



LEMBAR PENGESAHAN

Judul Tugas Akhir : Analisis Kata Kunci yang Relevan dengan Faktor Gas Rumah Kaca di Indonesia Menggunakan *Text Mining* dan Teori Luhn
Nama Mahasiswa : Kiagus Muhammad Arsyad
Nomor Induk Mahasiswa : 105219002
Program Studi : Ilmu Komputer
Fakultas : Sains dan Ilmu Komputer
Tanggal Lulus Sidang Tugas Akhir : 3 Agustus 2023

Jakarta, 31 Juli 2023

MENGESAHKAN

Pembimbing I

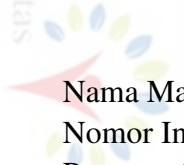
Pembimbing II

Dr. Tasmi, S.Si., M.Si.
NIP. 116109

Dr. Ariana Yunita
NIP. 116015

MENGETAHUI,
Ketua Program Studi

Ade Irawan, Ph.D
NIP. 116130



LEMBAR PERNYATAAN

Dengan ini saya menyatakan bahwa Tugas Akhir berjudul Analisis Kata Kunci yang Relevan dengan Faktor Gas Rumah Kaca di Indonesia Menggunakan *Text Mining* dan Teori Luhn ini adalah benar-benar merupakan hasil karya saya sendiri dan tidak mengandung materi yang ditulis oleh orang lain kecuali telah dikutip sebagai referensi yang sumbernya telah dituliskan secara jelas sesuai dengan kaidah penulisan karya ilmiah.

Apabila dikemudian hari ditemukan adanya kecurangan dalam karya ini, saya bersedia menerima sanksi dari Universitas Pertamina sesuai dengan peraturan yang berlaku.

Demi pengembangan ilmu pengetahuan, saya menyetujui untuk memberikan kepada Universitas Pertamina hak bebas royalti noneksklusif (*non-exclusive royalty-free right*) atas Tugas Akhir ini beserta perangkat yang ada. Dengan hak bebas royalti noneksklusif ini Universitas Pertamina berhak menyimpan, mengalih media/format-kan, mengelola dalam bentuk pangkatan data (*database*), merawat, dan mempublikasikan Tugas Akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya

Jakarta, 31 Juli 2023

Yang membuat pernyataan,

Kiagus Muhammad Arsyad

ABSTRAK

Kiagus Muhammad Arsyad. 105219002. Analisis Kata Kunci yang Relevan dengan Faktor Gas Rumah Kaca di Indonesia Menggunakan *Text Mining* dan Teori Luhn.

Penyebab Gas Rumah Kaca (GRK) menjadi isu global yang mendesak untuk ditangani, termasuk di Indonesia. Penelitian ini bertujuan untuk mengidentifikasi faktor penyebab GRK di Indonesia melalui analisis data abstrak dari studi literatur lima tahun terakhir. Pendekatan yang digunakan melibatkan teknik *text mining*, *clustering*, dan seleksi menggunakan teori Luhn. Dengan menggunakan algoritma K-Means dalam teknik *clustering*, data abstrak dikategorikan menjadi beberapa kluster berdasarkan kesamaan konten. Metode seleksi menggunakan teori Luhn juga mampu mengidentifikasi kata-kata kunci yang paling relevan dan mewakili relevansi GRK di Indonesia. Hasil analisis menunjukkan bahwa faktor-faktor pada aspek yang memiliki relevansi GRK di Indonesia meliputi diantaranya deforestasi, biodiesel, CO2/karbon, dan listrik. Pada penelitian ini juga memberikan pemahaman yang cukup tentang faktor-faktor yang berhubungan pada aspek GRK di Indonesia dan memberikan informasi penting bagi pembuat kebijakan, para ahli, para pembaca untuk mengembangkan strategi mitigasi yang efektif.

Kata kunci: faktor, gas rumah kaca, kata kunci, nlp, teori luhn, *text mining*

ABSTRACT

Kiagus Muhammad Arsyad. 105219002. Analysis of Relevant Keywords to Greenhouse Gas Factors in Indonesia Using Text Mining and Luhn's Theory.

The causes of Greenhouse Gas (GHG) emissions have become a pressing global issue to address, including in Indonesia. This study aims to identify the factors contributing to GHG emissions in Indonesia through the analysis of abstract data from literature studies of the past five years. The approach involves techniques such as text mining, clustering, and selection using Luhn's theory. By employing the K-Means algorithm in clustering, abstract data is categorized into several clusters based on content similarity. The Luhn's theory-based selection method also identifies the most relevant keywords that represent the relevance of GHG emissions in Indonesia. The analysis results indicate that factors related to the aspects relevant to GHG emissions in Indonesia include deforestation, biodiesel, CO₂/carbon, and electricity. This study provides a comprehensive understanding of the factors associated with the GHG aspect in Indonesia, offering essential insights for policymakers, experts, and readers to develop effective mitigation strategies.

Keywords: factors, greenhouse gas, keywords, luhn's theory, nlp, text mining

KATA PENGANTAR

Dengan memanjatkan segala puji syukur kehadirat Allah SWT yang telah melimpahkan rahmat, taufik, dan hidayah-Nya sehingga penulis dapat menyelesaikan skripsi berupa laporan Tugas Akhir (TA) yang berjudul **“Analisis Kata Kunci yang Relevan dengan Faktor Gas Rumah Kaca di Indonesia Menggunakan Text Mining dan Teori Luhn”**, sebagai salah satu persyaratan dalam menyelesaikan masa belajar di Program Studi S1 Ilmu Komputer, Fakultas Sains dan Ilmu Komputer Universitas Pertamina. Penulis menyadari bahwa laporan TA ini tidak akan terselesaikan tanpa adanya dukungan, bimbingan, bantuan, dan nasihat dari berbagai pihak selama penyusunan laporan TA ini. Pada kesempatan ini, penulis ingin menyampaikan terima kasih setulus-tulusnya kepada:

1. Bapak Ade Irawan, Ph.D selaku Ketua Program Studi Ilmu Komputer Universitas Pertamina dan juga sebagai dosen penguji pertama yang telah memberikan masukan dan saran pada skripsi ini.
2. Bapak Rangga Ganzar Noegraha, Ph.D selaku dosen penguji kedua yang telah memberikan masukan dan saran pada skripsi ini.
3. Ibu Dr. Tasmi, S.Si, M.Si. dan Ibu Dr. Ariana Yunita selaku dosen pembimbing pertama dan dosen pembimbing kedua pada skripsi ini yang telah memberikan bimbingan, motivasi, dan arahan selama penyusunan laporan tugas akhir dari awal hingga akhir saat ini.
4. Seluruh staf pengajar yang telah memberikan ilmu pengetahuan yang tak ternilai selama penulis menempuh pendidikan di Universitas Pertamina.
5. Seluruh *civitas academica* di lingkungan Universitas Pertamina yang telah menunjang layanan dan dukungan selama proses pembelajaran di Universitas Pertamina.
6. Keluarga, saudara, dan teman-teman penulis yang telah memberikan doa, kasih sayang, motivasi, dan dorongan di setiap langkah perjuangan penulis.

Penulis menyadari bahwa laporan TA ini belum sempurna dan banyak kekurangan, oleh karena itu, segala saran dan kritik yang membangun diharapkan oleh penulis sebagai penyempurnaan penulisan laporan ini di masa yang mendatang. Semoga laporan TA ini dapat memberikan manfaat bagi penulis dan para pembaca.

Jakarta, 31 Juli 2023

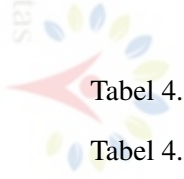
Kiagus Muhammad Arsyad

DAFTAR ISI

KATA PENGANTAR	iii
DAFTAR ISI	iv
DAFTAR TABEL	vi
DAFTAR GAMBAR	vii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah	2
1.4 Tujuan penelitian	2
1.5 Manfaat penelitian	2
BAB II TINJAUAN PUSTAKA	4
2.1 Gas Rumah Kaca (GRK)	4
2.2 <i>Text Mining</i>	5
2.2.1 <i>Natural Language Processing (NLP)</i>	5
2.2.2 <i>Term Frequency-Inverse Document Frequency (TF-IDF)</i>	6
2.2.3 <i>Principal Component Analysis (PCA)</i>	7
2.2.4 <i>K-Means</i>	7
2.3 Teknik Kitchenham	7
2.4 <i>Luhn's Theory</i>	8
BAB III METODE PENELITIAN	11
3.1 Tahapan Penelitian	11
3.2 <i>Data Collection</i>	12
3.3 Tahapan <i>Text Mining</i> dan Interpretasinya	14
BAB IV HASIL DAN PEMBAHASAN	17

4.1	<i>Text Preprocessing</i>	17
4.2	Hasil Pengelompokan Kata Kunci berdasarkan teknik <i>clustering</i>	18
4.3	Hasil Pemilihan Kata Kunci berdasarkan Teori Luhn	21
4.4	Hasil Interpretasi Kata Kunci Berdasarkan Teknik <i>Clustering</i> dan Seleksi Teori Luhn	24
BAB V	KESIMPULAN DAN SARAN	27
5.1	Kesimpulan	27
5.2	Saran	27
	DAFTAR PUSTAKA	29
LAMPIRAN A	Lampiran 1. Tren Total Emisi GRK Indonesia	33





DAFTAR TABEL

Tabel 4.1	Total jumlah kata setiap klaster	22
Tabel 4.2	Hasil Kata Kunci yang Terpilih berdasarkan teknik Clustering dan implementasi teori Luhn	24

1. Dilarang mengutip karya tulis ini, kecuali:
 - a. menyebutkan sumber sesuai kaidah keceandekiaan;
 - b. pengutipan hanya untuk keperluan pendidikan, penulisan karya ilmiah atau penelitian;
 - c. pengutipan tidak merugikan Universitas Pertamina.
2. Dilarang mempublikasikan atau memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa adanya izin dari Universitas Pertamina.



DAFTAR GAMBAR

Gambar 2.1	Diagram <i>Word Frequency</i>	9
Gambar 3.1	Diagram Alir Penelitian	11
Gambar 3.2	Tahapan <i>Data Collection</i> hingga hasil interpretasi	12
Gambar 3.3	Tahapan proses pemilihan data studi literatur	13
Gambar 3.4	Alur tahapan saat dilakukan <i>text mining</i> hingga hasil interpretasi	14
Gambar 4.1	Proses tahapan <i>text preprocessing</i> hingga hasil menggunakan data studi literatur	17
Gambar 4.2	Visualisasi hasil dari PCA	19
Gambar 4.3	<i>Elbow Plot</i> untuk menentukan nilai klaster dengan K-Means	20
Gambar 4.4	Metrik Evaluasi nilai K dengan <i>Davies-Bouldin index</i> dan <i>library ElbowVisualizer</i>	20
Gambar 4.5	Klasterisasi Pewarnaan dengan K-Means	21
Gambar 4.6	Representasi visual frekuensi data dengan <i>word cloud</i> pada <i>Cluster 0</i>	22
Gambar 4.7	Representasi visual frekuensi data dengan <i>word cloud</i> pada <i>Cluster 1</i>	22
Gambar 4.8	Representasi visual frekuensi data dengan <i>word cloud</i> pada <i>Cluster 2</i>	23
Gambar 4.9	Representasi visual frekuensi data dengan <i>word cloud</i> pada <i>Cluster 3</i>	23
Gambar 1.1	Total emisi GRK di Indonesia 2010-2019	33