

## DAFTAR ISI

ABSTRAK.....	i
ABSTRACT .....	ii
KATA PENGANTAR.....	iii
DAFTAR ISI .....	v
DAFTAR GAMBAR .....	vii
DAFTAR TABEL.....	viii
DAFTAR ISTILAH .....	ix
BAB I PENDAHULUAN .....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	3
1.3 Tujuan Penelitian .....	3
1.4 Batasan Masalah .....	3
1.5 Metodologi Penelitian .....	4
1.6 Sistematika Laporan.....	4
BAB II TINJAUAN PUSTAKA .....	6
2.1 <i>Named Entity Recognition</i> (NER).....	6
2.1.1 Ekstraksi <i>Content</i> .....	7
2.1.2 Praproses Teks ( <i>Preprocessing Text</i> ).....	8
2.1.3 Pengenalan Entitas .....	11
2.2 Jaringan Perceptron.....	12
2.2.1 Arsitektur Perceptron .....	13
2.2.2 Algoritma Pelatihan Perceptron .....	14
2.2.3 Contoh Kasus.....	16
2.3 Perceptron dan NER.....	18
2.3.1 Tahap Pelatihan.....	19
2.3.2 Tahap Pengenalan .....	22
2.4 Mikroblog ( <i>Twitter</i> ) .....	23
2.5 Mikroblog ( <i>Twitter</i> ) dan NER .....	24
2.6 Teknik Pengujian .....	26
2.7 Evaluasi Kinerja.....	26
BAB III METODOLOGI PENELITIAN .....	29
3.1 Desain Penelitian .....	29
3.2 Model Pengembang Sistem .....	31
3.3 Alat dan Bahan Penelitian .....	32

BAB IV HASIL PENELITIAN DAN PEMBAHASAN .....	34
4.1 Pengembangan Sistem.....	34
4.1.1 Analisa Kebutuhan.....	34
4.1.2 Perancangan Sistem .....	35
4.1.3 Implementasi .....	37
4.1.4 Integrasi dan Pengujian .....	39
4.2 Pembahasan Penelitian.....	40
4.2.1 Pembuatan Data Set .....	40
4.2.2 <i>Preprocessing Text</i> .....	43
4.2.3 Ekstraksi Fitur.....	44
4.2.4 Pelatihan ( <i>Training</i> ) .....	46
4.2.5 Pengujian ( <i>Testing</i> ) .....	47
4.2.6 Evaluasi .....	48
4.2.7 <i>Postprocessing Text</i> .....	49
4.3 Eksperimen .....	49
4.3.1 Skenario Eksperimen .....	49
4.3.2 Eksperimen Pertama .....	50
4.3.3 Eksperimen Kedua .....	55
4.3.4 Eksperimen Ketiga.....	57
BAB V KESIMPULAN DAN SARAN .....	59
5.1 Kesimpulan.....	59
5.2 Saran.....	59
DAFTAR PUSTAKA .....	61
LAMPIRAN A: Hasil Eksperimen Pengaruh Nilai <i>Threshold</i> .....	A
LAMPIRAN B: Contoh Unigram Yang Dipengaruhi Jenis Data .....	B
LAMPIRAN C: Data Set Nama Orang .....	C
LAMPIRAN D: Data Set Nama Lokasi .....	D
LAMPIRAN E: Data Set Gabungan .....	E
LAMPIRAN F: Contoh Hasil Keluaran Sistem .....	F
LAMPIRAN G: Detail Hasil Eksperimen Pertama .....	G
LAMPIRAN H: Detail Hasil Eksperimen Kedua .....	H
LAMPIRAN I: Detail Hasil Eksperimen Ketiga .....	I

## DAFTAR GAMBAR

<b>Gambar 2.1.</b> Tahapan pada sistem ekstraksi informasi (Meier, 2010) .....	7
<b>Gambar 2.2.</b> Hasil <i>tokenization</i> .....	8
<b>Gambar 2.3.</b> Contoh hasil <i>part-of-speech tagging</i> .....	10
<b>Gambar 2.4.</b> Pembatasan linear dengan <i>perceptron</i> .....	12
<b>Gambar 2.5.</b> Arsitektur jaringan perceptron .....	13
<b>Gambar 2.6.</b> Detail arsitektur jaringan perceptron .....	14
<b>Gambar 2.7.</b> Contoh NER pada tweet .....	25
<b>Gambar 2.8.</b> Bagan proses evaluasi sistem NER .....	27
<b>Gambar 3.1.</b> Desain penelitian .....	29
<b>Gambar 3.2.</b> Model <i>waterfall</i> .....	31
<b>Gambar 4.1.</b> Rancangan umum sistem .....	35
<b>Gambar 4.2.</b> Form hasil klasifikasi .....	39
<b>Gambar 4.3.</b> Grafik hasil eksperimen pertama pada data set nama orang .....	50
<b>Gambar 4.4.</b> Grafik hasil eksperimen pertama pada data set lokasi .....	51
<b>Gambar 4.5.</b> Grafik hasil eksperimen pertama pada data set gabungan .....	51
<b>Gambar 4.6.</b> Grafik hasil eksperimen kedua pada data set nama orang .....	55
<b>Gambar 4.7.</b> Grafik hasil eksperimen kedua pada data set nama lokasi .....	56
<b>Gambar 4.8.</b> Grafik hasil eksperimen ketiga pada data set nama orang .....	57
<b>Gambar 4.9.</b> Grafik hasil eksperimen ketiga pada data set nama lokasi .....	57

## DAFTAR TABEL

<b>Tabel 2.1.</b> Daftar <i>tagset</i> kelas kata .....	10
<b>Tabel 2.2.</b> Algoritma Perceptron .....	15
<b>Tabel 2.3.</b> Fitur kontekstual (Budi, Wahyudi, 2005).....	19
<b>Tabel 2.4.</b> Fitur morfologi (Budi, Wahyudi, 2005) .....	19
<b>Tabel 2.5.</b> Hasil praproses teks data <i>training</i> .....	20
<b>Tabel 2.6.</b> Daftar fitur yang digunakan.....	20
<b>Tabel 2.7.</b> Hasil konversi data <i>training</i> .....	21
<b>Tabel 2.8.</b> Proses pelatihan .....	21
<b>Tabel 2.9.</b> Hasil praproses teks pada data <i>testing</i> .....	22
<b>Tabel 2.10.</b> Hasil konversi dan klasifikasi data <i>testing</i> .....	22
<b>Tabel 2.11.</b> Kode hasil perbandingan objek.....	27
<b>Tabel 4.1.</b> Kelas pada proses <i>preprocessing text</i> .....	37
<b>Tabel 4.2.</b> Kelas pada proses ekstraksi fitur .....	37
<b>Tabel 4.3.</b> Kelas pada proses pelatihan ( <i>training</i> ).....	38
<b>Tabel 4.4.</b> Kelas pada proses pengujian ( <i>testing</i> ).....	38
<b>Tabel 4.5.</b> Kelas pada proses <i>postprocessing text</i> .....	38
<b>Tabel 4.6.</b> Bentuk pengujian .....	39
<b>Tabel 4.7.</b> Contoh data yang dikumpulkan .....	41
<b>Tabel 4.9.</b> Hasil pelabelan entitas.....	42
<b>Tabel 4.10.</b> Kelompok data set.....	42
<b>Tabel 4.11.</b> Contoh hasil <i>preprocessing text</i> .....	43
<b>Tabel 4.12.</b> Contoh hasil akhir dari <i>preprocessing text</i> .....	44
<b>Tabel 4.13.</b> Fitur yang digunakan.....	45
<b>Tabel 4.14.</b> Contoh hasil ekstraksi fitur .....	46
<b>Tabel 4.15.</b> Hasil perhitungan dan prediksi sistem .....	48
<b>Tabel 4.16.</b> Contoh hasil <i>postprocessing text</i> .....	49

## DAFTAR ISTILAH

<b>Istilah</b>	<b>Keterangan</b>
Entitas	Objek dari dunia nyata yang dapat dibedakan dengan objek lain.
Token	Unit terkecil dari sebuah objek.
Data set	Kumpulan dokumen atau teks.
Metadata	Data yang merepresentasikan tentang data tersebut.
<i>Tweet</i>	Teks yang di- <i>post</i> pada twitter
<i>Threshold</i>	Batas ambang
Data training	Data yang digunakan untuk proses pelatihan
Data testing	Data yang digunakan untuk proses pengujian
Fitur	Informasi untuk memprediksi objek
Korpus	Kumpulan teks sistematis berjumlah besar yang dapat disimpan dan diolah secara elektronik
<i>Unigram</i>	Satu kata sebelum entitas
<i>Bigram</i>	Dua kata sebelum entitas