

# Implementasi Metode *K-Means* Dalam Menentukan Mahasiswa Potensial *Drop Out*

Cut Lika Mestika Sandy

<sup>1</sup>Fakultas Komputer dan Multimedia, Universitas Islam Kebangsaan Indonesia, Aceh

\*Corresponding Email: likaclms@gmail.com

## ABSTRAK

*Data mining* adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode atau algoritma dalam *data mining* sangat bervariasi. Dalam penelitian ini, *data mining* digunakan untuk menentukan mahasiswa berpotensi *drop out* dengan menggunakan algoritma *K-Means*. Sistem ini mengelompokkan Mahasiswa di Jurusan Teknik Informatika Universitas Islam Kebangsaan Indonesia, Angkatan 2019 hingga 2022. Data yang dikelompokkan adalah nilai IPK (Indeks Prestasi Kumulatif) dan jumlah SKS. Data yang diperlukan diperoleh dari Jurusan di Teknik Informatika Universitas Islam Kebangsaan Indonesia. Hasil dari sistem ini berupa kelompok Mahasiswa berpotensi *drop out*. Sistem ini diharapkan dapat membantu pihak Jurusan untuk mengetahui lebih dini Mahasiswa yang berpotensi *drop out* dan dapat mengambil tindakan untuk mengantisipasinya.

**Kata kunci :** *Data Mining, Algoritma K-Means, Drop Out*

## ABSTRACT

*Data mining is a process to find the pattern or interesting piece of information in selected data using the technique or certain methods. Techniques, the method or algorithm in data mining varies. In this research, the data mining is used to determine potential students drop out with using logarithm K-means. This system groups students at the Informatics of Engineering at University Islam Kebangsaan Indonesian, A generation 2019 until 2022. Data that are grouped into is the GPA (Grade Point Average) and the number of credits. That is needed by majors at the Informatics of Engineering at University Islam Kebangsaan Indonesian. Results from the system was a student groups critical market drop out. This system is expected to be able to help the way to know more students early critical market drop out and can take action to anticipate it.*

**Keywords:** *Data Mining, Algorithm K-means, Drop out*

## 1. PENDAHULUAN

Sebuah Perguruan Tinggi harus terus meningkatkan kualitasnya demi mencetak anak bangsa yang cerdas dan tidak mudah dibodohi bangsa asing. Perguruan Tinggi yang ada di seluruh nusantara baiknya berlomba-lomba untuk meningkatkan mutu pendidikannya, tujuannya adalah menciptakan lulusan yang berguna bagi bangsa dan negara. Untuk menjalankan itu maka setiap Universitas harus mengadakan langkah perbaikan agar tetap dapat menjaga mutu dan nama baik. Ada beberapa hal yang harus ditegakkan yang mana salah satunya yaitu kualitas dan kuantitas kelulusan peserta didik.

*Drop out* adalah proses pencabutan status kemahasiswaan atas diri Mahasiswa, disebabkan oleh hal-hal tertentu yang telah ditentukan oleh Universitas. *Drop out* bisa digolongkan dalam beberapa bentuk, salah satunya *drop out Akademik* yaitu sistem yang diberlakukan kepada Mahasiswa karena tidak memenuhi ketentuan Akademik. *Drop out Akademik* ini berlaku kepada Mahasiswa yang selama beberapa semester berturut-turut mendapatkan nilai atau Indeks Prestasi (IP) semester rendah atau kurang dari standar yang ditetapkan Akademik. Selain IP alasan lain pemberlakuan *drop out* kepada Mahasiswa, karena tidak bisa memenuhi target SKS (Satuan Kredit Semester) yang ditentukan Perguruan Tinggi dalam kurun waktu

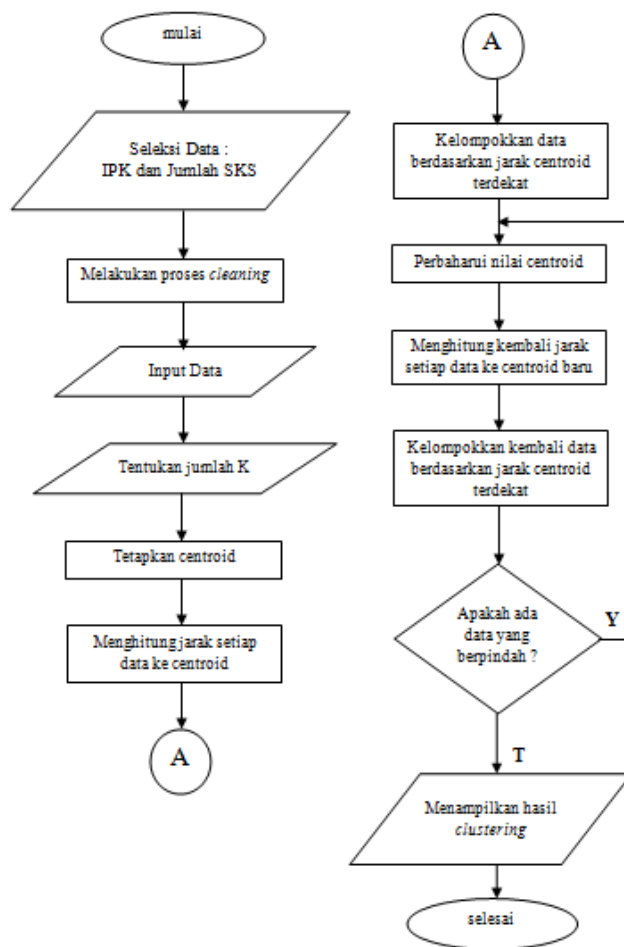
tertentu.

Ukuran prestasi Mahasiswa dapat dilihat dari Indeks Prestasi yang mencerminkan seluruh nilai yang diperoleh Mahasiswa sampai semester yang sedang berlangsung.

Algoritma K-Means memungkinkan untuk menemukan karakteristik-karakteristik yang sama dari Indeks Prestasi Mahasiswa dan dapat digunakan untuk memprediksi prestasi Mahasiswa dimasa depan. K-Means merupakan salah satu metode data clustering non hierarki yang berusaha mempartisi data ke dalam cluster sehingga data yang memiliki karakteristik sama dikelompokkan ke dalam satu cluster yang sama.

**2. METODE PENELITIAN**

Adapun langkah-langkah dalam penelitian ini ditampilkan pada gambar 1.



**Gambar 1.** Alur Penelitian

**3. HASIL DAN PEMBAHASAN**

**3.1 Perancangan Tabel**

Dalam sistem yang akan dibuat, digunakan beberapa tabel dalam basis data yang akan diaplikasikan pada Implementasi Metode K-Means *Clustering* dalam Menentukan Mahasiswa Potensial *Drop Out* yaitu sebagai berikut :

A. Tabel Admin

Nama Tabel : tb\_admin

**Tabel 1.** tb\_admin

Nama	Tipe	Ukuran	Keterangan
adminid	Varchar	20	Primary Key
adminpass	Varchar	20	
adminname	Varchar	50	

**B. Tabel Mahasiswa**

Nama Tabel : tb\_mhs

**Tabel 2.** tb\_mhs

Nama	Tipe	Ukuran	Keterangan
NIM	Varchar	10	Primary Key
NAMA	Varchar	50	
ANGKATAN	Varchar	5	
PROGRAM_STUDI	Varchar	50	
JUMLAH_SKS	Integer	5	
IPK_S_2	Double		
IPK_S_4	Double		
IPK_S_6	Double		
IPK_S_8	Double		

**C. Tabel Data Pusat Baru**

Nama Tabel : tb\_pusatbaru

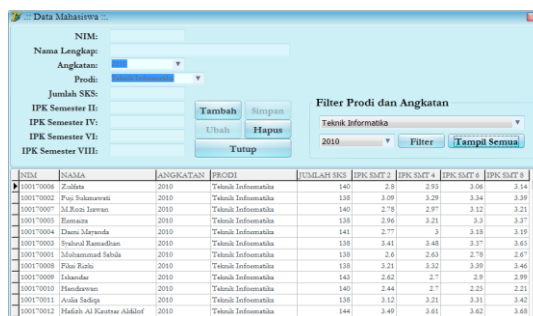
**Tabel 3.** tb\_pusatbaru

Nama	Tipe	Ukuran	Keterangan
ipk2	Double		
ipk4	Double		
ipk6	Double		
ipk8	Double		
sks	Integer	5	
clusterr	Char	5	
nim	Integer	10	

**3.2 Implementasi Sistem**

**A. Form Input Data Mahasiswa**

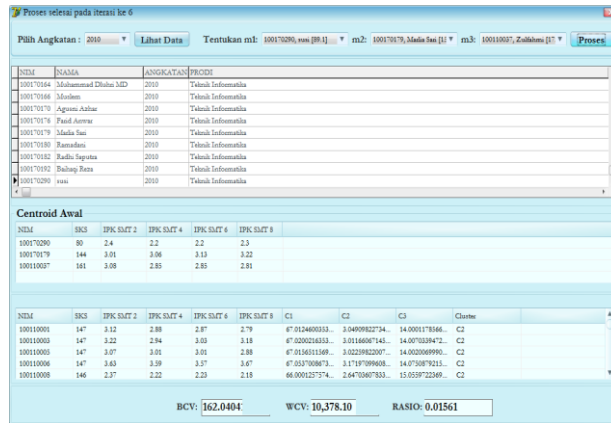
Form ini berfungsi untuk menambahkan, mengubah, dan menghapus data mahasiswa. Setelah selesai menginput data Mahasiswa, kembali ke halaman utama. Klik Proses kemudian K-means untuk memproses data. Berikut tampilan form input data Mahasiswa :



**Gambar 2.** Form Input Data Mahasiswa

**B. Form Proses**

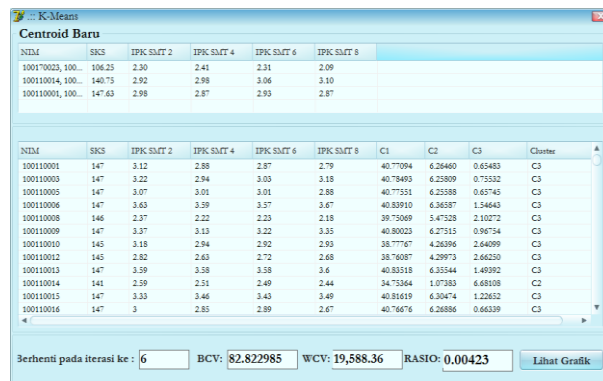
Form ini berfungsi untuk memproses data berdasarkan Angkatan yang dipilih, kemudian sistem akan kembali mengelompokkan data dengan *centroid* baru. Berikut tampilan form proses :



**Gambar 3.** Form Proses

**C. Form Tampilan Proses dengan Centroid Baru**

Form ini berfungsi untuk memproses data berdasarkan *centroid* baru dan menampilkan jumlah iterasi dalam proses *clustering*. Berikut tampilan form proses dengan *centroid* baru :



**Gambar 4.** Form Proses dengan *Centroid* Baru

**D. Form Tampilan Menu Data User**

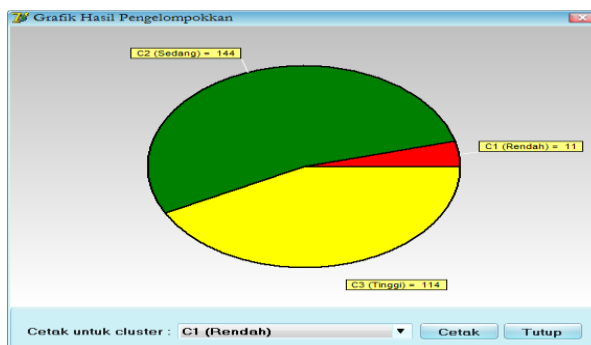
Form ini berfungsi untuk menginput admin name dan password. Berikut tampilan form menu data admin:



**Gambar 5.** Form Tampilan Menu Data Admin

E. Form Tampilan Grafik

Form ini berfungsi untuk menampilkan jumlah mahasiswa dalam masing-masing *cluster*. Berikut tampilan form grafik :



**Gambar 6.** Form Tampilan Grafik

**4. SIMPULAN**

Sistem ini mengelompokkan data Mahasiswa kedalam tiga cluster. Dimana C1 untuk *cluster* kategori rendah (berpotensi *drop out*), C2 untuk *cluster* kategori sedang, dan C3 untuk *cluster* kategori tinggi. Berdasarkan data yang di uji dan *centroid* awal yang dipilih, pada Angkatan 2019 terdapat 11 Mahasiswa yang dikelompokkan berpotensi *drop out*. Pada Angkatan 2020 terdapat 38 Mahasiswa yang dikelompokkan berpotensi *drop out*. Pada Angkatan 2021 terdapat 21 Mahasiswa yang dikelompokkan berpotensi *drop out*. Pada Angkatan 2012 terdapat 13 Mahasiswa yang dikelompokkan berpotensi *drop out*. Kelemahan algoritma K-Means yaitu, kelompok yang dihasilkan akan sangat bergantung pada tebakan K dan pemilihan nilai *centroid* awal. Jika nilai random untuk inisialiasi kurang baik, maka pengelompokkan yang dihasilkan pun akan menjadi kurang optimal. Selain itu K-Means juga terjebak pada kondisi local optimum dan tidak mampu menangani data yang menyimpang atau outlier.

**DAFTAR PUSTAKA**

- [1] Antares, J. (2020). Rancangan Sistem Informasi Kependudukan Berbasis Web Di Kantor Camat Medan Deli. Djtechno: Jurnal Teknologi Informasi, 1(2), 46-51.
- [2] Dinata, R. K., Novriando, H., Hasdyna, N., & Retno, S. (2020). Reduksi atribut menggunakan information gain untuk optimasi cluster algoritma k-means. J. Edukasi dan Penelit. Inform, 6(1), 48-53.
- [3] Rahayu, S., Alfeno, S., & Wahyono, K. N. (2019). Rancang Bangun Aplikasi Sistem Pengolahan Data Pembuatan Akta Tanah Pada Kantor Desa Cangkudu Tangerang. Innovative Creative and Information Technology, 5(1), 1-11.
- [4] Sujono, S. (2018). Penerapan Aplikasi Sistem Informasi Kependudukan Berbasis Web Pada Kantor Kepala Desa Puput Kec. Simpangkatis. Simetris: Jurnal Teknik Mesin, Elektro dan Ilmu Komputer, 9(1), 707-716.
- [5] Dinata, R. K., Safwandi, S., Hasdyna, N., & Azizah, N. (2020). Analisis K-Means Clustering pada Data Sepeda Motor. INFORMAL: Informatics Journal, 5(1), 10-17.
- [6] Puspita, D., Muslim, B., & Aminah, S. (2019). Aplikasi Pengelolaan Data Penduduk Dengan Pemrograman PHP (Studi Kasus: Kecamatan Suka Merindu). JURNAL ILMIAH BETRIK: Besemah Teknologi Informasi dan Komputer, 10(01), 7-16.
- [7] Sari, D. P., Retnosari, D. A., & Astuti, N. (2018). Aplikasi Sistem Pengolahan Data Pegawai pada Kantor Kelurahan Sukamulya Kabupaten Tangerang. SEMNASTEKNOMEDIA ONLINE, 6(1), 2-10.
- [8] Dinata, R. K., Safwandi, S., Hasdyna, N., & Mahendra, R. (2020). Kombinasi Algoritma Brute Force dan Stemming pada Sistem Pencarian Mashdar. CESS (Journal of Computer Engineering, System and Science), 5(2), 273-278.
- [9] Nugroho, A. H., & Rohimi, T. (2020). Perancangan Aplikasi Sistem Pengolahan DataPenduduk Dikelurahan Desa Kaduronyok Kecamatan Cisata, Kabupaten Pandeglang Berbasis Web.

- [10] Retno, S., Hasdyna, N., Mutasar, M., & Dinata, R. K. (2020). Algoritma Honey Encryption dalam Sistem Pendataan Sertifikat Tanah dan Bangunan di Universitas Malikussaleh. *INFORMAL: Informatics Journal*, 5(3), 87-95.
- [11] Erlinda, E. (2018). Pengolahan Data Sensus Penduduk Menggunakan Bahasa Pemrograman Php Berbasis Web Pada Kecamatan Bukit Sundi Kabupaten Solok. *Jurnal Teknologi dan Open Source*, 1(1), 46-57.
- [12] Dinata, R. K., Akbar, H., & Hasdyna, N. (2020). Algoritma K-Nearest Neighbor dengan Euclidean Distance dan Manhattan Distance untuk Klasifikasi Transportasi Bus. *ILKOM Jurnal Ilmiah*, 12(2), 104-111.
- [13] Maesaroh, S., Erliyani, I., & Ningsih, Y. F. (2020). Aplikasi Pengolahan Data Kependudukan Industri 4.0 Berbasis Web. *Journal Cerita*, 6(1), 95-105.
- [14] Hasdyna, N., & Dinata, R. K. (2020). Analisis Matthew Correlation Coefficient pada K-Nearest Neighbor dalam Klasifikasi Ikan Hias. *INFORMAL: Informatics Journal*, 5(2), 57-64.
- [15] Anraeni, S., Hasanuddin, T., Belluano, P. L. L., & Fadhiel, M. (2020). Sistem Informasi Pelayanan Administrasi Kependudukan Desa Pucak, Kecamatan Tompobulu, Kabupaten Maros. *Jurnal Ilmiah Ilmu Komputer Fakultas Ilmu Komputer Universitas Al Asyariah Mandar*, 6(2), 50-54.
- [16] Dinata, R. K., Retno, S., & Hasdyna, N. (2021). Minimization of the number of iterations in K-medoids clustering with purity algorithm. *Revue d'Intelligence Artificielle*, 35(3), 193-199.
- [17] Lestari, M. A., Tabrani, M., & Ayumida, S. (2018). Sistem Informasi Pengolahan Data Administrasi Kependudukan Pada Kantor Desa Pucung Karawang. *Jurnal Interkom: Jurnal Publikasi Ilmiah Bidang Teknologi Informasi dan Komunikasi*, 13(3), 14-21.
- [18] Dinata, R. K., Fajriana, F., Zulfa, Z., & Hasdyna, N. (2020). Klasifikasi Sekolah Menengah Pertama/Sederajat Wilayah Bireuen Menggunakan Algoritma K-Nearest Neighbors Berbasis Web. *CESS (Journal of Computer Engineering, System and Science)*, 5(1), 33-37.
- [19] Dinata, R. K., Hasdyna, N., Retno, S., & Nurfaumi, M. (2021). K-means algorithm for clustering system of plant seeds specialization areas in east Aceh. *ILKOM Jurnal Ilmiah*, 13(3), 235-243.
- [20] Dinata, R. K., Hasdyna, N., & Alif, M. (2021). Applied of Information Gain Algorithm for Culinary Recommendation System in Lhokseumawe. *Journal Of Informatics And Telecommunication Engineering*, 5(1), 45-52.
- [21] Nugroho, A., & Mulyana, A. (2018). Sistem Informasi Data Kependudukan Berbasis Web Di Kantor Kecamatan Cikarang Timur. *Jurnal SIGMA*, 9(2), 19-24.