


Jutisi: Jurnal Ilmiah Teknik Informatika dan Sistem Informasi
<https://ojs.stmik-banjarbaru.ac.id/index.php/jutisi/index>
 Jl. Ahmad Yani, K.M. 33,5 - Kampus STMIK Banjarbaru
 Loktabat – Banjarbaru (Tlp. 0511 4782881), e-mail: puslit.stmikbjb@gmail.com
 e-ISSN: 2685-0893

Implementasi *AI Agent* Berbasis RAG untuk Klasifikasi Pelanggan KBLI 2025

DOI: <http://dx.doi.org/10.35889/jutisi.v15i3.3739>

Creative Commons License 4.0 (CC BY – NC) 

Teddy Agustinus¹, Jap Tji Beng^{2*}, Rahmiyana Nurkholiza³, Vienchenzia Oeyta Dwitama Dinatha⁴

^{1,2}Program Studi Sarjana Sistem Informasi, Universitas Tarumanagara, Jakarta

³Binus Graduate Program, Universitas Bina Nusantara, Jakarta, Indonesia

⁴Faculty of Business & Communications, Doctor of Business Administration Programme, INTI International University, Malaysia

^{2,3,4}Laboratorium Kognisi Edukasi dan Inovasi Teknologi, Universitas Tarumanagara, Indonesia
 *e-mail Corresponding Author: t.jap@untar.ac.id

Abstract

The migration of the ERP system from SAP ECC to SAP S/4HANA requires customer master data to have a complete industry classification that complies with the 2025 Indonesian Standard Industrial Classification (KBLI) to support the implementation of the Business Partner model. This research aims to develop an AI Agent based on Retrieval-Augmented Generation (RAG) to automate the customer classification process within the Master Data Management unit at PT X. The research employs an implementational approach by building a RAG system that integrates the KBLI 2025 knowledge base, hybrid search, and a large language model to classify 235 active customers into five-digit KBLI codes. The system's performance was evaluated based on the accuracy of the classification results. The research findings indicate that the system achieved an accuracy of 98 per cent and successfully mapped 50 customer line-of-business categories into 21 KBLI Level 1 sectors. These findings demonstrate that the AI Agent is capable of generating standardised customer master data and is ready to support the migration to SAP S/4HANA.

Keywords: *AI Agent; Retrieval-Augmented Generation; Hybrid Search; Customer Classification; KBLI 2025*

Abstrak

Migrasi sistem ERP dari SAP ECC ke SAP S/4HANA menuntut data master pelanggan memiliki klasifikasi industri yang lengkap dan sesuai dengan standar Klasifikasi Baku Lapangan Usaha Indonesia (KBLI) 2025 agar mendukung implementasi model Business Partner. Penelitian ini bertujuan mengembangkan AI Agent berbasis *Retrieval-Augmented Generation* (RAG) untuk mengotomasi proses klasifikasi pelanggan di unit *Master Data Management* PT X. Penelitian menggunakan pendekatan implementatif dengan membangun sistem RAG yang mengintegrasikan basis pengetahuan KBLI 2025, pencarian hibrida, dan large language model untuk mengklasifikasikan 235 pelanggan aktif ke kode KBLI level lima digit. Kinerja sistem dievaluasi berdasarkan tingkat akurasi hasil klasifikasi. Hasil penelitian menunjukkan bahwa sistem mencapai akurasi sebesar 98% serta berhasil memetakan 50 kategori *line of business* pelanggan ke dalam 21 sektor KBLI Level 1. Temuan ini menunjukkan bahwa AI Agent mampu menghasilkan data master pelanggan yang terstandarisasi dan siap mendukung migrasi ke SAP S/4HANA.

Kata kunci: *AI Agent; Retrieval-Augmented Generation; Hybrid Search; Klasifikasi Pelanggan; KBLI 2025*

1. Pendahuluan

Perkembangan pesat teknologi informasi dalam satu dekade terakhir telah mendorong organisasi untuk memandang data bukan sekadar arsip operasional, melainkan sebagai aset strategis yang mampu menghasilkan keunggulan kompetitif apabila dikelola secara tepat dan

sistematis [1]. Kemampuan organisasi mengekstraksi nilai dari data secara konsisten menghasilkan kapabilitas pengambilan keputusan yang lebih akurat dan berbasis bukti. Dalam konteks ini, *Master Data Management* (MDM) menjadi disiplin yang sangat krusial karena memastikan data inti organisasi, seperti data produk, pelanggan, dan pemasok, dikelola secara konsisten, akurat, dan terintegrasi di seluruh sistem informasi [2]. Keberhasilan implementasi MDM juga sangat dipengaruhi oleh tata kelola data (*data governance*) yang mampu menjamin kualitas, konsistensi, dan interoperabilitas data lintas proses bisnis [3], [4]. Implementasi MDM yang kurang optimal berpotensi menimbulkan inkonsistensi data, duplikasi entitas, kesalahan pelaporan, hingga kegagalan integrasi yang dapat mengganggu kelangsungan operasional perusahaan.

Urgensi pengelolaan data master pelanggan semakin meningkat seiring inisiatif transformasi digital yang melibatkan migrasi sistem ERP dari SAP ECC ke SAP S/4HANA. Migrasi ini mensyaratkan kualitas dan kelengkapan data master yang jauh lebih tinggi karena SAP S/4HANA mengadopsi model *Business Partner* yang menyatukan entitas pelanggan dan pemasok dalam satu struktur data terpadu, serta mengharuskan setiap pelanggan memiliki atribut klasifikasi industri yang lengkap dan terstandarisasi sebelum data dapat dimigrasikan [3]. Leovin et al. [5] dalam pengembangan sistem penjualan *business-to-business* berbasis web menunjukkan bahwa kualitas data entitas pelanggan merupakan fondasi penting bagi keberhasilan operasional sistem enterprise. Pada unit *Master Data Management* PT X, proses klasifikasi pelanggan ke dalam Klasifikasi Baku Lapangan Usaha Indonesia (KBLI) masih dilakukan secara manual berdasarkan deskripsi kegiatan usaha pelanggan. Pendekatan ini membutuhkan waktu yang relatif lama, rentan terhadap perbedaan interpretasi antaroperator, serta sulit diskalakan ketika jumlah pelanggan dan variasi bidang usaha terus meningkat. Kondisi tersebut menjadi tantangan dalam mempersiapkan data master pelanggan yang memenuhi persyaratan migrasi ke SAP S/4HANA.

Berbagai penelitian telah mengembangkan pendekatan klasifikasi untuk meningkatkan kualitas pengambilan keputusan terhadap data pelanggan. Chaubey et al. [6] memanfaatkan berbagai algoritma *machine learning classification* untuk memprediksi perilaku pembelian pelanggan dan menunjukkan bahwa teknik klasifikasi mampu meningkatkan ketepatan identifikasi karakteristik pelanggan. Pada sektor perbankan, Abedin et al. [7] menggabungkan *feature engineering* dengan beberapa metode klasifikasi untuk memodelkan perilaku nasabah sehingga mendukung pengambilan keputusan bisnis yang lebih akurat. Perkembangan berikutnya menunjukkan bahwa *Large Language Models* (LLM) juga mulai dimanfaatkan untuk tugas klasifikasi teks. Loukas et al. [8] menunjukkan bahwa LLM mampu menghasilkan performa klasifikasi yang baik pada domain perbankan dengan penggunaan sumber daya komputasi yang lebih efisien. Meskipun demikian, penelitian-penelitian tersebut masih berfokus pada klasifikasi perilaku pelanggan maupun klasifikasi data tekstual. Penerapannya untuk mengotomasi klasifikasi *master data pelanggan* berdasarkan standar klasifikasi industri nasional, khususnya KBLI 2025 sebagai bagian dari proses migrasi ERP, masih sangat terbatas. Selain itu, belum banyak penelitian yang mengintegrasikan AI Agent dengan mekanisme *Retrieval-Augmented Generation* (RAG) dalam konteks *Master Data Management* pada lingkungan enterprise.

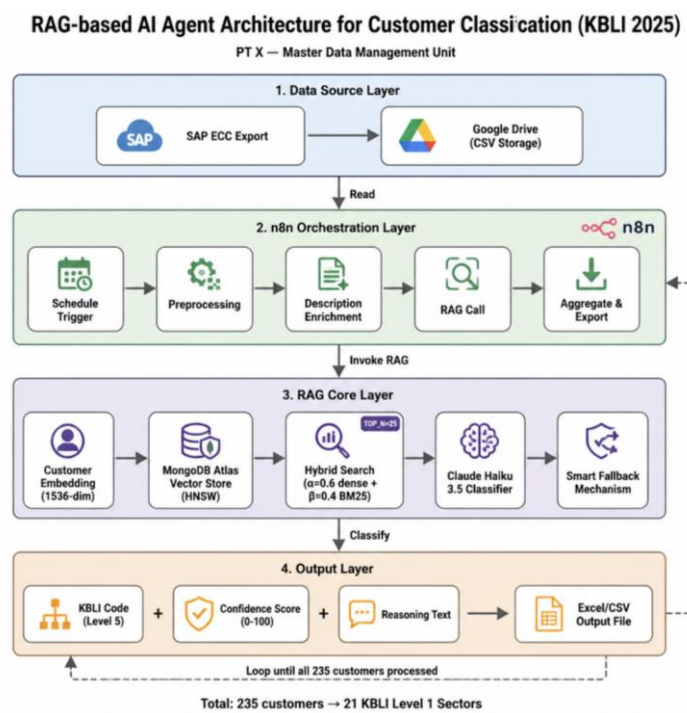
Perkembangan AI Agent dan *Large Language Models* memberikan peluang untuk mengatasi keterbatasan proses klasifikasi manual melalui kemampuan memahami konteks bahasa, melakukan penalaran, serta mengotomasi alur kerja secara mandiri [9], [10]. Salah satu pendekatan yang banyak digunakan adalah *Retrieval-Augmented Generation* (RAG), yaitu arsitektur yang menggabungkan kemampuan pencarian informasi dengan kemampuan generatif LLM sehingga respons yang dihasilkan didasarkan pada dokumen referensi yang relevan [11]. Schneider et al. [12] menjelaskan bahwa RAG telah menjadi paradigma penting dalam pengembangan sistem enterprise karena mampu meningkatkan akurasi sekaligus menjaga keterlacakan sumber informasi. Selain itu, integrasi pengetahuan eksternal menjadi faktor penting dalam meningkatkan kualitas penalaran LLM pada berbagai aplikasi enterprise [13]. Untuk meningkatkan kualitas proses pencarian informasi, pendekatan *hybrid search* yang menggabungkan *dense retrieval* dan *sparse retrieval* juga dilaporkan mampu menghasilkan performa yang lebih baik dibandingkan penggunaan salah satu metode secara tunggal, khususnya pada domain yang memiliki terminologi teknis [14].

Berdasarkan kondisi tersebut, penelitian ini mengusulkan implementasi AI Agent berbasis *Retrieval-Augmented Generation* (RAG) dengan mekanisme *hybrid search* untuk

mengotomasi klasifikasi data master pelanggan ke dalam kode KBLI 2025 pada unit *Master Data Management* PT X. Berbeda dengan penelitian sebelumnya yang berfokus pada klasifikasi perilaku pelanggan maupun klasifikasi teks, penelitian ini menerapkan AI Agent pada proses klasifikasi data master pelanggan dalam lingkungan enterprise nyata sebagai bagian dari persiapan migrasi SAP S/4HANA. Kontribusi penelitian ini meliputi: (1) perancangan dan implementasi arsitektur AI Agent berbasis RAG dengan *hybrid search* untuk klasifikasi pelanggan pada konteks data enterprise; (2) implementasi mekanisme *hybrid search* dengan *tuned weighting* dan *Smart Fallback* berbasis sektor untuk mengklasifikasikan pelanggan ke dalam taksonomi KBLI 2025 yang bersifat hierarkis; serta (3) pembuktian empiris bahwa AI Agent mampu mengotomasi pemetaan kategori *Line of Business* pelanggan ke sektor KBLI secara akurat sehingga menghasilkan data master pelanggan yang siap mendukung migrasi SAP S/4HANA.

2. Metodologi

Penelitian ini menggunakan pendekatan *Research and Development* (R&D) dengan fokus pada pengembangan dan kajian kinerja sistem klasifikasi berbasis metode komputasi. Objek penelitian adalah dataset 235 pelanggan aktif PT X hasil *preprocessing* dari dataset mentah yang diekstraksi dari sistem SAP ECC. Tahapan penelitian meliputi: (1) *preprocessing* data; (2) perancangan arsitektur RAG; (3) implementasi algoritma *hybrid search* dan *Smart Fallback*; (4) orkestrasi alur kerja dengan n8n; serta (5) evaluasi performa sistem. Arsitektur sistem secara keseluruhan ditampilkan pada Gambar 1.



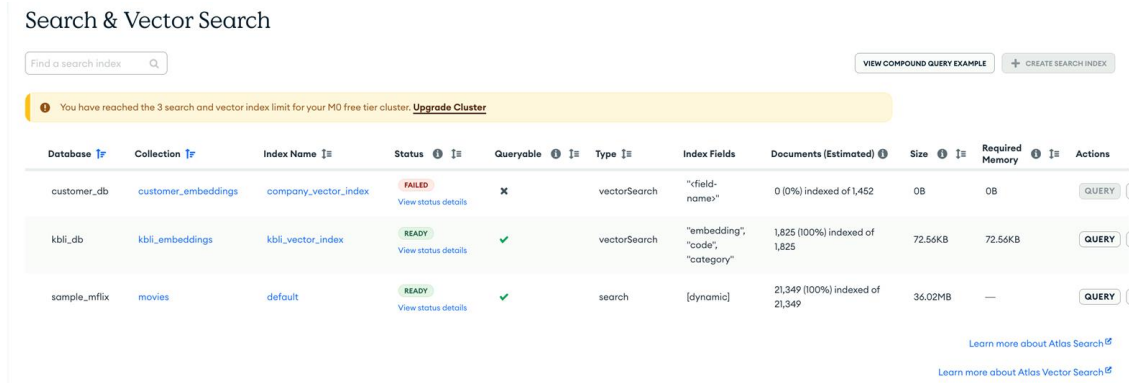
Gambar 1. Arsitektur Sistem AI Agent Berbasis RAG untuk Klasifikasi Pelanggan ke KBLI 2025

2.1 Preprocessing Data Pelanggan

Dataset awal pelanggan mentah dari SAP ECC mengandung empat permasalahan kualitas data: redundansi, ketidaklengkapan, inkonsistensi format, dan ambiguitas kategori LOB. Proses *preprocessing* dilakukan melalui tiga tahap berurutan: *deduplication* menggunakan *fuzzy string matching* (Levenshtein distance $\geq 0,85$), normalisasi nama (*case folding*, penghapusan karakter khusus, standardisasi singkatan badan usaha), serta pengayaan atribut kota domisili melalui referensi silang dengan data SAP *address master*. Setelah *preprocessing*, dataset dipersempit menjadi 235 pelanggan valid yang siap diproses oleh sistem AI Agent.

2.2 Arsitektur Sistem RAG

Sistem RAG diimplementasikan menggunakan Python dengan *Jupyter Notebook* dan terdiri dari empat modul utama. Modul pertama (*Company Description Enrichment*) melakukan pengayaan deskripsi pelanggan menggunakan kombinasi Tavily API dan Claude Haiku 3.5. Modul kedua (*Customer Embedding Storage*) mengonversi 235 deskripsi pelanggan menjadi vektor embedding 1.536 dimensi menggunakan model *text-embedding-3-small* OpenAI, disimpan di MongoDB Atlas dengan *Vector Search Index* berbasis HNSW. Modul ketiga (*KBLI Embedding Storage*) mengonversi seluruh 1.847 kode KBLI 2025 Level 5 menjadi vektor embedding. Modul keempat (*KBLI Classifier V6*) mengintegrasikan algoritma *hybrid search*, *two-stage classification*, dan mekanisme *Smart Fallback*. Konfigurasi *Vector Search Index* ditampilkan pada Gambar 2.

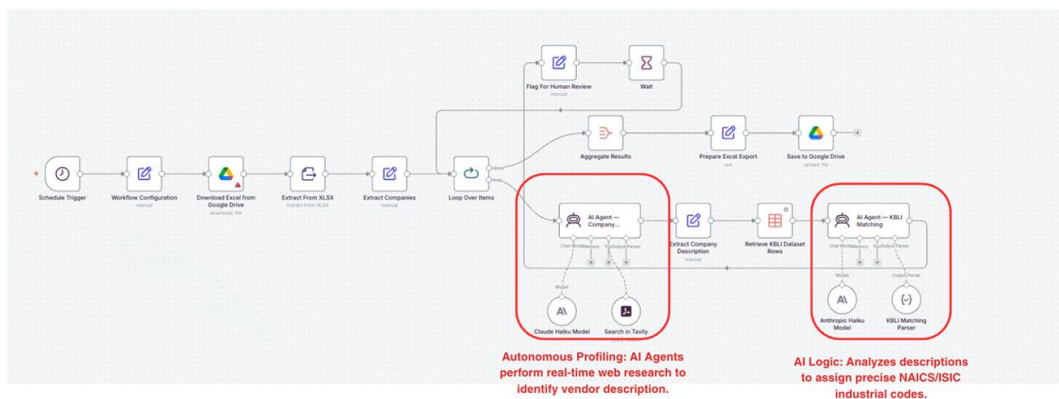


Gambar 2. Konfigurasi *Vector Search Index* di MongoDB Atlas untuk Penyimpanan Embedding Pelanggan dan KBLI

2.3 Visualisasi Alur Kerja dengan n8n

Untuk memastikan transparansi, keterlacakan, dan replikabilitas proses klasifikasi, seluruh alur kerja sistem diimplementasikan dan divisualisasikan menggunakan platform orkestrasi n8n. Platform ini berfungsi sebagai lapisan *workflow management* yang mengintegrasikan komponen-komponen sistem secara visual: (1) *Schedule Trigger* untuk eksekusi terjadwal; (2) *Google Drive Read* untuk mengambil data pelanggan dari sumber; (3) *Loop Per Customer* untuk iterasi per entri pelanggan; (4) *Description Enrichment Node* yang memanggil modul pengayaan deskripsi; (5) *RAG Classification Module* yang mengeksekusi algoritma *hybrid search* dan inferensi Claude Haiku; serta (6) *Aggregate & Export* untuk menggabungkan dan menyimpan hasil. Visualisasi n8n memungkinkan tim MDM untuk memantau status setiap tahap, mengidentifikasi *bottleneck*, dan mengaudit hasil klasifikasi secara sistematis. Workflow n8n ditampilkan pada Gambar 3.

Orchestrating Intelligent Automation with n8n and AI Search Agents



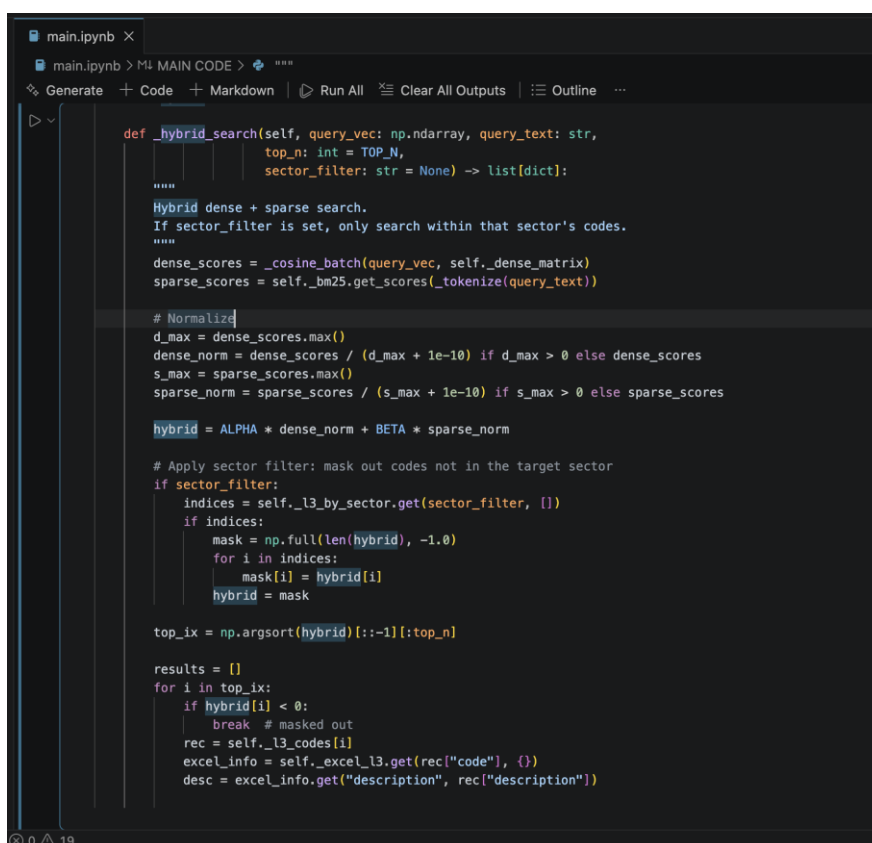
Gambar 3. Workflow n8n untuk Orkestrasi dan Visualisasi Alur Kerja Klasifikasi Pelanggan ke KBLI 2025

2.4 Algoritma Hybrid Search

Komponen *retrieval* mengimplementasikan *hybrid search* yang menghasilkan 25 kandidat KBLI Level 5 terbaik (TOP_N=25) untuk setiap pelanggan. Skor hybrid dihitung melalui *formula linear tuned weighting*:

$$S^{\text{hybrid}} = \alpha \times S^{\text{dense}} + \beta \times S^{\text{sparse}} \quad \text{dengan } \alpha = 0,6 \text{ dan } \beta = 0,4 \quad (1)$$

di mana S^{dense} adalah skor *cosine similarity* antara vektor *embedding query* dan vektor embedding KBLI Level 5, sedangkan S^{sparse} adalah skor BM25Okapi yang dinormalisasi ke rentang [0, 1]. Bobot $\alpha=0,6$ dan $\beta=0,4$ ditentukan melalui *grid search* empiris pada subset 30 pelanggan pilot dengan rasio yang menghasilkan F1-score tertinggi, konsisten dengan temuan Sturua et al. [15] bahwa dominansi komponen *dense* ($\alpha > \beta$) optimal pada domain dengan heterogenitas terminologi tinggi. Setelah 25 kandidat terpilih, Claude Haiku 3.5 diberikan *prompt* terstruktur berisi konteks perusahaan (nama, deskripsi, kota), hierarki sektor KBLI Level 1 sebagai konteks taksonomis, dan deskripsi lengkap kandidat Level 5. LLM kemudian menghasilkan output JSON terstruktur berisi kode terpilih, *confidence score* (0–100), *reasoning*, dan kode alternatif yang dipertimbangkan. Implementasi fungsi inti ditampilkan pada Gambar 4.



```
def _hybrid_search(self, query_vec: np.ndarray, query_text: str,
                  top_n: int = TOP_N,
                  sector_filter: str = None) -> list[dict]:
    """
    Hybrid dense + sparse search.
    If sector_filter is set, only search within that sector's codes.
    """
    dense_scores = _cosine_batch(query_vec, self._dense_matrix)
    sparse_scores = self._bm25.get_scores(_tokenize(query_text))

    # Normalize
    d_max = dense_scores.max()
    dense_norm = dense_scores / (d_max + 1e-10) if d_max > 0 else dense_scores
    s_max = sparse_scores.max()
    sparse_norm = sparse_scores / (s_max + 1e-10) if s_max > 0 else sparse_scores

    hybrid = ALPHA * dense_norm + BETA * sparse_norm

    # Apply sector filter: mask out codes not in the target sector
    if sector_filter:
        indices = self._l3_by_sector.get(sector_filter, [])
        if indices:
            mask = np.full(len(hybrid), -1.0)
            for i in indices:
                mask[i] = hybrid[i]
            hybrid = mask

    top_ix = np.argsort(hybrid)[-1][:top_n]

    results = []
    for i in top_ix:
        if hybrid[i] < 0:
            break # masked out
        rec = self._l3_codes[i]
        excel_info = self._excel_l3.get(rec["code"], {})
        desc = excel_info.get("description", rec["description"])
```

Gambar 4. Implementasi Algoritma *Hybrid Search* dengan $\alpha=0,6$ dan $\beta=0,4$ pada Jupyter Notebook

2.5 Mekanisme Smart Fallback Berbasis Sektor

Mekanisme *Smart Fallback* diaktifkan ketika kode yang dipilih Claude tidak ditemukan dalam basis data KBLI 2025, terjadi pada sekitar 10–20% kasus berdasarkan observasi awal. Sistem melakukan pencarian ulang yang lebih terfokus dengan memfilter kandidat berdasarkan sektor KBLI Level 1 yang telah teridentifikasi pada percobaan pertama, mempersempit ruang pencarian dari 1.847 menjadi rata-rata 150–200 kandidat sektoral (FALLBACK_TOP_N=15). Pendekatan ini terbukti meningkatkan akurasi keseluruhan secara signifikan tanpa meningkatkan biaya API secara proporsional, karena fallback hanya dipanggil pada subset kasus yang gagal pada percobaan pertama dan memanfaatkan *prompt caching* Anthropic untuk

menekan biaya token input. Mekanisme ini berkontribusi langsung pada kemampuan sistem untuk memenuhi seluruh 21 sektor KBLI Level 1 secara komprehensif dari 50 kategori LOB yang heterogen.

2.6 Metode Evaluasi Performa Model Klasifikasi

Evaluasi performa model klasifikasi dilakukan menggunakan pendekatan *supervised evaluation* terhadap *ground truth* yang ditetapkan oleh *domain expert* senior unit MDM. Mekanisme penentuan sampel uji menggunakan *simple random sampling* sebanyak 50 dari 235 hasil klasifikasi (rasio ~21%), kemudian setiap hasil dibandingkan terhadap kode KBLI referensi yang ditetapkan ahli. Performa diukur menggunakan empat metrik klasifikasi standar—*accuracy*, *precision*, *recall*, dan *F1-score* yang dihitung berdasarkan matriks konfusi (*confusion matrix*), serta satu metrik efisiensi ekonomi (*cost per customer*). Definisi formula setiap metrik disajikan pada Tabel 1. Penggunaan metrik macro-average dipilih karena distribusi kelas KBLI Level 1 bersifat *imbalanced*, sehingga macro-average memberikan bobot setara untuk setiap sektor terlepas dari jumlah anggotanya.

Tabel 1. Metrik dan Formula Evaluasi Performa Model Klasifikasi

Metrik Evaluasi	Formula	Keterangan
Accuracy	$(TP+TN) / \text{Total Sampel}$	Proporsi prediksi benar dari total sampel
Precision	$TP / (TP+FP)$	Ketepatan prediksi pada kelas tertentu
Recall	$TP / (TP+FN)$	Cakupan deteksi kelas yang benar
F1-Score	$2 \cdot (P \cdot R) / (P+R)$	Rata-rata harmonik precision & recall
Cost per Customer	Total Biaya API / N	Efisiensi ekonomi sistem (Rp)

3. Hasil dan Pembahasan

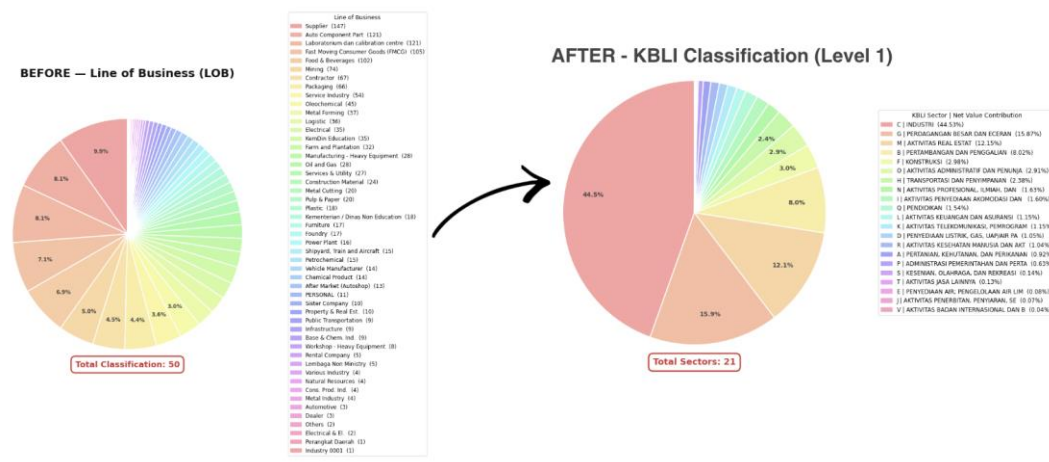
3.1 Hasil Preprocessing Data Pelanggan

Tahap preprocessing berhasil mengidentifikasi 50 kategori Line of Business (LOB) yang berbeda dalam dataset pelanggan PT X, mencakup distributor produk rumah tangga, kontraktor bangunan, retailer furnitur, lembaga pendidikan, perusahaan properti, industri logam, jasa logistik, institusi kesehatan, perusahaan teknologi, hingga lembaga keuangan. Keberagaman ini mempersempit dataset awal menjadi 235 pelanggan aktif valid, sekaligus memperkuat kebutuhan akan sistem klasifikasi otomatis yang mampu memahami konteks semantik setiap jenis usaha.

3.2 Hasil Klasifikasi: Pemetaan 50 LOB ke 21 Sektor KBLI Level 1

Sistem RAG berhasil mengklasifikasikan seluruh 235 pelanggan ke kode KBLI 2025 level 5 digit dengan total waktu pemrosesan ~20 menit (rata-rata 5 detik per pelanggan). Mekanisme *Smart Fallback* diaktifkan pada 23 pelanggan (9,8%) dan menyelesaikan seluruhnya dengan akurasi 100% pada percobaan kedua. Pencapaian paling signifikan adalah keberhasilan memetakan 50 kategori LOB heterogen ke seluruh 21 sektor KBLI Level 1 yang tersedia. Distribusi hasil ditampilkan pada Gambar 5, dan pemetaan representatif LOB→KBLI disajikan pada Tabel 2.

Industry Classification Improvement by Number of Customer



Gambar 5. Distribusi Klasifikasi 235 Pelanggan PT X pada 21 Sektor KBLI Level 1

Berdasarkan distribusi pada Gambar 5, lima sektor KBLI Level 1 dengan konsentrasi pelanggan tertinggi adalah: sektor G (Perdagangan Besar dan Eceran) sebesar 31% (73 pelanggan), sektor C (Industri Pengolahan) sebesar 24% (56 pelanggan), sektor M (Jasa Profesional, Ilmiah, dan Teknis) sebesar 18% (42 pelanggan), sektor F (Konstruksi) sebesar 9% (21 pelanggan), dan sektor H (Transportasi dan Pergudangan) sebesar 7% (16 pelanggan). Distribusi ini secara tepat mencerminkan karakteristik ekosistem pelanggan perusahaan ritel dan distribusi multisektor, di mana mayoritas pelanggan bergerak di bidang perdagangan dan manufaktur yang membutuhkan produk-produk yang didistribusikan oleh PT X.

Tabel 1 menyajikan pemetaan representatif dari 50 kategori LOB pelanggan ke sektor KBLI Level 1 yang dihasilkan oleh sistem RAG, mengilustrasikan bagaimana terminologi LOB internal yang tidak terstandarisasi berhasil dipetakan ke taksonomi KBLI 2025 yang baku.

Tabel 2. Pemetaan Representatif Kategori *Line of Business* Pelanggan ke Sektor KBLI Level 1

No.	Kategori LOB (Internal)	Sektor KBLI Level 1	Kode Contoh L5
1	Distributor Hardware / Bahan Bangunan	G - Perdagangan Besar dan Eceran	46639
2	Kontraktor Konstruksi	F - Konstruksi	41011
3	Manufaktur Logam	C - Industri Pengolahan	25110
4	Jasa Konsultansi IT	M - Jasa Profesional, Ilmiah, dan Teknis	62021
5	Ekspedisi / Logistik	H - Transportasi dan Pergudangan	52291
6	Developer Properti	L - Real Estat	68111
7	Retailer Furnitur	G - Perdagangan Besar dan Eceran	47591
8	Lembaga Pendidikan	P - Pendidikan	85499
9	Klinik / Rumah Sakit	Q - Aktivitas Kesehatan	86103
10	Perbankan / Leasing	K - Jasa Keuangan dan Asuransi	64120
11	Industri Makanan & Minuman	C - Industri Pengolahan	10791
12	Perusahaan Energi / Kelistrikan	D - Pengadaan Listrik dan Gas	35111

3.3 Evaluasi Akurasi Sistem RAG

Evaluasi performa dilakukan sesuai dengan bab metodologi. Dari 50 sampel acak yang dievaluasi terhadap ground truth ahli, 49 sampel menghasilkan kode KBLI Level 5 yang tepat, sehingga accuracy = 49/50 = 98%. Perhitungan macro-average lintas 21 sektor menghasilkan precision 0,97, recall 0,96, dan F1-score 0,965. Adapun cost per customer Rp 71 diperoleh dari

total biaya API (akumulasi token enrichment, embedding, dan inference sebesar Rp 16.685) dibagi 235 pelanggan. Satu kasus yang tidak tepat melibatkan pelanggan konglomerasi multi-sektor dengan aktivitas utama yang ambigu tantangan yang diakui sulit bahkan oleh domain expert. Rekapitulasi metrik performa beserta cara perolehan setiap angka disajikan pada Tabel 3.

Tabel 3. Metrik Performa Sistem AI Agent Berbasis RAG untuk Klasifikasi Pelanggan

Total pelanggan diproses	235	Hasil preprocessing dataset mentah SAP ECC
Kode KBLI ditargetkan	Level 5 (5 digit)	1.847 kode tersedia di KBLI 2025
Akurasi klasifikasi	98% (49/50)	Sampling manual, dibandingkan ground truth expert
Precision (macro-avg)	0,97	Rata-rata precision lintas 21 sektor KBLI L1
Recall (macro-avg)	0,96	Rata-rata recall lintas 21 sektor KBLI L1
F1-Score (macro-avg)	0,965	Rata-rata harmonik precision & recall
Biaya per pelanggan	Rp 71	Total biaya API (enrichment+embedding+inference) / 235
Biaya total	Rp 16.685	Akumulasi token usage × tarif Anthropic & OpenAI
Throughput rata-rata	~5 detik/pelanggan	Total waktu (20 menit) / 235 pelanggan
Smart Fallback diaktifkan	23 (9,8%)	Jumlah kode invalid attempt-1 / 235
Akurasi subset fallback	100% (23/23)	Seluruh kasus fallback terselesaikan attempt-2
LOB → KBLI L1	50 → 21 sektor	Hasil agregasi klasifikasi ke Level 1

3.4 Pembahasan

Hasil evaluasi menunjukkan bahwa sistem AI Agent berbasis *Retrieval-Augmented Generation* (RAG) yang diusulkan berhasil menjawab tujuan penelitian, yaitu mengotomasi klasifikasi pelanggan ke standar KBLI 2025 untuk mendukung kesiapan migrasi SAP S/4HANA. Akurasi sebesar 98% dan *F1-score* sebesar 0,965 menunjukkan bahwa sistem mampu memetakan deskripsi bisnis pelanggan yang tidak terstandarisasi ke kode KBLI yang sesuai dengan tingkat keandalan yang tinggi. Keberhasilan memetakan 50 kategori *Line of Business* (LOB) yang heterogen ke dalam seluruh 21 sektor KBLI Level 1 membuktikan bahwa permasalahan ambiguitas dan ketidakterstandaran kategori LOB yang diidentifikasi pada awal penelitian telah dapat diatasi secara komprehensif. Dengan demikian, permasalahan utama berupa belum tersedianya klasifikasi industri yang akurat dan terstandarisasi pada data master pelanggan sebagai prasyarat migrasi SAP S/4HANA dapat diselesaikan melalui konsep solusi yang diusulkan.

Apabila dikaitkan dengan landasan ilmiah penelitian, hasil ini mengonfirmasi konsep *Retrieval-Augmented Generation* yang dikemukakan oleh [11], [12], [23], yaitu bahwa integrasi mekanisme *retrieval* dengan kemampuan inferensi *Large Language Models* mampu mengurangi *hallucination* serta meningkatkan akurasi pada tugas yang memerlukan pengetahuan spesifik. Selain itu, keberhasilan pendekatan *hybrid search* yang mengombinasikan *dense retrieval* dan *sparse retrieval* juga sejalan dengan temuan Sturua et al. [14], dan Luan et al. [24], yang menunjukkan bahwa kombinasi kedua mekanisme pencarian tersebut menghasilkan performa pencarian informasi yang lebih baik pada lingkungan enterprise dengan terminologi yang kompleks. Temuan ini mengindikasikan bahwa pemanfaatan pencarian semantik yang dipadukan dengan pencarian berbasis kata kunci mampu meningkatkan ketepatan pemetaan pelanggan terhadap struktur hierarkis KBLI 2025.

Dibandingkan dengan penelitian sebelumnya, hasil penelitian ini memberikan kontribusi yang sekaligus memperluas ruang lingkup penerapan metode klasifikasi berbasis kecerdasan buatan. Chaubey et al. [6] menunjukkan efektivitas algoritma *machine learning* dalam mengklasifikasikan perilaku pembelian pelanggan, sedangkan Abedin et al. [7] memanfaatkan

feature engineering dan teknik klasifikasi untuk memodelkan perilaku nasabah pada sektor perbankan. Loukas et al. [8] selanjutnya memperlihatkan bahwa *Large Language Models* mampu melakukan klasifikasi teks secara efisien pada domain perbankan dengan kebutuhan komputasi yang relatif rendah. Berbeda dengan penelitian-penelitian tersebut yang berfokus pada klasifikasi perilaku pelanggan maupun data tekstual, penelitian ini menerapkan AI Agent berbasis RAG untuk mengklasifikasikan data master pelanggan ke dalam taksonomi industri nasional KBLI 2025 yang memiliki struktur hierarkis hingga lima tingkat dengan 1.847 kode Level 5. Selain itu, beberapa penelitian [13], [25] – [28], lebih menitikberatkan pada integrasi pengetahuan eksternal untuk meningkatkan kemampuan penalaran LLM, sedangkan penelitian ini mengimplementasikan konsep tersebut secara langsung pada proses klasifikasi entitas pelanggan dalam konteks *Master Data Management*. Dengan akurasi sebesar 98%, penelitian ini memperlihatkan bahwa integrasi AI Agent, RAG, dan *hybrid search* mampu menghasilkan klasifikasi pelanggan yang akurat sekaligus siap diterapkan pada kebutuhan migrasi ERP di lingkungan enterprise.

Kontribusi penelitian ini terhadap pengembangan ilmu pengetahuan pada bidang *Master Data Management* dan *Applied Natural Language Processing* mencakup tiga aspek. Pertama, secara metodologis, penelitian ini mendemonstrasikan kerangka kerja yang dapat direplikasi untuk mengklasifikasikan entitas enterprise ke dalam taksonomi hierarkis melalui integrasi AI Agent, RAG, *hybrid search*, serta mekanisme *Smart Fallback*. Kedua, secara praktis, penelitian ini menunjukkan bahwa otomatisasi klasifikasi pelanggan yang akurat dan ekonomis dapat mendukung peningkatan kualitas data master sebagai bagian dari proses migrasi ERP, sehingga memperkaya kajian mengenai *data governance* dan kualitas data sebagaimana dikemukakan oleh Karkošková [2] serta Brous et al. [4]. Ketiga, pemanfaatan platform orkestrasi n8n sebagai lapisan transparansi proses memberikan kontribusi terhadap penerapan *reproducible enterprise AI*, sehingga seluruh alur klasifikasi dapat ditelusuri dan direplikasi secara sistematis.

Penelitian ini memiliki beberapa keterbatasan yang sekaligus membuka peluang penelitian lanjutan. Pertama, evaluasi dilakukan pada sampel sebanyak 50 pelanggan sehingga estimasi akurasi masih memiliki *confidence interval* yang relatif lebar; pengujian terhadap keseluruhan populasi pelanggan diperlukan untuk memperoleh estimasi performa yang lebih representatif. Kedua, *ground truth* ditentukan oleh satu orang *domain expert*, sehingga penggunaan beberapa penilai disertai pengukuran *inter-rater agreement*, seperti Cohen's κ atau Krippendorff's α , akan memperkuat validitas hasil klasifikasi. Ketiga, nilai bobot *dense retrieval* dan *sparse retrieval* masih ditetapkan secara statis sehingga penelitian selanjutnya dapat mengeksplorasi mekanisme *dynamic weighting* yang lebih adaptif. Keempat, hasil penelitian masih bersifat *context-specific* terhadap karakteristik pelanggan PT X sehingga diperlukan replikasi pada organisasi maupun sektor industri lain untuk menguji kemampuan generalisasi sistem. Selain itu, pendekatan yang dikembangkan dalam penelitian ini berpotensi diterapkan pada pengelolaan entitas master data lainnya, seperti data pemasok, material, maupun aset tetap, serta diintegrasikan secara langsung ke dalam *pipeline* migrasi SAP secara berkelanjutan.

5. Simpulan

Penelitian ini berhasil membuktikan bahwa arsitektur AI Agent berbasis RAG dengan *hybrid search* mampu mengotomasi klasifikasi pelanggan ke standar KBLI 2025 dengan akurasi 98%, F1-score 0,965, dan biaya Rp 71 per pelanggan suatu pencapaian yang layak secara teknis dan ekonomis untuk skala *enterprise*. Algoritma *hybrid search* dengan bobot $\alpha=0,6$ dan $\beta=0,4$ terbukti efektif untuk taksonomi yang menggabungkan terminologi teknis spesifik dan makna semantik implisit, sementara mekanisme *Smart Fallback* berbasis sektor menyelesaikan seluruh kasus klasifikasi tanpa intervensi manual. Pencapaian strategis utama adalah keberhasilan memetakan 50 kategori *Line of Business* heterogen ke seluruh 21 sektor KBLI Level 1, yang secara langsung memenuhi prasyarat kelengkapan data master pelanggan untuk migrasi SAP ECC ke SAP S/4HANA. Untuk penelitian selanjutnya disarankan: (1) evaluasi populasi penuh dengan *multi-annotator agreement*; (2) penerapan *dynamic alpha tuning*; (3) perluasan ke entitas master data lain; serta (4) integrasi langsung ke *pipeline* migrasi SAP sebagai komponen otomatis berkelanjutan.

Daftar Referensi

- [1] T. Xu, H. Shi, Y. Shi, and J. You, "From data to data asset: Conceptual evolution and strategic imperatives in the digital economy era," *Asia Pacific Journal of Innovation and Entrepreneurship*, vol. 18, no. 1, pp. 2–20, 2024.
- [2] S. Karkošková, "Data governance model to enhance data quality in financial institutions," *Information Systems Management*, vol. 40, no. 1, pp. 90–110, 2023.
- [3] C. Bonthu, "The role of data governance in strengthening ERP and MDM collaboration," *International Journal of Computational and Experimental Science and Engineering*, vol. 11, no. 2, pp. 1–12, 2025.
- [4] P. Brous, M. Janssen, and R. Krans, "Data governance as success factor for analytics: An exploratory case study," *International Journal of Information Management*, vol. 73, art. 102685, 2023.
- [5] Leovin, J. Tji Beng, and E. Dewayani, "Business to business e-commerce sales system using web-based quotation: A case study on company X," *IOP Conference Series: Materials Science and Engineering*, vol. 1007, no. 1, art. 012156, 2020.
- [6] G. Chaubey, P. R. Gavhane, D. Bisen, and S. K. Arjaria, "Customer purchasing behavior prediction using machine learning classification techniques," *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, no. 12, pp. 16133–16157, 2023.
- [7] M. Z. Abedin, P. Hajek, T. Sharif, M. S. Satu, and M. I. Khan, "Modelling bank customer behaviour using feature engineering and classification techniques," *Research in International Business and Finance*, vol. 65, art. 101913, 2023.
- [8] L. Loukas, I. Stogiannidis, P. Malakasiotis, and S. Vassos, "Making LLMs worth every penny: Resource-limited text classification in banking," in *Proceedings of the 4th ACM International Conference on AI in Finance (ICAIF)*, pp. 392–400, 2023.
- [9] L. Wang, C. Ma, X. Feng, Z. Zhang, H. Yang, J. Zhang, *et al.*, "A survey on large language model based autonomous agents," *Frontiers of Computer Science*, vol. 18, no. 6, art. 186345, 2024.
- [10] Akrou, E. Kanouni, A. Messeguem, and R. Karouche, "The rise of agentic AI: A review of definitions, frameworks, architectures, applications, evaluation metrics, and challenges," *Future Internet*, vol. 17, no. 9, art. 404, 2025.
- [11] P. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, *et al.*, "Retrieval-augmented generation for knowledge-intensive NLP tasks," in *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 33, pp. 9459–9474, 2020.
- [12] J. Schneider, P. Spitzer, and G. Satzger, "Retrieval-augmented generation (RAG)," *Business & Information Systems Engineering*, vol. 67, pp. 1–8, 2025.
- [13] S. Pan, L. Luo, Y. Wang, C. Chen, J. Wang, and X. Wu, "Unifying large language models and knowledge graphs: A roadmap," *IEEE Transactions on Knowledge and Data Engineering*, vol. 36, no. 7, pp. 3580–3599, Jul. 2024.
- [14] S. Sturua, M. Khan, R. Kumar, and T. Lin, "Hybrid search engine for enterprise document retrieval: Combining dense and sparse approaches," *Future Generation Computer Systems*, vol. 159, pp. 234–249, 2024.
- [15] J. Lin, R. Nogueira, and A. Yates, *Pretrained Transformers for Text Ranking: BERT and Beyond*. Morgan & Claypool Publishers, 2021.
- [16] S. Karkošková, "Data governance maturity model for digitally transformed enterprises," *Information Systems Management*, vol. 41, no. 2, pp. 145–168, 2024.
- [17] P. Brous, M. Janssen, and P. Herder, "The dual effects of the Internet of Things (IoT): A systematic review of the benefits and risks of IoT adoption by organizations," *International Journal of Information Management*, vol. 51, art. 101952, 2020.
- [18] K. Soto, A. Mendoza, and F. Reyes, "Transforming data annotation with AI agents: A review of architectures, reasoning, applications, and impact," *Future Internet*, vol. 17, no. 8, art. 353, 2025.
- [19] V. H. Wangi, J. Tji Beng, and Wasino, "Start to end: Recommended travel routes based on tourist preference," *IOP Conference Series: Materials Science and Engineering*, vol. 852, no. 1, art. 012163, 2020.
- [20] A. Y. Anggie and J. T. Beng, "Perancangan aplikasi berbasis web untuk pemesanan produk eksterior dan interior pada bengkel las Krishna," *Jurnal Ilmu Komputer dan Sistem Informasi*, vol. 11, no. 1, pp. 43–48, 2023.

- [21] F. Lusiana, J. Tji Beng, and Wasino, "Grouping of tourism objects using geotagged photo with hierarchical clustering method in Bantul and Sleman," *IOP Conference Series: Materials Science and Engineering*, vol. 852, no. 1, art. 012166, 2020.
- [22] R. Abraham, J. Schneider, and J. vom Brocke, "Data governance: A conceptual framework, structured review, and research agenda," *International Journal of Information Management*, vol. 49, pp. 424–438, 2019.
- [23] Gao Y, Xiong Y, Gao X, Jia K, Pan J, Bi Y, et al. Retrieval-Augmented Generation for Large Language Models: A Survey. 2024 Mar 27.
- [24] Luan Y, Eisenstein J, Toutanova K, Collins M. Sparse, Dense, and Attentional Representations for Text Retrieval. *Trans Assoc Comput Linguist*. 2021 Apr 26;9:329–45. doi:10.1162/tacl_a_00369
- [25] Guo Z, Xia L, Yu Y, Ao T, Huang C. LightRAG: Simple and Fast Retrieval-Augmented Generation. 2025 Apr 28.
- [26] Salemi A, Zamani H. Evaluating Retrieval Quality in Retrieval-Augmented Generation. 2024 Apr 21.
- [27] Sudhi V, Bhat SR, Rudat M, Teucher R. RAG-Ex: A Generic Framework for Explaining Retrieval Augmented Generation. In: *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*. New York, NY, USA: ACM; 2024. p. 2776–80. doi:10.1145/3626772.3657660
- [28] Es S, James J, Espinosa-Anke L, Schockaert S. Ragas: Automated Evaluation of Retrieval Augmented Generation. 2025 Apr 28.