

PUBLICATIONS AND PROJECTS

- [Thesis, Ongoing] **Effect of Multiple Reasoning Paths**. Investigating how providing multiple valid reasoning trajectories for a single mathematical problem affects the reasoning performance and output diversity of language models trained through supervised fine-tuning, as well as its impact on reinforcement learning.
- [UGRIP, Ongoing] **LM Uncertainty Estimation**. Extended the LM-Polygraph framework to support chain-of-thought reasoning models and large-scale uncertainty evaluation. Benchmarked perplexity, predictive entropy, and sequence-level entropy as indicators of reasoning correctness across GSM8K, MMLU, and MedMCQA.
- [MRL 2025, Nov. 2025] **Entropy2Vec**. Developed a novel unsupervised framework to create dense and dynamic language representations from the cross-entropy of monolingual LMs (e.g., GPT-2). This method avoids the sparsity and static nature of traditional typological databases (like URIEL). The resulting vectors proved highly competitive with curated baselines on downstream NLP benchmarks (SIB-200, MASSIVE) and in reconstructing linguistic typologies, achieving a state-of-the-art LCA MAE score against the Glottolog tree.
- [Coursework, Jun. 2025] **Bilingual NLI**. Fine-tuned transformer-based language models for Indonesian and English natural language inference (NLI). Benchmarked BERT, RoBERTa, DeBERTa, and ModernBERT architectures, and evaluated the effectiveness of the Muon optimizer on multilingual NLI performance.
- [Coursework, May. 2025] **Cendekia-R1**. Reproduced DeepSeek-R1-Zero on a smaller scale using Qwen2.5-Math-1.5B. Built a mathematical reasoning dataset from university entrance examination problems (SIMAK UI, UTUL UGM) using Mistral OCR and Gemini-2.5-Pro for answer annotation. Trained the model with the [OpenR1](#) framework using Group Relative Policy Optimization (GRPO) objective.
- [Coursework, Dec. 2024] **AbstractFinder**. Developed and deployed a search engine for arXiv research abstracts. Built a FastAPI backend on Google Cloud Platform and a Next.js frontend on Vercel, and implemented a two-stage retrieval pipeline using BM25 (PyTerrier) and BERT-based reranking to improve search relevance across thousands of research papers.

EXPERIENCE

- Undergraduate Research Internship Program [MBZUAI](#) Jun. 2025 – Jul. 2025
- Took part in the "Benchmarking Uncertainty Quantification for Reasoning LLMs" project, supervised by Roman Vashurin, Maxim Panov, and Aleksandr Rubashevskii.
- Mentee of SeaCrowd Apprenticeship Program [SeaCrowd](#) Aug. 2024 – Aug. 2025
- Participated as a mentee in writing *Entropy2Vec* paper, which is a novel framework to derive a cross-lingual language representations using monolingual language models. Worked on the clustering algorithm to construct the language tree for comparison to Glottolog as the ground truth.

EDUCATION

- Bachelor of Computer Science [Universitas Indonesia](#) Aug. 2022 – Present
- 8th semester, **Cumulative GPA**: 3.81/4.00, with 156 study credits taken.

HIGHLIGHTED HONOURS

- [National, Nov. 2024] **Silver Medalist of Gemastik 17 Data Mining by Ministry of Education**. Competed against hundreds of teams from different universities across Indonesia. Built medical question-answering system that employs retrieval augmented generation to enhance reply accuracy.
- [Regional, Nov. 2023] **Finalist of Intercollegiate Competitive Programming Contest (ICPC) Regional Asia Jakarta**. Competed with some of the best CS students in Asia to solve 13 problems within the span of 5 hours.