

TUGAS AKHIR

PENCARIAN BOBOT SEBUAH TEKS UNTUK
PENGKATEGORIAN BERITA MENGGUNAKAN TF-IDF

(Sumber :<http://www.liputan6.com/>)



ARY FULCHOLIQ

1210651090

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH JEMBER
TAHUN 2018

TUGAS AKHIR

PENCARIAN BOBOT SEBUAH TEKS UNTUK
PENGKATEGORIAN BERITA MENGGUNAKAN TF-IDF

(Sumber :<http://www.liputan6.com//>)

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



ARY FUL CHOLIQ

1210651090

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH JEMBER
TAHUN 2018

Motto

“TETAP MENJADI DIRI SENDIRI DIDUNIA YANG TANPA HENTI
HENTINYA BERUSAHA MENGUBAHMU ADALAH PENCAPAIAN YANG
HEBAT”

~RALPH WALDO EMERSON ~

HALAMAN PERSETUJUAN SIDANG TUGAS AKHIR

**PENCARIAN BOBOT SEBUAH TEKS UNTUK
PENGKATEGORIAN BERITA MENGGUNAKAN TF-IDF**

Oleh :
ARY FUL CHOLIQ
1210651090

Telah disetujui bahwa Laporan Tugas Akhir ini untuk diajukan pada sidang
Tugas Akhir sebagai salah satu syarat kelulusan dan mendapatkan
gelar Sarjana Komputer (S.Kom)
di
Universitas Muhammadiyah Jember

Disetujui Oleh,

Pembimbing I

Ulya Anisatur, M.Kom

NPK. 12 03 705

HALAMAN PENGESAHAN

PENCARIAN BOBOT SEBUAH TEKS UNTUK PENGKATEGORIAN BERITA MENGGUNAKAN TF-IDF

ARY FUL CHOLIQ
1210651090

Telah mempertanggung jawabkan Laporan Tugas Akhirnya pada sidang Tugas Akhir tanggal 31 Juli 2018 salah satu syarat kelulusan dan mendapatkan gelar Sarjana Komputer (S.Kom) di Universitas Muhammadiyah Jember

Disetujui oleh,
Pembimbing I

Ulya Anisatur, M.Kom
NPK. 12 03 705

Penguji I

Penguji II

Ari Eko Wardoyo, S.T., M.Kom.
NIDN. 19750214 200501 1 001

Lutfi Ali Muharram, M.Si
NPK. 10.09.550

Mengesahkan,

Mengetahui,

Dekan Fakultas Teknik

Ketua Program Studi Teknik Informatika

Ir. Suhartinah, M.T.
NPK. 95 05 246

Yeni Dwi Rahayu, M.Kom.
NPK. 11 03 590

SURAT PERNYATAAN KEASLIAN

Saya yang bertanda tangan di bawah ini :

Nama : Aryful choliq

NIM : 1210651090

Institusi : S1 Teknik Informatika, Fakultas Teknik,
Universitas Muhammadiyah Jember.

Menyatakan dengan sebenar-benarnya bahwa segala pernyataan dalam laporan akhir saya yang berjudul **“PENCARIAN BOBOT SEBUAH TEKS UNTUK PENGKATAEGORIAN BERITA MENGGUNAKAN ALGORITMA TF-IDF”** merupakan gagasan dan hasil karya saya sendiri dengan arahan komisi pembimbing.

Semua data dan informasi yang digunakan telah dinyatakan secara jelas dan dapat diperiksa kebenarannya. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan dari penulis lain telah disebutkan dalam naskah dan dicantumkan dalam daftar pustaka di bagian akhir Laporan Akhir ini.

Jember, 14 Desember 2018

Ary ful Choliq
NIM : 1210651090

HALAMAN PERSEMBAHAN

Alhamdulillah wa syukurillah atas kehadiran ALLAH SWT, segala rahmat dan hidayah-Nya sampai saat ini saya masih sangat yakin dan percaya apa yang terjadi pada diri saya ini semua atas kehendak-Nya.

Dalam penelitian yang dituliskan ini, saya persembahkan tugas akhir ini untuk kedua orang tua yang sangat saya cintai dan sayangi. Tidak lupa pula saya ucapkan terimakasih untuk orang – orang yang sangat membantu dalam menyelesaikan penelitian ini, antara lain :

1. Ibu Ulya Anisatur M.Kom. selaku Dosen Pembimbing yang telah meluangkan waktu, pikiran dan memberikan ilmunya selama pengerjaan tugas akhir ini.
2. Bapak Ari Eko Wardoyo, S. T., M. Kom selaku Dosen Penguji I dan Bapak Lutfi Ali Muharram M.SI selaku Dosen Penguji II yang telah memberikan saran dan kritikan yang sangat membangun dalam penyelesaian tugas akhir ini.
3. Kedua orang tua tercinta bapak Sukawit dan ibu Suratmi saya ucapkan terimakasih yang tak terhingga, karena kasih sayang, doa, pengorbanan dan dukungan yang telah beliau – beliau berikan sehingga saya bisa meraih kelancaran dan kesuksesan tugas akhir ini.
4. Terima kasih kepada Adek ayu adek yang tidak bosan – bosannya memberi semangat, dukungan serta doa dalam pengerjaan tugas akhir ini.
5. Terima kasih juga untuk Bang daus, Erick, Satrio, Kopet, Tabung, Muklis, Bagus, Argo, Yoga, Mas gondrong yang selalu memberikan motivasi, doá dan semangat dalam mengerjakan Tugas akhir ini.
6. Terima kasih banyak kepada teman – teman mincing teman kalipepe atas dukungan dan semangatnya. Akhir kata, dengan segala kerendahan hati penulis menyadari bahwa pelaksanaan penelitian dan penyusunan laporan Tugas Akhir ini masih belum sempurna. Oleh karena itu, kritik dan saran

yang membangun sangat penulis harapkan. Semoga penyusunan Laporan Tugas Akhir ini bermanfaat bagi semua pihak.

Jember, 14 Desember 2018

Ary ful Choliq
NIM : 1210651090

KATA PENGANTAR

Assalamu 'alaikum Warrahmatullah Wabbarakatu

Segala puji bagi Allah Yang Maha Pengasih dan Penyayang yang telah melimpahkan rahmat dan hidayah-Nya, sehingga Laporan Tugas Akhir yang berjudul **“PENCARIAN BOBOT SEBUAH TEKS UNTUK PENGKATEGORIAN BERITA MENGGUNAKAN ALGORITMA TF-IDF”** dapat diselesaikan dengan baik. Adapun tujuan penulisan laporan ini adalah untuk memenuhi persyaratan dalam menyelesaikan Laporan Tugas Akhir Program Studi Teknik Informatika Universitas Muhammadiyah Jember.

Akhir kata, penulis menyadari bahwa pelaksanaan penelitian dan penyusunan laporan Tugas Akhir ini masih belum sempurna. Oleh karena itu, kritik dan saran yang membangun sangat penulis harapkan. Semoga penyusunan laporan Tugas Akhir ini bermanfaat bagi semua pihak.

Wassalamu 'alaikum Warrahmatullah Wabbarakatu

Jember, 14 Desember 2018

Ary Ful Choliq
NIM : 12106510990

DAFTAR ISI

HALAMAN JUDUL	i
MOTTO	ii
HALAMAN PERSETUJUAN	iii
HALAMAN PENGESAHAN	iv
SURAT PERNYATAAN KEASLIAN	v
ABSTRAK	vi
ABSTRACT	vii
HALAMAN PERSEMBAHAN	viii
KATA PENGANTAR	x
DAFTAR ISI	xi
DAFTAR GAMBAR	xiv
DAFTAR TABEL	xvi
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah	2
1.4 Tujuan Penelitian	2
1.5 Manfaat Penelitian	3
BAB II TINJAUAN PUSTAKA	4
2.1 Teks mining	4
2.2 Peringkasan Teks Otomatis	4
2.3 Steming Porter	8
2.4 Term Frequency-Invers Documen Frequency (TF-IDF)	9
2.5 Pengertian berita	14

BAB III METODE PENELITIAN	16
3.1 Identifikasi masalah	16
3.2 Tinjauan Pustaka	17
3.3 Metode Pengumpulan Data	17
3.4 Analisis Kebutuhan	17
3.5 Metode Pengembangan Sistem.....	18
3.6 Perhitungan Manual	19
3.6.1 Proses Tokenizing	19
3.6.2 Proses Stopword/filtering	21
3.6.3 Proses stemer	23
3.6.4 Proses pembobotan TF-IDF.....	24
BAB IV IMPLEMENTASI DAN PENGUJIAN	28
4.1 Spekifikasi Shoftwhare Dan Hardwhare Yang Di Gunakan	28
4.1.1 Spekifikasi Shoft Whare.....	28
4.1.2 Spekifikasi Hard whare	28
4.2 Implementasi program	28
4.2.1 Implementasi Teks PreProsesing	28
4.3 Implementasi Antar Muka	30
4.3.1 Halaman Utama.....	31
4.3.2 Halaman Pembobotan Berita	31
4.3.1 Analisa pengujian.....	41
BAB V KESIMPULAN DAN SARAN	44
5.1 Kesimpulan	44
5.2 Saran	44
DAFTAR PUSTAKA	45
LAMPIRAN	46
BIOGRAFI PENULIS	55

DAFTAR GAMBAR

Gambar 2.1 Flowchard Stopword	6
Gambar 2.2 stemer Porter	11
Gambar 3.1 metodologi Penelitan.....	16
Gambar 3.2 Waterfall.....	18
Gambar 4.1 Rangkaian Proses Teks Prosesing	29
Gambar 4.2 Halaman Utama	31

DAFTAR TABEL

Tabel 2.1 Kelompok Rule Pertama	9
Tabel 2.2 Kelompok Rule Ke Dua	9
Tabel 2.3 Kelompok Rule Ke Tiga.....	9
Tabel 2.4 Kelompok Rule Ke Empat	10
Tabel 2.5 Kelompok Rule Ke Lima.....	10
Tabel 3.1 Karakter Yang akan Di Hapus	20
Tabel 3.2 Ilustrasi Tokenizing	20
Tabel 3.4 Hasil Stopword	21
Tabel 3.6 Steming Proter	23
Tabel 3.7 Pembobotan Kata	24
Tabel 4.1 Katagori Berita	33
Tabel 4.2 Table Perhitungan TF-IDF	34
Tabel 4.3 Hasil Pengujian	41

DAFTAR PUSTAKA

Anggaradana, A. 2013. Aplikasi text mining untuk pencarian laporan tugas akhir menggunakan metode tf/idf dan vector space berbasis web. Skripsi. Universitas Pendidikan Ganesha.

An, J., Cha, M., Gummadi, K., dan Crowcroft, J. 2011. Media Landscape in Twitter, A New World Convention and Political Diversity, University of Cambridge.

Feldman, R., & Sanger, J. 2007. The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data. Cambridge University Press: New York.

Hovy, E. 2001. Automated Text Summarization. In R. Mitkov. (Eds) Handbook of computation linguistics. Oxford: Oxford University Press.

Kurniawan, Andre, & Toba, H. 2009. Pembuatan aplikasi bergerak temu ulang file elektronik berbahasa indonesia dengan memanfaatkan java cldc. Seminar Aplikasi Teknologi Informasi 2009.

Lioret, Elena. 2008. Text Summarization: An Overview, Dept. Lenguajes y Sistemas Informaticos Universidad de Alicante Alicante, Spain.

Mani, Inderjeet, dan Marbury, T, Mark. 1999. Advances in Automatic Text Summarization, Massachusetts Institute of Technology, Massachusetts, Amerika.

Manning, C. D., Raghavan, P., & Schutze, H. 2008. Introduction to Information Retrieval. Cambridge University Press: New York.

Rachmansyah. 2009. Stemmer Kata Bahasa Indonesia 2009. Elex Media

Komputindo. Jakarta.

Rajaraman, A. & Ullman, J. D. 2011. Mining of Massive Datasets. Cambridge University Press: New York.

Tokunaga., T. And Iwayama, M., 1994. Text Categorization Based On Weighted Inverse Document Frequency. Tokyo, Japan: Tokyo Institute of Technology.

Wardhana, Wisnu L. 2008. Peringkasan Multi-Dokumen Untuk Bahasa Indonesia Menggunakan Teknik Centroid-Based Summarization Dan Teknik K-MeansBased Summarization. Skripsi pada Fakultas Ilmu Komputer Universitas Indonesia. tidak diterbitkan.