



AI Parafrase Bahasa Indonesia Informal ke Formal dengan Sentence-BERT dan Gemini API

Maulana Farid Fatin, Muhamad Rian Hermawan, Anisa Nur Ramadhani, Hanif Sinaga Zaidan

Program Studi Bisnis Digital, Fakultas Ekonomi dan Bisnis, Universitas Ibn Khaldun Bogor

maulanafaridfatin@gmail.com, mrianh20@gmail.com, anisa.ninis1787@gmail.com, hanifzaidan@gmail.com

Abstrak

Penggunaan bahasa informal dalam komunikasi digital sering menimbulkan kesulitan ketika pengguna perlu berkomunikasi secara formal, khususnya pada konteks akademik dan profesional. Penelitian ini mengembangkan sistem AI untuk melakukan parafrase bahasa Indonesia informal menjadi formal menggunakan pendekatan Retrieval Augmented Style Transfer (RAST) yang mengombinasikan Sentence-BERT Similarity dan Gemini API. Sentence-BERT Similarity digunakan untuk menemukan contoh kalimat dalam dataset yang memiliki kemiripan makna dengan input pengguna, sedangkan Gemini API digunakan untuk menghasilkan kalimat formal melalui proses style transfer. Dataset yang digunakan terdiri dari 1.110 pasangan kalimat informal dan formal yang mencakup empat konteks komunikasi, yaitu dosen, admin kampus, LinkedIn, dan email lamaran kerja. Pada proses retrieval, sistem menggunakan tiga contoh dengan nilai cosine similarity tertinggi sebagai referensi dalam penyusunan prompt. Evaluasi kualitas output dilakukan secara manual oleh tiga evaluator terhadap 12 keluaran sistem menggunakan empat aspek, yaitu kesesuaian makna, formalitas, kelancaran bahasa, dan kesesuaian konteks dengan skala Likert 1–5. Hasil evaluasi menunjukkan nilai rata-rata keseluruhan sebesar 4,54 dengan kategori sangat baik. Pengujian pembandingan juga menunjukkan bahwa penggunaan Sentence-BERT sebagai metode retrieval menghasilkan keluaran yang lebih terarah terhadap konteks dibandingkan penggunaan Gemini API tanpa retrieval. Hasil penelitian menunjukkan bahwa kombinasi Sentence-BERT Similarity dan Gemini API dapat membantu menghasilkan kalimat formal yang tetap mempertahankan makna utama serta menyesuaikan gaya bahasa dengan konteks komunikasi pengguna.

Kata kunci: Artificial Intelligence, Natural Language Processing, Sentence-BERT Similarity, Style Transfer, Gemini API

Abstract

The use of informal language in digital communication often poses challenges when users need to communicate formally, particularly in academic and professional contexts. This study develops an AI system to paraphrase informal Indonesian into formal Indonesian using a Retrieval Augmented Style Transfer (RAST) approach that combines Sentence-BERT Similarity and the Gemini API. Sentence-BERT Similarity is employed to identify example sentences within the dataset that are semantically similar to the user's input, while the Gemini API is used to generate formal sentences through a style transfer process. The dataset comprises 1,110 pairs of informal and formal sentences covering four communication contexts: interactions with lecturers, campus administration, LinkedIn, and job application emails. During the retrieval process, the system utilizes the three examples with the highest cosine similarity scores as references for prompt construction. Output quality was manually evaluated by three evaluators across 12 system outputs based on four criteria semantic accuracy, formality, linguistic fluency, and contextual appropriateness using a 1–5 Likert scale. The evaluation yielded an overall average score of 4.54, placing it in the "very good" category. Comparative testing demonstrated that using Sentence-BERT for retrieval produced outputs that were more contextually aligned than those generated by the Gemini API without retrieval. The findings indicate that combining Sentence-BERT Similarity and the Gemini API facilitates the generation of formal sentences that preserve the core meaning while adapting the language style to the user's communication context.

Keywords: Artificial Intelligence, Natural Language Processing, Sentence-BERT Similarity, Style Transfer, Gemini API

1. Pendahuluan

Evolusi teknologi digital telah banyak mengubah cara masyarakat berkomunikasi. Platform digital seperti media sosial, email, dan aplikasi pesan instan telah menggantikan komunikasi tatap muka sebagai sarana komunikasi utama. Perubahan ini memfasilitasi pertukaran informasi namun juga berdampak pada cara orang menggunakan bahasa. Banyak jenis bahasa informal yang muncul sebagai akibat dari penggunaan bahasa yang singkat, cepat, dan praktis, dan semakin lazim dalam kehidupan sehari-hari. Karena pengguna sering mengabaikan

struktur kalimat, ejaan, dan aturan linguistik yang relevan, bahasa informal berkembang cukup cepat. Bahasa informal berkembang sebagai alat komunikasi yang dianggap lebih berhasil dan menyampaikan keintiman antar penutur. Namun, tidak semua situasi komunikasi memerlukan penggunaan bahasa santai, khususnya dalam lingkungan akademis dan profesional yang memerlukan bahasa yang lebih resmi dan sopan (Khairul & Perdana, 2024).

Salah satu keterampilan utama yang perlu dimiliki mahasiswa dalam lingkungan akademis adalah kemampuan menggunakan bahasa formal. Ketika mahasiswa terlibat dengan dosen, anggota staf, atau institusi akademik lainnya, keterampilan ini diperlukan. Selain sopan santun, menggunakan bahasa formal menunjukkan profesionalisme dan kemampuan komunikasi seseorang. Menurut Febriyanti dan Pitasari (2023), pelanggaran etika komunikasi mahasiswa dalam menghubungi dosen melalui WhatsApp masih ditemukan, terutama pada aspek penggunaan bahasa, waktu menghubungi, dan cara pengiriman pesan. Permasalahan yang sering muncul antara lain penggunaan bahasa yang kasar, kegagalan mengidentifikasi diri dengan benar, penggunaan akronim yang terlalu banyak, dan penyampaian gagasan yang tidak sesuai dengan standar akademik. Temuan tersebut menunjukkan bahwa sebagian mahasiswa masih mengalami kesulitan dalam menerapkan bahasa formal pada komunikasi akademik.

Meluasnya penggunaan bahasa yang tidak baku juga dipengaruhi oleh pertumbuhan media sosial. Eleison et al. (2022) menguraikan bagaimana meningkatnya penggunaan bahasa gaul atau bahasa non-standar di media sosial dapat menurunkan pemahaman komunikasi karena tidak semua pengguna memahami frasa tersebut. Singkatan, kesalahan ejaan, akronim, dan terminologi slang adalah beberapa di antara berbagai jenis bahasa informal yang dapat menjadi tantangan bagi sistem dan pengguna dari berbagai latar belakang untuk memahaminya. Sejalan dengan hal tersebut, Ardinata et al. (2024) menemukan bahwa proses normalisasi kata slang pada teks berbahasa Indonesia di media sosial masih menghadapi tantangan. Karena tidak semua istilah informal dapat dipetakan secara akurat ke bentuk baku menggunakan pendekatan berbasis *embedding* kata.

Ada banyak jenis informalitas dalam bahasa Indonesia sehari-hari, seperti kesalahan dalam susunan kata, ejaan, dan kosa kata. Mayoritas model bahasa Indonesia dibuat menggunakan data yang diungkapkan secara formal, sehingga menimbulkan masalah bagi pengembangan teknologi pemrosesan bahasa alami. Akibatnya, ketika menampilkan tulisan informal yang sering muncul dalam interaksi digital, kinerja sistem seringkali menurun (Wibowo et al., 2020).

Pendekatan baru untuk memecahkan permasalahan ini dimungkinkan oleh perkembangan kecerdasan buatan (AI). Kecerdasan buatan (AI) merupakan teknologi yang dapat meniru proses kognitif manusia untuk melakukan berbagai aktivitas secara otomatis, termasuk pemrosesan bahasa alami (*NLP*). Menurut Budi et al. (2024) teknologi *NLP* memungkinkan sistem komputer untuk memahami, menganalisis, dan menghasilkan bahasa manusia dengan lebih berhasil dan efisien. *NLP* saat ini banyak digunakan di berbagai aplikasi, termasuk terjemahan mesin, chatbots, analisis sentimen, dan sistem pendukung penulisan.

Kecerdasan buatan (AI) semakin banyak digunakan di bidang pendidikan. Ini dapat membantu mahasiswa dalam meningkatkan kualitas tulisan akademis mereka melalui sejumlah fungsi, termasuk pemeriksaan tata bahasa, konstruksi kalimat, pencarian referensi, dan perbaikan struktur tulisan. AI memungkinkan siswa mendapatkan umpan balik dengan cepat, sehingga meningkatkan efektivitas proses pembelajaran (Prawiro et al., 2025).

Menurut penelitian, penggunaan alat parafrase berbasis AI meningkatkan keterampilan menulis siswa. Siswa mengevaluasi bagaimana teknologi ini meningkatkan konten tertulis, menurunkan kesalahan tata bahasa, dan meningkatkan standar bahasa yang digunakan dalam pekerjaan akademis. Hasil ini menunjukkan potensi besar AI sebagai alat untuk meningkatkan kemahiran berbahasa (Syahnaz & Fithriani, 2023).

Temuan serupa juga dilaporkan oleh Tresna et al. (2026) yang menunjukkan bahwa penggunaan alat parafrase berbasis AI secara signifikan meningkatkan kemampuan menulis akademik mahasiswa dibandingkan kelompok yang tidak menggunakan alat tersebut.

Kemajuan penelitian dalam bidang transfer gaya teks menunjukkan bahwa transformasi gaya bahasa tidak hanya berfokus pada perubahan pemahaman, tetapi juga mempertahankan makna utama kalimat yang diproses. Tujuan dari transfer gaya adalah mengubah ciri-ciri linguistik suatu teks agar sesuai dengan gaya yang diinginkan

tanpa mengorbankan isinya. Strategi ini sangat penting karena memungkinkan sistem menghasilkan kalimat yang tetap mempertahankan isi yang dimaksud sekaligus lebih sesuai dengan situasi komunikasi tertentu.

Formality Style Transfer merupakan proses mengubah kalimat informal menjadi bentuk formal dengan tetap mempertahankan makna dari kalimat sumber (Liu et al., 2022). Proses ini tidak hanya berfokus pada perubahan pilihan kata, tetapi juga pada penyesuaian gaya bahasa agar sesuai dengan tingkat formalitas yang diinginkan.

Penelitian sebelumnya menunjukkan bahwa proses style transfer dapat mempertahankan makna kalimat sambil mengubah karakteristik linguistik seperti struktur kalimat dan pilihan kata (Bujnowski et al., 2020). Model berbasis GPT juga telah menunjukkan kemampuan untuk melakukan transformasi gaya bahasa informal ke formal dengan tetap mempertahankan makna asli pesan (De Rivero et al., 2023).

Penelitian mengenai text style transfer dalam bahasa Indonesia juga telah dilakukan dengan pendekatan unsupervised untuk mengubah tingkat formalitas teks. Sari dan Al Faridzi (2023) menunjukkan bahwa pendekatan style transfer dapat digunakan untuk mengubah gaya bahasa dengan tetap mempertahankan isi atau makna teks. Temuan tersebut menunjukkan bahwa style transfer memiliki potensi untuk menangani transformasi gaya bahasa pada teks berbahasa Indonesia.

Perkembangan Large Language Model (LLM) juga memungkinkan proses style transfer dilakukan melalui pendekatan berbasis prompting tanpa melakukan fine-tuning model untuk setiap gaya bahasa (Reif et al., 2022). Pendekatan tersebut menunjukkan bahwa model generatif dapat diarahkan untuk melakukan transformasi gaya bahasa melalui instruksi yang sesuai.

Namun, pendekatan tersebut belum memanfaatkan mekanisme retrieval berbasis semantic similarity untuk mengambil contoh kalimat yang relevan dari dataset dan menggunakannya sebagai referensi bagi model generatif. Berbagai penelitian menunjukkan bahwa model berbasis *Large Language Model (LLM)* maupun machine translation mampu menghasilkan kalimat formal dengan kualitas yang baik. Namun, pendekatan tersebut umumnya menghasilkan keluaran berdasarkan kemampuan generatif model tanpa secara eksplisit memanfaatkan contoh kalimat formal yang memiliki kemiripan makna dari dataset sebagai referensi. Kondisi tersebut dapat menyebabkan keluaran yang dihasilkan masih bersifat umum dan belum sepenuhnya mengikuti pola komunikasi pada konteks tertentu, seperti komunikasi dengan dosen, administrasi kampus, LinkedIn, dan email lamaran kerja.

Berdasarkan kondisi tersebut, terdapat research gap berupa kebutuhan akan pendekatan yang tidak hanya melakukan generasi bahasa formal, tetapi juga memanfaatkan contoh kalimat yang relevan secara semantik sebagai referensi gaya bahasa. Novelty penelitian ini terletak pada penerapan *Retrieval-Augmented Style Transfer (RAST)* yang mengintegrasikan *Sentence-BERT Similarity* sebagai mekanisme *retrieval* dan Gemini API sebagai model generatif. *Sentence-BERT Similarity* digunakan untuk mengambil tiga contoh kalimat dengan kemiripan makna tertinggi dari dataset, kemudian contoh tersebut dimasukkan ke dalam prompt Gemini API bersama input pengguna dan konteks komunikasi. Dengan mekanisme tersebut, penelitian ini tidak hanya berfokus pada peningkatan formalitas bahasa, tetapi juga mengarahkan keluaran agar lebih sesuai dengan karakteristik konteks komunikasi yang dipilih.

Penelitian ini bertujuan mengembangkan dan mengevaluasi sistem AI parafrase bahasa Indonesia informal ke formal menggunakan kombinasi *Sentence-BERT Similarity* dan Gemini API. Evaluasi dilakukan melalui pengujian keluaran pada empat konteks komunikasi serta perbandingan antara penggunaan Gemini API tanpa *retrieval* dan kombinasi *Sentence-BERT Similarity* dengan Gemini API. Pendekatan ini diharapkan dapat memberikan kontribusi terhadap pengembangan aplikasi *Natural Language Processing (NLP)* bahasa Indonesia, khususnya dalam penerapan *semantic similarity* dan *style transfer* untuk menghasilkan komunikasi formal yang lebih sesuai konteks.

2. Metode Penelitian

Penelitian ini menggunakan pendekatan *Research and Development (R&D)* untuk mengembangkan sistem AI parafrase bahasa informal ke formal. Sistem yang dikembangkan menggunakan *Sentence-BERT Similarity* dan Gemini API. *Sentence-BERT Similarity* digunakan untuk mencari contoh kalimat dalam dataset yang memiliki kemiripan makna dengan input pengguna, sedangkan Gemini API digunakan untuk menghasilkan kalimat formal

melalui proses *style transfer*. Konsep penggabungan mekanisme *retrieval* dengan model generatif dibahas secara sistematis oleh Murtiyoso et al. (2025) yang menunjukkan bahwa pendekatan retrieval-augmented generation memungkinkan model bahasa besar memperoleh pengetahuan tambahan dari sumber eksternal sebelum menghasilkan jawaban, sehingga keluaran yang dihasilkan lebih relevan dengan domain yang dituju.

Secara konsep, alur sistem ini menerapkan pendekatan *Retrieval Augmented Style Transfer (RAST)*, yaitu proses pencarian contoh relevan dari dataset sebelum teks diproses oleh model generatif. Contoh hasil pencarian tersebut digunakan sebagai referensi dalam penyusunan prompt agar Gemini API dapat menghasilkan parafrase formal yang lebih sesuai dengan konteks komunikasi pengguna.



2.1 Dataset

Dataset yang digunakan dalam penelitian ini berupa pasangan kalimat informal dan formal. Kalimat informal berfungsi sebagai input, sedangkan kalimat formal berfungsi sebagai referensi output. Dataset disusun dalam format CSV dengan beberapa kolom, yaitu id, kategori, *context*, *input_raw*, dan *output_formal*.

Kolom *input_raw* berisi kalimat informal, sedangkan *output_formal* berisi bentuk kalimat formal. Kolom *context* digunakan untuk membedakan konteks komunikasi, seperti dosen, admin kampus, LinkedIn, dan email lamaran kerja. Pembagian konteks ini bertujuan agar sistem dapat menghasilkan kalimat formal yang sesuai dengan kebutuhan pengguna.

Tabel 1. Distribusi Dataset Berdasarkan Konteks

<i>Context</i>	Nilai
Dosen	650
Admin Kampus	160
Email Lamaran	150
LinkedIn	150
Total data	1.110

2.2 Pra-Pemrosesan Data

Tahap pra-pemrosesan dilakukan untuk membersihkan dan menormalkan input sebelum diproses oleh sistem. Proses ini meliputi case folding, penghapusan spasi berlebih, pembersihan simbol yang tidak diperlukan, serta normalisasi beberapa kata informal dan singkatan. Contohnya, kata “ga” dinormalisasi menjadi “tidak”, “pa” menjadi “Pak”, dan “telat” menjadi “terlambat”.

Pra-pemrosesan diperlukan karena bahasa informal sering mengandung singkatan, kesalahan pengetikan, dan struktur kalimat yang tidak baku. Dengan proses ini, input pengguna menjadi lebih bersih sehingga lebih mudah dibandingkan dengan data dalam dataset.

2.3 Pemilihan Konteks Pesan

Setelah input diproses, pengguna memilih konteks pesan sesuai kebutuhan. Konteks yang digunakan dalam sistem ini meliputi dosen, admin kampus, LinkedIn, dan email lamaran kerja. Pemilihan konteks diperlukan karena setiap situasi komunikasi memiliki gaya bahasa yang berbeda.

Sebagai contoh, pesan kepada dosen membutuhkan bahasa yang sopan dan natural dengan sapaan “Pak” atau “Bu”. Sementara itu, email lamaran kerja membutuhkan bahasa yang lebih resmi, profesional, dan terstruktur.

2.4 Pencarian Contoh Terdekat dengan *Sentence-BERT Similarity*

Model *Sentence-BERT Similarity* yang digunakan dalam penelitian ini adalah *sentence-transformers/paraphrase-multilingual-MiniLM-L12-v2* sebagai model *Sentence-BERT* untuk menghasilkan *embedding* kalimat. Penelitian Holis et al. (2025) yang menerapkan model *paraphrase-multilingual-MiniLM-L12-v2* pada sistem tanya jawab akademik menunjukkan bahwa pendekatan *Sentence-BERT* dengan *cosine similarity* mampu mengenali pertanyaan secara semantik mirip meskipun disampaikan dengan formulasi kalimat yang berbeda. Model ini digunakan untuk membandingkan kemiripan makna antara input pengguna dan data dalam dataset menggunakan *cosine similarity*.

Dalam proses *text style transfer*, perubahan gaya bahasa perlu tetap mempertahankan informasi dan makna utama dari teks sumber. Kesamaan semantik antara teks sumber dan hasil transformasi menjadi salah satu aspek penting dalam menjaga informasi tersebut (Nouri, 2022). Berdasarkan prinsip tersebut, penelitian ini menggunakan *Sentence-BERT Similarity* untuk merepresentasikan kalimat ke dalam bentuk *embedding* dan mencari contoh kalimat dalam dataset yang memiliki kemiripan makna dengan input pengguna. Dalam sistem ini, tiga contoh dengan nilai *Similarity* tertinggi atau top-k = 3 digunakan sebagai referensi dalam penyusunan *prompt* sebelum diproses oleh Gemini API.

$$\text{CosSim}(A, B) = \frac{A \cdot B}{|A| \times |B|}$$

Keterangan :

A = *embedding* input pengguna

B = *embedding* dataset kalimat dalam dataset

cosine similarity digunakan untuk mengukur tingkat kemiripan semantik antara kalimat input dan kalimat dalam dataset. Nilai *Similarity* berada pada rentang 0 hingga 1, di mana nilai yang semakin mendekati 1 menunjukkan tingkat kemiripan makna yang semakin tinggi. Pada proses pencarian contoh terdekat, input pengguna terlebih dahulu dinormalisasi, kemudian diubah menjadi *embedding* menggunakan model *Sentence-BERT*. Selanjutnya, *embedding* input dibandingkan dengan *embedding* seluruh data pada dataset menggunakan *cosine similarity*. Sistem mengambil tiga contoh dengan nilai kemiripan tertinggi atau top-k = 3 sebagai referensi dalam penyusunan *prompt*. Contoh terdekat tersebut kemudian diberikan kepada Gemini API agar output parafrase formal lebih sesuai dengan pola bahasa dalam dataset.

2.5 Penyusunan *Prompt* Berbasis Dataset

Setelah contoh terdekat ditemukan, sistem menyusun *prompt* yang berisi input pengguna, hasil pra-pemrosesan, konteks pesan, dan contoh kalimat dari dataset. *Prompt* ini berfungsi sebagai instruksi kepada Gemini API agar output yang dihasilkan sesuai dengan konteks dan gaya bahasa formal yang diinginkan.

Dengan adanya contoh dari dataset, Gemini API tidak hanya menghasilkan kalimat formal secara umum, tetapi juga diarahkan untuk mengikuti pola bahasa formal yang sudah tersedia dalam dataset.

2.6 *Style transfer* Menggunakan Gemini API

Tahap utama dalam sistem ini adalah *Style transfer* menggunakan Gemini API. *Style transfer* merupakan proses mengubah gaya bahasa suatu teks dari satu bentuk ke bentuk lain tanpa menghilangkan makna utama. Dalam penelitian ini, *Style transfer* digunakan untuk mengubah bahasa informal menjadi bahasa formal.

Hariyadi et al. (2024) dalam penelitiannya yang mengintegrasikan Gemini API pada aplikasi berbasis web menunjukkan bahwa API tersebut mampu menghasilkan teks secara otomatis sesuai kebutuhan aplikasi yang dikembangkan, sehingga mengurangi waktu dan usaha yang diperlukan untuk penyusunan teks secara manual. Penelitian ini menggunakan model Gemini 2.5 Flash melalui Gemini API. *Prompt* disusun berdasarkan konteks komunikasi, input pengguna, serta tiga contoh kalimat hasil *retrieval* menggunakan *Sentence-BERT Similarity*. Parameter temperature diatur sebesar 0,3 untuk menghasilkan keluaran yang lebih konsisten, sedangkan parameter top-k menggunakan nilai default dari Gemini API.

Gemini API digunakan sebagai model berbasis Transformer untuk menghasilkan kalimat formal akhir. Model ini memproses *prompt* yang telah disusun sebelumnya, kemudian menghasilkan output yang lebih sopan, jelas, natural, dan sesuai konteks komunikasi.

2.7 Prosedur Evaluasi Sistem

Evaluasi sistem dilakukan untuk mengetahui kualitas keluaran parafrase bahasa Indonesia informal menjadi formal serta mengetahui pengaruh penggunaan mekanisme *retrieval* terhadap hasil generasi. Evaluasi dilakukan melalui dua tahap, yaitu evaluasi kualitas output secara manual dan pengujian pembandingan antara penggunaan Gemini API tanpa *retrieval* dengan kombinasi *Sentence-BERT Similarity* dan Gemini API.

Evaluasi kualitas hasil style transfer tidak hanya dapat dilakukan menggunakan metrik otomatis, tetapi juga dapat melibatkan penilaian manusia untuk menilai kualitas keluaran dari perspektif pengguna atau evaluator. Lai et al. (2022) menunjukkan pentingnya human judgement dalam evaluasi formality transfer karena kualitas keluaran tidak selalu dapat direpresentasikan secara memadai melalui metrik otomatis.

Berdasarkan pertimbangan tersebut, penelitian ini menggunakan evaluasi manual terhadap keluaran sistem. Evaluasi dilakukan oleh tiga evaluator terhadap 12 output yang mewakili empat konteks komunikasi, yaitu dosen, admin kampus, LinkedIn, dan email lamaran kerja. Setiap evaluator menilai seluruh output secara independen menggunakan skala Likert 1–5 berdasarkan empat aspek, yaitu kesesuaian makna, formalitas bahasa, kelancaran bahasa, dan kesesuaian konteks. Skor 1 menunjukkan kualitas yang sangat rendah, sedangkan skor 5 menunjukkan kualitas yang sangat baik. Nilai rata-rata setiap aspek dihitung berdasarkan seluruh penilaian evaluator untuk memperoleh gambaran kualitas keluaran sistem.

Selain evaluasi manual, dilakukan pengujian pembandingan untuk melihat perbedaan keluaran antara dua skenario. Skenario pertama menggunakan Gemini API secara langsung tanpa contoh hasil *retrieval*. Skenario kedua menggunakan *Sentence-BERT Similarity* untuk mengambil tiga contoh kalimat yang memiliki kemiripan semantik tertinggi dari dataset, kemudian contoh tersebut digunakan sebagai referensi dalam prompt Gemini API. Kedua skenario menggunakan input yang sama sehingga perbedaan keluaran dapat diamati secara deskriptif berdasarkan kesesuaian hasil terhadap konteks komunikasi.

Penilaian kualitas output digunakan untuk melihat sejauh mana sistem mampu mempertahankan makna utama, meningkatkan formalitas bahasa, menghasilkan kalimat yang lancar, serta menyesuaikan keluaran dengan konteks komunikasi yang dipilih. Nilai rata-rata dihitung menggunakan persamaan berikut:

Tabel 2. Skala Penilaian Evaluasi Kualitas Output

Skor	Keterangan
1	Sangat Buruk
2	Kurang Baik
3	Cukup
4	Baik
5	Sangat Baik

Penilaian ini menggunakan rumus :

$$\bar{x} = \frac{\sum x}{n}$$

Keterangan:

$\bar{x}$ = skor rata-rata

$\sum x$ = total seluruh skor evaluator

n = jumlah penilaian

Hasil evaluasi manual dan pengujian pembandingan selanjutnya digunakan sebagai dasar untuk menilai kemampuan sistem dalam menghasilkan parafrase formal yang tetap mempertahankan makna utama serta sesuai dengan konteks komunikasi pengguna.

2.8 Evaluasi Model Klasifikasi

Model klasifikasi kategori digunakan sebagai komponen pendukung untuk mengenali kategori pesan berdasarkan teks input. Model dibangun menggunakan kombinasi *TF-IDF Vectorizer* dan *Logistic Regression*. *TF-IDF* digunakan untuk mengubah teks menjadi representasi numerik, sedangkan *Logistic Regression* digunakan untuk melakukan klasifikasi berdasarkan representasi tersebut. Konfigurasi *TF-IDF* menggunakan *ngram_range=(1,2)* sehingga fitur yang digunakan mencakup unigram dan bigram. Model *Logistic Regression* menggunakan parameter *max_iter=1000*.

Dataset klasifikasi terdiri atas 888 data berlabel. Data dibagi menjadi data pelatihan dan data pengujian menggunakan fungsi *train_test_split* dengan *random_state=42*. Dengan pengaturan pembagian default, sebanyak 666 data (75%) digunakan sebagai data pelatihan dan 222 data (25%) digunakan sebagai data pengujian. Model dilatih menggunakan data pelatihan, kemudian digunakan untuk memprediksi kategori pada data pengujian.

Kinerja model dievaluasi menggunakan *accuracy*, *precision*, *recall*, dan *F1-score*. *Accuracy* digunakan untuk mengukur proporsi prediksi yang benar terhadap seluruh data pengujian. *Precision* mengukur ketepatan prediksi pada suatu kategori, *recall* mengukur kemampuan model dalam mengenali data yang termasuk dalam kategori tersebut, sedangkan *F1-score* merupakan nilai gabungan antara *precision* dan *recall*.

Perhitungan metrik evaluasi menggunakan persamaan berikut:

$$Accuracy = (TP + TN) / (TP + TN + FP + FN)$$

$$Precision = TP / (TP + FP)$$

$$Recall = TP / (TP + FN)$$

$$F1 - score = 2 \times (Precision \times Recall) / (Precision + Recall)$$

Keterangan: TP adalah *True Positive*, TN adalah *True Negative*, FP adalah *False Positive*, dan FN adalah *False Negative*.

Selain nilai pada masing-masing kategori, evaluasi juga menggunakan *macro average* dan *weighted average* untuk memberikan gambaran performa model secara keseluruhan.

2.9 Skenario Pengujian Pembandingan

Penelitian ini menambahkan skenario pengujian pembandingan untuk mengetahui pengaruh penggunaan *Sentence-BERT Similarity* terhadap hasil parafrase. Pengujian dilakukan dengan membandingkan dua skenario, yaitu Gemini API saja dan *Sentence-BERT* + Gemini API. Pada skenario Gemini API saja, input pengguna langsung diproses oleh Gemini API tanpa menggunakan contoh terdekat dari dataset. Sementara itu, pada skenario *Sentence-BERT* + Gemini API, sistem terlebih dahulu mencari tiga contoh kalimat terdekat dari dataset menggunakan *Sentence-BERT Similarity*, kemudian contoh tersebut digunakan sebagai referensi dalam penyusunan *prompt* sebelum diproses oleh Gemini API.

Tabel 3. Skenario Pengujian Pemanding Skenario

Skenario	Proses Pengujian	Tujuan
Gemini API	Input pengguna langsung diproses oleh Gemini API tanpa contoh terdekat dari dataset.	Mengetahui kualitas output Gemini API tanpa bantuan <i>retrieval</i> dataset.
<i>Sentence-BERT</i> + Gemini API	Input pengguna dicocokkan dengan dataset menggunakan <i>Sentence-BERT Similarity</i> , kemudian tiga contoh terdekat dimasukkan ke dalam <i>prompt</i> Gemini API.	Mengetahui pengaruh contoh dataset terhadap kualitas output parafrase formal.

Berdasarkan Tabel 3, pengujian pembandingan ini digunakan untuk melihat apakah penggunaan *Sentence-BERT Similarity* sebagai metode *retrieval* dapat membantu Gemini API menghasilkan output yang lebih sesuai dengan konteks komunikasi pengguna.

3. Hasil dan Pembahasan

3.1 Hasil Implementasi Sistem

Hasil dari penelitian ini adalah sistem AI parafrase yang dapat mengubah bahasa informal menjadi bahasa formal berdasarkan konteks komunikasi. Sistem ini dikembangkan dengan menggabungkan *Sentence-BERT Similarity* dan *Style transfer* berbasis Gemini API. Pengembangan sistem ini dilakukan menggunakan laptop Acer Swift SFX14-41G dengan sistem operasi Microsoft Windows 11 64bit, prosesor AMD Ryzen 7 5800U with Radeon Graphics, 8 core, dan 16 logical processor. Perangkat ini digunakan untuk menyusun kode program, mengelola dataset, melakukan pengujian sistem, serta mendukung proses deployment aplikasi berbasis web melalui Hugging Face Spaces.

Sentence-BERT Similarity berperan sebagai metode pencarian contoh terdekat dari dataset berdasarkan kemiripan makna. Contoh tersebut kemudian digunakan sebagai referensi dalam penyusunan *prompt*. Setelah itu, Gemini API digunakan untuk melakukan *Style transfer* agar input informal dapat diubah menjadi kalimat formal yang lebih sopan, jelas, dan sesuai konteks.

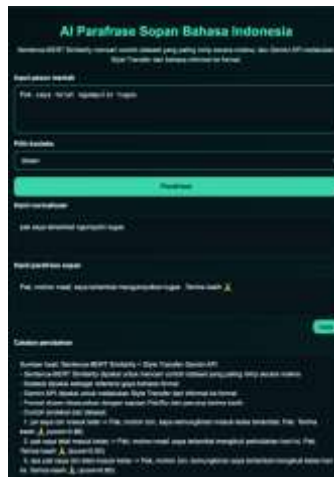
3.2 Hasil Pengujian Output Informal ke Formal

Pengujian dilakukan dengan memasukkan beberapa contoh kalimat informal ke dalam sistem. Hasil pengujian dapat dilihat pada tabel di bawah ini.

Tabel 4. Contoh Konteks Kalimat Informal dan Formal

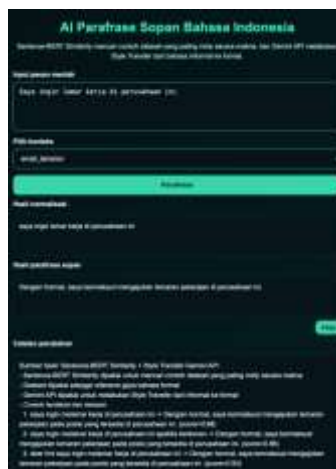
Konteks	Input Informal	Output Formal
Dosen	Pak saya telat ngumpul tugas	Pak, mohon maaf, saya terlambat mengumpulkan tugas. Terima kasih
Email Lamaran	Saya ingin lamar kerja di perusahaan ini	Dengan hormat, Saya bermaksud mengajukan lamaran pekerjaan di perusahaan ini.
Admin Kampus	Min KRS bisa diubah ga	Halo, mohon izin bertanya, apakah KRS masih dapat diubah? Terima kasih
LinkedIn	Kak saya lagi cari kerja nih, kalau ada info boleh kabarin ya.	Halo kak, saya sedang aktif mencari peluang kerja. Apabila ada informasi yang relevan, mohon dapat mengabari saya. Terima kasih

Berdasarkan Tabel 4, sistem juga telah diimplementasikan pada platform web berbasis Hugging Face Spaces. Hasil Implementasi ditampilkan pada Gambar 1 sampai 4.



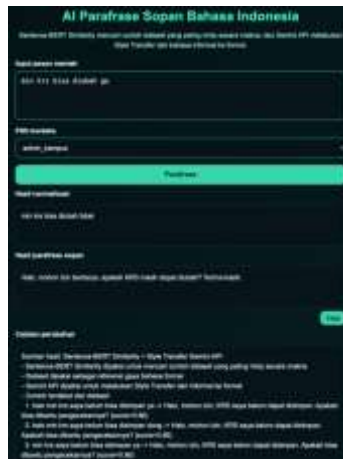
Gambar 1. Contoh Hasil Pengujian Konteks Dosen

Berdasarkan Gambar 1, sistem berhasil mengubah kalimat informal menjadi kalimat formal yang sesuai dengan konteks komunikasi akademik antara mahasiswa dan dosen. Pada proses ini, pengguna memilih kategori dosen sehingga sistem menghasilkan kalimat yang lebih sopan, santun, dan menghormati penerima pesan. Selain melakukan perubahan pada kata-kata informal, sistem juga memperbaiki struktur kalimat serta menambahkan unsur kesopanan yang umum digunakan dalam komunikasi akademik. Hasil tersebut menunjukkan bahwa kombinasi *Sentence-BERT Similarity* dan Gemini API mampu menghasilkan parafrase yang tetap mempertahankan makna asli pesan namun dengan tingkat formalitas yang lebih tinggi.



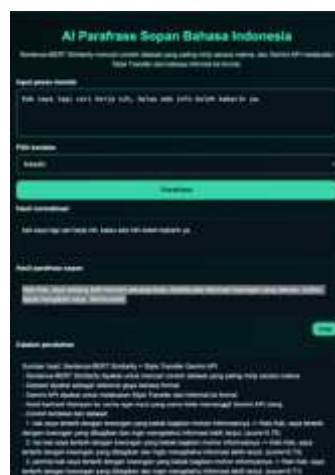
Gambar 2. Contoh Hasil Pengujian Konteks Email Lamaran

Berdasarkan Gambar 2, sistem mampu menghasilkan kalimat formal yang sesuai untuk kebutuhan email lamaran kerja. Pada konteks ini, bahasa yang digunakan lebih profesional, terstruktur, dan berorientasi pada komunikasi formal dalam dunia kerja. Kalimat hasil parafrase menunjukkan adanya penyesuaian gaya bahasa yang berbeda dibandingkan konteks dosen, yaitu penggunaan diksi yang lebih resmi dan penyampaian informasi yang lebih sistematis. Hal ini menunjukkan bahwa sistem tidak hanya melakukan parafrase secara umum, tetapi juga mempertimbangkan konteks komunikasi yang dipilih pengguna sebelum menghasilkan keluaran akhir.



Gambar 3. Contoh Hasil Pengujian Konteks Admin Kampus

Berdasarkan Gambar 3, sistem berhasil menghasilkan kalimat formal yang sesuai untuk komunikasi dengan pihak administrasi kampus. Output yang dihasilkan memiliki karakteristik lebih ringkas, jelas, dan informatif dibandingkan konteks email lamaran kerja. Sistem mampu menyesuaikan gaya bahasa agar tetap sopan tanpa menggunakan kalimat yang terlalu panjang. Hasil ini menunjukkan bahwa pendekatan berbasis konteks yang diterapkan pada sistem dapat membantu menghasilkan bentuk komunikasi yang lebih relevan sesuai kebutuhan pengguna. Dengan demikian, proses *style transfer* tidak hanya meningkatkan formalitas bahasa, tetapi juga menyesuaikan karakteristik komunikasi berdasarkan tujuan pesan yang dipilih.



Gambar 4. Contoh Hasil Pengujian Konteks LinkedIn

Berdasarkan Gambar 4, sistem mampu menghasilkan kalimat formal yang sesuai dengan konteks komunikasi profesional pada LinkedIn. Input informal diubah menjadi pesan yang lebih sopan, jelas, dan profesional tanpa menghilangkan maksud utama pengguna. Hasil ini menunjukkan bahwa sistem dapat menyesuaikan gaya bahasa untuk kebutuhan networking atau komunikasi profesional.

Secara keseluruhan, hasil pengujian menunjukkan bahwa sistem AI parafrase yang dikembangkan mampu mengubah kalimat informal menjadi bahasa formal pada berbagai konteks komunikasi.

3.3 Analisis *Sentence-BERT* Similarity

Sentence-BERT Similarity membantu sistem mencari contoh kalimat yang memiliki kemiripan makna dengan input pengguna. Hasil ini sejalan dengan penelitian Wang et al. (2020) yang menunjukkan bahwa proses *retrieval* mampu membantu model menghasilkan keluaran yang lebih sesuai dengan konteks.

Dengan *Sentence-BERT*, sistem dapat membandingkan makna kalimat dalam bentuk *embedding*. Hasil pencarian contoh terdekat kemudian digunakan sebagai referensi gaya bahasa formal. Hal ini membuat sistem lebih terarah karena Gemini API tidak hanya menerima input pengguna, tetapi juga menerima contoh formal yang relevan dari dataset.

Kelebihan *Sentence-BERT* dalam sistem ini adalah kemampuannya mengenali kemiripan makna, bukan hanya kemiripan kata. Namun, kualitas pencarian tetap bergantung pada variasi dataset. Jika dataset belum memiliki contoh yang cukup beragam, contoh yang ditemukan masih dapat kurang sesuai.

3.4 Analisis *Style transfer* Menggunakan Gemini API

Gemini API digunakan untuk melakukan *Style transfer* dari bahasa informal ke bahasa formal. Proses ini tidak hanya mengganti kata informal menjadi kata formal, tetapi juga memperbaiki struktur kalimat, menambahkan unsur kesopanan, dan menyesuaikan gaya bahasa berdasarkan konteks.

Pada konteks dosen, sistem menghasilkan kalimat yang lebih santun dan sesuai dengan komunikasi akademik. Pada konteks admin kampus, output dibuat lebih singkat dan jelas. Pada konteks LinkedIn dan email lamaran kerja, output dibuat lebih profesional dan terstruktur.

Dengan demikian, Gemini API berperan sebagai model generatif yang menyusun output akhir, sedangkan *Sentence-BERT Similarity* berperan sebagai metode pencarian referensi dari dataset. Kombinasi ini menunjukkan bahwa pendekatan *Retrieval Augmented Style Transfer* (RAST) dapat membantu sistem menghasilkan parafrase formal yang lebih sesuai dengan kebutuhan komunikasi pengguna.

3.5 Hasil Evaluasi Kualitas Output

Evaluasi kualitas parafrase dilakukan secara manual dengan melibatkan tiga evaluator. Setiap evaluator memberikan penilaian terhadap 12 output parafrase yang mewakili konteks dosen, email lamaran kerja, admin kampus dan LinkedIn. Perhitungannya adalah sebagai berikut.

Perhitungan rata-rata keseluruhan :

$$\bar{x} = \frac{4,83 + 4,53 + 4,28 + 4,53}{4} = 4,54$$

Tabel 5. Rekapitulasi Penilaian Evaluator

Aspek	Evaluator 1	Evaluator 2	Evaluator 3	Total Nilai	Rata-rata
Kesesuaian Makna	58	57	59	174	4,83
Formalitas	55	54	54	163	4,53
Kelancaran Bahasa	52	49	53	154	4,28
Kesesuaian Konteks	56	53	54	163	4,53

Tabel 6. Hasil Evaluasi Kualitas Output

Aspek Evaluasi	Skor Rata-rata	Interpretasi
Kesesuaian Makna	4,83	Sangat Baik
Formalitas	4,53	Sangat Baik
Kelancaran Bahasa	4,28	Baik
Kesesuaian Konteks	4,53	Sangat Baik
Rata-rata keseluruhan	4,54	Sangat Baik

Berdasarkan Tabel 5 dan Tabel 6, hasil evaluasi menunjukkan bahwa aspek kesesuaian makna memperoleh skor rata-rata tertinggi sebesar 4,83 dengan kategori sangat baik. Hasil ini menunjukkan bahwa sistem mampu mempertahankan makna utama dari kalimat sumber setelah dilakukan transformasi ke bahasa formal. Aspek formalitas memperoleh skor rata-rata sebesar 4,53 dan aspek kesesuaian konteks memperoleh skor rata-rata

sebesar 4,53, yang menunjukkan bahwa output sistem sudah sesuai dengan kaidah komunikasi formal dan konteks yang dipilih pengguna. Sementara itu, aspek kelancaran bahasa memperoleh skor rata-rata sebesar 4,28 dengan kategori baik. Secara keseluruhan, sistem memperoleh nilai rata-rata sebesar 4,54 yang termasuk dalam kategori sangat baik. Hasil ini menunjukkan bahwa pendekatan *Retrieval-Augmented Style transfer* menggunakan *Sentence-BERT Similarity* dan Gemini API mampu menghasilkan kalimat formal yang sesuai dengan konteks komunikasi akademik dan profesional.

3.6 Perbandingan Hasil Gemini API saja dan *Sentence-BERT* + Gemini API

Pengujian pembandingan dilakukan dengan menggunakan input yang sama pada dua skenario, yaitu Gemini API saja dan *Sentence-BERT* + Gemini API. Skenario Gemini API saja digunakan sebagai pembandingan karena input langsung diproses oleh Gemini API tanpa contoh formal dari dataset. Sementara itu, skenario *Sentence-BERT* + Gemini API merupakan metode utama dalam sistem yang dikembangkan, karena input pengguna terlebih dahulu dicocokkan dengan dataset menggunakan *Sentence-BERT Similarity* sebelum diproses oleh Gemini API.

Tabel 7. Perbandingan Output Gemini API saja dan *Sentence-BERT* + Gemini API

Konteks	Input Informal	Gemini API	<i>Sentence-BERT</i> + Gemini API
Dosen	Pa saya telat ngumpul tugas	Pak, mohon maaf sebelumnya, saya telat mengumpulkan tugas. Terima kasih 🙏	Pak, mohon maaf, saya terlambat mengumpulkan tugas. Terima kasih 🙏
Email Lamaran	Saya ingin kerja di perusahaan ini	Saya sangat tertarik untuk dapat bergabung dan berkontribusi di perusahaan ini.	Dengan hormat, saya bermaksud mengajukan lamaran pekerjaan di perusahaan ini.
Admin Kampus	Min KRS bisa diubah ga	Halo, mohon izin bertanya, apakah KRS masih dapat diubah? Terima kasih.	Halo, mohon izin bertanya, apakah KRS masih dapat diubah? Terima kasih.
LinkedIn	Kak saya lagi cari kerja nih, kalau ada info boleh kabarin ya.	Saya sedang aktif mencari peluang pekerjaan. Apabila Bapak/Ibu memiliki informasi terkait lowongan yang relevan, saya akan sangat menghargai jika dapat menginformasikannya. Terima kasih	Halo kak, saya sedang aktif mencari peluang kerja. Apabila ada informasi yang relevan, mohon dapat mengabari saya. Terima kasih

Berdasarkan Tabel 7, hasil Gemini API saja sudah mampu menghasilkan kalimat yang lebih formal dibandingkan input awal. Namun, pada beberapa konteks, output yang dihasilkan masih bersifat umum dan belum sepenuhnya mengikuti pola komunikasi yang sesuai dengan dataset. Hal ini terlihat pada konteks email lamaran, di mana Gemini API saja menghasilkan kalimat yang bersifat pernyataan ketertarikan, sedangkan *Sentence-BERT* + Gemini API menghasilkan kalimat yang lebih sesuai dengan format lamaran kerja.

Pada skenario *Sentence-BERT* + Gemini API, hasil parafrase cenderung lebih terarah karena sistem menggunakan contoh terdekat dari dataset sebagai referensi gaya bahasa formal. *Sentence-BERT Similarity* membantu sistem mencari contoh yang memiliki kemiripan makna dengan input pengguna, sedangkan Gemini API menyusun ulang kalimat menjadi bentuk formal akhir. Dengan demikian, *Sentence-BERT* tidak berperan sebagai model generatif, melainkan sebagai metode *retrieval* yang memperkuat penyusunan *prompt* agar output Gemini API lebih sesuai dengan konteks komunikasi.

Hasil perbandingan ini menunjukkan bahwa penggunaan *Sentence-BERT Similarity* dapat membantu meningkatkan kesesuaian output terhadap konteks tertentu, terutama pada konteks yang membutuhkan format bahasa lebih spesifik seperti email lamaran kerja. Dengan adanya *retrieval* dari dataset, sistem tidak hanya menghasilkan kalimat formal secara umum, tetapi juga menyesuaikan gaya bahasa berdasarkan contoh formal yang relevan.

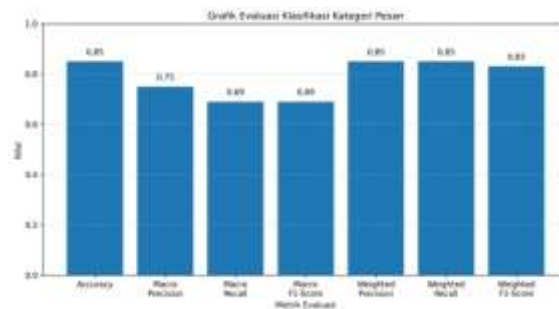
3.7 Hasil Klasifikasi Kategori Pesan

Selain pengujian utama terhadap kemampuan sistem dalam melakukan parafrase, penelitian ini juga melakukan analisis tambahan untuk mengevaluasi kemampuan model dalam mengenali kategori pesan berdasarkan teks input. Analisis klasifikasi ini digunakan sebagai evaluasi pendukung terhadap kemampuan pengenalan kategori pesan dan bukan sebagai komponen utama dalam penentuan konteks pada sistem. Pada proses utama sistem, konteks komunikasi tetap dipilih oleh pengguna sesuai dengan kebutuhan pesan. Berdasarkan

implementasi pada Google Colab, model klasifikasi dibangun menggunakan kombinasi *TF-IDF Vectorizer* dan *Logistic Regression*. *TF-IDF* digunakan untuk merepresentasikan teks dalam bentuk numerik, sedangkan *Logistic Regression* digunakan untuk melakukan klasifikasi berdasarkan representasi tersebut. Konfigurasi *TF-IDF* menggunakan *ngram_range*=(1,2) sehingga representasi mempertimbangkan unigram dan bigram, sedangkan *Logistic Regression* menggunakan *max_iter*=1000.

Data yang digunakan untuk klasifikasi berjumlah 888 data berlabel. Data dibagi menggunakan fungsi *train_test_split* dengan *random_state*=42, dengan 75% data sebanyak 666 data digunakan sebagai data pelatihan dan 25% sebanyak 222 data digunakan sebagai data pengujian. Model kemudian dilatih menggunakan data pelatihan dan digunakan untuk memprediksi kategori pada data pengujian.

Evaluasi dilakukan menggunakan *accuracy*, *precision*, *recall*, dan *F1-score*. Hasil pengujian terhadap 222 data menunjukkan bahwa model memperoleh akurasi sebesar 84,68%. Nilai macro average yang diperoleh adalah *precision* sebesar 0,75, *recall* sebesar 0,69, dan *F1-score* sebesar 0,69. Sementara itu, nilai *weighted average* menghasilkan *precision* sebesar 0,85, *recall* sebesar 0,85, dan *F1-score* sebesar 0,83.



Gambar 5. Grafik Metrik Evaluasi Klasifikasi

Berdasarkan Gambar 5, hasil tersebut menunjukkan bahwa model mampu melakukan klasifikasi kategori dengan tingkat ketepatan sebesar 84,68% pada data pengujian. Nilai *weighted F1-score* sebesar 0,83 juga menunjukkan performa yang cukup baik ketika mempertimbangkan jumlah data pada masing-masing kategori. Namun, nilai *macro F1-score* yang lebih rendah, yaitu 0,69, menunjukkan bahwa performa model belum merata pada seluruh kategori.

Perbedaan performa antar kategori terlihat pada hasil classification report. Beberapa kategori dengan jumlah data pengujian yang relatif banyak memperoleh performa yang baik, seperti *daftar ulang* dengan *F1-score* 0,94, *lamaran* sebesar 0,95, serta *krs*, *kirim cv*, *lowongan*, *networking*, *perubahan jadwal*, *presentasi*, dan *ruang kelas* yang memperoleh *F1-score* 1,00. Sebaliknya, beberapa kategori dengan jumlah data yang sedikit memperoleh nilai *F1-score* rendah atau bahkan 0, seperti *admin dosen*, *dosen wali*, *etika umum*, *jadwal ketemu*, dan *magang*. Kondisi tersebut menunjukkan bahwa jumlah sampel pada setiap kategori masih menjadi salah satu faktor yang memengaruhi kemampuan model dalam mengenali kategori pesan.

Berdasarkan hasil tersebut, model TF-IDF dan Logistic Regression dapat digunakan sebagai komponen pendukung untuk mengenali kategori pesan berdasarkan teks input. Hasil klasifikasi menunjukkan bahwa model memiliki kemampuan yang cukup baik dalam mengenali kategori pesan, meskipun performanya belum merata pada seluruh kategori. Evaluasi klasifikasi ini bersifat tambahan dan bukan merupakan komponen utama dalam proses parafrase. Pada proses utama, konteks komunikasi dipilih oleh pengguna, sedangkan Sentence-BERT Similarity digunakan untuk melakukan retrieval contoh yang relevan dan Gemini API digunakan untuk menghasilkan keluaran bahasa formal.

3.8 Pembahasan

Berdasarkan hasil pengujian, sistem AI parafrase yang dikembangkan mampu mengubah kalimat informal menjadi kalimat formal sesuai dengan konteks komunikasi yang dipilih pengguna. Hasil parafrase menunjukkan bahwa sistem tidak hanya melakukan penggantian kata informal ke bentuk formal, tetapi juga memperbaiki struktur kalimat serta menambahkan unsur kesopanan yang sesuai dengan konteks akademik dan profesional.

Sentence-BERT Similarity berperan dalam mencari contoh kalimat yang memiliki kemiripan makna dengan input pengguna. Contoh tersebut digunakan sebagai referensi dalam penyusunan *prompt* sehingga Gemini API dapat menghasilkan keluaran yang lebih sesuai dengan konteks komunikasi. Pendekatan *Retrieval-Augmented Style transfer (RAST)* membuat proses parafrase menjadi lebih terarah karena model generatif memperoleh referensi tambahan dari dataset sebelum menghasilkan output akhir.

Hasil pengujian pembandingan menunjukkan bahwa Gemini API saja sudah mampu menghasilkan output formal, tetapi hasilnya cenderung lebih umum karena tidak menggunakan contoh formal dari dataset. Sementara itu, skenario *Sentence-BERT* + Gemini API menghasilkan output yang lebih terarah karena sistem memanfaatkan contoh terdekat sebagai referensi gaya bahasa. Hal ini memperkuat peran *Sentence-BERT Similarity* sebagai metode *retrieval* yang mendukung proses *style transfer* berbasis Gemini API.

Hasil evaluasi kualitas output terhadap 12 output yang mewakili empat konteks komunikasi menunjukkan bahwa sistem memperoleh nilai rata-rata keseluruhan sebesar 4,54 dengan kategori sangat baik. Hasil tersebut menunjukkan bahwa sistem mampu mempertahankan makna utama kalimat sekaligus meningkatkan tingkat formalitas bahasa sesuai kebutuhan komunikasi pengguna.

4. Kesimpulan

Penelitian ini memberikan kontribusi berupa implementasi *Retrieval-Augmented Style transfer (RAST)* berbasis *Sentence-BERT Similarity* dan Gemini API untuk membantu transformasi bahasa informal ke formal pada bahasa Indonesia. Sistem yang dikembangkan mampu membantu pengguna menghasilkan kalimat yang lebih sopan, jelas, dan sesuai dengan konteks komunikasi akademik maupun profesional. *Sentence-BERT Similarity* digunakan untuk mencari contoh kalimat yang memiliki kemiripan makna dengan input pengguna, sedangkan Gemini API digunakan untuk menghasilkan keluaran formal melalui proses *style transfer*. Kombinasi kedua metode tersebut memungkinkan sistem mempertahankan makna utama pesan sekaligus menyesuaikan tingkat formalitas bahasa berdasarkan konteks yang dipilih pengguna. Hasil pengujian pembandingan menunjukkan bahwa penggunaan *Sentence-BERT Similarity* dapat membantu Gemini API menghasilkan output yang lebih sesuai dengan konteks dibandingkan penggunaan Gemini API saja. Hasil evaluasi kualitas output terhadap 12 output yang mewakili empat konteks komunikasi menunjukkan bahwa sistem memperoleh nilai rata-rata keseluruhan sebesar 4,54 dengan kategori sangat baik. Penelitian ini dapat menjadi dasar pengembangan aplikasi pendukung komunikasi formal berbahasa Indonesia pada lingkungan akademik maupun profesional. Selain membantu mahasiswa dalam menyusun pesan kepada dosen atau pihak administrasi kampus, pendekatan yang diusulkan juga berpotensi diterapkan pada komunikasi profesional, seperti email lamaran kerja maupun platform LinkedIn. Hasil penelitian ini masih memiliki keterbatasan pada jumlah evaluator yang terbatas serta evaluasi yang dilakukan pada sebagian sampel keluaran sehingga hasil penelitian belum sepenuhnya mewakili seluruh variasi data pada dataset. Oleh karena itu, penelitian selanjutnya dapat dilakukan dengan memperluas dataset, menambah jumlah evaluator, serta membandingkan hasil sistem dengan metode parafrase lainnya.

Reference

1. Ardinata, P. S., Permana, A. A., & Wijaya, W. (2024). IDENTIFIKASI DAN NORMALISASI TEKS SLANG DENGAN FASTTEXT PADA TWITTER DALAM BAHASA INDONESIA. *Jurnal Pendidikan Teknologi dan Kejuruan*, 21(1), 33–44. <https://doi.org/10.23887/jptkuniksha.v21i1.66381>
2. Budi, I. S., Putrayasa, I. B., Wisudariani, N., & Sudiana, I. N. (2024). PERAN DAN TANTANGAN PENGGUNAAN ARTIFICIAL INTELLIGENCE DALAM INOVASI PENGEMBANGAN KURIKULUM PEMBELAJARAN BAHASA INDONESIA MASA DEPAN. *LEARNING: Jurnal Inovasi Penelitian Pendidikan dan Pembelajaran*, 4(4), 1188–1194. <https://doi.org/10.51878/learning.v4i4.3767>
3. Bujnowski, P., Ryzhova, K., Choi, H., Witkowska, K., Piersa, J., Krumholz, T., & Beksa, K. (2020). An Empirical Study on Multi-Task Learning for Text Style Transfer and Paraphrase Generation. *Proceedings of the 28th International Conference on Computational Linguistics: Industry Track*, 50–63. <https://doi.org/10.18653/v1/2020.coling-industry.6>
4. De Rivero, M., Tirado, C., & Ugarte, W. (2023). Formalstyler: GPT-Based Model for Formal Style Transfer with Meaning Preservation. *SN Computer Science*, 4(6), 739. <https://doi.org/10.1007/s42979-023-02110-7>
5. Eleison, K. C., Hutahaean, S. U. I., Tampubolon, S. C., Panggabean, T. M., & Fitriyaningsih, I. (2022). An empirical evaluation of phrase-based statistical machine translation for Indonesia slang-word translator. *Indonesian Journal of Electrical Engineering and Computer Science*, 25(3), 1803. <https://doi.org/10.11591/ijeecs.v25.i3.pp1803-1813>
6. Febriyanti, B. D., & Pitasari, M. A. R. (2023). ETIKA MAHASISWA DALAM MENGHUBUNGI DOSEN. *FAJAR Jurnal Pendidikan Islam*, 2(2), 186–197. <https://doi.org/10.56013/fj.v2i2.1721>
7. Hariyadi, H., Irawan, J., Pranatawijaya, V. H., & Sari, N. N. K. (2024). Implementasi Gemini API untuk Generatif Teks Deskripsi Karya Otomatis dalam Aplikasi Pameran Berbasis Web dengan Metode Waterfall. *Jurnal Teknologi Informasi dan Terapan (J-TIT)*, 11(1). <https://doi.org/10.25047/jtit.v11i1.356>

8. Holis, R. M., Utomo, P. E. P., & Hutabarat, B. F. (2025). Semantic FAQ Chatbot Using SBERT (Sentence-BERT) and Cosine Similarity for Academic Services. *Brilliance: Research of Artificial Intelligence*, 5(2), 915–922. <https://doi.org/10.47709/brilliance.v5i2.7027>
9. Khairul, M. F. R., & Perdana, R. S. (2024). Arsitektur Sistem Percakapan Otomatis Berbahasa Indonesia dengan Normalisasi Bahasa Informal Menjadi Baku. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 11(5), 1009–1016. <https://doi.org/10.25126/jtiik.1077984>
10. Lai, H., Mao, J., Toral, A., & Nissim, M. (2022). Human Judgement as a Compass to Navigate Automatic Metrics for Formality Transfer. *Proceedings of the 2nd Workshop on Human Evaluation of NLP Systems (HumEval)*, 102–115. <https://doi.org/10.18653/v1/2022.humeval-1.9>
11. Liu, A., Wang, A., & Okazaki, N. (2022). Semi-Supervised Formality Style Transfer with Consistency Training. *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 4689–4701. <https://doi.org/10.18653/v1/2022.acl-long.321>
12. Murtiyoso, M., Tahyudin, I., & Berlilana, B. (2025). A Systematic Review of Retrieval-Augmented Generation for Enhancing Domain-Specific Knowledge in Large Language Models. *Sinkron*, 9(2), 969–977. <https://doi.org/10.33395/sinkron.v9i2.14824>
13. Nouri, N. (2022). Text Style Transfer via Optimal Transport. *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2532–2541. <https://doi.org/10.18653/v1/2022.naacl-main.182>
14. Prawiro, I. Y., Yuliasri, I., Pratama, H., & Yuliati, Y. (2025). Integrating AI and Technology in Project-Based Learning for Academic Writing: Student Teachers' Experiences from Indonesia. *Educational Process International Journal*, 19(1). <https://doi.org/10.22521/edupij.2025.19.580>
15. Reif, E., Ippolito, D., Yuan, A., Coenen, A., Callison-Burch, C., & Wei, J. (2022). A Recipe for Arbitrary Text Style Transfer with Large Language Models. *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, 837–848. <https://doi.org/10.18653/v1/2022.acl-short.94>
16. Sari, Y., & Al Faridzi, F. P. (2023). Unsupervised Text Style Transfer for Authorship Obfuscation in Bahasa Indonesia. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, 17(1), 23. <https://doi.org/10.22146/ijccs.79623>
17. Syahnaz, M., & Fithriani, R. (2023). Utilizing Artificial Intelligence-based Paraphrasing Tool in EFL Writing Class: A Focus on Indonesian University Students' Perceptions. *Scope: Journal of English Language Teaching*, 7(2), 210. <https://doi.org/10.30998/scope.v7i2.14882>
18. Tresna, I. C., Usman, I., Rachmah, N., & Siswono, T. (2026). Leveraging AI-Based Paraphrasing Tools to Enhance Academic Writing Skills in Higher Education: A Comparative Experimental Study. *Online Learning*, 30(1), 78–113. <https://doi.org/10.24059/olj.v30i1.4856>
19. Wang, Y., Wu, Y., Mou, L., Li, Z., & Chao, W. (2020). Formality Style Transfer with Shared Latent Space. *Proceedings of the 28th International Conference on Computational Linguistics*, 2236–2249. <https://doi.org/10.18653/v1/2020.coling-main.203>
20. Wibowo, H. A., Prawiro, T. A., Ihsan, M., Aji, A. F., Prasajo, R. E., Mahendra, R., & Fitriany, S. (2020). Semi-Supervised Low-Resource Style Transfer of Indonesian Informal to Formal Language with Iterative Forward-Translation. *2020 International Conference on Asian Language Processing (IALP)*, 310–315. <https://doi.org/10.1109/IALP51396.2020.9310459>