

CoopMind V1: Fine-Tuning Large Language Model dengan Pendekatan Low Rank Adaptation (LoRA) untuk Demokratisasi Pengetahuan Perkoperasian

Ryan Fadhilah Faizal Hakim¹, Catherine Vanya Pangemanan², Ery Supriyadi R.³
Universitas Koperasi Indonesia^{1,2,3}
catherinevanyapang@gmail.com²

ABSTRAK

Kualitas sumber daya manusia (SDM) yang rendah menjadi kendala fundamental dalam pengembangan koperasi di Indonesia, menyebabkan adanya kesenjangan kapabilitas (*capability gap*) yang menghambat realisasi potensi ekonomi koperasi. Menjawab tantangan ini, penelitian ini bertujuan untuk mengembangkan model bahasa terspesialisasi guna mendemokratisasi pengetahuan perkoperasian. Penelitian ini melakukan *fine-tuning* pada *Large Language Model (LLM) Gemma-3-1B-it* menggunakan pendekatan *Low-Rank Adaptation (LoRA)* yang efisien dari segi komputasi. Proses *fine-tuning* dilakukan pada *dataset* instruksional spesifik domain koperasi yang terdiri dari 900 sampel. Serangkaian eksperimen dijalankan dengan memvariasikan *hyperparameter LoRA*, yaitu *rank (r)* dan *alpha (α)*, untuk menemukan konfigurasi optimal. Evaluasi kuantitatif dilakukan dengan metrik *ROUGE*, *BLEU*, dan *BERTScore* untuk mengukur kualitas leksikal dan semantik, serta analisis efisiensi komputasi. Hasil penelitian menunjukkan bahwa semua model yang melalui proses *fine-tuning* menunjukkan peningkatan performa yang signifikan dibandingkan dengan model dasar (*base model*). Konfigurasi *LoRA* dengan *rank 16* dan *alpha 32* teridentifikasi sebagai model optimal, yang mencapai skor tertinggi pada sebagian besar metrik evaluasi tanpa menunjukkan gejala *overfitting*. Keberhasilan ini membuktikan bahwa metode *LoRA* sangat efektif dan efisien untuk mengadaptasi *LLM* pada domain spesifik. Model yang dihasilkan, *CoopMind*, berpotensi besar menjadi instrumen untuk menjembatani kesenjangan pengetahuan dan mendukung pengembangan SDM koperasi di Indonesia secara lebih luas dan merata.

Kata Kunci: Koperasi, Large Language Model, Fine-Tuning, Low Rank Adaptation (LoRA), Demokratisasi Pengetahuan

I. PENDAHULUAN

Koperasi, sesuai dengan amanat Undang-Undang No. 25 Tahun 1992, memegang peranan fundamental dalam struktur perekonomian nasional. Fungsinya tidak hanya terbatas pada peningkatan kesejahteraan anggota dan masyarakat (Pasal 4a), tetapi juga sebagai pilar fundamental yang memperkokoh perekonomian rakyat (Pasal 4c). Urgensi peran ini semakin terefleksi dalam kerangka visi Indonesia Emas 2045 dan program Asta Cita, di mana pemerintah melalui berbagai inisiatif, seperti "Koperasi Desa/Kelurahan Merah Putih", menginisiasi upaya revitalisasi peran strategis koperasi melalui pendampingan, edukasi, dan penguatan manajemen berbasis digital.

Namun, realitas empiris mengindikasikan bahwa keberhasilan program-program tersebut seringkali terhambat oleh problematika fundamental, yakni rendahnya kualitas sumber daya manusia (SDM). Penelitian oleh Dasuki R. (2021) mengafirmasi bahwa aset paling vital bagi suatu organisasi adalah SDM yang kompeten untuk mengelola entitas usaha secara profesional. Selaras dengan hal tersebut, temuan Toman S. (2019) mengindikasikan bahwa banyak koperasi mengalami kegagalan dalam menjaga keberlangsungan usahanya akibat inkompetensi SDM dalam tata kelola organisasi, yang seringkali bersumber dari pemahaman yang tidak

komprehensif mengenai konsep dan tujuan utama koperasi itu sendiri. Hal ini diperkuat oleh studi Lozano et al. (2016) yang menekankan signifikansi kesamaan persepsi antar anggota untuk mencapai tujuan kolektif berupa kesejahteraan. Meskipun intervensi melalui pendidikan dan pelatihan diakui sebagai solusi efektif untuk peningkatan kinerja SDM (Purwito et al., 2024), implementasinya masih dihadapkan pada kendala-kendala serius. Mujiyanti (2023) menyoroti bahwa banyak koperasi di Indonesia tidak memiliki akses yang memadai terhadap pendidikan perkoperasian, sehingga mengalami kesulitan dalam melakukan eskalasi kapabilitas para anggotanya. Kondisi demikian berkontribusi pada terciptanya sebuah kesenjangan kapabilitas (*capability gap*) yang menghambat realisasi potensi koperasi.

Sebagai respons terhadap tantangan aksesibilitas dan kualitas pendidikan tersebut, kemajuan teknologi menghadirkan suatu terobosan paradigmatik. *Artificial Intelligence (AI)*, khususnya *Generative AI* dalam bentuk *Large Language Models (LLM)*, memiliki potensi signifikan untuk mendemokratisasi pengetahuan perkoperasian. Demokratisasi ini dimaknai sebagai upaya melampaui metodologi pelatihan konvensional dengan menyediakan akses informasi yang bersifat personal, skalabel, dan mudah dijangkau oleh seluruh pemangku kepentingan koperasi, termasuk di wilayah-wilayah yang kurang terlayani (Aksor, 2024; Costa et al., 2024).

Secara operasional, aplikasi LLM dapat menjadi solusi pragmatis terhadap permasalahan SDM koperasi melalui beberapa mekanisme. Teknologi ini dapat berfungsi sebagai tutor virtual yang dapat diakses secara asinkron, yang mampu memberikan respons terhadap pertanyaan-pertanyaan spesifik mengenai regulasi, prinsip akuntansi koperasi, hingga penyusunan Anggaran Dasar/Anggaran Rumah Tangga (AD/ART). Lebih lanjut, LLM dapat menyajikan materi pelatihan dalam format dialog interaktif, menyediakan studi kasus simulatif untuk manajemen konflik atau analisis keuangan, serta memberikan asistensi kepada pengurus dalam penyusunan laporan dengan efisiensi yang lebih tinggi. Dengan demikian, LLM tidak hanya berfungsi sebagai medium diseminasi informasi, tetapi juga mendukung proses edukasi mandiri yang bersifat kontinu, memberdayakan setiap anggota untuk belajar sesuai dengan ritme dan kebutuhan individual, suatu manfaat transformatif yang dapat mendorong inklusivitas (Papadopoulos, 2024; Saouabe et al., 2024).

Meskipun demikian, untuk optimalisasi potensi tersebut, kapabilitas LLM generik dipandang tidak memadai. Diperlukan suatu proses adaptasi agar model dapat menginternalisasi terminologi, kerangka regulasi, dan nuansa spesifik dalam domain perkoperasian. Dalam konteks inilah, proses fine-tuning menjadi suatu tahapan yang krusial, yakni melatih lebih lanjut sebuah model pada korpus data khusus untuk mengadaptasinya pada domain yang terspesialisasi (Anisuzzaan, et al., 2024; Lu et al., 2025). Namun, implementasi *full fine-tuning* secara inheren menuntut sumber daya komputasi yang sangat besar dan mahal. Sebagai solusinya, pendekatan *Low-Rank Adaptation (LoRA)* yang diperkenalkan oleh Hu et al. (2021) hadir sebagai suatu metodologi fine-tuning yang dicirikan oleh efisiensi tinggi. *LoRA* secara signifikan mereduksi jumlah parameter yang perlu dilatih, sehingga menurunkan kebutuhan memori dan biaya komputasi secara drastis tanpa mengorbankan kualitas model. Efisiensi inilah yang menjadikan *LoRA* sebagai opsi yang memiliki relevansi dan viabilitas

tinggi untuk diimplementasikan dalam konteks pengembangan alat bantu bagi koperasi, di mana keterbatasan sumber daya kerap menjadi pertimbangan utama (Hu et al., 2021).

Hingga saat ini, ketersediaan LLM yang secara spesifik telah melalui proses *fine-tuning* untuk domain perkoperasian Indonesia belum teridentifikasi secara luas. Eksistensi kesenjangan teknologi (*technology gap*) ini menjadi justifikasi utama yang melandasi urgensi penelitian ini. Dengan demikian, tujuan utama dari penelitian ini adalah untuk melakukan *fine-tuning Large Language Model* dengan pendekatan *Low-Rank Adaptation (LoRA)* pada model *Gemma 3 1B Instruct*. Model ini akan dilatih lebih lanjut menggunakan *dataset* spesifik koperasi dengan tujuan fundamental untuk menjembatani kesenjangan kapabilitas yang ada melalui demokratisasi pengetahuan, sehingga informasi dan pemahaman mendalam mengenai koperasi dapat diakses dengan tingkat kemudahan, keluasan, dan pemerataan yang lebih tinggi.

II. TINJAUAN PUSTAKA

A. Problematika Fundamental dalam Pengembangan Sumber Daya Manusia Koperasi

Kualitas sumber daya manusia (SDM) diidentifikasi sebagai variabel determinan yang signifikan bagi keberhasilan operasional dan keberlanjutan institusional koperasi. Dasuki (2021) menggarisbawahi bahwa SDM yang memiliki kompetensi tinggi merupakan aset organisasional yang paling esensial dalam mendukung pencapaian tujuan secara profesional. Kendati demikian, berbagai studi secara konsisten mengindikasikan bahwa sektor koperasi di Indonesia dihadapkan pada tantangan inheren dalam aspek ini. Inkompetensi SDM dalam tata kelola organisasi secara efektif merupakan faktor kausal yang dominan terhadap disrupsi keberlangsungan koperasi (Toman, 2019). Akar permasalahan ini seringkali berhulu pada defisit pemahaman yang mendalam mengenai filosofi, prinsip, dan nilai-nilai perkoperasian. Defisit tersebut termanifestasi dalam praktik operasional sehari-hari, seperti kekeliruan dalam pengelolaan keuangan, penyelenggaraan Rapat Anggota Tahunan (RAT) yang tidak selaras dengan prosedur, hingga ketidakmampuan untuk berinovasi dalam layanan keanggotaan. Studi oleh Toman (2019) dan Lozano et al. (2016) secara konvergen mengindikasikan bahwa defisit pemahaman ini berimplikasi pada deviasi operasional dari tujuan fundamental koperasi, yakni peningkatan kesejahteraan anggota.

Intervensi melalui program pendidikan dan pelatihan konvensional telah diakui sebagai strategi yang krusial untuk mitigasi permasalahan tersebut (Purwito et al., 2024). Namun, implementasinya pada tataran praktis menghadapi berbagai hambatan signifikan. Model pelatihan ini seringkali bersifat sporadis, berbiaya tinggi, dan distribusinya tidak merata. Mujiyanti (2023) secara spesifik mengemukakan bahwa sebagian besar koperasi, khususnya yang berlokasi di wilayah dengan keterbatasan akses, tidak memperoleh eksposur yang memadai terhadap program-program pendidikan perkoperasian. Keterbatasan akses ini pada akhirnya melahirkan sebuah kesenjangan kapabilitas (*capability gap*) yang bersifat sistemik. Kesenjangan ini tidak hanya mencakup pengetahuan dasar perkoperasian, tetapi juga keterampilan modern yang krusial seperti literasi digital, manajemen keuangan kontemporer, dan strategi pemasaran di era digital. Akibatnya, pengurus dan anggota mengalami kesulitan

dalam mengakselerasi perolehan pengetahuan dan keterampilan yang relevan untuk mengelola entitas usaha secara efisien dan inovatif.

B. Demokratisasi Pengetahuan sebagai Paradigma yang Difasilitasi Kecerdasan Buatan

Sebagai respons langsung terhadap kesenjangan kapabilitas yang pada hakikatnya merupakan manifestasi dari hambatan akses terhadap pengetahuan evolusi teknologi Kecerdasan Buatan (AI) memperkenalkan sebuah paradigma transformatif. Konsep demokratisasi pengetahuan melalui AI merujuk pada upaya sistematis untuk memperluas aksesibilitas terhadap informasi dan wahana pembelajaran melampaui batas-batas institusi formal, sehingga mampu menjangkau audiens yang lebih heterogen secara inklusif (Aksoy, 2024; Costa et al., 2024). Dalam konteks koperasi, hal ini dapat diartikan sebagai pemberdayaan seorang pengurus di lokasi terpencil untuk memperoleh jawaban atas kueri kompleks mengenai regulasi perpajakan secara instan, tanpa harus menunggu interval waktu yang panjang untuk kunjungan seorang fasilitator. AI memiliki potensi signifikan untuk melampaui hambatan geografis dan sosio-ekonomi dengan menyediakan sumber daya pendidikan yang terpersonalisasi (*customized*) serta memfasilitasi ekosistem pembelajaran jarak jauh yang efektif, terutama bagi komunitas yang secara historis kurang terlayani (Aksor, 2024).

Transisi dari metodologi pengembangan SDM tradisional menuju pendekatan yang diperkaya oleh AI (*AI-enhanced*) memungkinkan pergeseran dari model pembelajaran episodik ke model pembelajaran kontinu dan sesuai permintaan (*on-demand*). Hal ini memungkinkan proses pengambilan keputusan yang berbasis data (*data-driven*) dan peningkatan efisiensi operasional secara substansial (Saouabe et al., 2024). Paradigma ini selaras dengan agenda untuk mendorong adopsi AI yang inklusif oleh pemangku kepentingan yang beragam, termasuk institusi pendidikan dan pemerintah, guna memastikan bahwa diseminasi manfaat teknologi terjadi secara ekuivalen (Costa et al., 2024). Dalam kerangka konseptual ini, AI diposisikan bukan sekadar sebagai instrumen pasif, melainkan sebagai fasilitator strategis aktif yang mampu menjembatani kesenjangan akses informasi yang selama ini menjadi kendala utama dalam pengembangan SDM koperasi.

C. *Supervised Fine-Tuning* (SFT)

Sebagai sub-bidang AI yang paling relevan untuk intervensi edukasi, *Generative AI*, khususnya *Large Language Models* (LLM), merepresentasikan sebuah lompatan kapabilitas teknologi yang berpotensi merevolusi metode penyampaian dalam sektor pendidikan. LLM merupakan model komputasional yang dilatih pada korpus data tekstual berskala masif, yang memberikannya kemampuan untuk melakukan komprehensi, sintesis, generasi, dan prediksi konten linguistik dengan tingkat kefasihan yang mendekati manusia.

Untuk mengadaptasi LLM pada tugas spesifik seperti *Question Answering instruksional*, metode *Supervised Fine-Tuning* (SFT) menjadi pendekatan standar. SFT secara esensial adalah proses “spesialisasi” di mana model dilatih pada dataset yang terdiri dari pasangan *prompt-response* (atau *instruction-response*) untuk mengajarkan model agar mengikuti instruksi dan menghasilkan jawaban yang relevan dan koheren (Li et al., 2021). Proses ini membentuk

perilaku model, mengubahnya dari generator teks umum menjadi asisten yang berorientasi pada tugas. Pemilihan SFT sebagai paradigma pelatihan utama didasarkan pada tujuannya yang secara eksplisit melatih model untuk menjadi “*helpful and harmless assistant*”, yang selaras dengan kebutuhan untuk menciptakan alat bantu edukasi yang andal dan aman (Li et al., 2023). Proses ini secara langsung mengoptimalkan kemampuan model dalam mematuhi format instruksional.

D. *Low-Rank Adaptation (LoRA)*

Untuk mencapai tingkat efektivitas yang optimal, performa LLM generik memerlukan adaptasi untuk mencapai pemahaman yang mendalam dan kontekstual pada domain spesifik. Proses ini secara teknis dikenal sebagai *fine-tuning* (Anisuzzaman, et al., 2024; Lu et al., 2025). Namun, *full fine-tuning* sangat tidak efisien. Sebagai respons, Hu et al. (2021) memperkenalkan *Low-Rank Adaptation (LoRA)*, sebuah metode *Parameter-Efficient Fine-Tuning (PEFT)* yang populer.

Pendekatan inovatif ini secara drastis mereduksi jumlah parameter yang dapat dilatih (dengan potensi reduksi hingga 10.000 kali) serta menurunkan kebutuhan memori GPU (hingga 3 kali lebih rendah) apabila dibandingkan dengan metode full fine-tuning, seraya mempertahankan atau bahkan melampaui kualitas model yang dihasilkan (Hu et al., 2021).

Prinsip kerja *LoRA* didasarkan pada hipotesis bahwa pengetahuan yang dibutuhkan untuk adaptasi domain baru berada dalam sub-ruang berperingkat rendah (*low-rank subspace*). Dengan kata lain, model tidak perlu mempelajari ulang seluruh bahasa, tetapi hanya perlu "disetir" sedikit ke arah domain baru. *LoRA* mengimplementasikan ini dengan membekukan bobot (*weights*) asli dari model prelatih ($W_0 \in R_d \times k$) dan menginjeksikan sepasang matriks dekomposisi peringkat rendah (*low-rank decomposition*) yang dapat dilatih ($A \in R_d \times k$ dan $B \in R_d \times k$) dalam lapisan-lapisan tertentu pada arsitektur Transformer. Dengan demikian, pembaruan bobot tidak dilakukan secara langsung pada W_0 , melainkan diaproksimasi melalui perkalian matriks B dan A. Proses forward pass untuk sebuah lapisan yang dimodifikasi *LoRA* dapat direpresentasikan dengan persamaan:

$$h = W_0 + \Delta W_x = W_{0x} + BA_x \quad (1)$$

Di mana h adalah output, x adalah input, dan r adalah rank dari matriks dekomposisi, yang merupakan sebuah hyperparameter krusial ($r \ll \min(d_k)$). Jumlah parameter yang dapat dilatih berkurang drastis dari $\times k$ menjadi hanya $r \times (d + k)$.

Pemilihan *hyperparameter* *LoRA* sangat menentukan keberhasilan adaptasi. *Hyperparameter* pertama adalah *rank (r)*, yang secara konseptual mendefinisikan kapasitas dari matriks adaptasi yang dapat dilatih. Pemilihan nilai r melibatkan pertukaran fundamental antara kapasitas ekspresif dan efisiensi. Dalam penelitian yang dilakukan oleh Chen et al. (2025) menemukan bahwa *LoRA* sensitif terhadap inisialisasi “ r ”, di mana peningkatan “ r ” awal melalui pre-warm up dapat meningkatkan kinerja hingga 5-10% dibandingkan inisialisasi acak, dengan peningkatan yang lebih besar pada rank yang lebih tinggi.

Hyperparameter kedua yang krusial adalah *alpha (lora_alpha)*, yang berfungsi sebagai faktor penskalaan untuk bobot adaptasi (ΔW). *Hyperparameter* ini memodulasi seberapa besar pengaruh matriks adaptasi yang baru dipelajari terhadap bobot asli yang telah dibekukan.

Praktik umum yang seringkali efektif adalah menyetarakan nilai *alpha* dengan nilai *rank*. Heuristik ini membantu menstabilkan proses pelatihan dengan memastikan bahwa bobot adaptasi memiliki skala magnitudo yang sebanding dengan bobot asli, sehingga mencegah pembaruan yang terlalu besar atau terlalu kecil di awal pelatihan yang dapat mengganggu stabilitas model (Hu et al., 2021).

Efektivitas LoRA juga sangat bergantung pada pemilihan modul target dalam arsitektur Transformer. Keputusan untuk menerapkan matriks adaptasi tidak dilakukan secara acak, melainkan pada lapisan-lapisan yang dianggap paling vital untuk pemrosesan informasi (Hu et al., 2021). Praktik yang umum direkomendasikan adalah menargetkan modul-modul kunci dari mekanisme atensi (*self-attention*) dan lapisan *feed-forward*. Lapisan atensi menentukan "fokus" atau relevansi antar-token dalam input, sementara lapisan *feed-forward* adalah tempat "pemrosesan" dan ekstraksi fitur dari informasi tersebut. Mai (2025) dalam penelitiannya menemukan bahwa memfokuskan LoRA pada modul-modul seperti lapisan perhatian (*attention layers*) dapat mempertahankan kinerja pada tugas-tugas seperti penalaran umum (misalnya, *HellaSwag*) dan matematis (*GSM8K*), serta mengurangi parameter yang dapat dilatih hingga 50%.

E. Kerangka Evaluasi Komprehensif untuk Model Generatif

Evaluasi performa LLM yang telah di-fine-tune memerlukan pendekatan multi-metrik untuk menangkap berbagai aspek kualitas generasi teks, karena tidak ada satu metrik pun yang dapat secara sempurna mengukur kualitas bahasa. Titik awal evaluasi adalah metrik loss pelatihan dan validasi, yang berfungsi sebagai indikator fundamental keberhasilan pembelajaran. Metrik ini mengukur tingkat kesalahan prediksi model, di mana penurunan kurva loss yang stabil mengindikasikan bahwa model berhasil menginternalisasi pola dari data pelatihan. Meskipun demikian, *loss* yang rendah tidak selalu berkorelasi langsung dengan kualitas output yang dirasakan manusia, sehingga fungsinya lebih sebagai penanda konvergensi pelatihan daripada sebagai ukuran kualitas absolut, yang mengharuskan penggunaan metrik evaluasi komplementer.

Selanjutnya, evaluasi beralih ke metrik berbasis tumpang tindih *n*-gram, seperti *ROUGE* dan *BLEU*. *ROUGE* (*Recall-Oriented Understudy for Gisting Evaluation*) secara khusus mengukur tumpang tindih leksikal dengan fokus pada *recall*, menilai seberapa banyak informasi dari teks referensi yang berhasil ditangkap oleh teks yang dihasilkan (Lin, 2024). Di sisi lain, *BLEU* (*Bilingual Evaluation Understudy*), yang awalnya dirancang untuk evaluasi terjemahan mesin, juga menggunakan prinsip tumpang tindih *n*-gram tetapi dengan penalti untuk presisi (Papineni et al., 2002). Meskipun kedua metrik ini sangat berguna untuk mengukur kesamaan pada level permukaan (*surface-level similarity*), keterbatasan utamanya adalah keabaian terhadap makna semantik. Keduanya dapat memberikan skor rendah untuk jawaban yang benar secara konseptual namun menggunakan kosakata yang berbeda (sinonim), sehingga tidak mampu menangkap kualitas parafrase atau kekayaan linguistik.

Untuk mengatasi kelemahan metrik berbasis *n*-gram, kerangka evaluasi modern mengintegrasikan metrik berbasis kesamaan semantik, dengan *BERTScore* sebagai salah satu yang paling menonjol (Zhang et al., 2019). Metrik ini tidak lagi bergantung pada pencocokan kata secara harfiah, melainkan memanfaatkan embedding kontekstual dari model Transformer

seperti *BERT* untuk menghitung kesamaan kosinus (*cosine similarity*) antara token dalam teks yang dihasilkan dan teks referensi (Zhang et al., 2019). Dengan demikian, *BERTScore* mampu menilai kesamaan makna, bahkan jika kata-kata yang digunakan berbeda secara leksikal. Metrik ini terurai menjadi *Precision* (seberapa relevan kata-kata yang dihasilkan), *Recall* (seberapa banyak ide dari referensi yang dicakup), dan *F1-score* (rata-rata harmonik keduanya), sehingga memberikan penilaian yang lebih bernuansa dan selaras dengan penilaian intuitif manusia terhadap kualitas.

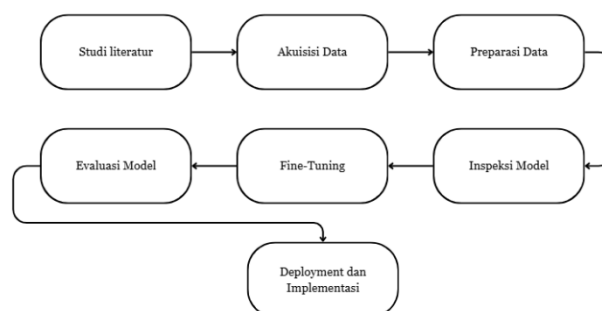
Di luar evaluasi kualitas linguistik, pemantauan sumber daya komputasi menjadi metrik sekunder yang esensial, terutama dalam konteks penelitian yang menggunakan PEFT. Pengukuran penggunaan memori dan utilitas GPU secara sistematis selama proses pelatihan dan inferensi berfungsi untuk memvalidasi klaim efisiensi dari metode yang digunakan. Data ini memberikan bukti kuantitatif mengenai kelayakan implementasi metode tersebut pada perangkat keras dengan sumber daya terbatas.

III. METODE PENELITIAN

A. Desain Penelitian

Penelitian ini mengadopsi pendekatan kuantitatif dengan desain eksperimental yang diarahkan untuk pengembangan dan evaluasi sebuah model bahasa yang terspesialisasi dalam domain perkoperasian Indonesia. Alur penelitian ini dirancang dengan suatu kerangka kerja sekuensial dan terstruktur, yang diinisiasi dengan fase studi literatur dan akuisisi data. Fase ini dilanjutkan dengan preparasi dan pemrosesan data untuk menghasilkan *dataset* yang representatif dan terstruktur. Tahap inti dari penelitian adalah konfigurasi dan pelatihan model, yang melibatkan serangkaian eksperimen melalui variasi *hiperparameter fine-tuning* guna menemukan konfigurasi optimal. Pasca-pelatihan, dilakukan evaluasi komparatif menggunakan serangkaian metrik kuantitatif. Selanjutnya, model yang menunjukkan performa optimal diimplementasikan dalam sebuah aplikasi prototipe. Fase terakhir adalah analisis hasil, di mana temuan dari evaluasi dianalisis untuk menarik inferensi mengenai efektivitas metodologi yang diterapkan serta kapabilitas model yang dihasilkan.

Alur kerja penelitian ini divisualisasikan dalam diagram alir pada Gambar 1 untuk memberikan gambaran yang jelas mengenai setiap tahapan dan keterhubungannya.



Gambar 1. Diagram Alir Desain Penelitian

B. Akuisisi dan Preparasi Data

Kualitas dan relevansi *dataset* merupakan variabel determinan bagi keberhasilan proses *fine-tuning*. Untuk menjamin akurasi dan konsistensi faktual, penelitian ini melakukan delimitasi terhadap sumber data. Basis pengetahuan fundamental model diekstraksi dari data primer yang bersumber dari buku-buku teks, panduan, dan materi otoritatif lainnya mengenai perkoperasian di Indonesia, guna menjamin landasan faktual yang dapat diverifikasi. Selanjutnya, untuk memperkaya variasi dan volume data latih tanpa mengorbankan akurasi, dilakukan augmentasi melalui data sintetik. Proses ini melibatkan pembuatan pasangan pertanyaan-jawaban baru dengan cara memparafrasekan sampel yang ada, mengajukan pertanyaan dari sudut pandang yang berbeda, dan menggabungkan konsep-konsep dari sumber primer. Pendekatan hibrida ini bertujuan untuk membangun basis pengetahuan yang kuat sekaligus meningkatkan kapabilitas generalisasi model terhadap berbagai gaya bahasa dan formulasi pertanyaan pengguna.

Selanjutnya, dilakukan penyusunan dataset, di mana data dari kedua sumber tersebut diproses menjadi format instruksional yang terdiri dari tiga komponen utama: *context* (potongan teks relevan), *input* (pertanyaan spesifik), dan *output* (jawaban ideal). Setelah penyusunan, struktur dan ukuran dataset ditetapkan. Dataset akhir terdiri dari 900 sampel yang kemudian didistribusikan secara proporsional menjadi 90% (810 sampel) untuk data latih (*training set*) dan 10% (90 sampel) untuk data validasi (*validation set*). Langkah terakhir adalah pemformatan untuk pelatihan, di mana setiap sampel data diformat ke dalam suatu struktur tekstual tunggal yang konsisten dengan menggabungkan *context* dan *input* sebagai instruksi pengguna, dan *output* sebagai respons model, menggunakan token-token pemisah yang spesifik.

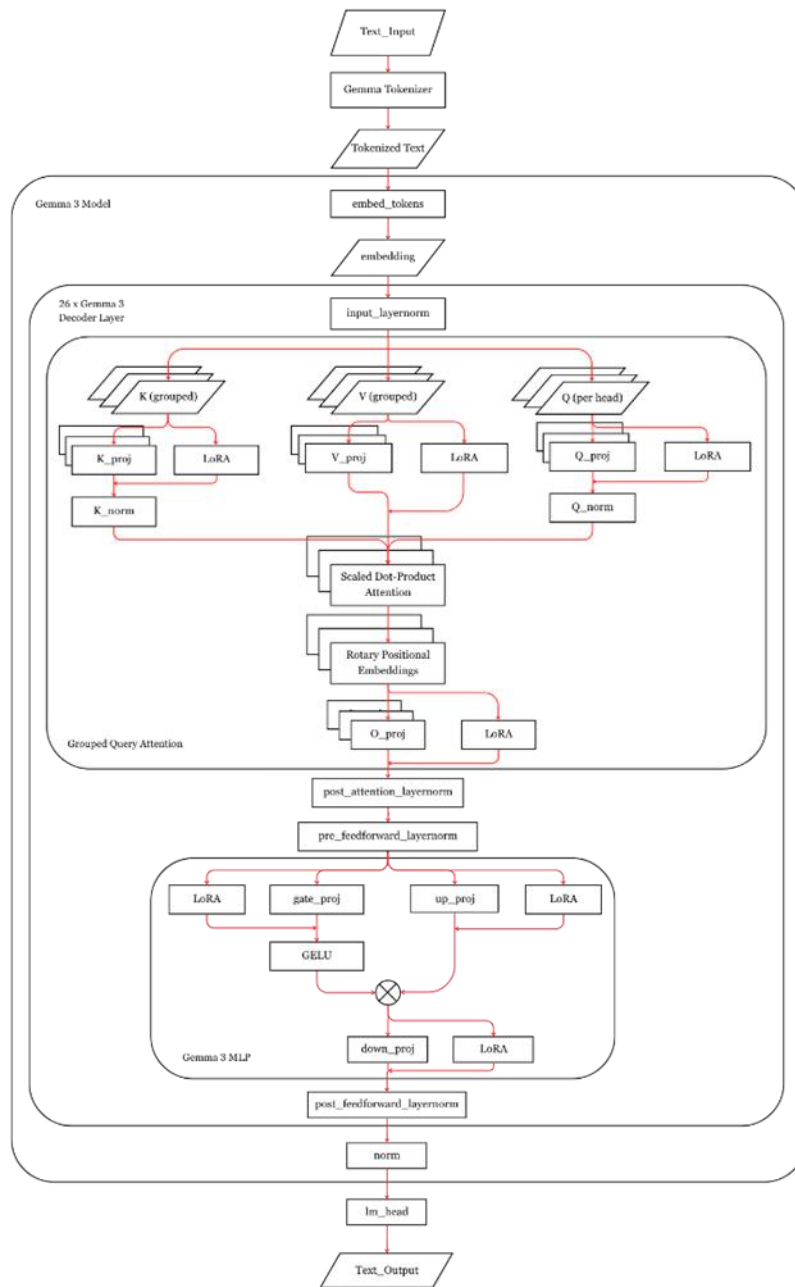
Tabel 1. Contoh Format Sampel Dataset Instruksional

Komponen	Contoh Isi
Konteks (Context)	Koperasi dan Masa Depan
Input (Pertanyaan)	Apa saja model koperasi baru yang akan muncul di masa depan (misalnya, data cooperative)?
Output (Jawaban)	Seiring dengan perubahan teknologi, sosial, dan ekonomi, model koperasi terus berevolusi untuk menjawab tantangan dan peluang baru. Di masa depan, kita akan melihat munculnya model-model koperasi inovatif yang menerapkan prinsip-prinsip kepemilikan bersama pada sektor-sektor yang sebelumnya
Format Akhir	<start_of_turn>user\nBerdasarkan konteks berikut, jawablah pertanyaan yang diberikan.\n\n####Konteks:\nKoperasi dan Masa Depan.\n\n####Pertanyaan: Apa saja model koperasi baru yang akan muncul di masa depan (misalnya, data cooperative)?<end_of_turn>\n<start_of_turn>model\nSeiring dengan perubahan teknologi, sosial, dan ekonomi, model koperasi terus berevolusi untuk menjawab tantangan dan peluang baru. Di masa depan, kita akan melihat munculnya model-model koperasi inovatif yang menerapkan prinsip-prinsip kepemilikan bersama pada sektor-sektor yang sebelumnya

C. Arsitektur Model dan Proses *Fine-Tuning*

Proses *fine-tuning* merupakan inti dari fase eksperimental penelitian ini. Tahapan ini diinisiasi dengan seleksi model dasar, di mana penelitian ini menggunakan *Gemma-3-1B-it*. Selanjutnya, metode adaptasi parameter-efisien diimplementasikan melalui *Low-Rank Adaptation* (LoRA). Untuk menemukan konfigurasi LoRA yang paling efektif untuk domain perkoperasian, penelitian ini merancang serangkaian skenario eksperimental dengan

memvariasikan *hyperparameter* kunci, yakni *rank* (r) dan *alpha* ($lora_alpha$). Konsep injeksi LoRA ke dalam arsitektur model dasar diilustrasikan pada Gambar 2.



Gambar 2. Arsitektur Injeksi LoRA pada Lapisan Transformer dan *Multi Layer Perceptron Model Gemma-3-1B it*

Tiga konfigurasi utama diuji untuk menganalisis dampak dari kapasitas adaptasi (ditentukan oleh r) dan skala pembelajaran (ditentukan oleh $lora_alpha$). Skenario-skenario ini dirangkum dalam Tabel 2. Dalam semua skenario, modul target untuk LoRA dipertahankan secara konstan, yaitu pada lapisan q_proj , k_proj , v_proj , o_proj , $gate_proj$, up_proj , dan $down_proj$.

Seluruh proses pelatihan dan eksperimen dalam penelitian ini dijalankan pada satu unit *Graphics Processing Unit* (GPU) NVIDIA T4 dengan VRAM sebesar 16 GB. Pemilihan

perangkat keras ini merefleksikan skenario penggunaan sumber daya komputasi yang relatif terjangkau dan umum tersedia, yang relevan dengan tujuan penelitian untuk mengembangkan solusi yang efisien. Untuk menjamin perbandingan yang adil antar skenario, proses pelatihan diorkestrasi menggunakan *SFTTrainer* dengan serangkaian *hyperparameter* yang dipertahankan secara konstan. Konfigurasi ini dirangkum dalam Tabel 3.

Tabel 2. Skenario Eksperimen Konfigurasi LoRA

Skenario	Rank (r)	Alpha (lora_alpha)	Rasio Alpha/Rank
1	16	32	2
2	32	64	2
3	64	128	2

Tabel 3. Konfigurasi Hyperparameter Pelatihan (Konstan)

Hyperparameter	Nilai	Hyperparameter	Nilai
per_device_train_batch_size	2	lr_scheduler_type	cosine
gradient_accumulation_steps	4	eval_strategy	steps
warmup_steps	15	eval_steps	5
max_steps	150	load_best_model_at_end	true
learning_rate	2e-04	metric_for_best_model	eval_loss
logging_steps	1	greater_is_better	false
optim	adamw_torch	seed	42
weight_decay	0.01		

D. Prosedur Evaluasi

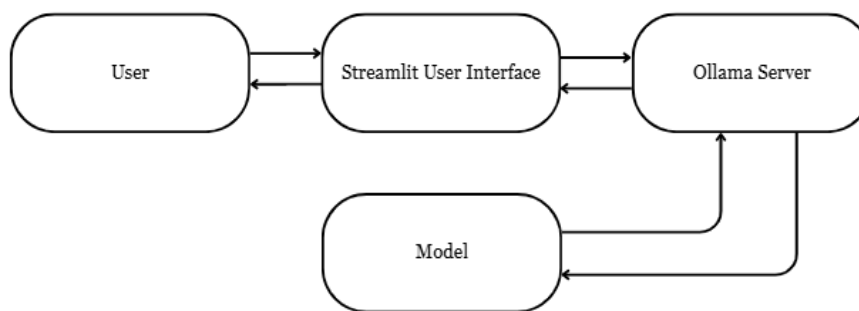
Evaluasi performa untuk setiap model hasil eksperimen dieksekusi secara kuantitatif menggunakan serangkaian metrik standar. Kombinasi dari metrik-metrik ini memfasilitasi suatu analisis komparatif yang holistik untuk menentukan konfigurasi terbaik. Rangkuman metrik evaluasi disajikan pada Tabel 4.

Tabel 4. Rangkuman Metrik Evaluasi

Kategori Metrik	Nama Metrik	Varian / Komponen	Tujuan Evaluasi
Pembelajaran Internal	<i>Loss</i>	<i>Training & Validation</i>	Memantau konvergensi dan keberhasilan proses optimasi model
Kesamaan Leksikal	ROUGE	ROUGE-L	Mengukur tumpang tindih n-gram (kata & frasa) dengan fokus pada <i>recall</i>
Kesamaan Leksikal	BLEU	-	Mengukur tumpang tindih n-gram dengan penalti untuk <i>precision</i>
Kesamaan Semantik	BERTScore	Precision, Recall, F1-Score	Mengukur kesamaan makna kontekstual antara output dan referensi
Efisiensi Komputasi	Penggunaan Sumber Daya	Memori GPU (Bytes), Utilitas GPU (%)	Memvalidasi efisiensi metode PERT yang digunakan

E. Implementasi dan *Deployment Model*

Pasca-identifikasi model dengan performa optimal melalui evaluasi komparatif, tahap berikutnya adalah implementasi dan *deployment*. Tahap ini bertujuan untuk mendemonstrasikan viabilitas praktis dari model yang telah dioptimalkan. Proses ini diinisiasi dengan penyajian model (*model serving*) pada lingkungan lokal menggunakan *Ollama*, yang dipilih disebabkan oleh kemudahan konfigurasinya dalam mengelola dan menyajikan LLM melalui API. Langkah berikutnya adalah pengembangan antarmuka pengguna (UI) berbasis web yang interaktif, yang direalisasikan menggunakan *Streamlit*. Keseluruhan arsitektur *deployment* prototipe ini tersusun atas dua komponen fundamental: *front-end* aplikasi *Streamlit* yang berkomunikasi dengan *back-end* yang dilayani oleh *Ollama*. Arsitektur sederhana ini diilustrasikan pada Gambar 3.



Gambar 3. Arsitektur *Deployment Prototype*

IV. HASIL PENELITIAN DAN PEMBAHASAN

A. Hasil Eksperimen

Performa dari ketiga skenario eksperimen dievaluasi dan dibandingkan dengan performa model dasar (*Gemma-3-1B-it tanpa fine-tuning*) pada tugas *zero-shot*. Hasil ini mencakup metrik pembelajaran internal (*loss, gradient norm, dan learning rate*), metrik kualitas teks generatif (*ROUGE, BLEU, dan BERTScore*), serta metrik efisiensi komputasional.

1. Analisis Kurva Pembelajaran Loss

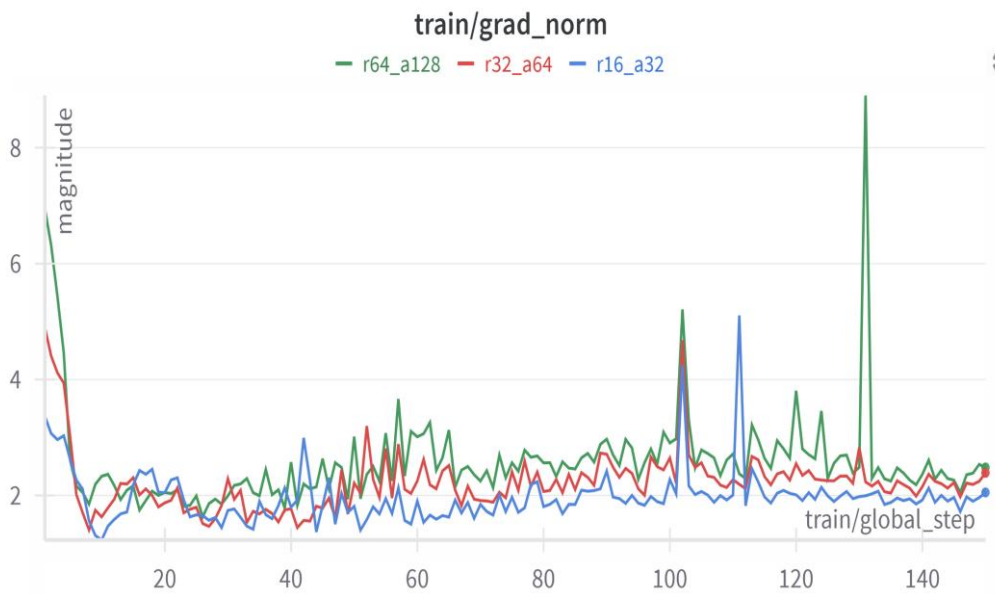
Kurva *training loss* dan *validation loss* dari ketiga skenario eksperimen dipantau selama 150 langkah pelatihan untuk mengobservasi dinamika pembelajaran model. Visualisasi dari kurva-kurva ini disajikan pada Gambar 4. Secara umum, ketiga skenario menunjukkan tren penurunan *loss* yang stabil, sebuah indikasi bahwa model pada setiap konfigurasi berhasil menginternalisasi pola dari *dataset* perkoperasian. Meskipun demikian, terdapat divergensi subtil dalam trajektori kurva validasi. Skenario 1 ($r=16$) dan Skenario 2 ($r=32$) menunjukkan kurva validasi yang menurun secara konsisten bersamaan dengan kurva pelatihan, menandakan proses generalisasi yang baik. Sebaliknya, Skenario 3 ($r=64$) menunjukkan indikasi awal fenomena *overfitting*, di mana kurva *validation loss* mulai mencapai titik jenuh (*plateau*) menjelang akhir pelatihan. Hal ini mengimplikasikan bahwa kapasitas model yang lebih tinggi cenderung melakukan memorisasi terhadap data latih ketimbang melakukan generalisasi pola pada data yang belum pernah dilihat.



Gambar 4. Grafik Perbandingan Kurva Training & Validation Loss

2. Analisis Stabilitas Pelatihan (*Gradient Form*)

Gradient norm mengukur magnitudo dari seluruh gradien dalam model. Gambar 5 menunjukkan bahwa nilai yang stabil dan tidak melonjak secara ekstrem seperti pada model pelatihan skenario 1 dan 2 mengindikasikan bahwa proses pelatihan berjalan dengan baik tanpa mengalami isu seperti *exploding gradients*. Sedangkan pada skenario 3 terjadi *exploding gradients* pada step 133 dengan nilai magnitudo 8,71.



Gambar 5. Grafik Perbandingan Gradient Norm selama Pelatihan

3. Analisis Learning Rate Schedule

Penelitian ini menggunakan penjadwal tingkat pembelajaran (*learning rate scheduler*) tipe kosinus. Penjadwal ini secara dinamis menyesuaikan learning rate selama pelatihan, dimulai dengan fase pemanasan (*warm-up*) singkat di mana *learning rate* meningkat secara *linear*, kemudian diikuti oleh peluruhan (*decay*) secara bertahap mengikuti kurva kosinus hingga mendekati nol di akhir pelatihan. Visualisasi dari jadwal ini disajikan pada Gambar 6. Penggunaan strategi ini bertujuan untuk menstabilkan pelatihan di awal (*fase warm-up*) dan membantu model untuk mencapai konvergensi yang lebih baik dan lebih halus menuju titik minimum pada fungsi *loss* (fase peluruhan kosinus).



Gambar 6. Grafik Learning Rate Schedule

4. Hasil Evaluasi Kualitas Teks Generatif

Evaluasi kuantitatif dieksekusi pada set data validasi untuk mengukur kualitas teks yang dihasilkan oleh model dasar dan model-model dari masing-masing skenario. Hasil komparatif dari evaluasi ini dirangkum secara detail dalam Tabel 5.

Tabel 5. Hasil Evaluasi Komparatif

Metrik	Base Model (Zero-Shot)	Skenario 1 ($r = 16, \alpha = 32$)	Skenario 2 ($r = 32, \alpha = 64$)	Skenario 3 ($r = 64, \alpha = 128$)
ROUGE-L	0,3521	0,7124	0,6932	0,6851
BLEU	0,2133	0,7036	0,6674	0,6687
BERTScore-P	0,8150	0,8691	0,8572	0,8456
BERTScore-R	0,8310	0,8764	0,8563	0,8591
BERTScore-F1	0,8229	0,8681	0,8568	0,8522

5. Analisis Efisiensi Komputasional

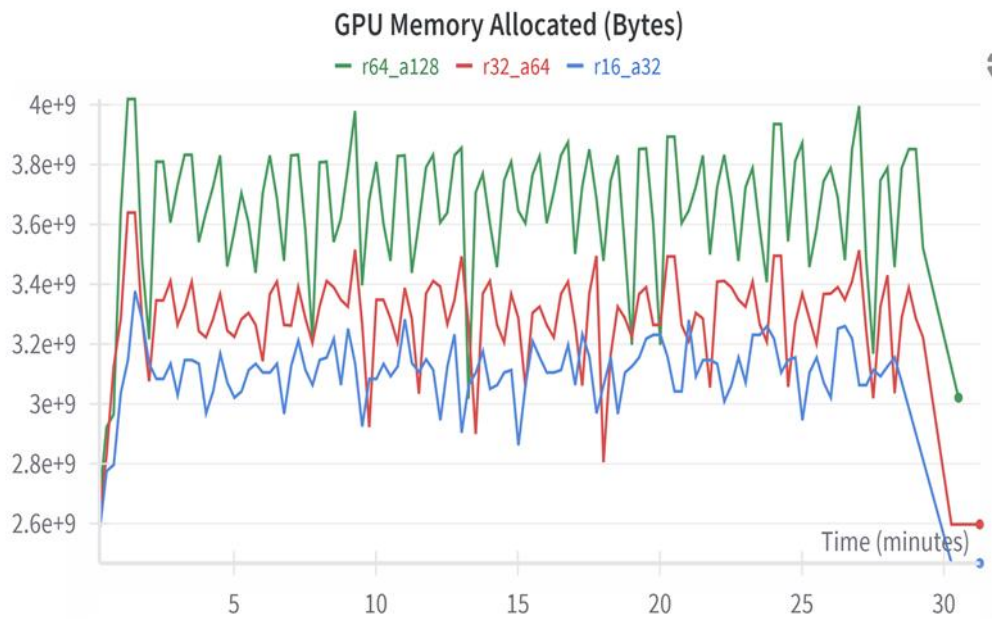
Selain evaluasi kualitas, efisiensi penggunaan sumber daya juga diukur untuk memvalidasi klaim efisiensi dari metode PEFT yang digunakan. Salah satu keunggulan utama LoRA adalah kemampuannya untuk secara drastis mengurangi jumlah parameter yang perlu dilatih dibandingkan dengan full fine-tuning. Tabel 6 menyajikan perbandingan jumlah parameter yang dapat dilatih untuk setiap skenario LoRA terhadap total parameter dari model dasar Gemma-3-1B-it.

Tabel 6. Perbandingan Efisiensi Parameter LoRA

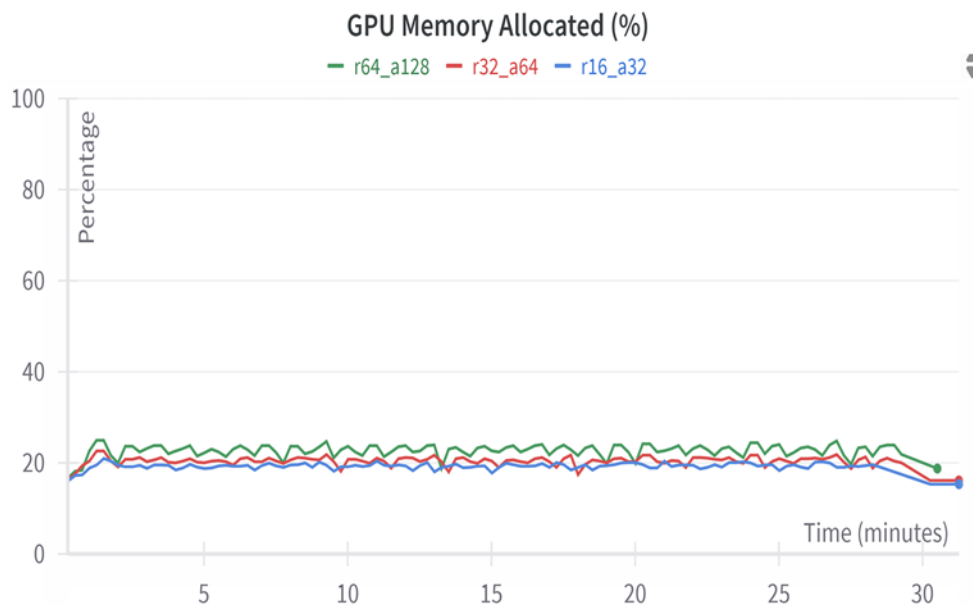
Skenario	Parameter Dapat Dilatih (Trainable)	Total Parameter Model	Persentase Dilatih
1 ($r = 16$)	13.045.760	1.012.931.721	1,29%
2 ($r = 32$)	26.091.520	1.012.931.721	2,57%
3 ($r = 64$)	52.183.040	1.012.931.721	5,15%

Data pada Tabel 6 secara kuantitatif menunjukkan efisiensi yang luar biasa dari metode LoRA. Bahkan pada skenario rank tertinggi ($r=64$), jumlah parameter yang dilatih hanya 5,51% dari total parameter model. Ini membuktikan bahwa adaptasi domain dapat dicapai dengan memodifikasi sebagian sangat kecil dari keseluruhan model.

Selanjutnya, penggunaan memori dan utilitas GPU selama proses pelatihan untuk setiap skenario disajikan dalam Gambar 7 dan 8.



Gambar 7. Grafik Penggunaan Memory GPU/VRAM (Bytes)



Gambar 8. Grafik Alokasi Memory GPU (%)

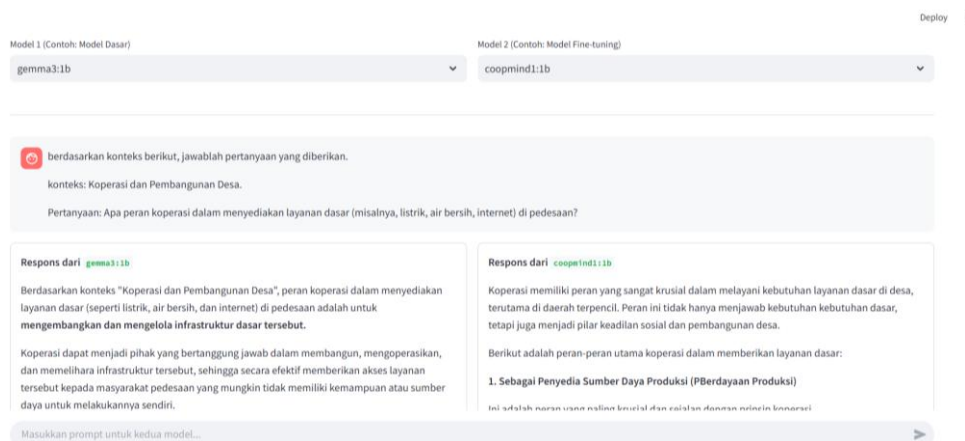
Hasil observasi yang disajikan pada Gambar 7 menunjukkan perbedaan dalam penggunaan Memory GPU/VRAM di mana model dengan $r=16$ menggunakan rata-rata sekitar $3,1 \times 10^9$ Bytes, $r=32$ menggunakan rata-rata sekitar $3,3 \times 10^9$ Bytes, dan $r=64$ menggunakan rata-rata sekitar $3,7 \times 10^9$ Bytes. Jika dipersentasikan, alokasi VRAM untuk ketiga model tersebut hanya berada dikisaran 20% dari total VRAM yang tersedia seperti pada Gambar 8.

B. Pembahasan

1. Analisis Perbandingan Skenario dan Penentuan Model Optimal

Analisis data pada Tabel 5 diawali dengan perbandingan antara performa model dasar (*base model*) dengan model-model yang telah melalui proses *fine-tuning*. Terlihat adanya peningkatan performa yang sangat signifikan pada seluruh metrik setelah proses *fine-tuning*. Sebagai contoh, skor *ROUGE-L* melonjak dari 0,3521 pada *base model* menjadi di atas 0,68 pada semua skenario *fine-tuning*. Peningkatan drastis ini secara kuantitatif membuktikan bahwa model dasar, meskipun memiliki kapabilitas bahasa yang umum, tidak mampu menjawab pertanyaan spesifik domain perkoperasian secara akurat tanpa adaptasi. Proses *fine-tuning* terbukti berhasil menginjeksikan pengetahuan domain spesifik ke dalam model.

Untuk memberikan ilustrasi kualitatif mengenai perbedaan performa ini, Gambar 8 menyajikan perbandingan langsung output yang dihasilkan oleh *base model* dan model hasil *fine-tuning* (Skenario 1) untuk pertanyaan yang sama, sebagaimana ditampilkan pada antarmuka prototipe *Streamlit* dengan prompt “berdasarkan konteks berikut, jawablah pertanyaan yang diberikan. konteks: Koperasi dan Pembangunan Desa. Pertanyaan: Apa peran koperasi dalam menyediakan layanan dasar (misalnya, listrik, air bersih, internet) di pedesaan?”.



Gambar 9. Perbandingan Output Model Base vs Model Coopmind (Fine-Tuned)

Selanjutnya, perbandingan antar skenario *fine-tuning* menunjukkan bahwa Skenario 1 ($r=16, \alpha=32$) secara konsisten menunjukkan superioritas performa pada sebagian besar metrik evaluasi. Skenario ini mencapai skor tertinggi pada seluruh varian *ROUGE* dan *BLEU*, serta mencatatkan skor *BERTScore* F1 tertinggi, yang menandakan tercapainya keseimbangan terbaik antara presisi dan kelengkapan pada level semantik.

Skenario 2 ($r=32$) menunjukkan performa yang kompeten namun secara konsisten berada di bawah Skenario 1. Hal ini dapat diinterpretasikan bahwa kapasitas adaptasi yang lebih rendah ($\text{rank}=16$) sudah cukup memiliki kemampuan untuk menangkap nuansa domain secara komprehensif. Di sisi lain, Skenario 3 ($r=64$), meskipun memiliki kapasitas adaptasi tertinggi, tidak menunjukkan peningkatan performa yang proporsional dan bahkan menunjukkan sedikit penurunan pada beberapa metrik kunci dibandingkan Skenario 2.

Observasi ini, ketika dikorelasikan dengan perilaku kurva validasi pada Gambar 4, memperkuat hipotesis mengenai terjadinya *overfitting*.

Dengan mempertimbangkan peningkatan performa yang masif dibandingkan model dasar serta keunggulan komparatifnya terhadap skenario lain, model dari Skenario 1 ($r=16, \alpha=32$) ditetapkan sebagai model dengan performa optimal untuk tahap implementasi lebih lanjut.

2. Implikasi Hasil Penelitian

Temuan yang dihasilkan dari penelitian ini memiliki serangkaian implikasi teoretis dan praktis yang signifikan. Pertama, hasil ini memvalidasi efektivitas metodologi LoRA untuk melakukan spesialisasi pada *Large Language Model* (LLM). Peningkatan performa yang dramatis dari *base model* ke model yang di-*fine-tune* menunjukkan bahwa adaptasi parameter-efisien merupakan strategi yang sangat efektif dan efisien untuk penguasaan domain spesifik koperasi.

Kedua, pencapaian skor *BERTScore* yang tinggi pada model optimal mengindikasikan bahwa model tersebut tidak hanya melakukan reproduksi leksikal, tetapi juga mampu melakukan komprehensi dan representasi makna secara kontekstual. Ini merupakan implikasi krusial yang mendukung tujuan utama penelitian untuk mendukung demokratisasi pengetahuan. Model yang dihasilkan berpotensi menjadi instrumen pendukung yang andal bagi para praktisi koperasi untuk memperoleh informasi yang relevan secara semantik, melampaui sekadar pencocokan kata kunci.

Ketiga, keberhasilan *deployment* prototipe menggunakan *Ollama* dan *Streamlit* menunjukkan viabilitas teknis dari hasil penelitian ini. Hal ini membuktikan bahwa model yang dikembangkan dapat diintegrasikan ke dalam aplikasi nyata dengan arsitektur yang relatif sederhana, sehingga menginisiasi potensi pengembangan instrumen digital yang dapat diakses secara luas oleh koperasi di seluruh Indonesia.

Keempat, Sejauh penelitian ini dibuat, model digunakan dalam skala terbatas. *Deploy* dan Inferensi baru hanya digunakan oleh Koperasi Keluarga Besar Universitas Koperasi Indonesia.

3. Keterbatasan Penelitian

Perlu diakui bahwa penelitian ini memiliki beberapa keterbatasan metodologis yang dapat menjadi landasan bagi investigasi di masa depan. Pertama, ukuran dan cakupan *dataset* yang berjumlah 900 sampel, meskipun telah dikurasi dengan baik, masih dapat dianggap relatif terbatas. Eskalasi volume data dan diversifikasi sumber (misalnya, studi kasus dari berbagai jenis koperasi) berpotensi meningkatkan robustitas dan kapabilitas model lebih lanjut. Kedua, prosedur evaluasi yang dilakukan bersifat sepenuhnya kuantitatif dan otomatis. Meskipun metrik otomatis memberikan proksi yang baik, evaluasi kualitatif yang melibatkan penilaian oleh ahli domain (pakar koperasi) dan pengguna akhir akan memberikan wawasan yang lebih mendalam mengenai aspek-aspek subtil seperti kejelasan, utilitas praktis, dan tingkat kepercayaan terhadap respons yang dihasilkan model. Ketiga, prototipe yang dikembangkan masih berada pada tahap *proof-of-concept*, deployment sederhana dan baru dilakukan

pengujian dalam skala terbatas serta belum melalui pengujian dalam skenario penggunaan nyata dengan beban kerja yang tinggi atau interaksi pengguna yang beragam.

V. KESIMPULAN

Penelitian ini berhasil membuktikan efektivitas metode *Low-Rank Adaptation* (LoRA) untuk melakukan *fine-tuning* pada *Large Language Model* (LLM) *Gemma-3-1B-it* guna menciptakan model bahasa yang terspesialisasi pada domain perkoperasian di Indonesia. Melalui serangkaian eksperimen, ditemukan bahwa konfigurasi LoRA dengan *rank* 16 dan *alpha* 32 menunjukkan performa optimal, secara signifikan melampaui kemampuan model dasar dalam menjawab pertanyaan spesifik domain dengan akurasi semantik dan leksikal yang tinggi. Validasi kuantitatif melalui metrik *ROUGE*, *BLEU*, dan *BERTScore*, serta keberhasilan implementasi prototipe, mengonfirmasi bahwa adaptasi model yang efisien ini mampu menjembatani kesenjangan pengetahuan dan berpotensi besar untuk mendemokratisasi akses informasi bagi para pemangku kepentingan koperasi, sekaligus menunjukkan viabilitas teknis untuk pengembangan aplikasi praktis di masa depan.

DAFTAR PUSTAKA

- AKSOY, S. I. (2024). Democratizing Education AI and OpenAI Models for Global Access to Knowledge (pp. 79–92). IGI Global. <https://doi.org/10.4018/979-8-3693-1666-5.ch004>
- Anisuzzaman, D.M., Malins, J.G., Friedman, P.A., & Attia, Z.I. (2024). Fine-Tuning Large Language Models for Specialized Use Cases. *Mayo Clinic Proceedings: Digital Health*, 3(1). <https://doi.org/10.1016/j.mcpdig.2024.11.005>
- Chen, Y., Wang, Y., Fu, S., Shen, L., Jing, Y., Tian, X., & Tao, D. (2025). HRP: High-Rank Preheating for Superior LoRA Initialization. *arXiv*. <https://doi.org/10.48550/arXiv.2502.07739>
- Costa, C. J., Aparício, M., Aparicio, S., & Aparício, J. T. (2024). The Democratization of Artificial Intelligence: Theoretical Framework. *Applied Sciences*, 14(18), 8236. <https://doi.org/10.3390/app14188236>
- Dasuki, R. E. (2021). Peningkatan Kapasitas Sumber Daya Manusia Koperasi dan Usaha Kecil Menengah Di Jawa Barat. *Jurnal Ilmiah Abdimas*, 2(1).
- Hu, E.J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., Wang, L., & Chen, W. (2021). LoRA: Low-Rank Adaptation of Large Language Models. *arXiv:2106.09685*.
- Li, S., Yan, J., Wang, H., Tang, Z., Ren, X., Srinivasan, V., & Jin, H. (2023). Instruction-following Evaluation through Verbalizer Manipulation. *arXiv*. <https://doi.org/10.48550/arXiv.2307.10558>
- Lin, C.-Y. (2004). ROUGE: A package for automatic evaluation of summaries. In *Text Summarization Branches Out* (pp. 74–81). Association for Computational Linguistics.
- Lozano, M. B., Martinez, B., & Pindado, J. (2016). Corporate governance, ownership and firm value: Drivers of ownership as a good corporate governance mechanism. *International Business Review*. <https://doi.org/10.1016/j.ibusrev.2016.04.005>
- Lu, W., Luu, R.K., & Buehler, M.J. (2025). Fine-tuning large language models for domain adaptation: exploration of training strategies, scaling, model merging and synergistic capabilities. *npj Computational Materials*, 11, Article number: 84. <https://doi.org/10.1038/s41524-025-01564-y>

- Mai, Q. (M.). (2025). Slimming Down LLMs Without Losing Their Minds. arXiv. <https://doi.org/10.48550/arXiv.2506.10885>
- Mujiyanti, S. (2023). Koperasi Indonesia dan Permasalahannya. *Jurnal Informatika Ekonomi Bisnis*. <https://doi.org/10.37034/infeb.v5i3.653>
- Papadopoulos, C. (2024). Large language models for autistic and neurodivergent individuals: Concerns, benefits and the path forward. *Neurodiversity*, 2. <https://doi.org/10.1177/27546330241301938>
- Papineni, K., Roukos, S., Ward, T., & Zhu, W.-J. (2002). Bleu: A method for automatic evaluation of machine translation. In P. Isabelle, E. Charniak, & D. Lin (Eds.), *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics* (pp. 311–318). Association for Computational Linguistics. <https://doi.org/10.3115/1073083.1073135>
- Purwito, L., Sucipto, Zulkarnain. (2024). Pelatihan Berbasis Kompetensi Sebagai Program Pengembangan Sumber Daya Manusia Bagi Pengelola Koperasi Serba Usaha Syariah Al Mizan Wlingi Di Kabupaten Blitar. *Community Development Journal*, 5(1), 2581 – 2586.
- Saouabe, A., Oualla, H., Mazar, M., & Messadia, M. (2024). Evolution of Human Resources Management Methods: From Traditional Approaches to Generative Artificial Intelligence. 1–5. <https://doi.org/10.1109/dasa63652.2024.10836655>
- Toman Sony, T. H. (2019). *Manajemen Koperasi*. Yrama Widya.
- Zhang, T., Kishore, V., Wu, F., Weinberger, K. Q., & Artzi, Y. (2019). BERTScore: Evaluating text generation with BERT. arXiv. <https://doi.org/10.48550/arXiv.1904.09675>

