



ITS
Institut
Teknologi
Sepuluh Nopember

TUGAS AKHIR - IS184853

**ANALISIS SENTIMEN MASYARAKAT INDONESIA
TERHADAP GATRA EKONOMI KETAHANAN
NASIONAL MENGGUNAKAN FUZZY ONTOLOGY-
BASED SEMANTIC KNOWLEDGE**

**SENTIMENT ANALYSIS OF INDONESIAN SOCIETY
TOWARDS ECONOMIC ASPECT OF NATIONAL
DEFENCE USING FUZZY ONTOLOGY-BASED
SEMANTIC KNOWLEDGE**

MUHAMAD FAIQ PURNOMO PUTRA
0521154000141

Dosen Pembimbing
Nur Aini Rakhmawati, S.Kom., M.Sc. Eng., Ph.D

DEPARTEMEN SISTEM INFORMASI
Fakultas Teknologi Informasi dan Komunikasi
Institut Teknologi Sepuluh Nopember
Surabaya 2019

TUGAS AKHIR - IS184853

**ANALISIS SENTIMEN MASYARAKAT INDONESIA
TERHADAP GATRA EKONOMI KETAHANAN
NASIONAL MENGGUNAKAN FUZZY
ONTOLOGY-BASED SEMANTIC KNOWLEDGE**

**MUHAMAD FAIQ PURNOMO PUTRA
05211540000141**

**Dosen Pembimbing
Nur Aini Rakhmawati, S.Kom., M.Sc. Eng., Ph.D**

**DEPARTEMEN SISTEM INFORMASI
Fakultas Teknologi Informasi dan Komunikasi
Institut Teknologi Sepuluh Nopember
Surabaya 2019**

Halaman ini sengaja dikosongkan

UNDERGRADUATE THESIS - IS184853

**INDONESIAN PUBLIC SENTIMENTS ANALYSIS
ON NATIONAL ECONOMICS RESISTANCE
USING FUZZY ONTOLOGY-BASED SEMANTIC
KNOWLEDGE**

**MUHAMAD FAIQ PURNOMO PUTRA
05211540000141**

Supervisor

Nur Aini Rakhmawati, S.Kom., M.Sc. Eng., Ph.D

**INFORMATION SYSTEM DEPARTMENT
Information Technology and Communication Faculty
Sepuluh Nopember Institute of Technology
Surabaya 2019**

Halaman ini sengaja dikosongkan

LEMBAR PENGESAHAN

**ANALISIS SENTIMEN MASYARAKAT
INDONESIA TERHADAP GATRA EKONOMI
KETAHANAN NASIONAL MENGGUNAKAN
FUZZY ONTOLOGY-BASED SEMANTIC
KNOWLEDGE**

TUGAS AKHIR

Disusun untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer
pada
Departemen Sistem Informasi
Fakultas Teknologi Informasi dan Komunikasi
Institut Teknologi Sepuluh Nopember

Oleh :

MUHAMAD FAIQ PURNOMO PUTRA
NRP. 05211540000141

Surabaya, Juni 2019

KEPALA

DEPARTEMEN SISTEM INFORMASI



Mahendrawathi Er, S.T., M.Sc., Ph.D.
NIP. 19761011 200604 2 001

Halaman ini sengaja dikosongkan

LEMBAR PERSETUJUAN

**ANALISIS SENTIMEN MASYARAKAT
INDONESIA TERHADAP GATRA EKONOMI
KETAHANAN NASIONAL MENGGUNAKAN
FUZZY ONTOLOGY-BASED SEMANTIC
KNOWLEDGE**

TUGAS AKHIR

Disusun untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer
pada

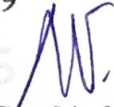
Departemen Sistem Informasi
Fakultas Teknologi Informasi dan Komunikasi
Institut Teknologi Sepuluh Nopember

Oleh :

MUHAMAD FAIQ PURNOMO PUTRA
NRP. 05211540000141

Disetujui Tim Penguji : Tanggal Ujian : 10 Juli 2019
Periode Wisuda : September 2019

Nur Aini Rakhmawati, S.Kom., M.Sc. Eng., Ph.D


(Pembimbing I)

Faisal Johan Atletiko, S.Kom, M.T


(Penguji I)

Irmasari Hafidz, S.Kom, M.Sc


(Penguji II)

ANALISIS SENTIMEN MASYARAKAT INDONESIA TERHADAP GATRA EKONOMI KETAHANAN NASIONAL MENGGUNAKAN FUZZY ONTOLOGY-BASED SEMANTIC KNOWLEDGE

Nama Mahasiswa : Muhamad Faiq Purnomo Putra
NRP : 05211540000141
Departemen : Sistem Informasi FTIK-ITS
Pembimbing I : Nur Aini Rakhmawati, S.Kom.,
M.Sc. Eng., Ph.D

ABSTRAK

Kampanye diantara dua kubu acap kali meramaikan media sosial yang telah menjadi target kampanye dimana total pengguna media sosial di Indonesia telah mencapai 130 juta pengguna. Memanfaatkan momentum ramainya media sosial pada tahun pemilu dan kampanye, penulis mencoba menggali sentimen masyarakat melalui Twitter terhadap gatra ekonomi dalam konsepsi ketahanan nasional menggunakan Fuzzy ontology-based semantic knowledge. Ontologi pada biasanya dianggap tidak terlalu efektif dalam mengekstrak informasi dari tweets, sehingga digunakanlah konsep Fuzzy-ontology based semantic knowledge.

Fuzzy ontology-based semantic knowledge merupakan salah satu cara analisis sentimen menggunakan pendekatan gabungan lexicon-based, ontologi, dan fuzzy logic untuk menghasilkan apakah suatu tweet dapat dikategorikan sebagai strong negative, negative, netral, positive, maupun strong positive.

Pada akhirnya, ontologi biasa tidak dapat mengklasifikasikan masuk kedalam sentimen apa sebuah tweet jika tweet tersebut memiliki lebih dari satu nilai SentiWord. Dari 2032 tweet bersentimen, terdapat 205 tweet yang memiliki lebih dari satu

nilai SentiWord sehingga diperlukan penerapan FuzzyDL untuk memecahkan permasalahan tersebut. Dengan menggunakan metode ini, didapatkan akurasi 78%, dengan tingkat presisi 93%, recall 73%, dan function measure 82%.

Kata Kunci: analisis sentimen, ontologi, sentiword, fuzzy logic, gatra ekonomi, ketahanan nasional, Twitter.

INDONESIAN PUBLIC SENTIMENTS ANALYSIS ON NATIONAL ECONOMICS RESISTANCE USING FUZZY ONTOLOGY-BASED SEMANTIC KNOWLEDGE

Name : Muhamad Faiq Purnomo Putra
NRP : 0521154000141
Department : Information System FTIK-ITS
Supervisor : Nur Aini Rakhmawati, S.Kom., M.Sc.
Eng., Ph.D

ABSTRACT

The campaign between the two camps often enlivened social media which has become the target of the campaign where the total number of social media users in Indonesia has reached 130 million users. Utilizing the momentum of the hectic social media in the election year and campaign, the author tries to explore the public sentiment through Twitter on gatra economy in the concept of national resilience using Fuzzy ontology-based semantic knowledge. Ontology is usually considered to be not very effective in extracting information from tweets, so the concept of Fuzzy-ontology based semantic knowledge is used. Fuzzy ontology-based semantic knowledge is one method of sentiment analysis using a combined approach of lexicon-based, ontology, and fuzzy logic to produce whether a tweet can be categorized as strong negative, negative, neutral, positive, or strong positive.

In the end, ordinary ontologies cannot classify what sentiment is a tweet if the tweet has more than one SentiWord value. Of the 2032 sentiment tweets, there are 205 tweets that have more than one SentiWord value, so FuzzyDL is needed to solve these problems. By using this method, an accuracy of 78% is obtained, with a precision level of 93%, a recall of 73%, and a function measure of 82%.

Keywords: sentiment analysis, ontology, sentiword, fuzzy logic, economic aspect, national security, Twitter.

KATA PENGANTAR

Puji dan syukur Penulis panjatkan atas kehadiran Allah SWT, karena atas rahmat, karunia, dan hidayah-Nya, tugas akhir ini telah berhasil diselesaikan sebagai salah satu syarat untuk menyelesaikan pendidikan untuk memperoleh gelar Sarjana Komputer pada Departemen Sistem Informasi ITS. Pada kesempatan ini Penulis ingin mengucapkan terima kasih kepada pihak-pihak yang membantu penyelesaian Tugas Akhir ini, yaitu:

1. Ibu Nur Aini Rakhmawati, S.Kom., M.Sc. Eng., Ph.D selaku Dosen Pembimbing, atas waktu dan kesabaran yang telah diberikan dalam membantu Penulis dengan menyediakan berbagai pemikiran dan petunjuk untuk memperbaiki kekurangan dalam penyelesaian Tugas Akhir ini;
2. Bapak Faisal Johan Atletiko, S.Kom, M.Kom, dan Irmasari Hafidz, S.Kom, M.Sc, selaku Dosen Penguji yang telah meluangkan waktu untuk menguji dan memberikan masukan yang berharga bagi Penulis;
3. Kedua orang tua Penulis, Sugeng Purnomo dan Almarhumah Emilia Sitompul yang selalu memberikan doa tanpa henti kepada Penulis baik dalam kesenangan maupun kesulitan. Karena doa dan dukungannya Penulis dapat menyelesaikan tugas akhir dan pendidikan hingga saat ini. Doa kalian tidak akan pernah lekang oleh waktu;
4. Kakak penulis, Rizkika Purliageng, Chairunissa Purliasih, Maulida Fitri Purliayu, Sofwan Rahman, dan Andi Prakoso serta keponakan Penulis, Story Damar Rahman, Syaira Alexandra Prakoso, dan Syailendra Arkatama Prakoso atas doa, dukungan, canda, dan tawa untuk Penulis, *best things in life are free*;

5. *Rumah Tangga Puker* atas *persekutuan* dan momen-momen bersamanya, Ojan, Bimo, dan Efrem;
6. Badan Pengurus Harian GenBI Korkom Surabaya, Aqila dan Galang, yang telah membantu penulis melewati masa- masa sulit, masa- masa kritis, teman bercerita, teman tumbuh dewasa;
7. *Persekutuan Jokopi* atas tempat menghilangkan penat, Ardo, Benny, Farchan, Ijul, Azzam, Dito, Farhan, dan Indro;
8. Fadel Say Audianto yang selalu ada saat orang lain tidak ada;
9. Kikik dan Erica, teman organisasi di departemen tercinta, atas semua kritik dan masukannya;
10. Teman sebimbangan dan teman satu lab, terkhusus Supri atas motivasinya yang begitu menyegarkan;
11. Teman- teman Lannister yang telah mendukung, menyemangati, menghibur, dan menemani penulis bahkan disaat ketidakjelasan melanda.

Penulis menyadari bahwa tugas akhir ini masih jauh dari kesempurnaan, sehingga kritik dan saran yang bersifat membangun sangat diharapkan. Akhir kata semoga laporan tugas akhir ini dapat bermanfaat bagi banyak pihak.

Surabaya, 14 Juli 2019

Muhamad Faiq Purnomo P.

DAFTAR ISI

ABSTRAK.....	i
ABSTRACT.....	iii
KATA PENGANTAR.....	v
DAFTAR ISI.....	vii
DAFTAR GAMBAR.....	xi
DAFTAR TABEL.....	xiii
DAFTAR KODE	xv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	3
1.3 Batasan Permasalahan.....	3
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	4
1.6 Relevansi.....	4
BAB II TINJAUAN PUSTAKA	5
2.1. Penelitian Terkait.....	5
2.2. Dasar Teori	9
2.2.1. Twitter.....	9
BAB III METODOLOGI.....	19
3.1 Arsitektur Sistem	19
3.2 Tahapan Pelaksanaan Tugas Akhir.....	20
3.1.1 Studi Literatur	21
3.1.2 Pengumpulan Data	21
3.1.3 Data Preprocessing	22
3.1.4 Pembuatan Ontologi	22
3.1.5 Ekstraksi Fitur.....	24

3.1.6	Analisis Sentimen dan Perhitungan Polaritas menggunakan Fuzzy Logic	25
3.1.7	Pengujian.....	26
BAB IV	PERANCANGAN.....	29
4.1	Akuisisi Data.....	29
4.1.1.	Akuisisi Data Username.....	29
4.1.2.	Akuisisi Data Tweets dari User	32
4.2.	Perancangan Pre-processing Dataset	33
4.2.1.	Perancangan Penghapusan Duplikasi Dataset.....	33
4.2.2.	Perancangan Filtering Dataset	33
4.2.3.	Perancangan Pembersihan Dataset.....	34
4.2.4.	Perancangan POS Tagger.....	35
4.3.	Perancangan Ontologi Gatra Ekonomi	36
4.3.1.	Perancangan Konsep Ontologi.....	36
4.3.2.	Perancangan Subclass Faktor Gatra Ekonomi	37
4.4.	Perancangan Ekstraksi Fitur.....	39
4.5.	Perancangan SentiWord Scoring	39
4.6.	Perancangan SWRL-Rule dan Fuzzy Ontology.....	39
BAB V	IMPLEMENTASI.....	43
5.1	Lingkungan Implementasi.....	43
5.2	Pembuatan Crawler Twitter	44
5.2.1	Crawler Hashtag.....	44
5.2.2	Crawler Username.....	49
5.2.3	Crawler Tweets	52
5.3	Pembuatan Filtering Dataset.....	56
5.3.1	Pembuatan Filtering berdasarkan Keyword.....	56
5.3.2	Penghapusan Data Duplikat.....	57
5.4	Pembuatan Preprocessing Dataset	57
5.4.1	Case Folding	57
5.4.2	Data Cleaning.....	58
5.4.3	Tokenizing	58
5.4.4	Stopwords Removal	59

5.4.5	Part-Of-Speech Tagging	60
5.5	Pembuatan Ontologi	60
5.5.1	Pencarian Fitur dengan Frekuensi Tertinggi.....	60
5.5.2	Ekstraksi Adjektiva.....	62
5.5.3	Pembuatan Ontologi menggunakan Protégé.....	63
5.6	Ekstraksi Fitur.....	72
5.6.1	Query MySQL dihubungkan dengan OWL.....	72
5.6.2	Pembuatan file excel terstruktur	73
5.6.3	Memasukkan instance ke dalam Ontologi	73
5.7	Penilaian polaritas menggunakan SentiWordNet ...	74
5.8	Pembuatan SWRL-Rule.....	75
5.8.1	Rule Klasifikasi Fitur.....	75
5.8.2	Rule Klasifikasi Opini	76
5.8.3	Rule SentiScore.....	76
5.8.4	Rule Klasifikasi Polaritas.....	76
5.9	Pembuatan Fuzzy Description Logic	77
5.10	Implementasi Owlready	80
5.11	Pengujian	81
BAB VI HASIL DAN PEMBAHASAN.....		83
6.1.	Hasil Data Crawling.....	83
6.1.3	Hasil Pengumpulan Tweets	87
6.2	Hasil Preprocessing Dataset.....	87
6.2.3	Hasil Penghapusan Data Duplikat	87
6.2.4	Hasil Filtering Keyword Gatra Ekonomi.....	88
6.3	Hasil Pembuatan Ontologi	90
6.4	Hasil Ekstraksi Fitur	91
6.5	Hasil Analisis Sentimen dan Perhitungan Polaritas	93
6.6	Hasil Analisis Pengujian	98
BAB VII KESIMPULAN DAN SARAN.....		101
7.1	Kesimpulan	101

7.2	Saran.....	102
	DAFTAR PUSTAKA.....	103
	LAMPIRAN A. Hasil Ontologi Lengkap.....	107
	LAMPIRAN B. SWRL.....	118
	BIODATA PENULIS.....	121

DAFTAR GAMBAR

Gambar 2.1 <i>Membership Function</i>	13
Gambar 3.1 Arsitektur Sistem	19
Gambar 3.2 Metodologi	21
Gambar 3.3 Contoh Ontologi pada OntoGraf.....	24
Gambar 3.4 Fuzzy Inference Layer untuk Perhitungan Polaritas	26
Gambar 4.1 Hashtag	30
Gambar 4.2 Flowchart <i>Crawling User</i>	32
Gambar 4.3 Flowchart <i>Crawling Tweets</i>	33
Gambar 4.4 Contoh kode yang telah dibersihkan.....	35
Gambar 4.5 Contoh Tagging	36
Gambar 4.6 Perancangan Konsep Ontologi.....	37
Gambar 4.7 Flowchart pembuatan Ontologi.....	38
Gambar 4.8 Ekstraksi Fitur	39
Gambar 4.9 SWRL Rule	40
Gambar 4.10 Flowchart Pengolahan Tweet Fuzzy	41
Gambar 5.1 Hasil <i>Crawl</i> Hashtag	47
Gambar 5.2 JSON Hasil <i>Crawl</i>	51
Gambar 5.3 Tabel Hasil <i>Crawl</i> pada MySQL	56
Gambar 5.4 Hasil <i>FreqDist Bigram</i>	62
Gambar 5.5 Owlready untuk <i>query</i> MySQL	72
Gambar 5.6 Pembuatan Data Terstruktur	73
Gambar 5.7 Contoh Tweet memiliki 2 SentiScore	77
Gambar 5.8 Diagram <i>Fuzzy</i> untuk Analisis Sentimen	78
Gambar 6.1 Hasil Ontologi	90
Gambar 6.2 <i>Fuzzy Concept</i>	95

Halaman ini sengaja dikosongkan

DAFTAR TABEL

Tabel 4.1 Akun Umpan.....	29
Tabel 4.2 <i>Keyword</i> Gatra Ekonomi	33
Tabel 5.1 Spesifikasi Perangkat Keras.....	43
Tabel 5.2 Daftar Library, Editor, dan Bahasa Pemrograman .43	
Tabel 5.3 <i>Class</i> Faktor Gatra Ekonomi	64
Tabel 5.4 Kata Opini dari Faktor Gatra Ekonomi.....	66
Tabel 5.5 Rincian jumlah Opini pada setiap kelas.....	71
Tabel 5.6 <i>SentiScore</i>	74
Tabel 6.1 Hashtag Umpan	83
Tabel 6.2 Hasil Pengumpulan <i>Username</i>	84
Tabel 6.3 Hasil Pengumpulan <i>Username</i> - Manual	85
Tabel 6.4 Hasil Pengumpulan <i>Username Verified</i>	86
Tabel 6.5 Hasil Pengumpulan <i>Username Unverified</i>	87
Tabel 6.6 Hasil Filtering Gatra Ekonomi.....	88
Tabel 6.7 Hasil Ekstraksi Fitur	91
Tabel 6.8 Jumlah <i>Instance</i> pada Kelas Sentimen.....	93
Tabel 6.9 Pola <i>SentiScore</i>	93
Tabel 6.10 Contoh <i>Tweet</i> yang memiliki lebih dari satu <i>SentiScore</i>	94
Tabel 6.11 Contoh <i>Decision Matrix</i>	94
Tabel 6.12 Pencarian Polaritas.....	96
Tabel 6.13 Hasil Akhir Analisis Sentimen	97
Tabel 6.14 Pengujian	98
Tabel 6.15 Kasus yang mengurangi nilai akurasi	98

Halaman ini sengaja dikosongkan

DAFTAR KODE

Kode 5.1 Potongan Kode Implementasi <i>API</i> untuk <i>Crawling</i>	44
Kode 5.2 Potongan Kode Implementasi <i>Crawler Tweets</i> pada <i>User</i>	46
Kode 5.3 Potongan Kode Membaca File CSV ke Python	48
Kode 5.4 Potongan Kode <i>FreqDist</i> untuk Mencari Frekuensi Kata Terbanyak.....	48
Kode 5.5 Potongan Kode <i>Tweepy</i> memanfaatkan <i>TwitterSearch</i>	50
Kode 5.6 Potongan Kode untuk Ekstraksi JSON ke SQL	52
Kode 5.7 Potongan Kode <i>TweetScrapper</i>	53
Kode 5.8 Potongan Kode untuk menghubungkan <i>TweetScrapper</i> ke MySQL.....	54
Kode 5.9 Potongan Kode untuk Memasukkan Hasil <i>Crawl</i> ke MySQL	55
Kode 5.10 Command Line untuk Iterasi tiap Username Perintah <i>TweetScrapper</i>	56
Kode 5.11 Query SQL untuk <i>Filtering Dataset</i> berdasarkan <i>Keyword</i> Gatra Ekonomi	57
Kode 5.12 Kode Drop Duplikat.....	57
Kode 5.13 Kode <i>Case Folding</i>	58
Kode 5.14 Kode <i>Cleaning</i>	58
Kode 5.15 Kode <i>Tokenizing</i>	58
Kode 5.16 Kode <i>Stopwords Removal</i>	59
Kode 5.17 Kode <i>POS Tagger</i>	60
Kode 5.18 Kode <i>Bigram</i>	61
Kode 5.19 Kode <i>FreqDist</i>	62
Kode 5.20 Kode Ekstraksi Adjektiva	63
Kode 5.21 <i>Syntax</i> Cellfie	74
Kode 5.22 Rule Klasifikasi Fitur	76
Kode 5.23 Rule Klasifikasi Opini.....	76
Kode 5.24 Rule <i>Assign SentiScore</i>	76
Kode 5.25 Rule Klasifikasi Polaritas	77
Kode 5.26 <i>Syntax</i> FuzzyDL Konsep Polaritas.....	78
Kode 5.27 <i>Syntax Data Property</i>	78
Kode 5.28 <i>Syntax Weighted Concept</i>	79
Kode 5.29 <i>Syntax</i> Pembuatan <i>Instance</i>	79

Kode 5.30 <i>Syntax</i> Defuzifikasi.....	80
Kode 5.31 Owlready Memuat Ontologi.....	80
Kode 5.32 Memuat Kelas ke Variabel.....	81
Kode 5.33 Print Hasil.....	81

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pada tahun 2019 ini , akan terjadi pesta demokrasi besar-besaran di Indonesia. Tahapan formal Pemilu 2019 yang meliputi pemilihan presiden dan wakilnya, anggota DPR serta DPD, telah resmi dimulai dengan kampanye, sejak 23 September, dan akan berpuncak pada hari pencoblosan, Jumat, 19 April 2019 [1]. Pilpres 2019 kali ini diramaikan oleh dua pasangan yang saling berkontestasi dan berkompetisi untuk mendapatkan simpati rakyat. Tingginya persaingan acap kali mendorong timbulnya kampanye negatif dan kampanye hitam untuk menjatuhkan lawan politik. Kampanye yang seharusnya menyebarkan visi, misi, dan program menjadi forum menyerang, menangkis, bahkan menyebarkan kabar palsu fitnah dan ujaran kebencian [2]. Kampanye ini pun ikut menyeret para tokoh-tokoh politik dan masyarakat awam untuk saling berkicau di Twitter membahas topik- topik seputar pemilu. Peraturan KPU Nomor 23 Tahun 2018 tentang Kampanye Pemilu, menyebutkan media sosial adalah kumpulan saluran komunikasi dalam jaringan internet yang digunakan untuk interaksi dan berbagi konten berbasis komunitas.

Pada lain sisi tidak dapat dipungkiri, total pengguna media sosial di Indonesia mencapai 130 juta pengguna, dengan rata-rata waktu yang dihabiskan adalah 3 jam 23 menit [3], tentu menjadi target kampanye yang sangat menarik bagi para kedua pasangan calon. Salah satu media sosial yang sering digunakan oleh masyarakat Indonesia adalah Twitter, dimana Indonesia menempati peringkat ketiga pengguna Twitter dengan mencapai 24,3 juta pengguna [4]. Dengan adanya sosial media ini, pemerintah juga dapat terbantu dikarenakan penggunaan sentiment analisis dapat membantu pemerintah dalam

memahami opini ataupun perasaan dari masyarakat pada umumnya [5]. Sentimen analisis dapat digunakan untuk mengetahui sentimen dari kejadian-kejadian penting seperti pemilihan umum, kebijakan pemerintah, dan lainnya.

Dan pada era generasi milenial ini, generasi yang akrab dengan media sosial, interaksi di media sosial semakin marak dan mulai mengkhawatirkan dengan munculnya bentuk- bentuk sentiment berbasis ras, golongan, dan agama, *hoax*, dan ujaran kebencian yang dapat mengarah kepada disintegrasi bangsa [2]. Pada lain sisi, integralitas nasional adalah prasyarat mutlak keberhasilan pencapaian cita-cita dan tujuan nasional adalah demikian vital bagi bangsa Indonesia sehingga harus secara jelas dijaga, dimana apabila kondisi seperti ini tidak dibina maka dapat mengarah pada terganggunya persatuan dan kesatuan bangsa [6]. Konsep ini dikenal dengan konsep ketahanan nasional dimana seluruh aspek kehidupan nasional dirinci menjadi Astagatra atau delapan aspek kehidupan nasional. Dimana dalam Astagatra terdapat Pancagatra yakni Ideologi, Politik, Ekonomi, Sosial Budaya, dan Pertahanan dan Keamanan.

Adapun misi nomor 2 dari paslon nomor urut 1 berbunyi “Struktur Ekonomi yang Produktif, Mandiri, dan Berdaya Saing” [7], sedangkan misi nomor 1 dari paslon nomor urut 2 adalah “Membangun perekonomian nasional yang adil, makmur, berkualitas, dan berwawasan lingkungan dengan mengutamakan kepentingan rakyat Indonesia melalui jalan politik-ekonomi sesuai pasal 33 dan 34 UUD 1945 [8].” Dilihat dari prioritas kedua misi dari paslon, jelas prioritas dari misi mereka adalah pada bagian ekonomi.

Pada lain sisi, telah berkembang pada rumpun ilmu komputer pengetahuan mengenai analisis sentimen. Dengan menggunakan analisis sentimen dengan data yang diambil dari media sosial, analisis dapat membantu pemerintah untuk mengukur seberapa baik atau buruk opini masyarakat umum

terhadap pemerintahan saat ini [5]. Berbagai cara pun telah berkembang untuk analisis sentimen, salah satunya adalah pendekatan *lexicon-based* menggunakan *fuzzy-ontology based semantic approach*. Metodologi ini berkembang dikarenakan ontologi biasa dianggap tidak terlalu efektif dalam mengekstrak informasi dari tweets. Dengan menggunakan metodologi ini, diharapkan dapat dilakukan analisis sentimen menggunakan data Twitter untuk melihat seberapa kuat ketahanan nasional Indonesia dalam gatra ekonomi.

1.2 Rumusan Masalah

Permasalahan yang diangkat dalam tugas akhir ini adalah sebagai berikut:

1. Bagaimana mengakuisisi data tweet dari *buzzer* politik dan tokoh masyarakat pada media sosial Twitter?
2. Bagaimana membuat konsepsi ketahanan nasional gatra ekonomi dibentuk dalam ontologi?
3. Bagaimana menganalisis sentimen masyarakat terhadap gatra ekonomi dalam konsepsi ketahanan nasional menggunakan fuzzy logic?

1.3 Batasan Permasalahan

Batasan masalah didalam pengerjaan tugas akhir ini adalah:

1. Data yang akan diambil adalah tweets dari akun-akun *buzzer* politik, dan tokoh masyarakat.
2. Data tweet yang diambil dimulai dari deklarasi capres dan cawapres dilakukan.
3. Gatra yang dipakai dalam Ketahanan Nasional pada penelitian ini hanya terbatas pada Gatra Ekonomi.
4. Ontologi yang dibuat didasari oleh faktor- faktor yang mempengaruhi Gatra Ekonomi pada buku

“Ketahanan Nasional” yang dikeluarkan oleh Lembaga Pertahanan Nasional.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah mengetahui sentimen masyarakat terhadap gatra ekonomi dalam konsepsi ketahanan nasional menggunakan pendekatan lexicon-based approach yakni menggunakan fuzzy-ontology based semantic knowledge. Dengan adanya analisis ini, dapat diketahui seberapa baik atau buruknya sentiment masyarakat terhadap gatra ekonomi dalam konsepsi ketahanan nasional.

1.5 Manfaat Penelitian

Manfaat yang diperoleh dengan dilakukan penelitian ini adalah:

1. Membantu pemerintah dalam mengetahui sentimen masyarakat terhadap gatra ekonomi, sehingga pemerintah dapat melakukan penyesuaian-penyesuaian yang harus dilakukan.
2. Menyediakan data yang dapat digunakan untuk menganalisis kondisi ketahanan nasional, khususnya dalam gatra ekonomi.

1.6 Relevansi

Tugas akhir ini berkaitan dengan mata kuliah Pengolahan Bahasa Alami, Sistem Cerdas, Desain Basis Data, Pemrograman Integratif, dan Konstruksi dan Pengujian Perangkat Lunak.

BAB II TINJAUAN PUSTAKA

2.1. Penelitian Terkait

2.1.1. Fuzzy ontology-based sentiment analysis of transportation and city feature reviews for safe traveling

Penulis	Farman Ali, Daehan Kwak, Pervez Khan, S.M Riazul Islam, Kye Hyun Kim, K.S. Kwak
Tahun	2018
Studi Kasus	Transportasi dan Fasilitas Umum
Objektif	Pada paper ini dijelaskan mengenai penggunaan <i>Fuzzy Logic</i> dengan <i>Ontology-based</i> untuk menganalisis sentimen ulsan transportasi dan fasilitas kota untuk memfasilitasi ITS (Intelligent Transport System) dan para pelancong. Paper ini menilai bahwa analisis sentimen menggunakan ontologi biasa tidak bisa mengekstrak informasi yang kabur (<i>blurred</i>) pada <i>tweet</i> dan ulasan, sehingga diperlengkap menggunakan Fuzzy ontology-based. Hasilnya presisi meningkat sebesar 23%.
Keterkaitan	Keterkaitan penelitian ini dengan Tugas Akhir ini adalah penggunaan Fuzzy Ontology-based untuk menganalisis sentimen pada Twitter yang berkaitan dengan Politik dan Ekonomi.

2.1.2. Rancang Bangun Aplikasi untuk Klasifikasi Post pada Sosial Media Pemerintah Daerah di Indonesia Menggunakan Support Vector Machine (SVM)

Penulis	Guntur Kondang Prakoso
Tahun	2018
Studi Kasus	Klasifikasi Post pada Sosial Media Pemerintah Daerah
Objektif	Tujuan dari penelitian ini adalah untuk mengkategorikan serta melihat jenis postingan seluruh Pemda di Indonesia pada media sosial Facebook, Twitter, dan Youtube guna memberikan kepada pemerintah daerah post seperti apa yang cocok dan mendapat banyak tanggapan oleh masyarakat. Menggunakan hasil crawling yang diklasifikasi menggunakan SVM, penelitian ini berhasil menciptakan sebuah aplikasi yang dapat melihat kategori postingan Pemda di Indonesia untuk mendukung <i>E-Government</i> .
Keterkaitan	Keterkaitan dengan penelitian ini dengan pengerjaan tugas akhir ini adalah pada data crawling di sosial media. Letak perbedaannya adalah jika pada penelitian ini mengkategorikan postingan menggunakan kategori yang sudah ada pada penelitian berjudul “Facebook usage in government – a cases study of information content” menggunakan klasifikasi SVM, maka pengerjaan tugas akhir ini melakukan

	crawling data yang bertujuan untuk mengetahui analisis sentimen masyarakat saat ini terhadap ketahanan nasional, khususnya dalam bidang Politik dan Ekonomi.
--	--

2.1.3. Pre Processing of Twitter's Data for Opinion Mining in Political Context

Penulis	Ratab Gull, Umar Shoaib, Saba Rasheed, Washma Abid, Beenish Zahoor
Tahun	2016
Studi Kasus	Melakukan <i>opinion mining</i> dengan bantuan analisis linguistik dan pengklasifikasian pada Twitter untuk mengetahui sentimen positif, negatif, dan netral untuk partai-partai politik di Pakistan.
Objektif	Tujuan dari penelitian ini adalah menampilkan hasil analisis sentimen pada Twitter (positif, negatif, dan netral) terkait dengan partai- partai politik yang berada di Pakistan. Penelitian ini menggunakan dua metode klasifikasi yaitu SVM dan Naïve Bayes. Pada akhir penelitian ini disimpulkan bahwa SVM melakukan performa yang lebih baik. Pada akhir penelitian ini juga dihasilkan sebuah visualisasi apakah sebuah partai politik mendapatkan sentimen yang positif, negatif, maupun netral.

Keterkaitan	Keterkaitan dengan pengerjaan tugas akhir ini adalah crawling data pada Twitter, penggalian opini dan pengolahan sentimen. Perbedaannya adalah pada tugas akhir ini tidak menggunakan SVM, dan tidak melihat sentimen terhadap partai politik, melainkan Politik dan Ekonomi.
--------------------	---

2.1.4. Twitter Ontology-Driven Sentiment Analysis

Penulis	Ioan Roxin, Ramona Paun, Camelia Delcea
Tahun	2015
Studi Kasus	Melakukan analisis sentimen pada data twitter menggunakan <i>semantic social media analysis</i> .
Objektif	Tujuan dari penelitian ini adalah membuat sebuah <i>semantic social media analysis</i> yang baru, dengan sebutan TweetOntoSense, yang menggunakan ontologi untuk memodelkan perasaan-perasaan yang kompleks seperti kebahagiaan, terkejut, cinta, ataupun kesedihan, sehingga tidak hanya melihat sentimennya berdasarkan positif, negatif, dan netral.
Keterkaitan	Keterkaitannya dengan penelitian tugas akhir ini adalah penggunaan ontologi dalam ekstraksi fitur- fitur. Pada penelitian ini fitur adalah emosi-emosi yang ada pada tweets sehingga dapat diekstraksi dan dapat diketahui emosi yang paling sering muncul.

2.2. Dasar Teori

2.2.1. Twitter

Twitter adalah sebuah website yang memungkinkan penggunaannya untuk menyebarkan kabar kehidupan mereka, pemikiran, dan berbagi gambar, tautan, ataupun video. Twitter memungkinkan organisasi maupun individu dengan status sosial yang tinggi untuk mempromosikan acara, kampanye, produk, ataupun jasa [5]. Dengan kemampuan menyediakan informasi secara real-time dan dapat mencapai banyak komunitas dan individu, para pengguna dapat mengetahui kabar mengenai suatu berita sehingga para pengguna dapat mengikuti perkembangan berita tersebut, ikut menyebarkan, dan ikut berpendapat.

Dari semua media sosial yang ada, Twitter telah berkembang menjadi platform untuk berita. Tidak seperti platform media sosial lainnya, dimana berita biasanya ditemui sebagai efek samping dari interaksi lainnya, para pengguna Twitter menggunakan platform tersebut sebagai tempat untuk melacak perkembangan berita. Oleh karena itu, platform *micro-blogging* ini lebih diminati oleh jurnalis dan media massa untuk melakukan *tweet* secara langsung berita-berita terkini yang sedang berlangsung di dunia nyata, atau untuk melakukan *tweet* dalam bentuk rangkaian secara berkelanjutan [9]. Penjelasan- penjelasan diatas merupakan dasar penulis memilih Twitter sebagai media sosial yang akan digunakan untuk penelitian ini.

2.2.2. Social Media Crawler

Social Media Crawler adalah sebuah aplikasi yang bertujuan untuk mengambil data-data yang ada pada media sosial. Media sosial yang nantinya akan diambil

datanya adalah Twitter dengan memanfaatkan Twitter API, dan bahasa pemrograman Python. Social Media Crawler melibatkan berbagai konsep dan komponen sebagai berikut:

a. Media Sosial

Istilah media sosial adalah istilah yang mencakup layanan berbasis internet ataupun *mobile* yang memungkinkan pengguna untuk berpartisipasi dalam pertukaran *online*, berkontribusi pada konten yang dibuat pengguna, ataupun bergabung ke sebuah komunitas *online* [10]. Salah satu contoh dari media sosial adalah Twitter, sebuah platform *micro-blogging* ini lebih diminati oleh jurnalis dan media massa untuk melakukan *tweet* secara langsung berita-berita terkini yang sedang berlangsung di dunia nyata, atau untuk melakukan *tweet* dalam bentuk rangkaian secara berkelanjutan [9].

b. Web Crawler

Web Crawler merupakan komponen substansial dari sebuah mesin pencari, dimana mereka dipakai sebagai pengumpul berbagai korpus yang nantinya akan diindeks oleh mesin pencari [11]. Crawling data adalah tahapan yang bertujuan untuk mengumpulkan atau mengunduh data dari suatu database [12].

c. Twitter API

Twitter API memungkinkan para pengembang untuk membuat perangkat lunak yang dapat terintegrasi dengan Twitter, contohnya adalah membantu perusahaan untuk dapat merespon kepada timbal balik pelanggannya. Tidak hanya itu, Twitter API juga memungkinkan para pengembang mendapatkan akses kepada data publik pada Twitter [13].

2.2.3 Analisis Sentimen

Semantic Orientation (SO) adalah pengukuran dari subjektivitas dan opini dari sebuah teks. Semantic Orientation biasanya menangkap faktor evaluatif (positif, netral, negatif) dan potensi atau kekuatan (derajat dari kata, frasa, kalimat, apakah negatif, netral, atau positif) terhadap subjek topik, orang, atau gagasan [14]. Ketika Semantic Orientation digunakan pada opini publik ataupun berita, contohnya ketika dipakai untuk menginterpretasikan *caption* pada Twitter, Semantic Orientation akan dapat membantu pemerintah untuk mengukur seberapa baik atau buruk opini masyarakat umum terhadap pemerintahan saat ini [5]. Pemerintah juga dapat mengetahui kelompok manakah yang mempunyai sentimen negatif sehingga pemerintah dapat memberikan perhatian yang lebih untuk memecahkan permasalahan tersebut.

Pada sisi lain, analisis dan ekstraksi otomatis dari Semantic Orientation dikenal sebagai Sentiment Analysis. Sentiment Analysis mengacu kepada metode umum untuk mengekstrak subjektivitas dan polaritas dari sebuah teks. Analisis sentimen merupakan bidang interdisipliner yang terdiri dari *Natural Language Processing*, *text analysis*, dan *computational linguistics* untuk mengidentifikasi sentimen dari sebuah teks. Web telah menjadi platform yang berkembang pesat bagi warganet untuk mengekspresikan sentimen dan emosi mereka dalam bentuk teks. Karena teks-teks yang termasuk sebagai opini sering kali terlalu banyak untuk diproses untuk membuat sebuah keputusan [15], maka Sentiment Analysis dapat digunakan untuk memecahkan permasalahan tersebut.

i. Ontologi

Didalam konteks ilmu komputer dan informasi, ontologi mendefinisikan seperangkat representasional primitif yang dapat digunakan untuk pengetahuan ataupun percakapan [16]. Primitif representasional antara lain kelas (*classes*) atau himpunan (*sets*), atribut (*attributes*) atau properti (*properties*) dan hubungan (*relationships*) atau relasi antara anggota kelas. Definisi primitif representasional mencakup informasi tentang makna dan batasan pada aplikasi yang konsisten secara logis. Dalam konteks sistem basis data, ontologi dapat dipandang sebagai tingkat abstraksi model data, analog terhadap model hierarkis dan relasional, tetapi dimaksudkan untuk memodelkan pengetahuan tentang individu, atributnya, dan hubungannya dengan individu lain.

Pada pemahaman lainnya, Ontologi adalah pengetahuan bersama mengenai domain yang spesifik diantara orang-orang dan sistem. Secara matematika, ontologi didefinisikan sebagai berikut:

$$\text{Ontology} = (C, P, R, V, V_c)$$

Pada persamaan diatas, notasi C, P, R, V, dan V_c adalah konsep (concepts), properti (properties), hubungan antar kelas (relationship of classes), nilai (value), dan batasan nilai pada properti (constraint values of properties) [17].

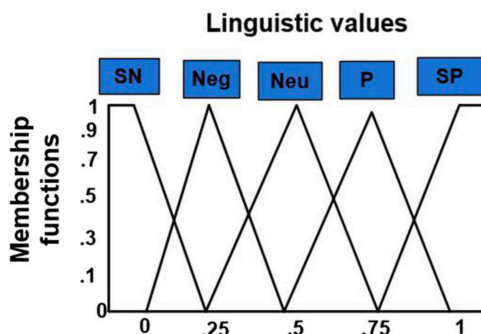
2.2.5 ***Fuzzy Ontology-based semantic knowledge***

Fuzzy Ontology-based adalah sebuah ontologi biasa yang diperlengkap menggunakan Fuzzy Logic dikarenakan ontologi biasa tidak terlalu efektif dikarenakan istilah-istilah yang blur pada konsep nilai [18]. Untuk melakukan hal ini, para pengembang dapat menggunakan *plug-in* pada Protégé OWL untuk mengubah ontologi biasa menjadi Fuzzy Ontology [19].

Konsep Fuzzy diperkenalkan pertama kali oleh Profesor Lotfi Zadeh dari University of California di Berkeley pada tahun 1965. Fuzzy logic pada umumnya digunakan pada masalah- masalah yang mengandung ketidakpastian (*uncertainty*), kegaduhan (*noise*), ketidaktepatan (*imprecise*), dan sebagainya [20]. Fuzzy Logic berbeda dengan Boolean Logic yang membedakan antara 0 dan 1, dalam Fuzzy Set, F, dalam semesta A dapat ditunjukkan oleh fungsi keanggotaan (*membership function*) μ_F , yang menghadirkan elemen 'A' dalam interval [0,1] [17].

$$\mu_F(A): A' \rightarrow [0,1]$$

Pada persamaan diatas A' milik A dan μ_F merepresentasikan fungsi keanggotaan dimana $A' \in A$. A' dikatakan sepenuhnya anggota dari A ketika $\mu_F(A) = 1$, dan dikatakan anggota sebagian (*partial membership*) dari A ketika $\mu_F(A)$ bernilai diantara nol dan satu [21].



Gambar 2.1 *Membership Function*

2.2.6 SentiWordNet

SentiWordNet merupakan kumpulan dari leksikal yang dibuat untuk mendukung klasifikasi sentimen dan proses-proses *opinion mining* pada aplikasi [22].

SentiWordNet berisikan klasifikasi sentimen-sentimen positif, negatif, ataupun netral dari keseluruhan *synset* WordNet. Pada tiap *synset* pada WordNet terdapat skor numerik dalam pengklasifikasian sentimen. Skor sentimen pada SentiWordNet terdiri atas skor Positif, Negatif, dan Objektif yang mengindikasikan nilai seberapa positif, negatif, dan objektif (netral) pada istilah-istilah yang terdapat dalam suatu *synset* [23].

2.2.7 **Buzzer Politik**

Istilah *buzzer* mengacu pada konsep *buzz marketing*, yaitu aktivitas atau kegiatan pemasaran suatu produk pada saluran media komunikasi untuk menciptakan gangguan [24]. *Buzzer* inilah yang nantinya akan membuat isu di media sosial, pada kasus ini Twitter, sampai menjadi perbincangan khalayak luas (menjadi *viral*). Memasuki tahun pemilu ini, akun *buzzer* menjadi faktor penting dalam meluaskan kepentingan politik, golongan *buzzer* politik ini secara sempit membela habis-habisan kepentingan politiknya [24]. Saling sahut antara *buzzer* politik acap kali membanjiri dan mengotori *timeline* pada media sosial. Umumnya, sebuah akun bisa menjadi *buzzer* jika minimal memiliki 3000 *follower*, dan memiliki tiga karakter dasar, yaitu: memiliki konten unik, relevan dan berguna, frekuensi *tweet* yang konsisten dan kualitas interaksi yang tinggi [25].

2.2.8 **Konsepsi Ketahanan Nasional**

Berdasarkan Konsepsi Ketahanan Nasional Indonesia seluruh aspek kehidupan nasional dirinci menjadi Astagatra atau delapan aspek kehidupan nasional, yang terdiri dari Trigatra atau tiga aspek alamiah dan Pancagatra atau lima aspek sosial [6]. Yang dimaksud dengan Pancagatra atau lima aspek sosial adalah aspek-

aspek kehidupan nasional yang menyangkut kehidupan- kehidupan dan pergaulan hidup manusia dalam bermasyarakat, berbangsa, dan bernegara dengan ikatan- ikatan, aturan- aturan, dan norma- norma tertentu. Gatra Ekonomi masuk kedalam Pancagatra yang meliputi: Ideologi, Politik, Ekonomi, Sosial Budaya dan Pertahanan Keamanan.

a. Gatra Ekonomi

Penyelenggaraan ekonomi diupayakan agar dapat meningkatkan kemakmuran seluruh rakyat secara selaras, adil, dan merata untuk mewujudkan perekonomian yang mandiri dan handal, dalam usaha bersama yang berasaskan ekonomi kekeluargaan, yaitu demokrasi ekonomi yang berlandaskan Pancasila dan UUD 1945 [26].Kegiatan ekonomi ialah kegiatan-kegiatan yang berkaitan dengan: produksi, distribusi, dan konsumsi barang dan jasa; usaha- usaha untuk meningkatkan taraf hidup masyarakat, secara individu maupun kelompok; cara-cara atau alat yang dipergunakan di dalam kehidupan manusia untuk memenuhi kebutuhan. Faktor- faktor yang secara obyektif berpengaruh terhadap stabilitas maupun kelangsungan kehidupan ekonomi bangsa adalah:

- Sifat keterbukaan perekonomian
Dengan sistem perekonomian yang dianut oleh suatu negara. Pada kasus ini, Indonesia pada halnya kebanyakan negara berkembang, menganut sistem ekonomi campuran (*the middle way system*) yang intinya berupa penerapan mekanisme pasar dengan modifikasi seperlunya serta pemberian insentif untuk mendorong produktivitas.
- Struktur Ekonomi

Struktur ekonomi Indonesia seperti terlihat pada komposisi sumbangan sektor-sektor ekonomi pada pembentukan Produk Domestik Bruto (PDB) mencerminkan seberapa jauh pengaruh- pengaruh tersebut memberi dampak terhadap masing-masing sektor apabila menghadapi tantangan, ancaman, hambatan, dan gangguan.

- Potensi dan pengelolaan sumber daya alam
Sebagai pernyataan indikatif, dapat dibenarkan bahwa Indonesia mempunyai kekayaan alam secara potensial yang besar dan beraneka ragam. Dalam buku Ketahanan Nasional yang dikeluarkan oleh Lemhanas, dalam masalah ini terfokus menjadi dua yaitu masalah pertanian, dan masalah energi.
- Potensi dan pengelolaan sumber daya manusia
Permasalahan pada bidang sumber daya manusia terbagi menjadi 3, yaitu (1) Masalah tenaga kerja, pertumbuhan penduduk yang cepat dan tidak diimbangi dengan perluasan kesempatan kerja menimbulkan pengangguran kelihatan dan tidak kelihatan; (2) Masalah kesempatan kerja, menyebabkan banyak tenaga kerja di daerah pedesaan tidak terserap secara penuh. Disamping kurangnya kesempatan kerja, sebagian besar tenaga kerja yang tersedia tidak memiliki keterampilan yang dibutuhkan untuk menopang pekerjaan yang dibebankan kepadanya; (3) Masalah kemiskinan, keadaan ketenagaan serta kesempatan kerja yang tidak memadai ditambah pula dengan tingkat pendapatan yang rendah erat hubungannya dengan masalah kemiskinan di Indonesia. Rendahnya tingkat pendapatan serta tidak dimilikinya keterampilan tertentu, disamping lapangan pekerjaan

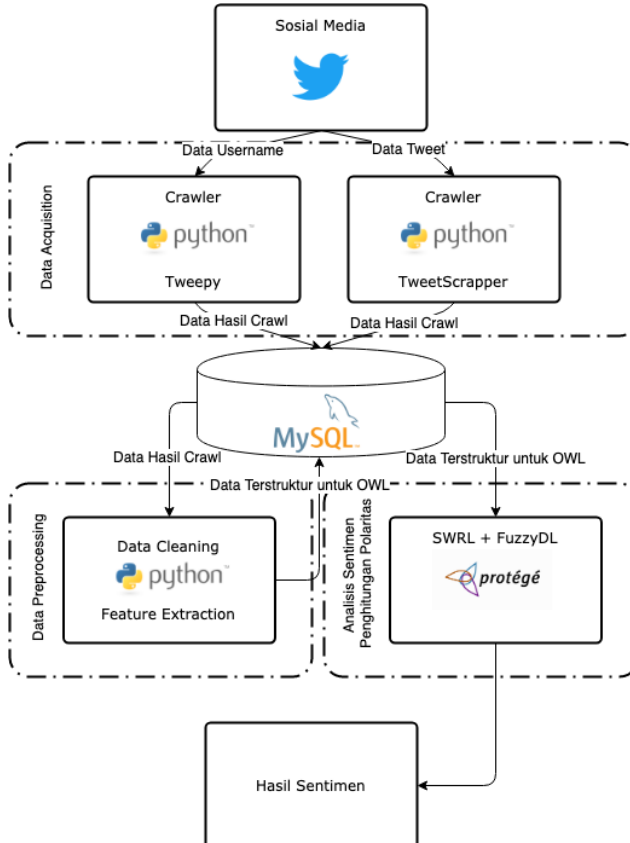
yang masih terbatas, merupakan faktor penyebab timbulnya kemiskinan di daerah perkotaan.

- Potensi dan pengelolaan sumber dana
Sumber dana baik dari dalam maupun luar negeri sangat penting bagi upaya meningkatkan pembangunan dan pengembangan ekonomi. Untuk memobilisasikan dana yang diperoleh untuk pembangunan nasional perlu terus dikembangkan berbagai kebijaksanaan dalam: (1) Sektor pengerahan dana, melalui kelembagaan seperti deposito berjangka, Tabanas dan Taska, Pasar Uang dan Pasar Modal serta Emisi Saham untuk Masyarakat; (2) Sektor investasi, ialah dengan Undang-undang Penanaman Modal Dalam Negeri dan Undang-undang Penanaman Modal Asing.
- Teknologi
Dalam kehidupan ekonomi, teknologi merupakan faktor penting bagi upaya peningkatan berbagai kegiatan ekonomi.
- Birokrasi dan sikap masyarakat
Sistem birokrasi yang berbelit-belit berarti tidak efisien dalam kegiatan serta kehidupan ekonomi, sehingga berdampak negatif, sedang pelaksanaan sistem birokrasi yang baik akan mendukung kegiatan serta kehidupan ekonomi nasional sehingga akan memberi dampak positif bagi kondisi serta perkembangan kehidupan bangsa, karena mampu menciptakan iklim usaha yang sehat dan dinamis.
- Manajemen
Kegiatan ekonomi memerlukan penerapan manajemen yang tepat dan memadai untuk meningkatkan produktivitas dan mutu produksi barang dan jasa.
- Infrastruktur

Kegiatan ekonomi berupa produksi dan distribusi/perdagangan akan terhambat, bahkan dapat macet tanpa adanya prasarana dan sarana yang memungkinkan kelancaran arus bahan, barang, dan jasa.

BAB III METODOLOGI

3.1 Arsitektur Sistem



Gambar 3.1 Arsitektur Sistem

Arsitektur sistem yang akan dibuat dapat dilihat pada Gambar 3.1. Adapun penjelasan dari arsitektur sistem adalah sebagai berikut:

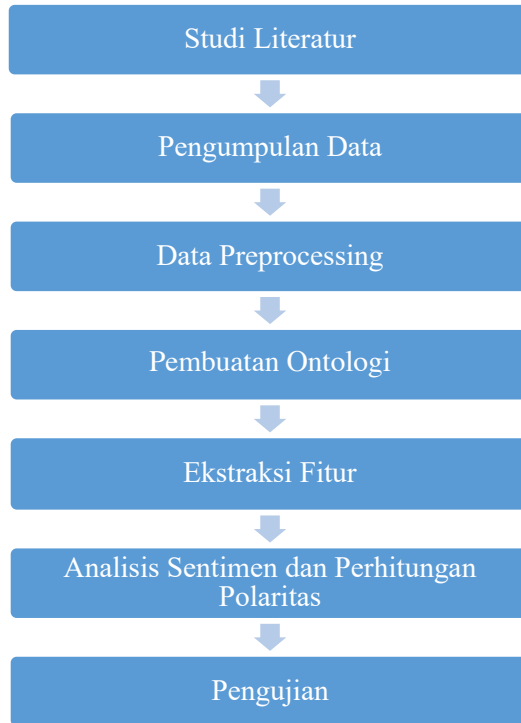
1. Sumber data pada tugas akhir ini adalah data pada Twitter.
2. Data pada Twitter diakuisisi melalui proses *crawling* menggunakan 2 cara, yang pertama menggunakan

Tweepy dan yang kedua digunakan TweetScraper. Hasil *crawling* data disimpan pada *database* MySQL.

3. Data yang telah berhasil diakuisisi kemudian dilakukan *preprocessing*, dimana data dibersihkan (filtrasi, penghapusan duplikasi, *case folding*, *tokenizing*, penghapusan *stopwords*, dll), kemudian dicari frekuensi *bigram* dan *trigram* untuk mengetahui fitur yang akan dijadikan *class* pada ontologi. Pada tahapan ini juga dilakukan ekstraksi adjektiva untuk data pembuatan *class* opini pada ontologi.
4. *Feature Extraction* dilakukan setelah mengetahui fitur yang didapatkan pada tahapan sebelumnya, dimana *tweet* yang memiliki fitur akan diekstrak, bersamaan dengan kata-kata yang mengandung opini.
5. Setelah fitur dan kata opini diekstrak, dengan menggunakan SWRL pada Protégé dan FuzzyDL, hasil sentimen dapat diketahui.

3.2 Tahapan Pelaksanaan Tugas Akhir

Pada subbab ini dijelaskan mengenai metodologi dalam pelaksanaan tugas akhir. Metodologi ini dapat dilihat pada gambar dibawah ini



Gambar 3.2 Metodologi

3.1.1 Studi Literatur

Pada tahap ini dilakukan pemahaman konsep, metode, dan teknik yang sesuai dengan bahasan dan permasalahan sehingga ditemukan sebuah cara yang menjadi solusi untuk penyelesaian permasalahan pada penyusunan tugas akhir ini. Adapun literatur yang digunakan dalam penelitian ini adalah literasi yang berkaitan dengan *Fuzzy ontology-based sentiment analysis*.

3.1.2 Pengumpulan Data

Pada tahapan pengumpulan data, digunakan *application programming interfaces (APIs)* pada situs sosial media Twitter. *Tweets* dapat diperoleh menggunakan *representational state transfer (REST) APIs* dan *Streaming APIs*. Menggunakan *REST APIs*, *tweets* dapat diperoleh sampai 3200 *tweets* terbaru. *REST APIs* memungkinkan penggunaanya untuk melakukan *query* untuk perolehan *tweets*, *query* ini mengandung seperangkat kata kunci, operator (AND dan OR), radius dan centroid (contoh: kecekalkaan AND (car OR vehicle)). Batas dari *REST APIs* adalah 350 *query* per 15 menit dan 3200 *tweets* dalam sekali *query*. *Query* yang dipakai adalah *query* berdasarkan kata kunci- kata kunci yang berkaitan dengan ekonomi.

3.1.3 Data Preprocessing

Pada tahap *pre-processing*, pengumpulan teks dan *pre-processing* korpus adalah tujuan utama sebelum penggalian sentimen. Tahap ini mengubah data yang telah terkumpul menjadi petunjuk sentimen sehingga dapat dengan mudah dalam ekstrak fitur dan kata-kata sentimen. *Tweets* diolah untuk menghilangkan konten-konten yang tidak diperlukan seperti simbol (@, #, tanggal, dsb) imbuhan (si, sebuah), dan teks di *stem* sehingga menjadi bentuk akarnya, seperti meminum, minuman, dibersihkan menjadi kata akarnya yaitu minum. Selanjutnya, teks ditandai (*tag*) untuk mengidentifikasi *parst of speech (POS)*. Lalu kemudian *tweets* dibagi menjadi kalimat-kalimat, yang kemudian dikonfirmasi sebagai sebuah klausa lengkap dengan kata benda dan frasa kata kerja. Pembagian ini memudahkan penandaan fitur pada *fuzzy-ontology based* menjadi mudah.

3.1.4 Pembuatan Ontologi

Ontologi yang dimaksud pada penelitian tugas akhir ini adalah deskripsi formal yang eksplisit tentang konsep dalam domain wacana (*domain of discourse*) (kelas/*class* atau konsep), properti dari setiap konsep yang menggambarkan berbagai fitur dan atribut konsep (slot atau properti/*properties*), dan pembatasan slot (*facets/pembatasan peran/role restriction*). Dalam pembuatan ontologi terdapat aturan- aturan dasar yang akan terkesan dogmatis, antara lain; (1) Tidak ada cara yang benar dalam memodelkan sebuah domain—akan ada banyak alternatif. Solusi terbaik akan selalu bergantung dengan aplikasi yang ada di pikiran dan perluasan yang dapat diantisipasi, (2) Pembuatan ontologi adalah proses yang iteratif, dan (3) Konsep dalam ontologi harus dekat dengan objek (fisik maupun logis) dan hubungan dengan peminatan domain. Hal ini kemungkinan adalah kata benda (objek) atau kata kerja (hubungan) dalam kalimat yang menggambarkan domain. Pada pembuatan ontologi ini akan mengacu kepada faktor- faktor dari Ekonomi yang sudah tertuliskan di dasar teori.

Pada penelitian ini penulis akan menggunakan Protégé sebagai alat untuk pembuatan ontologi. Pada penelitian Farman Ali dkk mengenai penggunaan *fuzzy ontology-based sentiment analysis of transportation and city feature reviews for safe travelling*, mereka membuat kelas- kelas fitur yang nantinya akan digunakan untuk pengekstrasian fitur pada data- data *tweets* yang dapat dilihat pada gambar- gambar dibawah ini.



Gambar 3.3 Contoh Ontologi pada OntoGraf

3.1.5 Ekstraksi Fitur

Tahapan ekstraksi fitur dilakukan setelah data preprocess dan pembuatan ontologi telah dilakukan. Melihat contoh pada penelitian yang dilakukan oleh Farman Ali dkk, *tweet* yang telah melalui tahapan *preprocess* dicocokkan dengan kelas ontologi (fitur) yang telah dibuat. Fitur yang cocok dengan frasa kata benda menunjukkan bahwa *tweet* cocok dengan fitur pada ontologi. Contoh, *park* adalah kata benda pada *tweet* “*Park is very clean*”. Pada ontologi *city*, karena *park* adalah *class* dari *city*, maka akan teridentifikasi bahwa frasa “*Park is very clean*” adalah frasa yang membahas tentang fitur *Park*. Selanjutnya, fitur- fitur yang terekstrak akan dipertimbangkan untuk diprediksi polaritasnya, dimana setiap fitur memiliki kata-kata polaritas yang berbeda-beda, dalam contoh ini fitur *Park*. Fitur *Park* yang diiringi kata-kata bersih dan bagus memiliki nilai polaritas

positif; berdebu dan kecil memiliki nilai polaritas negatif; dan sedang, normal, rata-rata adalah kata-kata opini yang memiliki nilai polaritas netral.

3.1.6 Analisis Sentimen dan Perhitungan Polaritas menggunakan Fuzzy Logic

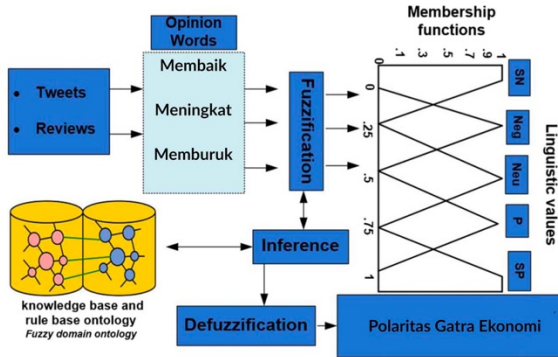
Bagian ini menjelaskan metodologi analisis sentimen yang digunakan, dilanjutkan dengan perhitungan polaritas menggunakan *fuzzy inference layer*.

a. Analisis Sentimen menggunakan SentiWord Scoring

Setelah fitur-fitur pada data hasil *crawling* telah diekstrak, maka kata-kata opini akan dianalisis sentimennya menggunakan SentiWordNet. SentiWordnet adalah sumber daya leksikal yang nantinya akan terhubung dengan setiap kata dan akan memiliki tiga skor numerik: positif, objektif, dan negatif [22]. Skor-skor ini menunjukkan seberapa dalamkah sebuah istilah masuk kedalam *synset* positif, netral, ataupun negatif. Setelah diolah menggunakan SentiWordNet, hasil skor akan digunakan untuk tahapan berikutnya yaitu perhitungan polaritas menggunakan *fuzzy inference layer*.

b. Perhitungan Polaritas berdasarkan Fuzzy Inference Layer

Setelah kata-kata opini dihitung nilai menggunakan SentiWordNet, selanjutnya nilai-nilai tersebut akan diolah menggunakan *fuzzy inference layer*.



Gambar 3.4 Fuzzy Inference Layer untuk Perhitungan Polaritas

Fuzzy inference layer mempunyai 4 komponen, yaitu bagian *fuzzification*, *inference*, *knowledge-base* dan *rule base*, dan yang terakhir *defuzzification*. Pada komponen *fuzzification*, dari kata-kata opini yang telah mendapatkan nilai pada Sentiword Scoring akan didapatkan *membership value*, dalam kata lain adalah suatu proses pengubahan nilai tegas (*crisp values*) ke dalam keanggotaan *fuzzy*. *Knowledge base* dan *rule base* adalah suatu bentuk aturan relasi/ implikasi if-then, contohnya adalah ketika nilai dari Sentiword didapatkan 0,1 maka akan masuk kedalam *value* Strong Negative. Pada *inference engine*, proses implikasi dalam menalar nilai masukan untuk menentukan nilai keluaran sebagai bentuk pengambilan keputusan. Sedangkan pada tahapan *defuzzification*, nilai-nilai yang dihasilkan oleh himpunan fuzzy akan dijadikan ke nilai tegas kembali (*crisp values*).

3.1.7 Pengujian

Untuk melihat kegunaan dari arsitektur yang digunakan, performa *fuzzy ontology* dihitung dengan tingkat akurasi. Tingkat akurasi dihitung dengan menghitung antara nilai

Precision (P), *Recall (R)*, *Accuracy (Ac)*, dan *Function Measure (FM)*. Secara matematika, *precision*, *recall*, *accuracy* dan *function measure* dapat dihitung menggunakan persamaan dibawah ini.

$$Precision (P) = \frac{TP}{(TP + FP)} \times 100\%$$

$$Recall (R) = \frac{TP}{(TP + FN)} \times 100\%$$

$$Accuracy (Ac) = \frac{(TP+TN)}{(TP+FP+FN+TN)}$$

$$Function Measure (FM) = \frac{2'P'R}{P + R}$$

Dimana “TP”, “FP”, “FN”, dan “TN” adalah “*true positive*”, “*false positive*”, “*false negative*”, dan “*true negative*”.

Halaman ini sengaja dikosongkan

BAB IV PERANCANGAN

4.1 Akuisisi Data

4.1.1. Akuisisi Data Username

Dalam melakukan analisis sentimen gatra ekonomi, data berupa tweets adalah obyek utama yang akan dianalisis. Tweet yang diambil adalah tweet yang berasal dari *buzzer* serta tokoh masyarakat, sehingga diperlukan dua tahapan pada akuisisi data ini. Adapun dua tahapan yang dimaksud adalah pengumpulan user dan pengumpulan tweets. Pengumpulan user adalah tahapan pencarian user pada Twitter yang masuk kedalam tokoh masyarakat dan *buzzer*, sedangkan pengumpulan tweets adalah pengumpulan dari tweets yang berasal dari user yang dimaksud pada tahapan pertama.

Pada tahapan pertama, pencarian user dilakukan dengan menggunakan hashtag yang paling sering digunakan oleh akun-akun umpan yang terdiri dari 13 user dengan followers terbanyak dari masing-masing kubu, akun-akun partai politik, dan 50 akun media dengan followers terbanyak. Untuk lebih lengkapnya dapat dilihat pada tabel dibawah ini.

Tabel 4.1 Akun Umpan

addiems	TRANSTV_CORP	MetroTVNewsRoom	RamliRizal	detikinet
budimandjtmiko	SCTV_	detikfinance	haikal_hassan	tictoc
Hary_Tanoe	VIVAcoid	SeputariNews	saididu	alineadotid
hanifdhakiri	liputan6dotcom	idntimes	ustad_alhabsyi	OfficialGTVID
rizariri	OfficialRCTI	zetizen_id	MardaniAliSera	langitRTV
IndraJPiliang	KompasTV	ochannelupdates	PriyoBudiS	BeritagarID
OlgaLy_DIA	MataNajwa	KontanNews	AndiArief_	PDI_Perjuangan

zuhairimisrawi	InfoValidKita	ILCtv1	panca66	PDemokrat
ariabima99	netmediatama	SonoraFM92	detikcom	Gerindra
andrinof_a_ch	KapanLagicom	CNNIndonesia	Metro_TV	DPPPartaiGaruda
andreOPA	IndosiarID	nusantaraneWSco	kompascom	Golkar5
TsamaraDKI	korantempo	TwitterID	TRANS7	PartaiHANURA
jarwokuat	tempodotco	TirtoID	kompasiana	PKSejahtera
tifsembiring	mediaindonesia	ndorokakung	jakpost	DPP_PKB
hnurwahid	antaranews	SINDOnews	Official_MNCTV	NasDem
fadlizon	republikaonline	PRFMnews	hariankompas	PartaiPerindo
Fahrihamzah	BBCIndonesia	Official_PAN	tribunnews	DPP_PPP
fahraidris	Bisniscom	rizkyberkarya	temponewsroom	psi_id



Indonesia is ready to welcome the new president Mr. @prabowo · Apr 7

Indonesia is ready to welcome the new president Mr. @prabowo

Today April, 7th 2019

Gelora Bung Karno, Jakarta

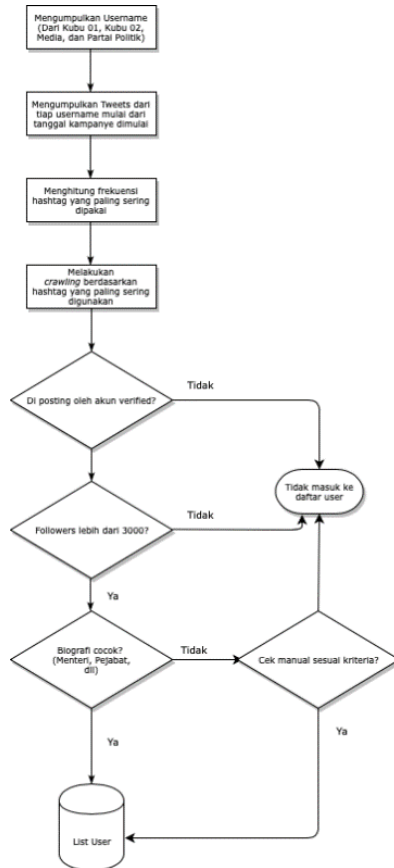
What a great Campaign ever in Indonesian History

#TheGreatCampaignOfPrabowo
#AkalSehatPutihkanGBK

Gambar 4.1 Hashtag

Dari daftar akun umpan diatas, dilakukan penggalian data dengan mengambil *tweets* dari akun tersebut semenjak tanggal dimulainya kampanye yaitu 23 September 2018. Penggalian data pada tahap ini menggunakan Tweepy, sebuah *library* dari Python yang digunakan untuk menghubungkan Python dengan TwitterAPI. Hasil *crawl* menggunakan Tweepy akan menghasilkan file berbentuk JSON, dimana didalam JSON

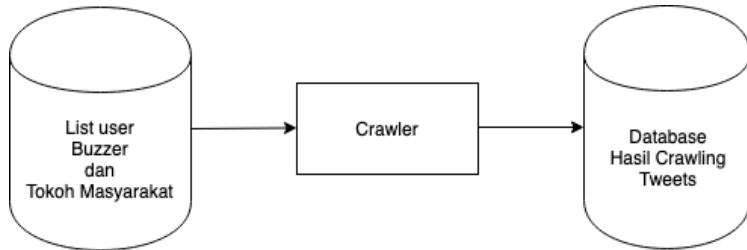
tersebut terdapat Objek *Entity*, yang didalamnya terdapat *Entity Hashtag*. *Entity hashtag* ini akan dikumpulkan sehingga diketahui *hashtag* apa saja yang paling sering digunakan. Setelah mendapatkan *hashtag* yang paling sering digunakan, maka selanjutnya akan kita lakukan *crawling* ulang dengan menggunakan *keyword hashtag* yang telah kita dapatkan. Dari tweet-tweet yang didapat, selanjutnya akan dilakukan pengecekan user apakah user tersebut masuk kedalam kategori *buzzer* ataupun tokoh masyarakat. Keseluruhan proses akuisisi data username dapat dilihat pada *flowchart* dibawah ini.



Gambar 4.2 Flowchart *Crawling User*

4.1.2. Akuisisi Data Tweets dari User

Setelah akun-akun telah didapat, selanjutnya dilakukan akuisisi data berupa tweets dari akun-akun tersebut. Tweets yang dikumpulkan adalah tweet yang di post pada saat masa kampanye, yaitu dimulai pada tanggal 23 September 2018 sampai 13 April 2019.



Gambar 4.3 Flowchart *Crawling Tweets*

4.2. Perancangan *Pre-processing Dataset*

4.2.1. Perancangan Penghapusan Duplikasi Dataset

Data yang telah diperoleh melalui proses crawling dipastikan kembali tidak terdapat duplikasi, oleh karena itu diperlukan penghapusan data duplikat. Hal ini dilakukan menggunakan query sederhana yang di eksekusi melalui MySQL.

4.2.2. Perancangan *Filtering Dataset*

Setelah dilakukan hasil crawling, perlu dilakukan filtering dataset sehingga dataset yang akan diolah lebih berkaitan dengan Gatra Ekonomi, sehingga pada tahapan selanjutnya data-data yang tidak diperlukan tidak ikut diolah yang dapat berpengaruh kepada kecepatan pemrosesan. Filtering dilakukan dengan berbagai macam query yang dilakukan pada MySQL, didasarkan pada kata-kata kunci faktor yang mempengaruhi Gatra Ekonomi pada buku Ketahanan Nasional.

Tabel 4.2 *Keyword Gatra Ekonomi*

No	Faktor	Keyword
1	Sifat Keterbukaan Perekonomian	<ul style="list-style-type: none"> • Ekonomi • Sistem Perekonomian • Sistem Ekonomi
2	Struktur Ekonomi	<ul style="list-style-type: none"> • Sektor Ekonomi • Struktur Ekonomi

		<ul style="list-style-type: none"> • Produk Domestik Bruto
3	Potensi dan Pengelolaan Sumber Daya Alam	<ul style="list-style-type: none"> • Sumber Daya Alam • Pertanahan • Energi
4	Potensi dan Pengelolaan Sumber Daya Manusia	<ul style="list-style-type: none"> • Sumber Daya Manusia • Tenaga Kerja • Kesempatan Kerja • Kemiskinan • Tingkat Pendapatan
5	Potensi dan Pengelolaan Sumber Daya Dana	<ul style="list-style-type: none"> • Sumber dana • Sektor Pengerahan Dana • Pasar Uang • Pasar Modal • Investasi • Penanaman Modal • Modal Dalam Negeri • Modal Asing
6	Birokrasi dan Sikap Masyarakat	<ul style="list-style-type: none"> • Sistem Birokrasi
7	Teknologi	<ul style="list-style-type: none"> • Teknologi
8	Manajemen	<ul style="list-style-type: none"> • Produktivitas • Mutu Produksi Barang • Mutu Jasa
9	Infrastruktur	<ul style="list-style-type: none"> • Infrastruktur • Prasarana • Sarana • Arus Bahan • Arus Barang • Arus Jasa

4.2.3. Perancangan Pembersihan *Dataset*

Data yang telah diperoleh melalui proses crawling masih kotor, yaitu masih terdapat banyak symbol- symbol dan karakter yang tidak diperlukan dalam proses selanjutnya. Maka dari itu, perlu dilakukan pembersihan data hasil crawling. Berikut adalah langkah-langkah yang diperlukan:

1. *Case folding*
2. *Data Cleaning*
3. *Tokenizing*
4. *Stopwords Removal*

Dibawah ini merupakan contoh dari data yang telah dibersihkan dan ditokenisasi.

```
["menjaga', 'budaya', 'zaman', 'modern',
'sulit', 'generasi', 'melek', 'teknologi',
'budaya', 'karakter', 'bawa']"
["maraknya', 'ojek', 'online', 'hasil',
'perkembangan', 'teknologi', 'digital',
'terhindarkan', 'teknologi', 'sepenuhnya',
'menggantikan', 'manusia', 'manusia',
'berguyub']"
["twit', 'hilang', 'masuk', 'time',
'line', 'huawei', 'telcom', 'china',
'menguasai', 'pasar', 'teknologi',
'telekomunikasi', 'ri', "'karyawan'",
'huawei', 'china', 'masuk', 'ri',
'didukung', 'izin', 'lengkap', 'mrka',
'hnya', 'bayar', 'suap', 'usd', 'orang',
'imigrasi', 'bandara']"
["hati', 'produk', 'aplikasi', 'made',
```

Gambar 4.4 Contoh kode yang telah dibersihkan

4.2.4. Perancangan POS Tagger

Perancangan POS Tagger menggunakan library NLTK dan CRF Tagger, dimana tagger yang dipakai adalah tagger berbahasa Indonesia yaitu `all_indo_man_tag_corpus_model.crf.tagger`.

Hasil dari tagging ini berupa list yang berisi kata-kata yang telah

ditokenisasi dan di tagging apakah sebuah kata termasuk kata benda (NN), kata benda (VB), konjungsi (CC), dll.

```
"[['menjaga', 'VB'], ('budaya', 'NN')],
[['budaya', 'NN'], ('zaman', 'NN')],
[['zaman', 'NN'], ('modern', 'JJ')],
[['modern', 'JJ'], ('sulit', 'JJ')].
```

Gambar 4.5 Contoh Tagging

4.3. Perancangan Ontologi Gatra Ekonomi

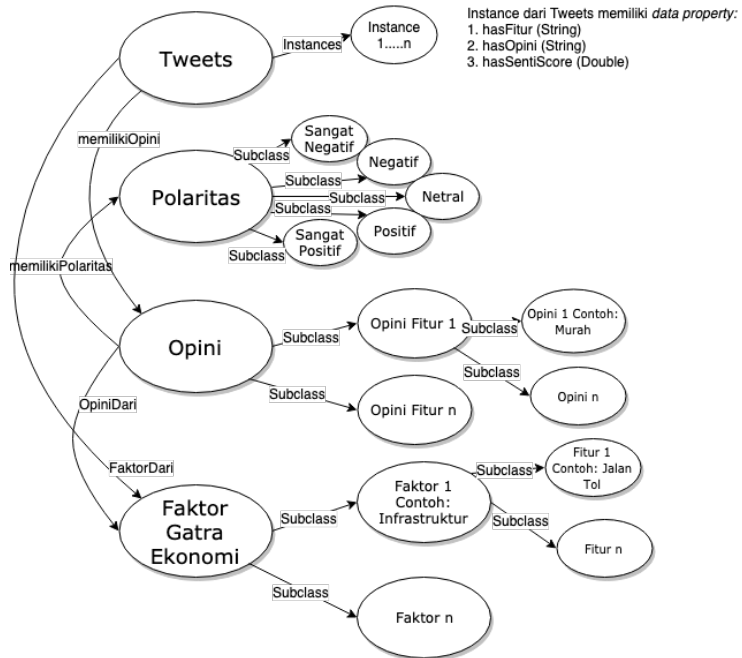
4.3.1. Perancangan Konsep Ontologi

Perancangan Ontologi Gatra Ekonomi dilakukan menggunakan perangkat lunak Protégé, dimana Ontologi yang dibuat berdasarkan buku Ketahanan Nasional. Pada buku Ketahanan Nasional dijelaskan bahwa Ketahanan Nasional pada Gatra Ekonomi dipengaruhi oleh delapan faktor yang secara obyektif berpengaruh terhadap stabilitas maupun kelangsungan kehidupan ekonomi bangsa, antara lain:

1. Sifat keterbukaan perekonomian (sistem ekonomi)
2. Struktur Ekonomi (Produk Domestik Bruto/ PDB, struktur)
3. Potensi dan pengelolaan sumber daya alam (Pertanahan, masalah energi)
4. Potensi dan pengelolaan sumber daya manusia (Tenaga kerja, pertumbuhan penduduk, kesempatan kerja, kemiskinan, lapangan pekerjaan)
5. Potensi dan pengelolaan sumber dana (Sektor penerahan dana, pasar uang, pasar modal, emisi saham, investasi)

6. Birokrasi dan sikap masyarakat (sistem birokrasi)
7. Teknologi
8. Manajemen (produktivitas, mutu produksi)
9. Infrastruktur (prasarana, sarana, kelancaran arus bahan, barang, dan jasa)

Dikarenakan tujuan dari tugas akhir ini adalah mengetahui sentimen masyarakat terhadap gatra ekonomi, maka konsep ontologi yang akan dibuat dapat dilihat dibawah ini.

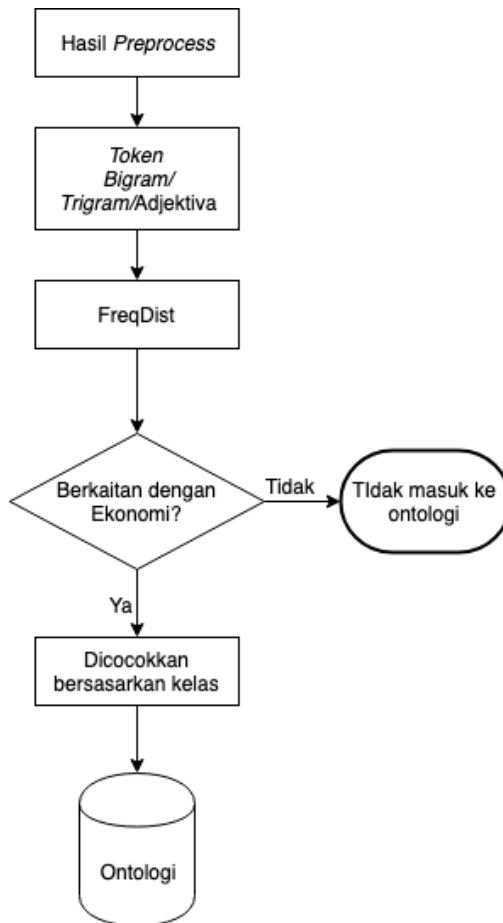


Gambar 4.6 Perancangan Konsep Ontologi

4.3.2. Perancangan *Subclass* Faktor Gatra Ekonomi

Untuk mengisi ontologi tersebut, diperlukan sebuah proses yang dapat menunjukkan token-token *bigram/trigram* yang paling

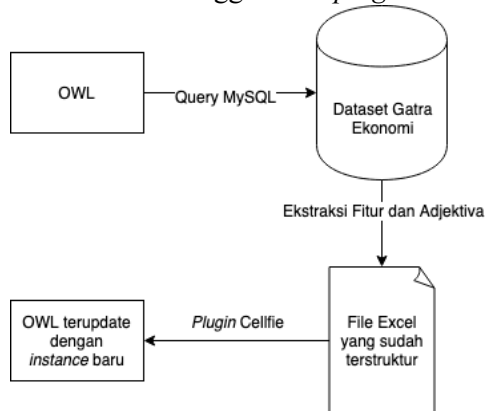
sering dibahas pada hasil data crawl. Untuk itu, diperlukan tokenisasi berupa *bigram/trigram* dan menggunakan library `nltk.probability FreqDist`. Dari hasil tersebut, *bigram/trigram* dengan frekuensi terbanyak dan berkaitan dengan gatra ekonomi dimasukkan ke dalam ontologi dengan dicocokkan dengan kelas yang telah dibuat di awal.



Gambar 4.7 Flowchart pembuatan Ontologi

4.4. Perancangan Ekstraksi Fitur

Ekstraksi fitur dilakukan dengan memanfaatkan hasil dari *keyword* pembuatan ontologi. Fitur-fitur digunakan dalam *query* MySQL pada *dataset* yang telah difilter menggunakan kata kunci gatra ekonomi. Jika pada suatu *tweet* terkandung suatu fitur, maka *tweet* tersebut diekstrak, dan kemudian diekstrak kata adjektivanya. Kemudian, data dibentuk dalam bentuk excel, untuk kemudian dimasukkan kedalam ontologi, khususnya kelas Tweets menggunakan *plugin* cellfie.



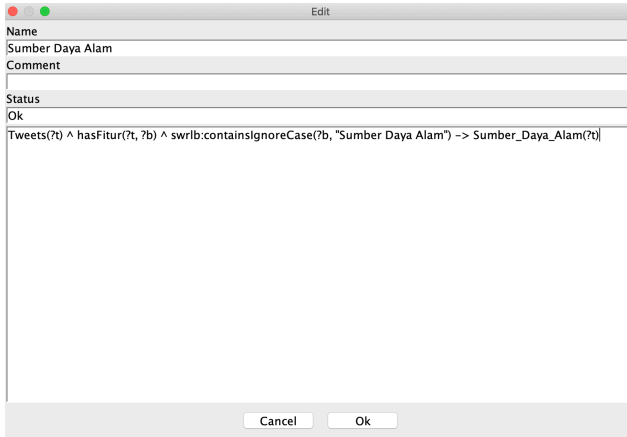
Gambar 4.8 Ekstraksi Fitur

4.5. Perancangan SentiWord Scoring

SentiWordNet merupakan kumpulan dari leksikal yang dibuat untuk mendukung klasifikasi sentimen dan proses-proses opinion mining pada aplikasi. SentiWord pada kasus ini digunakan untuk mengetahui nilai sentimen dari kata adjektiva dan kata adverbial. Penulis menggunakan SentiWordNet yang telah disesuaikan dengan Bahasa Indonesia yaitu Barasa yang didapatkan dari GitHub. Kata adjektiva yang telah diekstrak akan diberikan skor berdasarkan SentiWordNet tersebut, sehingga dapat diketahui polaritas dari sebuah kata.

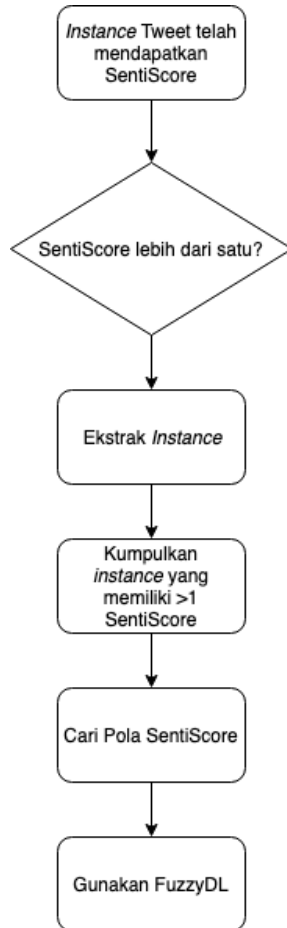
4.6. Perancangan SWRL-Rule dan Fuzzy Ontology

Perancangan SWRL dilakukan langsung pada aplikasi Protégé dengan memanfaatkan *plugin* SWRLTab. SWRL-Rule dibuat dengan *syntax* yang sesuai dengan penulisan SWRL.



Gambar 4.9 SWRL Rule

Sedangkan perancangan *Fuzzy Ontology* dibuat dengan menggunakan sebuah bahasa *description logic* yang telah dimodifikasi sehingga mengerti konsep *fuzzy*, yaitu FuzzyDL. FuzzyDL adalah sebuah *description logic* yang dibuat oleh peneliti dari Italia yang bernama Umberto Straccia. Dengan menggunakan FuzzyDL, dapat dibuat sebuah konsep yang dapat diubah menjadi bentuk ontologi. *FuzzyDL* digunakan ketika sebuah *instance* memiliki nilai sentiment lebih dari satu.



Gambar 4.10 Flowchart Pengolahan Tweet Fuzzy

Halaman ini sengaja dikosongkan

BAB V IMPLEMENTASI

5.1 Lingkungan Implementasi

Pengerjaan penelitian ini menggunakan komputer dengan spesifikasi berikut:

Tabel 5.1 Spesifikasi Perangkat Keras

Prosesor	Intel Core i5 3,1 GHz
Memory	8 GB 2133 MHz LPDDR3 SDRAM
GPU	Iris Plus Graphics 650
OS	MacOS Mojave 10.14.4
Arsitektur	64bit

Selain itu, terdapat beberapa library, database, dan bahasa pemrograman yang membantu pengerjaan penelitian ini sebagai berikut.

Tabel 5.2 Daftar Library, Editor, dan Bahasa Pemrograman

Webserver	Apache 2.4.27
Bahasa Pemrograman	Python 3.7
Database	MySQL 10.1.26
Programming Editor	Visual Studio Code 1.35.1
Ontology Editor	Protégé 5.5.0, Owlready 0.2, FuzzyDL
Library	<ul style="list-style-type: none">• Tweepy 3.5.0• Scrapy 1.6

	<ul style="list-style-type: none"> • Pandas 0.24.1 • Mysql-connector 8.0 • Jsonpickle 1.0 • Nltk 3.4.3 • Owlready 0.2
--	--

5.2 Pembuatan Crawler Twitter

5.2.1 Crawler Hashtag

Crawler hashtag dibuat dengan memanfaatkan Tweepy, sebuah *library* dari Python untuk memudahkan akses ke Twitter API. Tweepy digunakan pada pembuatan *crawler hashtag* dikarenakan memudahkan pengambilan data berupa JSON melalui Twitter API. Pada data JSON yang diambil, terdapat objek *entity* yang mengandung *hashtag*, yang pada proses ini digunakan untuk mengetahui *hashtag* apa yang paling sering digunakan.

```

1. consumer_key = 'sfpEcxsyXeUPdNydbq3fTmkEq'
2. consumer_secret = 'KOFgKNuQyqiqD4yZvtmyhtsYfOI3aiL
   EPdqTaHZanPCLEkK'
3. access_token='957424880131047424Q565IXfqB7ND0mKAdsa
   dnNlMq3djLUH'
4. access_secret = 'EXyUxDNAndjDjgXGEBp123tYc73lgn51l
   F2E12yDwyJppzIP'
5. auth = OAuthHandler(consumer_key, consumer_secret)
6. auth.set_access_token(access_token, access_secret)
7. api = tweepy.API(auth, wait_on_rate_limit=True)

```

Kode 5.1 Potongan Kode Implementasi API untuk Crawling

Kode diatas merupakan potongan kode untuk mengautentikasi klien Twitter API, jika sudah terautentikasi, maka Twitter API sudah dapat digunakan.

```

1. ...
2. def getTweets(string):
3.     startDate = dt.datetime(2018, 9, 23, 0, 0, 0)
4.     endDate = dt.datetime(2019, 4, 6, 0, 0, 0)
5.
6.     for t in tweepy.Cursor(api.user_timeline, screen_name=string, tweet_mode="extended").items(2000):
7.
8.         if t.created_at < endDate and t.created_at > startDate:
9.             f = open("jadiin1.csv", "a+")
10.            f.write("\n"+ str(t.id) + ";" + str(
11.                username) + ";" + str(t.full_text.encode('utf-8')) + ";" + str(t.created_at))
12.            print(username)"""
13.
14. def getPage(username):
15.     startDate = dt.datetime(2018, 9, 23, 0, 0, 0)
16.     endDate = dt.datetime(2019, 4, 6, 0, 0, 0)
17.
18.     for page in tweepy.Cursor(api.user_timeline, screen_name=username, tweet_mode="extended").pages():
19.
20.         for t in page:
21.             if t.created_at < endDate and t.created_at > startDate:
22.                 f = open("jadiin1.csv", "a+")
23.                 f.write("\n"+ str(t. [hashtag['text'] for hashtag in tweet['entities']['hashtags']]))
24.                 print(username)"""
25.
26. ...
27. def getTweets():
28.     df = pd.read_csv("username.csv", sep=";", error_bad_lines=False, encoding='cp1252',
29.                     warn_bad_lines=False)
30.     usernames = df.username
31.     keterangan = df.keterangan
32.     kubu = df.kubu
33.
34.     for user in usernames:
35.         util = getPage(user)

```

Kode 5.2 Potongan Kode Implementasi *Crawler Tweets* pada *User*

Kode diatas digunakan untuk menggunakan Twitter API untuk mengambil *tweets* dari username, implementasi dari 4.1.1 Akuisisi Data Username, dengan memanfaatkan library `api.user_timeline`, dan pengaturan `tweet_mode` diubah menjadi `"extended"`. Tanggal *tweets* yang akan diambil diatur sesuai dengan tanggal dimulainya kampanye yaitu 23 September 2018 dengan `startDate = dt.datetime(2018, 9, 23, 0, 0, 0)`. Hasil *crawl* akan dijadikan file berbentuk file csv yang memiliki satu kolom yaitu hashtag sebagaimana dapat dilihat pada gambar dibawah ini.

1	hashtag
2	['Jokowiliriana']
3	['Musisi01']
4	['KonserPutihBersatu', 'BarengJokowi']
5	['NgapakBarengJokowi']
6	['Coblos01SudahTerbukti', 'IndonesiaMaju']
7	['IsraMiraj']
8	['RabuPutihJokowiMenang']
9	['AmerikaBerSATU', 'RelawanJokowiNY']
10	['MerahPutihSatukanKita', 'TetapMerahPutih', 'IndonesiaMaju']
11	['OptimisIndonesiaMaju']
12	['SudahTamatBos', 'Coblos01SudahTerbukti']
13	['HtiSampahDunia']
14	['KonserPersembahanku', 'KeberagamanItulndah', 'IndahKurnia4RI', 'IndahKurnia4', 'IndahKurnia', 'IndonesiaMaju']
15	['GaspolJokowiJempol', '01IndonesiaMaju', '2019PilihJokowi']
16	['01TerusAja', '2019PilihJokowi']
17	['01JokowiLagi', '2019JokowiKyailMaruf']
18	['gaspoljokowijempol', '01jokowilagi', '2019jokowilagi']
19	['GaspolJokowiJempol']
20	['Ad']
21	['BantenMeluJokowiAmin']

Gambar 5.1 Hasil *Crawl* Hashtag

a. Pengolahan Frekuensi Hashtag

Setelah hashtag didapatkan, selanjutnya data diolah dengan menggunakan library `nltk` untuk mengetahui hashtag apakah yang memiliki frekuensi tertinggi.

```

1. import csv
2.
3. with open('Data_terbaru.csv') as csv_file:
4.     csv_reader = csv.reader(csv_file, delimiter=',',
5.                               )
6.     line_count = 0
7.     listHashtag = []
8.     for row in csv_reader:
9.         listHashtag.append(row[0])
10.        line_count +=1
11.
12. strTokens = ''.join(str(i) for i in listHashtag)
13. print(strTokens)
14. print(f'Processed {line_count} lines.')

```

Kode 5.3 Potongan Kode Membaca File CSV ke Python

Potongan kode diatas digunakan untuk membuka data hasil crawl yang berbentuk csv dan memasukkan tiap baris dari csv tersebut ke dalam listHashtag.

```

1. from nltk.tokenize import sent_tokenize, word_tokenize
2. symbols = [' ', '.', ':', ')', '[', ']', '"', "'",
3.            "?", '\\n', '\\n2']
4. symbols = set(symbols)
5. tokens = word_tokenize(strTokens)
6. fixTokens = [w for w in tokens if w not in symbols]
7.
8. from nltk.probability import FreqDist
9. fdist = FreqDist()
10. for word in fixTokens:
11.     fdist[word.lower()]+=1
12. fdist.most_common()

```

Kode 5.4 Potongan Kode FreqDist untuk Mencari Frekuensi Kata Terbanyak

Potongan kode diatas menggunakan *library* nltk, `tokens = word_tokenize(strTokens)` digunakan untuk membuat token-token kata, sedangkan `fixTokens = [w for w in tokens if w not in symbols]` digunakan untuk

menghilangkan symbol-simbol yang tidak diperlukan seperti `symbols = [',', '.', ':', ')', '[', ']', '"', "'", "?", '\\n', '\\n2']`. Selanjutnya digunakan *library* `nlk.probability` untuk melihat frekuensi kata yang paling sering muncul.

5.2.2 Crawler Username

Pada perancangan *crawler username*, masih digunakan Tweepy untuk memudahkan akses ke Twitter API. Hal ini dilakukan karena dengan menggunakan Twitter API, hasil *crawl* berupa JSON yang mempunyai hal yang dibutuhkan untuk pencarian username, yaitu objek user. Pada objek user terdapat username, id, description, dan jumlah *follower* yang dibutuhkan untuk mengetahui termasuk ke kategori apakah user tersebut: buzzer atau tokoh masyarakat. Kode yang digunakan memanfaatkan `search rest api`.

```

1. with open('data_' + fName, 'w') as f:
2.     while True:
3.         try:
4.             if (max_id <= 0):
5.                 if (not sinceId):
6.                     new_tweets = api.search(q=search
hQuery, count=tweetsPerQry, tweet_mode='extended')
7.
8.                 else:
9.                     new_tweets = api.search(q=search
hQuery, count=tweetsPerQry,
10.                                         since_i
11.                                         d=sinceId, tweet_mode='extended')
12.                 else:
13.                     if (not sinceId):
14.                         new_tweets = api.search(q=search
hQuery, count=tweetsPerQry,
15.                                         max_id=
16.                                         str(max_id - 1), tweet_mode='extended')
17.                     else:
18.                         new_tweets = api.search(q=search
hQuery, count=tweetsPerQry,
19.                                         max_id=
20.                                         str(max_id - 1),
21.                                         since_i
22.                                         d=sinceId, tweet_mode='extended')
23.                     if not new_tweets:
24.                         print("No more tweets found")
25.                         break
26.                     for tweet in new_tweets:
27.                         f.write(jsonpickle.encode(tweet._js
on, unpicklable=False) +
28.                                 '\n')
29.                         tweetCount += len(new_tweets)
30.                         print(searchQuery + " - Downloaded {0}
tweets".format(tweetCount))
31.                         max_id = new_tweets[-1].id
32.                     except tweepy.TweepError as e:
33.                         # Just exit if any error
34.                         print("some error : " + str(e))

```

Kode 5.5 Potongan Kode *Tweepy* memanfaatkan *TwitterSearch*

Potongan kode diatas menjelaskan pemanfaatan search api yang telah disediakan oleh Twitter API, dimana kita mengambil data

melewati Twitter Search, dengan memasukkan keyword tertentu, dalam kasus ini adalah *hashtag* yang telah didapat pada *crawler hashtag*. File yang dihasilkan berformat text yang berisikan baris-baris JSON.

```
{
  "contributors": null,
  "coordinates": null,
  "created_at": "Wed Apr 10 07:34:20 +0000 2019",
  "display_text_range": [
    0,
    148
  ],
  "entities": {
    "hashtags": [
      {
        "indices": [
          77,
          88
        ],
        "text": "JokowiLagi"
      },
      {
        "indices": [
          89,
          105
        ],
        "text": "01IndonesiaMaju"
      }
    ]
  }
}
```

Gambar 5.2 JSON Hasil Crawl

Dikarenakan data masih berupa bentuk JSON yang memiliki banyak objek, maka digunakanlah sebuah kode program sederhana yang dapat mengambil objek yang diperlukan, yaitu objek username, hashtag, verified, description, followers, id, filename, text, dan timestamp. Kode program dibawah merubah baris-baris JSON menjadi bentuk query untuk SQL, sehingga data dapat dimasukkan ke dalam database SQL.


```

1. with open('HASIL_' + st + '_' + filename + '.sql', 'a') as
   saveFile:
2.     for line in f:
3.         try:
4.             tweet = json.loads(line)
5.             if 'full_text' in tweet:
6.                 username = tweet['user']['s
   creen_name']
7.                 verified = str(tweet['user'
   ][ 'verified'])
8.                 hashtag = str([hashtag['tex
   t'] for hashtag in tweet['entities']['hashtags'])]
9.                 description = tweet['user']
   ['description']
10.                followers = tweet['user']['
   followers_count']
11.                id = tweet['id']
12.                created_at = tweet['created
   _at']
13.                text = tweet['full_text']
14.                ts = time.strftime('%Y-%m-
   %d %H:%M:%S',
15.                time.strptime(tweet['cr
   eated_at'],
16.                '%a %b %d %H:%M:%S
   +0000 %Y'))
17.                abc = username,verified,hashtag
   ,description,followers,id,filename,text,ts
18.                print(text)
19.                print('INSERT INTO tweetshashta
   gs VALUES {};'.format(abc))
20.                saveFile.write('INSERT INTO twi
   tter VALUES {};'.format(str(abc)))
21.                saveFile.write('\n')
22.            except: continue
23.            print('\n' + 'COMPLETE READ : {}'.f
   ormat(filename) + '\n')

```

Kode 5.6 Potongan Kode untuk Ekstraksi JSON ke SQL

5.2.3 Crawler Tweets

Pada proses meng-*crawl* tweet dari setiap username, penulis menggunakan metode *scraping* dengan menggunakan

TweetScrapper yang memanfaatkan library Scrapy. Pada tahapan ini digunakan Scrapy untuk mengakali limit dari TwitterAPI, dikarenakan Scrapy menggunakan TwitterSearch sehingga tidak akan terkena limit dari TwitterAPI. Berikut adalah potongan kode yang digunakan, dan merupakan implementasi dari 4.1.2.

```

1. def parse_tweet_item(self, items):
2.     for item in items:
3.         try:
4.             tweet = Tweet()
5.             tweet['usernameTweet'] = item.xpath(
6.                 ('./span[@class="username u-dir u-
7.                 textTruncate"]/b/text()').extract()[0]
8.             ID = item.xpath('./@data-tweet-
9.                 id').extract()
10.            if not ID:
11.                continue
12.            tweet['ID'] = ID[0]
13.            tweet['text'] = ' '.join(
14.                item.xpath('./div[@class="js-
15.                tweet-text-
16.                container"]/p//text()').extract()).replace(' # ',
17.                '#').replace(
18.                    '@ ', '@')
19.            if tweet['text'] == '':
20.                continue
21.            tweet['url'] = item.xpath('./@data-
22.                -permalink-path').extract()[0]
23.
24.            tweet['datetime'] = datetime.fromti
25.                mestamp(int(
26.                    item.xpath('./div[@class="stre
27.                    am-item-header"]/small[@class="time"]/a/span/@data-
28.                    time').extract()[
29.                        0])).strftime('%Y-%m-
30.                %d %H:%M:%S')

```

Kode 5.7 Potongan Kode TweetScrapper

Potongan kode diatas merupakan sebuah fungsi untuk mengambil username dari pengirim tweet, ID tweet, teks, url, dan datetime dari tweet.

```

1. class SavetoMySQLPipeline(object):
2.
3.     ''' pipeline that save data to mysql '''
4.     def __init__(self):
5.         # connect to mysql server
6.         user = "root"
7.         pwd = ""
8.         self.cnx = mysql.connector.connect(user=user
9.         r, password=pwd,
10.                                         host='localhost',
11.                                         database='tweets',
12.         buffered=True)
13.         self.cursor = self.cnx.cursor()
14.         self.table_name = "akun_unverified_orang6"
15.
16.         create_table_query = "CREATE TABLE `" + s
17.         elf.table_name + "` (\
18.             `ID` CHAR(20) NOT NULL,\
19.             `url` VARCHAR(140) NOT NULL,\
20.             `datetime` VARCHAR(22),\
21.             `text` VARCHAR(280),\
22.             `user_id` CHAR(20) NOT NULL,\
23.             `usernameTweet` VARCHAR(20) NOT NUL
24.             L\
25.         )"
26.
27.     try:
28.         self.cursor.execute(create_table_query)
29.
30.     except mysql.connector.Error as err:
31.         logger.info(err.msg)
32.     else:
33.         self.cnx.commit()

```

Kode 5.8 Potongan Kode untuk menghubungkan TwitterScrapper ke MySQL

Potongan kode diatas merupakan pipeline untuk menyambungkan data yang telah di *crawl* untuk dimasukkan ke dalam MySQL. Kode diatas akan otomatis membuat sebuah

tabel bernama `akun_unverified_orang6` jika belum terdapat tabel dengan nama yang sama.

```

1. def insert_one(self, item):
2.     ret = self.check_vals(item)
3.
4.     if not ret:
5.         return None
6.
7.     ID = item['ID']
8.     user_id = item['user_id']
9.     url = item['url']
10.    text = item['text']
11.    username = item['usernameTweet']
12.    datetime = item['datetime']
13.
14.    insert_query = 'INSERT INTO ' + self.table
15.    _name + ' (ID, url, datetime, text, user_id, userna
16.    meTweet )'
17.    insert_query += ' VALUES ( %s, %s, %s, %s,
18.    %s, %s)'
19.
20.    try:
21.        self.cursor.execute(insert_query, (
22.            ID,
23.            url,
24.            datetime,
25.            text,
26.            user_id,
27.            username
28.        ))
29.    except mysql.connector.Error as err:
30.        logger.info(err.msg)
31.    else:
32.        self.cnx.commit()

```

Kode 5.9 Potongan Kode untuk Memasukkan Hasil *Crawl* ke MySQL

Potongan kode diatas adalah kode untuk meng-insert hasil *scraping* ke dalam MySQL. Pada tabel MySQL, akan terdapat 6 kolom yaitu ID, url, datetime, text, user_id, dan username.

ID	url	datetime	text	user_id	username/Tweet
11160654	/TheArieA	2019-04-11 02:48:37	Meskipun kampret kuar2 infrastruktur tak bisa dimakan.	16460428	TheArieA/R
11177567	/IreneVier	2019-04-15 18:49:15	Setiap sen dolar utang LN yg digunakan tdk untuk program atau proyek terkait peningkatan ekspor,	80562916	IreneViena
11029472	/bayprio/a	2019-03-05 22:01:25	@musniumar @fahiraIdris @AntiMiras_ID	112584871	bayprio
11129165	/Nadiku18	2019-04-02 10:16:02	Waktu itu @jokowi minta @prabowo buktikan tuduhan bocor itu benar.	25878457	Nadiku18
1066452	/Nadiku18	2018-11-25 05:04:00	Di berita tak ucapan orang Perindo itu soal gaji tapi pendapatan ojol. Redaksi @detikcom sengaja p	25878457	Nadiku18
10548834	/Nadiku18	2018-10-24 05:35:35	Kemarin, orang yang main ke tempat kerja saya itu, pendapatan per bulannya lebih dari seratus juta.	25878457	Nadiku18
11160654	/TheArieA	2019-04-11 02:48:37	Meskipun kampret kuar2 infrastruktur tak bisa dimakan.	16460428	TheArieA/R
11114004	/TheArieA	2019-03-29 05:51:25	Tahukah anda transaksi saham setiap harinya? Rerata 6 sd 9 triliun per harinya.	16460428	TheArieA/R
11097665	/TheArieA	2019-03-24 17:38:52	orang naik MRT biasanya tidak langsung dari rumah keluar rumah langsung naik MRT	16460428	TheArieA/R
11097531	/TheArieA	2019-03-24 16:45:50	kalian pikir, jika negara bikin waduk, rigasi dan pinjam dana dari bank dunia, tidak ditung...	16460428	TheArieA/R
11093941	/TheArieA	2019-03-23 16:59:16	Pendidikan: rerata lama sekolah, jml buta huruf, rerata lulusan, Jumlah guru dst.	16460428	TheArieA/R
1108560C	/TheArieA	2019-03-21 09:44:53	Debat Pilpres 2014, JKW ingin meningkatkan produksi beras dan pendapatan petani dengan memba	16460428	TheArieA/R
10771179	/TheArieA	2019-03-19 01:34:34	Prioritas penoenuaan dana desa untuk prooram dan keoliatan bidano oembervayaan.masvarakat.de.	16460428	TheArieA/R

Gambar 5.3 Tabel Hasil *Crawl* pada MySQL

Kode Program TweetScrapper penulis dijalankan melalui terminal dengan potongan kode berikut:

```
1. (base) Faiqs-MacBook-Pro:TweetScrapper-
unverified faiqurnomo$ file="username.txt";while I
FS= read line;do scrapy crawl TweetScrapper -
a query="from:$line since:2018-09-23 until:2019-04-
13";done<"$file"
```

Kode 5.10 Command Line untuk Iterasi tiap Username Perintah TweetScrapper

Penulis menggunakan sebuah file txt yang berisikan nama-nama username dari hasil crawling username untuk meng-loop perintah `do scrapy crawl TweetScrapper -a query="from:$line since:2018-09-23 until:2019-04-13"`, dimana query yang dipakai pada TwitterSearch adalah `"from:$line since:2018-09-23 until:2019-04-13"`. Variabel `$line` adalah tiap username yang berasal dari file `username.txt`, dan `from:` merupakan query TwitterSearch untuk mengambil tweet yang dikirim dari, `since:` adalah batasan waktu kapan tweet tersebut dikirim, dalam hal ini penulis mengisinya dengan tanggal kampanye dimulai, dan `until:` adalah tanggal masa kampanye terakhir.

5.3 Pembuatan *Filtering Dataset*

5.3.1 Pembuatan *Filtering* berdasarkan *Keyword*

Filtering dataset digunakan untuk menyaring data-data tweet yang tidak berkaitan dengan Gatra Ekonomi agar tidak masuk ke dalam pengolahan selanjutnya. Selain tidak ada gunanya

untuk diolah, data- data tweet yang tidak berkaitan dengan Gatra Ekonomi akan memperbanyak data sehingga lebih memperlambat proses kedepannya. Oleh karena itu, Filtering Dataset untuk mendapatkan data- data yang berkaitan dengan Gatra Ekonomi diperlukan. Filtering dilakukan menggunakan query SQL sederhana dengan memanfaatkan Regular Expression. Query dibawah merupakan implementasi dari 4.2.1.

```
1. INSERT INTO akun_siap_filtered
2. SELECT * FROM akun_siap_full
3. WHERE text REGEXP "[[:<:]] (keyword Gatra Ekonomi)
   [[:>:]]"
```

Kode 5.11 Query SQL untuk *Filtering Dataset* berdasarkan *Keyword Gatra Ekonomi*

5.3.2 Penghapusan Data Duplikat

Setelah filtering dataset berdasarkan kata kunci dari Gatra Ekonomi, perlu dilakukan penghapusan data yang duplikat dengan memanfaatkan fitur *unique ID* pada MySQL menggunakan query sederhana yang merupakan implementasi dari 4.2.2.

```
1. alter ignore table akun_verified_full_filtered_1 ad
   d unique(ID)
```

Kode 5.12 Kode Drop Duplikat

5.4 Pembuatan *Preprocessing Dataset*

Potongan kode dibawah ini, semuanya merupakan implementasi dari tahap pembersihan dataset pada sub bab 4.2.3

5.4.1 *Case Folding*

Tahapan ini perlu dilakukan dimana merubah data hasil crawl menjadi huruf kecil semua, dengan tujuan penyamarataan format. Potongan kode yang dipakai adalah sebagai berikut.

```
1. lowerdocs = str(i.lower())
```

Kode 5.13 Kode Case Folding

5.4.2 Data Cleaning

Data dari hasil crawling mengandung banyak karakter yang tidak berguna dan dapat mengganggu hasil ekstraksi fitur. Oleh karena itu, dilakukan berbagai pembersihan data dengan cara menghilangkan berbagai tanda baca dan simbol-simbol maupun emoji. (Dihapus simbol2 dan tanda bacanya)

```
1. hapusws = lowerdocs.rstrip()
2. hapuslink1 = re.sub(r'pic\S+', '', hapusws)
3. hapuslink0 = re.sub(r'http\S.+', '', hapuslink1)
4. hapuslink = re.sub(r'\w+:\/{2}[\d\w-]+(\.[\d\w-
]+)*(?:\?:(?:\?[\^s/]*))*', '', hapuslink0)
5. hapushashtag = re.sub(r"#(\w+)", '', hapuslink)
6. hapusline = re.sub("\n|\r", '', hapushashtag)
7. hapususername = re.sub('@[\^s]+', '', hapusline)
8. hapussymbol = re.sub('[!@' '$, #&_*~.-:?"()%;/\'"
]', ' ', hapususername)
```

Kode 5.14 Kode Cleaning

5.4.3 Tokenizing

Tokenizing merupakan tahapan untuk memecah teks menjadi token-token kata. *Tokenizing* memisahkan teks berdasarkan spasi dan menjadikan kata-kata tersebut menjadi sebuah korpus untuk tahapan selanjutnya. Untuk melakukan tahapan *tokenizing*, penulis menggunakan library `nltk.tokenize` dengan `RegexTokenizer`.

```
1. tokens = tokenizer.tokenize(hapusemoji)
```

Kode 5.15 Kode Tokenizing

5.4.4 Stopwords Removal

Stopwords adalah kata-kata umum yang sering muncul dalam jumlah yang banyak dan dianggap tidak memiliki arti, seperti: tentang, dalam, saat, sama, dan masih banyak lagi. Jika sebuah kata masuk kedalam kata *stopwords*, maka kata tersebut harus dihilangkan. Untuk melakukan proses ini, cukup menggunakan kode program sederhana dengan menggunakan list *stopwords* Bahasa Indonesia yang dapat diunduh pada <https://github.com/stopwords-iso/stopwords-iso>. Ada beberapa penyesuaian *stopwords*, yaitu berupa penambahan beberapa kata dan bentuk kata yang berbeda, antara lain:

adlh	jangan
janganlah	jgn
jadi	jd
sdg	sdgkn
dr	dln

Untuk *stopwords* yang telah diperbarui selengkapnya dapat dilihat pada <https://github.com/faiqurnomo/StopWords-Indonesia>.

```

1. list_stopword = []
2.     with open('stopwords.txt') as stopwords:
3.         for line in stopwords:
4.             list_stopword.append(line.strip())
5.
6. stopped_tokens = [i for i in tokenWithoutInt if not
   i in list_stopword]
```

Kode 5.16 Kode *Stopwords Removal*

5.4.5 Part-Of-Speech Tagging

Setelah data telah berbentuk token- token, tahapan selanjutnya, token- token kata tersebut akan diberikan label berdasarkan jenis kata apakah kata tersebut. Untuk melakukan hal tersebut digunakan sebuah *library* yaitu CRFTagger, dengan memanfaatkan *tagger* berbahasa Indonesia. Kode 5.17 merupakan implementasi dari perancangan 4.2.4.

```
1. from nltk.tag import CRFTagger
2.
3. taggedlist_onegram = ct.tag_sents([hapusstring])
```

Kode 5.17 Kode POS Tagger

5.5 Pembuatan Ontologi

Pembuatan ontologi dilakukan menggunakan perangkat lunak yang dibuat oleh Stanford University, yaitu protégé. Protégé adalah platform *open source* yang menyediakan alat untuk membangun model domain dan aplikasi *knowledge-based* dengan ontologi. Sebelum membangunnya menggunakan aplikasi, diperlukan data- data yang dibutuhkan untuk melengkapi ontologi, *subclass* dan *entity* apa saja yang akan dimasukkan. Dalam hal ini, *subclass* adalah fitur yang didapatkan dari data yang telah diolah, dan *entity* berisi kata-kata sentiment berupa kata adverbial dan kata adjektiva yang telah didapat dari proses *POS Tagging*. Ketiga sub bab dibawah ini merupakan implementasi dari perancangan 4.3.

5.5.1 Pencarian Fitur dengan Frekuensi Tertinggi

Untuk menemukan fitur untuk dimasukkan ke dalam ontologi, diperlukan untuk mengetahui fitur apa saja yang paling sering dibicarakan dan yang berkaitan dengan ekonomi. Untuk menemukan fitur yang akan dimasukkan ke dalam ontologi yang akan dibuat, penulis menggunakan *bigram* pada hasil dari

preprocess dan menggunakan FreqDist untuk mengetahui fitur yang paling sering muncul.

```
1. from nltk import ngrams
2.
3. bigrams = list(ngrams(hapusstring,2))
4.
5. with open('bigrams_verified.csv', 'w') as f:
6.     writer = csv.writer(f) #lineterminator='\n'
7.
8.     for row in bigrams:
9.         writer.writerow([row])
10. f.close()
```

Kode 5.18 Kode *Bigram*

Potongan kode diatas digunakan untuk merubah token yang berupa 1-gram menjadi bigram, lalu hasilnya akan disimpan dalam bentuk csv. Selanjutnya, bigram akan dihitung frekuensi bigram yang paling sering muncul dengan menggunakan potongan kode dibawah.

```

1. import ast
2. import csv
3. ts_unverified = []
4. ts_verified = []
5.
6. with open('bigrams.csv', 'r') as csvFile:
7.     reader = csv.reader(csvFile)
8.     for row in reader:
9.         testarray = ast.literal_eval(str(row[0])) #
           MENGUBAH ROW MENJADI ARRAY
10.        ts_unverified.append(testarray)
11.
12. from nltk.probability import FreqDist
13. fdist = FreqDist()
14.
15. count=0
16. for i in range(0,len(ts_unverified)):
17.     for word in ts_unverified[i]:
18.         fdist[word]+=1
19.         count +=1
20.
21. print(fdist)
22. fdist.most_common()

```

Kode 5.19 Kode *FreqDist*

Dari potongan kode diatas, akan muncul hasil dari *FreqDist* pada bagan dibawah ini.

```

<FreqDist with 46378 samples and 74048 outcomes>
Out[27]:
[ (('pertumbuhan', 'ekonomi'), 392),
  (('pembangunan', 'infrastruktur'), 222),
  (('tenaga', 'kerja'), 133),
  (('era', 'sby'), 126),
  (('ekonomi', 'indonesia'), 121),
  (('sumber', 'daya'), 108),
  (('jalan', 'tol'), 91),
  (('membangun', 'infrastruktur'), 84),
  (('lapangan', 'kerja'), 82),
  (('era', 'jokowi'), 78),

```

Gambar 5.4 Hasil *FreqDist Bigram*

5.5.2 Ekstraksi Adjektiva

Ekstraksi Adjektiva dan Adverbial dilakukan setelah proses *POS Tagging* telah dilakukan. Menggunakan *for*, *if-else*, dan *append*

sederhana, kata adjektiva dan adverbial dapat diekstrak dengan mudah. Kode 5.5 merupakan implementasi dari perancangan 4.4.

```

1. import ast
2.
3.     ts = []
4.     adj = []
5.     noun = []
6.     with open('onegrams_'+user+'.csv', 'r') as
       csvFile:
7.         reader = csv.reader(csvFile)
8.         for row in reader:
9.             testarray = ast.literal_eva
l(str(row[0])) #MENGUBAH ROW MENJADI ARRAY
10.            ts.append(testarray)
11.            #count = 0
12.            for i in range(0, len(ts)):
13.                print(" ")
14.                print(ts[i][0])
15.                for j in range(0, len(ts[i][0])):
16.                    if ts[i][0][j][1] == 'RB':
17.                        #JIKA NOUN MAKA DI APPEND KE LIST adv
18.                        noun.append(ts[i][0
19.                        ][j][0])
20.                        print(ts[i][0][j][0
21.                        ])
22.                    if ts[i][0][j][1] == 'JJ':
23.                        #JIKA ADJ MAKA DI APPEND KE LIST adj
24.                        adj.append(ts[i][0
25.                        ][j][0])
26.                        print(ts[i][0][j][0
27.                        ])
28.                )
29.            )

```

Kode 5.20 Kode Ekstraksi Adjektiva

Hasil dari potongan kode diatas masih berupa data mentah, sehingga jika ingin mengetahui adjektiva dan adverbial yang paling sering muncul, perlu dilakukan penghitungan frekuensi.

5.5.3 Pembuatan Ontologi menggunakan Protégé

Dalam pembuatan ontologi Gatra Ekonomi, digunakan aplikasi Protégé, sebuah aplikasi *open source* editor ontologi yang dibuat oleh Stanford University. Pembuatan ontologi didasari

oleh hasil ekstraksi *bigram* dan *trigram*, dan juga ekstraksi adjektiva. Hasil dari ekstraksi ketiga hal tersebut difilter sesuai dengan faktor-faktor gatra ekonomi pada buku Ketahanan Nasional, dan adjektiva diambil yang berkaitan dengan istilah-istilah ekonomi.

Hasil *class* fitur dan *class* opini beserta dengan *subclassnya* dapat dilihat pada Tabel 5.3 dibawah ini.

Tabel 5.3 Class Faktor Gatra Ekonomi

No	Nama Class	Nama Subclass
1	Sifat Keterbukaan Perekonomian	1. Pertumbuhan Ekonomi
		2. Ekonomi Indonesia
		3. Perekonomian Indonesia
		4. Pembangunan Ekonomi
		5. Ekonomi RI
2	Struktur Ekonomi	6. Ekonomi Kreatif
		7. Kondisi Ekonomi
		8. Ekonomi Syariah
		9. Ekonomi Masyarakat
		10. Ekonomi Rakyat
		11. Ekonomi Nasional
		12. Ekonomi Kerakyatan
3	Potensi dan Pengelolaan Sumber Daya Alam	13. Lingkungan Hidup
		14. Sumber Daya Alam
		15. Energi Terbarukan
4	Potensi dan Pengelolaan Sumber Daya Manusia	16. Kualitas SDM
		17. Sumber Daya Manusia
		18. Tenaga Kerja

		19. Lapangan Kerja
		20. Angka Kemiskinan
5	Potensi dan Pengelolaan Sumber Dana	21. Dana Desa
		22. Pasar Modal
		23. Investasi Indonesia
		24. Realisasi Investasi
		25. Modal Asing
		26. Investasi Asing
6	Birokrasi dan Sikap Masyarakat	27. Kebijakan Ekonomi
		28. Paket Kebijakan
		29. Reformasi Birokrasi
7	Teknologi	30. Ekonomi Digital
		31. Teknologi Informasi
		32. Revolusi Industri
		33. Teknologi Digital
		34. Perkembangan Teknologi
		35. Perusahaan Teknologi
		36. Mobil Listrik
8	Manajemen	37. Perlindungan Konsumen
		38. Produksi Minyak
		39. Produksi Beras
9	Infrastruktur	40. Pembangunan Infrastruktur
		41. Proyek Infrastruktur
		42. Membangun Infrastruktur
		43. Infrastruktur Jokowi

	44. Infrastruktur Indonesia
	45. Jalan Tol
	46. MRT Jakarta
	47. Sarana Prasarana

Dari hasil ekstraksi fitur, terdapat 47 fitur yang berkaitan dengan ekonomi dengan perincian sebagai berikut:

1. Sifat Keterbukaan Perekonomian terdapat 5 fitur;
2. Struktur Ekonomi terdapat 7 fitur;
3. Potensi dan Pengelolaan Sumber Daya Alam terdapat 3 fitur;
4. Potensi dan Pengelolaan Sumber Daya Manusia terdapat 5 fitur;
5. Potensi dan Pengelolaan Sumber Dana terdapat 7 fitur;
6. Birokrasi dan Sikap Masyarakat terdapat 3 fitur;
7. Teknologi terdapat 7 fitur;
8. Manajemen terdapat 3 fitur;
9. dan Infrastruktur terdapat 8 fitur.

Dari setiap fitur kemudian dicari jumlah adjektiva yang paling banyak, lalu dikumpulkan berdasarkan faktor gatra ekonomi. Dari tahapan pada sub bab 5.5.2, dihasilkan opini-opini yang kemudian dirangkum dalam Tabel 5.4 dibawah ini.

Tabel 5.4 Kata Opini dari Faktor Gatra Ekonomi

<i>Class Opini</i>	<i>Subclass</i>
opini_Birokrasi SikapMasyarakat (Faktor No.6 pada Tabel 5.3)	birokrasi_aman birokrasi_asing birokrasi_bebas birokrasi_bersih birokrasi_gagal birokrasi_kuat birokrasi_optimal birokrasi_strategis

opini_Infrastruktur (Faktor No.9 pada Tabel 5.3)	infrastruktur_asing infrastruktur_cepat infrastruktur_efektif infrastruktur_efisien infrastruktur_gagal infrastruktur_gratis infrastruktur_konsisten infrastruktur_lancar infrastruktur_layak infrastruktur_macet infrastruktur_mahal infrastruktur_maju infrastruktur_murah infrastruktur_nasional infrastruktur_optimal infrastruktur_positif infrastruktur_produkatif infrastruktur_ramai infrastruktur_rendah infrastruktur_resmi infrastruktur_riil infrastruktur_rugi infrastruktur_sesuai infrastruktur_signifikan infrastruktur_strategis infrastruktur_sulit infrastruktur_terbaik infrastruktur_tertinggal
opini_Manajemen (Faktor No.8 pada Tabel 5.3)	Tidak ada opini yang kuat
opini_SifatKeterbukaanPerekonomian (Faktor No.1 pada Tabel 5.3)	sifatperekonomian_bagus sifatperekonomian_baik sifatperekonomian_buruk sifatperekonomian_gagal sifatperekonomian_inklusif sifatperekonomian_kebal

	sifatperekonomian_krisis sifatperekonomian_kuat sifatperekonomian_negatif sifatperekonomian_nyata sifatperekonomian_optimis sifatperekonomian_positif sifatperekonomian_rendah sifatperekonomian_sesuai sifatperekonomian_signifikan sifatperekonomian_stabil sifatperekonomian_sulit sifatperekonomian_terbaik sifatperekonomian_terbesar
opini_StrukturE konomi (Faktor No.2 pada Tabel 5.3)	struktur_ekonomi_asing struktur_ekonomi_bagus struktur_ekonomi_berat struktur_ekonomi_buruk struktur_ekonomi_cepat struktur_ekonomi_kuat struktur_ekonomi_lambat struktur_ekonomi_luas struktur_ekonomi_optimis struktur_ekonomi_partisipatif struktur_ekonomi_positif struktur_ekonomi_sesuai struktur_ekonomi_stabil struktur_ekonomi_sulit struktur_ekonomi_terbesar struktur_ekonomi_utama
opini_SumberD ayaAlam (Faktor No.3 pada Tabel 5.3)	sda_asing sda_canggih sda_cepat sda_efisien sda_kaya sda_krisis sda_kuat sda_luas

	sda_mahal sda_miskin sda_sesuai sda_sukses sda_sulit sda_terbaik
opini_SumberD ayaDana (Faktor No.5 pada Tabel 5.3)	sdd_asing sdd_bagus sdd_deras sdd_efektif sdd_gagal sdd_masif sdd_negatif sdd_optimal sdd_optimis sdd_positif sdd_produkatif sdd_rendah sdd_rugi sdd_sesuai sdd_signifikan sdd_stabil sdd_sukses sdd_terbaik sdd_terbesar sdd_tertinggi
opini_SumberD ayaManusia (Faktor No.4 pada Tabel 5.3)	sdm_aktif sdm_asing sdm_bagus sdm_baik sdm_buruk sdm_efektif sdm_gagal sdm_ilegal sdm_kuat sdm_layak sdm_lemah

	sdm_lokal sdm_luas sdm_maju sdm_nyata sdm_optimis sdm_positif sdm_produkatif sdm_puas sdm_rendah sdm_sesuai sdm_signifikan sdm_stabil sdm_sukses sdm_terampil sdm_terbaik sdm_terendah sdm_tertinggi
opini_Teknologi (Faktor No.7 pada Tabel 5.3)	teknologi_aktif teknologi_asing teknologi_baik teknologi_canggih teknologi_cepat teknologi_efisien teknologi_ilegal teknologi_inklusif teknologi_internal teknologi_khawatir teknologi_kompetitif teknologi_kondusif teknologi_lambat teknologi_layak teknologi_mahal teknologi_masif teknologi_murah teknologi_negatif teknologi_optimal teknologi_optimis

	teknologi_penuh teknologi_pesat teknologi_positif teknologi_produkatif teknologi_ramai teknologi_rugi teknologi_sepakat teknologi_sesuai teknologi_stabil teknologi_terampil teknologi_terbaik teknologi_terbesar teknologi_terkemuka teknologi_utama
--	--

Pada *subclass* opini, terdapat 167 opini dari 9 faktor gatra ekonomi, dengan rata-rata 18 opini per faktor. Adapun rinciannya adalah sebagai berikut:

Tabel 5.5 Rincian jumlah Opini pada setiap kelas

No	Faktor	Jumlah Opini
1	Sifat Keterbukaan Perekonomian	19
2	Struktur Ekonomi	16
3	Potensi dan Pengelolaan Sumber Daya Alam	14
4	Potensi dan Pengelolaan Sumber Daya Manusia	28
5	Potensi dan Pengelolaan Sumber Dana	20
6	Birokrasi dan Sikap Masyarakat	8
7	Teknologi	34

8	Manajemen	0
9	Infrastruktur	28

5.6 Ekstraksi Fitur

5.6.1 Query MySQL dihubungkan dengan OWL

Memfaatkan *library* Owlready pada Python, file OWL dapat dibaca dan dijadikan variabel pada Python. Sehingga hal ini memungkinkan Python mengekstraksi *dataset* dari MySQL menggunakan *keyword* pada file OWL. Kode dibawah merupakan implementasi dari perancangan 4.4.

```

1. from owlready2 import *
2. onto_path.append("/Users/faiqurnomo/Desktop/ontologi")
3. onto = get_ontology("/Users/faiqurnomo/Desktop/ontologi/GatraEkonomiMark2 copy.owl")
4. onto.load()
5.
6. import mysql.connector
7. from mysql.connector import Error
8.
9. connection = mysql.connector.connect(host='localhost',
10.                                     database='tweets',
11.                                     user='root',
12.                                     password='')
13.
14. FaktorGatraEkonomi = onto.FaktorGatraEkonomi
15. Tweets = onto.Tweets
16. Kebijakan_Ekonomi = onto.Kebijakan_Ekonomi
17.
18. import pandas as pd
19. from pandas import DataFrame
20.
21. sql_select_Query = "SELECT * FROM akun_SEMUA_SIAP_FILTERED where text regexp '[:<:;]+'Kebijakan_Ekonomi.comment[0]+'[:>:;]'"
22. cursor = connection.cursor()
23. cursor.execute(sql_select_Query)
24. df = DataFrame(cursor.fetchall())

```

Gambar 5.5 Owlready untuk query MySQL

5.6.2 Pembuatan file excel terstruktur

Untuk memasukkan data berupa *instance* ke dalam ontologi, diperlukan data yang terstruktur sehingga *plugin* cellfie dapat melakukan kerjanya dengan baik.

```

1. from owlready2 import *
2. onto_path.append("/Users/faiqurnomo/Desktop/ontolo
   gi")
3. onto = get_ontology("/Users/faiqurnomo/Desktop/ont
   ologi/GatraEkonomiMark2 copy.owl")
4. onto.load()
5.
6. import mysql.connector
7. from mysql.connector import Error
8.
9. connection = mysql.connector.connect(host='localhos
   t',
10.                                     database='tweets',
11.                                     user='root',
12.                                     password='')
13.
14. FaktorGatraEkonomi = onto.FaktorGatraEkonomi
15. Tweets = onto.Tweets
16. Kebijakan_Ekonomi = onto.Kebijakan_Ekonomi
17.
18. import pandas as pd
19. from pandas import DataFrame
20.
21. sql_select_Query = "SELECT * FROM akun_SEMUA_SIAP_F
   ILTERED where text regexp '[:<:~]" +Kebijakan_Ekono
   mi.comment[0]+"[:>:~]"
22. cursor = connection.cursor()
23. cursor.execute(sql_select_Query)
24. df = DataFrame(cursor.fetchall())

```

Gambar 5.6 Pembuatan Data Terstruktur

5.6.3 Memasukkan *instance* ke dalam Ontologi

Untuk memasukkan *instance* ke dalam Ontologi, digunakan sebuah *plugin* bawaan Protégé, yaitu Cellfie. Cellfie merubah *file excel* yang sudah terstruktur untuk kemudian dimasukkan ke dalam ontologi dengan *syntax* tertentu.

```

{"Collections":[{"sheetName":"Sheet1","startColumn":"A",
"endColumn":"A","startRow":"3","endRow":"+", "comment
":"","rule":"Individual: @A*\n  Types: Tweets\n
Facts: hasFitur @B*,\n\thasOpini @C*","active":true}]}

```

Kode 5.21 Syntax Cellfie

5.7 Penilaian polaritas menggunakan SentiWordNet

Pemberian skor pada opini masing- masing fitur dilakukan menggunakan SentiWordNet, sebuah kamus berisi skor yang menyatakan apakah sebuah kata memiliki sentiment positif, objektif, maupun negatif. Pada penelitian ini, penulis menggunakan dua versi SentiWordNet, yang pertama adalah Barasa, sebuah SentiWordNet berbahasa Indonesia, dan SentiWordNet bahasa Inggris. Penulis menggunakan dua SWN dikarenakan Barasa tidak memiliki glosarium, sehingga diperlukan SWN berbahasa Inggris untuk menemukan makna dari sebuah kata sehingga tidak salah dalam pemilihan skor dari sebuah sentimen. Pada contohnya, semisal ada sebuah opini “bagus”, pada Barasa maupun SentiWordNet, terdapat lebih dari satu kata “bagus”. Oleh karena itu, perlu dicocokkan dengan glosarium sehingga tepat dengan makna yang dimaksudkan. Selengkapnya dapat dilihat pada lampiran.

Tabel 5.6 SentiScore

opini_Infrastruktur	infrastruktur_asing	0,375
	infrastruktur_cepat	0,500
	infrastruktur_efektif	0,5
	infrastruktur_efisien	0,5
	infrastruktur_gagal	0,5
	infrastruktur_gratis	0,5
	infrastruktur_konsisten	0,5
	infrastruktur_lancar	0,5

	infrastruktur_layak	0,5
	infrastruktur_macet	0,125
	infrastruktur_mahal	0,5
	infrastruktur_maju	0,5
	infrastruktur_murah	0
	infrastruktur_optimal	0,75
	infrastruktur_positif	0,75
	infrastruktur_produkatif	0,625
	infrastruktur_ramai	0,375
	infrastruktur_rendah	0,125
	infrastruktur_resmi	0,5
	infrastruktur_riil	0,625
	infrastruktur_rugi	0
	infrastruktur_sesuai	0,5
	infrastruktur_signifikan	0,875
	infrastruktur_strategis	0,5
	infrastruktur_sulit	0,125
	infrastruktur_terbaik	1
	infrastruktur_tertinggal	0

5.8 Pembuatan SWRL-Rule

Pembuatan SWRL Rule dapat dikelompokkan menjadi 4 garis besar; klasifikasi fitur, klasifikasi opini, SentiScore, dan klasifikasi polaritas. Keempat sub bab dibawah ini merupakan implementasi dari perancangan 4.5 dan 4.6.

5.8.1 Rule Klasifikasi Fitur

Rule klasifikasi fitur digunakan untuk mengklasifikasikan sebuah *instance* dari kelas Tweets sehingga *instance* tersebut menjadi *instance* dari sebuah kelas FaktorGatraEkonomi.


```
1. Tweets(?t) ^ hasFitur(?t, ?b) ^ swrlb:containsIgnoreCase(?b, "pertumbuhan ekonomi") -
> Pertumbuhan_Ekonomi(?t)
```

Kode 5.22 Rule Klasifikasi Fitur

Rule diatas bermakna jika *instance* dari *class* Tweets memiliki *string* “pertumbuhan ekonomi” pada *data property* hasFitur, maka *instance* tersebut masuk kedalam *class* Pertumbuhan_Ekonomi.

5.8.2 Rule Klasifikasi Opini

Rule klasifikasi opini digunakan untuk mengklasifikasikan sebuah *instance* dari *class* Tweets sehingga menjadi *instance* dari *class* Opini.

```
1. SifatKeterbukaanPerekonomian(?t) ^ hasOpini(?t, ?b)
  ^ swrlb:containsIgnoreCase(?b, "kuat") -
> sifatperekonomian_kuat(?t)
```

Kode 5.23 Rule Klasifikasi Opini

Rule diatas bermakna jika *instance* dari *class* SifatKeterbukaanPerekonomian memiliki *string* “kuat” pada *data property* hasOpini, maka *instance* tersebut masuk kedalam *class* Opini sifatperekonomian_kuat.

5.8.3 Rule SentiScore

Rule SentiScore digunakan untuk memberikan *data property* ke sebuah *instance*.

```
1. sifatperekonomian_kuat(?t) -
> hasSentiScore(?t, "0.75"^^xsd:double)
```

Kode 5.24 Rule Assign SentiScore

Rule diatas bermakna jika sebuah *instance* adalah anggota dari *class* sifatperekonomian_kuat, maka *instance* tersebut memiliki *data property* hasSentiScore dengan nilai *double* 0.75.

5.8.4 Rule Klasifikasi Polaritas

Rule klasifikasi polaritas digunakan untuk mengklasifikasikan apakah sebuah *instance* memiliki polaritas sangat negatif, negatif, netral, positif, ataupun sangat positif, yang mengacu kepada penelitian yang dibuat oleh Farman Ali, dkk. [18]


```
1. opini_SifatKeterbukaanPerekonomian(?t) ^ hasSentiScore(?t, ?b) ^ swrlb:greaterThanOrEqual(?b, 0.75) -
> Sangat_Positif(?t)
```





Kode 5.25 Rule Klasifikasi Polaritas

Rule diatas bermakna jika sebuah *instance* memiliki *data property* *hasSentiScore* dengan nilai diatas atau sama dengan 0.75, maka *instance* tersebut masuk kedalam *class* *Polaritas Sangat_Positif*.

5.9 Pembuatan *Fuzzy Description Logic*

Fuzzy Description Logic atau *FuzzyDL* diterapkan jika sebuah *instance* memiliki lebih dari satu *data property* *hasSentiScore*. Pada ontologi biasa, jika sebuah *instance* memiliki lebih dari satu *data property* *hasSentiScore*, maka ontologi akan memasukkan *instance* tersebut kedalam dua kelas polaritas yang berbeda. Dengan menggunakan *FuzzyDL*, jika sebuah *instance* memiliki dua nilai *SentiScore* yang berbeda, maka konsep tersebut adalah konsep fuzzy, sehingga perlu nilai tersebut perlu difuzifikasi, melalui tahapan *inference*, kemudian didefuzifikasi sehingga menghasilkan nilai *crisp* untuk menentukan ke kelas manakah *instance* tersebut seharusnya berada.

Data property assertions 

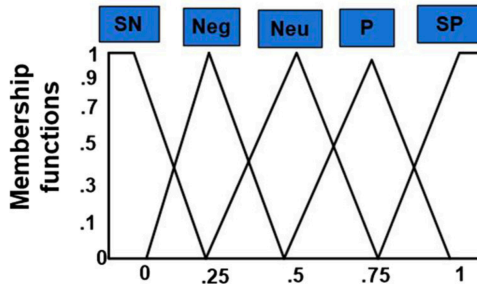
-  *hasSentiScore* "0.125"^^xsd:double
-  *hasSentiScore* "0.5"^^xsd:double
-  *hasFitur* "Pertumbuhan Ekonomi"^^xsd:string
-  *hasOpini* "terbesar, terendah"^^xsd:string

Gambar 5.7 Contoh Tweet memiliki 2 *SentiScore*

Syntax *FuzzyDL* dibawah ini digunakan untuk mendefinisikan nilai dari *Polaritas* dalam bentuk *fuzzy*.

```
(define-fuzzy-concept SangatNegatif left-shoulder(0,
1, 0, 0.25))
(define-fuzzy-concept Negatif triangular(0, 1, 0,
0.25, 0.5))
(define-fuzzy-concept Netral triangular(0, 1, 0.25,
0.5, 0.75))
(define-fuzzy-concept Positif triangular(0, 1, 0.5,
0.75, 1))
(define-fuzzy-concept SangatPositif right-shoulder(0,
1, 0.75, 1))
```

Kode 5.26 Syntax FuzzyDL Konsep Polaritas



Gambar 5.8 Diagram Fuzzy untuk Analisis Sentimen

Syntax FuzzyDL dibawah ini digunakan untuk mendefinisikan *data properties* `hasScoreAkhir`, dan `hasPolaritas1-4`. *Object properties* `hasScoreAkhir` memiliki *data type* `real` dengan batas 0-100, begitu pula dengan *data property* `hasPolaritas1-4`.

```
(range hasScoreAkhir *real* 0 100)
(range hasPolaritas1 *real* 0 100)
(range hasPolaritas2 *real* 0 100)
(range hasPolaritas3 *real* 0 100)
(range hasPolaritas4 *real* 0 100)
```

Kode 5.27 Syntax Data Property

Syntax dibawah digunakan untuk mendefinisikan konsep untuk proses defuzifikasi dengan menggunakan *weighted concept*, dimana setiap variabel memiliki beban yang dibagi rata.

```
(define-concept DoubleFuzzy (= hasScoreAkhir
((0.5*hasPolaritas1) + (0.5*hasPolaritas2))))
(define-concept TripleFuzzy (= hasScoreAkhir
((0.333*hasPolaritas1) + (0.334*hasPolaritas2) +
(0.333*hasPolaritas3))))
(define-concept QuadFuzzy (= hasScoreAkhir
((0.25*hasPolaritas1) + (0.25*hasPolaritas2) +
(0.25*hasPolaritas3) +(0.25*hasPolaritas4))))
```

Kode 5.28 Syntax Weighted Concept

Pada FuzzyDL juga dapat dilakukan pembuatan *instance* yang diinginkan, dengan meng-*assign* nilai polaritas, dan juga konsep defuzifikasi yang ingin digunakan ke *instance* tersebut sehingga dapat dihitung nilai *crisp* dari proses defuzifikasi dengan menggunakan *syntax* defuzzify-som?.

```
(instance OP (and DoubleFuzzy (some
hasPolaritas1 Netral) (some hasPolaritas2
Positif)))

(instance N-O-P (and TripleFuzzy (some
hasPolaritas1 Negatif) (some hasPolaritas2
Netral) (some hasPolaritas3 Positif)))

(instance N-O-P-SP (and QuadFuzzy (some
hasPolaritas1 Negatif) (some hasPolaritas2
Netral) (some hasPolaritas3 Positif) (some
hasPolaritas4 SangatPositif)))
```

Kode 5.29 Syntax Pembuatan Instance

```
(defuzzify-som? DoubleFuzzy OP hasScoreAkhir)
% 0.625
(defuzzify-som? TripleFuzzy N-O-P
hasScoreAkhir)
% 0.5
(defuzzify-som? QuadFuzzy N-O-P-SP
hasScoreAkhir)
% 0.625
```

Kode 5.30 Syntax Defuzifikasi

Hasil akhir yang didapatkan adalah sebuah nilai *crisp* dari hasil defuzifikasi. Pada hasil dari (defuzzify-som? QuadFuzzy N-O-P-SP hasScoreAkhir), didapatkan nilai 0.625, sehingga masuk ke dalam polaritas Positif.

5.10 Implementasi Owlready

Owlready merupakan sebuah *library* pada Python untuk membuka file OWL melalui Python. Pada penelitian ini, Owlready digunakan untuk mengetahui jumlah dari setiap kelas polaritas setelah SWRL telah dijalankan. Karena masih pada tahap pengembangan, Owlready masih mempunyai beberapa kekurangan seperti saat ini masih belum bisa menjalankan *reasoner*. Oleh karena itu, penulis memanfaatkan Owlready hanya untuk sebatas menggunakan memudahkan visualisasi.

```
1. from owlready2 import *
2. onto_path.append("path ontolgi")
3. onto = get_ontology("/Users/faiqurnomo/Desktop/ont
   ologi/INOIPML.owl")
4. onto.load()
5. namespace = onto.get_namespace("http://www.semantic
   web.org/faiqurnomo/ontologies/2019/5/untitled-
   ontology-68#")
```

Kode 5.31 Owlready Memuat Ontologi

Potongan kode diatas digunakan untuk memuat file ontology kedalam Python.

```

1. Tweets = onto.Tweets
2. SangatNegatif = onto.Sangat_Negatif
3. Negatif = onto.Negatif
4. Netral = onto.Netral
5. Positif = onto.Positif
6. SangatPositif = onto.Sangat_Positif
7. Polaritas = onto.Polaritas
8.
9. SN = SangatNegatif.instances()
10. N = Negatif.instances()
11. O = Netral.instances()
12. P = Positif.instances()
13. SP = SangatPositif.instances()
14. iPolaritas = Polaritas.instances()

```

Kode 5.32 Memuat Kelas ke Variabel

Potongan kode diatas digunakan untuk memanggil kelas dari ontologi dan memasukkannya kedalam variabel.

```

1. pSN = len(SN)
2. pN = len(N)
3. pO = len(O)
4. pP = len(P)
5. pSP = len(SP)
6.
7. print("Opini Sangat Negatif berjumlah: "+str(pSN))
8. print("Opini Negatif berjumlah: "+str(pN))
9. print("Opini Netral berjumlah: "+str(pO))
10. print("Opini Positif berjumlah: "+str(pP))
11. print("Opini Sangat Positif berjumlah: "+str(pSP))

```

Kode 5.33 Print Hasil

Potongan kode diatas digunakan untuk mengambil jumlah *instance* dari tiap kelas untuk diketahui berapa jumlah dari tiap-tiap kelas polaritas.

5.11 Pengujian

Pengujian dilakukan dengan melakukan perbandingan secara manual terhadap data yang berhasil di ekstrak melalui fitur-

fitur. Perbandingan dilakukan dengan melihat perbandingan antara nilai *Precision (P)*, *Recall (R)*, *Accuracy (Ac)*, dan *Function Measure (FM)*. Contoh hasil implementasi dapat dilihat pada tabel berikut.

BAB VI HASIL DAN PEMBAHASAN

6.1. Hasil Data *Crawling*

6.1.1. Hasil *Crawling Hashtag*

Dengan menggunakan 50 akun umpan, dihasilkan 219.854 *tweet* kotor dan setelah diolah terdapat 75.248 *tweet* yang mengandung *hashtag*. Dari 75.248 *tweet* terdapat 11.313 jenis *hashtag*, dan diambil 20 *hashtag* teratas yang berkaitan dengan kampanye pemilihan presiden yang dapat dilihat dibawah ini.

Tabel 6.1 Hashtag Umpan

No	Hashtag	Frekuensi
1	Pdiperjuangan	2813
2	Jokowilagi	2488
3	0lindonesiamaju	2442
4	debatpilpres2019	2262
5	Perindouupdate	1899
6	Nasdemantimahar	1690
7	Politiktanpamahar	1462
8	Partaiperindomenang	1327
9	pemilu2019	826
10	pilpres2019	661
11	2019jokowimarufamin	645
12	9perindo	498

13	partaiperindo	403
14	debatkeempatpilpres2019	393
15	debatcawapres2019	392
16	debatcapres2019	385
17	2019gantipresiden	358
18	2019pilihpk	325
19	debatpilpresliputan6	297
20	17aprilpilihnsdem	296

6.1.2. Hasil Pengumpulan *Username*

Pengumpulan *username* untuk diambil datanya dilakukan secara manual dan menggunakan *crawler*. Total *username* yang terkumpul adalah 425 *username* yang terkumpul dari tokoh masyarakat, media, jurnalis, organisasi, dan masyarakat umum. Rincian dapat dilihat pada bagan dibawah dan pada poin-poin dibawah ini.

Tabel 6.2 Hasil Pengumpulan *Username*

Tokoh	163
Masyarakat	
Media	80
Jurnalis	20
Organisasi	53
Masyarakat Umum	109

Total	425
-------	-----

a. Pengumpulan Manual Tokoh Masyarakat

Pengumpulan *username* tokoh masyarakat secara manual dilakukan penulis sebatas tokoh-tokoh yang secara jabatan politis, antara lain:

1. Pemerintahan (Presiden/ Wakil Presiden, Menteri, Pejabat setingkat Menteri, Gubernur/ Wakil Gubernur, dan Walikota/Wakil Walikota);
2. Calon Presiden dan Calon Wakil Presiden;
3. Anggota MPR, DPR, dan DPD;
4. Tokoh Partai Politik (Ketua, Sekretaris Jenderal, dan Ketua Fraksi).

Penulis mengambil dari berbagai sumber mengenai nama- nama dan mencari *username* pada Twitter secara manual. Hasil pengumpulan *username* dapat dilihat pada tabel dibawah ini.

Tabel 6.3 Hasil Pengumpulan *Username* - Manual

Pemerintahan	Presiden/Wapres	2
	Menteri	27
	Pejabat setingkat Menteri	5
	Gubernur/Wagub	11
	Walikota/Wakil Walikota	5
Capres dan Cawapres		3 (sudah termasuk di presiden)
Anggota MPR, DPR, dan DPD		7
Tokoh Parpol		20

Total	80
--------------	----

Dikarenakan adanya petahana pada pemilu kali ini, oleh karena itu akan ada duplikasi *username* sehingga diperlukan penghapusan duplikat.

b. Crawling Username

1) Username Verified

Setelah dilakukan crawling dengan menggunakan keyword yakni hashtag diatas, didapatkan 572.068 *tweets* dari 128.020 akun. Dari 128.020 *username*, terdapat 242 *username verified*, dan ketika disaring lagi 42 diantaranya merupakan *username* asing, sehingga terdapat 199 akun *verified* yang memenuhi syarat. Dari 199 akun *verified* ini, terdapat tokoh masyarakat, media, jurnalis, organisasi profit maupun non-profit, serta masyarakat umum.

Tabel 6.4 Hasil Pengumpulan Username Verified

Tokoh Masyarakat	44
Media	80
Jurnalis	17
Organisasi	53
Masyarakat Umum	5
Total	199

Pada *username verified* yang telah didapatkan, masih terdapat data yang tumpang tindih dengan data yang dikumpulkan pada poin sebelumnya yaitu secara manual, oleh karena itu harus dilakukan penghapusan duplikasi.

2) Username Unverified

Dari 128.020 *username*, terdapat 127.778 *username unverified*. Dari 127.778 dilakukan penyaringan dengan mengambil *username* dengan jumlah *followers* lebih dari 40.000, sehingga didapat 266 *username*. Dari 266 *username* dilakukan penyaringan lagi secara sehingga didapat 147 *username* perorangan yang terdiri dari tokoh masyarakat, jurnalis, dan masyarakat umum.

Tabel 6.5 Hasil Pengumpulan *Username Unverified*

Tokoh Masyarakat	39
Jurnalis	3
Masyarakat Umum	104
Total	146

6.1.3 Hasil Pengumpulan *Tweets*

Pengumpulan *tweets* dari tiap *username* yang telah terkumpul dilakukan menggunakan TweetScraper, sebuah aplikasi yang memanfaatkan *library* Scrapy. Dari total 425 *username* yang terkumpul, dilakukan *crawling* dari setiap tweet yang pernah di-*post* oleh setiap *username* semenjak tanggal 28 September 2018 hingga 5 April 2019. Dari keseluruhan proses *crawling*, terkumpul 2.322.075 *tweets* mentah dari 425 *username*. Dari 2.322.075 *tweets*, masih perlu diolah lagi sehingga hanya *tweets* yang berhubungan dengan ekonomi saja yang masuk ke pengolahan.

6.2 Hasil *Preprocessing Dataset*

6.2.3 Hasil Penghapusan Data Duplikat

Setelah dilakukan penghapusan data duplikat, dari 2.322.075 *tweets*, dihasilkan 2.213.171 *tweets*.

6.2.4 Hasil *Filtering Keyword* Gatra Ekonomi

Setelah dilakukan penghapusan data duplikat, dari 2.213.171 *tweets*, didapatkan 86.699 *tweets* yang berkaitan dengan gatra ekonomi dengan menggunakan kata kunci yang diambil dari buku Ketahanan Nasional. Setelah dilakukan *filtering*, perlu dilakukan penghapusan duplikat kembali untuk mengurangi hasil filter yang tumpang tindih sehingga dihasilkan 78.889 *tweets*.

Tabel 6.6 Hasil Filtering Gatra Ekonomi

No	Faktor	Keyword	Jumlah
1	Sifat Keterbukaan Perekonomian	Ekonomi	31.928
		Sistem Perekonomian	12
		Sistem Ekonomi	68
2	Struktur Ekonomi	Sektor Ekonomi	124
		Struktur Ekonomi	23
		Produk Domestik Bruto	39
		PDB	997
3	Potensi dan Pengelolaan Sumber Daya Alam	Sumber Daya Alam	464
		SDA	462
		Pertanahan	145
		Energi	4327
4	Potensi dan Pengelolaan Sumber Daya Manusia	Sumber Daya Manusia	299
		SDM	1644
		Tenaga Kerja	1285
		Kesempatan Kerja	58
		Kemiskinan	1556
		Tingkat Pendapatan	1
5	Potensi dan Pengelolaan Sumber Daya Dana	Sumber dana	118
		Sektor Pengerahan Dana	0
		Pasar Uang	59

		Pasar Modal	665
		Investasi	14331
		Penanaman Modal	113
		Modal Dalam Negeri	7
		Modal Asing	310
		Modal	4201
6	Birokrasi dan Sikap Masyarakat	Sistem Birokrasi	10
		Birokrasi	680
7	Teknologi	Teknologi	6564
8	Manajemen	Manajemen	1884
		Produktivitas	439
		Mutu Produksi	0
		Produksi	5131
		Mutu Jasa	0
9	Infrastruktur	Infrastruktur	7316
		Prasarana	203
		Sarana	1208
		Arus Bahan	0
		Arus Barang	28
		Arus Jasa	0
Total			86.699
Total akhir setelah penghapusan duplikasi			75.101

6.3 Hasil Pembuatan Ontologi



Gambar 6.1 Hasil Ontologi

6.4 Hasil Ekstraksi Fitur

Setelah dilakukan pembuatan ontologi, fitur-fitur pada ontologi digunakan untuk mengekstraksi fitur dari *dataset* hasil filter dari tahapan filtrasi hasil *crawl*. Dari 75.101 tweet, terekstrak 15.856 tweet sesuai dengan fitur pada ontologi. Rincian jumlah setiap fitur dapat dilihat pada tabel dibawah ini.

Tabel 6.7 Hasil Ekstraksi Fitur

No	Faktor	Fitur	Jumlah Tweet yang terekstraksi
1	Sifat Keterbukaan Perekonomian	Pertumbuhan Ekonomi	2221
		Ekonomi Indonesia	1022
		Perekonomian Indonesia	284
		Pembangunan Ekonomi	141
		Ekonomi RI	477
2	Struktur Ekonomi	Ekonomi Kreatif	527
		Kondisi Ekonomi	267
		Ekonomi Syariah	242
		Ekonomi Masyarakat	162
		Ekonomi Rakyat	150
		Ekonomi Nasional	162
		Ekonomi Kerakyatan	130
3	Potensi dan Pengelolaan Sumber Daya Alam	Lingkungan Hidup	269
		Sumber Daya Alam	395
		Energi Terbarukan	286
4	Potensi dan Pengelolaan Sumber Daya Manusia	Kualitas SDM	235
		Sumber Daya Manusia	277
		Tenaga Kerja	1192
		Lapangan Kerja	335

		Angka Kemiskinan	345
5	Potensi dan Pengelolaan Sumber Daya Dana	Dana Desa	291
		Pasar Modal	616
		Investasi Indonesia	36
		Realisasi Investasi	132
		Modal Asing	280
		Investasi Asing	180
6	Birokrasi dan Sikap Masyarakat	Kebijakan Ekonomi	251
		Paket Kebijakan	159
		Reformasi Birokrasi	164
7	Teknologi	Ekonomi Digital	569
		Teknologi Informasi	230
		Revolusi Industri	190
		Teknologi Digital	189
		Perkembangan Teknologi	184
		Perusahaan Teknologi	149
		Mobil Listrik	133
8	Manajemen	Perlindungan Konsumen	14
		Produksi Minyak	149
		Produksi Beras	143
9	Infrastruktur	Pembangunan Infrastruktur	1183
		Proyek Infrastruktur	538
		Membangun Infrastruktur	292
		Infrastruktur Jokowi	111
		Infrastruktur Indonesia	60
		Jalan Tol	341
		MRT Jakarta	119

	Sarana Prasarana	34
Total		15.856

6.5 Hasil Analisis Sentimen dan Perhitungan Polaritas

Setelah SWRL-Rule dijalankan, *instance* Tweet telah masuk kedalam kelas-kelas yang telah dibuat, dimulai dari pengklasifikasian Faktor, hingga polaritas. Dengan bantuan *library* Owlready pada Python, dapat dihitung jumlah dari tiap *instance* pada sebuah kelas, sehingga dihasilkan tabel hasil analisis sentimen dibawah ini.

Tabel 6.8 Jumlah *Instance* pada Kelas Sentimen

Class	Jumlah Instance	Jumlah Instance SentiScore = 1	Gap
Polaritas	2032	1827	205
Sangat Negatif	331	263	68
Negatif	405	339	66
Netral	694	591	103
Positif	254	182	72
Sangat Positif	527	452	75

Dengan ontologi biasa, jika ada sebuah *instance* yang memiliki lebih dari satu SentiScore, maka *instance* tersebut akan masuk ke dalam dua kelas yang berbeda. Untuk mengatasi masalah ini, digunakan konsep *fuzzy* dengan memanfaatkan FuzzyDL. Dari 2032 *tweet* yang memiliki sentimen, terdapat 205 *tweet* yang memiliki lebih dari satu SentiScore. Dua ratus lima *tweet* tersebut kemudian diekstrak, dan dicari polanya sehingga didapatkan rincian pada tabel dibawah ini.

Tabel 6.9 Pola *SentiScore*

Pola	Jumlah Instance	Pola	Jumlah Instance
------	-----------------	------	-----------------

O P	29	N P SP	4
SP SP	27	P SP SP	3
N O	22	N SP	2
SN O	20	SN SP SP	2
SN SN	18	SN P	2
O SP	15	N O SP	2
N P	14	SN SN N	2
SN SP	11	SN P SP	1
SN N	10	SN O SP	1
N O P	10	N O P SP SP	1
P SP	10	SN SN O	1

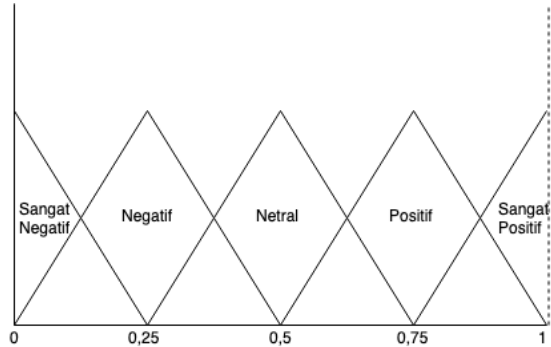
Contoh dari Tweet yang memiliki lebih dari satu *SentiScore* dapat dilihat pada Tabel 6.10 dibawah. Kata yang miring merupakan opini yaitu *terbesar* dan *terendah*, yang memiliki *SentiScore* yaitu 0,5 dan 0,125 secara berurutan, yang artinya memiliki dua nilai sentimen yaitu netral dan sangat negatif. Sehingga *tweet* pada Tabel 6.10 akan masuk kedalam pola SN-N pada Tabel 6.9.

Tabel 6.10 Contoh *Tweet* yang memiliki lebih dari satu *SentiScore*

pertumbuhan ekonomi indonesia perekonomian <i>terbesar</i> asia tenggara melambat mencapai persen angka <i>terendah</i>
--

Tabel 6.11 Contoh *Decision Matrix*

Pola	Sentimen 1	Sentimen 2
O P	Netral	Positif
N O	Negatif	Netral
SN O	Sangat Negatif	Negatif



Gambar 6.2 Fuzzy Concept

Pada *instance* pola O P, pola O P masuk kedalam kategori netral pada sentimen pertama, dan kategori positif pada sentimen kedua. Untuk mengetahui hasil dari nilai akhir campuran antara nilai netral dan positif, dilakukan beberapa langkah:

1. Definisikan atribut yang berhubungan dengan kriteria.
*(range Sentimen1 *real* 0 1)*
*(range Sentimen2 *real* 0 1)*
2. Definisikan atribut untuk nilai akhir.
*(range HasilAkhir *real* 0 1)*
3. Definisikan *concept*.
(define-concept DoubleFuzzy (= HasilAkhir (
*(0.5*Sentimen1) + (0.5*Sentimen2)))*
Syntax diatas memiliki arti bahwa HasilAkhir dari DoubleFuzzy adalah kombinasi linear antara nilai sentimen, dan dibobotkan dengan bobot yang sama berjumlah 1.
4. Deskripsikan masing- masing pola, sesuai dengan *decision matrix*.
(instance O-P (and DoubleFuzzy (some hasPolaritas1
Netral) (some hasPolaritas2 Positif)))

(instance N-O (and DoubleFuzzy (some hasPolaritas1 Negatif) (some hasPolaritas2 Netral)))
(instance SN-O (and DoubleFuzzy (some hasPolaritas1 SangatNegatif) (some hasPolaritas2 Netral)))

5. Pada akhirnya, digunakan proses defuzifikasi untuk melihat hasil nilai skalar.
(defuzzify-som? DoubleFuzzy O-P hasScoreAkhir)
(defuzzify-som? DoubleFuzzy N-O hasScoreAkhir)
(defuzzify-som? DoubleFuzzy SN-O hasScoreAkhir)

Setelah diolah menggunakan langkah diatas, dihasilkan nilai *crisp* dari defuzifikasi pola-pola pada Tabel 6.9, dengan rincian dapat dilihat pada Tabel 6.12.

Tabel 6.12 Pencarian Polaritas

Pola	Nilai Crisp	Polaritas
O P	0.625	Positif
SP SP	1	Sangat Positif
N O	0.375	Negatif
SN O	0.25	Sangat Negatif
SN SN	0	Sangat Negatif
O SP	0.75	Sangat Positif
N P	0.5	Netral
SN SP	0.5	Netral
SN N	0.125	Sangat Negatif
N O P	0.5	Netral
P SP	0.875	Sangat Positif

N P SP	0.6668	Positif
P SP SP	0.9168	Sangat Positif
N SP	0.75	Sangat Positif
SN SP SP	0.667	Positif
SN P	0.375	Negatif
N O SP	0.5833	Positif
SN SN N	0.0833	Sangat Negatif
SN P SP	0.5835	Positif
SN O SP	0.5	Netral
N O P SP SP	0.7	Positif
SN SN O	0.1665	Sangat Negatif

Sehingga, didapatkan hasil akhir analisis sentimen sebagai berikut.

Tabel 6.13 Hasil Akhir Analisis Sentimen

Class	Jumlah
Sangat Negatif	314
Negatif	392
Netral	627
Positif	192
Sangat Positif	509
Total	2032

Didapat jumlah dari tweet yang memiliki sentimen sangat negatif berjumlah 314, sentimen negatif berjumlah 392,

sentiment netral berjumlah 627. Sentimen positif berjumlah 192, dan yang terakhir sentimen sangat positif berjumlah 509.

6.6 Hasil Analisis Pengujian

Setelah dilakukan pengujian, didapatkan hasil seperti pada Tabel dibawah ini:

Tabel 6.14 Pengujian

Accuracy	0,78
Precision	93%
Recall	73%
Function Measure	0,82

Dapat dilihat pada tabel diatas, hasil dari *Accuracy* adalah sebesar 0,78, atau 78% akurat. Nilai *precision* sebesar 93%, dimana hal ini merupakan hal yang bagus. Nilai *accuracy* belum dapat maksimal dikarenakan banyaknya klasifikasi yang salah pada kelas sentimen negatif, sehingga mempengaruhi juga ke nilai *recall* yang kurang maksimal. Contoh klasifikasi negatif yang salah paling banyak terjadi adalah pada *bigram* “tenaga kerja” yang bertemu dengan opini “asing”.

Tabel 6.15 Kasus yang mengurangi nilai akurasi

Tweet
"[['kawasan', 'NN'), ('industri', 'NN'), ('morowali', 'RB'), ('serius', 'JJ'), ('kurangi', 'VB'), ('tenaga', 'NN'), ('kerja', 'NN'), ('asing', 'JJ)]]"
"[['(optimis', 'NN'), ('tenaga', 'NN'), ('kerja', 'NN'), ('asing', 'JJ'), ('menurun', 'VB'), ('contoh', 'NN'), ('sektor', 'NN'), ('wisata', 'NN'), ('bali', 'RB'), ('kerjakan', 'VB'), ('orang', 'NN'), ('asli', 'JJ'), ('bali', 'RB'), ('agus', 'JJ'), ('sari', 'VB)]]"
"[['(menemukan', 'VB'), ('tka', 'NN'), ('tenaga', 'NN'), ('kerja', 'NN'), ('asing', 'JJ'), ('ilegal', 'JJ'), ('detail', 'NN'), ('laporkan', 'VB)]]"

Contoh kasus dapat dilihat pada ketiga tweet ini fitur *bigram* “tenaga kerja” bertemu dengan opini “asing”. Asing memiliki nilai 0.375, sehingga *tweet* tersebut akan selalu bernilai negatif. Sehingga, pada kasus ini, pemilihan antara *bigram* dan *trigram* sangat mempengaruhi hasil dari akurasi.

Halaman ini sengaja dikosongkan

BAB VII KESIMPULAN DAN SARAN

7.1 Kesimpulan

Berdasarkan proses yang telah dilakukan dalam pengerjaan tugas akhir dengan judul “Analisis Sentimen Masyarakat Indonesia Terhadap Gatra Ekonomi Ketahanan Nasional Menggunakan *Fuzzy Ontology-Based Semantic Knowledge*” yang telah dilakukan, dapat disimpulkan sebagai berikut:

1. Penelitian ini membuktikan bahwa metode *crawling* data dari Twitter baik menggunakan Tweepy yang terhubung dengan TwitterAPI, dan menggunakan TweetScraper dapat dengan baik mengambil data dari *username-username* tokoh masyarakat, media, dan masyarakat umum selama masa kampanye Pemilu 2019 berlangsung.
2. Hasil dari pengumpulan tweets, dari total 2.213.171 *tweets* dari 425 *username* selama masa kampanye Pemilu 2019 berlangsung, terdapat 75.101 atau sekitar 0,03% dari total yang berbicara mengenai topik ekonomi.
3. Pada proses pembuatan ontologi, untuk menentukan fitur dari kelas-kelas faktor gatra ekonomi dapat dilakukan dengan melakukan pencarian frekuensi tertinggi pada token-token *bigram* maupun *trigram* yang kemudian diseleksi berdasarkan kaitannya dengan ekonomi. Sedangkan untuk menentukan opini, dapat dilakukan ekstraksi adjektiva dari *dataset* pada fitur-fitur, setelah sebelumnya dilakukan *pos tagging*.
4. Pada tahapan penilaian opini menggunakan SentiWordNet, SentiWordNet Bahasa Indonesia atau lebih dikenal sebagai Barasa memiliki kekurangan yaitu tidak memiliki glosarium sehingga masih harus mengandalkan SentiWordNet asli untuk pemahaman kata sehingga tidak salah memilih nilai untuk sebuah opini.
5. *Rule* SWRL dapat bekerja baik untuk mengklasifikasikan fitur dan polaritas, akan tetapi

FuzzyDL diperlukan untuk memecahkan permasalahan pada *instance* yang memiliki lebih dari satu nilai sentimen.

6. Hasil akurasi kurang maksimal dikarenakan adanya kesalahan yang terlalu banyak pada kelas klasifikasi negatif, sehingga sangat mempengaruhi nilai akurasi.

7.2 Saran

Dalam pengerjaan tugas akhir ini masih terdapat banyak kekurangan yang dapat diperbaiki. Beberapa saran ini diharapkan dapat membantu untuk penelitian kedepannya dapat lebih baik. Saran penulis untuk penelitian selanjutnya sebagai berikut:

1. Pada tahapan *POS Tagging*, perlu dicoba menggunakan *POS Tagger* berbahasa Indonesia lainnya untuk tingkat akurasi yang lebih tinggi.
2. Perlu dilakukan *sharing* dengan orang berlatar belakang ekonomi untuk verifikasi dan validasi model ontologi.
3. Perlu dicari cara lain pada proses pencarian fitur dan opini selain penghitungan frekuensi sehingga ontologi yang dihasilkan dapat lebih *fine-grained* untuk analisis sentimen.
4. Perlu dilakukan usaha penggabungan anotasi *fuzzy* pada ontologi sehingga FuzzyDL hanya digunakan untuk sebatas *reasoner* saja.
5. Pengujian dapat dilakukan oleh lebih dari satu orang untuk mengurangi subjektivitas hasil pengujian.

DAFTAR PUSTAKA

- [1] BBC, “Pesta sudah dimulai: Yang perlu Anda ketahui soal pemilu 2019,” BBC, 24 September 2018. [Online]. Available: [//www.bbc.com/indonesia/indonesia-45618212#orb-banner](http://www.bbc.com/indonesia/indonesia-45618212#orb-banner). [Diakses 1 Februari 2019].
- [2] M. H. Susanti, Setiajid dan N. W. Wardhani, “Kampanye pemilu 2019 dan Potensi Ancaman Disintegrasi Bangsa,” dalam *SIDING SEMINAR NASIONAL JURUSAN POLITIK DAN ARGANEGARAAN*, Semarang, 2018.
- [3] S. Kemp, “Digital in 2018: World’s internet users pass the 4 billion mark,” We Are Social Ltd, 30 January 2018. [Online]. Available: <https://wearesocial.com/blog/2018/01/global-digital-t-2018>. [Diakses 25 January 2019].
- [4] katadata, “Indonesia Pengguna Twitter Terbesar Ketiga di Dunia,” Katadata, 22 November 2016. [Online]. Available: [//databoks.katadata.co.id/datapublish/2016/11/22/indonesia-pengguna-twitter-terbesar-ketiga-di-dunia](http://databoks.katadata.co.id/datapublish/2016/11/22/indonesia-pengguna-twitter-terbesar-ketiga-di-dunia). [Diakses 25 January 2019].
- [5] V. Chang, “A proposed social network analysis platform for data analysis,” *Technological Forecasting & Social Change*, 30, pp. 57-58, 2018.
- [6] Lembaga Ketahanan Nasional, Ketahanan Nasional, Jakarta: Pustaka, 1997.
- [7] Tim Kampanye, “Visi, Misi, Program Aksi,” 2018. [Online]. Available: <https://jokowiamin.id/visi-misi/>. [Diakses 25 January 2019].
- [8] Tim Kampanye, “Misi,” 25 October 2018. [Online]. Available: <https://prabowo-sandi.com/misi/>. [Diakses 25 January 2019].
- [9] C. Orellana-Rodriguez dan M. T. Keane, “Attention to news dissemination on Twitter: A survey,” *Computer Science w*, vol. 29, pp. 74-79, 2018.

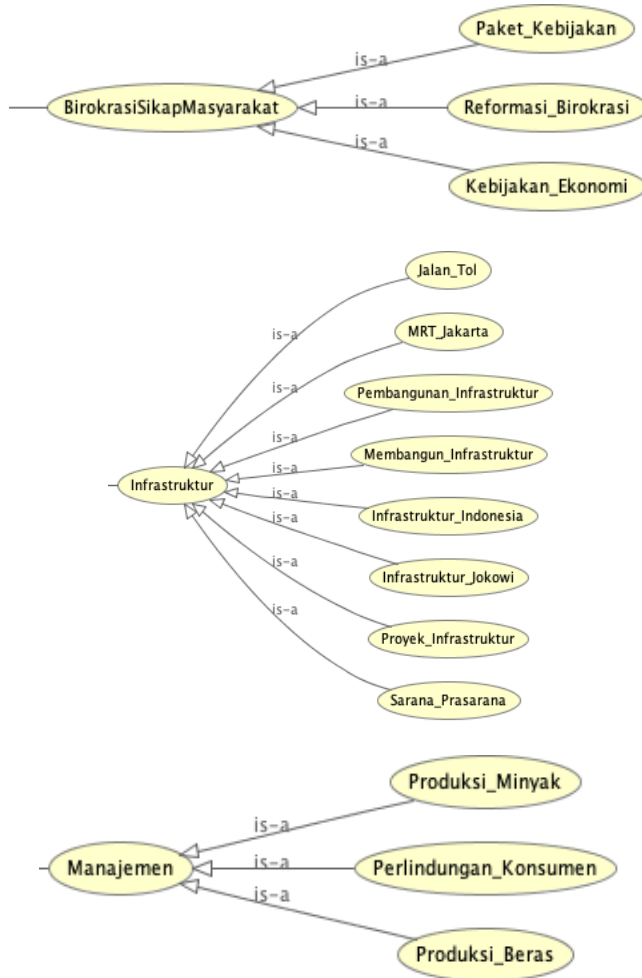
- [10] M. Dewing, *Social Media: An Introduction*, Ottawa: Library Parliament, 2012.
- [11] M. Najork, *Web Crawler Architecture: Encyclopedia of Database System*, Boston: Springer, 2009.
- [12] J. E. Sembodo, E. B. Setiawan dan A. Baizal, "Data Mining Otomatis pada Twitter," dalam *Indonesia Symposium on Computing*, Bandung, 2016.
- [13] Twitter, "About Twitter's APIs," [Online]. Available: [//help.twitter.com/en/rules-and-policies/twitter-api](https://help.twitter.com/en/rules-and-policies/twitter-api). (Accessed 20 January 2019).
- [14] M. Taboada, J. Brooke, M. Tofiloski, K. Voll dan M. Stede, "Con- Based Methods for Sentiment Analysis," *Association Computational Linguistics*, vol. 37, no. 2, pp. 267-307, 2009.
- [15] G. Vinodhini dan R. Chandrasekaran, "A comparative performance evaluation of neural network based approach for sentiment classification of online reviews," *Journal of King Saud University - Computer and Information Sciences*, vol. 28, no. 1, pp. 1-12, 2016.
- [16] T. Gruber, "Ontology," dalam *Encyclopedia of Databases*, Boston, Springer, 2009.
- [17] F. Ali, E. K. Kim dan Y.-G. Kim, "Type-2 fuzzy ontology-based opinion mining and information extraction: A proposal to enhance the hotel reservation system," *Applied Intelligence*, vol. 30, no. 3, pp. 481-500, 2015.
- [18] F. Ali, D. Kwak, P. Khan, S. R. Islam, K. H. Kim dan K. S. Kim, "Fuzzy ontology-based sentiment analysis of transportation and city feature reviews for safe traveling," *Transportation Research Part C: Emerging Technologies*, vol. 77, pp. 33-48, 2017.
- [19] F. Bobillo dan U. Straccia, "Fuzzy ontology representation in OWL 2," *International Journal of Approximate Reasoning*, vol. 52, no. 7, pp. 1073-1094, 2011.
- [20] S. Nurmaini, Saparudin, B. Tutuko dan A. P. P., "Pattern Recognition Approach for Swarm Robots Reactive Control with Fuzzy-Kohonen Networks and Particle Swarm Optimization"

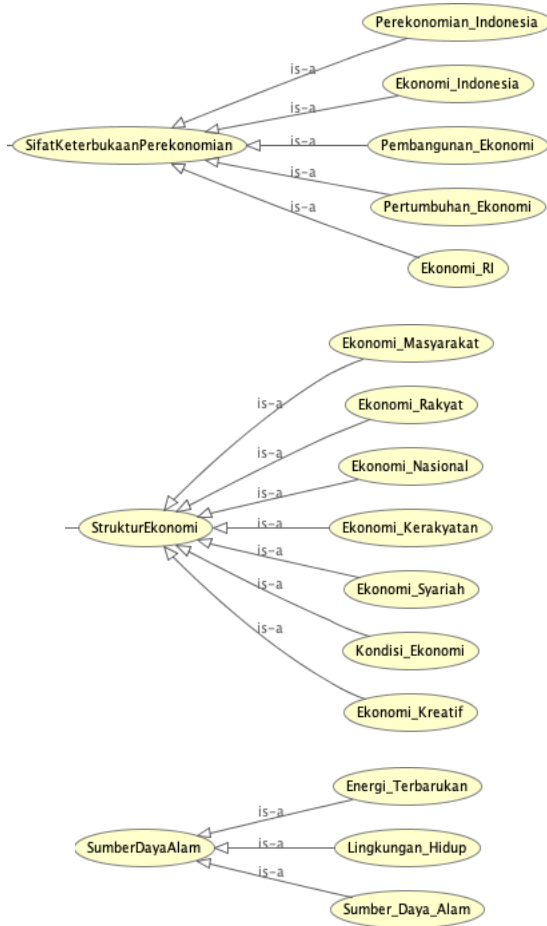
- ithm,” *The Journal of Telecommunication, Electronic and Inter Engineering*, vol. 8, pp. 155-160, 2016.
- [21] F. Ali, E. K. Kim dan Y.-G. Kim, “Type-2 fuzzy ontology-semantic knowledge for collision avoidance of autonomous water vehicles,” *Information Sciences*, vol. 295, pp. 441-2015.
- [22] S. Baccianella, A. Esuli dan F. Sebastiani, “TIWORDNET 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining,” *Proceedings of LREC*, 0, pp. 2200-2204, 2010.
- [23] S. Christina dan D. Ronaldo, “A Survey Of Sentiment Analysis On Bahasa Indonesia,” *Jurnal Teknologi Informasi*, vol. 3, pp. 69-73, 2018.
- [24] I. Syahputra, “Demokrasi Virtual dan Perang Siber di Media Sosial: Perspektif Netizen Indonesia,” *ASPIKOM*, vol. 3, pp. 457-2017.
- [25] D. Yuliah saridwi, “Pemanfaatan Twitter Buzzer Untuk Meningkatkan Partisipasi Pemilih Muda Dalam Pemilihan Umum,” *The Messenger*, vol. 7, pp. 41-48, 2015.
- [26] Lembaga Ketahanan Nasional, *Disiplin Nasional*, Jakarta: Pustaka, 1997.

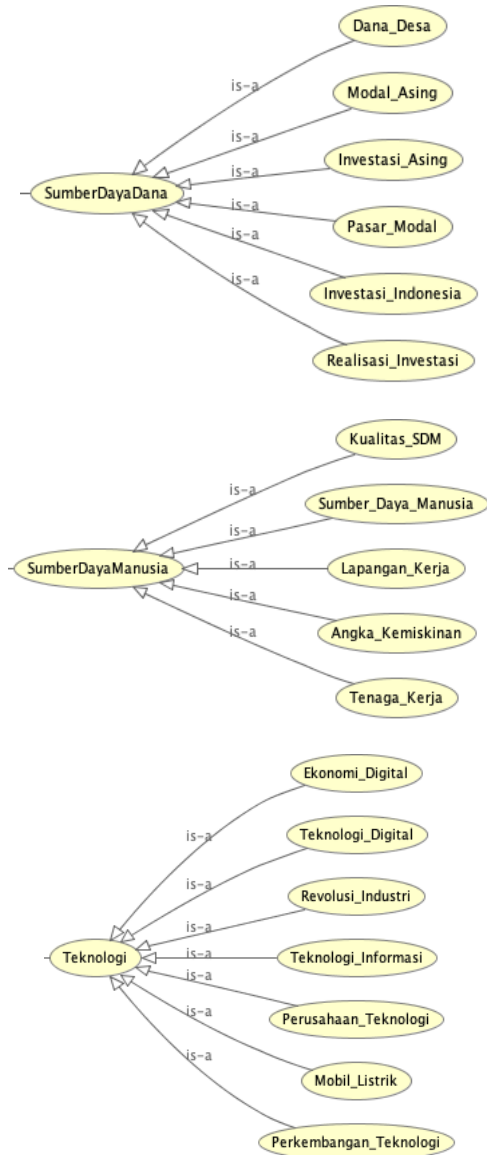
Halaman ini sengaja dikosongkan

LAMPIRAN A. Hasil Ontologi Lengkap

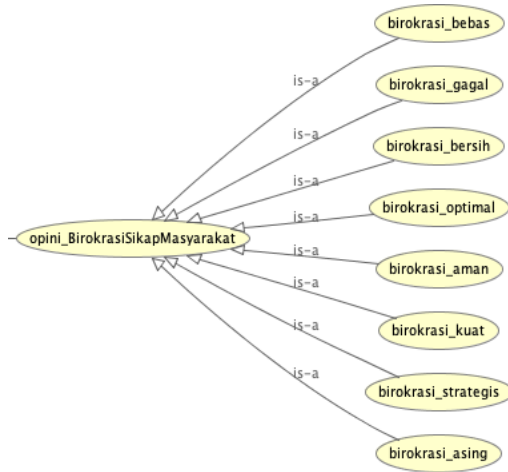
Detail dari FaktorGatra Ekonomi

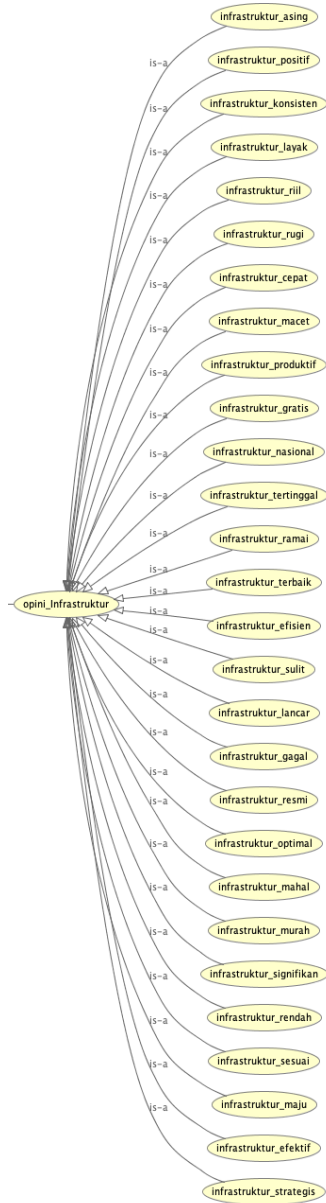


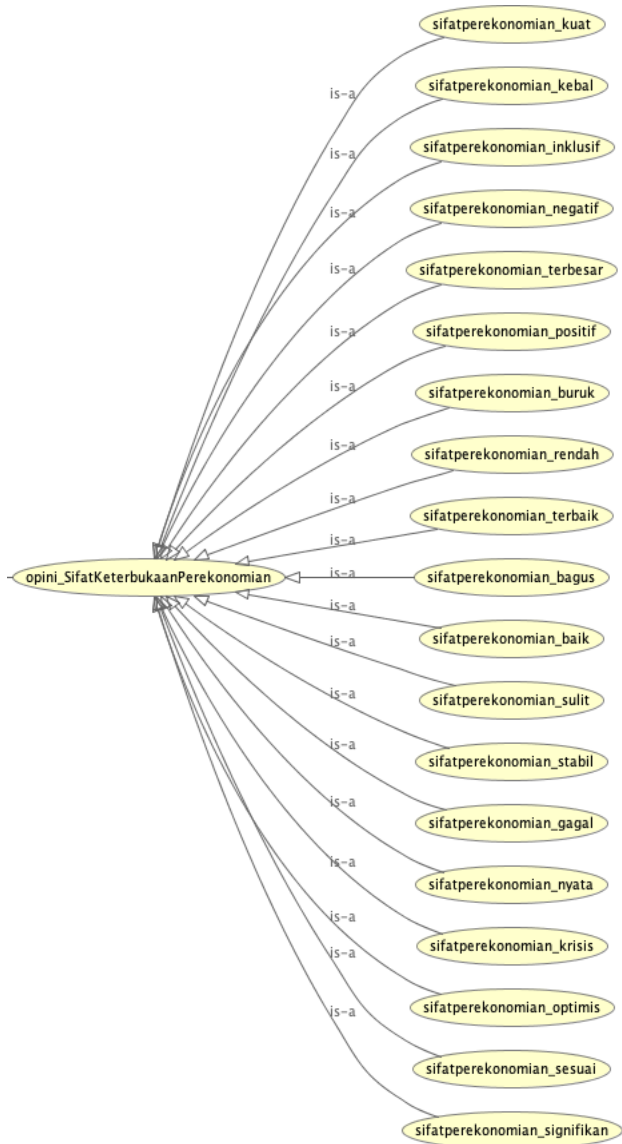


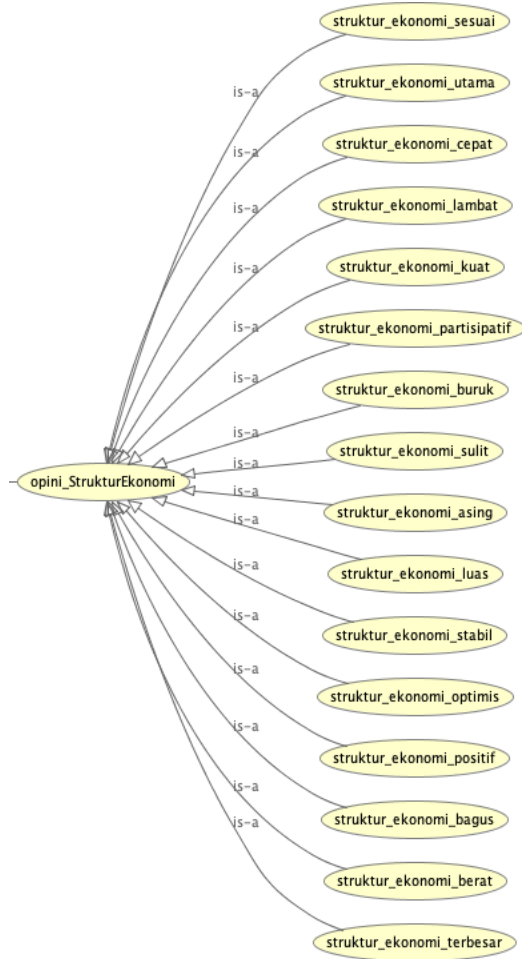


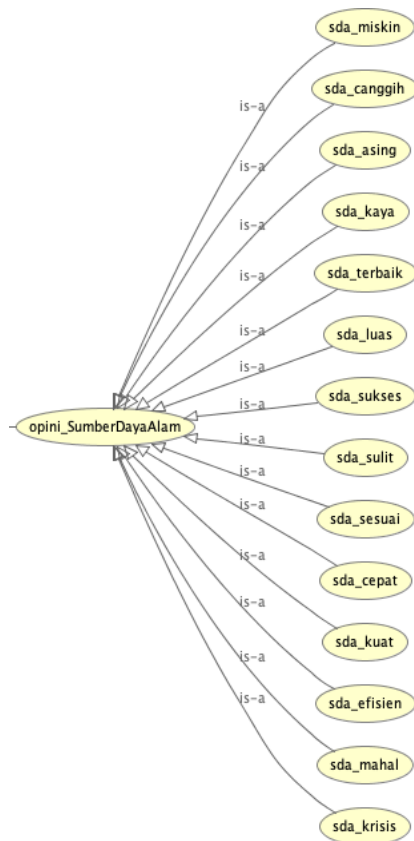
Detail dari kelas Opini

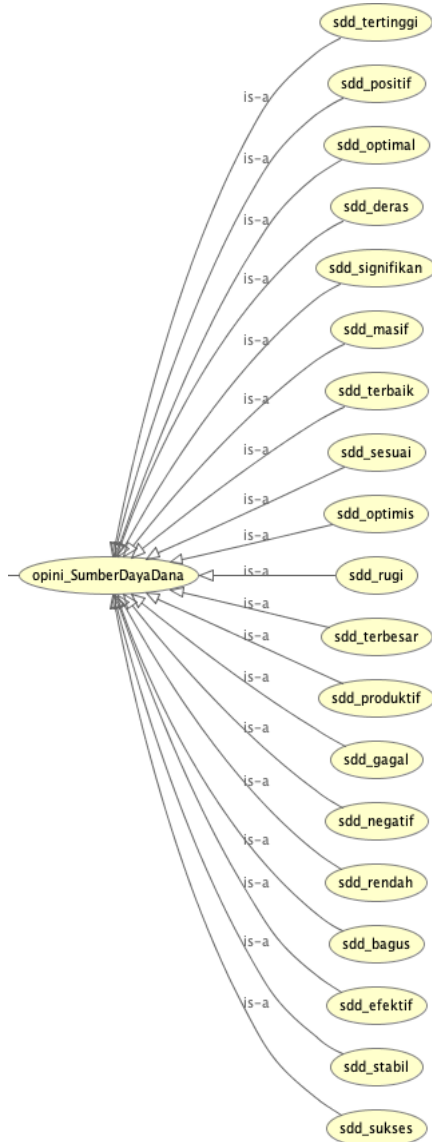


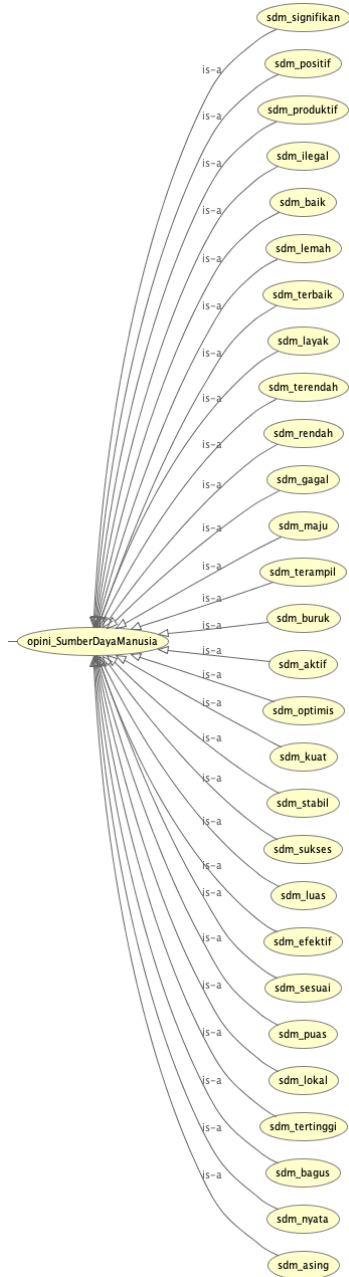


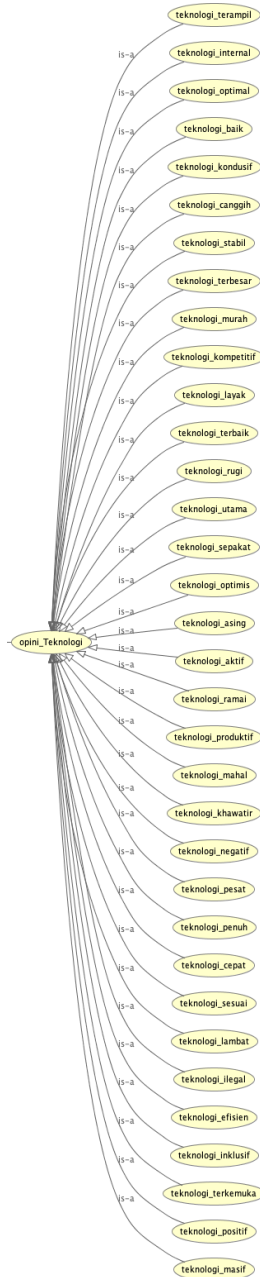












Halaman ini sengaja dikosongkan

LAMPIRAN B. SWRL

<https://drive.google.com/open?id=1kaWNE7-BY4oZqgUPrwfbA-yXEbBwS-9F>

Halaman ini sengaja dikosongkan

BIODATA PENULIS



Penulis lahir di Jakarta pada tanggal 8 September 1997. Merupakan anak terakhir dari empat bersaudara. Penulis telah menempuh beberapa pendidikan formal yaitu; SD Islam PB Soedirman, SMP Negeri 49 Jakarta, dan SMA Taruna Nusantara. Pada Tahun 2015 pasca kelulusan SMA, penulis melanjutkan pendidikan di Departemen Sistem Informasi Fakultas Teknologi Informasi dan Komunikasi – Institut Teknologi Sepuluh Nopember (ITS) Surabaya dan terdaftar sebagai mahasiswa dengan NRP 0521154000141

Selama menjadi mahasiswa, penulis mengikuti berbagai kegiatan kemahasiswaan seperti beberapa kepanitian serta pernah menjabat sebagai Kepala Departemen Hubungan Luar HMSI ITS 2017/2018. Penulis juga pernah menjabat sebagai Ketua Koordinator Komisariat di Komunitas Generasi Baru Indonesia Surabaya tahun 2018. Penulis juga tercatat sebagai penerima beasiswa Bank Indonesia tahun 2017 dan 2018.

Pada tahun keempat, karena penulis memiliki ketertarikan di bidang pengolahan data, maka penulis mengambil bidang minat Akuisisi Data dan Diseminasi Informasi (ADDI). Penulis dapat dihubungi melalui *email* di faiqpp@gmail.com. Adapun motto dari penulis adalah:

**~ Memberikan Karya Terbaik Bagi Masyarakat, Bangsa,
Negara, dan Dunia ~**