunan tugas akhir yang mana dapat disesuaikan dengan bidang yang diminati mahasiswa. Dalam penggalian informasi ini penulis akan menggunakan teknik *Data Mining*.

Proses pengelompokkan mahasiswa tingkat akhir dilakukan dengan cara memanfaatkan data- data nilai matakuliah dari pemetaan matakuliah bidang *system analyst* dan *database administrator*. Dengan mengelompokkan nilai mahasiswa dengan cara menghitung jarak terdekat antara data dengan titik tengah suatu *cluster*. Dalam pengelompokkan tersebut akan dihasilkan dua klaster yaitu mahasiswa yang mengambil tugas akhir *system analyst* dan *database administrator*. Sehingga dengan menggunakan metode *K-means Clustering* ini, mahasiswa mempunyai arahan dalam pengambilan topik tugas akhirnya.

Dari penelitian sebelumnya sudah banyak mengangkat topik dengan menggunakan algoritma *Data Mining*, seperti implementasi *Data Mining* dalam pengelompokkan jurusan yang diminati siswa SMK Negeri 1 Lolowa'u menggunakan metode *Clustering* (Eferoni Ndruru dan Riswan Limbong, 2018). Implementasi metode *K-means* dalam pemetaan kelompok mahasiswa melalui data aktivitas kuliah (Rosmini, et al. 2018).

1.2 Perumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan di atas, maka dapat dirumuskan dua permasalahan sebagai berikut:

1. Bagaimana metode *K-means Clustering* dapat membantu mahasiswa dalam pemilihan konsentrasi yang akan memudahkan mahasiswa mencari judul tugas akhir berdasarkan keahliannya?
2. Bagaimana penerapan metode *K-Means Clustering* dalam pengklasteran tugas akhir mahasiswa STMIK Indonesia Padang menggunakan RapidMiner Studio?

1.3 Batasan Masalah

Agar penelitian ini lebih terarah maka penelitian diberi batasan-batasan sebagai berikut:

1. Penelitian ini hanya menganalisa penerapan algoritma *K-means Clustering* dalam menentukan keahlian mahasiswa terhadap tugas akhir di STMIK Indonesia Padang.
2. Teknik yang digunakan adalah klusterisasi.
3. Dalam menguji kebenaran hasil dari klusterisasi, penulis menggunakan sebuah aplikasi RapidMiner Studio.

1.4 Tujuan Penelitian

Tujuan penelitian yang ingin diperoleh penulis yaitu:

1. Menggali atau mendapatkan pengetahuan baru dari data nilai mahasiswa tingkat akhir di STMIK Indonesia Padang untuk menentukan topik tugas akhir.
2. Mengimplementasikan proses algoritma *K-means Clustering* terhadap nilai mahasiswa tingkat akhir.
3. Menghasilkan informasi yang dapat digunakan oleh pihak Kaprodi dalam membantu mahasiswa tingkat akhir untuk menentukan topik tugas akhir.
4. Membantu pengelompokkan mahasiswa tingkat akhir di STMIK Indonesia Padang untuk menentukan topik tugas akhir sesuai keahliannya.

1.5 Manfaat Penelitian

Manfaat yang bisa diperoleh dari penelitian ini adalah:

1. Memberikan beberapa pengetahuan untuk menjadi pedoman kepada mahasiswa dalam menentukan topik tugas akhir sesuai dengan keahliannya.
2. Diharapkan dengan penerapan metode ini mahasiswa terbantu dan dapat memudahkan dalam menentukan judul tugas akhir.

1.6 Sistematika Penulisan

Agar penelitian ini terarah dan sistematis, penulis membuat kerangka tulisan yang dituangkan dalam sistematika penulisan. Sistematika yang digunakan dalam penulisan ini adalah sebagai berikut:

BAB I PENDAHULUAN

Pada bab ini menjelaskan tentang latar belakang masalah, perumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penelitian.

BAB II LANDASAN TEORI

Pada bab ini membahas tentang landasan teori yang meliputi konsep dan penerapan *Knowledge Discovery in Databases (KDD)*, *Data Mining*, *Clustering*, *Algoritma K-means Clustering*.

BAB III METODOLOGI PENELITIAN

Pada bab ini akan dijelaskan secara detail tentang kerangka kerja penelitian serta metodologi-metodologi yang digunakan oleh peneliti dalam penentuan topik tugas akhir mahasiswa berdasarkan keahliannya di STMIK Indonesia Padang.

BAB IV ANALISA DAN PERANCANGAN

Pada bagian ini akan dibahas tentang analisis sistem yang sedang berjalan dan perancangan sistem yang diusulkan.

BAB V IMPLEMENTASI DAN HASIL

Pada bab ini akan dilakukan pengujian dari hasil analisa dengan aplikasi yang telah dipilih dan dengan prinsip kerja dari algoritma *K-means Clustering*.

BAB VI KESIMPULAN DAN SARAN

Pada bab penutup ini akan dikemukakan beberapa kesimpulan dari hasil penelitian yang didapat. Selanjutnya, beberapa saran setelah melakukan penelitian dalam menentukan topik tugas akhir mahasiswa berdasarkan keahlian menggunakan algoritma *K-means Clustering*.