

***Query Expansion Pada Mesin Pencari Menggunakan Lexical  
Database Dan Vector Space Model***

*Diajukan Sebagai Syarat Untuk Menyelesaikan  
Pendidikan Program Strata-1 Pada Jurusan Teknik Informatika  
Fakultas Ilmu Komputer Universitas Sriwijaya*



Oleh:

Abdul Halim  
NIM : 09021381419090

**Jurusan Teknik Informatika  
FAKULTAS ILMU KOMPUTER UNIVERSITAS SRIWIJAYA  
2018**

## LEMBAR PENGESAHAN TUGAS AKHIR

*Query Expansion Pada Mesin Pencari Menggunakan Lexical Database  
Dan Vector Space Model!*

Oleh :

ABDUL HALIM  
NIM : 09021381419090

Palembang, 18 Desember 2019

Pembimbing I,



Novi Yusliani, M.T.  
NIP. 198211082012122001

Pembimbing II,



Yoppy Sazaki, M.T.  
NIPUS 19740602012101201

Mengetahui,  
Ketua Jurusan Teknik Informatika



Rifkie Primartha, M.T.  
NIP. 197706012009121004

## TANDA LULUS UJIAN SIDANG TUGAS AKHIR

Pada hari Jum'at, 16 November 2018 telah dilaksanakan ujian sidang tugas akhir oleh Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.

Nama : Abdul Halim  
NIM : 09021381419090  
Judul : *Query Expansion Pada Mesin Pencari Menggunakan Lexical Database Dan Vector Space Model*

1. Pembimbing I

Novi Yusliani, MT  
NIP. 198211082012122001

2. Pembimbing II

Yoppy Sazaki, M.T.  
NIPUS. 19740602012101201

3. Pengaji I

M. Fachrurrozi, MT  
NIP. 198005222008121002

4. Pengaji II

Mastura Diana Marieska, M.T  
NIP. 198603212018032001

Mengetahui,  
Ketua Jurusan Teknik Informatika,

Rifkie Primartha, M.T  
NIP. 197706012009121004

## **HALAMAN PERNYATAAN**

Yang bertanda tangan di bawah ini :

Nama : Abdul Halim  
NIM : 09021381419090  
Program Studi : Teknik Informatika Bilingual  
Judul Skripsi : *Query Expansion Pada Mesin Pencari Menggunakan Lexical Database Dan Vector Space Model*

Hasil Pengecekan Software *iThenticate/Turnitin* : **17%**

Menyatakan bahwa Laporan Projek saya merupakan hasil karya sendiri dan bukan hasil penjiplakan/plagiat. Apabila ditemukan unsur penjiplakan/plagiat dalam laporan projek ini, maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya sesuai dengan ketentuan yang berlaku.

Demikian, pernyataan ini saya buat dengan sebenarnya dan tidak ada paksaan oleh siapapun.

Palembang, 18 Desember 2018



(Abdul Halim)

NIM. 09021381419090

***QUERY EXPANSION PADA MESIN PENCARI MENGGUNAKAN LEXICAL  
DATABASE DAN VECTOR SPACE MODEL***

**Oleh :**  
**Abdul Halim**  
**09021381419090**

**ABSTRACT**

The rapid increase of information flow makes it easier for users to browse and search to get information quickly, relevantly, and according to the needs of users, but often users do not get information that suits their needs because users cannot express their information needs properly or search engines that produce less relevant information. This research creates a framework for building an information retrieval system that implements Query Expansion and Vector Space Models. The Expansion query is used so that the query term input from the user can be expanded by taking synonyms from Lexical Database so as to create new query terms that match the initial queries that aim to improve search results, and Vector Space Models as modeling calculations performed on query vectors and documents. The implementation of the Query Expansion on search engines using the Lexical Database and Vector Space Model successfully results in better recall and precision values, which are 10% -15% higher and 8.63% execution times shorter than search engines that do not implement the Query Expansion. Therefore, this research can help improve search engine capabilities in providing better search results.

**Keywords** : search engine, vector space model, Query Expansion, Lexical Database, information retrieval system, recall, precision

***QUERY EXPANSION PADA MESIN PENCARI MENGGUNAKAN LEXICAL  
DATABASE DAN VECTOR SPACE MODEL***

**Oleh :**  
**Abdul Halim**  
**09021381419090**

**ABSTRAK**

Peningkatan arus informasi yang sangat cepat mempermudah pengguna dalam *browsing* dan *searching* untuk mendapatkan informasi secara cepat, relevan, dan sesuai kebutuhan yang diinginkan pengguna, namun seringkali pengguna tidak mendapatkan informasi yang sesuai dengan kebutuhannya karena pengguna tidak dapat mengekspresikan kebutuhan informasinya dengan baik atau mesin pencari yang menghasilkan informasi yang kurang relevan. Penelitian ini membuat sebuah kerangka kerja untuk membangun sebuah *information retrieval system* yang mengimplementasikan *Query Expansion* dan *Vector Space Model*. *Query Expansion* digunakan agar *query term* masukan dari pengguna dapat diperluas dengan mengambil sinonim dari *Lexical Database* sehingga menciptakan *query term* baru yang sesuai dengan *query awal* yang bertujuan untuk meningkatkan hasil pencarian, dan *Vector Space Model* sebagai permodelan perhitungan yang dilakukan pada vektor *query* dan dokumen. Penelitian ini berhasil memberikan hasil nilai *recall* dan *precision* yang lebih baik, yaitu 10%-15% lebih tinggi dan waktu eksekusi yang lebih singkat 8.63% daripada mesin pencari yang tidak menerapkan *Query Expansion*. Oleh karena itu, penelitian ini dapat membantu dalam meningkatkan kemampuan mesin pencari dalam memberikan hasil pencarian yang lebih baik.

**Kata Kunci :** mesin pencari, *vector space model*, *Query Expansion*, *Lexical Database*, *information retrieval system*, *recall*, *precision*

## DAFTAR ISI

Halaman

<b>HALAMAN JUDUL .....</b>	i
<b>HALAMAN PENGESAHAN .....</b>	ii
<b>HALAMAN TANDA LULUS UJIAN SIDANG TUGAS AKHIR .....</b>	iii
<b>HALAMAN PERNYATAAN .....</b>	iv
<b>MOTTO DAN PERSEMBAHAN .....</b>	v
<b>ABSTRACT .....</b>	vi
<b>ABSTRAK .....</b>	vii
<b>KATA PENGANTAR .....</b>	viii
<b>DAFTAR ISI .....</b>	x
<b>DAFTAR TABEL .....</b>	xv
<b>DAFTAR GAMBAR .....</b>	xvii
<b>DAFTAR LAMPIRAN .....</b>	xix

### **BAB I PENDAHULUAN**

1.1 Pendahuluan .....	I-1
1.2 Latar Belakang .....	I-1
1.3 Rumusan Masalah .....	I-4
1.4 Tujuan Penelitian .....	I-4
1.5 Manfaat Penelitian .....	I-5
1.6 Batasan Masalah .....	I-5
1.7 Sistematika Penulisan .....	I-5
1.8 Kesimpulan .....	I-7

### **BAB II KAJIAN LITERATUR**

2.1 Pendahuluan .....	II-1
2.2 Landasan Teori .....	II-1
2.2.1 Mesin Pencari .....	II-1

2.2.2 <i>Crawler</i> .....	II-3
2.2.3 <i>Pre-Processing</i> .....	II-4
2.2.4 <i>Query Expansion</i> .....	II-5
2.2.5 <i>Lexical Database</i> .....	II-6
2.2.6 <i>TF-IDF</i> .....	II-6
2.2.7 <i>Vector Space Model</i> .....	II-8
2.2.8 <i>Cosine Similarity</i> .....	II-8
2.2.9 <i>Recall</i> dan <i>Precision</i> .....	II-8
2.2.9.1 <i>Recall</i> .....	II-8
2.2.9.2 <i>Precision</i> .....	II-8
2.2.10 <i>Rational Unified Process (RUP)</i> .....	II-8
2.3 Penelitian Lain yang Relevan .....	II-13
2.4 Kesimpulan .....	II-14

### **BAB III METODOLOGI PENELITIAN**

3.1 Pendahuluan .....	III-1
3.2 Unit Penelitian .....	III-1
3.3 Metode Pengumpulan Data .....	III-1
3.3.1 Jenis Data.....	III-1
3.3.2 Sumber Data .....	III-1
3.3.3 Teknik Pengumpulan Data .....	III-2
3.4 Tahapan Penelitian .....	III-2
3.4.1 Blok Diagram Tahapan Penelitian.....	III-3
3.4.2 Blok Diagram Skema Umum Perangkat Lunak .....	III-4
3.5 Metode Pengembangan Perangkat Lunak .....	III-6
3.6 Penjadwalan Penelitian.....	III-10
3.7 Kesimpulan.....	III-14

### **BAB IV PENGEMBANGAN PERANGKAT LUNAK**

4.1 Pendahuluan .....	IV-1
4.2 Fase Insepsi .....	IV-1

4.2.1 Permodelan Bisnis .....	IV-1
4.2.2 Kebutuhan Sistem.....	IV-2
4.2.2.1 Fitur Praproses.....	IV-3
4.2.2.2 Fitur <i>Query Expansion</i> .....	IV-3
4.2.2.3 Fitur <i>Vector Space Model</i> .....	IV-3
4.2.2.3 Fitur <i>Search Engine Result Page</i> .....	IV-3
4.2.3 Analisis dan Desain .....	IV-4
4.2.3.1 Analisis Perangkat Lunak.....	IV-4
4.2.3.3 Desain Perangkat Lunak.....	IV-6
4.3 Fase Elaborasi.....	IV-15
4.3.1 Permodelan Bisnis .....	IV-16
4.3.3.1 Perancangan Data .....	IV-16
4.3.3.2 Perancangan Antarmuka.....	IV-16
4.3.2 Diagram <i>Sequence</i> .....	IV-19
4.4 Fase Konstruksi .....	IV-22
4.4.1 Diagram Kelas .....	IV-22
4.4.2 Implementasi .....	IV-24
4.4.2.1 Implementasi Kelas .....	IV-24
4.4.2.2 Implementasi Antarmuka .....	IV-25
4.5 Fase Transisi .....	IV-28
4.5.1 Permodelan Bisnis .....	IV-28
4.5.2 Rencana Pengujian .....	IV-28
4.5.3.1 Rencana <i>Use Case</i> Melakukan Pencarian .....	IV-28
4.5.3.2 Rencana <i>Use Case</i> Menampilkan Detail Hasil Pencarian .....	IV-29
4.5.3.2 Rencana <i>Use Case</i> Menampilkan Halaman <i>About</i>	IV-30
4.5.3 Implementasi .....	IV-30
4.5.3.1 Pengujian <i>Use Case</i> Melakukan Pencarian .....	IV-30
4.5.3.2 Pengujian <i>Use Case</i> Menampilkan Detail Hasil Pencarian .....	IV-32
4.5.3.2 Pengujian <i>Use Case</i> Menampilkan Halaman	

<i>About</i> .....	IV-33
4.6 Kesimpulan .....	IV-34

## **BAB V HASIL DAN ANALISIS PENELITIAN**

5.1 Pendahuluan .....	V-1
5.2 Hasil Percobaan Penelitian .....	V-1
5.2.1 Skenario Pengujian Pertama.....	V-2
5.2.2 Skenario Pengujian Kedua .....	V-5
5.2.3 Skenario Pengujian Ketiga .....	V-8
5.2.4 Skenario Pengujian Keempat .....	V-11
5.2.5 Skenario Pengujian Kelima .....	V-14
5.2.6 Skenario Pengujian Keenam.....	V-17
5.3 Hasil Pengujian.....	V-20
5.4 Analisa Hasil Penelitian.....	V-22
5.5 Kesimpulan.....	V-23

## **BAB VI KESIMPULAN DAN SARAN**

6.1 Pendahuluan .....	VI-1
6.2 Kesimpulan .....	VI-1
6.3 Saran .....	VI-2

<b>DAFTAR PUSTAKA</b> .....	xx
<b>LAMPIRAN 1</b> .....	L1-1
<b>LAMPIRAN 2</b> .....	L2-1
<b>LAMPIRAN 3</b> .....	L3-1

## DAFTAR TABEL

	Halaman
Tabel III-1 Kegiatan Pengembangan Perangkat Lunak RUP .....	III-6
Tabel III-2 Jadwal Penelitian .....	III-11
Tabel IV-1 Definisi Aktor <i>Use Case</i> .....	IV-6
Tabel IV-2 Definisi <i>Use Case</i> .....	IV-7
Tabel IV-3 Skenario Melakukan Pencarian .....	IV-9
Tabel IV-4 Skenario <i>Use Case</i> Menampilkan Detail Hasil Pencarian .....	IV-10
Tabel IV-5 Skenario Use Case Menampilkan halaman <i>About</i> .....	IV-11
Tabel IV-6 Implementasi Kelas .....	IV-24
Tabel IV-7 Rencana Pengujian <i>Use Case</i> Melakukan Pencarian .....	IV-28
Tabel IV-8 Rencana Pengujian <i>Use Case</i> Menampilkan Detail Hasil Pencarian .....	IV-29
Tabel IV-9 Rencana Pengujian <i>Use Case</i> Melakukan Pendekripsi.....	IV-30
Tabel IV-10 Pengujian <i>Use Case</i> Melakukan Pencarian .....	IV-31
Tabel IV-11 Rencana Pengujian <i>Use Case</i> Menampilkan Detail Hasil Pencarian .....	IV-32
Tabel IV-12 Rencana Pengujian <i>Use Case</i> Menampilkan Halaman <i>About</i> .....	IV-33
Tabel V-1 Hasil percobaan nilai <i>Threshold</i> pada nilai <i>Cosine Similarity</i> pengujian pertama .....	V-3
Tabel V-2 Hasil pengujian <i>query</i> “gorengan” pada sistem QE.....	V-4
Tabel V-3 Hasil pengujian query “gorengan” pada sistem non-QE .....	V-5
Tabel V-4 Hasil percobaan nilai <i>Threshold</i> pada nilai <i>Cosine Similarity</i> pengujian kedua .....	V-6
Tabel V-5 Hasil pengujian <i>query</i> “sate” pada sistem QE .....	V-7
Tabel V-6 Hasil pengujian query “sate” pada sistem non-QE.....	V-8
Tabel V-7 Hasil percobaan nilai <i>Threshold</i> pada nilai Cosine Similarity pengujian ketiga .....	V-9
Tabel V-8 Hasil pengujian <i>query</i> “ketupat” pada sistem QE.....	V-10

Tabel V-8 Hasil pengujian <i>query</i> “ketupat” pada sistem QE.....	V-10
Tabel V-9 Hasil pengujian <i>query</i> “ketupat” pada sistem non-QE .....	V-11
Tabel V-10 Hasil percobaan nilai <i>Threshold</i> pada nilai <i>Cosine Similarity</i> pengujian kempat .....	V-12
Tabel V-11 Hasil pengujian query “resep membuat kue bolu” pada sistem QE .....	V-13
Tabel V-12 Hasil pengujian query “resep membuat kue bolu” pada sistem non-QE.....	V-14
Tabel V-13 Hasil percobaan nilai <i>Threshold</i> pada nilai <i>Cosine Similarity</i> pengujian kelima .....	V-15
Tabel V-14 . Hasil pengujian query “resep kue brownies” pada sistem QE .....	V-16
Tabel V-15 . Hasil pengujian query “resep kue brownies” pada sistem non-QE.....	V-17
Tabel V-16 Hasil percobaan nilai <i>Threshold</i> pada nilai <i>Cosine Similarity</i> pengujian keenam .....	V-19
Tabel V-17 Hasil pengujian <i>query</i> “resep opor enak” pada sistem QE .....	V-20
Tabel V-18 Hasil pengujian <i>query</i> “resep kue brownies” pada sistem non-QE	V-21

## DAFTAR GAMBAR

Halaman

Gambar II-1 Arsitektur dasar Mesin Pencari secara umum .....	II-2
Gambar II-2 Representasi dokumen dan <i>query</i> pada ruang vektor .....	II-8
Gambar II-3 Arsitektur <i>Rational Unified Process</i> .....	II-12
Gambar II-4 Arsitektur <i>Rational Unified Process</i> .....	II-35
Gambar III-1 Blok Diagram Tahapan Penelitian .....	III-3
Gambar III-2 Diagram Alur Proses Pengujian Data .....	III-4
Gambar III-3 Blok Diagram Skema Umum Perangkat Lunak .....	III-8
Gambar IV-1 Diagram <i>Use Case</i> .....	IV-6
Gambar IV-2 Diagram Aktivitas <i>Use Case</i> Melakukan Pencarian.....	IV-13
Gambar IV-3 Diagram Aktivitas <i>Use Case</i> Menampilkan Detail Hasil Pencarian.....	IV-14
Gambar IV-4 Diagram Aktivitas <i>Use Case</i> Menampilkan halaman <i>About</i> ....	IV-15
Gambar IV-5 Rancangan Antarmuka <i>Homepage</i> .....	IV-17
Gambar IV-6 Rancangan Antarmuka halaman Search Engine Result Page (SERP) .....	IV-17
Gambar IV-7 Rancangan Antarmuka halaman Detail Hasil Pencarian .....	IV-18
Gambar IV-8 Rancangan Antarmuka halaman <i>About</i> .....	IV-18
Gambar IV-9 Diagram <i>sequence</i> Melakukan Pencarian .....	IV-20
Gambar IV-10 Diagram <i>sequence</i> Menampilkan Detail Hasil Pencarian .....	IV-21
Gambar IV-11 Diagram <i>sequence</i> Menampilkan halaman <i>About</i> .....	IV-21
Gambar IV-12 Diagram Kelas Perangkat Lunak .....	IV-23
Gambar IV-13 Antarmuka <i>Homepage</i> .....	IV-26
Gambar IV-14 Antarmuka halaman <i>Search Engine Result Page</i> (SERP).....	IV-26
Gambar IV-15 Antarmuka halaman Detail Hasil Pencarian.....	IV-27
Gambar IV-16 Antarmuka halaman <i>About</i> .....	IV-27
Gambar V-1 Grafik perbandingan nilai recall pada sistem yang menggunakan query expansion dan tanpa query expansion.....	IV-22

Gambar V-2 Grafik perbandingan nilai precision pada sistem yang menggunakan query expansion dan tanpa query expansion..... IV-22

## **DAFTAR LAMPIRAN**

## **BAB I**

### **PENDAHULUAN**

#### **1.1 Pendahuluan**

Pada bab ini membahas latar belakang masalah, rumusan masalah, tujuan penelitian, manfaat penelitian, dan batasan masalah, serta menjelaskan secara umum tentang keseluruhan penelitian. Dimulai dengan pendahuluan yang menjelaskan tentang latar belakang masalah dimana metode yang digunakan dapat menyelesaikan kasus pada penelitian yang telah dilakukan sebelumnya. Setelah mengetahui permasalahan, maka dari permasalahan tersebut akan dilakukan penelitian untuk mengatasi permasalahan yang ada dan menghasilkan solusi dari permasalahan tersebut.

#### **1.2 Latar Belakang**

Menurut Davie and Weeks (1995) dalam Stolt (1997), tahun 1982 pertumbuhan informasi meningkat dua kali lipat setiap 5 tahun. Tahun 1988 diprediksi informasi meningkat dua kali lipat setiap 2,2 tahun dan tahun 1992 berubah lagi menjadi setiap 1,6 tahun, kecenderungan ini akan selalu berubah dan saat ini terjadi peningkatan informasi dua kali lipat setiap tahun. Peningkatan arus informasi yang sangat cepat mempermudah pengguna dalam *browsing* dan *searching* untuk mendapatkan informasi secara cepat, relevan, dan sesuai kebutuhan yang diinginkan (Irmawati, 2017). Namun sering kali pengguna

mencari sebuah artikel atau dokumen dengan memasukkan sebuah *keyword* dan pengguna pun tidak mendapatkan informasi yang sesuai dengan kebutuhannya. Menurut Diva dan Sony (2012), penyebab pengguna tidak mendapatkan informasi yang sesuai dengan kebutuhannya dari sebuah mesin pencari karena *query* masukan yang dimasukkan oleh pengguna tidak tepat dan pengguna tidak mampu untuk merepresentasikan kebutuhan informasi yang diinginkan ke dalam bentuk *query*.

Salah satu metode yang sering digunakan dalam mengukur relevansi pada sebuah sistem *information retrieval* adalah *Vector Space Model*. Dalam pengembangan metode tersebut salah satunya dapat dilakukan dengan cara melakukan perluasan *query* atau *Query Expansion*. *Query Expansion* atau perluasan *query* merupakan proses me-reformulasikan kembali *query* awal yang dimasukkan oleh pengguna dengan melakukan penambahan beberapa *term* atau kata pada *query* demi meningkatkan perfoma dalam proses pencarian kembali informasi. Dalam konteks mesin pencari berbasis web, hal ini termasuk evaluasi *input* pengguna dan memperluas *query* pencarian untuk mendapatkan dokumen yang cocok dengan *query* (Nugroho, 2009).

Metode *Query Expansion* dan *Vector Space Model* pada mesin pencari telah banyak diteliti oleh beberapa peneliti sebelumnya diantaranya adalah penelitian yang dilakukan oleh Broto dan Gunawan pada tahun 2015 yaitu menerapkan *query expansion* yang bertujuan untuk memperluas *query* pengguna demi menghasilkan daftar ayat-ayat yang relevan yang memiliki kesamaan topik dengan *query* pengguna. Broto dan Gunawan berhasil meningkatkan jumlah ayat-

ayat yang ditemukan dikarenakan adanya penambahan keyword pada *query* dan didapatkan nilai precision sebesar 75% pada *query* yang diekspansi. Penelitian lainnya yaitu *The Design of Lexical Database for Indonesian Language* oleh Gunawan, D. Dan A. Amalia pada tahun 2017. Pada penelitiannya mereka mengusulkan desain *lexical database* untuk bahasa indonesia berdasarkan dari KBBI (Kamus Besar Bahasa Indonesia) edisi 4, Kateglo, dan Struktur dari *Wordnet*. mereka menggunakan fitur yang ada di KBBI dan Kateglo sebagai sumber utama kata-kata bahasa Indonesia untuk membangun sebuah *lexical database*. Dari penelitian ini mereka berhasil menghasilkan sebuah pondasi struktur baru dalam membangun sebuah *lexical database* untuk bahasa indonesia. Penelitian lainnya yaitu *Query Expansion Dengan Menggabungkan Metode Ruang Vektor Dan Wordnet Pada Sistem Information Retrieval* oleh Susetyo Adi Nugroho pada tahun 2009. Pada penelitiannya Nugroho (2009) menggunakan *Query Expansion* untuk memperluas *query* masukan pengguna dengan mengambil sinonim dari *WordNet* dan menghasilkan kesimpulan bahwa *Query Expansion* berhasil meningkatkan jumlah dokumen yang diterima sistem.

Pada penelitian ini akan menerapkan *Query Expansion* pada mesin pencari menggunakan *Lexical Database* dan *Vector Space Model* yang diharapkan dapat mengatasi masalah dalam pencarian konten yang tepat dan relevan dengan *keyword* pencarian.

### 1.3 Rumusan Masalah

Berdasarkan uraian tentang latar belakang rumusan, masalah pada penelitian ini yaitu bagaimana menerapkan *Query Expansion* menggunakan *Lexical Database* dan metode *Vector Space Model* digunakan pada mesin pencari, maka *Research Question* dari penelitian ini adalah:

- a) Bagaimana mengimplementasikan dan cara kerja *Query Expansion* menggunakan *Lexical Database* dan metode *Vector Space Model* pada mesin pencari?
- b) Bagaimana hasil dari implementasi *Query Expansion* menggunakan *Lexical Database* dan metode *Vector Space Model* pada mesin pencari?

#### **1.4 Tujuan Penelitian**

Tujuan dari penelitian ini adalah:

- a) Menghasilkan mesin pencari yang menggunakan *Query Expansion* dengan *Lexical Database* dan *Vector Space Model*.
- b) Mengetahui nilai *recall* dan *precision*, dan waktu eksekusi yang dihasilkan oleh mesin pencari yang menggunakan *Query Expansion* dengan *Lexical Database* dan *Vector Space Model*.

#### **1.5 Manfaat Penelitian**

Manfaat dari penelitian ini adalah untuk mempelajari *Query Expansion* pada mesin pencari menggunakan *Lexical Database* dan *Vector Space Model*, dan

mempermudah pengguna mesin pencari dalam mencari artikel atau dokumen yang diperlukan, serta dapat digunakan sebagai referensi untuk penelitian selanjutnya.

### **1.6 Batasan Masalah**

Batasan-batasan yang ditetapkan dalam pengembangan perangkat lunak penelitian ini adalah:

- a) Contoh dokumen atau artikel yang digunakan dalam penelitian diambil dari mesin pencari *bing.com* dengan menggunakan aplikasi *scraper* dan mesin *crawler* sederhana. Dokumen yang digunakan berjumlah 150 dokumen yang berkategori masakan.
- b) *Query* yang dimasukkan ke sistem menggunakan Bahasa Indonesia.
- c) Bahasa pemrograman yang digunakan adalah PHP.

### **1.7 Sistematika Penulisan**

Sistematika penulisan tugas akhir ini mengikuti standar penulisan tugas akhir yang ditetapkan oleh Fakultas Ilmu Komputer Universitas Sriwijaya, yaitu sebagai berikut :

## **BAB I PENDAHULUAN**

Bab ini akan membahas tentang latar belakang, rumusan masalah, tujuan dan manfaat penelitian, batasan masalah atau ruang lingkup, metodologi penelitian, dan sistematika penulisan.

## **BAB II KAJIAN LITERATUR**

Bab ini membahas seluruh dasar-dasar teori yang digunakan mulai dari definisi sistem, informasi mengenai domain, dan semua yang digunakan pada tahapan analisis, perancangan, dan implementasi.

## **BAB III METODOLOGI PENELITIAN**

Bab ini membahas mengenai tahap-tahap yang akan diterapkan pada penelitian. Setiap rencana dari tahapan penelitian dideskripsikan secara rinci berdasarkan kerangka kerja. Dilanjutkan dengan perancangan manajemen proyek dalam pelaksanaan penelitian.

## **BAB IV PENGEMBANGAN PERANGKAT LUNAK**

Bab ini membahas perancangan dan lingkungan implementasi, berupa analisis dari masalah yang dihadapi dalam penelitian serta perancangan perangkat lunak.

## **BAB V HASIL DAN ANALISIS PENELITIAN**

Bab ini membahas implementasi dari hasil analisis dan perancangan yang sudah dilakukan sebelumnya. Hasil analisis berupa kesimpulan yang dapat

diambil dari penelitian. Melakukan pengujian perangkat lunak dan pengujian data penelitian

## **BAB VI KESIMPULAN DAN SARAN**

Bab ini berisi semua kesimpulan dari uraian-uraian yang telah dibahas sebelumnya, dan saran yang diharapkan dapat berguna untuk pengembangan penelitian lebih lanjut.

### **1.8 Kesimpulan**

Ada enam bab yang dibahas dalam penelitian ini. Bab 1 membahas gagasan dasar yang diajukan mengenai *Query Expansion* pada mesin pencari menggunakan *Lexical Database* dan metode *Vector Space Model*. Hal ini penting untuk memahami konsep dasar tentang apa yang akan dikerjakan. Latar belakang masalah dibahas pada bagian 1.2. Pernyataan masalah telah dijelaskan sehingga solusi untuk memecahkan masalah dapat diidentifikasi. Selain itu, ada dua tujuan yang ingin dicapai dalam mencapai tujuan penelitian. Batasan masalah pada penelitian ini juga diberikan dengan jelas. Pada bagian akhir, berisi alasan mengapa penting untuk melakukan penelitian ini.

Bab 2 akan membahas tinjauan dan kajian literatur yang berkaitan dengan penelitian. Bab 3 adalah keseluruhan metodologi penelitian untuk merinci

kerangka penelitian untuk mengembangkan sistem. Bab 4 akan melanjutkan tahap dari penelitian ini yaitu memberikan gambaran bagaimana perangkat lunak dikembangkan sehingga dapat menghasilkan hasil beserta analisisnya seperti yang akan dituliskan pada Bab 5. Pada bagian akhir yaitu Bab 6 menjadi kesimpulan dari penelitian yang diajukan dan juga menerima saran dari semua pihak terkait penelitian ini agar dapat diperbaiki dan dikembangkan lebih lanjut.

## **DAFTAR PUSTAKA**

- Alva, Pooja., Vinay Hegde. 2016. *Hidden Markov model for POS tagging in Word Sense Disambiguation*, International Conference on Computational Systems and Information Systems for Sustainable Solutions.
- Imbar, Radiant Victor., Adelia, Mewati Ayub, Alexander Rehatta. 2014 *Implementasi Cosine Similarity dan Algoritma Smith-Waterman untuk Mendeteksi Kemiripan Teks*. Jurnal Informatika, Vol. 10 No. 1, Juni 2014: 31 – 42.
- Karmayasa, O., Mahendra, I. B.. 2012. *Implementasi Vector Space Model dan Beberapa Notasi Metode Term Frequency Inverse Document Frequency (TFIDF) pada Sistem Temu Kembali Informasi*. Jurnal Elektronik Ilmu Komputer Universitas Udayana, Vol. 1, No. 1
- Khan, Javed Ahmad. (2014) *Comparative Study of Information Retrieval Models Used in Search Engine*. IEEE International Conference on Advances in Engineering & Technology Research (ICAETR - 2014), August 01-02, 2014.
- Liddy, E.D. 2001. *Natural Language Processing*, Encyclopedia of Library and Information Science. New York:Marcel Decker, Inc.
- Lv, Lin-Tao., Li-Ping Chen and Hong-Fang Zhou. 2008. *An Improved Topic Relevance Algorithm For Vertical Search Engines*. Proceedings of the 2008 International Conference on Wavelet Analysis and Pattern Recognition, Hong Kong, 30-31 Aug.

Nugroho, S. A. (2009). Query Expansion Dengan Menggabungkan Metode Ruang Vektor Dan Wordnet Pada Sistem Information Retrieval. *4 JURNAL INFORMATIKA, VOLUME 5 NOMOR 1, APRIL 2009*

Qiu, Y. And Rfe, HP.1993. *Concept-based query expansion*. SIGIR '93, page 160-169.

Saadah, Munjiah Nur., Rigga Widar Atmagi., Dyah S. Rahayu., Agus Zainal Arifin. 2013 *Sistem Temu Kembali Dokumen Teks dengan Pembobotan Tf-Idf Dan LCS*. Jurnal Ilmiah Teknologi Informasi, Volume 11, Nomor 1, Januari 2013 : 17 – 20.

Stolt, Hakan. Agents, 1997. *Filter and Search Engines : An evaluating survey on technologies for effective search for information from internet resources*. Department of Computing Science. Umea University. Graduation Thesis.  
Nugroho, S. A. (2009). Query Expansion Dengan Menggabungkan Metode Ruang Vektor Dan Wordnet Pada Sistem Information Retrieval. *4 JURNAL INFORMATIKA, VOLUME 5 NOMOR 1, APRIL 2009*

Usharani, J., Iyakutti, K.. (2013). *A Genetic Algorithm Based on Cosine Similrity for Relevant Document Retrieval*. International Journal of Engineering Research & Technology (IJERT), Vol. 2, Issue 2.

Wibowo, Adi., Andreas Handojo and Albert Halim. (2011). *Application of Topic Based Vector Space Model with WordNet*. International Conference on Uncertainty Reasoning and Knowledge Engineering. 2011.