TUGAS AKHIR

APLIKASI TEKS MINING UNTUK AUTOMASI PENCARIAN KALIMAT INTI DALAM DOKUMEN TUNGGAL BERBAHASA INDONESIA DENGAN METODE *Term Frecuency-Invers Document Frecuenfy* (TF-IDF)

Diajukan Untuk Memenuhi Persyaratan Guna Meraih Gelar Sarjana Komputer Teknik Informatika Universitas Muhammadiyah Jember

Oleh :

Rifkillah Suaibi

NIM : 1210651183

FAKULTAS TEKNIK

JURUSAN TEKNIK INFORMATIKA

UNIVERSITAS MUHAMMADIYAH JEMBER

2018

HALAMAN PENGESAHAN

APLIKASI TEKS MINING UNTUK AUTOMASI PENCARIAN KALIMAT INTI DALAM DOKUMEN TUNGGAL BERBAHASA INDONESIA DENGAN METODE *Term Frecuency-Invers Document Frecuenfy* (TF-IDF)

RIFKILLAH SUAIBI
1210651183

Telah mempertanggung jawabkan Laporan Tugas Akhir pada Sidang Tugas Akhir Tanggal 02 Agustus 2018 sebagai salah satu syarat kelulusan dan mendapatkan gelar Sarjana Komputer (S.Kom) di Universitas Muhammadiyah Jember

Disetujui Oleh,

Dosen Penguji :	**Dosen Pembimbing :**
Penguji I	**Pembimbing I**
Lutfi Ali Muharom, M. Si **NPK. 10 09 550**	**Ir. Dewi Lusiana, MT** **NPK. 04 10 624**
Penguji II	**Pembimbing II**
Ulya Anisatur R, S.Kom., M.Kom **NPK. 12 03 705**	**Daryanto, M. Kom** **NPK. 11 03 589**

Jember, 02 Agustus 2018
Universitas Muhammadiyah Jember

Mengesahkan, **Dekan Fakultas Teknik**	**Mengetahui,** **Ketua Program Study Teknik Informatika**
Ir. Suhartinah, M.T **NPK. 95 05 264**	**Yeni Dwi Rahayu, S.ST, M.Kom** **NPK. 11 03 590**

KATA PENGANTAR

Puji syukur kehadirat allah SWT atas limpahan rahmat dan hidayah sehingga penelitian dengan judul **"APLIKASI TEKS MINING UNTUK AUTOMASI PENCARIAN KALIMAT INTI DALAM DOKUMEN TUNGGAL BERBAHASA INDONESIA DENGAN METODE *Term Frecuency-Invers Document Frecuenfy* (TF-IDF)"** dapat terselesaikan. Penulis menyadari bahwa dalam penelitian maupun penulisan laporan ini banyak pihak yang telah membantu menyelesaikannya. Maka dari itu, saya ucapkan terima kasih kepada:

1. Ibu Ir. Suhartinah, ST., MT., selaku Dekan Fakultas Teknik Universitas Muhammadiyah Jember.
2. Ibu Yeni Dwi Rahayu, S.ST., M.Kom., selaku Ketua Jurusan Teknik Informatika Fakultas Teknik Universitas Muhammadiyah Jember.
3. Ibu Ir. Dewi Lusiana, MT. dan Bapak Daryanto, M. Kom., selaku dosen pembimbing yang telah banyak memberi petunjuk dan arahan.
4. Bapak Lutfi Ali Muharom, M. Si. dan Ibu Ulya Anisatur R, S.Kom., M.Kom., Selaku dosen penguji.

Jember, 02 Agustus 2018

Penulis

DAFTAR ISI

HALAMAN SAMPUL i

HALAMAN PENGESAHAN ii

HALAMAN PERNYATAAN iii

ABSTRAK iv

ABSTRACT v

MOTO vi

HALAMAN PERSEMBAHAN vii

UNGKAPAN TERIMAKASIH viii

KATA PENGANTAR ix

DAFTAR ISI x

DAFTAR GAMBAR xii

DAMTAR TABEL xiii

BAB I PENDAHULUAN 1

1.1 Latar Belakang 1

1.2 Rumusan Masalah 1

1.3 Batasan Masalah 2

1.4 Tujuan Penelitian. 2

1.5 Manfaat Penelitian. 2

BAB II TINJAUAN PUSTAKA 3

2.1 Kalimat Inti 3

2.2 Dokumen 3

2.2.1 Artikel 4

2.3 Teks MIning 5

2.3.1 Casefolding 6

2.3.2 *Tokenizing* 7

2.3.3 *Stoplist* 8

2.3.4 *Stemming* 9

2.4 *Term Frequency-Invers Document Frequency* 16

BAB III METODE PENELITIAN 19

3.1 Desain Sistem **19**
3.2 Menginput Dokumen **20**
3.3 *Teks Mining* **20**
3.3.1 *Scanning* **20**
3. 3.2 *Casefolding* **20**
3. 3.3 *Split* **21**
3 3.4 *Tokenezing* **21**
3. 3.5 *Stoplist* **22**
3. 3.6 *Stemming* **22**
3.4 Menghitung TF dan IDF **23**
3.5 Menghitung Bobot (*W*) Setiap Kata Pada Setiap Kalimat **24**
3.6 Menjumlah dan Mengurutkan Bobot (*W*) Setiap Kalimat **26**
3.7 Dokumen dengan Bobot (*W*) Terbesar Sebagai Kalimat Inti **26**
3.8 Flowchart Aplikasi **27**
BAB IV HASIL DAN PEMBAHASAN **28**
4.1 Dataset **28**
4.2 Praproses **28**
BAB V KESIMPULAN DAN SARAN **36**
5.1 KESIMPULAN **36**
5.2 SARAN **36**
DAFTAR PUSTAKA **37**

DAFTAR GAMBAR

Gambar 2.1 *Text Mining*..... 6

Gambar 2.2 *Stemming* 16

Gambar 3.1 Desain Sistem 19

Gambar 3.2 *Flowchat* Aplikasi 27

Gambar 4.1 Sistem Pencarian Otomatis Kalimat inti 34

Gambar 4.2 Sistem Pencarian Otomatis Kalimat inti 35

Gambar 4.3 Sistem Pencarian Otomatis Kalimat inti 35

DAFTAR TABEL

Tabel 2.1 Contoh *Casefolding* 7

Tabel 2.2 Contoh *Tokenezing* 7

Tabel 2.3 Contoh *Stoplist* 8

Tabel 2.4 Contoh *Stemming* 9

Tabel 2.5 Hapus Awalan Jika Ditemukan 11

Tabel 2.6 Kata Menentukan Awalan Untuk Kata Yang Diawali Dengan "te_" 12

Tabel 2.7 Jenis Awalan Berdasarkan Tipe Awalannya 12

Tabel 2.8 Aturan Untuk *Infectinal Particel* 13

Tabel 2.9 Aturan Untuk *Infectinal Particel Pronon* 13

Tabel 2.10 Aturan Untuk *First Order Deri Vational Prefix* 13

Tabel 2.11 Aturan Untuk *Second Order Derivational Prefix* 14

Tabel 2.12 Aturan Untuk *Derivional Prefix* 14

Tabel 3.1 Pemilihan Dokumen Menjadi Kalimat 21

Tabel 3.2 Proses *Tokenezing* 21

Tabel 3.3 Proses *Filtering* 22

Tabel 3.4 Proses Pembentukan *Stemming* 22

Tabel 3.5 Perhitungan *Tf-Idf* 23

Tabel 3.6 Perhitungan dan Pengurutan Nilai Kumulatif *W* 24

Tabel 4.1 Perhitungan *Tf-Idf* 28

Tabel 4.2 Perhitungan Bobot (*W*) Serta Kata (*Term*) Dalam Setiap Kalimat 31

Tabel 4.3 Pengurutan Nilai (*W*) 33

DAFTAR PUSTAKA

Arifin, A Z., (2002), *Penggunaan Digital Tree Hibrida pada Aplikasi Information Retrieval untuk Dokumen Berita*, Jurusan Teknik Informatika, FTIF, Institut Teknologi Sepuluh Nopember, Surabaya.

Gottschalk, L., (1986), *Understanding History: A Primer of Historical Method* (terjemahan Nugroho Notosusanto). UI Press,Jakarta.

Grossman, D., Ophir, F., (1998), *Information Retrieval : Algorithm and Heuristics*. Kluwer Academic Publisher,USA.

Jones K, S., (1998), *Charting a New Course: Natural Language Processing and Information Retrieval : Essays in Honour of Karen Spärck Jones.* Penerbit: Springer,New York.

Kunchady, M., (2006), *Text Mining Application Programming.* Thomson Learning Inc. ISBN 1- 58450-460-9,USA. Mustaqhfiri, M., Abidin Z., Kusumawati,R.(2011).

Peringkasan Teks Otomatis Berbahasa Indonesia Menggunakan Metode Maximum Marginal Relevance.Ejournal Matics,4,4,135-147. Dimbali 5 Januarii 2012 dari basis data saintek.

Robertson, S., (2004). "*Understanding Inverse Document Frequency: On theoretical arguments for IDF*", Journal of Documentation,Vol.60, no.5, pp. 503- 520,New York.

Widjono, H.S., (2007). "*Bahasa Indonesia Mata Kuliah Pengembangan Kepribadian di PT".* Penerbit : grasindo, Jakarta.

DAFTAR RIWAYAT HIDUP

DATA PRIBADI

Nama Lengkap	: Rifkillah Suaibi
Tempat, Tanggal Lah	: Situbondo, 01 Mei 1995
Jenis Kelamin	: Laki-Laki
Kewarganegaraan	: Indonesia
Agama	: Islam
Status	: Belum Menikah
Alamat	: jl. Talang sari , besuki situbondo jawa timur
No. Telpon	: 081803210896
Email	: rifkillahs@gmail.com

PENDIDIKAN FORMAL

1. SD Negeri 2 Bloro - Besuki : Tahun 2000 - 2006
2. SMP Negeri 1 Besuki : Tahun 2006 - 2009
3. MAN 1 SITUBONDO: Tahun 2009 - 2012
4. Masih Mahasiswa

PENDIDIKAN INFORMAL

1. Basic Language English (Toefle) 2016

RIWAYAT PEKERJAAN

-

Jember, 10 Agustus 2018

Rifkillah Suaibi