

**PENERAPAN METODE *FINITE MIXTURE PARTIAL LEAST SQUARE*
DALAM ANALISIS *STRUCTURAL EQUATION MODELING* PADA DATA
PENDIDIKAN DI INDONESIA TAHUN 2022**

**NAJLAH FAUZIAH
H051201076**



**PROGRAM STUDI STATISTIKA DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN
MAKASSAR
2024**

**PENERAPAN METODE *FINITE MIXTURE PARTIAL LEAST SQUARE*
DALAM ANALISIS *STRUCTURAL EQUATION MODELING* PADA DATA
PENDIDIKAN DI INDONESIA TAHUN 2022**

NAJLAH FAUZIAH
H051201076



Skripsi

Diajukan sebagai salah satu syarat untuk mencapai gelar Sarjana Statistika

Program Studi Statistika

pada

**DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN
MAKASSAR
2024**

SKRIPSI

PENERAPAN METODE *FINITE MIXTURE PARTIAL LEAST SQUARE*
DALAM ANALISIS *STRUCTURAL EQUATION MODELING* PADA DATA
PENDIDIKAN DI INDONESIA TAHUN 2022

NAJLAH FAUZIAH
H051201076

Skripsi,

telah dipertahankan di hadapan Panitia Ujian Sarjana Statistika pada 16 Agustus
2024 dan dinyatakan telah memenuhi syarat kelulusan

pada

Program Studi Statistika
Departemen Statistika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Hasanuddin
Makassar

Mengesahkan
Pembimbing tugas akhir,



Sitti Sahriaman, S.Si., M.Si.
NIP. 19881018 201504 2 002

Mengetahui,
Ketua Program Studi,



Dr. Anna Samiyati, S.Si., M.Si.
NIP. 19770808 200501 2 002

PERNYATAAN KEASLIAN SKRIPSI DAN PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa, skripsi berjudul "Penerapan Metode *Finite Mixture Partial Least Square* Dalam Analisis *Structural Equation Modeling* Pada Data Pendidikan di Indonesia Tahun 2022" adalah benar karya saya dengan arahan dari pembimbing skripsi saya (Sitti Sahriman S St., M Si.) Karya ilmiah ini belum diajukan dan tidak sedang diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka skripsi ini. Apabila di kemudian hari terbukti atau dapat dibuktikan bahwa sebagian atau keseluruhan skripsi ini adalah karya orang lain, maka saya bersedia menerima sanksi atas perbuatan tersebut berdasarkan aturan yang berlaku.

Dengan ini saya melimpahkan hak cipta (hak ekonomis) dari karya tulis saya berupa skripsi ini kepada Universitas Hasanuddin.

Makassar, 16 Agustus 2024



Nailan Fauziah
NIM. H051201076

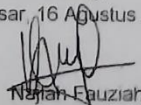
UCAPAN TERIMA KASIH

Puji syukur kehadiran Allah SWT atas segala rahmat dan karunia-Nya yang telah memberikan kekuatan serta kesehatan sehingga penulis dapat menyelesaikan skripsi "**Penerapan Metode *Finite Mixture Partial Least Square* dalam Analisis *Structural Equation Modeling* Pada Data Pendidikan di Indonesia Tahun 2022**" dengan baik. Penuh rasa syukur skripsi ini dapat diselesaikan penulis yang dibimbing dan diarahkan secara langsung oleh Ibu **Sitti Sahriman, S.Si., M.Si.** Penulis mengucapkan terima kasih kepada Ibu **Prof. Dr. Dr. Georgina Maria Tinungki, M.Si.** dan Bapak **Siswanto S.Si., M.Si.** yang telah meluangkan waktunya untuk memberikan saran serta kritikan dalam penyusunan skripsi ini. Terima kasih pula penulis ucapkan kepada **Pimpinan Universitas Hasanuddin Pihak Fakultas Matematika dan Ilmu Pengetahuan Alam, Ketua Departemen Statistika seluruh dosen dan staf Departemen Statistika** yang telah memberikan ilmu serta fasilitas selama proses penyusunan skripsi ini.

Penulis ingin menyampaikan rasa terima kasih secara tulus kepada kedua orang tua tercinta. Ayahanda **Muhammad Yusuf** dan Ibunda **Suriaty Hafid** yang telah memberikan doa dukungan, serta kasih sayang tak terhingga sepanjang perjalanan pendidikan penulis. Terima kasih telah mengizinkan penulis untuk melanjutkan pendidikan Sarjana di luar Pulau Jawa. Kebaikan, kesabaran, dan pengorbanan orang tua memberikan inspirasi untuk terus berusaha dalam setiap langkah yang diambil. Terima kasih untuk abang **Faiz**, kak **Diana Amasya Nur Ilham**, tante **Ramlah** dan om **Haris** yang selalu memberikan semangat serta doa agar skripsi ini terselesaikan dengan lancar. Terima kasih pula kepada keluarga penulis di Makassar: almarhumah tante **Rusni**, om **Haruna**, tante **Suami**, om **Muis Yanti Aiya Karina** beserta keluarga atas keberadaan dan perhatian yang diberikan selama penulis menjalani pendidikan di Universitas Hasanuddin.

Penulis sangat mengucapkan terima kasih kepada teman-teman **Statistika 2020** dan **Pois20n** terkhusus **Azal, Shafa, Liza, Linda, Nahda, Irma, Ara, Uci, Ayu, Razy, Kur, Ryan, Izzul, Rani, Ira, Isra, Ayu Afrinah, Aliyah, Dania, Tiwi, Dwini, Fadlan, Faldi, Edward, Sabil, Febi, Nahdi, Ayuni, Febi, Isti, Heri, Hakam, Divia, Inung**, teman seperbimbingan **Daya, Nurfa, Via, Laurine**, dan **Rahmi** teman KKN **Fadel, Wafiah, Nanda, Nisa, Khaeril, Alief** yang saling memberikan dukungan, semangat, serta keceriaan selama proses perkuliahan hingga penyusunan skripsi. Terima kasih juga untuk **Nurul Mujahida** selaku teman seperantauan dari Bekasi yang selalu menemani dalam suka maupun duka dan menjadi tempat berbagi cerita. Kehadiran teman-teman memberikan kenangan indah dan persahabatan yang tidak terlupakan. Penulis menyadari masih terdapat kekurangan dalam penyusunan skripsi ini, baik dari penulisan maupun analisis yang dilakukan. Penulis berharap dengan adanya skripsi ini dapat bermanfaat bagi semua pihak di masa mendatang.

Makassar, 16 Agustus 2024


Nurhan Fauziah

ABSTRAK

NAJLAH FAUZIAH. **Penerapan Metode *Finite Mixture Partial Least Square* dalam Analisis *Structural Equation Modeling* Pada Data Pendidikan di Indonesia Tahun 2022** (dibimbing oleh Sitti Sahriman, S.Si., M.Si.)

Latar Belakang. Pendidikan yang berkualitas merupakan faktor kunci dalam pembangunan negara dan kemajuan masyarakat. Sarana dan prasarana pendidikan, kegiatan peserta didik, partisipasi, serta hasil dan capaian pendidikan merupakan faktor dalam pendidikan yang dinamakan variabel laten. Untuk menganalisis hubungan kompleks antar variabel laten tersebut, diperlukan metode multivariat, yaitu *Structural Equation Modeling* (SEM) dengan estimasi parameter berbasis varians, seperti *Partial Least Square* (PLS). Dalam PLS-SEM, analisis yang dihasilkan memiliki heterogenitas pada data sehingga dibutuhkan metode *Finite Mixture Partial Least Square* (FIMIX-PLS) dengan membagi segmen hingga mendapatkan segmen terbaik pada pengelompokan wilayah dalam data berdasarkan kriteria statistik. **Tujuan.** Penelitian ini bertujuan memperoleh hubungan antar variabel laten yang berkaitan dengan pendidikan di Indonesia Tahun 2022 dan memperoleh pengelompokan wilayah di Indonesia. **Metode.** Analisis data dilakukan dengan metode PLS-SEM untuk mengetahui hubungan antar variabel laten, kemudian membagi segmen hingga mendapatkan segmen terbaik menggunakan metode FIMIX-PLS. **Hasil.** Hasil analisis menunjukkan bahwa setiap variabel indikator memiliki hubungan yang kuat terhadap variabel latennya dan dari enam koefisien jalur terbentuk, terdapat dua hubungan antar variabel laten yang signifikan. Dalam pembagian segmen dilakukan sebanyak enam kali yang menghasilkan jumlah segmen terbaik sebanyak dua segmen dengan nilai *Akaike's Information Criterion* (AIC) sebesar 133.558, lebih rendah dibandingkan segmen lain dan nilai *Normed Entropy* (EN) lebih tinggi sebesar 0.960. **Kesimpulan.** Hubungan antar variabel laten yang signifikan adalah sarana dan prasarana pendidikan (ξ_1) terhadap hasil dan capaian pendidikan (η_3) serta kegiatan peserta didik (η_1) terhadap partisipasi (η_2), kemudian pengelompokan wilayah di Indonesia terbentuk menjadi dua segmen dengan 28 provinsi dalam segmen satu dan 6 provinsi dalam segmen dua.

Kata Kunci: *Pendidikan, Structural Equation Modeling, Partial Least Square, Heterogenitas, Finite Mixture.*

ABSTRACT

NAJLAH FAUZIAH. **Application of Finite Mixture Partial Least Square Method in Structural Equation Modeling Analysis on Education Data in Indonesia 2022** (supervised by Sitti Sahriman, S.Si., M.Si.)

Introduction. Quality education is a key factor in a country's development and society's progress. Educational facilities and infrastructure, learner activities, participation, educational outcomes and achievements are factors in education that are called latent variables. To analyze the complex relationship between latent variables in education, a multivariate method is needed, namely Structural Equation Modeling (SEM) with variance-based parameter estimation, such as Partial Least Square (PLS). In PLS-SEM, the resulting analysis has heterogeneity in the data, so the Finite Mixture Partial Least Square (FIMIX-PLS) method is applied by dividing the segments to get the best segment on grouping regions in the data based on statistical criteria. **Purpose.** This study aims to obtain the relationship between latent variables related to education in Indonesia in 2022 and obtain regional groupings in Indonesia. **Methods.** Data analysis was carried out using the PLS-SEM method to determine the relationship between latent variables and then the segments were divided to get the best segment using the FIMIX-PLS method. **Results.** The results of the analysis showed that each indicator variable had a strong relationship with its latent variable and of the six path coefficients formed, there were two significant relationships between latent variables. In the division of segments, which was carried out six times, the best number of segments obtained was two with an Akaike's Information Criterion (AIC) value of 133.558, lower than other segments and a higher Normed Entropy (EN) value of 0.960. **Conclusion.** The significant relationships between latent variables were educational facilities and infrastructure (ξ_1) to educational outcomes and achievements (η_3) and learner activities (η_1) to participation (η_2), then the grouping of regions in Indonesia was formed into two segments with 28 provinces in segment one and 6 provinces in segment two.

Keywords: *Education, Structural Equation Modeling, Partial Least Square, Heterogeneity, Finite Mixture.*

DAFTAR ISTILAH

Istilah	Arti dan Penjelasan
Variabel	Faktor yang menentukan perubahan
Variabel laten	Variabel yang tidak dapat diukur secara langsung
Variabel indikator	Variabel yang dapat diukur secara langsung dan digunakan untuk menggambarkan variabel laten
Eksogen	Variabel yang tidak dipengaruhi dari variabel lain
Endogen	Variabel yang dipengaruhi dari variabel lain
Heterogenitas	Perbedaan yang terjadi dalam suatu kelompok pada data
Model pengukuran	Model yang menjelaskan mengenai hubungan antara variabel indikator terhadap variabel laten yang diwakilinya
Model struktural	Model yang menjelaskan mengenai hubungan antar variabel laten
Model indikator reflektif	Model indikator saling berkaitan dan menggambarkan satu variabel laten yang sama
Estimasi	Perkiraan, penilaian; Proses yang dilakukan untuk menentukan nilai atau ukuran dari sesuatu yang belum diketahui secara pasti
Parameter	Ukuran seluruh populasi dalam penelitian yang harus diperkirakan atau nilai numerik yang menggambarkan karakteristik dari suatu populasi
Hipotesis	Sebuah pernyataan atau dugaan bersifat sementara yang perlu diuji untuk mengetahui kebenarannya
Iterasi	Proses melakukan suatu langkah secara berulang-ulang
Segmen	Bagian atau kelompok dari keseluruhan data yang dibagi karena adanya perbedaan
<i>Standard error</i>	Ukuran statistik yang menunjukkan seberapa akurat data sampel mewakili keseluruhan populasi

DAFTAR LAMBANG/SINGKATAN

Lambang/Singkatan	Arti dan Penjelasan
ξ	Variabel laten eksogen
η	Variabel laten endogen
X	Variabel indikator pada variabel laten eksogen
Y	Variabel indikator pada variabel laten endogen
λ	Nilai <i>loading factor</i>
Λ_x	Matriks bagi <i>loading factor</i> variabel laten eksogen
Λ_y	Matriks bagi <i>loading factor</i> variabel laten endogen
δ	Kesalahan pengukuran eksogen
ε	Kesalahan pengukuran endogen
B	Matriks koefisien pengaruh variabel laten endogen ke variabel lainnya
Γ	Matriks koefisien hubungan antara variabel laten eksogen dengan endogen
ζ	Vektor residual
ρ_k	Proporsi <i>mixing</i> kelas laten k dengan $\rho_k > 0$ dan $\sum_{k=1}^K \rho_k = 1$
Ψ_k	Matriks berukuran $m \times m$ untuk kelas laten k yang mengandung varians regresi
k	Kelas atau segmen k dengan $k = 1, 2, 3, \dots, K$
τ_{mk}	Vektor koefisien regresi
ω_{mk}	Sel berukuran $m \times m$ dari ψ_k
c	Konstanta
N_k	Jumlah parameter
P_{ik}	Peluang observasi ke- i pada kelas ke- k
i	observasi ke- i ($i = 1, 2, \dots, I$)
Y_{mi}	Nilai observasi regresor untuk regresi m , individu i untuk variabel endogen
X_{mi}	Nilai observasi regresor untuk regresi m , individu i untuk variabel eksogen
A_m	Jumlah dari variabel eksogen sebagai regresor dalam regresi m
ρ_c	Reliabilitas komposit
μ	Rata-rata
y	Estimasi dari variabel laten
e	<i>error</i>
SEM	<i>Structural Equation Modeling</i>
PLS-SEM	<i>Partial Least Square Structural Equation Modeling</i>
FIMIX-PLS	<i>Finite Mixture Partial Least Square</i>
AIC	<i>Akaike's Information Criterion</i>
BIC	<i>Bayesian Information Criterion</i>
CAIC	<i>Consistent Akaike's Information Criterion</i>
EN	<i>Normed Entropy</i>
AVE	<i>Average Variance Extracted</i>

DAFTAR ISI

	Halaman
HALAMAN JUDUL.....	i
PERNYATAAN KEASLIAN SKRIPSI DAN PELIMPAHAN HAK CIPTA.....	vii
UCAPAN TERIMA KASIH.....	ix
ABSTRAK.....	xi
ABSTRACT.....	xiii
DAFTAR ISTILAH.....	xv
DAFTAR LAMBANG/SINGKATAN.....	xvii
DAFTAR ISI.....	xix
DAFTAR TABEL.....	xxi
DAFTAR GAMBAR.....	xxiii
DAFTAR LAMPIRAN.....	xxv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Batasan Masalah.....	2
1.3 Tujuan Penelitian.....	3
1.4 Manfaat Penelitian.....	3
1.5 Teori.....	3
1.5.1 Analisis <i>Structural Equation Modeling</i>	3
1.5.2 <i>Partial Least Square Structural Equation Modeling</i>	6
1.5.3 <i>Finite Mixture Partial Least Square</i>	13
1.5.4 Pendidikan di Indonesia.....	16
BAB II METODOLOGI PENELITIAN.....	19
2.1 Sumber Data.....	19
2.2 Variabel Penelitian.....	19
2.3 Tahapan Analisis.....	21
BAB III HASIL DAN PEMBAHASAN.....	23
3.1 Statistik Deskriptif Data Pendidikan di Indonesia Tahun 2022.....	23
3.2 Analisis <i>Partial Least Square Structural Equation Modeling</i> Pada Data Pendidikan di Indonesita Tahun 2022.....	24
3.2.1 Konseptualisasi Model, Mengonstruksi dan Konversi Diagram Jalur ke Sistem Persamaan.....	24
3.2.2 Estimasi Parameter.....	27
3.2.3 Mengevaluasi Model.....	28
3.2.4 <i>Resampling Bootstrap</i>	30
3.3 Analisis <i>Finite Mixture Partial Least Square</i>	33
BAB IV KESIMPULAN DAN SARAN.....	37
4.1 Kesimpulan.....	37
4.2 Saran.....	37
DAFTAR PUSTAKA.....	39
LAMPIRAN.....	43

DAFTAR TABEL

Tabel	Halaman
1. Variabel Penelitian	19
2. Statistik Deskriptif Variabel Indikator.....	23
3. Nilai <i>Loading Factor</i>	28
4. Nilai <i>Cross Loading</i>	29
5. Nilai AVE Variabel Laten.....	30
6. Nilai Reliabilitas Komposit.....	30
7. <i>Resampling Bootstrap</i>	31
8. Koefisien Jalur	32
9. Nilai Koefisien Determinasi Variabel Laten	33
10. Nilai Skor Faktor Setiap Provinsi.....	34
11. Kriteria Statistik Setiap Segmen.....	34
12. Nilai Probabilitas Setiap Provinsi	34
13. Koefisien Jalur dengan FIMIX-PLS	36

DAFTAR GAMBAR

Gambar	Halaman
1. Konstruksi Diagram Jalur.....	24
2. Hasil Diagram Jalur.....	27

DAFTAR LAMPIRAN

Lampiran	Halaman
1. Data Pendidikan di Indonesia Tahun 2022	45
2. Nilai Skor Faktor dengan FIMIX-PLS	46
3. Pengelompokan Wilayah Provinsi di Indonesia	47
4. Nilai t tabel	48
5. Hasil <i>Syntax</i> R Studio	49
6. Riwayat Hidup Penulis	54

BAB I PENDAHULUAN

1.1 Latar Belakang

Pendidikan merupakan prasyarat penting untuk pembangunan suatu negara. Standar pendidikan menentukan suatu negara menjadi maju atau tidak. Suatu negara akan mengalami kemajuan secara pesat apabila memiliki kualitas pendidikan yang baik. Sebaliknya, apabila kualitas pendidikan buruk, maka negara tersebut terhambat dalam bersaing dengan negara lain. Peningkatan kualitas pendidikan termasuk bagian dari proses yang terpadu dengan peningkatan kemampuan serta keterampilan manusia sehingga generasi penerus bangsa harus mampu dan siap bersaing di era globalisasi yang sedang berlangsung (Nurliani dkk., 2016). Keberhasilan dari pendidikan dapat dilihat berdasarkan faktor-faktor, seperti sarana dan prasarana pendidikan, kegiatan peserta didik, partisipasi, serta hasil dan capaian pendidikan di Indonesia. Faktor tersebut merupakan variabel laten yang berisikan variabel indikator (Mukhaiyar dkk., 2023). Oleh karena itu, untuk mengetahui hubungan antar variabel laten tidak dapat menggunakan analisis regresi klasik sehingga dibutuhkan suatu metode, yaitu *Structural Equation Modeling*.

Structural Equation Modeling (SEM) dikembangkan oleh Karl Joreskog, Keesling dan Walley. SEM merupakan analisis statistik multivariat yang digunakan untuk memecahkan suatu masalah dalam penelitian dengan variabel eksogen dan variabel endogen adalah variabel laten. Secara simultan, SEM dapat menguji banyak hubungan yang kompleks. Hubungan yang terbentuk ini terdiri dari variabel eksogen dan variabel endogen dengan jumlah masing-masing lebih dari satu variabel (Ningsi dan Agustina, 2018). SEM memiliki dua model persamaan, yakni model pengukuran dan model struktural. Model pengukuran merupakan model yang menjelaskan hubungan antara variabel indikator terhadap variabel latennya sedangkan model struktural merupakan model yang menjelaskan hubungan antar variabel laten (Mattjik dan Sumertajaya, 2011). Analisis SEM memiliki metode pemodelan persamaan struktural yang menggunakan pendekatan berbasis varians, yakni *Partial Least Square Structural Equation Modeling*.

Partial Least Square Structural Equation Modeling (PLS-SEM) pertama kali diperkenalkan oleh Wold. PLS-SEM merupakan teknik untuk membentuk model prediktif yang terdiri dari banyak faktor dan memiliki keterkaitan erat antara faktor-faktor tersebut. Pada PLS-SEM, model struktural dimanfaatkan untuk menguji atau mengevaluasi hubungan sebab-akibat melalui pengujian hipotesis menggunakan model prediktif sedangkan model pengukuran dimanfaatkan untuk menguji validitas dan keandalan suatu model (Ngunadi dan Anondho, 2018). PLS-SEM termasuk metode yang kuat dan fleksibel sebab dapat menggunakan berbagai macam skala pengukuran data, seperti nominal, ordinal, interval, dan rasio serta tidak membutuhkan asumsi data yang normal (Tohari dkk., 2021). Umumnya, penggunaan PLS-SEM dalam analisis melibatkan pengolahan suatu kumpulan data lengkap yang secara tidak langsung mengasumsikan bahwa data tersebut berasal dari varians yang bersifat homogen. Namun, asumsi tersebut kerap tidak sesuai dengan situasi

sebenarnya sebab data masih memiliki variasi atau heterogenitas yang tidak dapat diabaikan sehingga memengaruhi analisis yang telah dibuat.

Heterogenitas terjadi sebab terdapat perbedaan karakteristik yang signifikan dalam data. Mengetahui adanya heterogenitas pada data tidak dapat dilakukan dengan PLS-SEM sehingga hal tersebut membutuhkan suatu metode, yaitu *Finite Mixture Partial Least Square* (FIMIX-PLS). Metode ini dilakukan dengan membagi segmen hingga mendapatkan segmen yang terbaik pada pengelompokan wilayah dalam data. Hal tersebut diperoleh berdasarkan kriteria statistik dalam FIMIX-PLS, yaitu nilai *Akaike's Information Criterion* (AIC), *Bayesian Information Criterion* (BIC), *Consistent Akaike's Information Criterion* (CAIC), dan *Normed Entropy* (EN). Semakin rendah nilai AIC, BIC, dan CAIC, semakin baik segmen yang terbentuk. Dari semua nilai tersebut, nilai EN paling penting sebab apabila mendekati 1, maka data dapat dipisahkan dengan lebih jelas ke dalam segmen-segmen yang berbeda.

Penelitian mengenai PLS-SEM telah dilakukan oleh Siwalette dkk (2022) untuk mengetahui faktor-faktor yang memengaruhi terhadap keputusan pembelian secara *online* dengan hasil faktor kepercayaan dan kualitas pelayanan berpengaruh signifikan. Selain itu, Dash dan Paul (2021) melakukan perbandingan antara metode *Covariance Based Structural Equation Modeling* (CB-SEM) dengan PLS-SEM untuk penelitian dalam ilmu sosial dan peramalan teknologi dengan hasil metode PLS-SEM lebih baik dalam menangani model reflektif dan formatif dibandingkan metode CB-SEM. Peneliti Anggita dkk (2019) juga melakukan penelitian dalam pemodelan kemiskinan rumah tangga di Indonesia tahun 2017 dengan metode FIMIX-PLS yang menghasilkan jumlah segmen terbaik sebanyak dua segmen dengan nilai AIC sebesar 200.883 dan nilai EN sebesar 0.964.

Dalam penelitian mengenai hubungan antar variabel laten, sering kali terdapat dugaan heterogenitas pada data. Dugaan heterogenitas terjadi pada sampel yang diambil berasal dari populasi yang tidak sama sehingga hal tersebut perlu dilakukan pemisahan kelas atau segmentasi. Oleh karena itu, penelitian ini akan melakukan penerapan metode *Finite Mixture Partial Least Square* dalam analisis *Structural Equation Modeling* pada data pendidikan di Indonesia tahun 2022.

1.2 Batasan Masalah

Batasan masalah dalam penelitian ini adalah sebagai berikut.

1. Penelitian ini berfokus pada data pendidikan di Indonesia tahun 2022 dengan variabel laten, yaitu sarana dan prasarana pendidikan, kegiatan peserta didik, partisipasi, serta hasil dan capaian pendidikan berdasarkan Badan Pusat Statistik tahun 2022
2. Model indikator yang digunakan merupakan model indikator reflektif
3. Dalam mengatasi heterogenitas dengan pengelompokan wilayah (segmentasi), perhitungan segmen dilakukan sebanyak enam kali.
4. Variabel indikator berkaitan dengan jenjang pendidikan dasar serta menengah, yaitu Sekolah Dasar, Sekolah Menengah Pertama, dan Sekolah Menengah Atas

1.3 Tujuan Penelitian

Tujuan dalam penelitian ini adalah sebagai berikut.

1. Memperoleh hubungan antar variabel laten yang berkaitan dengan pendidikan berdasarkan metode *Finite Mixture Partial Least Square* dalam analisis *Structural Equation Modeling* pada data pendidikan di Indonesia tahun 2022
2. Memperoleh pengelompokan wilayah Indonesia yang didapatkan berdasarkan faktor pendidikan pada tahun 2022 dengan metode *Finite Mixture Partial Least Square* dalam analisis *Structural Equation Modeling*.

1.4 Manfaat Penelitian

Manfaat dalam penelitian ini adalah sebagai berikut.

1. Sebagai sarana untuk peneliti dalam menerapkan pengetahuan yang diperoleh mengenai penerapan metode *Finite Mixture Partial Least Square* dalam analisis *Structural Equation Modeling* pada data pendidikan di Indonesia tahun 2022
2. Sebagai salah satu referensi untuk penelitian selanjutnya yang berhubungan dengan data pendidikan di Indonesia dengan penerapan metode *Finite Mixture Partial Least Square* dalam analisis *Structural Equation Modeling*.

1.5 Teori

1.5.1 Analisis *Structural Equation Modeling*

Structural Equation Modeling (SEM) merupakan teknik analisis multivariat yang dikembangkan guna menutupi keterbatasan yang dimiliki oleh model analisis sebelumnya yang telah digunakan secara luas dalam penelitian statistika (Sasongko dkk., 2016). Metode ini termasuk keluarga model statistik yang digunakan untuk menjelaskan relasi antara banyak variabel. Dalam hal ini, SEM menguji struktur relasi yang ada dalam bentuk beberapa persamaan, seperti pada regresi berganda. SEM banyak digunakan dalam analisis pada penelitian sosial, perilaku, pendidikan, kesehatan, pemasaran dan ekonomi. Metode SEM memungkinkan peneliti untuk memiliki metode yang komprehensif dalam menghitung dan menguji model teori yang ada. Oleh karena itu, syarat utama menggunakan SEM adalah membangun suatu model hipotesis yang terdiri dari model struktural dan model pengukuran dalam bentuk diagram jalur yang berdasarkan justifikasi teori (Yamin, 2009).

Analisis data dengan menggunakan SEM memiliki kemampuan untuk mengestimasi hubungan antar variabel yang bersifat *multiple relationship*. Hubungan ini dibentuk dalam model struktural atau hubungan antara konstruk dependen dan independen. Kemudian, memiliki kemampuan pula untuk menggambarkan pola hubungan antara konstruk laten dan variabel manifes atau variabel indikator (Putlely dkk., 2021). Selain itu, metode ini dapat menguji hubungan kausalitas, validitas, dan reliabilitas secara bersama-sama dan dapat melihat pengaruh antar variabel secara langsung atau tidak langsung (Sasongko dkk., 2016). SEM bukan hanya membahas tentang hubungan kausalitas atau langsung dan tidak langsung pada variabel yang diamati bisa terdeteksi, tetapi juga komponen-komponen yang berkontribusi terhadap pembentukan konstruksi itu dapat ditentukan besarnya. Dengan demikian, hubungan kausalitas antara variabel atau konstruk menjadi lebih informatif, lengkap, dan akurat.

1.5.1.1 Model Persamaan *Structural Equation Modeling*

Model persamaan struktural (SEM) terdapat dua model, yaitu model pengukuran dan model struktural.

a. Model Pengukuran

Model pengukuran sering disebut dengan *outer relation* atau *measurement model* yang mendefinisikan bagaimana setiap blok indikator berhubungan dengan variabel latennya. Model pengukuran memiliki dua model indikator, yaitu model reflektif dan model formatif. Model reflektif merupakan hubungan antara variabel laten dan variabel manifes (indikator), jika terjadi perubahan variabel laten, maka akan mengakibatkan perubahan pada variabel manifes. Model reflektif dikenal sebagai model faktor utama. Sebaliknya, pada model formatif, jika terjadi perubahan variabel manifes, maka tidak akan terjadi perubahan pada variabel laten (Otok dkk., 2024). Model umum dari *outer model* adalah (Tohari dkk., 2021)

1. Model reflektif variabel eksogen, yakni variabel eksogen berpengaruh terhadap variabel endogen.

$$\mathbf{x} = \Lambda_x \boldsymbol{\xi} + \boldsymbol{\delta} \quad (1)$$

$$\begin{bmatrix} x_1 \\ \vdots \\ x_q \end{bmatrix} = \begin{bmatrix} \lambda_{x_{11}} & \lambda_{x_{12}} & \dots & \lambda_{x_{1n}} \\ \lambda_{x_{21}} & \lambda_{x_{22}} & \dots & \lambda_{x_{2n}} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{x_{q1}} & \lambda_{x_{q2}} & \dots & \lambda_{x_{qn}} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix} + \begin{bmatrix} \delta_1 \\ \vdots \\ \delta_q \end{bmatrix}$$

2. Model reflektif variabel endogen, di mana model ini terdapat variabel dependen yang dipengaruhi oleh variabel laten lainnya.

$$\mathbf{y} = \Lambda_y \boldsymbol{\eta} + \boldsymbol{\varepsilon} \quad (2)$$

$$\begin{bmatrix} y_1 \\ \vdots \\ y_p \end{bmatrix} = \begin{bmatrix} \lambda_{y_{11}} & \lambda_{y_{12}} & \dots & \lambda_{y_{1m}} \\ \lambda_{y_{21}} & \lambda_{y_{22}} & \dots & \lambda_{y_{2m}} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{y_{p1}} & \lambda_{y_{p2}} & \dots & \lambda_{y_{pm}} \end{bmatrix} \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_m \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_p \end{bmatrix}$$

dengan,

- $\mathbf{x}_{(q \times 1)}$: vektor variabel manifes/indikator eksogen
- $\mathbf{y}_{(p \times 1)}$: vektor variabel manifes/indikator endogen
- Λ_x dan Λ_y : matriks bagi *loading factor* (λ) eksogen dan endogen
- $\boldsymbol{\xi}_{(n \times 1)}$: vektor variabel laten eksogen
- $\boldsymbol{\eta}_{(m \times 1)}$: vektor variabel laten endogen
- $\boldsymbol{\delta}_{(q \times 1)}$: vektor kesalahan pengukuran eksogen
- $\boldsymbol{\varepsilon}_{(p \times 1)}$: vektor kesalahan pengukuran endogen
- p : jumlah indikator variabel endogen
- q : jumlah indikator variabel eksogen
- m : jumlah laten variabel endogen
- n : jumlah laten variabel eksogen

Model pengukuran dapat digunakan untuk menilai validitas dan reliabilitas model. Uji validitas dilakukan untuk mengetahui kemampuan instrumen penelitian mengukur apa yang seharusnya diukur. Kemudian, uji reliabilitas digunakan untuk mengukur konsistensi alat ukur dalam mengukur suatu konsep atau dapat juga digunakan untuk mengukur konsistensi responden dalam menjawab item pertanyaan dalam kuesioner atau instrumen penelitian (Abdillah dan Jogiyanto, 2009).

Evaluasi model pengukuran adalah evaluasi hubungan antara konstruk dengan indikatornya. evaluasi model pengukuran yang bersifat reflektif, yaitu evaluasi *convergent validity* dilihat dari *item reliability* (indikator validitas) yang ditunjukkan oleh nilai *loading factor*. Selanjutnya, Evaluasi *discriminant validity* dilakukan dengan dua tahap, yaitu melihat nilai *cross loading* dan membandingkan nilai akar AVE dengan korelasi antar konstruk. Kriteria dalam *cross loading* adalah bahwa setiap indikator yang mengukur konstraknya haruslah berkorelasi lebih tinggi dengan konstraknya dibandingkan dengan konstruk lainnya dan Evaluasi *composite reliability* dilakukan untuk mengukur konsistensi internal alat ukur. *Composite reliability* digunakan untuk mengukur nilai sesungguhnya reliabilitas suatu konstruk, apakah mempunyai reliabilitas yang baik atau tidak (Kasmawati dkk., 2022).

b. Model Struktural

Model struktural menunjukkan hubungan antar variabel laten berdasarkan *substantive theory*. Tanpa kehilangan sifat umumnya, diasumsikan bahwa variabel laten dan indikatornya atau variabel manifes di skala *zero means* dan unit varian sama dengan satu sehingga parameter lokasi (parameter konstanta) dapat dihilangkan dari model. Model umumnya adalah sebagai berikut (Tohari dkk., 2021).

$$\eta = B\eta + \Gamma\xi + \zeta \quad (3)$$

$$\begin{bmatrix} \eta_1 \\ \vdots \\ \eta_m \end{bmatrix} = \begin{bmatrix} 0 & 0 & \dots & 0 \\ \beta_{21} & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ \beta_{m1} & \dots & \beta_{m(m-1)} & 0 \end{bmatrix} \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_m \end{bmatrix} + \begin{bmatrix} \gamma_{11} & \gamma_{12} & \dots & \gamma_{1n} \\ \gamma_{21} & \gamma_{22} & \dots & \gamma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{m1} & \gamma_{m2} & \dots & \gamma_{mn} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix} + \begin{bmatrix} \zeta_1 \\ \vdots \\ \zeta_m \end{bmatrix}$$

dengan,

$\xi_{(n \times 1)}$: vektor variabel laten eksogen

$\eta_{(m \times 1)}$: vektor variabel laten endogen

$B_{(m \times m)}$: matriks koefisien pengaruh variabel laten endogen ke variabel lainnya

$\Gamma_{(m \times n)}$: matriks koefisien hubungan antara variabel laten eksogen dengan endogen

$\zeta_{(m \times 1)}$: vektor residual

m : banyak laten variabel endogen

n : banyak laten variabel eksogen

Setelah model pengukuran terpenuhi, maka dilakukan pengujian terhadap model struktural. Pengujian ini dilakukan dengan dua kriteria, yaitu pengujian *goodness of fit model* dan uji signifikansi hubungan konstruk. Pengujian *goodness of fit model* dilakukan dengan melihat nilai koefisien determinasi. Uji signifikansi dilakukan untuk menentukan apakah hipotesis yang dibuat dalam penelitian akan diterima atau ditolak. Uji signifikansi pada penelitian ini dilakukan dengan melihat nilai koefisien parameter dan nilai signifikansi t-statistik (Kasmawati dkk., 2022).

1.5.1.2 Analisis Jalur

Analisis jalur atau *path analysis* merupakan suatu pendekatan penelitian yang memeriksa pengaruh dari variabel-variabel yang telah diasumsikan sebagai hasil dari perlakuan terhadap variabel tersebut, baik secara langsung maupun tidak langsung secara bersama-sama atau mandiri (Sudaryono, 2011). Metode ini dikembangkan oleh Sewal Wright pada tahun 1934 yang dirancang sebagai pendekatan untuk menguji hipotesis mengenai hubungan sebab akibat dengan memanfaatkan konsep korelasi. Dalam analisis jalur, pengujian proporsi teoritis hubungan sebab akibat dapat dilakukan tanpa memanipulasi langsung terhadap variabel yang telah ditetapkan. Manipulasi yang dimaksud merujuk pada memberikan perlakuan pada suatu variabel tertentu dalam pengukurannya (Sarwono, 2011).

Analisis jalur termasuk bentuk terapan dari analisis multiregresi sebab menggunakan diagram jalur (*path diagram*) yang kompleks. Diagram ini berguna untuk mengetahui pengaruh dari variabel yang diukur dengan melihat nilai dari koefisien jalur (*path coefisients*). Metode ini dapat dijelaskan sebagai suatu prosedur untuk menghitung koefisien dari serangkaian persamaan struktural linear mengenai hubungan sebab akibat yang telah diasumsikan. Maka, dengan digunakannya diagram jalur (*path diagram*), dapat diketahui akibat secara langsung maupun tidak langsung dari satu variabel ke variabel lainnya (Sudaryono, 2011).

1.5.1.3 Analisis Faktor Konfirmatori

Analisis faktor konfirmatori atau *confirmatory factor analysis* (CFA) merupakan metode statistik yang diterapkan dalam analisis faktor untuk menguji serta memvalidasi model faktor yang telah diajukan sebelumnya. CFA bertujuan untuk mengukur sejauh mana data teramati secara konsisten sesuai dengan struktur konseptual atau teoretis yang telah diusulkan. Model CFA menggambarkan operasionalisasi variabel atau konstruksi penelitian menjadi indikator yang dapat diukur. Hal ini dirumuskan dalam bentuk persamaan atau diagram jalur tertentu (Shaleh, 2018).

Pada CFA terdapat *first order confirmatory factor analysis* dan *second order confirmatory factor analysis*. Dalam *first order* CFA, variabel laten diukur berdasarkan indikator yang diukur secara langsung sedangkan *second order* CFA, variabel laten tidak dapat diukur secara langsung melalui variabel indikatornya (Efendi dan Purnomo, 2012). Model umum untuk CFA adalah sebagai berikut (Bollen, 1989).

$$x = \Lambda_x \xi + \delta \quad (4)$$

dengan,

$x_{(q \times 1)}$: vektor variabel manifes/indikator

Λ_x : matriks bagi *loading factor* (λ)

$\xi_{(n \times 1)}$: vektor variabel laten

$\delta_{(p \times 1)}$: vektor kesalahan pengukuran

1.5.2 Partial Least Square Structural Equation Modeling

Metode *Partial Least Square* (PLS) merupakan *soft* model yang dapat menjelaskan struktur keragaman data. PLS dapat dilihat sebagai bentuk yang saling berkaitan

dengan *Prinsip Component Regression* (PCR). Model yang dihasilkan oleh metode ini mengoptimalkan hubungan antara dua kelompok variabel (Bilfarsah, 2005). Metode ini dilakukan untuk menguji teori serta data yang lemah, seperti jumlah sampel yang kecil maupun adanya masalah normalisasi data. PLS-SEM dapat digunakan pada semua jenis skala data, seperti nominal, ordinal, interval, dan rasio. Ukuran sampel pada PLS-SEM harus sama atau lebih besar dari sepuluh (10) kalinya jumlah terbanyak dari variabel manifes atau indikator yang digunakan untuk mengukur sebuah konstruk (Marliana, 2020).

Pada dasarnya, PLS-SEM digunakan untuk mengembangkan teori dalam eksplorasi penelitian dengan memfokuskan pada penjelasan varians variabel dependen ketika menguji model. Prosedur estimasi PLS-SEM adalah berdasar metode regresi *Ordinary Least Square* (OLS) daripada menggunakan *Maximum Likelihood* (ML) yang biasanya digunakan pada prosedur estimasi pada CB-SEM. PLS-SEM menggunakan data yang tersedia untuk mengestimasi hubungan *path model* dengan tujuan untuk meminimumkan varians residual dari konstruk endogen. Dengan kata lain, estimasi koefisien PLS-SEM memaksimalkan nilai R^2 dari konstruk endogen. PLS-SEM adalah model yang dipilih ketika tujuan penelitian, yaitu untuk pengembangan teori dan prediksi konstruk (menjelaskan varians). Oleh karena alasan tersebut, PLS-SEM dipilih sebagai pendekatan SEM berbasis varians. PLS-SEM dapat digunakan pada ukuran sampel kecil dan model yang kompleks serta tidak membutuhkan asumsi dari distribusi data. Sebagai tambahan, PLS-SEM dapat dengan mudah menangani model pengukuran reflektif dan formatif. PLS-SEM tidak memiliki ukuran *goodness of fit model* yang memadai sehingga pengujian teori dan *confirmation* terbatas (Hair dkk., 2014).

1.5.2.1 Pemodelan *Partial Least Square Structural Equation Modeling*

Metode *Partial Least Square* (PLS) dapat digunakan dalam pemodelan struktural berbentuk indikator reflektif maupun formatif. Model indikator reflektif dirancang berdasarkan teori uji klasik yang memperkirakan bahwa variasi skor pengukuran konstruk bergantung pada *true score* yang ditambah dengan kesalahan atau eror. Ciri dari model tersebut adalah arah hubungan kausalitas dari variabel laten ke indikator, antar indikator diharapkan saling berkorelasi, menghilangkan satu indikator dari model pengukuran tidak berubah arti dari variabel laten dan menghitung kesalahan atau eror pada tingkat indikator.

Dalam pemodelan *Partial Least Square Structural Equation Modeling* (PLS-SEM), langkah awal yang dilakukan adalah melakukan konseptualisasi model, di mana dilakukan pengembangan dan pengukuran variabel laten (Jaya dan Sumertajaya, 2008). Hal ini terdapat dua model yang dibentuk, yaitu model pengukuran dan model struktural (Alfa dkk., 2017).

1.5.2.2 Konstruksi Diagram Jalur

Diagram jalur dikonstruksi apabila telah dilakukan perancangan pada model pengukuran (*outer model*) dan model struktural (*inner model*). Diagram ini menggambarkan hubungan antar variabel laten dengan indikatornya menggunakan

path models. Hal ini memudahkan untuk melihat hubungan dari variabel tersebut (Irwan dan Adam, 2015).

1.5.2.3 Konversi Diagram Jalur ke Sistem Persamaan

Konversi ini bertujuan untuk mengetahui kekuatan pengaruh antar variabel laten yang dijelaskan pada model, yakni efek langsung dan tidak langsung (Hamidiana dkk., 2016). Dalam konversi ini dilakukan terhadap persamaan model pengukuran (*outer model*) dan model struktural (*inner model*). Apabila konversi ini dilakukan terhadap model pengukuran (*outer model*), maka persamaan model tersebut ditunjukkan pada Persamaan (1) dan Persamaan (2). Apabila konversi ini dilakukan terhadap model struktural, maka ditunjukkan pada Persamaan (3).

1.5.2.4 Estimasi Parameter

Metode pendugaan parameter atau estimasi dalam metode *Partial Least Square* (PLS) menggunakan metode kuadrat terkecil (*Least Square Methods*). Estimasi dalam PLS-SEM bertujuan untuk menghasilkan bobot atau skor komponen terbaik dari variabel laten endogen yang memungkinkan prediksi optimal terhadap hubungan antara variabel laten dengan variabel manifes atau indikator. Proses perhitungan dilakukan dengan proses iterasi, di mana iterasi akan berhenti jika telah mencapai kondisi konvergen. Batas konvergen dalam proses iterasi adalah sebagai berikut (Jaya dan Sumertajaya, 2008).

$$\left(\left(\frac{\lambda_{jh}^* - \lambda_{jh}}{\lambda_{jh}} \right) \right) \leq 10^{-5}$$

Proses iterasi dalam PLS terdapat tiga tahap, setiap tahap menghasilkan estimasi. Berikut langkah-langkah estimasi parameter *Partial Least Square* (PLS) (Irwan dan Adam, 2015).

1. Estimasi bobot

Estimasi bobot bertujuan untuk menciptakan skor (*score factor*) pada variabel laten. Estimasi bobot λ_{jh} diperoleh dari mode A dan mode B. Mode A dirancang untuk memperoleh estimasi bobot dengan model indikator reflektif sedangkan mode B dirancang untuk memperoleh estimasi bobot dengan model indikator formatif.

a) Model indikator reflektif (mode A)

Model indikator reflektif memiliki arah hubungan kausalitas dari variabel laten ke indikator, antar indikator diharapkan saling berkorelasi. Adapun persamaan dalam model indikator reflektif adalah sebagai berikut.

$$x_{jh} = \lambda_{jh}\xi_j + \delta_{jh} \quad (5)$$

Pada persamaan tersebut, variabel eksogen dilambangkan simbol ξ (Ksi), bobot λ_{jh} merupakan koefisien regresi dari ξ_j dalam regresi sederhana yang memuat variabel eksogen x_{jh} .

Estimasi ini didapatkan dengan meminimumkan jumlah kuadrat error δ_{jh} .

$$\begin{aligned} x_{jh} &= \lambda_{jh}\xi_j + \delta_{jh} \\ \delta_{jh} &= x_{jh} - \lambda_{jh}\xi_j \end{aligned}$$

$$\sum_{j=1}^J \delta_{jh}^2 = \sum_{j=1}^J (x_{jh} - \lambda_{jh} \xi_j)^2 \quad (6)$$

Selanjutnya, melakukan penurunan jumlah kuadrat error δ_{jh} terhadap λ_{jh} pada Persamaan (6).

$$\begin{aligned} \frac{\partial \sum_{j=1}^J \delta_{jh}^2}{\partial \lambda_{jh}} &= 0 \\ \frac{\partial \sum_{j=1}^J (x_{jh} - \lambda_{jh} \xi_j)^2}{\partial \lambda_{jh}} &= 0 \\ \sum_{j=1}^J 2(x_{jh} - \lambda_{jh} \xi_j)(-\xi_j) &= 0 \\ \sum_{j=1}^J -2\xi_j(x_{jh} - \lambda_{jh} \xi_j) &= 0 \\ \sum_{j=1}^J \xi_j(x_{jh} - \lambda_{jh} \xi_j) &= 0 \\ \sum_{j=1}^J (x_{jh} \xi_j) - \sum_{j=1}^J (\lambda_{jh} \xi_j^2) &= 0 \\ \sum_{j=1}^J (x_{jh} \xi_j) &= \sum_{j=1}^J (\lambda_{jh} \xi_j^2) \\ \sum_{j=1}^J (x_{jh} \xi_j) &= \lambda_{jh} \sum_{j=1}^J (\xi_j^2) \\ \lambda_{jh} &= \frac{\sum_{j=1}^J (x_{jh} \xi_j)}{\sum_{j=1}^J (\xi_j^2)} \\ \hat{\lambda}_{jh} &= E \left[\frac{\sum_{j=1}^J (x_{jh} \xi_j)}{\sum_{j=1}^J (\xi_j^2)} \right] = \frac{E[\sum_{j=1}^J (x_{jh} \xi_j)]}{E[\sum_{j=1}^J (\xi_j^2)]} \end{aligned}$$

sehingga, didapatkan persamaan estimasi untuk mode A terhadap variabel eksogen, yaitu

$$\hat{\lambda}_{jh} = \frac{cov(x_{jh}, \xi_j)}{var(\xi_j)} \quad (7)$$

Hal ini berlaku juga untuk mode A terhadap variabel endogen yang memiliki persamaan sebagai berikut.

$$y_{jh} = \lambda_{jh} \eta_j + \varepsilon_{jh} \quad (8)$$

Pada persamaan tersebut, variabel endogen dilambangkan simbol η , bobot λ_{jh} merupakan koefisien regresi dari η_j dalam regresi sederhana yang memuat variabel endogen y_{jh} .

Estimasi untuk variabel endogen, didapatkan dengan meminimumkan jumlah kuadrat error ε_{jh} .

$$\begin{aligned} y_{jh} &= \lambda_{jh} \eta_j + \varepsilon_{jh} \\ \varepsilon_{jh} &= y_{jh} - \lambda_{jh} \eta_j \\ \sum_{j=1}^J \varepsilon_{jh}^2 &= \sum_{j=1}^J (y_{jh} - \lambda_{jh} \eta_j)^2 \end{aligned} \quad (9)$$

Selanjutnya, melakukan penurunan jumlah kuadrat error ε_{jh} terhadap λ_{jh} pada Persamaan (9).

$$\begin{aligned} \frac{\partial \sum_{j=1}^J \varepsilon_{jh}^2}{\partial \lambda_{jh}} &= 0 \\ \frac{\partial \sum_{j=1}^J (y_{jh} - \lambda_{jh} \eta_j)^2}{\partial \lambda_{jh}} &= 0 \end{aligned}$$

$$\begin{aligned}
\sum_{j=1}^J 2(y_{jh} - \lambda_{jh}\eta_j)(-\eta_j) &= 0 \\
\sum_{j=1}^J -2\eta_j(y_{jh} - \lambda_{jh}\eta_j) &= 0 \\
\sum_{j=1}^J \eta_j(y_{jh} - \lambda_{jh}\eta_j) &= 0 \\
\sum_{j=1}^J (y_{jh}\eta_j) - \sum_{j=1}^J (\lambda_{jh}\eta_j^2) &= 0 \\
\sum_{j=1}^J (y_{jh}\eta_j) &= \sum_{j=1}^J (\lambda_{jh}\eta_j^2) \\
\sum_{j=1}^J (y_{jh}\eta_j) &= \lambda_{jh} \sum_{j=1}^J (\eta_j^2) \\
\lambda_{jh} &= \frac{\sum_{j=1}^J (y_{jh}\eta_j)}{\sum_{j=1}^J (\eta_j^2)} \\
\hat{\lambda}_{jh} &= E \left[\frac{\sum_{j=1}^J (y_{jh}\eta_j)}{\sum_{j=1}^J (\eta_j^2)} \right] = \frac{E[\sum_{j=1}^J (y_{jh}\eta_j)]}{E[\sum_{j=1}^J (\eta_j^2)]}
\end{aligned}$$

Maka, persamaan estimasi untuk mode A terhadap variabel endogen, yaitu

$$\hat{\lambda}_{jh} = \frac{cov(y_{jh}, \eta_j)}{var(\eta_j)} \quad (10)$$

2. Estimasi Jalur

Estimasi jalur bertujuan untuk menghubungkan antar variabel laten dan estimasi *loading factor* dengan indikatornya. Estimasi *loading factor* merupakan koefisien *outer model*, dilambangkan λ . Koefisien jalur dilambangkan dengan β dan γ , di mana koefisien β merupakan koefisien penghubung antar variabel laten endogen, sedangkan koefisien γ merupakan koefisien penghubung antara variabel laten eksogen ke variabel endogen.

3. Estimasi rata-rata (*means*)

Estimasi ini untuk variabel indikator dan variabel laten. Tahapan ini untuk memperoleh skor rata-rata dan konstanta untuk variabel laten yang digunakan sebagai parameter dengan mencerminkan sifat hubungan kausalitas dan nilai rata-rata sampel yang diperoleh (Jaya dan Sumertajaya, 2008). Estimasi rata-rata μ_j didapatkan melalui persamaan (Otok dkk., 2024).

$$\begin{aligned}
\xi_j &= y_j + \mu_j + e_j \\
\xi_j - \mu_j &= y_j + e_j
\end{aligned} \quad (11)$$

diketahui bahwa nilai $y_j = \sum_{h=1}^J \hat{\lambda}_{jh}(x_{jh} - \bar{x}_{jh})$ sehingga

$$\begin{aligned}
\xi_j - \mu_j &= \sum_{h=1}^J \hat{\lambda}_{jh}(x_{jh} - \bar{x}_{jh}) \\
\xi_j - \mu_j &= \sum_{h=1}^J (\hat{\lambda}_{jh}x_{jh} - \hat{\lambda}_{jh}\bar{x}_{jh})
\end{aligned}$$

$$\begin{aligned}
\xi_j &= \sum_{h=1}^J (\hat{\lambda}_{jh}x_{jh}) \\
\mu_j &= \sum_{h=1}^J \hat{\lambda}_{jh}\bar{x}_{jh}
\end{aligned} \quad (12)$$

1.5.2.5 Resampling Bootstrap

Metode *bootstrapping* merupakan alat untuk membantu mengurangi ketidakandalan yang berhubungan dengan kesalahan penggunaan distribusi normal dan penggunaannya. Pada proses *bootstrap*, dilakukan pengambilan sampel secara *resampling with replacement*. PLS dengan sampel kecil menggunakan metode *resampling bootstrap standard error* untuk menilai tingkat signifikansi dan memperoleh kestabilan estimasi model pengukuran dan model struktural dengan mencari estimasi dari *standard error*. Adapun rumus *bootstrap standard error* dari $\hat{\theta}$ yang dihitung dengan standar deviasi dari B perulangan adalah sebagai berikut (Mukhaiyar dkk., 2023).

$$\widehat{se}(\hat{\theta}_B) = \sqrt{\frac{\sum_{b=1}^B (\hat{\theta}_{(b)}^* - \hat{\theta}_{(\cdot)}^*)^2}{B-1}} = \sqrt{\widehat{var}_F(\hat{\theta}^*)} \quad (13)$$

$$\hat{\theta}_{(\cdot)}^* = \frac{\sum_{b=1}^B \hat{\theta}_{(b)}^*}{B}$$

dengan,

B : jumlah *resampling* yang berukuran n dengan *resampling* dari *plug-in estimate* F (ukuran sampel *bootstrapping*)

$\hat{\theta}_{(b)}^*$: statistik data asli

$\hat{\theta}_{(\cdot)}^*$: dihitung dari sampel ulang ke- b untuk $b = 1, 2, \dots, B$

1.5.2.6 Evaluasi Model

Evaluasi model pada PLS-SEM terdapat dua tahap, yaitu evaluasi model pengukuran (*outer model*) dan evaluasi model struktural (*inner model*). Evaluasi model dapat dilakukan pada model indikator reflektif. Evaluasi ini bertujuan untuk mengetahui bahwa model tersebut telah valid dan reliabel.

1. Evaluasi Model Pengukuran

Dalam evaluasi model pengukuran pada model indikator reflektif, dilakukan pengujian validitas konvergen (*convergent validity*), validitas diskriminan (*discriminant validity*), reliabilitas komposit (*composite reliability*), dan *Average Variance Extracted* (AVE).

a. Validitas Konvergen

Validitas konvergen merupakan validitas yang menegaskan bahwa alat ukur suatu konstruk harus memiliki korelasi yang kuat dan signifikan sehingga dapat digunakan untuk mengukur suatu hubungan antara variabel laten dan variabel yang teramati dalam model pengukuran reflektif. Validitas ini berdasarkan pada nilai *loading factor* lebih besar dari 0.70 untuk menjadi ideal, tetapi nilai *loading factor* sebesar 0.50 hingga 0.60 masih dapat diterima. Apabila nilai *loading factor* dibawah 0.50, maka indikator tidak valid dan harus dikeluarkan dari model. Hal ini disebabkan indikator tidak cukup baik digunakan untuk mengukur variabel laten.

b. Validitas Diskriminan

Validitas diskriminan menunjukkan bahwa pengukur konstruk yang berbeda seharusnya tidak berkorelasi tinggi. Hal ini dapat diketahui berdasarkan nilai *cross*

loading dari variabel manifes terhadap masing-masing dalam variabel laten. Indikator yang baik dapat diketahui dengan melihat korelasi antara variabel laten dengan setiap indikator lebih besar daripada korelasi dengan variabel laten lainnya (Ghozali dan Latan, 2014).

Metode lain untuk menguji validitas diskriminan adalah membandingkan nilai akar kuadrat dari *Average Variance Extracted* (AVE) untuk setiap konstruk dengan nilai korelasi antar konstruk dalam model. Validitas diskriminan dapat terpenuhi apabila nilai AVE lebih besar dari 0,50 (Alfa dkk., 2017). Adapun persamaan rumus AVE adalah sebagai berikut.

$$AVE = \frac{\sum \lambda_i^2}{\sum \lambda_i^2 + \sum var(\varepsilon_i)}$$

$$AVE = \frac{\sum \lambda_i^2}{\sum \lambda_i^2 + \sum (1 - \lambda_i^2)} \quad (14)$$

di mana, λ_i adalah *loading factor* dari variabel indikator.

c. Reliabilitas Komposit

Variabel laten dikatakan mempunyai reliabilitas yang baik apabila nilai *composite reliability* lebih besar dari 0.70. Nilai *composite reliability* yang tinggi menunjukkan nilai konsistensi dari masing-masing indikator dalam mengukur variabel laten. Adapun persamaan rumus dari reliabilitas komposit atau *composite reliability* adalah sebagai berikut.

$$\rho_c = \frac{(\sum \lambda_i)^2}{(\sum \lambda_i)^2 + \sum (1 - \lambda_i^2)} \quad (15)$$

2. Evaluasi Model Struktural

Model struktural (*inner model*) dievaluasi untuk melihat hubungan antar konstruk laten (Irwan dan Adam, 2015). Pengujian evaluasi model struktural adalah

a. Uji Signifikansi

Uji signifikansi bertujuan untuk mengetahui pengaruh antara variabel manifes/indikator terhadap variabel laten dan hubungan antar variabel laten. Perumusan hipotesis pada uji signifikansi adalah sebagai berikut.

H_0 : Variabel manifes/indikator tidak berpengaruh secara signifikan terhadap variabel laten

H_1 : Variabel manifes/indikator berpengaruh secara signifikan terhadap variabel laten

H_0 : Hubungan antar variabel laten tidak berpengaruh signifikan

H_1 : Hubungan antar variabel laten berpengaruh signifikan

Statistik uji yang digunakan adalah

$$t_{hitung} = \frac{b_j}{s(b_j)} \quad (16)$$

dengan,

b_j : nilai taksiran untuk β_j

$s(b_j)$: *standard error* untuk b_j

Kriteria pengujianya adalah apabila diperoleh nilai t_{hitung} lebih besar dari t_{tabel} , maka H_0 ditolak sehingga variabel manifes/indikator berpengaruh secara signifikan

terhadap variabel latennya dan hubungan antar variabel laten berpengaruh signifikan (Alfa dkk., 2017).

b. *R-Squares* (R^2)

Pengujian *R-Squares* dilakukan dengan mengukur tingkat *goodnes of fit* suatu model struktural. Hal ini untuk mengetahui besar pengaruh antar variabel laten. Nilai *R-Squares* 0.75; 0.50; 0.25 atau 0.35; 0.15; 0.02 dapat disimpulkan bahwa model kuat, moderate, dan lemah (Sholiha dan Salamah, 2015).

1.5.3 *Finite Mixture Partial Least Square*

Metode PLS-SEM digunakan dalam menganalisis suatu data yang lengkap dan mengasumsikan bahwa data tersebut berasal dari satu populasi yang homogen. Asumsi mengenai homogenitas dalam karakteristik data sering kali tidak sesuai dengan kenyataan. Ketidampungan untuk memperhitungkan heterogenitas dapat menjadi ancaman bagi validitas hasil dari metode PLS-SEM sehingga menyebabkan kesimpulan yang salah.

Heterogenitas dalam data dapat diamati dan tidak dapat diamati. Heterogenitas dapat diamati apabila terdapat perbedaan antara dua atau lebih kelompok data berkaitan dengan karakteristik yang dapat diamati, seperti usia, jenis kelamin, atau asal negara sehingga karakteristik ini digunakan untuk membagi data ke dalam kelompok-kelompok pengamatan yang terpisah. Sebaliknya, Heterogenitas tidak dapat diamati terjadi apabila ada perbedaan signifikan dalam hubungan model antara kelompok data dan sumber perbedaan ini tidak dapat ditelusuri kembali ke karakteristik yang dapat diamati (Hair dkk., 2016).

Kuadrat terkecil parsial campuran hingga atau *Finite Mixture Partial Least Square* merupakan pendekatan campuran terbatas dengan mengasumsikan bahwa data berasal dari beberapa segmen atau subpopulasi. Setiap segmen dimodelkan secara terpisah dan keseluruhan populasi merupakan campuran dari segmen-segmen tersebut. Dalam data yang mengalami heterogenitas dapat menggunakan FIMIX-PLS dengan membagi segmen hingga mendapatkan segmen yang terbaik untuk data tersebut (Sarstedt dkk., 2022).

1.5.3.1 Model Persamaan *Finite Mixture Partial Least Square*

Finite Mixture Partial Least Square diterapkan dengan memanfaatkan skor prediksi variabel laten dan menjelaskan tentang hubungan dalam model struktural (*inner model*). Asumsi dalam FIMIX-PLS adalah apabila unit observasi telah dipisahkan menurut segmennya, maka kasus heterogenitas tidak akan terjadi dalam model struktural. Adapun persamaan yang terbentuk berdasarkan hubungan yang ada pada model struktural adalah sebagai berikut.

$$B\eta_i + \Gamma\xi_i = \zeta_i \quad (17)$$

Estimasi FIMIX-PLS didasarkan pada asumsi bahwa heterogenitas terjadi pada model struktural dengan asumsi η_i berdistribusi campuran berhingga (*finite mixture*) dengan fungsi kepadatan multivariat normal $f_{i|k}(\cdot)$, maka

$$\eta_i \sim \sum_{k=1}^K \rho_k f_{i|k}(\eta_i | \xi_i, B_k, \Gamma_k, \Psi_k) \quad (18)$$

dengan,

- $\boldsymbol{\eta}_i$: vektor variabel endogen pada model struktural dengan $i = 1, 2, \dots, I$
 $\boldsymbol{\xi}_i$: vektor variabel laten eksogen pada model struktural dengan $i = 1, 2, \dots, I$
 ρ_k : proporsi *mixing* kelas laten k dengan $\rho_k > 0$ dan $\sum_{k=1}^K \rho_k = 1$
 $f_{i|k}(\cdot)$: peluang untuk kasus i *given* kelas k dan parameter (\cdot)
 \mathbf{B}_k : matriks koefisien jalur pada model struktural untuk kelas laten k yang menunjukkan hubungan antar variabel laten endogen dengan ukuran adalah $m \times m$
 $\boldsymbol{\Gamma}_k$: matriks koefisien jalur pada model struktural untuk kelas laten k yang menunjukkan hubungan antara variabel laten eksogen dengan endogen berukuran $m \times j$
 $\boldsymbol{\Psi}_k$: matriks $m \times m$ untuk kelas laten k yang mengandung varians regresi
 I : jumlah total kasus/observasi
 i : kasus/observasi i ($i = 1, 2, 3, \dots, I$)
 K : jumlah keseluruhan kelas
 k : kelas atau segmen k dengan $k = 1, 2, 3, \dots, K$

Selanjutnya, melakukan substitusi hasil $f_{i|k}(\boldsymbol{\eta}_i | \boldsymbol{\xi}_i, \mathbf{B}_k, \boldsymbol{\Gamma}_k, \boldsymbol{\Psi}_k)$, maka menghasilkan persamaan

$$\boldsymbol{\eta}_i = \sum_{k=1}^K \rho_k \left[\frac{|\mathbf{B}_k|}{M\sqrt{2\pi}\sqrt{|\boldsymbol{\Psi}_k|}} \right] e^{-\frac{1}{2}(\mathbf{B}_k\boldsymbol{\eta}_i + \boldsymbol{\Gamma}_k\boldsymbol{\xi}_i)' \boldsymbol{\Psi}_k^{-1} (\mathbf{B}_k\boldsymbol{\eta}_i + \boldsymbol{\Gamma}_k\boldsymbol{\xi}_i)} \quad (19)$$

Dalam FIMIX-PLS, estimasi model mengikuti fungsi *likelihood* yang dimaksimumkan dengan algoritma *Expectation Maximization* (EM). Algoritma ini terdiri dari E-*step* yang digunakan dalam estimasi parameter untuk menghasilkan fungsi ekspektasi *likelihood* dan M-*step* yang memaksimumkan ekspektasi *log-likelihood* dari E-*step* untuk menghitung parameter. Adapun persamaan yang menunjukkan fungsi *log-likelihood* adalah sebagai berikut.

$$L = \prod \prod [\rho_k f_{i|k}(\boldsymbol{\eta}_i | \boldsymbol{\xi}_i, \mathbf{B}_k, \boldsymbol{\Gamma}_k, \boldsymbol{\Psi}_k)]^{z_{ik}} \quad (20)$$

$$\ln L = \sum_{i=1}^I \sum_{k=1}^K z_{ik} \ln(f(\boldsymbol{\eta}_i | \boldsymbol{\xi}_i, \mathbf{B}_k, \boldsymbol{\Gamma}_k, \boldsymbol{\Psi}_k)) + \sum_{i=1}^I \sum_{k=1}^K z_{ik} \ln(\rho_k) \quad (21)$$

Nilai ekspektasi pada persamaan di atas dihitung dalam E-*step* dengan z_{ik} bernilai 1 jika i masuk kelas ke- k dan jika lainnya bernilai 0 (nol). Ukuran segmen ρ_k , parameter $\boldsymbol{\xi}_i, \mathbf{B}_k, \boldsymbol{\Gamma}_k, \boldsymbol{\Psi}_k$ dari fungsi probabilitas bersyarat. Adapun nilai ekspektasi untuk z_{ik} mengikuti teorema bayes adalah sebagai berikut.

$$E(z_{ik}) = P_{ik} = \frac{\rho_k f_{i|k}(\boldsymbol{\eta}_i | \boldsymbol{\xi}_i, \mathbf{B}_k, \boldsymbol{\Gamma}_k, \boldsymbol{\Psi}_k)}{\sum_{k=1}^K \rho_k f_{i|k}(\boldsymbol{\eta}_i | \boldsymbol{\xi}_i, \mathbf{B}_k, \boldsymbol{\Gamma}_k, \boldsymbol{\Psi}_k)} \quad (22)$$

dengan, P_{ik} merupakan peluang keanggotaan observasi i untuk kelas- k .

Pada persamaan *log-likelihood*, nilai ρ_k dihitung dengan rata-rata dari nilai ekspektasi P_{ik} yang dihasilkan dari E-*step* sebelumnya, maka

$$\rho_k = \frac{\sum_{i=1}^I P_{ik}}{I} \quad (23)$$

Parameter optimal untuk $\mathbf{B}_k, \boldsymbol{\Gamma}_k$, dan $\boldsymbol{\Psi}_k$ ditentukan oleh regresi OLS independen. Estimator *Maximum Likelihood* (ML) untuk koefisien dan varians diasumsikan dengan prediksi *Ordinary Least Square* (OLS). Adapun persamaan

untuk mendapatkan parameter regresi variabel laten endogen dan eksogen adalah sebagai berikut.

$$Y_{mi} = \eta_{mi} \quad (24)$$

dengan Y_{mi} adalah nilai observasi regresor untuk regresi m , individu i

$$X_{mi} = (E_{mi}, N_{mi})' \quad (25)$$

dengan X_{mi} adalah nilai observasi regresor untuk regresi m , individu i

$$E_{mi} = \begin{cases} \{\xi_1, \dots, \xi_{A_m}\}, A_m \geq 1, a_m = 1, \dots, A_m \wedge \xi_{a_m}, \text{ regresor dari } m \\ \emptyset \text{ lainnya} \end{cases} \quad (26)$$

dengan,

A_m : Jumlah variabel eksogen sebagai regresor dalam regresi m

a_m : variabel eksogen a_m ($a_m = 1, 2, \dots, A_m$)

$$N_{mi} = \begin{cases} \{\eta_1, \dots, \eta_{B_m}\}, B_m \geq 1, b_m = 1, 2, \dots, B_m \wedge \eta_{b_m}, \text{ regresor dari } m \\ \emptyset \text{ lainnya} \end{cases} \quad (27)$$

dengan,

B_m : Jumlah variabel endogen sebagai regresor dalam regresi m

b_m : variabel endogen b_m ($b_m = 1, 2, \dots, B_m$)

Selanjutnya, bentuk persamaan OLS untuk τ_{mk} dan ω_{mk} adalah sebagai berikut.

$$\tau_{mk} = \left((\gamma_{a_m} mk), (\beta_{b_m} mk) \right)' \quad (28)$$

$$\tau_{mk} = [\sum_{i=1}^I P_{ik} (X'_{mi} X_{mi})]^{-1} [\sum_{i=1}^I P_{ik} (X'_{mi} Y_{mi})]$$

dan

$$\omega_{mk} = \text{cell } (m \times m) \text{ of } \psi_k \quad (29)$$

$$\omega_{mk} = \frac{\sum_{i=1}^I P_{ik} (Y_{mi} - X_{mi} \tau_{mk}) (Y_{mi} - X_{mi} \tau_{mk})'}{IP_{ik}}$$

dengan,

τ_{mk} : vektor koefisien regresi

ω_{mk} : sel berukuran $m \times m$ dari ψ_k

M-step menentukan proporsi *mixing* baru ρ_k dan regresi OLS independen digunakan pada iterasi E-step selanjutnya untuk meningkatkan nilai P_{ik} . Algoritma EM akan berhenti apabila tidak ada perubahan yang substansial dalam nilai LnL dari satu iterasi ke iterasi berikutnya (Ringle dkk., 2010).

1.5.3.2 Kriteria Statistik *Finite Mixture Partial Least Square*

Dalam FIMIX-PLS, jumlah segmen terbaik tidak dapat ditentukan sebab distribusi model campuran tidak asimtotik, seperti *chi-square* dan validitas untuk rasio uji *likelihood*. Untuk mengetahui segmen terbaik dapat dilihat dari kriteria statistik pada FIMIX-PLS, yaitu

a. *Akaike's Information Criterion*

Persamaan rumus untuk menghitung AIC adalah sebagai berikut.

$$AIC_k = -2LnL + cN_k \quad (30)$$

dengan,

c : konstanta

N_k : jumlah parameter

$$N_k = (K - 1) + KR = KQ \quad (31)$$

dengan, R adalah jumlah variabel prediktor pada model regresi *inner model*

b. *Bayesian Information Criterion*

Persamaan rumus untuk menghitung BIC adalah sebagai berikut.

$$BIC_k = -2LnL + \ln IN_k \quad (32)$$

c. *Consisten Akaike's Information Criterion*

Persamaan rumus untuk menghitung CAIC adalah sebagai berikut.

$$CAIC_k = -2LnL + (\ln(I) + 1)N_k \quad (33)$$

d. *Normed Entropy*

Nilai *Normed Entrophy* memiliki nilai antara 0 sampai 1, semakin tinggi nilai tersebut, maka kualitas pemisahan semakin baik dan model dapat diinterpretasikan. Kriteria ini yang paling penting daripada yang lain (Ghozali dan Latan, 2014). Persamaan rumus untuk menghitung nilai *Normed Entrophy* adalah sebagai berikut (Hahn dkk., 2002).

$$EN_k = 1 - \frac{[\sum_{i=1}^I \sum_{k=1}^K -P_{ik} \ln(P_{ik})]}{I \ln K} \quad (34)$$

dengan,

EN_k : *Normal Entrophy*, ukuran relatif antara 0 – 1

P_{ik} : peluang observasi ke- i pada kelas ke- k

k : kelas atau segmen ($k = 1, 2, \dots, K$)

i : observasi ke- i ($i = 1, 2, \dots, I$)

1.5.4 Pendidikan di Indonesia

Pendidikan merupakan suatu proses menciptakan setiap individu yang berperan penting dalam masyarakat dengan cara mendidik, mengembangkan, serta menyediakan pengetahuan dan pemahaman yang lebih luas. Pendidikan dianggap sebagai proses seumur hidup dengan tujuan meningkatkan martabat manusia melalui penyediaan pengetahuan dan keterampilan yang dibutuhkan untuk meningkatkan produktivitas (Sutarman, 2007). Pemerintah masih berupaya keras untuk meningkatkan kualitas pendidikan. Salah satu fokus utama pemerintah dalam meningkatkan mutu pendidikan di Indonesia adalah memastikan bahwa kualitas pendidikan merata di seluruh wilayah sehingga setiap daerah memiliki standar kualitas yang setara dengan daerah lainnya.

Keberhasilan pemerataan pendidikan diukur berdasarkan pada keberagaman karakteristik tiap daerah di Indonesia. Beberapa daerah masih membutuhkan berbagai peningkatan dalam hal sumber daya manusia, fasilitas pendidikan, dan infrastruktur. Pendidikan adalah cermin dari kualitas sumber daya manusia di suatu daerah tertentu. Kemajuan pendidikan dapat berdampak pada perkembangan masyarakat di setiap wilayah. Semakin tinggi tingkat pendidikan yang diperoleh secara merata dan terjamin dalam jangka waktu yang lebih lama, semakin berkualitas pula sumber daya manusia yang dihasilkan (Alifah, 2021).

Menurut Badan Pusat Statistik (2022), pendidikan harus memiliki sarana dan prasarana yang baik agar anak memiliki kesempatan bersekolah sehingga dapat meningkatkan pengetahuan yang menjadikan kehidupan lebih sejahtera. Perubahan

dari masa pandemi ke pasca pandemi mengharuskan peralihan proses pembelajaran dari sistem pembelajaran di rumah kembali menjadi sistem pembelajaran tatap muka di sekolah setelah dua tahun tanpa kegiatan belajar mengajar secara langsung sehingga penting untuk menilai kesiapan sarana dan prasarana sekolah dalam mendukung pembelajaran serta menghadapi tantangan pada masa peralihan ini. Oleh karena itu, keberhasilan proses pembelajaran tidak terlepas dari kehadiran seorang guru. Kehadiran guru yang profesional sangat memengaruhi dalam proses belajar mengajar menjadi lebih efektif dan efisien.

Pada era digital saat ini, Teknologi Informasi dan Komunikasi (TIK) menjadi bagian yang tidak terpisahkan dari aktivitas seluruh penduduk, termasuk peserta didik di Indonesia. Kegiatan belajar mengajar diharuskan mengakses internet dan menggunakan komputer sebab dua hal tersebut paling banyak digunakan, baik di sekolah maupun di rumah. Internet dan komputer dapat memudahkan guru dengan murid dalam proses pembelajaran berlangsung. Upaya dalam meningkatkan suatu pendidikan ditunjukkan dengan partisipasi sekolah. Partisipasi sekolah diketahui berdasarkan kelompok usia dan jenjang pendidikan untuk mempermudah suatu pelaksanaan program pendidikan yang lebih tepat sasaran. Salah satu indikator partisipasi sekolah, yakni Angka Pendidikan Murni (APM) digunakan untuk melihat ketidakmerataan partisipasi pendidikan. Program pendidikan yang baik akan memiliki hasil dan capaian sesuai dengan melihat jumlah murid di setiap jenjang pendidikan, seperti SD, SMP, dan SMA.

BAB II METODOLOGI PENELITIAN

2.1 Sumber Data

Dalam penelitian ini, data diambil melalui buku resmi publikasi Badan Pusat Statistik (BPS), yakni buku Statistik Pendidikan 2022 dan *website* BPS. Buku dan *website* ini dapat diakses melalui laman www.bps.go.id. Jenis data tersebut merupakan jenis data sekunder dengan bentuk data *cross section* berdasarkan Lampiran 1. Data mengenai pendidikan di Indonesia yang berisikan 34 provinsi didapatkan dari hasil Survei Sosial Ekonomi Nasional (Susenas) Maret 2022. Survei Sosial Ekonomi Nasional (Susenas) adalah survei rumah tangga yang mencakup berbagai variabel dengan ruang lingkup sangat luas, termasuk seluruh aspek sosial dan ekonomi dari penduduk.

2.2 Variabel Penelitian

Penelitian ini menggunakan empat variabel laten, yaitu sarana dan prasarana pendidikan, kegiatan peserta didik, partisipasi sekolah serta hasil pendidikan, di mana variabel tersebut akan diukur dengan indikator yang ada. Variabel laten eksogen dalam penelitian ini adalah sarana dan prasarana pendidikan, sedangkan variabel laten endogen terdiri dari kegiatan peserta didik, partisipasi, hasil pendidikan. Variabel yang digunakan dalam penelitian ini disajikan secara rinci pada Tabel berikut.

Tabel 1. Variabel Penelitian

Variabel Laten	Variabel Indikator
X_1	Jumlah guru pada jenjang pendidikan Sekolah Dasar (SD) di bawah Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi menurut provinsi, 2022
X_2	Jumlah guru pada jenjang pendidikan Sekolah Menengah Pertama (SMP) di bawah Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi menurut provinsi, 2022
X_3	Jumlah guru pada jenjang pendidikan Sekolah Menengah Atas (SMA) di bawah Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi menurut provinsi, 2022

Tabel 1. Variabel Penelitian

Variabel Laten	Variabel Indikator	
Kegiatan Peserta Didik (η_1)	Y_1	Persentase peserta didik umur 5 – 24 tahun yang menggunakan komputer selama 3 bulan terakhir menurut provinsi dan jenis kelamin berdasarkan perkotaan dan pedesaan, 2022
	Y_2	Persentase peserta didik umur 5 – 24 tahun yang mengakses internet selama 3 bulan terakhir menurut provinsi dan jenis kelamin berdasarkan perkotaan dan pedesaan, 2022
	Y_3	Persentase peserta didik umur 5 – 24 tahun yang mengakses internet selama 3 bulan terakhir menurut provinsi dan jenjang pendidikan SMP/Sederajat berdasarkan perkotaan dan pedesaan, 2022
Partisipasi (η_2)	Y_4	Angka Partisipasi Murni (APM) menurut provinsi dan jenjang pendidikan untuk SMP/Sederajat, 2022
	Y_5	Angka Partisipasi Murni (APM) menurut provinsi dan jenjang pendidikan untuk SMA/Sederajat, 2022

Tabel 1. Variabel Penelitian

Variabel Laten		Variabel Indikator
Hasil dan Capaian Pendidikan (η_3)	Y_6	Jumlah murid pada jenjang pendidikan Sekolah Dasar (SD) di bawah Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi menurut provinsi, 2022
	Y_7	Jumlah murid pada jenjang pendidikan Sekolah Menengah Pertama (SMP) di bawah Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi menurut provinsi, 2022
	Y_8	Jumlah murid pada jenjang pendidikan Sekolah Menengah Atas (SMA) di bawah Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi menurut provinsi, 2022

2.3 Tahapan Analisis

Tahapan analisis yang dilakukan dalam penelitian ini dengan menggunakan bantuan *software* SmartPLS versi 3.2.9 dan R Studio adalah sebagai berikut.

1. Melakukan statistik deskriptif di setiap variabel manifes atau indikator pendidikan di Indonesia tahun 2022
2. Melakukan konseptualisasi model, yakni model struktural dan model pengukuran. Dalam penelitian ini, terdapat satu variabel laten eksogen sarana dan prasarana pendidikan (ξ_1), tiga variabel laten endogen kegiatan peserta didik (η_1), partisipasi (η_2), hasil dan capaian pendidikan (η_3) serta 11 variabel indikator, yaitu $X_1, X_2, X_3, Y_1, Y_2, Y_3, Y_4, Y_5, Y_6, Y_7, Y_8$.
3. Mengonstruksi diagram jalur yang menunjukkan pola hubungan sebab akibat antara variabel eksogen dan variabel endogen
4. Mengonversi diagram jalur ke dalam persamaan. Konversi ini dilakukan terhadap persamaan model struktural dan model pengukuran
5. Mengestimasi parameter model dengan melakukan tiga tahap iterasi, yaitu estimasi bobot, estimasi jalur, serta estimasi rata-rata dan parameter lokasi
6. Mengevaluasi model struktural dan model pengukuran untuk mendapatkan parameter yang signifikan. Apabila model telah valid dan reliabel, maka melanjutkan analisis selanjutnya. Jika tidak, kembali melakukan konstruksi diagram jalur

7. Pengujian hipotesis menggunakan *resampling bootsrap*
8. Menggunakan nilai skor faktor dari variabel laten pada model struktural hasil estimasi jalur dalam FIMIX-PLS
9. Melakukan pembagian segmen sebanyak enam kali pada FIMIX-PLS
10. Mengidentifikasi hasil dari jumlah segmen yang sesuai
11. Menganalisis dan mengestimasi jumlah segmen terbaik dari masing-masing segmen berdasarkan nilai kriteria statistik pada FIMIX-PLS, yaitu AIC, BIC, CAIC, dan terutama nilai *Normed Entrophy* (EN)
12. Mengelompokkan wilayah sesuai dengan segmen yang terbentuk dengan mengetahui nilai probabilitas yang didapatkan
13. Menginterpretasi hasil yang didapatkan.