

ABSTRAK

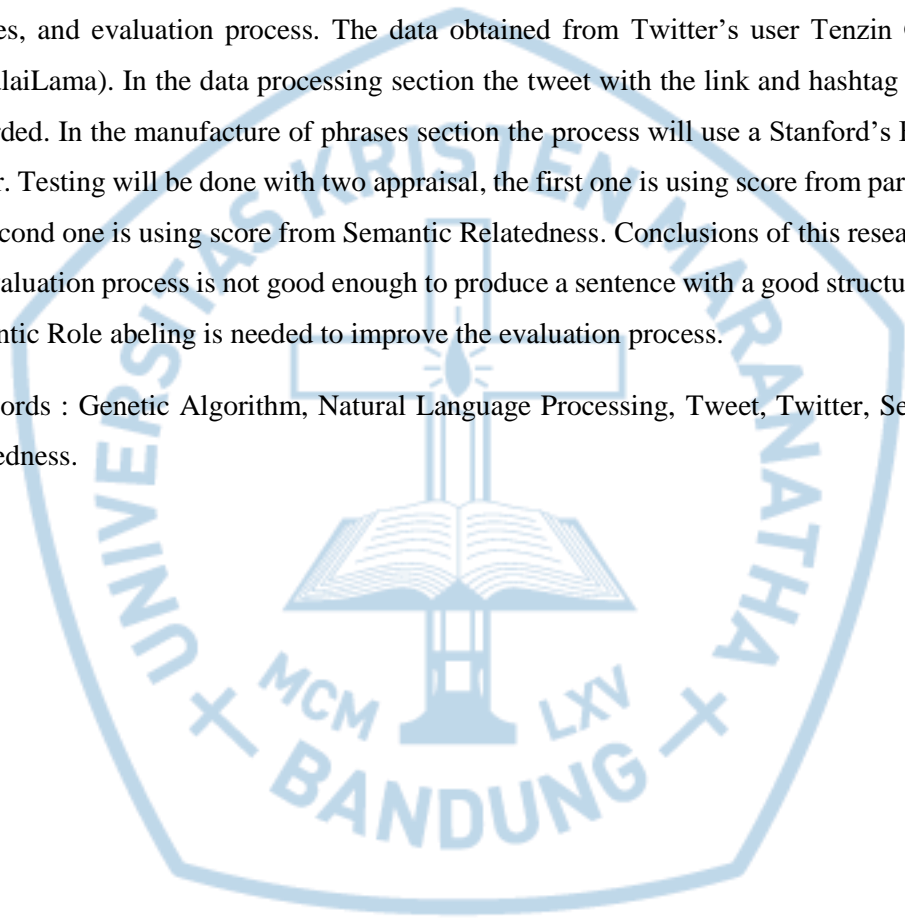
Dengan munculnya berbagai media sosial, banyak orang yang menuliskan bermacam-macam hal, salah satunya memberikan menuliskan motivasi. Dengan demikian, dibuatlah penelitian untuk membuat sebuah kalimat yang memiliki makna positif berdasarkan kalimat-kalimat bermakna positif yang dibagikan oleh pengguna media sosial. Pembentukan kalimat dilakukan dengan menggunakan algoritma genetika. Proses pembangunan kalimat pada penelitian ini dibagi menjadi empat bagian utama, yaitu pengambilan data, pengolahan data, pemecahan data menjadi frasa, dan proses algoritma genetika. Data yang dipakai pada penelitian ini berasal dari sebuah akun *Twitter* milik Tenzin Gyatzo (@DalaiLama). Proses pengolahan data pada penelitian ini adalah membuang semua *tweet* yang memiliki tautan atau *hashtag*. Dalam proses pemecahan data menjadi frasa digunakan sebuah *parser* untuk Bahasa Inggris milik *Stanford University*. Pengujian atau evaluasi pada penelitian ini dilakukan dengan menggabungkan dua penilaian yaitu nilai dari *parser* dan nilai *Semantic Relatedness*. Simpulan dari penelitian ini proses evaluasi yang dilakukan belum cukup untuk menghasilkan kalimat yang memiliki struktur yang benar. Perlu digunakan metode *Semantic Role Labeling* untuk menyempurnakan proses evaluasi pada penelitian ini.

Kata Kunci : Algoritma Genetika, Pemrosesan Bahasa Alami, *Twiter*, *Tweet*, *Semantic Relatedness*.

ABSTRACT

With the emergence of social media, a lot of people write a various things, one of them give some motivation. Therefore, made a research about create a sentence that has a positive meaning based by positive sentence that contained in social media. Generating process is performed using genetic algorithm. Generating process in this research were divided into 4 main section, there is collecting data, data processing, manufacture of phrases, and evaluation process. The data obtained from Twitter's user Tenzin Gyatzo (@DalaiLama). In the data processing section the tweet with the link and hashtag will be discarded. In the manufacture of phrases section the process will use a Stanford's English Parser. Testing will be done with two appraisal, the first one is using score from parser and the second one is using score from Semantic Relatedness. Conclusions of this research are the evaluation process is not good enough to produce a sentence with a good structure. The Semantic Role labeling is needed to improve the evaluation process.

Keywords : Genetic Algorithm, Natural Language Processing, Tweet, Twitter, Semantic Relatedness.

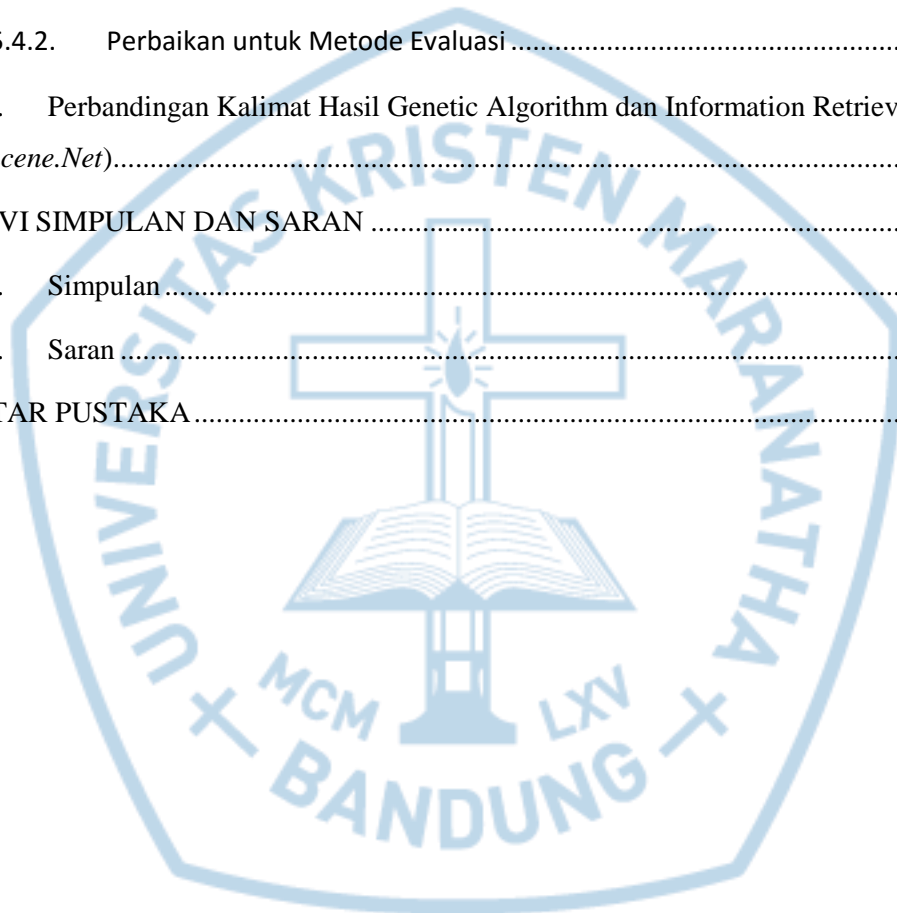


DAFTAR ISI

LEMBAR PENGESAHAN	i
PERNYATAAN ORISINALITAS LAPORAN PENELITIAN	ii
PERNYATAAN PUBLIKASI LAPORAN PENELITIAN	iii
PRAKATA	iv
ABSTRAK.....	v
ABSTRACT	vi
DAFTAR ISI	vii
DAFTAR GAMBAR.....	x
DAFTAR TABEL.....	xii
DAFTAR SOURCE CODE.....	xiii
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang.....	1
1.2. Rumusan Masalah.....	2
1.3. Tujuan Pembahasan	2
1.4. Batasan Masalah	2
1.5. Sumber Data	2
1.6. Sistematika Penyajian.....	2
BAB II LANDASAN TEORI.....	4
2.1. Algoritma Genetika.....	4
2.1.1. Operasi Genetika	4
2.1.2. Operasi Evolusi.....	6
2.2. Pemrosesan Bahasa Alami.....	6
2.2.1. Definisi	6
2.2.2. Stanford English Parser.....	6
2.3. <i>Twitter</i>	7
2.3.1. Definisi	7
2.3.2. <i>Tweetinvi</i>	8

2.4.	<i>Information Retrieval</i>	8
2.4.1.	Definisi	8
2.4.2.	<i>Apache Lucene.Net</i>	9
2.5.	<i>Semantic Relatedness</i>	9
2.5.1.	Definisi	9
2.5.2.	<i>WordNet Similarity for Java (WS4J)</i>	9
BAB III ANALISIS DAN RANCANGAN SISTEM.....		11
3.1.	Konsep Perbedaan <i>Information Retrieval</i> Dan <i>Creative Generative Language Processing</i>	11
3.2.	Analisa Statistik <i>Tweet</i> DalaiLama.....	11
3.3.	Rancangan Eksperimen.....	11
3.3.1.	Pengambilan Data.....	12
3.3.2.	Pengolahan Data.....	12
3.3.3.	Pemecahan Data Menjadi Frasa	13
3.3.4.	Eksperimen	14
3.4.	Desain Perangkat Lunak.....	18
3.4.1.	Pemodelan Perangkat Lunak.....	18
3.4.2.	Desain Antarmuka.....	25
BAB IV PENGEMBANGAN PERANGKAT LUNAK.....		28
4.1.	Implementasi Modul / <i>Class</i>	28
4.1.1.	<i>Class Diagram</i>	28
4.1.2.	Implementasi <i>Method</i> dan Algoritma.....	32
4.2.	Implementasi Antarmuka.....	33
4.2.1.	Implementasi <i>Form</i> Halaman Utama	34
4.2.2.	Implementasi <i>Form Generate Sentence</i>	34
4.2.3.	Implementasi <i>Form Get Tweet</i>	35
BAB V PEMBAHASAN DAN UJI COBA HASIL PENELITIAN		36

5.1.	Penyempurnaan.....	36
5.1.1.	Kamus Sinonim <i>Offline</i>	36
5.1.2.	Penyempurnaan Pada Proses <i>Semantic</i>	36
5.2.	Pengujian Terhadap Jumlah Generasi.....	37
5.3.	Perbandingan Penilaian Menggunakan <i>Semantic Relatedness</i>	38
5.4.	Hipotesis dan Perbaikan untuk Metode Evaluasi.....	39
5.4.1.	Hipotesis	39
5.4.2.	Perbaikan untuk Metode Evaluasi	39
5.5.	Perbandingan Kalimat Hasil Genetic Algorithm dan Information Retrieval (<i>Lucene.Net</i>).....	40
BAB VI SIMPULAN DAN SARAN		41
6.1.	Simpulan	41
6.2.	Saran	41
DAFTAR PUSTAKA		42



DAFTAR GAMBAR

Gambar 2.1 Proses Kawin Silang	5
Gambar 2.2 Proses Mutasi	5
Gambar 2.3 Contoh Pohon Hasil Penyatuan Kata	7
Gambar 3.1 Contoh <i>Tweet</i> Yang Memiliki <i>Hashtag</i> dan tautan	12
Gambar 3.2 Contoh <i>Tweet</i> Yang Dapat Digunakan	13
Gambar 3.3 Gambar Hasil <i>Parsing</i> Menggunakan <i>Stanford Parser</i>	13
Gambar 3.4 Frasa Yang Dihasilkan Dan Nilai Dari <i>Sub-Tree</i>	14
Gambar 3.5 Kalimat yang Dikawin Silangkan	16
Gambar 3.6 Hasil Kawin Silang	16
Gambar 3.7 Kalimat yang Dimutasi	17
Gambar 3.8 Hasil Mutasi	17
Gambar 3.9 <i>Use Case Diagram</i> Sistem	18
Gambar 3.10 <i>Activity Diagram</i> Penarikan Data <i>Tweet</i>	20
Gambar 3.11 <i>Activity Diagram</i> Pemilihan <i>Tweet</i>	21
Gambar 3.12 <i>Activity Diagram</i> Generate Kalimat	23
Gambar 3.13 <i>Activity Diagram</i> Pembuatan Kalimat Acak	24
Gambar 3.14 Desain Antarmuka <i>Form</i> Halaman Utama	25
Gambar 3.15 Desain Antarmuka <i>Form Generate Sentence : Configuration</i>	25
Gambar 3.16 Desain Antarmuka <i>Form Generate Sentence : Detail</i>	26
Gambar 3.17 Desain Antarmuka <i>Form Get Tweet</i>	27
Gambar 4.1 Struktur <i>Class Individu</i>	28
Gambar 4.2 Struktur <i>Class StanfordParser</i>	29
Gambar 4.3 Struktur <i>Class Synonym</i>	29
Gambar 4.4 Struktur <i>Class Tweet</i>	30
Gambar 4.5 Struktur <i>Class LuceneClasses</i>	31
Gambar 4.6 Struktur <i>Class GeneticAlgorithm</i>	31
Gambar 4.7 Implementasi Antarmuka <i>Form</i> Halaman Utama	34
Gambar 4.8 Implementasi Antarmuka <i>Form Generate Sentence : Configuration</i>	34
Gambar 4.9 Implementasi Antarmuka <i>Form Generate Sentence : Detail</i>	35
Gambar 4.10 Implementasi Antarmuka <i>Form Get Tweet</i>	35
Gambar 5.1 Grafik Perbandingan Nilai Fitness dengan Generasi 1	37
Gambar 5.2 Grafik Perbandingan Nilai Fitness dengan Generasi 2	37
Gambar 5.3 Hasil Generasi ke 5, 10, dan 15	38

Gambar 5.4 Hasil Generasi ke 5,10 dan 15 Tanpa Nilai *Semantic* 39
Gambar 5.5 Contoh *Semantic Role Labeling*..... 40
Gambar 5.6 Perbandingan Kalimat *Genetic Algorithm* dan *Information Retrieval*..... 40



DAFTAR TABEL

Tabel 3.1 Deskripsi <i>Use Case Diagram</i> Menarik Data <i>Tweet</i>	19
Tabel 3.2 Deskripsi <i>Use Case Diagram</i> Generate Kalimat	19



DAFTAR SOURCE CODE

Source Code 4.1 <i>Source Code</i> untuk Pemecahan Kalimat Menjadi Frasa	33
Source Code 4.2 <i>Source Code</i> untuk Proses Genetika.....	33

