

ABSTRAK

Evolusi dari hadirnya Internet sebagai perpustakaan digital telah merubah cara memproses suatu dokumen data-teks. Sistem pencarian informasi data-teks saat ini sangatlah diperlukan. Suatu sistem yang cepat, akurat dan terpercaya untuk mencari dokumen yang dibutuhkan oleh pengguna.

Saat ini sistem pencarian informasi data-teks dibangun berdasarkan konsep ruang vektor. Kecocokan antara query dan database dapat dipecahkan dengan oprasi vektor sederhana.

Dalam Tugas Akhir ini akan direalisasikan sistem pencarian informasi data-teks menggunakan Model Ruang Vektor. Dokumen *database* dan *query* akan direpresentasikan sebagai vektor.

Berdasarkan percobaan yang dilakukan, penggunaan model ruang vektor dapat menyederhanakan suatu sistem pencarian data-teks. Pencarian dapat berjalan dengan cepat akurat dan terpercaya.

ABSTRACT

Evolution of the Internet as digital library has transformed the way of processing text-based document. An information retrieval now becomes very importance thing. Users need an information retrieval system that can work fast, accurate and reliable.

Recently information retrieval systems are developed based on concept of vector space. Relevancy between query and database are identified via simple vector operation.

In this final project that will be realized Information Retrieval System using Vector Space Model. The database documents and the query will be represented as vector.

Based on the result of examination, Vector Space Model can simplify Information Retrieval and the final result of the system is accurate and reliable.

KATA PENGANTAR

Puji dan Syukur kepada Allah SWT atas karunia-Nya, sehingga penulis dapat menyelesaikan Laporan Tugas Akhir dengan judul “**Sistem Pencarian Informasi Data-Teks Menggunakan Model Ruang Vektor**”.

Laporan Tugas Akhir ini adalah salah satu syarat yang harus dipenuhi untuk memperoleh gelar Sarjana Strata Satu (S1) di Fakultas Teknik Jurusan Teknik Elektro Universitas Kristen Maranatha.

Penulis menyadari bahwa penyusunan Laporan Tugas Akhir ini tidak dapat selesai dengan baik tanpa bantuan berbagai pihak. Oleh karena itu penulis ingin menyampaikan ucapan terima kasih yang sedalam-dalamnya, terutama kepada :

1. Orangtua dan saudara atas semua dukungan yang tiada henti dan tanpa pamrih.
2. Bapak Riko Arlando ST., MT., selaku pembimbing yang telah bersedia meluangkan waktu untuk memberikan bimbingan dan pengarahan.
3. Bapak Ir. Aan Darmawan MT., selaku Ketua Jurusan Teknik Elektro Universitas Kristen Maranatha.
4. Ibu Ir. Anita Supartono M.Sc., selaku Koordinator Tugas Akhir.
5. Seluruh karyawan dan staff tata usaha Jurusan Teknik Elektro Universitas Kristen Maranatha.
6. Seluruh karyawan dan staff perpustakaan Fakultas Teknik Universitas Kristen Maranatha.
7. Tia Indriana tercinta dan keluarga, yang selalu memberikan doa dan dorongan dalam mengerjakan Tugas Akhir ini.
8. Teman-teman penulis, Ivan Rony, Poltak Putra Manik, Irwan Muthalib, Dodi Simanjuntak ST, Yudi Toding, Ronny Turangan ST, Ari Fadil Nasution ST, dan Yudhie Suherdani yang selalu bersedia untuk membantu dalam penyusunan Tugas Akhir ini.
9. Teman-teman sesama bimbingan dan teman-teman lain yang sangat berarti namun tidak mungkin disebutkan satu per satu.

Laporan ini dibuat dengan harapan dapat memberi manfaat bagi yang membutuhkannya. Namun demikian, penulis menyadari sepenuhnya bahwa kesempurnaan bukanlah milik manusia, oleh karena itu saran dan kritik akan sangat membantu untuk perbaikan di masa mendatang.

Bandung, Februari 2007

Penulis

DAFTAR ISI

| | |
|---|------|
| ABSTRAK | i |
| ABSTRACT | ii |
| KATA PENGANTAR | iii |
| DAFTAR ISI | v |
| DAFTAR TABEL | vii |
| DAFTAR GAMBAR | viii |
| BAB I PENDAHULUAN | |
| I.1 Latar Belakang | 1 |
| I.2 Identifikasi Masalah | 1 |
| I.3 Tujuan | 1 |
| I.4 Pembatasan Masalah | 1 |
| I.5 Sistematika Penulisan | 2 |
| BAB II DASAR TEORI | |
| II.1 Metode “String“ dan Boolean | 3 |
| II.2 Model Ruang Vektor | 4 |
| II.2.1 <i>Term</i> | 4 |
| II.2.2 <i>Dokumen Database</i> | 5 |
| II.2.3 <i>Query</i> | 5 |
| II.4 Kemiripan Dokumen | 5 |
| BAB III PERANCANGAN SISTEM Pencarian Dokumen | |
| DATA-TEKS MEGGUNAKAN MODEL RUANG VEKTOR | |
| III.1 Proses Input <i>Database</i> dan <i>Term</i> | 8 |
| III.2 Proses Input <i>Query</i> | 10 |
| III.3 Perhitungan Nilai Kemiripan | 11 |
| BAB IV DATA PENGAMATAN DAN ANALISA | |
| IV.1 Percobaan 1 | 13 |
| IV.2 Percobaan 2 | 15 |
| IV.3 Percobaan 3 | 17 |
| IV.4 Percobaan 4 | 19 |

| | | |
|---|---------------------------|-----|
| IV.5 | Percobaan 5 | 21 |
| IV.6 | Hasil Analisa Keseluruhan | 23 |
| BAB V KESIMPULAN DAN SARAN | | |
| V.1 | Kesimpulan | 24 |
| V.2 | Saran | 24 |
| DAFTAR PUSTAKA | | |
| LAMPIRAN A LISTING PROGRAM | | A-1 |
| LAMPIRAN B FILE DATABASE DAN QUERY | | B-1 |

DAFTAR TABEL

| | | |
|-------------|---|----|
| Tabel IV.1 | Frekuensi kemunculan term pada database, percobaan ke-1. | 14 |
| Tabel IV.2 | Frekuensi kemunculan <i>term</i> pada <i>query</i> , percobaan ke-1. | 14 |
| Tabel IV.3 | Frekuensi kemunculan term pada database, percobaan ke-2. | 16 |
| Tabel IV.4 | Frekuensi kemunculan term pada query, percobaan ke-2. | 16 |
| Tabel IV.5 | Frekuensi kemunculan <i>term</i> pada <i>database</i> , percobaan ke-3. | 18 |
| Tabel IV.6 | Frekuensi kemunculan <i>term</i> pada <i>query</i> , percobaan ke-3. | 18 |
| Tabel IV.7 | Frekuensi kemunculan <i>term</i> pada <i>database</i> , percobaan ke-4. | 19 |
| Tabel IV.8 | Frekuensi kemunculan <i>term</i> pada <i>query</i> , percobaan ke-4. | 19 |
| Tabel IV.9 | Frekuensi kemunculan <i>term</i> pada <i>database</i> , percobaan ke-5. | 20 |
| Tabel IV.10 | Frekuensi kemunculan <i>term</i> pada <i>query</i> , percobaan ke-5. | 20 |

DAFTAR GAMBAR

| | | |
|--------------|---|----|
| Gambar II.1 | Sudut θ adalah sudut antara vektor dokumen dan vektor <i>query</i> . | 6 |
| Gambar III.1 | Sistem pencarian dokumen data-teks sederhana. | 8 |
| Gambar III.2 | Langkah pembentukan matriks \bar{D} . | 9 |
| Gambar III.3 | Elemen matriks \bar{D} . | 10 |
| Gambar III.4 | Langkah pembentukan matriks | 11 |
| Gambar III.5 | Diagram Alir Perangkat Lunak Sistem Pencarian Informasi Data-Teks | 12 |
| Gambar IV.1 | Hasil nilai kemiripan pada percobaan ke-1. | 15 |
| Gambar IV.2 | Hasil nilai kemiripan pada percobaan ke-2. | 17 |
| Gambar IV.3 | Hasil nilai kemiripan pada percobaan ke-3 | 19 |
| Gambar IV.4 | Hasil nilai kemiripan pada percobaan ke-4 | 21 |
| Gambar IV.5 | Hasil nilai kemiripan pada percobaan ke-5 | 23 |