

PENERAPAN TEKNIK *STEMMING* PADA BAHASA JAWA NGOKO BERBASIS ALGORITMA PORTER

Heniz Cahyaning Kartika, Wiwik Suharso, S. Kom, M. Kom
Program Studi Teknik Informatika, Fakultas Teknik,
Universitas Muhammadiyah Jember
E-mail: heniz.kartika@yahoo.com

ABSTRAK

Mencari suatu dokumen yang relevan berdasarkan kebutuhan atau keinginan membutuhkan waktu yang lama, biasanya pengguna mencari informasi tertentu dengan membaca kata demi kata atau dokumen demi dokumen. Dalam Teknologi Informatika terdapat ilmu yang disebut dengan *Information Retrieval* (Temu Kembali Informasi). Didalam *Information Retrieval* terdapat ilmu yang bertujuan untuk menemukan informasi baru yang sebelumnya tidak terungkap dan mencari kata-kata yang dapat mewakili apa yang ada dalam dokumen sehingga dapat dilakukan analisa keterhubungan antar dokumen, dengan memproses dan menganalisa data dalam jumlah besar berupa teks atau dokumen yang disebut dengan *Text Mining*. Dokumen yang dibutuhkan di sini adalah dokumen berbahasa Jawa, di dalam sebuah dokumen tersebut akan dicari kata dasarnya. Hampir sama seperti bahasa Indonesia, di dalam bahasa Jawa juga terdapat imbuhan awalan, akhiran dan sisipan. Bahasa Jawa yang di bagi menjadi beberapa spesifikasi, yang di pilih ini adalah Bahasa Jawa Ngoko, karena banyak di jumpai dokumen-dokumen menggunakan bahasa Jawa Ngoko. Algoritma yang di gunakan adalah Algoritma Stemming, Algoritma Stemming adalah proses mengubah suatu kata bentukan menjadi kata dasar. Algoritma Stemming masih di bagi menjadi beberapa algoritma, salah satunya Algoritma Porter yang khusus menghilangkan ahiran di dalam sebuah kata berimbuhan berbahasa Inggris dan menghilangkan akhiran dan awalan dalam sebuah kata berimbuhan berbahasa Indonesia. Pada Tugas Akhir ini membahas tentang penerapan Stemming bahasa Jawa Ngoko berbasis algoritma porter bahasa Indonesia, dalam penerapan ini bentuk aplikasi yang dapat dijalankan di komputer sehingga mempermudah pengguna dalam pencarian bentuk dasar suatu kata.

Kata kunci: *Text Mining, Algoritma Stemming, Algoritma Porter, Bahasa Jawa Ngoko.*

PENERAPAN TEKNIK *STEMMING* PADA BAHASA JAWA NGOKO BERBASIS ALGORITMA PORTER

Henis Cahyaning Kartika, Wiwik Suharso, S. Kom, M. Kom
Program Studi Teknik Informatika, Fakultas Teknik,
Universitas Muhammadiyah Jember
E-mail: heniz.kartika@yahoo.com

ABSTRACT

Looking for a document based on the needs or desires relevant takes a long time, usually users looking for specific information by reading word for word or document by document. In Information Technology are science called the Information Retrieval. Information contained in the retrieval science that aims to discover new information not previously disclosed and search for words that can represent what is in the document so that it can be analyzed connectivity between document, with processing and analyzing large amounts of data in the form of text or a document called by Text Mining. Documents required here is the Java language documents, in a document that will be searched said basically. Almost the same as Indonesia, in the Java language are also affixes prefixes, suffixes and inserts. Java language is divided into several specifications, which in this select is Ngoko Java language, since many documents encountered in the Java language using Ngoko. Algorithms in use is Stemming Algorithm, Stemming Algorithms is a process of converting a word and eliminate suffix and prefix in a word Indonesia language. In this final project is about the implementation of the Java language Ngoko Stemming algorithm based Indonesia porter, in this application form application that can be run on computers that allows users to search the basic form of a word.

Key words: Text Mining, Stemming Algorithm, Algorithm Porter, Ngoko Java language.

DAFTAR ISI

HALAMAN JUDUL	
LEMBAR PENGESAHAN.....	i
PERNYATAAN	ii
KATA PENGANTAR	iii
MOTTO.....	v
ABSTRAK	vi
ABSTRACT	vii
DAFTAR ISI	viii
DAFTAR TABEL	xi
DAFTAR GAMBAR	xii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah	2
1.4 Tujuan.....	2
1.5 Manfaat.....	2
BAB II TINJAUAN PUSTAKA	3
2.1 Bahasa Jawa.....	3
2.1.1 Bahasa Jawa Ngoko	3
2.1.2 Bahasa Jawa Madya.....	3
2.1.3 Bahasa Jawa Krama.....	3
2.2 Bahasa Jawa Ngoko	4
2.2.1 Ngoko Lugu.....	4
2.2.2 Ngoko Andhap.....	4
2.2.3 Tembung Andap (imbuhan)	5
2.3 Text Mining	6
2.3.1 Case Folding.....	7
2.3.2 Tokenizaing	7
2.3.3 Wordlist/ Stoplist.....	7

2.3.4 Stemming	8
2.4 Stemming	8
2.5 Algoritma Porter Bahasa Indonesia	11
2.6 Dokumen	13
2.6.1 Field	13
BAB III METODE PENELITIAN	15
3.1 Flowchat Rancangan Penelitian	15
3.2 Implementasi Sistem Stemming	16
3.2.1 Case Folding	17
3.2.2 Tokenizing	18
3.2.3 Filtering	19
3.2.4 Stemming	19
3.3 Proses Stemming Bahasa Jawa Ngoko	21
3.4 Desain Tampilan Perangkat Lunak	23
BAB IV HASIL DAN PEMBAHASAN	25
4.1 Skenario Pengujian	25
4.1.1 Input Berupa Kata	25
4.1.2 Input Berupa Kalimat	26
4.1.3 Input Berupa Paragraf	27
4.1.4 Input Berupa Dokumen	29
4.2 Analisis Pengujian	32
4.2.1 Hasil Pengujian Kata	32
4.2.2 Hasil Pengujian Kalimat	33
4.2.3 Hasil Pengujian Paragraf	35
4.2.4 Hasil Pengujian Dokumen	37
BAB V KESIMPULAN DAN SARAN	40
5.1 Kesimpulan	40
5.2 Saran	40
DAFTAR PUSTAKA	41
LAMPIRAN 1	43
LAMPIRAN 2	49

LAMPIRAN 3	52
LAMPIRAN 4	58
LAMPIRAN 5	64
BIODATA.....	76

DAFTAR TABEL

Tabel. 2.1 Awalan dan Akhiran Bahasa Jawa Ngoko.....	5
Tabel 3.1 Hasil Tokenizing	18
Tabel 3.2 Stemming	20
Tabel 4.1 Stemming Perkata	25
Tabel 4.2 Stemming Kalimat.....	26
Tabel 4.3 Stemming Paragraf.....	27
Tabel 4.4 Stemming Dokumen.....	29
Tabel 4.5 Tingkat Keberhasilan dan Tingkat Kegagalan.....	38

DAFTAR GAMBAR

Gambar 2.1 Block diagram Algoritma Porter Bahasa Indonesia	12
Gambar 3.1 Flowchat Rancangan Penelitian	15
Gambar 3.2 Proses Stemming Bahasa Jawa Ngoko	21
Gambar 3.3 Tampilan Proses Stemming.....	23
Gambar 3.4 Tampilan Hasil Stemming.....	24
Gambar 4.1 Grafik Tingkat Keberhasilan	38
Gambar 4.2 Grafik Tingkat Kegagalan.....	39