



PERBANDINGAN ANALISIS KOMPONEN UTAMA DENGAN PARADIGMA PENCILAN SEL DAN PENCILAN BARIS

MAULANA RAFI ARDIAN



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

- 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 University

IPB University



IPB University
— Bogor Indonesia —

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.



PERNYATAAN MENGENAI SKRIPSI DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa skripsi dengan judul “Perbandingan Analisis Komponen Utama dengan Paradigma Sel dan Pencilan Baris” 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, Juli 2024

Maulana Rafi Ardian



ABSTRAK

MAULANA RAFI ARDIAN. Perbandingan Analisis Komponen Utama dengan Paradigma Pencilan Sel dan Pencilan Baris. Dibimbing oleh NGAKAN KOMANG KUTHA ARDANA dan HADI SUMARNO.

Big Data merupakan data berukuran besar yang volumenya akan terus bertambah dengan terdiri dari berbagai jenis data tertentu. Hal ini mengakibatkan kerumitan dalam mengolah data sehingga perlu adanya cara khusus dalam mengelola *Big Data*. Analisis Komponen Utama (AKU) adalah teknik yang digunakan untuk menyederhanakan suatu data. Penelitian ini membahas penerapan metode Analisis Komponen Utama (AKU) dalam tiga varian: AKU-Klasik, AKU-Klasik dengan deteksi pencilan menggunakan *Detect Deviating Cells* (DDC), dan AKU-Makro. Penelitian ini akan berfokus untuk mengevaluasi dan membandingkan ketahanan ketiga metode tersebut terhadap adanya pencilan sel dan pencilan baris dalam data dengan proporsi pencilan sebesar 0%, 5%, dan 10%. Analisis ini dilakukan dengan mengukur nilai eigen, variansi yang dijelaskan oleh komponen, serta proporsi kumulatif variansi yang dijelaskan oleh komponen utama dalam dataset. Analisis lebih lanjut pada nilai kumulatif variansi menunjukkan bahwa meskipun terjadi penurunan nilai eigen dengan peningkatan proporsi pencilan, AKU-Makro menunjukkan efektivitas lebih baik dibandingkan AKU-Klasik dan AKU-Klasik menggunakan DDC.

Kata Kunci: Analisis Komponen Utama, AKU-Makro, DDC, pencilan

ABSTRACT

MAULANA RAFI ARDIAN. A Comparison of Principal Component Analysis with Cell Outlier Paradigm and Row Outlier Paradigm. Supervised by NGAKAN KOMANG KUTHA ARDANA and HADI SUMARNO.

Big Data is large-sized data whose volume continues to increase, consisting of various types of specific data. This results in complexity in data processing, requiring special methods to manage Big Data. Principal Component Analysis (PCA) is a technique used to simplify data. This research discusses the application of Principal Component Analysis (PCA) method in three variants: Classic PCA, Classic PCA with outlier detection using Detect Deviating Cells (DDC), and Macro PCA. This research will focus on evaluating and comparing the robustness of these three methods against cell outliers and row outliers in data with outlier proportions of 0%, 5%, and 10%. This analysis is conducted by measuring eigenvalues, variance explained by components, and cumulative proportion of variance explained by principal components in the dataset. Further analysis on cumulative variance values indicates that despite a decrease in eigenvalues with an increase in outlier proportions, Macro PCA demonstrates better effectiveness compared to Classic PCA and Classic PCA using DDC.

Keywords: DDC, Macro PCA, outliers, Principal Component Analysis



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.



PERBANDINGAN ANALISIS KOMPONEN UTAMA DENGAN PARADIGMA PENCILAN SEL DAN PENCILAN BARIS

MAULANA RAFI ARDIAN

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.

Penguji pada Ujian Skripsi:

Prof. Dr. Ir. I Wayan Mangku, M.Sc.

Judul Skripsi : Perbandingan Analisis Komponen Utama dengan Paradigma
Sel dan Pencilan Baris
Nama : Maulana Rafi Ardian
NIM : G54180043

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, M.S.
NIP 19631228 198903 2 001



Tanggal Ujian: 20 Juni 2024

Tanggal Lulus:



PRAKATA

Puji dan syukur penulis panjatkan kepada Allah *subhaanahu wa ta'ala* atas segala karunia-Nya sehingga karya ilmiah ini berhasil diselesaikan. Tema yang dipilih dalam penelitian ini ialah pemodelan matematika dengan judul "Perbandingan Analisis Komponen Utama dengan Paradigma Pencilan Sel dan Pencilan Baris". Proses penulisan skripsi ini tidak terlepas dari doa, dukungan, motivasi, saran, dan bantuan dari berbagai pihak yang terlibat. Oleh karena itu, penulis mengucapkan terima kasih kepada:

1. Kedua orang tua tercinta yaitu Bapak Abdul Wahid dan Ibu Eni Mustafidah, Saudara Maulana Nadiful Afkar, Nauraifa Elita Putriwahdani, dan Maulana Rangga Nurilasyfar serta seluruh keluarga yang telah memberikan doa, motivasi, dan kasih sayang yang tak terhingga kepada penulis.
2. Bapak Ir. Ngakan Komang Kutha Ardana, M.Sc. selaku dosen pembimbing 1 dan Dr. Ir. Hadi Sumarno, M.S. selaku dosen pembimbing 2 yang telah memberikan berbagai saran, masukan, dan perhatiannya kepada penulis selama proses penulisan skripsi ini.
3. Seluruh dosen, staf pengajar, dan civitas akademika Departemen Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, IPB University yang telah memberikan dukungan kepada penulis.
4. Seluruh teman-teman Matematika angkatan 2018 yang telah berdedikasi dan memberikan banyak bantuan kepada penulis.

Semoga karya ilmiah ini bermanfaat bagi pihak yang membutuhkan dan bagi kemajuan ilmu pengetahuan.

Bogor, Juli 2024

Maulana Rafi Ardian

DAFTAR ISI

| | |
|---|-----|
| DAFTAR TABEL | xii |
| DAFTAR GAMBAR | xii |
| DAFTAR LAMPIRAN | xii |
| I PENDAHULUAN | 1 |
| 1.1 Latar Belakang | 1 |
| 1.2 Perumusan Masalah | 2 |
| 1.3 Tujuan Penelitian | 2 |
| 1.4 Manfaat Penelitian | 2 |
| II TINJAUAN PUSTAKA | 3 |
| 2.1 Analisis Komponen Utama | 3 |
| 2.2 Nilai Eigen dan Vektor Eigen | 4 |
| 2.3 <i>Detect Deviating Cells</i> | 5 |
| 2.4 Pencilan | 5 |
| 2.5 AKU-Makro | 6 |
| III METODE | 7 |
| IV HASIL DAN PEMBAHASAN | 8 |
| 4.1 Membangkitkan Data | 8 |
| 4.2 Penambahan Pencilan | 10 |
| 4.3 Perbandingan AKU-Klasik, AKU-Klasik dengan DDC, dan AKU-Makro | 14 |
| 4.4 <i>Detect Deviating Cells</i> | 16 |
| 4.5 Analisis AKU-Makro pada Pencilan Baris dan Pencilan Sel | 18 |
| V SIMPULAN DAN SARAN | 22 |
| 5.1 Simpulan | 22 |
| 5.2 Saran | 22 |
| DAFTAR PUSTAKA | 23 |
| LAMPIRAN | 25 |
| RIWAYAT HIDUP | 35 |



DAFTAR TABEL

| | | |
|---|---|----|
| 1 | Penerapan AKU-Klasik, AKU-Klasik dengan DDC, dan AKU-Makro pencilan 0% | 14 |
| 2 | Penerapan AKU-Klasik, AKU-Klasik dengan DDC, dan AKU-Makro pencilan 5% | 15 |
| 3 | Penerapan AKU-Klasik, AKU-Klasik dengan DDC, dan AKU-Makro pencilan 10% | 16 |
| 4 | Penerapan AKU-Makro pada pencilan 0%, 5%, dan 10% | 18 |

DAFTAR GAMBAR

| | | |
|----|---|----|
| 1 | Perbandingan pencilan baris dan sel | 6 |
| 2 | Visualisasi variabel acak dengan <i>boxplot</i> | 8 |
| 3 | Visualisasi variabel acak dengan <i>correlation plot</i> | 9 |
| 4 | Visualisasi variabel acak dengan proporsi 5% pencilan | 10 |
| 5 | Visualisasi variabel acak dengan proporsi 10% pencilan | 11 |
| 6 | <i>Correlation plot</i> variabel acak dengan proporsi 5% pencilan | 12 |
| 7 | <i>Correlation plot</i> variabel acak dengan proporsi 10% pencilan | 13 |
| 8 | Perbandingan peningkatan pencilan dari 0%, 5%, dan 10% menggunakan <i>Detect Deviating Cell</i> | 17 |
| 9 | Perbandingan penerapan DDC dan <i>Residual Map</i> AKU-Makro pada dataset pencilan 5% | 19 |
| 10 | Perbandingan penerapan DDC dan <i>Residual Map</i> AKU-Makro pada dataset pencilan 10% | 20 |

DAFTAR LAMPIRAN

| | | |
|----|--|----|
| 1 | Dataset 0% Pencilan | 26 |
| 2 | Dataset 5% Pencilan | 27 |
| 3 | Dataset 10% Pencilan | 28 |
| 4 | Statistik Deskriptif Dataset 0% Pencilan | 29 |
| 5 | Statistik Deskriptif Dataset 5% Pencilan | 29 |
| 6 | Statistik Deskriptif Dataset 10% Pencilan | 29 |
| 7 | Korelasi Variabel Dataset 0% Pencilan | 30 |
| 8 | Korelasi Variabel Dataset 5% Pencilan | 30 |
| 9 | Korelasi Variabel Dataset 10% Pencilan | 30 |
| 10 | Kode Sumber: AKU-Klasik | 31 |
| 11 | Kode Sumber: AKU-Klasik dengan DDC | 31 |
| 12 | Kode Sumber: AKU-Makro | 32 |
| 13 | Indeks Pencilan AKU-Klasik dengan DDC pada dataset 5% | 32 |
| 14 | Indeks Pencilan AKU-Klasik dengan DDC pada dataset 10% | 33 |

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