



**Sistem Rekomendasi Lowongan Pekerjaan pada Perusahaan Outsourcing  
Menggunakan Algoritma Term Frequency Inverse Document Frequency dan  
Cosine Similarity**

*TUGAS AKHIR*

Imam Farisi  
41519110030

UNIVERSITAS  
MERCU BUANA  
PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS MERCU BUANA  
JAKARTA  
2021



**Sistem Rekomendasi Lowongan Pekerjaan pada Perusahaan Outsourcing  
Menggunakan Algoritma Term Frequency Inverse Document Frequency dan  
Cosine Similarity**

*Tugas Akhir*

Diajukan Untuk Melengkapi Salah Satu Syarat  
Memperoleh Gelar Sarjana Komputer

Oleh:

Imam Farisi  
41519110030

UNIVERSITAS  
PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS MERCU BUANA  
JAKARTA  
2021

## LEMBAR PERNYATAAN ORISINALITAS

Yang bertanda tangan dibawah ini:

NIM : 41519110030  
Nama : Imam Farisi  
Judul Tugas Akhir : Sistem Rekomendasi Lowongan Pekerjaan pada Perusahaan Outsourcing Menggunakan Algoritma Term Frequency Inverse Document Frequency dan Cosine Similarity.

Menyatakan bahwa Laporan Tugas Akhir saya adalah hasil karya sendiri dan bukan plagiat. Apabila ternyata ditemukan didalam laporan Tugas Akhir saya terdapat unsur plagiat, maka saya siap untuk mendapatkan sanksi akademik yang terkait dengan hal tersebut.

Jakarta, 21 Agustus 2021

  
Imam Farisi



## SURAT PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR

Sebagai mahasiswa Universitas Mercu Buana, saya yang bertanda tangan di bawah ini :

Nama Mahasiswa : Imam Farisi  
NIM : 41519110030  
Judul Tugas Akhir : Sistem Rekomendasi Lowongan Pekerjaan pada Perusahaan Outsourcing Menggunakan Algoritma Term Frequency Inverse Document dan Cosine Similarity.

Dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Noneksklusif** (*None-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul diatas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti/Noneksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat dan mempublikasikan tugas akhir saya.

Selain itu, demi pengembangan ilmu pengetahuan di lingkungan Universitas Mercu Buana, saya memberikan izin kepada Peneliti di Lab Riset Fakultas Ilmu Komputer, Universitas Mercu Buana untuk menggunakan dan mengembangkan hasil riset yang ada dalam tugas akhir untuk kepentingan riset dan publikasi selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 21 Agustus 2021



Imam Farisi

UNIVERSITAS  
MERCU BUANA

**SURAT PERNYATAAN LUARAN TUGAS AKHIR**

Sebagai mahasiswa Universitas Mercu Buana, saya yang bertanda tangan di bawah ini :

Nama Mahasiswa : Imam Farisi  
 NIM : 41519110030  
 Judul Tugas Akhir : Sistem Rekomendasi Lowongan Pekerjaan pada Perusahaan Outsourcing Menggunakan Algoritma Term Frequency Inverse Document Frequency dan Cosine Similarity.

Menyatakan bahwa :

1. Luaran Tugas Akhir saya adalah sebagai berikut :

No	Luaran	Jenis		Status	
1	Publikasi Ilmiah	Jurnal Nasional Tidak Terakreditasi			Diajukan
		Jurnal Nasional Terakreditasi		✓	
		Jurnal International Tidak Bereputasi			Diterima
		Jurnal International Bereputasi			
Disubmit/dipublikasikan di :	Nama Jurnal	Matrik : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer			
	ISSN	2476-9843			
	Link Jurnal	<a href="https://journal.universitasbumigora.ac.id/index.php/matrik">https://journal.universitasbumigora.ac.id/index.php/matrik</a>			
	Link File Jurnal Jika Sudah di Publish				

2. Bersedia untuk menyelesaikan seluruh proses publikasi artikel mulai dari submit, revisi artikel sampai dengan dinyatakan dapat diterbitkan pada jurnal yang dituju.
3. Diminta untuk melampirkan scan KTP dan Surat Pernyataan (Lihat Lampiran Dokumen HKI), untuk kepentingan pendaftaran HKI apabila diperlukan

Demikian pernyataan ini saya buat dengan sebenarnya. S

**MERCU BUANA**

Jakarta, 21 Agustus 2021



Imam Farisi

## LEMBAR PERSETUJUAN PENGUJI

NIM : 41519110030  
Nama : Imam Farisi  
Judul Tugas Akhir : Sistem Rekomendasi Lowongan Pekerjaan pada Perusahaan Outsourcing Menggunakan Algoritma Term Frequency Inverse Document Frequency dan Cosine Similarity.

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 28 Juli 2021



## LEMBAR PERSETUJUAN PENGUJI

NIM : 41519110030  
Nama : Imam Farisi  
Judul Tugas Akhir : Sistem Rekomendasi Lowongan Pekerjaan pada Perusahaan Outsourcing Menggunakan Algoritma Term Frequency Inverse Document Frequency dan Cosine Similarity.

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 28 Juli 2021

  
(Hery Derajad Wijaya, S.Kom, MM)

UNIVERSITAS  
MERCU BUANA

## LEMBAR PERSETUJUAN PENGUJI

NIM : 41519110030  
Nama : Imam Farisi  
Judul Tugas Akhir : Sistem Rekomendasi Lowongan Pekerjaan pada Perusahaan Outsourcing Menggunakan Algoritma Term Frequency Inverse Document Frequency dan Cosine Similarity.

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 28 Juli 2021



(Wawan Sunaryan, S.Kom, MT)

UNIVERSITAS  
MERCU BUANA

## LEMBAR PENGESAHAN

NIM : 41519110030  
Nama : Imam Farisi  
Judul Tugas Akhir : Sistem Rekomendasi Lowongan Pekerjaan pada Perusahaan Outsourcing Menggunakan Algoritma Term Frequency Inverse Document Frequency dan Cosine Similarity.

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 28 Juli 2021

Menyetujui,

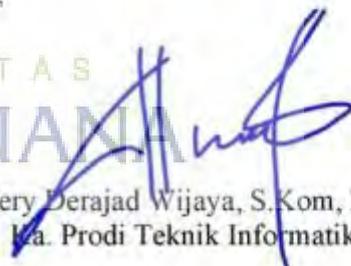


(Dr. Ir. Eliyani)  
Dosen Pembimbing

Mengetahui,



(Wawan Gunawan, S.Kom, MT)  
Koord Tugas Akhir Teknik Informatika



(Hery Derajad Wijaya, S.Kom, MM)  
Ka. Prodi Teknik Informatika

## ABSTRAK

Nama : Imam Farisi  
NIM : 41519110030  
Pembimbing TA : Dr. Ir. Eliyani  
Judul : Sistem Rekomendasi Lowongan Pekerjaan pada Perusahaan Outsourcing Menggunakan Algoritma Term Frequency Inverse Document Frequency dan Cosine Similarity.

Perusahaan outsourcing menjadi salah satu alternatif yang bisa dimanfaatkan oleh pencari kerja dalam mendapatkan sebuah pekerjaan. Pekerjaan yang didapatkan akan disesuaikan berdasarkan daftar riwayat hidup (CV) oleh bagian human resource development (HRD). Penelitian ini bertujuan untuk membantu bagian HRD dalam memetakan antara daftar riwayat hidup pencari kerja dengan perusahaan yang sesuai agar tidak membacanya secara manual. Sistem yang akan dibangun dalam penelitian ini menggunakan algoritma TF-IDF untuk pembobotan kata dan metode cosine similarity untuk mendapatkan kesesuaian antara daftar riwayat hidup dengan perusahaan. Data yang diolah oleh sistem berbentuk file pdf dengan kerangka yang sudah ditentukan. Kerangka tersebut terdiri dari tiga bagian yaitu pendidikan, kemampuan dan pengalaman dari pencari kerja. Hasil uji coba dengan pengujian confusion matrix didapatkan akurasi sebesar 76%.

Kata kunci:

sistem rekomendasi, lowongan pekerjaan, tf-idf, cosine similarity, text mining

UNIVERSITAS  
MERCU BUANA

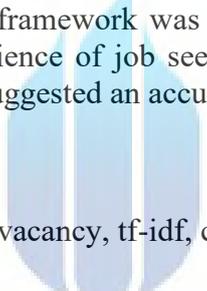
## ABSTRACT

Name : Imam Farisi  
Student Number : 41519110030  
Counsellor : Dr. Ir. Eliyani  
Title : Job Vacancies Recommendation System in Outsourcing Companies using Term Frequency Inverse Document Frequency and Cosine Similarity.

Outsourcing company is considered as one of the alternatives which may be used by job seekers for the purpose of finding any employment. The work obtained shall be adjusted by human resources development (HRD) based on curriculum vitae. This research was aimed to help HRD department with the mapping between job seeker's curriculum vitae and relevant companies so as to avoid manual reading. A system was established in this study by using TF-IDF algorithm for terms weighting and cosine similarity method to find out any conformity between curriculum vitae and the company. The data was processed by a system in the form of pdf file with a specified framework. The framework was composed of three sections namely education, ability, and experience of job seekers. The result of a test conducted using confusion matrix test suggested an accuracy of 76%.

Key words:

recommendation system, job vacancy, tf-idf, cosine similarity, text mining



UNIVERSITAS  
MERCU BUANA

## KATA PENGANTAR

Puji syukur kita panjatkan ke hadirat Allah subhanahu wa ta'ala yang telah melimpahkan kasih dan sayang-Nya sehingga penulis bisa menyelesaikan tugas akhir dengan tepat waktu yang berjudul “Sistem Rekomendasi Lowongan Pekerjaan pada Perusahaan Outsourcing Menggunakan Algoritma Term Frequency Inverse Document Frequency dan Cosine Similarity”. Tujuan dari penyusunan tugas akhir ini guna memenuhi salah satu syarat untuk bisa menempuh ujian sarjana komputer pada Fakultas Ilmu Komputer Program Studi Teknik Informatika di Universitas Mercu Buana. Penulis menyadari bahwa tanpa bantuan dan bimbingan dari Ibu Dr. Ir. Eliyani penulis akan kesusahan dalam mengerjakan tugas akhir ini. Oleh karena itu, penulis mengucapkan terima kasih kepada :

1. Ibu Dr. Ir. Eliyani, selaku Dosen Pembimbing Tugas Akhir Program Studi Informatika Universitas Mercu Buana.
2. Bapak Hery Derajad Wijaya, S.Kom, MM., selaku Kepala Program Studi Informatika, Universitas Mercu Buana.
3. Bapak Wawan Gunawan, S.Kom, M.T., selaku Koordinator Tugas Akhir Program Studi Informatika Universitas Mercu Buana
4. Kedua orang tua dan keluarga yang selalu memberikan doa, semangat dan dukungan kepada penulis
5. Teman-teman seperjuangan yang mau berbagi ilmu dan informasi terkait pengerjaan tugas akhir.
6. Beserta semua pihak yang telah memotivasi dan ikut memberikan bantuannya kepada penulis yang namanya tidak dapat penulis sebutkan satu per satu.

Akhir kata, penulis berharap semoga laporan tugas akhir ini bermanfaat bagi kita semua. Amin

Jakarta, 30 Juni 2021

Imam Farisi

## DAFTAR ISI

<b>HALAMAN SAMPUL</b> .....	<b>i</b>
<b>HALAMAN JUDUL</b> .....	<b>i</b>
<b>LEMBAR PERSETUJUAN PENGUJI</b> .....	<b>v</b>
<b>ABSTRAK</b> .....	<b>ix</b>
<b>ABSTRACT</b> .....	<b>x</b>
<b>KATA PENGANTAR</b> .....	<b>xi</b>
<b>DAFTAR ISI</b> .....	<b>xii</b>
<b>NASKAH JURNAL</b> .....	<b>1</b>
<b>KERTAS KERJA</b> .....	<b>14</b>
<b>BAB 1. LITERATUR REVIEW</b> .....	<b>15</b>
<b>BAB 2. ANALISIS DAN PERANCANGAN</b> .....	<b>20</b>
<b>BAB 3. SOURCE CODE</b> .....	<b>39</b>
<b>BAB 4. DATASET</b> .....	<b>90</b>
<b>BAB 5. TAHAPAN EKSPERIMEN</b> .....	<b>97</b>
<b>BAB 6. HASIL SEMUA EKSPERIMEN</b> .....	<b>108</b>
<b>DAFTAR PUSTAKA</b> .....	<b>114</b>
<b>LAMPIRAN DOKUMEN HAKI</b> .....	<b>116</b>
<b>LAMPIRAN KORESPONDENSI</b> .....	<b>117</b>

## NASKAH JURNAL

# Sistem Rekomendasi Lowongan Pekerjaan pada Perusahaan Outsourcing Menggunakan Algoritma Term Frequency Inverse Document Frequency dan Cosine Similarity

Imam Farisi<sup>1</sup>, Eliyani<sup>2</sup>

Universitas Mercu Buana, Indonesia

Article Info	ABSTRAK
<p><b>Genesis Artikel:</b> Diterima, xxx xx xxxx Direvisi, xxx xx xxxx Disetujui, xxx xx xxx</p>	<p>Perusahaan <i>outsourcing</i> menjadi salah satu alternatif yang bisa dimanfaatkan oleh pencari kerja dalam mendapatkan sebuah pekerjaan. Pekerjaan yang didapatkan akan disesuaikan berdasarkan daftar riwayat hidup oleh bagian <i>human resource development (HRD)</i>. Penelitian ini bertujuan untuk membantu bagian HRD dalam memetakan antara daftar riwayat hidup pencari kerja dengan perusahaan yang sesuai agar tidak membacanya secara manual. Sistem yang akan dibangun dalam penelitian ini menggunakan algoritma <i>TF-IDF</i> untuk pembobotan kata dan metode <i>cosine similarity</i> untuk mendapatkan kesesuaian antara daftar riwayat hidup dengan perusahaan. Data yang diolah oleh sistem berbentuk <i>file pdf</i> dengan kerangka yang sudah ditentukan. Kerangka tersebut terdiri dari tiga bagian yaitu pendidikan, kemampuan dan pengalaman dari pencari kerja. Hasil uji coba dengan pengujian confusion matrix didapatkan akurasi sebesar 76%.</p>
<p><b>Kata Kunci:</b> <i>Sistem Rekomendasi Lowongan Pekerjaan Sumber Daya Manusia TF-IDF Cosine Similarity</i></p>	<p><b>ABSTRACT</b></p>
<p><b>Keywords:</b> <i>Recommendation System Job Vacancy Human Resource TF-IDF Cosine Similarity</i></p>	<p><i>Outsourcing company is considered as one of the alternatives which may be used by job seekers for the purpose of finding any employment. The work obtained shall be adjusted by human resources development (HRD) based on curriculum vitae. This research was aimed to help HRD department with the mapping between job seeker's curriculum vitae and relevant companies so as to avoid manual reading. A system was established in this study by using TF-IDF algorithm for terms weighting and cosine similarity method to find out any conformity between curriculum vitae and the company. The data was processed by a system in the form of pdf file with a specified framework. The framework was composed of three sections namely education, ability, and experience of job seekers. The result of a test conducted using confusion matrix test suggested an accuracy of 76%.</i></p>
	<p style="text-align: right;"><i>This is an open access article under the CC BY-SA license.</i></p> 
<p><b>Penulis Korespondensi:</b> Imam Farisi, Fakultas Ilmu Komputer, Universitas Mercu Buana, Email: imamfarisi@gmail.com</p>	

## 1. PENDAHULUAN

Pekerjaan bisa didapatkan dengan berbagai cara antara lain dengan mendatangi langsung ke perusahaan untuk melihat informasi pengumuman lowongan pekerjaan [1], dengan melamar melalui portal karir atau dengan melamar melalui perusahaan *outsourcing*. Perusahaan *outsourcing* menjadi salah satu pilihan yang bisa dimanfaatkan untuk melanjutkan karir pekerjaan di beberapa industri besar yang tersebar di beberapa kota [2]. Tugas utama perusahaan ini adalah menyalurkan tenaga kerja ke perusahaan rekanan [3]. Manfaat dari perusahaan *outsourcing* menurut [4] adalah memungkinkan perusahaan rekanan fokus pada bisnis intinya dan bisa menghemat biaya operasional perusahaan karena tidak perlu merekrut karyawan sendiri.

Dalam mencari tenaga kerja, perusahaan dibantu oleh bagian *Human Resource Development* (HRD). HRD bertanggung jawab penuh dalam melakukan proses seleksi tenaga kerja, proses wawancara sampai ke proses akhir [5]. Namun bagian HRD sering kesulitan dalam melakukan proses seleksi terkait penyesuaian daftar riwayat hidup (CV) tenaga kerja dengan perusahaan rekanan. Hal ini dikarenakan HRD melakukan pembacaan manual CV kemudian mencari kesesuaian dengan spesifikasi perusahaan rekanan secara manual juga. Oleh karena itu dibuatlah sebuah sistem rekomendasi pekerjaan yang bisa membantu memberikan rekomendasi pekerjaan yang sesuai antara CV tenaga kerja dan spesifikasi perusahaan rekanan.

Penelitian terkait yang sudah dilakukan sebelumnya juga membahas mengenai hal ini antara lain penelitian [6] membahas tentang *hot topic detection* menggunakan algoritma *TF-IDF* yang menghasilkan akurasi sebesar 78,36%. Kemudian penelitian [7] yang membahas tentang *classification of shindi headline news* menggunakan *TF-IDF Text Analysis Schema* dengan hasil akurasi dari *Multinomial NB* (82%), *Linear SVC* (84%), *Logistic Regression* (83%), *MLP Classifier* (84%), *SGD Classifier* (82%), *Random Forest Classifier* (83%) dan *Ridge Classifier* (83%). Penelitian lain juga seperti pada [8] membahas tentang *research paper classification* menggunakan algoritma *TF-IDF* dan *LDA Schemes* menghasilkan *F-score* sekitar 0,9.

Pada penelitian ini data CV diambil sebanyak 100 dan data spesifikasi perusahaan rekanan diambil sebanyak 25 lowongan pekerjaan dari perusahaan *outsourcing* di PT XYZ. CV yang diambil berbahasa inggris dengan kerangka yang sudah ditentukan berbentuk *file pdf*. Spesifikasi perusahaan rekanan dan CV hanya yang berhubungan dengan Informasi dan Teknologi (IT). Metode yang digunakan sebagai pembobotan adalah *Term Frequency and Inverse Document Frequency (TF-IDF)* dan *cosine similarity* yang digunakan untuk mendapatkan kesesuaian antar dokumen [9].

## 2. METODE PENELITIAN

### 2.1 Metode Pengumpulan Data

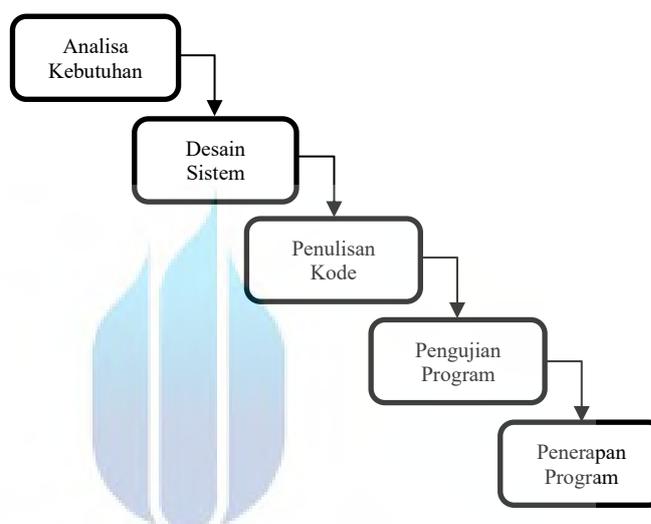
Dalam melakukan penelitian diperlukan data dan informasi yang nantinya digunakan sebagai bahan referensi untuk pembuatan penelitian.

#### 2.1.1 Wawancara

Wawancara merupakan salah satu teknik pengumpulan data yang dilakukan melalui tatap muka langsung dengan narasumber dengan cara tanya jawab secara langsung [10]. Wawancara dilakukan dengan bagian HRD PT XYZ yang menangani langsung pemetaan CV pencari kerja dengan perusahaan rekanan yang membuka lowongan kerja.

## 2.2 Metode Pengembangan Sistem

Metode pengembangan sistem yang digunakan dalam penelitian ini adalah dengan menggunakan metode *System Development Life Cycle (SDLC) Waterfall*. SDLC merupakan sebuah proses perubahan atau pembuatan sebuah sistem yang diperlukan dalam proses pembangunan sistem informasi berdasarkan kebutuhan pengguna [4-5] seperti pada gambar 1.



Gambar 1. *System Development Life Cycle Waterfall* [13]

### 2.2.1 Analisa Kebutuhan

Pada bagian ini dilakukan analisa terhadap kebutuhan sistem dengan melakukan wawancara tentang permasalahan yang dihadapi oleh bagian HRD ketika melakukan pemetaan CV dengan perusahaan rekanan. HRD akan memetakan secara manual dengan cara membaca satu persatu dan membandingkan dengan kebutuhan perusahaan rekanan yang sesuai. Jika terdapat kebutuhan baru, maka akan mengulangi proses membaca secara manual untuk menemukan perusahaan yang sesuai dengan CV tenaga kerja.

### 2.2.2 Desain Sistem

Pada bagian ini dilakukan pembuatan rancangan sistem antara lain menggunakan *use case diagram* untuk menggambar tugas dari masing-masing pengguna dan *entity relationship diagram* untuk menggambarkan hubungan antar data berdasarkan objek-objek dalam basis data.

### 2.2.3 Penulisan Kode Program

Pada bagian ini dilakukan pembuatan kode program menggunakan bahasa pemrograman *Java* dengan menggunakan *framework Spring Boot* dan *database PostgreSQL*. Kode program yang dihasilkan akan disesuaikan dengan desain sistem yang sudah dibuat pada bagian desain sistem.

#### 2.2.4 Pengujian Program

Pada bagian ini dilakukan uji coba menggunakan pengujian *black box* untuk menemukan kesalahan yang terjadi terhadap tampilan dan fungsi-fungsi yang ada di program [14] serta menguji tingkat kinerja suatu sistem rekomendasi dengan menggunakan pengujian *confusion matrix* [15].

#### 2.2.5 Penerapan Program

Pada bagian ini dilakukan penerapan program dengan menjalankan program yang telah berhasil diuji di dalam *tomcat server* sehingga bisa digunakan oleh bagian HRD.

### 2.3 Preprocessing

*Preprocessing* merupakan tahap pertama yang dilakukan untuk mempersiapkan data sebelum melakukan proses selanjutnya [16]. Dengan melakukan tahap ini data akan lebih terstruktur [17] dikarenakan proses ini telah menghilangkan beberapa data yang tidak diperlukan seperti angka dan karakter selain huruf. Dalam tahap ini terbagi menjadi beberapa bagian antara lain *case folding*, *stopword removal*, *stemming* dan *tokenizing*.

#### 2.3.1 Case Folding

*Case folding* merupakan tahap untuk merubah semua huruf di dalam dokumen menjadi huruf kecil semua [18]. Pada tahap ini dokumen akan dilakukan *case folding* dengan fungsi *toLowerCase()* di pemrograman *Java*.

#### 2.3.2 Stopword Removal

*Stopword removal* merupakan tahap untuk menghapus kata yang sering digunakan namun tidak memiliki makna yang penting di dalam dokumen [19]. Pada tahap ini dilakukan *stopword removal* dengan cara membuat sebuah daftar *list stopwords* di pemrograman *Java*. Kemudian daftar tersebut dibandingkan dengan dokumen, jika suatu kata di dokumen sama dengan daftar *list stopwords*, maka kata tersebut akan dihapus. Tujuan proses ini adalah untuk mengurangi jumlah kata dalam *dataset* yang akan mempengaruhi terhadap kecepatan dan performa kinerja sistem [20].

#### 2.3.3 Stemming

*Stemming* merupakan tahap untuk merubah kata berimbuhan menjadi kata dasar [21]. Pada tahap ini dokumen akan dilakukan *stemming* menggunakan *library apache opennlp*.

#### 2.3.4 Tokenizing

*Tokenizing* merupakan tahap untuk memisahkan kalimat menjadi potongan-potongan yang disebut token [22]. Pada tahap ini dokumen akan dilakukan *tokenizing* dengan fungsi *split()* di pemrograman *Java*.

### 2.4 Term Frequency Inverse Document Frequency (TF-IDF)

Metode *TF-IDF* merupakan metode untuk menghitung bobot setiap kata yang paling umum digunakan pada *information retrieval* dengan cara pemberian bobot hubungan suatu kata (term) terhadap dokumen [23]. Jadi semakin besar kemunculan suatu kata dalam dokumen, maka semakin besar pula bobotnya yang

akan memberikan nilai kesesuaian yang semakin besar. Bobot dari masing-masing dokumen dengan *query* bisa dihitung menggunakan rumus seperti pada (1) dan (2) :

$$IDF = \left( \frac{D}{DF} \right) \quad (1)$$

$$W = TF * \log (IDF) \quad (2)$$

Keterangan pada rumus 1 :

- IDF : nilai *Inversed Document Frequency*
- D : jumlah semua dokumen
- DF : jumlah dokumen yang mengandung kata (term)

Keterangan pada rumus 2 :

- W : bobot dokumen
- TF : jumlah kata yang muncul di dokumen
- DF : jumlah dokumen yang mengandung kata (term)
- IDF : nilai *Inversed Document Frequency*

## 2.5 Cosine Similarity

*Cosine similarity* merupakan ukuran kesamaan antara dua buah vektor dalam sebuah ruang dimensi yang didapat dari nilai cosinus sudut dari perkalian dua buah vektor yang dibandingkan [24]. Nilai cosine similarity ini hanya antara 0 dan 1 dan jika nilai kesamaan dari dua buah vektor dikatakan mirip ketika nilai dari cosine similarity adalah 1, sedangkan nilai 0 menandakan nilai kesamaannya tidak mirip. Untuk rumus perhitungan cosine similarity bisa dilihat di persamaan (1) dan cosine similarity dengan TF-IDF di persamaan (2) :

$$similarity(d_j, q_k) = \frac{\sum_{i=1}^n (td_{ij} \times tq_{ik})}{\sqrt{\sum_{i=1}^n td_{ij}^2} \times \sqrt{\sum_{i=1}^n tq_{ik}^2}} \quad (1)$$

$$similarityTFIDF(d_i, q_j) = \frac{\sum (wd_i \times wq_j)}{\sqrt{wd_i^2} \times \sqrt{wq_j^2}} \quad (2)$$

Keterangan pada rumus 1:

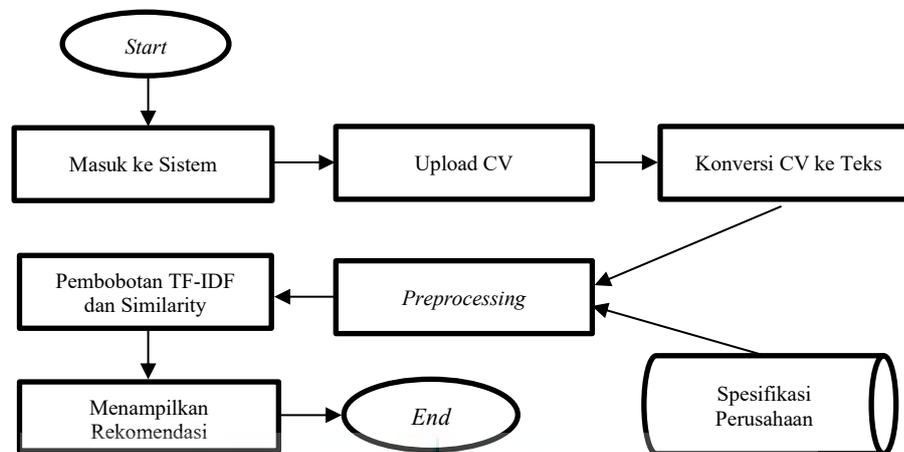
- $similarity(d_j, q_k)$  : tingkat kesamaan dokumen ke-j dengan query ke-k
- $td_{ij}$  : term ke-i dalam vektor untuk dokumen ke-j
- $td_{ik}$  : term ke-i dalam vektor untuk dokumen ke-k
- n : jumlah term yang unik

Keterangan pada rumus 2:

- $similarityTFIDF(d_j, q_k)$ : tingkat kesamaan dokumen ke-i dengan query ke-j
- $wd_i$  : bobot dokumen ke-i
- $wq_j$  : bobot query ke-j

## 2.6 Alur Proses Sistem Rekomendasi Lowongan Pekerjaan

Berikut merupakan alur proses sistem rekomendasi lowongan pekerjaan menggunakan diagram alir yang disajikan pada gambar 2 :



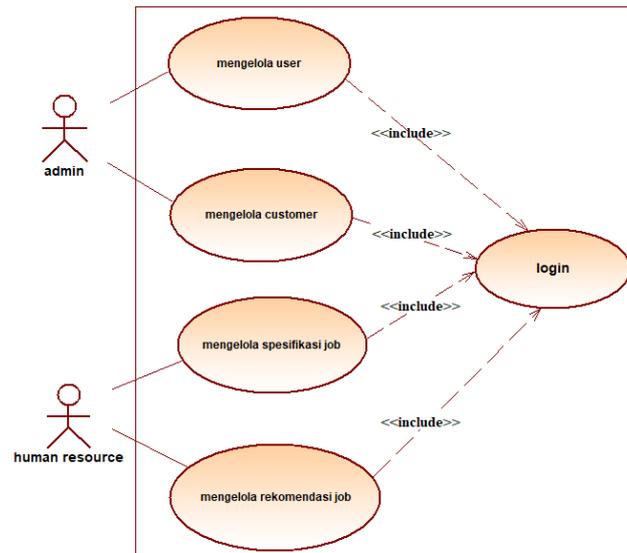
Gambar 2. Alur proses sistem rekomendasi lowongan pekerjaan

Pertama bagian HRD masuk ke sistem terlebih dahulu dengan melakukan *login* ke dalam sistem. Setelah berhasil langkah selanjutnya adalah mulai melakukan pengunggahan CV dan sistem melakukan konversi dari *file pdf* ke dalam bentuk teks. Sistem akan melakukan *query* data spesifikasi perusahaan rekanan di *database* yang selanjutnya akan dilakukan *preprocessing*. Kemudian dilakukan pembobotan dengan *TF-IDF* dan *cosine similarity* untuk mendapatkan kesesuaian antar dokumen. Setelah proses selesai maka sistem akan menampilkan hasil rekomendasi di dalam aplikasi..

### 3. HASIL DAN ANALISIS

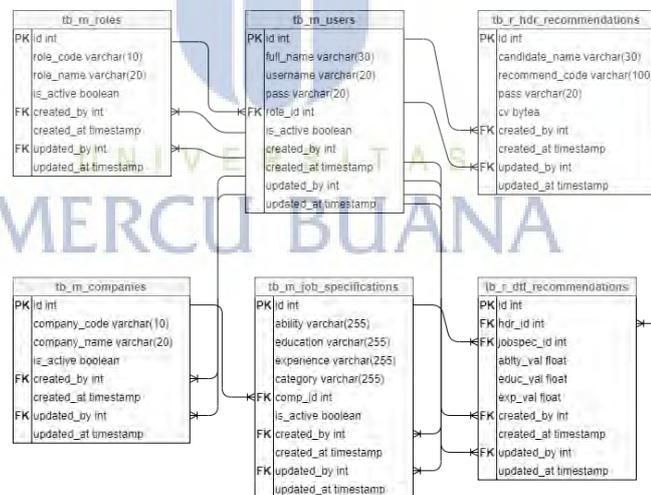
Hasil dari penelitian ini adalah rekomendasi lowongan pekerjaan dengan proses awal *preprocessing* terhadap dokumen CV dan dokumen spesifikasi perusahaan rekanan yang selanjutnya dilakukan proses pembobotan dengan *TF-IDF* dan perhitungan *cosine similarity* untuk mendapatkan tingkat kesamaan dokumen.

Pada pembuatan sistem rekomendasi terdapat dua pengguna yaitu admin dan HRD. Tugas admin antara lain mengelola pengguna dan perusahaan, sedangkan HRD yaitu mengelola spesifikasi dan rekomendasi pekerjaan. Untuk gambar tugas dari masing-masing pengguna digambarkan dengan *usecase diagram* pada gambar 3:



Gambar 3. Use Case Diagram

Gambar 4. menunjukkan rancangan relasi antar data dalam Database PostgreSQL. Terdapat enam tabel yang saling berhubungan antara lain `tb_m_roles`, `tb_m_users`, `tb_r_hdr_recommendations`, `tb_m_companies`, `tb_m_job_specifications` dan `tb_r_dtl_recommendations`.



Gambar 4. Entity Relationship Diagram (ERD)

Untuk proses rekomendasi pekerjaan, *human resource* melakukan pengunggahan *file* CV. CV ini mempunyai kerangka khusus seperti pada gambar 5. Kerangka ini terdapat tiga bagian utama yaitu pendidikan, kemampuan dan pengalaman kerja yang telah disesuaikan dengan standar CV di PT XYZ.

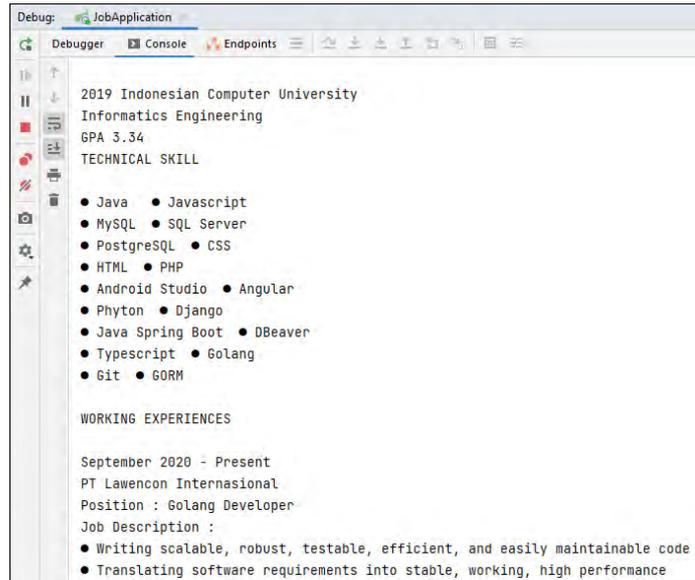
[ Full Name ]	
<b>PERSONAL BACKGROUND</b>	
Age	: [ age ]
Gender	: [ Male or Female ]
<b>FORMAL EDUCATION</b>	
2012	[ University name ]
	[ Departement name ]
	[ GPA ]
<b>TECHNICAL SKILLS</b>	
• [ Skill 1 ]	• [ Skill 2 ]
<b>WORKING EXPERIENCES</b>	
[ start date – end date ]	
[ company name ]	
Position	: [ position ]
<b>PROJECT LIST</b>	
Project Name	: [ project name ]
Project Description	: [ project description ]
Technology	: [ technology ]

Gambar 5. Kerangka daftar riwayat hidup (CV)

Proses pengunggahan terdiri dari 7 tahapan antara lain pembacaan *file pdf*, konversi *file pdf* ke teks, ambil data spesifikasi perusahaan dari *database*, *preprocessing*, pembobotan dengan *TF-IDF*, kemudian perhitungan tingkat kemiripan dokumen dengan *cosine similarity*. Pada gambar 6. merupakan proses pembacaan dan konversi dokumen yang diunggah dengan menggunakan *library apache pdfbox* ke dalam bentuk teks.



Gambar 6. Flowchart konversi PDF ke bentuk teks



Gambar 7. Contoh hasil konversi PDF ke bentuk teks

Pada gambar 8. merupakan tahapan pengambilan data spesifikasi perusahaan di dalam *database* dengan melakukan *query* ke tabel *tb\_m\_job\_specifications* yang kolom *is\_activenya* bernilai *true*.

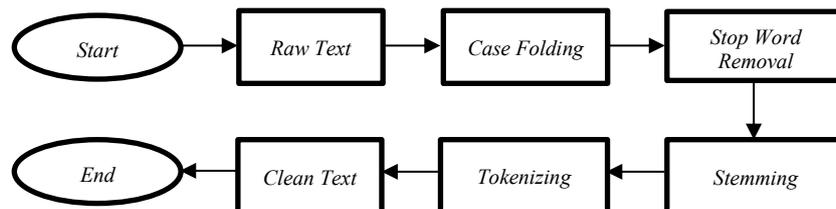


Gambar 8. Flowcart Pengambilan Data Spesifikasi Perusahaan

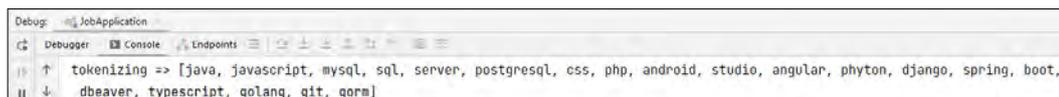
id	job category	job ability	job education	job experience
1	iOS Developer	iOS SDK, Java, HTML, javascript, CSS, knowledge in Object Oriented Programming and	all	Mobile application development either Objective C or Swift
2	Webmethods Engineer	Strong knowledge of API concept and microservices, web method	all	Working knowledge of MS SQL Server, and/or Oracle relational databases configur
3	Scrum Master	Scrum, sprint, Strong MS Office Skills	all	Extensive experience working in an Agile environment, Experience Tracking and rep
4	System Engineer	UNIX, apache, Java, RDBMS, Application Script, Oracle, shell script, Docker, Kubernetes	all	Understanding Docker, Kubernetes is a must, Understand IT software system develo
5	.NET Developer	.NET, ASP.NET, VB, C#, SQL Server, JQUERY, MVC, HTML, CSS, Javascript, Bootstrap, T	all	Experienced in SQL Server/Postgresql and Stored Procedure (SP), Having knowledge
6	IT Solution Analyst	Programming, SDLC, Software Development, Java, PHP, Full stack, ASP.net, .NET, Visi	Informatic Engineering, Information System	Experience with software development, creating FSD, TSD, Database Relationship D
7	Android Developer	Java, JS, RDBMS, Mobile Development, Android SDK, Web Service api, kotlin	all	Android skill development using java or kotlin, working knowledge of the general
8	React Native Developer	REST, MVC, ES6, React Native and reactjs	all	Experience with test-driven development (TDD) and agile methodologies, Experie
9	Business Analyst	SDLC, Analytical, .NET, VB, ASP.NET, Java, SQL Server, BI	Informatic Engineering, Information System	Have experience with Big Data and with other systems such as ODBC, MSSQL, mysql,
10	Java Developer	J2EE MVC, Hibernate, Angular, React JS HTML, JS, CSS, OOP, SQL & RDBMS	Informatic Engineering	Develop, testing, analyze and manage Java Web Application and supporting tools, i
11	Front End Developer	HTML, CSS, Javascript, Jquery, Progressive Web Applications (PWA), Vue	all	Experience working with apis, especially indexeddb, grpc and Protobuf
12	Full Stack Developer	Linux / Apache web servers, PHP, Java, mysql/postgresql/ Maria DB/mongodb databa	all	Proficient understanding of code versioning tools, such as Github/ Gitlab, Web ser
13	SQL Developer	T-SQL programming, oracle, HTML and javascript	all	Experience with Transparent Data Encryption (TDE), Have experience with ETL Platfo
14	PHP Developer	PHP MVC, Laravel, CI, Angular, HTML, JS, NodeJS, CSS, OOP, SQL & RDBMS	all	Work/Project based on PHP Tech Platform
15	Technical Writer (IT)	Jira and confluence, Microsoft Office, SDLC	Informatic Engineering, Information System	Write easy-to-understand user interface text, online help and developer guides, An
16	IT Help Desk	Application Delivery, Implementation, Software, SQL, RDBMS, mysql, postgresql, Tro	Informatic Engineering, Information System	Troubleshoot and Diagnose Problems, Communicate Effectively to help customers fi
17	.NET Developer	.NET, ASP.NET, VB, C#, SQL Server, JQUERY, MVC, HTML, CSS, Javascript, Bootstrap, T	all	Experienced in SQL Server/Postgresql and Stored Procedure (SP), Having knowledge
18	Android Developer	Java, JS, RDBMS, Mobile Development, Android SDK, Web Service api, kotlin	all	Android skills development using java or kotlin, working knowledge of the general
19	Java Developer	J2EE MVC, Hibernate, Angular, React JS HTML, JS, CSS, OOP, SQL & RDBMS	Informatic Engineering	Develop, testing, analyze and manage Java Web Application and supporting tools, i
20	Software Tester	Automatic Testing, Whitebox & Blackbox, Software Tester, SDLC, SQL, RDBMS, Linux	all	Have experience in automation testing tools, palatka, Selenium, spium and others,
21	Golang Developer	Golang, grpc, protobuf	all	Familiar with golang and web framework for REST API, familiar with grpc protobu
22	Vue Developer	Vuejs, pwa	all	Familiar with vuejs framework using pwa
23	Java Developer	J2EE MVC, Hibernate, Angular, React JS HTML, JS, CSS, OOP, SQL & RDBMS	Informatic Engineering	Develop, testing, analyze and manage Java Web Application and supporting tools, i
24	Node.js Developer	Nodejs, HTML, CSS, Javascript and/or library such as jquery, UI, or prototype, Angu	all	Preferably having knowledge of microservices, solid understanding and demonstr
25	Java Developer	J2EE MVC, Hibernate, Angular, React JS HTML, JS, CSS, OOP, SQL & RDBMS	all	Develop, testing, analyze and manage Java Web Application and supporting tools, i

Gambar 9. Hasil query ke table *tb\_m\_job\_specifications*

Pada gambar 10 dan gambar 11. merupakan tahapan *preprocessing* yang terdiri dari *case folding*, *stemming*, *stop word removal* dan *tokenizing*.



Gambar 10. Flowcart Preprocessing



Gambar 11. Contoh hasil preprocessing

Tahap berikutnya adalah pembobotan kata dengan *TF-IDF* pada tabel 1. untuk menghitung bobot kata dalam dokumen. Pembobotan dilakukan terhadap tiga bagian yaitu pendidikan, kemampuan dan pengalaman kerja. Berikut contoh perhitungan *TF-IDF* antara *CV (query)* dan dua dokumen spesifikasi perusahaan (d1 dan d2).

Tabel 1. Perhitungan TF-IDF

term	tf			df	log(idf)	tf-idf	
	q	d1	d2			d1	d2
echo	1	0	0	1	0,477	0	0
golang	1	1	0	2	0,176	0,176	0
java	1	0	1	2	0,176	0	0,176
testing	0	0	1	1	0,477	0	0,477
api	1	1	0	2	0,176	0,176	0
bank	1	0	0	1	0,477	0	0
grpc	1	1	0	2	0,176	0,176	0
gorm	1	1	0	2	0,176	0,176	0
spring	1	0	1	2	0,176	0	0,176
postgresql	1	0	0	1	0,477	0	0
rest	1	1	0	2	0,176	0,176	0
protobuf	1	1	0	2	0,176	0,176	0
css	1	0	1	2	0,176	0	0,176
android	1	0	0	1	0,477	0	0
angular	1	0	0	1	0,477	0	0
php	1	0	0	1	0,477	0	0
react	1	0	0	1	0,477	0	0
mysql	1	0	0	1	0,477	0	0
django	1	0	0	1	0,477	0	0
dbeaver	1	0	0	1	0,477	0	0
javascript	1	0	0	1	0,477	0	0

Setelah pembobotan selesai, proses selanjutnya adalah menghitung tingkat kemiripan antar dokumen menggunakan *cosine similarity* pada tabel 2.

Tabel 2. Perhitungan Cosine Similarity

term	tf-idf		akar kuadrat tf-idf		
	(q * d1)	(q * d2)	q	d1	d2
echo	0	0	0,477	0	0
golang	0,031	0	0,176	0,176	0
java	0	0,031	0,176	0	0,176
testing	0	0	0,000	0	0,477
api	0,031	0	0,176	0,176	0
bank	0	0	0,477	0	0
grpc	0,031	0	0,176	0,176	0
gorm	0,031	0	0,176	0,176	0
spring	0	0,031	0,176	0	0,176
postgresql	0	0	0,477	0	0
rest	0,031	0	0,176	0,176	0
protobuf	0,031	0	0,176	0,176	0
css	0	0,031	0,176	0	0,176
android	0	0	0,477	0	0
angular	0	0	0,477	0	0
php	0	0	0,477	0	0

react	0	0	0,477	0	0
mysql	0	0	0,477	0	0
django	0	0	0,477	0	0
dbeaver	0	0	0,477	0	0
javascript	0	0	0,477	0	0
<b>Total</b>	<b>0,186</b>	<b>0,093</b>	<b>6,833</b>	<b>1,057</b>	<b>1,005</b>
					<i>Cosine Similarity (q, d1)</i>
					<b>0,026</b>
					<i>Cosine Similarity (q, d2)</i>
					<b>0,014</b>

Berdasarkan hasil perhitungan *cosine similarity* di atas didapatkan hasil untuk persamaan dokumen 1 dengan *query* menghasilkan nilai 0,026 sedangkan hasil persamaan untuk dokumen 2 dan *query* adalah 0,014. Maka dapat disimpulkan bahwa dokumen yang paling dekat dengan *query* adalah dokumen 1, karena semakin besar nilai *cosinus*, maka semakin mirip dokumen tersebut. Hasil dari proses *cosine similarity* akan ditampilkan di dalam aplikasi dengan menampilkan dua keterangan yaitu “sesuai”, “dipertimbangkan” atau “tidak sesuai”. Untuk keterangan dari masing-masing kategori berdasarkan nilai *cosine similarity* bisa dilihat pada tabel 3.

Tabel 3. Keterangan Nilai Kategori

Keterangan	Pendidikan	Kemampuan	Pengalaman
Sesuai	nilai > 0	nilai > 0	nilai > 0
Sesuai	0	nilai > 0	nilai > 0
Sesuai	nilai > 0	0	nilai > 0
Sesuai	nilai > 0	nilai > 0	0
Dipertimbangkan	nilai > 0	0	0
Dipertimbangkan	0	nilai > 0	0
Dipertimbangkan	0	0	nilai > 0
Tidak sesuai	0	0	0

Untuk menguji akurasi pada sistem rekomendasi ini digunakan metode *confusion matrix* dengan melakukan perhitungan antara hasil dari sistem rekomendasi dan hasil uji secara manual dengan membaca satu persatu CV pada tabel 4.

Tabel 4. Perbandingan Hasil dari Sistem Rekomendasi dan Hasil Uji Manual

CV	Jumlah Hasil Uji Manual	Jumlah Hasil Rekomendasi	Selisih
Sesuai	19	22	3
Dipertimbangkan	3	3	0
Tidak Sesuai	3	0	3

Total	25	25	6
-------	----	----	---

Dari hasil pengujian pada tabel 5, maka akurasi dapat dihitung dengan metode *confusion matrix* dengan tabel 3 berikut :

Tabel 5. Pengujian Confusion Matrix

n = 25	<b>Aktual Benar</b>	<b>Aktual Salah</b>
<b>Prediksi Benar</b>	TP : 19	FP : 0
<b>Prediksi Salah</b>	FN : 6	TN : 0
	25	0

$$\begin{aligned}
 \text{Akurasi} &= ((TP + TN) / (TP + FP + FN + TN)) * 100\% \\
 &= ((19 + 0) / (19 + 0 + 6 + 0)) * 100\% \\
 &= (19 / 25) * 100\% \\
 &= 76\%
 \end{aligned}$$

#### 4. KESIMPULAN

Berdasarkan hasil penelitian di atas bisa disimpulkan bahwa penelitian ini bisa memberikan sebuah rekomendasi lowongan pekerjaan berupa tampilan yang berisi keterangan kesesuaian antara CV tenaga kerja dan kebutuhan perusahaan rekanan yang bisa membantu bagian HRD dalam pekerjaannya. Dalam pengujiannya hasil akurasi yang didapat dari pengujian *confusion matrix* adalah 76 persen. Penulis berharap untuk penelitian yang lebih lanjut bisa mengembangkan penelitian ini dengan melakukan *filter* pada data spesifikasi perusahaan sehingga akurasi yang didapat bisa lebih bagus lagi.

#### REFERENSI

- [1] S. Yuliani and R. Aries, "Aplikasi Gerai Layanan Informasi Kerja (Glik) (Studi Kasus: Dinas Tenaga Kerja Dan Transmigrasi Provinsi Jawa Barat)," *Maj. Ilm. UNIKOM*, vol. 13, no. 01, pp. 93–100, 2015, doi: 10.34010/miu.v13i01.15.
- [2] R. Fahmi, N. Meiji, and N. Hadi, "DINAMIKA PEKERJA MUDA OUTSOURCING : KEGEMBIRAAN DAN KERENTANAN MENUJU RISIKO MASA DEPAN," *J. Antropol. Isu-Isu Sos. Budaya*, vol. 01, no. June, pp. 37–46, 2021.
- [3] P. Supriatin and N. Meliana, "Sistem Informasi Penggajian Karyawan Pada Pt . Essei Perbama," vol. 6, no. 1, pp. 142–135, 2021.
- [4] C. P. Andu, "Analisis Penyebab Karyawan Resign pada Project Perusahaan BPO (Business Process Outsourcing): Studi PT . Transcosmos Indonesia Project Tokopedia Semarang," *J. Represent.*, vol. 7, no. 01, pp. 1–9, 2021.
- [5] D. Damiyana and M. Maroghy, "Tugas dan Peran Human Resource Development (HRD) pada PT. Jakarta Teknologi Utama," *J. Mitra Manaj.*, vol. 5, no. 1, pp. 13–21, 2021, [Online]. Available: <http://e-jurnalmitramanajemen.com/index.php/jmm/article/view/125/69>.
- [6] Z. Zhu, J. Liang, D. Li, H. Yu, and G. Liu, "Hot Topic Detection Based on a Refined TF-IDF Algorithm," *IEEE Access*, vol. 7, no. January, pp. 26996–27007, 2019, doi: 10.1109/ACCESS.2019.2893980.

- [7] I. A. Kandhro *et al.*, "Classification of Sindhi Headline News Documents based on TF-IDF Text Analysis Scheme," *Indian J. Sci. Technol.*, vol. 12, no. 33, pp. 1–10, 2019, doi: 10.17485/ijst/2019/v12i33/146130.
- [8] S. W. Kim and J. M. Gil, "Research paper classification systems based on TF-IDF and LDA schemes," *Human-centric Comput. Inf. Sci.*, vol. 9, no. 1, 2019, doi: 10.1186/s13673-019-0192-7.
- [9] H. Kusniyati and A. A. Nugraha, "Analysis of Matric Product Matching Between Cosine Similarity with Term Frequency-Inverse Document Frequency (TF-IDF) and Word2Vec in PT. Pricebook Digital Indonesia," *Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol.*, vol. 6, no. 1, pp. 105–112, 2020, doi: 10.32628/cseit195672.
- [10] Malik and A. Juliansyah, "Kebijakan pengelolaan pariwisata desa sebagai potensi pendapatan apbdes," *J. Sos. Hum.*, vol. 01, no. 01, pp. 1–10, 2021.
- [11] D. Haeraini and H. Hartono, "Aplikasi Pengelolaan Dokumen pada Muat Kapal Berbasis Desktop ( Studi Kasus : PT KSJM ) Desktop-Based Document Management Application on Shiploding ( Case Study : PT KSJM )," *J. Sist. dan Teknol. Inf.*, vol. 09, no. 2, pp. 86–92, 2021, doi: 10.26418/justin.v9i2.42928.
- [12] Suwarno and R. R. Lim, "Perancangan Sistem Kasir Layanan Mandiri Berbasis Web Pada Supermarket Bless Dengan Metode Sdlc," *Conf. Manag. Business, Innov. Educ. Soc. Sci.*, vol. 1, no. 1, pp. 860–873, 2021.
- [13] A. Priamnistiko, Y. Handrianto, and S. Hadi Sukmana, "Rancang Bangun Sistem Informasi E-Learning Menggunakan Model Waterfall," *J. Inov. Inform.*, vol. 6, no. 1, pp. 50–57, 2021, doi: 10.51170/jii.v6i1.143.
- [14] F. Ardi and H. P. Putro, "Pengujian Black Box Aplikasi Mobile Menggunakan Katalon Studio (Studi Kasus: ACC Partner PT. Astra Sedaya Finance)," *Automata*, vol. 2, no. 1, 2021.
- [15] Danny, Lina, and A. Chris, "PREPARAT DENGAN CONVOLUTIONAL NEURAL," *J. Ilmu Komput. dan Sist. Inf.*, vol. 9, no. 1, pp. 48–52, 2021.
- [16] S. Khairunnisa, Adiwijaya, and S. Al Faraby, "Pengaruh Text Preprocessing terhadap Analisis Sentimen Komentar Masyarakat pada Media Sosial Twitter ( Studi Kasus Pandemi)," *J. Media Inform. Budidarma*, vol. 5, no. April, pp. 406–414, 2021, doi: 10.30865/mib.v5i2.2835.
- [17] R. D. Himawan and Eliyani, "Perbandingan Akurasi Analisis Sentimen Tweet terhadap Pemerintah Provinsi DKI Jakarta di Masa," *J. Edukasi dan Penelit. Inform.*, vol. 7, no. 1, pp. 58–63, 2021, [Online]. Available: <https://jurnal.untan.ac.id/>.
- [18] Oryza Habibie Rahman, G. Abdillah, and Agus Komarudin, "Klasifikasi Ujaran Kebencian pada Media Sosial Twitter Menggunakan Support Vector Machine," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 1, pp. 17–23, 2021, doi: 10.29207/resti.v5i1.2700.
- [19] Samsir, Ambiyar, U. Verawardina, F. Edi, and R. Watrianthos, "Analisis Sentimen Pembelajaran Daring Pada Twitter di Masa Pandemi COVID-19 Menggunakan Metode Naïve Bayes," *J. Media Inform. Budidarma*, vol. 5, no. 1, pp. 157–163, 2021, doi: 10.30865/mib.v5i1.2604.
- [20] A. B. Arifa, Gita Fadila Fitriana, and Ananda Rifkiy Hasan, "Temu Kembali Informasi pada Soal Ujian dengan Rencana Pembelajaran Menggunakan Vector Space Model," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 1, pp. 63–68, 2021, doi: 10.29207/resti.v5i1.2739.
- [21] F. D. Ananda and Y. Pristyanto, "Analisis Sentimen Pengguna Twitter Terhadap Layanan Internet Provider Menggunakan Algoritma Support Vector Machine," *J. Matrik*, vol. 20, no. 2, pp. 407–416, 2021, doi: 10.30812/matrik.v20i2.1130.
- [22] W. W. Kamal and C. I. Ratnasari, "Analisis Sentimen Ulasan Produk : Kajian Pustaka," *Automata*, vol. 2, no. 1, 2021.
- [23] R. Melita, V. Amrizal, H. B. Suseno, and T. Dirjan, "Penerapan Metode Term Frequency Inverse Document Frequency (Tf-Idf) Dan Cosine Similarity Pada Sistem Temu Kembali Informasi Untuk Mengetahui Syarah Hadits Berbasis Web (Studi Kasus: Hadits Shahih Bukhari-Muslim)," *J. Tek. Inform.*, vol. 11, no. 2, pp. 149–164, 2018, doi: 10.15408/jti.v11i2.8623.
- [24] I. Mawanta, T. S. Gunawan, and Wanayumini, "Uji Kemiripan Kalimat Judul Tugas Akhir dengan Metode Cosine Similarity dan Pembobotan TF-IDF," *J. Media Inform. Budidarma*, vol. 5, no. April, pp. 726–738, 2021, doi: 10.30865/mib.v5i2.2935

## KERTAS KERJA

### 1. Literatur Review

Di dalam Literatur Review ini ada beberapa hasil review jurnal yang terkait dengan penelitian. Literatur yang di review terdiri dari 10 jurnal nasional dan 7 jurnal internasional.

### 2. Analisa dan Perancangan

Analisa dan perancangan berisi penjelasan beserta gambar dari use case, user interface, activity diagram, sequence diagram, class diagram, data flow diagram dan ERD (Entity relationship Diagram).

### 3. Source Code

Source code ini berisi kumpulan pernyataan atau deklarasi bahasa pemrograman komputer. Program ini menggunakan bahasa *Java* dengan *framework Spring Boot*.

### 4. Dataset

Dataset yang digunakan dibagi menjadi tiga bagian yaitu *dataset* riwayat hidup berbentuk file pdf berbahasa inggris, *dataset* spesifikasi *job* dan *dataset* pengguna berbentuk record basis data.

### 5. Tahapan Eksperimen

Tahapan eksperimen merupakan penjelasan tahapan-tahapan eksperimen pada laporan ini yang menggunakan rekayasa perangkat lunak yaitu dengan pengujian blackbox dan confusion matrix.

### 6. Hasil Eksperimen

Hasil eksperimen bagian ini berisi hasil semua eksperimen baik yang berhasil maupun yang gagal. Sesuai metode maupun jenis pengujian yang digunakan pada laporan ini.