

**PENGEMBANGAN CHATBOT AKADEMIK BERBASIS LARGE
LANGUAGE MODEL DENGAN FINE-TUNED MISTRAL 7B INSTRUCT
V0.3 PADA DOMAIN AKADEMIK UNIVERSITAS JENDERAL
ACHMAD YANI YOGYAKARTA**

Ria kristi¹, Puji Winar Cahyo²

INTISARI

Latar Belakang: Unjaya telah menyediakan pedoman akademik digital, namun belum memiliki sistem untuk akses informasi yang cepat dan interaktif. Pencarian informasi yang bersifat spesifik atau sering ditanyakan masih dilakukan secara manual melalui daftar isi atau pencarian teks, yang kerap tidak langsung mengarah pada jawaban yang eksplisit.

Tujuan: Penelitian ini bertujuan melakukan fine-tuning LLM Mistral 7B Instruct v0.3 menggunakan data pedoman akademik Unjaya untuk membangun fondasi *chatbot* akademik yang relevan dan kontekstual, serta mengevaluasi kemampuannya dalam merespons pertanyaan secara akurat.

Metode Penelitian: Mengadopsi metodologi *Cross Industry Standard Process for Machine Learning with Quality Assurance* (CRISP-ML(Q)), yang terdiri dari enam tahapan utama: *business and data understanding*, *data preparation*, *modeling*, *evaluation*, *deployment*, serta *monitoring and maintenance*. Setiap tahapan disesuaikan dengan ruang lingkup penelitian yang mencakup proses *fine-tuning* dan evaluasi model.

Hasil: Penelitian menghasilkan model Mistral 7B Instruct v0.3 yang di-*fine-tune* menggunakan QLoRA pada domain Pedoman Akademik Unjaya. Pelatihan selama 1 *epoch* menghasilkan *training loss* sebesar 0.99 dan *validation loss* sebesar 0.94. Evaluasi dengan BERTScore menunjukkan skor F1 sebesar 0.6421, *precision* 0.6323, dan *recall* 0.6525. Model telah dipublikasikan di Hugging Face Hub dan disiapkan untuk integrasi dalam sistem *chatbot* akademik.

Kesimpulan: Model hasil *fine-tuning* menunjukkan potensi adaptasi terhadap domain Pedoman Akademik Unjaya, meskipun masih terdapat keterbatasan performa dan indikasi *hallucination*. Penelitian ini menjadi landasan awal pengembangan *chatbot* akademik berbasis LLM untuk mendukung sistem informasi di perguruan tinggi.

Kata-kunci: *Large Language Model*, *Fine-tuning*, Mistral 7B Instruct, QLoRA, *Chatbot* Akademik, Pedoman Akademik, Hugging Face

¹ Mahasiswa Informatika (S-1) Universitas Jenderal Achmad Yani Yogyakarta

² Dosen Informatika (S-1) Universitas Jenderal Achmad Yani Yogyakarta

**DEVELOPMENT OF AN ACADEMIC CHATBOT BASED ON LARGE
LANGUAGE MODEL WITH FINE-TUNED MISTRAL 7B INSTRUCT V0.3
ON THE ACADEMIC DOMAIN OF UNIVERSITAS JENDERAL ACHMAD
YANI YOGYAKARTA**

Ria Kristi¹, Puji Winar Cahyo²

ABSTRACT

Background: *Unjaya has provided a digital academic guideline, but does not yet have a system for fast and interactive information access. Searches for specific or frequently asked information are still conducted manually through the table of contents or text search, which often do not directly lead to explicit answers.*

Objective: *This research aims to fine-tune the Mistral 7B Instruct v0.3 LLM using data from Unjaya's academic guidelines to establish a foundation for a relevant and contextual academic chatbot, as well as to evaluate its ability to respond accurately to user queries.*

Method: *The research adopts the Cross Industry Standard Process for Machine Learning with Quality Assurance (CRISP-ML(Q)) methodology, which consists of six main phases: business and data understanding, data preparation, modeling, evaluation, deployment, and monitoring and maintenance. Each phase is tailored to the scope of the study, focusing on the fine-tuning and evaluation process of the model.*

Result: *The research produced a fine-tuned Mistral 7B Instruct v0.3 model using QLoRA, specifically adapted to the academic guideline domain of Unjaya. Training over one epoch resulted in a training loss of 0.99 and a validation loss of 0.94. Evaluation using BERTScore yielded an F1-score of 0.6421, precision of 0.6323, and recall of 0.6525. The model has been published on Hugging Face Hub and is ready for further integration in the development of an academic chatbot system.*

Conclusion: *The fine-tuned model demonstrates potential for adapting to the domain of Unjaya's Academic Guidelines, although performance limitations and indications of hallucination are still present. This research serves as an initial foundation for developing an LLM-based academic chatbot to support information systems in higher education institutions.*

Keywords: *Large Language Model, Fine-tuning, Mistral 7B Instruct, QLoRA, Academic Chatbot, Academic Guideline, Hugging Face*

¹ Student of Informatics (S-1) Universitas Jenderal Achmad Yani Yogyakarta

² Lecturer of Informatics (S-1) Universitas Jenderal Achmad Yani Yogyakarta