

### **WAJIB DI PERHATIKAN**

- Ukuran kertas: A4
- Margin: 3 cm (kiri), 3 cm (atas), 2.5 cm (kanan dan bawah)
- Font: Times New Roman
- Ukuran font:
  - Isi teks : 12 pt
  - Judul Subbab : 12pt, **Bold**
  - Judul Bab dan Cover : 13pt, **Bold**
- Spasi: 1.15 (untuk semua isi laporan)
- Spasi antar paragraf: 6 pt (after)
- Penomoran halaman: bawah tengah

**CareerPath AI: Sistem Rekomendasi Jalur Karir dan Pengembangan Skill  
Personal untuk Mahasiswa Berbasis Analisis Tren Industri**



**[Kelompok : 18]**

**Anggota Kelompok:**

- 1. JOSEPH PRATAMA LASE - 255150201111012**
- 2. [RAIHAN MIRZA FIRDAUS] – [255150207111107]**
- 3. MUHAMMAD HIDAYATULLOH HABIB AL MUHLISIN –  
255150201111025**
- 4. [Nama Anggota 4] – [NIM]**

**DEPARTEMEN TEKNIK INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS BRAWIJAYA  
2025**

**DAFTAR ISI**

<b>DAFTAR ISI .....</b>	<b>2</b>
<b>ABSTRAK.....</b>	<b>3</b>
<b>BAB I PENDAHULUAN .....</b>	<b>4</b>
1.1 Latar Belakang .....	4
1.2 Rumusan Masalah .....	5
1.3 Tujuan.....	5
1.4 Manfaat.....	6
Manfaat Teoritis (Bagi Ilmu Pengetahuan) .....	6
Manfaat Praktis (Bagi Masyarakat & Instansi).....	6
<b>BAB II TINJAUAN PUSTAKA .....</b>	<b>7</b>
<b>BAB III METODOLOGI DAN SOLUSI .....</b>	<b>10</b>
3.1 Metodologi Perancangan .....	10
3.2 Solusi.....	12
<b>BAB IV HIPOTESIS HASIL .....</b>	<b>14</b>
<b>DAFTAR PUSTAKA .....</b>	<b>19</b>
<b>LAMPIRAN .....</b>	<b>21</b>

Jangan lupa upgrade halamannya yak :) Semangatt !!!

## **ABSTRAK**

Pesatnya pertumbuhan ekonomi digital di Indonesia menciptakan permintaan tinggi akan talenta digital, namun dihadapkan pada tantangan kesenjangan keterampilan (*skills gap*) antara

kompetensi lulusan perguruan tinggi dan kebutuhan riil industri. Kondisi ini menyebabkan kebingungan bagi mahasiswa dalam merencanakan jalur karir dan memilih keterampilan yang relevan untuk dipelajari. Proposal ini merancang "CareerPath AI", sebuah sistem cerdas yang bertujuan untuk menjembatani kesenjangan tersebut melalui pendekatan berbasis data. Dengan memanfaatkan teknologi Ilmu Data, sistem ini menggunakan *Natural Language Processing* (NLP) untuk secara otomatis menganalisis dan mengekstrak daftar keterampilan yang paling dibutuhkan dari ribuan data lowongan pekerjaan online. Selanjutnya, sebuah *Recommendation Engine* akan mengolah informasi tersebut untuk menghasilkan peta jalan (roadmap) belajar yang dipersonalisasi bagi setiap mahasiswa sesuai dengan minat dan tujuan karir mereka. Diharapkan sistem ini dapat menjadi panduan yang akurat bagi mahasiswa untuk meningkatkan daya saing dan kesiapan kerja, sekaligus memberikan masukan berharga bagi institusi pendidikan dalam pengembangan kurikulum yang lebih relevan.

## **BAB I**

### **PENDAHULUAN**

#### **1.1 Latar Belakang**

Indonesia sedang mengalami akselerasi transformasi digital yang pesat, menciptakan permintaan eksponensial terhadap talenta digital yang kompeten. Laporan dari Google, Temasek, dan Bain & Company (2023) memproyeksikan bahwa ekonomi digital Indonesia akan terus tumbuh secara signifikan, menegaskan bahwa kebutuhan akan tenaga kerja di bidang teknologi informasi (TI) bukanlah tren sesaat, melainkan sebuah fondasi ekonomi masa depan. Pemerintah melalui Kementerian Komunikasi dan Informatika (Kominfo) juga menargetkan pemenuhan jutaan talenta digital untuk mendukung visi Indonesia Emas 2045, menandakan adanya peluang karir yang sangat luas bagi lulusan TI.

Namun, di tengah besarnya permintaan ini, muncul sebuah tantangan krusial yang dikenal sebagai **kesenjangan keterampilan** (*skills gap*). Laporan Bank Dunia (World Bank, 2021) menyoroti adanya ketidaksesuaian antara keterampilan yang diajarkan di institusi pendidikan dengan yang dibutuhkan oleh industri digital saat ini. Banyak perusahaan di Indonesia melaporkan kesulitan dalam menemukan kandidat yang memiliki kombinasi keterampilan teknis dan non-teknis yang relevan dengan perkembangan teknologi terkini. Akibatnya, banyak lulusan baru merasa kurang siap bersaing karena portofolio skill mereka belum selaras dengan ekspektasi pasar kerja.

Dari perspektif mahasiswa, kondisi ini menimbulkan kebingungan dan ketidakpastian. Terutama bagi mahasiswa di tahun-tahun awal, mereka dihadapkan pada pilihan spesialisasi, bahasa pemrograman, dan *tools* yang sangat beragam tanpa panduan yang jelas mengenai mana yang paling prospektif. Laporan "Future of Jobs" dari World Economic Forum (2023) secara konsisten menyebutkan bidang seperti *AI and Machine Learning Specialists*, *Data Analysts and Scientists*, dan *Cybersecurity Specialists* sebagai peran dengan pertumbuhan tercepat. Akan tetapi, kurikulum standar seringkali tidak dapat beradaptasi secepat laju perubahan di industri, sehingga mahasiswa memerlukan panduan proaktif untuk pengembangan diri di luar kelas.

Oleh karena itu, dibutuhkan sebuah solusi inovatif yang dapat menjembatani kesenjangan ini. "CareerPath AI" dirancang sebagai sistem cerdas yang memberikan rekomendasi jalur karir dan pengembangan skill yang personal dan berbasis data. Dengan menganalisis tren terkini langsung dari ribuan lowongan pekerjaan di industri digital, sistem ini bertujuan untuk memberikan peta jalan yang jelas dan relevan bagi mahasiswa, memastikan setiap upaya pembelajaran yang mereka lakukan selaras dengan kebutuhan nyata industri dan mempersiapkan mereka untuk menjadi talenta digital yang unggul dan siap kerja.

## 1.2 Rumusan Masalah

1. Bagaimana sebuah sistem dapat menentukan daftar skill yang paling dibutuhkan oleh industri digital saat ini berdasarkan data lowongan kerja?
2. Bagaimana menyajikan hasil analisis tren industri ke dalam bentuk rekomendasi langkah-langkah belajar yang personal untuk setiap mahasiswa?
3. Bagaimana menerapkan teknologi Data Science untuk secara otomatis menganalisis ribuan lowongan pekerjaan dan menghasilkan rekomendasi skill yang cocok bagi mahasiswa?

## 1.3 Tujuan

1. Membangun model yang mampu mengidentifikasi dan merangking daftar skill teknis maupun non-teknis yang paling dibutuhkan industri digital berdasarkan analisis kumpulan data lowongan kerja.
2. Merancang antarmuka pengguna (user interface) yang dapat menyajikan data tren industri dalam bentuk rekomendasi peta jalan (roadmap) belajar yang personal dan mudah dipahami oleh mahasiswa.
3. Mengimplementasikan teknologi Data Science, khususnya Natural Language Processing (NLP) dan Recommendation Engine, untuk mengotomatisasi proses analisis data lowongan kerja dan menghasilkan rekomendasi yang akurat.

## 1.4 Manfaat

### Manfaat Teoritis (Bagi Ilmu Pengetahuan)

- Memberikan kontribusi pada bidang rekayasa perangkat lunak dan ilmu data, khususnya dalam penerapan teknik *Natural Language Processing* (NLP) untuk analisis teks berbahasa Indonesia dalam konteks deskripsi pekerjaan.
- Menjadi model acuan dan studi kasus bagi penelitian selanjutnya mengenai pengembangan sistem rekomendasi (*recommendation engine*) di bidang teknologi pendidikan (*EdTech*) dan perencanaan karir.
- Menyediakan wawasan metodologis tentang cara mengolah dan menerjemahkan data tren industri yang tidak terstruktur menjadi sebuah model pengetahuan yang terstruktur untuk panduan akademis.

### Manfaat Praktis (Bagi Masyarakat & Instansi)

- **Bagi Mahasiswa:**
  - Mendapatkan panduan karir yang jelas, personal, dan berbasis data, sehingga mengurangi kebingungan dalam memilih jalur spesialisasi.
  - Membantu mahasiswa dalam mengambil keputusan yang lebih strategis terkait pemilihan mata kuliah, pelatihan, dan sertifikasi yang relevan dengan target karir.
  - Meningkatkan daya saing dan kesiapan kerja (*employability*) lulusan baru saat memasuki dunia industri.
- **Bagi Perguruan Tinggi:**
  - Menyediakan data dan wawasan terkini mengenai tren kebutuhan *skill* di industri digital yang dapat digunakan sebagai bahan evaluasi dan pengembangan kurikulum.
  - Membantu meningkatkan relevansi program studi dengan kebutuhan pasar kerja, sehingga dapat meningkatkan reputasi dan tingkat penyerapan lulusan.
- **Bagi Industri:**
  - Secara tidak langsung, membantu mengurangi kesenjangan keterampilan (*skills gap*) antara lulusan baru dan kebutuhan industri.

- Mempermudah proses rekrutmen talenta tingkat pemula (*entry-level*) karena kandidat telah memiliki pemahaman dan keterampilan yang lebih selaras dengan kebutuhan perusahaan.

## **BAB II**

### **TINJAUAN PUSTAKA**

#### **2.1 Teori atau Teknologi yang Digunakan**

##### **2.1.1 Natural Language Processing (NLP)**

Natural Language Processing (NLP) adalah cabang dari kecerdasan buatan yang memungkinkan komputer untuk memahami, menafsirkan, dan menghasilkan bahasa manusia. Dalam sistem CareerPath AI, NLP digunakan untuk mengekstraksi kata kunci keterampilan (skills) dari deskripsi pekerjaan. Teknik seperti tokenization, part-of-speech tagging, dan named entity recognition digunakan untuk mengenali istilah penting yang menunjukkan keahlian teknis maupun non-teknis. Menurut Jurafsky dan Martin (2023), NLP memiliki peran vital dalam analisis teks karena dapat mengubah data bahasa alami menjadi representasi numerik yang dapat dianalisis secara otomatis.

##### **2.1.2 Data Science**

Data Science merupakan bidang multidisiplin yang menggabungkan statistika, analisis data, serta kecerdasan buatan untuk mengekstraksi wawasan dan pola dari kumpulan data besar. Dalam konteks penelitian ini, Data Science berperan penting dalam mengolah data lowongan pekerjaan yang bersifat tidak terstruktur agar dapat diubah menjadi informasi yang bermakna. Menurut Provost dan Fawcett (2020), proses utama dalam Data Science meliputi pengumpulan data, pembersihan data, analisis, visualisasi, dan interpretasi hasil untuk mendukung pengambilan keputusan berbasis bukti (*data-driven decision making*).

### 2.1.3 Recommendation System

Sistem rekomendasi (*Recommendation System*) merupakan teknologi yang digunakan untuk memberikan saran personal berdasarkan preferensi pengguna atau pola data tertentu. Menurut Ricci et al. (2022), sistem rekomendasi umumnya dibagi menjadi tiga pendekatan utama: *content-based filtering*, *collaborative filtering*, dan *hybrid model*. Pada penelitian ini, pendekatan *content-based* digunakan untuk menyesuaikan rekomendasi skill dengan tren industri yang ditemukan dari data lowongan kerja.

### 2.1.4 Machine Learning

Machine Learning merupakan teknik komputasi yang memungkinkan sistem belajar dari data tanpa harus diprogram secara eksplisit. Algoritma pembelajaran seperti *classification* dan *clustering* digunakan untuk mengelompokkan keterampilan berdasarkan kemiripan dan frekuensinya dalam data industri. Misalnya, penelitian *A Study on using data clustering for feature extraction to improve the quality of classification* (2021) menunjukkan bahwa clustering dapat dimanfaatkan sebagai fitur tambahan untuk meningkatkan akurasi *classification*. Selain itu, pentingnya generalisasi model dari data pelatihan ke kondisi nyata juga ditekankan dalam literatur terkini (Liu, Shen, He, Zhang, Xu, Yu, & Cui, 2021), mengingat distribusi data industri dapat berubah dinamis.

## 2.2 Proyek-Proyek Sejenis sebagai Pemanding

Salah satu proyek yang memiliki tujuan serupa dengan penelitian ini adalah SkillRec yang dikembangkan oleh Ong dan Lim pada tahun 2023. Proyek tersebut berupaya membantu pengguna mengenali keterampilan yang relevan dengan dunia kerja melalui analisis data lowongan pekerjaan. Sistem ini memanfaatkan ribuan data deskripsi pekerjaan, kemudian menerapkan model *machine learning* yang mampu mengenali pola hubungan antara posisi kerja dan keterampilan yang paling sering disebut. Dari hasil proses tersebut, SkillRec dapat memberikan daftar keterampilan yang umumnya dibutuhkan dalam suatu bidang pekerjaan.

Meskipun memiliki arah yang hampir sama, CareerPath AI hadir dengan pendekatan yang lebih dekat dengan kebutuhan individu. Jika SkillRec hanya menampilkan hasil analisis dalam skala luas dan bersifat umum, CareerPath AI dirancang untuk memberikan rekomendasi yang bersifat personal. Setiap pengguna akan memperoleh saran keterampilan dan langkah pembelajaran yang disesuaikan dengan profil, minat, serta tujuan kariernya masing-masing.



Perbedaan lain terletak pada cakupan dan konteks data. SkillRec menggunakan data berbahasa Inggris yang bersumber dari pasar kerja global, sedangkan CareerPath AI berfokus pada data industri di Indonesia yang sering kali menggunakan bahasa campuran. Hal ini menjadikan sistem CareerPath AI lebih adaptif terhadap karakteristik lokal, termasuk istilah pekerjaan dan tren keterampilan yang khas di pasar tenaga kerja nasional.

Selain itu, CareerPath AI tidak hanya berhenti pada tahap mengenali keterampilan populer, tetapi juga membantu pengguna memahami bagaimana cara menguasainya melalui peta pembelajaran yang terarah. Dengan kombinasi analisis data industri dan sistem rekomendasi berbasis *Natural Language Processing*, CareerPath AI tidak sekadar memberi informasi, tetapi juga menyediakan panduan nyata untuk pengembangan karier.

### 2.3 Literatur Akademik

Bagian literatur akademik ini mengacu pada jurnal, prosiding, dan buku ilmiah yang terbit dalam lima tahun terakhir, yang membahas penerapan Artificial Intelligence (AI), khususnya *Data Science*, *Machine Learning*, dan *Natural Language Processing (NLP)*, dalam menganalisis data pekerjaan serta memetakan keterampilan digital yang dibutuhkan industri.

Menurut Provost dan Fawcett (2020), *Data Science* adalah disiplin multidisiplin yang menggabungkan statistik, pemrograman, dan analisis data untuk mengekstraksi wawasan dari kumpulan data besar. Pendekatan ini mendasari banyak penelitian modern yang bertujuan mengubah data mentah menjadi informasi yang dapat mendukung pengambilan keputusan berbasis bukti (*data-driven decision making*).

Dalam konteks pemetaan keterampilan kerja, Rahmawati et al. (2021) melakukan analisis terhadap data lowongan pekerjaan di bidang teknologi menggunakan teknik *text mining*. Penelitian mereka menemukan bahwa keterampilan seperti *Python*, *SQL*, dan *Data Analysis* merupakan kompetensi yang paling sering muncul dan sangat diminati oleh industri. Hasil ini menunjukkan bahwa pendekatan berbasis analisis teks mampu mengungkap kebutuhan keterampilan secara sistematis.

Sementara itu, Ong dan Lim (2023) melalui proyek *SkillRec* memperkenalkan sistem rekomendasi keterampilan yang menggunakan algoritma *machine learning* untuk mempelajari hubungan antara posisi kerja dan keterampilan yang sering muncul pada deskripsi lowongan. Sistem ini berhasil memberikan rekomendasi keterampilan yang relevan berdasarkan pola data. Namun, penelitian tersebut masih berfokus pada konteks global dan belum memperhatikan karakteristik pasar kerja lokal.

Laporan *The State of Skills: A Global View from Burning Glass Institute and Wiley* (Ehlers & Eigbrecht, 2024) menunjukkan bahwa keterampilan seperti AI/ML, Cloud Computing, Manajemen Produk, dan Media Sosial kini menjadi kebutuhan lintas

industri dan lokasi, bukan hanya di sektor teknologi saja. Ini mendukung kebutuhan sistem seperti CareerPath AI agar memperhitungkan tren skill global dan mengadaptasikannya pada konteks lokal.

Dari kajian terhadap berbagai literatur tersebut, dapat disimpulkan bahwa penelitian-penelitian sebelumnya telah berhasil memanfaatkan *AI* dan *Data Science* untuk memahami kebutuhan keterampilan di dunia kerja. Namun, sebagian besar masih berfokus pada konteks global, bersifat umum, dan belum memberikan rekomendasi yang disesuaikan dengan karakteristik pengguna maupun konteks lokal Indonesia. Oleh karena itu, penelitian ini mengembangkan CareerPath AI, sebuah sistem berbasis kecerdasan buatan yang dirancang untuk menganalisis data lowongan kerja berbahasa Indonesia dan memberikan rekomendasi keterampilan serta peta pembelajaran yang bersifat personal, adaptif, dan relevan dengan pasar kerja nasional.

## **BAB III METODOLOGI DAN SOLUSI**

### **3.1 Metodologi Perancangan**

#### **3.1.1 Pendekatan Penelitian**

Penelitian ini menggunakan pendekatan Research and Development (R&D) dengan metode prototyping untuk merancang dan mengembangkan sistem CareerPath AI. Pendekatan ini dipilih karena memungkinkan proses pengembangan yang cepat dan fleksibel, serta dapat menyesuaikan diri dengan kebutuhan pengguna. Penelitian ini dibagi menjadi beberapa tahapan:

- Studi Literatur – untuk mengumpulkan dan memahami teknik serta model yang telah digunakan dalam penelitian sebelumnya.
- Pengumpulan dan Analisis Data – mengumpulkan data lowongan kerja dari berbagai sumber online seperti LinkedIn, Kalibrr, dan JobStreet menggunakan web scraping atau API.
- Perancangan Sistem – merancang struktur dan komponen utama sistem yang akan digunakan.
- Implementasi dan Pengujian – membangun versi awal sistem (prototipe) dan melakukan uji coba untuk mengecek kualitas rekomendasi yang diberikan.

#### **3.1.2 Tahap Perancangan**

Berikut adalah tahapan perancangan sistem CareerPath AI yang dijelaskan dalam bentuk flowchart :



Penjelasan setiap tahapan:

#### 1. Pengumpulan data

Data lowongan kerja diambil dari berbagai portal pekerjaan seperti LinkedIn, Kalibrr, dan JobStreet.

Data tersebut diambil dengan cara web scraping atau menggunakan API bila tersedia.

#### 2. Pembersihan dan preprocessing data

Data yang dikumpulkan akan dicuci terlebih dahulu.

Langkah ini meliputi menghapus data yang duplikat atau tidak relevan, mengubah teks menjadi huruf kecil, menghilangkan tanda baca, memisahkan kata-kata, serta menghilangkan kata-kata umum (stopwords).

#### 3. Ekstraksi keterampilan dengan NLP

Penggunaan teknik NLP untuk menemukan keterampilan dalam teks lowongan kerja, yaitu: Named Entity Recognition (NER) – untuk mengenali kata-kata yang menyatakan keterampilan

Part-of-Speech Tagging – untuk mengidentifikasi kata benda yang menunjukkan keterampilan

Keyword Matching – menggunakan kamus keterampilan teknologi informasi yang telah disusun sebelumnya

#### 4. Pembentukan database keterampilan

Menyimpan daftar keterampilan, frekuensi kemunculan, kategori seperti Data Science dan Web Development, serta tingkat kesulitan keterampilan tersebut.

#### 5. Analisis tren dan peringkat keterampilan

Melakukan penghitungan frekuensi dan pertumbuhan permintaan keterampilan, serta menganalisis tren berdasarkan lokasi, level pengalaman, dan jenis industri.

#### 6. Sistem rekomendasi

Menggunakan metode Content-Based Filtering untuk mencocokkan profil mahasiswa dengan keterampilan yang relevan.

Input dari mahasiswa meliputi minat karir, keterampilan yang telah dikuasai, dan tingkat pendidikan.

#### 7. Pembuatan roadmap pembelajaran

Menyusun urutan belajar berdasarkan hubungan keterampilan yang saling terkait, serta merekomendasikan sumber belajar seperti kursus online, buku, atau proyek.

#### 8. Tampilan visual di dashboard

Menampilkan grafik tren keterampilan, roadmap belajar interaktif, serta rekomendasi karir yang sesuai.

#### 9. Evaluasi dan umpan balik

Melakukan pengujian terhadap akurasi rekomendasi menggunakan metrik precision dan recall.

Selain itu, juga dilakukan kuesioner kepuasan pengguna (mahasiswa).

### 3.1.3 Tools yang Digunakan

No	Tools yang digunakan	Fungsi
1	LinkedIn Job Trends	Mengambil data tren pekerjaan dan skill yang dibutuhkan industri.
2	Surprise Library	Membangun sistem rekomendasi skill berdasarkan preferensi mahasiswa lain.
3	VS Code	Menulis dan mengelola kode proyek CareerPath AI.
4	Hugging Face Transformers	Menggunakan model NLP untuk menganalisis deskripsi pekerjaan, CV, dan minat mahasiswa secara otomatis.
5	GitHub Pages / Netlify	Hosting frontend web aplikasi CareerPath AI secara gratis.

### 3.2 Solusi

### 3.2.1 Penjelasan Solusi Utama

Sistem CareerPath AI adalah alat bantu rekomendasi karir dan pembelajaran yang dirancang khusus untuk mahasiswa. Sistem ini memanfaatkan data lowongan kerja secara real-time untuk memberikan arahan yang tepat sesuai dengan kebutuhan industri digital di Indonesia.

Solusi ini bertujuan untuk:

- Memahami apa saja keterampilan yang paling dicari di pasar kerja digital saat ini.
- Memberikan rekomendasi belajar yang terstruktur dan sesuai dengan minat, tingkat kemampuan, dan tujuan karir pengguna.
- Membantu lulusan lebih relevan dengan kebutuhan industri, sehingga lebih mudah diterima kerja.

### 3.2.2 Cara Kerja Solusi

Sistem secara rutin mengumpulkan data lowongan kerja dari berbagai sumber yang dapat dipercaya.

Selanjutnya, mesin NLP (Natural Language Processing) menganalisis deskripsi pekerjaan tersebut untuk menemukan keterampilan yang dicari. Data keterampilan tersebut kemudian diproses, diurutkan, dan dinilai berdasarkan seberapa sering dan tingkat kebutuhannya. Setelah itu, sistem memproses profil pengguna, termasuk minat, keterampilan yang sudah dimiliki, dan tujuan karir.

Akhirnya, sistem menghasilkan jalur belajar yang mencakup:  
- Urutan materi belajar  
- Estimasi waktu yang dibutuhkan  
- Sumber belajar yang direkomendasikan  
- Proyek praktis untuk memperkuat portofolio

### 3.2.3 Manfaat Solusi

Bagi mahasiswa:

- Memudahkan dalam memilih jalur karir yang tepat
- Menghemat waktu dengan fokus pada keterampilan yang benar-benar dibutuhkan
- Meningkatkan kepercayaan diri dan kesiapan menghadapi dunia kerja

Bagi perguruan tinggi:

- Mendapatkan pemahaman tentang tren industri untuk memperbarui kurikulum
- Meningkatkan tingkat penyerapan lulusan
- Menjaga posisi sebagai institusi pendidikan yang responsif terhadap kebutuhan industri

Bagi industri:

- Mengurangi biaya pelatihan karyawan baru
- Memperoleh kandidat dengan kualitas yang lebih baik

### 3.2.4 Batasan solusi

- Bergantung pada keberadaan dan kualitas data lowongan kerja yang diambil.
- Rekomendasi perlu diperbarui secara berkala karena tren industri berubah cepat.
- Sistem hanya mempertimbangkan faktor teknis, sementara pencapaian dalam karir juga

dipengaruhi oleh soft skill dan jaringan.  
- Rekomendasi bersifat umum, sehingga mungkin kurang tepat untuk kasus yang sangat spesifik.

## **BAB IV**

### **HIPOTESIS HASIL**

#### **4.1. Prediksi Keluaran Utama**

Berdasarkan rancangan sistem yang telah dijabarkan pada Bab III, dihipotesiskan bahwa solusi CareerPath AI akan berfungsi sesuai desain awal, baik dari sisi arsitektur teknis maupun nilai fungsional pengguna.

Secara teknis, sistem ini mengimplementasikan pipeline Natural Language Processing (NLP) yang bertugas mengekstraksi dan memfilter kata kunci keterampilan dari kumpulan data deskripsi lowongan kerja. Tahap ini mencakup tokenization, part-of-speech tagging, dan named entity recognition untuk mengenali istilah yang mengindikasikan keterampilan teknis maupun non-teknis yang relevan di industri digital.

Output utama dari proses ini berupa ranked skill list, daftar keterampilan yang telah diidentifikasi dan diberi skor berdasarkan tingkat relevansi dan frekuensi kemunculannya dalam data industri. Data hasil ekstraksi ini kemudian dikonsumsi oleh modul content-based recommendation engine yang bertugas menyesuaikan hasil rekomendasi dengan profil pengguna.

Model pembelajaran mesin yang digunakan berada pada tahap baseline MVP, yakni algoritma Logistic Regression dan Random Forest secara paralel. Kedua model ini bertugas mempelajari pola hubungan antara atribut pengguna (minat, jurusan, pengalaman) dengan tren keterampilan yang ditemukan dari hasil analisis NLP.

Dari sisi fungsional, sistem diharapkan mampu menampilkan hasil analisis tersebut melalui interactive dashboard yang memvisualisasikan personal learning roadmap. Antarmuka ini berfungsi menampilkan rekomendasi keterampilan, urutan prioritas pembelajaran, serta tips pengembangan diri yang dihasilkan dari sistem analisis berbasis data.

Oleh karena itu, CareerPath AI diharapkan dapat berperan sebagai prototipe yang mampu membantu mahasiswa memahami arah pengembangan keterampilan mereka secara terukur dan kontekstual terhadap kebutuhan industri digital. Struktur modular dan desain yang lightweight juga memastikan sistem ini siap dikembangkan menuju fase beta dengan integrasi model deep learning embeddings di masa mendatang.

#### **4.2. Prediksi Pencapaian Tujuan**

Penulis berhipotesis bahwa sistem CareerPath AI mampu mencapai standar tujuan yang dijelaskan pada Bab I. Setiap komponen yang dikembangkan adalah fungsi inti yang dapat dioperasikan untuk menguji validitas konsep dan performa awal sistem.

- **Skill Identification Engine**

Modul NLP dibangun atas fondasi ekstraksi data keterampilan teknis dan non-teknis dari deskripsi lowongan kerja dengan tingkat akurasi yang tinggi. Proses ini mengintegrasikan proses tokenization, part-of-speech tagging, dan named entity recognition untuk mengenali modus istilah keterampilan. Pada tahap MVP, sistem menggunakan TF-IDF vectorization sebagai representasi fitur dan model Logistic Regression untuk mengidentifikasi relevansi keterampilan terhadap kategori pekerjaan. Hasil akhirnya adalah daftar keterampilan yang telah melampaui kalkulasi frekuensi dan bobot konteks.

- **Interactive Dashboard**

Komponen antarmuka dikembangkan dengan pendekatan data-driven visualization, yang menampilkan hasil analisis NLP dalam format learning roadmap. Peta ini menghubungkan antara bidang karier, keterampilan utama, serta keterampilan pendukung yang perlu dikuasai. Secara fungsional, antarmuka ini menjadi jembatan antara pengguna dan sistem analitik, sehingga mahasiswa dapat memahami jalur pengembangan keterampilan secara personal dan kontekstual.

- **Personalization Engine**

Modul rekomendasi berdasarkan content-based filtering ini dirancang untuk memasang kesesuaian profil pengguna dan tren industri. Dengan menerapkan algoritma content-based filtering, sistem menilai kesesuaian keterampilan berdasarkan fitur pengguna seperti bidang minat, latar belakang pendidikan, dan kemampuan yang sudah dimiliki. Pendekatan machine learning umum seperti Random Forest diprioritaskan untuk estimasi sederhana relevansi antar fitur sehingga lebih efisien secara komputasi serta mudah dikembangkan, ditambah dengan Logistic Regression. Akhirnya, secara perlahan diintegrasikan ke dalam deep learning melalui jaringan neuron.

Implementasi kasar model Random Forest dengan neural network (Python, Pandas, scikit-learn, TensorFlow/Keras, Matplotlib, NumPy):

## # SETUP DATA

```
# Skill dataset hasil ekstraksi NLP dari deskripsi lowongan kerja
```

```
# Tujuan:
```

```
# (1) Klasifikasi skill: in_demand (1) atau tidak (0)
```

```
# (2) Prediksi nilai demand_score (0..1) untuk ranking skill
```

```
# Fitur utama:
```

```
# skill_token_count, job_count_with_skill, avg_job_level,
```

```
# skill_cooccurrence_score, industry_vector_*, demand_score,
```

```
# in_demand
```

```

# Target:

# in_demand (binary) --> klasifikasi

# demand_score (float) --> regresi

# RF CLASSIFIER

from sklearn.ensemble import RandomForestClassifier

from sklearn.model_selection import train_test_split

from sklearn.metrics import accuracy_score, roc_auc_score,
confusion_matrix

x = df_skills.drop(columns=['skills', 'demand_score',
'in_demand']).fillna(0)

y = df_skills['in_demand']

train_X, val_X, train_Y, val_Y = train_test_split(X, Y,
test_size=0.25, stratify=y, random_state=42)

rf = RandomForestClassifier(n_estimators=200, min_samples_leaf=5,
random_state=1)

rf.fit(train_X, train_y)

train_pred = rf.predict(train_X)

val_pred = rf.predict(val_X)

val_proba = rf.predict_proba(val_X)[:,1]

acc_train = accuracy_score(train_y, train_pred)

acc_val = accuracy_score(val_y, val_pred)

roc_val = roc_auc_score(val_y, val_proba)

print(f"Train_acc={acc_train:.4f} | Val_acc={acc_val:.4f} | ROC-
AUC={roc_val:.4f}")

# NEURAL NETWORK REGRESSOR

from tensorflow import keras

from tensorflow.keras import layers

```



```

from sklearn.preprocessing import StandardScaler

from sklearn.metrics import mean_squared_error, r2_score

X = df_skills.drop(columns=['skill','in_demand']).fillna(0).values

Y = df_skills['demand_score'].values

scaler = StandardScaler()

X_scaled = scaler.fit_transform(X)

train_X, val_X, train_y, val_y = train_test_split(X_scaled, y,
test_size=0.25, random_state=42)

dnn = keras.Sequential([

layers.Dense(128, activation='relu',
input_shape=[train_X.shape[1]]),layers.Dropout(0.2),
layers.Dense(64, activation='relu'), layers.Dense(1,
activation='linear')

])

dnn.compile(optimizer='adam', loss='mse', metrics=['mae']) history =
dnn.fit(train_X, train_y, validation_data=(val_X,val_y), epochs=80,
batch_size=64, verbose=1)

val_pred = dnn.predict(val_X).flatten() mse =
mean_squared_error(val_y, val_pred) r2 = r2_score(val_y, val_pred)
print(f"Val_MSE={mse:.4f} | R2={r2:.4f}")

```

## **# LOSSES**

# Random Forest:

# - Loss internal: Gini Impurity

# Neural Network:

# - Loss: Mean Squared Error (MSE)

## **# DEMAND TABLE EXAMPLE**

```
import pandas as pd
```

```
sample_output = pd.DataFrame({ 'Skill': ['Python', 'SQL',
'Communication', 'TensorFlow', 'Excel', 'Data Visualization'],
```

```
'Demand_Score': [0.91, 0.85, 0.74, 0.72, 0.68, 0.63], 'In_Demand':  
[1, 1, 1, 1, 1, 0], 'Recommendation': [ 'Core Skill - Prioritize  
mastery + project implementation', 'Core Skill - Strengthen database  
understanding', 'Soft Skill - Practice presentation & teamwork',  
'Tech Skill - Build model demo projects', 'Support Skill -  
Complement analysis workflow', 'Watchlist - Learn basic tools later'  
] }) sample_output
```

## # PIPELINE SUMMARY

```
# 1. Extract job text -> tokenize -> detect skills  
  
# 2. Aggregate per-skill: count, cooccurrence, industry_vector,  
avg_level  
  
# 3. Compute demand_score = normalize(count + weight + tfidf)  
  
# 4. in_demand = (demand_score >= 0.6) ? 1 : 0  
  
# 5. Train RFClassifier -> predict in_demand  
  
# 6. Train NeuralNet Regressor -> predict demand_score  
  
# 7. Combine both -> personalized skill roadmap
```

## 4.3. Kesesuaian dengan Kajian Pustaka

Pendekatan yang diterapkan pada CareerPath AI sejalan dengan teori text mining dan information extraction yang dijelaskan oleh Jurafsky & Martin (2023), di mana tahapan seperti tokenization, entity recognition, dan analisis frekuensi kata digunakan untuk mengekstraksi informasi bermakna dari data teks berskala besar.

Selain itu, rancangan sistem rekomendasi berbasis konten pada proyek ini selaras dengan penelitian Ong & Lim (2023) melalui SkillRec, yang memanfaatkan analisis data lowongan kerja dan model pembelajaran mesin untuk mengidentifikasi keterampilan industri yang relevan.

CareerPath AI memperluas pendekatan tersebut dengan menambahkan elemen personalisasi pengguna dan konteks industri digital Indonesia, sehingga memperkuat validitas teoritis sekaligus menawarkan solusi praktis dalam menjembatani kesenjangan antara dunia pendidikan dan kebutuhan industri modern.

## DAFTAR PUSTAKA

Harus menggunakan APA Style

Elcholiqi, A., & Musdholifah, A. (2023). Prototipe integrasi chatbot dalam sistem informasi akademik berbasis web menggunakan metode Natural Language Processing (studi kasus di IST AKPRIND Yogyakarta). *Jurnal Script: Jurnal Komputer dan Informatika*, 11(2), 123–131.

<https://ejournal.akprind.ac.id/index.php/script/article/view/4805>

Ramli, M., & Arifin, M. (2025). Aplikasi chatbot berbasis Dialogflow dan NLP untuk pelayanan informasi akademik pada Fakultas Ilmu Komputer Universitas Tomakaka. *Asian Business Research Journal (ABRJ)*, 6(1), 44–52.

<https://www.journal.stieamsir.ac.id/index.php/abrj/article/view/610>

Ko, H., Lee, S., Park, Y., & Choi, A. (2022). A Survey of Recommendation Systems: Recommendation Models, Techniques, and Application Fields. *Electronics*, 11(1), 141.

<https://www.mdpi.com/2079-9292/11/1/141>

Zangerle, E., Bauer, C., & others. (2022). Evaluating Recommendation Systems: Survey and Framework. *ACM Computing Surveys*, 55(8), Article 170.

<https://dl.acm.org/doi/full/10.1145/3556536>

Author(s). (2021). A Study on using data clustering for feature extraction to improve the quality of classification. *Knowledge and Information Systems*, 63, 1771-1805.

<https://link.springer.com/article/10.1007/s10115-021-01572-6>

Liu, J., Shen, Z., He, Y., Zhang, X., Xu, R., Yu, H., & Cui, P. (2021). Towards Out-Of-Distribution Generalization: A Survey. *arXiv*.

<https://arxiv.org/abs/2108.13624>

Ong, Y. H., & Lim, K. H. (2023). SkillRec: A data-driven approach to job skill recommendation for career insights. *Proceedings of the 2023 5th International*

Conference on Computer, Automation and Artificial Intelligence (ICCAE 2023).  
Singapore University of Technology and Design.  
[https://people.sutd.edu.sg/~kwanhui\\_lim/publications/2023-ICCAE-skillRecom.pdf](https://people.sutd.edu.sg/~kwanhui_lim/publications/2023-ICCAE-skillRecom.pdf)

Provost, F., & Fawcett, T. (2020). Data science for business: What you need to know about data mining and data-analytic thinking (2nd ed.). O'Reilly Media.  
Rahmawati, A., Pratama, D., & Nugroho, S. (2021). Analisis lowongan pekerjaan untuk pemetaan keterampilan digital. Jurnal Teknologi Informasi dan Komputer, 7(2), 145–152.

Ehlers, U.-D., & Eigbrecht, L. (2024). The state of skills: A global view from Burning Glass Institute and Wiley. In Ehlers, U.-D., & Eigbrecht, L. (Eds.), Creating the University of the Future (Chapter 25, pp. 509-512). Springer VS.  
[https://doi.org/10.1007/978-3-658-42948-5\\_25](https://doi.org/10.1007/978-3-658-42948-5_25)

## **LAMPIRAN**

Wajib melampirkan :

- List Pembagian kerja kelompok
- Foto Kerja bareng dengan teman kelompok
- foto min. 2 kali dengan mentor (2 kali konsultasi dengan mentor).