

**TUGAS AKHIR**

***TOPIC MODELING TUGAS AKHIR MAHASISWA  
FAKULTAS INFORMATIKA MENGGUNAKAN  
LATENT DIRICHLET ALLOCATION***



**NATHANAEL NOVARIAN  
17102112**

**PROGRAM STUDI S1 TEKNIK INFORMATIKA  
FAKULTAS INFORMATIKA  
INSTITUT TEKNOLOGI TELKOM PURWOKERTO  
2023**

**TUGAS AKHIR**

***TOPIC MODELING TUGAS AKHIR MAHASISWA  
FAKULTAS INFORMATIKA MENGGUNAKAN LATENT  
DIRICHLET ALLOCATION***

***TOPIC MODELING OF INFORMATICS FACULTY FINAL  
PROJECT USING LATENT DIRICHLET ALLOCATION***

Disusun Sebagai Salah Satu Syarat untuk Memperoleh Gelar Sarjana Komputer



**NATHANAEL NOVARIAN**

**17102112**

**PROGRAM STUDI S1 TEKNIK INFORMATIKA  
FAKULTAS INFORMATIKA  
INSTITUT TEKNOLOGI TELKOM PURWOKERTO  
2023**

Lembar Persetujuan Pembimbing

***TOPIC MODELING TUGAS AKHIR MAHASISWA  
FAKULTAS INFORMATIKA MENGGUNAKAN  
LATENT DIRICHLET ALLOCATION***

***TOPIC MODELING OF INFORMATICS FACULTY  
FINAL PROJECT USING LATENT DIRICHLET  
ALLOCATION***

Dipersiapkan dan Disusun Oleh

**Nathanael Novarian  
17102112**

**Fakultas Informatika  
Institut Teknologi Telkom Purwokerto  
Pada Tanggal:**

Pembimbing I,



(Siti Khomsah, S.Kom.,M.Cs)  
NIDN. 0517108101

Pembimbing II,



(Amalia Beladonna Arifa, S.Pd.,  
M.Cs)  
NIDN. 0606019201


**TOPIC MODELING TUGAS AKHIR MAHASISWA  
FAKULTAS INFORMATIKA MENGGUNAKAN  
LATENT DIRICHLET ALLOCATION  
TOPIC MODELING OF INFORMATICS FACULTY  
FINAL PROJECT USING LATENT DIRICHLET  
ALLOCATION**

Disusun oleh:

**NATHANAEL NOVARIAN  
17102112**

Telah Diujikan dan Dipertahankan dalam Sidang Ujian Tugas Akhir  
Pada Kamis, 16 Februari 2023.

Penguji I,

  
(Sudianto, S.Pd., M.Kom)

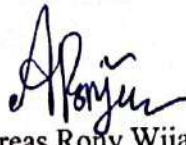
NIDN. 0605049301

Penguji II,

  
(Nia Annisa Ferani T., S.Si.,  
M.Sc)

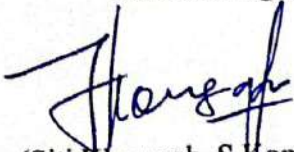
NIDN. 0630049203

Penguji III,

  
(Andreas Rony Wijaya,  
S.Mat., M.Sc)

NIDN. 0630059602

Pembimbing I,

  
(Siti Khomsah, S.Kom., M.Cs)  
NIDN. 051708101

Pembimbing II,

  
(Amalia Beladinna Arifa, S.Pd., M.Cs)  
NIDN. 0606019201

Dekan,

  
(Auliya Buchanuddin, S.Si., M. Kom)  
NIK. 19820008

## HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR

Yang Bertandatangan di bawah ini,

Nama Mahasiswa : Nathanael Novarian

NIM : 17102112

Program Studi : Teknik Informatika

Menyatakan bahwa Tugas Akhir dengan judul berikut:

**TOPIC MODELING TUGAS AKHIR MAHASISWA FAKULTAS  
INFORMATIKA MENGGUNAKAN LATENT DIRICHLET  
ALLOCATION**

Dosen Pembimbing I: Siti Khomsah, S.Kom.,M.Cs

Dosen Pembimbing II: Amalia Beladonna Arifa, S.Pd., M.Cs

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Institut Teknologi Telkom Purwokerto maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan, dan penelitian Saya Sendiri, tanpa bantuan pihak lain kecuali arahan dari Tim Dosen Pembimbing.
3. Dalam Karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggungjawab Saya, bukan tanggungjawab Institut Teknologi Telkom Purwokerto.
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.

Purwokerto, Februari 2023,

Yang Menyatakan,



(Nathanael Novarian)


## KATA PENGANTAR

Puji syukur penulis panjatkan kehadiran Tuhan Yang Maha Esa atas karuniaNya dan atas segala limpahan rahmat Nya, sehingga penulis dapat menyelesaikan laporan penyusunan sidang tugas akhir ini dengan baik.

Penulis mengucapkan terima kasih kepada pihak yang telah membantu dalam penelitian ini, untuk itu penulis dalam kesempatan ini mengucapkan terima kasih kepada:

1. Bapak Dr. Arfianto Fahmi, S.T., M.T., IPM selaku Rektor Institut Teknologi Telkom Purwokerto
2. Bapak Auliya Burhanuddin, S.Si., M. Kom selaku Dekan Fakultas Informatika.
3. Ibu Amalia Beladinna Arifa, S.Pd., M.Cs selaku Kaprodi Fakultas Informatika Institut Teknologi Telkom Purwokerto, Wali Dosen, serta Dosen Pembimbing kedua atas segala bimbingan dan saran kepada penulis sehingga penulis dapat menyelesaikan skripsi ini.
4. Ibu Siti Khomsah, S.Kom., M.Cs. selaku Dosen Pembimbing pertama yang telah memberikan bimbingan, arahan, kesabaran, dan dorongan kepada penulis sehingga penyusunan skripsi ini berjalan dengan baik
5. Orang tua dan keluarga tercinta atas segala bantuan, bimbingan, dorongan serta doa yang diberikan kepada penulis selama penyusunan skripsi.
6. Desi Natalia BR. Sidauruk, S.Kom. atas kesabaran, bantuan, dorongan serta doa yang diberikan kepada penulis selama penyusunan skripsi.

Purwokerto, 7  
Februari 2023

  
Nathanael Novarian

## DAFTAR ISI

Lembar Persetujuan Pembimbing.....	i
Lembar Pengesahan Penguji.....	ii
HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR.....	iii
KATA PENGANTAR.....	iv
DAFTAR ISI .....	v
DAFTAR TABEL .....	vii
DAFTAR GAMBAR.....	viii
ABSTRAK .....	ix
ABSTRACT .....	x
BAB I.....	1
1.1    Latar Belakang Masalah.....	1
1.2    Perumusan Masalah .....	2
1.3    Pertanyaan Penelitian.....	2
1.4    Batasan Masalah .....	3
1.5    Tujuan Penelitian .....	3
1.6    Manfaat Penelitian .....	3
BAB II .....	4
2.1    Kajian Pustaka Penelitian.....	4
2.2    Dasar Teori .....	9
2.2.1 Clustering.....	9
2.2.2 Pemodelan Topik .....	9
2.2.3 Text Preprocessing.....	10
2.2.4 Latent Dirichlet Allocation .....	11

2.2.5 Term Frequency-Inverse Document Frequency .....	12
BAB III .....	14
3.1 Subyek dan Obyek Penelitian .....	14
3.2 Alat dan Bahan Penelitian .....	14
3.2.1 Alat Penelitian .....	14
3.2.2 Bahan Penelitian .....	14
3.3 Diagram Alir Penelitian .....	15
3.3.1 Studi Literatur .....	15
3.3.2 Pengumpulan Data .....	15
3.3.3 Text Preprocessing .....	15
3.3.4 Vektorisasi dengan TF-IDF .....	17
3.3.5 Pembangunan Model LDA .....	20
3.3.6 Evaluasi Model dan Analisis Hasil .....	21
BAB IV .....	22
4.1 Pengumpulan Data .....	22
4.2 Text Preprocessing .....	22
4.3 Vektorisasi .....	25
4.4 Pembangunan Model LDA .....	28
4.5 Evaluasi Model dan Analisis Hasil .....	31
BAB V .....	40
5.1 Kesimpulan .....	40
5.2 Saran .....	40
DAFTAR PUSTAKA .....	41
LAMPIRAN 1. SOURCE CODE DAN DATASET .....	44



## DAFTAR TABEL

Tabel 2.1. Tabel Kajian Pustaka.....	6
Tabel 3 1. Hasil Lower Case .....	16
Tabel 3 2. Hasil Remove Punctuation. ....	17
Tabel 3 3. Hasil Stopword Removal.....	17
Tabel 3 4. Contoh Hasil Text Processing .....	18
Tabel 3 5. Contoh Hasil TF. ....	18
Tabel 3 6. Contoh Hasil TF-IDF .....	19
Tabel 4. 1. Contoh Data Abstraksi. ....	22
Tabel 4. 2. Contoh Lower Case. ....	23
Tabel 4. 3 Contoh Remove Punctuation. ....	23
Tabel 4. 4. Contoh Isi Variabel stop_words. ....	24
Tabel 4. 5. Contoh Hasil Stopword Removal. ....	24
Tabel 4. 6. Contoh Hasil Vektorisasi.....	25
Tabel 4. 7. Hasil Pemodelan LDA.....	28

## DAFTAR GAMBAR

Gambar 3.1. Tahapan Penelitian.....	15
Gambar 3.2. Text Preprocessing.....	16
Gambar 4. 1. Diagram Hasil Coherence Value .....	27
Gambar 4. 2. Hasil Coherence Value. ....	28
Gambar 4. 3. Hasil Visualisasi pyLDAvis Topik 1.....	32
Gambar 4. 4. Hasil Visualisasi pyLDAvis Topik 2.....	33
Gambar 4. 5. Hasil Visualisasi pyLDAvis Topik 3.....	34
Gambar 4. 6. Hasil Visualisasi pyLDAvis Topik 4.....	35
Gambar 4. 7. Hasil Visualisasi pyLDAvis Topik 5.....	36
Gambar 4. 8. Hasil Visualisasi pyLDAvis Topik 6.....	37
Gambar 4. 9. Hasil Visualisasi pyLDAvis Topik 7.....	38
Gambar 4. 10. Hasil Visualisasi pyLDAvis Topik 8.....	39