

Penggerombolan Hasil Ujian Nasional SMA Menggunakan K-Rataan Samar*

Nouval Habibie¹, Akbar Rizki^{2‡}, Pika Silvianti^{3ⁿ}

^{1,2,3}Department of Statistics, IPB University, Indonesia

[‡]corresponding author: akbar.ritzky@apps.ipb.ac.id

Copyright © 2020 Nouval Habibie, Akbar Rizki, and Pika Silvianti. This is an open-access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

National examination scores can be a basis for the government to make a mapping of education quality in order to increase it. The mapping can be done by using fuzzy cluster analysis. The objective of this experiment is to cluster districts/cities in Indonesia based on national examination score in natural and social science in 2014/2015 until 2017/2018 school year by using the fuzzy *c*-means method. The evaluation criteria that will be used are the standard deviation ratio, silhouette coefficient, and Xie Beni index. The best cluster size is two clusters, A and B. The clustering result shows cluster A has a higher mean from each subject than cluster B. Therefore, cluster A will be categorized as good, whereas cluster B as bad. The proportion of districts/cities that belong to cluster A decreased each year. The final cluster result can be determined by the mean of its degree of membership from those four school years. The analysis results show that the distribution of education quality is dominated in Java Island and squatter cities. East Nusa Tenggara, West Sulawesi, Central Sulawesi, and North Kalimantan don't have any districts/cities belong to cluster A.

Keywords: cluster analysis, fuzzy *c*-means, mapping, national examination score.

1 Pendahuluan

Pemerintah memiliki tanggung jawab menentukan kebijakan dan standar mutu pendidikan nasional serta memiliki peran untuk mengevaluasi penyelenggaraan pendidikan dalam rangka pengendalian mutu pendidikan nasional. Salah satu cara mengevaluasi standar pendidikan dasar dan menengah secara nasional adalah hasil capaian Ujian Nasional (UN). Menurut Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 4 Tahun 2018 tentang Penilaian Hasil Belajar oleh Satuan Pendidikan dan Penilaian Hasil Belajar oleh Pemerintah, pemerintah menggunakan hasil UN sebagai alat pengendali dan pemetaan mutu pendidikan secara nasional, pendorong peningkatan mutu pendidikan, dan bahan pertimbangan dalam seleksi penerimaan peserta didik baru pada jenjang yang lebih tinggi. Pemetaan mutu pendidikan berdasarkan hasil UN tersebut dapat dilakukan pada kabupaten atau kota se-Indonesia sehingga dapat menjadi dasar bagi pemerintah untuk melakukan evaluasi. Pemetaan tersebut dapat dilakukan menggunakan analisis gerombol.

Secara umum, analisis gerombol memiliki dua metode, yaitu hierarki dan nonhierarki. Ukuran pusat gerombol harus ditentukan terlebih dahulu pada metode nonhierarki sehingga objek akan digerombolkan berdasarkan jaraknya dengan pusat gerombol. Menurut Ross (2010), metode nonhierarki terdapat dua jenis, yaitu secara tegas (*crisp*) dan secara samar (*fuzzy*). Perbedaan kedua metode ini terletak pada asumsi yang digunakan sebagai dasar pengalokasian objek pada setiap gerombol. Penggerombolan secara tegas tidak memungkinkan satu objek masuk dalam dua gerombol atau lebih, sedangkan penggerombolan secara samar dapat memberikan suatu objek masuk pada setiap gerombol dengan nilai atau derajat keanggotaan tertentu sehingga mendapatkan hasil yang lebih teliti (Yang, 1993).

Penelitian terkait hasil UN telah dilakukan oleh beberapa peneliti, antara lain Pangestu (2018) menggunakan analisis regresi peubah ganda, Nugroho (2016) menggunakan analisis regresi spasial Durbin, Intendia (2012) menggunakan klasifikasi dengan metode pohon keputusan serta regresi logistik, dan Mongi (2014) menggunakan analisis gerombol untuk menggerombolkan kabupaten/kota di Provinsi Jawa Barat. Kemudian, penelitian terkait metode k-rataan samar, antara lain penelitian Agustin & Djuraidah (2010) dalam penggerombolan daerah tertinggal di Indonesia mengatakan bahwa keunggulan metode k-rataan samar adalah nilai derajat keanggotaan tidak berubah meskipun terdapat pencilan. Serban & Moldovan (2006) mengatakan bahwa metode k-rataan samar memiliki performa yang lebih baik daripada k-rataan (tegas) pada aspek penambangan. Putri (2017) melakukan simulasi metode k-rataan samar pada data yang mengandung derau (*noise*) sebanyak 10% lalu menghasilkan rata-rata keakuratan sebesar 99.42%. Khairunnisa (2017) juga melakukan simulasi perbandingan metode k-rataan samar dan tegas dengan rancangan pada peubah yang berkorelasi, hasilnya metode k-rataan samar memiliki performa yang lebih baik.

Penelitian ini bertujuan menggerombolkan kabupaten/kota di Indonesia berdasarkan hasil UN pada jurusan Ilmu Pengetahuan Alam (IPA) dan Ilmu Pengetahuan Sosial (IPS) pada tahun ajaran 2014/2015 sampai 2017/2018 menggunakan metode k-rataan samar. Jurusan IPA dan IPS dipilih karena hampir seluruh kabupaten/kota di Indonesia memiliki kedua jurusan tersebut. Tahun ajaran 2014/2015 sampai 2017/2018 dipilih karena ujian nasional berbasis komputer (UNBK) telah diberlakukan di Indonesia. Metode k-rataan samar dapat mengatasi pencilan dan perbedaan hasil UN antardaerah

yang tidak merata. Selain pencilan pada nilai mata pelajaran, metode k-rataan samar juga dapat mengatasi korelasi yang kuat pada nilai-nilai mata pelajaran. Solusi gerombol akhir dapat ditentukan menggunakan rata-rata dari derajat keanggotaan metode samar sebagai simpulan akhir mutu pendidikan dalam empat tahun tersebut. Hasil penggerombolan dapat digunakan untuk membuat pemetaan mutu pendidikan sebagai salah satu cara mengevaluasi kualitas pendidikan di Indonesia. Tujuan penelitian ini adalah menggerombolkan kabupaten/kota di Indonesia berdasarkan hasil UN SMA jurusan IPA dan IPS pada tahun ajaran 2014/2015 sampai 2017/2018 menggunakan analisis gerombol dengan metode k-rataan samar. Hasil penggerombolan tersebut digunakan untuk membuat pemetaan sebaran mutu pendidikan. Pemetaan mutu pendidikan tersebut dapat digunakan untuk mengetahui penyebaran dan pemerataan kualitas pendidikan di Indonesia.

2 Metodologi

2.1 Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang berasal dari situs web Badan Penelitian dan Pengembangan Kementerian Pendidikan dan Kebudayaan (Kemendikbud). Data terdiri dari delapan gugus data atau skenario. Satuan data yang digunakan adalah kabupaten atau kota se-Indonesia yang berjumlah 514 data, akan tetapi terdapat kabupaten/kota yang tidak memiliki satuan pendidikan yang mengikuti UN. Tabel 1 memberikan ilustrasi mengenai data yang digunakan. Tabel 1 memberikan ilustrasi mengenai data yang digunakan.

Tabel 1: Ilustrasi gugus data yang digunakan

Gugus data	Jurusan	Tahun ajaran	Jumlah data
1	IPA	2014/2015	510
2	IPA	2015/2016	511
3	IPA	2016/2017	512
4	IPA	2017/2018	513
5	IPS	2014/2015	513
6	IPS	2015/2016	513
7	IPS	2014/2015	513
8	IPS	2017/2018	513

Setiap gugus data memiliki peubah kode kabupaten atau kota, nama kabupaten atau kota, jumlah satuan pendidikan, jumlah peserta, mata pelajaran wajib, mata pelajaran pilihan, dan rerata. Mata pelajaran wajib terdiri atas Bahasa Indonesia, Bahasa Inggris, dan Matematika. Mata pelajaran pilihan jurusan IPA terdiri atas Fisika, Kimia, dan Biologi. Jurusan IPS terdiri atas Ekonomi, Geografi, dan Sosiologi. Daftar peubah yang digunakan secara ringkas terdapat pada Tabel 2

2.2 Prosedur Analisis Data

Tahapan analisis data yang dilakukan dalam penelitian ini ialah tahap praproses, eksplorasi data, tahap penggerombolan, tahap evaluasi, dan tahap pemetaan hasil

Tabel 2: Daftar peubah yang tersedia.

Kode peubah	Nama peubah	Gugus data
Kode	ID kabupaten atau kota	1-8
Nama	Nama kabupaten atau kota	1-8
JST	Jumlah satuan pendidikan	1-8
PES	Jumlah peserta	1-8
Mata pelajaran wajib		
IND	Bahasa Indonesia	1-8
ING	Bahasa Inggris	1-8
MAT	Matematika	1-8
Mata pelajaran pilihan jurusan IPA		
EKO	Ekonomi	5-8
FIS	Fisika	1-4
KIM	Kimia	1-4
BIO	Biologi	1-4
Mata pelajaran pilihan jurusan IPS		
EKO	Ekonomi	5-8
SOS	Sosiologi	5-8
GEO	Geografi	5-8

gerombol.

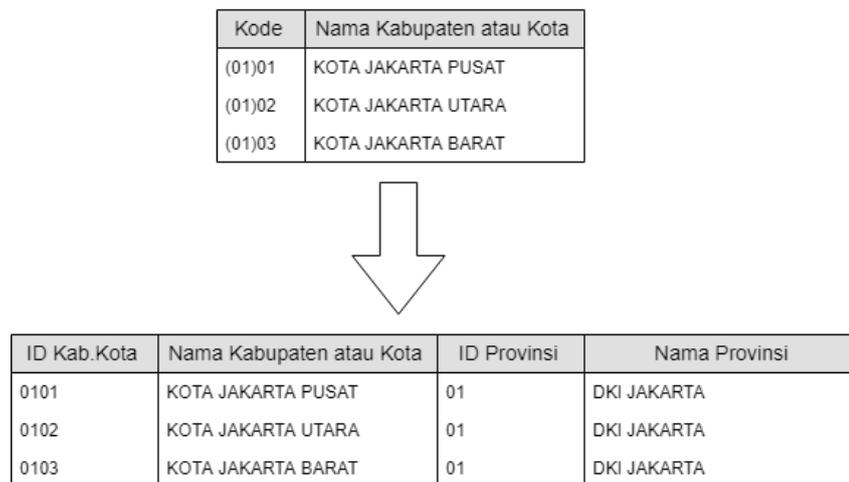
Tahap Praproses

Tahap praproses terdiri atas dua bagian, yaitu sebagai berikut:

1. Melakukan pemeriksaan pada data hilang. Data hilang tersebut akan dibuang apabila terdapat pada peubah jumlah satuan pendidikan atau jumlah peserta. Data tersebut dibuang karena kabupaten/kota pada gugus data tersebut belum memiliki peserta UN. Di samping itu, data hilang juga dapat terjadi pada peubah mata pelajaran pilihan, seperti Fisika, Kimia, Biologi, Ekonomi, Sosiologi, dan Geografi. Hal tersebut terjadi karena siswa dapat memilih salah satu dari tiga mata pelajaran pilihan sesuai jurusannya sejak tahun ajaran 2016/2017. Oleh karena itu, terdapat kabupaten/kota yang tidak ada satu pun siswa di dalamnya memilih salah satu mata pelajaran pilihan.
2. Membuat peubah baru provinsi yang berasal dari dua digit awal kode kabupaten atau kota dengan kode ID provinsi. Gambar 1 menunjukkan ilustrasi pembentukan peubah kode provinsi tersebut.
3. Membuat peubah baru rata-rata mata pelajaran pilihan dengan kode PIL.

Tahap Eksplorasi data

Tahap selanjutnya adalah melakukan analisis deskriptif dan eksplorasi data pada keenam peubah mata pelajaran UN, peubah rerata hasil UN (RAT), dan rata-rata mata pelajaran pilihan (PIL) untuk melihat karakteristik serta gambaran umum mengenai data. Eksplorasi data ini digunakan untuk melihat sebaran data dan mendeteksi adanya pencilan dalam data terutama pada peubah mata pelajaran UN. Selanjutnya, analisis



Gambar 1: Ilustrasi pembuatan kode provinsi

korelasi akan digunakan untuk melihat hubungan linier antarpeubah pada setiap gugus data.

Tahap Penggerombolan

Sebelum melakukan analisis gerombol, penentuan rentang ukuran gerombol optimum ditentukan menggunakan koefisien siluet (Rousseeuw, 1987), indeks Xie Beni (Xie & Beni, 1991), dan rasio simpangan Baku (R) (Kalkstein *et al.*, 1986). Analisis gerombol digunakan untuk melakukan pemetaan dengan mengelompokkan kabupaten/kota berdasarkan nilai atau hasil capaian UN semua siswa pada kabupaten/kota tersebut. Peubah yang akan digunakan pada tahap ini adalah mata pelajaran Bahasa Indonesia (IND), Bahasa Inggris (ING), Matematika (MAT), dan rata-rata hasil UN mata pelajaran pilihan (PIL). Analisis gerombol dilakukan untuk masing-masing gugus data sehingga terdapat delapan kali analisis. Algoritme optimasi iterasi yang digagas Bezdek (1981) adalah sebagai berikut:

1. Menentukan ukuran gerombol c yang akan digunakan.
2. Menentukan nilai parameter pembobot keanggotaan m . Hasil penelitian Klawonn & Hoppner (2006) mengatakan bahwa nilai m yang paling baik dan sering digunakan adalah $m = 2$.
3. Menentukan matriks partisi awal $\mathbf{U}^{(r=0)}$. Matriks \mathbf{U} merupakan matriks yang berisi nilai derajat keanggotaan berukuran $c \times n$. Matriks \mathbf{U} secara matematis didefinisikan sebagai berikut (Gan *et al.*, 2007)

$$\mathbf{U}_{c \times n}^r = \begin{pmatrix} \mu_{11} & \mu_{12} & \cdots & \mu_{1n} \\ \mu_{21} & \mu_{22} & \cdots & \mu_{2n} \\ \vdots & \vdots & \vdots & \ddots \\ \mu_{c1} & \mu_{c2} & \cdots & \mu_{cn} \end{pmatrix}$$

dengan r merupakan proses iterasi ke- r , $\mu_{ij} \in [0, 1]$, $\sum_{i=1}^c \mu_{ij} = 1$, $\sum_{j=1}^n \mu_{ij} > 0$, $1 \leq i \leq c$, $1 \leq j \leq n$. Nilai derajat keanggotaan untuk matriks partisi awal dibangkitkan secara acak.

4. Menghitung nilai pusat masing-masing gerombol v .
5. Menghitung jarak objek dengan pusat gerombol menggunakan jarak Euclid.
6. Menghitung kembali nilai derajat keanggotaan, untuk melakukan pembaruan matriks partisi U .
7. Menghitung nilai selisih mulak antara matriks partisi awal dengan matriks partisi iterasi setelahnya ($|U^{(r)} - U^{(r+1)}|$) kemudian membandingkannya dengan batas toleransi yang telah ditetapkan.
8. Mengulangi langkah 4 sampai 7. Jika $|U^{(r)} - U^{(r+1)}| \leq \varepsilon$ atau iterasi telah mencapai nilai maksimum yang telah ditetapkan maka proses dihentikan.

Tahap evaluasi Prosedur tahap penggerombolan akan dilakukan sebanyak 100 ulangan dengan menetapkan bibit (*set seed*) 1-100. Pengulangan dan penetapan bibit ini dilakukan karena terdapat proses pembangkitan bilangan acak untuk matriks partisi awal sehingga hasil gerombol yang didapatkan dapat berbeda-beda ketika menjalankan program. Hasil penggerombolan akan dievaluasi menggunakan perbandingan antara simpangan baku di dalam gerombol (S_w) dengan simpangan baku antargerombol (S_B). Nilai perbandingan atau rasio (R) terkecil merupakan kriteria yang terbaik.

Setelah mendapatkan penggerombolan terbaik, bibit (*seed*) yang akan digunakan untuk melakukan pemetaan ialah bibit yang memiliki nilai R minimum. Perangkat lunak yang akan digunakan pada penelitian ini adalah R Studio dengan versi 3.4. Paket yang digunakan: *e1071* untuk analisis gerombol, *tidyverse* untuk praproses data, *factoextra*, *ggplot* untuk visualiasi data, dan *sf* untuk membaca format data *shp* untuk pemetaan.

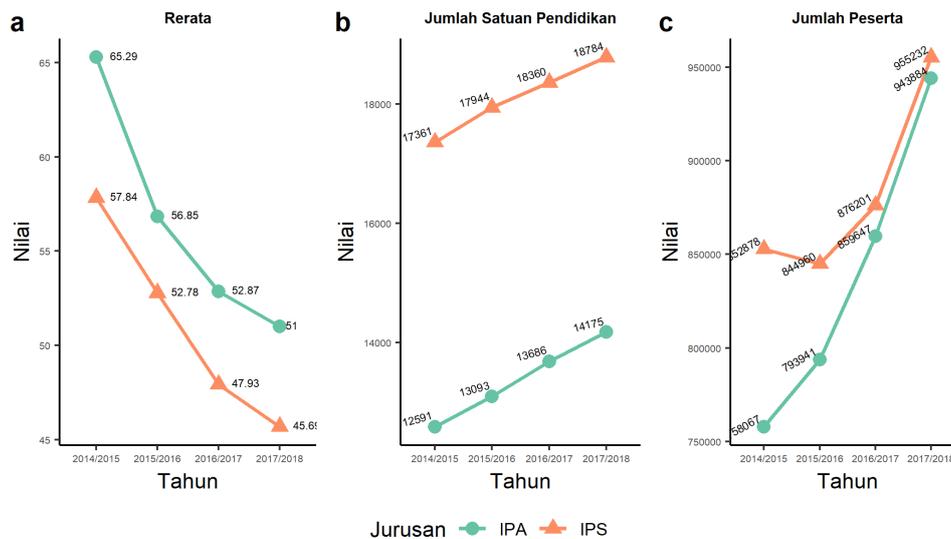
Tahap pemetaan Setelah mendapatkan hasil dari semua gugus data, solusi akhir gerombol akan ditentukan berdasarkan nilai rata-rata derajat keanggotaan dari tahun ajaran 2014/2015 sampai 2017/2018. Selanjutnya, gerombol yang didapatkan akan dibuat pemetaan (geovisualisasi).

3 Hasil dan Pembahasan

3.1 Deskripsi data

Analisis deskriptif dan eksplorasi data dapat memberikan gambaran mengenai data hasil UN jurusan IPA dan IPS pada tahun ajaran 2014/2015 sampai 2017/2018. Gambar 2 menunjukkan diagram garis rerata hasil UN, jumlah satuan pendidikan, dan jumlah peserta pada tahun ajaran 2014/2015 sampai 2017/2018 secara nasional. Rerata hasil UN jurusan IPA lebih tinggi dibandingkan dengan IPS, namun rerata hasil UN jurusan IPA mengalami penurunan dari 65.29 sampai 51, sedangkan IPS mengalami penurunan dari 57.64 sampai 48.69. Jumlah satuan pendidikan jurusan IPA sebanyak 14 175 pada tahun ajaran 2017/2018, sedangkan IPS sebanyak 18 784 pada tahun yang sama. Jurusan IPA memiliki satuan pendidikan yang lebih sedikit dibandingkan IPS setiap tahunnya. Jumlah peserta UN IPA pun lebih sedikit daripada IPS, namun sejak tahun ajaran 2016/2017 jumlah peserta UN IPA meningkat relatif hampir sama dengan IPS.

Nilai tengah mata pelajaran Biologi, Fisika, dan Kimia mengalami penurunan dari tahun ajaran 2014/2015 sampai 2017/2018 (Gambar 3). Kabupaten atau kota yang



Gambar 2: (a) Rerata UN secara nasional, (b) jumlah satuan pendidikan secara nasional, (c) jumlah peserta secara nasional

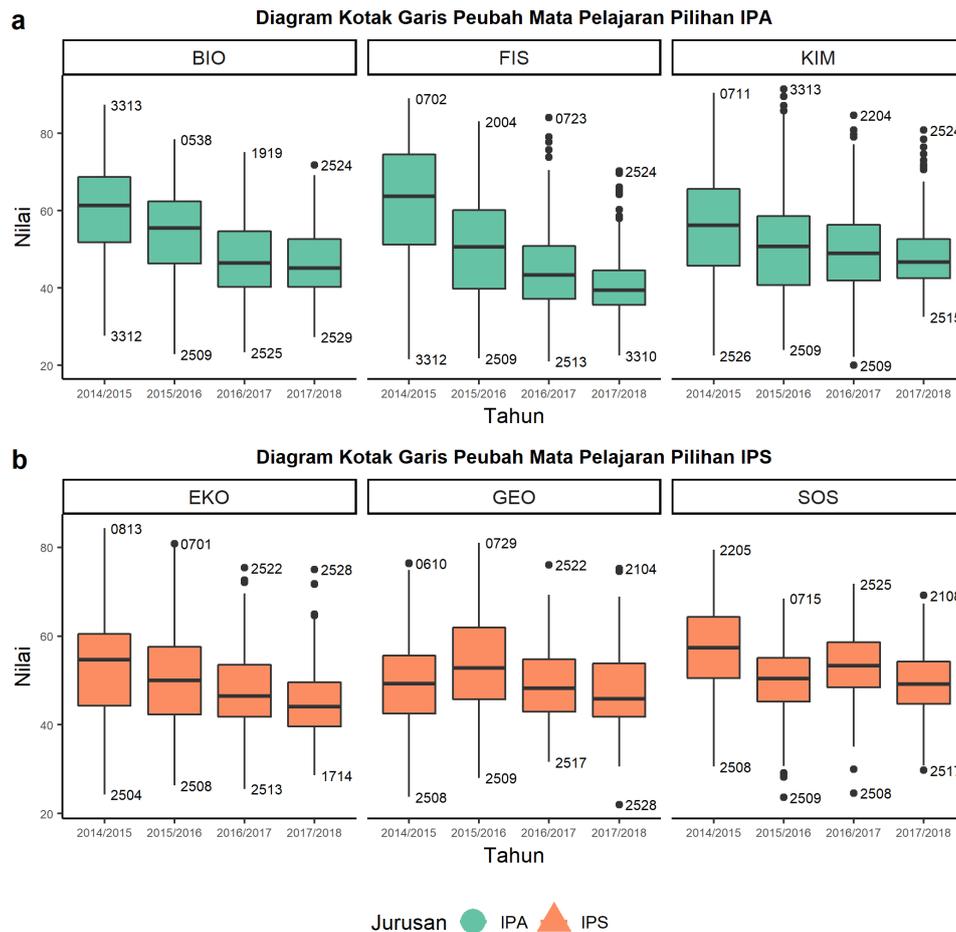
memiliki nilai terendah di ketiga mata pelajaran tersebut banyak di antaranya merupakan kabupaten yang berada di Provinsi Papua. Tidak seperti jurusan IPA, Nilai tengah mata pelajaran Ekonomi, Sosiologi, dan Geografi mengalami fluktuasi. Nilai rentang keenam mata pelajaran tersebut mengecil sejak tahun ajaran 2016/2017. Hal ini terjadi karena siswa memilih satu mata pelajaran UN. Kebijakan ini secara eksploratif mengurangi ukuran penyebaran nilai.

Jurusan IPA memiliki nilai tengah yang lebih tinggi dibandingkan IPS untuk ketiga mata pelajaran wajib dan rata-rata mata pelajaran pilihan Gambar 4. Namun begitu, jurusan IPS pada tahun ajaran 2016/2017 dan 2017/2018 IPS memiliki nilai tengah rata-rata mata pelajaran pilihan yang lebih tinggi, yaitu saat kebijakan pemilihan mata pelajaran UN. Nilai semua mata pelajaran baik jurusan IPA dan IPS terus mengalami penurunan, akan tetapi nilai rentang atau batas maksimum dan minimum juga mengalami penurunan. Artinya, penyebaran mutu pendidikan secara eksploratif lebih baik dibandingkan sebelumnya.

Beberapa gugus data memiliki sebaran yang cukup simetrik pada beberapa peubah. Contohnya ialah pada peubah PIL dan Bahasa Inggris. Walaupun begitu, Bahasa Inggris memiliki pencilan atas pada jurusan IPS tahun ajaran 2016/2017 dan 2017/2018. Bahasa Indonesia terlihat konsisten memiliki pencilan bawah, kabupaten/kota yang termasuk dalam pencilan bawah tersebut harus diteliti lebih lanjut karena sebaran data cukup baik. Hal sebaliknya terjadi pada Bahasa Indonesia, Matematika memiliki banyak pencilan atas, artinya sebaran data berkumpul pada nilai yang rendah.

3.2 Analisis gerombol

Penggerombolan dilakukan sebanyak seratus kali dengan nilai bibit (*seed*) satu sampai seratus pada perangkat lunak R. Hal ini dilakukan karena matriks partisi awal dibangkitkan secara acak. Penentuan bibit tersebut agar mendapatkan hasil gerombol yang paling optimum dan konsisten. Bibit yang akan dipilih ialah bibit yang memiliki nilai rasio simpangan baku terkecil. Gambar 5 menunjukkan nilai rata-rata dan rentang dalam diagram kotak garis untuk koefisien siluet, indeks validitas Xie Beni, dan rasio



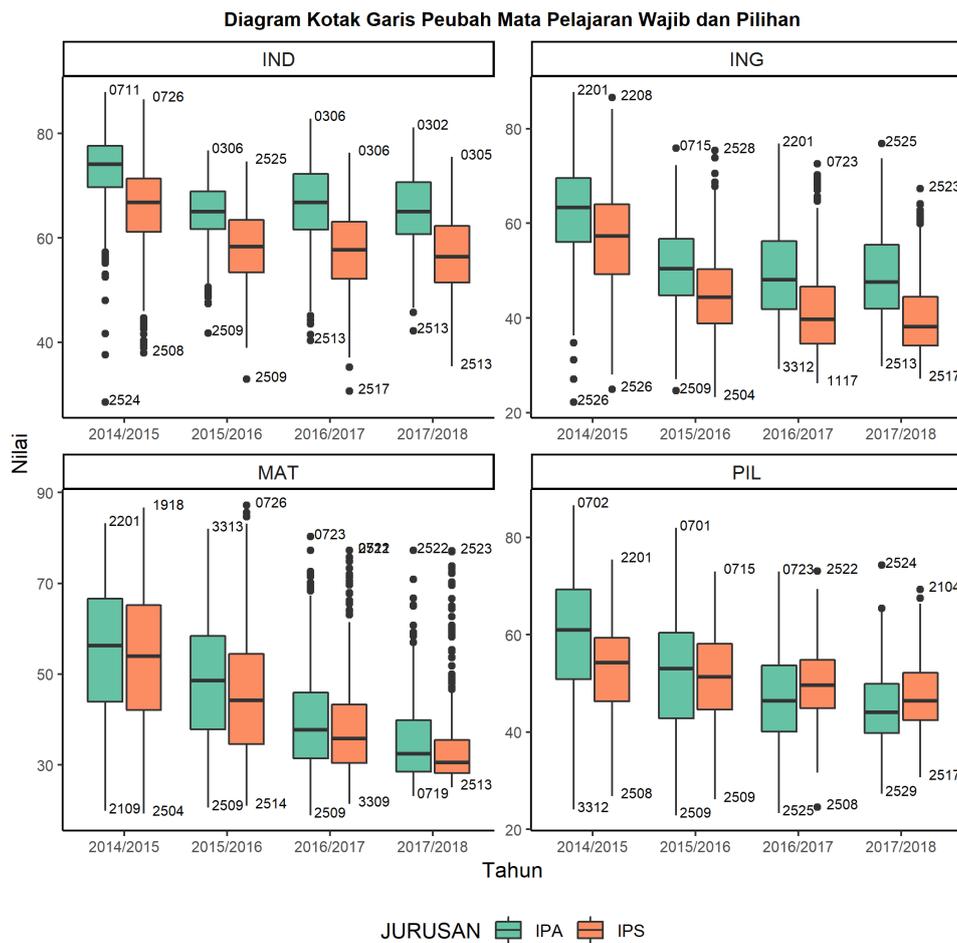
Gambar 3: Sebaran nilai mata pelajaran pilihan

simpangan baku dari 100 pengulangan. Hal tersebut dilakukan untuk menentukan ukuran gerombol yang akan digunakan dalam analisis.

Informasi yang diperoleh dari gambar di atas adalah nilai tengah baik rata-rata (bentuk bintang) maupun median fungsi objektif dan rasio simpangan baku mengalami penurunan ketika ukuran gerombol bertambah. Fungsi objektif dan rasio simpangan baku mengukur selisih jarak objek ke pusat gerombol. Oleh karena itu, bertambahnya ukuran gerombol nilai akan membuat kedua indeks validitas tersebut berkurang. Ukuran dua gerombol pada fungsi objektif dan rasio simpangan baku memiliki nilai rata-rata dan median yang sama sehingga sebaran kedua indeks tersebut berbentuk genta. Koefisien siluet dan Xie Beni pada ukuran dua gerombol adalah yang terbaik karena koefisien siluet maksimum dan Xie Beni minimum. Ukuran gerombol terbaik yang akan dipilih ialah dua gerombol.

Pengerombolan Dua ukuran gerombol

Gerombol yang dihasilkan berupa bilangan bulat positif, tetapi bilangan bulat positif tersebut tidak sama untuk setiap analisisnya. Sebagai contoh, gugus data 1 pada gerombol 1 tidak sama dengan gerombol 1 pada gugus data 2. Oleh karena itu, gerombol akan diurutkan terlebih dahulu sehingga menjadi skala ordinal dengan label huruf alfabet untuk memudahkan pelabelan. Nilai pusat gerombol untuk masing-masing gugus data dapat dilihat pada Gambar 6

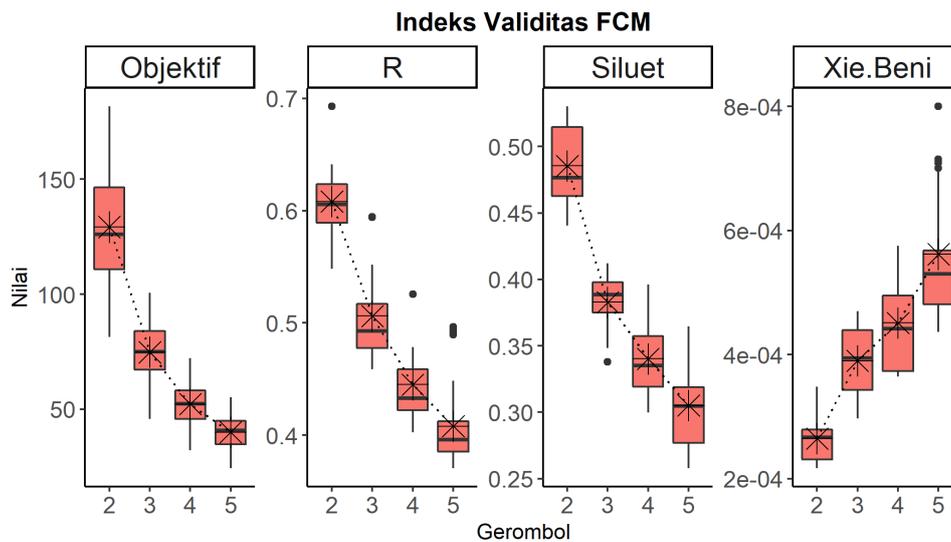


Gambar 4: Sebaran nilai mata pelajaran

Gambar 6 memberikan informasi bahwa gerombol A konsisten memiliki nilai rata-rata yang lebih tinggi dibandingkan gerombol B pada semua peubah dan gugus data. Oleh sebab itu, gerombol A dapat dicirikan dengan hasil UN yang baik, sedangkan gerombol B memiliki hasil UN yang kurang baik. Persentase atau proporsi kabupaten atau kota dalam tiap gerombol dapat dilihat pada Gambar 7

Salah satu kelebihan metode gerombol samar adalah derajat keanggotaan. Solusi akhir gerombol untuk tiap kabupaten/kota dapat ditentukan menggunakan nilai rata-rata dari derajat keanggotaan pada tahun ajaran 2014/2015 sampai 2017/2018. Hal ini dilakukan untuk mengevaluasi mutu pendidikan dalam empat tahun terakhir. Agar dapat mendapatkan informasi dari satuan amatan yang lebih luas, proporsi anggota gerombol dapat diperbesar menjadi provinsi dan wilayah pulau di Indonesia. Proporsi gerombol antarwilayah ini dapat dilihat pada Gambar 8

Gambar 8 menunjukkan bahwa Pulau Jawa memiliki proporsi gerombol A yang terbesar dengan angka 79.68% untuk jurusan IPA dan 78.81% untuk IPS. Pulau Kalimantan, Maluku, Sumatera, Papua, dan Maluku memiliki proporsi 37.5%, 38.1%, 39.61%, dan 41.46%, Keempat pulau tersebut memiliki proporsi gerombol yang merata untuk jurusan IPA. Nusa Tenggara merupakan pulau yang memiliki proporsi kabupaten atau kota dengan gerombol A yang paling sedikit, yaitu 24.39%. Berbeda dengan IPA, proporsi jurusan IPS tersebar tidak merata. Nusa Tenggara masih merupakan pulau dengan gerombol A paling sedikit, yaitu hanya 18.91%. Selanjutnya adalah proporsi



Gambar 5: Indeks validitas seratus kali penggerombolan

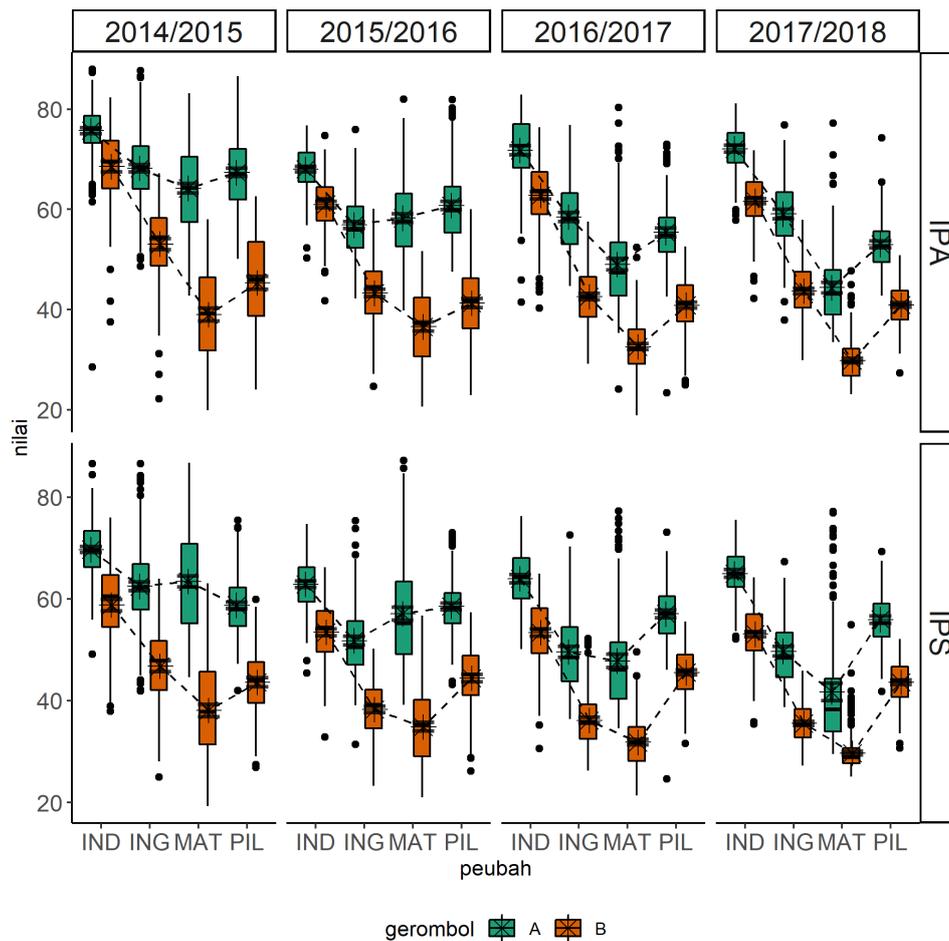
antarprovinsi yang dapat dilihat pada Gambar 9

Provinsi DKI Jakarta, DI Yogyakarta, dan Bali pada jurusan IPA memiliki proporsi gerombol A 100%. Semua kabupaten/kota dalam tiga provinsi tersebut telah tergolong baik. Sulawesi Tengah, Sulawesi Barat, Nusa Tenggara Timur, dan Kalimantan Utara merupakan provinsi yang sama sekali tidak memiliki kabupaten atau kota yang termasuk dalam gerombol A. Jurusan IPS memiliki proporsi provinsi gerombol A 100% hanya DKI Jakarta dan DI Yogyakarta. Bali hanya memiliki 2 dari 3 kabupaten atau kota yang termasuk dalam gerombol A. Sulawesi Barat, Kalimantan Utara, dan Bengkulu merupakan tiga provinsi yang tidak memiliki sama sekali kabupaten atau kota yang termasuk ke dalam gerombol A. Nusa Tenggara Timur hanya memiliki 4% saja yang termasuk dalam gerombol A. Berdasarkan penejelas sebelumnya, kabupaten atau kota yang termasuk dalam gerombol B perlu dilakukan pemeriksaan dan pengevaluasian lebih lanjut. Provinsi-provinsi yang memiliki proporsi gerombol B lebih banyak daripada gerombol A harus melakukan perbaikan. Daerah-daerah penyebaran gerombol A dan B dapat dilihat melalui peta mutu pendidikan pada Gambar 10

Proporsi kabupaten/kota yang termasuk gerombol A yang dikategorikan baik mengalami penurunan sejak tahun ajaran 2014/2015 sampai 2017/2018 baik pada jurusan IPA maupun IPS. Pembagian menjadi 2 gerombol lebih ketat karena hanya mengategorikan baik dan kurang baik. Jika dilihat melalui peta mutu pendidikan akhir dari Gambar 10, gerombol A terlihat dominan di Pulau Jawa, terutama Pulau Jawa bagian Timur. Pulau Sumatera terlihat dominan di Sumatera Utara, Pulau Kalimantan di Kalimantan Timur, Pulau Sulawesi di Sulawesi Selatan. Kabupaten atau kota yang masuk dalam gerombol A tersebut masih cenderung berada di kota-kota besar.

4 Simpulan

Ukuran gerombol yang disarankan menggunakan rasio simpangan baku, koefisien siluet, dan indeks Xie Beni adalah dua ukuran gerombol, yaitu A dan B. Gerombol A memiliki nilai rata-rata lebih tinggi daripada gerombol B pada mata pelajaran wajib dan pilihan. Gerombol A terus mengalami penurunan dari tahun ajaran 2014/2015. Artinya,

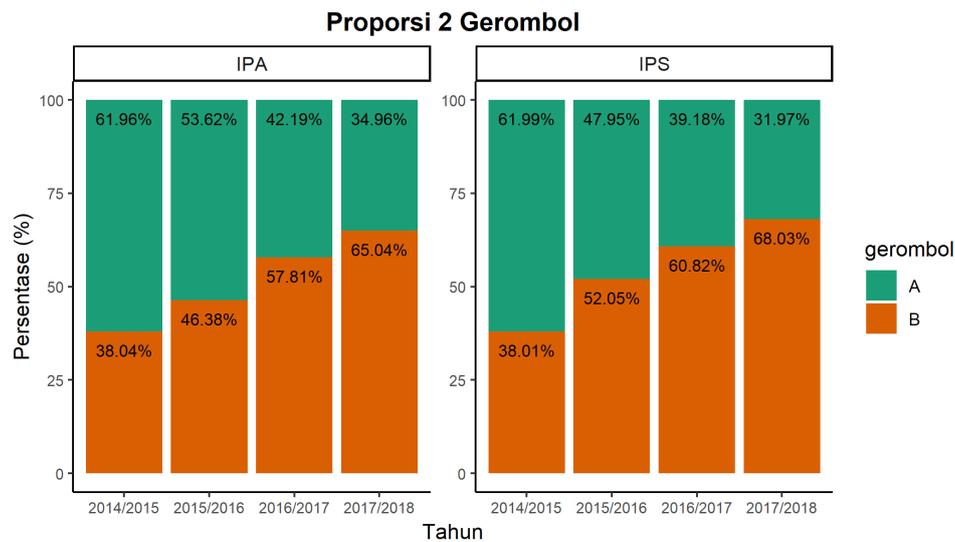


Gambar 6: Ilustrasi sebaran gerombol

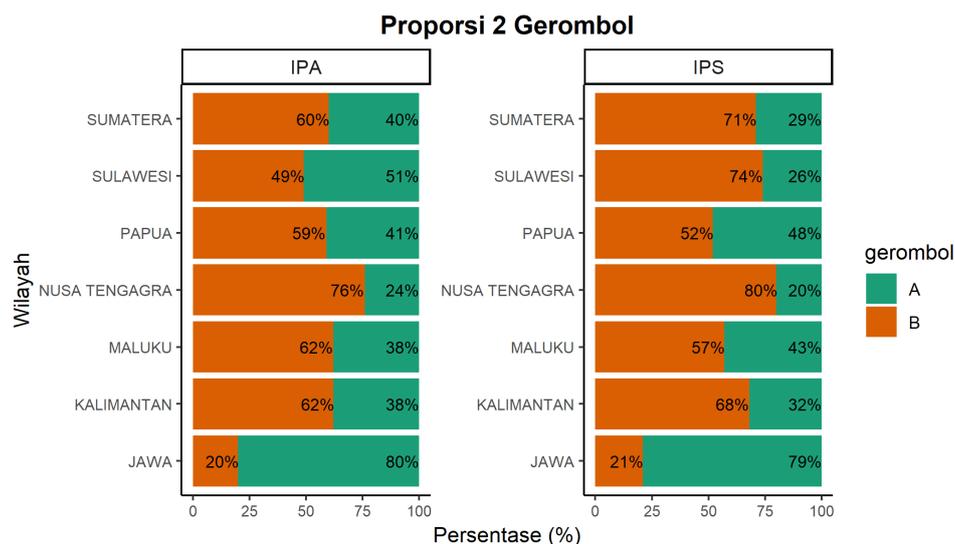
kabupaten/kota yang termasuk dalam kategori baik semakin berkurang. Hal ini juga diperkuat dengan melihat penyebaran gerombol pada peta mutu pendidikan berdasarkan hasil UN di Indonesia. Proporsi kabupaten/kota yang termasuk dalam gerombol A didominasi oleh Pulau Jawa. Namun demikian, masih terdapat wilayah di Pulau Jawa yang memiliki sedikit proporsi kabupaten/kotanya termasuk dalam gerombol A, yaitu Jawa Barat dan Banten. DKI Jakarta dan Yogyakarta merupakan dua provinsi yang memiliki seratus persen gerombol A baik jurusan IPA maupun IPS. Nusa Tenggara Timur, Sulawesi Tengah, Sulawesi Barat, dan Kalimantan Utara merupakan provinsi yang sama sekali tidak memiliki kabupaten/kota yang termasuk dalam gerombol A. Hasil analisis menunjukkan penyebaran mutu pendidikan di Indonesia masih belum merata dan terfokus pada Pulau Jawa, terutama di kota-kota besar.

Daftar Pustaka

- Agustin, T. & Djuraidah, A. (2010). Penggerombolan daerah tertinggal di Indonesia dengan fuzzy k-rataan. *Forum Statistika dan Komputasi*, 15(1): 22–27.
- Bezdek, J. C. (1981). *Pattern Recognition with Fuzzy Objective Function Algorithm*. New York(US): Plenum Press.



Gambar 7: Persentase anggota gerombol tiap gugus data



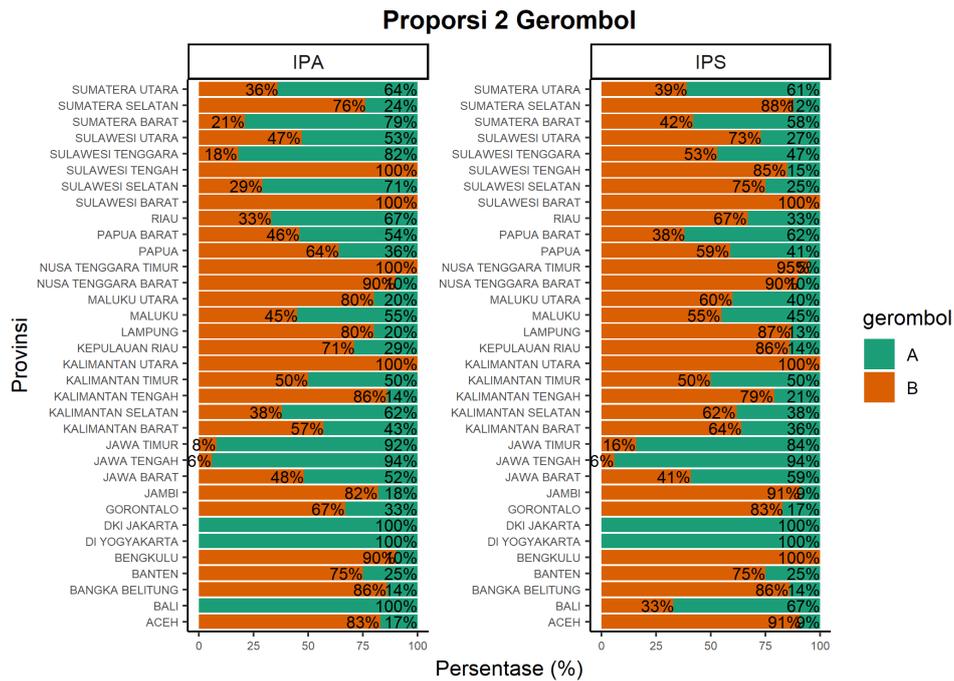
Gambar 8: Persentase pembagian solusi akhir anggota gerombol per wilayah

Gan, G., Ma, C., & Wu, J. (2007). *Data Clustering: Theory, Algorithms, and Applications*. Virginia(US): American Statistical Association and the Society for Industrial and Applied Mathematics.

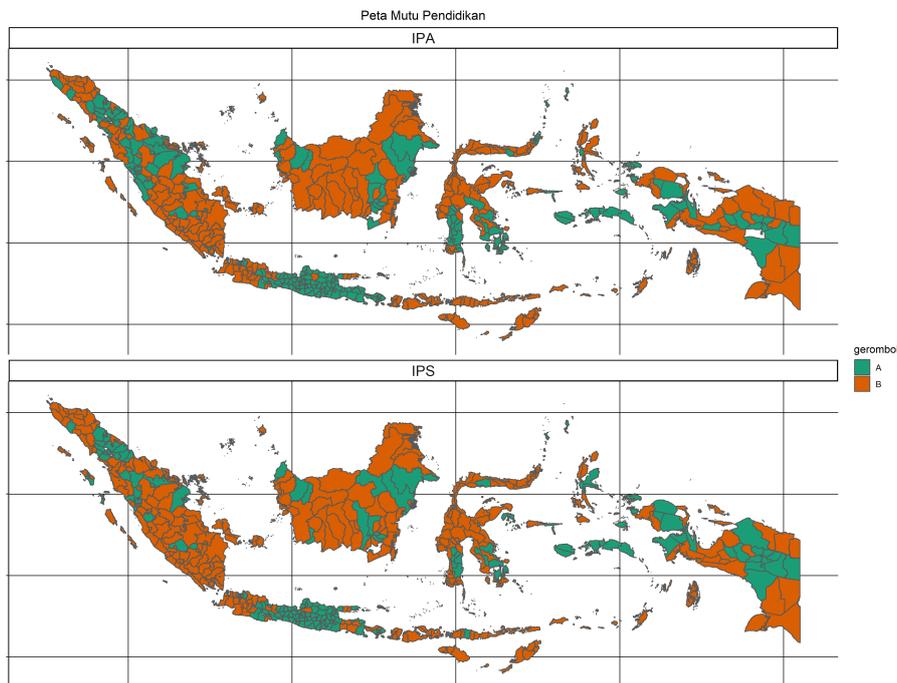
Intendia, F. (2012). *Penentuan indikator standar pelayanan minimal mutu sekolah dengan CART dan regresi logistik [skripsi]*. Institut Pertanian Bogor, Bogor(ID).

Kalkstein, L., Tan, G., & Skindolv, J. (1986). An evaluation of three procedures for use in synoptic climatological classification. *Journal of climate and applied meteorology*, 26(2): 717–730.

Khairunnisa (2017). *Kajian perbandingan metode k-means dan fuzzy c-means untuk menggerombolkan kecamatan di Provinsi Jawa Barat berdasarkan data kemiskinan [tesis]*. Institut Pertanian Bogor, Bogor(ID).



Gambar 9: Persentase pembagian solusi akhir anggota gerombol per provinsi



Gambar 10: Peta mutu pendidikan

Klawonn, F. & Hoppner, F. (2006). A comparison of clustering techniques in aspect mining. *Studia Univ. Babeş-Bolyai, Informatica*, 51(1): 69–78.

Mongi, C. E. (2014). *Penggerombolan dan pemetaan kabupaten/kota di Provinsi Jawa Barat berdasarkan nilai Ujian Nasional SMA dan akreditasi sekolah [tesis]*. Institut Pertanian Bogor, Bogor(ID).

- Nugroho, S. (2016). *Kajian Spasial Durbin Model (SDM) untuk menelaah faktor-faktor yang mempengaruhi nilai Ujian Nasional (UN) di madrasah aliyah negeri (MAN) [tesis]*. Institut Pertanian Bogor, Bogor(ID).
- Pangestu, W. A. (2018). *Penerapan Regresi Peubah Ganda untuk melihat hubungan antara akreditasi dan Ujian Nasional tingkat SMA [skripsi]*. Institut Pertanian Bogor, Bogor(ID).
- Putri, D. H. (2017). *Kajian kekekaran metode penggerombola fuzzy k-rataan pada data yang mengandung noise [skripsi]*. Institut Pertanian Bogor, Bogor(ID).
- Ross, T. J. (2010). *Fuzzy Logic with Engineering Application. Third Edition*. New York(US): John Wiley and Sons.
- Rousseeuw, P. (1987). Silhouettes: a graphical aid to interpretation and validation of cluster analysis. *Journal of Computational and Mathematics*, 20: 53–66.
- Serban, G. & Moldovan, G. S. (2006). A comparison of clustering techniques in aspect mining. *Studia Univ. Babeş-Bolyai, Informatica*, 51(1): 69–78.
- Xie, X. & Beni, G. (1991). A validity measure for fuzzy clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(8): 841–847.
- Yang, M. S. (1993). A survey of fuzzy clustering. *Mathl. Comput. Modelling*, 18(1): 1–16.