



EFEKTIVITAS REDUKSI DIMENSI DATA PADA ANALISIS KOMPONEN UTAMA DENGAN PENANGANAN PENCILAN SEL

SYAMMIRA DHIFA MAULIA



**DEPARTEMEN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT PERTANIAN BOGOR
BOGOR
2024**



PERNYATAAN MENGENAI SKRIPSI DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa skripsi dengan judul “Efektivitas Reduksi Dimensi Data pada Analisis Komponen Utama dengan Penanganan Pencilan Sel” adalah karya saya dengan arahan dari dosen pembimbing dan belum 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 di bagian akhir skripsi ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Institut Pertanian Bogor.

Bogor, Agustus 2024

Syammira Dhifa Maulia
G5401201059

Hak Cipta milik IPB University

IPB University

Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

ABSTRAK

SYAMMIRA DHIFA MAULIA. Efektivitas Reduksi Dimensi Data pada Analisis Komponen Utama dengan Penanganan Pencilan Sel. Dibimbing oleh NGAKAN KOMANG KUTHA ARDANA dan HADI SUMARNO.

Data kompleks seringkali mengandung pencilan signifikan, yang terbagi menjadi pencilan baris dan pencilan sel. Penelitian ini berfokus pada pencilan sel, yang dapat menyebabkan kesalahan analisis jika tidak ditangani dengan baik. Analisis Komponen Utama (AKU) klasik sering terpengaruh oleh pencilan, sehingga diperlukan metode pengembangan seperti AKU Macro dan AKU Kekar. AKU Macro mampu menangani pencilan sel, pencilan baris, dan data yang hilang, sementara AKU Kekar menggabungkan *Projection Pursuit* dengan *Minimum Covariance Determinant*. Metode *Detect Deviating Cells* (DDC) secara khusus dapat mendeteksi dan menangani pencilan sel. Berdasarkan ukuran kesesuaian dalam menganalisis komponen utama, AKU Macro menunjukkan hasil yang lebih efektif dibandingkan metode lain. Selain itu, analisis dengan plot tebaran *score distance* dan *orthogonal distance* menunjukkan bahwa AKU Macro lebih efektif dalam mendeteksi pencilan. Hasil penelitian ini menegaskan bahwa AKU Macro meningkatkan keandalan analisis data yang kompleks, menawarkan pendekatan yang lebih *robust* dalam mengelola pencilan.

Kata kunci: AKU Kekar, AKU Klasik, AKU Macro, DDC, pencilan

ABSTRACT

SYAMMIRA DHIFA MAULIA. Effectiveness of Dimensionality Reduction in Principal Component Analysis with Cellwise Outlier Handling. Supervised by NGAKAN KOMANG KUTHA ARDANA and HADI SUMARNO.

Complex data often contains significant outliers, which can be categorized into rowwise and cellwise outliers. This study focuses on cellwise outliers, which can cause severe analytical errors if not properly addressed. Classical Principal Component Analysis (PCA) is highly susceptible to outliers, necessitating advanced methods such as Macro PCA and Robust PCA. Macro PCA can handle cellwise outliers, rowwise outliers, and missing data, while Robust PCA integrates Projection Pursuit with the Minimum Covariance Determinant. The Detect Deviating Cell (DDC) method is specifically designed to detect and manage cellwise outliers. Based on fit measures in principal component analysis, Macro-PCA is superior to other methods. Furthermore, analysis using score distance and orthogonal distance scatter plots demonstrates that Macro PCA is more effective in outlier detection. This research confirms that Macro PCA significantly enhances the reliability of complex data analysis, providing a robust approach to outlier management.

Keywords: Classical PCA, DDC, Macro PCA, outliers, Robust PCA



Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

© Hak Cipta milik IPB, tahun 2024
Hak Cipta dilindungi Undang-Undang

Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah, dan pengutipan tersebut tidak merugikan kepentingan IPB.

Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apa pun tanpa izin IPB.

EFEKTIVITAS REDUKSI DIMENSI DATA PADA ANALISIS KOMPONEN UTAMA DENGAN PENANGANAN PENCILAN SEL

SYAMMIRA DHIFA MAULIA

Skripsi
sebagai salah satu syarat untuk memperoleh gelar
Sarjana Matematika pada
Program Studi Matematika

**DEPARTEMEN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT PERTANIAN BOGOR
BOGOR
2024**



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Judul Skripsi : Efektivitas Reduksi Dimensi Data pada Analisis Komponen Utama dengan Penanganan Pencilan Sel

Nama : Syammira Dhifa Maulia

NIM : G5401201059

Disetujui oleh

Pembimbing 1:

Ir. Ngakan Komang Kutha Ardana, M.Sc.



Pembimbing 2:

Dr. Ir. Hadi Sumarno, M.S.



Diketahui oleh

Ketua Departemen Matematika:

Dr. Ir. Endar Hasafah Nugrahani, MS.

NIP 19631228 198903 2 001



Tanggal Ujian: 2 Agustus 2024

Tanggal Lulus:



PRAKATA

Puji dan syukur penulis panjatkan kepada Allah subhanahu wa ta'ala atas segala karunia-Nya sehingga karya ilmiah yang berjudul “Efektivitas Reduksi Dimensi Data pada Analisis Komponen Utama dengan Penanganan Pencilan Sel” berhasil diselesaikan. Terima kasih penulis ucapkan kepada:

1. Ayahanda Maulana, Ibunda Maslia, adik-adik Fara Diva Maulia dan Ferry Fadillah Maulia atas dukungan, motivasi, serta doa yang selalu diberikan kepada penulis tiada henti-hentinya.
2. Ungkapan terima kasih juga disampaikan kepada Bapak Ir. Ngakan Komang Kutha Ardana, M.Sc. dan Bapak Dr. Ir. Hadi Sumarno, M.S. selaku pembimbing, yang dengan sabar memberikan bimbingan, pengarahan, saran, dan ilmu dalam menyelesaikan karya ilmiah ini.
3. Bapak Prof. Dr. Ir. I Wayan Mangku, M.Sc. selaku penguji yang telah banyak memberi arahan dan saran kepada penulis.
4. Seluruh Bapak/Ibu dosen dan staf Departemen Matematika IPB yang telah memberikan ilmu serta bantuannya kepada penulis selama proses perkuliahan.
5. Dita Dervinta, Salsabila Fitri Imni, Aqilah Fitri Chania, Renda Sulistyorini Padma Putri, Alifia Nabila Hakim, Annisa Permata Sari, Ana Nur Fitria, Desi Novianti, Tazkia Nur Hafizha, dan Rizka Dwi Andriani selaku sahabat kuliah yang telah memberikan semangat, motivasi, serta menemani penulis dalam suka dan duka selama masa perkuliahan.
6. Tania Chandra Prastiwi dan Hafidz Al Ubaidah sebagai rekan seperbimbingan atas motivasi dan energi positifnya yang sangat membangun.
7. Regita Zulkarnaen dan Geng Setujuh selaku sahabat penulis yang telah banyak memberikan semangat dalam menyelesaikan karya ilmiah ini.
8. Keluarga besar Departemen Matematika 57 sebagai teman seperjuangan selama masa perkuliahan.
9. Semua pihak yang tidak dapat penulis sebutkan satu per satu, yang telah membantu dalam proses penulisan karya ilmiah ini.

Semoga karya ilmiah ini bermanfaat bagi semua pihak yang membutuhkan dan bagi kemajuan ilmu pengetahuan. Penulis mohon maaf atas segala kekurangan dan kesalahan yang terdapat dalam karya ilmiah ini.

Bogor, Agustus 2024

Syammira Dhifa Maulia



DAFTAR ISI

DAFTAR TABEL	x
DAFTAR GAMBAR	x
DAFTAR LAMPIRAN	xi
I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Tujuan	1
II TINJAUAN PUSTAKA	2
2.1 Analisis Komponen Utama	2
2.2 Analisis Komponen Utama Kekar	3
2.3 Analisis Komponen Utama Macro	4
2.4 Pencilan Sel	4
2.5 Peta Pencilan	5
2.6 <i>Detect Deviating Cells (DDC)</i>	6
III METODE	8
3.1 Alat dan Data	8
3.2 Tahapan Analisis Data	8
IV HASIL DAN PEMBAHASAN	9
4.1 Profil Antar Variabel	9
4.2 Korelasi Antar Variabel	11
4.3 Proporsi Kumulatif Ragam	14
4.4 Deteksi Pencilan	16
V SIMPULAN DAN SARAN	26
5.1 Simpulan	26
5.2 Saran	26
DAFTAR PUSTAKA	27
LAMPIRAN	29
RIWAYAT HIDUP	56

Hak Cipta Dilindungi Undang-undang
 1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



DAFTAR TABEL

1	Proporsi kumulatif ragam AKU Klasik, AKU Kekar, dan AKU Macro data tanpa pencilan	14
2	Proporsi kumulatif ragam AKU Klasik, AKU Kekar, dan AKU Macro data pencilan 10%	15
3	Proporsi kumulatif ragam AKU Klasik, AKU Kekar, dan AKU Macro data pencilan 20%	15
4	Proporsi kumulatif ragam AKU Klasik, AKU Kekar, dan AKU Macro data <i>Rice</i>	16

DAFTAR GAMBAR

1	Pencilan baris	5
2	Pencilan sel	5
3	Peta pencilan	5
4	<i>Boxplot</i> data tanpa pencilan	9
5	<i>Boxplot</i> data dengan pencilan 10%	9
6	<i>Boxplot</i> data dengan pencilan 20%	10
7	<i>Boxplot</i> data <i>Rice</i>	10
8	Hubungan berdasarkan korelasi	11
9	Matriks korelasi data tanpa pencilan	11
10	Matriks korelasi data dengan pencilan 10%	12
11	Matriks korelasi data dengan pencilan 20%	13
12	Matriks korelasi data <i>Rice</i>	13
13	Plot tebaran AKU Klasik dan deteksi pencilan pada data tanpa pencilan	16
14	Plot tebaran AKU Kekar dan deteksi pencilan pada data tanpa pencilan	17
15	Plot tebaran AKU Macro dan deteksi pencilan pada data tanpa pencilan	17
16	Deteksi pencilan sel dari DDC data tanpa pencilan	18
17	Plot tebaran AKU Klasik dan deteksi pencilan pada data proporsi pencilan 10%	18
18	Plot tebaran AKU Kekar dan deteksi pencilan pada data proporsi pencilan 10%	19
19	Plot tebaran AKU Macro dan deteksi pencilan pada data proporsi pencilan 10%	19
20	Deteksi pencilan sel dari DDC data pencilan 10%	20
21	Plot tebaran AKU Klasik dan deteksi pencilan pada data proporsi pencilan 20%	21
22	Plot tebaran AKU Kekar dan deteksi pencilan pada data proporsi pencilan 20%	21
23	Plot tebaran AKU Macro dan deteksi pencilan pada data proporsi pencilan 20%	22
24	Deteksi pencilan sel dari DDC data pencilan 20%	22
25	Plot tebaran AKU Klasik dan deteksi pencilan data <i>Rice</i>	23
26	Plot tebaran AKU Kekar dan deteksi pencilan data <i>Rice</i>	23
27	Plot tebaran AKU Macro dan deteksi pencilan data <i>Rice</i>	24

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

28	Deteksi pencilan sel dari DDC data <i>Rice</i>	24
----	--	----

DAFTAR LAMPIRAN

1	Statistika deskriptif variabel data proporsi pencilan 0%	30
2	Statistika deskriptif variabel data proporsi pencilan 10%	31
3	Statistika deskriptif variabel data proporsi pencilan 20%	32
4	Statistika deskriptif variabel data <i>Rice</i>	33
5	Korelasi antar variabel data proporsi pencilan 0%	34
6	Korelasi antar variabel data proporsi pencilan 10%	35
7	Korelasi antar variabel data proporsi pencilan 20%	36
8	Korelasi antar variabel data <i>Rice</i>	37
9	Koordinat <i>bad leverage point</i> pada AKU Kekar, dan AKU Macro pada data proporsi pencilan 10%	38
10	Koordinat <i>bad leverage point</i> pada AKU Kekar dan AKU Macro pada data proporsi pencilan 20%	39
11	Koordinat <i>bad leverage point</i> pada AKU Kekar dan AKU Macro pada data <i>Rice</i>	40
12	Indeks sel pencilan pada <i>Detect Deviating Cells</i> data proporsi pencilan 0%	41
13	Indeks sel pencilan pada <i>Detect Deviating Cells</i> data proporsi pencilan 10%	42
14	Indeks sel pencilan pada <i>Detect Deviating Cells</i> data proporsi pencilan 20%	44
15	Indeks sel pencilan pada <i>Detect Deviating Cells</i> data <i>Rice</i>	46
16	Data proporsi pencilan 0%	47
17	Data proporsi pencilan 10%	49
18	Data proporsi pencilan 20%	51
19	Data <i>Rice</i>	53

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.

2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.