



Penerapan Algoritma K-Means Dalam Pengelompokan Rasio Angka Partisipasi Kasar di Tingkat Pendidikan Perguruan Tinggi Menurut Provinsi

Fernandi Simanjourang*, Riki Winanjaya, Fitri Rizki

STIKOM Tunas Bangsa, Pematangsiantar, Sumatera Utara, Indonesia

Email: ^{1,*}simanjourangfernandi@gmail.com, ²winanjaya85@gmail.com, ³30jihjan@gmail.com

Email Penulis Korespondensi: simanjourangfernandi@gmail.com

Abstrak—Rasio Angka Partisipasi Kasar Pendidikan merupakan pemegang peranan yang sangat penting dalam pembangunan bangsa dan negara, karena dari sanalah kecerdasan dan kemampuan bahkan watak bangsa dimasa akan datang banyak di tentukan oleh pendidikan pada saat ini. Penelitian ini membahas tentang pengelompokan banyaknya rasio angka partisipasi kasar menurut provinsi yang ada di Indonesia. Metode yang digunakan adalah Datamining dengan algoritma K-means Clustering. Dengan menggunakan metode ini data data yang diperoleh dapat dikelompokkan menjadi 3 cluster. Penelitian ini menggunakan data sekunder yaitu data diperoleh melalui media perantara yang terekam di situs badan pusat statistik dengan alamat url : <http://www.bps.go.id>. Hasil yang didapat dalam penelitian ini adalah pengelompokan rasio angka partisipasi kasar pendidikan yang di kelompokkan menjadi 3 cluster yaitu cluster tertinggi, sedang dan cluster terendah. Dalam penelitian ini diharapkan dapat memberikan masukan kepada pihak pemerintah terkait, agar lebih memperhatikan provinsi yang termasuk dalam cluster terendah untuk menanggulangi tingkat kualitas perguruan tinggi.

Kata Kunci: Angka Partisipasi Kasar Pendidikan; K-Means; Cluster

Abstract—The Ratio of Gross Participation Rate of Education is a very important role holder in the development of the nation and state, because that's where the intelligence and ability and even the character of the nation in the future is determined by education at this time. This study discusses the grouping of the number of ratios of gross participation rates by province in Indonesia. The method used is datamining with the K-means clustering algorithm. Using this method the data obtained can be grouped into 3 clusters. This study uses secondary data that is data obtained through intermediary media recorded on the website of the central statistics agency with the url address: <http://www.bps.go.id>. The results obtained in this study are the grouping of the ratio of gross participation rates of education grouped into 3 clusters, namely the highest cluster and the lowest cluster. In this study is expected to provide input to the relevant government, in order to pay more attention to the provinces that are included in the lowest cluster to overcome the level of quality of universities.

Keywords: Gross Participation Rate of Education; K-Means; Cluster

1. PENDAHULUAN

Rasio Angka Partisipasi Kasar Pendidikan berperan sebagai dasar dalam membentuk kualitas manusia yang mempunyai tingkat skill yang tinggi. Angka partisipasi kasar yaitu dalam menentukan penduduk usia sekolah dilakukan dengan perbandingan antara siswa pada jenjang pendidikan tertentu dengan penduduk usia sekolah tertentu tanpa melihat ketepatan usia dan dinyatakan dalam persentase [1]. Seberapa jauh keberhasilan pemerintah dalam usaha di sektor pendidikan dapat dilihat melalui salah satu indikator yang dapat digunakan adalah jumlah mahasiswa aktif dengan bermacam-macam usia yang menempuh pendidikan dalam tingkat tertentu terhadap penduduk sesuai dengan usia yang berkaitan dengan pendidikan tersebut. Menurut data dari situs resmi Pemerintah beralamatkan <https://bps.go.id>. Diperoleh persentase Rasio Angka Partisipasi Kasar (APK) Perempuan dan Laki-Laki di tingkat pendidikan perguruan tinggi menurut provinsi tahun 2018 sampai 2020. Penelitian ini dilakukan pengelompokan terhadap minimnya tingkat perguruan tinggi yang memiliki kualitas mulai dari tahun 2018 sampai 2020 dengan menggunakan metode *K-means*. Data-data yang diperoleh dikelompokkan menjadi 3 *Cluster* mulai dari terendah, sedang dan tertinggi. *K-means* merupakan salah satu algoritma dalam data mining yang bisa digunakan untuk melakukan pengelompokan/clustering suatu data. Dalam metode K-Means, pengelompokan data dilakukan dengan cara mengelompokkan atribut data ke dalam beberapa cluster berdasarkan kemiripan dari data tersebut. Kemudian, kemiripan suatu kelompok data diukur menggunakan suatu metode pengukuran jarak [2]. Algoritma K-Means sangat terkenal karena kemudahan dan kemampuannya untuk mengcluster data yang besar dan data outlier dengan sangat cepat. Dalam algoritma K-Means, setiap data harus termasuk ke cluster tertentu dan bisa dimungkinkan bagi setiap data yang termasuk cluster tertentu pada suatu tahapan proses, pada tahapan berikutnya berpindah kecluster lainnya. Algoritma K-means banyak diterapkan oleh beberapa peneliti dalam kasus pengelompokan seperti pada penelitian terdahulu yang dilakukan oleh [3] yang berjudul “Analisa Metode K-Means Pada Pengelompokan Perguruan Tinggi Menurut Provinsi Berdasarkan Fasilitas Yang Dimiliki Desa”. Membuat cluster dari kualitas pendidikan SMA di Indonesia. Untuk melakukan penilaian terhadap hasil pengelompokan perguruan tinggi menurut fasilitas yang dimiliki desa dapat menerapkan metode clustering K-means. Peneliti lain juga menerapkan metode K-Means dalam menentukan lokasi promosi produk minuman penurun kolesterol dengan hasil persentase keakuratan sebesar 95% [4]. Untuk menciptakan lulusan yang berkualitas serta mampu bersaing dan memiliki skill yang dapat berkompeten di tingkat nasional maupun internasional, maka sangat diperlukan perguruan tinggi yang memiliki kualitas dan Rasio Angka Partisipasi Kasar Pendidikan yang tinggi. Penulis melihat bahwa masih banyak perguruan tinggi yang kurang memiliki kualitas sehingga sangat sulit



menciptakan lulusan yang mampu bersaing di tingkat nasional maupun internasional. Sehingga penulis melakukan penelitian untuk mencapai tingkat perguruan tinggi yang memiliki kualitas tinggi. Berdasarkan permasalahan diatas, maka penulis menerapkan algoritma k-means dalam pengelompokan rasio angka partisipasi kasar di tingkat pendidikan perguruan tinggi menurut provinsi. Hasil dari penelitian diharapkan dapat di jadikan masukan untuk Pemerintah dalam hal membenahi atau meningkatkan efektifitas dalam memperbaiki kualitas perguruan tinggi yang ada di provinsi Indonesia.

2. METODOLOGI PENELITIAN

2.1 Data Mining

Data Mining atau pengembangan data dapat didefinisikan sebagai proses seleksi, eksplorasi dan pemodelan dari sejumlah besar data untuk menemukan pola atau kecenderungan yang biasanya tidak disadari keberadaannya [5]. Data mining seringkali digunakan secara bergantian untuk menjelaskan proses penggalian informasi tersembunyi dalam suatu basis data yang besar [6]. Data Mining merupakan proses ataupun kegiatan untuk mengumpulkan data yang berukuran besar kemudian mengekstraksi data tersebut menjadi informasi – informasi yang nantinya dapat digunakan [7].

2.2 Algoritma K-Means

K-Means merupakan salah satu metode pengelompokan data nonhierarki (sekatan) yang dapat mempartisi data kedalam bentuk dua kelompok ataupun lebih. [8]. Tujuan dari pengelompokan yaitu untuk meminimalkan dari fungsi objektif yang diset dalam proses pengelompokan. Algoritma *K-Means* dimulai dengan pembentukan partisi kluster diawal kemudian secara iteraktif partisi cluster ini diperbaiki hingga tidak terjadi perubahan yang signifikan pada partisi cluster[9]. Berikut adalah Alur penyelesaian dari metode *K-Means* [10]:

- Tentukan berapa jumlah kluster yang ingin ditetapkan pusat cluster k.
- Menggunakan jarak euclidean dan kemudian menghitung setiap data ke pusat cluster.

$$d_{ik} = \sqrt{\sum_j^m (C_{ij} - C_{kj})^2} \quad (1)$$

- Kategorikan data ke dalam cluster dengan jarak yang terpendek dengan menggunakan persamaan

$$\text{Min } \sum_k^k = d_{ik} = \sqrt{\sum_j^m (C_{ij} - C_{kj})^2} \quad (2)$$

- Menghitung pusat cluster dengan menggunakan persamaan

$$C_{kj} = \frac{\sum_{i=1}^p X_{ij}}{p} \quad (3)$$

Dengan : $x_{ij} \in$ Kluster ke – k p = banyak member kluster ke – k

Silahkan ulangi langkah dua sampai empat sehingga sudah tidak ada lagi data yang berpindah ke kluster yang lain.

3. HASIL DAN PEMBAHASAN

Pengumpulan data dalam penelitian ini dilakukan dengan memperoleh Data diperoleh dari situs resmi Pemerintah yang beralamatkan <https://www.bps.go.id> atau Badan Pusat Statistik yang terdiri dari 34 provinsi yang ada di Indonesia. Penelitian ini dilakukan berdasarkan data yang terdiri dari 34 Provinsi yang dijadikan sebagai alternatif dengan 3 kriteria yaitu mulai tahun 2018 sampai dengan 2020.

Tabel 1. Data Rasio Angka Partisipasi Kasar (APK) di Tingkat Perguruan Tinggi Menurut Provinsi

| No | Provinsi | 2018 | 2019 | 2020 |
|----|----------------------|--------|--------|--------|
| 1 | Aceh | 119,55 | 119,84 | 107,36 |
| 2 | Sumatera Utara | 124,15 | 122,32 | 115,41 |
| 3 | Sumatera Barat | 117,78 | 123,63 | 133,35 |
| 4 | Riau | 100,31 | 128,94 | 100,22 |
| 5 | Jambi | 118,32 | 118,29 | 109,22 |
| 6 | Sumatera Selatan | 124,26 | 109,65 | 122,14 |
| 7 | Bengkulu | 142,45 | 145,87 | 118,84 |
| 8 | Lampung | 128,53 | 119,21 | 125,37 |
| 9 | Kep. Bangka Belitung | 153,04 | 113,88 | 112,94 |
| 10 | Kep. Riau | 88,63 | 91,49 | 134,90 |
| 11 | Dki Jakarta | 94,42 | 108,29 | 95,71 |
| 12 | Jawa Barat | 106,81 | 109,49 | 102,11 |
| 13 | Jawa Tengah | 110,70 | 98,47 | 129,61 |
| 14 | Di Yogyakarta | 110,16 | 96,79 | 96,18 |
| 15 | Jawa Timur | 119,35 | 103,47 | 107,17 |

| No | Provinsi | 2018 | 2019 | 2020 |
|-----|---------------------|--------|--------|--------|
| 16 | Banten | 112,69 | 109,75 | 88,36 |
| 17 | Bali | 110,48 | 96,96 | 97,37 |
| 18 | Nusa Tenggara Barat | 104,80 | 92,79 | 99,60 |
| 19 | Nusa Tenggara Timur | 122,41 | 107,66 | 101,38 |
| ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... |
| 33 | Papua Barat | 126,51 | 92,76 | 102,45 |
| 34 | Papua | 97,78 | 92,82 | 101,38 |

1) Menentukan Jumlah *Cluster*

Jumlah *Cluster* yang digunakan pada data akta kelahiran sebanyak 3 *Cluster*. *Cluster* tersebut diantaranya Tinggi dari Target (*C1*), Sedang (*C2*) dan Rendah (*C3*) berdasarkan data Angka Partisipasi Kasar (APK) Rasio Angka Partisipasi Kasar Pendidikan yang diolah dari BPS tahun 2018, 2019 dan 2020 dengan jumlah 34 provinsi.

2) Tentukan *Centroid*

Menentukan pusat awal *Cluster* (*Centroid*) ditentukan secara manual atau *acak* yang diambil dari data yang ada dalam *range*. Nilai *Cluster* 1 diambil dari data dari data nilai tertinggi, nilai *Cluster* 2 diambil dari data sedang dan Nilai *Cluster* 3 diambil dari data yang paling rendah.

Tabel 2. Centroid

| Provinsi | 2018 | 2019 | 2020 |
|--|--------|--------|--------|
| Di ambil data ke-7 sebagai pusat cluster ke-1 | 142,45 | 145,87 | 118,84 |
| Di ambil data ke-13 sebagai pusat cluster ke-2 | 110,70 | 98,47 | 129,61 |
| Di ambil data ke-18 sebagai pusat cluster ke-3 | 104,80 | 92,79 | 99,60 |

Untuk menghitung jarak antara titik *Centroid* dengan titik tiap objek menggunakan *Euclidian Distance* menggunakan persamaan (1).

$$D_{x1,c1} = \sqrt{(119,55 - 142,45)^2 + (119,84 - 145,87)^2 + (107,36 - 118,84)^2} = 36,52$$

$$D_{x2,c1} = \sqrt{(124,15 - 142,45)^2 + (122,32 - 145,87)^2 + (115,41 - 118,84)^2} = 30,02$$

$$D_{x3,c1} = \sqrt{(117,78 - 142,45)^2 + (123,63 - 145,87)^2 + (133,35 - 118,84)^2} = 36,24$$

$$D_{x4,c1} = \sqrt{(100,31 - 142,45)^2 + (128,94 - 145,87)^2 + (100,22 - 118,84)^2} = 49,08$$

$$D_{x5,c1} = \sqrt{(118,32 - 142,45)^2 + (118,29 - 145,87)^2 + (109,22 - 118,84)^2} = 37,88$$

Dan seterusnya sampai diperoleh hasil dari $D_{x34,c3}$.

3) Sehingga didapat tabel jarak dari *Centroid* dan mencari nilai minimal dari ketiga *Centroid*. Tabel Jarak dari *Centroid* adalah sebagai berikut :

Tabel 3. Jarak Centroid Iterasi ke-1

| C1 | C2 | C3 | Jarak Centroid | C1 | C2 | C3 |
|---------|---------|---------|----------------|-----|-----|-----|
| 36,5207 | 32,0946 | 31,7724 | C3 | | | 1 |
| 30,0209 | 30,8442 | 38,6833 | C1 | 1 | | |
| 36,2459 | 26,4034 | 47,5252 | C2 | | 1 | |
| 49,0827 | 43,5907 | 36,4330 | C3 | | | 1 |
| 37,8874 | 29,4389 | 30,4234 | C2 | | 1 | |
| 40,6652 | 19,0963 | 34,2199 | C2 | | 1 | |
| 0,0000 | 58,0587 | 67,8615 | C1 | 1 | | |
| 30,7760 | 27,6773 | 43,8774 | C2 | | 1 | |
| 34,2099 | 48,0420 | 54,3124 | C1 | 1 | | |
| 78,1774 | 23,7442 | 38,8491 | C2 | | 1 | |
| 65,2237 | 38,8675 | 19,0559 | C3 | | | 1 |
| 53,6060 | 29,8801 | 17,0068 | C3 | | | 1 |
| 58,0587 | 0,0000 | 31,1074 | C2 | | 1 | |
| 62,9680 | 33,4765 | 7,5117 | C3 | | | 1 |
| 49,6745 | 24,5637 | 19,5722 | C3 | | | 1 |
| ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... |
| 57,8220 | 31,9410 | 21,8963 | C3 | | | 1 |
| 71,5162 | 31,5560 | 7,2422 | C3 | | | 1 |

Selanjutnya dalam metode *K-Means*, perhitungan berhenti apabila *Cluster* pada iterasi yang dihasilkan sama pada iterasi sebelumnya. Maka selanjutnya mencari *Cluster* pada iterasi selanjutnya sampai nilai Iterasinya sama. Untuk



mencari nilai *Centroid* selanjutnya dengan menggunakan *Centroid* baru pada Iterasi ke-1 dengan menjumlahkan nilai sesuai yang tertera pada *Cluster* di tabel diatas. Adapun *Centroid* baru untuk mencari *Cluster* selanjutnya adalah dengan menjumlahkan nilai yang terpilih pada *Cluster* tersebut kemudian memmbagikannya sebanyak jumlah nilai sebagai berikut :

$$C_{1,2018} = (124,15 + 142,45 + 153,04 + 195,63 + 129,71 + 126,50)/6 = 145,24$$

$$C_{2,2018} = (117,78 + 118,32 + 124,26 + 128,53 + 88,63 + 110,70 + 87,48 + 92,21 + 104,09 + 136,66 + 119,62)/11 = 111,662$$

$$C_{3,2018} = (119,55 + 100,31 + 94,42 + 106,81 + 110,16 + 119,35 + 112,69 + 110,48 + 104,80 + 122,41 + 111,47 + 132,27 + 106,72 + 123,30 + 105,49 + 126,51 + 97,78)/17 = 112,031$$

$$C_{1,2019} = 1427,24$$

$$C_{2,2019} = 111,00$$

$$C_{3,2019} = 104,402$$

$$C_{1,2020} = 112,277$$

$$C_{2,2020} = 124,87$$

$$C_{3,2020} = 101,73$$

Berikut, data *Centroid* baru Iterasi ke-1 adalah sebagai berikut:

Tabel 4. *Centroid* Baru Iterasi Ke-1

| Centroid | 2018 | 2019 | 2020 |
|----------|---------|---------|---------|
| C1 | 145,247 | 127,24 | 112,277 |
| C2 | 111,662 | 113,003 | 124,865 |
| C3 | 112,031 | 104,402 | 101,734 |

Dengan menggunakan langkah – langkah yang sama seperti sebelumnya untuk menentukan Jarak dari *Centroid* dengan menggunakan *Centroid* baru Iterasi ke-1, maka berikut hasil Jarak dari *Centroid* :

Tabel 5. Jarak *Centroid* Iterasi ke-2

| C1 | C2 | C3 | Jarak Centroid | C1 | C2 | C3 |
|----------|-------------|----------|----------------|-----|-----|-----|
| 27,18919 | 20,38167436 | 18,07036 | C3 | | | 1 |
| 21,88821 | 18,22558303 | 25,59276 | C2 | | 1 | |
| 34,80712 | 14,91169293 | 37,44839 | C2 | | 1 | |
| 46,55703 | 31,46837886 | 27,23578 | C3 | | | 1 |
| 28,53929 | 17,80637206 | 16,98491 | C3 | | | 1 |
| 29,10556 | 13,31852348 | 24,3624 | C2 | | 1 | |
| 19,94933 | 45,43650359 | 54,19951 | C1 | 1 | | |
| 22,70161 | 17,98111116 | 32,40675 | C2 | | 1 | |
| 15,48114 | 43,07133622 | 43,55679 | C1 | 1 | | |
| 70,67761 | 33,07497825 | 42,59479 | C2 | | 1 | |
| 56,71778 | 34,19840087 | 19,01405 | C3 | | | 1 |
| 43,5408 | 23,53062056 | 7,299754 | C3 | | | 1 |
| 48,18329 | 15,31783169 | 28,53163 | C2 | | 1 | |
| 49,16685 | 32,98428847 | 9,606235 | C3 | | | 1 |
| 35,52081 | 21,51999397 | 9,164998 | C3 | | | 1 |
| 44,02094 | 36,6645006 | 14,41838 | C3 | | | 1 |
| 48,45419 | 31,85538908 | 8,764963 | C3 | | | 1 |
| ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... |
| 40,45363 | 33,65544701 | 18,59294 | C3 | | | 1 |
| 59,63689 | 33,93543769 | 18,36686 | C3 | | | 1 |

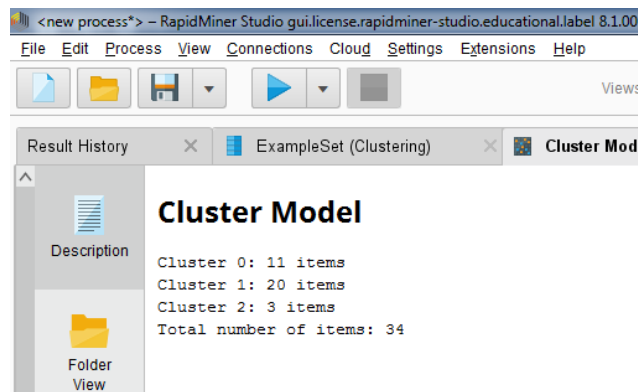
Perulangan dilakukan hingga diperoleh hasil cluster yang sama antara iterasi 3 dan iterasi 4 sehingga perhitungan berhenti pada iterasi 4. Berikut adalah hasil cluster pada iterasi 4:

Tabel 6. Jarak *Centroid* Iterasi ke-4

| C1 | C2 | C3 | Jarak Centroid | C1 | C2 | C3 |
|----------|-------------|----------|----------------|----|----|----|
| 37,14474 | 20,12378614 | 16,75646 | C3 | | | 1 |
| 32,32562 | 17,01564135 | 24,11682 | C2 | | 1 | |
| 44,31278 | 12,7149867 | 36,03722 | C2 | | 1 | |
| 55,84656 | 31,66124797 | 26,90981 | C3 | | | 1 |
| 38,65475 | 17,6685567 | 15,59421 | C3 | | | 1 |
| 38,72871 | 13,03504945 | 22,95011 | C2 | | 1 | |

| C1 | C2 | C3 | Jarak Centroid | C1 | C2 | C3 |
|----------|-------------|----------|----------------|-----|-----|-----|
| 22,18657 | 43,46044069 | 52,83656 | C1 | 1 | | |
| 32,34801 | 16,31803312 | 30,89846 | C2 | | 1 | |
| 16,24101 | 42,19235169 | 42,43371 | C1 | 1 | | |
| 80,6046 | 34,85298807 | 42,45774 | C2 | | 1 | |
| 66,13501 | 35,73831373 | 19,89402 | C3 | | | 1 |
| 53,1352 | 24,86781464 | 7,353994 | C3 | | | 1 |
| 57,69591 | 16,95190672 | 27,74465 | C2 | | 1 | |
| 57,6272 | 34,64454785 | 11,01799 | C3 | | | 1 |
| 44,59432 | 22,64446682 | 8,109761 | C3 | | | 1 |
| 51,92644 | 37,62908372 | 15,11299 | C3 | | | 1 |
| 56,99527 | 33,52361617 | 10,13449 | C3 | | | 1 |
| ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... |
| 47,5166 | 34,81386421 | 18,55889 | C3 | | | 1 |
| 68,8912 | 36,03853487 | 19,49838 | C3 | | | 1 |

Berdasarkan Tabel 6, *Cluster 2 (C1)* memiliki 3 data yang diartikan bahwa provinsi pertama adalah kategori yang Angka Partisipasi Kasar (Apk) Rasio Angka Partisipasi Kasar Pendidikan Yang Diolah Dari BPS Tahun 2018, 2019 Dan 2020 dari Jumlah 34 Provinsi dengan kategori tertinggi. *Cluster 0 (C2)* memiliki 11 data dan *Cluster 1 (C3)* memiliki 20 data yang diartikan bahwa provinsi ketiga adalah kategori yang Angka Partisipasi Kasar (Apk) Rasio Angka Partisipasi Kasar Pendidikan Yang Diolah Dari BPS Tahun 2018, 2019 Dan 2020 dari Jumlah 34 Provinsi dengan katgori rendah. Setelah diperoleh hasil akhir dari pengolahan data secara manual menggunakan algoritma k-means, selanjutnya dilakukan pengolahan data menggunakan tools rapidminer dengan melakukan proses pembentukan permodelan yaitu melakukan *importing* data dan menghubungkan dengan operator k-means. Berikut adalah hasil dari pengolahan data menggunakan tools rapidminer.



Gambar 1. Kelompok Angka Partisipasi Kasar (Apk) Rasio Angka Partisipasi Kasar Pendidikan yang termasuk *cluster 0*

Hasil pada *cluster 0* (Sedang) berjumlah 11 provinsi Angka Partisipasi Kasar (APK) Rasio Angka Partisipasi Kasar Pendidikan Yang Diolah Dari BPS Tahun 2018, 2019 Dan 2020 dari 34 Provinsi. Ini menunjukkan bahwa adanya beberpa provinsi yang Angka Partisipasi Kasar (APK) Rasio Angka Partisipasi Kasar Pendidikan yang dengan pratisipasi skala sedang. Dan perlu adanya penyuluhan dalam meningkat Angka Partisipasi Kasar (Apk) Rasio Angka Partisipasi Kasar Pendidikan tersebut.

4. KESIMPULAN

Berdasarkan hasil penelitian yang dilakukan, penulis menarik beberapa kesimpulan diantaranya: Jumlah Cluster yang digunakan pada data akta Angka Partisipasi Kasar (Apk) sebanyak 3 Cluster. Cluster tersebut diantaranya dengan jenis Angka Partisipasi Kasar (Apk) Rasio Angka Partisipasi Kasar Pendidikan probilitas tinggi dengan kategori Menengah (C1), dengan jenis Angka Partisipasi Kasar (Apk) Rasio Angka Partisipasi Kasar Pendidikan probilitas sedang dengan kategori kecil (C2) dan dengan jenis Angka Partisipasi Kasar (Apk) Rasio Angka Partisipasi Kasar Pendidikan probilitas rendah dengan kategori kecil (C3) berdasarkan data angka partisipasi kasar (apk) rasio angka partisipasi kasar pendidikan yang diolah dari BPS tahun 2018, 2019 dan 2020 dengan jumlah 34 provinsi.

REFERENCES

[1] Istiqomah Anny, Sukidin, And S. Pudjo, “Analisis Partisipasi Pendidikan Pada Masyarakat Miskin Dusun Gumuk Limo Desa Nogosari Kecamatan Rambipuji Kabupaten Jember,” *J. Pendidik. Ekon.*, Vol. 12, Pp. 227–235, 2018, Doi:



- 10.19184/Jpe.V12i2.8553.
- [2] T. Imandasari, E. Irawan, A. P. Windarto, And A. Wanto, “Algoritma Naive Bayes Dalam Klasifikasi Lokasi Pembangunan Sumber Air,” *Semin. Nas. Ris. Inf. Sci.*, 2019, Doi: 10.30645/Senaris.V1i0.81.
 - [3] M. A. Amri, A. P. Windarto, A. Wanto, I. S. Damanik, And S. Informasi, “Analisis Metode K-Means Pada Pengelompokan Perguruan Tinggi,” Vol. 3, Pp. 674–679, 2019, Doi: 10.30865/Komik.V3i1.1677.
 - [4] J. P. Informatika, J. T. Hutabarat, I. Pendahuluan, A. D. Mining, And B. A. K-Means, “Penerapan Algoritma K-Means Untuk Menentukan Lokasi Promosi Produk Minuman Penurun Kolesterol Nutriva Benecol Pada Pt . Perkasa / Kalbe Nutritional,” Vol. 17, Pp. 465–472, 2018.
 - [5] N. A. Hasibuan *Et Al.*, “Implementasi Data Mining Untuk Pengaturan Layout,” Vol. 4, No. 4, Pp. 6–11, 2017.
 - [6] Y. Mardiyanto, “Data Mining : Klasifikasi Menggunakan Algoritma C4.5,” *J. Edik Inform.*, Vol. 2, No. 2, Pp. 213–219, 2017.
 - [7] C. Fadlan, S. Ningsih, And A. P. Windarto, “Penerapan Metode Naive Bayes Dalam Klasifikasi Kelayakan Keluarga Penerima Beras Rastra,” *Jutim*, Vol. 3, No. 1, Pp. 1–8, 2018.
 - [8] M. H. Adiyanto And Y. Desnelita, “Jurnal Nasional Teknologi Dan Sistem Informasi Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan Pada RSUD Pekanbaru,” Vol. 01, Pp. 17–24, 2019.
 - [9] F. L. Sibuea, A. Sapta, S. Informasi, And S. Royal, “Pemetaan Siswa Berprestasi Menggunakan Metode K-Means Clustering,” Vol. Iv, No. 1, 2017.
 - [10] I. Parlina, A. P. Windarto, A. Wanto, And M. R. Lubis, “Memanfaatkan Algoritma K-Means Dalam Menentukan Pegawai Yang Layak Mengikuti Assessment Center Untuk Clustering Program Sdp,” *Cess (Journal Comput. Eng. Syst. Sci.)*, Vol. 3, No. 1, Pp. 87–93, 2018.