

# DATA MINING EFEKTIVITAS PEMBELAJARAN DARING MENGUNAKAN ALGORITMA K-MEANS

Miftahul Huda  
Nasir

Universitas Islam Negeri Sultan Aji Muhammad Idris Samarinda  
Universitas Islam Negeri Sultan Aji Muhammad Idris Samarinda

miftahulh2@gmail.com  
muhammadnasir@iain-samarinda.ac.id

## Abstrak

Pandemi Covid-19 yang muncul di dunia pada tahun 2019 dan mulai masuk di Indonesia pada awal tahun 2020 telah mengubah kondisi banyak sektor salah satunya pendidikan. Kebijakan pemerintah untuk mengatasi covid yang juga rentan untuk siswa-siswi di sekolah dengan melakukan pembelajaran daring atau belajar dari rumah. Hal ini menimbulkan banyak pro dan kontra dari banyak pihak, siswa, guru, bahkan orang tua. Masa-masa pembelajaran tatap muka juga dijadikan bahan perbandingan oleh masyarakat terutama di daerah sekitar Kalimantan Timur. Tingkat efektivitas belajar telah menjadi pertanyaan dikarenakan tidak banyak orang tua yang dapat mendampingi anaknya belajar selaknyaknya guru di sekolah, yang memang sudah terdidik dan terlatih menyampaikan materi agar dapat dimengerti. Pengukuran efektivitas juga dapat menggunakan salah satu algoritma k-means dari data mining, dengan melakukan pengelompokan data dan dibagi menjadi cluster, sehingga dapat diketahui hasil analisa untuk pengambilan keputusan. Penelitian ini melakukan penyebaran kuesioner kepada 30 siswa-siswi secara acak di sekitar Kalimantan Timur. Kemudian dilakukan pengolahan data menggunakan k-means dan diperoleh hasil dari iterasi terakhir kategori kurang puas sebesar 4,24, kategori puas 1,41, dan angka 0,00 pada kategori netral.

**Kata kunci:** data mining, k-means, pembelajaran daring, covid-19

## Abstract

*The Covid-19 pandemic that emerged in the world in 2019 and began to enter Indonesia in early 2020 has changed the conditions of many sectors, one of which is education. The government's policy to deal with covid, which is also vulnerable for students at school, is by doing online learning or studying from home. This raises many pros and cons from many parties, students, teachers, and even parents. Face-to-face learning periods are also used as comparison material by the community, especially in the area around East Kalimantan. The level of learning effectiveness has become a question because there are not many parents who can accompany their children to study like teachers at school, who are already educated and trained to deliver material so that it can be understood. Measurement of effectiveness can also use one of the k-means algorithms from data mining, by grouping the data and dividing it into clusters, so that the results of the analysis for decision making can be known. This study distributed questionnaires to 30 students randomly around East Kalimantan. Then the data processing was carried out using k-means and the results obtained from the last iteration for the less satisfied category were 4.24, the satisfied category was 1.41, and the number was 0.00 in the neutral category.*

**Keywords:** data mining, k-means, online learning, covid-19

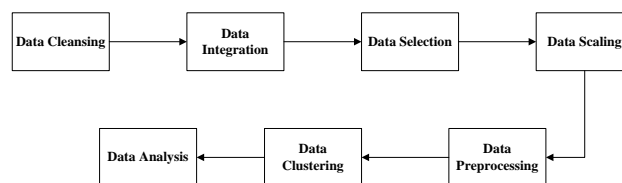
## PENDAHULUAN

Pada Oktober-Desember 2020 penambahan kasus harian covid di Kalimantan Timur sebanyak 66-418 kasus dan total kasus kematian sebanyak 743 kasus dan terjadi lonjakan sebesar 60,6% pada bulan Desember Wabah covid-19 secara fluktuatif dan belum stabil di seluruh provinsi di Indonesia sehingga masih diterapkannya PSBB (Pembatasan Sosial Berskala Besar) di beberapa provinsi besar. Wabah Covid-19 telah mempengaruhi banyak sektor baik kesehatan, ekonomi, lingkungan, tidak terkecuali pendidikan. Saat awal pandemi, pemerintah memutuskan kebijakan pembelajaran dari rumah atau daring. Sehingga siswa dapat mengikuti pembelajaran dari rumah dan didampingi oleh orang tua atau mengambil lembar tugas disekolah. Setelah berjalan setahun lebih sejak awal 2020 masuk ke Indonesia, kebijakan ini menimbulkan pro dan kontra dari segi guru dan siswa. Pembelajaran daring dinilai kurang efektif apabila dibandingkan dengan pembelajaran tatap muka. Dari sisi siswa, secara tidak langsung, siswa dituntut agar memiliki gadget dan akses internet untuk menunjang pembelajaran. Sementara latar belakang ekonomi siswa berbeda-beda. Dari sisi guru atau pengajar, dituntut untuk beradaptasi dengan tools yang mendukung untuk bisa menyampaikan materi dengan baik dan dapat dimengerti oleh siswa. Namun fakta dilapangan, guru mengalami kewalahan dan memberikan tugas. Karena banyak guru dan mata pelajaran yang ada, siswa pun kewalahan dalam mengerjakan tugas yang menumpuk dan dengan deadline yang singkat. Tidak jarang orang tua siswalah yang mengerjakan tugas-tugas anaknya atau siswa menggunakan mesin pencarian instan seperti google. Hal ini menyebabkan siswa malas berpikir dan pembelajaran daring menjadi kurang efektif. Maka dari itu, penelitian ini akan mengukur efektifitas pembelajaran daring selama pandemi dengan metode data mining k-means dan mengambil random sampling dari 30 siswa-siswi sekolah menengah atas di Kalimantan Timur. Hasil dari penelitian ini juga dapat dimanfaatkan oleh pemerintahan sekita Kalimantan Timur sebagai bahan acuan evaluasi pembelajaran daring selama tahun 2020/2021.

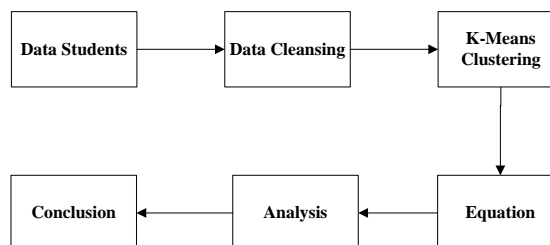
Pengukuran efektifitas pembelajaran kepada mahasiswa random sampling yang terbagi menjadi tiga cluster pada metode k-means mendapatkan hasil dimana cluster pertama yang menyatakan tidak setuju sangat tinggi, namun pada cluster ketiga pernyataan setuju lebih tinggi daripada cluster pertama. Sementara dengan menggunakan metode k-means pada cluster sangat tidak setuju terhadap pembelajaran online menunjukkan angka tertinggi yaitu 241 siswa. Metode k-means dengan toeri euclidean distance untuk memilih jarak antar data dan ke pusat cluster juga dapat digunakan untuk mengukur efektivitas pembelajaran online pada masa covid-19.

## METODE

Alur data mining yang digunakan dalam penelitian ini dapat dilihat pada gambar 1. Dan metodologi pengolahan data pada gambar 2.



Gambar 1 Alur Data Mining



Gambar 2 Metodologi Pengolahan Data

### A. Data Mining

Data yang berkapasitas besar atau big data tidak akan efisien apabila diolah dengan alat tradisional agar mendapat informasi yang tepat dan akurat. Menggali informasi dan pola dari big data untuk menemukan korelasi dan menemukan kesimpulan merupakan fungsi dari data mining. Cara kerja data mining meliputi pengumpulan data, ekstraksi, analisis dan statistik data. Dari proses tersebut akan ditemukan informasi yang dapat digunakan untuk mengambil keputusan. Data mining termasuk dalam machine learning dimana kecerdasan buatan yang membuat model dengan data empiris.

### B. K-Means

Metode dari data mining, salah satu yaitu k-means, dimana metode ini termasuk dalam metode clustering yang mengelompokkan data menjadi satu atau bisa dalam beberapa cluster. Algoritma ini juga lebih unggul untuk mengolah data dibandingkan algoritma lainnya. K-means juga dikategorikan sebagai non-hierarchical clustering, cara kerja dari metode ini dengan membagi data berdasarkan karakteristik yang sama dan yang tidak sama. Urutan pengelompokan data pada algoritma ini yaitu :

- Menentukan Jumlah cluster
- Inisialisasi awal dan pusat cluster random
- Menghitung jarak dengan rumus euclidean distance

$$d(x, y) = \sqrt{\sum_{i=1}^n (X_i - Y_i)^2}$$

$d(x, y)$  = ukuran ketidakmiripan

$X_i$  = variable data

$Y_i$  = variable pada titik pusat

- Menghitung pusat cluster baru dengan anggota baru dengan menghitung rata-rata obyek cluster atau dengan median.

Menghitung ulang jarak tiap obyek dengan pusat cluster sampai cluster stagnan, proses selesai.

### HASIL

Penelitian ini memperoleh data dari hasil kuesioner dengan 10 point pertanyaan, yang telah diisi oleh 30 siswa sekolah menengah atas secara random dari beberapa sekolah di Kalimantan Timur. Dari data tersebut akan dilakukan pengolahan data kemudian dilakukan clustering.

Tabel 1 Data Hasil Kuesioner

No	Nama Siswa	Pernyataan	
		Puas	Tidak Puas
1	Dinda Salsabila	5	5
2	Taufiq Sabilal Aziz	6	4
3	Imam Razaan Hibatullah	2	8
4	Novi Sulistyawati	5	5
5	Irma Ermayanti	4	6
6	Hana Husna Azalia	7	3
....			
....			
24	Gaelan Faren Azka	2	8
25	Ranesa	4	6
26	Meyra Delisha	5	5
27	Varisha Adeeva	6	4
28	Marenda Endah	5	5
29	Meidy Bagus	7	3
30	Surya Setyawan	2	8

Dari hasil kuesioner diatas maka dapat dijabarkan yaitu, jumlah data 30, dengan tiga cluster dan dua atribut (Puas dan Tidak Puas)

Iterasi ke-1

- Menentukan centroid awal dengan acak, data yang dipilih adalah (Taufiq Sabilal Aziz, Novi Sulistyawati dan Hana Husna Azalia) :

$$S^1=(6,4), S^2=(5,5), S^3=(7,3)$$

- Hitung centroid terdekat, yaitu : Dinda Salsabila (5,5). Lakukan Perhitungan dengan rumus :

$$d(S^1) = \sqrt{(5-6)^2 + (5-4)^2} = 1,41$$

$$d(S^2) = \sqrt{(5-5)^2 + (5-5)^2} = 0,00$$

$$d(S^3) = \sqrt{(5-7)^2 + (5-3)^2} = 0,00$$

Dan seterusnya sampai data ke-n

Apabila nilai centroid sudah tidak berubah dan sama nilainya dengan iterasi terakhir, maka proses iterasi berhenti. Dasi hasil tersebut dihasilkan tiga cluster yaitu, tinggi, sedang dan rendah.

Tabel 2 Data Hasil Iterasi 1

No	Pernyataan		C1	C2	C3	Jarak Terdekat
	Puas	Tidak Puas				
1	5	5	1,41	0	0	1,41
2	6	4	0	1,41	0	1,41
3	2	8	0	4,24	0	4,25
4	5	5	1,41	0	0	1,41
5	4	6	0	1,41	0	1,41
6	7	3	1,41	0	0	1,41
....	....	....	....	....	....	....
....	....	....	....	....	....	....
24	2	8	0	4,24	0	4,24
25	4	6	0	1,41	0	1,41
26	5	5	1,41	0	0	1,41
27	6	4	0	1,41	0	1,41
28	5	5	1,41	0	0	1,41
29	7	3	1,41	0	0	1,41
30	2	8	0	4,24	0	4,24

Tabel 3 Data Clustering

No	Nama Siswa	Cluster
1	Dinda Salsabila	C1
2	Taufiq Sabilal Aziz	C2
3	Imam Razaan Hibatullah	C2
4	Novi Sulistyawati	C1
5	Irma Ermayanti	C2
6	Hana Husna Azalia	C1
....		
....		
24	Gaelan Faren Azka	C2
25	Ranesa	C2
26	Meyra Delisha	C1
27	Varisha Adeeva	C2
28	Marenda Endah	C1
29	Meidy Bagus	C1
30	Surya Setyawan	C2

Implementasi dari penelitian ini juga dapat dilihat pada gambar 3 berikut :



Gambar 3 Halaman Utama

DATA OBJEK			DATA CLUSTER	
Data Excel: Browse... No file selected. Import Delete All			Cluster	Centroid Awal
Nama Objek	Data	Tambahkan	Input	Tambahkan
Hana Husna Azalia	3,7	Hapus Data	Cluster 1	6,4 Hapus Data
Irma Ermayanti	4,6	Hapus Data	Cluster 2	5,5 Hapus Data
Novi Sulistyawati	5,5	Hapus Data	Cluster 3	7,3 Hapus Data
Imam Razaan Hibatullah	2,8	Hapus Data		
Taufiq Sablal Aziz	3,7	Hapus Data		
Dinda Salsabila	5,5	Hapus Data		

Gambar 4 Data Clustering

Hasil Proses Clustering Algoritma K-Means Efektivitas Pembelajaran Daring					
ITERASI 1					
Objek	Data 1	Data 2	Cluster 1	Cluster 2	Cluster 3
Objek 1	5	5	Null	Null	OK
Objek 2	6	4	Null	Null	OK
Objek 3	2	8	Null	OK	Null
Objek 4	5	5	Null	Null	OK
Objek 5	4	6	Null	OK	Null
Objek 6	7	3	Null	Null	OK

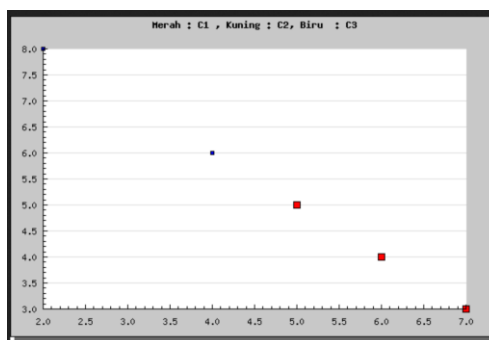
  

ITERASI 2					
Objek	Data 1	Data 2	Cluster 1	Cluster 2	Cluster 3
Objek 1	5	5	Null	Null	OK
Objek 2	6	4	Null	Null	OK
Objek 3	2	8	Null	OK	Null
Objek 4	5	5	Null	Null	OK
Objek 5	4	6	Null	OK	Null
Objek 6	7	3	Null	Null	OK

Cluster 1 -> 7	3	Cluster 2 -> 3	7	Cluster 3 -> 5,75	4,25
----------------	---	----------------	---	-------------------	------

Gambar 5 Hasil Clustering



Gambar 6 Diagram Hasil Clustering

**KESIMPULAN**

Dari hasil penelitian dapat disimpulkan bahwa, tingkat efektivitas pembelajaran dari selama masa pandemic covid-19 pada siswa-siswi SMA di Kalimantan Timur diperoleh angka tertinggi untuk ketidakpuasan sebesar 4,24, angka 1,41 untuk mewakili kategori puas, dan 0,00 untuk netral. Dalam hal ini dapat diketahui bahwa pembelajaran daring di Kalimantan Timur kurang efektif sehingga perlu adanya evaluasi lebih lanjut.

**DAFTAR PUSTAKA**

Mufidah, Firdausya Nur, dkk, 2021, Perbedaan Kasus Covid-19 Periode Maret-Mei 2020 dan Oktober-Desember 2020 di Kalimantan Timur, Jurnal Verdure, Vol.3, No.2, Tahun 2021, Hal 57-69, ISSN: 2714-8696.  
 Seftiya, Aulia dan Khemasili Kosala, 2021, Epidemiologi Karakteristik Pasien Covid-19 di Kalimantan Utara, Jurnal Sains dan Kesehatan, Vol 3, No.5, p-ISSN: 2303-0267, e-ISSN: 2407-6082 .

- Novidianto, RadityA dan Andrea Tri R.D, 2020, Analisis Klaster Kasus Aktif Covid-19 Menurut Provinsi di Indonesia Berdasarkan Data Deret Waktu, *Jurnal Aplikasi Statistika & Komputasi Statistik*, V.12.2.2020, ISSN 2086-4132.
- Anwar, Y., Maulana, M. F., dkk, 2020, Adaptasi Masyarakat Terhadap COVID-19 di Desa Genting Tanah, Kecamatan Kembang Janggut, Kalimantan Timur, *Jurnal Pendidikan Geografi* 7, Vol 7 No.2 Hal 11-23.
- Wegasari, Kurnia, dkk, 2021, Dampak Pandemi Covid-19 Terhadap Proses Pembelajaran Online di SDN Cabean 3 Demak, *Jurnal Penelitian*, Volume 15, Nomor 1, Februari 2021.
- Sianipar, Kristin D R, dkk 2020, Penerapan Algoritma K-Means dalam Menentukan Tingkat Kepuasan Pembelajaran Online Pada Masa Pandemi Covid-19, *Jurnal Teknologi Informasi*, Vol 4, No.1, Juni 2020, P-ISSN 2580-7927, E-ISSN 2615-2738.
- Lesmana, Sianipar, dkk, 2020, Penerapan K-Means dalam Efektivitas Pembelajaran E-Learning pada Masa Pandemi Covid-19, *Seminar Nasional Informatika 2020*, ISSN: 1979-2328.
- Nawawi, M Syukron dan Dudih Gustian, 2020, Algoritma K-Means Dalam Penentuan Tingkat Kepuasan Siswa Terhadap Pembelajaran Online di Masa Pandemi Covid-19, *Seminar Nasional Informatika 2020*, ISSN: 1979-2328.
- Arhami, Muhammad dan Muhammad Nasir, 2020, *Data Mining*, Yogyakarta : Penerbit Andi.
- Darmansah dan Ni Wayan Wardani, 2021, Analisis Persebaran Penularan Virus Corona di Provinsi Jawa Tengah Menggunakan Metode K-Means Clustering, *Jurnal Teknik Informatika dan Sistem Informasi*, VOL. 8, No.1, Maret 2021, Hal 105-117, ISSN 2407-4322.
- Wahyudi, Mochamad, dkk, 2020, *Data Mining : Penerapan Algoritma K-Means Clustering dan K-Medoids Clustering*, Yayasan Kita Menulis.
- Indraputra, .R.A dan R. Fitriana, 2020, K-Means Clustering Data COVID-19, *Jurnal Teknik Industri*, Vol.10 No.3, ISSN 2622-5131.
- Liliana, Dewi Y, dkk, 2021, Data Mining untuk Prediksi Status Pasien Covid-19 dengan Pengklasifikasi Naïve Bayes, *Jurnal Multinetics* Vol.7 No.1.
- Solichin, Achmad dan Khansa Khairunnisa, 2020, Klaterisasi Persebaran Virus Corona (Covid-19) di DKI Jakarta Menggunakan Metode K-Means, *Fountain of Informatics Journal*, Volume 5, No.2, ISSN: 2541-4313.
- Sindi, Sukma, dkk, 2020, Analisis Algoritma K-Medoids Clustering Dalam Pengelompokan Penyebaran Covid-19 di Indonesia, *Jurnal Teknologi Informasi*, Vol 4, No.1, P-ISSN 2580-7927.