Fine-Tuning Large Language Models for Medical Physics: Domain Adaptation, Statistical Evaluation, and Physics-Informed Analysis

Aulia Octaviani[a,1,2,*]

[a]*Independent Researcher, Yogyakarta, Indonesia*

Abstract

We present a reproducible pipeline to fine-tune open Large Language Models (LLMs) for the medical physics domain (radiation therapy, diagnostic imaging, nuclear medicine, and radiation safety). Using instruction-formatted Q&A with parameter-efficient fine-tuning (LoRA/QLoRA) and 4-bit quantization, we adapt Gemma/Llama-3 class models on Google Colab T4. The resulting model achieves *perplexity* ≈ 5.49 on a held-out set and *domain MCQ accuracy* 66.7% (2/3) on a small physics checklist. We integrate core physical theory (exponential attenuation, half-value layer, interaction cross-sections), radiotherapy planning principles (IMRT), PET system components, and radiation protection (time–distance–shielding) to guide qualitative error analysis. The paper contributes: (i) a compact, open, and auditable procedure; (ii) a physics-informed evaluation rubric; and (iii) portfolio-ready artifacts (notebook, model card, plots) demonstrating readiness for PhD-level research at the interface of AI and medical physics.

Keywords: Medical Physics, Radiation Therapy, Diagnostic Imaging, Nuclear Medicine, LLM Fine-Tuning, LoRA/QLoRA, Statistical Evaluation

1. Introduction

Large Language Models (LLMs) have shown strong generalization in open-domain NLP, yet domain-specific adaptation remains essential for safety-critical fields such as medical physics. Tasks include explaining physical interactions (photoelectric, Compton, pair production), computing attenuation and half-value layer (HVL), outlining IMRT planning steps, describing PET components, and summarizing radiation safety principles. We address these with parameter-efficient fine-tuning on domain Q&A.

Contributions. (1) A reproducible Colab-based fine-tuning pipeline using LoRA/QLoRA and 4-bit quantization; (2) an evaluation suite combining standard NLP metrics (loss, perplexity) and physics-aware checks; (3) detailed theory and equations to ground error analysis in domain physics.

2. Theory Background in Medical Physics

2.1. X-ray Interaction with Matter

For a narrow monoenergetic beam of initial intensity I_0 passing through thickness x,

$$I(x) = I_0\, e^{-\mu x}, \qquad \mu = \text{linear attenuation coefficient (cm}^{-1}\text{)}. \tag{1}$$

Mass attenuation coefficient μ/ρ (with density ρ) separates material dependence. The *half-value layer* (HVL) is

$$\text{HVL} = \frac{\ln 2}{\mu}. \tag{2}$$

Dominant interaction regimes (photon energy E and atomic number Z dependent): photoelectric $\sigma_{\text{pe}} \propto Z^n/E^3$ ($n \approx 3$–4), Compton (approximately Z-independent per electron; Klein–Nishina), and pair production for $E > 1.022\,\text{MeV}$.

2.2. Image Quality and SNR (Poisson Limits)

For quantum-limited detection, counts N are approximately Poisson with $\text{Var}(N) = N$; thus $\text{SNR} \approx N/\sqrt{N} = \sqrt{N}$. Scatter and detector blur degrade contrast-to-noise ratio (CNR); anti-scatter grids improve primary-to-scatter ratio at a dose cost.

2.3. IMRT Planning Essentials

IMRT uses multiple beamlets with modulated fluence to solve an inverse problem:

$$\min_{\boldsymbol{w}\geq 0} \mathcal{L}_{\text{plan}}(\mathbf{D}\boldsymbol{w}), \quad \mathcal{L}_{\text{plan}} = \sum_{v\in\text{PTV}} \phi_{\text{target}}\Big(D_v; D_{\text{presc}}\Big) + \sum_{o\in\text{OAR}} \phi_{\text{oar}}(D_o; D_{\text{lim}}), \tag{3}$$

where $\boldsymbol{w}$ are beamlet weights, $\mathbf{D}$ is the dose influence matrix, PTV is the planning target volume, and OAR denotes organs at risk. Dose constraints appear as penalties or hard bounds; DVH metrics (e.g., D_{95}) summarize coverage.

2.4. PET Physics

A PET system detects near-coincident 511 keV annihilation photons. The prompt rate R_p and randoms R_r obey

$$\text{NECR} = \frac{R_t^2}{R_t + R_s + kR_r}, \tag{4}$$

*Corresponding author. Email: auliaoctavvia@gmail.com. GitHub: github.com/aoctavia

[1] Alumna (M.Sc. Physics), Universitas Sebelas Maret (UNS), 2023.

[2] Maintainer, *physics-archives* (open-source). `https://github.com/aoctavia`

with R_t true coincidences, R_s scatter, and $k \approx 1$–2 accounting for variance inflation; time-of-flight (TOF) improves SNR.

2.5. Radiation Protection

The *time–distance–shielding* triad governs dose reduction. If exposure rate is $\dot{H}_0$ at distance r_0, at distance r (point source),

$$\dot{H}(r) = \dot{H}_0 \left(\frac{r_0}{r}\right)^2 e^{-\mu x}. \tag{5}$$

Absorbed dose D (Gy) is energy per unit mass; *equivalent dose* accounts for radiation weighting w_R; *effective dose* $E = \sum_T w_T H_T$ with tissue weights w_T.

3. Methods

3.1. Data and Instruction Formatting

We curate compact Q&A covering: attenuation/HVL, interaction mechanisms, IMRT planning, PET components, and radiation safety. Each sample is formatted as:

$$\text{Prompt: } \mathsf{Instruction\ about\ medical\ physics} \quad \Rightarrow \quad \text{Response: aligned, concise, correct.} \tag{6}$$

Train/validation split uses stratified sampling to preserve topic coverage.

3.2. Model, Quantization, and LoRA

We adopt instruction-tuned backbones (`google/gemma-1.1-7b-it` or `unsloth/llama-3.2-8b`) with 4-bit quantization (NF4) to fit Colab T4 memory. LoRA introduces a low-rank update

$$\mathbf{W}' = \mathbf{W} + \alpha\, \mathbf{A}\mathbf{B}^\top, \qquad \mathbf{A} \in \mathbb{R}^{d\times r},\ \mathbf{B} \in \mathbb{R}^{k\times r},\ r \ll \min(d,k), \tag{7}$$

applied to attention and MLP projections (e.g., q, k, v, o, up, down, gate). We used rank r=16, α=16, dropout 0.

3.3. Training Objective and Metrics

Given tokenized sequence $\mathbf{x}_{1:T}$ with next-token targets $\mathbf{y}_{1:T}$, the cross-entropy loss is

$$\mathcal{L}_{\mathrm{CE}} = -\sum_{t=1}^{T} \log p_\theta(y_t \mid \mathbf{x}_{\le t}), \tag{8}$$

and *perplexity* is

$$\mathrm{PPL} = \exp\!\left(\frac{1}{T}\sum_{t=1}^{T} -\log p_\theta(y_t \mid \mathbf{x}_{\le t})\right). \tag{9}$$

For domain checks we report MCQ accuracy on a physics mini-benchmark. Further optional metrics include Expected Calibration Error (ECE) and BLEU/ROUGE for long responses.

3.4. Training Configuration and Compute

We train for 100 steps (demo) with batch size 1 and gradient accumulation 4, learning rate 2×10^{-4} (cosine schedule), bf16 where supported. Colab T4 (16 GB) suffices with 4-bit quantization.

4. Experiments

4.1. Setup

Backbone: `gemma-1.1-7b-it` (primary) and `llama-3.2-8b` (unsloth variant). **LoRA:** rank 16, targets $\{q, k, v, o, \text{up}, \text{down}, \text{gate}\}$. **Dataset:** compact instruction Q&A (train n=3, val n=1 in the minimal demo; larger sets planned).

4.2. Monitoring

Weights & Biases (wandb) tracked loss and evaluation; logs are synced locally when WAN unavailable.

5. Results

5.1. Language Modeling

Final training loss ≈ 0.4743; validation loss ≈ 1.7037; computed perplexity ≈ 5.49.

5.2. Domain MCQ

Three canonical questions (absorbed dose unit, IMRT acronym, I-131 half-life) yielded 2/3 correct (66.7%). The absorbed-dose unit error (answered Sv instead of Gy) highlights the need for unit-grounding augmentation.

5.3. Inference Samples (Exact Text Not Truncated)

Photoelectric Effect (Imaging).. *"The photoelectric effect describes how an X-ray photon transfers* all *of its energy to a bound inner-shell electron, ejecting it (photoelectron) and leaving a vacancy that relaxes via characteristic X-rays/Auger electrons. Its probability scales roughly as Z^3/E^3, dominating at lower energies and high-Z materials, thereby enhancing subject contrast in diagnostic radiology (e.g., with contrast agents)."*

HVL at 100.0000 keV in Al (Concept).. *"The half-value layer (HVL) is the thickness that reduces intensity to one half:* $\mathrm{HVL} = \ln 2/\mu$. *For 100.0000 keV photons in Al, use tabulated μ or μ/ρ and density to compute a numerical HVL. Attenuation follows $I(x) = I_0 e^{-\mu x}$; beam quality and buildup must be considered in broad-beam geometry."*

IMRT Planning (Principles).. *"Define PTV and OAR on CT/MRI, prescribe dose, set constraints, and solve the inverse planning problem (fluence modulation) with MLC-delivered beamlets. Evaluate DVH metrics (D_{95}, OAR limits), perform QA, and deliver with image guidance to ensure geometric accuracy."*

PET Components.. *"Cyclotron (isotope production), radiopharmaceutical injection/dose preparation, and detector ring with fast scintillators/SiPMs enabling TOF coincidence detection and list-mode acquisition."*

Radiation Safety.. *"Apply time–distance–shielding; optimize protocols (ALARA), secure storage/transport/disposal of sources, monitor with dosimetry, and train staff regarding biological effects and emergency procedures."*

6. Error Analysis (Physics-Informed)

The model occasionally confuses *absorbed* and *equivalent* dose units (Gy vs Sv). We propose: (i) unit-focused contrastive pairs, (ii) formula-anchored prompting (e.g., append (1)–(2) when asking HVL), and (iii) retrieval augmentation from vetted sources (AAPM task group reports).

7. Safety, Ethics, and Intended Use

This LLM is intended for *education and research support*, not clinical decision-making. Outputs should be verified by qualified medical physicists. We recommend guardrails: refusal for dosage prescriptions, highlighting uncertainty, and linking to references.

8. Limitations and Future Work

Current data volume is small; we plan to scale instruction pairs, incorporate retrieval (RAG) with citation, add unit-consistency regularizers, expand evaluations (nDCG/Recall for retrieval, ECE for calibration), and explore multimodal (text+image) fusion for CT/MRI/PET contexts.

9. Reproducibility and Resources

Code & Model Card: https://github.com/aoctavia/FT-LLM
https://huggingface.co/<<your-hf-username>>/medical-physics-llm
Notebook (Colab): <<colablinktothenotebook>>
Artifacts: training logs, final checkpoints, model_info.json, and plots under figures/.

Acknowledgements

We thank open-source contributors of Transformers, PEFT, TRL, bitsandbytes, Unsloth, and the medical physics community for publicly available educational resources.

References

[1] J.H. Hubbell, Photon mass attenuation and energy-absorption coefficients, *Int. J. Appl. Radiat. Isot.* 33 (1982) 1269–1290.

[2] American Association of Physicists in Medicine (AAPM), Task Group Reports (various), *AAPM Reports*, 2001–2025.

[3] T. Bortfeld, IMRT: a review and preview, *Phys. Med. Biol.* 51 (2006) R363.

[4] S.R. Cherry, J.A. Sorenson, M.E. Phelps, *Physics in Nuclear Medicine*, 3rd ed., Saunders, 2003.

[5] E.J. Hu, et al., LoRA: Low-Rank Adaptation of Large Language Models, arXiv:2106.09685, 2021.

[6] T. Dettmers, et al., LLM.int8(): 8-bit Matrix Multiplication for Transformers at Scale, arXiv:2208.07339, 2022.

Appendix A. Practical HVL Computation (Worked Example)

Given 100.0000 keV photons and tabulated μ/ρ for Al, compute $\mu = (\mu/\rho)\rho$; then $\mathrm{HVL} = \ln 2/\mu$. For broad-beam setups include scatter buildup; report narrow-beam and broad-beam HVL distinctly.