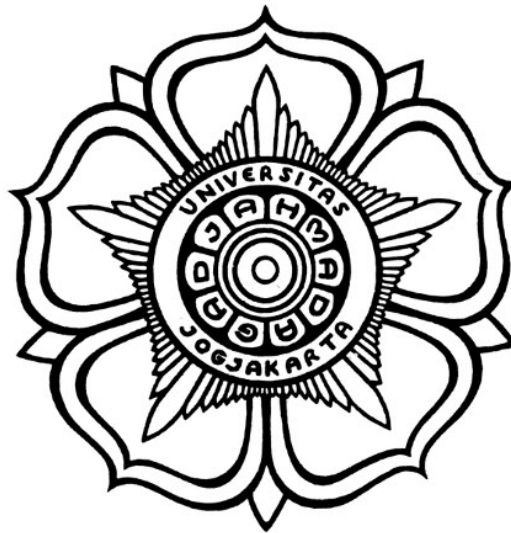


«JUDUL SKRIPSI»

SKRIPSI



THE SUSTAINABLE DEVELOPMENT GOALS
Industry, Innovation and Infrastructure
Affordable and Clean Energy
Climate Action

Disusun oleh:

«QORNAIN AJI»
«21/481767/TK/53170»

PROGRAM SARJANA PROGRAM STUDI «TEKNOLOGI INFORMASI»
DEPARTEMEN TEKNIK ELEKTRO DAN TEKNOLOGI INFORMASI
FAKULTAS TEKNIK UNIVERSITAS GADJAH MADA
YOGYAKARTA
«TAHUN PENDADARAN»

HALAMAN PENGESAHAN

«JUDUL SKRIPSI»

SKRIPSI

Diajukan Sebagai Salah Satu Syarat untuk Memperoleh
Gelar Sarjana Teknik
pada Departemen Teknik Elektro dan Teknologi Informasi
Fakultas Teknik
Universitas Gadjah Mada

Disusun oleh:

«QORNAIN AJI»
«21/481767/TK/53170»

Telah disetujui dan disahkan

Pada tanggal

Dosen Pembimbing I

Dosen Pembimbing II

Dosen Pembimbing 1, S.T., M.Eng., PhD.

«NIP xxxxxx»

Dosen Pembimbing 2, S.T., M.Eng., PhD.

«NIP xxxxxx»

PERNYATAAN BEBAS PLAGIASI

Saya yang bertanda tangan di bawah ini :

Nama :
NIM :
Tahun terdaftar :
Program : Sarjana
Program Studi :
Fakultas : Teknik Universitas Gadjah Mada

Menyatakan bahwa dalam dokumen ilmiah Skripsi ini tidak terdapat bagian dari karya ilmiah lain yang telah diajukan untuk memperoleh gelar akademik di suatu lembaga Pendidikan Tinggi, dan juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang/lembaga lain, kecuali yang secara tertulis disitasi dalam dokumen ini dan disebutkan sumbernya secara lengkap dalam daftar pustaka.

Dengan demikian saya menyatakan bahwa dokumen ilmiah ini bebas dari unsur-unsur plagiasi dan apabila dokumen ilmiah Skripsi ini di kemudian hari terbukti merupakan plagiasi dari hasil karya penulis lain dan/atau dengan sengaja mengajukan karya atau pendapat yang merupakan hasil karya penulis lain, maka penulis bersedia menerima sanksi akademik dan/atau sanksi hukum yang berlaku.

Yogyakarta, tanggal-bulan-tahun

Materai Rp10.000

(Tanda tangan)

Nama Mahasiswa
NIM

HALAMAN PERSEMBAHAN

Tuliskan kepada siapa skripsi ini dipersembahkan!

contoh

KATA PENGANTAR

Contoh: Puji syukur ke hadirat Allah SWT atas limpahan rahmat, karunia, serta petunjuk-Nya sehingga tugas akhir berupa penyusunan skripsi ini telah terselesaikan dengan baik. Dalam hal penyusunan tugas akhir ini penulis telah banyak mendapatkan arahan, bantuan, serta dukungan dari berbagai pihak. Oleh karena itu pada kesempatan ini penulis mengucapkan terima kasih kepada:

1. Orang 1 yang telah
2. Orang 2 yang telah
3. <isi dengan nama orang lainnya>

Akhir kata penulis berharap semoga skripsi ini dapat memberikan manfaat bagi kita semua, aamiin.

Catatan: setiap nama yang dituliskan boleh disertai dengan alasan berterima kasih.

DAFTAR ISI

HALAMAN PENGESAHAN	ii
PERNYATAAN BEBAS PLAGIASI	iii
HALAMAN PERSEMBAHAN	iv
KATA PENGANTAR	v
DAFTAR ISI	vi
DAFTAR TABEL	ix
DAFTAR GAMBAR	x
DAFTAR SINGKATAN.....	xi
INTISARI.....	xiii
ABSTRACT	xiv
BAB I Pendahuluan	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan Penelitian	3
1.4 Batasan Penelitian	3
1.4.1 Objek Penelitian.....	3
1.4.2 Metode Penelitian	4
1.4.3 Waktu dan Tempat Penelitian.....	4
1.4.4 Populasi dan Sampel	4
1.4.5 Variabel	4
1.4.6 Keterbatasan Penelitian	4
1.5 Manfaat Penelitian	5
1.6 Sistematika Penulisan.....	5
BAB II Tinjauan Pustaka dan Dasar Teori	7
2.1 Tinjauan Pustaka	7
2.1.1 Peningkatan Umpan Balik LLM dengan <i>Direct Preference Optimization</i> (DPO).....	8
2.1.2 Peningkatan <i>Feedback</i> LLM dengan <i>Prompt Engineering</i> Berbasis <i>Learning Analytics</i>	10
2.1.3 <i>Reinforcement Learning from Human Feedback</i> (RLHF) untuk Peningkatan Kualitas Respons	11
2.1.4 Regulasi Diri dalam Pendidikan Pemrograman Berbasis LLM dengan <i>Learning Analytics</i>	12
2.1.5 Dashboard <i>Real-time</i> untuk Orkestrasi Kelas Berbasis Kanban	14
2.1.6 Perbandingan Metode Peningkatan Kualitas Respons LLM	16
2.1.7 Metode Evaluasi Umpan Balik LLM dalam Pendidikan	19

2.1.8	Evaluasi Ahli Manusia	20
2.2	Hipotesis Penelitian	21
2.3	Dasar Teori	22
2.3.1	<i>Self Regulated Learning</i> (SRL) dan Peran Umpan Balik	22
2.3.2	<i>Prompt Engineering</i>	24
2.3.3	<i>Context Engineering</i>	24
2.3.4	Perbedaan <i>Prompt Engineering</i> dan <i>Context Engineering</i>	25
2.3.5	<i>Learning Analytics</i> dan Cakupannya	25
2.4	Penelitian Terkait	27
2.4.1	Studi Menggabungkan LLM dengan <i>Learning Analytics</i>	27
2.4.2	Pendekatan Rekayasa dalam <i>Learning Analytics</i>	28
2.4.3	Menghubungkan Hasil <i>Learning Analytics</i> dengan <i>Prompt Context-Aware</i> pada LLM.....	29
2.5	Analisis Perbandingan Metode	30
2.6	Pertanyaan Tugas Akhir (Jika Perlu).....	30
BAB III Metode Penelitian.....		31
3.1	Alat dan Bahan Tugas akhir (Opsional).....	31
3.1.1	Alat Tugas akhir.....	31
3.1.2	Bahan Tugas akhir	31
3.2	Metode yang Digunakan.....	32
3.3	Alur Tugas Akhir	32
3.4	Etika, Masalah, dan Keterbatasan Penelitian (Opsional)).....	32
BAB IV Hasil dan Pembahasan.....		33
4.1	Pembahasan Tujuan 1 dengan Hasil Penelitian 1 (Ubah Judul Sesuai dengan Hal yang Hendak dibahas)	33
4.2	Pembahasan Tujuan 1 dengan Hasil Penelitian 2 (Ubah Judul Sesuai dengan Hal yang Hendak dibahas)	33
4.3	Pembahasan Tujuan 2 dengan Hasil Penelitian 3 (Ubah Judul Sesuai dengan Hal yang Hendak dibahas)	33
4.4	Perbandingan Hasil Penelitian dengan Hasil Terdahulu	33
BAB V Tambahan (Opsional)		34
BAB VI Kesimpulan dan Saran.....		35
6.1	Kesimpulan.....	35
6.2	Saran.....	35
DAFTAR PUSTAKA.....		36
LAMPIRAN		L-1
L.1	Isi Lampiran.....	L-1
L.2	Panduan Latex.....	L-2
L.2.1	Syntax Dasar	L-2

	L.2.1.1	Penggunaan Sitasi	L-2
	L.2.1.2	Penulisan Gambar	L-2
	L.2.1.3	Penulisan Tabel	L-2
	L.2.1.4	Penulisan formula.....	L-2
	L.2.1.5	Contoh list.....	L-3
L.2.2		Blok Beda Halaman.....	L-3
	L.2.2.1	Membuat algoritma terpisah	L-3
	L.2.2.2	Membuat tabel terpisah.....	L-3
	L.2.2.3	Menulis formula terpisah halaman.....	L-4
L.3		Format Penulisan Referensi	L-6
	L.3.1	Book	L-6
	L.3.2	Handbook.....	L-8
L.4		Contoh Source Code	L-10
	L.4.1	Sample algorithm	L-10
	L.4.2	Sample Python code	L-11
	L.4.3	Sample Matlab code	L-12

DAFTAR TABEL

Tabel 2.1	Komparasi penelitian peningkatan kualitas umpan balik LLM di pendidikan	17
Tabel 2.2	Perbandingan <i>Prompt Engineering</i> dan <i>Context Engineering</i>	25
Tabel L.1	Tabel ini	L-2
Tabel L.2	Contoh tabel panjang	L-4

DAFTAR GAMBAR

Gambar L.1	Contoh gambar.	L-2
------------	---------------------	-----

DAFTAR SINGKATAN

[SAMPLE]

b	=	bias
$K(x_i, x_j)$	=	fungsi kernel
y	=	kelas keluaran
C	=	parameter untuk mengendalikan besarnya pertukaran antara penalti variabel slack dengan ukuran margin
L_D	=	persamaan Lagrange dual
L_P	=	persamaan Lagrange primal
\mathbf{w}	=	vektor bobot
\mathbf{x}	=	vektor masukan
ANFIS	=	Adaptive Network Fuzzy Inference System
ANSI	=	American National Standards Institute
DAG	=	Directed Acyclic Graph
DDAG	=	Decision Directed Acyclic Graph
HIS	=	Hue Saturation Intensity
QP	=	Quadratic Programming
RBF	=	Radial Basis Function
RGB	=	Red Green Blue
SV	=	Support Vector
SVM	=	Support Vector Machines

INTISARI

Intisari ditulis menggunakan bahasa Indonesia dengan jarak antar baris 1 spasi dan maksimal 1 halaman. Intisari sekurang-kurangnya berisi tentang latar belakang dan tujuan penelitian, metodologi yang digunakan, hasil penelitian, kesimpulan dan implikasi, dan Kata kunci yang berhubungan dengan penelitian.

Kata Kunci ditulis maksimal 5 kata yang paling berhubungan dengan isi skripsi. Silakan mengacu pada ACM / IEEE *Computing classification* jika Anda adalah mahasiswa Sarjana TI <http://www.acm.org/about/class/> atau mengacu kepada IEEE keywords http://www.ieee.org/documents/taxonomy_v101.pdf jika Anda berasal dari Prodi Sarjana TE.

Kata kunci : Kata kunci 1, Kata kunci 2, Kata kunci 3, Kata kunci 4, Kata kunci 5

Contoh Abstrak Teknik Elektro:

"Penelitian ini bertujuan untuk mengembangkan sistem pengendalian suhu ruangan dengan menggunakan microcontroller. Metodologi yang digunakan adalah desain sirkuit, implementasi sistem pengendalian, dan pengujian performa. Hasil penelitian menunjukkan bahwa sistem pengendalian suhu ruangan yang dikembangkan mampu mengendalikan suhu ruangan dengan akurasi sebesar $\pm 0,5^{\circ}\text{C}$. Kesimpulan dari penelitian ini adalah sistem pengendalian suhu ruangan yang dikembangkan efektif dan efisien.

Kata kunci: microcontroller, sistem pengendalian suhu, akurasi."

Contoh Abstrak Teknik Biomedis:

"Penelitian ini bertujuan untuk mengevaluasi keefektifan prototipe alat pemantau denyut jantung berbasis elektrokardiogram (ECG) untuk pasien jantung. Metodologi yang digunakan meliputi desain dan pembuatan prototipe, pengujian dengan pasien, dan analisis data. Hasil penelitian menunjukkan bahwa prototipe alat pemantau denyut jantung berbasis ECG memiliki akurasi yang baik dan mampu memantau denyut jantung pasien secara efektif. Kesimpulan dari penelitian ini adalah prototipe alat pemantau denyut jantung berbasis ECG merupakan solusi yang efektif dan efisien untuk memantau pasien jantung.

Kata kunci: elektrokardiogram, alat pemantau denyut jantung, akurasi."

Contoh Abstrak Teknologi Informasi:

"Penelitian ini bertujuan untuk mengevaluasi keamanan dan privasi pengguna aplikasi media sosial terpopuler. Metodologi yang digunakan meliputi analisis kebijakan privasi dan pengaturan keamanan, pengujian penetrasi, dan survei pengguna. Hasil penelitian menunjukkan bahwa beberapa aplikasi media sosial memiliki kebijakan privasi yang kurang jelas dan rendahnya tingkat keamanan. Kesimpulan dari penelitian ini adalah pentingnya meningkatkan kebijakan privasi dan tingkat keamanan pada aplikasi media sosial untuk melindungi privasi dan data pengguna.

Kata kunci: media sosial, keamanan, privasi, pengguna."

ABSTRACT

Abstract ditulis italic (miring) menggunakan bahasa Inggris dengan jarak antar baris 1 spasi dan maksimal 1 halaman. Abstract adalah versi Bahasa Inggris dari intisari. Abstract dapat ditulis dalam beberapa paragraf. Baris pertama paragraph harus menjorok ke dalam sekitar 1 cm. Tidak dsarankan menggunakan mesin penerjemah melainkan tulis ulang.

Keywords : Keyword 1, Keyword 2, Keyword 3, Keyword 4, Keyword 5

BAB I

PENDAHULUAN

1.1 Latar Belakang

Dunia pendidikan tinggi saat ini menuntut mahasiswa untuk bisa belajar secara mandiri. Namun banyak dari mereka belum siap menghadapi tuntutan tersebut. Hal ini menimbulkan kesenjangan yang serius dalam keterampilan Self-Regulated Learning (SRL), yang dapat menghambat keberhasilan belajar dan mengganggu pemerataan akses pendidikan [1, 2]. Tantangan ini justru diperburuk oleh teknologi digital yang menjadi ciri khas pembelajaran masa kini. Meskipun menawarkan fleksibilitas, teknologi tersebut juga menghadirkan distraksi yang signifikan dan menuntut disiplin diri yang lebih besar [3]. Kegagalan institusi dalam memberikan dukungan yang memadai untuk pengembangan SRL bukan hanya masalah akademik, melainkan juga masalah sosial yang berdampak pada prospek karier jangka panjang mahasiswa dan efektivitas sistem pendidikan [1].

Untuk mengatasi kesenjangan SRL, Zimmerman mengemukakan kerangka kerja yang mencakup tahap perencanaan (*textitforethought*), pelaksanaan (*textitperformance*), dan refleksi diri (*textitself-reflection*) di mana SRL merupakan gabungan strategi kognitif, motivasional, dan perilaku [1, 4]. Individu dengan kemampuan regulasi diri yang efektif terbentuk melalui penetapan tujuan, perancangan strategi, pemantauan diri secara aktif, dan evaluasi diri yang bijaksana. Sejumlah besar penelitian mengonfirmasi hubungan kuat antara keterampilan ini dan pencapaian akademik [1, 2]. Namun, pengembangannya kerap terhambat oleh hambatan personal, kontekstual, dan sosial—mulai dari manajemen waktu yang buruk dan distraksi digital hingga rendahnya kesadaran metakognitif—sehingga intervensi terstruktur diperlukan untuk membimbing mahasiswa menjalani proses yang kompleks ini [3, 5].

Metode tradisional yang paling efektif untuk mengembangkan keterampilan ini adalah umpan balik personal dari ahli manusia. Umpan balik berkualitas tinggi yang berfokus tidak hanya pada jumlah tugas yang diselesaikan, tetapi juga pada proses belajar dan regulasi diri mahasiswa merupakan katalis kuat untuk perbaikan [15, 16] [6, 7]. Namun standar emas ini menghadapi krisis skalabilitas dengan meningkatnya jumlah mahasiswa dan kebutuhan dukungan yang beragam membuat model bimbingan satu-satu sulit diterapkan dalam skala besar. Hal ini diperkuat dengan studi-studi yang dilakukan dan menunjukkan bahwa tekanan anggaran mendorong ukuran kelas yang semakin besar, sehingga interaksi personal antara dosen dan mahasiswa sangat terbatas [8, 9]. Ini menciptakan kebutuhan mendesak akan solusi inovatif berbasis teknologi yang dapat melengkapi keahlian manusia.

Salah satu teknologi tersebut adalah penggunaan Large Language Models (LLM). LLM dapat digunakan sebagai tutor virtual yang mampu memberi umpan balik yang cepat dan konsisten. LLM telah digunakan di berbagai penelitian dalam konteks pendidikan dengan bermacam terobosan. Temuan terbaru memperlihatkan bahwa ketika LLM digunakan untuk penilaian dan umpan balik, *artificial intelligence* (AI) cenderung murah hati dalam memberikan nilai sehingga nilainya lebih tinggi daripada nilai yang diberikan ahli manusia. Sementara itu penilaian antar rekan dan dosen lebih rendah dan sesuai dengan performa mahasiswa. Umpan balik AI sering terstruktur, tetapi tetap memerlukan pengawasan ahli manusia agar relevan secara pedagogis sehingga pendekatan hibrida (penggabungan AI dan ahli manusia) direkomendasikan [10]. Pendekatan hibrida direkomendasikan karena dapat menggabungkan kecepatan dan skalabilitas AI dengan penguasaan konteks, pemahaman empati dan nuansa pedagogis dari penilaian manusia [10]. Anjuran penulis terhadap pendekatan hibrida menegaskan bahwa respons LLM masih membutuhkan pengawasan pakar agar relevan. Di sisi lain, upaya standarisasi evaluasi tutor-AI menunjukkan bahwa meskipun model terkini seperti GPT-4 kuat dalam menjawab pertanyaan, mereka cenderung terlalu cepat memberikan jawaban kepada mahasiswa dan kurang menuntun proses. Fakta ini menunjukkan bahwa LLM belum ideal sebagai tutor dan belum dapat menggantikan ahli manusia dalam penilaian kualitas pedagogis [11].

Oleh karena itu, skripsi ini berfokus pada peningkatan kualitas performa respons LLM sebagai pemberi umpan balik agar mendekati mutu ahli manusia secara kuantitatif dan kualitatif. Peningkatan dilakukan melalui kerangka evaluasi hibrida yang menggabungkan acuan ahli manusia sebagai *ground truth* dan penilaian kualitas yang menilai dimensi-dimensi pedagogis relevan [10, 11]. Untuk mendukung proses pengukuran tersebut, skripsi ini memanfaatkan jejak proses belajar yang terstruktur. Data pembelajaran mahasiswa yang dipresentasikan ke dalam bentuk Papan Kanban digital dipakai sebatas alat operasional untuk menata dan mengekstraksi data proses (perpindahan kartu, checklist, waktu pengerjaan) yang dapat dipetakan ke fase SRL (perencanaan–monitoring–refleksi) [12]. Penggunaan Papan Kanban hanya sebagai salah satu instrumen pengaya data agar evaluasi dan perbaikan kualitas umpan balik LLM bisa berjalan lebih terarah. Fokus kebaruan terletak pada rangkaian evaluasi dan perbaikan *prompt* LLM yang menargetkan tiga fungsi pedagogis utama yakni feedback kognitif, dukungan motivasional, dan dukungan apresiatif dengan tujuan praktis yaitu menyamai kualitas ahli manusia pada metrik kuantitatif sekaligus memenuhi standar kualitas kualitatif [6, 7, 10, 11]. Dengan membandingkan keluaran LLM dan pakar ahli, studi ini menilai peningkatan performa LLM agar dapat menjadi pertimbangan dalam penggunaan sebagai agen pedagogis pelengkap mahasiswa yang masif dan personal.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dipaparkan, penelitian ini mengidentifikasi Secara spesifik, penelitian ini berfokus pada pengembangan umpan balik yang menargetkan tiga fungsi pedagogis: *feedback*, *motivation*, dan *appreciation*. Rumusan masalah yang diangkat dalam penelitian ini adalah sebagai berikut:

1. Bagaimana strategi optimal untuk meningkatkan performa LLM dalam menghasilkan umpan balik yang mencakup ketiga fungsi pedagogis tersebut?
2. Bagaimana kualitas umpan balik yang dihasilkan LLM terhadap proses SRL mahasiswa dibandingkan dengan umpan balik ahli manusia?
3. Bagaimana pengaruh besar ukuran model LLM terhadap kualitas umpan balik yang dihasilkan?

1.3 Tujuan Penelitian

Tujuan penelitian ini adalah sebagai berikut:

1. Merancang skema *prompt* serta memberikan metode *learning analytics* yang efektif untuk meningkatkan kualitas umpan balik LLM agar mendekati ahli manusia.
2. Mengevaluasi secara empiris apakah adanya peningkatan kualitas umpan balik LLM atas proses SRL mahasiswa baik secara kuantitatif maupun kualitatif dengan umpan balik ahli manusia.
3. Melihat pengaruh besar ukuran model LLM terhadap kualitas umpan balik yang dihasilkan.

1.4 Batasan Penelitian

Penelitian ini berfokus untuk melihat peningkatan performa dari LLM ketika diberikan perlakuan berupa metode *context engineering* yang memanfaatkan data *learning analytics*. Upaya untuk menyederhanakan proses ini dilakukan dengan cara memberikan batasan penelitian yang jelas dan ditetapkan oleh peneliti dengan mempertimbangkan aspek-aspek yang relevan. Berikut adalah batasan-batasan tersebut:

1.4.1 Objek Penelitian

- (a) Ruang lingkup mencakup analisis kemampuan model bahasa besar pada tugas berbahasa Indonesia dengan konteks pembelajaran berbasis papan Kanban.
- (b) Data pembelajaran yang digunakan berupa data sintetis yang disusun untuk memicu respons dari model dan dari pakar manusia. Pemakaian data sintetis dinilai memadai karena keterbatasan waktu pengambilan data ril serta karena fokus utama penelitian adalah perbandingan kualitas respons model dan respons pakar.

- (c) Model yang diteliti merupakan model yang tersedia secara *open source* dan diakses melalui Groq API. Dua model yang digunakan adalah Llama 3.1 8B Instant 128k dan Llama 3.3 70B Versatile 128k.

1.4.2 Metode Penelitian

Penelitian menggunakan rancangan eksperimen komparatif tanpa pelatihan ulang pada model. Seluruh prosedur mengandalkan rekayasa konteks yang mencakup perancangan templat *prompt* penetapan instruksi dan penyediaan konteks dari data Kanban berbahasa Indonesia. Langkah penelitian meliputi penyusunan skenario soal dan konteks berbasis Kanban pembuatan respons oleh kedua model melalui Groq API pengumpulan respons pakar manusia pada skenario yang sama serta evaluasi kuantitatif dan penilaian kualitatif oleh pakar. Reprodusibilitas dijaga melalui pencatatan parameter *decoding* penetapan *seed* serta pelaporan lingkungan komputasi.

1.4.3 Waktu dan Tempat Penelitian

Penelitian dilaksanakan pada Juni 2025 hingga Oktober 2025. Proses perancangan dan analisis dilakukan dengan memanfaatkan layanan Groq untuk inferensi model.

1.4.4 Populasi dan Sampel

Populasi penelitian mencakup seluruh respons atas skenario berbahasa Indonesia yang relevan dengan konteks pembelajaran berbasis Kanban. Sampel terdiri atas sekumpulan skenario tugas yang dipilih secara purposif beserta pasangan respons model dan respons pakar manusia. Jumlah *rater* pakar jumlah skenario serta pembagian data ditetapkan agar memadai untuk analisis kuantitatif dan penilaian kualitatif.

1.4.5 Variabel

- (a) **Variabel bebas** mencakup jenis model Llama 3.1 8B Instant 128k dan Llama 3.3 70B Versatile 128k serta strategi rekayasa konteks yang digunakan pada *prompt* dan penyajian *learning analytics* berbasis Kanban .
- (b) **Variabel terikat** adalah kualitas keluaran yang diukur dengan metrik kuantitatif seperti *accuracy precision recall macro F1* serta ukuran kesesuaian teks yang relevan.
- (c) **Variabel kendali** meliputi penggunaan bahasa Indonesia pada seluruh data karakteristik skenario Kanban, parameter seperti *temperature* dan *top-p random seed* serta prosedur evaluasi yang seragam.

1.4.6 Keterbatasan Penelitian

1. Data pembelajaran berbasis papan Kanban yang digunakan merupakan data sintetis yang disusun untuk keperluan pengambilan data dari pakar manusia dan model. Da-

ta sintetis dipilih karena keterbatasan waktu pengambilan data ril serta karena fokus penelitian berada pada perbandingan respons.

2. Penelitian hanya menggunakan model yang telah tersedia dan bersifat *open source* guna efisiensi waktu dan biaya.
3. Model diakses melalui Groq API sebagai sarana inferensi.
4. Model yang digunakan terbatas pada dua model yaitu Llama 3.1 8B Instant 128k dan Llama 3.3 70B Versatile 128k.
5. Tidak dilakukan pelatihan ulang pada model sehingga seluruh prosedur bersifat rekayasa konteks.
6. Seluruh kegiatan dilakukan pada Juni 2025 hingga Oktober 2025.
7. Analisis difokuskan pada kemampuan berbahasa Indonesia sehingga seluruh data yang digunakan berbahasa Indonesia.
8. Sumber data utama berasal dari keluaran model dan respons pakar manusia yang dianalisis sedangkan data pembelajaran Kanban berfungsi sebagai pendukung penyusunan konteks.
9. Luaran penelitian berfokus pada identifikasi metode optimisasi yang paling baik berdasarkan pengukuran kuantitatif yang selanjutnya diuji kepada pakar manusia untuk penilaian kualitatif.

1.5 Manfaat Penelitian

Penelitian ini menawarkan wawasan bagaimana penerapan LLM sebagai pendamping dalam proses SRL bagi mahasiswa bisa menjadi alternatif yang efektif dalam meningkatkan kualitas pembelajaran. Selain itu, penelitian ini juga dapat menjadi dasar untuk penelitian lebih lanjut untuk peningkatan performa respons LLM terutama pada bidang pedagogis sehingga seiring berjalannya waktu dapat memberikan hasil yang lebih optimal.

1.6 Sistematika Penulisan

Sistematika penulisan berisi pembahasan apa yang akan ditulis di setiap bab. Sistematika pada umumnya berupa paragraf yang setiap paragraf mencerminkan bahasan setiap Bab. Contoh:

Bab I membahas tentang pendahuluan yang berisi latar belakang, perumusan masalah dan tujuan penelitian.

Bab II berisi tentang metodologi penelitian yang terdiri dari desain penelitian, sumber data, Teknik pengumpulan data dan Teknik analisis data.

Dan seterusnya.

BAB II

TINJAUAN PUSTAKA DAN DASAR TEORI

Bab ini menyajikan tinjauan komprehensif yang disusun secara sistematis untuk membangun landasan penelitian. Tinjauan dimulai dengan analisis terhadap penelitian-penelitian terdahulu, dilanjutkan dengan komparasi kritis, dan diakhiri dengan pembahasan dasar teori yang relevan.

Pertama, dibahas terlebih dahulu kumpulan penelitian sebelumnya yang berfokus pada upaya meningkatkan kualitas umpan balik dari Large Language Models (LLMs) dalam konteks pendidikan. Termasuk di dalamnya adalah berbagai strategi optimasi yang telah diusulkan serta metode-metode evaluasi yang digunakan untuk mengukur efektivitas umpan balik yang dihasilkan. Kedua, dilakukan komparasi mendalam terhadap penelitian-penelitian tersebut untuk mengidentifikasi kesenjangan yang ada di dalam literatur. Analisis komparatif ini membantu memetakan posisi penelitian ini dalam lanskap penelitian yang lebih luas. Ketiga, dibahas metode-metode evaluasi yang telah digunakan dalam studi-studi sebelumnya. Hal ini mencakup evaluasi baik dari segi kuantitatif maupun kualitatif. Keempat, berdasarkan identifikasi kesenjangan tersebut, disajikan dasar teori yang melandasi penelitian ini. Pembahasan mencakup peran LLM dalam dunia pendidikan, kerangka teoritis Self-Regulated Learning (SRL), konsep Prompt Engineering dan Context Engineering, pendekatan Learning Analytics beserta metode-metode analitik yang relevan untuk konteks pembelajaran, serta metode evaluasi yang dapat digunakan di dalam penelitian ini.

Melalui struktur ini, bab ini tidak hanya memberikan fondasi teoretis yang kuat, tetapi juga memperjelas kontribusi penelitian dalam mengisi celah yang teridentifikasi dari penelitian-penelitian sebelumnya.

2.1 Tinjauan Pustaka

Bagian ini menyajikan tinjauan terhadap penelitian-penelitian sebelumnya yang menjadi landasan bagi pengembangan sistem umpan balik LLM dalam pendidikan. Tinjauan ini mencakup dua fokus utama yaitu strategi peningkatan kualitas umpan balik LLM dan metode evaluasi yang digunakan.

Melalui analisis komparatif terhadap berbagai pendekatan yang ada, diidentifikasi pola, tren, dan kesenjangan dalam literatur. Pemetaan ini tidak hanya memberikan konteks bagi penelitian ini, tetapi juga memperjelas kontribusi unik yang ditawarkan dalam mengatasi keterbatasan-keterbatasan yang masih ada pada penelitian-penelitian sebelumnya.

2.1.1 Peningkatan Umpan Balik LLM dengan *Direct Preference Optimization* (DPO)

Para penulis yaitu, Juliette Woodrow, Sanmi Koyejo, dan Chris Piech (Stanford), mengajukan makalah “*Improving Generative AI Student Feedback: Direct Preference Optimization with Teachers in the Loop*” pada EDM 2025 [13]. Mereka memotret persoalan klasik pada umpan balik berbasis LLM yaitu, walau model mampu menghasilkan teks yang koheren, gaya, ketelitian, dan terminologi, sering kali tidak selaras dengan preferensi pengajar dan konteks spesifik mata kuliah. Penulis menyebutnya “*Feedback Alignment Challenge*” yaitu tantangan menyesuaikan keluaran AI dengan ketelitian, terminologi, dan gaya respons yang diharapkan tim pengajar, sehingga *feedback* terasa “*course-specific*”, bukan *feedback* generik.

Solusi yang ditawarkan adalah Direct Preference Optimization (DPO) dengan guru/TA di dalam loop. Inti dari DPO adalah selama proses penilaian, sistem menampilkan dua kandidat umpan balik AI untuk setiap jawaban mahasiswa, asisten dosen dapat memilih, mengedit, atau menulis ulang. Pilihan ini menjadi pasangan preferensi (menang-kalah) yang kemudian dipakai untuk *fine-tuning* model menggunakan DPO, tanpa perlu membangun reward model seperti pada *Reinforcement Learning from Human Feedback* (RLHF) sehingga alurnya lebih sederhana dan langsung menekan probabilitas keluaran yang disukai dibanding yang tidak disukai. Penulis merancang pipa tiga tahap—koleksi preferensi saat grading, pelatihan DPO antar tugas, dan inferensi untuk tugas berikutnya—mewujudkan sistem self-improving yang semakin selaras dengan preferensi pengajar dari satu *assignment* ke *assignment* berikutnya. Dalam implementasinya, mereka melakukan *fine-tuning* pada model Llama 3.1 8B Instruct menggunakan Hugging Face DPOTrainer. Proses pelatihan berlangsung sekitar tujuh jam pada tiga GPU A6000 dengan rata-rata 1.408 pasangan preferensi per gelombang. Pada tahap noticing, yaitu pengumpulan observasi atas pekerjaan mahasiswa, mereka memanfaatkan rubrik, tes otomatis, dan juga bantuan model LLM lain (GPT-4o). Hasil dari tahap ini kemudian disusun ke dalam prompt terstruktur agar model dapat menghasilkan respons yang lebih tepat sasaran.

Penulis menguji pendekatan ini pada dua skenario. Pertama, studi terkontrol buta yakni digunakan 10 evaluator ahli (staf pengajar) membandingkan umpan balik dari beberapa model pada kiriman mahasiswa yang dianonimkan di mana mereka tidak mengetahui identitas model dan menilai preferensi, *Insightfulness*, serta *Naturalness*. Kedua, deployment nyata di dua penyelenggaraan mata kuliah universitas besar (≈ 330 dan ≈ 289 mahasiswa; >10 *teaching assistants* (TA) per semester), di mana setiap generasi model dilatih dari preferensi tugas sebelumnya lalu dipakai pada tugas berikutnya. Penulis juga mencatat potensi *confounding variable* dalam deployment (mis. setiap soal dinilai oleh satu TA sehingga preferensi pribadi bisa memengaruhi pemilihan *feedback*; variasi soal

antar generasi menyulitkan isolasi efek model).

Untuk evaluasi, penulis memadukan penilaian manusia secara buta dan penilaian otomatis berbasis *critic model*. Pada sisi otomatis, mereka mengadopsi kerangka Scarlatos et al. [14] guna menilai lima dimensi yaitu, *Correctness*, *Not-revealing-the-answer*, *Suggestions*, *Positive tone*, serta skor total. Kerangka ini juga harusnya memuat “*diagnosing errors*” namun tidak difokuskan di hasil tabel karena tidak sejalan dengan tugas utama studi ini. Di luar itu, penulis memperkenalkan *critic model* baru dengan tiga dimensi yaitu *Assertiveness*, *Accuracy*, dan *Helpfulness* yang lebih umum untuk berbagai tipe jawaban (benar/salah). Kedua kerangka ini menggunakan GPT-4o sebagai evaluator otomatis; prompts dan kode penilaian dibuka untuk umum.

Hasil studi terkontrol menunjukkan bahwa model DPO lebih disukai dibanding baseline kuat GPT-4o dengan persentase 56,8% berbanding 40,2% sementara 3% responden memilih netral. Pada metrik *Insightfulness*, DPO juga memperoleh skor lebih tinggi secara signifikan dibanding GPT-4o (rata-rata 3,13 berbanding 2,79) maupun model SFT (2,82). Sebaliknya, pada metrik *Naturalness* tidak ditemukan perbedaan signifikan antar model. Ulasan kualitatif evaluator menegaskan bahwa alasan memilih DPO adalah karena tingkat kekhususan yang lebih tinggi, relevansi terhadap masalah, nada yang mendorong, serta saran yang lebih dapat ditindaklanjuti. Sementara itu, alasan utama dipilihnya GPT-4o adalah kejelasan dan keringkasan, sementara DPO terkadang dinilai terlalu panjang atau agak canggung dari segi gaya.

Pada evaluasi otomatis (kerangka Scarlatos), DPO melampaui GPT-4o pada 4/5 metrik *Correctness* ($0,932 \pm 0,004$ vs $0,862 \pm 0,006$), *Not-revealing-the-answer* ($0,980 \pm 0,002$ vs $0,973 \pm 0,003$), *Positive tone* ($0,989 \pm 0,002$ vs $0,905 \pm 0,005$), dan skor total ($0,724 \pm 0,004$ vs $0,664 \pm 0,005$). Satu-satunya area yang lebih baik pada GPT-4o adalah *Suggestions* (GPT-4o $0,229 \pm 0,007$ vs DPO $0,163 \pm 0,006$), menandakan bahwa keluaran GPT-4o relatif lebih banyak memberi saran eksplisit, sementara DPO unggul dalam ketepatan, menjaga agar tidak membocorkan jawaban, dan mempertahankan nada positif. Di *critic model* baru, sebaran titik menunjukkan klaster DPO lebih konsisten di area akurasi dan helpfulness yang tinggi, sedangkan GPT-4o menampilkan variasi lebih besar berikut klaster pada wilayah rendah akurasi dan rendah helpfulness; secara kuantitatif, umpan balik GPT-4o 8,73 kali lebih mungkin tergolong unhelpful dan 7 kali lebih mungkin inaccurate dibanding DPO (tingkat assertiveness relatif sebanding). Hasil deployment memperlihatkan kelayakan integrasi di kelas besar dan penguatan preferensi TA terhadap DPO dalam seting riil, seraya membuka jalan untuk monitoring otomatis dan eksplorasi awal fairness (gender) yang dapat diperluas ke demografi lain.

2.1.2 Peningkatan *Feedback* LLM dengan *Prompt Engineering* Berbasis *Learning Analytics*

Artikel ini ditulis oleh Lucas Jasper Jacobsen dan Kira Elena Weber (Universitt Hamburg) [15]. Tujuan utamanya ada dua: (1) mengidentifikasi jenis prompt seperti apa yang dibutuhkan agar LLM dapat menghasilkan umpan balik berkualitas tinggi di pendidikan guru, dan (2) membandingkan kualitas umpan balik LLM dengan umpan balik manusia dari kelompok novice (calon guru/preservice teachers) dan expert. Latar masalah yang mereka tekankan: studi empiris yang langsung membandingkan kualitas umpan balik LLM vs manusia masih jarang, dan belum ada manual/kerangka berbasis teori untuk menilai kualitas prompt di konteks pendidikan tinggi, padahal umpan balik berkualitas penting namun sumber daya manusia dan waktu dosen sering terbatas.

Sebagai respons, penulis mengembangkan sebuah manual prompt berbasis teori. Manual ini merangkum kategori-kategori kualitas prompt (mencakup aspek konteks/peran–audiens–medium, misi, kejernihan dan spesifikasi—misalnya format, kekonkretan, spesifik domain, dan logika instruksi). Berdasarkan manual tersebut, mereka merancang tiga versi prompt dengan kualitas rendah–menengah–tinggi. Seluruh generasi umpan balik LLM menggunakan ChatGPT-4, dengan alasan kesesuaian praktik nyata di pendidikan tinggi saat penelitian dilakukan. Gagasannya sederhana: jika mutu prompt dapat disistematiskan, kualitas umpan balik LLM akan lebih konsisten dan bisa dibandingkan secara adil dengan umpan balik manusia.

Tugas yang dipakai berupa learning goal di topik geometri (mengenal segitiga siku-siku dan teorema Pythagoras), lengkap dengan tiga jenis kesalahan yang umum pada perumusan tujuan pembelajaran. Dari tiga kualitas prompt (rendah–menengah–tinggi), penulis menghasilkan 20 umpan balik per prompt (total 60 keluaran LLM) yang kemudian dikodekan oleh tiga coder terlatih. Untuk perbandingan antar penyedia umpan balik, kelompok novice dan expert juga diminta menulis umpan balik dengan menggunakan prompt terbaik (kualitas tinggi), sehingga perbandingan dengan LLM berlangsung pada kondisi instruksi yang optimal. Desain ini memungkinkan isolasi efek “kualitas prompt” pada kinerja LLM sekaligus membangun baseline manusia yang kuat pada kondisi yang sama.

Kualitas umpan balik dianalisis dengan skema kodifikasi yang mencakup dimensi-dimensi pedagogis seperti concreteness (kriteria dan penjelasan), activation (mis. keberadaan pertanyaan pemandu), empathy (nada, penggunaan orang pertama), specificity, serta errors (kesalahan isi). Analisis statistik dilakukan untuk membandingkan LLM vs novice vs expert, dengan pengujian post hoc (Bonferroni) pada beberapa subkategori. Dengan pendekatan ini, penulis tidak hanya menilai “siapa yang lebih baik”, tetapi juga pada aspek apa LLM unggul/kurang, sehingga relevan bagi desain intervensi dan prompting ke depan.

Studi tersebut menunjukkan bahwa hanya prompt berkualitas tinggi yang secara konsisten memicu LLM menghasilkan umpan balik bermutu, sehingga desain prompt menjadi faktor kunci. Dibanding penilai pemula, LLM unggul pada hampir semua subkategori, sementara pada aspek afektif tertentu seperti valensi dan penggunaan orang pertama keunggulan itu tidak tampak. Dibanding pakar, LLM melampaui pada explanation, questions, dan specificity dengan selisih rata-rata (M_{diff}) masing-masing 0,46 (CI 95%: 0,17–0,74; $p < 0,001$), 0,50 (CI 95%: 0,07–0,93; $p < 0,05$), dan 0,96 (CI 95%: 0,52–1,41). Dari sisi efisiensi, ChatGPT-4 mampu menghasilkan sekitar 49 umpan balik dalam waktu yang sama ketika seorang pakar hanya menyusun satu, sehingga LLM berpotensi menjadi alternatif yang skalabel untuk umpan balik berkualitas asalkan perancangan prompt dilakukan dengan tepat.

Penulis menekankan bahwa kualitas prompt harus diajarkan sebagai keterampilan bagi pendidik; manual yang mereka kembangkan dapat dijadikan panduan praktis. Namun, mereka juga mengingatkan keterbatasan—misalnya, halusinasi masih mungkin terjadi pada prompt tertentu, dan persepsi siswa terhadap umpan balik LLM serta transfer ke konteks/mata kuliah lain perlu diteliti lebih lanjut. Dengan kata lain, LLM berpotensi menjadi alternatif berkualitas dan efisien untuk sebagian peran feedback pakar, tetapi desain prompt serta pengawasan pedagogis tetap penting agar umpan balik selaras dengan kebutuhan dan konteks kursus.

2.1.3 *Reinforcement Learning from Human Feedback (RLHF)* untuk Peningkatan Kualitas Respons

Makalah berjudul “*Training language models to follow instructions with human feedback*” ini ditulis oleh Long Ouyang, Jeff Wu, Xu Jiang, Diogo Almeida, Carroll L. Wainwright, dkk. [16], sebagai proyek tim Alignment di OpenAI. Mereka memulai dari pengamatan bahwa memperbesar model bahasa tidak otomatis membuatnya lebih mampu mengikuti niat pengguna; model besar tetap bisa menghasilkan keluaran yang tidak benar, toksik, atau tidak membantu sehingga tidak selaras (misaligned) dengan kebutuhan pengguna. Paper ini menawarkan arah untuk menyelaraskan model dengan niat pengguna pada beragam tugas melalui penyelarasan berbasis umpan balik manusia (fine-tuning with human feedback).

Masalah utamanya adalah ketidakselarasan tujuan pelatihan model bahasa (memprediksi token berikutnya dari web) dengan tujuan penggunaan di dunia nyata (mengikuti instruksi secara membantu, jujur, dan aman). Konsekuensinya, meski bisa diprompt untuk banyak tugas, model sering “membuat fakta”, bias/toksik, atau tidak mengikuti instruksi pengguna. Tim menekankan kebutuhan menyelaraskan model agar helpful-honest-harmless dalam penggunaan sehari-hari. Solusi mereka adalah pipeline RLHF (reinforcement learning from human feedback) dengan tiga tahap: (1) Supervised fine-

tuning (SFT) pada demonstrasi yang ditulis pelabel tentang perilaku model yang diinginkan; (2) pelatihan reward model (RM) dari perbandingan/ranking beberapa keluaran model oleh pelabel; (3) optimasi kebijakan dengan PPO menggunakan RM sebagai sinyal ganjaran. Dataset dikumpulkan dari prompt yang ditulis pelabel dan prompt nyata yang disubmit melalui OpenAI API, sehingga mencerminkan distribusi penggunaan riil. Hasilnya adalah keluarga model InstructGPT.

Secara operasional, RM akhir diinisialisasi dari GPT-3 6B yang telah di-*fine-tune* pada beragam dataset NLP publik; pelatihan RM dilakukan satu *epoch* pada himpunan latih, dan mereka menemukan pelatihan sensitif terhadap jumlah *epoch* (multi-*epoch* cepat *overfit*), sementara *learning rate* tidak terlalu sensitif. Untuk model RLHF, kebijakan diinisialisasi dari model SFT dengan campuran 10% data pra-latih (membantu stabilitas PPO), kemudian dilatih sebanyak 256,000 episode yang mencakup sekitar 31,000 *prompt* unik (setelah penyaringan PII dan deduplikasi). *Batch* per iterasi sebanyak 512 (dipecah menjadi 8 *minibatch* berukuran 64) dan hanya satu *inner epoch* per iterasi; *warmup* pada 10 iterasi pertama, ditambah KL reward mengikuti Stiennon dkk. ($\beta = 0,02$). Rincian ini memperlihatkan setup yang hati-hati agar PPO stabil pada skala 1,3B/6B/175B.

Evaluasi utama menggunakan preferensi manusia di atas distribusi prompt API (mencerminkan use-case riil), serta metrik safety seperti truthfulness dan toxicity (dengan regresi kinerja minimal pada dataset NLP publik). Mereka juga memformalkan tujuan evaluasi selaras dengan kerangka *helpful, honest, harmless*. Visualisasi utama (Figur 1) menampilkan seberapa sering keluaran model lebih disukai dibanding *baseline* GPT-3 175B (SFT).

Secara konsisten, InstructGPT menunjukkan perbaikan yang berarti: model 1.3B mereka lebih sering dipilih manusia dibanding GPT-3 175B, meski 100 kali lebih kecil; ada peningkatan truthfulness dan penurunan toksisitas, sambil menjaga regresi kinerja pada dataset publik tetap minimal. Ini memperlihatkan bahwa penyelarasan dengan umpan balik manusia lebih menentukan daripada sekadar memperbesar parameter. Penulis tetap mencatat keterbatasan model yang masih bisa melakukan kesalahan sederhana. Namun, arah ini dinilai menjanjikan untuk menyelaraskan model dengan niat pengguna di banyak aplikasi.

2.1.4 Regulasi Diri dalam Pendidikan Pemrograman Berbasis LLM dengan *Learning Analytics*

Makalah *Design of AI-Powered Tool for Self-Regulation Support in Programming Education* karya Huiyong Li dan Boxuan Ma dari Kyushu University [17], disajikan pada CHI 2025 Workshop: Augmented Educators and AI, berangkat dari tantangan integrasi AI di pendidikan pemrograman. Banyak alat berbasis LLM yang digunakan mahasiswa berjalan di luar LMS institusional seperti Moodle. Keterpisahan ini memutus keterkait-

an penting dengan konteks perkuliahan, misalnya materi kuliah, rincian soal, dan hasil eksekusi kode. Dampaknya, umpan balik AI sering kurang selaras dengan tujuan dan ruang lingkup mata kuliah, sementara dosen kesulitan menelusuri dampak jangka panjang interaksi mahasiswa dengan AI karena tidak ada jejak data proses yang terkumpul secara sistematis di lingkungan LMS. Pada saat yang sama, sebagian besar riset LLM di domain ini masih memfokuskan akuisisi pengetahuan (seperti generasi kode dan perbaikan bug) sehingga perhatian pada keterampilan SRL seperti perencanaan, monitoring, dan refleksi—relatif kurang. Kekhawatiran lain adalah ketergantungan mahasiswa pada jawaban instan LLM, yang berisiko mengurangi upaya kognitif mendalam yang dibutuhkan pemula untuk membangun kemandirian belajar.

Untuk menjawab celah tersebut, penulis merancang CodeRunner Agent, sebuah alat LLM yang menempel langsung pada ekosistem Moodle. Ia memadukan *lecture viewer* untuk penyajian materi, plugin CodeRunner guna menjalankan dan menilai kode secara otomatis, serta *xAPI logging* ke Learning Record Store sehingga setiap interaksi penting—mulai dari akses materi, percobaan *test case*, hingga permintaan bantuan yang terekam sebagai data proses. Di sisi pedagogis, dukungan SRL dioperasionalkan lewat model PPESS (*Planning, Program creation, Error correction, Self-monitoring, Self-reflection*) yang menurunkan kerangka SRL Zimmerman khusus untuk pembelajaran pemrograman. Dengan pendekatan ini, mahasiswa bisa meminta bantuan sesuai fase SRL yang sedang dijalani, sehingga umpan balik AI tidak acak, melainkan terarah pada kebutuhan regulasi diri saat itu.

Arsitektur CodeRunner Agent menonjol karena mengandalkan dua mesin konteks. Pertama, LACE (*Learning Analytics Context Engine*) yang merangkum metrik keterlibatan dan performansi dari jejak xAPI, seperti waktu yang dihabiskan, frekuensi upaya, tingkat keberhasilan, dan pola kesalahan, lalu menyuntikkan ringkasan tersebut ke *prompt* agar respons LLM memanfaatkan sinyal perilaku SRL nyata. Kedua, KCE (*Knowledge Context Engine*) yang mengelola basis pengetahuan materi kuliah dan latihan, termasuk konsep kunci, prasyarat, tingkat kesulitan, solusi, dan *typical mistakes*, sehingga umpan balik AI selaras kurikulum dan menghindari kebocoran solusi langsung. Dosen dapat mengunggah materi, memutakhirkan basis pengetahuan, dan menyetel parameter konteks agar gaya dan standar umpan balik “menjadi milik mata kuliah” alih-alih generik.

Skenario penggunaan disiapkan untuk kelas pemrograman pemula. Mahasiswa menulis kode di CodeRunner, menjalankan *test cases*, lalu mengakses panel LLM-based Support dan memilih fase SRL yang relevan, misalnya *error correction* ketika menghadapi pesan kesalahan. Permintaan itu seperti identitas latihan, waktu, dan tipe bantuan tercatat sebagai xAPI. Dengan demikian terbentuk alur *end-to-end* dari log pembelajaran, rekayasa konteks, dan respons LLM yang dapat dianalisis untuk penyelarasan pedagogis

dan perbaikan berkelanjutan. Pada tahap ini, makalah berposisi sebagai paper rancangan; penulis merencanakan *pilot* jangka pendek dan eksperimen satu semester lintas beberapa kelas untuk memvalidasi rancangan pada situasi nyata.

Rencana evaluasi memanfaatkan log xAPI untuk melihat perubahan pola percobaan, akurasi, waktu penyelesaian, serta pergeseran strategi SRL ketika dukungan AI diaktifkan. Di luar metrik kuantitatif, penulis juga merencanakan studi kelas guna menilai kualitas, relevansi, dan keamanan pedagogis dari umpan balik LLM, termasuk apakah agen berhasil menjaga SRL (memberi *scaffold* dan pertanyaan pemandu) alih-alih langsung membocorkan jawaban. Pendekatan evaluasi ganda ini menggabungkan analitik perilaku yang objektif dan penilaian pedagogis yang kualitatif ditujukan untuk memastikan alat tidak hanya berfungsi, tetapi bermanfaat bagi kemandirian belajar mahasiswa.

Kontribusi utama paper ini adalah rancangan terintegrasi yang menghubungkan data proses LMS dengan rekayasa konteks dan *scaffolding* SRL berbasis LLM di lingkungan yang benar-benar digunakan kampus. Pemetaan PPESS memberi kerangka operasional agar bantuan AI selalu terkait fase SRL yang tepat, sementara LACE dan KCE memastikan umpan balik kontekstual, selaras kurikulum, dan mendorong kemandirian. Hasil yang ditekankan bukan angka evaluasi, melainkan kelayakan implementasi dan janji untuk kelas besar yang ingin memanfaatkan analitik pembelajaran serta dukungan AI tanpa mengorbankan tujuan pengembangan SRL. Validasi empiris disiapkan melalui *pilot* dan studi semester penuh pada tahap berikutnya.

2.1.5 Dashboard *Real-time* untuk Orkestrasi Kelas Berbasis Kanban

Makalah ini ditulis oleh Sven Strickroth, Melanie Kreidenweis, dan Andreas Götzfried. Mereka menyoroti tantangan nyata pada pembelajaran berbasis tugas di kelas yang bersifat kolaboratif. Ketika beberapa kelompok bergerak pada tugas berbeda dalam waktu yang sama, guru sulit memiliki pandangan menyeluruh tentang status pekerjaan setiap kelompok. Versi awal alat mereka, AgileBoard4Teaching, memang memungkinkan perencanaan dan pelaksanaan dengan papan Kanban per kelompok, tetapi berjalan sepenuhnya *offline*. Guru perlu menyalurkan berkas ke semua siswa, berkeliling memeriksa satu per satu, serta tidak memiliki ringkasan progres kelas secara langsung. Di sisi lain, solusi profesional seperti Trello tidak ideal untuk sekolah karena fitur dan kebijakan privasi. Kebutuhan akan *dashboard real-time* menjadi titik mula rancangan mereka.

Solusi yang diusulkan adalah memperluas AgileBoard4Teaching menjadi sistem berjejaring dengan arsitektur *client-server*. Di sisi guru tersedia *authoring mode* untuk menyusun tugas dan putaran kerja lalu menggandakan papan untuk banyak kelompok. Di sisi siswa tersedia antarmuka Kanban tiga kolom *ToDo*, *In Progress*, *Done* dengan unggah hasil dan permintaan *review*. Inti pembeda adalah *dashboard* guru yang menampilkan kotak ringkas untuk tiap kelompok. Guru dapat melihat jumlah tugas per kolom,

daftar pekerjaan yang menunggu persetujuan, serta sinyal seperti *raise hand*. Tersedia pula pengiriman pesan massal, penetapan *timer*, dan jeda tampilan siswa. Secara teknis, sistem menggunakan Java Servlets dan MariaDB di *backend*, penyimpanan papan sebagai JSON, komunikasi WebSocket untuk sinkronisasi perubahan secara langsung, dan opsi *deployment* via Docker. Mekanisme bergabung untuk siswa menggunakan kode grup agar tetap ramah privasi tanpa akun individual.

Penulis melakukan dua studi evaluasi. Pertama, studi lapangan di kelas TIK tingkat 8 pada topik *CAD* selama tiga minggu dengan dua sesi per minggu. Desainnya *within-subjects*. Satu minggu pertama menggunakan versi *offline*, lalu beralih ke versi jaringan untuk membandingkan pengalaman secara adil. Total 18 siswa dalam 9 kelompok berpartisipasi. Pengumpulan data meliputi observasi langsung, kuesioner fitur, *System Usability Scale* (SUS), wawancara dengan guru, serta analisis *log* interaksi server. Kedua, studi simulasi berupa lokakarya dua jam yang diikuti 18 guru dan *trainer* berpengalaman. Pada lokakarya ini, prototipe dipakai untuk mengorkestrasi kegiatan sehingga peserta mengalami peran “siswa” sekaligus menilai kegunaan alat.

Temuan dari studi lapangan menunjukkan pergeseran preferensi yang kuat ke versi jaringan. Siswa menyatakan alur kerja lebih lancar karena tidak perlu impor dan simpan manual. Fitur seperti *raise hand*, pesan, *timer*, dan *pause* dinilai membantu koordinasi. SUS siswa mencapai 84 yang tergolong sangat baik dan SUS guru 92 yang termasuk kategori pengalaman terbaik. Analisis *log* memperlihatkan intensitas penggunaan tinggi di kelas serta pemanfaatan di luar kelas untuk persiapan dan *review*. Guru menyampaikan bahwa *dashboard* paling berguna untuk meninjau progres banyak kelompok sekaligus dan menyiapkan pertemuan berikutnya. Selama sesi tatap muka, guru tidak selalu menatap *dashboard* karena fokus pada pendampingan langsung. Beberapa saran muncul seperti tampilan yang lebih ringkas untuk kelas besar, pembuatan putaran baru secara cepat, serta statistik waktu dan progres.

Pada studi simulasi, para guru mengapresiasi sinkronisasi waktu nyata dan kemudahan kloning papan untuk banyak kelompok. SUS rata-rata 69 yang tetap masuk kategori baik. Skor ini dipengaruhi kejadian *bug* yang membuat sebagian peserta terkendala masuk ke putaran kedua. Meski demikian, mayoritas peserta menyatakan ingin mencoba alat ini di kelas mereka. Masukan tambahan menyangkut kebutuhan versi terkelola yang stabil dan peningkatan visibilitas opsi komunikasi.

Secara keseluruhan, kontribusi utama makalah adalah desain dan prototipe *dashboard real-time* yang mendukung orkestrasi kelas *agile* berbasis Kanban, lengkap dengan bukti kegunaan dari dua konteks evaluasi. Solusi ini menjembatani celah antara alat perencanaan yang kuat dan kebutuhan pemantauan langsung di kelas. Hasilnya menyiratkan kesiapan untuk dilanjutkan ke tahap produksi dan pengayaan fitur *learning analytics* seperti waktu penyelesaian dan deviasi dari target. Penulis menekankan pen-

tingnya menjaga keseimbangan antara informasi yang cukup dan beban kognitif guru agar *dashboard* tetap dapat dibaca sekejap sekaligus memberi sinyal tindakan yang jelas.

2.1.6 Perbandingan Metode Peningkatan Kualitas Respons LLM

Dapat dilihat, Tabel 2.1 merangkum karya kunci yang mewakili dua arus utama yakni penyesuaian *model* melalui *fine-tuning* misalnya *DPO* dan *RLHF* serta penyesuaian *konteks* melalui *prompt* dan *context engineering* serta orkestrasi kelas berbasis Kanban.

Tabel 2.1. Komparasi penelitian peningkatan kualitas umpan balik LLM di pendidikan

Studi	Tujuan	Pendekatan Inti	Pelaksanaan Penelitian	Temuan Utama
Woodrow et al. (EDM 2025) [13]	Menyelaraskan gaya dan isi feedback AI dengan preferensi pengajar	<i>Direct Preference Optimization</i> dengan guru/TA di dalam loop. Preferensi menang-kalah dari dua kandidat umpan balik dipakai untuk fine-tuning	Mata kuliah besar. Siklus antartugas. Observasi pekerjaan mahasiswa diringkas ke prompt terstruktur	Studi terkontrol menunjukkan DPO lebih disukai daripada GPT-4o. Skor insight lebih tinggi. <i>Naturalness</i> setara. Evaluasi otomatis juga lebih baik pada empat dari lima dimensi. DPO kadang lebih panjang
Jacobsen & Weber (2025) [15]	Menentukan desain prompt yang memicu feedback LLM berkualitas serta membandingkan LLM dengan manusia	Manual prompt berbasis teori. Tiga kualitas prompt rendah menengah tinggi. LLM dibanding novice dan expert pada tugas geometri	Feedback pada learning goal dan miskonsepsi umum. Semua pihak memakai prompt berkualitas tinggi untuk perbandingan adil	Prompt berkualitas tinggi menghasilkan feedback terbaik. LLM unggul novice di hampir semua subkategori. LLM melampaui expert pada explanation questions specificity. Efisiensi LLM jauh lebih tinggi
Ouyang et al. (InstructGPT, 2022) [16]	Menyelaraskan model agar mengikuti instruksi yang membantu jujur dan aman	<i>RLHF</i> tiga tahap. SFT pada demonstrasi manusia. Reward model dari perbandingan keluaran. PPO untuk mengoptimalkan kebijakan	Berbagai tugas umum. Prompt dari pelabel dan pengguna API. Bukan konteks kursus spesifik	Model kecil yang telah selaras lebih sering dipilih manusia dibanding GPT-3 yang jauh lebih besar. Truthfulness naik dan toksisitas turun. Kerangka selaras tugas umum bukan feedback proses belajar
Li & Ma (CHI 2025 WS) [17]	Merancang alat LLM yang mendukung SRL di programan serta tetap terhubung ke LMS	Rekayasa konteks ganda. LACE untuk ringkasan learning analytics berbasis xAPI. KCE untuk pengetahuan kurikulum. Pemetaan SRL ke PPESS	Moodle dengan CodeRunner. Log xAPI merekam akses materi upaya tes dan bantuan. Fase SRL dipilih pengguna untuk memicu bantuan yang tepat	Paper rancangan. Rencana uji lapangan dan studi satu semester. Menekankan kelayakan integrasi. Fokus pada dukungan SRL dan kontrol kebocoran jawaban. Belum ada angka hasil akhir
Strickroth et al. (2025) [18]	<i>Dashboard real-time</i> untuk orkestrasi kelas Kanban agar guru memantau progres banyak kelompok	Ekstensi AgileBoard4Teaching ke sistem <i>client-server</i> . <i>Dashboard</i> guru. Sinkronisasi <i>WebSocket</i> . Penyimpanan papan JSON. Opsi <i>deployment</i> Docker. Mekanisme kode grup ramah privasi	Kelas TIK tingkat 8 topik CAD. 18 siswa 9 kelompok. Desain <i>within-subjects offline</i> ke jaringan. Lokakarya dua jam dengan 18 guru/trainer. Data observasi kuesioner SUS wawancara dan log server	Preferensi kuat ke versi jaringan. Alur kerja lebih lancar. SUS siswa 84 dan SUS guru 92. Lokakarya SUS 69. Masukan lanjutan meliputi tampilan ringkas kelas besar pembuatan putaran cepat serta statistik waktu dan progres

Tabel 2.1 secara garis besar menjelaskan karya terdahulu yang berpusar ke dua arus utama. Arus pertama berfokus pada penyesuaian *model* melalui *fine-tuning*, misalnya *Direct Preference Optimization* (DPO) [13] dan *Reinforcement Learning from Human Feedback* (RLHF) [16]. Arus kedua menekankan penyesuaian *konteks* melalui perancangan *prompt* dan lingkungan interaksi beserta pemanfaatan data proses pembelajaran, sebagaimana tampak pada [15] dan [17]. Di luar itu, [18] menawarkan orkestrasi kelas berbasis papan Kanban dengan *dashboard real-time* yang kaya sinyal proses sehingga menyediakan landasan operasional untuk memantau dan menindaklanjuti progres kelompok secara langsung.

Dari sisi tujuan intervensi, DPO dan RLHF mengubah parameter model agar keluaran lebih selaras dengan preferensi manusia [13,16]. Sebaliknya, [15] memperlihatkan bahwa *prompt* berkualitas tinggi dapat mendorong kualitas umpan balik tanpa perlu pelatihan ulang, sementara [17] menekankan rekayasa konteks yang ditopang oleh *learning analytics*.

Menariknya, [18] secara tidak langsung menyediakan fitur yang mendukung guru untuk melakukan tugas *learning analytics*. Hal ini dapat dilihat dari kemampuan *dashboard* dalam menampilkan seperti jumlah tugas per kolom, daftar pekerjaan, dan waktu yang dihabiskan oleh siswa pada setiap tugas. Dengan demikian, guru dapat memantau aktivitas pembelajaran secara real-time dan mengambil tindakan yang diperlukan. Penelitian ini memosisikan diri pada jalur kedua, yaitu optimalisasi konteks, sejalan dengan batasan metodologis bahwa tidak dilakukan *fine-tuning*.

Untuk perspektif sumber daya dan kelayakan, pendekatan DPO/RLHF menuntut data preferensi atau label serta infrastruktur pelatihan. Hal tersebut tidak selaras dengan rancangan studi ini yang memanfaatkan model *open source* melalui Groq API tanpa pelatihan ulang. Pendekatan *prompt/context engineering* lebih ringan dijalankan, mudah direplikasi, dan sesuai dengan keterbatasan komputasi maupun waktu. Penelitian [18] memanfaatkan *dashboard real-time* dan jejak proses Kanban yang memungkinkan pemantauan banyak kelompok sekaligus, pengumpulan sinyal proses yang autentik, dan pengambilan tindakan cepat oleh pengajar. Komponen ini belum tercakup pada studi yang berfokus pada *prompt* semata. Beberapa hal yang ada di penelitian Strickroth dkk. dapat dimanfaatkan untuk memperkaya konteks umpan balik LLM.

Kontrol pedagogis tetap dipertahankan melalui penilaian kualitatif yang akan ditinjau langsung oleh pakar pendidikan. Keluaran LLM yang berupa *feedback* ditinjau dan dinilai melalui rubrik penilaian yang telah dipersiapkan. Dengan cara ini, kualitas pedagogis tetap terjaga dan sejalan dengan kebutuhan untuk dilakukan penilaian oleh pakar manusia.

Dengan demikian, penelitian ini mendapatkan celah perbaruan di mana pendekat-

an *prompting* dan *context engineering* yang murah sumber daya dan biaya digabungkan dengan *learning analytics* yang berasal dari papan Kanban untuk memperkaya konteks umpan balik LLM. Kombinasi tersebut belum pernah dibahas sebelumnya oleh literatur yang ada. Celah penelitian ini akan dievaluasi secara kuantitatif untuk mengukur kualitas keluaran LLM terhadap acuan pakar serta secara kualitatif melalui penilaian pakar pendidikan untuk menjaga kualitas pedagogis. Melalui batasan-batasan yang telah ditetapkan, rancangan ini menawarkan solusi praktis dan efektif. Pilihan metodologis ini konsisten dengan batasan penelitian, memanfaatkan data proses untuk memperkuat konteks *prompt*, menjaga peran guru sebagai pengarah kualitas, dan menghadirkan kesiapan implementasi di lapangan.

2.1.7 Metode Evaluasi Umpan Balik LLM dalam Pendidikan

Evaluasi kualitas teks yang dihasilkan oleh *Large Language Models* (LLM) merupakan tantangan kompleks yang memerlukan pendekatan multi-aspek. Secara historis, metrik evaluasi didominasi oleh metode berbasis tumpang tindih leksikal seperti BLEU dan ROUGE. Metrik ini bekerja dengan menghitung kesamaan *n-gram* (urutan kata) antara teks yang dihasilkan dan teks referensi (*ground truth*)[1]. Namun, ketergantungan pada pencocokan permukaan ini memiliki kelemahan signifikan; metrik ini gagal memahami kesetaraan makna (semantik) dan sering kali memberikan skor rendah pada parafrasa yang valid, ketika dua kalimat memiliki arti yang sama tetapi menggunakan kosakata yang berbeda[2]. Berbagai studi telah menunjukkan bahwa korelasi metrik ini dengan penilaian manusia sering kali rendah, terutama dalam tugas-tugas yang menuntut keragaman ekspresi seperti simplifikasi teks atau dialog[3].

Keterbatasan ini mendorong pergeseran paradigma menuju metrik evaluasi yang mampu menangkap kesamaan semantik. Pada penelitian ini, evaluasi yang akan dilakukan dibatasi pada dua faktor yaitu kesamaan semantik dan kualitas relevansinya terhadap konteks. Untuk pendekatan kesamaan semantik, metrik berbasis *embedding* kontekstual seperti BERTScore dan BARTScore menjadi pilihan utama [1][2]. Lalu, penilai (dikenal sebagai *LLM-as-a-Judge*) dapat memberikan penilaian relevansinya terhadap kriteria tertentu untuk teks yang dihasilkan. Metrik berbasis *embedding* seperti BERTScore dan BARTScore tidak lagi membandingkan kata secara literal, melainkan representasi vektornya yang kaya akan konteks, sehingga lebih selaras dengan penilaian kualitas oleh manusia[1]. Di sisi lain, pendekatan *LLM-as-a-Judge* memanfaatkan kemampuan pemahaman bahasa dari model canggih seperti GPT-4 untuk memberikan penilaian kualitatif yang bernuansa, meniru proses evaluasi oleh manusia pada skala yang lebih besar[4].

Dalam paper ini [1] Metrik ini bekerja dengan mengubah setiap token dalam kalimat kandidat dan referensi menjadi *embedding* kontekstual menggunakan model BERT. Selanjutnya, kesamaan makna diukur dengan menghitung kemiripan kosinus (*cosine si-*

ilarity) antara setiap pasang token, lalu secara serakah (*greedy matching*) mencocokkan token dari satu kalimat ke token yang paling mirip di kalimat lainnya untuk menghasilkan skor presisi, *recall*, dan F1. Keunggulan utamanya adalah kemampuannya untuk mengenali sinonim dan parafrasa, sehingga memberikan evaluasi yang lebih kuat terhadap kesetaraan makna dibandingkan metrik *n-gram*.

Sementara itu, BARTScore [1] menawarkan filosofi yang berbeda secara fundamental dengan membingkai evaluasi sebagai "tugas generasi teks". Alih-alih hanya membandingkan *embedding*, BARTScore menggunakan model seq2seq (sequence-to-sequence) untuk menghitung probabilitas sebuah teks dihasilkan dari teks lainnya. Misalnya, untuk mengukur presisi, ia akan menghitung seberapa mungkin model BART akan menghasilkan teks referensi jika diberi teks kandidat sebagai sumber. Skor yang lebih tinggi (*log-probability* yang lebih mendekati nol) menandakan kualitas yang lebih baik. Fleksibilitas ini memungkinkan BARTScore untuk menilai berbagai aspek seperti kefasihan, koherensi, dan bahkan faktualitas hanya dengan mengubah input sumber dan targetnya.

Untuk mengatasi subjektivitas dan bias dalam pendekatan *LLM-as-a-Judge*, kerangka kerja terstruktur seperti G-Eval telah dikembangkan [1]. G-Eval menggunakan LLM penilai (misalnya, GPT-4) tetapi memandunya dengan penalaran *Chain-of-Thought* (CoT). Alih-alih hanya meminta skor akhir, G-Eval terlebih dahulu menginstruksikan LLM untuk menghasilkan 1 langkah-langkah evaluasi berdasarkan kriteria yang diberikan pengguna, lalu mengikuti langkah-langkah tersebut untuk sampai pada sebuah skor. Pendekatan ini membuat proses penilaian lebih transparan, dapat dikontrol, dan terbukti mencapai korelasi yang lebih tinggi dengan penilaian manusia dibandingkan metode LLM-penilai yang tidak terstruktur. Pada penelitian [2], dilakukan juga penilaian umpan balik kode pemrograman yang dihasilkan oleh model open-source. Seperti G-Eval, penilai LLM diminta untuk menilai berdasarkan kriteria yang telah ditentukan. Sehingga dapat disimpulkan bahwa pendekatan *LLM-as-a-Judge* yang terstruktur dengan kriteria jelas meningkatkan keandalan penilaian.

2.1.8 Evaluasi Ahli Manusia

Untuk menilai kualitas umpan balik yang dihasilkan oleh LLM secara komprehensif, evaluasi manusia yang terstruktur menjadi standar emas, karena mampu menangkap nuansa pedagogis dan psikologis yang terlewat oleh metrik otomatis. Penilaian ini dapat ditingkatkan menggunakan rubrik yang didasarkan pada prinsip-prinsip teoretis yang telah mapan. Mengacu pada Shute (2008), kriteria fundamental yang harus dinilai adalah apakah umpan balik LLM bersifat jelas, akurat, dan relevan dengan konteks tugas yang diberikan. Penilai manusia dapat secara langsung memverifikasi apakah instruksi yang diberikan spesifik dan apakah informasi faktualnya benar.

Selanjutnya, dengan mengadopsi model dari Hattie dan Timperley (2007), penilai dapat mengevaluasi kelengkapan struktural umpan balik tersebut. Rubrik penilaian akan memeriksa apakah LLM mampu menjawab tiga pertanyaan kunci: (1) Feed Up: Apakah umpan balik mengklarifikasi tujuan akhir (Where am I going?); (2) Feedback: Apakah ia memberikan analisis tentang kinerja saat ini terhadap tujuan tersebut (How am I going?); dan (3) Feed Forward: Apakah ia menawarkan saran yang konkret dan dapat ditindaklanjuti untuk perbaikan (Where to next?).

Terakhir, dampak motivasional dari umpan balik, sebuah aspek yang sulit diukur secara otomatis, dapat dinilai melalui lensa Self-Determination Theory (Deci & Ryan, 1985) dan prinsip penguatan positif (Fredrickson, 2001). Penilai manusia akan mengevaluasi apakah bahasa yang digunakan LLM mendukung otonomi (memberikan pilihan, bukan mendikte), menumbuhkan rasa kompetensi (mengakui usaha dan kemajuan), dan menggunakan nada yang apresiatif untuk meningkatkan self-efficacy. Dengan demikian, evaluasi manusia yang berlandaskan teori ini tidak hanya mengukur kebenaran teknis dari umpan balik, tetapi juga efektivitasnya dalam memotivasi dan memberdayakan pengguna secara konstruktif.

Secara keseluruhan, penilaian kualitas umpan balik LLM oleh ahli manusia menggabungkan evaluasi objektif berdasarkan rubrik yang dikategorikan menjadi tiga jenis yaitu penilaian secara *feedback*, *motivation support*, dan *appreciation support*. Setiap kategori dievaluasi menggunakan skala likert (1-5) untuk mengukur berbagai dimensi kualitas umpan balik dari perspektif pedagogis dan psikologis. Dengan cara ini, evaluasi manusia menyediakan gambaran yang kaya dan bernuansa tentang efektivitas umpan balik LLM dalam konteks pendidikan.

2.2 Hipotesis Penelitian

Berdasarkan tujuan penelitian yang telah diuraikan dan metode evaluasi kuantitatif (*BERTScore*, *BARTScore*, dan *LLM-as-a-judge*) serta kualitatif (Penilaian skala Likert oleh pakar) yang dipilih, maka hipotesis penelitian ini dirumuskan sebagai berikut:

Hipotesis Efektivitas Metode(*Context Engineering* dengan *Learning Analytics*)

H0₁: Tidak ada perbedaan kualitas antara metode *context engineering* digabungkan dengan LA dan *Baseline*

H1₁: Ada perbedaan kualitas antara metode *context engineering* digabungkan dengan LA dan *Baseline*

Hipotesis Kualitas Metode (*Context Engineering* dengan *Learning Analytics*) Menurut Ahli Manusia

H0₂: Hasil evaluasi dari ahli manusia menunjukkan metode *context engineering* digabungkan dengan LA tidak layak digunakan

H1₂: Hasil evaluasi dari ahli manusia menunjukkan metode *context engineering* digabungkan dengan LA layak digunakan

Hipotesis Pengaruh Ukuran Model

H0₃: Ukuran model tidak berpengaruh pada kualitas

H1₃: Model Bahasa berukuran lebih besar menghasilkan kualitas lebih tinggi

2.3 Dasar Teori

Bagian ini menyajikan landasan konseptual yang menopang rancangan dan metode penelitian. Pertama, diringkas prinsip *Self-Regulated Learning* (SRL)—termasuk peran umpan balik dalam fase perencanaan, pelaksanaan, dan refleksi—serta kaitannya dengan tiga fungsi pedagogis yang dinilai pada penelitian ini: *feedback* kognitif, dukungan motivasional, dan dukungan apresiatif. Kedua, diuraikan konsep *prompt engineering* sebagai praktik perumusan instruksi bagi LLM dan *context engineering* sebagai pengelolaan muatan informasi/konteks yang dipasok ke model; perbedaan keduanya serta implikasinya bagi desain tutor-AI juga dibahas untuk menegaskan peran masing-masing. Ketiga, dipaparkan *learning analytics* (LA) yang mencakup definisi, ruang lingkup, dan jenis data proses (mis. log LMS, urutan aktivitas, *time-on-task*, perpindahan tugas, beban/WIP) beserta cara ekstraksinya. Pada akhirnya, dijelaskan bagaimana sinyal-sinyal LA tersebut direkayasa menjadi konteks *prompt* agar LLM menghasilkan umpan balik yang lebih terstruktur, *actionable*, koheren, dan berempati, sekaligus menekankan kebutuhan evaluasi hibrida (acuan pakar + metrik kuantitatif–kualitatif) untuk memastikan kualitas respons LLM mendekati mutu pakar manusia.

2.3.1 *Self Regulated Learning* (SRL) dan Peran Umpan Balik

Self-Regulated Learning (SRL) merujuk pada kapasitas mahasiswa untuk secara aktif merencanakan, memantau, dan mengevaluasi proses belajarnya sendiri, melalui siklus perencanaan, pelaksanaan, dan refleksi yang saling terkait [1], [2]. Pada fase perencanaan, mahasiswa menetapkan tujuan serta strategi. Lalu, pada fase pelaksanaan, mereka menerapkan strategi kognitif dan melakukan pemantauan metakognitif (mis.

manajemen waktu, monitoring pemahaman). Kemudian, pada fase refleksi, mahasiswa melibatkan penilaian diri terhadap kemajuan dan hasil untuk menyesuaikan strategi berikutnya [1], [2]. Berbekal mekanisme siklikal ini, SRL menggabungkan dimensi kognitif, motivasional, dan perilaku yang bersama-sama menentukan kualitas belajar mandiri di pendidikan tinggi [1], [2].

Di konteks pendidikan tinggi, SRL berperan krusial karena lingkungan belajar menuntut kemandirian, terutama pada pembelajaran digital (daring atau campuran) yang memerlukan disiplin diri lebih tinggi serta kemampuan mengelola distraksi. Kajian terkini menunjukkan bahwa SRL berkorelasi dengan kinerja akademik, keterlibatan, dan ketahanan studi. Sebaliknya, kekurangan SRL di pembelajaran daring atau campuran dapat memperlebar kesenjangan capaian yang diperoleh mahasiswa [3]. Dengan demikian, pengembangan kompetensi SRL termasuk penetapan tujuan, strategi, monitoring, dan refleksi—menjadi prasyarat penting untuk keberhasilan mahasiswa di ekosistem belajar modern [2], [3].

Kemajuan *Large Language Models* (LLM) membuka peluang penyediaan bimbingan pengarahan (*feedback*) yang personal, cepat, dan interaktif untuk menopang SRL. Studi eksperimental di konteks pemrograman (pendidikan tinggi) menunjukkan *feedback* berbasis LLM dapat meningkatkan strategi metakognitif, motivasi, dan performa mahasiswa dibanding kondisi tanpa *feedback* dari AI [6]. Riset pengembangan menunjukkan asisten belajar berbasis LLM yang dirancang dengan prinsip-prinsip pedagogis mampu menyajikan *feedback* yang scaffolded (bertahap), memicu *self-explanation*, dan menuntun refleksi [7]. Dengan dialog natural, LLM dapat membantu ketiga fase SRL yaitu, (i) perencanaan (menyusun tujuan dan rencana aksi), (ii) pelaksanaan (memberi hint/pertanyaan pemandu saat buntu), dan (iii) refleksi (merangkum kekuatan-kelemahan dan langkah tindak lanjut) [6], [7]. Agar efektif dan aman, *feedback* LLM sebaiknya disejajarkan dengan praktik pedagogis berbasis bukti (fokus pada proses, *actionability*, dan nada empatik) dan dievaluasi kualitasnya secara sistematis.

Dalam kerangka SRL, efektivitas umpan balik dipahami melalui fokus bertingkat pada *task*, *process*, dan *self-regulation*, serta dampaknya pada keyakinan diri dan ketekunan belajar [4]. Menurut artikel jurnal ini [4], disebutkan bahwa *feedback* kognitif efektif bila spesifik, mendiagnosis miskonsepsi, dan menyarankan langkah perbaikan yang terstruktur. Lalu, **dukungan motivasional** menjaga *self-efficacy* (keyakinan diri), menetapkan tujuan yang menantang namun realistis, serta mendorong komitmen rencana aksi. Terakhir, pemberian apresiasi yang spesifik terhadap usaha/progres dapat berkontribusi pada afek positif dan keterlibatan yang disertai arahan perbaikan agar tidak menjadi pujian kosong [4]. Oleh karena itu, penelitian ini menilai kualitas respons LLM pada ketiga fungsi tersebut karena ketiganya (*feedback*, *motivation*, dan *appreciation*) secara langsung mendukung fase perencanaan–pelaksanaan–refleksi dalam SRL [1], [2], [4].

2.3.2 *Prompt Engineering*

Prompt engineering merupakan bidang kajian baru yang berfokus pada perancangan, penyempurnaan, dan implementasi prompt atau instruksi yang diberikan kepada model bahasa besar LLM untuk mengarahkan keluaran model tersebut [1]. Dengan kata lain, *prompt engineering* adalah “mekanisme pengendali” (*steering mechanism*) yang digunakan pengguna kecerdasan buatan generatif untuk menyusun *prompt* sedemikian rupa agar menghasilkan keluaran yang diinginkan [2]. Konsep ini menjadi semakin penting seiring munculnya LLM populer seperti ChatGPT, karena kemampuan merancang *prompt* yang efektif memungkinkan pengguna untuk mendapatkan jawaban yang lebih akurat, relevan, dan sesuai kebutuhan dari model AI [1].

Sebagai suatu keterampilan, *prompt engineering* dapat dipelajari dan dikembangkan. Para peneliti menuliskan bahwa *prompt engineering* sudah mulai diajarkan secara formal, misalnya di perguruan tinggi, karena dianggap sebagai *skillset* baru yang dibutuhkan di era integrasi AI [2]. Intinya, *prompt engineering* menekankan seni merumuskan pertanyaan atau perintah dalam bahasa natural sedemikian rupa sehingga model AI memahami konteks tugas dan memberikan respon optimal sesuai keinginan pengguna.

2.3.3 *Context Engineering*

Context engineering adalah disiplin yang berkembang untuk melengkapi dan melampaui *prompt engineering*. Jika *prompt engineering* berfokus pada apa yang dikatakan kepada model pada satu waktu, maka *context engineering* berfokus pada apa saja yang diketahui model saat *prompt* tersebut diberikan. Menurut survei terbaru, *context engineering* didefinisikan sebagai disiplin formal yang melampaui sekadar perancangan *prompt* dan mencakup optimisasi sistematis dari informasi konteks yang disuplai ke LLM [3]. Kinerja LLM sangat ditentukan oleh informasi kontekstual yang diberikan selama inferensi, sehingga *context engineering* berusaha memastikan model menerima *payload* informasi yang paling relevan dan berguna.

Dalam praktiknya, *context engineering* mencakup teknik-teknik seperti *retrieval-augmented generation* (RAG) untuk mengambil pengetahuan eksternal, penggunaan *memory systems* agar model mengingat interaksi sebelumnya, integrasi alat tambahan (*tool-integrated reasoning*), hingga koordinasi *multi-agent* AI [3]. Berbeda dengan *prompt engineering* yang hanya memanipulasi teks *prompt*-nya, *context engineering* mengatur lingkungan dan konteks bagi model, misalnya mengumpulkan data pendukung, menyusun instruksi sistem, mendefinisikan aturan, serta memasukkan memori atau informasi historis sebelum *prompt* diberikan. Tujuannya adalah memaksimalkan kualitas jawaban AI dengan menghadirkan konteks informasi yang optimal dalam batas panjang konteks model tersebut [3]. Dengan demikian, *context engineering* menjadikan AI lebih andal dan *context-aware*, tidak semata-mata mengandalkan satu perintah teks statis.

2.3.4 Perbedaan *Prompt Engineering* dan *Context Engineering*

Meskipun terkait, *prompt engineering* dan *context engineering* memiliki fokus dan pendekatan yang berbeda. Berikut adalah perbedaan utama keduanya [3]:

Tabel 2.2. Perbandingan *Prompt Engineering* dan *Context Engineering*

Aspek	<i>Prompt Engineering</i>	<i>Context Engineering</i>
Lingkup Input	Memandang <i>prompt</i> sebagai teks statis tunggal yang menjadi masukan model.	Memandang konteks secara dinamis dan terstruktur, terdiri dari berbagai komponen informasi (instruksi, pengetahuan, memori, status pengguna) yang dirangkai bersama.
Informasi dan Memori	Informasi bersifat tetap dalam teks <i>prompt</i> ; interaksi biasanya tanpa memori jangka panjang (setiap <i>prompt</i> berdiri sendiri).	Mempertimbangkan komponen memori atau <i>state</i> secara eksplisit, memungkinkan model memanfaatkan informasi dari interaksi sebelumnya (<i>stateful</i>).
Pendekatan Optimasi	Mengandalkan percobaan manual dan iteratif; sering dianggap sebagai suatu seni (<i>art</i>) coba-coba.	Menerapkan pendekatan yang sistematis dan ilmiah (<i>scientific</i>); melibatkan optimisasi di level sistem secara menyeluruh.
Skalabilitas dan Kompleksitas	Cenderung rapuh untuk <i>prompt</i> yang panjang dan kompleks; kurang terstruktur untuk skenario kompleks.	Dirancang untuk mengelola konteks yang besar dan beragam secara modular (segmentasi, hirarki, kompresi).

Secara ringkas, *prompt engineering* menitikberatkan cara merumuskan pertanyaan kepada AI, sedangkan *context engineering* menitikberatkan apa saja informasi pendukung yang disertakan agar AI dapat menjawab dengan lebih baik. *Context engineering* mengubah fokus dari sekadar “seni” merangkai prompt menjadi “ilmu” mengelola informasi untuk optimalisasi sistem AI [3]. Kedua pendekatan ini saling melengkapi: prompt yang baik diperlukan, namun penyajian konteks yang tepat akan semakin meningkatkan kemampuan model dalam menyelesaikan tugas secara andal.

2.3.5 *Learning Analytics* dan Cakupannya

Learning analytics adalah bidang kajian yang memanfaatkan data pendidikan untuk memahami dan meningkatkan proses pembelajaran. Salah satu definisi yang umum dikutip menyatakan bahwa learning analytics merupakan pengukuran, pengumpulan, analisis, dan pelaporan data tentang peserta didik serta konteksnya, dengan tujuan memahami dan mengoptimalkan pembelajaran serta lingkungan di mana pembelajaran tersebut ber-

langsung [4]. Dalam praktiknya, learning analytics mencakup penggunaan berbagai sumber data (misalnya data aktivitas pada Learning Management System, nilai akademik, log interaksi, dan sebagainya) untuk memperoleh wawasan yang dapat ditindaklanjuti oleh pendidik, lembaga, ataupun peserta didik sendiri.

Cakupan learning analytics sangat luas, meliputi beragam skenario dan kepenyertaan pendidikan. Pada tingkat institusi, learning analytics digunakan untuk memonitor kinerja dan retensi mahasiswa. Studi menunjukkan banyak institusi pendidikan menerapkan learning analytics terutama untuk mengidentifikasi mahasiswa berisiko tinggi (misalnya berpotensi drop-out atau gagal studi) dan mengambil langkah intervensi dini guna meningkatkan retensi [4]. Dengan analisis data yang tepat, institusi dapat bersikap lebih preventif daripada reaktif. Model prediktif dapat menandai siswa yang membutuhkan bantuan, sehingga advisor atau dosen dapat memberikan bimbingan sebelum masalah berkembang lebih jauh [4]. Hasilnya, learning analytics berkontribusi pada peningkatan keberhasilan studi melalui intervensi yang tepat sasaran.

Di sisi lain, cakupan learning analytics juga mencakup tingkat kelas dan individu. Bagi pendidik, analitik pembelajaran dapat dipakai untuk mengevaluasi efektivitas pengajaran dan kualitas interaksi di kelas. Misalnya, data pola akses e-learning dan keterlibatan siswa dapat dianalisis untuk mengetahui bagian materi mana yang sulit dipahami atau siapa siswa yang kurang aktif, sehingga pendidik dapat menyesuaikan strategi pengajaran. Bagi pelajar, dashboard learning analytics bisa memberikan umpan balik personal tentang progres belajar mereka, membantu self-reflection dan pengambilan keputusan belajar yang lebih baik.

Sebagai bidang multidisiplin, learning analytics memadukan teknik dari data science, machine learning, dan pendidikan. Penerapannya mencakup analisis deskriptif (melaporkan apa yang terjadi dalam pembelajaran), diagnostik (mengapa sesuatu terjadi), prediktif (memprediksi hasil atau risiko di masa depan, seperti kemungkinan drop-out), hingga preskriptif (memberikan rekomendasi intervensi untuk perbaikan). Dengan kata lain, cakupan learning analytics tidak hanya terbatas pada pelaporan statistik, tetapi juga pengambilan keputusan berbasis data dalam pendidikan.

Secara keseluruhan, learning analytics berperan sebagai alat bantu strategis bagi institusi pendidikan untuk meningkatkan kualitas pembelajaran. Dari meningkatkan outcome belajar siswa, mengoptimalkan penggunaan teknologi pendidikan, menurunkan tingkat putus studi, hingga meningkatkan personalisasi pembelajaran, semuanya termasuk dalam ruang lingkup learning analytics [4]. Dengan dukungan literatur ilmiah yang terus berkembang, learning analytics menawarkan kerangka berbasis data untuk memahami proses belajar mengajar secara lebih mendalam dan mendorong inovasi dalam praktik pendidikan.

2.4 Penelitian Terkait

2.4.1 Studi Menggabungkan LLM dengan Learning Analytics

Beberapa penelitian terbaru telah mengintegrasikan Large Language Models (LLM) dengan teknik learning analytics untuk meningkatkan pemberian umpan balik dan tutoring. Misalnya, TutorLLM (Li dkk., 2024) menggabungkan model knowledge tracing (KT) dengan retrieval-augmented generation (RAG) dalam LLM. TutorLLM memanfaatkan data jejak belajar tiap siswa (dari model KT BERT-based) serta pencarian konteks relevan, sehingga LLM dapat memberikan rekomendasi belajar yang dipersonalisasi sesuai status pengetahuan siswa dan mengurangi halusinasi [1]. Hasil evaluasi TutorLLM menunjukkan peningkatan kepuasan pengguna 10% dan skor kuis 5% dibanding penggunaan LLM umum tanpa personalisasi [1].

Studi LLM-KT (Wang dkk., 2025) juga menunjukkan pendekatan serupa dengan fokus pada knowledge tracing. LLM-KT menggunakan kerangka plug-and-play untuk menyelaraskan LLM dengan tugas KT, memanfaatkan kemampuan penalaran dan pengetahuan dunia luas dari LLM sekaligus memasukkan rekaman interaksi siswa secara sekuensial seperti pada model KT tradisional [2]. Melalui penyisipan konteks pertanyaan dan urutan jawaban historis ke dalam LLM, model hybrid ini mencapai kinerja state-of-the-art dalam memprediksi jawaban siswa di beberapa dataset KT klasik [2]. Ini menunjukkan potensi framework yang memadukan kekuatan LLM (fleksibilitas bahasa dan penalaran) dengan analitik pembelajaran tradisional (pelacakan pengetahuan sekuensial).

Selain itu, Reddig dkk. (2025) mengevaluasi penggunaan GPT-4 untuk memberikan umpan balik personal dalam konteks Intelligent Tutoring System (ITS). Studi ini menganalisis data kesalahan siswa dalam tutor aljabar lalu *mem-*prompt GPT-4 untuk menghasilkan hint/koreksi spesifik terhadap kesalahan tersebut [3]. Hasilnya, model LLM cukup efektif mendiagnosis berbagai kesalahan siswa dan menghasilkan umpan balik relevan; namun sekitar 35% hint yang dihasilkan terlalu umum, kurang tepat, atau justru langsung memberi jawaban [3]. Temuan ini menekankan bahwa meskipun LLM dapat meningkatkan skala dan personalisasi umpan balik, diperlukan mekanisme kontrol kualitas agar saran yang menyesatkan tidak ditampilkan pada siswa [3]. Menariknya, Reddig dkk. juga menguji kemampuan LLM untuk secara otomatis menilai kualitas umpan balik yang dibuatnya, sebagai upaya menggabungkan praktik tutoring terbukti dengan fleksibilitas LLM demi memperoleh “best of both worlds” [3].

Penelitian lain mengindikasikan efektivitas umpan balik berbasis LLM bergantung pada keterlibatan pelajar. Thomas dkk. (2025) menemukan bahwa dalam skenario pelatihan tutor, peserta yang memilih melihat penjelasan feedback dari LLM (ChatGPT) mendapat skor post-test lebih tinggi daripada yang tidak menggunakannya, dengan effect size moderat (0.3) pada beberapa lesson [4]. Ini menunjukkan potensi LLM untuk

memberikan umpan balik seefektif tutor manusia dalam konteks tertentu, asalkan penggunaannya terarah dan sesuai kebutuhan.

2.4.2 Pendekatan Rekayasa dalam *Learning Analytics*

Dari sisi pendekatan rekayasa dalam learning analytics, beberapa karya menyoroti pentingnya merancang sistem yang dapat mengumpulkan dan menganalisis data proses belajar secara efisien, termasuk data “Kanban” atau pelacakan tugas belajar. Sebagai contoh, sebuah studi memanfaatkan papan Kanban untuk orkestrasi kelas sebagai alat self-regulated learning; papan tersebut menampilkan kartu tugas siswa dalam kolom “to do / in progress / done” untuk memvisualisasikan kemajuan mereka [1]. Data status tugas di Kanban ini memberikan informasi real-time tentang progres dan keterlibatan siswa yang dapat dimanfaatkan guru. Bahkan, dikembangkan prototipe alat digital Kanban yang dirancang khusus untuk pendidikan (tidak bergantung pada Trello dkk. karena isu privasi) agar guru dapat menyiapkan template board dan mereplikasi ke banyak kelompok belajar [1].

Ke depan, platform Kanban edukasi semacam itu berpotensi diperluas dengan kapabilitas learning analytics langsung. Misalnya, Winter dkk. (2022) merancang alat Kanban kelas yang rencananya dilengkapi analytics berupa statistik otomatis, identifikasi “tugas yang macet”, rekomendasi khusus, serta dashboard guru untuk memantau semua papan siswa secara bersamaan [1]. Integrasi semacam ini menunjukkan pendekatan rekayasa di mana data pelacakan tugas mikro (misal, berapa lama tugas tertahan di kolom tertentu) dikumpulkan dan dianalisis guna memberi wawasan tindakan: guru dapat segera melihat siapa yang butuh bantuan (tugasnya lama di “in progress”), atau sistem bisa menyarankan intervensi pada tugas spesifik. Perekaman data proses semacam ini juga bermanfaat untuk penelitian edukasi dan evaluasi metode belajar, asalkan sesuai aturan privasi [1].

Lebih umum lagi, data proses (process data) dalam learning analytics, yaitu rekaman langkah-langkah atau interaksi saat siswa mengerjakan tugas yang dipandang kian penting untuk memahami pola belajar. Mazzullo dkk. (2023) menekankan bahwa pemanfaatan data proses (misal urutan klik, waktu respon, transisi antar tugas) dapat menghasilkan insight yang lebih interpretable tentang perilaku belajar dibanding sekadar nilai akhir [2]. Pendekatan rekayasa LA yang baik perlu mampu menangkap data proses ini dan menyajikannya secara informatif. Tantangan yang diidentifikasi meliputi bagaimana mendesain sistem LA dengan landasan pedagogis kuat, memastikan insight mudah dipahami, prediksi akurat, dan umpan balik yang dapat ditindaklanjuti [2]. Dengan kata lain, arsitektur LA harus dirancang tidak hanya untuk mengumpulkan data (termasuk data “kanban” tugas) tetapi juga untuk menganalisis dan menyajikannya dalam bentuk yang mendukung pengambilan keputusan pembelajaran.

2.4.3 Menghubungkan Hasil *Learning Analytics* dengan Prompt *Context-Aware* pada LLM

Bidang ini semakin berkembang menuju penggabungan insight dari learning analytics dengan prompt engineering yang context-aware untuk meningkatkan kualitas output LLM. Idennya adalah menjadikan LLM lebih “sadar konteks” situasi belajar individu, sehingga respon yang dihasilkan lebih tepat sasaran. Beberapa studi telah mengeksplorasi pendekatan ini.

Venugopalan dkk. (LAK 2025) mempelajari skenario hybrid tutoring di mana sistem tutoring tradisional digabung dengan LLM untuk mendukung orang tua yang membimbing anaknya belajar matematika. Mereka menggunakan prompt engineering dengan memasukkan konteks real-time dari ITS (misal: langkah pengerjaan siswa, kesalahan yang dilakukan, hint yang sudah diberikan) ke prompt LLM, ditambah contoh praktik tutoring yang baik (few-shot). Hasilnya, LLM (Llama 3) mampu menghasilkan rekomendasi percakapan bagi orang tua yang dinilai membantu meningkatkan dukungan metakognisi siswa (misal mendorong anak menjelaskan pikirannya) []. Studi ini menunjukkan bahwa memberikan konteks situasional (dari analytics ITS) dalam prompt dapat meng-align output LLM dengan strategi pedagogis yang diinginkan.

Demikian pula, TutorLLM yang disebut sebelumnya memanfaatkan contextual prompt secara otomatis: modul “Scraper” dalam arsitekturnya mengambil konten teks terkait materi belajar selama sesi online, lalu informasi itu bersama prediksi model KT tentang kelemahan siswa disuplai ke LLM (GPT-4) untuk menghasilkan jawaban yang akurat dan sesuai tingkat pemahaman siswa []. Dengan menambah pengetahuan latar kontekstual ke prompt, TutorLLM terbukti mengurangi risiko halusinasi dan meningkatkan relevansi jawaban terhadap kebutuhan spesifik siswa []. Ini adalah contoh konkret bagaimana hasil learning analytics (dalam hal ini profil pengetahuan siswa dari KT dan materi relevan yang di-retrieve) dapat tersambung langsung ke LLM sehingga responnya terpersonalisasi dan kontekstual.

Pendekatan serupa tampak pada studi Reddig dkk. (2025) tadi: alih-alih LLM bekerja “kosong”, mereka mem-prompt GPT-4 dengan input berupa deskripsi kesalahan spesifik yang dilakukan siswa pada tutor algebra, seolah memberikan konteks masalah dan kesalahan untuk diperbaiki []. Strategi prompting yang context-aware ini memungkinkan GPT-4 memberikan feedback yang lebih tepat karena “tahu” kesalahan mana yang terjadi. Reddig dkk. juga membahas pentingnya format prompt dan kriteria penilaian keluaran agar LLM tidak memberi jawaban terlalu langsung atau keliru []. Hal ini menggarisbawahi perlunya prompt engineering berbasis insight LA: misalnya menentukan informasi apa saja yang harus disertakan dalam prompt (error siswa, riwayat upaya, dll.) dan bagaimana meminta LLM memberikan keluaran (misal hint yang membimbing, bukannya solusi instan) []. Penelitian Stamper dkk. (2024) bahkan menyusun kerangka

sistematis untuk merancang prompt LLM berbasis prinsip umpan balik ITS yang sudah teruji, seperti menentukan kapan LLM diberi trigger untuk menilai jawaban siswa, informasi apa yang perlu disertakan dalam prompt input ke LLM, dan jenis keluaran umpan balik apa yang diharapkan []. Pendekatan teoritis ini memastikan penggunaan LLM selaras dengan teori pendidikan (mis. umpan balik segera, spesifik, mendorong self-explanation) sehingga kualitas output-nya lebih terjamin.

Secara keseluruhan, tren riset ini menunjukkan upaya menggabungkan kekuatan learning analytics dengan kecerdasan generatif LLM. Learning analytics menyediakan data-driven insights tentang proses dan status belajar (misal: apa yang sudah dikuasai, kesalahan yang sering muncul, progress task di Kanban, dll.), sementara LLM dapat memanfaatkan konteks tersebut untuk menghasilkan umpan balik atau bimbingan yang adaptif. Mazzullo dkk. (2023) menyatakan bahwa implementasi LLM dalam LA dapat membantu memberikan insight yang lebih mudah dipahami, umpan balik yang lebih cepat dan actionable, peningkatan personalisasi, serta mendukung tugas-tugas guru secara lebih luas []. Dengan mengaitkan hasil analitik (misal deteksi kebuntuan tugas atau prediksi kemampuan) ke dalam prompt atau modul pengetahuan LLM, diharapkan output AI menjadi lebih relevan konteks dan berkualitas tinggi dalam mendukung pembelajaran. Ini juga area kebaruan yang bisa peneliti tekankan: membedakan bagaimana pendekatan mereka menyambungkan LA dan LLM secara lebih erat atau efektif dibanding studi-studi terdahulu seperti TutorLLM, LLM-KT, atau Reddig dkk. di atas.

2.5 Analisis Perbandingan Metode

Di dalam tinjauan pustaka hasil akhirnya adalah analisis secara kualitatif atau pun secara kuantitatif kelebihan dan kekurangan metode jika dikaitkan dengan masalah, batasan-batasan masalah dan solusi yang diinginkan. Analisis kuantitatif tidak wajib tetapi mempunyai nilai tambah di dalam tugas akhir saudara. Bagian ini menjelaskan kenapa metode tersebut dipilih dan uraikan dengan lebih jelas metode pelaksanaan tugas akhir yang ingin Anda lakukan.

2.6 Pertanyaan Tugas Akhir (Jika Perlu)

Pertanyaan tugas akhir bersifat opsional dan dapat ditambahkan untuk menekankan hal-hal yang hendak diketahui dari tugas akhir berdasar pada tujuan tugas akhir. Pertanyaan tugas akhir dikenal dengan RQ (*Research Question*) dan harus memiliki keterkaitan dengan RO (*Research Objective*). Satu RO dapat memiliki satu atau lebih dari satu RQ.

BAB III

METODE PENELITIAN

Bab ini menjelaskan metode atau cara yang digunakan dalam penelitian ini untuk mencapai maksud dan tujuan seperti yang tertulis dalam sub-bab 1.3 [jika diinginkan, kalian dapat menuliskan Kembali tujuan penelitian yang ingin dicapai di sini].

3.1 Alat dan Bahan Tugas akhir (Opsional)

3.1.1 Alat Tugas akhir

Alat-alat yang digunakan pada tugas akhir ini berupa perangkat keras maupun perangkat lunak sebagai sarana pendukung antara lain. Kemukakan secara detail sesuai dengan kebutuhan tugas akhir dan juga tambahkan spesifikasi minimum sehingga peneliti lain yang hendak melakukan hal yang sama bisa melakukannya :

1. *Notebook* dengan spesifikasi minimum sistem operasi Windows 8, *processor* Intel Core i3 2330M CPU @ 2,2 GHz, memori 4GB DDR3, grafis NVIDIA GeForce GT 610 (4GB), hardisk 500GB. Pada tugas akhir ini digunakan Windows 10, Intel Core i7 4570M CPU, Memori 4GB DDR 3, grafis Intel HD4300.
2. *Smartphone* dengan spesifikasi tipe minimum, OS Android OS v4.1.2 (Jelly Bean), CPU Dual-core 800 MHz, GPU Mali-400, Internal 4 GB, 768 MB RAM. Pada tugas akhir ini digunakan
3. *Game creation platform* versi 3.3.2 untuk Stencyl dan Construct2.
4. CORELDRAW X7, Tiled dan GIMP 2

3.1.2 Bahan Tugas akhir

Bahan tugas akhir adalah segala sesuatu yang bersifat fisik atau digital yang digunakan untuk kebutuhan tugas akhir. Bahan tugas akhir dapat berupa:

1. Bahan habis pakai. Bahan yang digunakan untuk tugas akhir. Sebagai contoh mungkin dibutuhkan kertas transparansi, baterai, atau yang lain
2. Bahan yang berupa data atau informasi yang menjadi dataset tugas akhir. Dataset tugas akhir dapat berupa:
 - Dataset pihak lain yang diperoleh dengan izin atau dalam lisensi yang diizinkan untuk digunakan secara langsung
 - Dataset pihak pertama yang disusun sendiri melalui kuisisioner, observasi, atau interview

- Dokumen panduan yang mengacu pada standar, hasil tugas akhir, atau artikel yang disitasi dan digunakan.

3.2 Metode yang Digunakan

Bagian ini membahas metode atau cara yang akan digunakan dalam penelitian, tahapan penerapan metode, dan desain penelitian (misalnya apakah penelitian akan menggunakan eksperimen di Laboratorium atau di lapangan, misalkan saja penelitian biomedis atau penelitian alat ukur hama yang dapat dilakukan di laboratorium ataupun di lapangan, atau menggunakan metode survei (misalnya untuk teknologi Informasi), studi kasus, atau analisis dengan perangkat lunak (ETAP, LTSpice, dst), atau *prototyping* (pembuatan perangkat keras).

Bagian ini juga membahas bagaimana data [akan] dianalisis, apakah dengan membandingkan keluaran beberapa alat ukur, membandingkan dengan standar atau bagaimana.

3.3 Alur Tugas Akhir

Menguraikan prosedur yang akan digunakan dan jadwal atau alur penyelesaian setiap tahap. Alur penelitian ini dapat disajikan dalam bentuk diagram. Diagram dapat disusun dengan aturan yang baik semisal menggunakan *flowchart*. Aturan dan tutorial pembuatan *flowchart* dapat dilihat di <http://ugm.id/flowcharttutorial>. Setelah menggambarkannya, penulis wajib menjelaskan langkah-langkah setiap alur tugas akhir dalam sub bab tersendiri sesuai dengan kebutuhan.

3.4 Etika, Masalah, dan Keterbatasan Penelitian (Opsional)

Bagian ini membahas pertimbangan etis penelitian dan [potensi] masalah serta keterbatasannya. Jika menyangkut penelitian dengan makhluk hidup, maka dibutuhkan adanya *ethical clearance*, di bagian ini hal itu akan dibahas. Demikian juga tentang keterbatasan ataupun masalah yang akan timbul.

BAB IV

HASIL DAN PEMBAHASAN

Berikut ini adalah yang perlu diperhatikan untuk mengisi bab hasil dan pembahasan:

1. Setiap rumusan masalah boleh memiliki lebih dari 1 tujuan.
2. Setiap subbab harus spesifik menjawab setiap tujuan yang dituliskan.
3. Setiap rumusan masalah boleh dijawab dengan 1 subbab atau lebih.

Berikut ini adalah contoh sub bab untuk menjelaskan tujuan penelitian.

4.1 Pembahasan Tujuan 1 dengan Hasil Penelitian 1 (Ubah Judul Sesuai dengan Hal yang Hendak dibahas)

Sub bab pertama adalah membahas tujuan penelitian pertama dengan hasil penelitian ke-1. Dapat ditambahkan beberapa sub bab jika diperlukan.

4.2 Pembahasan Tujuan 1 dengan Hasil Penelitian 2 (Ubah Judul Sesuai dengan Hal yang Hendak dibahas)

Sub bab kedua adalah membahas tujuan penelitian pertama dengan hasil penelitian ke-2. Sub bab ini merupakan contoh tambahan sub bab pertama.

4.3 Pembahasan Tujuan 2 dengan Hasil Penelitian 3 (Ubah Judul Sesuai dengan Hal yang Hendak dibahas)

Sub bab ketiga adalah membahas tujuan penelitian kedua. Dapat ditambahkan beberapa sub bab jika diperlukan.

4.4 Perbandingan Hasil Penelitian dengan Hasil Terdahulu

Pembahasan penutup dapat menjelaskan mengenai kelebihan hasil pengembangan / penelitian dan kekurangan dibandingkan dengan skripsi atau penelitian terdahulu atau perbandingan terhadap produk lain yang ada di pasaran. Penulis dapat menggunakan tabel untuk membandingkan secara gamblang dan menjelaskannya.

BAB V

TAMBAHAN (OPSIONAL)

Anda boleh menambahkan Bab jika diperlukan. Jumlah Bab tidak harus sesuai dengan *template*.

Bab tambahan ini diperlukan jika hasil penelitian untuk menjawab tujuan cukup panjang atau terdiri dari banyak sub bab. Mahasiswa boleh menjawab 1 tujuan penelitian dengan 1 bab.

BAB VI

KESIMPULAN DAN SARAN

6.1 Kesimpulan

Kesimpulan dapat diawali dengan apa yang dilakukan dengan tugas akhir ini lalu dilanjutkan dengan poin-poin yang menjawab tujuan penelitian, apakah tujuan sudah tercapai atau belum, tentunya berdasarkan data ataupun hasil dari Bab pembahasan sebelumnya. Dalam beberapa hal, kesimpulan dapat juga berisi tentang temuan/*findings* yang Anda dapatkan setelah melakukan pengamatan dan atau analisis terhadap hasil penelitian.

Kesimpulan menjawab seberapa jauh rumusan masalah tercapai berdasarkan hasil penelitian. Semua rumusan masalah harus disimpulkan berdasarkan data penelitian.

6.2 Saran

Saran berisi hal-hal yang bisa dilanjutkan dari penelitian atau skripsi ini, yang belum dilakukan karena batasan permasalahan. Saran bukan berisi saran kepada sistem atau pengguna, tetapi saran diberikan kepada aspek penelitian yang dapat dikembangkan dan ditambahkan di penelitian atau skripsi selanjutnya.

Catatan: Mahasiswa perlu melihat sinkronisasi antara rumusan masalah, tujuan, metode, hasil penelitian, dan kesimpulan.

DAFTAR PUSTAKA

- [1] C. Liu, Z. A. Bakar, and Q. Xu, “Self-regulated learning and academic achievement in higher education: A decade systematic review,” *International Journal of Research and Innovation in Social Science (IJRISS)*, vol. 9, no. 03, pp. 4488–4504, 2025. [Online]. Available: <https://dx.doi.org/10.47772/IJRISS.2025.90300358>
- [2] S. Xiao, K. Yao, and T. Wang, “The relationships of self-regulated learning and academic achievement in university students,” *SHS Web of Conferences*, vol. 60, p. 01003, 2019, pHECSS2018. [Online]. Available: <https://doi.org/10.1051/shsconf/20196001003>
- [3] E. Chitra, N. Hidayah, M. Chandratilake, and V. D. Nadarajah, “Self-regulated learning practice of undergraduate students in health professions programs,” *Frontiers in Medicine*, vol. 9, p. 803069, 2022. [Online]. Available: <https://doi.org/10.3389/fmed.2022.803069>
- [4] A. Akdeniz, “Exploring the impact of self-regulated learning intervention on students’ strategy use and performance in a design studio course,” *International Journal of Technology and Design Education*, pp. 1–35, 2022, advance online publication. [Online]. Available: <https://doi.org/10.1007/s10798-022-09798-3>
- [5] F. Nyman, ““you’re not learning skills—you’re just realizing what you can do”: a preliminary study of self-regulation in higher education,” *Frontiers in Education*, vol. Volume 9 - 2024, 2024. [Online]. Available: <https://www.frontiersin.org/journals/education/articles/10.3389/feduc.2024.1418297>
- [6] L. F. Yang, Y. Liu, and Z. Xu, “Examining the effects of self-regulated learning-based teacher feedback on english-as-a-foreign-language learners’ self-regulated writing strategies and writing performance,” *Frontiers in Psychology*, vol. Volume 13 - 2022, 2022. [Online]. Available: <https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2022.1027266>
- [7] T. Bock, E. Thomm, J. Bauer, and B. Gold, “Fostering student teachers’ research-based knowledge of effective feedback,” *European Journal of Teacher Education*, vol. 47, no. 2, pp. 389–407, 2024. [Online]. Available: <https://doi.org/10.1080/02619768.2024.2338841>
- [8] H. E. Schaffer, K. R. Young, E. W. Ligon, and D. D. Chapman, “Automating individualized formative feedback in large classes based on a directed concept graph,” *Frontiers in Psychology*, vol. 8, p. 260, 2017, article 260. [Online]. Available: <https://doi.org/10.3389/fpsyg.2017.00260>
- [9] A. Pardo, J. Jovanović, S. Dawson, D. Gašević, and N. Mirriahi, “Using learning analytics to scale the provision of personalised feedback,” *British Journal of Educational Technology*, vol. 50, no. 1, pp. 128–138, Jan. 2019. [Online]. Available: <https://doi.org/10.1111/bjet.12592>
- [10] M. Usher, “Generative ai vs. instructor vs. peer assessments: a comparison of grading and feedback in higher education,” *Assessment & Evaluation in*

- Higher Education*, 2025, published online: Apr. 9, 2025. [Online]. Available: <https://doi.org/10.1080/02602938.2025.2487495>
- [11] K. K. Maurya, K. V. A. Srivatsa, K. Petukhova, and E. Kochmar, “Unifying ai tutor evaluation: An evaluation taxonomy for pedagogical ability assessment of llm-powered ai tutors,” *arXiv preprint arXiv:2412.09416*, 2025, version 2. [Online]. Available: <https://arxiv.org/abs/2412.09416>
 - [12] S. Strickroth, M. Kreidenweis, and Z. Wurm, “Learning from agile methods: Using a kanban board for classroom orchestration,” in *Proceedings of the 25th International Conference on Interactive Collaborative Learning (ICL 2022)*, ser. Lecture Notes in Networks and Systems, M. E. Auer, W. Pachatz, and T. Rüttmann, Eds., vol. 633. Springer, 2022, pp. 68–79. [Online]. Available: https://doi.org/10.1007/978-3-031-26876-2_7
 - [13] J. Woodrow, S. Koyejo, and C. Piech, “Improving generative ai student feedback: Direct preference optimization with teachers in the loop,” in *Proceedings of the 2025 International Conference on Educational Data Mining (EDM 2025)*. International Educational Data Mining Society, 2025, p. —. [Online]. Available: <https://educationaldatamining.org/EDM2025/proceedings/2025.EDM.short-papers.166/index.html>
 - [14] A. Scarlatos, D. Smith, S. Woodhead, and A. Lan, “Improving the validity of automatically generated feedback via reinforcement learning,” in *Artificial Intelligence in Education*, A. M. Olney, I.-A. Chounta, Z. Liu, O. C. Santos, and I. I. Bittencourt, Eds. Cham: Springer Nature Switzerland, 2024, pp. 280–294.
 - [15] L. J. Jacobsen and K. E. Weber, “The promises and pitfalls of large language models as feedback providers: A study of prompt engineering and the quality of ai-driven feedback,” *AI*, vol. 6, no. 2, 2025. [Online]. Available: <https://doi.org/10.3390/ai6020035>
 - [16] L. Ouyang, J. Wu, X. Jiang, D. Almeida, C. L. Wainwright, P. Mishkin, C. Zhang, S. Agarwal, K. Slama, A. Ray, J. Schulman, J. Hilton, F. Kelton, L. Miller, M. Simens, A. Askell, P. Welinder, P. Christiano, J. Leike, and R. Lowe, “Training language models to follow instructions with human feedback,” 2022. [Online]. Available: <https://arxiv.org/abs/2203.02155>
 - [17] H. Li and B. Ma, “Design of ai-powered tool for self-regulation support in programming education,” 2025. [Online]. Available: <https://arxiv.org/abs/2504.03068>
 - [18] S. Strickroth, M. Kreidenweis, and A. Götzfried, “Supporting agile classroom orchestration with a live teacher dashboard,” in *Futureproofing Engineering Education for Global Responsibility (ICL 2024)*, ser. Lecture Notes in Networks and Systems, vol. 1260. Cham: Springer, 2025, pp. 38–50.

Catatan: Daftar pustaka adalah apa yang dirujuk atau disitasi, bukan apa yang telah dibaca, jika tidak ada dalam sitasi maka tidak perlu dituliskan dalam daftar pustaka.

LAMPIRAN

L.1 Isi Lampiran

Lampiran bersifat opsional bergantung hasil kesepakatan dengan pembimbing dapat berupa:

1. Bukti pelaksanaan Kuesioner seperti pertanyaan kuesioner, resume jawaban responden, dan dokumentasi kuesioner.
2. Spesifikasi Aplikasi atau Sistem yang dikembangkan meliputi spesifikasi teknis aplikasi, tautan unduh aplikasi, manual penggunaan aplikasi, hingga screenshot aplikasi.
3. Cuplikan kode yang sekiranya penting dan ditambahkan.
4. Tabel yang terlalu panjang yang masih diperlukan tetapi tidak memungkinkan untuk ditayangkan di bagian utama skripsi.
5. Gambar-gambar pendukung yang tidak terlalu penting untuk ditampilkan di bagian utama. Akan tetapi, mendukung argumentasi/pengamatan/analisis.
6. Penurunan rumus-rumus atau pembuktian suatu teorema yang terlalu panjang dan terlalu teknis sehingga Anda berasumsi bahwa pembaca biasa tidak akan menelaah lebih lanjut. Hal ini digunakan untuk memberikan kesempatan bagi pembaca tingkat lanjut untuk melihat proses penurunan rumus-rumus ini.

LAMPIRAN

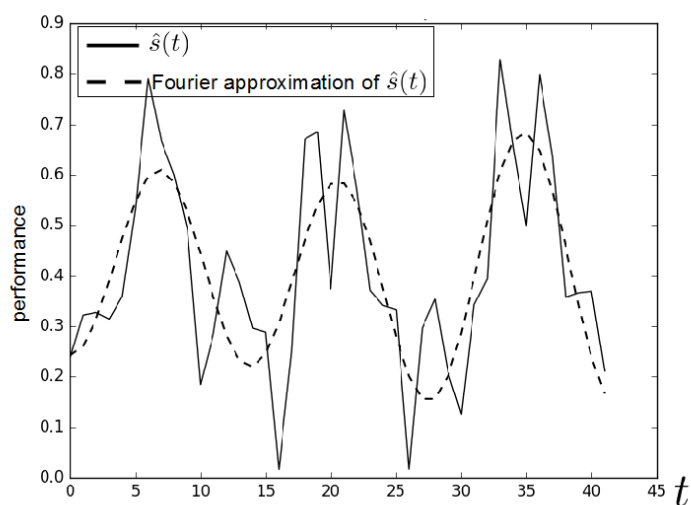
L.2 Panduan Latex

L.2.1 Syntax Dasar

L.2.1.1 Penggunaan Sitasi

Contoh penggunaan sitasi [?, ?] [?] [?] [?] [?, ?]

L.2.1.2 Penulisan Gambar



Gambar L.1. Contoh gambar.

Contoh gambar terlihat pada Gambar L.1. Gambar diambil dari [?].

L.2.1.3 Penulisan Tabel

Tabel L.1. Tabel ini

ID	Tinggi Badan (cm)	Berat Badan (kg)
A23	173	62
A25	185	78
A10	162	70

Contoh penulisan tabel bisa dilihat pada Tabel L.1.

L.2.1.4 Penulisan formula

Contoh penulisan formula

$$L_{\psi_z} = \{t_i \mid v_z(t_i) \leq \psi_z\} \quad (1)$$

Contoh penulisan secara *inline*: $PV = nRT$. Untuk kasus-kasus tertentu, kita membutuhkan perintah "mathit" dalam penulisan formula untuk menghindari adanya jeda saat penulisan formula.

Contoh formula **tanpa** menggunakan "mathit": $PVA = RTD$

Contoh formula **dengan** menggunakan "mathit": $PVA = RTD$

L.2.1.5 Contoh list

Berikut contoh penggunaan list

1. First item
2. Second item
3. Third item

L.2.2 Blok Beda Halaman

L.2.2.1 Membuat algoritma terpisah

Untuk membuat algoritma terpisah seperti pada contoh berikut, kita dapat memanfaatkan perintah *algstore* dan *algrestore* yang terdapat pada paket *algcompatible*. Pada dasarnya, kita membuat dua blok algoritma dimana blok pertama kita simpan menggunakan *algstore* dan kemudian di-restore menggunakan *algrestore* pada algoritma kedua. Perintah tersebut dimaksudkan agar terdapat kesinamungan antara kedua blok yang sejatinya adalah satu blok.

Algorithm 1 Contoh algorima

```
1: procedure CREATESET( $v$ )  
2:   Create new set containing  $v$   
3: end procedure
```

Pada blok algoritma kedua, tidak perlu ditambahkan caption dan label, karena sudah menjadi satu bagian dalam blok pertama. Pembagian algoritma menjadi dua bagian ini berguna jika kita ingin menjelaskan bagian-bagian dari sebuah algoritma, maupun untuk memisah algoritma panjang dalam beberapa halaman.

```
4: procedure CONCATSET( $v$ )  
5:   Create new set containing  $v$   
6: end procedure
```

L.2.2.2 Membuat tabel terpisah

Untuk membuat tabel panjang yang melebihi satu halaman, kita dapat mengganti kombinasi *table* + *tabular* menjadi *longtable* dengan contoh sebagai berikut.

Tabel L.2. Contoh tabel panjang

header 1	header 2
foo	bar
foo	bar
foo	bar
foo	bar
foo	bar
foo	bar
foo	bar
foo	bar
foo	bar
foo	bar
foo	bar
foo	bar

L.2.2.3 Menulis formula terpisah halaman

Terkadang kita butuh untuk menuliskan rangkaian formula dalam jumlah besar sehingga melewati batas satu halaman. Solusi yang digunakan bisa saja dengan memindahkan satu blok formula tersebut pada halaman yang baru atau memisah rangkaian formula menjadi dua bagian untuk masing-masing halaman. Cara yang pertama mungkin akan menghasilkan alur yang berbeda karena ruang kosong pada halaman pertama akan diisi oleh teks selanjutnya. Sehingga di sini kita dapat memanfaatkan *align* yang sudah diatur dengan mode *allowdisplaybreaks*. Penggunaan *align* ini memungkinkan satu rangkaian formula terpisah berbeda halaman.

Contoh sederhana dapat digambarkan sebagai berikut.

$$\begin{aligned}
 x &= y^2 \\
 x &= y^3 \\
 a + b &= c \\
 x &= y - 2 \\
 a + b &= d + e \\
 x^2 + 3 &= y \\
 a(x) &= 2x
 \end{aligned}
 \tag{2}$$

$$b_i = 5x$$

$$10x^2 = 9x$$

$$2x^2 + 3x + 2 = 0$$

$$5x - 2 = 0$$

$$d = \log x$$

$$y = \sin x$$

LAMPIRAN

L.3 Format Penulisan Referensi

Penulisan referensi mengikuti aturan standar yang sudah ditentukan. Untuk internasionalisasi DTETI, maka penulisan referensi akan mengikuti standar yang ditetapkan oleh IEEE (*International Electronics and Electrical Engineers*). Aturan penulisan ini bisa diunduh di <http://www.ieee.org/documents/ieeecitationref.pdf>. Gunakan Mendeley sebagai *reference manager* dan *export* data ke format Bibtex untuk digunakan di Latex.

Berikut ini adalah sampel penulisan dalam format IEEE:

L.3.1 Book

Basic Format:

- [1] J. K. Author, "Title of chapter in the book," in Title of His Published Book, xth ed. City of Publisher, Country: Abbrev. of Publisher, year, ch. x, sec. x, pp. xxx-xxx.

Examples:

- [1] B. Klaus and P. Horn, Robot Vision. Cambridge, MA: MIT Press, 1986.
- [2] L. Stein, "Random patterns," in Computers and You, J. S. Brake, Ed. New York: Wiley, 1994, pp. 55-70.
- [3] R. L. Myer, "Parametric oscillators and nonlinear materials," in Nonlinear Optics, vol. 4, P. G. Harper and B. S. Wherret, Eds. San Francisco, CA: Academic, 1977, pp. 47-160.
- [4] M. Abramowitz and I. A. Stegun, Eds., Handbook of Mathematical Functions (Applied Mathematics Series 55). Washington, DC: NBS, 1964, pp. 32-33.
- [5] E. F. Moore, "Gedanken-experiments on sequential machines," in Automata Studies (Ann. of Mathematical Studies, no. 1), C. E. Shannon and J. McCarthy, Eds. Princeton, NJ: Princeton Univ. Press, 1965, pp. 129-153.
- [6] Westinghouse Electric Corporation (Staff of Technology and Science, Aerospace Div.), Integrated Electronic Systems. Englewood Cliffs, NJ: Prentice-Hall, 1970.
- [7] M. Gorkii, "Optimal design," Dokl. Akad. Nauk SSSR, vol. 12, pp. 111-122, 1961 (Transl.: in L. Pontryagin, Ed., The Mathematical Theory of Optimal Processes. New York: Interscience, 1962, ch. 2, sec. 3, pp. 127-135).
- [8] G. O. Young, "Synthetic structure of industrial plastics," in Plastics, vol. 3,

Polymers of Hexadromicon, J. Peters, Ed., 2nd ed. New York: McGraw-Hill, 1964, pp. 15-64.

L.3.2 Handbook

Basic Format:

- [1] Name of Manual/Handbook, x ed., Abbrev. Name of Co., City of Co., Abbrev. State, year, pp. xx-xx.

Examples:

- [1] Transmission Systems for Communications, 3rd ed., Western Electric Co., Winston Salem, NC, 1985, pp. 44-60.
- [2] Motorola Semiconductor Data Manual, Motorola Semiconductor Products Inc., Phoenix, AZ, 1989.
- [3] RCA Receiving Tube Manual, Radio Corp. of America, Electronic Components and Devices, Harrison, NJ, Tech. Ser. RC-23, 1992.

Conference/Prosiding

Basic Format:

- [1] J. K. Author, "Title of paper," in Unabbreviated Name of Conf., City of Conf., Abbrev. State (if given), year, pp.xxx-xxx.

Examples:

- [1] J. K. Author [two authors: J. K. Author and A. N. Writer] [three or more authors: J. K. Author et al.], "Title of Article," in [Title of Conf. Record as], [copyright year] © [IEEE or applicable copyright holder of the Conference Record]. doi: [DOI number]

Sumber Online/Internet

Basic Format:

- [1] J. K. Author. (year, month day). Title (edition) [Type of medium]. Available: [http://www.\(URL\)](http://www.(URL))

Examples:

- [1] J. Jones. (1991, May 10). Networks (2nd ed.) [Online]. Available: <http://www.atm.com>

Skripsi, Tesis dan Disertasi

Basic Format:

- [1] J. K. Author, "Title of thesis," M.S. thesis, Abbrev. Dept., Abbrev. Univ., City of Univ., Abbrev. State, year.

[2] J. K. Author, "Title of dissertation," Ph.D. dissertation, Abbrev. Dept., Abbrev. Univ., City of Univ., Abbrev. State, year.

Examples:

[1] J. O. Williams, "Narrow-band analyzer," Ph.D. dissertation, Dept. Elect. Eng., Harvard Univ., Cambridge, MA, 1993. [2] N. Kawasaki, "Parametric study of thermal and chemical nonequilibrium nozzle flow," M.S. thesis, Dept. Electron. Eng., Osaka Univ., Osaka, Japan, 1993

LAMPIRAN

L.4 Contoh Source Code

L.4.1 Sample algorithm

Algorithm 2 Kruskal's Algorithm

```
1: procedure MAKESET( $v$ )
2:   Create new set containing  $v$ 
3: end procedure
4:
5: function FINDSET( $v$ )
6:   return a set containing  $v$ 
7: end function
8:
9: procedure UNION( $u, v$ )
10:  Unites the set that contain  $u$  and  $v$  into a new set
11: end procedure
12:
13: function KRUSKAL( $V, E, w$ )
14:   $A \leftarrow \{\}$ 
15:  for each vertex  $v$  in  $V$  do
16:    MakeSet( $v$ )
17:  end for
18:  Arrange  $E$  in increasing costs, ordered by  $w$ 
19:  for each  $(u, v)$  taken from the sorted list do
20:    if FindSet( $u$ )  $\neq$  FindSet( $v$ ) then
21:       $A \leftarrow A \cup \{(u, v)\}$ 
22:      Union( $u, v$ )
23:    end if
24:  end for
25:  return  $A$ 
26: end function
```

L.4.2 Sample Python code

```
1 import numpy as np
2
3 def incmatrix(genl1, genl2):
4     m = len(genl1)
5     n = len(genl2)
6     M = None #to become the incidence matrix
7     VT = np.zeros((n*m,1), int) #dummy variable
8
9     #compute the bitwise xor matrix
10    M1 = bitxormatrix(genl1)
11    M2 = np.triu(bitxormatrix(genl2),1)
12
13    for i in range(m-1):
14        for j in range(i+1, m):
15            [r,c] = np.where(M2 == M1[i,j])
16            for k in range(len(r)):
17                VT[(i)*n + r[k]] = 1;
18                VT[(i)*n + c[k]] = 1;
19                VT[(j)*n + r[k]] = 1;
20                VT[(j)*n + c[k]] = 1;
21
22    if M is None:
23        M = np.copy(VT)
24    else:
25        M = np.concatenate((M, VT), 1)
26
27    VT = np.zeros((n*m,1), int)
28
29    return M
```

L.4.3 Sample Matlab code

```
1 function X = BitXorMatrix(A,B)
2 %function to compute the sum without charge of two vectors
3
4 %convert elements into unsigned integers
5 A = uint8(A);
6 B = uint8(B);
7
8 m1 = length(A);
9 m2 = length(B);
10 X = uint8(zeros(m1, m2));
11 for n1=1:m1
12     for n2=1:m2
13         X(n1, n2) = bitxor(A(n1), B(n2));
14     end
15 end
```