

Klasterisasi Mutu Pendidikan SMA di Indonesia dan Analisis Pola Hubungan Antar Standar Nasional Pendidikan pada Masing-Masing Klaster Menggunakan SEM-PLS

Arif Khoirul Anam, Agnes Tuti Rumiati dan Vita Ratnasari
Departemen Statistika, Institut Teknologi Sepuluh Nopember (ITS)
email: agnes_tuti@statistika.its.ac.id

Abstrak—Mutu pendidikan yang baik akan mampu meningkatkan kualitas sumber daya manusia yang ada. Pada penelitian ini akan dilakukan pengelompokan mutu pendidikan jenjang SMA di Indonesia menggunakan metode Fuzzy C-Means dan dilanjutkan dengan analisis pola hubungan antar standar nasional pendidikan menggunakan metode Structural Equation Modelling – Partial Least Square. Hasil klaster terbaik membagi mutu pendidikan menjadi empat kelompok. Sebagian besar Sekolah Menengah Atas di Indonesia sudah tergolong cukup baik dan perbandingan status sekolah negeri dan swasta pada masing-masing klaster tidak begitu signifikan. Mutu pendidikan di Indonesia terbilang belum merata karena masih ada beberapa provinsi yang memiliki proporsi sekolah cukup banyak pada klaster terendah. Pada analisis SEM terdapat beberapa pola hubungan yang tidak signifikan pada masing-masing klaster. Adanya pola hubungan yang tidak signifikan bukan berarti tidak bagus, namun kemungkinan tanpa adanya pengaruh dari variabel lain, variabel tersebut sudah baik.

Kata Kunci—Fuzzy C-Means, Klaster, Sekolah Menengah Atas, Standar Nasional Pendidikan, *Structural Equation Modelling*

I. PENDAHULUAN

SETIAP tahun pendidikan di Indonesia terus mengalami perkembangan menuju arah yang lebih baik. Selama periode 2010 hingga 2017, Harapan Lama Sekolah di Indonesia telah meningkat sebesar 1,56 tahun, sementara Rata-Rata Lama Sekolah bertambah sebesar 0,64 tahun (BPS, 2018). Namun, pada dasarnya bukan hanya faktor tersebut saja yang menjadi acuan majunya pendidikan di Indonesia. Kemajuan dalam bidang pendidikan di Indonesia belum tersebar secara merata. Banyak daerah terpencil yang masih belum memiliki pendidikan layak, terutama wilayah Indonesia bagian timur. Rata-rata angka buta huruf di kawasan tersebut masih diatas 5%, dengan persentase tertinggi penduduk yang buta huruf, yaitu Provinsi Papua sebesar 36,31%, diikuti oleh Provinsi Nusa Tenggara Barat (16,48%) dan Provinsi Sulawesi Barat (10,33%). Sementara provinsi di Indonesia Timur lain yang masih memiliki persentase buta huruf diatas 5%, yaitu Nusa Tenggara Timur (10,13%), Gorontalo (5,05%) Sulawesi Tenggara (6,76%) dan Papua Barat (7,35%) [1]

Jenjang SMA menjadi tahapan yang sangat vital bagi seorang siswa dalam menentukan masa depannya. Oleh karena itu, para siswa harus mendapatkan pelayanan yang

baik agar bisa mencapai standar yang diinginkan. Jumlah sekolah jenjang SMA yang terpetakan pada tahun 2018 sebanyak 12.682 sekolah atau sekitar 85,90% dari sebanyak 14.762 sekolah. Sedangkan sebanyak 2.082 sekolah atau 14,10% belum terpetakan karena data tidak tersedia. Standar kompetensi lulusan SMA tahun 2018 sebesar 6,39, yaitu masuk pada tahapan menuju SNP level 4 dan terbilang cukup baik. Capaian SNP yang kurang baik, yaitu pada standar pendidik dan tenaga kependidikan sebesar 3,54 (menuju SNP level 2) serta standar sarana dan prasarana sebesar 4,39 (menuju SNP level 3)[2].

Klasterisasi mutu pendidikan perlu dilakukan untuk mengetahui sekolah mana saja yang perlu mendapat perhatian lebih. Selain itu, evaluasi Standar Nasional Pendidikan juga perlu dilakukan untuk mengetahui apakah sistem penilaian yang dilakukan sudah tepat atau belum. Penelitian sebelumnya dilakukan oleh Karti (2013), yaitu membandingkan metode Fuzzy C-Means dan α -C-Means untuk mengelompokkan kabupaten/kota di Jawa Timur berdasarkan indikator pencapaian T3 untuk Sekolah Menengah Kejuruan (SMK) dan diperoleh model terbaik menggunakan Fuzzy C-Means [3]. Fatimah (2018) melakukan penelitian tentang pengelompokan mutu pendidikan jenjang Sekolah Menengah Pertama (SMP) di Jawa Timur berdasarkan Standar Nasional Pendidikan (SNP) dengan membandingkan metode Hierarki, K-Means dan Fuzzy C-Means. Hasilnya metode pengelompokan Fuzzy C-Means menjadi yang terbaik [4]. Amalinda (2013) meneliti tentang pengaruh pengelolaan sekolah terhadap kompetensi lulusan berdasarkan SNP di Sekolah Menengah Pertama (SMP) Surabaya menggunakan metode Structural Equation Modelling – Partial Least Square. Hasilnya diketahui bahwa antar Standar Nasional Pendidikan saling berpengaruh terhadap standar pengelolaan [5]. Pada penelitian ini akan dibagi menjadi dua kasus, yaitu klasterisasi mutu pendidikan SMA di Indonesia menggunakan metode Fuzzy C-Means dan analisis pola hubungan antar Standar Nasional Pendidikan jenjang SMA di Indonesia pada masing-masing cluster menggunakan Metode Structural Equation Modelling – Partial Least Square.

Berdasarkan latar belakang yang telah diuraikan sebelumnya, maka didapatkan beberapa rumusan masalah, yaitu bagaimana karakteristik Standar Nasional Pendidikan SMA di Indonesia, bagaimana klasterisasi mutu pendidikan

SMA di Indonesia dan bagaimana pola hubungan antar Standar Nasional Pendidikan SMA di Indonesia. Klasterisasi mutu pendidikan SMA di Indonesia menggunakan metode Fuzzy C-Means. Sementara itu, pola hubungan antar Standar Nasional Pendidikan jenjang SMA di Indonesia menggunakan Metode Structural Equation Modelling – Partial Least Square.

Penelitian ini diharapkan dapat digunakan sebagai informasi mengenai kelompok optimum yang terbentuk untuk klasterisasi mutu pendidikan dan pola hubungan antar Standar Nasional Pendidikan Sekolah Menengah Atas (SMA) di Indonesia. Selain itu juga menjadi informasi pihak terkait untuk pertimbangan pengambilan kebijakan dalam meningkatkan mutu pendidikan di Indonesia.

II. TINJAUAN PUSTAKA

A. Standar Nasional Pendidikan

Standar Nasional Pendidikan adalah kriteria minimal tentang sistem pendidikan di seluruh wilayah hukum Negara Kesatuan Republik Indonesia. Standar Nasional Pendidikan terdiri dari kompetensi lulusan, isi, proses, pendidikan dan tenaga kependidikan, sarana dan prasarana, pengelolaan, pembiayaan pendidikan serta penilaian pendidikan. Standar Nasional Pendidikan berfungsi sebagai dasar dalam perencanaan, pelaksanaan, dan pengawasan pendidikan dalam rangka mewujudkan pendidikan nasional yang bermutu dan bertujuan menjamin mutu pendidikan nasional dalam rangka mencerdaskan kehidupan bangsa dan membentuk watak serta peradaban bangsa yang bermartabat. Standar Nasional Pendidikan disempurnakan secara terencana, terarah, dan berkelanjutan sesuai dengan tuntutan perubahan kehidupan lokal, nasional, dan global oleh Badan Standar Nasional Pendidikan (BSNP). Standar Nasional Pendidikan diturunkan ke dalam indikator hingga subindikator untuk dapat dibangun ukuran pencapaian mutunya. Delapan standar yang diatur dalam peraturan diturunkan menjadi 29 indikator dan 103 subindikator. Penurunan ke dalam indikator dan subindikator dilakukan berdasarkan kajian terhadap peraturan yang berlaku dan akreditasi [2].

Setiap Standar Nasional Pendidikan memiliki beberapa indikator dan sub indikator. Nilai komposit berkelanjutan mulai dari level pertanyaan hingga keseluruhan standar digunakan untuk menghitung nilai dua sumber data pencapaian mutu, yaitu data primer dan data sekunder. Perhitungan nilai komposit sub indikator menggunakan persamaan 2.1.

$$\sum^n (\alpha_{t,u,v,(n)} X_{t,u,v,(n)}) + \sum^m (\alpha_{t,u,v,(m)} Y_{t,u,v,(m)}) \quad (2.1)$$

dimana,

$X_{t,u,v,(n)}$: nilai data primer ke- n untuk sub indikator ke- v pada indikator ke- u pada standar ke- t

$Y_{t,u,v,(m)}$: nilai data sekunder ke- m untuk sub indikator ke- v pada indikator ke- u pada standar ke- t

$\alpha_{t,u,v}$: bobot nilai data untuk sub indikator ke- v pada indikator ke- u pada standar ke- t

Capaian indikator pada umumnya merupakan rata-rata aritmatik dari sub indikator dalam indikator yang sama. Hal tersebut berlaku pada sebagian besar indikator. Secara

matematis, perhitungan nilai komposit capaian indikator dituliskan pada persamaan 2.2.

$$T_{t,u} = \sum^v (\beta_{t,u,(v)} U_{t,u,(v)}) \quad (2.2)$$

dimana,

$T_{t,u}$: capaian indikator ke- u pada standar ke- t

$\beta_{t,u,(v)}$: bobot nilai sub indikator ke- v pada indikator ke- u pada standar ke- t

$U_{t,u,v}$: capaian sub indikator ke- v pada indikator ke- u pada standar ke- t

Nilai akhir capaian standar dihitung menggunakan persamaan 2.3.

$$S_t = \sum^u Y_{t,(u)} T_{t,(u)} \quad (2.3)$$

dimana,

S_t : capaian standar ke- t

$Y_{t,(u)}$: bobot nilai indikator ke- u pada standar ke- t dengan $v = \{1,2, \dots, k\}$, $u = \{1,2, \dots, l\}$, dan $t = \{1,2, \dots, 8\}$.

Angka capaian sub indikator hingga standar disajikan pada skala angka 0 – 7. Angka capaian SNP baik pada level sub indikator hingga standar dikategorikan sesuai tahapan pencapaian SNP. Penentuan rentang batas atas dan bawah mengadopsi metode fungsi *mean*- standar deviasi dan metode manual menyesuaikan konsep tahapan pencapaian SNP. Sekolah dikatakan sudah mencapai Standar Nasional Pendidikan ketika memiliki nilai 6,67 hingga 7,00.

B. Imputasi Regresi Linier

Imputasi merupakan metode yang digunakan untuk memprediksi data yang hilang pada suatu kumpulan data. Terdapat beberapa metode imputasi data yang sering digunakan, yaitu menggunakan rata-rata, median, modus, regresi, dan lain sebagainya. Analisis regresi adalah analisis yang digunakan untuk mengetahui dan mempelajari suatu model hubungan antara variabel respon (dependen) dan variabel prediktor (independen). Secara umum, bentuk persamaan linier sederhana regresi adalah $Y=a+bX$, dimana Y merupakan variabel dependen dan X merupakan variabel independen. Sementara itu, a adalah intersep titik potong kurva sumbu Y dan b adalah slope kurva linier [6]. Pada penelitian ini, yang bertindak sebagai variabel dependen adalah variabel yang memiliki missing value sedangkan yang menjadi variabel independen adalah variabel lainnya yang memiliki nilai atau tidak missing.

C. Principal Component Analysis

Analisis komponen utama bertujuan untuk mereduksi data dan menginterpretasikannya, meskipun dari n buah variabel asal dapat diturunkan n buah komponen utama untuk menerangkan keragaman total sistem, namun sering kali keragaman total itu dapat diterangkan secara memuaskan oleh sejumlah kecil komponen utama [7]. Komponen utama didefinisikan sebagai kombinasi linier dari n variabel asal yang dinyatakan dalam bentuk persamaan matriks sebagai berikut:

$$C_{n \times 1} = A_{n \times n} X_{n \times 1} \quad (2.4)$$

Dimana A merupakan matriks konstanta, C dan X adalah matriks variabel baru dan matriks variabel asal. Persamaan (2.4) dapat dinyatakan sebagai berikut:

$$\begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_n \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} \quad (2.5)$$

Ada dua cara yang digunakan dalam analisis komponen utama, yaitu dengan menggunakan matriks ragam peragam dan matriks korelasi.

D. Fuzzy C-Means

XGBoost *Fuzzy Cluster Means* (FCM) atau lebih dikenal dengan *Fuzzy C-Means* merupakan salah satu metode *non-hierarchical clustering* yang membuat partisi optimasi. Pada metode partisi FCM, sebuah objek dapat dikategorikan kedalam sebuah *cluster*, kemudian dikeluarkan lagi karena memiliki kedekatan dengan *cluster* lainnya. Setiap titik dalam suatu *cluster* menggunakan metode *Fuzzy C-Means* ditentukan oleh derajat keanggotaan [8]. Konsep dasar *Fuzzy C-Means* adalah sebagai berikut:

1. Menentukan pusat *cluster* sebagai penanda lokasi *mean* untuk masing-masing *cluster*.
2. Memperbaiki pusat *cluster* dan derajat keanggotaan setiap titik data secara berulang hingga pusat *cluster* menuju lokasi yang tepat. Perulangan berdasarkan minimalisasi fungsi objektif yang menggambarkan jarak dari titik data terhadap pusat *cluster* yang terboboti oleh derajat keanggotaan titik data tersebut.

Fungsi keanggotaan (u_{ij}) diartikan sebagai seberapa besar kemungkinan suatu data bisa menjadi anggota dalam sebuah *cluster*. Pada setiap *cluster* yang terbentuk secara acak, masing-masing data mempunyai nilai keanggotaan sebagai berikut

$$U^{(t)} = \begin{bmatrix} u_{11} & u_{12} & \dots & u_{1n} \\ u_{21} & u_{22} & \dots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ u_{c1} & u_{c2} & \dots & u_{cn} \end{bmatrix}, u_{ij} \in [0,1] \quad (2.6)$$

dimana,

- u_{ij} : nilai keanggotaan pada *cluster* ke- i ($i = 1,2, \dots, c$) dan objek ke- j ($j = 1,2, \dots, n$) dengan $\sum_{i=1}^c u_{ij} = 1, j = 1,2, \dots, n$ dan $0 < \sum_{i=1}^n u_{ij} < n, i = 1,2, \dots, c$
- $U^{(t)}$: matriks fungsi keanggotaan pada iterasi ke- t

Besarnya pusat *cluster* ke- i yang akan menandai lokasi rata-rata untuk masing-masing *cluster* dapat diperoleh dengan persamaan 2.7.

$$v_i = \frac{\sum_{j=1}^n (u_{ij})^m x_j}{\sum_{j=1}^n (u_{ij})^m} \quad (2.7)$$

dengan:

- v_i : pusat *cluster* ke- i , dimana $i = 1,2, \dots, c$
- m : pangkat pembobot

Besarnya pangkat pembobot dapat mempengaruhi fungsi u_{ij} dengan nilai $m > 1$. Pada kondisi awal, pusat *cluster* belum akurat dikarenakan masih terbentuk secara acak. Perlu dilakukan perulangan perbaikan nilai pusat *cluster* dan nilai keanggotaan untuk mendapatkan *cluster* yang tepat, sehingga pada iterasi ke- $(t+1)$ nilai keanggotaan objek ke- j pada *cluster* ke- i dituliskan menggunakan persamaan 2.8.

$$u_{ij}^{(t+1)} = \left[\sum_{k=1}^n \left(\frac{\|x_j - v_i\|^2}{\|x_j - v_k\|^2} \right)^{\frac{1}{m}} \right]^{-1} \quad (2.8)$$

dimana,

$$\|x_j\| = \sqrt{x_{j1}^2 + x_{j2}^2 + \dots + x_{jp}^2}$$

$$\|v_i\| = \sqrt{v_{i1}^2 + v_{i2}^2 + \dots + v_{ip}^2}$$

$$\|v_k\| = \sqrt{v_{k1}^2 + v_{k2}^2 + \dots + v_{kp}^2}$$

Setelah pusat *cluster* diperbaiki, langkah selanjutnya adalah memperbarui fungsi objektif J_m yang dapat dituliskan dengan persamaan 2.9.

$$J_m(U, V) = \sum_{j=1}^n \sum_{i=1}^c u_{ij}^m \|x_j - v_i\|^2 \quad (2.9)$$

Iterasi akan berhenti apabila $J_m < \epsilon$ dan akan dilakukan iterasi kembali apabila $J_m > \epsilon$. Berikut merupakan algoritma dalam *Fuzzy C-Means* [9]:

1. Menentukan beberapa hal sebagai berikut:
 - Jumlah *cluster* (c) dengan $2 \leq c \leq n$
 - Pangkat pembobot (m) dengan $m > 1$
 - Maksimum iterasi
 - Iterasi awal ($t = 1$)
 - Nilai *error* ($\epsilon = 0$)
2. Menentukan keanggotaan U_{ij} dimana $\sum_{i=1}^c u_{ij} = 1$
3. Menghitung pusat *cluster* v_i dimana $i = 1,2, \dots, c$
4. Memperbarui keanggotaan *Fuzzy*
5. Menghitung fungsi objektif

Apabila $J_m < \epsilon$, maka iterasi berhenti. Akan tetapi apabila $J_m > \epsilon$, maka dilakukan iterasi kembali mulai langkah 3.

E. Structural Equation Modelling

Structural equation modeling adalah analisis statistika yang digunakan untuk mengukur hubungan variabel laten dengan cara menggabungkan antara analisis regresi berganda, analisis faktor, dan analisis jalur. SEM termasuk kedalam teknik multivariat yang dapat menunjukkan bagaimana cara merepresentasikan suatu seri hubungan kasual dalam suatu diagram jalur yang kompleks [10][11].

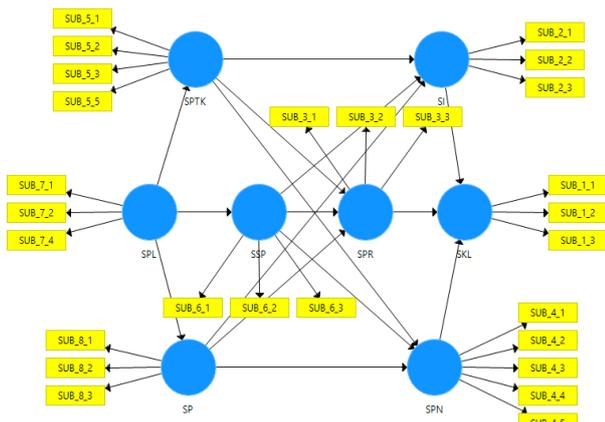
Estimasi parameter dalam model struktural SEM pada umumnya menggunakan struktur kovarians yang sering dikenal dengan Model Struktur Kovarians (MSK) atau *Covariance Based SEM* (CB-SEM) dan lebih populer dikenal dengan model LISREL (*Linier Structural Relationships*). Pengukuran estimasi dengan metode ini memiliki syarat bahwa variabel-variabel pengamatan harus kontinu, berdistribusi normal multivariat dan memiliki jumlah sampel yang besar [11].

Dalam SEM terdapat dua model, yaitu Model Struktural (*Structural Model/Inner Model*) dan Model Pengukuran (*Measurement Model/Outer Model*). Model struktural menggambarkan hubungan yang ada pada variabel laten. Hubungan yang ada umumnya berbentuk persamaan regresi linear. Beberapa persamaan regresi nantinya akan membentuk persamaan simultan variabel-variabel laten [12]. Persamaan model struktural dituliskan pada persamaan berikut.

$$\eta = \beta \eta + \Gamma \xi + \zeta \quad (2.10)$$

dimana,

- η : vektor variabel laten endogen
- ξ : vektor variabel laten eksogen
- β : matriks koefisien yang menunjukkan pengaruh variabel laten endogen terhadap variabel lainnya



Gambar 1. Model Konseptual.

Γ : matriks koefisien yang menunjukkan hubungan η dengan ξ
 ζ : vektor variabel residual

Model pengukuran mencakup hubungan antara variabel laten dengan indikator-indikatornya. Semua indikator yang dihubungkan dengan variabel latennya disebut sebagai satu blok. Setiap satu blok minimal terdiri dari satu indikator. Pola hubungan antara variabel laten dan indikator yaitu refleksif dan formatif. Pola hubungan refleksif yaitu hubungan dimana indikator merupakan refleksi variasi dari variabel latennya. Hubungan formatif yaitu pola hubungan dimana indikator membentuk variabel laten atau seolah-olah mempengaruhi variabel latennya [11]. Persamaan model pengukuran dituliskan pada persamaan 2.11 dan 2.12.

$$x = \Lambda_x \xi + \delta \tag{2.11}$$

$$y = \Lambda_y \eta + \varepsilon \tag{2.12}$$

dimana x merupakan indikator yang berkaitan atau merupakan efek dari variabel laten eksogen, sedangkan y adalah indikator yang berkaitan dengan variabel laten endogen. Λ_x dan Λ_y adalah matriks *loading* yang merupakan koefisien regresi sederhana. Sementara itu, δ dan ε adalah residual pengukuran dan indikator.

III. METODOLOGI

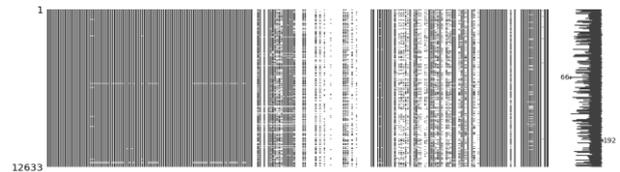
A. Sumber Data

Penelitian ini menggunakan data Standar Nasional Pendidikan jenjang Sekolah Menengah Atas (SMA) di Indonesia tahun 2018. Data tersebut merupakan data sekunder yang diperoleh dari Direktorat Jenderal Pendidikan Dasar dan Menengah Kementerian Pendidikan dan Kebudayaan Republik Indonesia.

B. Variabel Penelitian

Variabel penelitian yang digunakan untuk klusterisasi mutu pendidikan Sekolah Menengah Atas (SMA) di Indonesia adalah 8 Standar Nasional Pendidikan yang telah ditetapkan oleh Kementerian Pendidikan dan Kebudayaan Republik Indonesia seperti disajikan pada Tabel 1.

Analisis pola hubungan antar Standar Nasional Pendidikan Sekolah Menengah Atas (SMA) di Indonesia dilakukan dengan menggunakan variabel indikator Standar Nasional Pendidikan sebagaimana Tabel 2.



Gambar 2. Visualisasi Missing Value.

Tabel 1. Variabel Klusterisasi Mutu Pendidikan

Variabel	Keterangan	Skala
X ₁	Standar Kompetensi Lulusan	Rasio
X ₂	Standar Isi	Rasio
X ₃	Standar Proses	Rasio
X ₄	Standar Penilaian	Rasio
X ₅	Standar Pendidik dan Tenaga Pendidik	Rasio
X ₆	Standar Sarana dan Prasarana	Rasio
X ₇	Standar Pengelolaan	Rasio
X ₈	Standar Pembiayaan	Rasio

Model konseptual analisis SEM-PLS pada penelitian ini menggunakan delapan Standar Nasional Pendidikan (SNP) sebagai variabel laten seperti disajikan pada Tabel 2. Sementara itu, kerangka konseptual SEM-PLS pada penelitian ini disajikan pada Gambar 1.

C. Struktur Data

Struktur data yang digunakan untuk klusterisasi mutu pendidikan Sekolah Menengah Atas (SMA) di Indonesia dituliskan pada Tabel 3.

D. Langkah Penelitian

Langkah analisis digunakan untuk menggambarkan langkah-langkah penelitian yang akan dilakukan secara urut. Langkah analisis yang digunakan adalah sebagai berikut:

1. Merumuskan masalah dan studi literatur.
2. Mengumpulkan data.
3. Melakukan eksplorasi data untuk mengetahui karakteristik dari data Standar Nasional Pendidikan.
4. Klusterisasi mutu pendidikan Sekolah Menengah Atas (SMA) di Indonesia menggunakan metode *Fuzzy C-Means*.
5. Melakukan analisis pola hubungan antar Standar Nasional Pendidikan pada setiap *cluster* menggunakan metode *Structural Equation Modelling – Partial Least Square* dengan langkah-langkah sebagai berikut:
 - a. Merancang pengembangan dan pengukuran terhadap variabel laten dan indikator.
 - b. Membuat *path diagram* yang menjelaskan pola hubungan antara variabel laten dengan indikatornya.
 - c. Melakukan spesifikasi model untuk mendeskripsikan *path diagram* ke dalam model yang menggambarkan hubungan antar variabel.
 - d. Melakukan estimasi parameter berdasarkan model pengukuran dan model struktural.
 - e. Melakukan evaluasi model pengukuran dan model struktural.
 - f. Melakukan *resampling bootstrapping*
6. Menarik kesimpulan dan memberikan saran berdasarkan hasil terbaik untuk mengelompokkan mutu pendidikan Sekolah Menengah Atas (SMA) di Indonesia dan pola hubungan antar Standar Nasional Pendidikan Sekolah Menengah Atas (SMA) di Indonesia

Tabel 2.
Variabel Indikator Standar Nasional Pendidikan

Variabel Laten	Indikator	Keterangan
Kompetensi Lulusan	SUB_1_1	Lulusan memiliki kompetensi pada dimensi sikap
	SUB_1_2	Lulusan memiliki kompetensi pada dimensi pengetahuan
	SUB_1_3	Lulusan memiliki kompetensi pada dimensi keterampilan
Isi Pendidikan	SUB_2_1	Perangkat pembelajaran sesuai rumusan kompetensi lulusan
	SUB_2_2	Kurikulum Tingkat Satuan Pendidikan dikembangkan sesuai prosedur
	SUB_2_3	Sekolah melaksanakan kurikulum sesuai ketentuan
Proses Pembelajaran	SUB_3_1	Sekolah merencanakan proses pembelajaran sesuai ketentuan
	SUB_3_2	Proses pembelajaran dilaksanakan dengan tepat
	SUB_3_3	Pengawasan dan penilaian otentik dilakukan dalam proses pembelajaran
Penilaian Pendidikan	SUB_4_1	Aspek penilaian sesuai ranah kompetensi
	SUB_4_2	Teknik penilaian obyektif dan akuntabel
	SUB_4_3	Penilaian pendidikan ditindak lanjuti
	SUB_4_4	Instrumen penilaian menyesuaikan aspek
	SUB_4_5	Penilaian dilakukan mengikuti prosedur
Pendidik dan Tenaga Kependidikan	SUB_5_1	Ketersediaan dan kompetensi guru sesuai ketentuan
	SUB_5_2	Ketersediaan dan kompetensi kepala sekolah sesuai ketentuan
	SUB_5_3	Ketersediaan dan kompetensi tenaga administrasi sesuai ketentuan
	SUB_5_4	Ketersediaan dan kompetensi laboran sesuai ketentuan
	SUB_5_5	Ketersediaan dan kompetensi pustakawan sesuai ketentuan
Sarana dan Prasarana	SUB_6_1	Kapasitas daya tampung sekolah memadai
	SUB_6_2	Sekolah memiliki sarana dan prasarana pembelajaran yang lengkap dan layak
	SUB_6_3	Sekolah memiliki sarana dan prasarana pendukung yang lengkap dan layak
Pembiayaan	SUB_8_1	Sekolah memberikan layanan subsidi silang
	SUB_8_2	Beban operasional sekolah sesuai ketentuan
	SUB_8_3	Sekolah melakukan pengelolaan dana dengan baik
Pengelolaan	SUB_7_1	Sekolah melakukan perencanaan pengelolaan
	SUB_7_2	Program pengelolaan dilaksanakan sesuai ketentuan
	SUB_7_3	Kepala sekolah berkinerja baik dalam melaksanakan tugas kepemimpinan
	SUB_7_4	Sekolah mengelola sistem informasi manajemen

IV. ANALISIS DAN PEMBAHASAN

A. Karakteristik Data

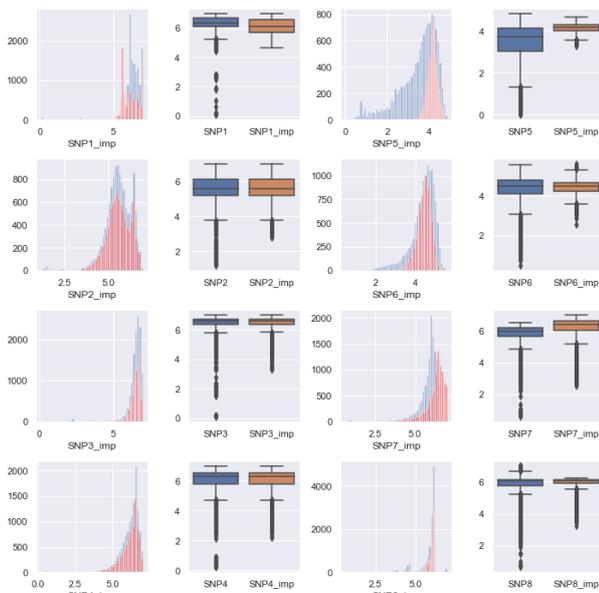
Total terdapat 13.939 sekolah jenjang SMA yang terdaftar di Kementerian Pendidikan dan Kebudayaan Indonesia. Namun, hanya 90,63% yang memiliki nilai Standar Nasional Pendidikan. Maka dari itu, pada penelitian ini hanya akan menggunakan data yang memiliki nilai Standar Nasional pendidikan, yaitu sebanyak 12.633 sekolah.

Perhitungan nilai Standar Nasional Pendidikan Indonesia berdasarkan bobot pada masing-masing sub indikator. Semua sub indikator diketahui memiliki *missing value*. Bahkan ada yang tidak ada datanya sama sekali, yaitu sebanyak 36 sub indikator. Visualisasi adanya *missing value* pada setiap variabel ditunjukkan pada Gambar 2. Titik-titik warna putih menunjukkan bahwa data observasi adalah *missing*. *Missing value* setiap variabel digambarkan dengan garis vertikal. Semakin banyak warna putih maka semakin banyak juga data yang *missing*.

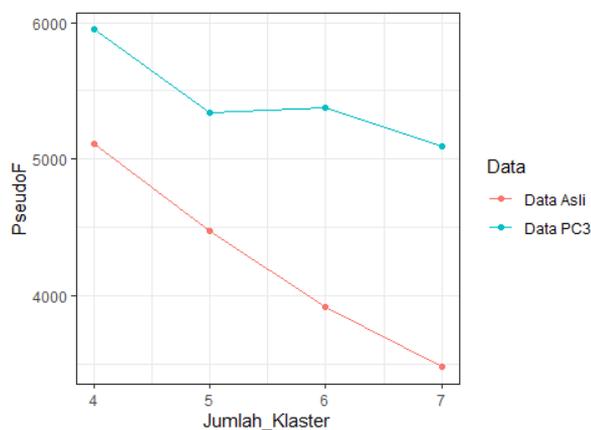
Banyaknya *missing value* pada data menyebabkan nilai Standar Nasional Pendidikan yang diperoleh tidak begitu

baik. Sekolah dengan data yang lebih lengkap cenderung memiliki nilai Standar Nasional Pendidikan yang lebih tinggi. Masalah tersebut kemudian diatasi dengan melakukan imputasi pada *missing value*. Imputasi data pada penelitian ini akan menggunakan metode regresi. Metode regresi dapat digunakan untuk mengatasi data yang saling berhubungan dan tidak berkelipatan. Selain itu, variansi yang didapatkan dari hasil imputasi juga akan lebih baik. hal tersebut tidak diperoleh ketika menggunakan imputasi *mean* atau median.

Pada penelitian ini, data yang akan diimputasi adalah data yang berasal dari variabel dengan jumlah *missing value* dibawah 80%. hal tersebut dilakukan karena jika semakin banyak terdapat *missing value* lalu diimputasi, maka hasil imputasi yang didapatkan tidak terlalu baik dan tidak representatif. Imputasi data akan dilakukan pada setiap indikator Standar Nasional Pendidikan, sehingga data akan dibagi menjadi 29 dataset baru. Variabel yang memiliki *missing value* akan menjadi respon dan variabel sisanya akan menjadi prediktor, sehingga akan diperoleh model regresi yang akan dilakukan untuk imputasi data pada variabel respon tersebut.



Gambar 3. Perbandingan SNP Sebelum dan Sesudah Imputasi.

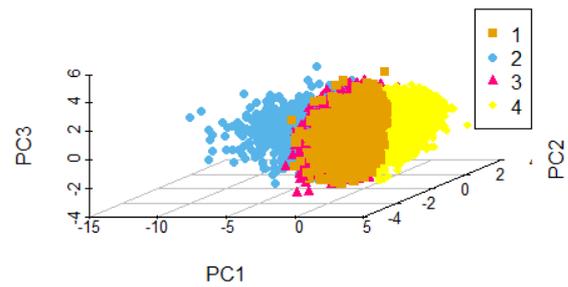


Gambar 4. Pemilihan Jumlah Kluster Terbaik.

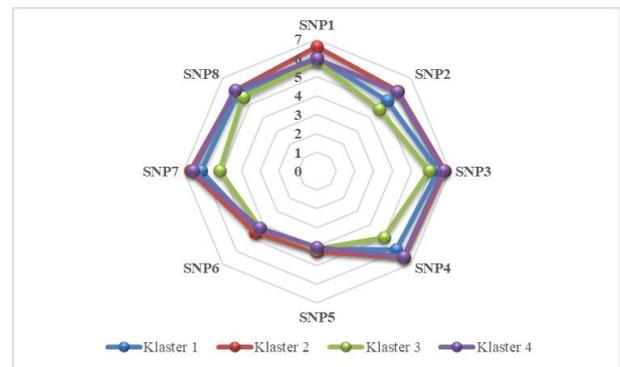
Setelah dilakukan imputasi pada data, nilai Standar Nasional Pendidikan cenderung memiliki varians yang lebih kecil dan mengurangi adanya *outlier*. secara umum, distribusi data tidak berubah secara drastis kecuali pada standar kelima, yaitu standar pendidik dan tenaga kependidikan seperti ditunjukkan pada Gambar 3 Standar ketujuh atau standar pengelolaan juga mengalami pergeseran ke kanan yang menunjukkan adanya pergeseran nilai yang lebih besar setelah dilakukan imputasi. Kedua hal tersebut kemungkinan dipengaruhi oleh indikator pada masing-masing standar yang tidak diikuti perhitungan sebelumnya.

Korelasi antar Standar Nasional Pendidikan jenjang SMA di Indonesia relatif kecil. Namun ada beberapa variabel yang memiliki korelasi cukup tinggi hingga mencapai 0,8. Beberapa standar yang memiliki korelasi tinggi, antara lain Standar Isi, Standar Proses, Standar Penilaian, dan Standar Pengelolaan. Adanya korelasi cukup tinggi tersebut bisa berpengaruh terhadap analisis yang akan dilakukan.

Setelah dilakukan *pre-processing data*, maka perlu dilakukan eksplorasi data. Hal ini dilakukan untuk memperoleh informasi-informasi penting yang terdapat pada data. Dalam statistika, eksplorasi data biasa disebut statistika dekriptif. Pada umumnya, statistika deskriptif menjelaskan ukuran pemusatan dan penyebaran data serta visualisasi data.



Gambar 5. Hasil Kluster.



Gambar 6. Perbandingan Rata-Rata SNP Tiap Kluster.

Nilai *mean* dan median setiap variabel tidak terpaut terlalu jauh yang menunjukkan bahwa kemungkinan data berdistribusi normal. Hal tersebut diperkuat dengan nilai varians yang relatif kecil, yaitu dibawah 0,5 pada semua variabel. *Range* data bisa dilihat melalui nilai maksimum dan minimum. Berdasarkan Tabel 4, diketahui bahwa *range* data cukup besar, kecuali pada variabel SNP5 dan SNP6. Hal tersebut mengindikasikan adanya *skewness* dan terdapat *outlier* pada data.

B. Klasterisasi Mutu Pendidikan

Adanya korelasi yang tinggi antar variabel yang akan digunakan dalam analisis kluster dapat menyebabkan kebaikan model berkurang. Oleh karena itu, pada penelitian ini akan dilakukan reduksi variabel menggunakan *Principal Component Analysis* (PCA). Dilakukannya reduksi pada variabel diharapkan dapat meningkatkan kebaikan model kluster yang terbentuk.

Sebelum melakukan analisis PCA, terlebih dahulu mengidentifikasi kecukupan data dan dependensi antar variabel menggunakan uji KMO dan *Bartlett*. Berdasarkan uji yang dilakukan, didapatkan nilai KMO sebesar 0,86. Nilai tersebut lebih dari 0,5 sehingga variabel-variabel yang digunakan telah memenuhi asumsi kecukupan data.

Hipotesis *Bartlett Test*:

$$H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2$$

H_1 : minimal ada satu pasang σ_k^2 yang tidak sama

Berdasarkan uji *Bartlett* diketahui bahwa nilai *p-value* yang didapatkan sebesar 0,00 yang berarti diperoleh keputusan tolak H_0 , sehingga data memiliki hubungan (korelasi) yang signifikan (dependen). Asumsi kecukupan

Tabel 3.
Struktur Data Pemetaan Mutu Pendidikan SMA

Sekolah	X ₁	X ₂	...	X _s
1	X _{1,1}	X _{2,1}	...	X _{s,1}
2	X _{1,2}	X _{2,2}	...	X _{s,2}
3	X _{1,3}	X _{2,3}	...	X _{s,3}
⋮	⋮	⋮	⋮	⋮
n	X _{1,n}	X _{2,n}	...	X _{s,n}

Tabel 4.
Statistika Deskriptif

Variabel	Mean	Median	Varians	Max	Min
SNP1	6,1494	6,0935	0,2543	7,0000	4,6317
SNP2	5,5972	5,5925	0,4448	6,9825	2,7807
SNP3	6,5180	6,5797	0,1155	7,0000	3,2895
SNP4	6,1095	6,2822	0,4006	7,0000	2,1998
SNP5	4,1407	4,1523	0,0507	4,6667	3,2709
SNP6	4,4273	4,4625	0,1124	5,5696	2,5586
SNP7	6,2614	6,4189	0,3712	7,0000	2,5250
SNP8	5,9489	6,0806	0,1360	6,2222	3,2402

Tabel 5.
Pola Hubungan SNP

Klaster	Pola Hubungan Tidak Signifikan
Pertama	SPN-SKL
	SPTK-SI
	SPTK-SPN
	SPTK-SPR
Kedua	SPN-SKL
	SPL-SPTK
Ketiga	SPTK-SI
	SPTK-SPN
	SPTK-SPR
	SP-SPN
Keempat	SPL-SPTK
	SPL-SSP
	SPN-SKL
	SSP-SPR

data dan dependensi telah terpenuhi sehingga dapat dilakukan PCA.

Berdasarkan nilai *eigenvalue*, terbentuk 2 komponen utama dengan syarat memiliki *eigenvalue* lebih dari satu. Namun, kedua komponen tersebut hanya mampu menjelaskan 59,07% variabilitas variabel yang ada. Nilai tersebut terbilang cukup kecil sehingga pemilihan banyaknya komponen PCA pada penelitian ini berdasarkan persentase varians yang mampu dijelaskan. Persentase kumulatif 3 komponen utama diketahui sebesar 71,18%, dimana angka tersebut sudah cukup baik karena mampu menjelaskan sebagian besar varians keseluruhan variabel. Oleh karena itu, pada penelitian ini diputuskan untuk menggunakan 3 komponen utama tersebut. Komponen pertama (PC1) terdiri dari Standar Isi (SNP2), Standar Proses (SNP3), Standar Penilaian (SNP4), dan Standar Pengelolaan (SNP7). Selanjutnya, komponen kedua (PC2) terdiri dari Standar Kompetensi Lulusan (SNP1) dan Standar Sarana dan Prasarana (SNP6). Sisanya, Standar Pendidik dan Tenaga

Kependidikan (SNP5) dan Standar Pembiayaan (SNP8) masuk kedalam komponen ketiga (PC3).

Setelah melakukan reduksi variabel, maka akan dilakukan analisis klaster menggunakan *Fuzzy C-Means*, dimana akan dibandingkan menggunakan data asli dengan data hasil PCA. Jumlah klaster yang dibandingkan ada sebanyak 4 klaster, yaitu 4 hingga 7 klaster. Pemilihan jumlah klaster terbaik berdasarkan nilai *PseudoF* hasil klaster. Semakin tinggi nilai *PseudoF* maka semakin baik dan menunjukkan jumlah klaster optimum.

Berdasarkan Gambar 4, diketahui bahwa menggunakan nilai *PseudoF* didapatkan jumlah klaster terbaik sebanyak 4 kelompok. Selain itu, baik menggunakan data asli maupun data hasil reduksi diperoleh hasil terbaik yang sama, yaitu empat klaster. Namun, klaster menggunakan data PCA memiliki nilai *PseudoF* yang lebih tinggi dibandingkan menggunakan data asli. Hal ini membuktikan bahwa reduksi variabel dapat meningkatkan kebaikan dari suatu klaster. Visualisasi hasil klaster disajikan pada Gambar 5.

Gambar 6 menunjukkan perbandingan rata-rata nilai Standar Nasional Pendidikan pada masing-masing klaster atau kelompok. Secara umum klaster kedua memiliki rata-rata yang paling tinggi dibandingkan dengan lainnya. Selanjutnya berturut-turut diikuti oleh klaster keempat, pertama, dan ketiga. Berdasarkan informasi tersebut dapat diketahui bahwa sekolah-sekolah yang masuk kedalam klaster kedua merupakan sekolah yang memiliki SNP lebih baik dibandingkan lainnya.

C. Analisis SEM Standar Nasional Pendidikan Jenjang SMA di Indonesia

Setelah didapatkan hasil terbaik dari analisis klaster Standar Nasional Pendidikan di Indonesia, maka selanjutnya akan dilakukan analisis *Structural Equation Modelling – Partial Least Square* (SEM-PLS) pada masing-masing klaster. Diagram jalur yang digunakan pada kedua klaster sama, yaitu seperti pada Gambar 1.

Analisis SEM-PLS melalui beberapa tahapan, diantaranya melakukan evaluasi model pengukuran dan model struktural. Kemudian setelah semua terpenuhi dilanjutkan dengan *bootstrapping* untuk mengetahui pola hubungan antar Standar Nasional Pendidikan.

Berdasarkan hasil analisis diketahui bahwa Kekuatan prediksi yang dimiliki oleh keseluruhan variabel laten endogen pada setiap klaster terbilang lemah. Hal ini ditunjukkan dengan nilai R-Square seluruh variabel laten sangat kecil dibawah 0,25. Pada masing-masing juga terdapat beberapa pola hubungan yang tidak signifikan. Selengkapnya disajikan pada Tabel 5. Adanya pola hubungan yang tidak signifikan bukan berarti tidak bagus, namun kemungkinan tanpa adanya pengaruh dari variabel lain, variabel tersebut sudah baik.

V. KESIMPULAN DAN SARAN

Berdasarkan analisis yang telah dilakukan, diperoleh kesimpulan sebagai berikut Data mengandung banyak *missing value* sehingga perlu diatasi terlebih dahulu sebelum melakukan analisis. Sementara itu, masing-masing SNP memiliki nilai varians yang kecil yang bisa jadi menunjukkan adanya pemerataan pendidikan di Indonesia. Jumlah klaster

terbaik menggunakan metode Fuzzy C-Means adalah sebanyak 4. Klaster kedua memiliki rata-rata yang paling tinggi dibandingkan dengan lainnya. Selanjutnya berturut-turut diikuti oleh klaster keempat, pertama, dan ketiga.

Pada SEM klaster pertama terdapat empat pola hubungan, yaitu antara Standar Penilaian dengan Standar Kompetensi Lulusan serta Standar Pendidik dan Tenaga Kependidikan dengan Standar Isi, Standar Penilaian, dan Standar Sarana Prasarana. Selanjutnya, pada klaster kedua hanya terdapat satu pola hubungan yang tidak signifikan, yaitu antara Standar Penilaian dengan Standar Kompetensi Lulusan. Kemudian, pada klaster ketiga bahwa terdapat empat pola hubungan yang tidak signifikan, yaitu antara Standar Pengelolaan dengan Standar Pendidik dan Tenaga Kependidikan serta Standar Pendidik dan Tenaga Kependidikan dengan Standar Isi, Standar Penilaian, dan Standar Sarana Prasarana. Sementara itu, pada klaster keempat terdapat lima pola hubungan yang tidak signifikan, yaitu antara Standar Pembiayaan dengan Standar Penilaian, Standar Pengelolaan dengan Standar Pendidik dan Tenaga Kependidikan, Standar Pengelolaan dengan Standar Sarana Prasarana, Standar Penilaian dengan Standar Kompetensi Lulusan, serta Standar Sarana Prasarana dengan Standar Proses. Adanya pola hubungan yang tidak signifikan bukan berarti tidak bagus, namun kemungkinan tanpa adanya pengaruh dari variabel lain, variabel tersebut sudah baik.

Penulis memberikan saran untuk pemerintah dan penelitian selanjutnya dalam meningkatkan mutu pendidikan di Indonesia sebagai berikut perlu melakukan pendataan yang lebih baik agar dapat dianalisis dengan baik dan mendapatkan hasil yang akurat, pada penelitian selanjutnya diharapkan dapat memberikan perlakuan yang lebih lengkap terhadap adanya missing value sehingga didapatkan hasil yang lebih akurat, meningkatkan mutu pendidikan khususnya untuk

sekolah-sekolah yang masuk kedalam klaster kedua agar memiliki mutu yang lebih baik serta perlu memperhatikan pola hubungan antar standar nasional pendidikan untuk meningkatkan mutu pendidikan di Indonesia.

DAFTAR PUSTAKA

- [1] BPS, *Indeks Pembangunan Manusia (IPM) Indonesia pada tahun 2017*. Jakarta: Badan Pusat Statistik Republik Indonesia, 2018.
- [2] Dikdasmen, *Pemetaan Mutu Pendidikan Sekolah Menengah Atas (SMA) 2018*. Jakarta: Kementerian Pendidikan dan Kebudayaan, 2018.
- [3] H. Karti, "Pengelompokan Kabupaten/Kota di Provinsi Jawa Timur Berdasarkan Indikator Pendidikan SMA/SMK/MA dengan Metode C-Means dan Fuzzy C-Means.," Institut Teknologi Sepuluh Nopember, 2013.
- [4] K. Fatimah, "Pengelompokan Mutu Pendidikan SMP Dan Pengaruh Indikator Standar Nasional Pendidikan Terhadap Mutu Kelulusan SMP," Institut Teknologi Sepuluh Nopember, 2018.
- [5] C. Amalinda, "Pengaruh Pengelolaan Sekolah Terhadap Kompetensi Lulusan Berdasarkan Standar Nasional Pendidikan di Sekolah Menengah Pertama Surabaya Menggunakan Metode Struktural Equation Modelling – Partial Least Square," Institut Teknologi Sepuluh Nopember, 2017.
- [6] N. Draper and H. Smith, *Analisis Regresi Terapan.*, 2nd ed. Jakarta: Gramedia Pustaka Utama, 1992.
- [7] V. Gasparsz, *Teknik Analisis Dalam Penelitian Percobaan 1*. Bandung: Tarsito, 1995.
- [8] J. F. H. Jr, W. C. Black, B. J. Babin, and R. E. Anderson, *Multivariate Data Analysis*, 7th ed. United States: Pearson, 2010.
- [9] X. Wang, Y. Wang, and L. Wang, "Improving fuzzy c-means clustering based on feature-weight learning," *Pattern Recognit. Lett.*, vol. 25, no. 10, 2004, doi: 10.1016/j.patrec.2004.03.008.
- [10] B. Swasono, S. Widjaja, A. Zubaydi, Z. Yuliadi, and V. Ratnasari, *A Strategic of Labour Productivity to Support Shipyard Competitiveness in Partial Least Square (PLS) Path Analysis: PLS Algorithm and Bootstrapping*. Manado: International Symposium on Ocean Science, Technology and Policy., 2009.
- [11] I. N. Afifah, "Analisis Structural Equation Modelling dengan FIMIX-PLS Studi Kasus Struktur Model Kemiskinan di Provinsi Jawa Tengah Tahun 2011.," Institut Teknologi Sepuluh Nopember, 2014.
- [12] K. A. Bollen, *Structural Equation With Latent Variables*. New York: John Wiley & Sons, Inc., 1989.