

Entry Kartu Bimbingan Tugas Akhir



NIM : 2011500846

Nama Mahasiswa : Mila Yuliani

Dosen Pembimbing : Indra

Judul Skripsi : KLAUSTERISASI INFORMASI BENCANA ALAM PADA DATA PLATFORM X MENGGUNAKAN METODE MAXIMUM CAPTURING DENGAN JACCARD SIMILARITY DAN EKSTRAKSI NAMED ENTITY RECOGNITION (NER) BERBASIS RULE-BASED

Hari / Tanggal / Jam : 20/09/2024 04:06

Isikan Materi :

SIMPAN

Kartu Bimbingan Mahasiswa

No	Tanggal	Materi	Action
1	13-MAR-2024 03:30:00	Pembahasan Topik TA	EDIT DELETE
2	19-MAR-2024 02:30:00	Pemantapan Topik TA dan penentuan Metode serta pembahasan dataset yang akan digunakan	EDIT DELETE
3	06-APR-2024 10:30:00	Melakukan pemaparan logika metode yang akan digunakan	EDIT DELETE
4	09-APR-2024 10:25:00	Pembahasan Rumus dan Perhitungan Manual terhadap dataset yang sudah di crawling	EDIT DELETE
5	13-MAY-2024 11:40:00	Pemaparan pendahuluan, latar belakang masalah dan studi literatur	EDIT DELETE
6	28-MAY-2024 12:25:00	Demo Program, revisi pendahuluan dan studi literatur laporan skripsi	EDIT DELETE
7	22-JUN-2024 10:50:00	Pemaparan Metodologi penelitian yang akan digunakan serta hasil dan pembahasan pada laporan skripsi	EDIT DELETE
8	28-JUN-2024 12:10:00	Pemaparan kesimpulan dan revisi laporan	EDIT DELETE



UNIVERSITAS BUDI LUHUR

FAKULTAS TEKNOLOGI INFORMASI | FAKULTAS EKONOMI & BISNIS | FAKULTAS ILMU SOSIAL & ILMU POLITIK
FAKULTAS TEKNIK | FAKULTAS ILMU KOMUNIKASI

Kampus Pusat : Jl. Raya Ciledug - Petukangan Utara - Jakarta Selatan 12260
Telp : 021-5853753 (hunting), Fax : 021-5853489, <http://www.budiluhur.ac.id>

BERITA ACARA SIDANG PENDADARAN TUGAS AKHIR

S/UBL/FTI/0189/VII/24

Pada hari ini, Senin 15 Juli 2024 telah dilaksanakan Ujian Sidang Pendadaran Tugas Akhir sebagai berikut:

Judul : KLASTERISASI INFORMASI BENCANA ALAM PADA DATA PLATFORM X MENGGUNAKAN METODE MAXIMUM CAPTURING DENGAN JACCARD SIMILARITY DAN EKSTRAKSI NAMED ENTITY RECOGNITION (NER) BERBASIS RULE-BASED

Nama : Mila Yuliani
NIM : 2011500846
Dosen Pembimbing : Dr. Indra, S.Kom., M.T.I

Berdasarkan penilaian pada Presentasi + Demo, Penulisan, Penguasaan Materi, Penguasaan Program maka Mahasiswa tersebut di atas dinyatakan:

LULUS

dengan nilai angka: 85 huruf: A

Mahasiswa tersebut di atas wajib menyerahkan hasil perbaikan tulisan Tugas Akhir dalam bentuk terjilid sesuai dengan Panduan Perbaikan Tugas Akhir, selambat-lambatnya Senin 29 Juli 2024.

Panitia Penguji:

1. Ketua Wahyu Pramusinto, S.Kom., M.Kom.
2. Anggota Painem, S.Kom., M.Kom.
3. Moderator Dr. Indra, S.Kom., M.T.I

Keterangan:

Nilai Huruf: A:85-100 A-:80-84,99 B+:75-79,99 B:70-74,99 B-:65-69,99 C:60-64,99 D:40-59,99 E-:0-39,99

**KLASTERISASI INFORMASI BENCANA ALAM PADA DATA
PLATFORM X MENGGUNAKAN METODE *MAXIMUM
CAPTURING* DENGAN JACCARD SIMILARITY DAN
EKSTRAKSI *NAMED ENTITY RECOGNITION* (NER)
BERBASIS *RULE-BASED***

TUGAS AKHIR



Oleh:

**MILA YULIANI
NIM : 2011500846**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS BUDI LUHUR
JAKARTA
2024**

**KLASTERISASI INFORMASI BENCANA ALAM PADA DATA
PLATFORM X MENGGUNAKAN METODE *MAXIMUM
CAPTURING* DENGAN *JACCARD SIMILARITY* DAN
EKSTRAKSI *NAMED ENTITY RECOGNITION (NER)*
BERBASIS *RULE-BASED***

**Diajukan untuk memenuhi salah satu persyaratan memperoleh
Gelar Sarjana Komputer (S.Kom)**



Oleh:

**MILA YULIANI
NIM : 2011500846**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS BUDI LUHUR
JAKARTA
2023/2024**



**PROGRAM STUDI TEKNIK
INFORMATIKA FAKULTAS
TEKNOLOGI INFORMASI UNIVERSITAS
BUDI LUHUR**

LEMBAR PENGESAHAN

Nama : Mila Yuliani
Nomor Induk Mahasiswa: 2011500846
Program Studi : Teknik Informatika
Bidang Peminatan : Artificial Intelligence
Jenjang Studi : Strata 1
Judul : *KLASTERISASI INFORMASI BENCANA ALAM
PADA DATA PLATFORM X MENGGUNAKAN
METODE *MAXIMUM CAPTURING* DENGAN
JACCARD SIMILARITY DAN EKSTRAKSI *NAMED
ENTITY RECOGNITION (NER)* BERBASIS *RULE-
BASED**



Laporan Tugas Akhir ini telah disetujui, disahkan dan direkam secara elektronik sehingga tidak memerlukan tanda tangan tim penguji.

Jakarta, Senin 15 Juli 2024

Tim Penguji:

Ketua : Wahyu Pramusinto, S.Kom., M.Kom.
Anggota : Painem, S.Kom., M.Kom.
Pembimbing : Dr. Indra, S.Kom., M.T.I
Ketua Program Studi : Dr. Indra, S.Kom., M.T.I

PERNYATAAN TIDAK PLAGIAT

SURAT PERNYATAAN TIDAK PLAGIAT DAN PERSETUJUAN PUBLIKASI

Saya yang bertanda tangan dibawah ini :

Nama : Mila Yuliani
NIM : 2011500846
Program Studi : Teknik Informatika
Bidang Peminatan : Artificial Intelligence
Jenjang Studi : Strata 1
Fakultas : Teknologi Informasi

menyatakan bahwa TUGAS AKHIR yang berjudul:

KLASTERISASI INFORMASI BENCANA ALAM PADA DATA TWITTER
MENGGUNAKAN METODE MAXIMUM CAPTURE WITH JACCARD
SIMILANTY DAN EKSTRAKSI NAMED ENTITY RECOGNITION BERBASIS RULE-BASED

Merupakan :

1. Karya tulis saya sebagai laporan tugas akhir yang asli dan belum pernah diajukan untuk mendapatkan gelar akademik apapun, baik di Universitas Budi Luhur maupun di perguruan tinggi lainnya.
2. Karya tulis ini bukan saduran / terjemahan, dan murni gagasan, rumusan dan pelaksanaan penelitian / implementasi saya sendiri, tanpa bantuan pihak lain, kecuali arahan pembimbing akademik dan pembimbing di organisasi tempat riset.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat yang telah ditulis atau dipublikasikan orang lain, kecuali secara tertulis dengan dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan dicantumkan dalam daftar pustaka.
4. Saya menyerahkan hak milik atas karya tulis ini kepada Universitas Budi Luhur, dan oleh karenanya Universitas Budi Luhur berhak melakukan pengelolaan atas karya tulis ini sesuai dengan norma hukum dan etika yang berlaku.

Pernyataan ini saya buat dengan sesungguhnya dan apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka saya bersedia menerima sanksi akademik berupa pencabutan gelar yang telah diperoleh berdasarkan karya tulis ini, serta sanksi lainnya sesuai dengan norma di Universitas Budi Luhur dan Undang-Undang yang berlaku.

Jakarta, 13 Juli 2019



Mila Yuliani

ABSTRAK

KLASTERISASI INFORMASI BENCANA ALAM PADA DATA PLATFORM X MENGUNAKAN METODE *MAXIMUM CAPTURING* DENGAN *JACCARD SIMILARITY* DAN EKSTRAKSI *NAMED ENTITY RECOGNITION (NER)* BERBASIS *RULE-BASED*

Oleh: MILA YULIANI (2011500846)

Bencana alam di Indonesia, seperti banjir, gempa bumi, tanah longsor merupakan fenomena yang sering terjadi dan memerlukan penanganan dan respons yang cepat serta efektif. Di era digital saat ini, media sosial termasuk platform X telah menjadi sumber utama informasi terkini mengenai bencana alam. Namun, volume informasi yang besar dan cepat dari platform ini seringkali menghasilkan data yang kompleks, sehingga menjadi sulit untuk diidentifikasi dan dikelompokkan secara akurat. Penelitian ini bertujuan untuk mengimplementasikan metode klusterisasi menggunakan algoritma *Maximum Capturing* untuk mengelompokkan bencana alam dan mendeteksi lokasi terjadinya bencana alam menggunakan teknik *Named Entity Recognition* berbasis *rule-based* dari data tweet di platform X. Data yang digunakan dalam sistem ini berjumlah 300 tweet. Sistem ini menggunakan algoritma *Maximum Capturing* dengan perhitungan *Jaccard Similarity* untuk menentukan nilai kesamaan antar tweet yang digunakan dalam pembentukan kluster. Hasil penelitian menunjukkan bahwa data dapat dikelompokkan berdasarkan jenis bencana banjir, gempa bumi dan tanah longsor. Deteksi lokasi bencana juga dapat dilakukan meskipun masih terdapat beberapa entitas yang masih kurang akurat. Hasil klusterisasi dilakukan pengujian menggunakan *purity* dan menghasilkan nilai *purity* sebesar 78.00% yang menunjukkan tingkat kesamaan dokumen yang relatif baik.

Kata kunci: Klusterisasi, Bencana Alam, *Maximum Capturing*, *Named Entity Recognition*, *rule-based*

xiii+89 halaman; 63 gambar; 30 tabel

KATA PENGANTAR

Dengan menyebut Nama Allah Yang Maha Pengasih lagi Maha Penyayang. Puji dan syukur penulis panjatkan karena atas Rahmat yang diberikan Nya sehingga Tugas Akhir ini dapat selesai sebagaimana yang diharapkan. Maksud dan tujuan penyusunan Tugas Akhir ini adalah untuk memenuhi salah satu syarat dalam menyelesaikan jenjang Strata 1 (S1) Fakultas Teknologi Informasi pada program studi Teknik Informatika di Universitas Budi Luhur.

Dalam proses penyelesaian tugas akhir ini, penulis ingin menyampaikan terima kasih yang sebesar-besarnya kepada berbagai pihak yang telah memberikan dukungan dan bantuannya, terutama kepada:

1. Allah SWT, Tuhan semesta alam, yang selalu memberikan rahmat, nikmat, dan hidayah-Nya sehingga penulis dapat menyelesaikan penyusunan Tugas Akhir ini dengan baik.
2. Kepada Bapak Raji dan Ibu Wasri tercinta, serta keluarga besar yang selalu memberikan doa, kasih sayang, dan dukungan.
3. Kepada Bapak Prof. Dr. Agus Setyo Budi, M.Sc., selaku Rektor Universitas Budi Luhur.
4. Kepada Bapak Dr. Ir. Achmad Solichin, S.Kom., M.T.I., selaku Dekan Fakultas Teknologi Informasi Universitas Budi Luhur.
5. Kepada Bapak Dr. Indra, S.Kom, M.T.I, selaku Ketua Program Studi Teknik Informatika, Fakultas Teknologi Informasi Universitas Budi Luhur serta Dosen Pembimbing Tugas Akhir yang selalu memberikan arahan, motivasi, bimbingan, dan masukan selama penyusunan Tugas Akhir ini.
6. Sahabat-sahabat penulis, Ailza Zandra, Firda Nur Syahrani dan Nurul Nabila yang sudah menemani berposes selama 4 tahun perkuliahan ini.
7. Teman-teman seperjuangan dan kaka tingkat saya, teman-teman kelas unggulan dan masih banyak lainnya yang selalu memotivasi serta memberikan masukan serta saran kepada penulis.
8. Serta terima kasih kepada semua pihak yang telah membantu dalam menyelesaikan laporan Tugas Akhir ini, baik secara langsung maupun tidak langsung.

Pada akhirnya, penulis menyadari bahwa dalam penulisan Tugas Akhir ini masih terdapat kekurangan. Oleh karena itu, penulis sangat mengharapkan kritik dan saran yang membangun dari berbagai pihak demi perbaikan menuju kesempurnaan Tugas Akhir ini. Semoga Tugas Akhir ini dapat menjadi acuan bagi rekan-rekan mahasiswa, khususnya di Universitas Budi Luhur.

Jakarta, Juli 2024

Penulis

DAFTAR TABEL

Table 2.1 Daftar Fitur Kontektual (Budi et al., 2015).....	11
Table 2.2 Daftar fitur Morphological (Budi et al., 2015).....	12
Table 2.3 Daftar fitur Part-of Speech (Budi et al., 2015).....	13
Table 2.4 Studi Literatur	14
Table 3.1 Sampel Dataset Tweet.....	18
Table 3.2 Perolehan Data Crawling	21
Table 3.3 Tokenisasi dan pemilihan fitur (feature assignment)	25
Table 3.4 Hasil Name Tagging	26
Table 3.5 datacraw	29
Table 3.6 proses	30
Table 3.7 kamus	31
Table 3.8 model.....	31
Table 3.9 clusters	31
Table 3.10 analisis.....	31
Table 3.11 hybrid_tfidf	32
Table 3.12 purity results.....	32
Table 3.13 ner	33
Table 3.14 ner testing.....	33
Table 4.1 Sample Data Bencana	43
Table 4.2 Hasil Nilai Jaccard Similarity	44
Table 4.3 Table Hasil Cluster Data Sample	47
Table 4.4 Table Hasil Analisis Data Sample	47
Table 4.5 Table Hasil Analisis Klaster	48
Table 4.6 Hasil Pembobotan TF.....	49
Table 4.7 Hasil Pembobotan IDF.....	50
Table 4.8 Hasil Pembobotan Kata.....	51
Table 4.9 Hasil Pembobotan Kalimat	52
Table 4.10 Hasil Ekstraksi Lokasi	55
Table 4.11 Hasil Pengujian Purity Per kluster	79
Table 4.12 Hasil Pengujian Dokumen	80



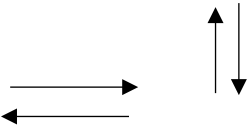

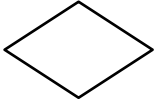
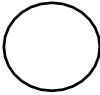

DAFTAR GAMBAR

Gambar 2.1 Arsitektur Indonesian Named Entity Recognition	11
Gambar 2.2 Model NER	14
Gambar 3.1 Tahapan Metode	20
Gambar 3.2 Proses Preprocessing	21
Gambar 3.3 Class Diagram	28
Gambar 3.4 Logical Record Structure	29
Gambar 3.5 Rancangan Menu	34
Gambar 3.6 Rancangan Layar Menu Dashboard	34
Gambar 3.7 Rancangan Layar Menu Import Data	35
Gambar 3.8 Rancangan Layar Menu Preprocessing	35
Gambar 3.9 Rancangan Layar Menu Clustering	36
Gambar 3.10 Rancangan Layar Menu Maksimum Capturing	36
Gambar 3.11 Rancangan Layar Menu Kalimat Utama	37
Gambar 3.12 Rancangan Layar Menu Rule based NER	37
Gambar 3.13 Rancangan Layar Menu Pengujian	38
Gambar 3.14 Rancangan Layar Menu Visualisasi	38
Gambar 4.1 Deployment Diagram	39
Gambar 4.2 Case Folding	40
Gambar 4.3 Hapus URL	41
Gambar 4.4 Hapus mention	41
Gambar 4.5 Hapus hashtag	41
Gambar 4.6 Hapus Spasi Berlebih	41
Gambar 4.7 Hapus non-Alfabet	42
Gambar 4.8 Normalisasi	42
Gambar 4.9 Tokenizing	42
Gambar 4.10 Stopword Removal	42
Gambar 4.11 Stemming	43
Gambar 4.12 Hasil uji coba NER	54
Gambar 4.13 Flowchart Proses Keseluruhan Sistem	56
Gambar 4.14 Flowchart Proses Import Data	56
Gambar 4.15 Flowchart Proses Preprocessing	57
Gambar 4.16 Flowchart Proses Case Folding	58
Gambar 4.17 Flowchart Proses Cleansing	58
Gambar 4.18 Flowchart Proses Normalisasi	59
Gambar 4.19 Flowchart Proses Tokenizing	59
Gambar 4.20 Flowchart Proses Stopword Removal	60
Gambar 4.21 Flowchart Proses Stemming	60
Gambar 4.22 Flowchart Proses Jaccard Similarity	61
Gambar 4.23 Flowchart Proses Maksimum Capturing	61
Gambar 4.24 Flowchart Proses Kalimat Utama	62
Gambar 4.25 Flowchart Proses NER Rule Based	63
Gambar 4.26 Flowchart Proses Pengujian	63
Gambar 4.27 Flowchart Menu Aplikasi	64
Gambar 4.28 Flowchart Menu Dashboard	65
Gambar 4.29 Flowchart Menu Import Data	65

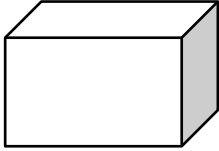
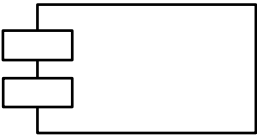


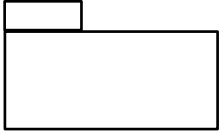
Gambar 4.30 Flowchart Menu Preprocessing	66
Gambar 4.31 Flowchart Menu Jaccard Similarity	66
Gambar 4.32 Flowchart Menu Maximum Capturing	67
Gambar 4.33 Flowchart Menu Kalimat Utama	67
Gambar 4.34 Flowchart Menu NER Rule Based	68
Gambar 4.35 Flowchart Menu Pengujian	68
Gambar 4.36 Flowchart Menu Visualisasi	69
Gambar 4.37 Tampilan Layar Dashboard	81
Gambar 4.38 Tampilan Layar Import Data	82
Gambar 4.39 Tampilan Layar Preprocessing	82
Gambar 4.40 Tampilan Layar Nilai Jaccard Similarity	83
Gambar 4.41 Tampilan Layar Hasil Clustering	83
Gambar 4.42 Tampilan Layar Maximum Capturing	84
Gambar 4.43 Tampilan Layar Kalimat Utama	84
Gambar 4.44 Tampilan Layar Hasil Kalimat Utama	85
Gambar 4.45 Tampilan Layar NER Rule Based	85
Gambar 4.46 Tampilan Layar Pengujian	86
Gambar 4.47 Tampilan Layar Visualisasi	86

DAFTAR SIMBOL

1. Simbol Activity Diagram

No.	Simbol	Nama	Keterangan
1		Terminator	Simbol yang menyatakan awal atau akhir program.
2		Off-Page Reference	Simbol untuk keluar masuk atau penyambung proses lembar program yang berbeda.
3		Flow	Simbol yang digunakan untuk menggabungkan antara symbol dengan symbol lain.
4		Input/Output	Simbol yang menyatakan proses input atau output tanpa tergantung peralatan.
5		Decision	Simbol yang menunjukkan kondisi tertentu yang akan menghasilkan dua kemungkinan jawaban seperti true atau false.
6		On-Page Reference	Simbol untuk keluar masuk atau penyambung proses lembar kerja yang sama.
7		Process	Simbol yang menyatakan suatu proses yang dilakukan program.

2. Simbol Deployment Diagram

No.	Simbol	Nama	Keterangan
1		Node	Node merupakan <i>hardware</i> atau <i>software</i> yang digunakan pada sistem.
2		Component	Suatu komponen yang terdapat pada infrastruktur sistem.
3		Link	Menggambarkan relasi
4		Dependency	Ketergantungan antar node, arah panah mengarah pada node yang dipakai.
5		Package	<i>Package</i> merupakan sebuah bungkus yang terdiri dari satu atau lebih node.

DAFTAR ISI

PERNYATAAN TIDAK PLAGIAT	ii
ABSTRAK	iii
KATA PENGANTAR.....	iv
DAFTAR TABEL	v
DAFTAR GAMBAR	vi
DAFTAR SIMBOL.....	viii
DAFTAR ISI.....	x
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	3
1.3. Batasan Masalah	3
1.4. Tujuan	3
1.5. Manfaat	3
1.6. Sistematika Penulisan.....	4
BAB II LANDASAN TEORI.....	4
2.1. <i>Text Mining</i>	5
2.2. Platform X.....	5
2.3. Pengumpulan Data	5
2.4. <i>Preprocessing</i>	5
2.4.1. <i>Case Folding</i>	6
2.4.2. <i>Cleaning</i>	6
2.4.3. Normalisasi	6
2.4.4. <i>Tokenizing</i>	6
2.4.5. <i>Stopword Removal</i>	6
2.5. <i>Clustering</i>	7
2.6. <i>Jaccard Similarity</i>	7
2.7. <i>Maximum Capturing</i>	7
2.8. <i>Hybrid TF-IDF</i>	8
2.9. Pengujian.....	9
2.10. <i>Named Entity Recognition (NER)</i>	10
2.11. Studi Literatur	14
BAB III METODOLOGI PENELITIAN.....	18
3.1. Data Penelitian	18

3.2.	Metode Pembanding.....	19
3.3.	Penerapan Metode.....	20
3.3.1.	Pengambilan Data (<i>Crawling</i>).....	21
3.3.2.	<i>Preprocessing</i>	21
3.3.3.	<i>Jaccard Similarity</i>	23
3.3.4.	Metode <i>Maximum Capturing</i>	24
3.3.5.	<i>Hybrid TF-IDF</i>	24
3.3.6.	<i>Named Entity Recognition (NER)</i>	25
3.4.	Rancangan Pengujian.....	26
3.5.	Rancangan Basis Data.....	27
3.5.1.	<i>Class Diagram</i>	27
3.5.2.	<i>Logical Record Structure</i>	28
3.5.3.	Spesifikasi Basis Data.....	29
3.6.	Rancangan Menu.....	33
3.7.	Rancangan Layar.....	34
3.7.1.	Rancangan Layar Menu <i>Dashboard</i>	34
3.7.2.	Rancangan Layar Menu <i>Import Data</i>	34
3.7.3.	Rancangan Layar Menu <i>Preprocessing</i>	35
3.7.4.	Rancangan Layar Menu <i>Clustering</i>	35
3.7.5.	Rancangan Layar Menu <i>Maksimum Capturing</i>	36
3.7.6.	Rancangan Layar Menu Kalimat Utama.....	37
3.7.7.	Rancangan Layar Menu <i>Rule based NER</i>	37
3.7.8.	Rancangan Layar Menu Pengujian.....	38
3.7.9.	Rancangan Layar Menu Visualisasi.....	38
BAB IV HASIL DAN PEMBAHASAN.....		39
4.1.	Lingkungan Percobaan.....	39
4.1.1.	Spesifikasi <i>Software</i>	39
4.1.2.	Spesifikasi <i>Hardware</i>	39
4.1.3.	<i>Deployment Diagram</i>	39
4.2.	Implementasi Metode.....	39
4.2.1.	Tahap Pengumpulan Data.....	40
4.2.2.	Tahap <i>Preprocessing</i>	40
4.2.3.	Tahap Penentuan Nilai <i>Jaccard Similarity</i>	43
4.2.4.	Pembentukan Klaster Metode <i>Maximum Capturing</i>	44
4.2.5.	Pembobotan <i>Hybrid TF-IDF</i>	48
4.2.6.	<i>NER Rule Based</i>	53

4.3.	<i>Flowchart</i>	55
4.3.1.	<i>Flowchart</i> Proses Keseluruhan Sistem.....	56
4.3.2.	<i>Flowchart</i> Proses Import Data	56
4.3.3.	<i>Flowchart</i> Proses <i>Preprocessing</i>	57
4.3.4.	<i>Flowchart</i> Proses <i>Jaccard Similarity</i>	61
4.3.5.	<i>Flowchart</i> Proses <i>Maximum Capturing</i>	61
4.3.6.	<i>Flowchart</i> Proses Kalimat Utama	62
4.3.7.	<i>Flowchart</i> Proses <i>NER Rule Based</i>	63
4.3.8.	<i>Flowchart</i> Proses Pengujian.....	63
4.3.9.	<i>Flowchart</i> Menu Aplikasi	64
4.3.10.	<i>Flowchart</i> Menu <i>Dashboard</i>	65
4.3.11.	<i>Flowchart</i> Menu <i>Import Data</i>	65
4.3.12.	<i>Flowchart</i> Menu <i>Preprocessing</i>	66
4.3.13.	<i>Flowchart</i> Menu <i>Jaccard Similarity</i>	66
4.3.14.	<i>Flowchart</i> Menu <i>Maximum Capturing</i>	67
4.3.15.	<i>Flowchart</i> Menu Kalimat Utama	67
4.3.16.	<i>Flowchart</i> Menu <i>NER Rule Based</i>	68
4.3.17.	<i>Flowchart</i> Menu Pengujian.....	68
4.3.18.	<i>Flowchart</i> Menu Visualisasi	69
4.4.	Algoritma	69
4.6.1.	Algoritma Proses Keseluruhan.....	69
4.6.2.	Algoritma Proses <i>Import Data</i>	70
4.6.3.	Algoritma Proses <i>Preprocessing</i>	70
4.6.4.	Algoritma Proses <i>Jaccard Similarity</i>	72
4.6.5.	Algoritma Proses <i>Maximum Capturing</i>	73
4.6.6.	Algoritma Proses Kalimat Utama	73
4.6.7.	Algoritma Proses <i>NER Rule Based</i>	74
4.6.8.	Algoritma Proses Pengujian.....	74
4.6.9.	Algoritma Menu Aplikasi	75
4.6.10.	Algoritma Menu <i>Dashboard</i>	75
4.6.11.	Algoritma Menu <i>Import Data</i>	75
4.6.12.	Algoritma Menu <i>Preprocessing</i>	76
4.6.13.	Algoritma Menu <i>Jaccard Similarity</i>	76
4.6.14.	Algoritma Menu <i>Maximum Capturing</i>	77
4.6.15.	Algoritma Menu Kalimat Utama.....	77
4.6.16.	Algoritma Menu <i>NER Rule Based</i>	78

4.6.17.	Algoritma Menu Pengujian.....	78
4.6.18.	Algoritma Menu Visualisasi.....	79
4.5.	Pengujian.....	79
4.6.	Tampilan Layar Aplikasi.....	81
4.6.1.	Tampilan Layar <i>Dashboard</i>	81
4.6.2.	Tampilan Layar <i>Import Data</i>	82
4.6.3.	Tampilan Layar <i>Preprocessing</i>	82
4.6.4.	Tampilan Layar <i>Jaccard Similarity</i>	83
4.6.5.	Tampilan Layar <i>Maximum Capturing</i>	84
4.6.6.	Tampilan Layar Kalimat Utama.....	84
4.6.7.	Tampilan Layar Pengujian.....	86
4.6.8.	Tampilan Layar Visualisasi.....	86
BAB V PENUTUP.....		87
5.1.	Kesimpulan.....	87
5.2.	Saran.....	87
Daftar Pustaka.....		88