

**ANALISIS *RESPONSE BASED UNITS SEGMENTATION PARTIAL
LEAST SQUARE* (REBUS PLS) PADA DATA EKONOMI
TERHADAP INDEKS PEMBANGUNAN MANUSIA**

**ANALYSIS OF *RESPONSE BASED UNITS SEGMENTATION
PARTIAL LEAST SQUARE* (REBUS PLS) ON ECONOMIC DATA
ON THE HUMAN DEVELOPMENT INDEX**

ESRA ROMBEALLO



**PROGRAM STUDI MAGISTER STATISTIKA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN
MAKASSAR
2023**

**ANALISIS *RESPONSE BASED UNITS SEGMENTATION PARTIAL*
LEAST SQUARE (REBUS PLS) PADA DATA EKONOMI
TERHADAP INDEKS PEMBANGUNAN MANUSIA**

Tesis

Sebagai salah satu syarat untuk mencapai gelar magister

Program Studi Magister Statistika

Disusun dan diajukan oleh

ESRA ROMBEALLO

H062212006

Kepada

**PROGRAM STUDI MAGISTER STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN
MAKASSAR
2023**

TESIS**ANALISIS RESPONSE BASED UNITS SEGMENTATION PARTIAL
LEAST SQUARE (REBUS PLS) PADA DATA EKONOMI TERHADAP
INDEKS PEMBANGUNAN MANUSIA****ESRA ROMBEALLO****H062212006**

Telah dipertahankan di hadapan Panitia Ujian yang dibentuk dalam
rangka Penyelesaian Program Studi Magister Statistika Fakultas
Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin
pada tanggal 7 Agustus 2023
dan dinyatakan telah memenuhi syarat kelulusan

Menyetujui

Pembimbing Utama

Pembimbing Pendamping


Dr. Dr. Georgina M. Tinungki, M.Si

NIP. 19620926 198702 2 001


Dr. Anna Islamiyati, S.Si., M.Si

NIP. 19770808 200501 2 002

Ketua Program Studi
Magister StatistikaDekan Fakultas Matematika dan
Ilmu Pengetahuan Alam
Universitas Hasanuddin
Dr. Dr. Georgina M. Tinungki, M.Si

NIP. 19620926 198702 2 001


Dr. Eng. Amiruddin, M.Si

NIP. 19720515 199702 1 002

PERNYATAAN KEASLIAN TESIS DAN PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa, tesis berjudul *Analisis Response Based Units Segmentation Partial Least Square (REBUS PLS) pada Data Ekonomi terhadap Indeks Pembangunan Manusia* adalah benar karya saya dengan arahan dari komisi pembimbing (Dr. Dr. Georgina Maria Tinungki, M.Si dan Dr. Anna Islamiyati, S.Si., 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 tesis ini. Sebagian dari isi tesis ini telah dipublikasikan di Jurnal (International Research Journal of Advanced Engineering and Science (IRJAES) ISSN: 2455-9024 Vol. 8, Issue 3, Juli - 2023, Pages: 18-22) sebagai artikel dengan judul "*Response-Based Units Segmentation Partial Least Square (REBUS PLS) Analysis on Human Development Data in Indonesia*".

Dengan ini saya melimpahkan hak cipta dari karya tulis saya berupa tesis ini kepada Universitas Hasanuddin.

Makassar, 7 Agustus 2023

Yang menyatakan,




Esra Rombeallo
NIM. H062212006

UCAPAN TERIMA KASIH

Puji syukur kehadiran Tuhan Yang Maha Esa atas anugerah dan kasih karunia-Nya sehingga penulis dapat menyusun dan menyelesaikan tesis ini. Penulis menyadari sepenuhnya bahwa apa yang dikemukakan dalam tesis ini masih jauh dari kesempurnaan yang merupakan sebagai akibat dari keterbatasan kemampuan serta berbagai kesulitan yang penulis hadapi dalam penyusunan tesis ini.

Penulis memanjatkan doa kepada Tuhan Yang Maha Esa agar menganugerahkan rahmat-Nya kepada pihak yang banyak membantu dalam penyelesaian tesis ini. Penulis juga percaya tesis ini dapat selesai bukan hanya dengan kekuatan pikiran penulis semata akan tetapi karena bantuan dari berbagai pihak juga, baik selama proses perkuliahan bahkan sampai proses pengerjaan tesis di Program Magister Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Hasanuddin. Namun demikian, penulis dengan senang hati menerima kritik dan saran yang bersifat membangun dari pembaca karya tulis ini demi sempurnanya tesis ini.

Terima kasih yang tak terhingga kepada kedua orang tua tercinta dan saudaraku atas doa yang tak pernah putus, dukungan serta segala kebaikan mereka yang sampai kapan pun takkan pernah bisa terbalaskan atas kasih sayang yang tiada henti dalam penyelesaian tesis ini. Selanjutnya, saya ingin menyampaikan juga rasa hormat dan terima kasih kepada:

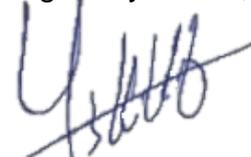
1. **Prof. Dr. Ir. Jamaluddin Jompa, M.Sc.** selaku Rektor Universitas Hasanuddin.
2. **Dr. Eng. Amiruddin, M.Si.** selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam beserta seluruh jajarannya.
3. **Dr. Anna Islamiyati, S.Si., M.Si.** selaku Ketua Departemen Statistika yang sekaligus pembimbing pertama yang selalu memberikan ilmu, dukungan, dan motivasi serta kemudahan kepada penulis dalam berbagai hal selama menjalani Pendidikan di Departemen Statistika.

4. **Dr. Dr. Georgina Maria Tinungki, M.Si.** selaku Ketua Program Studi Magister Statistika sekaligus sebagai pembimbing utama yang telah bersabar dan bersedia meluangkan waktu untuk membimbing penulis dan memberikan masukan dalam penyelesaian tesis ini.
5. **Sri Astuti Thamrin, S.Si., M.Stat., Ph.D., Dr. Erna Tri Herdiani, S.Si., M.Si.** dan **Dr. Nirwan Ilyas, M.Si.** selaku penguji penulis yang telah bersedia memberikan masukan-masukan dan arahan dalam penyusunan tesis.
6. Teristimewa kepada kedua orang tua tercinta Ayah (**Alm. Pelipus Bunga**) dan Ibu (**Damaris Batan**) yang tidak pernah lelah mendidik, mendoakan, memberi dukungan, bahkan tidak pernah mengeluh senantiasa bekerja keras untuk membiayai kuliahku dan memenuhi seluruh kebutuhanku selama menempuh studi.
7. Saudaraku (**Zet Nataniel, Robertus, Sarlota, Risna, Yakobus Rombeallo dan Gelora Tandirerung**) yang selalu mendukung, membiayai dan mendoakan selama studi.
8. Saudara tak sedarah mahasiswa program magister statistika angkatan 2022 diangkat kelima magister statistika.

Semoga Tuhan Yang Maha Esa memberikan berkat, kasih dan hikmat-Nya atas segala kebaikan yang telah diberikan kepada penulis dan semoga penulisan tesis ini bermanfaat bagi perkembangan ilmu pengetahuan dan teknologi, khususnya dalam dunia statistika dan data sains.

Makassar, 7 Agustus 2023

Yang menyatakan,



Esra Rombeallo
NIM. H062212006

ABSTRAK

ESRA ROMBEALLO. **Analisis *Response Based Units Segmentation Partial Least Square (REBUS PLS)* pada Data Ekonomi terhadap Indeks Pembangunan Manusia** (dibimbing oleh Dr. Dr. Georgina Maria Tinungki, M.Si dan Dr. Anna Islamiyati, S.Si., M.Si).

Partial Least Square (PLS) merupakan bagian dari *Structural Equation Modeling* yang dapat digunakan untuk berbagai skala pengukuran data, ukuran sampel yang kecil, dan tidak memerlukan asumsi distribusi. Dalam PLS yang umumnya diasumsikan bahwa data bersifat homogen. Asumsi ini tidak rasional karena data dapat berasal dari beberapa segmen. Permasalahan ini dinamakan dengan heterogenitas. Metode yang dapat mengatasi masalah tersebut yaitu *Response Based Units Segmentation Partial Least Square (REBUS PLS)*. *Response Based Units Segmentation Partial Least Square* merupakan metode yang dapat mensegmentasi unit observasi dan mendeteksi heterogenitas pada persamaan struktural maupun persamaan pengukuran. Salah satu bidang yang sering menggunakan banyak variabel dalam analisis *Response Based Units Segmentation* yang merupakan bagian dari *Partial Least Squares (PLS)* adalah ekonomi dan pembangunan manusia diantaranya sumber daya manusia, kesehatan dan kemiskinan. Penelitian ini bertujuan memperoleh hasil dan model segmentasi pada data ekonomi terhadap indeks pembangunan manusia dengan REBUS PLS. Variabel yang digunakan yaitu 4 variabel laten dan 14 variabel manifest. Hasil segmentasi pada penelitian ini yaitu membentuk 2 model lokal yang terdiri dari 9 provinsi dan 25 provinsi. Model lokal 1 yaitu $\eta_1 = 0,841 \xi$, $\eta_2 = -0,367 \xi + 1,229 \eta_1$, dan $\eta_3 = -0,248 \xi + 0,193 \eta_1 - 0,939 \eta_2$ sedangkan model lokal 2 yaitu $\eta_1 = 0,364 \xi$, $\eta_2 = -0,170 \xi - 0,727 \eta_1$, dan $\eta_3 = -0,286 \xi + 0,263 \eta_1 - 0,414 \eta_2$. Kesimpulan bahwa REBUS PLS mampu mendeteksi heterogenitas pada model SEM PLS dengan nilai GoF untuk model global sebesar 0,601, model lokal 1 sebesar 0,856, dan model lokal 2 sebesar 0,432 yang masuk dalam kategori GoF besar. Nilai GoF yang tinggi juga menunjukkan model struktural dan pengukuran yang lebih baik.

Kata Kunci: Ekonomi, Indeks Pembangunan Manusia, *Partial Least Square*, *Response Based Units Segmentation*, *Structural Equation Modeling*.

ABSTRACT

ESRA ROMBEALLO. **Analysis of Response Based Units Segmentation Partial Least Square (REBUS PLS) on Economic Data on the Human Development Index** (supervised by Dr. Dr. Georgina Maria Tinungki, M.Si and Dr. Anna Islamiyati, S.Si., M.Si).

Partial Least Square (PLS) is part of *Structural Equation Modeling*, which can be used for various data measurement scales, small sample sizes, and does not require distributional assumptions. In PLS, it is generally assumed that the data is homogeneous. This assumption is irrational because data can come from several segments. This problem is called heterogeneity. The method to overcome this problem is *Response Based Units Segmentation Partial Least Square* (REBUS PLS). *Response Based Units Segmentation Partial Least Square* is a method that can segment observation units and detect heterogeneity in structural equations and measurement equations. One area that often uses many variables in *Response Based Units Segmentation* analysis which is part of *Partial Least Squares* (PLS) is economics and human development, including human resources, health and poverty. This study aims to obtain results and segmentation models on economic data on the human development index with REBUS PLS. The variables used are four latent variables and 14 manifest variables. The segmentation results in this study formed two local models consisting of 9 provinces and 25 provinces. Local model 1 is $\eta_1 = 0,841 \xi$, $\eta_2 = -0,367 \xi + 1,229 \eta_1$, and $\eta_3 = -0,248 \xi + 0,193 \eta_1 - 0,939 \eta_2$ while local model 2 is $\eta_1 = 0,364\xi$, $\eta_2 = -0,170\xi - 0,727 \eta_1$, and $\eta_3 = -0,286 \xi + 0,263 \eta_1 - 0,414 \eta_2$. The conclusion is that REBUS PLS can detect heterogeneity in the PLS SEM model with a GoF value for the global model of 0,601, local model 1 of 0,856, and local model 2 of 0,432, which fall into the large GoF category. A high GoF value also indicates a better structural and measurement model.

Keywords: Economy, Human Development Index, *Partial Least Square*, *Response Based Units Segmentation*, *Structural Equation Modeling*.

DAFTAR ISI

HALAMAN JUDUL	i
PERNYATAAN PENGAJUAN	ii
LEMBAR PENGESAHAN	iii
PERNYATAAN KEASLIAN	iv
UCAPAN TERIMA KASIH	v
ABSTRAK	vi
ABSTRACT	vii
DAFTAR ISI	viii
DAFTAR GAMBAR	x
DAFTAR TABEL	xi
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Batasan Masalah	4
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian.....	5
BAB II TINJAUAN PUSTAKA	6
2.1 <i>Partial Least Square Structural Equation Modeling</i> (PLS SEM)	6
2.1.1 Spesifikasi Model	6
2.1.2 Estimasi Parameter.....	8
2.1.3 Evaluasi Model.....	8
2.1.4 <i>Goodness Of Fit</i>	10
2.1.5 Uji Signifikansi.....	11
2.2 <i>Response Based Units Segmentation Partial Least Square</i> (REBUS PLS)	12
2.3 Kerangka Konseptual	15
BAB III METODOLOGI PENELITIAN	16
3.1 Data Penelitian	16
3.2 Variabel Penelitian	16
3.3 Metode Analisis Data	17
BAB IV HASIL DAN PEMBAHASAN	19
4.1 Konseptualisasi Model Pengukuran dan Model Struktural	19
4.2 Konstruksi Diagram Jalur	19
4.3 Mengkonversi Diagram Jalur ke Sistem Persamaan	20

4.4 Estimasi Parameter	21
4.4.1 Estimasi Bobot (<i>Weight Estimate</i>)	21
4.4.2 Estimasi Jalur.....	23
4.5 Evaluasi Model	26
4.5.1 Evaluasi Model Pengukuran.....	26
4.5.2 Evaluasi Model Struktural.....	30
4.6 Goodness of Fit.....	31
4.7 Pengujian Signifikansi	31
4.7.1 Pengujian Signifikansi Model Pengukuran	31
4.7.2 Pengujian Signifikansi Model Struktural	32
4.8 Penerapan <i>Response Based Units Segmentations</i> (REBUS) PLS	33
4.8.1 Menghitung <i>communality</i> dan residual struktural dari tiap unit model global.....	33
4.8.2 Membentuk <i>hierarchical cluster analysis</i> berdasarkan <i>communality</i> dan residual struktural	33
4.8.3 Menghitung nilai CM (<i>Closeness Measure</i>) pada model lokal dan mengelompokkan berdasarkan nilai CM	34
4.8.4 Mengestimasi model pada setiap segmen (model lokal)	36
4.8.5 Interpretasi Hasil REBUS PLS	36
4.9 Model <i>Response Based Units Segmentations</i> (REBUS) PLS.....	39
4.9.1 Interpretasi Model Lokal 1	39
4.9.2 Interpretasi Model Lokal 2	40
BAB V KESIMPULAN DAN SARAN.....	41
5.1 Kesimpulan	41
5.2 Saran	41
DAFTAR PUSTAKA.....	42
LAMPIRAN.....	42

DAFTAR GAMBAR

Gambar 3.1 Kerangka Variabel Penelitian	18
Gambar 3.2 Diagram Alir Penelitian	20
Gambar 4.1 Diagram Jalur dengan Aplikasi SmartPLS	22
Gambar 4.2 Dendogram dari Outer dan Inner Residual	37

DAFTAR TABEL

Tabel 4.1 <i>Convergent validity</i> ξ dengan indikatornya	29
Tabel 4.2 <i>Convergent validity</i> η_1 dengan indikatornya	29
Tabel 4.3 <i>Convergent validity</i> η_2 dengan indikatornya	29
Tabel 4.4 <i>Convergent validity</i> η_3 dengan indikatornya	29
Tabel 4.5 <i>Discriminant validity</i> ξ dengan indikatornya	30
Tabel 4.6 <i>Discriminant validity</i> η_1 dengan indikatornya	30
Tabel 4.6 <i>Discriminant validity</i> η_1 dengan indikatornya	30
Tabel 4.7 <i>Discriminant validity</i> η_2 dengan indikatornya	31
Tabel 4.8 <i>Discriminant validity</i> η_3 dengan indikatornya	31
Tabel 4.9 Nilai <i>R-square</i>	34
Tabel 4.10 Nilai <i>R Square</i> dan <i>Communality</i>	34
Tabel 4.11 Nilai <i>T-statistics</i> model pengukuran.....	35
Tabel 4.12 Nilai <i>T-statistics</i> model struktural.....	36
Tabel 4.13 Segmen provinsi berdasarkan analisis <i>cluster</i> metode Ward's ...	38
Tabel 4.14 Nilai <i>closeness measure</i> (CM) untuk tiap segmen.....	38
Tabel 4.15 Segmen provinsi REBUS PLS berdasarkan nilai Closeness Measure.....	39
Tabel 4.16 Nilai koefisien jalur model global dan model lokal	40
Tabel 4.17 <i>Standardized Loadings</i> model global dan 2 model lokal.....	41
Tabel 4.18 Nilai Q^2 dan <i>GoF</i> model global dan 2 model lokal	42

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pada teknik regresi, model penelitiannya dibangun berdasarkan satu variabel dependen dan beberapa variabel independen. Ketika model penelitian menggunakan lebih dari satu variabel dependen, dibutuhkan metode analisis lain. Metode yang bisa menyelesaikan permasalahan tanpa harus membuat beberapa persamaan regresi karena menganalisis secara terpisah merupakan hal yang kurang tepat (Hamid & Anwar, 2019). Metode ini, disebut *Structural Equation Modeling* (SEM), memungkinkan peneliti untuk secara bersamaan memodelkan dan memperkirakan hubungan kompleks antara beberapa variabel dependen dan independen (Akter et al., 2017). Konsep yang dipertimbangkan biasanya tidak dapat diamati dan diukur secara tidak langsung dengan beberapa indikator (Hair Jr. et al., 2021). Salah satu metode yang bisa digunakan dalam menganalisis model persamaan jalur adalah *Structural Equation Modeling* (SEM). *Structural Equation Modeling* (SEM) merupakan salah satu analisis statistik yang dapat digunakan untuk menyelesaikan permasalahan yang *complicated* yaitu mampu melakukan estimasi hubungan antar variabel yang bersifat *multiple relationship* dengan output berupa model pengukuran dari sejumlah indikator dan model struktural yang tersusun dari sejumlah variabel laten (Anggita et al., 2019). SEM menghubungkan dua atau lebih konsep kompleks variabel laten, masing-masing diukur melalui sejumlah indikator yang dapat diamati yang dikenal sebagai variabel manifest (Vinzi et al., 2008).

Menurut Fornell dan Bookstein dalam (Hamid & Anwar, 2019) terdapat dua jenis SEM, yaitu *Covariance-Based Structural Equation Modeling* (CB SEM) dan *Partial Least Squares Structural Equation Modeling* (PLS SEM). CB SEM menuntut basis teori yang kuat, memenuhi berbagai asumsi parametrik, dan memenuhi uji kelayakan model (*Goodness of Fit*). PLS SEM tidak memerlukan asumsi yang ketat, baik mengenai sebaran dari peubah pengamatan maupun dari ukuran contoh yang tidak harus besar (Akter et al., 2017). Metode PLS SEM memungkinkan bagi peneliti untuk menggunakan indikator-indikator yang bersifat reflektif (peubah laten menjelaskan peubah penjas) ataupun formatif (peubah penjas menjelaskan peubah laten) (Hamid & Anwar, 2019). PLS SEM telah menjadi metode yang populer memperkirakan model jalur (kompleks) dengan variabel laten dan hubungannya (Sarstedt et al., 2017).

Partial Least Square (PLS) merupakan metode analisis yang powerful dan fleksibel karena sifatnya yang dapat digunakan untuk berbagai skala pengukuran data, dapat digunakan untuk ukuran sampel yang kecil, dan tidak mengharuskan berdistribusi normal multivariat (Wold dalam Patih & Halistin, 2020). Penelitian terkait PLS SEM diantaranya Surya et al., (2020) melakukan penelitian tentang analisis faktor yang memengaruhi brand loyalti gojek dengan efek mediator menggunakan PLS SEM. Mulyani (2021) menggunakan PLS SEM untuk menganalisis pentingnya kualitas kehidupan kerja, stres kerja, dan kepuasan kerja dalam meningkatkan komitmen organisasi guru. Penelitian lain dilakukan (Ningsi & Agustina, 2018) menggunakan PLS SEM untuk menganalisis kepuasan pelanggan atas kualitas produk dan pelayanan.

Dalam PLS SEM yang umumnya diasumsikan bahwa data bersifat homogen atau responden berasal dari karakteristik yang sama . Asumsi ini tidak rasional karena data dapat berasal dari beberapa segmen atau unit yang berbeda dari responden yang dikumpulkan (Mehmetoglu, 2011). Menurut Reagan (2016) pada penelitian PLS SEM melibatkan banyak variabel dan indikator, sehingga pada penggunaan data sekunder dan di ambil dari berbagai sumber yang berbeda maka hal ini dapat berpengaruh pada hasil analisis. Oleh karena hal tersebut, jika tanpa memperhatikan homogenitas tersebut maka akan menghasilkan analisis yang bias (Becker et al., 2013). Permasalahan ini dinamakan dengan heterogenitas yang sangat berpengaruh terhadap hasil analisis dan kesimpulan. Ada 2 jenis heterogenitas yaitu heterogenitas teramati (*observed heterogeneity*) dan heterogenitas tidak teramati (*unobserved heterogeneity*) (Mehmetoglu, 2011). Heterogenitas teramati didasarkan pada informasi yang telah diketahui sebelumnya tentang jumlah segmen. Misal berdasarkan demografik seperti jenis kelamin, pendapatan dan umur. Sedangkan heterogenitas tidak teramati, tidak diketahui informasi karakteristik mengenai banyaknya segmen, sehingga pengelompokkan responden berdasarkan performance dari model tersebut dan dinamakan *response-based classification* (Ikhsan, 2021). Mengabaikan heterogenitas tidak teramati akan membuat estimasi parameter menjadi bias, menimbulkan kesalahan inferensia tipe I dan tipe II, serta menghasilkan kesimpulan yang tidak valid (Jedidi et al., 1997).

Metode untuk heterogenitas tidak teramati ini dapat dilakukan dengan *Response Based Units Segmentation Partial Least Square* (REBUS PLS) yang diperkenalkan oleh Trinchera (2007) dan Vinzi et al. (2008). REBUS-PLS dapat

mensegmentasi unit observasi sekaligus mengestimasi parameter dari masing-masing model lokal pada segmen yang terbentuk (Trinchera, 2007). Dengan menggunakan REBUS PLS, unit observasi dikelompokkan (clustering) berdasarkan kesamaan performance di dalam model dan sekaligus mengestimasi parameter dari masing-masing kelompok yang terbentuk. REBUS PLS memiliki kelebihan-kelebihan misalnya: pada REBUS PLS mampu mendeteksi heterogenitas pada persamaan structural maupun persamaan pengukuran. Kelebihan lainnya adalah REBUS PLS tidak memerlukan asumsi distribusi baik pada variabel manifest maupun variabel laten (Trinchera, 2007).

Beberapa penelitian menggunakan REBUS PLS yang pernah dilakukan, diantaranya (Vinzi et al., 2008) dalam bidang sosial untuk memodelkan segmen negara-negara terkait kualitas hidup dengan menggunakan REBUS PLS. Vinzi et al. (2010) memodelkan segmen dalam bidang marketing. Assaker et al., (2018) meneliti heterogenitas menggunakan REBUS PLS studi kasus tentang sumber daya manusia dan kinerja perusahaan di restoran kelas atas. Fauzi et al. (2018), *clustering partial lest square* pada indeks prestasi dosen (LAI) berdasarkan persepsi mahasiswa UPN "Veteran" Surabaya menggunakan REBUS PLS. (Pratiwi et al., 2020) menggunakan REBUS PLS untuk menganalisis dan pengelompokan wilayah di provinsi Jawa Tengah. Salsabila et al. (2021) menggunakan REBUS PLS untuk mengelompokkan prakiraan curah hujan bulanan. Penelitian-penelitian tersebut membandingkan nilai r kuadrat total (r -square) dan *goodness of fit* antara model global PLS SEM dengan model lokal yang dihasilkan oleh REBUS PLS.

Pembangunan manusia merupakan model pembangunan yang menurut *United Nations Development Programme* (UNDP) ditujukan untuk memperluas pilihan-pilihan yang dapat ditumbuhkan melalui upaya pemberdayaan penduduk. Pembangunan manusia menggunakan pengukuran yang sudah dikenalkan oleh UNDP pada tahun 1990, yaitu Indeks Pembangunan Manusia (IPM) (Suntono & A'laa, 2021). Pertumbuhan ekonomi ialah proses terbangunnya perekonomian yang semakin meningkat dari barang ataupun produk dalam suatu negara. Terlepas dari mana yang lebih penting, yang pasti pertumbuhan ekonomi sangat penting dan dibutuhkan. Sebab, tanpa pertumbuhan tidak akan terjadi peningkatan kesejahteraan, kesempatan kerja, produktivitas dan distribusi pendapatan (Senewe et al., 2021).

Pada tahun 2017, indeks pembangunan manusia Indonesia mengalami peningkatan dibanding tahun sebelumnya yang mencapai 70,81 atau tumbuh sebesar 0,90 persen (Statistics, 2020). Berdasarkan hasil perhitungan IPM menggunakan metode baru, angka IPM Indonesia Tahun 2020 sebesar 71,94 sedangkan nilai IPM pada tahun 2019 sebesar 71,92. Pada tahun 2020, IPM tertinggi untuk level provinsi dicapai oleh DKI Jakarta sebesar 80,77. Sedangkan untuk capaian IPM terendah adalah provinsi Papua sebesar 60,44 (Statistics, 2020). Disparitas capaian indeks pembangunan manusia pada provinsi di Indonesia tersebut menimbulkan pertanyaan apakah ada pengaruh ekonomi dengan capaian indeks pembangunan manusia khususnya bila indeks pembangunan manusia bisa dijabarkan berdasarkan dimensi penyusunnya.

Berdasarkan uraian tersebut, dalam penelitian ini penulis ingin menganalisis menggunakan REBUS PLS dengan judul penelitian “**Analisis Response Based Units Segmentation Partial Least Square (REBUS PLS) pada Data Ekonomi terhadap Indeks Pembangunan Manusia**”.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, rumusan masalah yang akan didapatkan solusi pada penelitian ini adalah:

1. Bagaimana hasil segmentasi REBUS PLS?
2. Bagaimana model segmentasi pada data ekonomi terhadap indeks pembangunan manusia dengan REBUS PLS?

1.3 Batasan Masalah

Agar tidak menimbulkan penafsiran yang lebih luas, maka dalam analisis penelitian ini permasalahannya dibatasi pada :

1. Estimasi parameter menggunakan *ordinal least square*.
2. Jumlah segmen REBUS PLS adalah 2 segmen.
3. Data ekonomi terhadap indeks pembangunan manusia di 34 provinsi di Indonesia pada tahun 2020 menggunakan REBUS PLS.

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah di atas, maka penelitian ini bertujuan untuk:

1. Memperoleh hasil segmentasi REBUS PLS.
2. Membentuk model segmentasi pada data ekonomi terhadap indeks pembangunan manusia dengan REBUS PLS.

1.5 Manfaat Penelitian

Hasil penelitian ini diharapkan dapat berguna untuk menambah wawasan keilmuan dalam menerapkan metode *Response Based Units Segmentation Partial Least Square* (REBUS PLS).

BAB II

TINJAUAN PUSTAKA

2.1 *Partial Least Square Structural Equation Modeling (PLS SEM)*

Metode *Partial Least Square Structural Equation Modeling (PLS SEM)* merupakan suatu metode yang memaksimumkan ragam dalam peubah tak bebas yang dijelaskan oleh peubah bebas (Hair et al., 2012). PLS SEM memiliki dua jenis peubah (variabel) yaitu variabel laten dan variabel manifest. Variabel laten biasa disebut sebagai variabel abstrak atau variabel yang tidak dapat diukur, contohnya adalah perilaku orang, perasaan dan motivasi (Reagan, 2016). Variabel laten memiliki dua jenis yaitu variabel laten eksogen (ξ) dan variabel laten endogen (η). Variabel laten eksogen dapat disebut sebagai variabel bebas dalam suatu persamaan sedangkan variabel laten endogen merupakan variabel terikat pada suatu persamaan. Sedangkan variabel manifest merupakan variabel yang dapat diamati atau dapat diukur secara empiris (Sarstedt et al., 2017).

2.1.1 *Spesifikasi Model*

Spesifikasi lengkap dari model PLS SEM yang terdiri dari model struktural dan model pengukuran yaitu:

a. *Inner Model*

Inner model atau model struktural adalah model yang menggambarkan hubungan - hubungan yang ada di antara variabel - variabel laten (Sanchez, 2013). Persamaan *inner model* dapat ditulis sebagai berikut (Kartikasari, 2016):

$$\eta_{mx1} = \beta_{mxn}\eta_{nx1} + \gamma_{mxn}\xi_{nx1} + \zeta_{mx1} \quad (2.1)$$

atau

$$\begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_m \end{bmatrix} = \begin{bmatrix} \beta_{11} & \beta_{12} & \dots & \beta_{1n} \\ \beta_{21} & \beta_{22} & \dots & \beta_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{m1} & \beta_{m2} & \dots & \beta_{mn} \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{bmatrix} + \begin{bmatrix} \gamma_{11} & \gamma_{12} & \dots & \gamma_{1n} \\ \gamma_{21} & \gamma_{22} & \dots & \gamma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{m1} & \gamma_{m2} & \dots & \gamma_{mn} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{bmatrix} + \begin{bmatrix} \zeta_1 \\ \zeta_2 \\ \vdots \\ \zeta_m \end{bmatrix}$$

Dimana :

η = vektor variabel laten endogen (dependen), berdimensi $mx1$

β = matriks koefisien jalur variabel laten endogen, berdimensi mxn

Γ = matriks koefisien jalur yang menghubungkan variabel laten endogen dan eksogen, berdimensi mxn

ξ = vektor variabel laten eksogen, berdimensi mxn

ζ = vektor residual, berdimensi $mx1$

b. *Outer Model*

Outer model atau model pengukuran menspesifikasi hubungan antara suatu peubah laten dan indikator (Sanchez, 2013). Hubungan yang dibentuk salah satunya adalah hubungan refleksif, memiliki arah hubungan kausalitas dari peubah laten ke indikator dengan antar indikator diharapkan saling berkorelasi.

Outer model pada hubungan refleksif adalah (Kartikasari, 2016):

$$X_{qx1} = \Lambda_{X_{qxn}} \xi_{nx1} + \delta_{qx1} \quad (2.2)$$

Atau

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_q \end{bmatrix} = \begin{bmatrix} \lambda_{x11} & \lambda_{x12} & \dots & \lambda_{x1n} \\ \lambda_{x21} & \lambda_{x22} & \dots & \lambda_{x2n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{xq1} & \lambda_{xq2} & \dots & \lambda_{xqn} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{bmatrix} + \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_q \end{bmatrix}$$

$$Y_{px1} = \Lambda_{Y_{pxm}} \eta_{mx1} + \varepsilon_{px1} \quad (2.3)$$

Atau

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{bmatrix} = \begin{bmatrix} \lambda_{y11} & \lambda_{y12} & \dots & \lambda_{y1m} \\ \lambda_{y21} & \lambda_{y22} & \dots & \lambda_{y2m} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{yp1} & \lambda_{yp2} & \dots & \lambda_{ypm} \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_m \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_p \end{bmatrix}$$

Dimana :

Y : vektor variabel indikator untuk variabel laten endogen, berdimensi $px1$

Λ_Y : matriks koefisien Y terhadap η (*loading factor*), berdimensi pxm

η : vektor variabel laten endogen, berdimensi $mx1$

ε : vektor galat pengukuran Y , berdimensi $px1$

X : vektor variabel indikator untuk variabel laten eksogen, berdimensi $qx1$

Λ_X : matriks koefisien X terhadap ξ (*loading factor*), berdimensi qxn

ξ : vektor variabel laten eksogen, berdimensi $nx1$

δ : vektor galat pengukuran X , berdimensi $qx1$

Persamaan (2.1),(2.2) dan (2.3) diasumsikan bahwa

- $\zeta, \varepsilon, \delta$ tidak saling berkorelasi
- $E(\zeta) = 0, E(\varepsilon) = 0, \text{ dan } E(\delta) = 0$
- $cov(\xi) = \Phi, cov(\zeta) = \psi, cov(\varepsilon) = \theta_\varepsilon \text{ dan } cov(\delta) = \theta_\delta$
- $E(\xi) = 0 \text{ dan } E(\eta) = 0$
- δ (Delta) tidak berkorelasi dengan ξ (Ksi)
- ε (Epsilon) tidak berkorelasi dengan η (Eta)
- ζ (Zeta) tidak berkorelasi dengan ξ

- h. β mempunyai nilai nol dalam diagonal utamanya dan di bawah diagonal utamanya atau di atas diagonal utamanya
- i. $I_m - \beta$ merupakan matriks *nonsingular*

2.1.2 Estimasi Parameter

Estimasi parameter dalam PLS SEM menggunakan metode kuadrat terkecil (*Ordinary Least Square*) maka persoalan identifikasi model tidak menjadi masalah untuk model rekursif (model yang mempunyai satu arah kausalitas). Proses perhitungan dilakukan dengan cara iterasi dan akan berhenti saat sudah menemukan titik konvergen (Muliati & Handoyo, 2019). Dalam setiap prosedur iterasi, misalkan $S = 1, 2, 3, \dots$, konvergensi diperiksa dengan membandingkan *outer weight* pada setiap langkah S terhadap *outer weight* pada langkah $S-1$ (Sanchez, 2013). Pendugaan parameter pada model PLS meliputi (Henseler et al., 2009):

- a. Estimasi bobot (*Weight estimate*), untuk mendapatkan skor (*score factor*) variabel laten
- b. Estimasi jalur (*Path estimate*)
Estimasi jalur menghubungkan antar variabel laten (koefisien jalur) yaitu koefisien beta (β) dan gamma (γ) dan antara variabel laten dengan indikatornya yaitu estimasi *loading factor* yang merupakan koefisien *outer model* yaitu lambda (λ).

2.1.3 Evaluasi Model

Evaluasi model dalam PLS SEM (Hamid & Anwar, 2019) terdiri dari :

- a. Evaluasi *Outer Model*

Evaluasi *outer model*, dimulai dari tahapan uji validitas konstruk yang terdiri dari validitas konvergen, dengan memperhatikan nilai *loading factor*, nilai AVE, dan validitas diskriminan yang ditunjukkan oleh nilai *cross loading*. Kemudian, tahap kedua, yaitu pengujian reliabilitas ditunjukkan dengan nilai *composite reliability*.

- 1. Uji Validitas Konstruk

Dalam SEM-PLS, tahapan uji validitas konstruk terdiri dari dua tahap, yaitu:

- a) Validitas Konvergen

Tahap ini memiliki dua kriteria nilai yang akan dievaluasi, yaitu nilai *loading factor* dan nilai *average variance extracted* (AVE). Nilai *loading factor*, indikator dapat dinyatakan valid jika nilai *loading factor* $> 0,5$. Nilai *average variance extracted* (AVE), ukuran *convergent validity* yang baik memiliki AVE $> 0,5$.

b) Validitas Diskriminan

Pada tahapan ini ada dua kriteria nilai yang akan dievaluasi, yaitu nilai *cross loading* dan nilai korelasi antarkonstruksi laten,

- 1) Nilai *cross loading*, jika korelasi antara variabel laten dengan setiap indikatornya (variabel manifes) lebih besar daripada korelasi dengan variabel laten lainnya, maka variabel laten tersebut dapat dikatakan memprediksi indikatornya lebih baik daripada variabel laten lainnya (*cross loading* > 0,7).
- 2) Nilai korelasi antarkonstruksi laten, Apabila nilai akar kuadrat AVE lebih tinggi daripada nilai korelasi di antara variabel laten, maka *discriminant validity* dapat dianggap tercapai atau *discriminant validity* dapat dikatakan tercapai apabila nilai AVE > 0,5. Cara untuk menghitung nilai AVE adalah sebagai berikut :

$$AVE = \frac{\sum_{i=1}^n \lambda_i^2}{\sum_{i=1}^n \lambda_i^2 + \sum_{i=1}^n (1 - \lambda_i^2)} \quad (2.4)$$

2. Uji Reliabilitas

Tujuan uji *composite reliability* adalah untuk mengukur konsistensi dari masing-masing indikator terhadap konstruksinya. Suatu kelompok indikator dikatakan memiliki reliabilitas komposit yang baik jika nilai *composite reliability* > 0.7 (Vinzi, et al., 2010). *Composite reliability* diperoleh melalui rumus sebagai berikut:

$$\rho_c = \frac{(\sum_{i=1}^n \lambda_i)^2}{(\sum_{i=1}^n \lambda_i)^2 + \sum_{i=1}^n (1 - \lambda_i^2)} \quad (2.5)$$

Dimana, λ_i adalah *loading factor* dan ρ_c adalah *composite reliability*.

b. Evaluasi *Inner Model*

Ukuran-ukuran statistik yang digunakan untuk mengevaluasi *inner model* adalah :

1. R-Squared (R^2)

Pengujian *R-squared* (R^2) merupakan cara untuk mengukur tingkat kebaikan suatu model struktural. Nilai *R-squared* (R^2) digunakan untuk menilai seberapa besar pengaruh variabel laten independen tertentu terhadap variabel laten dependen. *R-square* diperoleh melalui rumus sebagai berikut :

$$R^2 = \sum_{h=1}^H \hat{\beta}_{jh} \text{cor}(X_{jh}, Y_j) \quad (2.6)$$

Menurut (Pratiwi et al., 2020), hasil R^2 sebesar 0,67 mengindikasikan bahwa model dikategorikan baik. Hasil R^2 sebesar 0,33 mengindikasikan bahwa model

dikategorikan moderat. Sedangkan hasil R^2 sebesar 0,19 mengindikasikan bahwa model dikategorikan lemah.

2. Signifikansi *Path Coefficient*

Pengujian signifikansi hubungan pada model struktural dilakukan dengan menggunakan statistik t yang diperoleh setelah *path coefficient* diestimasi. Algoritma PLS SEM menghasilkan nilai *path coefficient* (ρ) yang distandardisasi dengan rentang nilai -1 sampai dengan $+1$. Nilai *path coefficient* yang mendekati nilai $+1$ menunjukkan hubungan positif yang kuat sedangkan jika nilai nya mendekati -1 menunjukkan hubungan negatif yang kuat. Statistik t untuk *path coefficient* (ρ) antara variabel laten endogen ke- i dan ke- j dengan standard se_{ij}^* serta nilai t test (*critical ratio*) yang diperoleh dari proses *bootstrapping* $> 1,96$ ($\alpha = 5\%$) dengan formula:

$$t = \frac{\rho_{ij}}{se_{ij}^*} \quad (2.7)$$

3. *Q Square Predictive Relevance*

Prediction relevance (Q^2) atau dikenal dengan Stone-Geisser's. *Q square predictive relevance* yang bertujuan untuk mengukur seberapa baik nilai yang dihasilkan dari model dan estimasinya. Jika *Q square* > 0 maka dapat dikatakan model memiliki *predictive relevance* yang baik, dan sebaliknya jika *Q square* < 0 maka model dianggap kurang memiliki *predictive relevance* (Muliati & Handoyo, 2019). Formula *Prediction relevance* (Q^2) adalah sebagai berikut:

$$Q^2 = 1 - (1 - R^2) \quad (2.8)$$

2.1.4 *Goodness Of Fit*

Untuk menguji model secara keseluruhan adalah dengan *Goodness of Fit* (*GoF*), yaitu merupakan ukuran tunggal yang digunakan untuk memvalidasi performa gabungan antara model pengukuran dan model struktural (Tenenhaus *et al* dalam Vinzi *et al.*, 2010). *GoF* kecil jika bernilai 0,1; *GoF* moderat yang bernilai 0,25; dan *GoF* besar dengan nilai 0,36. Nilai *GoF* ini diperoleh dari mengalikan antara *average communalities index* dengan nilai *average R²*, dengan formula sebagai berikut:

$$GoF = \sqrt{\frac{\sum_{j=1}^J \sum_{q=1}^{p_j} Cor^2(x_{qj}, \hat{\xi}_j)}{\sum_{j=1}^J p_j}} \times \frac{R^2(\{\hat{\xi}_j^*, \{\hat{\xi}_j^* \text{ menjelaskan } \hat{\xi}_j^*\})}{J^*} \quad (2.9)$$

Dimana :

J = jumlah variabel laten dalam model

J^* = jumlah variabel laten endogen dalam model

j^* = mengindikasikan blok endogen awal

$Cor^2(x_{qj}, \hat{\xi}_j)$ = korelasi antara variabel manifest ke- q dari blok ke- j dan variabel laten yang berkaitan

$R^2(\hat{\xi}_{j^*}, \{\hat{\xi}_j's \text{ menjelaskan } \hat{\xi}_{j^*}\})$ = nilai R^2 dari regresi yang menghubungkan variabel laten endogen ke- j^* ke variabel laten penjelasnya

2.1.5 Uji Signifikansi

Pengujian signifikansi dilakukan dengan prosedur *bootstrapping*. *Bootstrapping* adalah pendekatan nonparametrik untuk mengestimasi ketepatan dari estimasi parameter PLS. Metode *bootstrapping* memungkinkan penggunaan jumlah data yang sedikit pada regresi dengan *partial least square*. Hal ini terjadi karena pada metode ini dilakukan *resampling* yang memungkinkan data dapat terdistribusi normal.

1. Uji Hipotesis *Outer Model*

Pada *outer model*, pengujian parameter λ menggunakan hipotesis sebagai berikut:

$H_0 : \lambda_h = 0$, (*loading factor* tidak signifikan dalam mengukur variabel laten)

$H_1 : \lambda_h \neq 0$ (*loading factor* signifikan dalam mengukur variabel laten)

dengan $h = 1, 2, 3, \dots, n$ merupakan variabel manifest/indikator.

Untuk statistik uji yang digunakan adalah statistik uji t (*T-statistics*) atau *critical ratio* sebagai berikut:

$$t = \frac{\lambda_h}{Se^*(\lambda_h)} \quad (2.10)$$

dimana $Se^*(\lambda_h)$ adalah standard error bootsrap dari koefisien λ_h .

H_0 ditolak jika nilai *t-statistic* lebih besar dari nilai kritis z pada *2-tailed* antara lain 1,65 (pada taraf signifikansi 10%), 1,96 (pada taraf signifikansi 5%), dan 2,58 (pada taraf signifikansi 1%) yang menunjukkan *loading factor* tersebut signifikan, sehingga dapat dikatakan valid.

2. Uji Hipotesis *Inner Model*

Pada *inner model*, pengujian hipotesis dilakukan terhadap parameter beta (β) dan parameter gamma (γ).

Parameter beta (β) merupakan parameter yang menunjukkan pengaruh antar variabel endogen. Pengujian parameter β menggunakan hipotesis sebagai berikut:

$H_0: \beta_{ji} = 0$, (variabel laten endogen tidak signifikan dalam mengukur variabel laten endogen lainnya)

$H_1: \beta_{ji} \neq 0$, (variabel laten endogen signifikan dalam mengukur variabel laten endogen lainnya)

dengan :

$j = 1,2,3, \dots, l$ merupakan banyaknya variabel laten endogen dependen, dan $i = 1,2,3, \dots, m$ merupakan banyaknya variabel laten endogen independen.

Dengan statistik uji t (*T-statistics*) atau *critical ratio*:

$$t = \frac{\beta_{ji}}{se^*(\beta_{ji})} \quad (2.11)$$

Dimana $se^*(\beta_{ji})$ merupakan *standard error bootstrap* dari koefisien β_{ji} . Pada level signifikansi 5% dengan tes dua arah (*two-tailed*) keputusan yang diambil adalah menolak H_0 saat ukuran dari nilai empiris *t statistics* yang dihasilkan $|T - statistics| > 1,96$ atau nilai *p - value* $< \alpha (0,05)$.

Parameter gamma (γ) merupakan parameter yang berhubungan dengan pengaruh variabel laten eksogen terhadap variabel laten endogen. Pengujian parameter γ menggunakan hipotesis sebagai berikut:

$H_0: \gamma_{ji} = 0$, (variabel laten eksogen tidak signifikan dalam mengukur variabel laten eksogen lainnya)

$H_1: \gamma_{ji} \neq 0$, (variabel laten eksogen signifikan dalam mengukur variabel laten eksogen lainnya)

Dengan:

$j = 1,2,3, \dots, l$ merupakan banyaknya variabel laten eksogen dependen, dan $i = 1,2,3, \dots, m$ merupakan banyaknya variabel laten eksogen independen.

Dengan statistik uji t (*T-statistics*) atau *critical ratio*:

$$t = \frac{\gamma_{ji}}{se^*(\gamma_{ji})} \quad (2.12)$$

Dimana $se^*(\gamma_{ji})$ merupakan *standard error bootstrap* dari koefisien γ_{ji} . Pada level signifikansi 5% dengan tes dua arah (*two-tailed*) keputusan yang diambil adalah menolak H_0 saat ukuran dari nilai empiris *t statistics* yang dihasilkan $|T - statistics| > 1,96$ atau nilai *p - value* $< \alpha (0,05)$.

2.2 Response Based Units Segmentation Partial Least Square (REBUS PLS)

REBUS PLS diperkenalkan oleh Trinchera (2007) dan Esposito et al (2008) untuk mendeteksi heterogenitas yang berasal dari model struktural dan model

pengukuran. REBUS PLS adalah algoritma iteratif yang dapat mengelompokkan atau mensegmentasi unit observasi sekaligus mengestimasi parameter dari masing-masing model lokal pada segmen yang terbentuk tanpa membutuhkan segala jenis asumsi distribusi baik pada variabel manifest atau pada variabel laten (Trinchera, 2007). Sumber-sumber heterogenitas tersebut diperoleh melalui jarak suatu unit amatan dalam model. Dalam REBUS PLS, jarak suatu unit dalam model didefinisikan dengan memperhatikan bentuk model yang terbentuk berdasarkan *residual* pada model struktural dan model pengukuran untuk seluruh peubah laten. Ukuran untuk menilai jarak antara unit dan model dalam REBUS PLS dihitung dari residual model *communality* yaitu menggunakan *Closeness Measure* (CM). Unit-unit dalam suatu gerombol diperbaharui dengan mempertimbangkan nilai *Closeness Measurement* (CM) (Mehmetoglu, 2011).

Inti pada algoritma REBUS PLS adalah pada ukuran kedekatan atau yang disebut sebagai *closeness measure*. Menurut (Vinzi et al., 2008), untuk melihat jarak antara sebuah unit dengan model adalah dengan menggunakan *closeness measure index* (CM index), yang merupakan struktur *goodness of fit index* (GOF index) yang dihitung dari residual model *communality*. Mengikuti struktur GOF, CM yang dipilih didasarkan pada residual dari model komunalitas (yaitu regresi variabel manifest atas variabel laten masing-masing) dan model struktural (regresi variabel laten endogen atas variabel laten penjelas masing-masing).

Rumus untuk menghitung residual struktural sebagai berikut:

$$f_{ij^*g} = \hat{\xi}_{ij^*g} - y_{ij^*g} \quad (2.13)$$

Dimana : $y_{ij^*g} = \sum_{j=1}^J \text{pada } j^* \beta_{jj^*g} \hat{\xi}_{ijg}$ dan $\hat{\xi}_{ijg} = \sum_{q=1}^{p_j} w_{qjg} x_{iqj}$. β_{jj^*g} adalah koefisien jalur yang menghubungkan variabel laten penjelas j dengan variabel laten endogen j^* , w_{qjg} merupakan *weight outer*, x_{iqj} adalah nilai dari tiap variabel manifest dan $\hat{\xi}_{ij^*g}$ nilai variabel laten endogen j^* . Rumus untuk menghitung residual komunal sebagai berikut:

$$e_{iqgj} = x_{iqj} - \hat{x}_{iqgj} \quad (2.14)$$

Dimana: $\hat{x}_{iqgj} = \lambda_{qgj} \hat{\xi}_{ijg}$. λ_{qgj} merupakan nilai *loading factor* variabel manifest.

CM didefinisikan sebagai berikut (Vinzi et al., 2008):

$$CM_{ig} = \sqrt{\frac{\sum_{j=1}^J \sum_{q=1}^{p_j} [e_{iqgj}^2 / \text{Com}(\hat{\xi}_{qj}, x_{jq})]}{\left(\frac{\sum_{i=1}^N \sum_{j=1}^J \sum_{q=1}^{p_j} [e_{iqgj}^2 / \text{Com}(\hat{\xi}_{qj}, x_{jq})]}{(n_g - m_g - 1)} \right)}} \times \frac{\sum_{j^*=1}^{J^*} [f_{ij^*g}^2 / R^2(\hat{\xi}_{j^*}, \{\hat{\xi}_{j^*s} \text{ menjelaskan } \hat{\xi}_{j^*}\})]}{\left(\frac{\sum_{i=1}^N \sum_{j^*=1}^{J^*} [f_{ij^*g}^2 / R^2(\hat{\xi}_{j^*}, \{\hat{\xi}_{j^*s} \text{ menjelaskan } \hat{\xi}_{j^*}\})]}{(n_g - m_g - 1)} \right)} \quad (2.15)$$

Dimana :

$Com(\hat{\xi}_{qj}, x_{jq})$	= indeks komunal indikator ke-q dari blok ke-j pada variabel laten ke-g
e_{iqjg}	= ukuran residual model untuk unit ke-i pada variabel laten ke-g yang merujuk variabel manifest ke-q di dalam blok ke-j
f_{ij^*g}	= residual model struktural pada unit ke-i pada variabel laten ke-g yang merujuk variabel endogen ke-j*
n_g	= jumlah unit variabel laten ke-g
m_g	= jumlah dimensi (karena semua variabel manifest bersifat reflektif maka nilainya selalu sama dengan 1)

Menurut (Trinchera, 2007) penggunaan *closeness measure* seperti pada persamaan (2.13) sebagai kriteria untuk mengelompokkan unit observasi mempunyai dua kelebihan. Pertama, *unobserved heterogeneity* dapat dideteksi pada model pengukuran dan model struktural. Kedua, karena *closeness measure* didefinisikan berdasarkan struktur dari *Goodness of Fit (GoF) index*, maka local model yang terbentuk akan menunjukkan nilai GoF dan R^2 yang lebih besar.

Langkah pertama dari algoritma REBUS PLS yaitu perhitungan model global menggunakan PLS SEM biasa. Kemudian dihitung *communality* dan residual struktural dari masing-masing unit dari model global. Jumlah kelas (K) diperoleh selama iterasi berturut-turut dan komposisi awal kelas diperoleh berdasarkan hasil *hierarchical cluster analysis*. Selanjutnya yaitu dihitung model lokal pada tiap kelas/segmen yang terbentuk dengan menggunakan PLS SEM biasa. Parameter spesifik segmen yang dihitung pada langkah sebelumnya digunakan untuk menghitung *communality* dan residual struktural dari model lokal sehingga nilai CM setiap unit dari model lokal dapat diperoleh. Selanjutnya, setiap unit observasi dimasukkan ke dalam kelas yang menunjukkan nilai CM yang lebih kecil. Setelah komposisi kelas diperbarui, sejumlah K model lokal baru diperkirakan.

Dalam mengatasi *unobserved heterogeneity*, ini berarti menentukan kelompok atau *cluster* dari unit-unit observasi dengan perilaku dan *performance* yang sama atau serupa. Pada Algoritma REBUS PLS, *clustering* tahap pertama dilakukan pada unit observasi berdasarkan *residual structural* dan *communality* dari global model dan *clustering* tahap kedua dilakukan pada unit observasi berdasarkan nilai CM. Menurut (Trinchera, 2007), *clustering* pertama dilakukan hanya untuk memperoleh kelas/kelompok awal dari unit observasi dan *clustering*

yang kedua berdasarkan asumsi bahwa unit observasi dengan perilaku dan *performance* yang sama berasal dari kelas/kelompok yang sama pula.

Analisis cluster merupakan teknik multivariat yang mempunyai tujuan utama untuk mengelompokkan objek-objek berdasarkan karakteristik yang dimilikinya. Pengelompokkan pada metode cluster berdasarkan pada kemiripan atau kedekatan yang kemudian terbentuk dalam satu kelompok (cluster). Sehingga dikelompokkan sesuai karakter yang sama berdasarkan metode hirarki dan pemusatan data menggunakan metode Ward's, kemudian dianalisis per klaster yang terbentuk menggunakan PLS SEM.

2.3 Kerangka Konseptual

