



ANALISIS DATA PUBLIKASI DENGAN PENDEKATAN BIBLIOMETRIKS DAN *TEXT CLUSTERING* UNTUK MENENTUKAN KRITERIA TINGKAT KEPARAHAN KARHUTLA

MUHAMMAD IQBAL ALVIAN WIDIYARTO



**DEPARTEMEN ILMU KOMPUTER
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.



@Hak cipta milik IPB University

IPB University

PERNYATAAN MENGENAI SKRIPSI DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa skripsi dengan judul “Analisis Data Publikasi Dengan Pendekatan Bibliometriks dan *Text Clustering* Untuk Menentukan Kriteria Tingkat Keparahan Karhutla” 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

Muhammad Iqbal Alvian Widiyarto
G6401201066

- 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.



ABSTRAK

MUHAMMAD IQBAL ALVIAN WIDIYARTO. Analisis Data Publikasi dengan Pendekatan Bibliometriks dan *Text Clustering* Untuk Menentukan Tingkat Keparahan Karhutla. Dibimbing oleh HARI AGUNG ADRIANTO dan IMAS SUKAESIH SITANGGANG.

Kebakaran hutan dan lahan (karhutla) di Indonesia telah menjadi isu lingkungan selama dua dekade terakhir, dengan dampaknya yang melibatkan berbagai aspek. Penilaian dampak kebakaran terutama dilakukan untuk penelitian, perencanaan rehabilitasi hutan, konservasi, dan penegakan hukum. Namun, standar penilaian dampak kebakaran yang digunakan sebagai acuan pengendalian karhutla masih sangat beragam. Penelitian ini bertujuan untuk mengidentifikasi kriteria untuk mengukur tingkat keparahan kebakaran hutan dengan menerapkan algoritma Bibliometriks dan *text mining*. Selain itu, analisis bibliometriks akan dibandingkan dengan hasil dari analisis *clustering*. Penelitian ini membandingkan beberapa algoritma *clustering* seperti *Regular K-Means*, *Bisecting K-Means*, dan *unweighted pair group method with arithmetic mean* (UPGMA) *Clustering* untuk menganalisis data artikel literatur dari tahun 2013 hingga 2023. Setiap algoritma dievaluasi menggunakan *silhouette score* dan jumlah *cluster* yang berbeda. *Regular K-Means* menghasilkan 10 *cluster* dengan *silhouette score* 0,138. *Bisecting K-Means* menghasilkan *silhouette score* 0,067 dengan jumlah *cluster* sebanyak 9. UPGMA *Clustering* menghasilkan *silhouette score* 0,138 dengan 10 *cluster*. Analisis bibliometriks menunjukkan *burned area* dan *peatland* menjadi kriteria yang muncul pada hasil bibliometriks dengan masing-masing berada di kuadran yang beririsan antara 1 dan 4, dan kuadran 1 yang memiliki arti bahwa kriteria terkait merupakan tema yang penting untuk penelitian dikarenakan banyaknya jumlah penelitian terkait dan terdapat hubungan yang besar dengan penelitian lainnya dan kuadran 4 memiliki arti bahwa kriteria terkait merupakan kriteria yang tidak banyak diteliti akan tetapi banyak hubungannya dengan penelitian lainnya. Analisis hasil *clustering* menunjukkan kriteria dan indikator berupa *burned area*, *peatland*, dan *vegetation* sebagai kriteria dan indikator terbanyak untuk hasil dari metode Bibliometriks dan *clustering*.

Kata Kunci: Algoritma K-Means, *Clustering*, Bibliometriks, Kebakaran hutan dan lahan, *Text mining*

ABSTRACT

MUHAMMAD IQBAL ALVIAN WIDIYARTO. Publication data analyst with Bibliometrics and Text Clustering Approach to Identify the Criteria of Forest and Land Fire Severity. Supervised by HARI AGUNG ADRIANTO dan IMAS SUKAESIH SITANGGANG.

Forest and land fires in Indonesia have been an environmental issue for the past two decades, with their impacts involving various aspects. Fire impact assessments are mainly conducted for research, forest rehabilitation planning, conservation and law enforcement. However, fire impact assessment standards used

as a reference for forest and land fire control are still very diverse. This research aims to identify criteria for measuring the severity of forest fires by applying Bibliometrics and text mining algorithms. In addition, the bibliometric analysis will be compared with the results of clustering analysis. This research compares several clustering algorithms such as Regular K-Means, Bisecting K-Means, and unweighted pair group method with arithmetic mean (UPGMA) Clustering to analyze literature article data from 2013 to 2023. Each algorithm was evaluated using different silhouette scores and number of clusters. Regular K-Means resulted in 10 clusters with a silhouette score of 0.138. Bisecting K-Means produced a silhouette score of 0.067 with a cluster count of 9. UPGMA Clustering produced a silhouette score of 0.138 with 10 clusters. Bibliometric analysis shows that burned area and peatland are the criteria that appear in the bibliometric results with each being in quadrants that intersect between 1 and 4, and quadrant 1 which means that related criteria are important themes for research due to the large number of related studies and there is a large relationship with other studies and quadrant 4 means that related criteria are criteria that are not widely researched but have a lot to do with other studies. Analysis of the clustering results shows criteria and indicators such as burned area, peatland, and vegetation as the most criteria and indicators for the results of the Bibliometrics and clustering methods.

Keywords: Clustering, Forest and Land Fire, Bibliometrics, K-Means Algorithm, Text Mining

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

© 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.

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.

ANALISIS DATA PUBLIKASI DENGAN PENDEKATAN BIBLIOMETRIKS DAN *TEXT CLUSTERING* UNTUK MENENTUKAN KRITERIA TINGKAT KEPARAHAN KARHUTLA

MUHAMMAD IQBAL ALVIAN WIDIYARTO

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

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



@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.

Penguji pada ujian Skripsi:

Muhammad Asyhar Agmalaro, S.Si, M.Kom

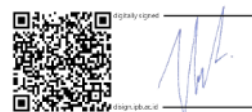
Judul : Analisis Data Publikasi Dengan Pendekatan Bibliometriks dan
Text Clustering Untuk Mengukur Kriteria Tingkat Keparahan
Karhutla

Nama : Muhammad Iqbal Alvian Widiyarto

NIM : G6401201066

Disetujui oleh

Pembimbing 1:
Hari Agung Adrianto, S.Kom, M.Si., Ph.D.



Pembimbing 2:
Prof. Dr. Imas Sukaesih Sitanggang, S.Si., M.Kom.



Diketahui oleh

Ketua Departemen Ilmu Komputer:
Dr. Sony Hartono Wijaya, S.Kom., M.Kom.
19810809 200812 1 002



Tanggal Ujian:
23 Juli 2024

Tanggal Lulus:



PRAKATA

Puji dan syukur penulis panjatkan kepada Allah SWT atas segala karunia-Nya sehingga karya ilmiah ini berhasil diselesaikan. Tema yang dipilih dalam penelitian yang dilaksanakan sejak bulan November 2023 sampai bulan Juli 2024 ini ialah kebakaran hutan dan lahan (karhutla), dengan judul “Analisis Data Publikasi Dengan Pendekatan Bibliometriks dan *Text Clustering* Untuk Mengukur Kriteria Tingkat Keparahan Karhutla”. Terima kasih penulis ucapkan kepada kepada orang tua tercinta yaitu Bapak Helmi Nur Widiyanto dan Ibu Neni Sofiani serta dosen pembimbing Prof. Dr. Imas Sukaesih Sitanggang, S.Si., M.Kom dan Hari Agung Adrianto, S.Kom., M.Si., Ph.D yang telah membimbing dan banyak memberi saran. Ucapan terima kasih juga disampaikan kepada pembimbing akademik, moderator seminar, dan penguji luar komisi pembimbing.

Ungkapan terima kasih penulis juga disampaikan kepada teman-teman Kom ST10, Dcampus, Etherone 57, rekan-rekan AURI dan FPI, beserta orang-orang baik yang sudah menemani dan kebersamai selama kuliah di IPB. Semoga karya ilmiah ini bermanfaat bagi pihak yang membutuhkan dan bagi kemajuan ilmu pengetahuan

(@Hak cipta milik IPB University)

DAFTAR ISI

ABSTRAK	ii
DAFTAR TABEL	vii
DAFTAR GAMBAR	vii
I. PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan	3
1.4 Manfaat	3
1.5 Ruang Lingkup	4
II. TINJAUAN PUSTAKA	4
2.1 Kebakaran Hutan dan Lahan	4
2.2 <i>Literature Review</i>	4
2.3 <i>Text Mining</i>	6
2.4 <i>Clustering</i>	7
2.5 Bibliometriks	9
2.6 <i>Linkage</i>	10
III. METODE	13
3.1 Data Penelitian	13
3.2 Tahapan Penelitian	15
3.3 Lingkungan Pengembangan	16
IV. HASIL DAN PEMBAHASAN	16
4.1 Mengumpulkan Artikel Literatur	16
4.2 Ekstraksi Data	17
4.3 Analisis Bibliometriks	18
4.4 <i>Clustering</i>	20
4.5 Evaluasi dan Komparasi Metode	25
4.5 Analisis Kriteria dan Indikator Tingkat Keparahan Karhutla	30
V. SIMPULAN DAN SARAN	33
5.1 Simpulan	33
5.2 Saran	33
DAFTAR PUSTAKA	35
LAMPIRAN	38



DAFTAR TABEL

1	Tools pendukung <i>literature review</i>	5
2	Term dan frekuensi 10 teratas	17
3	Frekuensi tema penelitian teratas	19
4	Komparasi formula <i>linkage</i> algoritma UPGMA	23
5	Karakteristik dari <i>cluster</i> algoritma <i>regular</i> K-Means	25
6	Karakteristik dari <i>cluster</i> algoritma <i>bisecting</i> K-Means	27
7	Karakteristik dari <i>cluster</i> algoritma UPGMA <i>clustering</i>	29
8	Analisis kriteria tingkat keparahan Karhutla	31

DAFTAR GAMBAR

1	Algoritma <i>Regular</i> K-Means (Banerjee <i>et al.</i> 2015)	7
2	Algoritma <i>bisecting</i> K-Means (Banerjee <i>et al.</i> 2015)	8
3	Algoritma UPGMA (Yoshida <i>et al.</i> 2022)	9
4	Kuadran thematic map dalam analisis bibliometriks (Yu dan Munoz 2020)	10
5	Diagram alur penelitian	13
6	Frekuensi artikel per tahun yang digunakan	16
7	Penelusuran dengan Publish or Perish	17
8	<i>Wordcloud</i> kumpulan artikel literatur	18
9	Hasil <i>text extraction</i>	18
10	Hasil <i>thematic map network</i> dengan menggunakan Biblioshiny	19
11	pembagian kuadran hasil analisis bibliometriks untuk setiap <i>cluster</i>	20
12	<i>Silhouette score</i> dari algoritma <i>regular</i> K-Means	21
13	Sebaran <i>cluster</i> dari algoritma <i>regular</i> K-Means	21
14	<i>Silhouette score</i> dari algoritma <i>Bisecting</i> K-Means	22
15	Sebaran <i>cluster</i> dari algoritma <i>Bisecting</i> K-Means	23
16	<i>silhouette score</i> dari algoritma UPGMA <i>Clustering</i>	24
17	Sebaran <i>cluster</i> dari algoritma UPGMA <i>clustering</i>	24