



SKRIPSI

KLASIFIKASI KASUS KEKERASAN PEREMPUAN DI KABUPATEN JEMBER MENGGUNAKAN LATENT DIRICHLET ALLOCATION DAN ZERO-SHOT CLASSIFICATION

ATHAYA AQILAH

NPM 20081010306

DOSEN PEMBIMBING

Fetty Tri Anggraeny, S.Kom., M.Kom.

Muhammad Muharrom Al Haromainy, S.Kom, M.Kom

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA
SURABAYA
2025**



SKRIPSI

**KLASIFIKASI KASUS KEKERASAN
PEREMPUAN DI KABUPATEN JEMBER
MENGUNAKAN LATENT DIRICHLET
ALLOCATION DAN ZERO-SHOT
CLASSIFICATION**

ATHAYA AQILAH
NPM 21081010306

DOSEN PEMBIMBING

Fetty Tri Anggraeny, S.Kom., M.Kom.

Muhammad Muharrom Al Haromainy, S.Kom., M.Kom.

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA
SURABAYA
2025**

Halaman ini sengaja dikosongkan

LEMBAR PENGESAHAN

KLASIFIKASI KASUS KEKERASAN PEREMPUAN DI KABUPATEN JEMBER MENGGUNAKAN LATENT DIRICHLET ALLOCATION DAN ZERO-SHOT CLASSIFICATION

Oleh:
ATHAYA AQILAH
NPM. 21081010306

Telah dipertahankan dihadapan dan diterima oleh Tim Penguji Skripsi Prodi Informatika
Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jawa Timur Pada
tanggal 01 September 2025.

Menyetujui

Fetty Tri Anggraeny, S.Kom., M.Kom.

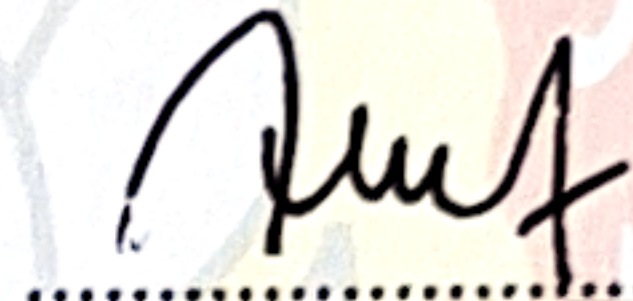
NIP. 19820211 202121 2 005



(Pembimbing I)

M. Muharrom Al Haromainy, S.Kom., M.Kom.

NIP. 19950601 202203 1 006



(Pembimbing II)

Henni Endah Wahanani, ST, M.Kom.

NIP. 19780922 202121 2 005



(Ketua Penguji)

Retno Mumpuni, S.Kom, M.Sc.

NPT. 172198 70 716054



(Anggota Penguji)

Mengetahui
Dekan Fakultas Ilmu Komputer


Prof. Dr. Ir. Novirina Hendrasarie, MT
NIP. 196811261994032001

Halaman ini sengaja dikosongkan

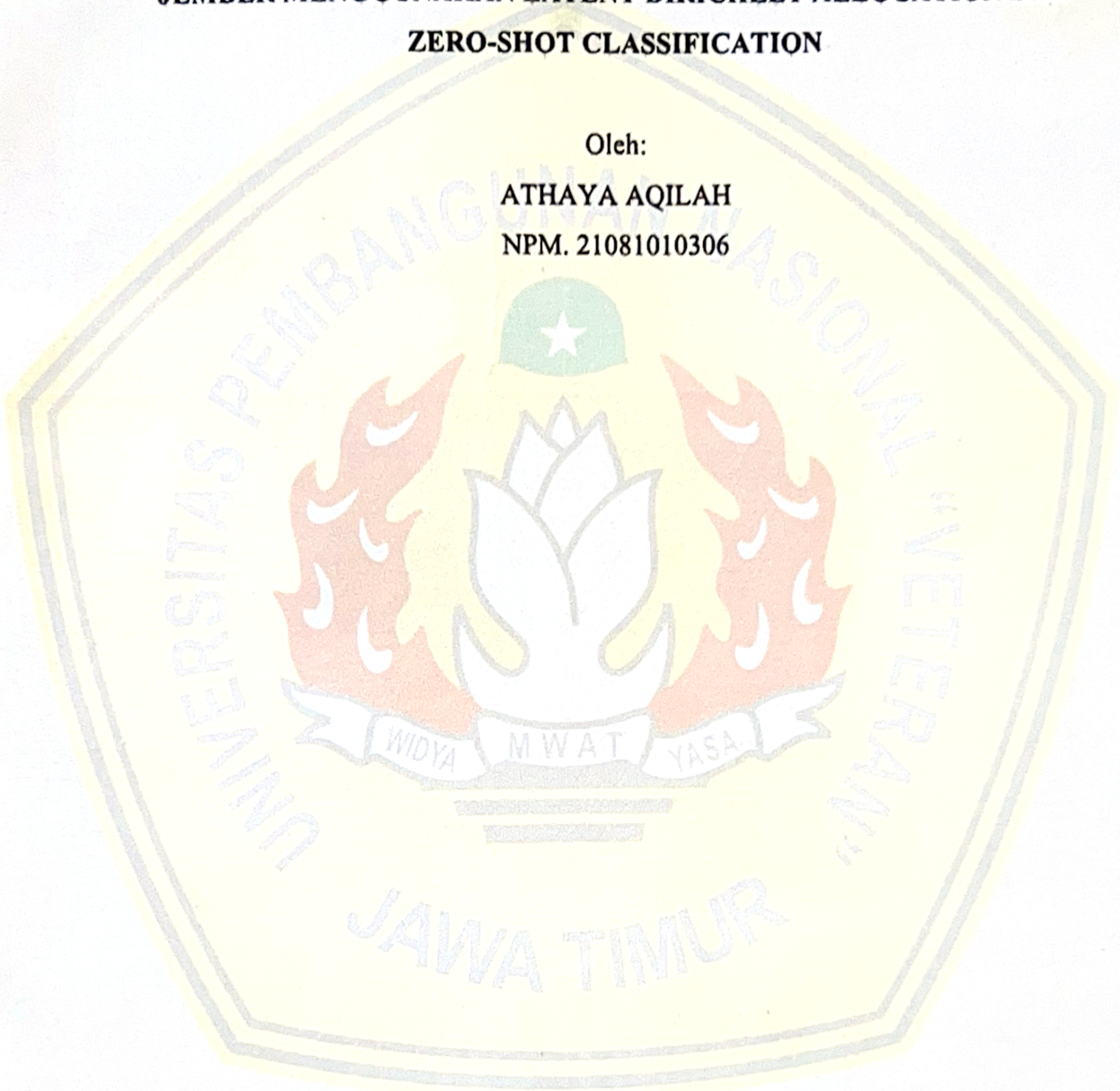
LEMBAR PERSETUJUAN

**KLASIFIKASI KASUS KEKERASAN PEREMPUAN DI KABUPATEN
JEMBER MENGGUNAKAN LATENT DIRICHLET ALLOCATION DAN
ZERO-SHOT CLASSIFICATION**

Oleh:

ATHAYA AQILAH

NPM. 21081010306



**Menyetujui,
Koordinator Program Studi Informatika
Fakultas Ilmu Komputer**

Fetty Tri Anggraeny S.Kom. M.Kom

NIP. 19820211 2021212 005

Halaman ini sengaja dikosongkan

SURAT PERNYATAAN BEBAS PLAGIASI

Saya yang bertanda tangan dibawah ini:

Nama : Athaya Aqilah
NPM : 21081010306
Program : Sarjana (S1)
Program Studi : Informatika
Fakultas : Ilmu Komputer

Menyatakan bahwa dalam dokumen ilmiah Skripsi ini tidak terdapat bagian dari karya ilmiah lain yang telah diajukan untuk memperoleh gelar akademik di suatu lembaga Pendidikan Tinggi, dan juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang/lembaga lain, kecuali yang secara tertulis disitasi dalam dokumen ini dan disebutkan secara lengkap dalam daftar pustaka.

Dan saya menyatakan bahwa dokumen ilmiah ini bebas dari unsur-unsur plagiasi. Apabila dikemudian hari ditemukan indikasi plagiat pada Skripsi ini, saya bersedia menerima sanksi sesuai dengan peraturan perundang-undangan yang berlaku.

Demikian surat pernyataan ini saya buat dengan sesungguhnya tanpa ada paksaan dari siapapun juga dan untuk dipergunakan sebagaimana mestinya.

Surabaya, 10 September 2025

Yang Membuat Pernyataan



Athaya Aqilah

NPM. 21081010306

Halaman ini sengaja dikosongkan

ABSTRAK

Nama Mahasiswa / NPM : Athaya Aqilah / 21081010306
Judul Skripsi : Klasifikasi Kasus Kekerasan Perempuan di
Kabupaten Jember Menggunakan Latent Dirichlet
Allocation dan Zero-Shot Classification
Dosen Pembimbing : 1. Fetty Tri Anggraeny, S.Kom., M.Kom.
2. M. Muharrom Al Haromainy, S.Kom., M.Kom.

Penelitian ini bertujuan untuk mengklasifikasikan kasus kekerasan terhadap perempuan di Kabupaten Jember menggunakan pendekatan Latent Dirichlet Allocation (LDA) dan Zero-Shot Classification (ZSC). Data penelitian berupa 300 laporan naratif korban yang diperoleh dari Dinas Pemberdayaan Perempuan, Perlindungan Anak, dan Keluarga Berencana (DP3AKB) Kabupaten Jember periode 2018–2024. Proses penelitian meliputi preprocessing teks, pemodelan topik dengan LDA, serta klasifikasi laporan menggunakan ZSC berbasis XLM-RoBERTa. Hasil evaluasi menunjukkan bahwa LDA mampu menghasilkan 16 topik utama dengan nilai coherence score terbaik sebesar 0,65, yang menggambarkan pola kekerasan fisik, psikis, seksual, dan verbal. Sementara itu, ZSC berhasil mengklasifikasikan dokumen ke dalam empat kategori kekerasan dengan performa yang cukup baik, ditunjukkan oleh nilai akurasi 90%, presisi 90%, recall 90%, dan F1-score 90%. Temuan ini membuktikan bahwa kombinasi LDA dan ZSC efektif dalam membantu lembaga terkait mengelola laporan kekerasan secara sistematis, serta dapat menjadi dasar pengembangan sistem analisis berbasis web untuk mendukung kebijakan perlindungan perempuan di tingkat daerah.

Kata kunci : Kekerasan Terhadap Perempuan, *Latent Dirichlet Allocation*, *Zero-Shot Classification*, NLP, *Bayesian Optimization*.

Halaman ini sengaja dikosongkan

ABSTRACT

Student Name / NPM : Athaya Aqilah / 21081010306
Thesis Title : Classification of Violence Against Women Cases
in Jember Regency Using Latent Dirichlet
Allocation and Zero-Shot Classification
Advisor : 1. Fetty Tri Anggraeny, S.Kom., M.Kom.
2. M. Muharrom Al Haromainy, S.Kom., M.Kom.

This study aims to classify cases of violence against women in Jember Regency using Latent Dirichlet Allocation (LDA) and Zero-Shot Classification (ZSC). The dataset consisted of 300 narrative reports collected from the Department of Women's Empowerment, Child Protection, and Family Planning (DP3AKB) of Jember during 2018–2024. The research process involved text preprocessing, topic modeling with LDA, and classification using ZSC based on XLM-RoBERTa. The evaluation results indicate that LDA successfully generated 16 main topics with the best coherence score of 0.65, reflecting patterns of physical, psychological, sexual, and verbal violence. Meanwhile, ZSC effectively classified documents into four violence categories, achieving an accuracy of 90%, precision of 90%, recall of 90%, and an F1-score of 90%. These findings demonstrate that the integration of LDA and ZSC is effective in supporting local agencies to systematically manage violence reports and provide a reliable foundation for developing a web-based analytical system to strengthen women's protection policies at the regional level.

Keywords: Violence Against Women, Latent Dirichlet Allocation, Zero-Shot Classification, NLP, Bayesian Optimization.

Halaman ini sengaja dikosongkan

KATA PENGANTAR

Puji syukur kehadiran Allah SWT atas segala rahmat, hidayah dan karunia-Nya kepada penulis sehingga skripsi dengan judul **“KLASIFIKASI KASUS KEKERASAN PEREMPUAN DI KABUPATEN JEMBER MENGGUNAKAN LATENT DIRICHLET ALLOCATION DAN ZERO-SHOT CLASSIFICATION”** dapat terselesaikan dengan baik.

Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer pada Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional “Veteran” Jawa Timur. Penyusunan skripsi ini tidak lepas dari bantuan, bimbingan, dan dukungan dari berbagai pihak. Oleh karena itu, penulis mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Ibu Prof. Dr. Ir. Novirina Hendrasarie, MT, selaku Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur.
2. Ibu Fetty Tri Anggraeny, S.Kom., M.Kom., selaku Ketua Program Studi Informatika Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur.
3. Ibu Fetty Tri Anggraeny, S.Kom., M.Kom. dan Bapak M. Muharrom Al Haromainy, S.Kom., M.Kom., selaku Dosen Pembimbing yang telah memberikan arahan, bimbingan, serta motivasi kepada penulis selama proses penyusunan skripsi ini.
4. Ibu Henni Endah Wahanani, ST. M.Kom. dan Ibu Retno Mumpuni, S.Kom., M.Sc., selaku Dosen Penguji yang telah memberikan masukan, kritik, serta saran yang sangat berharga demi kesempurnaan skripsi ini.
5. Seluruh dosen Program Studi Informatika yang telah memberikan ilmu, bimbingan, dan inspirasi selama proses perkuliahan yang menjadi bekal penting dalam penyusunan skripsi ini.
6. Kedua orang tua tercinta, ayah dan ibu, yang senantiasa memberikan doa, semangat, kasih sayang, serta dukungan moral dan material yang tak ternilai harganya.
7. Kepada kakak, Septia Rana Salsabila, yang telah menjadi sumber motivasi dan kekuatan.

8. Seseorang yang istimewa dalam hidup penulis, Adam Rifqi Permana, yang selalu setia mendampingi, memberi motivasi, dan semangat tanpa henti. Kehadiran dan dukunganmu menjadi kekuatan besar yang membantu penulis melewati masa-masa sulit selama proses penyusunan skripsi ini.
9. Azila Lailannafisa dan Bintang Tiara Pramesti, selaku sahabat saya yang selalu menemani saya dari awal perkuliahan hingga akhir.
10. Sahabat terdekat yang lainnya, Arab, Farras, Bagas, Aqiilah, Hilmi, Mahen, dan Samuel, yang turut memberikan motivasi dan mewarnai hari saya selama di perkuliahan.
11. Kepada Santi Lestari, yang setia menemani penulis di saat-saat susah.
12. Kepada teman-teman terdekat penulis lainnya yang senantiasa memberikan semangat dikala penulis dalam keadaan sedih serta selalu mengajak penulis untuk bermain agar tidak terlalu stress memikirkan masalah yang dihadapi.
13. Semua pihak yang tidak bisa disebutkan satu-persatu yang telah banyak membantu memberikan pemikiran demi kelancaran dan keberhasilan penulisan skripsi ini.

Penulis menyadari bahwa di dalam penyusunan skripsi ini banyak terdapat kekurangan. Untuk itu kritik dan saran yang membangun dari semua pihak sangat diharapkan demi kesempurnaan penulisan skripsi ini. Akhirnya, dengan segala keterbatasan yang penulis miliki semoga laporan ini dapat bermanfaat bagi semua pihak umumnya dan penulis pada khususnya.

Surabaya, 10 September 2025

Penulis,



Athaya Aqilah
NPM. 21081010306

DAFTAR ISI

LEMBAR PENGESAHAN	iii
LEMBAR PERSETUJUAN	v
SURAT PERNYATAAN BEBAS PLAGIASI.....	vii
ABSTRAK	ix
ABSTRACT	xi
KATA PENGANTAR.....	xiii
DAFTAR ISI.....	xv
DAFTAR GAMBAR.....	xix
DAFTAR TABEL.....	xxi
BAB I PENDAHULUAN	1
1.1 Latar Belakang	2
1.2 Rumusan Masalah	5
1.3 Tujuan Penelitian	5
1.4 Manfaat Penelitian.....	5
1.5 Batasan Masalah.....	6
BAB II TINJAUAN PUSTAKA.....	8
2.1 Penelitian Pendahulu.....	9
2.2 Landasan Teori.....	12
2.2.1 Kekerasan Terhadap Perempuan	12
2.2.2 Dinas Pemberdayaan Perempuan Perlindungan Anak Kabupaten Jember ..	
.....	13
2.2.3 Natural Language Processing (NLP).....	14
2.2.4 <i>Topic Modeling</i>	15
2.2.5 <i>Bag of Words</i> (BoW).....	15
2.2.6 <i>Bayesian Optimization</i>	16

2.2.7 Gibbs Sampling.....	16
2.2.8 Latent Dirichlet Allocation (LDA).....	17
2.2.9 Zero-Shot Classification (ZSC).....	19
2.2.10 Coherence Score.....	20
2.2.11 Confusion Matrix.....	21
BAB III DESAIN DAN IMPLEMENTASI SISTEM.....	23
3.1 Tahapan Penelitian.....	23
3.2 Pengumpulan Dataset.....	25
3.3 Data Preprocessing.....	26
3.3.1 Case Folding.....	26
3.3.2 Cleansing Data.....	27
3.3.3 Tokenizing.....	28
3.3.4 Stopword Removal.....	30
3.3.5 Stemming.....	31
3.3.5 Pembagian Dataset.....	33
3.4 Hyperparameter Tuning.....	34
3.5 Implementasi Metode.....	35
3.6 Evaluasi.....	42
3.7 Skenario Pengujian.....	43
3.8 Perancangan Arsitektur Website.....	44
BAB IV HASIL DAN PEMBAHASAN.....	47
4.1 Spesifikasi Hardware Dan Software.....	47
4.2 Import Library.....	47
4.3 Pengumpulan Dataset.....	50
4.4 Data Preprocessing.....	51
4.4.1 Case Folding.....	51

4.4.2 <i>Cleansing</i>	53
4.4.3 <i>Tokenizing</i>	54
4.4.4 <i>Stopword Removal</i>	56
4.4.5 <i>Stemming</i>	57
4.5 Arsitektur Model	59
4.5.1 <i>Latent Dirichlet Allocation (LDA)</i>	60
4.5.2 <i>Zero-Shot Classification</i>	61
4.6 Pelatihan Model	65
4.6.1 Pelatihan Model LDA	66
4.6.2 Pelatihan Model ZSC	67
4.7 Skenario Pengujian.....	69
4.7.2 Skenario Pengujian <i>Bayesian Optimization</i>	69
4.7.2 Skenario Perbandingan Model	73
4.8 Evaluasi Model.....	74
4.8.1 Evaluasi Hasil LDA	74
4.8.2 Visualisasi <i>Word Cloud</i>	79
4.8.3 Evaluasi Hasil Zero-Shot Classification	91
4.8.4 Confusion Matrix	93
4.9 Implementasi Website.....	94
BAB V PENUTUP	98
5.1 Kesimpulan	99
5.2 Saran.....	100
DAFTAR PUSTAKA	101
LAMPIRAN	104

Halaman ini sengaja dikosongkan

DAFTAR GAMBAR

Gambar 2. 1 Logo DP3AKB.....	13
Gambar 2. 2 Struktur Organisasi Dinas DP3AKB.....	14
Gambar 2. 3 Diagram LDA.....	18
Gambar 3. 1 Alur Penelitian.....	23
Gambar 3. 2 Alur Data Preprocessing.....	26
Gambar 3. 3 Alur Case Folding	26
Gambar 3. 4 Alur Cleansing Data	27
Gambar 3. 5 Alur Data Tokenizing.....	29
Gambar 3. 6 Alur Data Stopword Removal	30
Gambar 3. 7 Alur Data Stemming.....	32
Gambar 3. 8 Alur Hyperparameter Tuning	34
Gambar 3. 9 Alur LDA	35
Gambar 3. 10 Alur <i>Gibbs Sampling</i>	38
Gambar 3. 11 Alur Zero-Shot Classification	40
Gambar 3. 12 Flowchart Arsitektur Website	44
Gambar 3. 13 Wireframe Website	45
Gambar 4. 1 Word Cloud Topik 1	79
Gambar 4. 2 Word Cloud Topik 2	80
Gambar 4. 3 Word Cloud Topik 3	81
Gambar 4. 4 Word Cloud Topik 4	81
Gambar 4. 5 Word Cloud 4.5.....	82
Gambar 4. 6 Word Cloud Topik 6	82
Gambar 4. 7 Word Cloud Topik 7	83
Gambar 4. 8 Word Cloud Topik 8	84
Gambar 4. 9 Word Cloud Topik 9	85
Gambar 4. 10 Wordcloud Topik 10	85
Gambar 4. 11 Word Cloud Topik 11	86
Gambar 4. 12 Word Cloud Topik 12	87
Gambar 4. 13 Word Cloud Topik 13	88
Gambar 4. 14 Word Cloud Topik 14	88

Gambar 4. 15 Word Cloud Topik 15.....	89
Gambar 4. 16 Word Cloud Topik 16.....	90
Gambar 4. 17 Confussion Matrix	93
Gambar 4. 18 Tampilan Halaman Input	96
Gambar 4. 19 Tampilan Halaman Hasil	96
Gambar 4. 20 Tampilan Halaman Riwayat	97

DAFTAR TABEL

Tabel 2. 1 Confusion Matrix	21
Tabel 3. 1 Sampel Data Case Folding.....	27
Tabel 3. 2 Sampel Cleansing Data	28
Tabel 3. 3 Sampel Data Tokenizing	29
Tabel 3. 4 Sampel Data Stopword Removal	31
Tabel 3. 5 Sampel Data Stemming.....	33
Tabel 3. 6 Rasio Pembagian Dataset.....	33
Tabel 3. 7 Tabel Insialisasi Topik Acak.....	36
Tabel 3. 8 Tabel BoW	37
Tabel 3. 9 Sampel Hasil LDA	39
Tabel 3. 10 Skenario Pengujian LDA	44
Tabel 4. 1 Hasil Case Folding.....	52
Tabel 4. 2 Hasil Cleansing	54
Tabel 4. 3 Hasil Tokenizing	55
Tabel 4. 4 Hasil Stopword Removal	57
Tabel 4. 5 Hasil Stemming.....	59
Tabel 4. 6 Parameter Terbaik	70
Tabel 4. 7 Hasil Pengujian ZSC.....	71
Tabel 4. 8 Perbandingan Model	73
Tabel 4. 9 Hasil LDA Topik 0-2	75
Tabel 4. 10 Hasil LDA Topik 3-5	75
Tabel 4. 11 Hasil LDA Topik 6-8	76
Tabel 4. 12 Hasil LDA Topik 9-11	77
Tabel 4. 13 Hasil LDA Topik 12-14.....	78
Tabel 4. 14 Hasil LDA Topik 15	78
Tabel 4. 15 Hasil Label Manual.....	91
Tabel 4. 16 Hasil Klasifikasi ZSC.....	92
Tabel 4. 17 Metrik Evaluasi	94

Halaman ini sengaja dikosongkan