

BAB I

PENDAHULUAN

A. Latar Belakang

Pengangguran masih menjadi permasalahan di berbagai negara, termasuk Indonesia. Berdasarkan data dari Badan Pusat Statistik (BPS), tingkat pengangguran terbuka (TPT) di Indonesia pada Agustus 2023 mencapai 5,32% [1]. Bahkan, tingkat pengangguran pada tahun 2023 lebih tinggi dibandingkan dengan tahun 2019 yang hanya 5,23%. Ditinjau berdasarkan tingkat pendidikan, lulusan perguruan tinggi yang menganggur mencapai 5,18%, angka ini lebih tinggi dibandingkan dengan tingkat pengangguran pada lulusan SD (2,56%) dan SMP (4,78%) [1]. Hal ini menunjukkan bahwa meskipun lulusan perguruan tinggi memiliki tingkat pendidikan yang lebih tinggi, mereka masih menghadapi tantangan dalam mendapatkan pekerjaan yang sesuai.

Salah satu kendala utama dalam pencarian kerja adalah keterbatasan akses terhadap informasi lowongan yang relevan dengan keahlian dan minat pencari kerja. Banyak pencari kerja mengalami kesulitan dalam menemukan informasi lowongan yang benar-benar sesuai dengan keahlian yang dimiliki [2]. Di sisi lain, perusahaan juga mengalami kesulitan dalam menjaring kandidat yang sesuai dengan kebutuhan posisi, karena sistem perekrutan yang digunakan masih terbatas dalam mencocokkan keterampilan kandidat dengan kualifikasi yang dibutuhkan. Oleh karena itu, diperlukan sistem yang dapat membantu dalam pencarian lowongan pekerjaan antara pencari kerja dan perusahaan.

Salah satu sistem yang digunakan dalam proses rekrutmen adalah sistem *Applicant Tracking System* (ATS). Sistem ini, berfungsi untuk mengelola dan menyaring lamaran pekerjaan secara otomatis. Namun, pemanfaatan ATS masih memiliki kelemahan dalam memahami profil kandidat secara menyeluruh. Sistem ini, melakukan pencocokan berbasis kata kunci tanpa mempertimbangkan relevansi konteks dari keseluruhan latar belakang kandidat. Akibatnya, kandidat potensial tidak mendapatkan rekomendasi yang sesuai, meskipun pencari kerja memiliki keterampilan yang relevan dengan pekerjaan yang tersedia [3].

Perusahaan PT XYZ, mempunyai sistem ATS yang telah diterapkan untuk mendukung proses rekrutmen tenaga profesional di bidang IT. Namun, sistem yang tersedia saat ini belum memiliki fitur rekomendasi, sehingga pencari kerja harus mencari lowongan secara manual. Hal ini menyebabkan banyak calon pekerja mengalami kesulitan dalam mencari informasi lowongan pekerjaan yang sesuai dengan keahlian pencari kerja, sehingga menunjukkan pentingnya pengembangan sistem yang tepat [4].

Sebagai solusi, sistem rekomendasi pekerjaan mulai dikembangkan dalam mencocokkan kandidat dengan lowongan pekerjaan. Salah satu teknologi utama yang mendukung pengembangan sistem ini adalah *Natural Language Processing* (NLP). NLP memungkinkan sistem untuk menganalisis teks pada profil kandidat dan deskripsi pekerjaan untuk mengidentifikasi keterampilan serta melakukan pencocokan secara semantik. Melalui NLP, informasi pencari kerja seperti posisi, keahlian, dan pengalaman kerja dapat diekstraksi sehingga meningkatkan relevansi hasil rekomendasi [5].

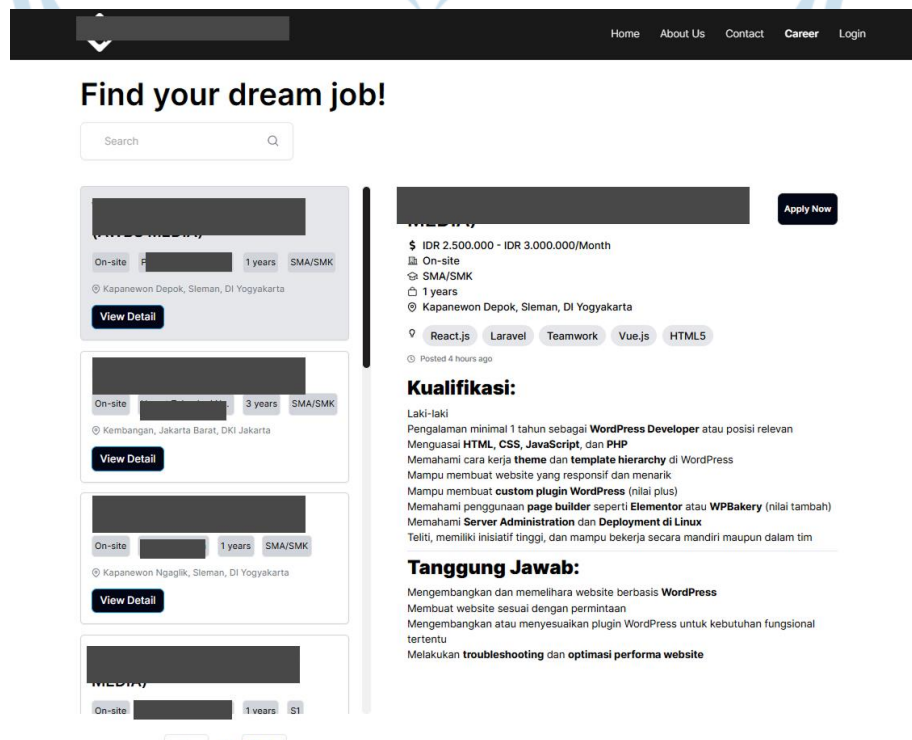
NLP mampu memahami keterampilan pencari kerja dalam konteks yang lebih luas dengan memproses teks secara semantik. Hasil pemrosesan NLP, dapat digunakan dalam metode *Content-Based Filtering* (CBF), yang merupakan salah satu pendekatan utama dalam sistem rekomendasi pekerjaan. Metode CBF bekerja dengan mencocokkan lowongan pekerjaan berdasarkan kesamaan dengan profil pencari kerja, seperti keterampilan, pengalaman, dan minat yang dimiliki. Dengan demikian, CBF dapat membantu sistem rekomendasi, termasuk pada sistem berbasis ATS, untuk memberikan rekomendasi pekerjaan dengan profil dan preferensi pengguna [6] [7].

Alasan utama penggunaan *Content-Based Filtering* (CBF) dalam kasus PT XYZ adalah karena CBF tidak memerlukan data dari pengguna lain untuk memberikan rekomendasi. Hal ini sangat bermanfaat ketika pengguna baru bergabung dengan sistem atau ketika data pengguna lain tidak tersedia. Dalam kondisi tersebut, sistem tetap dapat memberikan rekomendasi berdasarkan informasi yang terdapat dalam profil pengguna, seperti keterampilan, pengalaman, dan minat yang dimiliki [8]. Berdasarkan uraian tersebut, penelitian ini akan

mengembangkan sistem rekomendasi pekerjaan berbasis *Content-Based Filtering* pada *website* ATS di PT XYZ untuk memberikan rekomendasi pekerjaan antara kandidat dengan lowongan pekerjaan.

B. Rumusan Masalah

PT XYZ telah mengembangkan sistem *Applicant Tracking System* (ATS) untuk membantu pencari kerja menemukan peluang karier. Sistem ini masih memiliki kekurangan pada tidak adanya fitur rekomendasi lowongan pekerjaan. Saat ini, pencari kerja masih harus mencari lowongan pekerjaan secara manual. Seperti yang ditampilkan pada Gambar 1.1, pencarian manual belum memberikan pengalaman yang personal bagi pengguna, padahal sistem telah menyimpan data penting dalam profil pencari kerja seperti keterampilan, pengalaman, dan preferensi kerja. Oleh karena itu, penelitian ini bertujuan mengembangkan sistem rekomendasi berbasis *Content-Based Filtering* (CBF) yang didukung teknologi *Natural Language Processing* (NLP) untuk mencocokkan profil pencari kerja dengan lowongan.



Gambar 1.1. Tampilan Sistem ATS PT XYZ

Berdasarkan rumusan masalah tersebut, maka pertanyaan penelitian pada tugas akhir ini adalah:

1. Bagaimana membangun *dataset* lowongan pekerjaan dari Glints dan pencari kerja dari Google Form yang berfokus pada bidang IT?
2. Bagaimana mengembangkan sistem rekomendasi pekerjaan dengan pendekatan *Content-Based Filtering* (CBF) pada *Applicant Tracking System* (ATS) di PT XYZ?

C. Batasan Masalah

1. *Dataset* lowongan pekerjaan yang digunakan hanya berasal dari situs Glints, sedangkan data pencari kerja dikumpulkan melalui kuesioner Google Form.
2. Pengumpulan *dataset ground truth* dilakukan melalui penyebaran kuesioner yang menampilkan judul lowongan dan nama perusahaan, kemudian responden diminta untuk memilih pekerjaan yang sesuai dengan preferensi pencari kerja sebagai acuan evaluasi sistem rekomendasi.
3. Sistem rekomendasi hanya difokuskan untuk bidang pekerjaan Teknologi Informasi (IT). Lowongan dari sektor lain tidak termasuk dalam cakupan pengolahan data dan rekomendasi.
4. Atribut yang digunakan dalam metode *Content-Based Filtering* meliputi *title*, *minimal education*, *skills*, dan *description* dari *dataset* lowongan pekerjaan, serta *preference role*, *education*, *skills*, dan *about* dari *dataset* pencari kerja.
5. Sistem menggunakan metode *Content-Based Filtering* dengan *cosine similarity*, yang dikombinasikan dengan representasi teks TF-IDF, *Word Embedding*, dan BERT.
6. Evaluasi model rekomendasi dilakukan menggunakan metrik ranking seperti *Precision@k*, *NDCG@k*, dan *MAP@k*.
7. Model yang telah dikembangkan akan diintegrasikan ke dalam *website* pada *environment development*.

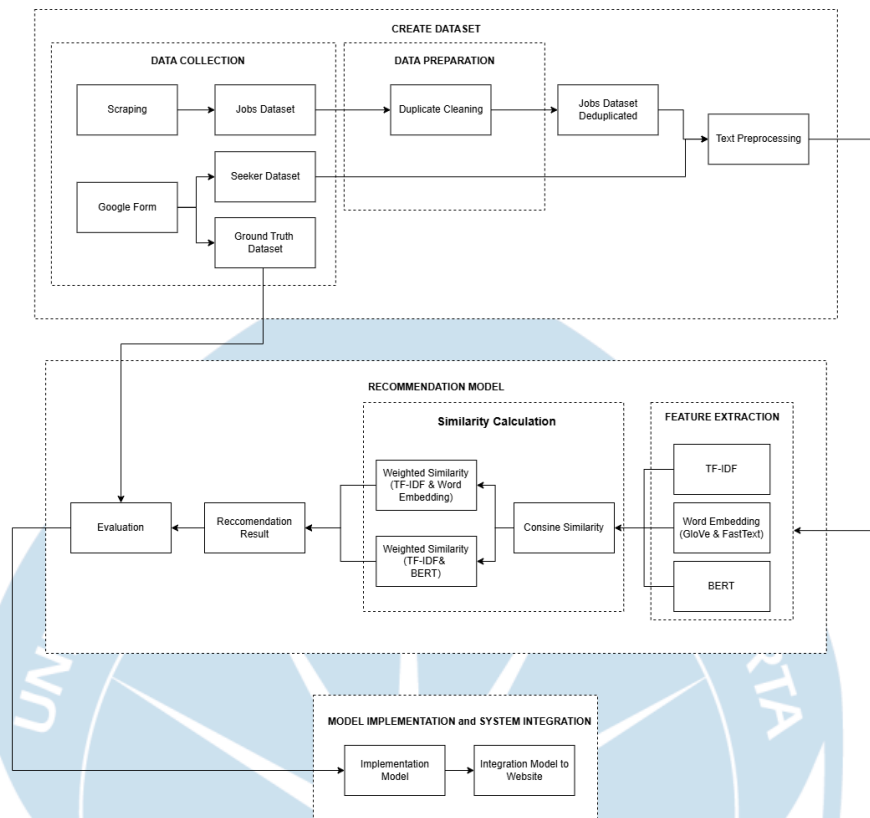
8. Sistem rekomendasi pada *website* ATS PT XYZ hanya menampilkan hasil secara langsung tanpa menerapkan mekanisme *caching* untuk peningkatan performa.

D. Tujuan Penelitian

1. Membangun *dataset* lowongan pekerjaan dari Glints dan pencari kerja dari Google Form yang berfokus pada bidang IT.
2. Mengembangkan sistem rekomendasi pekerjaan pada *Applicant Tracking System* (ATS) di PT XYZ dengan menerapkan pendekatan *Content-Based Filtering* (CBF).

E. Metode Penelitian

Metode penelitian ini terdiri dari tiga komponen utama, yaitu pembuatan *dataset*, pengembangan model rekomendasi, dan tahap integrasi model kedalam *website* ATS milik PT XYZ. Masing-masing komponen mencakup serangkaian tahapan, dimulai dari proses pengumpulan data hingga implementasi sistem rekomendasi pada *website* PT XYZ. Alur lengkap dari *pipeline* penelitian ini dapat dilihat pada Gambar 1.2.



Gambar 1.2. Pipeline Rekomendasi Pekerjaan

1. Membuat *Dataset*

a. Pengumpulan *Dataset*

Dataset dalam penelitian ini berasal dari dua sumber utama, yaitu data lowongan pekerjaan dan data pencari kerja. Data lowongan pekerjaan diperoleh melalui proses *web scraping* menggunakan pustaka Puppeteer, sedangkan data pencari kerja dan data *ground truth* dikumpulkan secara primer melalui Google Form.

b. Persiapan *Dataset*

Tahap persiapan dilakukan dengan menghapus entri duplikat secara semantik menggunakan metode *fuzzy string matching* pada deskripsi dan atribut lowongan pekerjaan untuk memastikan setiap data bersifat unik.

c. *Text Preprocessing*

Tahap ini bertujuan untuk membersihkan dan menyiapkan data teks dari lowongan pekerjaan dan profil pencari kerja agar dapat dianalisis oleh sistem rekomendasi. *Text preprocessing* dilakukan melalui delapan tahapan, yaitu *data cleaning*, *lowercasing*, *custom phrases*, *custom synonym*, *tokenization*, *stopword removal*, *pos tagging*, *lemmatization*, dan *stemming*.

2. Membuat Model Rekomendasi

a. *Feature Extraction*

Proses ekstraksi fitur dilakukan dengan menggunakan metode TF-IDF, *Word Embedding*, dan BERT. Ketiga metode ini digunakan untuk mengubah data teks menjadi representasi numerik yang dapat dianalisis oleh sistem.

b. *Similarity Calculation*

Perhitungan kemiripan dilakukan untuk mengukur tingkat kesesuaian antara profil pencari kerja dan deskripsi lowongan pekerjaan. Tahapan ini menggunakan metode *cosine similarity* sebagai dasar pengukuran. Selain itu, dilakukan juga perhitungan *weighted similarity* yang menggabungkan dua representasi teks, yaitu TF-IDF dengan *word embedding*, serta TF-IDF dengan BERT, guna membandingkan efektivitas masing-masing pendekatan.

c. Menampilkan Hasil Rekomendasi

Sistem menampilkan rekomendasi pekerjaan yang diurutkan secara *descending* berdasarkan skor kemiripan tertinggi, lalu menyajikan sejumlah *top-k* hasil paling relevan kepada setiap pengguna.

d. Evaluasi Model

Evaluasi dilakukan dengan membandingkan hasil rekomendasi terhadap data *ground truth* yang diperoleh dari penilaian manual

pengguna. Pengukuran kinerja sistem menggunakan metrik Precision@10, MAP@10, dan NDCG@10.

3. Implementasi Model dan Integrasi Sistem

a. Implementaion Model

Model sistem rekomendasi dikemas dalam bentuk REST API menggunakan *framework* FastAPI. API ini memungkinkan *backend* sistem untuk berkomunikasi dengan model menggunakan format JSON.

b. Integrasi Model ke dalam *Website* ATS

Integrasi model ke dalam *website* ATS yang dikembangkan menggunakan Next.js untuk *frontend* dan NestJS untuk *backend* dengan basis data *PostgreSQL*.

F. Sistematika Penulisan

Sistematika penulisan Tugas Akhir ini disusun secara runtut dan terstruktur guna memaparkan setiap tahapan penelitian secara menyeluruh, yang dijabarkan sebagai berikut:

1. BAB I: Pendahuluan

Bab ini menjelaskan mengenai latar belakang permasalahan yang mendasari penelitian, rumusan masalah, batasan masalah, tujuan penelitian, serta metode penelitian yang digunakan. Bab ini juga menyajikan gambaran umum alur penelitian dalam bentuk *pipeline* sistem rekomendasi pekerjaan.

2. BAB II: Tinjauan Pustaka

Bab ini membahas penelitian-penelitian sebelumnya yang relevan dengan topik penelitian ini. Selain itu, juga disajikan perbandingan metode dan pendekatan dari studi terdahulu dengan penelitian yang dilakukan saat ini.

3. BAB III: Landasan Teori

Bab ini menyajikan teori-teori dasar serta konsep-konsep pendukung yang berkaitan langsung dengan penelitian, mencakup pengertian, rumus-rumus, dan contoh penerapan metode yang digunakan.

4. BAB IV: Pembangunan *Dataset*

Bab ini menjelaskan proses pembuatan dan pengolahan *dataset* yang digunakan dalam penelitian, mulai dari pengumpulan *dataset* hingga tahap tahap persiapan data.

5. BAB V: Pengembangan Model Sistem Rekomendasi

Bab ini mencakup tahapan pengembangan model rekomendasi, mulai dari proses ekstraksi fitur, perhitungan kemiripan, menampilkan hasil rekomendasi, hingga evaluasi hasil rekomendasi menggunakan metrik evaluasi.

6. BAB VI: Implementasi Model dan Integrasi Sistem

Bab ini menjelaskan proses implementasi model ke dalam bentuk REST API, serta integrasinya ke dalam sistem *website* ATS yang dikembangkan. Disajikan pula gambaran alur integrasi dan hasil pengujian sistem pada *website* ATS PT XYZ.

7. BAB VII: Penutup

Bab ini berisi kesimpulan dari hasil penelitian yang telah dilakukan serta saran-saran pengembangan lebih lanjut untuk meningkatkan kinerja dan relevansi sistem rekomendasi di masa mendatang.