



**KOMBINASI METODE *VECTOR SPACE MODEL* DAN TEKNIK
*HIERARCHICAL AGGLOMERATIVE CLUSTERING SINGLE
 LINKAGE* DALAM RANCANG BANGUN *INFORMATION
 RETRIEVAL* PADA PERPUSTAKAAN DIGITAL**

TUGAS AKHIR

**Diajukan Sebagai Salah Satu Syarat
 Untuk Memperoleh Gelar Sarjana Teknik Pada
 Jurusan Teknik Informatika**

Oleh :

**WICE MILJAYA INDAH PUTRI
 11051201994**



Hak Cipta milik UIN Suska Riau

Islamic University of Sultan Syarif Kasim Riau

**FAKULTAS SAINS DAN TEKNOLOGI
 UNIVERSITAS ISLAM NEGERI SULTAN SYARIF KASIM RIAU
 PEKANBARU
 2016**

- Hak Cipta Dilindungi Undang-Undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

**KOMBINASI METODE VECTOR SPACE MODEL DAN
TEKNIK *HIERARCHICAL AGGLOMERATIVE CLUSTERING*
SINGLE LINKAGE DALAM RANCANG BANGUN
INFORMATION RETRIEVAL PADA PERPUSTAKAAN
DIGITAL**

WICE MILJAYA INDAH PUTRI

11051201994

Tanggal Sidang : 08 Januari 2016

Periode Wisuda : Februari 2016

Jurusan Teknik Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Sultan Syarif Kasim Riau
Jl. Soebrantas No. 155 Pekanbaru

ABSTRAK

Mesin pencarian (*Search Engine*) merupakan alat bantu yang efektif dalam membantu manusia menemukan informasi yang mereka inginkan. Namun sering kali penelusuran dokumen dari hasil mesin pencarian tidak relevan dengan apa yang diinginkan sehingga *user* terpaksa harus membaca satu persatu dokumen yang ditemukembalikan. Untuk itu diperlukan mesin pencarian dengan implementasi sistem temu kembali informasi menggunakan hasil *clustering* yang membantu menemukembalikan dokumen dalam bentuk kelompok-kelompok dokumen yang memiliki kemiripan dengan *query* sehingga akan menghemat waktu dan tenaga yang diperlukan. Metode yang digunakan adalah metode *Vektor Space Model* dan *Hierarchical Agglomerative Clustering Single Linkage*. Sistem akan menghitung bobot TF-IDF serta kemiripan *query* dengan dokumen yang selanjutnya membentuk satu *cluster* menggunakan perhitungan *hierarchical agglomerative clustering single linkage* untuk mendapatkan kelompok dokumen yang memiliki kemiripan dengan *query*. Hasil dari pengujian memperlihatkan bahwa *query* dengan *cluster* yang sudah terbentuk memiliki optimasi pencarian yang lebih bagus baik dari segi kemiripan dokumen hingga waktu *retrieve* dibandingkan *query* dengan *cluster* yang belum terbentuk.

Kata Kunci: *hierarchical agglomerative clustering single linkage*, mesin pencarian, *query*, *vector space model*

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

KATA PENGANTAR



Assalamu'alaikum Wr. Wb

Puji syukur saya ucapkan atas kehadiran ALLAH SWT. yang telah memberikan seluruh rahmat dan karuniaNya sehingga pada akhirnya penulis dapat menyelesaikan laporan Tugas Akhir ini sebagai salah satu syarat untuk memenuhi persyaratan akademis dalam rangka meraih gelar sarjana di program studi Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau. Shalawat dan salam selalu disampaikan kepada Baginda Nabi Muhammad SAW yang telah membawa umatnya dari zaman kegelapan menuju zaman yang terang benderang seperti saat sekarang ini. Adapun judul Laporan Tugas Akhir “Kombinasi Metode *Vector Space Model* dan *Teknik Hierarchical Agglomerative Clustering Single Linkage* dalam Rancang Bangun Information Retrieval pada Perpustakaan Digital”.

Dalam penyelesaian Laporan Tugas Akhir ini penulis mendapat bimbingan, bantuan, dan dukungan yang sangat berarti dari berbagai pihak. Untuk itu penulis mengucapkan banyak terimakasih kepada :

1. Bapak Prof. Dr. Munzir Hitami, MA selaku Rektor Universitas Islam Negeri Sultan Syarif Kasim Riau.
2. Bapak Dr. Hartono, M.Pd. selaku Dekan Fakultas Sains dan Teknologi.
3. Ibu Fadhillah Syafria S.T, M.Kom. selaku Dosen Pembimbing yang telah banyak meluangkan waktu, tenaga dan pikiran dalam membimbing dan memberikan petunjuk yang sangat berharga dalam penulisan laporan ini.
4. Bapak Febiyanto, M.Kom dan Ibu Yusra, ST, MT. selaku Dosen Penguji I dan Penguji II yang telah memberikan masukan dan pengarahan demi selesainya Laporan ini.
5. Ibu Fadhillah Syafria ST, M.Kom. selaku Koordinator Tugas Akhir Jurusan Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau..


Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

6. Ibu dan Bapak dosen TIF yang telah banyak memberikan ilmunya kepada penulis.
7. Untuk keluarga penulis yang tercinta Ayahanda Masril. A, Ibunda Jahyanti, Adik Penulis; Yofal Meiyu Ferdimas, Dela Hantrini Prillia dan Manisa Laila yang telah banyak memberikan dukungan moril dan materil serta do'a restu kepada penulis sehingga dapat menyelesaikan Tugas Akhir ini.
8. Untuk sahabat-sahabat penulis Untuk Mona Munawaroh, Juliana, Sechreni Nur Sartika. Dan Ade Elvira terimakasih karena selalu memberi semangat dan dukungan untuk penulis sehingga bisa menyelesaikan laporan Tugas Akhi ini.
9. Untuk teman-teman seperjuangan Initial D angkatan 2010 yang tidak bisa penulis sebutkan nama satu persatu yang selalu mendukung dalam pengerjaan kerja praktek penulis.
10. Semua pihak yang terlibat baik langsung maupun tidak langsung dalam penelitian tugas akhir ini yang tidak dapat penulis sebutkan satu persatu.

Semoga laporan ini dapat bermanfaat bagi penulis khususnya maupun bagi pembaca pada umumnya. Dalam penulisan laporan Tugas Akhir ini, penulis menyadari masih terdapat kekurangan baik dari segi pembahasan maupun dari segi penggunaan kata-kata. Untuk itu penulis mengharapkan adanya kritik maupun saran yang bersifat membangun atau bertujuan untuk menyempurnakan isi dari laporan tugas akhir ini. Semoga ALLAH SWT memberikan balasan yang setimpal atas jasa pihak-pihak yang membantu di atas dan semoga laporan tugas akhir ini bermanfaat bagi kita semua. Amin.

Wassalamu'alaikumWr.,Wb.

Pekanbaru, 08 Januari 2016

WICE MILJAYA INDAH

DAFTAR ISI

| | Halaman |
|---|----------------|
| HALAMAN COVER | i |
| LEMBAR PERSETUJUAN | ii |
| LEMBAR PENGESAHAN | iii |
| LEMBAR HAK ATAS KEKAYAAN INTELEKTUAL | iv |
| LEMBAR PERNYATAAN | v |
| LEMBAR PERSEMBAHAN | vi |
| ABSTRAK | vii |
| ABSTRACT | viii |
| KATA PENGANTAR | ix |
| DAFTAR ISI | xi |
| DAFTAR GAMBAR | xiv |
| DAFTAR RUMUS | xvi |
| DAFTAR TABEL | xvii |
| | |
| BAB I PENDAHULUAN | |
| 1.1 Latar Belakang | I-1 |
| 1.2 Rumusan Masalah | I-2 |
| 1.3 Batasan Masalah..... | I-3 |
| 1.4 Tujuan Penelitian | I-3 |
| 1.5 Sistematika Penulisan | I-3 |
| | |
| BAB II LANDASAN TEORI | |
| 2.1 Perpustakaan Digital (<i>Digital Library</i>)..... | II-1 |
| 2.2 Sistem Temu Kembali Informasi (<i>Information Retrieval</i>)..... | II-2 |
| 2.2.1 Arsitektur Sistem Temu Kembali Informasi | II-5 |
| 2.2.2 Tahapan Proses Sistem Temu Kembali Informasi | II-6 |
| 2.3 <i>Vector Space Model</i> | II-11 |
| 2.4 <i>Hierarchical Agglomerative Clustering</i> | II-16 |
| 2.4.1 Metode <i>Hierarchical Agglomerative Clustering</i> | II-18 |
| 2.4.2 <i>Single Linkage</i> | II-19 |
| 2.5 <i>Precision & Recall</i> | II-20 |

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

| | | |
|-----|--------------------------|-------|
| 2.6 | Penelitian Terkait | II-21 |
|-----|--------------------------|-------|

BAB III METODOLOGI PENELITIAN

| | | |
|-------|-------------------------------------|-------|
| 3.1 | Identifikasi Masalah | III-2 |
| 3.2 | Studi Pustaka | III-2 |
| 3.3 | Pengumpulan Data | III-2 |
| 3.4 | Analisa..... | III-2 |
| 3.4.1 | Analisa Data | III-2 |
| 3.4.2 | Penerapan Metode..... | III-3 |
| 3.5 | Perancangan | III-4 |
| 3.5.1 | Perancangan Database..... | III-4 |
| 3.5.2 | Perancangan Interface Website | III-4 |
| 3.6 | Pengujian..... | III-4 |
| 3.7 | Kesimpulan Dan Saran..... | III-4 |

BAB IV ANALISA DAN PERANCANGAN

| | | |
|-------|---|-------|
| 4.1 | Analisa Permasalahan | IV-1 |
| 4.2 | Analisa Sistem..... | IV-2 |
| 4.2.1 | Analisa Masukan..... | IV-3 |
| 4.2.2 | Analisa Proses | IV-4 |
| 4.2.3 | Analisa Keluaran..... | IV-4 |
| 4.3 | Analisa Sistem Temu Kembali Informasi | IV-4 |
| 4.3.1 | <i>Preprocessing</i> | IV-6 |
| 4.3.2 | Pembobotan TF-IDF | IV-8 |
| 4.3.3 | Menghitung Similaritas Query dan Dokumen | IV-10 |
| 4.3.4 | Membangun Cluster (<i>Hierarchical Agglomerative Clustering Method Single Linkage</i>)..... | IV-13 |
| 4.4 | Analisa Data Sistem | IV-15 |
| 4.4.1 | <i>Context Diagram</i> | IV-15 |
| 4.4.2 | <i>Data Flow Diagram Level 1</i> | IV-15 |
| 4.4.3 | <i>Data Flow Diagram Level 2</i> | IV-17 |
| 4.4.4 | <i>Entity Relationship Diagram (ERD)</i> | IV-21 |
| 4.5 | Perancangan <i>Database</i> | IV-22 |

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

4.6 Perancangan *Interface* Sistem Temu Kembali Informasi... .IV-28

BAB V IMPLEMENTASI DAN PENGUJIAN

| | | |
|-------|---|------|
| 5.1 | Implementasi | V-1 |
| 5.1.1 | Lingkungan Implementasi..... | V-1 |
| 5.1.2 | Hasil Implementasi | V-2 |
| 5.2 | Pengujian. | V-9 |
| 5.2.1 | Pengujian Blackbox..... | V-9 |
| 5.2.2 | Pengujian Algoritma dan Akurasi Sistem..... | V-11 |
| 5.2.3 | Pengujian Optimasi Waktu Pencarian..... | V-14 |
| 5.3 | Kesimpulan Pengujian..... | V-16 |

BAB VI PENUTUP

| | | |
|-----|------------------|------|
| 6.1 | Kesimpulan | VI-1 |
| 6.2 | Saran..... | VI-1 |

DAFTAR PUSTAKA xviii

LAMPIRAN A.....A-1