

---

## CLUSTERING KELOMPOK BELAJAR SISWA BERDASARKAN HASIL UJIAN SEKOLAH MENGGUNAKAN ALGORITMA K-MEANS

Mohammad Syaefudulloh<sup>1)\*</sup>, Ahmad Faqih<sup>2)</sup>, Fadhil Muhammad Basysyar<sup>3)</sup>

<sup>1,2)</sup> Teknik Informatika, STMIK IKMI Cirebon, Kota Cirebon, Jawa Barat

<sup>3)</sup> Sistem Informasi, STMIK IKMI Cirebon, Kota Cirebon, Jawa Barat

email: [moh.syaef@gmail.com](mailto:moh.syaef@gmail.com)<sup>1)</sup>, [ahmadfaqih367@gmail.com](mailto:ahmadfaqih367@gmail.com)<sup>2)</sup>,  
[fadhil.m.basysyar@gmail.com](mailto:fadhil.m.basysyar@gmail.com)<sup>3)</sup>

### Abstrak

Pemerintah memiliki tanggung jawab menentukan kebijakan dan standar mutu pendidikan nasional serta memiliki peran untuk mengevaluasi penyelenggaraan pendidikan dalam rangka pengendalian mutu pendidikan nasional. Evaluasi terhadap prestasi siswa di sekolah sangat bervariasi dilihat dari hasil capaian ujian nasional. Hal ini terlihat dari hasil ujian dari berbagai sekolah yang diraih. Sekolah beserta guru harus memperhatikan secara serius terhadap pembentukan kelompok belajar siswa, agar terbentuk kelompok belajar yang efektif. Ini akan memungkinkan semua siswa untuk menunjukkan kemampuan superior mereka. Untuk mengetahui bagaimana klasterisasi kualitas pendidikan SMA Negeri 1 Astanajapura berdasarkan hasil nilai ujian sekolah. Penelitian yang dilakukan menggunakan machine learning dengan algoritma *K-Means* dengan dataset sampel atau data sekunder berasal dari SMA Negeri 1 Astanajapura Kelas XII yang akan dijadikan bahan penelitian ini. Hasil penelitian ini mengungkapkan klaster siswa, yaitu berprestasi tinggi, berprestasi, dan berprestasi rendah. Klasterisasi yang dihasilkan dari penelitian ini  $k = 4$  adalah terdapat 145 siswa yang di kelompokkan ke dalam klaster 0 dengan nilai DBi 0.763. Hasil evaluasi dari algoritma *K-Means* menghasilkan *cluster* dengan nilai sangat baik dan baik, Hasil penelitian ini dapat di implementasikan untuk pembentukan *cluster* kelompok belajar siswa kelas XII di SMA 1 Astanajapura berdasarkan hasil dari nilai ujian sekolah dengan menerapkan algoritma *K-Means* merupakan langkah yang tepat.

**Kata Kunci :** Data Mining, Kualitas, Ujian Nasional, Algoritma K-Means.

### Abstract

*The government has the responsibility of determining national education quality policies and standards and has a role to evaluate the implementation of education in the framework of national education quality control. Evaluation of student achievement in school varies greatly judging from the results of national exam achievements. This can be seen from the exam results from various schools achieved. Schools and teachers must pay serious attention to the formation of student learning groups, in order to form effective learning groups. This will allow all students to demonstrate their superior abilities. To find out how to cluster the quality of education of Sma Negeri 1 Astanajapura based on the results of school test scores. Research conducted using machine learning with K-Means algorithm with sample datasets or secondary data comes from SMA Negeri 1 Astanajapura Class XII*



*which will be used as this research material. The results of this study revealed a cluster of students, namely high achievers, achievers, and low achievers. The clustering resulting from this study  $k = 4$  is that there are 145 students who are grouped into cluster 0 with a DBi score of 0.763. The evaluation results of the K-Means algorithm produce clusters with excellent and good grades. The results of this study can be implemented for the formation of a group of class XII students at SMA 1 Astanajapura based on the results of school test scores by applying the K-Means algorithm is the right step.*

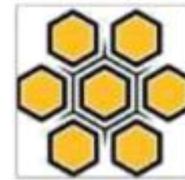
**Keywords:** *Clustering, Quality, National Exams, K-Means Algorithm.*

## PENDAHULUAN

Pendidikan merupakan penopang kemajuan peradaban (masyarakat). Semakin tinggi *index* pembangunan manusianya maka semakin maju pula peradabannya begitu pula sebaliknya peradaban akan mundur jika sumber daya manusianya rendah.[1]

Berkaitan dengan pendidikan pemerintah sudah menjamin bahwa semua orang mempunyai hak yang sama untuk mendapatkan Pendidikan yang layak. Jaminan oleh pemerintah tersebut direalisasikan dengan dibangunnya berbagai macam sarana pendidikan tingkat SD, SMP, SMA dan Sekolah Tinggi baik yang di selenggarakan oleh negara maupun oleh masyarakat dan semua jenjang dibiayai oleh negara kecuali perguruan tinggi. SMA Negeri 1 Astanajapura adalah salah satu instansi Pendidikan menengah atas yang berlokasi di wilayah Jawa Barat kabupaten Cirebon dengan alamat jalan Kh. Wahid Hasyim No.67. Kurikulum SMA Negeri 1 Astanajapura kabupaten Cirebon menggunakan kurtilas tahun 2013 untuk kelas XII. Rencana Pelaksanaan Pembelajaran (RPP) pada Sekolah Menengah Atas Negeri 1 Astanajapura berupa aktivitas kongkrit yg dilaksanakan secara sedikit demi sedikit dilakukan sang pengajar pada mendampingi peserta didik.

Hal terpenting yg wajib diperhatikan bahwa aktivitas pembelajaran difokuskan pada peserta didik. Pengajar berperan menjadi pendamping atau fasilitator. Berikut merupakan Rencana Pelaksanaan Pembelajaran (RPP) kurikulum 2013 kelas XII disusun dari buku dimana didalamnya memuat beberapa hal, mencakup bukti diri mata pelajaran, baku kompetensi, kompetensi dasar, materi pokok, aktivitas studi, indikasi, penilaian, alokasi waktu, dan asal belajar. Berikut ini merupakan Silabus kurikulum 2013 untuk jenjang Pendidikan kelas XII : PPKN, Matematika, Bahasa Inggris, Seni Musik, Seni Tari, Seni Teater, Penjas, Bahasa & sastra Indonesia, Biologi, Fisika, Kimia, Geografi, Sejarah, Ekonomi. Prestasi belajar siswa sangat dipengaruhi oleh peranan guru dalam proses pembelajaran. Ada kebutuhan mendesak untuk kinerja belajar siswa. Nilai siswa dapat digunakan sebagai dasar intervensi awal ketika siswa mungkin tidak dapat mencapai hasil belajarnya. Sambil mengubah strategi pembelajaran untuk mempromosikan keragaman siswa. Hasil belajar merupakan ukuran kinerja seorang peserta didik setelah menyelesaikan serangkaian kegiatan pembelajaran dalam bentuk skala penilaian. Untuk menentukan bobot atau metrik performa pembelajaran, Anda perlu



membuat metrik untuk memastikan efektivitas dan keandalannya. Nilai-nilai ini berfungsi sebagai deskripsi kinerja pembelajaran selama periode waktu tertentu.[2]

Hasil Ujian Nasional (UN) sebagai indikator dalam mengukur pencapaian lulusan siswa di sekolah. Ujian Nasional (UN) dapat digunakan oleh sekolah sebagai acuan untuk menetapkan kebijakan sekolah ke depan. Tujuan penelitian ini untuk mencari model algoritma untuk memprediksi nilai PBB dan akurasi.[3]

Kualitas hasil belajar di sekolah sangat bervariasi, sehingga pemahaman mereka berbeda. Hal ini terlihat dari berbagai hasil rapor siswa. Hal ini harus dijadikan sebagai catatan bagi sekolah, Pembentukan kelompok dimaksudkan agar seluruh siswa dapat menunjukkan kemampuan unggulannya. Hasil penelitian ini menghasilkan tiga cluster siswa yang berprestasi tinggi, berprestasi sedang, berprestasi rendah.[4]

## METODE PENELITIAN

### 1. Pengumpulan Data

Pada tahapan pengumpulan data yang berkaitan dengan penelitian dilakukan dengan melakukan observasi untuk mendapatkan data yang jelas dan akurat. Data yang di kumpulkan untuk dijadikan bahan penelitian ini adalah data nilai hasil ujian sekolah.

### 2. Pengolahan Data

Proses analisis data hasil nilai ujian sekolah pada penelitian ini melalui penerapan algoritma k-means clustering. Hasil proses pengolahan data ini diharapkan mendapatkan hasil cluster

siswa yang baik berdasarkan nilai ujian sekolah.

### 3. Menganalisa Hasil

Menganalisis hasil yang akan dilakukan untuk melihat yang telah dicapai dari tahapan sebelumnya dan pada tahapan ini akan dilakukan serangkaian pengujian apakah system yang dirancang sudah sesuai dengan model yang dirancang

### 4. Evaluasi Hasil

Evaluasi ini digunakan untuk menentukan keputusan akhir atau menarik hasil dari penelitian dengan garis besar untuk mengelompokan siswa, dan memberi saran untuk penelitian selanjutnya.

## HASIL DAN PEMBAHASAN

### A. Data Selection

Merupakan seleksi data dari variabel-variabel target untuk melakukan pemetaan kelompok belajar siswa berdasarkan hasil nilai ujian sekolah di SMA Negeri 1 Astanajapura. Sedangkan variabel-variabel adalah kelas, nama, nilai PAI, PKn, B\_IND, MTK\_WJB, SEJ\_IND, B\_INGG, SENI\_BDP, PJOK, PKWU, B\_SUN, MTK\_PGEO, BIO\_SEJ, FIS\_SOS, KIM\_EKO, LM\_SOS\_SIO.

### B. Preprocessing Data

Proses ini merupakan tahapan dalam penelitian yang digunakan dalam melakukan pembersihan data, menghilangkan duplikasi data dan menghilangkan data yang tidak konsisten.

### C. Transformasi Data

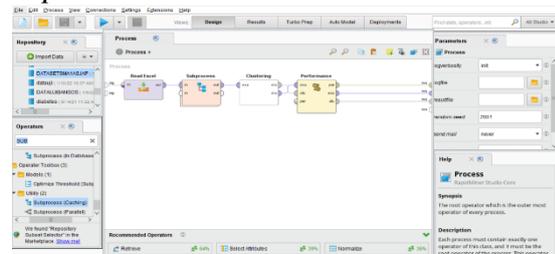
Agar data dalam penelitian ini dapat diproses sesuai dengan algoritma yang di gunakan yaitu K-Means maka data perlu dilakukan transformasi. Atribut yang akan



digunakan diberi label mengikuti kondisi data pada atribut tersebut. Data dalam penelitian ini terdiri dari 215 record hasil transformasi data.

#### D. Penerapan Algoritma Naïve Bayes

Model proses yang dibangun dalam penelitian ini menggunakan *software RapidMiner Studio*.



Gambar 1. Model Proses Algoritma *K-Means*

Tabel 1 Hasil Percobaan Dengan Nilai Parameter Max Run/Iterations = 10

k	K-means	
	Avg. within centroid distance	DBi
2	102.735	1.303
3	95.656	0.949
4	89.642	0.763
5	77.014	1.052
6	67.624	1.123
7	63.549	1.204
8	59.926	1.391
9	57.512	1.450
10	54.910	1.534

#### Cluster Model

```
Cluster 0: 145 items
Cluster 1: 1 items
Cluster 2: 1 items
Cluster 3: 68 items
Total number of items: 215
```

Gambar 2. Cluster Algoritma *K-Means*

Dilihat dari hasil gambar 3 bahwa nilai DBi = 0.763 maka hasil *cluster* model terbaik adalah pada k=4

#### E. Tahapan Evaluasi dan Interpretasi PerformanceVector

```
PerformanceVector:
Avg. within centroid distance: 89.642
Avg. within centroid distance_cluster_0: 87.952
Avg. within centroid distance_cluster_1: 0.000
Avg. within centroid distance_cluster_2: 0.000
Avg. within centroid distance_cluster_3: 95.883
Davies Bouldin: 0.763
```

Gambar 3. *Describe Performance Vector*

#### F. Evaluasi Performance

Dari nilai k = 2 s.d k = 10, didapat nilai DBi yang terkecil 0,763 dengan nilai k=4, Pengelompokan akan semakin baik jika indeksnya semakin kecl. Untuk mencari anggota group dari k dilihat dari hasil selisih jarak centroid rata-rata.Tabel 2.

Tabel 2. Perhitungan Selisih Cluster

K-means	Cluster 0	Cluster 3
Avg. within centroid distance	89.642	89.642
Avg. within centroid distance cluster	89.952	95.883
Selisih	310	6.241

#### KESIMPULAN

Berdasarkan hasil temuan dalam penelitian ini bahwa :

- Melalui penggunaan metode *clustering k-means*, dan dibantu menggunakan metode *elbow* untuk penentuan *cluster* menjadi optimal. Maka memudahkan untuk mengelompokan siswa berdasarkan hasil belajar siswa selama di sekolah, maka hasil dari pengelompokannya menghasilkan 4 *cluster* yang berprestasi akademik sangat baik, baik,



- cukup baik, kurang baik sebagai berikut:
- a. *Cluster* dengan prestasi akademik sangat baik dan baik terdapat pada *cluster* 0 (145 siswa) dengan nilai rata-rata 83-88
  - b. *Cluster* yang memiliki prestasi akademik cukup baik itu terdapat pada *cluster* 03 (68 siswa) dengan jumlah nilai rata-rata 79-83
2. Evaluasi hasil pengelompokan menggunakan metode k-means terhadap hasil belajar siswa terhadap cluster yang memiliki prestasi akademiknya kurang maka perlu untuk dilakukan pengarahan agar dapat meningkatkan hasil belajarnya.
- UCAPAN TERIMA KASIH**
- Terima kasih kepada semua pihak yang telah membantu dan terlibat dalam penyusunan jurnal ini.
- DAFTAR PUSTAKA**
- [1] Metode K-Means Clustering.” *Jointecs (Journal Of Information Technology And Computer Science)* 3(2): 229–34.
  - [2] Asriningtias, Salnan, And Elta Sonalitha. 2018. *Seminar Nasional Sistem Informasi*: 1179–87.
  - [3] C, Dengan Algoritma, Naïve Bayes, And D A N Svm. 2019. 13(1): 16–25.
  - [4] Juliartha, Berlian, Martin Putra, Dwi Ariani, And Finda Yuniarti. 2020. “Analisis Hasil Belajar Mahasiswa Dengan Clustering Menggunakan Metode K-Means.” *Poros Teknik* 12(2): 49–58.
  - [5] Kunci, Kata. 2021. “Tinjauan Pustaka Sistematis : Prediksi Prestasi Belajar Peserta Didik Dengan.” 1.
  - [6] Mahendra, Iwan. 2018. “Implementasi Kebijakan Pendataan Pemilih Dalam Pemilihan Umum Kepala Daerah Kota Malang 2013 ( Policy Implementation Of Voters Data Collection In Mayor Election Of Malang City.” 8: 28–36.
  - [7] Metode, Buku, Penelitian Kuantitatif, And Kualitatif Dan. 2017. “Sugiyono Metode Penelitian Kuantitatif Kualitatif Dan R D Download.”
  - [8] Natalia, Shefia, Br Sembiring, Hendryan Winata, And Sri Kusnasari. 2022. “Pengelompokan Prestasi Siswa Menggunakan Algoritma K-Means.” 1: 31–40.
  - [9] Penelitian, Metode. 2018. “Model Prediksi Nilai Ujian Nasional Siswa Smp Dengan Regresi Linear Berganda 1,2.” : 176–80.
  - [10] Rahmawati, And Toni Arifin. 2020. *Jurnal Responsif: Riset Sains Dan Informatika* 2(2): 184–90.
  - [11] Rijali, Ahmad. 2019. “Analisis Data Kualitatif.” *Alhadharah: Jurnal Ilmu Dakwah* 17(33): 81.
  - [12] Sapta, Andy, Sibuea Larasati Fitri Mustika, And Sembiring Ardiansyah Muhammad. 2018. *Journal Of Science And Social Research* 1(February): 73–79.
  - [13] K. Muhammad Susilo, Rezki Kurniati, “Rancang Bangun Website Toko Online Menggunakan Metode Waterfall,” *InfoTekJar (Jurnal Nas. Inform. dan Teknol. Jaringan)*, vol. 2, no. 2, pp. 98–105, 2018, doi: 10.30743/infotekjar.v2i2.171.