



**RANCANG BANGUN APLIKASI *CHATBOT*
MENGUNAKAN TEKNOLOGI OCR DAN MODEL
LLAMA UNTUK REKOMENDASI PEKERJAAN**

SKRIPSI

GERRY SATRIA HALIM 2107411028

**PROGRAM STUDI TEKNIK INFORMATIKA
JURUSAN TEKNIK INFORMATIKA DAN KOMPUTER
POLITEKNIK NEGERI JAKARTA**

2025



**RANCANG BANGUN APLIKASI *CHATBOT*
MENGUNAKAN TEKNOLOGI OCR DAN MODEL
LLAMA UNTUK REKOMENDASI PEKERJAAN**

SKRIPSI

**Dibuat untuk Melengkapi Syarat-Syarat yang Diperlukan
untuk Memperoleh Diploma Empat Politeknik**

GERRY SATRIA HALIM

2107411028

**PROGRAM STUDI TEKNIK INFORMATIKA
JURUSAN TEKNIK INFORMATIKA DAN KOMPUTER
POLITEKNIK NEGERI JAKARTA**

2025



SURAT PERNYATAAN BEBAS PLAGIARISME

Saya yang bertanda tangan di bawah ini:

Nama : Gerry Satria Halim
NIM : 2107411028
Jurusan/Program Studi : Teknik Informatika dan Komputer/Teknik Informatika
Judul Skripsi : Rancang Bangun Aplikasi *Chatbot* Menggunakan Teknologi OCR dan Model *LLaMA* Untuk Rekomendasi Pekerjaan

Menyatakan dengan sebenarnya bahwa skripsi ini benar-benar merupakan hasil dari karya dari saya sendiri, bebas dari peniruan terhadap karya dari orang lain. kutipan pendapat dan tulisan orang lain ditunjuk sesuai dengan cara-cara penulisan karya ilmiah yang berlaku.

Apabila di kemudian hari ini terbukti atau dapat dibuktikan bahwa dalam skripsi ini terkandung ciri-ciri plagiat dan bentuk-bentuk peniruan lain yang dianggap melanggar peraturan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Depok, 10 Juni 2025

Yang membuat pernyataan



Gerry Satria Halim

NIM 2107411028

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

LEMBAR PENGESAHAN

Skripsi diajukan oleh:

Nama : Gerry Satria Halim
NIM : 2107411028
Program Studi : Teknik Informatika
Judul Skripsi : Rancang Bangun Aplikasi *Chatbot* Menggunakan Teknologi OCR dan Model *LLaMA* Untuk Rekomendasi Pekerjaan

Telah diuji oleh tim penguji dalam Sidang Skripsi pada Hari Senin, Tanggal 23 Bulan Juni, Tahun 2025 dan dinyatakan **LULUS**.

Disahkan oleh

Pembimbing I : Rizki Elisa Nalawati S.T., M.T. (Rizki)

Penguji I : Dr. Dewi Yanti Liliana, S.Kom., M.Kom (Dewi)

Penguji II : Zahra Azizah, S.Kom., M.I.S. (Zahra)

Penguji III : Susana Dwi Yulianti, M.Kom. (Susana)

Mengetahui

Jurusan Teknik Informatika dan Komputer
Ketua



Dr. Anita Hidayati, S.Kom., M.Kom.
NIP. 197908032003122003

© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



KATA PENGANTAR

Puji syukur penulis panjatkan kehadirat Tuhan Yang Maha Esa, karena berkat rahmat dan karunia-Nya, penulis dapat menyelesaikan laporan akhir skripsi yang berjudul “Rancang Bangun Aplikasi *Chatbot* Menggunakan Teknologi OCR dan Model *LLaMA* untuk Rekomendasi Pekerjaan.” Penulisan laporan akhir skripsi ini dilakukan dalam rangka memenuhi persyaratan untuk mencapai gelar Sarjana Terapan pada Program Studi Teknik Informatika di Politeknik Negeri Jakarta.

Dalam proses penyusunan laporan akhir skripsi ini, penulis menyadari sepenuhnya bahwa terselesaikannya laporan ini tidak terlepas dari dukungan, motivasi, semangat, serta bimbingan dari berbagai pihak, baik yang bersifat moril maupun materil. Oleh karena itu, penulis ingin menyampaikan ucapan terima kasih kepada:

1. Ibu Dr. Anita Hidayati, S.Kom., M.Kom., selaku Ketua Jurusan Teknik Informatika dan Komputer.
2. Ibu Euis Oktavianti, S.Si., M.TI., selaku Kepala Program Studi Teknik Informatika.
3. Ibu Rizki Elisa Nalawati S.T., M.T., selaku dosen pembimbing yang telah dengan sabar membimbing, mengarahkan, serta membantu dalam proses penyelesaian skripsi ini.
4. Seluruh Bapak/Ibu guru atau dosen yang sudah membimbing, mengajarkan, dan mendidik penulis sehingga penulis memiliki pengetahuan yang baik.
5. Orang tua serta seluruh keluarga penulis yang selalu memberikan dukungan, doa, serta semangat selama penulis mengerjakan skripsi.
6. Semua teman seperjuangan penulis yang telah memberikan bantuan, semangat dan dukungan kepada penulis selama mengerjakan skripsi.
7. Semua pihak yang telah membantu secara langsung maupun tidak langsung dalam penyelesaian skripsi ini.

Penulis berharap laporan ini dapat memberikan manfaat bagi semua pihak, baik dari segi penelitian maupun pendidikan. Penulis menyadari bahwa penyusunan laporan akhir skripsi ini masih terdapat banyak kekurangan. Oleh karena itu, penulis sangat mengharapkan kritik dan saran yang membangun agar laporan ini menjadi lebih baik.

Hak Cipta :
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Depok, 10 Juni 2025

Penulis,

Gerry Satria Halim





© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

SURAT PERNYATAAN PERSETUJUAN PUBLIKASI
SKRIPSI UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademik Politeknik Negeri Jakarta, saya bertanda tangan di bawah ini:

Nama : Gerry Satria Halim

NIM : 2107411028

Jurusan/Program Studi : Teknik Informatika dan Komputer/Teknik Informatika

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Politeknik Negeri Jakarta Hak Bebas Royalti Non-Eksklusif atas karya ilmiah saya yang berjudul:

**RANCANG BANGUN APLIKASI *CHATBOT* MENGGUNAKAN
TEKNOLOGI OCR DAN MODEL *LLAMA* UNTUK REKOMENDASI
PEKERJAAN**

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Non-Eksklusif ini Politeknik Negeri Jakarta berhak menyimpan, mengalihmediakan /formatkan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan skripsi saya tanpa meminta izin dari saya selama tetap mencantumkan nama saya sebagai penulis pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Depok, 10 Juni 2025

Yang Menyatakan,



Gerry Satria Halim

NIM 2107411028



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

RANCANG BANGUN APLIKASI *CHATBOT* MENGGUNAKAN TEKNOLOGI OCR DAN MODEL *LLAMA* UNTUK REKOMENDASI PEKERJAAN

ABSTRAK

Kesulitan dalam merancang perencanaan karier sering kali muncul karena kurangnya pemahaman mengenai potensi diri, minimnya informasi terkait peluang karier, serta kebingungan dalam memilih jalur yang sesuai. Hal ini mengakibatkan banyak individu, khususnya mahasiswa, mengambil keputusan karier yang tidak optimal. Berdasarkan survei awal pada 32 responden, ditemukan bahwa 34,4% merasa ragu dalam menentukan pilihan pekerjaan, dengan tantangan terbesar adalah merasa kemampuan yang dimiliki belum cukup (53,1%). Sebanyak 68,8% responden menyatakan kebutuhan akan adanya rekomendasi pekerjaan yang relevan sesuai minat dan kompetensi. Untuk mengatasi permasalahan tersebut, penelitian ini mengembangkan sebuah aplikasi chatbot berbasis Android yang mampu memberikan rekomendasi pekerjaan dalam Bahasa Indonesia. Sistem ini mengimplementasikan beberapa model kecerdasan buatan, yaitu klasifikasi dokumen CV menggunakan Long Short-Term Memory (LSTM), ekstraksi kata kunci dengan KeyBERT, OCR dengan EasyOCR, dan Large Language Model menggunakan Llama 3 70B untuk menghasilkan respons rekomendasi. Hasil pengujian menunjukkan model klasifikasi CV mencapai akurasi sebesar 98,75%, dengan nilai precision 0.97, recall 1.0, dan F1-score 0.988. Selain itu, pengujian kebergunaan aplikasi menghasilkan skor System Usability Scale (SUS) sebesar 76.58 dan Net Promoter Score (NPS) sebesar 52.63, yang mengindikasikan bahwa aplikasi ini dapat diterima dengan baik dan memiliki tingkat rekomendasi yang tinggi dari pengguna.

Kata Kunci : *Android, Chatbot, Deep Learning, Large Language Model, Natural Language Processing.*



DAFTAR ISI

SURAT PERNYATAAN BEBAS PLAGIARISME.....	i
LEMBAR PENGESAHAN	ii
KATA PENGANTAR	iii
SURAT PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS	v
ABSTRAK	vi
DAFTAR ISI	vii
DAFTAR TABEL	x
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN	xii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah.....	3
1.4 Tujuan dan Manfaat	3
1.4.1 Tujuan Penelitian	3
1.4.2 Manfaat Penelitian	4
1.5 Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA.....	6
2.1 Pencari Pekerjaan	6
2.2 Chatbot	6
2.3 Android	7
2.4 Large Language Model (LLM).....	7
2.4.1 GPT	7
2.4.2 LLaMA	8
2.4.3 Mistral	9
2.5 Optical Character Recognition (OCR).....	10
2.6 Long Short Term Memory (LSTM).....	10
2.7 Senternce Transformer.....	11
2.8 AI Project Lifecycle.....	12
2.8.1 Problem Definition.....	12
2.8.2 Data Acquisition and Preparation	12
2.8.3 Model Development and Training	13
2.8.4 Model Evaluation and Refinement	13
2.8.5 Deployment	13
2.8.6 ML Ops.....	13
2.9 Unified Modeling Language (UML).....	13
2.9.1 Use Case Diagram	14
2.9.2 Activity Diagram	15
2.10 Explainable AI (XAI).....	16
2.10.1 Local Interpretable Model-agnostic Explanations (LIME).....	16
2.10.2 SHapley Addictive exPlanations (SHAP)	17
2.11 Hallucination AI.....	17
2.12 FastAPI.....	19
2.13 Black Box Testing.....	19

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

2.14	System Usability Scale (SUS)	19
2.15	Net Promoter Score (NPS)	20
2.16	Penelitian Terdahulu	21
BAB III METODE PENELITIAN		26
3.1	Rancangan Penelitian	26
3.2	Tahapan Penelitian	26
3.2.1	Pendefinisian Masalah (<i>Problem Definition</i>)	27
3.2.2	Pengumpulan Kebutuhan Pengguna (<i>Requirement Gathering</i>)	27
3.2.3	Desain Sistem (<i>System Design</i>)	28
3.2.4	Pengembangan Model (<i>Model Development</i>)	28
3.2.5	Implementasi Sistem (<i>System Implementation</i>)	34
3.2.6	Pengujian Sistem (<i>System Testing and Evaluation</i>)	34
3.2.7	<i>Deployment</i>	35
3.2.8	Penyusunan Laporan Akhir (<i>Final Report</i>)	35
3.3	Objek Penelitian	36
BAB IV HASIL DAN PEMBAHASAN		37
4.1	<i>Problem Definition</i>	37
4.2	<i>Requirement Gathering</i>	37
4.2.1	Studi Literatur	37
4.2.2	Survei Kuesioner	38
4.2.3	Analisis Kebutuhan Aplikasi	39
4.3	<i>System Design</i>	44
4.3.1	<i>Use Case Diagram</i>	45
4.3.2	<i>Activity Diagram</i>	45
4.3.3	<i>Database</i>	53
4.3.4	<i>Architecture Application</i>	54
4.3.5	<i>Wireframe Design</i>	55
4.4	<i>Model Development</i>	55
4.4.1	<i>Data Acquisition and Preparation</i>	56
4.4.2	<i>Model Development and Training</i>	63
4.4.3	<i>Model Evaluation and Refinement</i>	75
4.5	<i>System Implementation</i>	81
4.5.1	Halaman Landing	82
4.5.2	Halaman Sign Up	83
4.5.3	Halaman Login	84
4.5.4	Halaman Home	85
4.5.5	Halaman Chat	86
4.5.6	Halaman Profil	87
4.6	<i>System Testing</i>	87
4.6.1	<i>Black Box Testing</i>	88
4.6.2	<i>System Usability Scale (SUS)</i>	93
4.6.3	<i>Net Promoter Score (NPS)</i>	97
4.7	<i>Deployment</i>	100
BAB V PENUTUP		102
5.1	Kesimpulan	102
5.2	Saran	103
DAFTAR PUSTAKA		104
DAFTAR RIWAYAT HIDUP PENULIS		109



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

LAMPIRAN110



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



DAFTAR TABEL

Tabel 2.1 Simbol <i>Use Case Diagram</i>	14
Tabel 2.2 Simbol <i>Activity Diagram</i>	15
Tabel 2.3 Standar Interpretasi Skor dan <i>Grade SUS</i>	20
Tabel 2.4 Parameter <i>Net Promoter Score</i>	21
Tabel 2.5 Penelitian Terdahulu	21
Tabel 4.1 Referensi Teknologi dari Sumber Literatur	38
Tabel 4.2 Kebutuhan Dataset Klasifikasi CV dan Non-CV	40
Tabel 4.3 Kebutuhan Dataset Kategori Pekerjaan	41
Tabel 4.4 Kebutuhan Model	41
Tabel 4.5 Kebutuhan <i>Pre-trained Model</i>	42
Tabel 4.6 Kebutuhan Fungsional	43
Tabel 4.7 Kebutuhan Non-Fungsional	44
Tabel 4.8 Konfigurasi <i>hyperparameter tuning</i>	66
Tabel 4.9 Hasil Eksperimen <i>Hyperparameter Tuning</i>	67
Tabel 4.10 Metode Ekstraksi <i>Keyword</i>	70
Tabel 4.11 Hasil Percobaan Metode Ekstraksi <i>Keyword</i>	71
Tabel 4.12 Model <i>Sentence Transformer</i>	72
Tabel 4.13 Hasil Percobaan Model <i>Sentence Transformer</i>	72
Tabel 4.16 Contoh Hasil <i>Prompt LLM</i>	75
Tabel 4.15 Parameter Model Klasifikasi Dokumen CV dan Non-CV	76
Tabel 4.16 Hasil Evaluasi <i>Confusion Matrix</i>	76
Tabel 4.17 Perbandingan Hasil Implementasi EasyOCR	79
Tabel 4.18 Percobaan Nilai <i>Temperature</i>	80
Tabel 4.19 Skenario <i>Black Box Testing</i>	88
Tabel 4.20 Hasil Pengujian <i>Black Box</i>	91
Tabel 4.21 Pertanyaan SUS	93
Tabel 4.22 Hasil SUS Bagian 1	94
Tabel 4.23 Hasil SUS Bagian 2	95
Tabel 4.24 Nilai SUS	96
Tabel 4.25 Pertanyaan NPS	98
Tabel 4.26 Hasil NPS	98
Tabel 4.27 Perhitungan Nilai NPS	99

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



DAFTAR GAMBAR

Gambar 2.1	Arsitektur GPT (Sumber: GeeksforGeeks 2024)	8
Gambar 2.2	Arsitektur <i>LlaMA</i> (Sumber: Yaadav, 2024)	9
Gambar 2.3	Arsitektur <i>Mistral</i> 7B (Sumber: Smith, 2024)	10
Gambar 2.4	Arsitektur model LSTM (Sumber: Calzone, 2022)	11
Gambar 2.5	Tahapan <i>AI Project Lifecycle</i> (Sumber: Saltz, 2023)	12
Gambar 3.1	Tahapan Penelitian	27
Gambar 3.2	Tahapan Pengumpulan dan Persiapan Dataset Dokumen CV	29
Gambar 3.3	Tahapan Pengumpulan dan Persiapan Dataset Kategori Pekerjaan	31
Gambar 4.1	<i>Use Case Diagram</i>	45
Gambar 4.2	<i>Activity Diagram Sign Up</i>	46
Gambar 4.3	<i>Activity Diagram Login</i>	47
Gambar 4.4	<i>Activity Diagram Chat History</i>	48
Gambar 4.5	<i>Activity Diagram Mendapatkan Rekomendasi Pekerjaan</i>	49
Gambar 4.6	<i>Activity Diagram Delete Chat</i>	50
Gambar 4.7	<i>Activity Diagram Delete All Chat</i>	51
Gambar 4.8	<i>Activity Diagram Logout</i>	52
Gambar 4.9	<i>ER Diagram</i>	53
Gambar 4.10	Arsitektur Aplikasi	54
Gambar 4.11	Ilustrasi Pengumpulan data	57
Gambar 4.12	Ilustrasi Ekstraksi Teks	58
Gambar 4.13	Ilustrasi Pelabelan Data	58
Gambar 4.14	Alur Pelatihan Model Klasifikasi Dokumen	64
Gambar 4.15	Alur Implementasi Teknologi OCR	69
Gambar 4.16	Alur <i>Generate Respons LLM</i>	74
Gambar 4.17	Implementasi Metode LIME	78
Gambar 4.18	Implementasi Metode SHAP	78
Gambar 4.19	Halaman <i>Landing</i>	82
Gambar 4.20	Halaman <i>Register</i>	83
Gambar 4.21	Halaman <i>Login</i>	84
Gambar 4.22	Halaman <i>Home</i>	85
Gambar 4.23	Halaman <i>Chat</i>	86
Gambar 4.24	Halaman <i>Profil</i>	87

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritrik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

DAFTAR LAMPIRAN

Lampiran 1 <i>Prompt LLM</i>	110
Lampiran 2 <i>Wireframe Design</i>	111
Lampiran 3 Hasil Survei Identifikasi Permasalahan Pengguna	112
Lampiran 4 Dokumentasi Pengujian Aplikasi	113
Lampiran 4 Dokumentasi Validasi Data Aplikasi	114
Lampiran 5 Hasil <i>System Usability Scale</i> (SUS)	115
Lampiran 6 Hasil <i>Net Promoter Score</i> (NPS)	119





Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

BAB I PENDAHULUAN

1.1 Latar Belakang

Kesulitan dalam merancang perencanaan karier sering kali muncul karena kurangnya pemahaman mengenai potensi diri, minimnya informasi terkait peluang karir, serta kebingungan dalam memilih jalur yang sesuai dengan minat, dan bakat mereka. Akibatnya, tidak sedikit dari mereka yang mengambil keputusan karir secara tidak optimal, yang dapat berdampak pada ketidaksesuaian antara pekerjaan yang dijalani dengan kemampuan dan minat yang dimiliki (Luid et al., 2024).

Berdasarkan hasil survei yang dikumpulkan dengan menggunakan kuesioner pada penelitian ini, sebanyak 34,4 % dari 32 responden, yang sebagian besar merupakan mahasiswa merasa masih ragu dalam menentukan pilihan pekerjaan yang mereka inginkan. Tantangan terbesar yang mereka hadapi dalam memilih pekerjaan adalah merasa bahwa kemampuan yang dimiliki belum cukup untuk menentukan jenis pekerjaan yang sesuai, dengan persentase sebesar 53,1 %. Dari permasalahan tersebut, sebanyak 68,8 % responden ingin mendapatkan rekomendasi pekerjaan yang relevan sesuai dengan minat dan kompetensi yang dimiliki. Untuk mengatasi permasalahan tersebut, diperlukan sebuah inovasi teknologi yang mampu memberikan rekomendasi pekerjaan berdasarkan preferensi individu. Beberapa teknologi yang dapat digunakan untuk menangani permasalahan ini antara lain adalah sistem pendukung keputusan dan *chatbot*. Namun, pada penelitian ini akan dikembangkan sebuah aplikasi *chatbot* yang berfungsi sebagai sistem rekomendasi pekerjaan.

Chatbot merupakan salah satu penerapan dari *Natural Language Processing* yang memiliki keunggulan dalam kemudahan implementasi, fleksibilitas dalam penggunaan, serta kemampuan yang semakin meningkat dalam meniru percakapan manusia. Namun, di sisi lain, *chatbot* masih memiliki kelemahan, yaitu kurang mampu memahami konteks percakapan, untuk mengatasi keterbatasan ini, model yang lebih kecil, fleksibel, dan tidak terlalu bergantung pada domain tertentu akan sangat bermanfaat (Caldarini, Jaf and McGarry, 2022). Salah satu model fleksibel



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

yang dapat dijadikan sebagai sebuah pilihan adalah *Large Language Model* (LLM), contohnya seperti GPT, *Mistral*, dan *LLaMA* (Hou and Lian, 2024). Model GPT memiliki kelebihan dalam memahami dan menghasilkan teks yang menyerupai tulisan manusia, namun memiliki kekurangan dalam memahami konteks yang lebih dalam dan cenderung menghasilkan data yang tidak akurat (Brown et al., 2020).

Model *Mistral* memiliki kelebihan dalam pemrosesan data linguistik, namun memiliki kekurangan akurasi linguistik yang membuat model ini perlu dilakukan penyempurnaan lebih lanjut. Sementara itu, model *LLaMA* memiliki kelebihan dalam tingkat adaptabilitas dan efisiensi yang tinggi dalam memproses berbagai kumpulan data linguistik. namun memiliki kekurangan dalam kemungkinan inkonsistensi dalam kinerja di berbagai bahasa (Hou and Lian, 2024). Selain itu, model *LLaMA* ini bersifat *open source* dan juga memiliki model yang ringan untuk dijalankan di berbagai perangkat seperti perangkat *mobile* (Meta, 2022). Oleh karena itu, penelitian ini menggunakan model *LLaMA* untuk mengembangkan *chatbot* yang dapat menghasilkan rekomendasi pekerjaan berdasarkan input teks dari pengguna. Selain itu, teknologi *Optical Character Recognition* (OCR) juga diterapkan untuk mengekstraksi data teks dari CV yang diunggah oleh pengguna dalam bentuk gambar atau dokumen berformat .pdf.

Penelitian ini bertujuan untuk membuat sebuah aplikasi *chatbot* berbasis android yang dapat memberikan rekomendasi pekerjaan yang inputnya dapat berupa teks, gambar, atau dokumen dengan format .pdf untuk mengatasi permasalahan kurangnya pemahaman mengenai potensi diri, minimnya informasi terkait peluang pekerjaan, serta kebingungan dalam memilih pekerjaan yang sesuai dengan minat, dan bakat. Dengan memanfaatkan model *LLaMA*, aplikasi ini diharapkan dapat membantu setiap individu yang membutuhkan sebuah rekomendasi pekerjaan berdasarkan minat, dan bakat yang dimiliki.

1.2 Rumusan Masalah

Berdasarkan uraian latar belakang di atas, perumusan masalah dalam penelitian ini adalah:



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

- a. Bagaimana membangun aplikasi *chatbot* berbasis android yang dapat memberikan rekomendasi pekerjaan ?
- b. Bagaimana implementasi model *LLaMA* dan teknologi OCR untuk membangun *chatbot* yang dapat memberikan rekomendasi pekerjaan berdasarkan input pengguna yang dapat berupa teks, gambar, atau dokumen dengan format .pdf ke dalam aplikasi berbasis android ?

1.3 Batasan Masalah

Agar penelitian ini lebih terarah, terdapat beberapa batasan masalah yang ditetapkan, yaitu:

- a. Aplikasi *chatbot* dibuat berbasis android dengan menggunakan bahasa pemrograman Kotlin.
- b. Model *Large Language Model* (LLM) yang digunakan adalah model *LLaMA*.
- c. Teknologi *Optical Character Recognition* (OCR) yang diimplementasikan pada aplikasi menggunakan *library* EasyOCR.
- d. Aplikasi hanya mengolah input berupa teks, gambar dengan format .jpg, .jpeg, .png, dan dokumen dengan format .pdf.
- e. Dokumen yang dapat diproses merupakan dokumen CV dan memiliki maksimal 2 halaman.
- f. Hasil yang dikeluarkan oleh aplikasi berupa respons teks yang berisikan rekomendasi pekerjaan.

1.4 Tujuan dan Manfaat

Tujuan dan manfaat dari penelitian ini menjelaskan apa yang ingin dicapai serta manfaat yang dapat diperoleh dari hasil penelitian ini.

1.4.1 Tujuan Penelitian

Berikut adalah tujuan dari penelitian ini, yakni:

- a. Membangun aplikasi *chatbot* berbasis android yang dapat memberikan rekomendasi pekerjaan.



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

- b. Melakukan implementasi model *LLaMA* dan teknologi OCR untuk membangun *chatbot* yang dapat memberikan rekomendasi pekerjaan berdasarkan input pengguna yang dapat berupa teks, gambar dengan format .jpg, .jpeg, .png, atau dokumen dengan format .pdf ke dalam aplikasi berbasis android.

1.4.2 Manfaat Penelitian

Berikut adalah manfaat dari penelitian ini, yakni:

- a. Bagi pencari kerja, hasil penelitian ini diharapkan dapat membantu dalam pengambilan keputusan yang lebih tepat terkait jenis pekerjaan yang sesuai dengan kemampuan dan minat yang dimiliki.
- b. Bagi peneliti, hasil penelitian ini diharapkan dapat menjadi acuan dalam pengembangan aplikasi *chatbot* yang memanfaatkan implementasi model *LLaMA* dan teknologi OCR.

1.5 Sistematika Penulisan

Sistematika penulisan ini dibuat untuk memudahkan dalam melakukan penulisan laporan skripsi, maka perlu dibuat sebuah sistematika penulisan yang tepat dan benar. Sistematika penulisan dibagi menjadi beberapa bab di antaranya adalah sebagai berikut:

BAB I PENDAHULUAN

Bab ini membahas mengenai latar belakang penelitian, perumusan dan batasan masalah, tujuan yang ingin dicapai, manfaat penelitian, serta sistematika dalam penulisan laporan.

BAB II TINJAUAN PUSTAKA

Bab ini memuat pembahasan mengenai penelitian-penelitian sebelumnya serta teori-teori yang berkaitan dan mendukung penelitian yang dilakukan.

BAB III METODE PENELITIAN

Bab ini membahas mengenai metode yang digunakan dalam penelitian, yang mencakup pembahasan mengenai rancangan penelitian, tahapan penelitian, serta objek penelitian.



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

BAB IV HASIL DAN PEMBAHASAN

Bab ini membahas mengenai perancangan aplikasi *chatbot* yang memberikan rekomendasi pekerjaan dengan implementasi model *LLaMA* dan teknologi OCR beserta dengan hasil analisis pengujian yang dilakukan.

BAB V PENUTUP

Bab ini membahas mengenai kesimpulan akhir dari penelitian dan saran yang akan sangat berguna untuk penelitian selanjutnya.



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

BAB II

TINJAUAN PUSTAKA

2.1 Pencari Pekerjaan

Pencari kerja merupakan individu yang termasuk dalam angkatan kerja, baik yang sedang tidak memiliki pekerjaan maupun yang sudah bekerja namun memiliki keinginan untuk mendapatkan pekerjaan baru atau berpindah ke tempat kerja lain. Mereka secara aktif melakukan upaya pencarian kerja, baik dengan mendaftarkan diri pada lembaga atau instansi yang bertugas dalam penempatan tenaga kerja, maupun dengan mengajukan lamaran secara langsung kepada perusahaan atau pemberi kerja yang membuka lowongan (Ningsih and Abdullah, 2021). Setiap pencari kerja memiliki latar belakang kemampuan, keterampilan, tingkat pendidikan, serta karakter pribadi yang beragam. Perbedaan-perbedaan ini sering kali tidak selaras dengan kualifikasi dan kebutuhan dari lowongan pekerjaan yang tersedia, sehingga tidak semua pelamar kerja akan sesuai atau memenuhi syarat untuk diterima mengisi posisi yang dibutuhkan (Rohaniyah and Purnamasari, 2023).

2.2 Chatbot

Chatbot merupakan teknologi yang dirancang untuk berkomunikasi dengan manusia menggunakan bahasa alami dengan tujuan meniru percakapan layaknya manusia, baik melalui teks maupun suara (Babu and Boddu, 2024). *Chatbot* menggunakan pembelajaran mesin untuk dapat memahami input pertanyaan dari pengguna dan memberikan respons yang sesuai dengan input pertanyaan (Zulhilmi and Perdana, 2024). *Chatbot* memiliki keunggulan dalam kemudahan implementasi, fleksibilitas dalam penggunaan, serta kemampuan yang semakin meningkat dalam meniru percakapan manusia (Caldarini, Jaf and McGarry, 2022). Namun, di sisi lain, *chatbot* masih memiliki kelemahan, yaitu kurang mampu memahami konteks percakapan, untuk mengatasi keterbatasan ini, model yang lebih kecil, fleksibel, dan tidak terlalu bergantung pada domain tertentu akan sangat bermanfaat (Caldarini, Jaf and McGarry, 2022).



2.3 Android

Android adalah sistem operasi yang dirancang untuk perangkat *mobile* dengan basis linux, yang terdiri dari sistem operasi, *middleware*, serta aplikasi (Winarni et al., 2024). Android menawarkan platform terbuka bagi para pengembang yang ingin membuat aplikasi mereka sendiri agar dapat digunakan pada berbagai jenis perangkat *mobile*, karena bersifat *open source* (Arsi et al., 2023). Sistem operasi ini mendukung berbagai jenis perangkat, termasuk *smartphone*, *tablet*, *smartwatch*, *smart TV*, dan perangkat *Internet of Things (IoT)*.

2.4 Large Language Model (LLM)

Large Language Models (LLM) merupakan model bahasa yang dilatih menggunakan sejumlah besar data untuk memahami dan memproses bahasa alami secara mendalam, dengan cakupan pemahaman bahasa dan pengetahuan yang luas (Herwanza et al., 2024). LLM bekerja berdasarkan arsitektur *Transformer*, yang diperkenalkan oleh Vaswani et al. dalam makalah "*Attention Is All You Need*" pada tahun 2017. Arsitektur ini menggunakan mekanisme *self-attention* untuk memproses hubungan antar kata dalam sebuah kalimat tanpa memerlukan jaringan berulang. Hal ini menjadikan LLM sebagai alat yang sangat kuat untuk berbagai aplikasi dalam *Natural Language Processing (NLP)*, seperti ringkasan teks, pembuatan kode, dan *chatbot*. Berikut adalah beberapa model *Large Language Models (LLM)*:

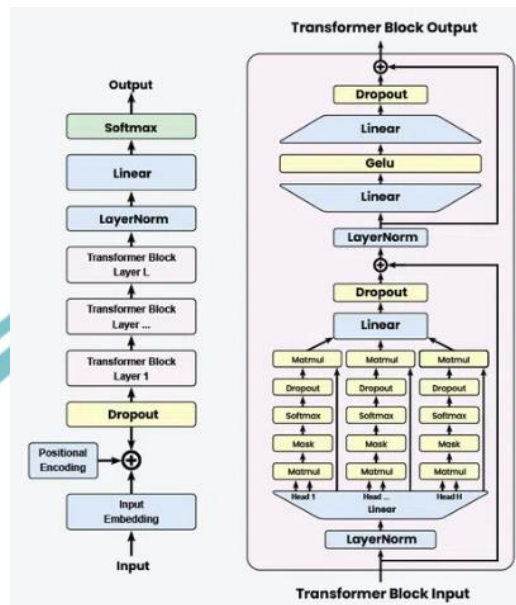
2.4.1 GPT

GPT adalah serangkaian model bahasa besar yang dikembangkan oleh OpenAI. Model ini dirancang untuk memahami dan menghasilkan teks yang menyerupai bahasa manusia. Versi terbaru, GPT-4, dirilis pada Maret 2023 dan memiliki kemampuan *multimodal*, memungkinkan pemrosesan input berupa teks dan gambar (OpenAI, 2023). Namun, detail spesifik mengenai arsitektur dan jumlah parameter GPT-4 tidak dipublikasikan oleh OpenAI. GPT memiliki kelebihan dalam memahami dan menghasilkan teks yang menyerupai tulisan manusia, namun memiliki kekurangan dalam memahami konteks yang lebih dalam dan cenderung

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

menghasilkan data yang tidak akurat (Brown et al., 2020). Arsitektur dari model GPT dapat dilihat pada gambar 2.1.



Gambar 2.1 Arsitektur GPT (Sumber: GeeksforGeeks 2024)

2.4.2 LLaMA

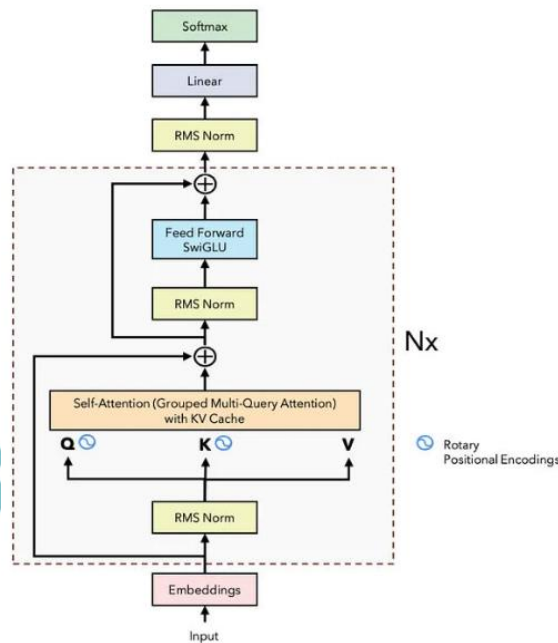
LLaMA adalah keluarga model bahasa besar yang dikembangkan oleh Meta AI. Versi pertama, *LLaMA 1*, diperkenalkan pada Februari 2023 dengan ukuran parameter mulai dari 7 miliar hingga 65 miliar (Touvron et al, 2023). *LLaMA 2*, dirilis pada Juli 2023, menawarkan peningkatan kinerja dengan ukuran model hingga 70 miliar parameter (Touvron et al, 2023). *LLaMA 3*, yang dirilis pada April 2024, memperluas ukuran model hingga 405 miliar parameter dan dilatih menggunakan 15 triliun token (Touvron et al, 2023). Model-model *LLaMA* dirancang untuk berbagai aplikasi pemrosesan bahasa alami dan tersedia untuk penggunaan komersial dengan lisensi tertentu. *LLaMA* memiliki kelebihan dalam tingkat adaptabilitas dan efisiensi yang tinggi dalam memproses berbagai kumpulan data linguistik. namun memiliki kekurangan dalam kemungkinan inkonsistensi dalam kinerja di berbagai bahasa (Hou and Lian, 2024). Selain itu, model *LLaMA* ini bersifat *open source* dan juga memiliki model yang ringan untuk dijalankan di berbagai perangkat seperti perangkat *mobile* (Meta, 2022). Arsitektur dari model *LLaMA* dapat dilihat pada gambar 2.2.

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



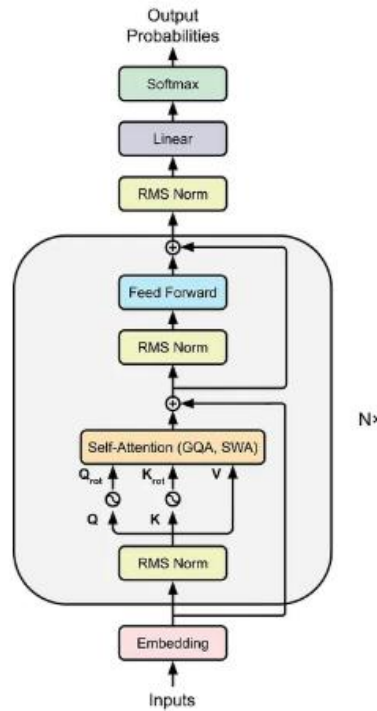
Gambar 2.2 Arsitektur *LLaMA* (Sumber: Yaadav, 2024)

2.4.3 *Mistral*

Mistral adalah serangkaian model bahasa besar yang dikembangkan oleh *Mistral AI*, sebuah perusahaan rintisan yang berbasis di Paris. Model pertama mereka, *Mistral 7B*, dirilis pada September 2023 dengan 7,3 miliar parameter dan dirancang untuk efisiensi tinggi. *Mistral 7B* menggunakan arsitektur *Transformer* dengan mekanisme perhatian yang dioptimalkan untuk kinerja dan skalabilitas. Model ini dirilis di bawah lisensi Apache 2.0, memungkinkan penggunaan dan modifikasi secara bebas. Arsitektur dari model *Mistral 7B* dapat dilihat pada gambar 2.3.

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



Gambar 2.3 Arsitektur *Mistral* 7B (Sumber: Smith, 2024)

2.5 Optical Character Recognition (OCR)

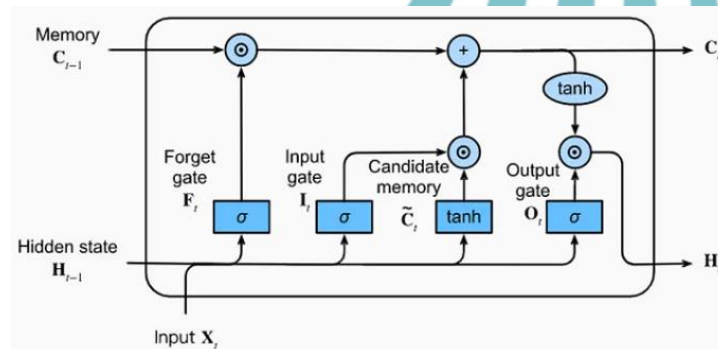
Optical Character Recognition (OCR) adalah teknologi yang berfungsi untuk mendeteksi teks dalam gambar dan mengkonversinya menjadi data yang dapat diolah oleh komputer (Salsabila et al., 2024). *Optical Character Recognition* (OCR) menggunakan teknik pemrosesan citra dan kecerdasan buatan untuk mengenali karakter dari berbagai format, seperti dokumen cetak, gambar, maupun tulisan tangan. Teknologi OCR banyak dimanfaatkan dalam berbagai aplikasi, seperti digitalisasi dokumen, pemindaian tanda terima, otomatisasi entri data, serta penerjemahan teks dari gambar.

Terdapat beberapa *library* yang dapat digunakan untuk mengimplementasikan teknologi OCR, di antaranya adalah Tesseract OCR, Keras OCR, dan Easy OCR (Nazeem et al, 2024).

2.6 Long Short Term Memory (LSTM)

Long Short Term Memory (LSTM) adalah sebuah algoritma yang dikenal memiliki keunggulan dalam membangun model prediktif secara akurat. LSTM merupakan

salah satu varian dari *Recurrent Neural Network* (RNN), yaitu metode yang dirancang khusus untuk menangani dan memproses data berurutan (*sequence data*) (Rizkillloh and Widiyanesti, 2022). *Recurrent Neural Network* (RNN) sering mengalami permasalahan seperti *vanishing gradient* dan *exploding gradient*. Untuk mengatasi hal tersebut, dikembangkanlah *Long Short Term Memory* (LSTM) yang dirancang secara khusus guna menangani kendala hilangnya gradien saat pelatihan pada jaringan RNN. Arsitektur LSTM terdiri dari satu unit yang disebut unit memori (*LSTM unit*). Unit ini terdiri dari empat *feedforward neural network*, masing-masing memiliki lapisan input dan output. Semua neuron input terhubung dengan neuron output, membentuk empat lapisan yang sepenuhnya terhubung., seperti yang ditunjukkan pada Gambar 2.4.



Gambar 2.4 Arsitektur model LSTM (Sumber: Calzone, 2022)

Tiga dari empat jaringan tersebut bertugas untuk memilih informasi *Forget gate* digunakan untuk menghapus informasi, *Input gate* digunakan untuk menyimpan informasi baru, *Output gate* digunakan untuk menggunakan informasi dalam memori. *Candidate memory* memiliki fungsi yang sama dengan *Input Gate* yaitu untuk menyimpan informasi baru. Pada setiap unit LSTM menerima tiga vektor masukan yakni *Cell state* (memori jangka panjang) dari waktu sebelumnya, *Hidden state* (memori jangka pendek) dari waktu sebelumnya, dan input eksternal dari waktu saat ini.

2.7 Senterne Transformer

SBERT atau *Sentence-Transformers* merupakan pengembangan dari BERT yang dirancang untuk menghasilkan representasi kontekstual pada tingkat kalimat,

paragraf, maupun gambar (Samosir et al, 2022). Kemampuan ini memungkinkan SBERT mencapai performa tinggi dalam berbagai tugas, seperti pengelompokan teks (*clustering*), *cross-encoder*, pencarian kalimat paralel, pencarian dan identifikasi parafrase, pencarian semantik, *retrieve & rerank*, pencarian gambar, serta ringkasan teks.

2.8 AI Project Lifecycle

AI Project Lifecycle merupakan sebuah rangkaian proses yang terstruktur dan menyeluruh yang digunakan untuk membangun dan mengelola proyek kecerdasan buatan dari awal hingga akhir (Azimah et al., 2022). Setiap tahap memiliki peran penting dalam menangani tantangan tertentu dan memastikan model AI memenuhi tujuan yang diinginkan. Setiap tahapannya dapat dilihat pada gambar 2.5.



Gambar 2.5 Tahapan *AI Project Lifecycle* (Sumber: Saltz, 2023)

2.8.1 Problem Definition

Tahap ini merupakan tahapan awal dimana pada tahap ini mendefinisikan sebuah permasalahan yang ingin diselesaikan dengan menggunakan pendekatan AI.

2.8.2 Data Acquisition and Preparation

Tahap ini merupakan tahap pengumpulan data dari berbagai sumber yang nantinya data tersebut akan digunakan untuk model AI yang akan dikembangkan.

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



2.8.3 Model Development and Training

Tahap ini merupakan tahap membangun model AI dan melatihnya dengan data yang telah disiapkan. Proses pada tahap ini dilakukan secara iteratif hingga model dapat menghasilkan tingkat akurasi yang baik. Pada tahap ini juga meliputi pemilihan model, penyesuaian parameter dan pelatihan model.

2.8.4 Model Evaluation and Refinement

Pada tahap ini, model AI yang telah dibangun dan dilatih akan diuji dengan menggunakan data yang belum pernah dilihat atau test data untuk menilai kelayakan dari model untuk diimplementasikan.

2.8.5 Deployment

Pada tahap ini, model yang telah diuji dan layak untuk digunakan akan diterapkan pada lingkungan produksi, proses ini meliputi integrasi dengan aplikasi, atau API.

2.8.6 ML Ops

Pada tahap ini, model yang telah diterapkan pada lingkungan produksi akan dipantau untuk memastikan model yang diterapkan dapat berjalan dengan konsisten.

2.9 Unified Modeling Language (UML)

Unified Modeling Language (UML) adalah bahasa pemodelan visual yang berfungsi untuk mendeskripsikan, merancang, dan menyampaikan rancangan suatu sistem secara terstruktur dan sistematis. UML telah menjadi standar internasional yang dikembangkan oleh *Object Management Group (OMG)* dan banyak digunakan dalam industri pengembangan perangkat lunak untuk membantu dalam proses perancangan sistem (Pranoto et al, 2024). Penggunaan UML memungkinkan para pengembang perangkat lunak untuk memvisualisasikan sistem dengan cara yang terstruktur dan mudah dipahami. Pendekatan ini dapat mempercepat proses pengembangan, meminimalkan potensi kesalahan, serta berkontribusi dalam meningkatkan mutu sistem yang dibangun. Pada penelitian ini digunakan beberapa diagram UML untuk membantu dalam perancangan sistem, yakni *use case diagram*, *activity diagram*, dan *sequence diagram*.

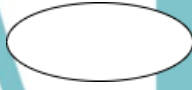
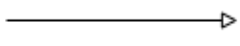
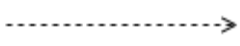
Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

2.9.1 Use Case Diagram

Use case diagram adalah salah satu alat bantu dalam pengembangan sistem yang dibuat untuk menggambarkan hubungan antara *actor* yang berinteraksi dengan sistem. Diagram ini umumnya dimanfaatkan untuk mengilustrasikan fungsi-fungsi sistem secara *high-level* serta ruang lingkup sistem secara menyeluruh, sekaligus mencerminkan berbagai harapan terhadap sistem dari perspektif pihak luar (Wayahdi and Ruziq, 2023). Dalam *use case diagram*, terdapat beberapa simbol standar yang digunakan untuk merepresentasikan elemen-elemennya. Simbol-simbol yang umum digunakan untuk pembuatan *use case diagram* dapat dilihat pada tabel 2.1.

Tabel 2.1 Simbol *Use Case Diagram*

Simbol	Nama	Keterangan
 Actor	Aktor	Entitas yang berinteraksi dengan sistem.
	<i>Use Case</i>	Aktivitas yang dapat dilakukan oleh aktor pada sistem.
	<i>Association</i>	Hubungan antara aktor dengan use case.
	<i>Include</i>	Suatu use case termasuk bagian dari use case lain.
	<i>Extend</i>	Suatu use case dapat diperluas dengan use case lain.
	<i>System</i>	Sistem yang dikembangkan.
	<i>Generalization</i>	Satu aktor atau use case merupakan generalisasi dari yang lain.
	<i>Dependency</i>	Ketergantungan antara elemen-elemen diagram.

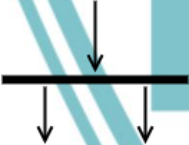
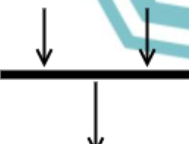
Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

2.9.2 Activity Diagram

Activity diagram merupakan diagram yang menggambarkan alur aktivitas atau proses yang dapat digunakan untuk mengendalikan urutan tindakan yang dijalankan, maupun untuk merepresentasikan alur pertukaran data. Biasanya, *activity diagram* digunakan baik untuk perancangan sistem pada tingkat tinggi seperti pemodelan proses bisnis maupun pada tingkat rendah, seperti mendeskripsikan algoritma yang akan diimplementasikan (Wayahdi and Ruziq, 2023). Simbol-simbol yang umum digunakan pembuatan *activity diagram* dapat dilihat pada tabel 2.2.

Tabel 2.2 Simbol *Activity Diagram*

Simbol	Nama	Keterangan
	Status Awal (<i>Initial</i>)	Menunjukkan awal dari proses atau aktivitas.
	Aktivitas	Tindakan atau pekerjaan spesifik yang dilakukan oleh sistem atau aktor.
	<i>Decision</i>	Titik percabangan alur proses berdasarkan kondisi atau keputusan tertentu.
	<i>Merge</i>	Menggabungkan dua atau lebih alur alternatif menjadi satu alur tunggal.
	<i>Fork</i>	Titik di mana sebuah aktivitas bercabang menjadi dua atau lebih aktivitas paralel.
	<i>Join</i>	Penggabungan dari aktivitas paralel menjadi satu alur lanjutan.
	Status Akhir (<i>Final</i>)	Menunjukkan akhir dari proses atau aktivitas.

Hak Cipta :

- Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
- Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

	Swimlane	Mengelompokkan aktivitas berdasarkan peran, unit organisasi, atau komponen sistem
---	----------	---

2.10 Explainable AI (XAI)

Explainable AI atau XAI merupakan cabang baru dari *Artificial Intelligence* (AI) yang muncul sebagai respons terhadap kebutuhan akan transparansi dan pemahaman dalam sistem *machine learning*, dengan tujuan untuk memberikan penjelasan yang jelas dan dapat dimengerti terkait cara kerja model AI (Nuraeda et al., 2024). Kehadiran XAI memungkinkan perubahan pendekatan dari model *black box* menjadi model *glass box* dengan memberikan penjelasan yang mudah dipahami oleh manusia. Prinsip-prinsip yang diterapkan dalam XAI bertujuan untuk merancang strategi yang mampu menghasilkan model yang lebih transparan tanpa mengorbankan performa atau akurasi tinggi dari sistem yang dikembangkan (Putri et al, 2024). Berikut adalah teknik yang dapat digunakan untuk menjelaskan hasil prediksi dari model *machine learning*:

2.10.1 Local Interpretable Model-agnostic Explanations (LIME)

Local Interpretable Model-agnostic Explanations atau LIME merupakan metode interpretasi yang bertujuan untuk memberikan penjelasan secara lokal terhadap prediksi dari model *black box* dengan pendekatan yang tidak tergantung pada jenis model yang digunakan (*model-agnostic*) (Nuraeda et al, 2024). Metode ini bekerja secara *post-hoc*, artinya penjelasan diberikan setelah model selesai dibangun atau dilatih. Formulasi LIME dapat dilihat pada persamaan (1).

$$\varepsilon(x) = \operatorname{argmin}_{g \in G} L(f, g, \pi_x) + \Omega(g) \quad (1)$$

x merupakan *instance* data tertentu yang ingin diketahui penjelasan dari hasil prediksinya. Formulasi ini dibuat untuk menemukan model lokal (g) yang mampu merepresentasikan model kompleks (f) secara akurat pada *instance* x , dengan tujuan meminimalkan fungsi kerugian (L) dan fungsi kompleksitas (Ω). Dalam hal ini, model kompleks (f) merujuk pada model *machine learning* yang digunakan untuk

menghasilkan prediksi terhadap data. Dengan persamaan ini dapat membantu untuk mengevaluasi keputusan yang diambil oleh model (Nuraeda et al, 2024).

2.10.2 SHapley Addictive exPlanations (SHAP)

Shapley Additive Explanations atau SHAP adalah sebuah metode yang digunakan untuk memberikan penjelasan atas prediksi individual dengan menerapkan konsep teori permainan, sehingga memungkinkan model *machine learning* untuk mengoptimalkan peran setiap fitur (Mumpuni, 2024). SHAP dapat mengevaluasi *output* model dengan menilai seberapa besar pengaruh masing-masing fitur, serta menghitung nilai yang merepresentasikan kontribusi fitur tersebut terhadap hasil prediksi model (Mumpuni, 2024). Penjelasan nilai SHAP dijelaskan pada persamaan (2).

$$g(z) = \phi_0 + \sum_{j=1}^M \phi_j z_j \quad (2)$$

$g(z)$ merupakan nilai prediksi yang dihasilkan pada model, ϕ_0 merupakan nilai konstanta, dan M merupakan jumlah fitur atau elemen dalam vektor j

2.11 Hallucination AI

Hallucination AI merupakan fenomena halusinasi dalam sistem kecerdasan buatan (AI), khususnya pada model pembelajaran mesin seperti *Large Language Models* (LLM), terjadi ketika AI menghasilkan informasi yang keliru, tidak akurat, atau sepenuhnya fiktif. Meskipun jawaban yang diberikan tampak masuk akal dan meyakinkan, sebenarnya tidak didasarkan pada data atau fakta yang valid (Tarumingkeng, 2024). Jenis halusinasi AI adalah sebagai berikut:

- Intrinsic Hallucination* adalah kesalahan yang terjadi saat *output* dari model secara langsung bertolak belakang dengan informasi yang ada pada *input*.
- Extrinsic Hallucination* adalah kesalahan yang terjadi saat *output* yang tidak dapat diverifikasi atau tidak ditemukan pada *input*.

Untuk mengevaluasi atau mendeteksi halusinasi pada model dapat dilakukan beberapa cara diantaranya adalah:

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

- a. Menggunakan evaluator manusia untuk memeriksa apakah output halusinasi atau tidak.
- b. Menggunakan sistem pengecekan otomatis untuk membandingkan output dengan referensi atau basis data menggunakan algoritma tertentu. Adapun beberapa metodenya yaitu *BLEU*, *ROUGE*, *METEOR*, *FactCC*, *GPTScore*, *BERTScore*.
- c. Menggunakan alat *fact checking* seperti *Google Fact Check Explorer*
- d. Benchmark dataset dengan menggunakan dataset seperti *TruthfulQA*, *FEVER*, *SciFact*.

Strategi atau teknik mitigasi yang dapat diterapkan untuk mengatasi halusinasi yang dapat terjadi pada model AI terutama pada penggunaan *Large Language Model* (LLM) adalah sebagai berikut.

- a. *Retrieval-Augmented Generation* (RAG)

Retrieval-Augmented Generation (RAG) adalah sebuah arsitektur yang mengintegrasikan proses pencarian informasi dengan pembuatan teks (Albert and Voutama, 2025). Konsep ini diperkenalkan sebagai cara untuk meningkatkan kualitas respons yang dihasilkan oleh model bahasa alami dengan memanfaatkan data dari sumber eksternal (Albert and Voutama, 2025). Melalui pendekatan ini, model bahasa dapat terlebih dahulu mencari informasi yang sesuai dari dokumen sebelum menyusun jawaban, sehingga hasil yang diberikan menjadi lebih akurat dan relevan.

- b. *Fine-Tuning* dengan data terverifikasi.

Fine-tuning merupakan teknik yang digunakan untuk menyesuaikan parameter dari model yang sudah dilatih sebelumnya agar lebih sesuai dengan dataset target yang memiliki karakteristik khusus atau spesifik (Aryanti et al, 2024).

- c. *Reinforcement Learning from Human Feedback* (RLHF).

Reinforcement Learning from Human Feedback merupakan teknik pelatihan model dimana model diberikan *feedback* dari manusia berdasarkan kualitas *output* dari model.



2.12 FastAPI

FastAPI merupakan salah satu web *framework* berbasis Python yang dirancang untuk membangun *Application Programming Interface* (API). *Framework* ini menerapkan konsep *Asynchronous Server Gateway Interface* (ASGI), yang memungkinkan setiap *endpoint* API pada FastAPI dapat diakses dan dijalankan secara asinkron. FastAPI sangat cocok digunakan dalam proyek *machine learning*, terutama untuk membangun platform *Machine Learning as a Service* (MLaaS) (Suryotomo et al, 2024).

2.13 Black Box Testing

Black Box Testing merupakan metode pengujian kualitas perangkat lunak yang berfokus pada pemeriksaan fungsionalitas sistem tanpa mengetahui struktur internalnya (Wijaya and Astuti, 2021). Tujuan dari pengujian ini adalah untuk mendeteksi fungsi yang tidak berjalan dengan semestinya, kesalahan pada antarmuka pengguna, kesalahan dalam struktur data, masalah performa, serta kesalahan saat proses inisialisasi dan terminasi sistem.

2.14 System Usability Scale (SUS)

System Usability Scale atau SUS merupakan metode evaluasi yang digunakan untuk menilai tingkat *usability* atau kemudahan penggunaan suatu produk atau sistem (Arjiansa and Sutabri, 2023). Pendekatan ini menggunakan kuesioner berisi 10 pertanyaan dengan skala penilaian 5 poin, yang dirancang untuk mengukur aspek seperti kepuasan pengguna, tingkat kompleksitas, serta efisiensi penggunaan. Hasil dari penilaian SUS dapat dimanfaatkan untuk memahami sejauh mana sistem mudah digunakan dan memberikan masukan perbaikan yang dibutuhkan guna meningkatkan pengalaman pengguna. Rumus perhitungan skor pada soal bernomor ganjil dapat dilihat pada persamaan (3) dan rumus perhitungan skor pada soal bernomor genap dapat dilihat pada persamaan (4).

$$\text{Nilai Score Soal}_{\text{Ganjil}} = \text{Score Soal}_{\text{Ganjil}} - 1 \quad (3)$$

$$\text{Nilai Score Soal}_{\text{Genap}} = 5 - \text{Score Soal}_{\text{Genap}} \quad (4)$$

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Rumus untuk menghitung score SUS dari soal dapat dilihat pada persamaan (5).

$$SUS = (Nilai\ Score\ Soal_{Ganjil} + Nilai\ Score\ Soal_{Genap}) \times 2.5 \quad (5)$$

Skor akhir SUS yang didapat kemudian dapat diinterpretasikan untuk memahami tingkat kebergunaan sistem secara kualitatif. Tabel 2.3 menyajikan standar interpretasi skor SUS yang umum digunakan, termasuk *grade* dan peringkat adjektifnya.

Tabel 2.3 Standar Interpretasi Skor dan *Grade* SUS

<i>SUS Score</i>	<i>Grade</i>	<i>Adjective Rating</i>
> 80.3	A	<i>Excellent</i>
68 - 80.3	B	<i>Good</i>
68	C	<i>Okay</i>
51 - 68	D	<i>Poor</i>
< 51	F	<i>Awful</i>

2.15 Net Promoter Score (NPS)

Net Promoter Score atau NPS merupakan metode pengukuran yang dikembangkan oleh Fred Reichheld pada tahun 2003 melalui *Harvard Business Review* untuk menilai sejauh mana pelanggan bersedia merekomendasikan suatu produk atau layanan (Fadlil et al, 2022). NPS digunakan untuk mengukur tingkat kepuasan dan loyalitas pelanggan dengan membagi responden menjadi tiga kategori:

- Promoter* (nilai 9–10): Pelanggan yang sangat puas dan bersedia merekomendasikan produk kepada orang lain.
- Passive* (nilai 7–8): Pelanggan yang puas namun tidak cukup antusias untuk merekomendasikan, serta berpotensi beralih ke produk lain.
- Detractor* (nilai 0–6): Pelanggan yang kecewa dan tidak akan merekomendasikan produk, bahkan bisa memberikan citra negatif terhadap perusahaan.

Rumus NPS dapat dilihat pada persamaan (6):

$$NPS = \%Promoter - \%Detractor \quad (6)$$

Rumus untuk menghitung nilai $\%Promoter$ dapat dilihat pada persamaan (7) dan $\%Detractor$ (8).

$$\%Promoter = \frac{\text{Responden yang memberi skor 9-10 pada produk}}{\text{Jumlah Responden yang memilih produk}} \times 100 \% \quad (7)$$

$$\%Detractor = \frac{\text{Responden yang memberi skor 0-6 pada produk}}{\text{Jumlah Responden yang memilih produk}} \times 100 \% \quad (8)$$

Jika jumlah *promoter* meningkat dan jumlah *detractor* menurun, maka tingkat loyalitas konsumen terhadap suatu produk akan semakin tinggi. Tabel parameter *net promoter score* dapat dilihat pada tabel 2.4.

Tabel 2.4 Parameter *Net Promoter Score*

<i>Net Promoter Score</i>	<i>Ukuran Net Promoter Score</i>
NPS < 0	Tidak Loyal
NPS = 0 - 50	Loyal
NPS > 50	Sangat Loyal

2.16 Penelitian Terdahulu

Pada penelitian terdahulu ini, diambil beberapa penelitian sejenis yang digunakan untuk menganalisis dan memperluas pembahasan dari beberapa penelitian yang telah dilakukan sebelumnya dan mengevaluasi serta meningkatkan kualitas dari penelitian yang sedang dilakukan. Beberapa referensi penelitian sejenis yang dijadikan sebagai acuan ditampilkan pada tabel 2.5.

Tabel 2.5 Penelitian Terdahulu

Judul	Permasalahan	Metode/Algoritma	Hasil Penelitian
Implementasi <i>Artificial</i>	Banyaknya wisatawan,	Menggunakan metode	Penelitian ini berhasil

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

<i>Intelligence</i> Dalam Aplikasi <i>Chatbot</i> Untuk Rekomendasi Wisata Pantai Di Batam Dengan Metode <i>Feedforward Neural Network</i> (Ramadhan, Masril and Caniago, 2024).	terutama yang baru pertama kali mengunjungi Batam, sering kali merasa kebingungan dalam memilih destinasi wisata. Selain itu, keterbatasan informasi yang lengkap dan akurat mengenai tempat-tempat wisata kerap menjadi kendala dalam pengambilan keputusan. Beragamnya pilihan atraksi dan aktivitas di Batam juga dapat membingungkan wisatawan dalam menentukan tujuan yang paling sesuai.	<i>Feedforward Neural Network</i> yang berfungsi untuk memetakan data inputan ke data keluaran.	mengembangkan model dengan performa sangat baik, yang ditunjukkan melalui nilai <i>macro precision</i> sebesar 79.71%, <i>macro recall</i> sebesar 80%, dan <i>macro F1 score</i> sebesar 79.76%. Kemudian dengan penggunaan <i>hyperparameter</i> berupa <i>batch size</i> sebesar 10, <i>learning rate</i> sebesar 0.002, sebanyak 600 <i>epoch</i>, dan optimizer ADAM, model berhasil mencapai tingkat akurasi yang tinggi sebesar 97.1% serta nilai <i>loss</i> yang rendah, yakni 1.54%. Namun terdapat, berapa skenario di mana model belum sepenuhnya mampu mengidentifikasi <i>intent</i> pengguna, sehingga diperlukan optimasi lebih lanjut atau penambahan data pelatihan.
Implementasi <i>Chatbot</i> Untuk Rekomendasi <i>Coffee Shop</i> Di Kota	Pertumbuhan <i>coffee shop</i> di Kota Semarang yang meningkat pesat serta beragamnya	Menggunakan <i>Framework</i> Rasa karena memiliki fleksibilitas dalam mengelola dialog.	Penelitian ini telah berhasil membuat <i>chatbot</i> yang dapat membantu pengunjung <i>coffee</i>



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Semarang (Cahyawati and Hadiono, 2024).	pilihan sajian kopi yang tersedia. Hal ini menjadikan kebutuhan akan rekomendasi yang sesuai dengan preferensi pelanggan semakin menjadi prioritas.		<i>shop</i> di Kota Semarang menggunakan <i>framework</i> Rasa yang memberikan keuntungan dalam pembuatan sistem yang responsif dan fleksibel, sementara kolaborasi dengan Telegram memudahkan akses pengguna. Namun, disamping itu, hasil rekomendasi pada <i>chatbot</i> terbatas pada dataset yang digunakan, sehingga dataset perlu diperluas untuk mendapatkan data rekomendasi lain.
Krishna – <i>The Career Guidance ChatBot</i> (Raut, Sharma and Kharbade, 2021).	Siswa seringkali merasa telah memilih karier yang salah dan kemudian menyesalinya, namun mereka tidak banyak berbuat tentang hal itu. Dari permasalahan tersebut, penelitian ini bertujuan untuk membuat <i>chatbot</i> yang membantu siswa dalam menentukan pendidikan dan karier mereka.	Menggunakan Algoritma <i>Natural Language Processing</i> dan beberapa <i>library</i> seperti <i>NLP</i> , <i>Chatterbot</i> , <i>Numpy</i> , <i>pytz</i> , dan <i>nlk.tokenize</i> .	Penelitian ini telah berhasil mengembangkan <i>chatbot</i> yang membantu siswa dalam menentukan pilihan pendidikan dan karier. Namun, terdapat beberapa fitur yang masih dapat dikembangkan atau ditambahkan pada <i>chatbot</i> , seperti kemampuan untuk menganalisis berbagai faktor yang mempengaruhi pemilihan jalur atau bidang karier yang dipilih oleh siswa.



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

ASPIREAI: <i>AI Chatbot for Career Guidance</i> (Dhamdhere et al., 2024).	Adanya tantangan yang semakin kompleks yang dihadapi oleh setiap individu dalam merencanakan dan menentukan jalur karier mereka di tengah lanskap profesional yang berkembang pesat. Oleh karena itu, untuk mengatasi permasalahan tersebut, dikembangkan sebuah <i>chatbot</i> yang diberi nama Aspire AI.	Menggunakan model <i>Large Language Model</i> (LLM) untuk membangun Aspire AI.	Penelitian ini telah berhasil mengembangkan <i>chatbot</i> yang menunjukkan efektivitas luar biasa dalam memberikan wawasan karier yang dipersonalisasi, saran pencarian kerja, dan rekomendasi pendidikan. Untuk pengembangan selanjutnya, penelitian ini dapat ditingkatkan kembali dengan menyempurnakan pemahaman kontekstualnya
<i>Career Connect – A Personalized Job Recommendation AI Chatbot</i> (Prakash et al., 2024).	Permasalahan dalam penelitian ini adalah metode bimbingan karier tradisional sering kali tidak mampu memberikan bantuan secara efektif serta mendukung para profesional dalam mempercepat perkembangan karier. Akibatnya, jumlah pelamar menjadi terbatas, yang berdampak pada ketidakefisienan dalam proses perekrutan.	Menggunakan beberapa model seperti <i>Llama 3</i> , <i>Mixtral</i> , dan <i>Gemma</i> sehingga memungkinkan pengguna untuk memilih model yang ingin digunakan. Kemudian untuk mengekstrak teks dari dokumen CV digunakan <i>library PDFPlumber</i> .	Penelitian ini telah berhasil mengembangkan <i>chatbot</i> yang mampu memberikan rekomendasi pekerjaan berdasarkan dokumen CV yang diunggah oleh pengguna. Namun, masih terdapat peluang pengembangan lebih lanjut, salah satunya dengan menerapkan dukungan multibahasa. Fitur ini akan memungkinkan



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

			pengguna yang tidak berbicara bahasa Inggris untuk berinteraksi dengan <i>chatbot</i> dalam bahasa pilihan mereka, sehingga dapat meningkatkan aksesibilitas dan pengalaman pengguna secara lebih luas.
--	--	--	---

Berdasarkan lima penelitian sebelumnya yang terdapat pada Tabel 2.5, terdapat kesamaan dalam pengembangan aplikasi, yaitu *chatbot* yang berfungsi untuk memberikan sebuah rekomendasi. Namun, masing-masing penelitian tersebut menggunakan metode yang berbeda dalam proses pengembangan *chatbot* untuk menghasilkan rekomendasi terhadap topik permasalahan yang diangkat. Pada penelitian ini dirancang sebuah aplikasi *chatbot* yang dapat memberikan rekomendasi mengenai pekerjaan berdasarkan minat dan kemampuan pengguna. Berbeda dengan lima penelitian sebelumnya, Penelitian ini memanfaatkan *Optical Character Recognition* (OCR) yang digunakan untuk mengekstraksi data teks yang ada pada input pengguna yang berupa dokumen yang berformat .pdf atau gambar, serta menggunakan model *LLaMA* untuk menghasilkan rekomendasi pekerjaan berdasarkan hasil ekstraksi data teks yang didapatkan dari input pengguna.



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

BAB III METODE PENELITIAN

3.1 Rancangan Penelitian

Penelitian ini berfokus pada perancangan aplikasi berbasis android yang mengimplementasikan teknologi *chatbot* dengan menggunakan model *LLaMA* serta menerapkan model *Optical Character Recognition* (OCR) untuk memberikan rekomendasi pekerjaan kepada pengguna. Tujuan dari penelitian ini adalah membangun sebuah aplikasi *chatbot* yang dapat membantu pengguna dalam pengambilan keputusan untuk memilih pekerjaan yang tepat sesuai dengan minat dan kompetensi pengguna. Selain itu, penelitian ini juga bertujuan untuk menguji dan menganalisis penerapan teknologi OCR dan model *LLaMA* dalam menghasilkan *output* berupa rekomendasi pekerjaan, sehingga dapat memberikan kontribusi dan masukan yang berguna bagi penelitian selanjutnya

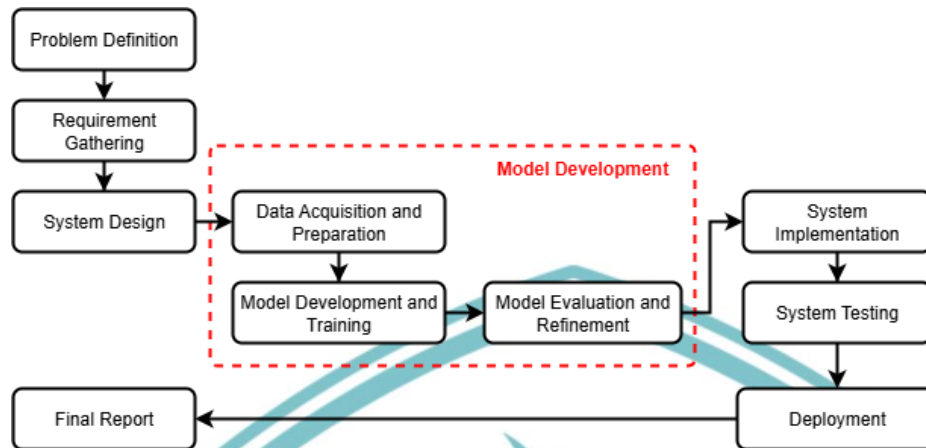
Metode penelitian yang digunakan adalah metode penelitian eksperimental dengan pendekatan kuantitatif, yang berfokus pada kegiatan eksperimen dalam perancangan aplikasi dan pengujian kinerja dari model yang dibuat. Kemudian, metode pengembangan yang diadopsi untuk mengembangkan aplikasi *chatbot* pada penelitian ini adalah *AI Project Lifecycle*, yang terdiri dari *Problem Definition*, *Data acquisition and preparation*, *Model development and training*, serta *Model evaluation and refinement*, dan *Deployment*.

3.2 Tahapan Penelitian

Tahapan dalam penelitian ini mengadopsi pendekatan metode *AI Project Lifecycle*. Siklus hidup pengembangan proyek ini dibagi menjadi beberapa tahapan yang akan dilakukan secara berurutan dan sistematis yang dapat dilihat pada gambar 3.1.

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



Gambar 3.1 Tahapan Penelitian

Berikut adalah penjelasan dari setiap tahapan penelitian pada gambar 3.1.

3.2.1 Pendefinisian Masalah (*Problem Definition*)

Tahap ini dilakukan melalui observasi terhadap penelitian-penelitian terdahulu yang dipublikasikan dalam jurnal ilmiah. Tujuannya adalah untuk menemukan serta mendefinisikan permasalahan yang dapat diselesaikan melalui penerapan teknologi. Hasil dari observasi ini menunjukkan bahwa terdapat beberapa isu utama yang sering dialami oleh individu, khususnya pencari kerja dan mahasiswa, antara lain:

- a. Kurangnya pemahaman terhadap potensi diri
- b. Minimnya informasi terkait pilihan karier
- c. Kebingungan dalam memilih jalur yang sesuai dengan minat dan bakat

Permasalahan-permasalahan tersebut seringkali menyebabkan ketidaksesuaian antara pekerjaan yang dipilih dengan kemampuan atau kompetensi individu. Oleh karena itu, diperlukan sebuah solusi yang dapat membantu mengarahkan pengguna dalam memahami potensi diri serta menentukan pilihan pekerjaan yang tepat.

3.2.2 Pengumpulan Kebutuhan Pengguna (*Requirement Gathering*)

Untuk memperoleh informasi yang lebih mendalam mengenai kebutuhan dan preferensi dari target pengguna yakni individu yang sedang mencari pekerjaan, pada tahap ini, dilakukan proses pengumpulan data melalui dua teknik, yang bertujuan untuk menghasilkan solusi yang sesuai, metode yang digunakan diantaranya:



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

a. Studi Literatur

Studi dilakukan terhadap jurnal-jurnal ilmiah untuk memperoleh wawasan konseptual serta memahami pendekatan-pendekatan yang telah dilakukan dalam penelitian sebelumnya.

b. Survei Kuesioner

Survei dibuat menggunakan Google Form dan ditujukan kepada pencari kerja, mahasiswa tingkat akhir, serta pencari magang, dengan target sebanyak 20 responden. Tujuan dari survei ini adalah untuk menggali kebutuhan dan kendala pengguna yang berkaitan dengan pemilihan pekerjaan. Selain itu, melalui survei kuesioner yang dilakukan, diperoleh pula informasi mengenai kebutuhan fungsional dan non-fungsional dari aplikasi yang akan dikembangkan.

3.2.3 Desain Sistem (*System Design*)

Pada tahap ini dilakukan perancangan desain aplikasi *chatbot* yang mencakup beberapa aspek utama, antara lain perancangan sistem dan alur kerja yang divisualisasikan menggunakan diagram UML, perancangan arsitektur sistem, perancangan basis data (*database*), serta perancangan antarmuka pengguna (*User Interface*) dengan menggunakan tools Figma.

3.2.4 Pengembangan Model (*Model Development*)

Tahapan pengembangan model menggunakan pendekatan *AI Project Lifecycle* diawali dengan tahap *Data Acquisition and Preparation*, *Model Development and Training*, serta *Model Evaluation and Testing*. Berikut adalah penjelasan dari masing-masing tahapan tersebut.

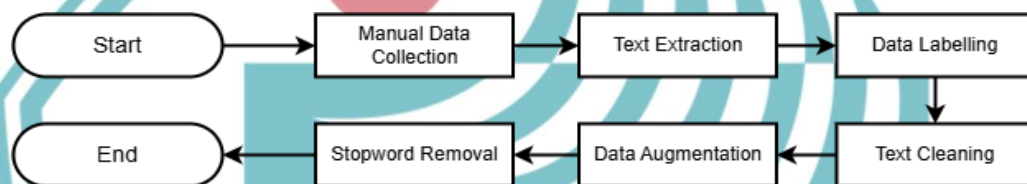
3.2.4.1 *Data Acquisition and Preparation*

Tahap ini merupakan langkah awal dalam pengembangan model yang akan diterapkan pada aplikasi chatbot. Pada tahap ini dilakukan proses pengumpulan dan persiapan data sebagai dasar dalam membangun dataset yang dibutuhkan oleh model kecerdasan buatan (AI) yang digunakan dalam penelitian. Adapun jenis dataset yang digunakan terdiri dari dataset dokumen resume dan dataset kategori

pekerjaan. Proses pengumpulan dan persiapan masing-masing dataset tersebut dijelaskan sebagai berikut.

a. **Pengumpulan dan Persiapan Dataset Dokumen *Resume***

Dataset ini berisi kumpulan dokumen yang diklasifikasikan ke dalam dua label, yaitu CV (*Curriculum Vitae*) dan Non-CV. Dataset ini diperlukan untuk proses klasifikasi dokumen, guna membedakan dokumen yang merupakan CV dengan dokumen lainnya. Namun, karena belum tersedia dataset publik yang secara khusus menyediakan dokumen dengan label CV dan Non-CV, maka pada tahap ini dataset tersebut dibuat secara mandiri dengan mengumpulkan dan melabeli data secara manual. Tahapannya dapat dilihat pada gambar 3.2.



Gambar 3.2 Tahapan Pengumpulan dan Persiapan Dataset Dokumen CV

Berikut adalah penjelasan dari setiap tahapan pada pengumpulan dan persiapan dataset dokumen CV yang ditampilkan pada gambar 3.2.

1. ***Manual Data Collection***

Pada tahap ini, dilakukan pengumpulan data berupa dokumen PDF yang terdiri dari dokumen CV dan Non-CV. Proses pengumpulan dilakukan secara manual dengan mengunduh dokumen dari situs Canva atau sumber dokumen publik lainnya yang tersedia di internet. Dokumen CV dan Non-CV yang diperoleh akan disimpan pada folder yang berbeda yakni folder CV dan Non-CV untuk mempermudah proses pelabelan data.

2. ***Text Extraction***

Pada tahap ini, digunakan *library* PyMuPDF untuk mengekstrak teks dari dokumen PDF yang telah dikumpulkan. Proses ekstraksi ini bertujuan untuk mengubah isi dokumen yang berupa file PDF menjadi data teks mentah yang dapat diolah lebih lanjut.

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

3. *Data Labeling*

Pada tahap ini, data teks yang telah diekstrak akan diberi label secara manual (*manual annotation*). Pelabelan dilakukan dengan mengkategorikan teks ke dalam 2 kelas, yakni CV dan Non-CV. Dokumen yang tersimpan dalam folder CV akan dilabeli sebagai CV, sedangkan dokumen dalam folder Non-CV akan dilabeli sebagai Non-CV.

4. *Text Cleaning*

Pada tahap ini, dilakukan proses pembersihan teks dari karakter atau elemen yang tidak relevan sebelum dilakukan augmentasi data.

5. *Data Augmentation*

Tahap ini dilakukan untuk menambah variasi dataset dengan menerjemahkan teks menggunakan *library* deep-translator, sehingga dataset menjadi lebih banyak dan beragam.

6. *Stopword Removal*

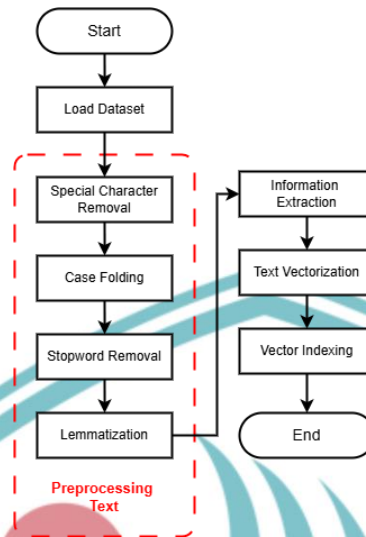
Menghapus *stopword* atau kata-kata umum yang tidak memiliki makna dari teks yang ada pada dataset sehingga saat data digunakan untuk melatih model maka model lebih fokus pada kata-kata yang lebih informatif.

b. **Pengumpulan dan Persiapan Dataset Kategori Pekerjaan**

Dataset ini memuat daftar keterampilan (*skills*) beserta kategori pekerjaan yang relevan. Dataset ini digunakan sebagai dasar untuk membangun *vector database* yang memungkinkan pencarian semantik berdasarkan kemiripan makna. Pada tahap ini, dilakukan proses pembuatan *vector database* menggunakan dataset keterampilan tersebut. Adapun tahapan-tahapan yang dilakukan adalah ditampilkan pada gambar 3.3.

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



Gambar 3.3 Tahapan Pengumpulan dan Persiapan Dataset Kategori Pekerjaan

Berikut adalah penjelasan dari setiap tahapan pada pengumpulan dan persiapan dataset kategori pekerjaan yang ditampilkan pada gambar 3.3.

1. Load Dataset

Dataset keterampilan dan kategori pekerjaan dimuat ke dalam sistem untuk diproses lebih lanjut.

2. Data Preprocessing

Pada proses ini, dilakukan pembersihan dan standarisasi teks guna meningkatkan kualitas data. Proses ini mencakup:

- *Special Character Removal*, yakni menghapus karakter khusus yang tidak relevan.
- *Case Folding*, yakni mengubah seluruh teks menjadi huruf kecil untuk konsistensi.
- *Stopword Removal*, yakni menghilangkan kata-kata umum yang tidak memiliki makna penting dalam analisis.
- *Lemmatization*, yakni mengubah kata ke bentuk dasarnya untuk menyatukan variasi kata yang memiliki makna serupa.

3. Information Extraction

Pada proses ini, dilakukan proses ekstraksi informasi penting berupa keterampilan dari teks dengan memanfaatkan model spaCy yang



dikombinasikan dengan fitur *Entity Ruler* untuk digunakan dalam proses representasi vektor.

4. *Text Vectorization*

Proses ini dilakukan dengan mengubah data teks menjadi representasi vektor menggunakan model *Sentence Transformer* dengan arsitektur distiluse-base-multilingual-cased-v2.

5. *Vector Indexing*

Pada proses ini, vektor yang dihasilkan kemudian diindeks ke dalam *vector database* dengan menggunakan chromaDB agar dapat digunakan dalam pencarian berbasis kemiripan semantik secara efisien.

3.2.4.2 *Model Development and Training*

Pada tahap ini, dilakukan pemilihan dan pengembangan berbagai model yang digunakan dalam sistem. Setiap model memiliki fungsi spesifik dan saling mendukung satu sama lain, di antaranya:

a. **Model Validasi Dokumen**

Model ini berfungsi untuk mendeteksi apakah dokumen yang diunggah merupakan sebuah *Curriculum Vitae* (CV) atau bukan. Model dibangun menggunakan algoritma *Long Short-Term Memory* (LSTM), yaitu jenis jaringan saraf tiruan yang unggul dalam memahami susunan dan konteks kata dalam teks. Untuk memastikan bahwa parameter yang digunakan pada model sudah optimal, dilakukan proses *hyperparameter tuning* pada model LSTM yang digunakan untuk mengklasifikasikan dokumen sebagai CV atau Non-CV. Tujuan dari penyesuaian ini adalah untuk menemukan kombinasi parameter terbaik sehingga model dapat memberikan performa yang optimal terhadap dataset yang digunakan.

b. **Model Optical Character Recognition (OCR)**

Untuk mengekstrak teks dari dokumen berbasis gambar, seperti PDF hasil scan, digunakan *library* EasyOCR. *Library* ini mendukung berbagai bahasa dan memiliki performa baik dalam mendeteksi teks dari dokumen tidak terstruktur.

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Namun, dalam penelitian ini dilakukan penyesuaian terhadap dokumen gambar sebelum diproses oleh EasyOCR. Penyesuaian ini bertujuan untuk meningkatkan efisiensi waktu pemrosesan, sehingga proses ekstraksi teks dapat berjalan lebih cepat dan akurat.

c. **Model Ekstraksi Keyword**

Model ekstraksi *keyword* digunakan untuk mengambil *keyword* penting dari teks yang telah diperoleh melalui proses ekstraksi teks sebelumnya. *Keyword* ini diambil untuk membantu LLM untuk dapat memberikan rekomendasi pekerjaan. Untuk menentukan metode ekstraksi kata kunci yang paling sesuai, dilakukan pengujian terhadap beberapa pendekatan, yaitu:

- KeyBERT, yang menggunakan model *Sentence Transformer* untuk memahami konteks kalimat dan memilih kata kunci yang paling representatif.
- YAKE (*Yet Another Keyword Extractor*), yaitu metode berbasis statistik yang cepat dan tidak membutuhkan pelatihan sebelumnya.
- NER (*Named Entity Recognition*) dari spaCy, yang secara otomatis mengenali entitas penting dalam teks seperti nama, organisasi, dan lokasi.

d. **Large Language Model (LLM)**

Model ini diimplementasikan untuk memberikan rekomendasi pekerjaan berdasarkan *keyword* yang dihasilkan. Model LLM yang digunakan pada penelitian ini adalah model *LLaMA*. Model *LLaMA* ini dipilih karena merupakan LLM *open-source* yang dapat diadaptasi sesuai kebutuhan sistem. Selain itu, untuk mendapatkan rekomendasi pekerjaan dari model LLM ini dilakukan *prompt engineering* terhadap model *LLaMA* guna meningkatkan kualitas respon yang dihasilkan dalam memberikan rekomendasi pekerjaan berdasarkan kata kunci yang telah diekstraksi.

3.2.4.3 **Model Evaluation and Refinement**

Pada tahap ini dilakukan analisis terhadap eksperimen pengembangan model, meliputi evaluasi hasil model validasi dokumen sebelum dan sesudah proses *hyperparameter tuning*. Apabila hasil setelah *tuning* menunjukkan peningkatan

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

performa, maka model validasi dokumen yang digunakan pada aplikasi adalah model yang telah melalui proses *hyperparameter tuning*. Selain itu, untuk memahami cara kerja model LSTM dalam memprediksi dokumen CV atau Non-CV, digunakan metode SHAP (*SHapley Additive exPlanations*) dan LIME (*Local Interpretable Model-Agnostic Explanations*). Kedua metode ini bertujuan menjelaskan hasil prediksi model LSTM, sehingga meningkatkan kepercayaan terhadap model yang digunakan.

Selain itu, pada tahap ini juga dilakukan evaluasi terhadap LLM yang digunakan untuk mendeteksi halusinasi pada model *LLaMA*. Metode evaluasi yang diterapkan adalah *Manual Evaluation* menggunakan dataset *TruthfulQA*. Jika LLM mengalami halusinasi pada beberapa pertanyaan, maka dilakukan penyesuaian variabel *temperature* serta penambahan konteks pada *prompt* untuk mencegah halusinasi pada respons LLM. Konteks tambahan ini diambil dari *vector database* yang berisi data terpercaya yang berisikan kategori pekerjaan dan *skill* untuk meningkatkan respons yang dikeluarkan dari model.

3.2.5 Implementasi Sistem (*System Implementation*)

Setelah model kecerdasan buatan (AI) siap digunakan, tahap ini berfokus pada pengembangan aplikasi *chatbot* berbasis Android. Aplikasi dikembangkan menggunakan bahasa pemrograman Kotlin melalui lingkungan pengembangan Android Studio. Selain itu, dilakukan pula perancangan dan pembuatan basis data menggunakan MySQL untuk menyimpan data yang dibutuhkan oleh sistem.

Pada tahap ini juga dikembangkan API menggunakan FastAPI, yang berfungsi sebagai jembatan untuk mengakses model AI dan basis data. Tahapan ini mencakup proses integrasi antara model AI, API, dan aplikasi *chatbot*, sehingga seluruh komponen dapat bekerja secara terpadu dan dapat diakses oleh pengguna melalui aplikasi Android.

3.2.6 Pengujian Sistem (*System Testing and Evaluation*)

Pada tahap ini dilakukan pengujian sistem untuk menguji fungsionalitas dan kelayakan aplikasi *chatbot* yang telah dikembangkan. Pengujian bertujuan



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

memastikan bahwa semua fitur berfungsi sesuai dengan kebutuhan dan aplikasi dapat digunakan dengan baik oleh pengguna. Pengujian dilakukan dengan menggunakan tiga metode utama, yaitu:

- a. *Black-box testing* untuk memverifikasi fungsionalitas fitur.
- b. *System Usability Scale* (SUS) untuk mengukur tingkat kemudahan penggunaan (*usability*) sebuah sistem. Evaluasi dilakukan melalui kuesioner yang terdiri atas sepuluh pernyataan, di mana responden diminta untuk memberikan penilaian berdasarkan skala *Likert* (1-5).
- c. *Net Promoter Score* (NPS) untuk mengukur tingkat kepuasan dan loyalitas pengguna terhadap aplikasi. Dalam pengujian ini, pengguna diminta menjawab pertanyaan dengan skala 1–10. Berdasarkan nilai yang diberikan, responden akan dikelompokkan ke dalam kategori Promotor (9–10), Pasif (7–8), dan Detraktor (1–6).

Metode SUS dan NPS dilaksanakan dengan menggunakan instrumen berupa formulir kuesioner yang dibuat melalui platform Google Form. Formulir tersebut disebarakan kepada sejumlah responden yang telah menggunakan aplikasi *chatbot* untuk memperoleh data kuantitatif yang dapat dianalisis lebih lanjut.

3.2.7 Deployment

Pada tahap ini, API yang telah dikembangkan akan di-*tunnel* menggunakan Ngrok agar dapat diakses secara publik. Selanjutnya, aplikasi *chatbot* akan di-build menjadi sebuah file dengan ekstensi .apk sehingga dapat diinstal pada perangkat Android oleh pengguna. File .apk tersebut kemudian akan dipublikasikan melalui GitHub dan dapat diunduh secara publik melalui tautan yang disediakan.

3.2.8 Penyusunan Laporan Akhir (*Final Report*)

Pada tahap ini dilakukan penyusunan laporan hasil penelitian, yang mencakup seluruh proses pengembangan aplikasi mulai dari tahap *problem identification* hingga tahap *deployment*. Laporan ini juga berisi analisis mengenai eksperimen yang telah dilakukan selama tahap implementasi.



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

3.3 Objek Penelitian

Objek penelitian ini adalah aplikasi *chatbot* berbasis android, yang dirancang untuk memberikan rekomendasi pekerjaan berdasarkan input pengguna dalam bentuk teks, gambar dengan format .jpg, .jpeg, .png, atau dokumen berformat .pdf. Objek penelitian ini dipilih berdasarkan tujuan dari penelitian ini yakni membangun sebuah aplikasi yang dapat memberikan rekomendasi pekerjaan kepada pengguna berdasarkan minat dan kemampuannya. Aplikasi *chatbot* ini akan dilakukan analisa dan evaluasi untuk mengukur kinerja dan keakuratannya dalam memberikan rekomendasi pekerjaan.

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

BAB IV

HASIL DAN PEMBAHASAN

4.1 Problem Definition

Berdasarkan hasil studi literatur dan observasi terhadap beberapa penelitian sebelumnya, ditemukan sejumlah permasalahan yang dapat dijadikan dasar dalam penelitian ini. Permasalahan tersebut antara lain:

- a. Kurangnya pemahaman terhadap potensi diri, yang menyebabkan individu kesulitan dalam mengidentifikasi kekuatan dan kelemahan pribadinya.
- b. Minimnya informasi terkait pilihan karier, sehingga individu tidak memiliki gambaran yang jelas mengenai berbagai jalur profesi dan persyaratan yang diperlukan.
- c. Kebingungan dalam memilih jalur yang sesuai dengan minat dan bakat, yang kerap mengarah pada keputusan karier yang tidak tepat.

Permasalahan-permasalahan di atas seringkali mengakibatkan ketidaksesuaian antara pekerjaan yang dipilih dengan kemampuan atau kompetensi individu (Luid et al., 2024). Ketidaksesuaian ini dapat berujung pada rendahnya kepuasan kerja, dan kurangnya produktivitas (Hariyanto et al., 2024).

4.2 Requirement Gathering

Untuk memahami kebutuhan pengguna serta mengeksplorasi berbagai pendekatan yang dapat digunakan dalam menyelesaikan permasalahan, dilakukan proses pengumpulan data melalui tiga metode, yaitu studi literatur, survei kuesioner, dan analisis kebutuhan aplikasi berdasarkan data yang telah diperoleh. Berikut adalah proses pengumpulan data melalui tiga metode tersebut.

4.2.1 Studi Literatur

Studi dilakukan terhadap jurnal-jurnal ilmiah untuk memperoleh wawasan konseptual serta memahami pendekatan-pendekatan yang telah dilakukan dalam penelitian sebelumnya. Berdasarkan hasil dari studi literatur didapatkan beberapa solusi yang dapat menyelesaikan masalah rekomendasi pekerjaan.

Tabel 4.1 Referensi Teknologi dari Sumber Literatur

Penulis	Judul Jurnal	Hasil Literatur
Sajay Prakash, Koushik Sundar, Eashaan Manohar, Ram Prakash L (2024)	<i>Career Connect – A Personalized Job Recommendation AI Chatbot.</i>	<i>Chatbot</i> dapat memberikan rekomendasi pekerjaan dengan baik dan menghasilkan rekomendasi yang sangat interaktif kepada pengguna. Selain itu, rekomendasi yang dihasilkan bisa sangat beragam.
Dimas Wahyu Wibowo, Mustika Mentari, Abdallah Darussalam Candra, Amin Anis Kuddah, Rahardhiyan Wahyu Putra (2020)	Sistem Pendukung Keputusan Berbasis Web Rekomendasi Pekerjaan Bagi Lulusan JTI Polinema Dengan Metode SAW.	Sistem pendukung keputusan dapat memberikan rekomendasi pekerjaan kepada pengguna berdasarkan preferensi pengguna. Namun, pendekatan ini kurang fleksibel, karena rekomendasi pekerjaan bergantung pada daftar pekerjaan yang ada pada database sistem.
Crismastiana Koloman, Raihan Maulana, Raisya Dwi Zahra Putri, Wahyu Abadi Harahap (2023)	Sistem Rekomendasi Pekerjaan di bidang IT Menggunakan Algoritma <i>Content-Based Filtering</i> .	Sistem rekomendasi dapat memberikan rekomendasi pekerjaan kepada pengguna berdasarkan riwayat data pengguna.

Berdasarkan hasil studi literatur pada tabel 4.1, solusi untuk mengatasi permasalahan yang didefinisikan pada penelitian ini adalah merancang aplikasi *chatbot* yang dapat memberikan rekomendasi pekerjaan berdasarkan preferensi pengguna.

4.2.2 Survei Kuesioner

Survei kuesioner yang dilakukan dalam penelitian ini memperoleh sebanyak 32 responden, yang mayoritas merupakan mahasiswa. Tujuan dari survei ini adalah untuk memperkuat topik yang diangkat dalam penelitian serta mengidentifikasi kebutuhan fungsional dan non-fungsional dari aplikasi yang akan dikembangkan. Hasil survei ini juga digunakan untuk mendukung dan memperjelas permasalahan yang telah didefinisikan sebelumnya. Berdasarkan hasil survei, diperoleh beberapa temuan penting sebagai berikut:

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

- a. Sebanyak 34,4% responden menyatakan bahwa mereka masih merasa ragu dalam menentukan pilihan pekerjaan yang sesuai dengan diri mereka.
- b. Sebanyak 53,1% responden mengungkapkan bahwa tantangan terbesar dalam memilih pekerjaan adalah perasaan bahwa kemampuan yang mereka miliki belum cukup.
- c. Sebanyak 68,8% responden menyatakan keinginan untuk mendapatkan rekomendasi pekerjaan yang relevan dengan minat dan kompetensi yang mereka miliki.

4.2.3 Analisis Kebutuhan Aplikasi

Analisis kebutuhan dilakukan untuk mengidentifikasi elemen-elemen yang dibutuhkan dalam pengembangan sistem, baik dari sisi pengembangan model kecerdasan buatan maupun dari sisi aplikasi yang akan digunakan oleh pengguna akhir. Berikut adalah kebutuhan pengembangan model dan kebutuhan pengembangan aplikasi pada penelitian ini.

4.2.3.1 Kebutuhan Pengembangan Model

Dalam pengembangan model kecerdasan buatan untuk sistem yang mampu memberikan rekomendasi pekerjaan berdasarkan input pengguna, baik dalam bentuk teks maupun dokumen berformat PDF, JPG, JPEG, dan PNG, terdapat sejumlah komponen utama yang perlu dipersiapkan. Komponen-komponen tersebut meliputi dataset yang digunakan untuk melatih model, model *deep learning* yang sesuai, serta *pre-trained model* yang mendukung proses klasifikasi, ekstraksi informasi, dan pemberian rekomendasi secara otomatis.

a. Kebutuhan Dataset

Dalam mengembangkan sebuah model kecerdasan buatan, diperlukan dataset yang mendukung proses pelatihan, pengujian, dan validasi model tersebut (Hermawan, 2024). Kualitas dari model yang dihasilkan sangat bergantung pada kualitas dataset yang digunakan (Hermawan, 2024). Pada penelitian ini, dibutuhkan dua jenis dataset utama, yaitu:

1. Dataset Klasifikasi CV dan Non-CV

Dataset ini digunakan untuk melatih model dalam membedakan antara dokumen yang termasuk *Curriculum Vitae* (CV) dan dokumen yang bukan CV. Dataset ini terdiri dari dua kelas utama, yaitu:

- CV merupakan dokumen yang memuat informasi pribadi, riwayat pendidikan, pengalaman kerja, dan keterampilan individu.
- Non-CV merupakan dokumen dengan format dan isi yang tidak termasuk dalam kategori CV, seperti laporan, artikel, atau dokumen umum lainnya.

Tabel 4.2 menampilkan kebutuhan dataset klasifikasi CV dan Non-CV.

Tabel 4.2 Kebutuhan Dataset Klasifikasi CV dan Non-CV

Kebutuhan	Deskripsi Kebutuhan
Bahasa	Dataset mencakup teks dokumen dalam bahasa Indonesia dan bahasa Inggris.
Jumlah data	Target jumlah data adalah sebanyak 100 dokumen untuk setiap kelas (CV dan Non-CV).
Isi data	Isi data meliputi teks dari dokumen serta label jenis dokumen (CV atau Non-CV).

Dataset ini mencakup teks dokumen dalam bahasa Indonesia dan bahasa Inggris, mengingat dokumen CV dapat disusun menggunakan salah satu dari kedua bahasa tersebut. Dataset ini akan dibagi menjadi data *training*, data *validation*, dan data *testing* untuk melatih dan mengevaluasi model dalam mengklasifikasi jenis dokumen CV dan Non-CV. Data *training* akan digunakan untuk melatih model, data *validation* digunakan untuk mengevaluasi model saat proses pelatihan berlangsung, dan data *testing* digunakan untuk menguji model secara keseluruhan setelah proses pelatihan selesai.

2. Dataset Kategori Pekerjaan

Dataset kategori pekerjaan memuat informasi mengenai berbagai jenis *skills* (kemampuan) serta kategori pekerjaan yang relevan. Dataset ini digunakan untuk membentuk *vector database*, yang kemudian berperan dalam

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

menambahkan konteks pada *prompt* yang dikirimkan ke model *Large Language Model* (LLM), sehingga rekomendasi pekerjaan yang diberikan menjadi lebih relevan. Tabel 4.3 menampilkan kebutuhan dataset keterampilan dan pekerjaan.

Tabel 4.3 Kebutuhan Dataset Kategori Pekerjaan

Kebutuhan	Deskripsi Kebutuhan
Bahasa	Dataset mencakup data berbahasa Indonesia atau Inggris.
Jumlah data	Target jumlah data adalah sebanyak 2000 data dengan variasi kategori pekerjaan yang luas.
Isi data	Isi data memuat informasi mengenai keterampilan (<i>skills</i>) beserta kategori pekerjaan yang relevan.

b. Kebutuhan Model

Model *deep learning* yang dibangun dalam penelitian ini dirancang untuk mengklasifikasikan jenis dokumen yang diunggah oleh pengguna ke dalam dua kategori, yaitu CV dan Non-CV. Model ini berperan penting dalam tahap awal proses rekomendasi pekerjaan, karena hanya dokumen yang diklasifikasikan sebagai CV yang akan diproses lebih lanjut untuk ekstraksi informasi dan pemberian rekomendasi. Tabel 4.4 menampilkan kebutuhan model pada penelitian ini.

Tabel 4.4 Kebutuhan Model

Kebutuhan	Deskripsi Kebutuhan
Tujuan Model	Mengklasifikasikan dokumen pengguna ke dalam dua kelas, yakni CV dan Non-CV.
Bahasa	Model dilatih dengan menggunakan data berbahasa Inggris dan Indonesia.
Format Input	Tokenisasi teks menggunakan <i>tokenizer</i> dengan padding hingga panjang maksimal 300.
Format Output	Label kelas (0 = Non-CV, 1 = CV)
Kebutuhan Akurasi	Akurasi klasifikasi minimal $\geq 85\%$ untuk evaluasi model

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

c. Kebutuhan *Pre-trained Model*

Dalam pengembangan sistem rekomendasi pekerjaan ini, beberapa *pre-trained model* digunakan untuk memanfaatkan kemampuan bawaan mereka dalam melakukan tugas-tugas spesifik, seperti ekstraksi teks, pengambilan kata kunci, dan pemahaman konteks. Pada penelitian ini, pemilihan *pre-trained model* didasarkan pada kebutuhan fungsional sistem yang ditampilkan pada tabel 4.5.

Tabel 4.5 Kebutuhan *Pre-trained Model*

Jenis <i>Pre-trained Model</i>	Nama Model	Fungsi
OCR	EasyOCR	Mengekstraksi teks dari file gambar dan PDF
Ekstraksi Kata Kunci	KeyBERT	Mengambil kata kunci dari teks hasil ekstraksi
Model <i>Embedding</i>	paraphrase-multilingual-MiniLM-L12-v2	Mengubah teks menjadi vektor <i>embedding</i> yang merepresentasikan makna semantik kalimat.
<i>Large Language Model</i>	<i>LLaMA</i>	Memberikan rekomendasi pekerjaan berbasis <i>prompt</i> dan konteks pengguna.

Pemilihan model didasarkan pada kecocokan dengan kebutuhan fungsional sistem serta keunggulan performa dan efisiensi masing-masing model.

- EasyOCR dipilih karena kemampuannya yang handal dalam mengekstraksi teks dari berbagai format dokumen gambar dan PDF, termasuk dokumen dengan tata letak kompleks.
- KeyBERT digunakan untuk ekstraksi kata kunci karena memanfaatkan *embedding* berbasis *transformer* sehingga dapat menangkap konteks kata lebih baik dibandingkan metode YAKE, dan NER.
- Model *embedding* paraphrase-multilingual-MiniLM-L12-v2 dipilih karena mendukung banyak bahasa dan dapat merepresentasikan makna semantik teks dengan baik dibandingkan model *semantic similarity* lainnya seperti distiluse-base-multilingual-cased-v2.

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



- *LLaMA* sebagai *Large Language Model* digunakan untuk memberikan rekomendasi pekerjaan berbasis pemahaman konteks dan *prompt*, sehingga mampu menghasilkan saran yang relevan dan personal kepada pengguna.

4.2.3.2 Kebutuhan Pengembangan Aplikasi

Aplikasi *chatbot* untuk rekomendasi pekerjaan yang dikembangkan dalam penelitian ini berbasis android dan dirancang untuk memberikan hasil rekomendasi yang interaktif kepada pengguna dalam mencari pekerjaan yang relevan dengan kompetensi mereka. Aplikasi ini dirancang dengan menyesuaikan kebutuhan fungsional dan non-fungsional yang diperoleh dari hasil survei serta analisis kebutuhan aplikasi. Adapun kebutuhan fungsional dan non-fungsional dari aplikasi *chatbot* dijelaskan sebagai berikut.

a. Kebutuhan Fungsional

Kebutuhan fungsional menggambarkan fitur utama yang harus dimiliki aplikasi agar dapat berjalan sesuai tujuan. Tabel 4.6 menampilkan kebutuhan fungsional dari aplikasi *chatbot* rekomendasi pekerjaan.

Tabel 4.6 Kebutuhan Fungsional

Kebutuhan	Deskripsi
Daftar (<i>Register</i>)	Pengguna dapat membuat akun baru dengan memasukkan informasi yang dibutuhkan.
Masuk (<i>Login</i>)	Pengguna dapat masuk ke dalam aplikasi menggunakan akun yang telah terdaftar.
Buat Chat Baru	Pengguna dapat membuat percakapan baru untuk memulai interaksi dengan <i>chatbot</i> rekomendasi kerja.
Pengelolaan Riwayat Chat	Pengguna dapat mengelola riwayat chat, seperti menambahkan, menghapus, dan melihat percakapan sebelumnya.
Input Aplikasi	Pengguna dapat mengirimkan input berupa teks, gambar, atau dokumen PDF untuk dianalisis oleh sistem.

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



b. Kebutuhan Non-Fungsional

Kebutuhan non-fungsional menggambarkan standar kualitas atau kinerja yang harus dipenuhi oleh aplikasi. Tabel 4.7 menampilkan kebutuhan non-fungsional dari aplikasi *chatbot* rekomendasi pekerjaan.

Tabel 4.7 Kebutuhan Non-Fungsional

Kebutuhan	Deskripsi
<i>User Interface</i>	Aplikasi harus memiliki tampilan antarmuka yang menarik dan enak dipandang.
<i>User Experience</i>	Aplikasi harus mudah digunakan, termasuk bagi pengguna yang belum terbiasa menggunakan aplikasi.
Performa Aplikasi	Aplikasi harus mampu memproses input dan memberikan hasil rekomendasi dalam waktu kurang dari 60 detik (≤ 60 detik)
Kompatibilitas	Aplikasi harus mendukung minimal Android 8.0 (Oreo) dan versi yang lebih baru.

Kebutuhan non-fungsional yang telah dirangkum pada tabel 4.7 berperan penting dalam memastikan kualitas keseluruhan dari aplikasi yang dikembangkan. Meskipun tidak secara langsung mempengaruhi fungsi utama aplikasi, elemen-elemen seperti antarmuka pengguna, kemudahan penggunaan, performa sistem, dan kompatibilitas sangat menentukan tingkat kenyamanan dan kepuasan pengguna dalam menggunakan aplikasi.

4.3 System Design

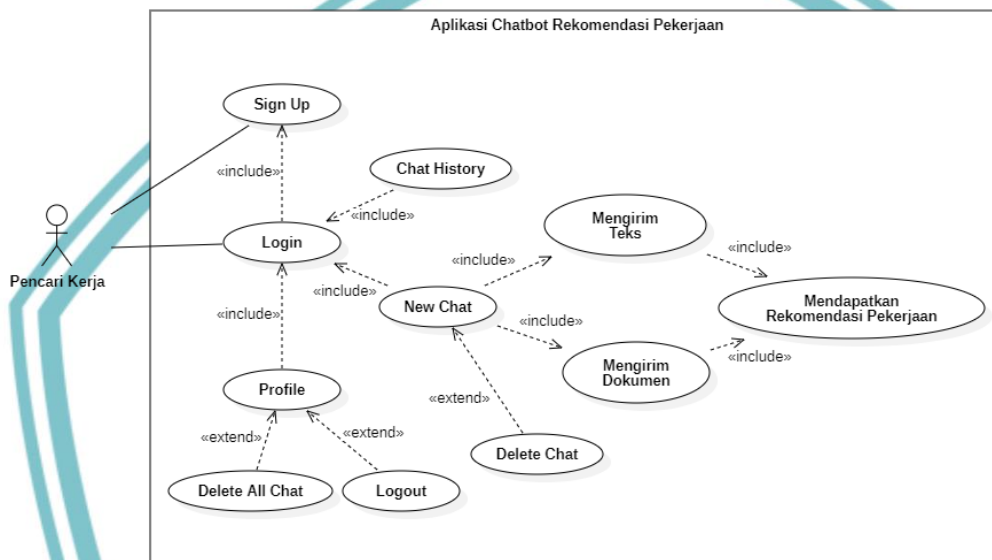
Setelah kebutuhan sistem berhasil diidentifikasi, tahap selanjutnya adalah melakukan perancangan atau desain sistem. Tahap ini bertujuan untuk menyusun rancangan sistem berdasarkan kebutuhan aplikasi yang telah ditentukan sebelumnya. Desain sistem berperan sebagai acuan utama dalam proses implementasi, guna memastikan pengembangan sistem berjalan sesuai dengan tujuan yang telah ditetapkan. Adapun rancangan sistem yang dibuat dapat dilihat sebagai berikut.

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

4.3.1 Use Case Diagram

Use case diagram pada penelitian ini disusun berdasarkan kebutuhan fungsional aplikasi yang telah diidentifikasi pada tahap sebelumnya. Diagram ini menggambarkan hubungan antara aktor, yaitu pengguna, dengan berbagai aksi atau fungsi yang dapat dilakukan dalam aplikasi *chatbot*. *Use case diagram* untuk aplikasi *chatbot* rekomendasi pekerjaan ditampilkan pada Gambar 4.1.



Gambar 4.1 *Use Case Diagram*

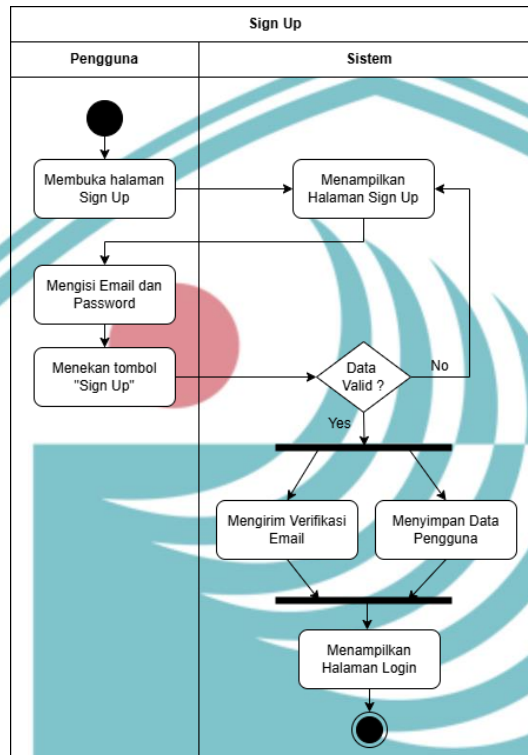
Berdasarkan gambar 4.1 dapat dilihat bahwa aktor pada aplikasi ini adalah pengguna umum, dimana pengguna aplikasi dapat melakukan pendaftaran akun dan *login*. Setelah berhasil *login*, pengguna dapat membuat percakapan baru, mengirim pesan berupa teks, mengunggah dokumen, dan mendapatkan rekomendasi pekerjaan. Selain itu, pengguna juga dapat melihat riwayat percakapan, melihat profil, menghapus percakapan, menghapus semua riwayat percakapan, dan *logout*.

4.3.2 Activity Diagram

Activity diagram adalah diagram yang digunakan untuk menggambarkan alur aktivitas atau proses dalam sistem. Diagram ini menunjukkan urutan aktivitas, keputusan, serta interaksi antara pengguna dan sistem dalam menjalankan suatu proses. *Activity diagram* dibuat berdasarkan *use case diagram*, sehingga membantu

dalam menggambarkan secara visual bagaimana alur logika dari sebuah fitur atau fungsionalitas sistem dijalankan dari awal hingga akhir.

4.3.2.1 Sign Up



Gambar 4.2 Activity Diagram Sign Up

Gambar 4.2 merupakan *activity diagram* yang menjelaskan alur proses pendaftaran akun baru oleh pengguna. Proses diawali dengan pengguna membuka halaman *sign up*. Selanjutnya, sistem akan menampilkan halaman *sign up*. Pengguna kemudian mengisi *email* dan *password*, lalu menekan tombol "*Sign Up*". Sistem akan melakukan validasi terhadap data yang telah diinputkan. Jika data dinyatakan valid, sistem akan mengirimkan verifikasi *email* dan menyimpan data pengguna. Setelah proses tersebut selesai, sistem akan menampilkan halaman *login*, yang menandakan bahwa proses pendaftaran akun telah berhasil.

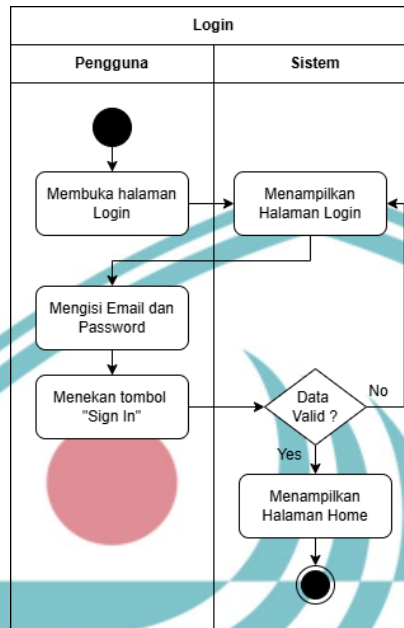
Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumuk dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

4.3.2.2 Login

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

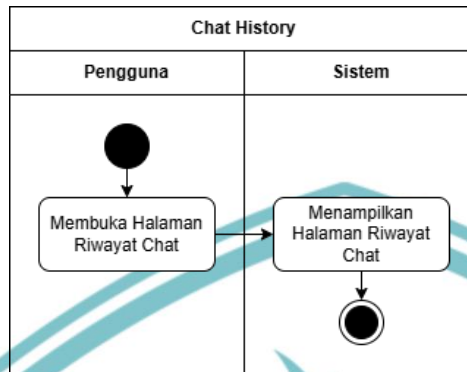


Gambar 4.3 Activity Diagram Login

Gambar 4.3 merupakan *activity diagram* yang menjelaskan alur proses *login* pengguna ke dalam sistem. Proses diawali dengan pengguna membuka halaman *login*. Selanjutnya, sistem akan menampilkan halaman tersebut. Pengguna kemudian memasukkan *email* dan *password* yang telah terdaftar, lalu menekan tombol *sign in*. Sistem akan melakukan validasi dengan mencocokkan kredensial yang dimasukkan dengan data yang tersimpan di *database*. Jika data yang dimasukkan valid, sistem akan menampilkan halaman *home*. Hal ini menandakan bahwa proses *login* telah berhasil dan pengguna mendapatkan akses untuk menggunakan fungsionalitas sistem.



4.3.2.3 Melihat Riwayat Percakapan (*Chat History*)



Gambar 4.4 *Activity Diagram Chat History*

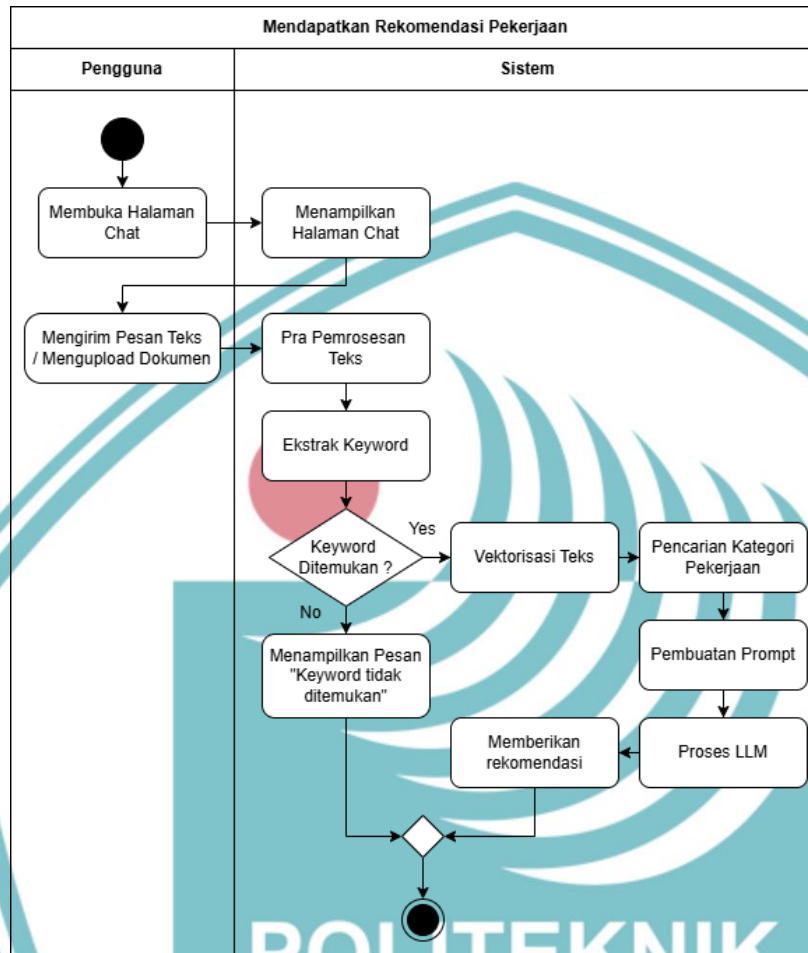
Gambar 4.4 merupakan *activity diagram* yang menggambarkan alur proses pengguna dalam melihat riwayat *chat* pada sistem. Proses dimulai ketika pengguna membuka halaman riwayat *chat* melalui antarmuka aplikasi. Setelah aksi tersebut dilakukan, sistem akan merespons dengan menampilkan halaman riwayat *chat* yang berisi daftar percakapan atau interaksi sebelumnya antara pengguna dan *chatbot*.

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

POLITEKNIK
NEGERI
JAKARTA

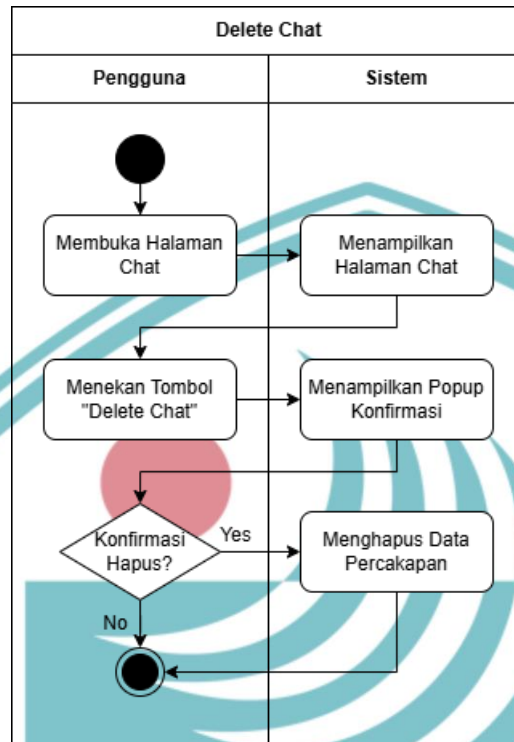
4.3.2.4 Mendapatkan Rekomendasi Pekerjaan



Gambar 4.5 *Activity Diagram* Mendapatkan Rekomendasi Pekerjaan

Gambar 4.5 merupakan *activity diagram* yang menjelaskan alur proses inti dari sistem, yaitu untuk mendapatkan rekomendasi pekerjaan. Proses ini diawali oleh pengguna yang membuka halaman *chat*. Sistem akan menampilkan halaman *chat* tersebut. Selanjutnya, pengguna berinteraksi dengan sistem dengan cara mengirim pesan teks atau mengunggah dokumen. Sistem akan melakukan pra-pemrosesan teks yang diterima, kemudian mengekstrak *keyword*. Sistem akan memeriksa apakah *keyword* ditemukan. Jika *keyword* tidak ditemukan, sistem akan menampilkan pesan "*Keyword tidak ditemukan*". Jika *keyword* ditemukan, sistem akan melanjutkan dengan vektorisasi teks, pencarian kategori pekerjaan, pembuatan prompt, dan proses LLM (*Large Language Model*). Setelah proses LLM, sistem akan memberikan rekomendasi pekerjaan kepada pengguna.

4.3.2.5 Menghapus Percakapan (*Delete Chat*)



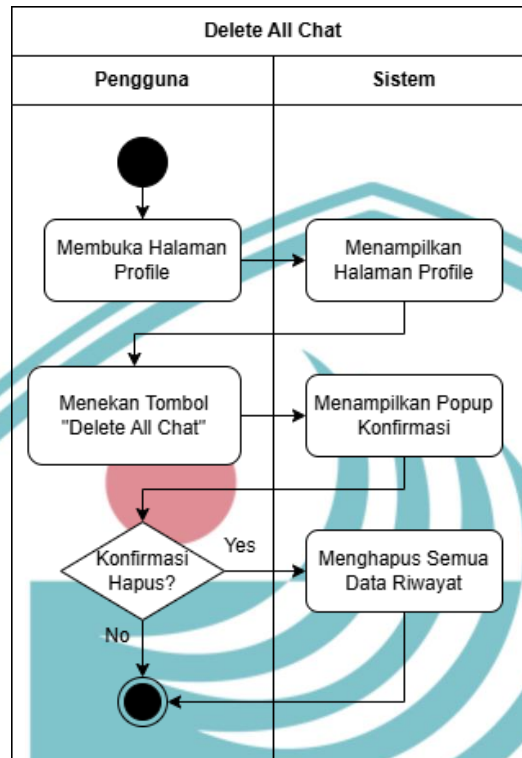
Gambar 4.6 Activity Diagram Delete Chat

Gambar 4.6 merupakan *activity diagram* yang menjelaskan alur proses pengguna saat menghapus riwayat percakapan. Alur dimulai ketika pengguna yang berada di halaman *chat* menekan tombol "*Delete Chat*". Sistem kemudian akan menampilkan *popup* konfirmasi sebagai langkah preventif untuk menghindari penghapusan yang tidak disengaja. Pengguna selanjutnya memberikan keputusan pada pilihan konfirmasi tersebut. Apabila pengguna menyetujui untuk menghapus, maka sistem akan melanjutkan proses dengan menghapus data percakapan terkait dari *database*. Proses ini berakhir setelah data berhasil dihapus secara permanen dari sistem.

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

4.3.2.6 Menghapus Semua Riwayat Percakapan (*Delete All Chat*)



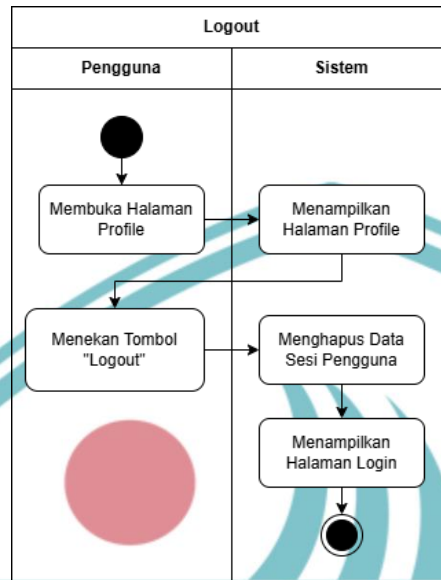
Gambar 4.7 Activity Diagram Delete All Chat

Gambar 4.7 merupakan *activity diagram* yang menjelaskan alur proses untuk menghapus keseluruhan riwayat percakapan pengguna. Proses ini dapat diakses oleh pengguna melalui halaman *profile*. Pengguna kemudian menekan tombol "*Delete All Chat*" untuk memulai proses. Mengingat tindakan ini bersifat permanen dan akan menghapus semua data, sistem akan menampilkan *popup* konfirmasi untuk memastikan keputusan pengguna. Apabila pengguna memberikan persetujuan pada *popup* tersebut, sistem akan melanjutkan eksekusi dengan menghapus seluruh data riwayat percakapan milik pengguna dari *database*. Alur ini memberikan pengguna kontrol penuh untuk mereset riwayat interaksinya dengan sistem secara menyeluruh.

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

4.3.2.7 Logout



Gambar 4.8 Activity Diagram Logout

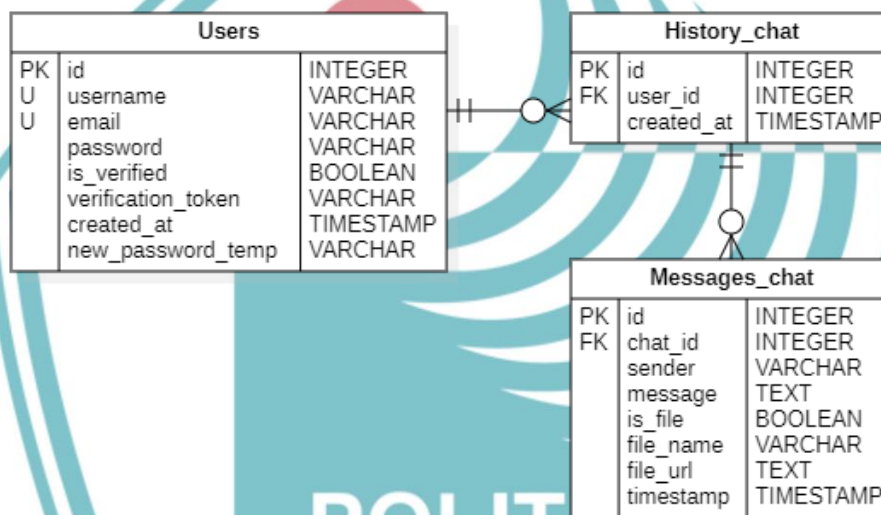
Gambar 4.8 merupakan *activity diagram* yang menjelaskan alur proses pengguna ketika keluar dari sistem (*logout*). Proses ini diawali oleh pengguna yang mengakses halaman *profile*. Pengguna kemudian menekan tombol "*logout*" untuk mengakhiri sesi aktifnya. Setelah tombol ditekan, sistem akan segera memproses permintaan tersebut dengan menghapus data sesi pengguna yang sedang aktif dari sistem. Selanjutnya, sistem akan mengalihkan pengguna kembali ke halaman *login*. Hal ini menandakan bahwa pengguna telah berhasil keluar dari sistem secara aman dan sesi sebelumnya telah diakhiri.

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

4.3.3 Database

Database yang digunakan dalam penelitian ini adalah MySQL. Database ini dipilih karena kemudahan integrasinya dengan berbagai teknologi *backend* yang berfungsi sebagai jembatan antara aplikasi *android* dan server. Sebagai database, MySQL menyimpan data dalam format tabel yang terdiri dari baris dan kolom, yang memungkinkan relasi yang jelas dan konsisten antar data. Untuk memvisualisasikan struktur dan hubungan antar tabel tersebut, digunakan *entity relationship diagram* (ERD) yang disajikan pada Gambar 4.9.



Gambar 4.9 ER Diagram

Gambar 4.9 di atas menggambarkan diagram relasi entitas atau ERD pada penelitian ini. Struktur diagram ini terdiri dari tiga entitas utama yang menggambarkan tabel penyimpanan data dari aplikasi. Ketiga entitas tersebut adalah *users* untuk menyimpan data profil pengguna, *history_chat* untuk mengelompokkan setiap sesi percakapan, dan *messages_chat* untuk menyimpan detail dari setiap pesan yang dikirim oleh pengguna maupun *chatbot*.

Dari diagram tersebut, dapat diuraikan relasi antar entitas sebagai berikut. Satu pengguna (*users*) dapat memiliki banyak riwayat percakapan (*history_chat*). Namun, setiap riwayat percakapan hanya dapat dimiliki oleh satu pengguna. Ini menunjukkan adanya relasi *one-to-many* antara tabel *users* dan *history_chat*.

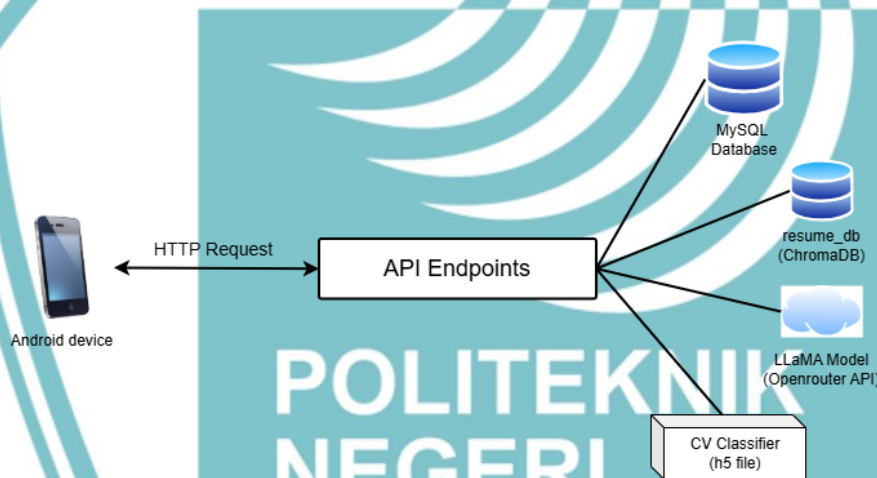
Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Selanjutnya, satu riwayat percakapan (*history_chat*) dapat terdiri dari banyak pesan (*messages_chat*), namun setiap pesan hanya tercatat dalam satu riwayat percakapan. Hal ini juga menunjukkan adanya relasi *one-to-many* antara *history_chat* dan *messages_chat*.

Terakhir, diagram ini juga mengindikasikan sifat opsional dari hubungan tersebut. Seorang pengguna bisa saja belum memiliki riwayat percakapan contohnya adalah pengguna yang baru mendaftar, yang menunjukkan bahwa keberadaan data di tabel *history_chat* bersifat opsional bagi *users*. Namun sebaliknya, setiap catatan pada *history_chat* wajib terhubung ke satu *users*, dan setiap pesan pada *messages_chat* wajib terhubung ke satu *history_chat*.

4.3.4 Architecture Application



Gambar 4.10 Arsitektur Aplikasi

Gambar 4.10 menyajikan arsitektur sistem yang dirancang untuk aplikasi *chatbot* rekomendasi pekerjaan. Arsitektur ini mengadopsi pendekatan *client-server*, yang memisahkan antara lapisan antarmuka pengguna (*client*) dan lapisan pemrosesan serta data (*server-side*). Komponen dalam arsitektur ini mencakup:

- Aplikasi android, yang berfungsi sebagai antarmuka pengguna untuk mengirimkan input berupa teks atau dokumen.
- *API endpoints* yang dibangun menggunakan FastAPI, sebagai penghubung antara aplikasi dan berbagai layanan pemrosesan.

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

55

- Database MySQL, yang digunakan untuk menyimpan data pengguna, riwayat chat, dan pesan.
- Model CV Classifier, yang bertugas untuk mengidentifikasi apakah dokumen yang diunggah merupakan CV atau bukan.
- *Vector database* dibuat dengan ChromaDB yang berfungsi untuk mencari kategori pekerjaan yang relevan berdasarkan keterampilan yang diekstrak dari CV.
- Model *LLaMA*, yang digunakan untuk menghasilkan respons berupa rekomendasi pekerjaan berdasarkan informasi keterampilan dan kategori pekerjaan yang diperoleh sebelumnya.

Semua komponen pemrosesan ini dapat diakses oleh aplikasi android melalui API *endpoints*. Arsitektur ini dirancang untuk menggambarkan bagaimana komponen yang dibuat untuk membangun aplikasi *chatbot* dapat saling terintegrasi dan berkomunikasi.

4.3.5 Wireframe Design

Wireframe merupakan sketsa atau gambaran awal dari tata letak sebuah antarmuka pengguna dari aplikasi berdasarkan fitur-fitur yang telah ditentukan sebelumnya. Tujuan dibuatnya *wireframe design* adalah untuk memberikan gambaran atau struktur dasar dari halaman aplikasi yang akan dibuat. Desain *wireframe* pada penelitian ini, terdiri dari halaman *landing*, halaman pendaftaran akun, halaman *login*, halaman *home*, halaman *chat*, dan halaman *profile* yang dapat dilihat pada Lampiran 2.

4.4 Model Development

Setelah tahap pengumpulan dan analisis data yang menghasilkan identifikasi kebutuhan sistem, tahap selanjutnya adalah pengembangan model. Proses ini mengikuti tahapan dalam *AI project lifecycle*, yang mencakup *data acquisition and preparation*, *model development and training*, serta *model evaluation and refinement*.



4.4.1 Data Acquisition and Preparation

Sebelum mengembangkan model, dataset yang diperlukan untuk proses pelatihan harus disiapkan terlebih dahulu. Pada tahap ini dilakukan pencarian dataset yang relevan dengan kebutuhan penelitian, dilanjutkan dengan proses pengumpulan data apabila dataset yang dibutuhkan belum tersedia. Setelah itu, data yang telah dikumpulkan akan dipersiapkan agar dapat digunakan untuk melatih model yang akan diimplementasikan dalam penelitian ini. Berdasarkan kebutuhan dataset, terdapat dua jenis dataset yang diperlukan untuk dikumpulkan dan disiapkan dalam penelitian ini, yaitu dataset yang berisi teks berlabel CV dan Non-CV, serta dataset yang memuat keterampilan dan kategori pekerjaan. Berikut adalah proses pengumpulan dan persiapan dataset resume serta pengumpulan dan persiapan dataset kategori pekerjaan.

4.4.1.1 Pengumpulan dan Persiapan Dataset Resume

Dataset berlabel CV dan Non-CV cukup sulit ditemukan secara publik. Oleh karena itu, dalam penelitian ini, dataset tersebut akan dikembangkan secara mandiri untuk melatih model klasifikasi jenis dokumen CV dan Non-CV. Proses pengumpulan dan pembuatan dataset *resume* dilakukan dengan lima tahapan yang meliputi *Manual Data Collection*, *Text Extraction*, *Data Labelling*, *Text Cleaning*, *Data Augmentation*, dan *Stopword Removal*. Berikut adalah proses pengumpulan dan persiapan dataset *resume*.

1. Pengumpulan Data secara Manual (*Manual Data Collection*)

Pada tahap ini, dilakukan proses pengumpulan data secara manual dengan mencari dokumen CV dan Non-CV yang masing-masing dibatasi maksimal dua halaman. Pengumpulan dokumen CV diperoleh dengan mengunduh dokumen dari kumpulan data *resume* yang tersedia pada platform Kaggle, yang dapat diakses melalui tautan: <https://www.kaggle.com/snehaanbhawal/resume-dataset>. Namun, dataset tersebut hanya menyediakan dokumen CV berbahasa Inggris. Untuk memenuhi kebutuhan dataset yang juga mencakup dokumen berbahasa Indonesia, dilakukan pengumpulan tambahan melalui situs Canva dengan

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



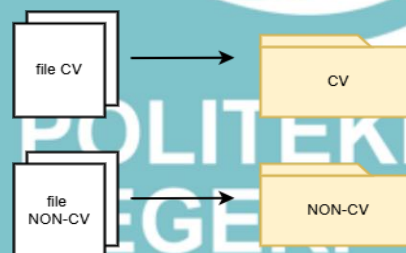
Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

memanfaatkan *template* dokumen CV berbahasa Indonesia dan Inggris yang tersedia secara publik dan memungkinkan untuk digunakan sebagai bagian dari dataset.

Selain itu, pengumpulan dokumen Non-CV dilakukan dengan mencari dokumen yang tidak termasuk dalam kategori *Curriculum Vitae* dan dibatasi maksimal dua halaman. Contoh dokumen Non-CV meliputi dokumen tugas pelajaran, sertifikat, surat izin sakit, surat keterangan, surat edaran, artikel, surat penawaran kerja, surat permohonan cuti, dan dokumen sejenis lainnya. Sumber dokumen Non-CV ini diperoleh dari dokumen publik yang tersedia di internet serta dari *template* yang disediakan oleh situs Canva, yang memungkinkan untuk digunakan sebagai bagian dari dataset.

Setelah kedua jenis dokumen terkumpul, dokumen CV disimpan dalam folder bernama CV, sedangkan dokumen Non-CV disimpan dalam folder bernama Non-CV, untuk memudahkan proses pelabelan dan pelatihan model klasifikasi dokumen sesuai dengan ilustrasi pada gambar 4.11.



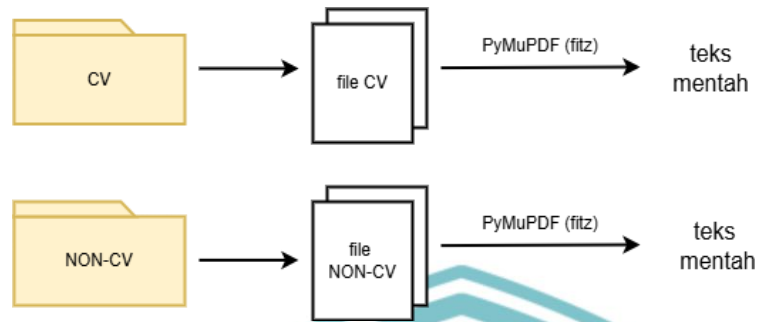
Gambar 4.11 Ilustrasi Pengumpulan data

2. Ekstraksi Teks (*Text Extraction*)

Tahap selanjutnya adalah mengekstrak isi dokumen PDF menjadi bentuk teks mentah yang dapat diolah secara komputasional. Proses ekstraksi dilakukan menggunakan *library* PyMuPDF (*fitz*), yang memungkinkan pengambilan teks dari setiap halaman dalam dokumen PDF. Proses ekstraksi teks ini dapat dilihat pada gambar 4.12.

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



Gambar 4.12 Ilustrasi Ekstraksi Teks.

3. Pelabelan Data (*Data Labelling*)

Setelah teks berhasil diekstrak, proses pelabelan data dilakukan secara manual berdasarkan klasifikasi folder tempat asal dokumen. Metode pelabelan ini dikenal sebagai *manual annotation*. Ilustrasi tahapan ini dapat dilihat pada gambar 4.13.



Gambar 4.13 Ilustrasi Pelabelan Data

Berdasarkan gambar 4.13, pelabelan data didasari oleh lokasi penyimpanan data dokumen berdasarkan tahap pengumpulan data secara manual, dimana label 1 akan diberikan untuk dokumen yang berasal dari folder CV, sedangkan label 0 akan diberikan untuk dokumen yang berasal dari folder Non-CV. Hasil pelabelan ini kemudian dikompilasi dalam sebuah file CSV dengan dua kolom utama, yaitu *text* yang berisi isi dokumen dalam bentuk *string* teks dan label yang berisi label klasifikasi. Dataset akhir yang dihasilkan terdiri dari total 200

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

data, dengan rincian 100 data dengan label 1 (CV) dan 100 data dengan label 0 (Non-CV).

4. Pembersihan Teks (*Text Cleaning*)

Untuk memperbanyak jumlah data pada dataset, dilakukan proses augmentasi data. Namun, sebelum proses augmentasi diterapkan pada teks dalam dataset, dilakukan terlebih dahulu tahap pembersihan teks (*text cleaning*). Proses pembersihan ini bertujuan untuk menghilangkan elemen-elemen yang tidak relevan sehingga teks menjadi lebih terstruktur dan siap untuk diolah lebih lanjut. Pada tahapan ini, terdapat beberapa langkah dalam melakukan pembersihan teks yang dilakukan pada penelitian ini, diantaranya adalah sebagai berikut.

- a. Konversi ke huruf kecil
Seluruh karakter dalam teks diubah menjadi huruf kecil untuk menyamakan representasi kata.
- b. Penghapusan angka
Seluruh angka di dalam teks dihapus dengan menggunakan ekspresi reguler karena dianggap tidak memberikan kontribusi yang signifikan dalam klasifikasi dokumen.
- c. Penghapusan tanda baca
Tanda baca seperti titik, koma, tanda seru, dan simbol lainnya dihapus dengan menggunakan ekspresi reguler untuk menyederhanakan teks.
- d. Penghapusan spasi berlebih
Spasi ganda, baris baru, atau tab diubah menjadi satu spasi tunggal dengan menggunakan ekspresi reguler, dan spasi di awal serta akhir teks dihapus.

5. Augmentasi Data (*Data Augmentation*)

Setelah teks dibersihkan, dilakukan proses augmentasi data untuk meningkatkan variasi dan jumlah data pelatihan. Teknik augmentasi yang digunakan dalam penelitian ini adalah penerjemahan teks (*translation-based*

augmentation), yaitu dengan menerjemahkan teks asli ke bahasa lain. Proses ini dilakukan dengan bantuan *library* deep-translator dan mendukung berbagai bahasa, seperti bahasa inggris dan bahasa indonesia. Dengan pendekatan ini, struktur kalimat dapat bervariasi tanpa mengubah makna. Jumlah dataset akhir setelah dataset dilakukan augmentasi adalah 400 data, dengan rincian 200 data dengan label 1 (CV) dan 200 data dengan label (Non-CV).

6. Penghapusan *Stopword* (*Stopword Removal*)

Penghapusan *stopword* merupakan bagian dari proses pembersihan lanjutan yang bertujuan untuk menghilangkan kata-kata umum yang tidak memiliki kontribusi signifikan terhadap makna teks, seperti "dan", "yang", "adalah", dan sebagainya. *Stopword* sering kali dianggap sebagai *noise* dalam pemrosesan teks, terutama saat digunakan untuk analisis atau pelatihan model *machine learning*. Dalam penelitian ini, *stopword removal* dilakukan menggunakan *library Natural Language Toolkit* (NLTK) untuk bahasa inggris dan daftar *stopword* dari *library* Sastrawi untuk bahasa indonesia, sehingga teks yang dihasilkan lebih ringkas dan relevan untuk digunakan dalam pelatihan model klasifikasi.

4.3.1.2 Pengumpulan dan Persiapan Dataset Kategori Pekerjaan

Dataset yang memuat informasi mengenai keterampilan dan kategori pekerjaan yang dibutuhkan dalam penelitian ini diperoleh melalui platform Kaggle, dan dapat diakses melalui tautan: [kaggle.com/datasets/snehaanbhawal/resume-dataset](https://www.kaggle.com/datasets/snehaanbhawal/resume-dataset). Berdasarkan dokumentasi dari pemilik dataset, data tersebut dikumpulkan dari situs www.livecareer.com dan telah dilengkapi dengan informasi mengenai keterampilan serta kategori pekerjaan. Dataset ini digunakan untuk menambahkan konteks pada *prompt* teks yang akan dikirimkan ke LLM, sehingga respons rekomendasi pekerjaan yang dihasilkan menjadi lebih relevan dan akurat. Untuk mempersiapkan data dalam dataset ini sebagai konteks tambahan pada *prompt* LLM, terlebih dahulu dilakukan proses pemrosesan terhadap dataset tersebut sebelum dimasukkan ke dalam *vector database*.

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



- Hak Cipta :**
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
 2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Proses persiapan dataset serta pembuatan *vector database* meliputi beberapa tahapan, diantaranya *Load Dataset*, *Data Preprocessing*, *Information Extraction*, *Text Vectorization*, dan *Vector Indexing*. Tahapan *Data Preprocessing* yang dilakukan mencakup proses *Special Character Removal*, *Case Folding*, *Stopword Removal*, dan *Lemmatization*. Setelah proses *preprocessing* selesai, dilanjutkan dengan ekstraksi informasi, transformasi teks menjadi vektor, dan penyimpanan vektor ke dalam *vector database*. Berikut adalah proses persiapan dataset serta pembuatan *vector database*.

1. *Load Dataset*

Proses ini merupakan proses awal, di mana dataset yang digunakan berasal dari sumber publik yang tersedia di platform Kaggle. Dataset ini terdiri dari beberapa kolom, di antaranya kolom *Resume_str* yang berisi teks lengkap dari dokumen *resume* atau *curriculum vitae* dalam format *string*, serta kolom *category* yang menunjukkan kategori pekerjaan yang sesuai dengan isi *resume* tersebut.

2. *Data Preprocessing*

Sebelum mengambil informasi mengenai keterampilan pengguna dari kolom *resume_str*, dilakukan tahap *data preprocessing* untuk membersihkan dan menyiapkan teks agar lebih mudah dianalisis. Tahapan *preprocessing* yang dilakukan pada dataset ini meliputi beberapa langkah berikut:

a. *Special Character Removal*

Langkah pertama adalah menghapus karakter-karakter khusus dan tanda baca yang tidak relevan dari teks dalam kolom *resume_str*. Penghapusan ini juga mencakup eliminasi tautan seperti *http*, *https*, dan *www*, yang umumnya tidak memiliki kontribusi bermakna terhadap analisis keterampilan.

b. *Case Folding*

Setelah karakter khusus dihapus, seluruh teks dalam kolom *resume_str* dikonversi menjadi huruf kecil (*lowercase*). Proses ini dilakukan untuk

menyamakan representasi kata, sehingga variasi dalam penggunaan huruf kapital tidak mempengaruhi hasil analisis.

c. *Stopword Removal*

Selanjutnya, dilakukan penghapusan *stopword*, yaitu kata-kata umum yang sering muncul namun tidak membawa makna penting dalam konteks analisis, seperti “dan”, “yang”, “the”, “is”, dan sebagainya. Langkah ini membantu mengurangi *noise* dalam data.

d. *Lemmatization*

Tahap terakhir adalah *lemmatization*, yaitu proses mengubah setiap kata dalam teks menjadi bentuk dasarnya (*lemma*). Tujuannya adalah untuk menyederhanakan variasi bentuk kata serta mengurangi jumlah fitur unik, sehingga kata-kata dengan akar yang sama dapat dianalisis sebagai satu entitas yang konsisten.

3. *Information Extraction*

Setelah dilakukan proses *preprocessing* terhadap teks pada kolom *resume_str*, tahap selanjutnya adalah melakukan *information extraction* untuk mengambil informasi penting berupa kemampuan (*skills*) dari setiap dokumen *resume*. Proses ekstraksi ini dilakukan dengan memanfaatkan model spaCy yang dikombinasikan dengan fitur *entity ruler*. Model spaCy yang digunakan dalam tahap ini adalah *en_core_web_lg*, yaitu model bahasa *pre-trained* berukuran besar dari spaCy yang dilengkapi dengan representasi vektor dan dukungan pengenalan entitas bernama NER (*Named Entity Recognition*). Untuk mengidentifikasi keterampilan, digunakan sebuah *entity ruler* yang diinisialisasi melalui file pola *jz_skill_patterns.jsonl*, yang dapat diperoleh dari repositori github.com/kushal-022/Resume-Parser-using-Spacy. File tersebut berisi daftar pola keterampilan yang umum ditemukan dalam *resume*.

Selain itu, sistem ini juga memiliki kemampuan untuk mengenali kategori pekerjaan dengan menambahkan entitas tambahan dari data kategori pekerjaan yang terdapat dalam kolom *category* pada dataset. Setiap kategori ditambahkan sebagai pola dalam *entity ruler*, yang selanjutnya diproses oleh model untuk

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

mendeteksi entitas dengan label *job-category*. Setelah keterampilan diambil, dilakukan penghapusan duplikasi pada data keterampilan yang didapatkan dan memastikan daftar keterampilan yang diperoleh bersifat unik untuk setiap dokumen.

4. *Text Vectorization*

Tahap ini merupakan proses untuk mengubah teks dari kolom *resume_str* menjadi representasi numerik dalam bentuk vektor yang dapat dipahami oleh mesin. Pada tahap ini digunakan model *sentence transformer* dengan arsitektur *distiluse-base-multilingual-cased-v2* dari pustaka *sentence-transformers*. Model ini mampu menghasilkan *sentence embedding* berkualitas tinggi dari teks dalam berbagai bahasa, termasuk bahasa Indonesia dan Inggris. Setiap teks resume diproses satu per satu, kemudian dikodekan menjadi sebuah vektor. Output dari proses ini adalah sebuah *array* numerik berdimensi tetap, yang merepresentasikan makna semantik dari teks tersebut.

5. *Vector Indexing*

Setelah teks dikonversi menjadi vektor, tahapan berikutnya adalah menyimpan informasi tersebut ke dalam *vector database*. Dalam hal ini, digunakan *chromaDB*, sebuah *vector database* yang dirancang untuk menyimpan dan mengelola *embedding* secara terstruktur. Proses *indexing* dilakukan dengan cara menambahkan setiap *embedding* ke dalam koleksi bernama "resumes" pada database lokal *resume_db*. Setiap data yang dimasukkan terdiri dari ID unik, teks asli, *embedding* vektornya, serta metadata tambahan seperti kategori pekerjaan (*category*) dan daftar keterampilan (*skills*). *Vector database* ini berfungsi sebagai mesin pencari berbasis semantik yang digunakan untuk menemukan kategori pekerjaan berdasarkan kesamaan semantik pada *skill* yang diinputkan pengguna dengan *skill* yang ada pada *vector database*.

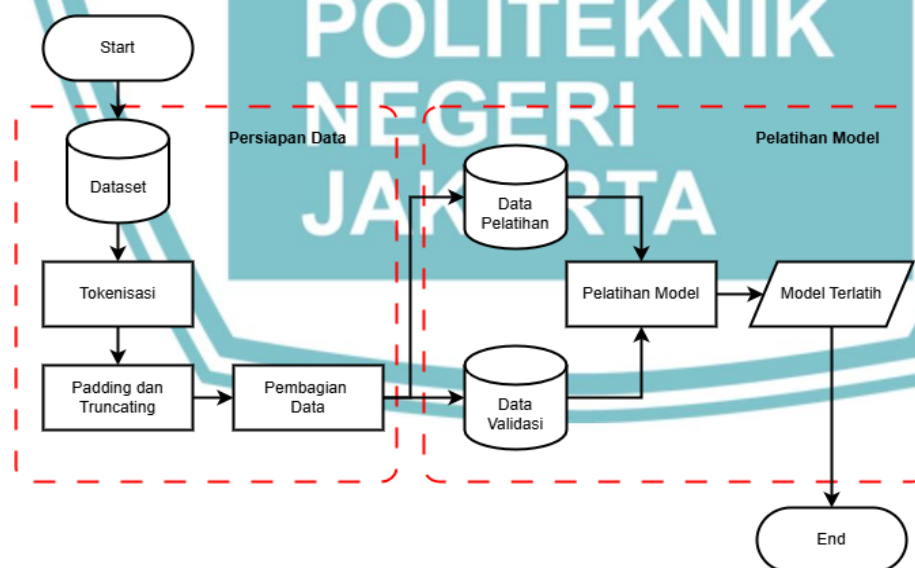
4.4.2 *Model Development and Training*

Setelah dataset yang akan digunakan telah dipersiapkan, langkah selanjutnya adalah mengembangkan dan melatih model yang merupakan komponen utama dari sistem.

Pertama, dibangun model klasifikasi untuk membedakan antara dokumen *curriculum vitae* (CV) dan dokumen non-CV. Selanjutnya, diimplementasikan pustaka KeyBERT yang didukung oleh model *sentence transformer* untuk mengekstrak kata kunci (*keyword*) yang merepresentasikan kemampuan atau keterampilan yang terdapat dalam dokumen CV. Terakhir, dilakukan proses *prompt engineering*, yaitu perancangan *prompt* atau instruksi yang tepat untuk diberikan kepada *large language model*, agar mampu memberikan rekomendasi pekerjaan yang relevan berdasarkan informasi dari CV pengguna.

4.4.2.1 Pengembangan Model Klasifikasi Dokumen CV dan Non-CV

Model yang dikembangkan untuk mengenali apakah dokumen yang diunggah merupakan *curriculum vitae* (CV) atau bukan, dibangun menggunakan pendekatan *deep learning* menggunakan metode *long short-term memory* (LSTM), yaitu salah satu jenis jaringan saraf tiruan yang efektif dalam memproses dan menganalisis data berbasis teks. Dataset yang digunakan untuk melatih model merupakan dataset yang digunakan untuk melatih model ini adalah dataset CV dan Non-CV yang telah disiapkan pada tahap sebelumnya. Proses pelatihan model klasifikasi dokumen CV dan Non-CV ditampilkan pada gambar 4.14.



Gambar 4.14 Alur Pelatihan Model Klasifikasi Dokumen

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Berdasarkan gambar 4.14, terdapat 2 tahapan pada proses pelatihan model yang digambarkan pada garis berwarna merah. Tahapan pertama berfokus pada pemrosesan data awal, sedangkan tahapan kedua mencakup pelatihan model menggunakan LSTM.

a. Persiapan data

Sebelum dilakukan pelatihan model, data teks perlu dikonversi menjadi representasi numerik agar dapat diproses oleh model. Proses ini meliputi:

- Tokenisasi, dimana setiap kata dalam teks diubah menjadi indeks bilangan bulat menggunakan *Tokenizer* dari Keras. *Tokenizer* dikonfigurasi dengan jumlah maksimal kata sebanyak 5000 dan token khusus untuk kata yang tidak dikenal (<OOV>).
- *Padding* dan *truncating* untuk menyamakan panjang seluruh urutan kata menjadi 300 token. Ini dilakukan menggunakan fungsi *pad_sequences* dengan metode *padding* dan *truncating* di bagian akhir urutan, sehingga seluruh input memiliki dimensi yang sama.
- Pembagian Data, dimana dataset dibagi menjadi dua bagian, yaitu data pelatihan sebanyak 0.8 dan data validasi sebanyak 0.2. Data pelatihan digunakan untuk melatih model, sedangkan data validasi digunakan untuk memantau kinerja model selama pelatihan.

b. Pelatihan model

Setelah pembagian data, selanjutnya adalah membangun dan melatih model klasifikasi dokumen CV dan Non-CV. Model yang digunakan adalah *long short-term memory* (LSTM), yaitu jenis jaringan saraf tiruan yang efektif dalam memproses data urutan (*sequential data*), khususnya teks. Arsitektur model yang dirancang pada penelitian ini terdiri atas beberapa lapisan, yaitu:

- *Embedding layer* yang berfungsi untuk mengubah setiap kata yang sudah di tokenisasi menjadi vektor berdimensi 128. *Embedding* dilakukan terhadap 5000 kata paling umum dalam data.



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

- *LSTM layer*, dimana pada model ini menggunakan dua lapisan LSTM berturut-turut, dengan jumlah unit masing-masing 64 dan 32. Lapisan pertama mengembalikan urutan agar dapat diteruskan ke lapisan LSTM berikutnya.
- Untuk mencegah *overfitting*, lapisan *dropout* dengan nilai 0.3 ditambahkan setelah LSTM pertama. Lapisan ini secara acak mengabaikan sebagian neuron selama pelatihan.
- *Dense layer (output layer)* merupakan lapisan terakhir yang terdiri dari satu neuron dengan fungsi aktivasi *sigmoid*. Lapisan ini digunakan untuk melakukan klasifikasi biner (CV atau Non-CV).

Model dikompilasi menggunakan fungsi *loss binary crossentropy*, *optimizer* Adam dengan *learning rate* sebesar 0.001, dan metrik evaluasi akurasi. Untuk mencegah *overfitting*, digunakan teknik *early stopping* yang memantau nilai *validation loss (val_loss)*. Pelatihan akan dihentikan jika tidak ada penurunan nilai *val_loss* selama 2 *epoch* berturut-turut. Model dilatih selama maksimum 50 *epoch* dengan ukuran *batch* sebesar 32. Setelah pelatihan selesai, model dievaluasi menggunakan data uji (data validasi). Hasil evaluasi menunjukkan bahwa model mampu mencapai akurasi sebesar 60% dalam mengklasifikasikan dokumen sebagai CV atau Non-CV.

Karena hasil akurasinya berada dibawah kebutuhan akurasi model yakni diatas 85 %, maka dilakukan *hyperparameter tuning* pada model untuk meningkatkan performa model. *Hyperparameter tuning* merupakan proses pengaturan ulang nilai-nilai parameter yang mengontrol proses pelatihan model, dengan tujuan untuk memperoleh konfigurasi terbaik yang dapat meningkatkan akurasi dan generalisasi model terhadap data baru (Wibowo and Isnain, 2025). Konfigurasi *tuning* pada model ini ditampilkan pada tabel 4.8.

Tabel 4.8 Konfigurasi *hyperparameter tuning*

Aspek <i>Tuning</i>	Parameter <i>Tuning</i>
Jumlah Unit pada <i>LSTM Layer</i>	[32, 64, 128]

Tingkat <i>Dropout</i>	[0.3, 0.5, 0.7]
<i>Learning Rate</i>	[0.001, 0.0005, 0.0001]
<i>Batch Size</i>	[16, 32, 64]
<i>Optimizer</i>	['adam', 'rmsprop', 'adamax']

Berdasarkan tabel 4.8, aspek yang diuji dalam proses *hyperparameter tuning* meliputi:

- Jumlah unit pada *LSTM layer* untuk melihat pengaruhnya terhadap kemampuan model dalam menangkap pola sekuensial dalam data teks.
- Tingkat *dropout* untuk menemukan keseimbangan antara mencegah *overfitting* dan mempertahankan performa pelatihan.
- Learning rate* untuk mengoptimalkan kecepatan dan stabilitas pelatihan model.
- Batch Size* untuk mengetahui konfigurasi terbaik dalam mempercepat proses pelatihan tanpa mengorbankan performa.
- Optimizer* untuk mencari *optimizer* yang terbaik yang dapat digunakan sesuai kasus penelitian.

Setiap kombinasi nilai *hyperparameter* tersebut akan diuji secara bertahap menggunakan pendekatan eksperimental. Hasil dari setiap kombinasi akan menggunakan data validasi yang sama, serta menggunakan teknik *early stopping* untuk menghentikan pelatihan jika model tidak menunjukkan peningkatan. Kombinasi parameter dengan performa terbaik akan dipilih sebagai konfigurasi akhir model. Tujuan dari proses ini adalah untuk meningkatkan akurasi model agar memenuhi kebutuhan sistem, yaitu di atas 85%. Hasil 10 eksperimen terbaik dari proses *hyperparameter tuning* dapat dilihat pada tabel 4.9.

Tabel 4.9 Hasil Eksperimen *Hyperparameter Tuning*

No	<i>LSTM Units</i>	<i>Dropout</i>	<i>Batch Size</i>	<i>Learning Rate</i>	<i>Optimizer</i>	<i>Accuracy</i>
1	64	0.7	32	0.001	<i>adamax</i>	0.9875
2	32	0.7	64	0.001	<i>adamax</i>	0.975

Hak Cipta :

- Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
- Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

3	32	0.5	32	0.001	<i>adamax</i>	0.975
4	32	0.7	16	0.001	<i>adamax</i>	0.9625
5	32	0.3	16	0.001	<i>adamax</i>	0.9625
6	64	0.7	32	0.0005	<i>adam</i>	0.9625
7	64	0.3	16	0.001	<i>adamax</i>	0.9625
8	32	0.5	64	0.0005	<i>adamax</i>	0.9625
9	128	0.3	32	0.0005	<i>adamax</i>	0.9625
10	128	0.3	16	0.001	<i>adamax</i>	0.95

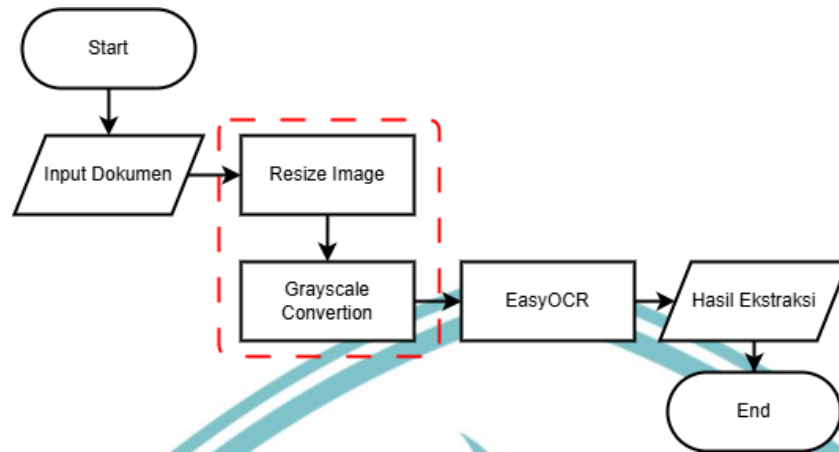
Hasil dari proses *hyperparameter tuning* pada tabel 4.9 menunjukkan bahwa konfigurasi terbaik adalah penggunaan LSTM sebanyak 64 unit, *dropout* sebesar 0.7, *batch size* 32, *learning rate* 0.001, dan *optimizer adamax*. Konfigurasi ini menghasilkan akurasi validasi sebesar 98.75%, yang memenuhi syarat akurasi minimum sistem yaitu di atas 85%.

4.4.2.2 Implementasi Teknologi *Optical Character Recognition* (OCR)

Teknologi *Optical Character Recognition* atau OCR yang diimplementasikan untuk mengekstrak teks dari dokumen atau gambar digunakan pustaka EasyOCR yang merupakan pustaka python untuk melakukan OCR yaitu proses mengenali dan mengekstrak teks dari gambar atau dokumen yang dipindai. Ketika penggunaan pustaka ini, ditemukan bahwa waktu yang dibutuhkan untuk melakukan ekstraksi teks adalah sekitar kurang lebih 1 menit. untuk mengoptimalkan penggunaan pustaka ini maka dilakukan pengoptimalan dengan melakukan *preprocessing* pada gambar yang diinputkan oleh pengguna. Alur implementasi teknologi OCR ditampilkan pada gambar 4.15.

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



Gambar 4.15 Alur Implementasi Teknologi OCR

Gambar 4.15 menjelaskan alur proses dalam pengolahan dokumen menggunakan teknologi OCR. Proses ini diawali dengan input dokumen, yang berupa file dokumen gambar atau hasil konversi halaman PDF ke dalam format gambar. Kemudian, dilakukan tahap *preprocessing* gambar yang terdiri dari dua komponen utama, yakni:

1. *Resize Image*

Gambar akan disesuaikan ukurannya berdasarkan nilai batas maksimal area piksel. Hal ini bertujuan untuk mengurangi waktu proses OCR dan menghindari beban komputasi yang tinggi pada gambar beresolusi besar. Proses ini dilakukan menggunakan metode interpolasi *Image.Resampling.LANCZOS*.

2. *Grayscale Conversion*

Setelah gambar di-*resize*, gambar dikonversi menjadi *grayscale* untuk menghilangkan informasi warna yang tidak diperlukan. Konversi ini juga bertujuan untuk menyederhanakan data input sehingga proses ekstraksi karakter menjadi lebih efisien dan akurat.

Setelah *preprocessing* selesai, gambar kemudian diproses menggunakan EasyOCR yang akan membaca isi teks dari gambar yang telah diproses dan mengembalikan hasil ekstraksi dalam bentuk *string*.

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

4.4.2.3 Implementasi Metode *Keyword Extraction*

Setelah teks berhasil diekstrak, langkah selanjutnya adalah melakukan ekstraksi kata kunci (*keyword extraction*) untuk memperoleh informasi penting dari isi dokumen. Proses ini bertujuan untuk mengidentifikasi informasi yang relevan, seperti keterampilan (*skills*) yang tercantum dalam *curriculum vitae* (CV). Informasi yang diperoleh dari tahap ini akan digunakan sebagai konteks tambahan oleh sistem untuk menghasilkan rekomendasi pekerjaan melalui model LLM. Untuk menentukan metode ekstraksi kata kunci yang paling sesuai, dilakukan eksperimen dan evaluasi terhadap beberapa pendekatan ekstraksi, sebagaimana ditunjukkan pada Tabel 4.10.

Tabel 4.10 Metode Ekstraksi *Keyword*

Kode Metode	Metode Ekstraksi Keyword
EK-1	KeyBERT + Model <i>Sentence Transformer</i>
EK-2	<i>Yet Another Keyword Extractor</i> (YAKE)
EK-3	<i>Named Entity Recognition</i> (NER)

Berdasarkan tabel 4.10 dilakukan percobaan dengan mencoba tiga metode yakni pustaka KeyBERT yang dikombinasikan dengan model *sentence transformer*, metode *Yet Another Keyword Extractor* atau YAKE, dan metode *Named Entity Recognition* atau NER dengan menggunakan model spaCy. Pada metode KeyBERT, digunakan model semantik distilase-base-multilingual-cased-v2 sebagai dasar perhitungan kemiripan semantik antar kalimat dan telah mendukung multilingual. Sementara itu, untuk metode NER dengan spaCy, digunakan model bahasa berukuran besar, yaitu *en_core_web_lg*, yang mampu mengenali entitas-entitas penting dalam teks.

Percobaan dilakukan dengan menggunakan satu dokumen CV yang sama sebagai sampel. Tujuan dari percobaan ini adalah untuk menentukan metode ekstraksi kata kunci yang paling relevan terhadap isi dokumen. Evaluasi dilakukan dengan mempertimbangkan tingkat relevansi kata kunci yang dihasilkan serta kemampuannya dalam menangkap keterampilan utama yang diperlukan untuk

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

mendukung proses rekomendasi pekerjaan. Hasil dari percobaan berbagai metode ekstraksi kata kunci tersebut disajikan pada Tabel 4.11.

Tabel 4.11 Hasil Percobaan Metode Ekstraksi *Keyword*

Kode Metode	Hasil Ekstraksi <i>Keyword</i>
EK-1	['skills google', 'mobile development', 'website mobile', 'mobile application', 'google tokopedia', 'maulana south', 'computer engineering', 'google', 'muhamad fazril', 'mobile']
EK-2	['politeknik negeri', 'negeri jakarta', 'muhamad fazril', 'manajemen waktu', 'teknik komputer', 'aplikasi mobile', 'komputer politeknik', 'fazril maulana', 'mfazrilmaulanagmailcom muhamad', 'fazril linkedin']
EK-3	[('MUHAMAD FAZIL MAULANA', 'ORG'), ('pengembangan aplikasi mobile', 'ORG'), ('dalam basis data', 'ORG'), ('dapat bekerja dengan', 'ORG'), ('serta beberapa sertifikat pelatihan', 'ORG'), ('BANGKIT ACADEMY', 'ORG'), ('GOOGLE', 'ORG'), ('TOKOPEDIA', 'ORG'), ('GOJEK &', 'ORG'), ('Mobile Development', 'ORG'), ('bersama tim', 'ORG'), ('SMK GRAFIKA', 'ORG'), ('Teknik Komputer & Jaringan', 'ORG'), ('Berperan', 'ORG'), ('Mengubah', 'ORG'), ('UI', 'ORG'), ('pada', 'ORG'), ('untuk fungsionalitas', 'ORG'), ('Berkolaborasi', 'ORG'), ('HIMATIK', 'ORG'), ('Bertanggung', 'ORG'), ('Instagram', 'ORG'), ('lainnya', 'ORG'), ('Framework and Library: Bootstrap', 'ORG'), ('Laravel 10/11', 'ORG'), ('CSS', 'ORG')]

Hasil percobaan metode ekstraksi kata kunci yang ditunjukkan pada Tabel 4.11 menunjukkan bahwa penggunaan pustaka KeyBERT yang dikombinasikan dengan model *sentence transformer* distiluse-base-multilingual-cased-v2 memberikan hasil yang lebih baik dibandingkan metode lainnya. Hal ini karena model tersebut mampu menghasilkan kata kunci yang lebih relevan dengan keterampilan yang tercantum dalam dokumen CV. Meskipun demikian, model ini masih menghasilkan beberapa kata kunci yang tidak berkaitan langsung dengan keterampilan. Untuk menemukan hasil yang lebih baik, dilakukan percobaan lanjutan dengan menggunakan beberapa model *sentence transformer* berbasis *semantic similarity*. Model *sentence transformer* yang diuji coba ditunjukkan pada tabel 4.12.

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Tabel 4.12 Model *Sentence Transformer*

Kode Model	Model <i>Sentence Transformer</i>
ST-1	distiluse-base-multilingual-cased-v2
ST-2	paraphrase-multilingual-MiniLM-L12-v2
ST-3	paraphrase-multilingual-mpnet-base-v2

Percobaan lanjutan dilakukan untuk mengevaluasi performa model ekstraksi kata kunci dengan mengganti model *sentence transformer* yang digunakan oleh KeyBERT. Model yang diuji ditunjukkan pada tabel 4.12, yaitu distiluse-base-multilingual-cased-v2, paraphrase-multilingual-mpnet-base-v2 dan paraphrase-multilingual-MiniLM-L12-v2. Dalam proses ekstraksinya, pustaka KeyBERT dirancang untuk mengambil maksimal 10 kata kunci terbaik dari teks hasil ekstraksi, yang terdiri dari satu hingga dua kata dalam setiap frasa. Hasil dari percobaan model *sentence transformer* ditampilkan pada tabel 4.13.

Tabel 4.13 Hasil Percobaan Model *Sentence Transformer*

Kode Model	Hasil Ekstraksi
ST-1	'skills google', 'mobile development', 'website mobile', 'mobile application', 'google tokopedia', 'maulana south', 'computer engineering', 'google', 'muhamad fazril', 'mobile'
ST-2	'mobile development', 'mobile application', 'skills google', 'database developing', 'developing applications', 'data management', 'website mobile', 'application data', 'android applications', 'studying programming'
ST-3	'mobile development', 'application data', 'studying programming', 'mobile application', 'data management', 'developing applications', 'developing implementing', 'mobile', 'mfazrimaulanagmailcom muhamad', 'mfazrimaulanagmailcom'

Hasil percobaan model *sentence transformer* yang ditunjukkan pada Tabel 4.13 menunjukkan bahwa model paraphrase-multilingual-MiniLM-L12-v2 yang digunakan bersama dengan pustaka KeyBERT memberikan hasil yang paling baik

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

dibandingkan model lainnya. Berdasarkan temuan tersebut, paraphrase-multilingual-MiniLM-L12-v2 dipilih sebagai model *sentence transformer* yang digunakan bersama pustaka KeyBERT, karena memberikan hasil paling baik dalam konteks penelitian ini.

4.4.2.4 Perancangan *Prompt Engineering* untuk model *LLaMA*

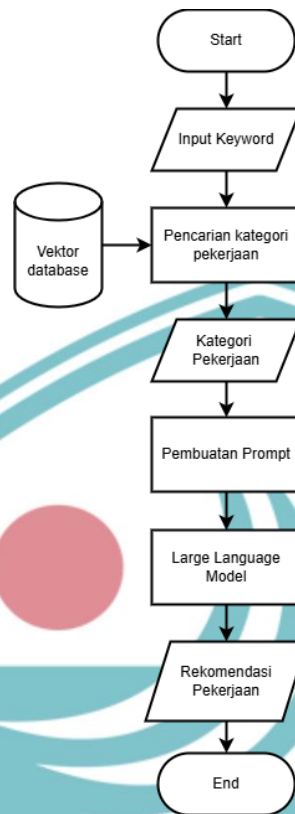
Pada penelitian ini, model *Large Language Model* (LLM) digunakan untuk menghasilkan rekomendasi pekerjaan berdasarkan informasi yang diperoleh dari dokumen CV. Model LLM yang digunakan adalah *LLaMA* 3.3 dengan ukuran 70B parameter, yang diakses melalui layanan OpenRouter untuk memperoleh API-nya. Untuk memastikan bahwa model LLM memberikan respons yang spesifik dan relevan terhadap permintaan rekomendasi pekerjaan, dilakukan teknik *prompt engineering*. Teknik ini bertujuan untuk merancang *prompt* yang tepat agar model dapat memberikan respons rekomendasi pekerjaan ketika pengguna memasukkan teks atau dokumen CV ke dalam sistem. *Prompt* yang dibuat untuk menghasilkan rekomendasi pekerjaan berdasarkan input pengguna dapat dilihat pada Lampiran 1.

Berdasarkan *prompt* pada lampiran 1 diperlukan 2 variabel yang perlu dimasukkan ke dalam *prompt* yakni *skills*, dan *categories*. Variabel *skills* ini didapatkan dari hasil ekstraksi *keyword* pada teks, dan variabel *categories* didapatkan dari pencarian data kategori pekerjaan pada *vector database* yang berisi data. Alur bagaimana LLM dapat memberikan respons berupa rekomendasi pekerjaan ditampilkan pada gambar 4.16.



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



Gambar 4.16 Alur *Generate Respons* LLM

Gambar 4.16 menunjukkan alur proses bagaimana *Large Language Model* (LLM) menghasilkan rekomendasi pekerjaan berdasarkan *keyword* hasil ekstraksi dari teks. Proses ini dimulai dengan *keyword* yang diinputkan ke dalam suatu fungsi untuk melakukan pencarian pada *vector database* yang berisi data keterampilan dan kategori pekerjaan. Jika terdapat keterampilan dalam basis data yang memiliki kemiripan dengan *keyword* yang dicari, maka kategori pekerjaan yang sesuai akan dihasilkan sebagai output. Selanjutnya, dilakukan proses pembuatan *prompt* dengan menggunakan konteks berupa *keyword* keterampilan dan kategori pekerjaan yang telah diperoleh. *Prompt* yang telah terbentuk kemudian diberikan kepada *large language model* untuk diproses, sehingga model dapat menghasilkan rekomendasi pekerjaan yang relevan. Contoh hasil dari *prompt* LLM pada lampiran 1 dapat dilihat pada tabel 4.16.

Tabel 4.16 Contoh Hasil *Prompt* LLM

Hasil Ekstraksi Keyword	“Melukis”
Hasil Pencarian Kategori Pekerjaan	“Seni dan Desain”
Hasil Prompt	Halo, <i>Job Seeker</i>! Berdasarkan keahlian yang Anda miliki, kami telah menemukan beberapa pekerjaan yang mungkin cocok untuk Anda. Kategori Pekerjaan: Seni dan Desain Kami merekomendasikan Anda untuk mempertimbangkan pekerjaan berikut: Desainer Grafis Deskripsi: Membuat desain visual untuk berbagai keperluan seperti iklan, brosur, dan media sosial. Alasan Cocok: Keahlian Anda dalam melukis menunjukkan kemampuan dalam mengolah warna dan bentuk, yang sangat penting dalam desain grafis.

4.4.3 Model Evaluation and Refinement

Evaluasi dilakukan dengan menganalisis hasil eksperimen yang diperoleh selama tahap pelatihan dan pengujian model. Tujuan dari evaluasi ini adalah untuk memastikan bahwa model yang digunakan benar-benar mampu memberikan hasil yang akurat, efisien, dan relevan sesuai dengan kebutuhan sistem. Selain itu, tahap ini juga digunakan untuk mengidentifikasi potensi kelemahan atau kekurangan dari model yang ada, sehingga dapat dilakukan perbaikan atau penyempurnaan.

4.4.3.1 Evaluasi Model Klasifikasi Dokumen CV dan Non-CV

Pada tahap pengembangan model, dilakukan proses *hyperparameter tuning* untuk mengoptimalkan performa model klasifikasi dokumen CV dan Non-CV. Untuk mengetahui sejauh mana peningkatan performa setelah *tuning* dilakukan, evaluasi model dilakukan dengan menggunakan *confusion matrix*, yang mencakup pengukuran terhadap metrik akurasi (*accuracy*), presisi (*precision*), *recall*, dan *F1-*

score. Parameter yang digunakan sebelum proses *hyperparameter tuning*, serta parameter yang dipilih setelah proses *tuning*, dapat dilihat pada Tabel 4.15.

Tabel 4.15 Parameter Model Klasifikasi Dokumen CV dan Non-CV

Parameter Model	Sebelum <i>Tuning</i>	Setelah <i>Tuning</i>
Jumlah Unit pada <i>LSTM Layer 1</i>	64	64
Jumlah Unit pada <i>LSTM Layer 2</i>	32	32
Tingkat <i>Dropout</i>	0.3	0.7
<i>Learning Rate</i>	0.001	0.001
<i>Batch Size</i>	32	32
<i>Optimizer</i>	adam	adamax

Hasil evaluasi sebelum dan sesudah proses *hyperparameter tuning* dapat dilihat pada Tabel 4.16.

Tabel 4.16 Hasil Evaluasi *Confusion Matrix*

Variabel Evaluasi	Sebelum tuning	Setelah tuning
<i>Accuracy</i>	60 %	98,75 %
<i>Precision</i>	1,0	0,97
<i>Recall</i>	0,27	1,0
<i>F1-score</i>	0,4286	0,988

Berdasarkan hasil evaluasi yang ditunjukkan pada Tabel 4.16, dapat dilihat bahwa proses *hyperparameter tuning* memberikan dampak signifikan terhadap performa model klasifikasi dokumen CV.

Sebelum dilakukan *tuning*, model memang menunjukkan nilai *precision* yang sangat tinggi (1.0), namun nilai *recall* yang rendah (0.27) menunjukkan bahwa model gagal mendeteksi sebagian besar dokumen CV secara benar. Hal ini menyebabkan *F1-score* yang merupakan rata-rata harmonis dari *precision* dan

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

recall juga rendah, yaitu sebesar 0.4286. Nilai *accuracy* pun masih berada di angka 60 %, yang menandakan model belum optimal secara keseluruhan.

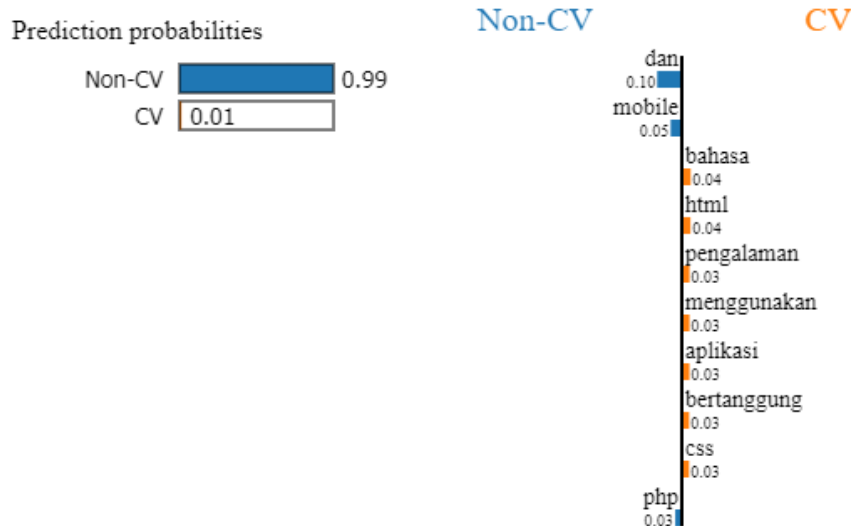
Setelah dilakukan proses *tuning* terhadap beberapa parameter seperti jumlah unit LSTM, *dropout rate*, *batch size*, *learning rate*, dan jenis *optimizer*, performa model meningkat secara drastis. Hasil evaluasi menunjukkan *accuracy* mencapai 0.9875, yang berarti model mampu mengklasifikasikan dokumen CV dan Non-CV dengan sangat baik. Selain itu, nilai *recall* meningkat menjadi 1.0, menunjukkan bahwa semua dokumen CV berhasil terklasifikasi dengan benar, tanpa ada yang terlewat. *Precision* juga tetap tinggi di angka 0.97, dan *f1-score* meningkat tajam menjadi 0.988, menunjukkan keseimbangan antara *precision* dan *recall* yang sangat baik.

Hasil ini mengindikasikan bahwa model LSTM yang telah di-*tuning* sangat efektif dalam membedakan antara dokumen CV dan Non-CV, serta layak digunakan dalam sistem sebagai tahap validasi dokumen sebelum proses lanjutan seperti ekstraksi teks atau rekomendasi pekerjaan dilakukan. Oleh karena itu, parameter yang digunakan pada model adalah jumlah unit pada LSTM layer pertama sebanyak 64 unit, layer kedua sebanyak 32 unit, nilai *dropout* sebesar 0,7, *batch size* 32, *learning rate* 0,001, dan optimizer yang digunakan adalah Adamax.

Selain itu, untuk melihat transparansi dari hasil klasifikasi model klasifikasi model akan diterapkan teknik *eXplainable AI* untuk melihat kata yang memiliki kontribusi besar dalam menghasilkan hasil klasifikasi. Metode *eXplainable AI* yang diterapkan adalah metode SHAP (*SHapley Additive exPlanations*) dan LIME (*Local Interpretable Model-Agnostic Explanations*).

a. Implementasi metode LIME

Metode LIME digunakan untuk memahami bagaimana model memberikan prediksi terhadap suatu input dengan mempelajari perubahan output berdasarkan variasi input secara lokal. Visualisasi dari penerapan metode LIME dapat dilihat pada Gambar 4.17, yang menunjukkan kata-kata mana yang dianggap penting oleh model saat mengambil keputusan.



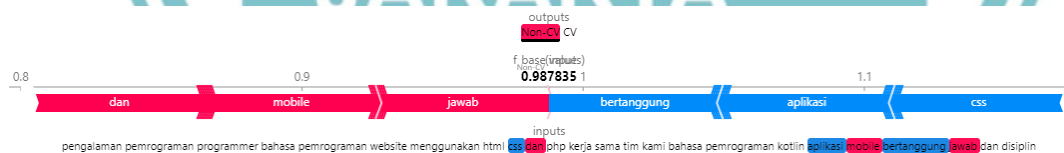
Text with highlighted words

pengalaman pemrograman programmer bahasa pemrograman website menggunakan html css dan php kerja sama tim kami bahasa pemrograman kotlin aplikasi mobile bertanggung jawab dan disiplin

Gambar 4.17 Implementasi Metode LIME

b. Implementasi metode SHAP

Metode SHAP menghitung kontribusi setiap kata terhadap prediksi model dengan pendekatan teori permainan. Hasil dari metode SHAP memberikan pemahaman global maupun lokal terhadap perilaku model. Visualisasi penerapan metode ini disajikan pada Gambar 4.18, yang memperlihatkan seberapa besar kontribusi masing-masing kata terhadap klasifikasi akhir.



Gambar 4.18 Implementasi Metode SHAP

4.4.3.2 Evaluasi Implementasi OCR

Hasil perbandingan implementasi EasyOCR dengan *preprocessing* gambar dengan tanpa *preprocessing* gambar dapat dilihat pada tabel 4.17.

Tabel 4.17 Perbandingan Hasil Implementasi EasyOCR

Implementasi	Waktu Proses	Kualitas Hasil OCR
Tanpa <i>Preprocessing</i>	1 menit 6 detik	Baik
Dengan <i>Preprocessing</i>	21 detik	Cukup Baik

Berdasarkan hasil pada Tabel 4.19, penerapan *preprocessing* seperti *resize* dan konversi ke *grayscale* berhasil mempercepat proses ekstraksi OCR secara signifikan dari 66 detik menjadi 21 detik. Meskipun kualitas teks hasil ekstraksi sedikit menurun, tetapi hasil tetap terbaca. Hal ini menunjukkan bahwa *preprocessing* dapat mengoptimalkan kinerja OCR, khususnya pada dokumen dengan ukuran besar atau kondisi pencahayaan yang kurang baik.

4.4.3.3 Evaluasi Model *Large Language Model*

Model *large language model* (LLM) seperti *LLaMA* memiliki kemampuan menghasilkan teks yang sangat alami, namun tidak selalu akurat. Salah satu tantangan utama dalam penggunaan LLM adalah halusinasi AI, yaitu kondisi ketika model menghasilkan informasi yang tidak benar atau tidak berdasar, tetapi disampaikan dengan keyakinan tinggi seolah-olah benar. Halusinasi ini terjadi karena model LLM berusaha mengisi kekosongan informasi berdasarkan pola dari data pelatihan, tanpa kemampuan untuk melakukan verifikasi fakta secara aktual.

Untuk mengevaluasi potensi halusinasi pada model *LLaMA* yang digunakan dalam penelitian ini, dilakukan pengujian dengan *benchmark dataset* bernama TruthfulQA dengan menggunakan teknik *manual annotation*. Dataset ini dirancang khusus untuk menguji keakuratan dan kejujuran jawaban yang dihasilkan oleh model bahasa besar. Sebanyak 80 sampel data diambil dari dataset TruthfulQA dan digunakan untuk mendeteksi kemungkinan terjadinya halusinasi pada model *LLaMA* yang digunakan. Hasil dari deteksi halusinasi AI ini menunjukkan dari 80 sampel data pertanyaan dengan jawaban salah dan benar, model *LLaMA* berhasil menjawab pertanyaan dengan benar sebanyak 50 soal, dan salah menjawab sebanyak 30 soal.

Dari hasil evaluasi menggunakan dataset TruthfulQA, dapat diketahui bahwa model *LLaMA* memiliki kemungkinan untuk menghasilkan jawaban yang salah atau tidak faktual. Hal ini menunjukkan bahwa fenomena halusinasi AI masih dapat terjadi pada model, terutama ketika konteks yang diberikan tidak cukup jelas atau spesifik. Untuk memitigasi potensi halusinasi pada output model, dilakukan penyesuaian terhadap *system prompt* dan parameter *temperature* dalam proses generasi bahasa. Parameter *temperature* berfungsi untuk mengontrol tingkat keacakan dan kreativitas model saat menghasilkan teks.

Adapun *system prompt* yang digunakan dalam pengaturan permintaan ke model LLM adalah sebagai berikut:

"Kamu adalah asisten karir. Jika tidak yakin 100%, katakan "Saya tidak begitu yakin dengan rekomendasi saya". Jawab pertanyaan dengan akurat. Jika tidak tahu, katakan tidak tahu dan jangan mengarang data."

Kemudian, untuk mengidentifikasi nilai *temperature* yang paling sesuai, dilakukan eksperimen dengan beberapa nilai parameter: 0, 0.1, 0.5, dan 1. Evaluasi dilakukan dengan membandingkan respons yang dihasilkan model terhadap pertanyaan yang sama, untuk menilai sejauh mana *temperature* memengaruhi kualitas, dan stabilitas jawaban. Hasil dari percobaan nilai *temperature* dapat dilihat pada tabel 4.18.

Tabel 4.18 Percobaan Nilai *Temperature*

Pertanyaan	Apa itu pekerjaan ?
Nilai <i>Temperature</i>	Respons Jawaban LLM
0	Pekerjaan adalah aktivitas yang dilakukan oleh seseorang untuk menghasilkan uang atau pendapatan, serta memenuhi kebutuhan hidupnya.
0.1	Pekerjaan adalah aktivitas yang dilakukan oleh seseorang untuk menghasilkan uang atau pendapatan, serta memenuhi kebutuhan hidupnya.
0.5	Pekerjaan adalah aktivitas yang dilakukan oleh seseorang untuk memperoleh penghasilan atau imbalan dalam bentuk uang atau barang.

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

1	Pekerjaan adalah aktivitas atau kegiatan yang dilakukan oleh seseorang untuk menghasilkan uang, barang, atau jasa yang berguna bagi masyarakat. Pekerjaan dapat berupa pekerjaan formal, seperti bekerja di kantor atau pabrik, atau pekerjaan informal, seperti berjualan atau berwirausaha.
---	---

Berdasarkan hasil pada tabel 4.18, diketahui bahwa pada nilai *temperature* 0 dan 0.1, model memberikan jawaban yang identik, yaitu ringkas, dan langsung. Respons pada nilai ini cenderung minim variasi. Hal ini menunjukkan bahwa penggunaan *temperature* rendah sangat cocok untuk konteks yang membutuhkan jawaban yang tetap. Pada nilai *temperature* 0.5, model mulai menunjukkan variasi dalam pemilihan kata dan struktur kalimat, misalnya, terdapat perbedaan dalam istilah yang digunakan seperti “penghasilan atau imbalan dalam bentuk uang atau barang”. Sementara itu, pada nilai *temperature* 1, respons yang dihasilkan menjadi lebih panjang, deskriptif, dan mencakup penjelasan tambahan seperti contoh pekerjaan formal dan informal. Meskipun jawaban terlihat lebih lengkap, penggunaan *temperature* tinggi juga berpotensi menimbulkan halusinasi atau penyimpangan dari fakta jika tidak diawasi dengan baik.

Berdasarkan hasil eksperimen tersebut, diputuskan bahwa nilai *temperature* sebesar 0.1 adalah yang paling tepat untuk digunakan dalam sistem ini. Nilai tersebut memberikan respons yang jelas, fokus, dan minim variasi yang tidak perlu, sehingga sangat sesuai untuk menghasilkan rekomendasi pekerjaan yang akurat.

4.5 System Implementation

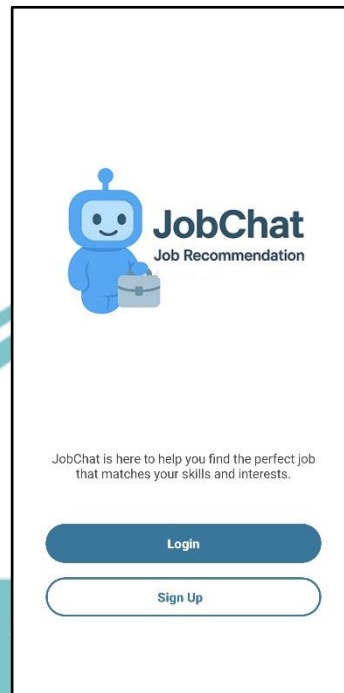
Kode program hasil implementasi perancangan aplikasi *chatbot* rekomendasi pekerjaan dapat diakses melalui repositori github yang dipublikasi. Repositori tersebut dapat diakses melalui tautan: <https://github.com/GerrysatriaH/JobChatApp>



4.5.1 Halaman *Landing*

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



Gambar 4.19 Halaman *Landing*

Halaman *landing* yang ditampilkan pada gambar 4.19 merupakan tampilan yang pertama kali muncul saat pengguna membuka aplikasi "JobChat". Pada halaman ini, ditampilkan beberapa elemen utama yang bertujuan untuk memperkenalkan aplikasi kepada pengguna. Terdapat logo aplikasi berupa robot berwarna biru yang memegang koper, nama aplikasi "JobChat", serta slogan "*Job Recommendation*" yang menjelaskan fungsi utama aplikasi. Di bawahnya, terdapat kalimat pengantar singkat yang menjelaskan tujuan aplikasi untuk membantu pengguna menemukan pekerjaan yang sesuai dengan keahlian dan minat mereka. Selain itu, tersedia dua tombol utama yakni, tombol "*Login*" untuk pengguna yang sudah memiliki akun dan tombol "*Sign Up*" bagi pengguna baru untuk mendaftarkan diri ke dalam sistem.



4.5.2 Halaman *Sign Up*

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Gambar 4.20 Halaman *Register*

Halaman *register* atau *sign up* merupakan halaman yang akan diakses pengguna ketika menekan tombol "*Sign Up*". Halaman ini, seperti yang ditampilkan pada gambar 4.20, berfungsi untuk membuat akun baru bagi pengguna yang belum terdaftar. Terdapat dua kolom input utama, yakni "*Email Address*" untuk pengguna memasukkan alamat email mereka, dan "*Password*" untuk membuat kata sandi. Untuk meningkatkan pengalaman pengguna, pada kolom *password* terdapat ikon mata yang berfungsi untuk menampilkan atau menyembunyikan kata sandi yang diketik. Di bawah kolom input, terdapat tombol "*Sign Up*" untuk melanjutkan proses pendaftaran. Bagi pengguna yang sudah memiliki akun, tersedia juga opsi "*Login*" untuk beralih ke halaman masuk.



4.5.3 Halaman *Login*

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Gambar 4.21 Halaman *Login*

Halaman *login* merupakan halaman yang akan diakses pengguna ketika menekan tombol "*Login*" pada halaman *landing*. Halaman ini, seperti yang ditampilkan pada Gambar gambar 4.21, berfungsi sebagai gerbang masuk bagi pengguna yang telah terdaftar. Terdapat dua kolom input utama, yakni "*Email Address*" untuk pengguna memasukkan alamat email mereka, dan "*Password*" untuk kata sandi. Untuk meningkatkan pengalaman pengguna, pada kolom *password* terdapat ikon mata yang berfungsi untuk menampilkan atau menyembunyikan kata sandi yang diketik. Di bawah kolom input, Di bawah kolom input, terdapat tautan "*Forgot password?*" untuk pengguna yang lupa kata sandi mereka, serta tombol "*Login*" untuk melanjutkan proses masuk. Bagi pengguna yang belum memiliki akun, tersedia juga opsi "*Sign Up*" untuk beralih ke halaman pendaftaran.

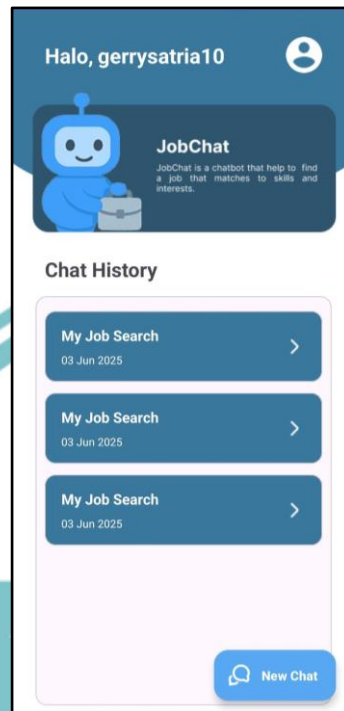


© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

4.5.4 Halaman *Home*

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



Gambar 4.22 Halaman Home

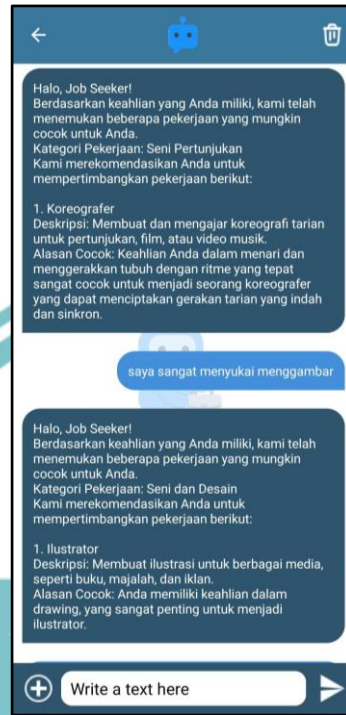
Setelah pengguna berhasil masuk ke dalam sistem, mereka akan dialihkan ke halaman *home*. Gambaran halaman tersebut dapat dilihat di Gambar 4.22. Halaman ini berfungsi sebagai halaman utama, dimana pengguna dapat melihat daftar riwayat percakapan sebelumnya pada bagian "*Chat History*". Untuk memulai interaksi baru dengan *chatbot* guna mendapatkan rekomendasi pekerjaan, pengguna dapat menekan tombol "*New Chat*" yang terletak di bagian bawah layar.

4.5.5 Halaman Chat

86

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



Gambar 4.23 Halaman Chat

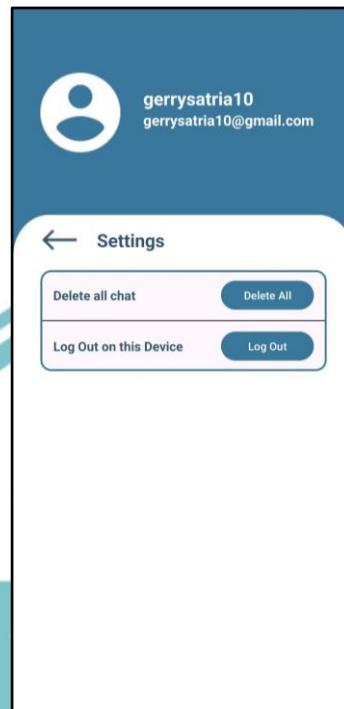
Setelah pengguna menekan tombol "New Chat", mereka akan diarahkan ke halaman percakapan dengan *chatbot*. Gambaran halaman ini dapat dilihat pada Gambar 4.23. Halaman ini adalah tempat pengguna berinteraksi langsung dengan sistem untuk mendapatkan rekomendasi pekerjaan. Terdapat kolom input teks di bagian bawah layar dengan tulisan "Write a text here" dimana pengguna dapat mengetikkan pesan teks atau mengunggah dokumen melalui tombol dengan ikon "+". Area utama layar akan menampilkan riwayat percakapan, di mana setiap balasan dari *chatbot* berisi rekomendasi pekerjaan yang sesuai dengan masukan dari pengguna.



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

4.5.6 Halaman Profil



Gambar 4.24 Halaman Profil

Halaman profil dapat diakses oleh pengguna melalui ikon profil di halaman *home*. gambaran halaman ini dapat dilihat pada gambar 4.24. Halaman ini menampilkan informasi akun pengguna seperti nama pengguna dan alamat email di bagian atas. Di bawahnya, terdapat menu "*Settings*" atau pengaturan yang berisi fungsionalitas untuk mengelola akun dan data. Terdapat dua opsi utama yang tersedia: "*Delete all chat*" yang memungkinkan pengguna untuk menghapus seluruh riwayat percakapan secara permanen, dan "*Log Out on this Device*" untuk keluar dari akun pada perangkat yang sedang digunakan.

4.6 System Testing

Aplikasi yang dikembangkan pada penelitian ini adalah aplikasi *chatbot* untuk memberikan rekomendasi pekerjaan berbasis android. Aplikasi ini diuji untuk mengevaluasi kinerjanya. Pengujian aplikasi ini terdiri dari tiga tipe pengujian, yakni *black box testing*, *system usability scale* (SUS), dan *net promoter score* (NPS). Pengujian secara *black box* akan dilakukan secara mandiri, sedangkan pengujian dengan SUS, dan NPS akan dilakukan oleh pengguna aplikasi.

4.6.1 Black Box Testing

a. Deskripsi Pengujian

Pengujian *black box* merupakan metode pengujian perangkat lunak yang berfokus pada evaluasi fungsionalitas sistem tanpa mengetahui struktur kode internal, logika, atau arsitektur di baliknya. Dalam konteks penelitian ini, pengujian dilakukan secara mandiri dengan memposisikan peneliti sebagai pengguna akhir aplikasi *chatbot*. Setiap fitur, mulai dari pendaftaran akun hingga penerimaan rekomendasi pekerjaan, diuji berdasarkan serangkaian skenario penggunaan yang telah dirancang untuk memastikan setiap fungsi berjalan sesuai dengan hasil yang diharapkan. Rincian lengkap dari skenario pengujian *black box* yang dilakukan disajikan pada Tabel 4.19.

Tabel 4.19 Skenario *Black Box Testing*

Kode	Kasus	Skenario	Harapan
BB-1-1	Pendaftaran Akun	1. Buka halaman pendaftaran (<i>sign up</i>). 2. Masukkan alamat <i>email</i> yang valid dan belum terdaftar. 3. Masukkan <i>password</i> . 4. Klik tombol “ <i>Sign Up</i> ”	Sistem menampilkan pesan sukses dan mengarahkan pengguna ke halaman <i>login</i> .
BB-1-2		1. Buka halaman pendaftaran (<i>sign up</i>). 2. Masukkan alamat <i>email</i> yang sudah terdaftar. 3. Masukkan <i>password</i> . 4. Klik tombol “ <i>Sign Up</i> ”	Sistem menampilkan pesan "Email sudah terdaftar" dan tetap di halaman pendaftaran.
BB-1-3		1. Buka halaman pendaftaran (<i>sign up</i>). 2. Kosongkan <i>email</i> dan <i>password</i> . 3. Klik tombol “ <i>Sign Up</i> ”	Sistem menampilkan pesan mengenai <i>field</i> yang tidak terisi dan tetap di halaman pendaftaran.
BB-2-1	Login	1. Buka halaman <i>login</i> . 2. Masukkan <i>email</i> dan <i>password</i> yang valid (sudah terdaftar). 3. Klik tombol “ <i>Login</i> ”	Sistem menampilkan pesan sukses dan mengarahkan pengguna ke halaman <i>home</i> .

Hak Cipta :

- Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
- Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

BB-2-2		1. Buka halaman <i>login</i>. 2. Masukkan <i>email</i> dan <i>password</i> yang tidak valid (salah/belum terdaftar). 3. Klik tombol "<i>Login</i>"	Sistem menampilkan pesan mengenai informasi data yang tidak valid.
BB-3-1	Rekomendasi dengan input teks	1. <i>Login</i> ke aplikasi. 2. Buka halaman <i>chat</i>. 3. Ketik pesan yang mengandung keterampilan (contoh: Saya suka melukis). 4. Kirim pesan.	Sistem berhasil memproses teks dan <i>chatbot</i> memberikan rekomendasi pekerjaan yang relevan.
BB-3-2		1. <i>Login</i> ke aplikasi. 2. Buka halaman <i>chat</i>. 3. Ketik pesan yang tidak mengandung keterampilan (contoh: halo, apa kabar?). 4. Kirim pesan.	<i>Chatbot</i> memberikan respon "Tidak ada kategori pekerjaan yang relevan"
BB-4-1	Rekomendasi dengan dokumen	1. <i>Login</i> ke aplikasi. 2. Buka halaman <i>chat</i>. 3. Unggah dokumen berformat .pdf yang merupakan dokumen CV. 4. Kirim pesan.	Sistem berhasil memproses CV dan <i>chatbot</i> memberikan rekomendasi pekerjaan yang relevan.
BB-4-2		1. <i>Login</i> ke aplikasi. 2. Buka halaman <i>chat</i>. 3. Unggah dokumen berformat .pdf yang bukan merupakan dokumen CV (contoh: surat keterangan sakit). 4. Kirim pesan.	Sistem berhasil memproses dokumen dan <i>chatbot</i> memberikan respon "Dokumen bukan merupakan dokumen CV".
BB-4-3		1. <i>Login</i> ke aplikasi. 2. Buka halaman <i>chat</i>. 3. Unggah dokumen berformat .pdf yang berisikan teks kosong. 4. Kirim pesan.	<i>Chatbot</i> memberikan respon "Dokumen tidak berisikan teks".
BB-4-4		1. <i>Login</i> ke aplikasi. 2. Buka halaman <i>chat</i>.	<i>Chatbot</i> memberikan respon "Format tidak

© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

		3. Unggah dokumen berformat yang tidak didukung (contoh: .docx, .doc, .exe, .mp4, mp3). 4. Kirim pesan.	mendukung".
BB-4-5		1. <i>Login</i> ke aplikasi. 2. Buka halaman <i>chat</i> . 3. Unggah dokumen berformat .pdf yang berisikan lebih dari 2 halaman. 4. Kirim pesan.	<i>Chatbot</i> memberikan respon "Dokumen lebih dari 2 halaman, tidak dapat diproses".
BB-5-1	Menghapus Percakapan	1. <i>Login</i> ke aplikasi. 2. Buka salah satu riwayat percakapan yang pernah dibuat. 3. Klik tombol dengan logo "tempat sampah" di kanan tampilan aplikasi. 4. Konfirmasi penghapusan pada <i>popup</i> yang muncul.	Seluruh percakapan di halaman <i>chat</i> terhapus dan mengarahkan pengguna ke riwayat percakapan.
BB-6-1	Menghapus Riwayat Percakapan	1. <i>Login</i> ke aplikasi. 2. Buka halaman profil. 3. Klik tombol " <i>Delete All Chat</i> ". 4. Konfirmasi penghapusan pada <i>popup</i> yang muncul.	Seluruh riwayat percakapan di halaman <i>home</i> terhapus.
BB-7-1	<i>Logout</i>	1. <i>Login</i> ke aplikasi. 2. Buka halaman profil. 3. Klik tombol " <i>Logout</i> ".	Sistem berhasil mengeluarkan pengguna dan mengarahkan kembali ke halaman <i>login</i> .

b. Data Hasil Pengujian

Pengujian *black box* dalam penelitian ini dilakukan melalui metode pengujian mandiri, di mana validasi fungsionalitas sistem dilaksanakan dengan cara membandingkan hasil aktual dari setiap pengujian dengan hasil yang telah diharapkan sebelumnya. Sebuah skenario pengujian dinyatakan berhasil atau layak

digunakan apabila hasil observasi sepenuhnya sesuai dengan ekspektasi. Sebaliknya, jika terdapat perbedaan antara hasil aktual dan yang diharapkan, maka skenario tersebut dianggap gagal dan mengindikasikan perlunya perbaikan pada fitur terkait. Data hasil pengujian dengan menggunakan *black box testing* ditampilkan pada tabel 4.20.

Tabel 4.20 Hasil Pengujian *Black Box*

Kode	Harapan	Pengamatan	Hasil
BB-1-1	Sistem menampilkan pesan sukses dan mengarahkan pengguna ke halaman <i>login</i> .	Sistem menampilkan pesan sukses dan mengarahkan pengguna ke halaman <i>login</i> .	Sesuai
BB-1-2	Sistem menampilkan pesan "Email sudah terdaftar" dan tetap di halaman pendaftaran.	Sistem menampilkan pesan "Email sudah terdaftar" dan tetap di halaman pendaftaran.	Sesuai
BB-1-3	Sistem menampilkan pesan mengenai <i>field</i> yang tidak terisi dan tetap di halaman pendaftaran.	Sistem menampilkan pesan mengenai <i>field</i> yang tidak terisi dan tetap di halaman pendaftaran.	Sesuai
BB-2-1	Sistem menampilkan pesan sukses dan mengarahkan pengguna ke halaman <i>home</i> .	Sistem menampilkan pesan sukses dan mengarahkan pengguna ke halaman <i>home</i> .	Sesuai
BB-2-2	Sistem menampilkan pesan mengenai informasi data yang tidak valid.	Sistem menampilkan pesan mengenai informasi data yang tidak valid.	Sesuai
BB-3-1	Sistem berhasil memproses teks dan <i>chatbot</i> memberikan rekomendasi pekerjaan yang relevan.	Sistem berhasil memproses teks dan <i>chatbot</i> memberikan rekomendasi pekerjaan yang relevan.	Sesuai
BB-3-2	<i>Chatbot</i> memberikan respon "Tidak ada kategori pekerjaan yang relevan"	<i>Chatbot</i> memberikan respon "Tidak ada kategori pekerjaan yang relevan"	Sesuai

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

BB-4-1	Sistem berhasil memproses CV dan <i>chatbot</i> memberikan rekomendasi pekerjaan yang relevan.	Sistem berhasil memproses CV dan <i>chatbot</i> memberikan rekomendasi pekerjaan yang relevan.	Sesuai
BB-4-2	Sistem berhasil memproses dokumen dan <i>chatbot</i> memberikan respon “Dokumen bukan merupakan dokumen CV”.	Sistem berhasil memproses dokumen dan <i>chatbot</i> memberikan respon “Dokumen bukan merupakan dokumen CV”.	Sesuai
BB-4-3	<i>Chatbot</i> memberikan respon "Dokumen tidak berisikan teks".	<i>Chatbot</i> memberikan respon "Dokumen tidak berisikan teks".	Sesuai
BB-4-4	<i>Chatbot</i> memberikan respon "Format tidak mendukung".	<i>Chatbot</i> memberikan respon "Format tidak mendukung".	Sesuai
BB-4-5	<i>Chatbot</i> memberikan respon "Dokumen lebih dari 2 halaman, tidak dapat diproses".	<i>Chatbot</i> memberikan respon "Dokumen lebih dari 2 halaman, tidak dapat diproses".	Sesuai
BB-5-1	Seluruh percakapan di halaman <i>chat</i> terhapus dan mengarahkan pengguna ke riwayat percakapan.	Seluruh percakapan di halaman <i>chat</i> terhapus dan mengarahkan pengguna ke riwayat percakapan.	Sesuai
BB-6-1	Seluruh riwayat percakapan di halaman <i>home</i> terhapus.	Seluruh riwayat percakapan di halaman <i>home</i> terhapus.	Sesuai
BB-7-1	Sistem berhasil mengeluarkan pengguna dan mengarahkan kembali ke halaman <i>login</i> .	Sistem berhasil mengeluarkan pengguna dan mengarahkan kembali ke halaman <i>login</i> .	Sesuai

Hasil pengujian *black box* yang ditampilkan pada tabel 4.19 menunjukkan bahwa seluruh skenario pengujian yang telah dirancang berhasil dieksekusi dan memberikan hasil sesuai dengan ekspektasi. Berdasarkan temuan ini, dapat disimpulkan bahwa fungsionalitas utama aplikasi telah berjalan dengan baik dan

mencapai tingkat kelayakan 100% dalam lingkup pengujian yang dilakukan. Meskipun demikian, perlu digaris bawahi bahwa validitas pengujian *black box* ini hanya terbatas pada cakupan skenario yang telah didefinisikan, dan tidak merepresentasikan keseluruhan potensi interaksi pengguna yang dapat terjadi dalam penggunaan di dunia nyata.

4.6.2 System Usability Scale (SUS)

a. Deskripsi Pengujian

Untuk mengukur tingkat kebergunaan (*usability*) dari aplikasi JobChat secara kuantitatif, penelitian ini menggunakan metode *system usability scale* (SUS). Pengujian ini melibatkan pengguna akhir yang diminta untuk berinteraksi langsung dengan aplikasi JobChat. Responden akan mencoba berbagai fungsionalitas utama, seperti melakukan pendaftaran, mencari rekomendasi pekerjaan dengan input teks dan dokumen, serta mengelola akun. Setelah sesi penggunaan selesai, setiap pengguna akan diminta untuk mengisi kuesioner SUS yang berisi 10 pernyataan. Daftar pertanyaan SUS tersebut dapat dilihat pada Tabel 4.21.

Tabel 4.21 Pertanyaan SUS

Kode	Pertanyaan
SUS-1	Saya pikir saya akan sering menggunakan aplikasi ini.
SUS-2	Saya merasa aplikasi ini terlalu rumit digunakan.
SUS-3	Saya merasa aplikasi ini mudah digunakan.
SUS-4	Saya pikir saya membutuhkan bantuan orang lain atau teknisi untuk dapat menggunakan aplikasi ini.
SUS-5	Saya merasa berbagai fungsi atau fitur pada aplikasi ini dapat berjalan dengan semestinya.
SUS-6	Saya merasa ada banyak hal yang tidak konsisten pada aplikasi ini.
SUS-7	Saya merasa kebanyakan orang akan mudah memahami aplikasi ini dengan sangat cepat.
SUS-8	Saya merasa aplikasi ini sangat membingungkan untuk digunakan.
SUS-9	Saya merasa tidak ada hambatan dalam menggunakan aplikasi ini.

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

SUS-10	Saya perlu beradaptasi terlebih dahulu sebelum bisa menggunakan sistem ini.
--------	---

Pertanyaan SUS yang ditampilkan pada tabel 4.21 menggunakan format jawaban yang menggunakan skala likert 5 poin yakni nilai 1 sampai 5. Nilai 1 menyatakan “Sangat tidak setuju”, sedangkan nilai 5 menyatakan “Sangat setuju”. Skor ini memberikan gambaran kuantitatif mengenai tingkat kebergunaan aplikasi, di mana skor yang lebih tinggi menunjukkan persepsi usabilitas yang lebih baik.

b. Data Hasil Pengujian

Pengumpulan data untuk pengujian SUS ini dilakukan dengan menyebarkan kuesioner yang dibuat menggunakan Google Form. Kuesioner tersebut diisi oleh total 19 responden. Responden ini merupakan pengguna akhir yang telah mencoba dan berinteraksi langsung dengan fungsionalitas aplikasi JobChat. Proses pengumpulan data dilakukan melalui dua cara, yaitu secara daring melalui penyebaran tautan Google Form dan secara luring dengan menemui beberapa pengguna secara langsung untuk memandu pengisian kuesioner. Hasil SUS dari kuesioner uji coba aplikasi chatbot dapat dilihat pada Tabel 4.22 dan 4.23.

Tabel 4.22 Hasil SUS Bagian 1

Responden	SUS-1	SUS-2	SUS-3	SUS-4	SUS-5
Responden-1	5	1	5	1	5
Responden-2	4	1	5	1	5
Responden-3	5	1	5	3	4
Responden-4	5	1	4	5	5
Responden-5	3	3	3	3	3
Responden-6	5	1	5	1	5
Responden-7	4	2	4	3	4
Responden-8	4	2	4	5	4
Responden-9	3	2	3	4	3

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Responden	SUS-1	SUS-2	SUS-3	SUS-4	SUS-5
Responden-10	3	1	5	1	5
Responden-11	3	1	5	5	5
Responden-12	4	2	4	2	4
Responden-13	3	3	3	3	3
Responden-14	4	2	5	3	5
Responden-15	4	1	5	1	4
Responden-16	3	1	5	1	5
Responden-17	5	1	5	1	5
Responden-18	4	1	5	1	5
Responden-19	4	1	5	1	5

Tabel 4.23 Hasil SUS Bagian 2

Responden	SUS-6	SUS-7	SUS-8	SUS-9	SUS-10
Responden-1	1	5	1	5	2
Responden-2	1	5	1	5	1
Responden-3	1	5	1	2	5
Responden-4	3	5	1	5	5
Responden-5	3	3	3	3	3
Responden-6	2	5	1	5	4
Responden-7	3	4	2	3	3
Responden-8	2	4	4	4	4
Responden-9	3	5	2	4	4
Responden-10	1	4	1	5	2
Responden-11	3	5	2	3	3
Responden-12	3	4	2	3	3
Responden-13	3	3	3	3	3

Responden	SUS-6	SUS-7	SUS-8	SUS-9	SUS-10
Responden-14	3	3	2	4	5
Responden-15	2	5	1	1	1
Responden-16	2	4	1	5	1
Responden-17	1	5	1	5	1
Responden-18	2	1	1	5	1
Responden-19	2	5	1	5	1

Data mentah yang disajikan pada kedua tabel hasil SUS di atas kemudian diolah untuk mendapatkan skor akhir SUS bagi setiap responden. Perhitungan skor ini dilakukan dengan mengikuti formula standar yang memisahkan antara pertanyaan positif (ganjil) dan negatif (genap). Proses perhitungannya dijelaskan dalam Persamaan (9), (10), dan (11) di bawah ini.

$$Score\ Soal_{Ganjil} = (Soal_1 + Soal_3 + Soal_5 + Soal_7 + Soal_9) - 5 \quad (9)$$

$$Score\ Soal_{Genap} = 25 - (Soal_2 + Soal_4 + Soal_6 + Soal_8 + Soal_{10}) \quad (10)$$

$$Nilai\ SUS = (Score\ Soal_{Ganjil} + Score\ Soal_{Genap}) \times 2.5 \quad (11)$$

Berdasarkan Persamaan (9), (10), dan (11) di atas, skor akhir SUS yang diperoleh dari masing-masing responden dapat dilihat pada Tabel 4.24.

Tabel 4.24 Nilai SUS

Responden	Nilai SUS
Responden-1	97.5
Responden-2	97.5
Responden-3	75
Responden-4	72.5
Responden-5	50

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Responden	Nilai SUS
Responden-6	90
Responden-7	65
Responden-8	57.5
Responden-9	57.5
Responden-10	90
Responden-11	67.5
Responden-12	67.5
Responden-13	50
Responden-14	65
Responden-15	82.5
Responden-16	90
Responden-17	100
Responden-18	85
Responden-19	95
Rata-Rata	76.58

Nilai SUS hasil perhitungan dari data yang ditampilkan pada Tabel 4.24 menunjukkan nilai SUS rata-rata sebesar 76.58. Berdasarkan standar interpretasi pada tabel 2.3, nilai tersebut masuk ke dalam peringkat adjektif "Good" dengan Grade "B". Hal ini menandakan bahwa aplikasi *chatbot* rekomendasi pekerjaan yang telah dikembangkan memiliki tingkat kebergunaan yang baik dan dapat diterima dengan mudah oleh pengguna akhir.

4.6.3 Net Promoter Score (NPS)

a. Deskripsi Pengujian

Pengujian *Net Promoter Score* (NPS) dilakukan untuk mengukur tingkat loyalitas dan kepuasan pengguna dengan cara mengetahui seberapa besar kemungkinan pengguna akan merekomendasikan aplikasi JobChat kepada teman atau kolega

mereka. Metode ini menggunakan satu pertanyaan inti yang meminta pengguna untuk memberikan peringkat pada skala 0 hingga 10. Jawaban dari pengguna kemudian dikelompokkan menjadi tiga kategori, yakni *detractors* (skor 0-6), *passives* (skor 7-8), dan *promoters* (skor 9-10). Skor NPS akhir dihitung dengan mengurangi persentase *detractors* dari persentase *promoters*. Pertanyaan yang diajukan kepada responden untuk pengujian NPS disajikan pada Tabel 4.25

Tabel 4.25 Pertanyaan NPS

Kode	Pertanyaan	Format Jawaban
NPS-1	Seberapa besar kemungkinan Anda merekomendasikan aplikasi ini kepada teman atau rekan Anda?	Skala Likert 1 - 10 1 = Sangat Tidak Merekomendasikan 10 = Sangat Merekomendasikan

b. Data Hasil Pengujian

Pengumpulan data untuk pengujian *Net Promoter Score* (NPS) dilakukan dengan metode yang sama seperti pada pengujian SUS. Kuesioner disebarkan menggunakan Google Form dan diisi oleh total 19 responden yang sama. Hasil SUS dari kuesioner uji coba aplikasi *chatbot* dapat dilihat pada Tabel 4.26.

Tabel 4.26 Hasil NPS

Responden	NPS-1
Responden-1	8
Responden-2	8
Responden-3	10
Responden-4	10
Responden-5	10
Responden-6	10
Responden-7	8

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Responden	NPS-1
Responden-1	8
Responden-2	8
Responden-3	10
Responden-8	8
Responden-9	7
Responden-10	7
Responden-11	9
Responden-12	8
Responden-13	8
Responden-14	7
Responden-15	9
Responden-16	9
Responden-17	10
Responden-18	10
Responden-19	9

Data mentah yang disajikan pada tabel hasil kuesioner NPS di atas kemudian diolah untuk menghitung skor akhir dengan menggunakan rumus pada persamaan (6). Proses kalkulasi dan hasil akhir dari nilai *Net Promoter Score* (NPS) untuk aplikasi ini dapat dilihat secara rinci pada Tabel 4.27.

Tabel 4.27 Perhitungan Nilai NPS

	1	2	3	4	5	6	7	8	9	10
NPS-1	0	0	0	0	0	0	3	6	4	6
Kategori	<i>Detractor</i>						<i>Passive</i>		<i>Promoter</i>	
Persentase	0 %						47.37 %		52.63 %	

Nilai NPS	52.63 %
-----------	---------

Hasil perhitungan yang tertera menunjukkan nilai *Net Promoter Score* (NPS) untuk aplikasi JobChat adalah sebesar 52.63 %.. Hal tersebut menunjukkan bahwa pengguna tidak hanya merasa puas, tetapi juga antusias untuk merekomendasikan aplikasi JobChat kepada orang lain. Ini adalah indikasi yang sangat positif dan menjadi validasi kuat bahwa aplikasi ini telah berhasil memenuhi kebutuhan pengguna.

4.7 Deployment

Tahap *deployment* atau penyebaran sistem merupakan fase akhir dalam siklus pengembangan di mana aplikasi dan layanan pendukungnya disiapkan agar dapat diakses oleh pengguna akhir untuk pengujian atau penggunaan. Proses *deployment* untuk proyek ini dibagi menjadi dua bagian utama, yakni *deployment* untuk *backend* API dan *deployment* untuk aplikasi *chatbot*.

a. Deployment API

Backend yang dikembangkan menggunakan FastAPI berjalan di lingkungan lokal (*localhost*). Secara *default*, server ini tidak dapat diakses dari jaringan internet eksternal, sehingga tidak dapat berkomunikasi dengan aplikasi android yang terpasang di perangkat pengguna. Untuk menjembatani hal ini, digunakan layanan *tunneling* bernama Ngrok. Ngrok berfungsi untuk membuat sebuah tunnel yang aman dari sebuah endpoint publik di internet ke server lokal yang berjalan di mesin pengembang. Dengan Ngrok, sebuah alamat URL publik yang unik akan dibuat dan semua permintaan yang masuk ke URL tersebut akan diteruskan secara aman ke server FastAPI di *localhost*. Alamat URL publik inilah yang kemudian dimasukkan ke dalam kode aplikasi android sebagai alamat API, memungkinkan aplikasi untuk berkomunikasi dengan layanan seperti model AI dan database aplikasi.

b. Deployment Aplikasi Chatbot

Aplikasi yang dikembangkan untuk platform Android akan di-*build* atau dikompilasi menjadi sebuah paket instalasi dengan format file .APK (*Android*

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Package Kit). File .APK ini berisi semua aset dan kode yang dibutuhkan agar aplikasi dapat berjalan di perangkat pengguna.

Untuk distribusi, file .APK yang telah jadi tersebut kemudian dipublikasikan melalui repositori publik di GitHub melalui tautan: <https://github.com/GerrysatriaH/JobChatApp>. Dengan demikian, pengguna akhir atau responden pengujian dapat mengunduh file .APK ini secara langsung melalui tautan yang disediakan pada repositori tersebut untuk kemudian dipasang pada perangkat android masing-masing.



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

BAB V PENUTUP

5.1 Kesimpulan

Penelitian mengenai "Rancang Bangun Aplikasi *Chatbot* Menggunakan Teknologi OCR dan Model *LLaMA* untuk Rekomendasi Pekerjaan" telah berhasil dilaksanakan. Hasil dari penelitian ini meliputi pengembangan beberapa model kecerdasan buatan dan sebuah aplikasi *chatbot* fungsional. Model klasifikasi dokumen CV yang dikembangkan menggunakan LSTM berhasil mencapai akurasi sebesar 98,75%. Selain itu, penelitian ini juga berhasil mengimplementasikan model ekstraksi kata kunci, OCR, dan *large language model* dengan menggunakan model *LLaMA 3 70B* ke dalam sebuah aplikasi *chatbot* berbasis *android*. Aplikasi ini dilengkapi dengan berbagai fitur, termasuk registrasi dan *login*, rekomendasi pekerjaan melalui input teks dan dokumen, serta manajemen percakapan.

Aplikasi *chatbot* yang telah dikembangkan ini telah melalui serangkaian pengujian, yaitu pengujian fungsionalitas dengan menggunakan pengujian *black box*, pengujian kebergunaan dengan menggunakan *system usability scale* (SUS) serta *net promoter score* (NPS) kepada pengguna akhir. Dari hasil pengujian tersebut, aplikasi JobChat berhasil mendapatkan nilai *system usability scale* (SUS) sebesar 76.58 dan *net promoter score* (NPS) sebesar 52.63. Nilai SUS menunjukkan bahwa aplikasi ini dinilai "Baik" dan mudah digunakan, sementara nilai NPS yang tergolong baik yang mengindikasikan bahwa pengguna sangat puas dan dapat merekomendasikan aplikasi ini kepada teman atau orang lain.

Meskipun hasil pengujian menunjukkan performa yang sangat baik, aplikasi dan model yang telah dikembangkan masih memiliki beberapa keterbatasan yang perlu diperhatikan. Keterbatasan utama terletak pada validasi dokumen, di mana sistem hanya mampu memproses file berformat PDF dengan batasan maksimal dua halaman. Selain itu, kecepatan respons dari *chatbot* bersifat variabel karena sangat dipengaruhi oleh faktor eksternal seperti kualitas koneksi internet dan jenis input dari pengguna ke *chatbot*. Kualitas rekomendasi yang dihasilkan juga memiliki dependensi yang kuat terhadap kata kunci yang diekstraksi dari input pengguna,



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

serta terbatas pada variasi kategori pekerjaan yang saat ini tersedia di dalam *vector database*.

5.2 Saran

Berdasarkan hasil penelitian dan pengujian aplikasi *chatbot* rekomendasi pekerjaan, berikut beberapa saran yang dapat dipertimbangkan untuk pengembangan lebih lanjut:

1. Penyempurnaan fitur rekomendasi, mengingat kualitas rekomendasi adalah inti dari aplikasi, disarankan untuk terus menyempurnakan model. Salah satu pendekatan yang bisa dieksplorasi adalah mengintegrasikan sumber data lowongan pekerjaan secara *real-time* melalui API dari platform pencarian kerja, sehingga rekomendasi yang diberikan selalu aktual dan relevan dengan pasar kerja saat ini.
2. Pengembangan fitur lanjutan, untuk meningkatkan nilai tambah aplikasi, dapat dipertimbangkan penambahan fitur-fitur baru. Beberapa fitur yang potensial antara lain fitur pelacakan lamaran kerja, dan fitur untuk membuat CV langsung di dalam aplikasi.
3. Pengujian skala besar dan iteratif, melakukan pengujian kepada jumlah pengguna yang lebih besar dengan latar belakang yang lebih beragam dapat memberikan wawasan yang lebih komprehensif mengenai kelebihan dan kekurangan aplikasi. Disarankan juga untuk melakukan pengujian secara berkala seiring dengan penambahan fitur baru untuk memastikan kualitas dan kebergunaan aplikasi tetap terjaga.

Dengan menerapkan saran-saran tersebut, diharapkan aplikasi *chatbot* rekomendasi pekerjaan ini dapat terus ditingkatkan kualitas, fungsionalitas, dan keandalannya, sehingga mampu memberikan manfaat bagi para pencari kerja dalam memilih pekerjaan yang ingin mereka dapatkan.



DAFTAR PUSTAKA

- Albert, G.D. and Voutama, A., 2025. Pengembangan chatbot berbasis PDF menggunakan local retrieval-augmented generation (RAG) dan Ollama. *Jurnal Informatika dan Teknik Elektro Terapan*, 13(2).
- Arjiansa, R.R. and Sutabri, T., 2023. Pengukuran Tingkat Kemudahan Pegawai Terhadap Penggunaan Layanan Aplikasi SIMRS Menggunakan Metode System Usability Scale (SUS) Pada Rumah Sakit Umum Daerah Sekayu. *Indonesian Journal of Multidisciplinary on Social and Technology*, 1(2), pp.115-120.
- Arsi, M. et al. (2023) 'Sistem Informasi Pencarian Jasa Tukang Berbasis Android (Studi Kasus: Bandar Lampung)', 4(1), pp. 67–74. Available at: <https://doi.org/10.33365/jtsi.v4i1.2441>.
- Aryanti, R., Agustiani, S., Wildah, S.K., Arifin, Y.T., Marlina, S. and Misriati, T., 2024. Optimisasi Model Deep Learning untuk Deteksi Penyakit Daun Tebu dengan Fine-Tuning MobileNetV2. *Journal of Informatics Management and Information Technology*, 4(4), pp.150-157.
- Azimah, F., Rizky, K. and Wardani, N., 2022. Sistem Pendeteksi Gejala Awal Covid-19 Dengan Penggunaan Metode Ai Project Cycle.
- Babu, A. and Boddu, S.B. (2024) 'BERT-Based Medical Chatbot: Enhancing Healthcare Communication through Natural Language Understanding', *Exploratory Research in Clinical and Social Pharmacy*, 13. Available at: <https://doi.org/10.1016/j.rcsop.2024.100419>.
- Brown, T.B. et al. (2020) Language Models are Few-Shot Learners. Available at: <https://commoncrawl.org/the-data/>.
- Cahyawati, A.F. and Hadiono, K., 2024. IMPLEMENTASI CHATBOT UNTUK REKOMENDASI COFFEE SHOP DI KOTA SEMARANG. *JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, 9(4), pp.1779-1788.
- Caldarini, G., Jaf, S. and McGarry, K. (2022) 'A Literature Survey of Recent Advances in Chatbots', *Information (Switzerland)*, 13(1). Available at: <https://doi.org/10.3390/info13010041>.
- Calzone, O. 2022. An Intuitive Explanation of LSTM - Ottavio Calzone - Medium. [online] Medium. Available at: https://medium-com.translate.goog/@ottaviocalzone/an-intuitive-explanation-of-lstm-a035eb6ab42c? x_tr_sl=en& x_tr_tl=id& x_tr_hl=id& x_tr_pto=imgs& x_tr_hist=true [Accessed 22 Mei 2025].

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



- Dhamdhere, A.Y. et al. (2024) 'ASPIREAI: AI Chatbot for Career Guidance', International Journal of Innovative Research in Engineering, 5(2), pp. 254–259. Available at: <https://doi.org/10.59256/ijire.20240502033>.
- Fadlil, A., Umar, R. and Juliansyah, F., 2022. Klasifikasi Loyalitas Pengguna Data Alumni Pada Forlap Dikti Menggunakan Metode Net Promotore Score. JURIKOM: Jurnal Riset Komputer, 9(3), pp.740-748.
- GeeksforGeeks, 2024. Introduction to Generative Pretrained Transformer (GPT). [online] GeeksforGeeks. Available at: <https://www.geeksforgeeks.org/artificial-intelligence/introduction-to-generative-pre-trained-transformer-gpt/> [Accessed 21 May 2025].
- Hariyanto, H., Kuat, T., Suprap, S., Rohmad, A.N., Faizin, A.N. and Hadi, H., 2024. Faktor-Faktor yang Mempengaruhi Pemilihan Karir pada Siswa SMK. Journal of Education Research, 5(4), pp.6247-6253.
- Hermawan, G., 2024. Memahami Peran Dataset dalam Penelitian Kecerdasan Buatan: Kualitas, Aksesibilitas, dan Tantangan. Preprint, Oct.
- Herwanza, N.A.M. et al. (2024) 'Penerapan Langchain Retriever dengan Model Chat Openai dalam Pengembangan Sistem Chatbot Hadis Berbasis Telegram', JTIM : Jurnal Teknologi Informasi dan Multimedia, 6(1), pp. 70–83. Available at: <https://doi.org/10.35746/jtim.v6i1.514>.
- Hou, G. and Lian, Q. (2024) 'Benchmarking of Commercial Large Language Models: ChatGPT, Mistral, and Llama'. Available at: <https://doi.org/10.21203/rs.3.rs-4376810/v1>.
- Luid, M.F.J.K. et al. (2024) 'Sistem Pakar Untuk Perencanaan Karier Dengan Algoritma Forward Chaining Berbasis Web', Diffusion: Journal of Systems and Information Technology, 4, pp. 115–125.
- Meta. 2022. Llama 3.2.URL <https://www.llama.com/> [accessed 11 Februari 2025].
- Mumpuni, E., 2024. Implementasi SHAP Pada Catboost Untuk Meningkatkan Akurasi Prediksi Temperatur Udara Di Kota Pekanbaru. Jurnal Studi Multidisipliner, 8(6).
- Nazeem, M., Anitha, R. and Navaneeth, S., 2024, December. Open-Source OCR Libraries: A Comprehensive Study for Low Resource Language. In Proceedings of the 21st International Conference on Natural Language Processing (ICON) (pp. 416-421).

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



Ningsih, W., & Abdullah, F. 2021. Analisis perbedaan pencari kerja dan lowongan kerja sebelum dan pada saat pandemi COVID-19 di Kota Malang. *Journal of Regional Economics Indonesia*, 2(1), pp. 42-56.

Nuraeda, N.R., Liebenlito, M. and Sutanto, T.E., 2024. Explainable Sentiment Analysis pada Ulasan Aplikasi Shopee Menggunakan Local Interpretable Model-agnostic Explanations. *The Indonesian Journal of Computer Science*, 13(3).

OpenAI. (2023). *GPT-4 Technical Report*. Retrieved from <https://arxiv.org/abs/2303.08774>

Prakash, S. et al. (2024) CAREER CONNECT-A Personalized Job Recommendation AI Chatbot. Available at: <https://ssrn.com/abstract=5035946>.

Pranoto, S., Sutiono, S. and Nasution, D., 2024. Penerapan UML Dalam Perancangan Sistem Informasi Pelaporan Dan Evaluasi Pembangunan Pada Bagian Administrasi Pembangunan Sekretariat Daerah Kota Tebing Tinggi. *Surplus: Jurnal Ekonomi dan Bisnis*, 2(2), pp.384-401.

Putri, N.M., Praseptiawan, M. and Untoro, M.C., 2024. Analisis Model Sistem Rekomendasi Kursus Mooc Dengan Metode Collaborative Filtering Dan Integrasi Explainable Ai. *Indonesian Journal of Business Intelligence (IJUBI)*, 7(1), pp.26-36.

Ramadhan, A.K., Masril, M.A. and Caniago, D.P. (2024) 'Implementasi Artificial Intelligence Dalam Aplikasi Chatbot Untuk Rekomendasi Wisata Pantai Di Batam Dengan Metode Feedforward Neural Network', *Jurnal Quancom*, 2(2), pp. 41–46.

Raut, R., Sharma, R. and Kharbade, R. (2021) 'Krishna-The Career Guidance ChatBot', *International Research Journal of Engineering and Technology* [Preprint]. Available at: www.irjet.net.

Rizkilloh, M.F. and Widiyanesti, S., 2022. Prediksi Harga Cryptocurrency Menggunakan Algoritma Long Short Term Memory (LSTM). *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 6(1), pp.25-31

Rohaniyah, S. and Purnamasari, A.I., 2023. Clustering Data Pencari Kerja Menurut Tingkat Pendidikan Menggunakan Algoritma K-Means. *Jurnal Minfo Polgan*, 12(1).

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

- Salsabila, N.S. et al. (2024) ‘PENERAPAN TEKNOLOGI OCR PLAT NOMOR UNTUK MENINGKATKAN EFISIENSI DAN KEAMANAN AKSES KENDARAAN’, Journal of Sciencetech Research and Development , 6, pp. 227–236
- Saltz, J. 2023. What is the AI Life Cycle? [online] Data Science Process Alliance. Available at: <https://www.datascience-pm.com/ai-lifecycle/> [Accessed 18 Mei 2025].
- Samosir, F.V.P., Toba, H. and Ayub, M., 2022. BESKlus: BERT Extractive Summarization with K-Means Clustering in Scientific Paper. Jurnal Teknik Informatika dan Sistem Informasi, 8(1), pp.202-217.
- Smith, B. 2024. Mistral 7B Explained: Towards More Efficient Language Models. [online] Medium. Available at: https://medium-com.translate.goog/data-science/mistral-7b-explained-towards-more-efficient-language-models-7f9c6e6b7251? x_tr_sl=en& x_tr_tl=id& x_tr_hl=id& x_tr_pto=imgs& x_tr_hist=true [Accessed 21 Mei 2025].
- Suryotomo, A.P., Akbar, B.M. and Husaini, R., 2024. Performance Analysis of FastAPI Framework on Lost Circulation Handling Management Application in Oil Well Drilling. Telematika: Jurnal Informatika dan Teknologi Informasi, 21(1), pp.110-121.
- Tarumingkeng, R.C., 2024. Halusinasi-AI. RUDYCT e-PRESS
- Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.A., Lacroix, T., Rozière, B., Goyal, N., Hambro, E., Azhar, F. and Rodriguez, A., 2023. Llama: Open and efficient foundation language models. arXiv preprint arXiv:2302.13971.
- Yaadav, V. 2024. Exploring and building the LLaMA 3 Architecture : A Deep Dive into Components, Coding, and Inference Techniques. [online] Medium. Available at: https://medium-com.translate.goog/@vi.ai_/exploring-and-building-the-llama-3-architecture-a-deep-dive-into-components-coding-and-43d4097cfbbb? x_tr_sl=en& x_tr_tl=id& x_tr_hl=id& x_tr_pto=imgs& x_tr_hist=true [Accessed 21 Mei 2025].
- Wayahdi, M.R. and Ruziq, F., 2023. Pemodelan sistem penerimaan anggota baru dengan unified modeling language (UML)(Studi kasus: Programmer Association of Battuta). Jurnal Minfo Polgan, 12(1), pp.1514-1521.
- Wibowo, D.W., Mentari, M., Chandra, A.D., Kuddah, A.A. and wahyu Putra, R., 2020. Sistem Pendukung Keputusan Berbasis Web Rekomendasi Pekerjaan

Bagi Lulusan JTI Polinema Dengan Metode SAW. JASIEK (Jurnal Aplikasi Sains, Informasi, Elektronika dan Komputer), 2(1), pp.68-79.

Wibowo, A. and Isnain, A.R., 2025. Implementasi Algoritma Machine Learning untuk Klasifikasi Suara Lingkungan: Implementation of Machine Learning Algorithm for Environmental Sound Classification. MALCOM: Indonesian Journal of Machine Learning and Computer Science, 5(2), pp.616-625.

Wijaya, Y.D. and Astuti, M.W., 2021. Pengujian Blackbox Sistem Informasi Penilaian Kinerja Karyawan Pt Inka (Persero) Berbasis Equivalence Partitions. Jurnal Digital Teknologi Informasi, 4(1), pp.22-26.

Winarni, A., Apriyanti, L. and Pedry, P., 2024, September. Aplikasi Pemesanan Lapangan Badminton Dengan Fitur Pembayaran Berbasis Android Di Kota Tanjungpinang. In Prosiding Seminar Nasional Ilmu Sosial dan Teknologi (SNISTEK) (Vol. 6, pp. 365-370).

Zulhilmi, A. and Perdana, R., 2024. Pengenalan Entitas Bernama Menggunakan Bi-LSTM pada Chatbot Bahasa Indonesia. Jurnal Teknologi Informasi dan Ilmu Komputer, 11(5), pp.965-970.

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



**POLITEKNIK
NEGERI
JAKARTA**



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

DAFTAR RIWAYAT HIDUP PENULIS



Gerry Satria Halim

Lahir di Cibinong, Bogor, 10 Agustus 2003. Lulus dari SDN 1 Cilodong pada tahun 2015, SMPN 1 Cibinong pada tahun 2018, dan SMAN 1 Cibinong pada tahun 2021. Saat ini sedang menempuh Pendidikan Sarjana Terapan pada Program Studi Teknik Informatika di Politeknik Negeri Jakarta. Tertarik pada bidang *machine learning*, *deep learning*, pengembangan aplikasi *mobile* dan aplikasi web.

**POLITEKNIK
NEGERI
JAKARTA**

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

LAMPIRAN

Lampiran 1 *Prompt* LLM

""

Berdasarkan keahlian yang dipunya seperti {skills} dalam kategori pekerjaan {categories}

berikan rekomendasi pekerjaan yang cocok, dan berikan jawabannya dengan format berikut:

"

Halo, Job Seeker!

Berdasarkan keahlian yang Anda miliki, kami telah menemukan beberapa pekerjaan yang mungkin cocok untuk Anda.

Kategori Pekerjaan: {categories}

Kami merekomendasikan Anda untuk mempertimbangkan pekerjaan berikut:

Nama Pekerjaan

Deskripsi Pekerjaan:

Alasan Cocok:

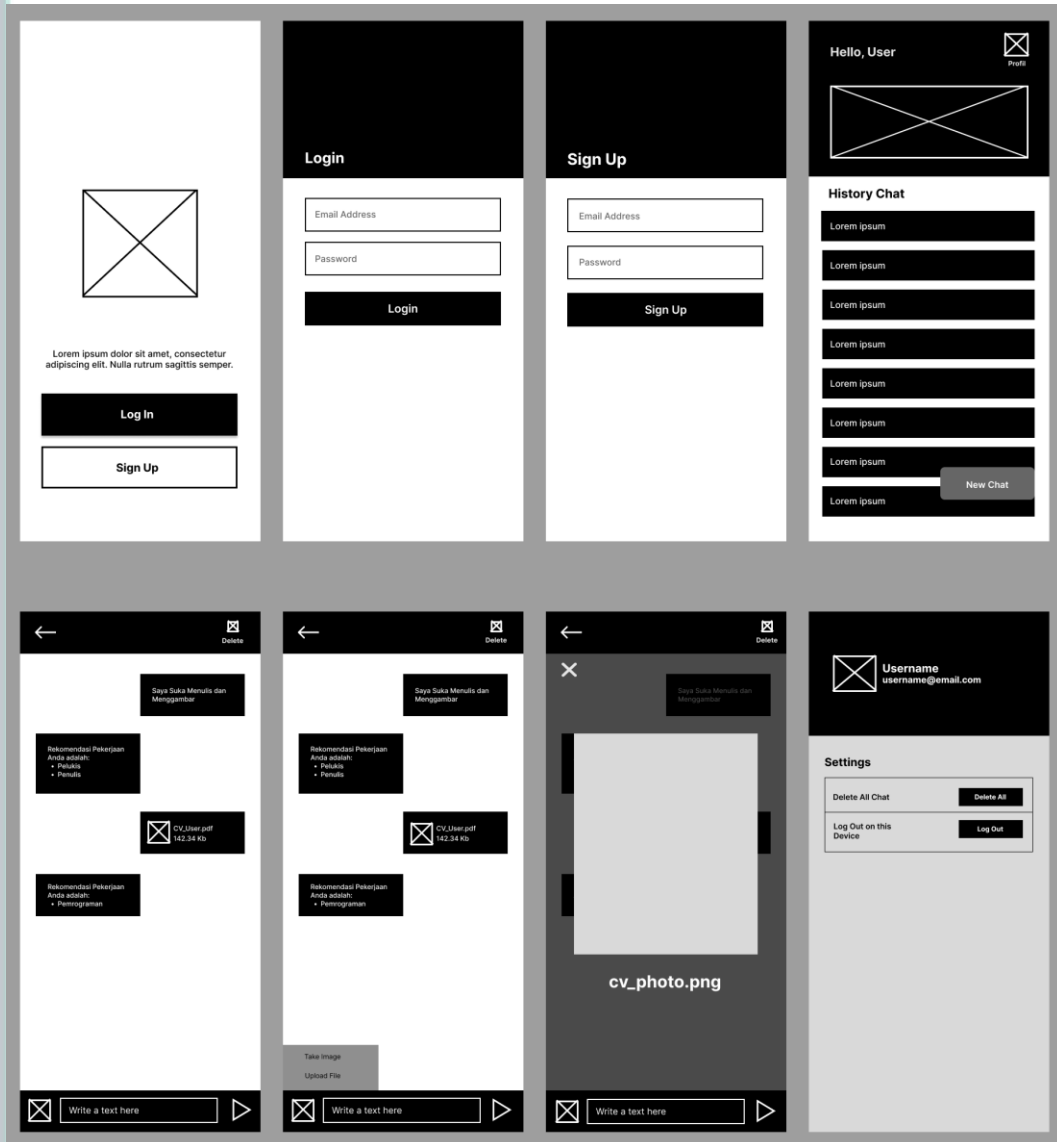
"

Jangan tambahkan format lain atau opini pribadi.

""

Hak Cipta :

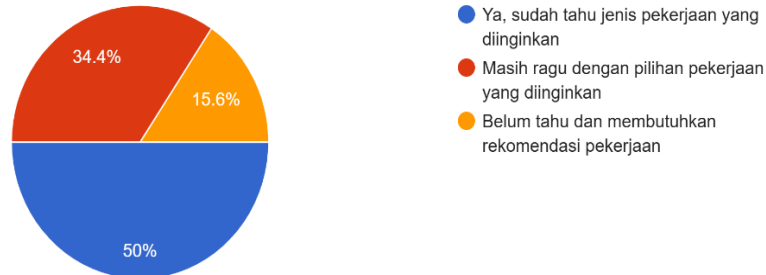
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



Lampiran 3 Hasil Survei Identifikasi Permasalahan Pengguna

Apakah Anda sudah memiliki gambaran tentang Jenis pekerjaan yang ingin Anda pilih?

32 responses



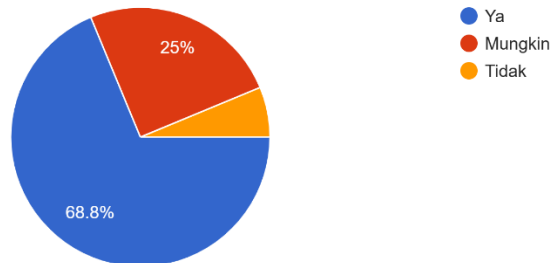
Apa tantangan terbesar yang Anda hadapi dalam memilih pekerjaan yang ingin dijalani?

32 responses



Apakah Anda ingin mendapatkan rekomendasi pekerjaan yang relevan dengan minat dan kompetensi yang anda miliki?

32 responses



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



POLITEKNIK
NEGERI
JAKARTA

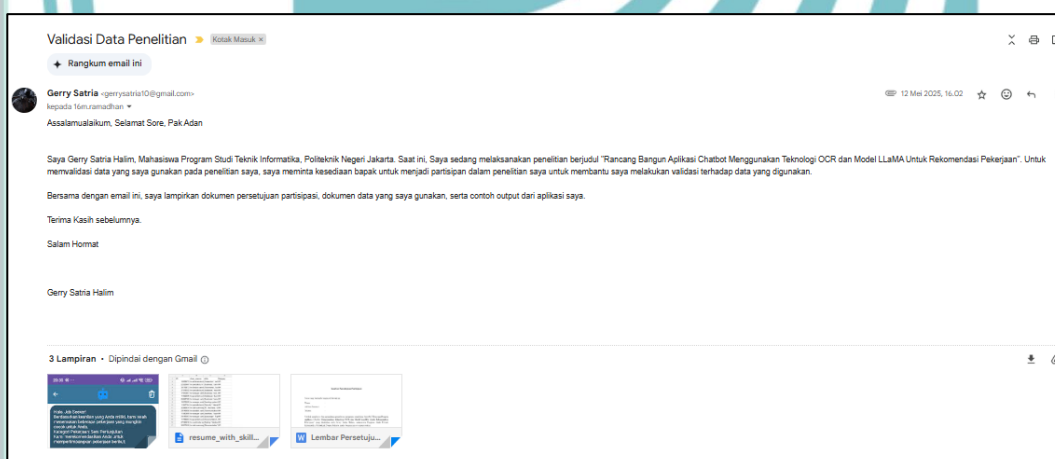
Lampiran 4 Dokumentasi Validasi Data Aplikasi



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



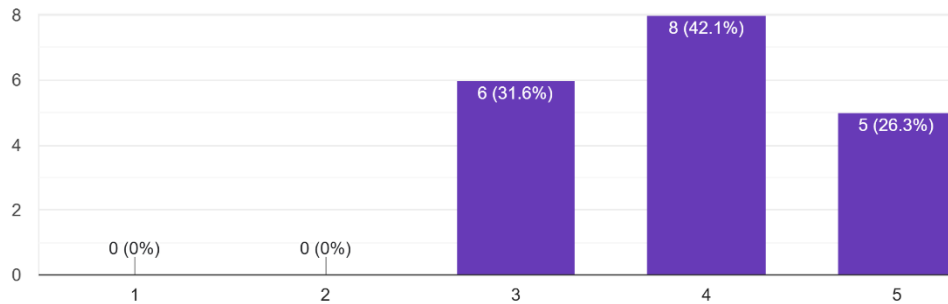
Lampiran 5 Hasil *System Usability Scale* (SUS)

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

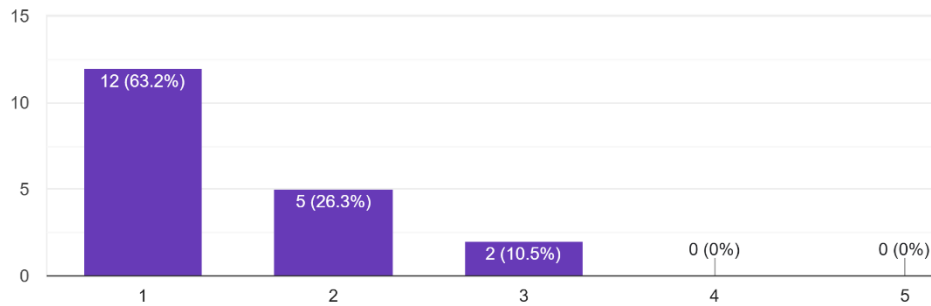
Saya pikir saya akan sering menggunakan aplikasi ini.

19 responses



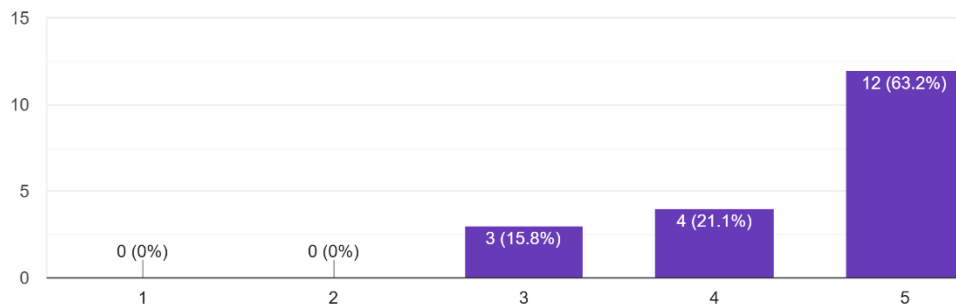
Saya merasa aplikasi ini terlalu rumit digunakan.

19 responses



Saya merasa aplikasi ini mudah digunakan.

19 responses

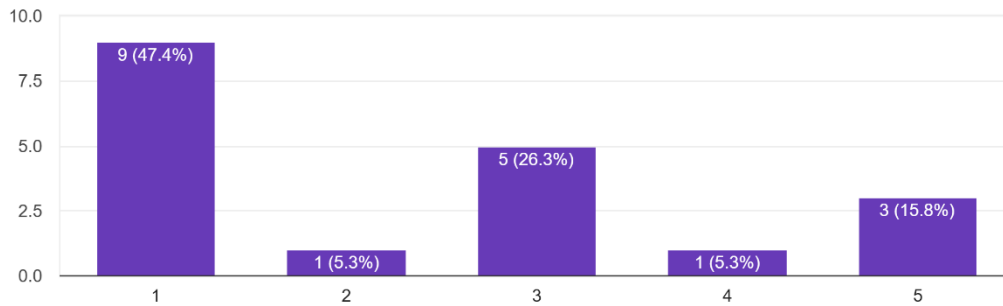


Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

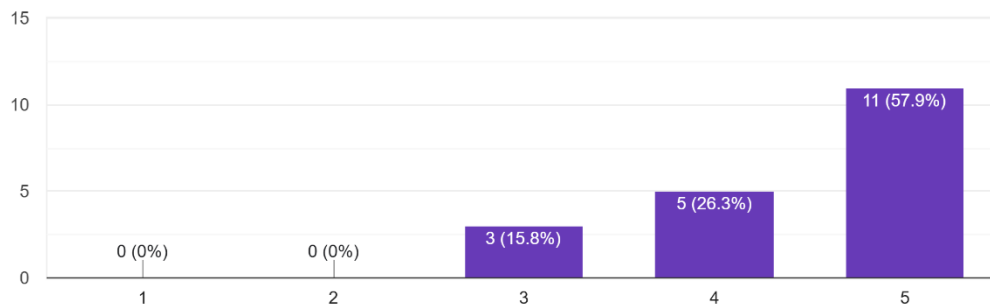
Saya pikir saya membutuhkan bantuan orang lain atau teknisi untuk dapat menggunakan aplikasi ini.

19 responses



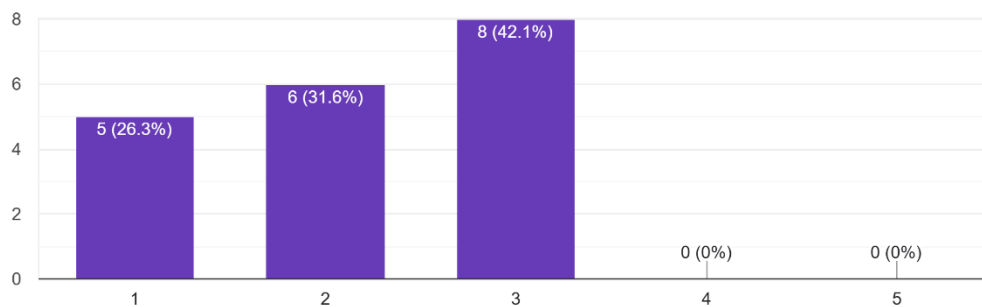
Saya merasa berbagai fungsi atau fitur pada aplikasi ini dapat berjalan dengan semestinya.

19 responses



Saya merasa ada banyak hal yang tidak konsisten pada aplikasi ini.

19 responses

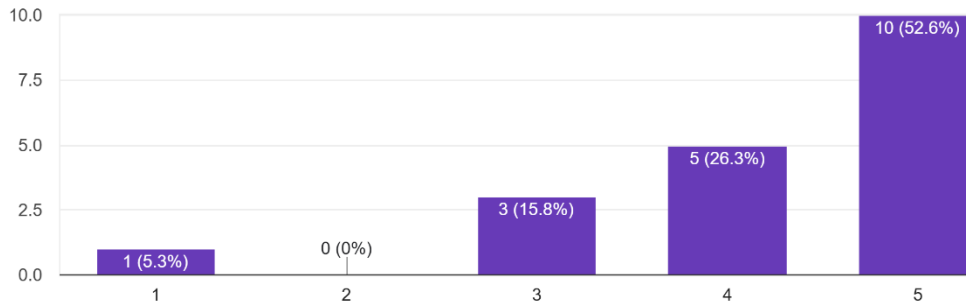


Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

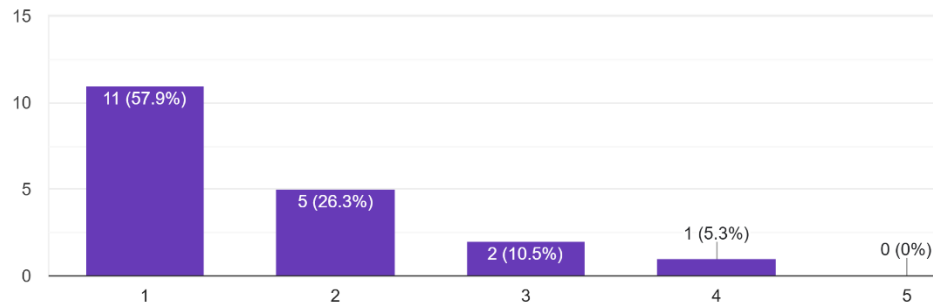
Saya merasa kebanyakan orang akan mudah memahami aplikasi ini dengan sangat cepat.

19 responses



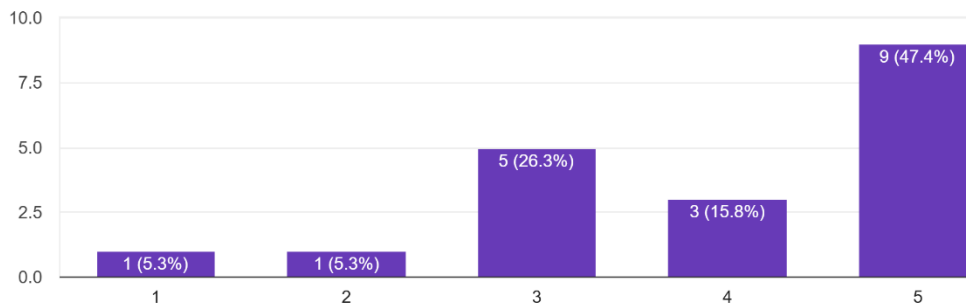
Saya merasa aplikasi ini sangat membingungkan untuk digunakan.

19 responses



Saya merasa tidak ada hambatan dalam menggunakan aplikasi ini.

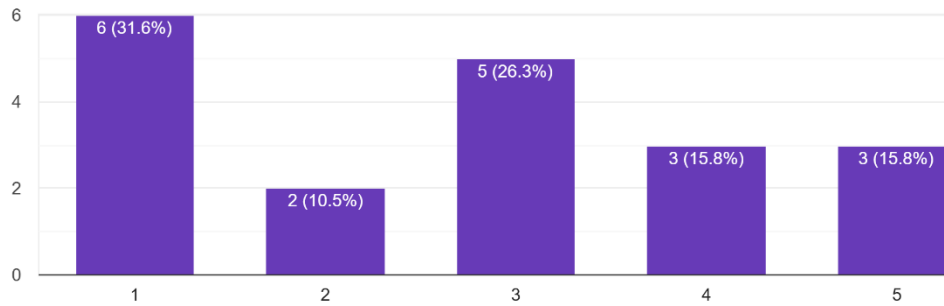
19 responses





Saya perlu beradaptasi terlebih dahulu sebelum bisa menggunakan sistem ini.

19 responses



Hak Cipta :

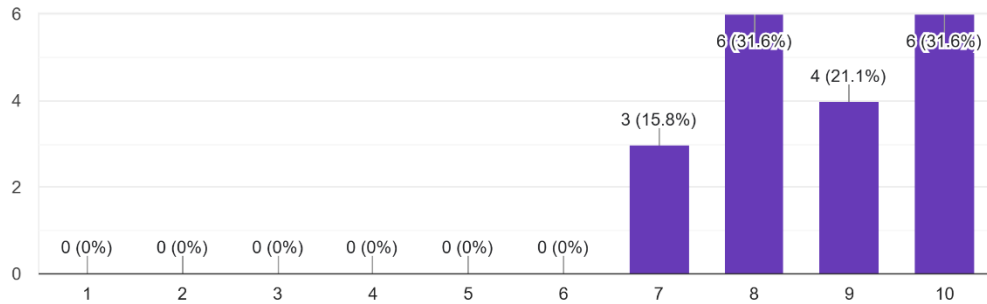
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



Lampiran 6 Hasil *Net Promoter Score* (NPS)

Seberapa besar kemungkinan Anda merekomendasikan aplikasi ini kepada teman atau rekan Anda?

19 responses



- Hak Cipta :**
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
 2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta