



**PERANCANGAN SISTEM *PARSING CV* DAN *SKILL MATCHING* BERBASIS NLP TERINTEGRASI WEBSITE
UNTUK MENDUKUNG PROSES REKRUTMEN: STUDI
KASUS PT. YIHONG NOVATEX INDONESIA**

**TUGAS AKHIR
APLIKATIF**

**ADE RAMA SAPUTRA
41521120036**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2025**



**PERANCANGAN SISTEM *PARSING* CV DAN *SKILL*
MATCHING BERBASIS NLP TERINTEGRASI WEBSITE
UNTUK MENDUKUNG PROSES REKRUTMEN: STUDI
KASUS PT. YIHONG NOVATEX INDONESIA**

**TUGAS AKHIR
APLIKATIF**

Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar Sarjana

ADE RAMA SAPUTRA

41521120036

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2025**

HALAMAN PERNYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama : Ade Rama Saputra
NIM : 41521120036
Program Studi : Teknik Informatika

Menyatakan dengan sebenar-benarnya bahwa Tugas Akhir berjudul:

“Perancangan Sistem Parsing CV Dan Skill Matching Berbasis NLP Terintegrasi Website Untuk Mendukung Proses Rekrutmen: Studi Kasus PT. Yihong Novatex Indonesia” adalah hasil karya saya sendiri, tidak mengandung unsur plagiarisme, pelanggaran hak cipta, atau konten ilegal dalam bentuk apapun dan tidak melanggar hukum atau pihak manapun.

Apabila di kemudian hari ditemukan pelanggaran terhadap pernyataan ini, saya bersedia menanggung seluruh konsekuensi hukum dan membebaskan Universitas Mercu Buana dari segala bentuk tuntutan hukum dan saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana.

Demikian pernyataan ini saya buat dengan sebenar-benarnya untuk digunakan sebagaimana mestinya.

Jakarta, 12 Agustus 2026


Ade Rama Saputra

PERNYATAAN SIMILARITY CHECK
/SIMILARITY CHECK STATEMENT

Saya yang bertanda tangan dibawah ini menyatakan, bahwa karya ilmiah yang ditulis oleh
/The undersigned, hereby declare that the scientific work written by

Nama /Name : ADE RAMA SAPUTRA
NIM /Student ID Number : 41521120036
Program Studi /Study Program : Teknik Informatika

Dengan Judul Tugas Akhir

/The title:

**“PERANCANGAN SISTEM PARSING CV DAN SKILL MATCHING BERBASIS
NLP TERINTEGRASI WEBSITE UNTUK Mendukung PROSES REKRUTMEN:
STUDI KASUS PT. YIHONG NOVATEX INDONESIA”**

telah dilakukan pengecekan *similarity* dengan sistem Turnitin pada tanggal:

/Similarity checks have been carried out with the Turnitin system on the date:

28 Juli 2026

dengan nilai persentase sebesar :

/with a percentage value of:

12%

dinyatakan memenuhi standar sesuai dengan ketentuan berlaku di **Fakultas Ilmu Komputer**
Universitas Mercu Buana. */declared to meet standards in accordance with applicable
regulations at the Faculty of Computer Science, Universitas Mercu Buana.*

File hasil cek similarity turnitin:

/Turnitin similarity report file

https://drive.google.com/file/d/1-d60YnhpuE2vo2HDR4PDzwJHpAYCYqmp/view?usp=drive_linkk



Jakarta, 28 Juli 2026
Admin Turnitin Fasilkom UMB



Agung Prawoto, S.Kom., B.Sc
NIK : 322970503

HALAMAN PENGESAHAN

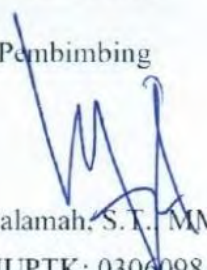
Tugas Akhir ini diajukan oleh:

Nama : Ade Rama Saputra
NIM : 41521120036
Fakultas/Program Studi : Fakultas Ilmu Komputer / Teknik Informatika
Judul Tugas Akhir : Perancangan Sistem Parsing CV dan Skill Matching
Berbasis NLP Terintegrasi Website untuk Mendukung
Proses Rekrutmen: Studi Kasus PT. Yihong Novatex
Indonesia

Telah berhasil dipertahankan pada sidang tanggal 12 Agustus 2026 dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer Universitas Mercu Buana.

Disahkan oleh:

Pembimbing


Umniy Salamah, S.T., MMSI.
NIDN/NUPTK: 0306098104

Jakarta, 12 Agustus 2026

Mengetahui,

Dekan Fakultas Ilmu Komputer



Dr. Bambang Jokonowo, S.Si., MTI
NIDN/NUPTK: 0320037002

Ketua Program Studi
Teknik Informatika



Dr. Hadi Santoso, S.Kom., M.Kom
NIDN/NUPTK: 0225067701

KATA PENGANTAR

Puji syukur kehadiran Tuhan yang Maha Esa, atas segala rahmat dan ridha-Nya sehingga penulis dapat menyelesaikan proposal penelitian yang merupakan salah satu persyaratan kelulusan Program Studi Strata Satu (S1) pada jurusan Teknik Informatika, Universitas Mercu Buana.

Penulis menyadari bahwa proposal penelitian ini masih jauh dari sempurna, karena kesempurnaan sejatinya hanya milik Tuhan yang Maha Esa. Oleh karena itu, saran dan masukan yang membangun senantiasa penulis terima dengan senang hati. Serta berkat dukungan, motivasi, bantuan, bimbingan, dan doa dari banyak pihak, penulis mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. Andi Adriansyah, M.Eng. selaku Rektor Universitas Mercu Buana.
2. Bapak Dr. Bambang Jokonowo, S.Si., MTI selaku Dekan Fakultas Ilmu Komputer.
3. Bapak Dr. Hadi Santoso, S.Kom., M.Kom. selaku Ketua Program Studi Teknik Informatika Universitas Mercubuana.
4. Ibu Umniy Salamah, S.T., MMSI., selaku dosen pembimbing yang telah memberikan pengarahan, motivasi, menyediakan waktu, tenaga, dan pikiran sehingga selama pembuatan proposal penelitian ini terjadwal dengan baik.
5. Teruntuk Ibu Armida dan Bapak Ramli, selaku orang tua saya yang selalu berdoa dan mendukung saya selama menjalani masa studi sebagai mahasiswa Universitas Mercubuana.
6. Teruntuk Silvia Apriani, dan Eliza Anis selaku kakak kandung saya, yang sudah mendukung dan memberikan semangat kepada saya selama menjalani kuliah.
7. Teruntuk Bapak Dianto Isnan Laksono Putra, selaku kakak ipar saya. Terimakasih karena sudah memberikan kesempatan dan dukungan untuk melakukan penelitian ini. Sehingga dapat ter-realisasikan. Semoga penelitian ini bisa memberikan manfaat yang seluas-luasnya dan mendapatkan pahala yang terus mengalir.
8. Teruntuk Nourma Goepita Ramadhita, seseorang yang terkasih. Terimakasih telah memberikan dukungan dan menemani dalam

menyelesaikan tugas akhir ini. Terimakasih telah kebersamai, mari selalu kebersamai setiap perjalanan kita.

9. Ade Rama Saputra, terimakasih sudah berusaha dan bertahan untuk menyelesaikan studi di Universitas Mercu Buana. Anda keren.

Akhir kata, penulis berharap semoga Tuhan yang Maha Esa membalas kebaikan dan selalu mencurahkan rahmat, hidayah, serta panjang umur kepada kita semua, aamiin. Terima Kasih.

Jakarta, 12 Agustus 2026

Ade Rama Saputra



HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR DI REPOSITORI UMB

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

Nama : Ade Rama Saputra
NIM : 41521120036
Fakultas/Program Studi : Ilmu Komputer/Teknik Informatika
Judul Tugas Akhir : Perancangan Sistem Parsing CV Dan Skill Matching Berbasis NLP Terintegrasi Website Untuk Mendukung Proses Rekrutmen: Studi Kasus PT. Yihong Novatex Indonesia

Demi pengembangan ilmu pengetahuan, dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Non-Eksklusif (*Non-exclusive Royalty-Free Right*)** atas karya ilmiah saya yang berjudul di atas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti Non-Eksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan Tugas Akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta. Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 12 Agustus 2026

Yang menyatakan,



Ade Rama Saputra

**PERANCANGAN SISTEM PARSING CV DAN SKILL MATCHING
BERBASIS NLP TERINTEGRASI WEBSITE UNTUK MENDUKUNG
PROSES REKRUTMEN: STUDI KASUS PT. YIHONG NOVATEX
INDONESIA**

ADE RAMA SAPUTRA

ABSTRAK

Proses *screening* Curriculum Vitae (CV) di PT. Yihong Novatex Indonesia masih dilakukan secara manual dengan membaca CV satu per satu dan mencocokkan keterampilan kandidat terhadap kebutuhan lowongan pekerjaan. Proses tersebut memerlukan waktu yang relatif lama serta berpotensi menimbulkan subjektivitas dalam proses seleksi awal. Penelitian ini bertujuan untuk merancang sistem *parsing* CV dan *skill matching* berbasis Natural Language Processing (NLP) yang diimplementasikan sebagai API dan diintegrasikan dengan *website prototype* rekrutmen guna mendukung proses *screening* awal kandidat. Penelitian ini menggunakan metode Research and Development (R&D) dengan pendekatan *mixed method*. Proses *parsing* CV dilakukan melalui ekstraksi teks dokumen PDF menggunakan *PDFPlumber*, dilanjutkan dengan ekstraksi informasi berbasis Regular Expression (Regex) serta normalisasi keterampilan menggunakan *synonym mapping*, sedangkan proses *skill matching* menerapkan mekanisme *Exact Match* dan algoritma *Fuzzy Matching* menggunakan metode *Token Set Ratio*. Pengujian dilakukan terhadap 64 CV yang berhasil diproses dari total 140 CV menggunakan tiga skenario lowongan pekerjaan, yaitu Operator Sablon, PPIC, dan *Quality Control*. Hasil pengujian menunjukkan bahwa sistem memperoleh nilai *accuracy* sebesar 96,88% pada skenario Operator Sablon, 73,44% pada skenario PPIC, dan 81,25% pada skenario *Quality Control*. Selain itu, hasil evaluasi pengguna menunjukkan tingkat penerimaan sistem yang sangat baik dengan persentase jawaban Sangat Setuju sebesar 91% dan Setuju sebesar 8%. Berdasarkan hasil tersebut, sistem yang dikembangkan mampu membantu proses *screening* kandidat secara lebih efektif, efisien, dan objektif, serta mendukung digitalisasi proses rekrutmen melalui integrasi layanan *parsing* CV dan *skill matching* pada *website prototype*.

Kata kunci: *parsing* CV, *skill matching*, Natural Language Processing, sistem rekrutmen, API

DESIGN AND DEVELOPMENT OF AN NLP-BASED CV PARSING AND SKILL MATCHING SYSTEM INTEGRATED WITH A RECRUITMENT WEBSITE: A CASE STUDY OF PT. YIHONG NOVATEX INDONESIA

ADE RAMA SAPUTRA

ABSTRACT

The initial screening of job applicants at PT. Yihong Novatex Indonesia is currently conducted manually by reviewing each Curriculum Vitae (CV) and comparing candidate skills with job requirements. This approach is time consuming and may lead to inconsistencies in the screening process. To address this issue, this study proposes an NLP-based CV parsing and skill matching system, implemented as an Application Programming Interface (API) and integrated into a recruitment website prototype to automate the early-stage candidate screening process. This research adopted the Research and Development (R&D) methodology with a mixed-methods approach. The proposed system extracts textual information from PDF-based CVs using PDFPlumber, identifies candidate information through Regular Expressions (Regex), normalizes extracted skills using a synonym mapping approach, and performs skill matching through an Exact Match mechanism followed by Fuzzy Matching based on the Token Set Ratio algorithm. The API integration enables the recruitment website to automatically process CV uploads, extract candidate information, perform skill matching, and generate candidate recommendations within a unified recruitment workflow. The system was evaluated using 64 valid CVs selected from a dataset of 140 CVs across three job scenarios: Screen Printing Operator, Production Planning and Inventory Control (PPIC), and Quality Control (QC). The evaluation achieved accuracy scores of 96.88%, 73.44%, and 81.25%, respectively. User acceptance testing also demonstrated positive results, with 91% of responses indicating Strongly Agree and 8% indicating Agree. These findings indicate that the proposed system effectively supports a faster, more efficient, and more objective candidate screening process while contributing to the digitalization of recruitment through the integration of CV parsing and skill matching services into a web-based recruitment platform.

Kata kunci: CV parsing, skill matching, Natural Language Processing, recruitment system, API

DAFTAR ISI

HALAMAN SAMPUL.....	0
HALAMAN JUDUL	i
HALAMAN PERNYATAAN KARYA SENDIRI	ii
HALAMAN SURAT KETERANGAN HASIL UJI TURNITIN	iii
HALAMAN PENGESAHAN	iv
KATA PENGANTAR.....	v
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR DI REPOSITORI UMB	vii
ABSTRAK	viii
ABSTRACT	ix
DAFTAR ISI.....	x
DAFTAR TABEL	xii
DAFTAR GAMBAR.....	xiii
DAFTAR LAMPIRAN	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah.....	3
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian.....	4
1.5 Sistematika Penulisan.....	5
BAB II TINJAUAN PUSTAKA.....	8
2.1 Teori Utama.....	8
2.2 Teori Pendukung	13
2.3 Gambaran Umum PT. Yihong Novatex Indonesia	16
2.4 Penelitian Terdahulu.....	16
2.5 Gap Penelitian	22
BAB III METODE PENELITIAN	24
3.1 Pendekatan Penelitian.....	24
3.2 Subjek Penelitian	25
3.3 Instrumen Penelitian.....	26
3.4 Teknik Pengumpulan Data	28
3.5 Analisis Data	28
3.6 Prosedur Penelitian.....	31

3.7	Evaluasi Hasil Penelitian.....	36
BAB IV PEMBAHASAN.....		38
4.1	Pembahasan.....	38
4.2	Alur Sistem.....	39
4.2.1	Alur Sistem Parsing CV	39
4.2.2	Alur Sistem Skill Matching.....	40
4.2.3	Alur Integrasi Sistem Sebagai API dengan Website Rekrutmen.	44
4.3	Hasil Sistem.....	45
4.3.1	Hasil Pengujian Dataset	45
4.3.2	Hasil Pengujian Parsing CV	49
4.3.3	Hasil Pengujian Skill Matching.....	51
4.3.4	Hasil Evaluasi Sistem.....	57
4.3.5	Hasil Integrasi Sistem.....	62
4.3.6	Hasil Evaluasi Pengguna.....	63
BAB V KESIMPULAN DAN SARAN		68
5.1	Kesimpulan.....	68
5.2	Saran.....	69
DAFTAR PUSTAKA.....		71
LAMPIRAN.....		73



DAFTAR TABEL

Tabel 2.4.1 Penelitian Terdahulu	16
Tabel 3.3.4.1 Format Pilihan Jawaban Kuesioner skala Likert	27
Tabel 3.3.4.2 Perhitungan Peresentase Jawaban Kuesioner skala Likert	27
Tabel 3.5.1 Analisis Data Kualitatif.....	29
Tabel 4.3.1.1 Hasil Pengujian Dataset	45
Tabel 4.3.1.2 Dokumen PDF Selain CV	48
Tabel 4.3.2.1 Pengujian Tahapan Parsing CV	49
Tabel 4.3.3.1 Hasil Pengujian Tahapan Skill Matching	51
Tabel 4.3.3.2 Skenario Pengujian	52
Tabel 4.3.4.1 Perbandingan Hasil Evaluasi Sistem dengan Ground Truth.....	57
Tabel 4.3.4.2 Hasil Perhitungan Confusion Matrix Operator Sablon.....	58
Tabel 4.3.4.3 Hasil Perhitungan Confusion Matrix QC.....	59
Tabel 4.3.4.4 Hasil Perhitungan Confusion Matrix PPIC.....	60
Tabel 4.3.6.1 Daftar Pertanyaan Kuesioner	64
Tabel 4.3.6.2 Hasil Penilaian dalam Persen.....	64



DAFTAR GAMBAR

Gambar 2.1.3.1 Rumus Algoritma Levenshtein Distance	10
Gambar 2.1.3.2 Perhitungan Token Set Ratio.....	11
Gambar 2.3.1 Logo PT. Yihong Novatex Indonesia.....	16
Gambar 3.1.1 Model Waterfall	24
Gambar 3.6.1 Flowchart Sistem.....	34
Gambar 3.6 2 Flowchart Ekstraksi dengan Regular Expression.....	35
Gambar 4.2.1.1 Flowchart Parsing CV	39
Gambar 4.2.2.1 Flowchart Skill Matching.....	41
Gambar 4.2.2.2 Data Judul Lowongan dan Required Skill.....	42
Gambar 4.2.2.3 Hasil Akhir Skill Matching	43
Gambar 4.2.3.1 Pengujian API dengan Method POST.....	44
Gambar 4.2.3.2 Response Pengujian API	45
Gambar 4.3.3.1 Sample Hasil Pengujian Skenario 1	53
Gambar 4.3.3.2 Sample Hasil Pengujian Skenario 2	54
Gambar 4.3.3.3 Sample Hasil Pengujian Skenario 3	54
Gambar 4.3.3.4 Hasil Pengujian Fuzzy Matching dengan Dummy CV 1	55
Gambar 4.3.3.5 Hasil Pengujian Fuzzy Matching dengan Dummy CV 2	56
Gambar 4.3.3.6 Hasil Penerapan Fuzzy Matching dengan Dataset	56
Gambar 4.3.5.1 Hasil Integrasi Sistem dengan Website Rekrutmen	62
Gambar 4.3.5.2 Hasil Integrasi Perhitungan Persentase Sistem dengan Website Rekrutmen	63
Gambar 4.3.6.1 Chart Bar Hasil Evaluasi Pengguna	67



DAFTAR LAMPIRAN

Lampiran 1. Kartu Asistensi	73
Lampiran 2. Curriculum Vitae	74
Lampiran 3. Sertifikasi BNSP	75
Lampiran 4. Daftar Pertanyaan Wawancara Kepada staf HRD	76
Lampiran 5. Proses Rekrutmen Manual di PT. Yihong Novatex Indonesia	77
Lampiran 6. Surat Permohonan Observasi Data kepada Dinas Ketenagakerjaan (Disnaker) Kabupaten Cirebon	78
Lampiran 7. Surat Tembusan Permohonan Observasi Data kepada Badan Kesatuan Bangsa dan Politik Kabupaten Cirebon	79
Lampiran 8. Surat Balasan Observasi Data PT. Yihong Novatex Indonesia	80
Lampiran 9. Piagam Penghargaan Dinas Ketenagakerjaan Kabupaten Cirebon .	81
Lampiran 10. Surat Pernyataan & Pengalihan HAKI	82

