

ABSTRAK

Penelitian ini bertujuan untuk memudahkan pencarian dokumen-dokumen yang memiliki hubungan antar kata, bukan hanya pencarian pada judul dokumen saja akan tetapi dapat juga mencari dari isi dokumen tersebut. Dalam penerapannya penelitian ini menggunakan metode *Latent Semantic Indexing* (LSI) sebagai dasar metodenya. *Latent Semantic Indexing* (LSI) merupakan metode pengindeksan yang menggunakan teknik matematika yang disebut *Singular Value Decomposition* (SVD) untuk mengidentifikasi pola dalam hubungan antar istilah. *Singular Value Decomposition* (SVD) adalah faktorisasi matriks real atau matriks kompleks. Sebagai data untuk penelitian ini menggunakan data tugas akhir dari jurusan S1 Teknik Informatika Universitas Kristen Maranatha dari tahun 2003 hingga 2013. Uji coba dilakukan dengan menggunakan beberapa *query* yang berbeda dan dengan total dokumen yang dibaca antara 10 dokumen dan 20 dokumen. Setelah uji coba dilakukan, dapat dihasilkan pencarian dokumen yang relevan sesuai dengan kata kunci yang dimasukkan, dan pemberian rekomendasi kata berbeda antara pembacaan 10 dokumen dan 20 dokumen.

Kata Kunci : *Latent Semantic Indexing*, pencarian, dokumen, *Singular Value Decomposition*.

ABSTRACT

This study aims to facilitate the search for documents which have realtionships between words, not just a search on the title of the document but also from the content of the document. In application of this study using Latent Semantic Indexing (LSI) as the base method. Latent Semantic Indexing (LSI) is an indexing method that uses a mathematical technique called Singular Value Decomposition (SVD) to identify patters in the relationships between terms. Singular Value Decomposition (SVD) is a matrix factorization of the real or complex matrix. As the data for this study using data from the final assignment of S1 Computer Science Maranatha Christian Universty from 2003 to 2013. Trial was performed using several different queries and total documents to read between 10 documents and 20 documents. After the trial is done, relevant documents can be generated according to the entered keywords, and recommendations differ between 10 documents and 20 documents to read.

Keywords : Latent Semantic Indexing, searching, documents, Singular Value Decomposition.

DAFTAR ISI

LEMBAR PENGESAHAN	i
PERNYATAAN ORISINALITAS LAPORAN PENELITIAN	ii
PERNYATAAN PUBLIKASI LAPORAN PENELITIAN.....	iii
PRAKATA.....	iv
ABSTRAK.....	v
<i>ABSTRACT</i>	vi
DAFTAR ISI.....	vii
DAFTAR GAMBAR	x
DAFTAR TABEL	xi
DAFTAR NOTASI/ LAMBANG.....	xiii
DAFTAR SINGKATAN	xvi
DAFTAR RUMUS.....	xvii
DAFTAR KODE PROGRAM.....	xviii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	1
1.3 Tujuan	2
1.4 Batasan Masalah.....	2
1.5 Sistematika Penyajian	3
BAB II KAJIAN TEORI.....	4
2.1 <i>Latent Semantic Indexing</i> (LSI)	4
2.2 <i>Singular Value Decomposition</i> (SVD).....	4
2.3 <i>Stopwrods</i>	10
2.4 <i>Indri</i>	11
2.5 Total Reciprocal Rank.....	13
BAB III ANALISIS DAN DISAIN	14
3.1 Arsitektur Perangkat Lunak	14
3.2 <i>Use Case</i>	15
3.3 <i>Use Case Skenario</i>	16

3.3.1	<i>Use Case</i> Cari Dokumen Tugas Akhir	16
3.3.2	<i>Use Case</i> Rekomendasi.....	16
3.3.3	<i>Use Case</i> Indexing.....	17
3.3.4	<i>Use Case</i> Refinement Query	17
3.4	<i>Activity Diagram</i>	18
3.4.1	<i>Activity Diagram</i> Cari Dokumen Tugas Akhir	18
3.4.2	<i>Activity Diagram</i> Melihat Topik Terkait.....	19
3.4.3	<i>Activity Diagram</i> Indexing	19
3.4.4	<i>Activity Diagram</i> Refinement Query	20
3.5	<i>Entity Relationship Diagram</i> (ERD).....	21
3.6	Gambaran Keseluruhan	22
3.6.1	Persyaratan Antarmuka Eksternal	22
3.6.2	Antarmuka dengan Pengguna	22
3.6.3	Antarmuka Perangkat Keras	23
3.6.4	Antarmuka Perangkat Lunak.....	23
3.6.5	Fitur-fitur Produk Perangkat Lunak	23
3.6.6	Fitur Rekomendasi	26
3.7	Disain Antarmuka	28
3.7.1	Main Form.....	28
3.7.2	Form Configuration.....	29
BAB IV	PENGEMBANGAN PERANGKAT LUNAK.....	30
4.1	Implementasi Class/Modul.....	30
4.2	Implementasi Penyimpanan Data.....	31
4.3	Implementasi Antarmuka	31
4.3.1	Menu Utama.....	31
4.3.2	Menu Konfigurasi	32
4.4	Pseudocode.....	33
4.5	Penjelasan Arsitektur Perangkat Lunak	34
4.5.1	Ubah Dokumen dari bentuk PDF menjadi <i>file text</i>	34
4.5.2	Membuang <i>Stopwords</i>	35
4.5.3	Melakukan Latent Semantic Indexing (LSI).....	37
4.5.4	Menyimpan Hasil LSI	40

4.5.5	Searching.....	41
BAB V TESTING DAN EVALUASI SISTEM		45
5.1	Rencana Pengujian	45
5.1.1	Pengujian 1.....	45
5.1.2	Pengujian Waktu	46
5.1.3	Pengujian Query	47
5.1.4	Pengujian Rekomendasi	86
5.1.5	Total Reciprocal Rank.....	93
5.2	Pelaksanaan Pengujian	93
5.2.1	Pengujian Menu Utama.....	93
5.2.2	Pengujian Menu Konfigurasi	95
BAB VI SIMPULAN DAN SARAN.....		96
6.1	Kesimpulan	96
6.2	Saran.....	96
DAFTAR PUSTAKA		98

DAFTAR GAMBAR

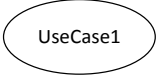
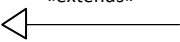
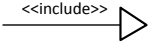
Gambar 2.1 Matriks A.....	6
Gambar 2.2 Menemukan Singular Value dan Eigenvalue	6
Gambar 2.3 Menghitung Singular Value	7
Gambar 2.4 Eigenvactor.....	7
Gambar 2.5 Normalisasi.....	7
Gambar 2.6 Matriks V.....	8
Gambar 2.7 Matriks U.....	8
Gambar 2.8 Reduksi SVD.....	9
Gambar 2.9 Vektor Query.....	10
Gambar 2.10 Cosinus Query	10
Gambar 2.11 Indexing Indri.....	11
Gambar 2.12 Searching Indri	12
Gambar 3.1 Arsitektur Perangkat Lunak	15
Gambar 3.2 <i>Use Case Diagram</i>	16
Gambar 3.3 <i>Activity Diagram</i> Cari Dokumen Tugas Akhir	18
Gambar 3.4 <i>Activity Diagram</i> Melihat Topik Terkait.....	19
Gambar 3.5 <i>Activity Diagram Indexing</i>	20
Gambar 3.6 Refinement Query	21
Gambar 3.7 <i>Entity Relationship Diagram (ERD)</i>	22
Gambar 3.8 Menu Utama.....	28
Gambar 3.9 Menu Konfiguration.....	29
Gambar 4.1 Class Diagram	30
Gambar 4.2 Implementasi Basis Data.....	31
Gambar 4.3 Menu Utama.....	31
Gambar 4.4 Menu Konfigurasi	32
Gambar 4.5 Arsitektur Perangkat Lunak	34
Gambar 4.6 Aplikasi PDFZilla 1.2	35
Gambar 5.1 Pengujian 1A.....	45
Gambar 5.2 Pengujian 1B	46

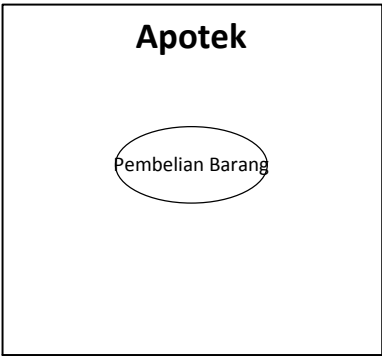
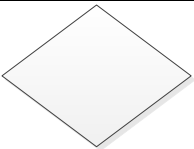
DAFTAR TABEL

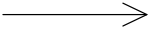
Tabel 2.1 Contoh TTR	13
Tabel 4.1 Contoh daftar stopwords	35
Tabel 5.1 Perbedaan waktu pembuatan indeks	46
Tabel 5.2 Perbandingan waktu pencarian	47
Tabel 5.3 Perbandingan Query 1.....	47
Tabel 5.4 Perbandingan Query 2.....	50
Tabel 5.5 Perbandingan Query 3.....	52
Tabel 5.6 Perbandingan Query 4.....	54
Tabel 5.7 Perbandingan Query 5.....	57
Tabel 5.8 Perbandingan Query 6.....	59
Tabel 5.9 Perbandingan Query 7.....	61
Tabel 5.10 Perbandingan Query 8.....	63
Tabel 5.11 Perbandingan Query 9.....	65
Tabel 5.12 Perbandingan Query 10.....	66
Tabel 5.13 Perbandingan Query 11.....	68
Tabel 5.14 Perbandingan Query 12.....	71
Tabel 5.15 Perbandingan Query 13.....	73
Tabel 5.16 Perbandingan Query 14.....	75
Tabel 5.17 Perbandingan Query 15.....	77
Tabel 5.18 Perbandingan Query 16.....	79
Tabel 5.19 Perbandingan Query 17.....	80
Tabel 5.20 Perbandingan Query 18.....	82
Tabel 5.21 Perbandingan Query 19.....	84
Tabel 5.22 Perbandingan Rekomendasi 1.....	86
Tabel 5.23 Perbandingan Rekomendasi 2.....	87
Tabel 5.24 Perbandingan Rekomendasi 3.....	87
Tabel 5.25 Perbandingan Rekomendasi 4.....	87
Tabel 5.26 Perbandingan Rekomendasi 5.....	88
Tabel 5.27 Perbandingan Rekomendasi 6.....	88
Tabel 5.28 Perbandingan Rekomendasi 7.....	88

Tabel 5.29 Perbandingan Rekomendasi 8.....	89
Tabel 5.30 Perbandingan Rekomendasi 9.....	89
Tabel 5.31 Perbandingan Rekomendasi 10.....	89
Tabel 5.32 Perbandingan Rekomendasi 11.....	90
Tabel 5.33 Perbandingan Rekomendasi 12.....	90
Tabel 5.34 Perbandingan Rekomendasi 13.....	90
Tabel 5.35 Perbandingan Rekomendasi 14.....	91
Tabel 5.36 Perbandingan Rekomendasi 15.....	91
Tabel 5.37 Perbandingan Rekomendasi 16.....	91
Tabel 5.38 Perbandingan Rekomendasi 17.....	92
Tabel 5.39 Perbandingan Rekomendasi 18.....	92
Tabel 5.40 Perbandingan Rekomendasi 19.....	92
Tabel 5.41 Hasil Total Reciprocal Rank.....	93
Tabel 5.42 Pengujian menu utama.....	93
Tabel 5.43 Pengujian menu konfigurasi.....	95

DAFTAR NOTASI/ LAMBANG

Jenis	Notasi/Lambang	Nama	Arti
UML		<i>Use Case</i>	Nama proses yang ada pada sistem
UML		<i>Actor</i>	Nama <i>actor</i> yang melakukan proses
UML		<i>Communication</i>	Hubungan antara proses (<i>use case</i>) dengan <i>actor</i>
UML		<i>Extends</i>	Pengembangan sebuah proses menjadi <i>use case</i> lain yang berbeda
UML		<i>Include</i>	Menandakan bahwa <i>use case</i> tersebut menggunakan <i>use case</i> yang lain

Jenis	Notasi/Lambang	Nama	Arti
UML		<i>System Boundary</i>	Kumpulan <i>use case</i> dalam sebuah sistem besar
ERD		Entitas	Bentuk dasar dari model data yang digunakan untuk melambangkan orang, tempat, benda, dll
ERD		Atribut	Deskripsi karakteristik dari sebuah entitas
ERD		Relasi	Hubungan antar entitas
ERD		<i>Communication</i>	Garis penghubung entitas dengan relasi
<i>Activity Diagram</i>		<i>Initial State</i>	Awal mula proses

Jenis	Notasi/Lambang	Nama	Arti
<i>Activity Diagram</i>		<i>Action State</i>	Proses yang dilakukan oleh sistem/pengguna
<i>Activity Diagram</i>		<i>Final State</i>	Akhir dari sebuah alur <i>activity</i>
<i>Activity Diagram</i>		<i>Control Flow</i>	Alur antar proses

DAFTAR SINGKATAN

LSI = *Latent Semantic Indexing*

SVD = *Singular Value Decomposition*

UML = *Unified Modeling Language*

ERD = *Entity Relationship Diagram*

DAFTAR RUMUS

Rumus 2.1 Rumus SVD	4
Rumus 2.2 Rumus reduksi dimensi sebesar k	9

DAFTAR KODE PROGRAM

Kode Program 4.1 Kode Program Stopwords.....	37
Kode Program 4.2 Kode Program Indexing.....	40
Kode Program 4.3 Save Index	41
Kode Program 4.4 Proses Searching.....	44