



**PENERAPAN ALGORITMA *SEM-GIBBS CO-CLUSTERING* UNTUK  
MENGATASI PERMASALAHAN DATA ORDINAL DALAM  
MENGELOMPOKKAN TIPE KEPRIBADIAN**

**SKRIPSI**

**untuk memenuhi persyaratan  
dalam menyelesaikan program sarjana strata-1 statistika**

Oleh  
**FENNY RAHMASARI**  
**NIM. 2011017120001**

**PROGRAM STUDI S-1 STATISTIKA  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS LAMBUNG MANGKURAT  
BANJARBARU  
FEBRUARI 2024**



**PENERAPAN ALGORITMA *SEM-GIBBS CO-CLUSTERING* UNTUK  
MENGATASI PERMASALAHAN DATA ORDINAL DALAM  
MENGELOMPOKKAN TIPE KEPRIBADIAN**

**SKRIPSI**

**untuk memenuhi persyaratan  
dalam menyelesaikan program sarjana strata-1 statistika**

Oleh  
**FENNY RAHMASARI**  
**NIM. 2011017120001**

**PROGRAM STUDI S-1 STATISTIKA  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS LAMBUNG MANGKURAT  
BANJARBARU  
FEBRUARI 2024**

# LEMBAR PENGESAHAN

## SKRIPSI

### PENERAPAN ALGORITMA *SEM-GIBBS CO-CLUSTERING* UNTUK MENGATASI PERMASALAHAN DATA ORDINAL DALAM MENGELOMPOKKAN TIPE KEPERIBADIAN

Oleh  
**Fenny Rahmasari**  
NIM. 2011017120001

Telah dipertahankan pada hari Kamis, tanggal 01-02-2024 dan disetujui dosen pembimbing dan dosen penguji sebagai berikut:

**Pembimbing I**



Yuana Sukmawaty, S.Si., M.Si.  
NIP. 198810152015042002

**Penguji I**



Nur Salam, S.Si., M.Sc.  
NIP. 197708132005011003

**Pembimbing II**



Selvi Annisa, S.Si., M.Si.  
NIP. 199212262022032016

**Penguji II**



Rahmi Fauzia, S.Psi., M.A.Psi.  
NIP. 197712222008122001

Banjarbaru, 15 Februari 2024

Mengetahui,

Koordinator Program Studi Statistika  
MIPA ULM



Prof. Dewi Anggraini, S.Si., M.App.Sci., Ph.D  
NIP. 198303282005012001

## PERNYATAAN

Dengan ini saya menyatakan bahwa dalam skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Banjarbaru, 15 Februari 2024



Fenny Rahmasari  
NIM. 2011017120001

PRODI STATISTIKA

## ABSTRAK

**Penerapan Algoritma *SEM-Gibbs Co-clustering* untuk Mengatasi Permasalahan Data Ordinal dalam Mengelompokkan Tipe Kepribadian**  
(Oleh: Fenny Rahmasari; Pembimbing: Yuana Sukmawaty dan Selvi Annisa, 2024; 83 halaman)

Tipe kepribadian penting diketahui agar membantu pengembangan diri, manajemen emosi, dan pemilihan karir. Dalam berbagai bidang kajian biasanya dilakukan pengelompokan kepribadian menggunakan Analisis *Clustering* yang dapat membantu dalam perumusan strategi dan perencanaan. Dalam psikologi, salah satu metode yang digunakan untuk mengetahui tipe kepribadian adalah *inventory* (kuesioner). Kuesioner biasanya menghasilkan data bertipe ordinal, sehingga dapat dianalisis menggunakan metode *SEM-Gibbs Co-clustering* yang merupakan algoritma khusus untuk data ordinal. Metode ini dapat mengelompokkan responden beserta unit pertanyaan secara simultan, sehingga dapat diketahui secara spesifik ciri kepribadian responden. Data penelitian ini adalah data hasil kuesioner kepribadian IPIP-BFM-50 yang didapatkan dari *Open-Source Psychometrics Project*. Berdasarkan nilai ICL-BIC yang didapatkan,  $(G, H)$  optimal untuk penelitian ini adalah  $(5,4)$ . Oleh karena itu, dihasilkan 5 Kelompok Responden yang dapat diidentifikasi karakteristiknya berdasarkan 4 Kelompok Unit Pertanyaan. Oleh karena itu, didapatkan hasil yaitu, Kelompok 1 berisikan responden yang memiliki sifat sangat terbuka, Kelompok 2 berisikan responden yang memiliki sifat sangat pencemas dan sangat terbuka, Kelompok 3 berisikan responden yang memiliki sifat pencemas dan terbuka, Kelompok 4 berisikan responden yang memiliki sifat terbuka, ramah, serta teliti; dan Kelompok 5 berisikan responden yang memiliki sifat sangat terbuka, sangat ramah, serta sangat teliti.

Kata Kunci: Data Ordinal, Analisis *Clustering*, *SEM-Gibbs Co-clustering*, Tipe Kepribadian

## ABSTRACT

**Application of SEM-Gibbs Co-clustering Algorithm to Solve Ordinal Data Problem in Personality Type Clustering** (by: Fenny Rahmasari; Mentor: Yuana Sukmawaty and Selvi Annisa, 2024; 83 pages)

Personality types are important to know to help with self-development, emotional management and career selection. In various fields of study, it is common to classify personalities using Clustering Analysis which can help in strategy formulation and planning. In psychology, one of the methods used to determine personality type is inventory/questionnaire. Questionnaires usually produce ordinal type data, so it can be analysed using the SEM-Gibbs Co-clustering method which is a special algorithm for ordinal data. This method can group respondents and question units together, so that specific personality traits of respondents can be known. This research data is the result of the IPIP-BFM-50 personality questionnaire obtained from the Open-Source Psychometrics Project. Based on the ICL-BIC value obtained, the optimal (G,H) for this study is (5,4). Therefore, the resulting 5 Respondent Groups can be identified based on 4 Groups of Question Units. So that the results obtained are, Group 1 contains respondents who have very intellect traits, Group 2 contains respondents who have very neuroticism and intellect traits, Group 3 contains respondents who have neuroticism and intellect traits, Group 4 contains respondents who have intellect, agreeableness, and conscientiousness traits; and Group 5 contains respondents who have very intellect, agreeableness, and conscientiousness traits.

Keywords: Ordinal Data, Clustering Analysis, *SEM-Gibbs Co-clustering*, Personality Types

## PRAKATA

Alhamdulillah puji syukur dipanjatkan kepada Allah *Subhanahuwata'ala* yang telah memberikan kesehatan, kekuatan, dan kemudahan kepada peneliti sehingga dapat menyelesaikan skripsi ini. Shalawat serta salam juga tidak lupa dihaturkan kepada Nabi Muhammad *Shallallahu'alaihi wassallam* beserta keluarga dan para sahabat yang telah membimbing dari zaman kegelapan menuju zaman yang penuh cahaya.

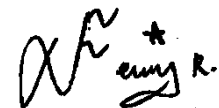
Skripsi yang berjudul "**Penerapan Algoritma SEM-Gibbs Co-clustering untuk Mengatasi Permasalahan Data Ordinal dalam Mengelompokkan Tipe Kepribadian**" ini disusun sebagai syarat wajib akademik yang harus ditempuh dalam menyelesaikan program sarjana di Program Studi S1-Statistika Fakultas Matematika Ilmu Pengetahuan Alam Universitas Lambung Mangkurat. Skripsi ini akan menjelaskan tentang hasil penelitian skripsi sebagai tugas akhir dengan topik yang diminati penulis.

Dalam penyusunan skripsi ini, penulis mengucapkan terimakasih kepada pihak-pihak yang telah berkontribusi dari awal hingga selesainya penelitian skripsi ini. Diantaranya:

1. Allah SWT, yang selalu memberikan saya kekuatan dan kelancaran dalam pengerjaan.
2. Orang tua yang telah memberikan doa, dukungan dan motivasi kepada penulis selama berkuliah sampai penyelesaian skripsi ini.
3. Ibu Yuana Sukmawaty, S.Si., M.Si. selaku Dosen Pembimbing I yang telah memberikan bimbingan, arahan, motivasi dan dukungan dalam penyelesaian skripsi ini.
4. Ibu Selvi Annisa, S.Si., M.Si. selaku Dosen Pembimbing II yang telah memberikan bimbingan, arahan, motivasi dan dukungan dalam penyelesaian skripsi ini.
5. Koordinator beserta seluruh jajaran dosen dan staf Program Studi S-1 Statistika FMIPA ULM yang telah memberikan banyak ilmu dan pengalaman selama kuliah.
6. Teman-teman yang telah memberikan masukan, pendapat, serta dukungan saat pengerjaan skripsi ini.

Skripsi ini jauh dari kata sempurna karena pasti banyak terdapat kekurangan dan keterbatasan. Untuk itu sangat diharapkan kritik dan saran membangun dari semua pihak. Semoga skripsi ini dapat bermanfaat bagi pembaca dalam menambah pengetahuan mengenai penelitian tugas akhir Program Studi S-1 Statistika FMIPA ULM.

Banjarbaru, Februari 2024



Fenny Rahmasari

## DAFTAR ISI

<b>HALAMAN JUDUL</b> .....	i
<b>LEMBAR PENGESAHAN</b> .....	ii
<b>PERNYATAAN</b> .....	iii
<b>ABSTRAK</b> .....	iv
<b>ABSTRACT</b> .....	v
<b>PRAKATA</b> .....	vi
<b>DAFTAR ISI</b> .....	vii
<b>DAFTAR TABEL</b> .....	ix
<b>DAFTAR GAMBAR</b> .....	x
<b>DAFTAR LAMPIRAN</b> .....	xi
<b>DAFTAR ISTILAH, LAMBANG, DAN SINGKATAN</b> .....	xii
<b>BAB I PENDAHULUAN</b> .....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	3
1.3 Tujuan Penelitian .....	3
1.4 Manfaat Penelitian.....	3
<b>BAB II TINJAUAN PUSTAKA</b> .....	5
2.1 Kajian Penelitian Terdahulu .....	5
2.2 Kajian Teori.....	9
2.2.1 Statistika Deskriptif.....	9
2.2.2 Analisis <i>Clustering</i> .....	10
2.2.3 <i>Co-Clustering</i> .....	11
2.2.4 Algoritma <i>SEM-Gibbs Co-clustering</i> .....	12
2.2.5 Teori Kepribadian.....	18
<b>BAB III METODE PENELITIAN</b> .....	21
3.1 Sumber Data.....	21
3.2 Variabel Penelitian.....	21
3.3 Prosedur Penelitian.....	22
<b>BAB IV HASIL DAN PEMBAHASAN</b> .....	24
4.1 Pengelompokkan dengan Algoritma <i>SEM-Gibbs Co-clustering</i> .....	24
4.1.1 <i>Preprocessing</i> dan Statistika Deskriptif.....	24
4.1.2 Inisialisasi Data .....	25
4.1.3 Menghitung ICL-BIC .....	27



4.2	Identifikasi Karakteristik Kelompok yang Terbentuk.....	28
4.2.1	Analisis Hasil Pengelompokan .....	28
4.2.2	Pemberian Label .....	31
<b>BAB V PENUTUP.....</b>		<b>32</b>
5.1	Kesimpulan.....	32
5.2	Saran.....	32
<b>DAFTAR PUSTAKA .....</b>		<b>34</b>
<b>LAMPIRAN.....</b>		<b>37</b>
<b>RIWAYAT HIDUP.....</b>		<b>71</b>

PRODI STATISTIKA

## DAFTAR TABEL

Tabel 1 Penelitian Terdahulu.....	5
Tabel 2 Jenis Metode <i>Clustering</i> .....	11
Tabel 3 Frekuensi Usia Responden .....	25
Tabel 4 Frekuensi Responden berdasarkan Jenis Kelamin .....	25
Tabel 5 Jumlah Responden di Tiap Kelompok.....	29
Tabel 6 Usia Responden Pada Kelompok 3 .....	29
Tabel 7 Jumlah Responden Kelompok 3 berdasarkan Jenis Kelamin.....	29
Tabel 8 Kelompok Unit Pertanyaan.....	30

PRODI STATISTIKA

## DAFTAR GAMBAR

Gambar 1 (a)Visualisasi Data Ordinal; (b)Hasil Co-Clustering Data Ordinal .....	13
Gambar 2 Distribusi BOS $p(X; \mu, \pi)$ .....	13
Gambar 3 Diagram Alir .....	23
Gambar 4 Perubahan Parameter $\pi$ Pada Iterasi SEM-Gibbs.....	26
Gambar 5 Perubahan Parameter $\pi$ (1,1) Pada Iterasi SEM-Gibbs .....	27
Gambar 6 Hasil <i>SEM-Gibbs Co-clustering</i> .....	28

PRODI STATISTIKA

## DAFTAR LAMPIRAN

Lampiran 1 Variabel Penelitian .....	37
Lampiran 2 Validitas dan Reliabilitas IPIP-BFM-50 .....	39
Lampiran 3 Frekuensi Jawaban Responden .....	42
Lampiran 4 Percobaan Jumlah Iterasi dan Persentase Sampel .....	44
Lampiran 5 Hasil Percobaan ICL-BIC .....	56
Lampiran 6 Hasil Pengelompokan Responden .....	59
Lampiran 7 Hasil Pengelompokan Pertanyaan .....	61
Lampiran 8 Karakteristik Kelompok .....	63
Lampiran 9 <i>Preprocessing</i> & Statistika Deskriptif di RStudio .....	65
Lampiran 10 Inisialisasi Data di RStudio .....	67
Lampiran 11 Menghitung ICL-BIC di RStudio .....	68
Lampiran 12 Pengelompokan dan Identifikasi Karakteristik .....	69
Lampiran 13 Karakteristik Responden Kelompok 3 .....	70

## DAFTAR ISTILAH, LAMBANG, DAN SINGKATAN

$\mu$	:	Parameter posisi
$\hat{\mu}$	:	Estimasi parameter posisi
$\pi$	:	Parameter presisi
$\hat{\pi}$	:	Estimasi parameter presisi
$x_{ij}$	:	Data matriks dengan $i$ baris dan $j$ kolom
$m$	:	Jumlah anggota parameter posisi
$v_{ig}$	:	Partisi dari $x_{ij}$ dengan $g$ adalah kelompok baris $i$
$w_{jh}$	:	Partisi dari $x_{ij}$ dengan $h$ adalah kelompok kolom $j$
$a_p$	:	Koefisien dalam fungsi polinomial
$\pi^p$	:	Parameter presisi berpangkat polinomial
$\alpha_{gh}$	:	Parameter fungsi kepadatan peluang $g$ baris dan $h$ kolom
$v_i$	:	Baris $i$ ( $i = 1, 2, \dots, G$ )
$\gamma$	:	Parameter partisi baris
$\gamma_g$	:	Proporsi dari $v_i$
$w_j$	:	Kolom $j$ ( $j = 1, 2, \dots, H$ )
$\rho$	:	Parameter partisi kolom
$\rho_h$	:	Proporsi dari $w_j$
$q$	:	Jumlah iterasi
$\propto$	:	Berbanding lurus
$\theta$	:	Parameter model <i>Latent Block</i>
$\hat{\theta}$	:	Estimasi akhir dari parameter model <i>Latent Block</i>
$\mu_{gh}$	:	Parameter posisi dengan $g$ baris dan $h$ kolom
$\pi_{gh}$	:	Parameter presisi dengan $g$ baris dan $h$ kolom
$N$	:	Total baris pada data
$J$	:	Total kolom pada data
BOS	:	<i>Binary Ordinal Search</i>
EM	:	<i>Expectation-Maximization</i>
SEM-Gibbs	:	<i>Stochastic-Expectation-Maximization-Gibbs</i>
SE-step	:	<i>Stochastic-Expectation-step</i>
M-step	:	<i>Maximization-step</i>
ICL-BIC	:	<i>Integrated Completed Likelihood-Bayesian Information Criterion</i>