

**EVALUASI LDA DAN LSA UNTUK TOPIK MODELING
TERKAIT BENCANA ALAM DI INDONESIA**

**SKRIPSI
(Jalur Scientist)**

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi S1 Informatika



Disusun oleh

**MUHAMAD GATOT SUPIADIN
20.11.3771**

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2023**

**EVALUASI LDA DAN LSA UNTUK TOPIK MODELING
TERKAIT BENCANA ALAM DI INDONESIA**

**SKRIPSI
(Jalur Scientist)**

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi S1 Informatika



Disusun oleh

MUHAMAD GATOT SUPIADIN
20.11.3771

Kepada

FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2023

HALAMAN PERSETUJUAN

JALUR SCIENTIST

EVALUASI LDA DAN LSA UNTUK TOPIK MODELING TERKAIT BENCANA
ALAM DI INDONESIA

Yang disusun dan diajukan oleh:

Muhamad Gatot Supiadin

20.11.3771

Telah disetujui oleh dosen pembimbing

Pada tanggal 08 Januari 2024

Dosen Pembimbing



Arif Dwi Laksono, M.Kom

NIK. 190302150

HALAMAN PENGESAHAN

JALUR SCIENTIST

EVALUASI LDA DAN LSA UNTUK TOPIK MODELING TERKAIT BENCANA ALAM DI INDONESIA

yang disusun dan diajukan oleh

Muhamad Gatot Supladin

20.11.3771

Telah dipertahankan di depan Dewan Penguji
pada tanggal 17 Januari 2024

Susunan Dewan Penguji

Nama Penguji

Windha Mega Pradnya D. M.Kom
NIK. 190302185

Arif Akbarul Huda, S.Si, M.Eng
NIK. 190302287

Arifiyanto Hadinegoro, S.Kom, MT
NIK. 190302289

Tanda Tangan



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 17 Januari 2024

DEKAN FAKULTAS ILMU KOMPUTER



Hanif Al Fatta, S.Kom., M.Kom.
NIK. 190302096

HALAMAN PERNYATAAN KONTRIBUSI KARYA

Yang bertanda tangan dibawah ini,

Nama Mahasiswa : **Muhamad Gatot Supiadin**

NIM : **20.11.3771**

Menyatakan bahwa karya dengan judul berikut:

EVALUASI LDA DAN LSA UNTUK TOPIK MODELING TERKAIT BENCANA ALAM DI INDONESIA

Dosen Pembimbing: **Ari Dwi Lukito, M.Kom**

1. Karya adalah benar-benar **ASLI** dan **BELUM PERNAH** diajukan untuk mendapatkan gelar akademik, baik di Universitas Amikom Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya merupakan gagasan, rumusan maupun penelitian yang orisinal dan **SAYA** memiliki **KONTRIBUSI** terhadap tersebut.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka atau Referensi pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab **SAYA**, bukan tanggung jawab Universitas Amikom Yogyakarta.
5. Pernyataan ini **SAYA** buat dengan sesungguhnya, apabila dikemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka **SAYA** bersedia menerima **SANKSI AKADEMIK** dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 04 Januari 2024

Yang menyatakan,



Muhamad Gatot Supiadin

HALAMAN PERSEMBAHAN

Segala puji syukur kepada Allah SWT, yang telah memberikan rahmat dan karunia-Nya sehingga penulis bisa sampai pada tahap ini dan menyelesaikan tugas akhir ini. Pada tugas akhir ini penulis persembahkan kepada :

1. Allah SWT yang telah memberikan rahmat dan hidayah sehingga penulis bisa sampai pada tahap ini.
2. Orang Tua penulis, Bapak Supyadin dan Ibu Masriyah yang selalu memberikan support dan doa, serta kasih sayang dan kebutuhan material yang dibutuhkan dalam menyelesaikan gelar ini.
3. Bapak Arif Dwi Laksito, selaku dosen pembimbing yang telah memberikan support dan banyak membimbing penulis dan masukan sehingga dapat menyelesaikan tugas akhir ini.
4. Bapak Tristanto, selaku dosen wali, yang telah memberikan pengertian dan dukungan dalam menjalani kegiatan perkuliahan.
5. Teman-teman seperjuangan Hanif Al-Irsyad, Muhammad Zhafran Ghaly, Benediktus Satria Mukti, Galih Purnomo, Yoga Irwanda, Satrio Arif, Bima Surya, Achmad Syarif yang telah memberikan wadah dan support dalam menjalani perkuliahan.
6. Segenap keluarga besar Universitas Amikom Yogyakarta yang telah memberikan wadah untuk menimba ilmu perkuliahan.
7. Bapak/Ibu dosen pengajar, yang telah memberikan ilmu dan cerita pengalamannya selama hampir 4 tahun .
8. Seluruh pihak yang telah membantu penulis dalam segala hal, yang tidak bisa penulis sebutkan satu persatu.

KATA PENGANTAR

Alhamdulillah penulis panjatkan puja dan puji syukur kehadiran Allah SWT, berkat rahmat dan hidayah-Nya penulis mampu menyelesaikan tugas akhir skripsi ini yang berjudul "*Evaluasi LDA dan LSA untuk Topik Modeling terkait Bencana Alam di Indonesia*".

Tugas akhir ini merupakan salah satu persyaratan yang dibutuhkan untuk kelulusan jenjang Program Sarjana Strata Satu (S1) Program Studi S1 Informatika pada Universitas Amikom Yogyakarta. Penulis ingin menyampaikan terima kasih yang sangat sungguh kepada :

1. Bapak Prof. Dr. M. Suyanto, M.M. selaku Rektor Universitas Amikom Yogyakarta.
2. Bapak Arif Dwi Laksito, M.Kom selaku dosen pembimbing.
3. Bapak Tristanto Ariaji, M.Kom, selaku dosen wali.
4. Seluruh pihak yang telah mendukung hingga terselesaikannya tugas akhir ini.

Akhir kata, semoga pembuatan tugas akhir penelitian ini dapat memberikan manfaat bagi pembaca dan peneliti lain serta menambah wawasan dan pengetahuan, khususnya dalam bidang topik modeling.

Yogyakarta, 04 Januari 2024

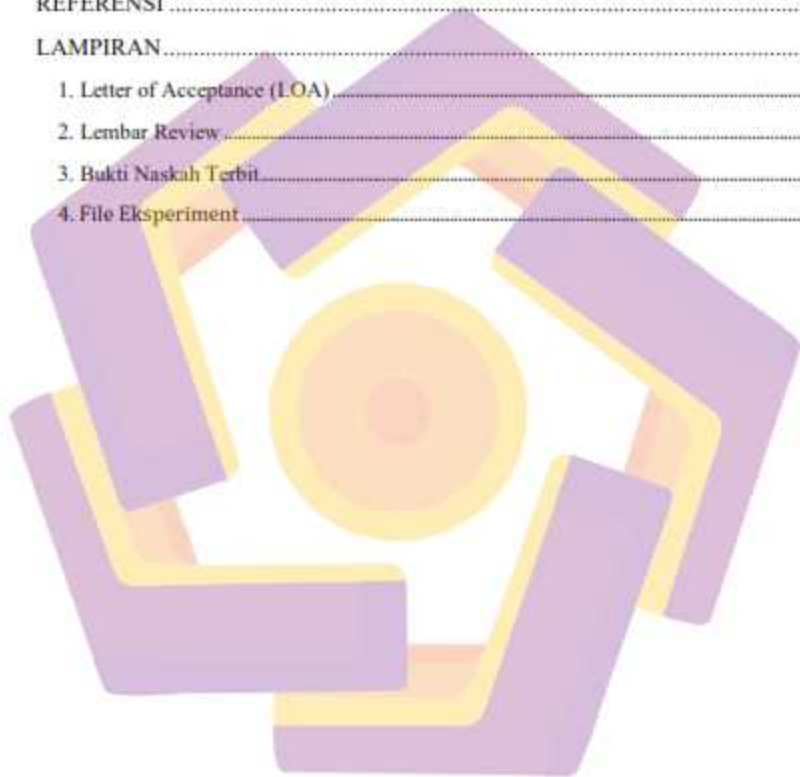


Muhamad Gatot Supiadin

DAFTAR ISI

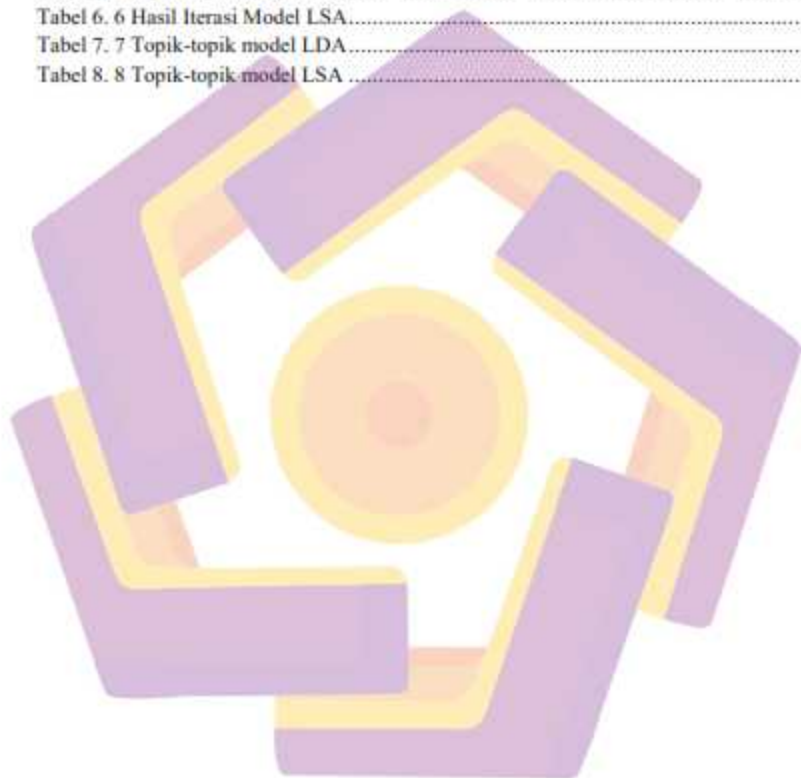
HALAMAN JUDUL.....	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN.....	iii
HALAMAN PERNYATAAN KONTRIBUSI KARYA.....	iv
HALAMAN PERSEMBAHAN.....	iv
DAFTAR ISI.....	vii
DAFTAR TABEL.....	ix
DAFTAR GAMBAR.....	x
INTISARI.....	1
<i>ABSTRACT</i>	2
BAB I PENDAHULUAN.....	3
1.1 Latar Belakang.....	3
BAB II TINJAUAN PUSAKA.....	4
BAB III METODE PENELITIAN.....	5
3.1 Metode Penelitian.....	5
3.2 Scraping Dataset.....	5
3.3 Prapemrosesan Data.....	5
3.4 TF-IDF.....	6
3.5 Latent Semantic Analysis / Indexing.....	6
3.6 Latent Dirichlet Allocation.....	7
3.7 Skor Koherensi.....	9
BAB IV HASIL DAN PEMBAHASAN.....	10
4.1 Pembahasan.....	10
4.2 Dataset Target.....	10
4.3 Scraping Dataset.....	10
4.4 Preprocessing Data.....	10
4.5 TF-IDF.....	12

4.6 Algoritma Pemodelan.....	13
4.7 Visualisasi Model.....	15
BAB V PENUTUP.....	17
5.1 KESIMPULAN.....	17
5.2 UCAPAN TERIMA KASIH.....	17
REFERENSI.....	18
LAMPIRAN.....	20
1. Letter of Acceptance (LOA).....	20
2. Lembar Review.....	20
3. Bukti Naskah Terbit.....	21
4. File Eksperimen.....	21



DAFTAR TABEL

Tabel 1. 1 Parameter LDA Model LDA grafis probabilistik	8
Tabel 2. 2 Data Sampel Sebelum dan Sesudah Pembersihan.....	11
Tabel 3. 3 LDA dengan TF-IDF	13
Tabel 4. 4 LSA dengan TF-IDF	13
Tabel 5.5 Hasil Iterasi Model LDA	14
Tabel 6. 6 Hasil Iterasi Model LSA.....	14
Tabel 7. 7 Topik-topik model LDA.....	14
Tabel 8. 8 Topik-topik model LSA	15



DAFTAR GAMBAR

Gambar 1. 1 Alur Penelitian.....	5
Gambar 2. 2 Model LDA grafis probabilistik.....	8
Gambar 3. 3 Alokasi Dirichlet Laten.....	9
Gambar 4. 4 Skor Koherensi.....	9
Gambar 5. 5 Visualisasi 20 kata yang paling sering muncul.....	12
Gambar 6. 6 Visualisasi sebagian besar kata dalam dataset.....	12
Gambar 7. 7 Visualisasi awan kata dari topik dalam model LDA.....	16
Gambar 8. 8 Visualisasi awan kata dari topik dalam model LSA.....	16
Gambar 9. 9 Letter of Acceptance (LOA).....	20
Gambar 10. 10 Lembar Review.....	20
Gambar 11. 11 Bukti Naskah Terbit.....	21



INTISARI

Pemodelan Topik adalah sebuah metode untuk menganalisis topik, dokumen, dan artikel dalam Pemrosesan Bahasa Alami. Algoritma LDA (Latent Dirichlet Allocation) dan LSA (Latent Semantic Analysis) banyak digunakan dalam pemodelan topik. Penelitian ini berfokus pada analisis artikel yang berkaitan dengan bencana alam dalam bahasa Indonesia.

Dataset untuk penelitian ini diperoleh melalui data scraping dari Google News, yang berfungsi sebagai wadah untuk beberapa artikel dan sumber berita online. Metode penelitian dibagi menjadi beberapa tahap: dataset scraping, preprocessing data, pemodelan topik dengan LDA dan LSA, visualisasi model, dan evaluasi model. Hasil penelitian menunjukkan bahwa kedua algoritma tersebut dapat menghasilkan topik pada dataset yang relevan dengan bencana alam di Indonesia seperti banjir, gempa bumi, longsor, tsunami dan lain sebagainya.

Dari hasil evaluasi menggunakan Coherence Scores, didapatkan bahwa LDA memiliki performa yang lebih baik dibandingkan dengan LSA dilihat dari Coherence Scores dalam memodelkan topik-topik yang berkaitan dengan Bencana Alam di Indonesia. Hasil evaluasi berdasarkan Coherence Scores dalam memodelkan topik terkait Bencana Alam di Indonesia, dengan nilai coherence score sebesar 0.4715 untuk LDA dan 0.4467 untuk LSA.

Kata Kunci: Topik Modeling, LDA, LSA, Dokument, Topik.

ABSTRACT

Topic Modeling is a method for analyzing topics, documents, and articles in Natural Language Processing. The LDA (Latent Dirichlet Allocation) and LSA (Latent Semantic Analysis) algorithms are widely used in topic modelling. This study focuses on analyzing articles related to natural disasters in the Indonesian language.

The dataset for this study was obtained through data scraping from Google News, which served as a container for several articles and online news sources. The research method is divided into several stages: dataset scraping, data preprocessing, topic modelling with LDA and LSA, model visualization, and model evaluation. The results of the research show that both algorithms can generate topics on datasets relevant to natural disasters in Indonesia such as floods, earthquakes, landslides, tsunamis and so on.

From the evaluation results using Coherence Scores, it is found that LDA has better performance than LSA seen from Coherence Scores in modeling topics related to Natural Disasters in Indonesia. Evaluation results, based on Coherence Scores, in modeling topics related to Natural Disasters in Indonesia, with a coherence score of 0.4715 for LDA and 0.4467 for LSA.

Keywords : Topic Modelling, LDA, LSA, Document, Topic.