

**INFORMATION RETRIEVAL PENCARIAN INFORMASI
DI YAYASAN DARUL ULUM AL CHOLILY
MENGUNAKAN METODE *TF-IDF***

SKRIPSI

Disusun sebagai salah satu syarat
untuk memperoleh gelar Sarjana Pendidikan
Program Studi Teknik Informatika

Oleh

Muhammad Tajul Mafakhir


NIM 2120200486

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS NAHDLATUL ULAMA SUNAN GIRI
2024**

PERNYATAAN KEASLIAN TULISAN

Saya menyatakan bahwa skripsi ini bebas plagiat, dan apabila dikemudian hari terbukti terdapat plagiat dalam skripsi ini, maka saya bersedia menerima sanksi sesuai ketentuan peraturan perundang-undangan.

Bojonegoro, 12 Juli 2024



Muhammad Tajul Mafakhir

Nim : 2120200486

HALAMAN PERSETUJUAN

Nama : Muhammad Tajul Mafakhir


NIM : 2120200486

Judul : *Information Retrieval* Pencarian Informasi Di Yayasan Darul Ulum Al
Cholily Menggunakan Metode *TF-IDF*

Telah disetujui dan dinyatakan memenuhi syarat untuk diajukan dalam sidang
skripsi.

Bojonegoro, 14 Marret 2024

Pembimbing I



Afta Ramadhan Zayn, M.Kom.

NIDN. 0708048903

Pembimbing II



Rizka Nur Faila, M.T.

NIDN. 0723019301

HALAMAN PENGESAHAN

Nama : Muhammad Tajul Mafakhir
NIM : 2120200486
Judul : *Information Retrieval* Pencarian Informasi Di Yayasan Darul
Ulum Al Cholily Menggunakan Metode *TF-IDF*

Telah dipertahankan dihadapan penguji pada tanggal 12 Juli 2024

Dewan Penguji I



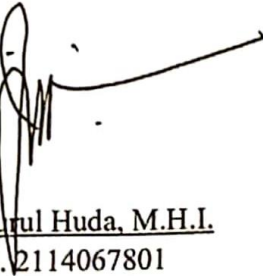
Ucta Pradema Sanjaya, M.Kom.
NIDN. 0729128903

Pembimbing I



Afta Ramadhan Zayn, M.Kom.
NIDN. 0708048903

Dewan Penguji II



Dr. Nurul Huda, M.H.I.
NIDN. 2114067801

Pembimbing II



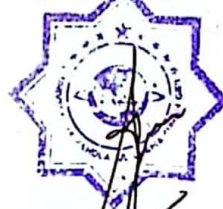
Rizka Nur Faila, M.T.
NIDN. 0723019301

Mengetahui,
Dekan Fakultas Sains dan
Teknologi



M. Jauhar Vikri, M.Kom
NIDN. 0712078803

Mengetahui,
Ketua Program Studi Teknik
Informatika



Mula Azung Barata, S.S.T., M.Kom
NIDN. 0711049301

“Sebaik-baik Manusia Adalah Yang Bermanfaat Bagi Orang Lain”

(HR. Ahmad)

PERSEMBAHAN

Dengan rasa syukur kepada Allah SWT, dengan telah diselesaikannya skripsi ini penulis mempersembahkannya kepada :

1. Kedua orang tua tercinta yang selalu memberikan semangat, dukungan, nasehat serta doa yang tiada terputus sehingga skripsi ini dapat selesai. Terimakasih Ibu Sri Purwati dan Bapak Moch. Zaenal Arifin atas segala yang engkau berikan.
2. Nita Cahyani, S.Si., M. Stat. selaku dosen pembimbing akademik yang telah membimbing jalannya perkuliahan.
3. Teman-teman seperjuangan Program studi Teknik Informatika 2020 yang telah ikut berjuang bersama-sama dari awal kuliah hingga terselesaikannya studi.

UNUGIRI

KATA PENGANTAR

Puji syukur kepada Allah SWT yang telah memberikan rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul “*INFORMATION RETRIEVAL* PENCARIAN INFORMASI DI YAYASAN DARUL ULUM AL CHOLILY MENGGUNAKAN METODE *TF-IDF*”. Skripsi ini disusun untuk memenuhi syarat memperoleh gelar Sarjana Komputer pada Program Studi Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Nahdlatul Ulama Sunan Giri Bojonegoro.

Penulis mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Bapak Dr. H. Yogi Prana Izza, Lc., M.A. selaku PLT Rektor Universitas Nahdlatul Ulama Sunan Giri Bojonegoro.
2. Bapak M. Jauhar Vikri, M.Kom., selaku Dekan Fakultas Sains dan Teknologi Universitas Nahdlatul Ulama Sunan Giri Bojonegoro.
3. Bapak Mula Agung Barata, S.S.T., M.Kom., selaku Ketua Program Studi Teknik Informatika.
4. Bapak Afta Ramadhan Zayn, M.Kom., selaku dosen pembimbing I yang telah meluangkan waktu, pikiran, dan tenaga untuk membimbing serta mendampingi penulis hingga skripsi ini dapat diselesaikan, serta selalu memberi motivasi kepada penulis.
5. Ibu Rizka Nur Faila, M.T., selaku dosen pembimbing II, yang telah meluangkan waktu, pikiran, dan tenaga untuk membimbing serta mendampingi penulis hingga skripsi ini dapat diselesaikan, serta tidak pernah berhenti memberi motivasi kepada penulis.
6. Bapak H. Ahmad Kholil, S.E., M.Pd., selaku Ketua Yayasan Darul Ulum Al Cholily.

Semoga segala bantuan dan kebaikan tersebut dibalas oleh Allah SWT dengan kebaikan yang lebih besar. Penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan dan memiliki banyak kekurangan.

Oleh karena itu, penulis meminta maaf apabila terdapat kesalahan dalam penulisan skripsi ini.

Bojonegoro, 12 Juli 2024

Penulis



UNUGIRI

ABSTRACT

Mafakhir, Muhammad Tajul, 2024. Information Retrieval at Yayasan Darul Ulum Al Cholily Using the TF-IDF Method. Thesis, Department of Informatics Engineering, Faculty of Science and Technology, Universitas Nahdlatul Ulama Sunan Giri. Main Supervisor Afta Ramadhan Zayn, M.Kom. and Co-Supervisor Rizka Nur Faila, M.T.

Keywords: Information Retrieval, TF-IDF, Data Management

Yayasan Darul Ulum Al Cholily, an educational institution established in 1937, requires a system to help manage documents supporting academic and administrative activities. This research aims to develop a web-based information retrieval system using the Term Frequency-Inverse Document Frequency (TF-IDF) method to facilitate document search based on keywords. Through analysis, design, implementation, and testing, the TF-IDF method is applied to identify and count the frequency of word occurrences in documents, thereby supporting the search function. The implementation results show that the system successfully indexes and searches documents with high accuracy. Testing keyword searches for "Rizqi" and "Karyawan" shows relevant and consistent results with manual calculations, although there are limitations in document title indexing. The system also displays documents in an appropriate relevance order based on TF-IDF scores. In conclusion, the developed TF-IDF-based search system enhances data management efficiency at Yayasan Darul Ulum Al Cholily and is expected to positively impact the quality of education provided.

UNUGIRI

ABSTRAK

Mafakhir, Muhammad Tajul, 2024. *Information Retrieval Pencarian Informasi di Yayasan Darul Ulum Al Cholily Menggunakan Metode TF-IDF*. Skripsi, Jurusan Teknik Informatika Fakultas Sains dan Teknologi Universitas Nahdlatul Ulama Sunan Giri. Pembimbing Utama Afta Ramadhan Zayn, M.Kom. dan Pembimbing Pendamping Rizka Nur Faila, M.T.

Kata Kunci: Pencarian Informasi, *TF-IDF*, Pengelolaan Data

Yayasan Darul Ulum Al Cholily, sebuah lembaga pendidikan yang berdiri sejak tahun 1937, memerlukan sistem untuk mengelola dokumen mendukung aktivitas akademik dan administratif. Penelitian ini bertujuan mengembangkan sistem pencarian informasi berbasis web dengan metode *Term Frequency-Inverse Document Frequency (TF-IDF)* untuk memudahkan pencarian dokumen berdasarkan kata kunci. Melalui analisis, perancangan, implementasi, dan pengujian, metode *TF-IDF* diterapkan untuk mengidentifikasi dan menghitung frekuensi kemunculan kata dalam dokumen, sehingga mendukung fungsi pencarian. Hasil implementasi menunjukkan bahwa sistem berhasil mengindeks dan mencari dokumen dengan akurasi tinggi. Pengujian pencarian kata kunci "Rizqi" dan "Karyawan" menunjukkan hasil yang relevan dan konsisten dengan perhitungan manual, meskipun terdapat keterbatasan dalam pengindeksan judul dokumen. Sistem ini juga menampilkan dokumen dengan urutan relevansi yang sesuai berdasarkan skor *TF-IDF*. Kesimpulannya, sistem pencarian berbasis *TF-IDF* yang dikembangkan meningkatkan efisiensi pengelolaan data di Yayasan Darul Ulum Al Cholily dan diharapkan dapat berdampak positif pada kualitas pendidikan yang disediakan.

DAFTAR ISI

PERNYATAAN KEASLIAN TULISAN	Error! Bookmark not defined.
HALAMAN PERSETUJUAN.....	iii
HALAMAN PENGESAHAN.....	Error! Bookmark not defined.
MOTTO	v
PERSEMBAHAN	vi
KATA PENGANTAR	vii
ABSTRACT.....	ix
ABSTRAK	x
DAFTAR ISI.....	xi
DAFTAR TABEL.....	xiii
DAFTAR GAMBAR	xv
DAFTAR LAMPIRAN.....	xvii
BAB I.....	1
PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Batasan Penelitian	4
1.4 Tujuan Penelitian.....	5
1.5 Manfaat Penelitian.....	5
BAB II.....	6
TINJAUAN PUSTAKA	6
2.1 Tinjauan Pustaka	6
2.2 Dasar Teori	14
2.2.1 <i>Information Retrieval</i>	14
2.2.2 <i>Term Frequency-Inverse Document Frequency (TF-IDF)</i>	16
2.2.3 <i>Unified Modeling Language (UML)</i>	21
2.2.4 Perangkat Lunak Berbasis WEB.....	25
BAB III	26
METODE PENELITIAN.....	26
3.1 Subjek dan Objek Penelitian	26
3.2 Jadwal Penelitian.....	26
3.3 Metode Penelitian.....	27

3.3.1 Analisis	27
3.3.2 Perancangan	29
3.3.3 Implementasi.....	36
3.3.4 Rancangan Pengujian.....	38
BAB IV	40
HASIL DAN PEMBAHASAN.....	40
4.1 Hasil Aplikasi	40
4.2 Implementasi <i>TF-IDF</i> dalam Program	41
4.3 Hasil Penghitungan Manual dan Aplikasi	47
4.4.1 Dokumen yang Digunakan	47
4.4.2 Perbandingan Penghitungan Manual dengan Aplikasi	50
4.4 Hasil Pengujian Aplikasi	76
4.5.1 Pengujian Pencarian Kata Kunci	76
4.5.2 Pengujian Pengindeksan Dokumen	79
4.5.3 Pengujian Penilaian Relevansi	82
KESIMPULAN DAN SARAN.....	85
5.1 Kesimpulan.....	85
5.2 Saran	85
DAFTAR PUSTAKA	87
LAMPIRAN.....	90



UNUGIRI

DAFTAR TABEL

Tabel 2.1 Tabel Data Penelitian Terkait.....	9
Tabel 2.2 Tiga alur utama <i>Information Retrieval</i> menurut Gerald J. & Mark T (1997).....	15
Tabel 2.3 Algoritma <i>TF-IDF (Term Frequency-Inverse Document Frequency)</i> ...19	
Tabel 2.4 Simbol-simbol <i>use case</i>	21
Tabel 2.5 Simbol-simbol <i>Class Diagram</i>	22
Tabel 2.6 Simbol-simbol <i>Activity Diagram</i>	23
Tabel 3.1 Jadwal penelitian.....	26
Tabel 3.2 Kebutuhan Fungsional.....	28
Tabel 3.3 Kebutuhan nonfungsional.....	28
Tabel 3.4 Kebutuhan Perangkat Lunak.....	28
Tabel 3.5 Penjelasan desain <i>User Interface</i>	36
Tabel 3.6 Implementasi.....	36
Tabel 3.7 Skenario Pengujian.....	38
Tabel 3.8 Rancangan Pengujian <i>Blackbox</i>	39
Tabel 4.1 Dokumen yang diunggah.....	47
Tabel 4.2 Tabel Skenario Pengujian Pencarian Kata.....	77
Tabel 4.3 Hasil Pencarian Kata Kunci “Rizqi”.....	77
Tabel 4.4 Hasil Pencarian Kata Kunci “Karyawan”.....	78
Tabel 4.5 Pengujian Pengindeksan Dokumen.....	79
Tabel 4.6 Pengujian Penilaian Relevansi.....	82
Tabel 4.7 Hasil Relevansi Kata “Rizqi”.....	82

Tabel 4.8 Hasil Relevansi Kata “Karyawan ”.....83



DAFTAR GAMBAR

Gambar 2.1 Proses <i>Information Retrieval</i> menurut Chowdhury (2010)	16
Gambar 2.2 <i>Flowchart</i> Algoritma <i>TF-IDF</i>	19
Gambar 3.1 <i>SDLC Waterfall</i>	27
Gambar 3.2 <i>Flowchart</i> algoritma <i>TF-IDF</i>	30
Gambar 3.3 <i>Use Case Diagram</i>	32
Gambar 3.4 <i>Class Diagram</i>	33
Gambar 3.5 <i>Activity Diagram</i>	34
Gambar 3.6 Desain <i>User Interface</i>	35
Gambar 4.1 Halaman Utama Aplikasi.....	40
Gambar 4.2 Form Unggah Dokumen.....	41
Gambar 4.3 <i>Field</i> Pencarian.....	41
Gambar 4.4 Form Hasil Pencarian.....	41
Gambar 4.5 Hasil Penghitungan Aplikasi pada Dokumen Miftakhussurur Rizqi Davit Pratama.txt.....	53
Gambar 4.6 Hasil Penghitungan Aplikasi pada Dokumen Muhammad Rizqi Al Ghifari.txt.....	55
Gambar 4.7 Hasil Penghitungan Aplikasi pada Dokumen Muhammad Rizqi Firdaus.txt.....	57
Gambar 4.8 Hasil Penghitungan Aplikasi pada Dokumen Muhammad Rizqi Ramadhani.txt.....	59
Gambar 4.9 Hasil Penghitungan Aplikasi pada Dokumen RIZQI MUBAROK.txt.....	61
Gambar 4.10 Hasil Penghitungan Aplikasi pada Dokumen SK Pembagian Tugas Ganjil 2018.txt.....	63

Gambar 4.11 Hasil Penghitungan Aplikasi pada Dokumen SK Pembagian Tugas Ganjil 2019.txt.....	64
Gambar 4.12 Hasil Penghitungan Aplikasi pada Dokumen SK Pembagian Tugas Ganjil 2020.txt.....	66
Gambar 4.13 Hasil Penghitungan Aplikasi pada Dokumen SK Pembagian Tugas Genap 2018.txt.....	68
Gambar 4.14 Hasil Penghitungan Aplikasi pada Dokumen SK Pembagian Tugas Genap 2019.txt.....	70
Gambar 4.15 Hasil Penghitungan Aplikasi pada Dokumen Elysia Ayu Kusumastuti.txt.....	71
Gambar 4.16 Hasil Penghitungan Aplikasi pada Dokumen Moh. Ainul Rizky.txt.....	73
Gambar 4.17 Hasil Penghitungan Aplikasi pada Dokumen Nauval Syafiq Al Hay Tammy.txt.....	75



UNUGIRI

DAFTAR LAMPIRAN

Lampiran 1. Surat Penelitian.....90



UNUGIRI