

Volume produk yang sangat besar dalam e-Katalog LKPP menimbulkan tantangan bagi instansi pemerintah untuk menemukan barang/jasa yang paling relevan dengan cepat dan tepat. Pencarian konvensional berbasis kata kunci sering kali kurang efektif menangani keragaman istilah dan konteks kebutuhan pengguna, terutama dalam domain pengadaan pemerintah yang memiliki kekhususan. Oleh karena itu, diperlukan sistem rekomendasi yang lebih cerdas dan kontekstual, yang mampu memahami maksud *query* secara mendalam dan menyajikan produk yang benar-benar sesuai. Penelitian ini bertujuan meningkatkan kualitas sistem rekomendasi produk pada e-Katalog pemerintah dengan memanfaatkan pendekatan *Retrieval-Augmented Generation* (RAG), yang menggabungkan pencarian semantik berbasis *embedding* dengan model *Large Language Model* (LLM) untuk menghasilkan keluaran narasi rekomendasi.

Solusi yang diusulkan meliputi pengembangan *pipeline* dua tahap. Pertama, model *SentenceTransformer* (all-mpnet-base-v2) di-*fine-tune* menggunakan Contrastive Learning pada data pasangan *query*–deskripsi produk e-Katalog. Proses ini menghasilkan representasi vektor semantik yang dioptimasi untuk domain pengadaan pemerintah, sehingga sistem dapat melakukan *semantic retrieval* yang lebih akurat dibanding model *pretrained* umum. Kedua, modul generatif dengan GPT-4 diintegrasikan untuk *query rewriting* dan generasi narasi rekomendasi. Model GPT-4 diberi instruksi khusus untuk menyusun ulang kalimat *query* pengguna menjadi lebih jelas dan spesifik, lalu menghasilkan rekomendasi dalam format narasi deskriptif. Narasi tersebut memuat penjelasan mengenai fitur utama produk yang direkomendasikan serta alasan kesesuaiannya dengan kebutuhan pengguna, termasuk pertimbangan konteks kebijakan. Kontribusi utama penelitian ini terletak pada integrasi *embedding* semantik yang disesuaikan domain dengan kemampuan generatif LLM dapat meningkatkan relevansi sekaligus interpretabilitas sistem rekomendasi e-Katalog pemerintah.

Hasil pengujian menunjukkan bahwa pendekatan ini berhasil meningkatkan kinerja sistem rekomendasi secara signifikan dibandingkan baseline. Model *SentenceTransformer* yang telah di-*fine-tune* mencapai skor evaluasi tinggi pada tugas *retrieval*: nilai ROC AUC $\sim 0,95$ dan *Average Precision* $\sim 0,82$ pada data uji, yang mengindikasikan kemampuan model dalam membedakan pasangan *query*–produk relevan versus tidak relevan sudah sangat baik. Dampak *fine-tuning* dan *query rewriting* terhadap skenario rekomendasi top-k juga positif: metrik *Precision@10* meningkat dari 0,345 (baseline tanpa *fine-tuning/rewriting*) menjadi 0,638, dan *nDCG@10* naik dari 0,727 menjadi 0,936. Temuan ini menggarisbawahi potensi penerapan RAG dalam e-Katalog dengan integrasi teknik *semantic retrieval* dan *generative explanation* terbukti efektif meningkatkan relevansi rekomendasi dan membantu pengguna memahami alasan di balik rekomendasi tersebut. Meskipun masih ditemui tantangan seperti kualitas data produk dan relevansi hasil pencarian produk yang perlu ditingkatkan.

Kata Kunci: sistem rekomendasi; e-katalog; *retrieval-augmented generation*; *embedding* semantik; contrastive learning; GPT-4

ABSTRACT

The vast volume of products listed in Indonesia's government e-Catalog (LKPP) presents a significant challenge for public institutions to efficiently and accurately identify the most relevant goods or services. Traditional keyword-based search methods often fall short in handling the diversity of terminology and the contextual needs of users, particularly within the specific requirements of government procurement. Consequently, a more intelligent and contextual recommendation system is needed, one capable of deeply understanding user intent and presenting truly relevant product options. This study aims to improve the quality of product recommendation systems in the government e-Catalog by employing a Retrieval-Augmented Generation (RAG) approach, which combines semantic search based on embeddings with Large Language Models (LLMs) to generate narrative recommendations.

The proposed solution consists of a two-stage pipeline. First, a SentenceTransformer model (all-mpnet-base-v2) is fine-tuned using Contrastive Learning on paired data of queries and product descriptions from the e-Catalog. This process yields semantically optimized vector representations tailored for the government procurement domain, enabling more accurate semantic retrieval than general pretrained models. Second, a generative module using GPT-4 is integrated to perform query rewriting and generate descriptive recommendation narratives. GPT-4 is specifically instructed to reformulate user queries into clearer and more specific expressions, followed by generating recommendation texts that explain the core features of the suggested products and the reasons for their relevance, considering policy-related aspects where applicable. The main contribution of this research lies in demonstrating that integrating domain-specific semantic embeddings with the generative capabilities of LLMs can enhance both the relevance and interpretability of the recommendation system in the government e-Catalog context.

Evaluation results show that this approach significantly improves recommendation system performance compared to the baseline. The fine-tuned SentenceTransformer model achieved high evaluation scores for retrieval tasks, with a ROC AUC of approximately 0.95 and an Average Precision of around 0.82, indicating strong capability in distinguishing between relevant and non-relevant query-product pairs. The impact of fine-tuning and query rewriting on top-k recommendation scenarios was also positive: Precision@10 increased from 0.345 (baseline) to 0.638, and nDCG@10 improved from 0.727 to 0.936. These findings highlight the potential of applying RAG in the e-Catalog domain, demonstrating that the integration of semantic retrieval and generative explanation effectively enhances recommendation relevance while helping users understand the rationale behind the recommendations. Nonetheless, challenges remain, such as improving product data quality and ensuring the continued relevance of retrieved results.

Keywords: recommendation system; e-catalog; retrieval-augmented generation; semantic embedding; contrastive learning; GPT-4