

ANALYSIS OF AIR QUALITY IN INDONESIA USING TENSOR DECOMPOSITION-BASED CLUSTERING

TIAS AMALIA SAFITRI



**DEPARTMENT OF STATISTICS
FACULTY OF MATHEMATICS AND NATURAL SCIENCES
IPB UNIVERSITY
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 mengumunkan 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 “*Analysis of Air Quality in Indonesia Using Tensor Decomposition-Based Clustering*” 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, Juni 2024

Tias Amalia Safitri
G1401201104

- 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.



ABSTRAK

TIAS AMALIA SAFITRI. Analisis Kualitas Udara di Indonesia Menggunakan Klustering Berbasis Dekomposisi Tensor. Dibimbing oleh ASEP SAEFUDDIN dan SACHNAZ DESTA OKTARINA.

Indonesia perlu mengatasi tingkat polusi udara yang tinggi. Untuk mengatasi dampaknya terhadap kesehatan manusia dan lingkungan, diperlukan upaya bersama dari semua pihak. Mengelompokkan wilayah berdasarkan kesamaan karakteristik kualitas udara akan membantu pemerintah menyesuaikan kebijakan dengan lebih efektif. Penelitian ini bertujuan untuk melakukan analisis berbasis kluster terhadap kualitas udara, tidak hanya berdasarkan komponen polutan, tetapi juga dalam konteks spasial dan temporal menggunakan dekomposisi tensor. Elemen-elemen kluster kemudian dianalisis dengan pola kejadian yang serupa untuk mendapatkan faktor-faktor potensial. Data kualitas udara yang mengandung waktu, wilayah, dan parameter kualitas udara diproses untuk membentuk sebuah tensor. Parameter kualitas udara yang ditangani meliputi karbon monoksida (CO), nitrogen dioksida (NO₂), sulfur dioksida (SO₂), ozon (O₃), dan partikel (PM₁₀ dan PM_{2.5}). Data diambil dari 31 stasiun pemantauan perwakilan di berbagai provinsi di Indonesia dari Oktober 2023 hingga Desember 2023. Dekomposisi tensor dengan tiga fitur laten di setiap dimensi digunakan untuk klustering k-medoid menggunakan evaluasi siluet dan siku. Dekomposisi Tucker digunakan sebagai metode umum dalam dekomposisi tensor karena kemampuannya untuk mengontrol peringkat dekomposisi. Ini berarti mengatur jumlah fitur laten atau komponen yang akan diekstraksi selama proses dekomposisi. Analisis menghasilkan jumlah kluster optimal untuk dimensi wilayah, waktu, dan polutan. Ditemukan bahwa empat kluster teridentifikasi berdasarkan dimensi wilayah. Sementara itu, dalam hal pola temporal, terdapat tiga kluster. Selain itu, tiga kluster teridentifikasi berdasarkan jenis polutan. Kluster wilayah yang paling mengkhawatirkan, termasuk Palembang dan Pontianak, memiliki konsentrasi PM₁₀ dan PM_{2.5} di atas standar aman.

Kata kunci: emisi karbon, klustering, kualitas udara, spasial, temporal

ABSTRACT

TIAS AMALIA SAFITRI. Analysis of Air Quality in Indonesia Using Tensor Decomposition-Based Clustering. Supervised by ASEP SAEFUDDIN and SACHNAZ DESTA OKTARINA.

Indonesia faces severe challenges in addressing high levels of air pollution. To anticipate and address the effect of air pollution on human health and the environment, all parties need concerted efforts. Grouping regions based on similarities in air quality will help the government tailor policies more effectively. This research aims to conduct a cluster-based analysis of air quality in terms of not only its pollutant compound but also its spatial- and temporal context using tensor decomposition. The cluster elements were then analyzed with similar occurrence patterns to obtain potential factors. Air quality data containing time, region, and air quality parameters are processed to construct a tensor. Air quality parameters being addressed are carbon monoxide (CO), nitrogen dioxide (NO₂), sulfur dioxide (SO₂), ozone (O₃), and particulate matter (PM₁₀ and PM_{2.5}). The data was retrieved from 31 representative monitoring stations across provinces in Indonesia from October 2023 to December 2023. Tensor decomposition with three latent features in each dimension was used for k-medoid clustering using the silhouette and elbow evaluation. Tucker decomposition is used for the general method in tensor decomposition because of its ability to control the decomposition rank. This means regulating the number of latent features or components to be extracted during the decomposition process. The analysis resulted optimal cluster numbers for region, time, and pollutant dimensions. It suggested that four clusters were identified based on region dimension. Meanwhile, in terms of temporal patterns, three clusters resulted. Furthermore, three clusters were identified based on pollutant types. The most concerning region cluster, including Palembang and Pontianak, have PM₁₀ and PM_{2.5} concentrations above the safe level.

Keywords: air quality, carbon emission, clustering, spatial, temporal



- 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.

ANALYSIS OF AIR QUALITY IN INDONESIA USING TENSOR DECOMPOSITION-BASED CLUSTERING

TIAS AMALIA SAFITRI

Undergraduate Thesis
To complete the requirement for graduation of
Bachelor Degree in
Statistics and Data Science Study Program

**DEPARTMENT OF STATISTICS
FACULTY OF MATHEMATICS AND NATURAL SCIENCES
IPB UNIVERSITY
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.

Tim Penguji pada Ujian Skripsi:
Dr. Farit Mochmad Afendi, S.Si., M.Si



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



Title : Analysis of Air Quality in Indonesia Using Tensor
Decomposition-Based Clustering
Name : Tias Amalia Safitri
NIM : G1401201104

@Hak cipta milik IPB University

Approved by

Main Supervisor:
Prof. Dr. Ir. Asep Saefuddin, M.Sc.



Co-Supervisor:
Sachnaz Desta Oktarina S.Stat., M.Agr.Sc., Ph.D



Acknowledged by

Head of Statistics Department
Dr. Bagus Sartono, S.Si., M.Si.
NIP 197804112005011002



Thesis Defense Date:
28th of June, 2024

Graduation Date:

PREFACE

All praise and gratitude to Allah SWT for the unending blessing so that this research can be completed. The topic of the research that has been conducted since December 2023 to July 2024 is building clustering analysis to analyze the characteristics among clusters in terms of time, region, and pollutant in Indonesia, titled “Analysis of Air Quality in Indonesia Using Tensor Decomposition-Based Clustering”. The completion of this scientific paper is due to various support and assistance from different parties. Therefore, the author expresses appreciation to:

1. Allah SWT for blessing me with good health and giving me strength
2. Author’s family for their unwavering support and encouragement
3. Prof. Dr. Asep Saefuddin. and Sachnaz Desta Oktarina, Ph.d. for their guidance and insights throughout the research development
4. The lecturers and administrative staff in the Department of Statistics for fostering an environment conducive to academic growth
5. Tanoto Foundation for financial support through the TELADAN Leadership and Scholarship Program
6. Author’s friends, particularly from the Statistics and Data Science Program Study Batch 57 for their unwavering support and suggestions.
7. All individuals and entities that have given all kinds of assistance in various way so that this research can be completed.

Hopefully, this research can be useful to many people and for a variety kind of purposes. Acknowledging the potential limitations and imperfections of this research, the author welcomes constructive feedback and recommendations with open arms. Such input will be greatly appreciated and duly considered.

Bogor, June 2024

Tias Amalia Safitri



TABLE OF CONTENTS

LIST OF TABLES	x
LIST OF FIGURES	x
LIST OF APPENDIXES	x
INTRODUCTION	1
1.1 Background	1
1.2 Objectives	2
LITERATURE REVIEW	3
2.1 Tensor Decomposition	3
2.2 K-Medoid Clustering	5
2.3 Air Quality Index	6
METHODOLOGY	8
3.1 Data	8
3.2 Analysis Procedure	9
RESULT AND DISCUSSION	11
4.1 Data Preprocessing and Exploration	11
4.2 Tensor Decomposition-Clustering	12
4.3 Region-Based Clustering	13
4.4 Time-Based Clustering	16
4.5 Pollutant-Based Clustering	17
CONCLUSION AND RECOMMENDATION	20
5.1 Conclusion	20
5.2 Recommendation	20
BIBLIOGRAPHY	21
APPENDICES	23
BIOGRAPHY	27