

SKRIPSI

**IMPLEMENTASI METODE TF-IDF UNTUK APLIKASI
PERINGKAS *DOCUMENT BASE***



DAVID KRISTIAWAN

Nomor Mahasiswa: 145410108

PROGRAM STUDI TEKNIK INFORMATIKA

SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER

AKAKOM

YOGYAKARTA

2020

SKRIPSI

**IMPLEMENTASI METODE TF-IDF UNTUK APLIKASI
PERINGKAS *DOCUMENT BASE***

**Diajukan sebagai salah satu syarat untuk menyelesaikan studi jenjang strata
satu (S1)**

Program Studi Sistem Informasi

Sekolah Tinggi Manajemen Informatika dan Komputer

Akakom

Yogyakarta

Disusun Oleh

DAVID KRISTIawan

Nomor Mahasiswa: 145410108

PROGRAM STUDI TEKNIK INFORMATIKA

SEKOLAH TINGGI MANAJEMAN INFORMATIKA DAN KOMPUTER

AKAKOM

YOGYAKARTA

2020

HALAMAN PERSETUJUAN

HALAMAN PERSETUJUAN

Judul : Implementasi Metode TF-IDF Untuk Aplikasi
Peringkat Document Base

Nama : David Kristiaawan

Nomor mahasiswa : 145410108

Program Studi : Teknik Informatika

Jenjang : Strata Satu (S1)

Tahun : 2020



Yogyakarta,2020

Mengetahui

Dosen Pembimbing

Femi Dwi Astuti, S.Kom., M.Cs

HALAMAN PENGESAHAN

HALAMAN PENGESAHAN

SKRIPSI

IMPLEMENTASI METODE TF-IDF UNTUK APLIKASI

PERINGKAS *DOCUMENT BASE*

Telah dipertahankan di depan Dewan Penguji Skripsi dan dinyatakan
diterima untuk memenuhi sebagai syarat guna memperoleh Gelar Sarjana
Komputer Sekolah Tinggi Manajemen Informatika dan Komputer

YOGYAKARTA

Yogyakarta,

Mengesahkan

Dewan Penguji

Tanda Tangan

1. Femi Dwi Astuti, S.Kom., M.Cs

2. Edi Faizal, S.T., M.Cs

Mengetahui

Ketua Program Studi Teknik Informatika



28 AUG 2020

Dini Fakta Sari, S.T., M.T.

HALAMAN PERSEMBAHAN DAN MOTTO

**Selama kesempatan keberhasilan tidak 0% masih ada
kesempatan untuk berhasil**

Saya mempersembahkan karya tulis ini beserta rasa trima kasih
saya untuk:

1. Tuhan yang Maha Esa yang telah ,memberkati saya dengan rahmat agar saya dapat menyelesaikan karya tulis ini.
2. Ayahanda Yohanes Sri Purnomo, ibunda Yustina Sri Wahyuniati, dan kakak saya Markus Alfian Novianto, yang telah memberikan doa, bimbingan, dan kasih sayang kepada saya yang sedang bekerja keras menyelesaikan karya tulis ini.
3. Kepada bude saya Maria Titik Maryati yang telah peduli dan selalu menanyakan keadaan saya kepada dosen pembimbing dan dosen lain yang dapat membantu saya.
4. Kepada seluruh sanak saudara saya yang telah membantu saya walaupun dari jauh.
5. Kepada teman baik saya yang selalu memberikan dukungan kepada saya.
6. Terakhir namun tidak kalah pentingnya, untuk diri saya yang pada akhirnya akan menjejakkan kaki saya ke dunia luar saya persembahkan karya tulis ini.

KATA PENGANTAR

Puji Syukur penulis panjatkan kepada Tuhan Yang Maha Esa, karena dengan rahmat dan karunia-Nya, penulis dapat menyelesaikan skripsi yang berjudul “Implementasi Metode TF-IDF Untuk Aplikasi Peringkasan *Document Base*”.

Skripsi ini dibuat dan diajukan untuk memenuhi syarat guna memperoleh gelar **Sarjana Teknik** pada Fakultas **Teknik Informatika** di **STMIK AKAKOM Yogyakarta**. Selain itu, tujuan dari penulisan skripsi ini adalah untuk memberikan pengetahuan kepada pembaca mengenai **peringkasan teks dengan metode TF-IDF**.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna karena adanya keterbatasan ilmu dan pengalaman yang dimiliki. Oleh karena itu, semua kritik dan saran yang bersifat membangun akan penulis terima dengan senang hati. Penulis berharap, semoga skripsi ini dapat bermanfaat bagi semua pihak yang memerlukan.

Yogyakarta, _____ 2020

Penulis

INTISARI

IMPLEMENTASI METODE TF-IDF UNTUK APLIKASI PERINGKAS *DOCUMENT BASE*

David Kristiawan

Sejak terjadinya revolusi digital, informasi berbasis elektronik lebih dibutuhkan dari pada informasi berbentuk padat/kertas. Padahal, informasi yang ada dalam internet tidak sepenuhnya benar. Dibutuhkan waktu yang lama untuk mencermati informasi yang kita dapatkan, apakah informasi yang dicari adalah akurat atau palsu. Karena itulah metode TF-IDF diambil sebagai inti dari penelitian ini.

Peringkasan Teks Otomatis (Automatic Text Summarization) merupakan pembuatan rangkuman dari sebuah sumber teks secara otomatis dengan menggunakan serta memanfaatkan sistem peringkasan teks yang dijalankan pada komputer. Sebuah sistem peringkasan teks diberi masukan berupa teks, kemudian sistem melakukan proses peringkasan, dan sistem menghasilkan keluaran berupa teks yang lebih singkat dari sumber teks aslinya. Hasil peringkasan mengandung poin-poin penting atau informasi utama dari teks sumber asli.

Metode *Term Frequency-Inverse Document Frequency* (TF-IDF) adalah cara pemberian bobot hubungan suatu kata terhadap dokumen. Metode ini menggabungkan dua konsep untuk perhitungan bobot, yaitu *Term frequency* (TF) merupakan frekuensi kemunculan kata (t) pada kalimat (d). *Document frequency* (DF) adalah banyaknya kalimat dimana suatu kata (t) muncul. Metode ini dirasa tepat dalam menggunakan prinsip peringkasan teks. Hasil dari penelitian ini menunjukkan bahwa nilai yang didapat dalam vektorisasi TF-IDF sangat berguna dalam peringkasan teks yang menampilkan kata-kata penting yang terdapat dalam berita.

Hasil dari penelitian ini menyimpulkan bahwa berdasarkan dari nilai vektorisasi TF-IDF, teks berita yang diringkas rata-rata memiliki kurang lebih 56,97% kata dari jumlah kata yang ada dalam teks berita yang belum diolah.

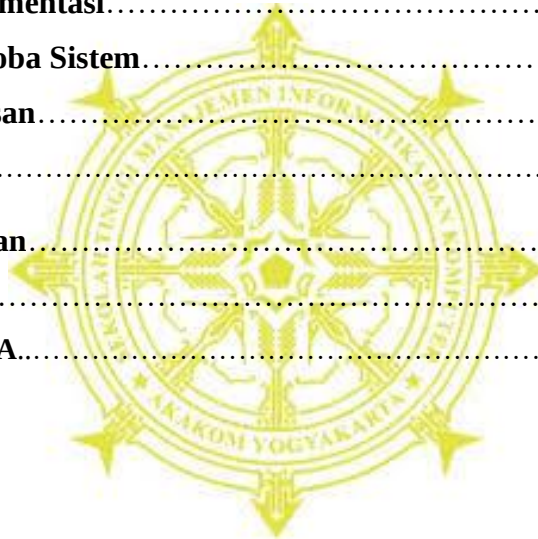
Kata kunci: *TF-IDF, informasi, kata, peringkasan*

DAFTAR ISI

Hal

SKRIPSI.....	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN.....	iii
HALAMAN PERSEMBAHAN DAN MOTTO.....	iv
KATA PENGANTAR.....	v
INTISARI.....	vi
DAFTAR ISI.....	vii
DAFTAR GAMBAR.....	ix
DAFTAR TABEL.....	x
BAB 1 PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	2
1.3 Ruang Lingkup.....	2
1.4 Tujuan Penelitian.....	3
1.5 Manfaat Penelitian.....	3
1.6 Sistematika penulisan.....	3
BAB 2 TINJAUAN PUSTAKA DAN DASAR TEORI.....	5
2.1 Tinjauan Pustaka.....	5
2.2 Landasan Teori.....	9
2.2.1 Metode TF-IDF.....	9
2.2.2 Porter Stemming.....	11
BAB 3 METODE PENELITIAN.....	12
3.1 Bahan/Data.....	12

3.2	Peralatan.....	12
3.2.1	Perangkat Keras.....	12
3.2.2	Perangkat Lunak.....	13
3.3	Prosedur Pengumpulan Data.....	13
3.4	Analisis dan Rancangan Sistem.....	13
3.4.1	Analisis Sistem.....	13
3.4.2	Analisis Kenutuhan.....	14
3.4.3	Rancangan Sistem.....	16
BAB 4 IMPLEMENTASI DAN PEMBAHASAN.....		18
4.1	Implementasi dan Uji Coba Sistem.....	18
4.1.1	Implementasi.....	18
4.1.2	Uji Coba Sistem.....	25
4.2	Pembahasan.....	31
BAB 5 PENUTUP.....		37
5.1	Kesimpulan.....	37
5.2	Saran.....	38
DAFTAR PUSTAKA.....		39



DAFTAR GAMBAR

Gambar 2. 1 contoh inverted index.....	6
Gambar 3. 1 alur diagram.....	16
Gambar 4. 1 program untuk input.....	18
Gambar 4. 2 import module.....	19
Gambar 4. 3 inialisasi module dan penerimaan input.....	20
Gambar 4. 4 preprocessing dan filtering.....	20
Gambar 4. 5 menyiapkan data yang dapat dipanggil kembali sebagai ringkasan.....	21
Gambar 4. 6 tokenizing dan stemming.....	22
Gambar 4. 7 vektorisasi TF-IDF.....	22
Gambar 4. 8 rata-rata nilai TF-IDF.....	23
Gambar 4. 9 menyusun kalimat mejadi ringkasan.....	24
Gambar 4. 10 ringkas.html.....	24
Gambar 4. 11 hasilringkas.html.....	24
Gambar 4. 12 input percobaan 1.....	25
Gambar 4. 13 output percobaan pertama.....	26
Gambar 4. 14 input percobaan 2.....	26
Gambar 4. 15 output percobaan kedua	27
Gambar 4. 16 input percobaan 3.....	27
Gambar 4. 17 output percobaan ketiga	28
Gambar 4. 18 input percobaan 4.....	28
Gambar 4. 19 output percobaan keempat	29
Gambar 4. 20 input percobaan 5.....	29
Gambar 4. 21 output percobaan kelima	30
Gambar 4. 22 input percobaan 6.....	30
Gambar 4. 23 output percobaan keenam	31

DAFTAR TABEL

Tabel 2. 1 daftar tinjauan pustaka.....	7
Tabel 2. 2 tabel frekuensi kata(term) dalam dokumen/kalumat(document).....	10
Tabel 3. 1 daftar perangkat keras.....	13
Tabel 3. 2 daftar perangkat lunak.....	13
Tabel 4. 1 contoh pperhitungan TF dalam percobaan pertama paragraf 1.....	32
Tabel 4. 2 hasil perhitungan TF-IDF secara lengkap.....	34

