

**ANALISIS TOPIK MODELLING AKUN TWITTER
DETIK.COM MENGGUNAKAN METODE LATENT
DIRICHLET ALLOCATION**

SKRIPSI

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



disusun oleh

ANGGA RENI DWI ASTUTI

19.11.2657

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2023

**ANALISIS TOPIK MODELLING AKUN TWITTER
DETIK.COM MENGGUNAKAN METODE LATENT
DIRICHLET ALLOCATION**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

ANGGA RENI DWI ASTUTI

19.11.2657

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2023

HALAMAN PERSETUJUAN

SKRIPSI

**ANALISIS TOPIK MODELLING AKUN TWITTER DETIK.COM
MENGUNAKAN METODE LATENT DIRICHLET ALLOCATION**

yang disusun dan diajukan oleh

Angga Reni Dwi Astuti

19.11.2657

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 17 Maret

Dosen Pembimbing,



Nuri Cahyono, M.Kom
NIK. 190302278

HALAMAN PENGESAHAN
SKRIPSI
ANALISIS TOPIK MODELLINGAKUN TWITTER DETIK.COM
MENGGUNAKAN METODE LATENT DIRICHLET ALLOCATION

yang disusun dan diajukan oleh

Angga Reni Dwi Astuti

19.11.2657

Telah dipertahankan di depan Dewan Penguji
pada tanggal 17 Maret 2023

Susunan Dewan Penguji

Nama Penguji

Andriyan Dwi Putra, M.Kom
NIK. 190302270

Anggit Ferdita Nugraha, S.T., M.Eng
NIK. 190302480

Nuri Cahyono, M.Kom
NIK. 190302278

Tanda Tangan



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 17 Maret 2023

DEKAN FAKULTAS ILMU KOMPUTER



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

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Angga Reni Dwi Astuti
NIM : 19.11.2657

Menyatakan bahwa Skripsi dengan judul berikut:

ANALISIS TOPIK MODELLING AKUN TWITTER DETIK.COM MENGUNAKAN METODE LATENT DIRICHLET ALLOCATION

Dosen Pembimbing : Nuri Cahyono, M.Kom

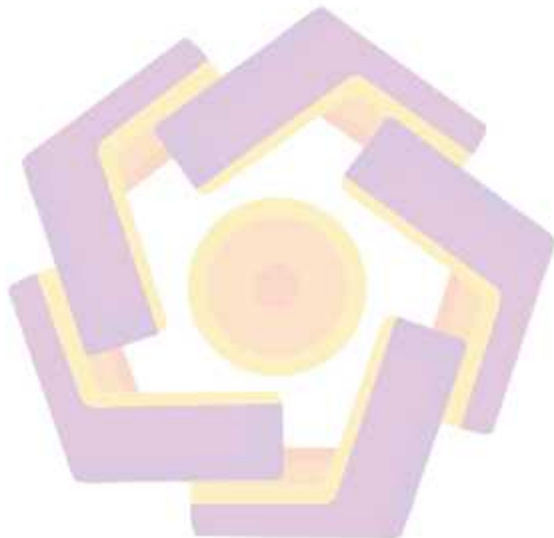
1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen 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 tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian 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, 17 Maret 2023

Yang Menyatakan,



Angga Reni Dwi Astuti



HALAMAN PERSEMBAHAN

Segala puji dan syukur kepada Tuhan Yesus Kristus yang telah memberikan rahmat, pertolongan dan anugerah-Nya melalui orang – orang yang membantu membimbing dan mendukung dengan berbagai cara sehingga penulis dapat menyelesaikan penulisan skripsi ini. Oleh karena itu pada kesempatan ini penulis ingin mempersembahkan skripsi yang telah penulis susun kepada :

1. Bapak dan Mam tercinta, yang selalu memberikan cinta, kasih sayang dan doa restu yang tidak ada hentinya kepada penulis, serta untuk ketersediaannya menjadi donatur terbesar dalam kehidupan penulis baik ilmu dan materi.
2. Terimakasih juga kepada kakak penulis yang juga menjadi donatur terbesar kedua selain kedua orang tua, dan untuk adik penulis yang selalu memberikan dukungan, selalu berdoa bersama, memberikan keceriaan serta semangat dalam menyelesaikan penulisan skripsi ini.
3. Tentunya terimakasih kepada bapak Nuri Cahyono, M.kom selaku dosen pembimbing penulis. Terimakasih sudah menjadi dosen pembimbing bagi penulis, untuk kesabarannya dan bimbingannya selama ini sehingga penulis dapat menyelesaikan penulisan ini berkat masukan dan juga bimbingan bapak selama ini.
4. Kepada teman – teman penulis, terimakasih kepada sela inum yang menjadi teman 24/7 penulis yang menjadi pendengar terbaik, terimakasih kepada Raihan, Hafid, Fahmi, Reza yang selalu mengerjakan dan memotivasi dalam menulis skripsi, Darmaji yang menjadi teman lintas jurusan penulis.
5. Grup ciwi – ciwi Informatika 01, yang sudah menjadi teman seperjuangan selama menempuh pendidikan di Universitas Amikom.
6. Semua orang-orang yang tidak dapat saya sebutkan nama – namanya satu persatu, penulis ucapkan terimakasih banyak atas dukungan dan suport kalian selama ini.

Penulis ucapkan terimakasih yang sebesar – besarnya untuk kalian semua. Mohon maaf jika ada salah kata atau perbuatan baik yang disengaja maupun tidak disengaja selama ini. Semoga Tuhan yang Maha Esa memberikan rahmat dan anugerah-Nya kepada kita semua. Serta semoga ilmu yang telah penulis dapat bermanfaat untuk banyak orang.



KATA PENGANTAR

Bagian ini berisi pernyataan resmi yang ingin disampaikan oleh penulis kepada pihak lain, misalnya ucapan terima kasih kepada Dosen Pembimbing, Tim Dosen Penguji, dan semua pihak yang terkait dalam penyelesaian skripsi termasuk orang tua dan penyanggah dana.

Segala puji dan syukur hanya bagi Tuhan Yesus Kristus, oleh karena anugerah-Nya yang melimpah, kemuliaan dan kasih setia yang besar akhirnya penulis dapat menyelesaikan proses penulisan skripsi yang berjudul “ Analisis Topic Modelling Akun Twitter Detik.Com Menggunakan Metode Latent Dirichlet Allocation”. Skripsi ini disusun sebagai salah satu syarat untuk mendapatkan gelar Sarjana Komputer dari Program Studi S1 Informatika, Fakultas Ilmu Komputer, Universitas Amikom Yogyakarta.

Pada kesempatan ini, penulis mengucapkan terimakasih kepada semua pihak yang telah memberikan dukungan, baik dalam proses penelitian, penyusunan, dan dukungan moral, kepada :

1. Bapak Prof. Dr. M. Suyanto, M.M selaku Rektor Universitas Amikom Yogyakarta.
2. Bapak Nuri Cahyono, M.Kom selaku dosen pembimbing yang telah membimbing penulis dalam proses penulisan skripsi.
3. Dosen penguji yang telah memberikan masukan terhadap penelitian ini.
4. Kedua orang tua dan keluarga inti penulis yang selalu memberikan doa, dukungan serta semangat kepada penulis.
5. Keluarga besar Informatika 01 angkatan 2019.
6. Serta semua pihak yang namanya tidak dapat disebutkan satu – persatu yang telah membantu dalam penyusunan skripsi ini.

Penulis menyadari bahwa masih terdapat banyak kekurangan dalam mengerjakan skripsi ini, sehingga penulis mengharapkan adanya saran dan juga kritik yang membangun dalam kesempurnaan skripsi ini. Akhir kata , penulis

berharap semoga skripsi ini dapat bermanfaat, baik bagi penulis dan juga untuk penelitian yang akan datang.

Yogyakarta, 17 Maret 2023

Angga Reni Dwi Astuti



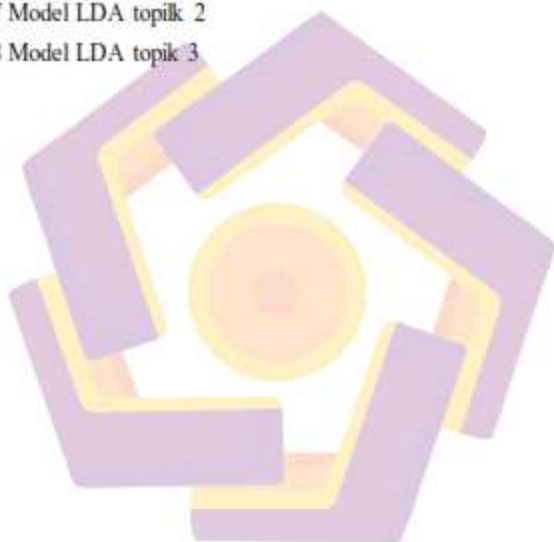
DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN.....	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI.....	iv
HALAMAN PERSEMBAHAN.....	vi
KATA PENGANTAR.....	viii
DAFTAR ISI.....	x
DAFTAR TABEL.....	xiii
DAFTAR GAMBAR.....	xiv
DAFTAR LAMPIRAN.....	xv
DAFTAR LAMBANG DAN SINGKATAN.....	xvi
INTISARI.....	xvii
ABSTRACT.....	xviii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	4
1.3 Batasan Masalah.....	4
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	4
1.6 Sistematika Penulisan.....	5
BAB I PENDAHULUAN.....	5
BAB II TINJAUAN PUSTAKA.....	5
BAB III METODE PENELITIAN.....	5
BAB IV HASIL DAN PEMBAHASAN.....	5
BAB V PENUTUP.....	5
BAB II TINJAUAN PUSTAKA.....	7
2.1 Studi Literatur.....	7
2.2 Dasar Teori.....	22
2.2.1 Media Sosial.....	22
2.2.2 Twitter.....	24
2.2.3 Portal Berita.....	28
2.2.4 Twint.....	28
2.2.5 Google Collab.....	28
2.2.6 Gensim.....	29
2.2.7 Python.....	29
2.2.8 Text Mining.....	30
2.2.9 Preprocessing.....	31

2.2.10 Topic Modelling	32
2.2.11 Latent Dirichlet Allocation (LDA).....	34
2.2.12 Coherence Topic	36
2.2.13 Word Cloud	37
BAB III METODE PENELITIAN	39
3.1 Objek Penelitian	39
3.2 Alur Penelitian	39
3.2.1 Pengumpulan data	40
3.2.2 Studi Pustaka	41
3.2.4 Preprocessing data	42
3.2.5 Pemodelan data.....	43
BAB IV HASIL DAN PEMBAHASAN.....	44
4.1 Pengumpulan data.....	44
4.2 Text processing.....	45
4.2.1 Case Folding.....	45
4.2.2 Stopword	47
4.2.3 Stemming	49
4.2.4 Tokenizing	50
4.3 Membangun model LDA.....	51
4.4 Visualisasi Pemodelan topic.....	51
4.4.1 Model LDA Topic 1	51
4.4.2 Model LDA Topic 2	53
4.4.3 Model LDA Topic 3.....	55
4.5 Coherence Topic.....	56
4.6 Word Cloud.....	57
BAB V PENUTUP.....	60
5.1 Kesimpulan.....	60
5.2 Saran	60
REFERENSI	62
LAMPIRAN.....	67

DAFTAR TABEL

Tabel 2.1. keaslian penelitian	13
Tabel 4.2 Case Fokding	46
Tabel 4.3 Stopword	48
Tabel 4.4 Stemming	49
Tabel 4.5 Tokenizing	50
Tabel 4.6 Model LDA topik 1	52
Tabel 4.7 Model LDA topik 2	54
Tabel 4.8 Model LDA topik 3	56



DAFTAR GAMBAR

Gambar 2.1. LDA	36
Gambar 3.2 Diagram Alir	40
Gambar 3.3 Scraping data	41
Gambar 4.4 Detikcom.csv	44
Gambar 4.5 LDA topik 1	53
Gambar 4.6 LDA topik 2	55
Gambar 4.7 LDA topik 3	56
Gambar 4.8 Coherence	57
Gambar 4.9 Word cloud 1	58
Gambar 4.10 Word cloud 2	58
Gambar 4.11 Word cloud New	59
Gambar 4.12 Word cloud New II	59



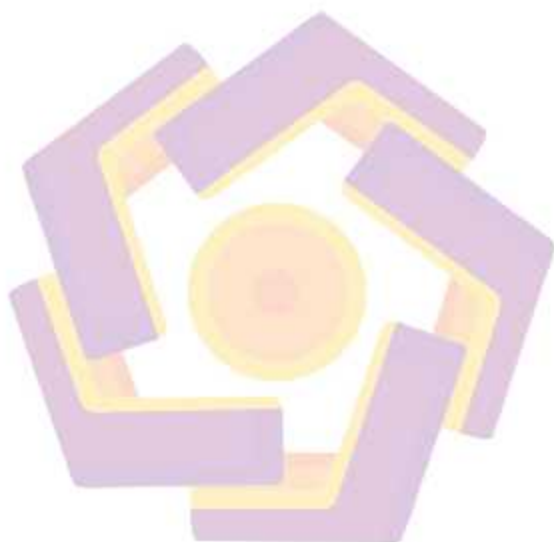
DAFTAR LAMPIRAN

Lampiran 1. Data mentah	67
Lampiran 2. Data setelah preprocessing	74



DAFTAR LAMBANG DAN SINGKATAN

LDA	Latent Dirichlet Allocation
SVM	Support Vector Machines



INTISARI

Kemajuan teknologi tidak diragukan lagi memiliki dampak besar pada media informasi. Salah satu dampak kemajuan teknologi adalah keberadaan media berita sebagai sumber informasi publik. Ada juga informasi daerah, baik dalam maupun luar negeri, dan tentunya ada berbagai pembahasan. Data berita dari portal berita online dapat dijadikan sebagai sumber informasi sekaligus sumber penelitian dan analisis. Tentu saja, portal berita mencakup semua jenis berita tentang berbagai topik tertentu. Mengidentifikasi topik yang sering dibahas di portal berita pasti akan memakan banyak waktu. Oleh karena itu, penelitian ini berfokus pada penerapan sistem Pemodelan topik untuk mengimplementasikan sistem keputusan topik berita dengan menggunakan metode *Latent Dirichlet Allocation* (LDA). Penelitian ini sukses menerapkan metode *Latent Dirichlet Allocation* (LDA) dalam menentukan topik berita, yang mana terdapat tiga topik kategori yang sering dibahas pada portal berita online Detik.com. Topik 1 berisi kejadian bencana alam, topik 2 berisi tokoh dan permasalahan politik, topik 3 berisi berita seputar piala dunia.

Kata kunci: Topik Modelling, *Latent Dirichlet Allocation*, Berita, Text Mining.



ABSTRACT

Technological advances have undoubtedly had a major impact on information media. One impact of technological progress is the existence of news media as a source of public information. There is also regional information, both domestic and foreign, and of course there are various discussions. News data from online news portals can be used as a source of information as a source of research and analysis. Of course, news portals cover all types of news on various topics. Identifying frequently discussed topics on news portal will definitely take a lot of time. Therefore, this research focuses on applying a topic modelling system to implement a news topic decision system using the Latent Dirichlet Allocation (LDA) method. This research successfully applies the Latent Dirichlet Allocation (LDA) method in determining news topics, of which there are three topic categories that are often discussed on the online news portal Detik.com. topic 1 contains natural disaster event, topik 2 contains political figures and issues, topik 3 contains news about the world cup.

Keyword: *Topic Modelling, Latent Dirichlet Allocation, News, Text Mining.*

