

**VISUALISASI VECTOR SPACE MODEL
PADA PENCARIAN BERITA-BERITA KEIMIGRASIAN
DI WILAYAH NUSA TENGGARA TIMUR
(STUDI KASUS : NUSA TENGGARA TIMUR)**

TUGAS AKHIR

NO.662/WM.FT.H6/T.ILKOM/TA/2019

**Diajukan Untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer**



Disusun Oleh :

JUVENTINHO JOSE M. DE ARAUJO
231 14 028

**PROGRAM STUDI ILMU KOMPUTER
FAKULTAS TEKNIK
UNIVERSITAS KATOLIK WIDYA MANDIRA
KUPANG
2019**

HALAMAN PERSETUJUAN

TUGAS AKHIR

NO.662/WM.FT.H6/T.ILKOM/TA/2019

VISUALISASI *VECTOR SPACE MODEL*
PADA PENCARIAN BERITA-BERITA KEIMIGRASIAN
DI WILAYAH NUSA TENGGARA TIMUR
(STUDI KASUS : NUSA TENGGARA TIMUR)

OLEH :

JUVENTINHO JOSE MARIA DE ARAUJO

NO. REG 231 14 028

PENGUJI I

Paskah A. Nani, ST. MT

PENGUJI II

Sisilia D. B. Mau, S.Kom. M.T

PENGUJI III

Natalia M. R. Mamulak, ST. MM

KETUA PELAKSANA

Natalia M. R. Mamulak, ST. MM

SEKRETARIS PELAKSANA

Alfry A. J. SinlaE, S.Kom. M.Cs

HALAMAN PENGESAHAN

TUGAS AKHIR

NO.662/WM.FT.H6/T.ILKOM/TA/2019

VISUALISASI *VECTOR SPACE MODEL*
PADA PENCARIAN BERITA-BERITA KEIMIGRASIAN
DI WILAYAH NUSA TENGGARA TIMUR
(STUDI KASUS : NUSA TENGGARA TIMUR)

OLEH :

JUVENTINHO JOSE MARIA DE ARAUJO

NO. REG 231 14 028

DIPERIKSA/DISETUJUI OLEH PEMBIMBING

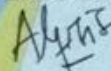
Di : Kupang

Tanggal : Februari 2019

DOSEN PEMBIMBING I :

DOSEN PEMBIMBING II :





Natalia M. R. Mamulak, ST. MM

Alfrv A. J. SinlaE, S.Kom. M.Cs

MENGETAHUI

KETUA PROGRAM STUDI
ILMU KOMPUTER
UNIKA WIDYA MANDIRA
KUPANG


Emiliana M. Meolbatak, ST. MT

MENGESAHKAN,

DEKAN FAKULTAS TEKNIK
UNIKA WIDYA MANDIRA
KUPANG


Patrisius Batarius, ST. MT

HALAMAN PERSEMBAHAN

Skripsi ini ku persembahkan untuk

TUHAN YESUS DAN BUNDA MARIA

Karena atas berkat dan rahmatnya saya masih diberikan nafas kehidupan untuk menikmati hidup yang indah ini

BAPA DAN MAMA, KAKAK DAN ADIK SERTA KELUARGA

Yang telah memberikan doa, dukungan dan motivasi

TEMAN – TEMAN INFORMATIKA ANGKATAN 2014

BAPAK DAN IBU DOSEN SERTA PEGAWAI TEKNIK INFORMATIKA

KALIAN SEMUA YANG SERING BERTANYA KAPAN WISUDA

TERIMA KASIH ATAS DOA KALIAN SEMUA

MOTTO

GOOD – GOD = 0

“Segala hal yang baik jika dilakukan tanpa mengandalkan Tuhan, hasilnya akan sia-sia”

“Selalu andalkan Tuhan dalam melakukan sesuatu”

PERNYATAAN ORISINALITAS

Yang bertanda tangan dibawah ini :

Nama : Juventinho Jose M. De Araujo

No. Registrasi : 231 14 028

Fakultas/Program Studi : Teknik/Ilmu Komputer

Dengan ini menyatakan bahwa karya tulis (Skripsi) dengan judul :

**VISUALISASI *VECTOR SPACE MODEL* PADA PENCARIAN BERITA-
BERITA KEIMIGRASIAN DI WILAYAH NUSA TENGGARA TIMUR**

Adalah benar-benar karya saya sendiri. Apabila di kemudian hari ditemukan penyimpangan, maka saya bersedia dituntut secara hukum.

Kupang, 14 Februari 2019



Juventinho J. M. De Araujo

KATA PENGANTAR

Puji syukur penulis sampaikan kepada Tuhan Yang Maha Esa atas berkat, kemampuan, dan kesehatan yang Tuhan berikan sehingga penulis dapat menyelesaikan Tugas Akhir ini.

Untuk selanjutnya penulis mengucapkan terima kasih kepada pihak-pihak yang telah membantu dalam penyelesaian Tugas Akhir ini, yaitu :

1. Tuhan Yesus yang selalu memberikan kesehatan bagi penulis sehingga dapat menyelesaikan tulisan ini.
2. Pater Dr. Philipus Tulle, SVD, selaku Rektor Universitas Katolik Widya Mandira Kupang.
3. Bapak Patrisius Batarius, ST. MT. selaku Dekan Fakultas Teknik Universitas Katolik Widya Mandira Kupang.
4. Ibu Emiliana Meolbatak, ST.MT. selaku Ketua Program Studi Ilmu Komputer.
5. Ibu Natalia M. R. Mamulak, ST. MM. selaku dosen pembimbing I yang telah bersedia mengorbankan waktu, tenaga dan pikiran dalam membimbing penulis dalam penyelesaian tulisan ini.
6. Bapak Alfry A. J. SinlaE, S.Kom. M.Cs. selaku dosen pembimbing II yang telah bersedia mengorbankan waktu, tenaga dan pikiran dalam membimbing penulis dalam penyelesaian tulisan ini.
7. Bapak Paskalis A. Nani, ST. MT selaku dosen penguji I.
8. Ibu Sisilia D. B. Mau, S.Kom. MT selaku dosen penguji II.
9. Bapak dan Ibu Dosen serta Pegawai Program Studi Ilmu Komputer Universitas Katolik Widya Mandira Kupang.
10. Kedua orang tua dan keluarga yang telah memberikan dukungan selama penyelesaian tulisan ini.

11. Teman-teman angkatan 2014 yang telah memberikan dukungan kepada penulis dalam penulisan Tugas Akhir ini.

Penulis menyadari bahwa tulisan ini masih jauh dari kesempurnaan, oleh karena itu penulis mengharapkan kritik dan saran yang membangun demi penyusunan laporan yang lebih baik. Akhir kata, penulis berharap kiranya tulisan ini dapat berguna bagi para pembaca.

Kupang, Februari 2019

Penulis,

DAFTAR ISI

HALAMAN JUDUL	i
<u>HALAMAN PERSETUJUAN</u>	ii
<u>HALAMAN PENGESAHAN</u>	iii
<u>HALAMAN PERSEMBAHAN</u>	iv
<u>MOTTO</u>	v
<u>PERNYATAAN ORISINALITAS</u>	vi
KATA PENGANTAR	vii
DAFTAR ISI	ix
DAFTAR GAMBAR	xi
DAFTAR TABEL	xii
<u>ABSTRAK</u>	xiii
ABSTRACT	xiv
BAB I PENDAHULUAN	
1.1. Latar Belakang	1
1.2. Rumusan Masalah	2
1.3. Tujuan Penelitian	2
1.4. Manfaat Penelitian	2
1.5. Batasan Masalah	2
1.6. Metodologi Penelitian	3
1.6.1. Analisis	3
1.6.2. Desain	4
1.6.3. Pengodean	4
1.6.4. Pengujian	4
1.6.5. <i>Maintenance</i>	4
1.7. Sistematika Penulisan	4
BAB II LANDASAN TEORI	
2.1. Penelitian Terdahulu	6
2.2. <i>Information Retrieval</i>	7
2.3. <i>Vector Space Model</i>	8
2.4. <i>Indexing</i>	10

2.5. Pembobotan tf-idf	11
BAB III PERANCANGAN SISTEM	
3.1. Rancangan Proses <i>Vector Space Model</i>	13
3.2. Rancangan Antarmuka Aplikasi	14
3.2.1. Rancangan Antarmuka Tampilan <i>Input Data</i>	14
3.2.2. Rancangan Antarmuka Tampilan Menu Utama	14
3.2.3. Rancangan Antarmuka Tampilan Indeks	15
3.2.4. Rancangan Antarmuka Tampilan Panjang Vektor	15
3.2.5. Rancangan Antarmuka Tampilan Lihat Dokumen	16
3.2.6. Rancangan Antarmuka Tampilan <i>Edit</i> Berita	16
3.2.7. Rancangan Antarmuka Tampilan Lihat Cache	17
3.3. Rancangan <i>Database</i>	17
3.3.1. Tabel Indeks	17
3.3.2. Tabel dokumen	18
3.3.3. Tabel vektor	18
3.3.4. Tabel <i>cache</i>	18
BAB IV IMPLEMENTASI SISTEM	
4.1. Implementasi Sistem	19
BAB V PENGUJIAN DAN ANALISIS HASIL	
5.1. Pengujian	22
5.2. Analisis Hasil	66
5.3. Visualisasi Data	66
BAB VI PENUTUP	
6.1. Kesimpulan	68
6.2. Saran Pengembangan	68
DAFTAR PUSTAKA	

DAFTAR GAMBAR

Gambar 1.1. Ilustrasi Model <i>Waterfall</i>	3
Gambar 2.1. Representasi Model Ruang Vektor	9
Gambar 2.2. Representasi Nilai Kata Dalam Dokumen.....	10
Gambar 3.1. <i>Flowchart</i> Dari Algoritma Vector Space Model	13
Gambar 3.2. Rancangan Antarmuka Tampilan <i>Input Data</i>	14
Gambar 3.3. Rancangan Antarmuka Tampilan Menu Utama	14
Gambar 3.4. Rancangan Antarmuka Tampilan Indeks	15
Gambar 3.5. Rancangan Antarmuka Tampilan Panjang Vektor	15
Gambar 3.6. Rancangan Antarmuka Tampilan Lihat Dokumen	16
Gambar 3.7. Rancangan Antarmuka Tampilan <i>Edit Berita</i>	16
Gambar 3.8. Rancangan Antarmuka Tampilan Lihat <i>Cache</i>	17
Gambar 4.1. Hasil Pencarian Terhadap <i>Query</i>	19
Gambar 4.2. Hasil Pencarian Terhadap <i>Query</i>	19
Gambar 4.3. Hasil Pencarian Terhadap <i>Query</i>	20
Gambar 4.4. Hasil Pencarian Terhadap <i>Query</i>	20
Gambar 5.1. Representasi Hasil Pencarian <i>Query</i>	67

DAFTAR TABEL

Tabel 4.1. Hasil Perhitungan Sistem	21
Tabel 5.1. Hasil <i>Index</i>	35
Tabel 5.2. Panjang Vektor	61
Tabel 5.3. <i>Index</i> Untuk <i>Query</i>	63
Tabel 5.4. Hasil Perhitungan Kemiripan Tiap Dokumen.....	65
Tabel 5.5. Perbandingan Hasil Perhitungan Manual Dengan Sistem.....	66

ABSTRAK

Kantor Imigrasi merupakan salah satu kantor pelayanan publik yang memiliki banyak kegiatan yang dilakukan. Kegiatan tersebut sering diliput dan dipublikasikan dalam bentuk tulisan berita. Berita-berita yang sudah dipublikasikan sangat banyak sehingga menyulitkan dalam mencari berita yang diinginkan. Salah satu cara untuk melakukan pencarian adalah dengan cara *ranking*. *Ranking* merupakan salah satu cabang ilmu dari *information retrieval*. *Information retrieval* yaitu menemukan materi (biasanya dokumen) dari sebuah kumpulan data yang tidak terstruktur (biasanya teks) untuk memenuhi kebutuhan informasi dari koleksi yang besar. Salah satu metode pencarian dokumen adalah dengan menggunakan *Vector Space Model* (VSM). VSM menggunakan konsep yang terdapat pada aljabar linear yaitu ruang vektor. Berdasarkan konsep yang digunakan tersebut, dikembangkan pemodelan dokumen untuk melakukan pencarian terhadap dokumen yang dibutuhkan. Representasi vektor kolom dimanfaatkan dalam pengkonversian dokumen *input*. Konsep lain yang digunakan adalah dengan mengetahui kedekatan antara dua buah vektor, yaitu dengan cara menghitung besarnya sudut yang terbentuk antara dua vektor dan kemudian diurutkan dari data yang memiliki besar sudut yang terkecil hingga yang terbesar yang menandakan urutan data hasil *ranking* dari yang paling relevan hingga tidak relevan. Salah satu algoritma pembobotannya adalah algoritma *tf-idf* yang dipengaruhi oleh frekuensi kemunculan kata pada tiap dokumen *online* dan frekuensi dari dokumen *online* yang memiliki kata tersebut. Pada penelitian ini dijelaskan tentang visualisasi VSM dalam pencarian dokumen pada berita-berita tentang keimigrasian wilayah Nusa Tenggara Timur.

Kata kunci : Imigrasi, *Information Retrieval*, TF-IDF, *Vector Space Model*, Visualisasi.

ABSTRACT

The Immigration Office is one of the public service offices that has many activities carried out. These activities are often covered and published in the form of news articles. The news that has been published is very much making it difficult to find the desired news. One way to search is by ranking. Ranking is one branch of knowledge from information retrieval. Information retrieval is to find material (usually documents) from a collection of unstructured data (usually text) to meet the information needs of a large collection. One method of searching documents is to use Vector Space Model (VSM). VSM uses a concept found in linear algebra, namely vector space. Based on the concept used, document modeling was developed to search the required documents. Column vector representation is used in converting input documents. Another concept used is to know the closeness between two vectors, namely by calculating the amount of angle formed between two vectors and then sorted from data that has the smallest to the largest angles that indicate the order of ranking data from the most relevant to not relevant. One of the weighting algorithms is the tf-idf algorithm which is influenced by the frequency of occurrence of words in each online document and the frequency of online documents that have the word. In this study explained the visualization of VSM in searching documents on news about immigration in the East Nusa Tenggara region.

Keyword : Imigration, Information Retrieval, TF-IDF, Vector Space Model, Visualization.