Empirical Optimization of Meta-Llama-3.2-3B for the Malaysian TVET Ecosystem: A Case Study in Sovereign AI Specialization

Norazuwa binti Salehudin*, Norwahida binti Saamri, Beny bin Yusmar

Kolej Komuniti Temerloh, Pahang, Malaysia

norazuwa@kkmen.edu.my; norwahida@kkmen.edu.my; beny@kkmen.edu.my

***Abstract*— The General-purpose frontier models frequently exhibit Out-of-Distribution (OOD) hallucinations when applied to localized institutional domains, such as the Malaysian POLYCC (Polytechnics and Community Colleges) and TVET Madani ecosystems. This study presents an empirical hyperparameter optimization of the Meta-Llama-3.2-3B Small Language Model (SLM) to bridge this domain knowledge gap. Utilizing a Supervised Fine-Tuning (SFT) pipeline on AWS SageMaker platform, we specialized the architecture using a high-density dataset of 2,700 synthetically distilled instruction-response pairs. Through a systematic 10-trial experimental sweep, we isolated the interaction between LoRA Rank (*r*), learning rate, and regularization. Performance was validated via a competitive leaderboard-driven evaluation against 50 hidden domain-specific questions. Our findings identify an optimal performance frontier (Trial 7), where a high-capacity configuration (*r*=256), a stable learning rate (5×10^{5}), and a precise dropout (0.03) yielded a winning 64% blind win rate. These results demonstrate that a 3B-parameter model can achieve high reasoning density and factual accuracy when its hyperparameter architecture is aligned with the logical complexity of the target domain. This study provides a validated technical roadmap for the deployment of sovereign, localized AI assistants within the Malaysian TVET ecosystem, establishing a scalable blueprint for broader institutional AI adoption.**

***Keywords*— Synthetic Data Generation, Large Language Models, LoRA, Knowledge Distillation, POLYCC**

I. INTRODUCTION

The rapid advancement of Large Language Models (LLMs) has reached a stage where general-purpose reasoning often conflicts with the need for high-precision domain-specific accuracy. While frontier models—the most advanced, large-scale systems currently available—are proficient in general conversation, they frequently exhibit Out-of-Distribution (OOD) errors when applied to localized institutional data [1]. In the context of the Malaysian POLYCC (Polytechnics and Community Colleges) system, these models lack the internal knowledge required to interpret specialized initiatives such as Technology-Enabled Collaborative Classroom (TECC), Maker Market, and the TVET Madani policy framework. This knowledge gap typically results in hallucinations, where the model generates factually incorrect procedural information that sounds linguistically convincing but is fundamentally inaccurate [2].

To mitigate these risks, this study focuses on the Meta-Llama-3.2-3B model. This is classified as a Small Language Model (SLM)—a compact architecture designed for high efficiency and localized deployment. Recent research indicates that SLMs, despite their smaller parameter counts, can achieve high reasoning density when properly specialized for niche tasks [3]. However, the process of specializing an SLM is highly sensitive to training configurations. We utilize a Supervised Fine-Tuning (SFT) approach where the model is trained on a curated dataset of 2,700 FAQ-formatted instruction pairs to align its outputs with official institutional standards.

Efficiency is achieved through Low-Rank Adaptation (LoRA), a technique that injects a small, trainable adapter into the model rather than updating all 3 billion parameters. While LoRA is resource-efficient, it introduces a significant risk of Catastrophic Forgetting, a phenomenon where the model unlearns its general language abilities while trying to absorb new specialized facts [4]. Preventing this failure requires precise calibration of hyperparameters—manual settings like the Rank (*r*) and Learning Rate that control how much influence the new data has over the model's core logic [5]. Current literature highlights that identifying the optimal configuration for these hyperparameters is a complex, empirical challenge that varies significantly by dataset [3].

This paper presents an Empirical Hyperparameter Study conducted on the AWS SageMaker platform within a competitive benchmarking environment. By analyzing ten distinct experimental trials, we identify the optimal performance frontier (or technical peak) of Llama-3.2-3B for Malaysian TVET domains. Performance was validated through a Leaderboard-Driven Evaluation, where the model was tested against 50 hidden domain-specific questions set by competition organizers. This blind testing ensures that the winning configuration based

*Corresponding Author

on the Blind Win Rate represents the true instructional alignment rather than simple memorization, providing a reproducible roadmap for specialized AI deployment in the public sector.

II. RELATED WORK

The specialization of Small Language Models (SLMs) for niche institutional domains involves a complex interplay between architectural capacity, optimization stability, and objective evaluation. This section reviews current literature regarding these three critical dimensions.

A. Parameter-Efficient Fine-Tuning (LoRA)

The transition from full-parameter fine-tuning to Parameter-Efficient Fine-Tuning (PEFT) has enabled the deployment of high-performance models on edge infrastructure. Low-Rank Adaptation (LoRA) remains the dominant methodology for this task, as it allows for the injection of domain-specific knowledge while keeping the core model weights frozen. This architectural constraint is vital for preventing catastrophic forgetting, a failure state where the model loses its foundational linguistic reasoning while attempting to absorb new facts [1].Recent research into the Meta-Llama-3.2-3B architecture suggests that while its parameter count is relatively low, it possesses significant reasoning density—the ability to perform complex logical inference per parameter—when the LoRA Rank (r) is properly scaled. Specifically, studies in technical domain adaptation indicate that higher ranks (e.g., r=256) are necessary to capture the structural nuances of procedural data without inducing instructional drift [2].

B. Optimization Dynamics and Hyperparameter Entanglement

Fine-tuning success is highly dependent on the calibration of hyperparameters, which emphasizes that the relationship between the Learning Rate and Alpha (α) determines the stability of the training gradient. An aggressive learning rate in SLMs can lead to model collapse, where the model's outputs become repetitive and lose the ability to generalize to unseen prompts [3]. Empirical evidence suggests that for instruction-based datasets, a conservative dropout rate (approximately 0.03) serves as the optimal range for effective regularization, providing enough setup to prevent overfitting on training FAQs while maintaining the flexibility required for zero-shot reasoning [4].

C. Competitive Evaluation and Leaderboard Dynamics

In high-stakes environments like the Malaysian POLYCC ecosystem, traditional metrics (e.g., ROUGE or BLEU) are increasingly viewed as insufficient because they measure word overlap rather than logical accuracy [5]. This has led to the rise of Leaderboard-Driven Development, where models are benchmarked using an LLM-as-a-Judge framework. In this paradigm, a superior model evaluates the candidate's response against a gold-standard reference, resulting in a Pairwise Win Rate [6].

Crucially, competition-based evaluations utilize blind test sets to mitigate the risk of data contamination. This ensures that the win rate represents the model's true capability for instructional alignment and generalization rather than its ability to memorize training patterns. For localized domains like TVET Madani, a blind win rate derived from hidden questions is currently considered the most robust validation of a model's operational reliability [5].

III. METHODOLOGY

The research methodology is structured as a dual-phase architectural process designed as Figure 1 to transform a general-purpose Meta-Llama-3.2-3B model into a domain-specialized expert for the Malaysian POLYCC ecosystem. This process balances linguistic diversity with rigorous technical optimization.

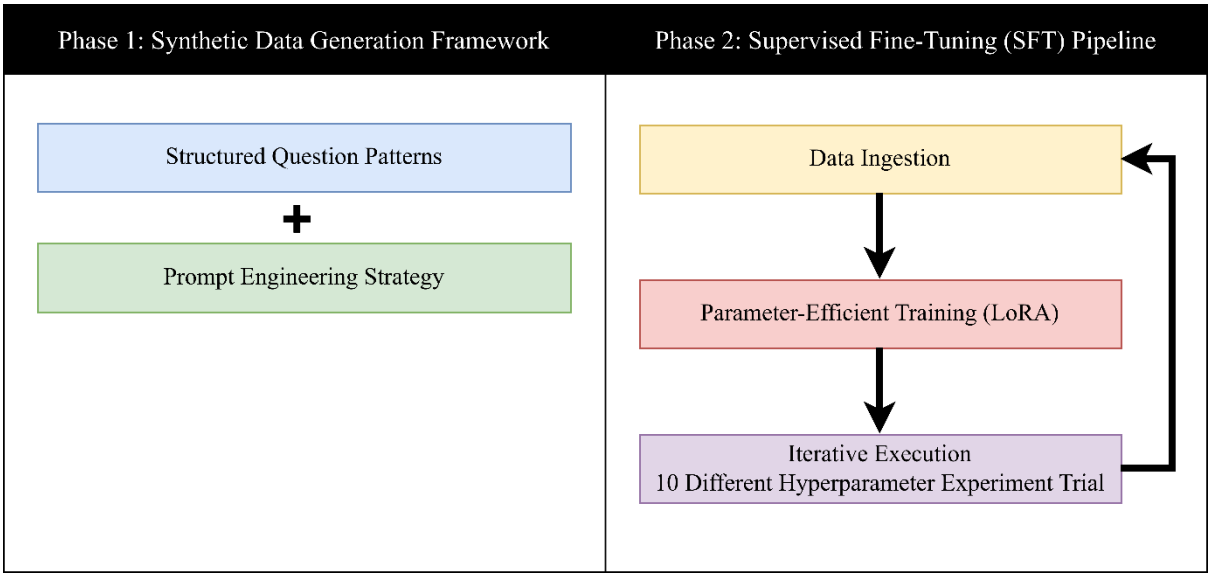

Fig. 1 Research methodology: from synthetic data generation framework to Supervised Fine-Tuning (SFT) pipeline

A. Phase 1: Synthetic Data Generation and Instruction-Response Synthesis

The success of domain-specific specialization is fundamentally dependent on the quality and variance of the instructional signal. To ensure the model developed deep structural reasoning rather than mere pattern memorization, we developed a framework centred on the systematic generation of high-fidelity Instruction-Response pairs.

1) Instructional Pattern Engineering: The 45 Structured Question Patterns engineered designed to encompass the full spectrum of institutional inquiry within the POLYCC ecosystem. These patterns covered direct policy lookups, comparative analysis of Community College versus Polytechnic setup, and hypothetical administrative scenarios. By diversifying these entry points into the data, we ensured the model learned the underlying logic of the whole TVET domain rather than static, one-dimensional answers [4].

2) Prompt Engineering Using Synthetic Instruction-Response Distillation Strategy: Each of the 45 patterns served as a seed for a Model-Based Synthesis process. We utilized a high-parameter frontier model (Gemini) as the generative engine to produce a 2,700-line FAQ-style dataset. Each entry in the dataset was structured as a distinct Instruction-Response pair. The FAQ format was strategically selected to simulate real-world conversational dynamics, where queries range from brief administrative lookups to complex procedural guidance.

This synthesis strategy provided two critical technical advantages:

- Linguistic Diversity in Instructions: The generative process produced multiple variations of how a user might phrase an inquiry (the Instruction). This variance is critical for the model to generalize effectively when faced with the 50 hidden questions during the competition's evaluation phase.
- Procedural Grounding in Responses: The Response component was constrained to maintain a formal Malaysian institutional tone and adhere to official "ground-truth" documentation. This ensured that the Llama-3.2-3B model adopted the specific reasoning and factual accuracy required for government-sector applications [7].

B. Phase 2: Supervised Fine-Tuning (SFT) Pipeline

Phase 2 represents the technical core of the study, where the synthesized Instruction-Response pairs were injected into the Meta-Llama-3.2-3B architecture using an iterative Supervised Fine-Tuning (SFT) pipeline.

1) Infrastructure and Data Ingestion: The 2,700 samples were tokenized and hosted on Amazon S3, creating a high-throughput data bridge to the AWS SageMaker JumpStart environment. The training was executed on an ml.g5.2xlarge instance, featuring the NVIDIA A10G Tensor Core GPU. This infrastructure ensured low-latency data streaming and provided the necessary VRAM to support high-capacity adapter configurations.

2) Parameter-Efficient Fine-Tuning (PEFT): To maintain the model's foundational reasoning and prevent catastrophic forgetting, we utilized Low-Rank Adaptation (LoRA) [1]. By freezing the original 3 billion parameters and training only the low-rank decomposition matrices (adapters), we maximized the model's ability to learn the POLYCC domain without degrading its general linguistic fluency. This method allowed for the use of high Rank (r=256) configurations that would otherwise exceed the memory constraints of full-parameter tuning on edge-class hardware.

3) Empirical Trial Execution and Leaderboard Evaluation: The study concluded with a rigorous execution phase consisting of 10 distinct hyperparameter trials. This iterative approach was designed to isolate specific variables—Rank (*r*), Alpha (α), Learning Rate, and Dropout—to identify the optimal performance frontier for the model. Performance was validated through a Leaderboard-Driven Evaluation. In this competitive framework, each model iteration was subjected to a Black-Box test consisting of 50 hidden domain-specific questions set by the competition organizers. The primary metric was the Pairwise Win Rate, representing the percentage of model responses that accurately met the institutional ground-truth requirements compared to a reference gold standard.

IV. EXPERIMENTAL

The experimental framework was structured to perform a systematic sensitivity analysis of the Meta-Llama-3.2-3B architecture. By isolating key hyperparameters through a controlled sweep, we aimed to identify the saturation point of domain-specific knowledge injection—the threshold where the model achieves maximum factual retention of POLYCC procedures without inducing gradient instability or catastrophic forgetting.

All computational trials were executed within the AWS SageMaker JumpStart environment utilizing an ml.g5.2xlarge instance. This configuration, featuring an NVIDIA A10G Tensor Core GPU with 24GB of VRAM, was selected to provide the necessary memory overhead for high-rank (*r*=256) adapter training. This hardware-software synergy allowed for high-throughput fine-tuning of the 2,700 instruction-response pairs, ensuring that the empirical results were derived from a stable and reproducible high-performance computing (HPC) environment.

TABLE I
SYSTEMATIC HYPERPARAMETER SEARCH SPACE

Hyperparameter	Value Tested	Rationale
Epoch	6, 7, 8	Evaluates the point of factual saturation vs. overfitting risk.
Learning Rate	2e-5, 3e-5, 5e-5, 7e-5	Determines the "step size" of weight updates in the loss landscape.
Lora R (*r)*	64, 128, 256	Defines the total capacity of the trainable adapter matrices.
Lora Alpha (α)	128, 256, 512	Scales the magnitude of the adapter's influence on base weights.
Lora Dropout	0.01, 0.03, 0.05	Acts as a regularizes to prevent memorization of synthetic noise.
Validation Split Ratio	0.2	Allocates 20% of data for internal testing to monitor accuracy.
Batch Size	4 (Eval and Train)	Ensures stable gradient updates on the NVIDIA A10G GPU.
Max Input Length	512	Provides a robust internal benchmark for generalization.

A. Analysis of Hyperparameter Interactions

1) LoRA Rank (r) and Structural Capacity: In this study, the LoRA Rank (r) served as the primary structural proxy for adapter capacity. We transitioned from a conservative rank of 64 to a high-capacity rank of 256 to evaluate the model's ability to encode the 45 complex instructional patterns within its weights. The empirical data suggests that technical domains—characterized by high-dimensional procedural logic—require higher ranks to avoid the bottleneck effect, where the adapter lacks the expressive width to store niche institutional facts [3]. This capacity was modulated by LoRA Alpha (α), which was scaled at a consistent 1:2 ratio ($\alpha = 2r$) in the winning configuration. This scaling is essential to ensure the magnitude of the learned weights is impactful enough to shift the model's internal knowledge base while maintaining numerical stability during training [5].

2) Learning Rate and the Stability Boundary: The modification of the Learning Rate was conducted to identify the Stability Boundary of the Llama-3.2 architecture. Given our constraint of 6 training epochs, we explored whether an aggressive LR (7 x 105) could accelerate convergence or if it would lead to gradient instability. The significant drop in performance in Trials 9 and 10 suggests that higher rates caused the model to overshoot the optimal local minima, resulting in a loss of semantic coherence. Conversely, the success of Trial 7 (5 x105) indicates that this value represents the optimal convergence point, where the model absorbs the 2,700 FAQ lines without destabilizing its foundational linguistic reasoning [1].

3) Regularization via Dropout: The LoRA Dropout was fine-tuned to identify the optimal regularization point that balances generalization with domain fidelity. A dropout rate that is too high prevents the model from absorbing the formal institutional tone required for government-sector applications. However, a rate that is too low (Trial 8) risks a state of over-regularization, where the model fails to capture the nuanced logic of the hidden 50 questions and instead resorts to pattern memorization. By optimizing Dropout to 0.03, we identified a

"Goldilocks" state: high enough to ensure the model could generalize to unseen prompts, but low enough to allow for the high-fidelity absorption of the TVET Madani and POLYCC procedural grounding.

V. RESULT AND DISCUSSION

The experimental evaluation of the Meta-Llama-3.2-3B model yields a compelling narrative on the relationship between hyperparameter architecture and domain-specific intelligence. The empirical data collected across ten distinct trials provide a mechanistic analysis of how variations in capacity, optimization speed, and regularization influenced the model's ability to internalize the complex landscape of Malaysian TVET governance.

TABLE II
CONSOLIDATED RESULTS OF HYPERPARAMETER TRIALS

Trial	Epoch	Learning Rate	Lora Rank	Lora Alpha	Dropout	Win Rate
1 Baseline	6	2e-5	256	512	0.05	58%
2	6	2e-5	128	256	0.05	44%
3	6	2e-5	64	128	0.05	32%
4	6	3e-5	256	512	0.05	60%
5	8	3e-5	256	256	0.05	46%
6	6	5e-5	256	512	0.05	60%
7	**6**	**5e-5**	**256**	**512**	**0.03**	**64%**
8	6	5e-5	256	512	0.01	54%
9	6	7e-5	256	512	0.05	50%
10	6	7e-5	256	512	0.03	50%

A. Model Capacity and Semantic Bottlenecks

A primary finding from our capacity-focused trials (Trial 1, Trial 2, and Trial 3) is that model intelligence in niche domains is strictly gated by the LoRA Rank (r). Transitioning from a rank of 64 to 256 resulted in a significant 26% absolute increase in the Win Rate. This suggests that for technical domains with dense, specialized terminology—such as the POLYCC ecosystem—low-rank adapters act as a semantic bottleneck. The 3-billion parameters of the base model require high-capacity adapter matrices (r=256) to successfully bridge the domain knowledge gap. Without sufficient rank, the model lacks the expressive width required to encode specialized facts alongside its general linguistic reasoning, leading to a failure in instructional alignment [3].

B. Regularization Efficiency and the Optimal Performance Frontier

A pivotal observation occurred in Trial 7, where the model achieved its optimal performance frontier with a 64% Win Rate. This success was not merely a function of optimization speed, but rather the strategic calibration of LoRA Dropout to 0.03.

While standard fine-tuning protocols often utilize a dropout of 0.05 (as seen in Trial 6), our data suggests that for high-density synthetic datasets, a slightly lower dropout allows the model to utilize its adapter capacity more efficiently without falling into knowledge rigidity. Conversely, further reducing dropout to 0.01 (Trial 8) triggered a performance regression (dropping to 54%). This confirms that minimal regularization leads to overfitting on the specific phrasing of the training set, causing the model to prioritize pattern memorization over the underlying policy logic required to answer the 50 hidden competition questions [2].

C. Gradient Instability and Knowledge Collapse

The final trials (Trial 9 and Trial 10) define the Stability Boundary of the Llama-3.2 architecture. By increasing the Learning Rate to $7x10^5$, we observed a significant performance degradation, with the Win Rate falling back to 50%. This phenomenon, characterized as Gradient Overshoot, occurs when weight updates are so aggressive that they destabilize the optimization path. In this state, the model effectively erases the foundational linguistic logic of the pre-trained weights in an attempt to converge too quickly on the new data [4]. This confirms that even with high-rank adapters and high-quality FAQ data, model reliability is ultimately subservient to the stability of the optimization path. Trial 7's Learning Rate of $5x10^5$ represents the maximum stable ceiling for this 3B architecture within the 6-epoch training window.

VI. CONCLUSIONS

This study has successfully defined the optimal performance frontier for specializing a domain-specific Large Language Model (LLM) within a localized institutional framework. Through a rigorous 10-trial experimental matrix conducted on the AWS SageMaker platform, we have demonstrated that achieving superior instructional alignment with the Meta-Llama-3.2-3B architecture requires a precise calibration of the Parameter-Efficient Fine-Tuning (PEFT) search space. Our results confirm that expert-level accuracy in the POLYCC domain is a function

of a specific "technical trifecta": high architectural capacity (r=256$), a stable optimization velocity (5x10^5), and finely tuned regularization (0.03 Dropout).

Our findings provide a critical technical blueprint for Malaysian higher education institutions and government agencies seeking to deploy sovereign, localized AI. We have shown that by optimizing the Supervised Fine-Tuning (SFT) pipeline, a 3B-parameter model can be transformed into a high-fidelity policy advisor that significantly mitigates Out-of-Distribution (OOD) hallucinations while maintaining the low-latency performance required for edge-class deployment. This empirical study proves that Small Language Models (SLMs) can effectively bridge the domain knowledge gap when their hyperparameter configurations are aligned with the high-density logic of institutional procedural data.

Future research will focus on the development of a hybrid architecture, integrating Retrieval-Augmented Generation (RAG) alongside this optimized fine-tuning foundation. By combining parametric knowledge with real-time document retrieval, we aim to exceed an 80% Win Rate, further narrowing the information deficit in the Malaysian public sector and establishing a SOTA standard for localized AI in the TVET Madani ecosystem.

ACKNOWLEDGEMENT

The authors wish to express their sincere gratitude to the Instructional and Digital Learning Division (BIPD), Department of Polytechnic and Community College Education (JPPKK), for their visionary role as the organizer of the competition that catalyzed this research. Their efforts in providing a high-performance benchmarking environment were instrumental in the success of our hyperparameter optimization study.

We also extend our deepest appreciation to the POLYCC ecosystem for providing the domain expertise and institutional documentation required to build the TVET Madani dataset. Finally, we acknowledge Amazon Web Services (AWS) for providing the SageMaker infrastructure and computational resources used throughout the ten experimental trials. This work stands as a testament to the power of collaborative innovation in the Malaysian public sector.

DECLARATION OF GENERATIVE AI USAGE

During the preparation of this work the author(s) used Gemini to idea generation, brainstorming, summarization, paraphrasing, research assistance and factchecking. The author(s) declare that they reviewed and edited the final output as needed and take(s) full responsibility for the content of the published article.

REFERENCES

[1] Y. Grattafiori *et al.*, "Building Domain-Specific Small Language Models via Guided Data Generation," in *Proc. AAAI Conf. Artif. Intell.*, 2025.

[2] I. Mansha, "Resource-Efficient Fine-Tuning of LLaMA-3.2-3B for Medical Chain-of-Thought Reasoning," *arXiv preprint arXiv:2510.05003*, Oct. 2025.

[3] S. E. Baek *et al.*, "Efficient Hyper-Parameter Search for LoRA via Language-aided Bayesian Optimization," *arXiv preprint arXiv:2602.11171*, Jan. 2026.

[4] M. Ahmed *et al.*, "Mitigating Catastrophic Forgetting in Fine-Tuned Large Language Models: An Experimental Study of LoRA and O-LoRA," *Artif. Intell. Digit. Technol.*, vol. 3, no. 1, pp. 52-61, 2026.

[5] C. Springer *et al.*, "(How) Learning Rates Regulate Catastrophic Overtraining," *arXiv preprint arXiv:2604.13627*, Apr. 2026.

[6] J. Li *et al.*, "LLM-as-a-Judge: A Comprehensive Survey on Large Language Model-Based Evaluation Methods," *arXiv preprint arXiv:2512.08472*, 2025.

[7] Z. Huang *et al.*, "HiRA: Parameter-efficient Hadamard High-Rank Adaptation for Large Language Models," in *Proc. 13th Int. Conf. Learn. Represent. (ICLR)*, 2025.