

Analisis Mutu Pendidikan Sekolah Menengah Atas Program Ilmu Alam di Jawa Tengah dengan Algoritme K-Means Terorganisir

Ridho Ananda ^{#1}, Auliya Burhanuddin ^{*2}

*Fakultas Teknologi Industri dan Informatika, Institut Teknologi Telkom Purwokerto
Jl. D.I. Panjaitan 128 Purwokerto, Jawa Tengah, Indonesia*

¹ ridho@ittelkom-pwt.ac.id

² aulya@ittelkom-pwt.ac.id

Accepted on 20-08-2019

Abstract

Banyaknya jumlah sekolah menengah atas (SMA) di Jawa Tengah dengan mutu pendidikan yang berbeda-beda menjadi kendala bagi stakeholder dalam mengambil kebijakan. Untuk mengatasinya, dibutuhkan metode dalam menganalisis data sekolah yang berkaitan dengan mutu pendidikan. Salah satu metode yang dapat digunakan adalah penggerombolan. Pada penelitian ini diterapkan metode penggerombolan dengan algoritme K-means serta kombinasi K-means dan Hirarki pada nilai ujian akhir nasional (UAN) program ilmu alam. Nilai UAN merupakan salah satu komponen penilaian mutu pendidikan. Penentuan banyak gerombol optimal digunakan Bayesian Information Criterion (BIC) dan diperoleh 5 gerombol optimal dengan BIC 221.45. Hasil penggerombolan terbaik berdasarkan nilai Silhouette diperoleh algoritme complete K-means dengan nilai 0.4537, sehingga hasil tersebut digunakan untuk menganalisis mutu pendidikan di Jawa Tengah. Berdasarkan hasil penggerombolan, diperoleh kesimpulan bahwa sekolah yang unggul banyak terdapat di kota Semarang dengan proporsi 12.76% dari seluruh sekolah unggul pada 35 wilayah di Jawa Tengah. Sedangkan sekolah terbanyak pada peringkat terendah di Boyolali dengan proporsi 9.03% dari seluruh sekolah peringkat terakhir pada 35 wilayah di Jawa Tengah. Lima wilayah yang perbedaan mutunya tidak merata ialah Banjarnegara, Demak, Kab. Pekalongan, Batang, dan Purwodadi. Sedangkan lima wilayah yang perbedaan mutunya paling merata adalah Wonosobo, Tegal, Semarang, dan Magelang.

Keywords: *Bayesian Information Criterion*, Hirarki, K-Means, Mutu pendidikan, Penggerombolan

I. PENDAHULUAN

Mutu pendidikan merupakan derajat keunggulan dalam pengelolaan pendidikan secara efektif dan efisien untuk memunculkan keunggulan akademis dan ekstrakurikuler pada peserta didik yang dinyatakan lulus [1]. Untuk mengetahui mutu pendidikan suatu sekolah maka diperlukan beberapa indikator yang mengembarkannya. Secara garis besar, ada tiga indikator yang menentukan mutu pendidikan, salah satunya adalah komponen penilaian [2]. Pemerintah Indonesia telah melakukan penilaian yang terstandarisasi dan diberikan dalam skala nasional untuk mengetahui mutu pendidikan formal yang diselenggarakan di seluruh wilayah [3][4]. Penilaian tersebut lebih dikenal dengan Ujian Akhir Nasional (UAN).

Di propinsi Jawa tengah terdapat hampir seribu sekolah jenjang sekolah menengah atas (SMA) program ilmu alam dengan mutu pendidikan yang berbeda-beda. Hal tersebut tampak dari rata-rata hasil UAN tahun 2019 yang variatif untuk setiap sekolah pada setiap wilayah di Jawa Tengah [5]. Jumlah yang cukup banyak tersebut menjadi kendala bagi stakeholder pendidikan saat mengambil kebijakan yang berhubungan dengan mutu pendidikan SMA program ilmu alam. Untuk mengatasi masalah tersebut, tentunya dibutuhkan suatu metode untuk menganalisis kondisi mutu pendidikan seluruh sekolah pada setiap wilayah di Jawa Tengah. Hasil analisis diharapkan dapat memberikan gambaran kondisi mutu pendidikan. Salah satu metode yang dapat diterapkan adalah metode gerombol.

Metode gerombol merupakan salah satu metode dalam data mining yang penggunaannya cukup intens di beberapa permasalahan seperti teknologi[6], pendidikan[7], ekonomi[8], dan bidang lainnya. Ada beberapa algoritme penggerombolan yang dapat digunakan untuk melakukan proses penggerombolan, salah satu algoritme yang paling terkenal adalah k-Means[9]. K-means merupakan algoritme yang memanfaatkan ukuran jarak Euclidean untuk menggerombolkan objek-objek dengan meminimumkan jarak antar objek pada gerombol sama serta memaksimalkan jarak setiap dua objek pada gerombol berbeda. Proses awal dari K-means mensyaratkan penentuan awal dari banyak gerombol dan centroid awal. Penentuan awal tersebut sangat berpengaruh terhadap hasil penggerombolan. Beberapa ilmuwan telah meneliti terkait penentuan jumlah gerombol dan centroid awal pada metode k-means. Hasil dari penelitian ialah penentuan banyak gerombol optimal dengan Bayesian Information Criterion (BIC) [10], centroid awal dapat diperoleh dari algoritme penggerombolan hirarki [11][12][13].

Berdasarkan uraian sebelumnya, maka penelitian ini bertujuan menganalisis mutu pendidikan SMA program ilmu alam pada setiap wilayah di Jawa Tengah dengan metode penggerombolan. Algoritme yang diimplementasikan adalah algoritme k-means terorganisir. K-means terorganisir pada penelitian ini merupakan algoritme k-means yang mana penentuan banyak gerombol awal berdasarkan kriteria BIC dan centroid awal berdasarkan algoritme penggerombolan hirarki. Hasil penggerombolan optimal kemudian dianalisis untuk mendapatkan informasi terkait mutu pendidikan. Hasil analisis merupakan informasi yang dapat digunakan oleh stakeholder dalam mengambil kebijakan.

II. RESEARCH METHOD

A. DATA

Data yang digunakan pada penelitian ini adalah laporan hasil UAN program ilmu alam tahun 2019 pada setiap wilayah di Jawa Tengah yang dikeluarkan secara resmi oleh Kementerian Pendidikan dan Kebudayaan [5]. Data yang diperoleh merupakan data lengkap berbentuk matriks dengan ukuran 786×6 dimana 786 merupakan objek berupa sekolah dan 6 merupakan mata pelajaran yang diujikan pada UAN tahun 2019 antara lain bahasa indonesia, bahasa inggris, matematika, fisika, kimia, dan biologi.

B. K-MEANS TERORGANISIR

K-Means terorganisir merupakan metode k-means yang diterapkan pada penelitian ini dimana banyak gerombol awal dan centroid awal ditentukan dengan menggunakan algoritme lain. Pada penelitian ini, gerombol awal ditentukan dengan kriteria BIC dengan asumsi data berdistribusi Gaussian [10]. Perhitungan BIC digunakan formula seperti pada Persamaan 1.

$$BIC = n \cdot \ln\left(\frac{RSS}{n}\right) + k * (p + 1) * \ln(n) \quad (1)$$

n merupakan banyak objek pada data, k merupakan banyak gerombol yang dibentuk, dan p merupakan banyak peubah pada data. RSS atau *Root sums square* merupakan jarak jumlah seluruh jarak objek terhadap pusat dari gerombol dimana objek itu berada. Centroid awal diperoleh dari algoritme penggerombolan hirarki yang kemudian dibandingkan dengan k-means secara langsung. Algoritme penggerombolan hirarki yang digunakan adalah *single linkage*, *complete linkage*, *average linkage*, *ward linkage*, dan *weighted linkage*. Selanjutnya

dari hasil k-means dan kombinasi lagoritme hirarki k-means, diambil hasil penggerombolan yang paling efektif.

C. UKURAN KEEFEKTIFAN

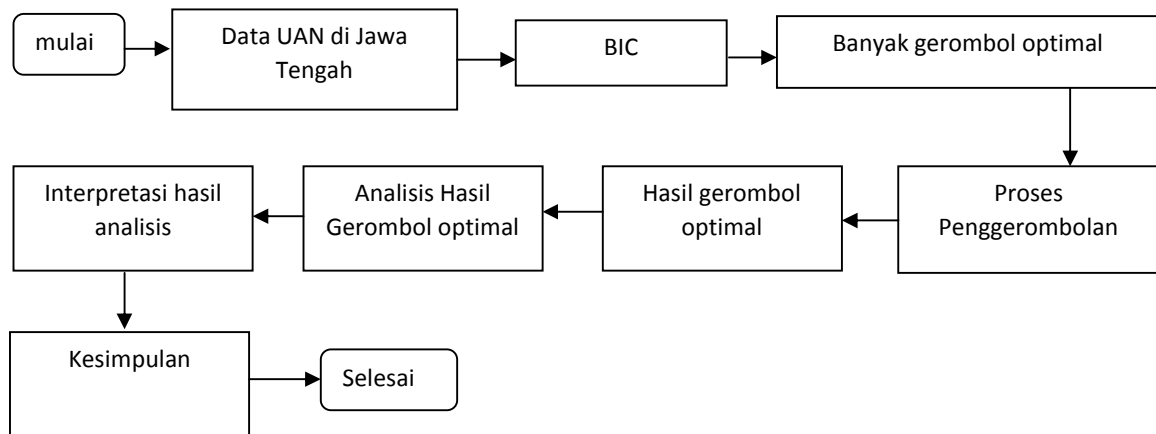
Ukuran keefektifan dibutuhkan untuk mengetahui kualitas dari hasil penggerombolan yang diperoleh. Ada dua validasi yang dapat digunakan pada metode penggerombolan, yaitu validasi internal dan validasi eksternal [14]. Pada penelitian ini digunakan validasi internal untuk mengetahui ukuran keefektifan dari metode penggerombolan. Penggunaan validasi internal dikarenakan tidak seperti validasi eksternal yang mensyaratkan adanya informasi gerombol awal, disamping ini validasi internal lebih akurat dibandingkan dengan validasi eksternal [15][16]. Validasi internal yang digunakan adalah Ukuran Silhouette yang dirumuskan seperti pada Persamaan 2.

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \tag{2}$$

s(i) merupakan nilai Silhouette untuk setiap objek-i. b(i) merupakan jarak rata-rata objek ke-I dengan seluruh objek pada gerombol terdekat. a(i) merupakan jarak jata-rata objek ke-i dengan seluruh objek dalam gerombol yang sama. Hasil dari seluruh s(i) kemudian dirata-rata untuk memperoleh ukuran Silhouette. Ukuran Silhouette memiliki nilai pada interval [-1,1] dimana nilai mendekati -1 menunjukkan hasil penggerombolan yang buruk dan nilai mendekati 1 menunjukkan hasil yang baik[17]. Silhouette merupakan ukuran keefektifan penggerombolan yang paling baik dibandingkan dengan beberapa ukuran keefektifan yang lain [18].

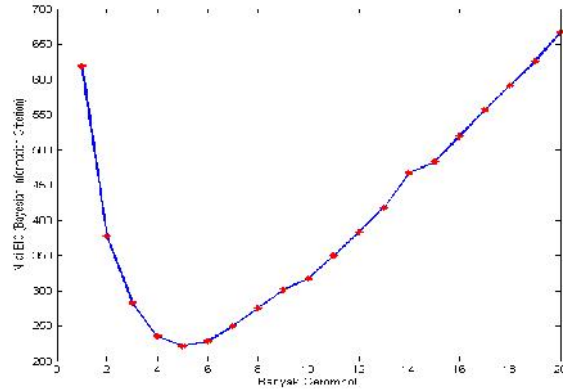
D. LANGKAH-LANGKAH PENELITIAN

Penelitian ini menggunakan Matlab 2014a dalam komputasinya. Langkah-langkah dalam penelitian ini terdiri dari persiapan data, pembentukan gerombol, dan analisis hasil gerombol. Pada langkah analisis hasil gerombol, prosedur analisis yang dilakukan antara lain melihat profil dari tiap gerombol yang terbentuk, sebaran sekolah pada hasil gerombol, sebaran sekolah untuk tiap wilayah di Jawa tengah pada hasil gerombol dan proporsinya, simpangan baku dari sebaran yang terjadi di setiap wilayah. Langkah-langkah penelitian selanjutnya dijabarkan dengan lebih sederhana pada diagram alir penelitian yang disajikan oleh Gambar 1.



Gambar 1. Diagram Alir Penelitian

III. HASIL

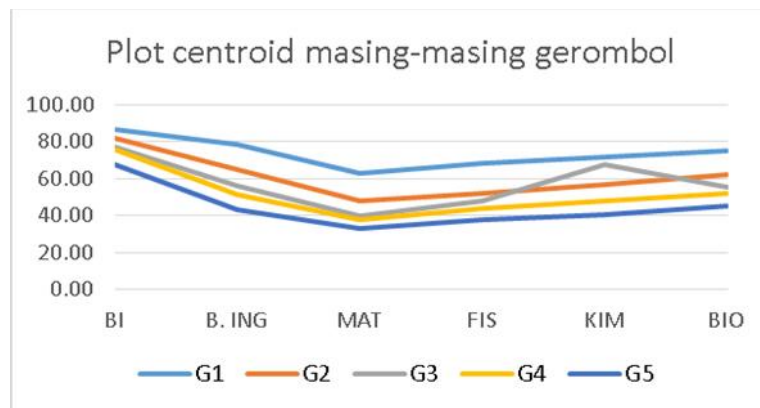


Gambar 2. Ukuran Bayesian Information Criterion

Berdasarkan kriteria BIC, diperoleh banyak gerombol optimal untuk data UAN SMA program ilmu alam pada setiap wilayah di Jawa Tengah ialah 5 gerombol. Hal tersebut diketahui ketika uji coba perhitungan BIC untuk 1 s.d. 20 gerombol, pada 5 gerombol, nilai BIC paling minimum yaitu sebesar 221.45, sebagaimana ditunjukkan pada Gambar 2. Selanjutnya dilakukan proses penggerombolan dengan algoritme k-means serta kombinasi k-means dengan algoritme penggerombolan hirarki. Nilai Silhouette dari setiap hasil penggerombolan ditunjukkan pada Tabel 1. Pada tabel tersebut, hasil penggerombolan optimal diperoleh pada algoritme *Complete k-means* dengan nilai Silhouette 0.537, sehingga dipilih hasil penggerombolan dengan algoritme tersebut.

Tabel 1. Ukuran Silhouette Untuk Setiap Algoritme

Algoritme	Kmeans	Single K-means	Complete K-means	Average K-means	Ward K-means	Weighted K-means
Silhouette	0.4245	0.437	0.4537	0.4292	0.4285	0.426



Gambar 3. Centroid Masing-Masing Gerombol

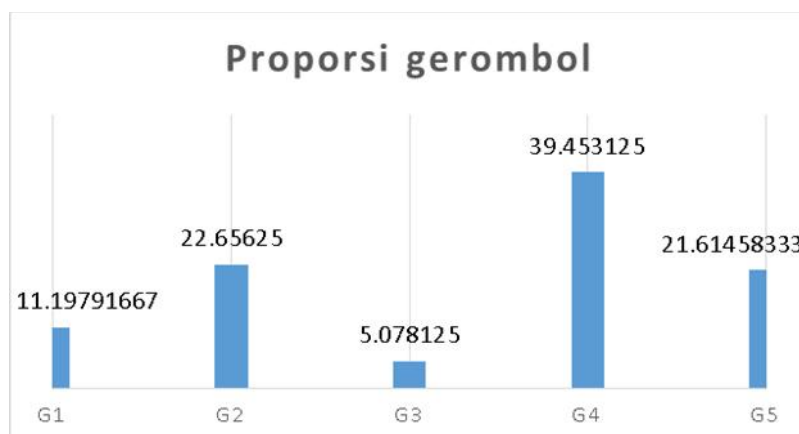
Hasil penggerombolan optimal yang dipilih terdapat 5 gerombol yang kemudian masing-masing disebut sebagai gerombol 1 (G1), gerombol 2 (G2), gerombol 3 (G3), gerombol 4 (G4), dan gerombol 5 (G5). Centroid dari kelima gerombol tersebut divisualisasikan pada ruang berdimensi dua seperti pada Gambar 3. Sebaran sekolah berdasarkan hasil penggerombolan yaitu 86 sekolah di G1, 174 sekolah di G2, 39 sekolah di G3, 303 sekolah di G5, dan 166 sekolah di G5, sebaran untuk tiap daerah dapat dilihat pada Tabel 2.

Tabel 2. Sebaran sekolah di setiap daerah untuk setiap gerombol

No	Kota	G1	G2	G3	G4	G5
1	SEMARANG	11 (12.79)	20 (11.49)	0 (0)	15 (4.95)	11 (6.62)
2	SURAKARTA	7 (8.13)	14 (8.04)	1 (2.56)	6 (1.98)	1 (0.6)
3	TEGAL	2 (2.32)	2 (1.14)	0 (0)	3 (0.99)	2 (1.2)
4	PEKALONGAN	3 (3.48)	2 (1.14)	0 (0)	5 (1.65)	1 (0.6)
5	SALATIGA	2 (2.32)	4 (2.29)	0 (0)	1 (0.33)	0 (0)
6	MAGELANG	4 (4.65)	3 (1.72)	2 (5.12)	2 (0.66)	0 (0)
7	BANYUMAS	5 (5.81)	6 (3.44)	3 (7.69)	20 (6.6)	3 (1.8)
8	BANJARNEGARA	1 (1.16)	2 (1.14)	0 (0)	10 (3.3)	0 (0)
9	CILACAP	2 (2.32)	10 (5.74)	2 (5.12)	14 (4.62)	11 (6.62)
10	PURBALINGGA	2 (2.32)	5 (2.87)	1 (2.56)	8 (2.64)	1 (0.6)
11	KEBUMEN	4 (4.65)	9 (5.17)	1 (2.56)	11 (3.63)	0 (0)
12	KAB. MAGELANG	6 (6.97)	10 (5.74)	1 (2.56)	12 (3.96)	2 (1.2)
13	PURWOREJO	3 (3.48)	9 (5.17)	0 (0)	5 (1.65)	1 (0.6)
14	TEMANGGUNG	1 (1.16)	4 (2.29)	0 (0)	8 (2.64)	3 (1.8)
15	WONOSOBO	1 (1.16)	5 (2.87)	4 (10.25)	4 (1.32)	2 (1.2)
16	BOYOLALI	2 (2.32)	6 (3.44)	0 (0)	10 (3.3)	8 (4.81)
17	KARANGANYAR	1 (1.16)	5 (2.87)	2 (5.12)	4 (1.32)	2 (1.2)
18	KLATEN	4 (4.65)	2 (1.14)	0 (0)	11 (3.63)	2 (1.2)
19	SRAGEN	4 (4.65)	2 (1.14)	0 (0)	5 (1.65)	7 (4.21)
20	SUKOHARJO	5 (5.81)	6 (3.44)	0 (0)	10 (3.3)	2 (1.2)
21	WONOGIRI	1 (1.16)	7 (4.02)	2 (5.12)	6 (1.98)	1 (0.6)
22	REMBANG	1 (1.16)	2 (1.14)	1 (2.56)	9 (2.97)	2 (1.2)
23	BLORA	1 (1.16)	3 (1.72)	0 (0)	7 (2.31)	6 (3.61)
24	JEPARA	1 (1.16)	3 (1.72)	4 (10.25)	12 (3.96)	11 (6.62)
25	KUDUS	2 (2.32)	7 (4.02)	2 (5.12)	11 (3.63)	3 (1.8)
26	PATI	1 (1.16)	7 (4.02)	3 (7.69)	14 (4.62)	10 (6.02)
27	KAB. TEGAL	1 (1.16)	2 (1.14)	1 (2.56)	9 (2.97)	5 (3.01)
28	DEMAK	0 (0)	3 (1.72)	0 (0)	7 (2.31)	15 (9.03)
29	PURWODADI	1 (1.16)	0 (0)	1 (2.56)	12 (3.96)	14 (8.43)
30	KENDAL	1 (1.16)	2 (1.14)	1 (2.56)	11 (3.63)	10 (6.02)
31	KAB. SEMARANG	2 (2.32)	6 (3.44)	2 (5.12)	10 (3.3)	2 (1.2)
32	BATANG	1 (1.16)	0 (0)	0 (0)	7 (2.31)	5 (3.01)
33	BREBES	0 (0)	3 (1.72)	2 (5.12)	10 (3.3)	11 (6.62)
34	KAB. PEKALONGAN	2 (2.32)	1 (0.57)	1 (2.56)	9 (2.97)	1 (0.6)
35	PEMALANG	1 (1.16)	2 (1.14)	2 (5.12)	5 (1.65)	11 (6.62)
	Jumlah	86 (100%)	174 (100%)	39 (100%)	303 (100%)	166 (100%)

IV. RESULT AND DISCUSSIONS

Gambar 3 memberikan informasi terkait profil masing-masing gerombol terhadap gerombol yang lain. Diketahui bahwa G1 memiliki nilai yang unggul di seluruh mata pelajaran yang diujikan, disusul G2 s.d. G5. Namun, hanya pada mata pelajaran kimia, G3 lebih unggul daripada G2. Apabila diurutkan berdasarkan peringkat maka sekolah di dalam G1 masuk dalam peringkat 1. Sedangkan sekolah di dalam G2, G3, G4, dan G5 berturut-turut masuk dalam peringkat 2,3,4, dan 5 berdasarkan tingkat mutu pendidikan. Gambar 4 memberikan informasi proporsi gerombol secara keseluruhan. Dari gambar tersebut diketahui bahwa sekolah di Jawa Tengah mayoritas masuk pada peringkat 4 untuk mutu pendidikan. Urutan kedua pada peringkat ke 2, kemudian peringkat ke 5, peringkat ke 1, dan terakhir peringkat ke 3. Sebanyak 33.86% sekolah di Jawa Tengah masuk dalam peringkat 1 atau peringkat 2. Sedangkan 61.06% sekolah di Jawa Tengah masuk dalam peringkat 4 atau 5.



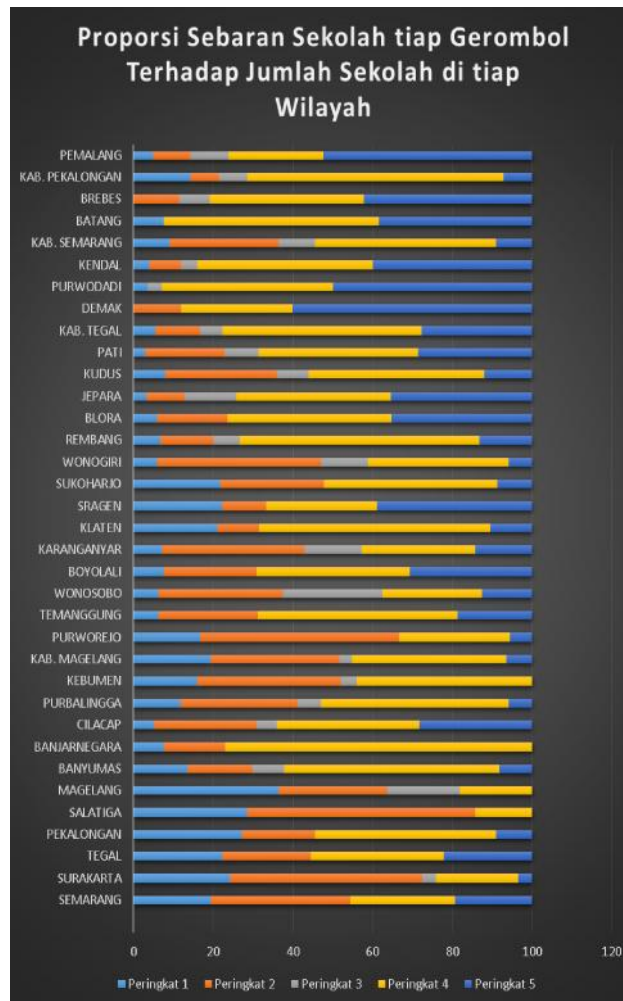
Gambar 4. Proporsi Gerombol Secara Keseluruhan

Hasil sebaran sekolah di tiap-tiap wilayah, berdasarkan Tabel 2 diketahui bahwa pada peringkat 1, 12.79% merupakan sekolah dari Semarang, 8.13% dari Surakarta, 6.97% dari Kab. Magelang, 5.81% dari Banyumas dan Sukoharjo, dan wilayah lainnya di bawah 5%. Pada wilayah Brebes dan Demak, tidak ada sekolah yang masuk pada peringkat 1. Pada peringkat 2, 11.49% dari Semarang, 8.04% dari Surakarta, 5.74% dari Kab. Magelang dan Cilacap, 5.17% dari Kebumen dan Purworejo, dan pada wilayah lainnya di bawah 5%. Pada wilayah Purwodadi dan Batang, tidak ada sekolah yang masuk pada peringkat 2. Pada peringkat 3, 10.25% dari Wonogiri dan Purbalingga, 7.69% dari Kebumen dan Blora, 5.12% dari Kudus, Temanggung, Magelang, Brebes, Sragen, Rembang, Kendal, dan Pemalang. Wilayah lain di bawah 5%. Pada wilayah Semarang, Kab. Magelang, Cilacap, Pati, Banyumas, Sukoharjo, Boyolali, Kab. Semarang, Karanganyar, Salatiga, Klaten, Banjarnegara, Kab. Pekalongan, dan Batang tidak ada sekolah yang masuk pada peringkat 3. Pada peringkat 4, 6.6% dari Kudus, sedangkan wilayah lain kurang dari 5%. Pada peringkat 5, 9.03% dari Boyolali, 8.43% dari Kab. Semarang, 6.62% dari Sragen, Wonogiri, Banjarnegara, Cilacap, dan Batang. 6.02% dari Karanganyar dan Banyumas. Sedangkan wilayah lainnya kurang dari 5%. Pada wilayah Kudus, Temanggung, Kendal, dan Brebes tidak ada sekolah yang masuk pada peringkat 5.

Tabel 3. Deskriptif dari Proporsi masing-masing peringkat

	Peringkat 1	Peringkat 2	Peringkat 3	Peringkat 4	Peringkat 5
Minimum	0	0	0	0.33	0
Maksimum	12.79	11.49	10.25	6.6	9.03
Rata-rata	2.85	2.85	2.85	2.86	2.85
Simpangan Baku	2.58	2.33	2.99	1.30	2.70

Berdasarkan informasi deskriptif pada setiap peringkat yang disajikan oleh Tabel 3, maka diperoleh informasi bahwa sekolah yang masuk pada peringkat 4 tersebar lebih merata di seluruh wilayah dibandingkan dengan peringkat lainnya. Hal tersebut dapat diketahui berdasarkan nilai simpangan baku pada peringkat 4 yang paling minimum. Sebaran yang tidak merata ada pada sekolah-sekolah yang masuk pada peringkat 3 berdasarkan nilai simpangan baku yang paling besar. Gambar 4 memberikan informasi mengenai proporsi dari penggerombolan sekolah berdasarkan peringkat masing-masing terhadap jumlah sekolah di masing-masing wilayah. Pada gambar tersebut diketahui bahwa lima wilayah yang perbedaan mutu pendidikannya paling tidak merata adalah Banjarnegara kemudian disusul oleh Demak, Kab. Pekalongan, Batang, dan Purwodadi. Hal tersebut dapat diketahui kuantitasnya berdasarkan nilai simpangan baku dari kelima wilayah tersebut berturut-turut 32.45, 25.14, 24.95, 24.69, dan 24.30. Sedangkan 5 wilayah yang paling merata dengan urutan Wonosobo, Karanganyar, Tegal, Semarang, dan Magelang. Hal tersebut juga dapat dilihat dari simpangan bakunya berturut-turut 10.27, 11.74, 12.17, 12.93, dan 13.48.



Gambar 5. Proporsi Sebaran Sekolah dalam Gerombol berdasarkan jumlah sekolah tiap wilayah

V. CONCLUSION

Berdasarkan uraian dari hasil dan diskusi, diperoleh beberapa kesimpulan diantaranya pada data hasil ujian nasional sekolah-sekolah di Jawa Tengah, jumlah gerombol optimal berdasarkan kriteria BIC ialah 5 gerombol dan hasil penggerombolan optimal dengan menggunakan algoritme *complete K-means* berdasarkan nilai Silhouette. Profil tiap gerombol memiliki mutu pendidikan yang berurutan dengan peringkat 1 pada gerombol pertama, peringkat 2 pada gerombol kedua, dan seterusnya hingga peringkat 5 pada gerombol ke 5. Mayoritas sekolah-sekolah di Jawa Tengah masuk pada peringkat 4 atau 5. Khusus untuk peringkat 4, sekolah-sekolah yang masuk dalam wilayah tersebut tersebar paling merata di setiap wilayah di Jawa Tengah. Sedangkan sekolah yang masuk ke peringkat 1 atau peringkat 2 sebesar 33.86% dan sisanya masuk dalam peringkat 3. Wilayah yang memiliki proporsi tertinggi di peringkat 1 adalah kota Semarang, artinya sekolah yang unggul banyak terdapat di kota Semarang. Sedangkan wilayah yang memiliki proporsi tertinggi di peringkat 5 adalah kota Boyolali. Hal tersebut mengartikan bahwa daerah Boyolali memiliki sekolah dengan peringkat 5 paling banyak. Berdasarkan proporsi sebaran penggerombolan, diketahui bahwa 5 wilayah yang

perbedaan mutu pendidikannya tidak merata adalah Banjarnegara, Demak, Kab. Pekalongan, Batang, dan Purwodadi, sedangkan yang paling merata adalah Wonosobo, Karanganyar, Tegal, Semarang, dan Magelang.

REFERENCES

- [1] Suti M., 2011., Strategi Peningkatan Mutu di Era Otonomi Pendidikan, *Jurnal MEDTEK*, 3(2).
- [2] Pakpahan R., 2012., Model Alternatif Ujian Akhir, *Jurnal Pendidikan dan Kebudayaan*, 18(2), pp. 121-131.
- [3] BNSP, "Prosedur operasi standar ujian nasional sekolah menengah pertama, madrasah tsanawiyahm sekolah menengah pertama luar biasa, sekolah menengah atas, madrasah aliyah, sekolah menengah atas luar biasa, dan sekolah menengah kejuruan tahun pelajaran 2010/2011", Badan Standar Nasional Pendidikan, Indonesia, (2011).
- [4] Sujita, "Analisis biplot untuk memetakan mutu sekolah yang sesuai dengan nilai ujian nasional", M.SI. Thesis, Dept. Matematika, Institut Pertanian Bogor, (2009).
- [5] Kementerian Pendidikan dan Kebudayaan, "Laporan Hasil Ujian Nasional", Jakarta (ID): Kemendikbud, (2019).
- [6] Hassan A.A., Shah W. M., Husein A. M., Talib M. S., Mohammed A. A. J., Iskandar M. F., 2019, Clustering Approach in Wireless Sensor Networks based on K-means: Limitations and Recommendations, *IJRTE*, 7(65), pp. 119-126.
- [7] Oyelade O. J., Oladipupo O. O., Obagbuwa I. C., 2010., Application of K-means Clustering Algorithm for Prediction of Students' Academic Performance, *IJCSIS*, 7(1), pp. 292-295.
- [8] Ong J. O., 2013., Implementasi Algoritma K-means Clustering untuk Menentukan Strategi Marketing President University, *Jurnal Ilmiah Teknik Industri*, 12(1), pp. 10-20.
- [9] MacQueen J. B., "Some Methods for Classification of High Dimensional Data Sets with Application to Reference Matching", in *Proceeding of 5th Berkeley Symposium on Mathematical Statistics and Probability*, 1967.
- [10] Priestley M. B., "Spectral analysis and Time Series", Academic Press, 1981.
- [11] Arai K., Barakbah A. R., 2007., Hierarchical K-Means: An Algorithm For Centroids Initialization For K-Means, *Rep. Fac. Sci. Engrg. Saga. Univ.*, 36(1), pp. 25-31.
- [12] Alfina T. Santosa B. Barakbah A. R., 2012, Analisa perbandingan metode Hierarchical Clustering, k-Means dan Gabungan keduanya dalam Cluster Data (Studi kasus: Problem kerja praktek jurusan teknik industri ITS), *Jurnal Teknik ITS*, 1, pp. 521-525.
- [13] Widiyanti W., "Sentroid awal metode k-Means dengan pendekatan metode berhirarki dalam pengelompokkan provinsi di Indonesia", Tesis, Departemen Matematika, FMIPA: IPB. 2017.
- [14] Rendon E., Abundez I. M. Gutierrez C., Zagal S. D., Arizmendi A. Quiroz E. M., Arzate H. E., "A comparison of internal and external cluster validation indexes," in *Proceeding of the 2011 American Conference on Applied Mathematics and the 5th WSEAS International Conference on Computer Engineering and Applications*, 2011.
- [15] Rendon E., Abudez I., Arizmendi A., Quiroz E. M., 2011, Internal Versus External Cluster Validation Indexes, *International Journal of Computers and Communications*, 5(1) 1, pp. 27-34.
- [16] Vendramin L., Campello R. J. G. B., Hruschka E. R., "On the comparison of relative clustering validity criteria," in *Proceedings of the 2009 SIAM International Conference on Data Mining*, 2009.
- [17] Baarsch J., Celebi M. E., "Investigation of Internal Validity Measures for K-means Clustering," in *Proceedings of the IMECS*, 2012.
- [18] Mamat A. R., Mohamed F. S., Mohamed M. A., Rawi N. M., Awang M. I., 2018, Silhouette Index for Determining Optimal K-means Clustering on Images in Different Color Models, *International Journal of Engineering and Technology*, 7(2.14), pp. 105-109.