

**PENGEMBANGAN SISTEM DETEKSI DAN TRANSFORMASI
KESANTUNAN TEKS *CHAT* MAHASISWA-DOSEN
MENGUNAKAN PENDEKATAN *FINE-TUNING*
INDOBERT DAN INDOT5**



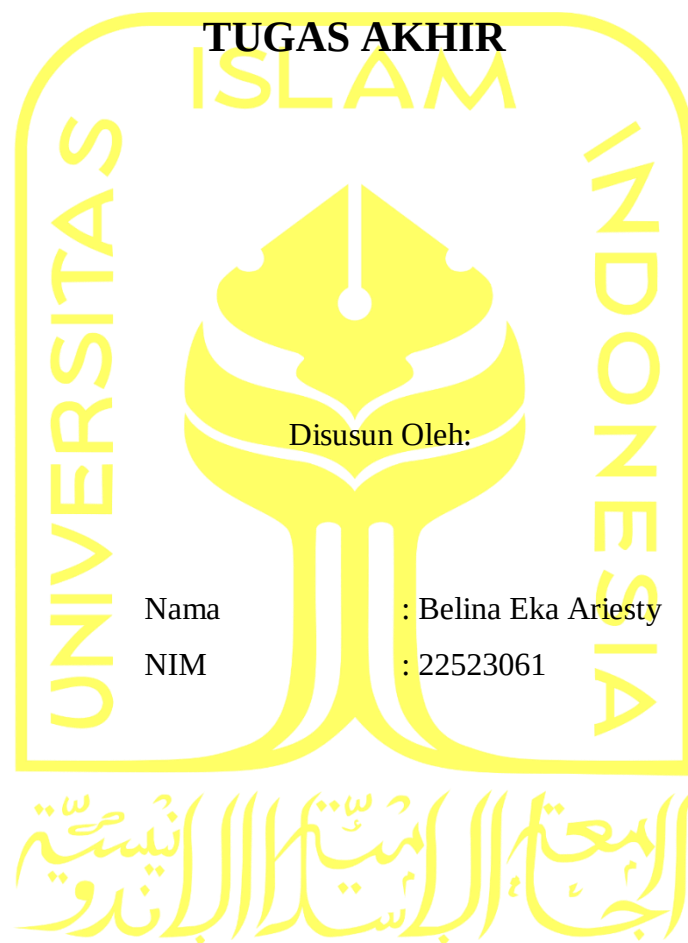
Disusun Oleh:

Nama : Belina Eka Ariesty
NIM : 22523061

**PROGRAM STUDI INFORMATIKA – PROGRAM SARJANA
FAKULTAS TEKNOLOGI INDUSTRI
UNIVERSITAS ISLAM INDONESIA
2026**

HALAMAN PENGESAHAN DOSEN PEMBIMBING

**PENGEMBANGAN SISTEM DETEKSI DAN TRANSFORMASI
KESANTUNAN TEKS *CHAT* MAHASISWA-DOSEN
MENGUNAKAN PENDEKATAN *FINE-TUNING*
INDOBERT DAN INDOT5**



Yogyakarta, 19 Mei 2026

Pembimbing,

(Chanifah Indah Ratnasari, S.Kom., M.Kom.)

HALAMAN PENGESAHAN DOSEN PENGUJI

**PENGEMBANGAN SISTEM DETEKSI DAN TRANSFORMASI
KESANTUNAN TEKS *CHAT* MAHASISWA-DOSEN
MENGUNAKAN PENDEKATAN *FINE-TUNING*
INDOBERT DAN INDOT5**

TUGAS AKHIR

Telah dipertahankan di depan sidang penguji sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer dari Program Studi Informatika – Program Sarjana di Fakultas Teknologi Industri Universitas Islam Indonesia

Yogyakarta, 19 Mei 2026

Tim Penguji

Chanifah Indah Ratnasari, S.Kom.,
M.Kom.

Anggota 1

Ahmad Fathan Hidayatullah, S.T.,
M.Cs., Ph.D.

Anggota 2

Ahmad Luthfi, Dr., S.Kom., M.Kom.

Mengetahui,

Ketua Program Studi Informatika – Program Sarjana

Fakultas Teknologi Industri

Universitas Islam Indonesia



(Dhomas Hatta Fudholi, S.T., M.Eng., Ph.D.)

HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR

Yang bertanda tangan di bawah ini:

Nama : Belina Eka Ariesty

NIM : 22523061

Tugas akhir dengan judul:

**PENGEMBANGAN SISTEM DETEKSI DAN TRANSFORMASI
KESANTUNAN TEKS *CHAT* MAHASISWA-DOSEN
MENGUNAKAN PENDEKATAN *FINE-TUNING*
INDOBERT DAN INDOT5**

Menyatakan bahwa seluruh komponen dan isi dalam tugas akhir ini adalah hasil karya saya sendiri. Apabila di kemudian hari terbukti ada beberapa bagian dari karya ini adalah bukan hasil karya sendiri, tugas akhir yang diajukan sebagai hasil karya sendiri ini siap ditarik kembali dan siap menanggung risiko dan konsekuensi apapun.

Demikian surat pernyataan ini dibuat, semoga dapat dipergunakan sebagaimana mestinya.

Yogyakarta, 28 April 2026



(Belina Eka Ariesty)

HALAMAN PERSEMBAHAN

Laporan tugas akhir ini saya dedikasikan untuk diri saya sendiri, yang telah berjuang dengan penuh kesabaran, ketekunan dan tidak menyerah hingga mencapai titik ini. Terima kasih kepada kedua orang tua tercinta dan adik saya yang selalu memberikan doa terbaik, dukungan, serta kasih sayang tiada henti, baik secara materiil maupun motivasi.

HALAMAN MOTO

“Maka sesungguhnya bersama kesulitan ada kemudahan. Sesungguhnya bersama kesulitan ada kemudahan.”

- Q.S Asy-Syarh 5-6

KATA PENGANTAR

Puji syukur penulis panjatkan kepada Allah Swt, atas segala rahmat dan karunia-Nya sehingga laporan tugas akhir ini dapat diselesaikan dengan baik. Laporan ini disusun sebagai salah satu syarat untuk menyelesaikan pendidikan sebagai bentuk penerapan ilmu yang telah diperoleh selama perkuliahan.

Dalam proses penyusunan laporan tugas akhir ini, penulis menyadari bahwa terdapat berbagai kendala nonilmiah yang dihadapi. Namun, berkat dukungan, doa, dan bantuan dari berbagai pihak, penulis dapat menyelesaikan laporan ini. Pada kesempatan ini, penulis ingin menyampaikan terima kasih kepada:

1. Kedua orang tua penulis, Mama dan Papa dan satu-satunya adik penulis yang selalu memberikan doa, dukungan, dan semangat untuk penulis.
2. Kedua nenek dan kakek penulis yang selalu menjadi bagian dari doa dan kenangan. Terima kasih sudah kebersamaan dan menyayangi penulis sejak kecil.
3. Keluarga besar dari pihak Papa dan Mama, terutama Om, kakak sepupu, dan bude yang telah memberikan dukungan dan menyayangi penulis sejak kecil hingga saat ini.
4. Prof. Fathul Wahid, S.T., M.SC., Ph.D., selaku rektor Universitas Islam Indonesia.
5. Bapak Dr. Raden Teduh Dirgahayu, S.T., M.S., selaku Ketua Jurusan Informatika Universitas Islam Indonesia.
6. Bapak Dhomas Hatta Fudholi, S.T., M.Eng., Ph.D., selaku Ketua Program Studi Informatika Program Sarjana Fakultas Teknologi Industri Universitas Islam Indonesia.
7. Ibu Chanifah Indah Ratnasari, S.Kom., M.Kom., selaku dosen pembimbing tugas akhir yang memberikan bimbingan dan masukan yang berharga.
8. Ibu Fani Eka Nurtjahjo, S.Psi., M.Psi., Psikolog yang telah membantu dalam proses pengerjaan skripsi ini, khususnya sebagai evaluator data penelitian penulis.
9. Bapak dan Ibu dosen Program Studi Informatika, yang telah berjasa dalam memberikan ilmu yang bermanfaat selama masa perkuliahan.
10. Seungcheol yang telah memberikan motivasi dan semangat selama proses penyusunan tugas akhir ini.
11. Sahabat penulis sejak kecil, Mayang, Puja, dan Firda yang selalu memberikan semangat, dan dukungan untuk penulis.

12. Sahabat penulis, HORE TEAM: Nisya, Sulthon, dan Ryan yang telah kebersamai penulis dari semester 3 hingga saat ini dan membantu penulis dalam penyusunan tugas akhir ini.
13. Teman Info Labris: Diena, Mifta, Nisya, Fath, Kasyiful, dan Raisha yang telah memberikan semangat dan saling memotivasi dalam penulisan tugas akhir ini.
14. Teman-teman penulis, Aghitsni, Anggun, Zahra, Bintang, Adit, Fafan, Aura, Widya, Kina, Novera, Umay, Fathya, Daffa Raihan, dan Alivah.
15. Teman-teman KKN Pakubalaho Ceriwi: Jia, Fira, Arfa, Vito, Fathur, Adil, Farras, dan Rahma, atas dukungan, kebersamaan, dan semangat yang telah diberikan selama proses penyusunan skripsi ini.
16. Seluruh pihak yang tidak dapat penulis sebutkan satu persatu, terima kasih atas semua dukungan baik secara langsung maupun tidak langsung dalam penyelesaian tugas akhir ini.

Penulis berharap laporan tugas akhir ini dapat memberikan manfaat dan menjadi referensi bagi pembaca di masa yang akan datang. Penulis menyadari bahwa laporan ini masih jauh dari kata sempurna, oleh karena itu penulis mengharapkan kritik dan saran yang membangun. Akhir kata, penulis mengucapkan terima kasih.

Yogyakarta, 28 April 2026



(Belina Eka Ariesty)

SARI

Penggunaan media komunikasi digital dalam lingkungan akademik antara mahasiswa dan dosen semakin meningkat, namun belum sepenuhnya diimbangi dengan penerapan kesantunan berbahasa. Hal ini berpotensi menimbulkan kesalahpahaman serta menurunkan kualitas komunikasi akademik. Selain itu, upaya dalam memperbaiki kesantunan masih dilakukan secara manual dan bergantung pada subjektivitas individu, sehingga belum terdapat standar yang konsisten. Penelitian ini bertujuan untuk membangun sebuah sistem berbasis *Natural Language Processing* yang mampu melakukan deteksi sekaligus transformasi kesantunan teks komunikasi akademik secara otomatis. Pendekatan yang digunakan adalah integrasi dua model, yaitu IndoBERT sebagai model klasifikasi dan IndoT5 yang digunakan untuk mengubah teks yang kurang santun menjadi lebih santun tanpa menghilangkan makna aslinya. *Dataset* penelitian ini terdiri dari 6.230 kalimat yang diperoleh dari Google Form, TikTok, dan ChatGPT, serta telah melalui proses pelabelan dan *preprocessing*. Evaluasi model klasifikasi dilakukan menggunakan *confusion matrix*, *accuracy*, *precision*, *recall*, dan *F1-score*, sedangkan model transformasi teks dievaluasi menggunakan *BLEU*, *ROUGE-1*, *ROUGE-2*, *ROUGE-L*, dan *METEOR*. Hasil penelitian menunjukkan bahwa IndoBERT mampu mencapai akurasi sebesar 97% dalam mengklasifikasikan tingkat kesantunan teks, sementara IndoT5 mampu menghasilkan transformasi kalimat yang lebih sesuai dengan kaidah komunikasi akademik. Sistem yang dikembangkan mampu mendeteksi dan memperbaiki kesantunan teks secara otomatis, serta diimplementasikan dalam bentuk aplikasi berbasis web.

Kata kunci: kesantunan teks, komunikasi akademik, NLP, IndoBERT, IndoT5

GLOSARIUM

<i>Accuracy</i>	Tingkat ketepatan model dalam memprediksi data dengan benar dibandingkan dengan seluruh data.
<i>BLEU</i>	Metrik evaluasi yang digunakan untuk mengukur tingkat kesamaan antara teks hasil generasi dengan teks referensi berdasarkan <i>n-gram</i> .
<i>Confusion Matrix</i>	Metode berbentuk matriks yang digunakan untuk mengukur kinerja model klasifikasi berdasarkan perbandingan antara hasil prediksi dan label sebenarnya.
<i>Decoder</i>	Bagian dari arsitektur transformer yang berfungsi untuk menghasilkan teks output berdasarkan representasi dari <i>encoder</i> .
<i>Deep Learning</i>	Metode pembelajaran mesin berbasis jaringan saraf tiruan berlapis yang mampu mempelajari pola kompleks dari data secara otomatis.
<i>Encoder</i>	Bagian dari arsitektur transformer yang berfungsi untuk mengubah teks input menjadi representasi vektor yang memahami konteks.
<i>F1-Score</i>	Metrik evaluasi yang merupakan rata-rata harmonik dari <i>precision</i> dan <i>recall</i> untuk mengukur keseimbangan performa model.
<i>Fine-Tuning</i>	Proses penyesuaian model yang telah dilatih sebelumnya (<i>pre-trained</i>) menggunakan <i>dataset</i> khusus agar sesuai dengan tugas tertentu.
IndoBERT	Model berbasis <i>transformer</i> yang dilatih khusus untuk Bahasa Indonesia dan digunakan untuk tugas klasifikasi teks.
IndoT5	Model berbasis <i>transformer</i> dengan pendekatan <i>text-to-text</i> yang digunakan untuk menghasilkan atau memperbaiki teks dalam Bahasa Indonesia.
<i>METEOR</i>	Metrik evaluasi yang digunakan untuk mengukur kesamaan makna antara teks hasil generasi dan teks referensi dengan mempertimbangkan sinonim dan struktur kalimat.
<i>Precision</i>	Metrik evaluasi yang menunjukkan tingkat ketepatan prediksi positif yang dihasilkan oleh model.
<i>Preprocessing</i>	Tahapan awal dalam pengolahan data teks yang bertujuan untuk membersihkan dan menyiapkan data sebelum digunakan dalam pemodelan.
<i>Recall</i>	Metrik evaluasi yang menunjukkan kemampuan model dalam menemukan seluruh data yang termasuk ke dalam kelas tertentu.

<i>ROUGE</i>	Metrik evaluasi yang digunakan untuk mengukur kesamaan antara teks hasil generasi dan teks referensi berdasarkan <i>recall</i> .
<i>Self-Attention</i>	Mekanisme dalam arsitektur transformer yang memungkinkan model memahami hubungan antar kata dalam satu kalimat secara kontekstual.
<i>Text-to-Text</i>	Pendekatan dalam pemrosesan bahasa alami yang mengubah suatu teks menjadi teks lain sebagai <i>output</i> dalam berbagai tugas NLP.
<i>Transformer</i>	Arsitektur <i>deep learning</i> yang menggunakan mekanisme <i>attention</i> untuk memahami hubungan antar kata dalam teks.

DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PENGESAHAN DOSEN PEMBIMBING.....	ii
HALAMAN PENGESAHAN DOSEN PENGUJI.....	iii
HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR.....	iv
HALAMAN PERSEMBAHAN	v
HALAMAN MOTO	vi
KATA PENGANTAR	vii
SARI	ix
GLOSARIUM.....	x
DAFTAR ISI	xii
DAFTAR TABEL.....	xiv
DAFTAR GAMBAR	xv
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	3
1.3 Batasan Masalah	4
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	5
1.6 Metodologi Penelitian	5
1.7 Sistematika Penulisan.....	6
BAB II KAJIAN PUSTAKA	8
2.1 Komunikasi Akademik Mahasiswa-Dosen	8
2.2 Kesantunan Berbahasa dalam Komunikasi Akademik	9
2.3 Komunikasi Digital	11
2.4 <i>Natural Language Processing</i> untuk Analisis Teks	12
2.5 <i>Deep Learning</i> dan <i>Transformer</i>	13
2.6 IndoBERT untuk Klasifikasi Teks.....	14
2.7 IndoT5 untuk Transformasi Teks	16
2.8 <i>Preprocessing</i> Teks.....	16
2.9 Evaluasi	17
2.9.1 Evaluasi Model Klasifikasi (IndoBERT)	17
2.9.2 Evaluasi Model Transformasi Teks (IndoT5).....	19
2.9.3 Evaluasi Manual.....	19
2.10 Penelitian Terdahulu	20
BAB III METODOLOGI PENELITIAN.....	24
3.1 Jenis dan Pendekatan Penelitian	24
3.2 Alur Penelitian	24
3.3 Pengumpulan Data	25
3.4 Pelabelan Data	27
3.5 <i>Preprocessing</i> Data	29
3.6 Pembangunan Model Klasifikasi Kesopanan.....	29
3.7 Pembangunan Model Transformasi Teks.....	31
3.8 Evaluasi Model	32
3.8.1 Evaluasi Model Klasifikasi.....	32
3.8.2 Evaluasi Model Transformasi.....	32
3.8.3 Evaluasi Manual.....	33
3.9 Deteksi Konteks Akademik.....	33

3.10	Integrasi Sistem.....	34
BAB IV HASIL DAN PEMBAHASAN		36
4.1	Pengumpulan Data	36
4.2	Pelabelan Data	38
4.3	<i>Preprocessing</i> Data	40
	4.3.1 Persiapan Data	40
	4.3.2 <i>Case Folding</i>	41
	4.3.3 Penghapusan URL.....	41
	4.3.4 Penghapusan <i>Mention</i>	42
	4.3.5 Penghapusan <i>Hashtag</i>	42
	4.3.6 Penghapusan Spasi Berlebih.....	43
	4.3.7 Penghapusan Karakter Khusus.....	43
	4.3.8 Penghapusan Kalimat yang Kurang dari Lima Karakter.....	44
4.4	Hasil Klasifikasi Kesopanan Menggunakan IndoBERT	45
	4.4.1 Data Split	45
	4.4.2 Proses Pelatihan Model IndoBERT	46
	4.4.3 Hasil Evaluasi Model IndoBERT	48
4.5	Hasil Transformasi Kesopanan Menggunakan IndoT5.....	52
	4.5.1 Implementasi IndoT5	52
	4.5.2 Evaluasi Model IndoT5	54
	4.5.3 Contoh Hasil Transformasi.....	55
4.6	Deteksi Konteks Akademik	56
	4.6.1 Implementasi Deteksi Konteks Akademik	57
	4.6.2 Output Sistem Deteksi Konteks	57
4.7	Integrasi Sistem.....	58
4.8	Evaluasi Manual	60
	4.8.1 Evaluasi Manual IndoBERT	60
	4.8.2 Evaluasi Manual IndoT5	62
	4.8.3 Analisis Keseluruhan Sistem	63
	4.8.4 Analisis Kasus Sulit	65
BAB V KESIMPULAN DAN SARAN		67
5.1	Kesimpulan.....	67
5.2	Saran.....	68
DAFTAR PUSTAKA		70
LAMPIRAN		75

DAFTAR TABEL

Tabel 2.1 Penelitian Terdahulu Klasifikasi Teks	20
Tabel 2.2 Penelitian Terdahulu Transformasi Teks	22
Tabel 3.1 Contoh Hasil Pelabelan.....	28
Tabel 3.2 Hasil Uji Coba Parameter IndoBERT.....	30
Tabel 4.1 Distribusi Sumber Data.....	36
Tabel 4.2 Distribusi Jenis Komunikasi	37
Tabel 4.3 Distribusi <i>Dataset</i> Berdasarkan Kategori Kesantunan	38
Tabel 4.4 Contoh <i>Dataset</i> Klasifikasi Kesopanan	39
Tabel 4.5 Contoh <i>Dataset</i> Pasangan Kalimat	39
Tabel 4.6 Hasil <i>Case Folding</i>	41
Tabel 4.7 Contoh Teks Hasil Penghapusan URL	42
Tabel 4.8 Contoh Perbandingan Teks Sebelum dan Sesudah Penghapusan <i>Mention</i>	42
Tabel 4.9 Contoh Teks Perbandingan Sebelum dan Sesudah Penghapusan <i>Hashtag</i>	43
Tabel 4.10 Contoh Teks Sebelum dan Sesudah Penghapusan Spasi Berlebih	43
Tabel 4.11 Contoh Teks Sebelum dan Sesudah Penghapusan Karakter Khusus	43
Tabel 4.12 Contoh Teks Sebelum dan Sesudah Penghapusan Kalimat Pendek.....	44
Tabel 4.13 Perbandingan <i>Dataset</i> Sesudah dan Sebelum <i>Preprocessing</i>	44
Tabel 4.14 Contoh Teks Sebelum dan Sesudah <i>Preprocessing</i>	45
Tabel 4.15 Distribusi <i>Dataset</i>	46
Tabel 4.16 Performa Model IndoBERT	50
Tabel 4.17 Contoh Hasil Prediksi Model IndoBERT	51
Tabel 4.18 Hasil Evaluasi Model IndoT5	54
Tabel 4.19 Contoh Hasil Tranformasi Teks IndoT5	56
Tabel 4.20 <i>Confusion Matrix</i> Evaluasi Manual	61
Tabel 4.21 Evaluasi Manual Model IndoT5.....	61
Tabel 4.22 Hasil Evaluasi Manual IndoT5.....	62

DAFTAR GAMBAR

Gambar 2.1 Arsitektur <i>Transformer</i>	14
Gambar 3.1 Alur Penelitian	25
Gambar 3.2 Alur Kerja Sistem	35
Gambar 4.1 Kode Program <i>Data Split</i>	46
Gambar 4.2 Alur Proses Pelatihan IndoBERT	47
Gambar 4.3 Kode Program Inisialisasi Model	47
Gambar 4.4 Kode Program Pengaturan <i>Training</i>	48
Gambar 4.5 <i>Confusion Matrix</i>	49
Gambar 4.6 Alur Transformasi Teks Pada Model IndoT5.....	52
Gambar 4.7 Kode Program Inisialisasi IndoT5	53
Gambar 4.8 Kode Program <i>Training</i> IndoT5	53
Gambar 4.9 Grafik Hasil Evaluasi IndoT5.....	54
Gambar 4.10 Antarmuka Sistem Deteksi dan Transformasi Kesantunan Komunikasi Akademik Berbasis Web	58

BAB I

PENDAHULUAN

1.1 Latar Belakang

Komunikasi akademik merupakan sebuah proses interaksi antara mahasiswa dan dosen dalam lingkungan pendidikan perguruan tinggi yang tidak hanya berfungsi untuk menyampaikan informasi, tetapi juga menekankan etika dan hubungan profesional dalam mendukung proses pembelajaran (Taslim et al., 2024). Kesantunan berbahasa merupakan salah satu aspek penting dalam komunikasi yang efektif dalam kajian pragmatik, yang merujuk pada penggunaan bahasa yang memperhatikan norma serta rasa hormat dalam interaksi sosial (Anugrawati & Syam, 2024).

Mahasiswa di Indonesia memiliki latar belakang budaya, sosial, dan ekonomi yang beragam, yang secara tidak langsung mempengaruhi pola komunikasi di lingkungan kampus. Perbedaan dialek dan persepsi dapat menjadi faktor penghambat terciptanya komunikasi yang efektif (Kurniadini & Putranto, 2025). Seiring dengan perkembangan teknologi digital, komunikasi akademik mengalami perubahan yang signifikan. Interaksi yang sebelumnya dilakukan secara tatap muka atau melalui surat formal kini banyak beralih ke media digital seperti WhatsApp, Telegram, ataupun email (Puspa et al., 2024). Media tersebut digunakan mahasiswa untuk berbagai keperluan akademik, seperti menyampaikan izin tidak menghadiri kelas, menanyakan tugas kuliah, hingga konsultasi pengerjaan skripsi.

Kemudahan komunikasi digital tersebut juga menimbulkan tantangan, terutama terkait perbedaan pemahaman etika dan kesantunan berbahasa. Hal ini dapat memicu kesalahpahaman apabila tidak memperhatikan konteks, pilihan kata, serta struktur bahasa yang sesuai (Mahas et al., 2025). Selain itu, masih terdapat keterbatasan pemahaman mahasiswa terhadap etika komunikasi digital. Reynaldi et al. (2024) menyatakan bahwa mahasiswa sering kali tidak memperhatikan kesantunan serta waktu dalam berkomunikasi dengan dosen. Selain itu, penggunaan bahasa yang tidak baku, bertele-tele, serta bahasa pergaulan yang populer masih sering ditemukan dalam pesan akademik. Kurangnya penerapan etika berdigital ini dapat menjadi permasalahan dalam penyampaian pesan melalui media komunikasi daring (Abidin & Wandu, 2023).

Fenomena tersebut juga didukung oleh data penelitian di IKIP Siliwangi yang menunjukkan bahwa sebanyak 13% mahasiswa tergolong kurang santun dan 21% mahasiswa tergolong tidak santun dalam berkomunikasi, termasuk melalui media digital (Aeni et al., 2025). Hal ini menunjukkan bahwa sebagian mahasiswa belum sepenuhnya menerapkan etika berbahasa yang baik dalam lingkungan akademik. Selain itu, aspek kesantunan berbahasa dalam komunikasi antara mahasiswa dan dosen masih menunjukkan adanya variasi dalam penerapannya. Beberapa bentuk yang sering ditemukan antara lain penggunaan kalimat yang terlalu singkat, penggunaan emoji seperti jempol, penggunaan bahasa tidak baku, serta penyampaian pesan tanpa memperkenalkan diri, tanpa salam atau tanpa permohonan maaf terlebih dahulu (Eksan et al., 2021).

Permasalahan tersebut menunjukkan bahwa masih diperlukan upaya untuk membantu mahasiswa dalam menyusun pesan akademik yang lebih santun dan sesuai dengan kaidah komunikasi akademik. Selama ini, proses perbaikan masih dilakukan secara manual berdasarkan pengalaman pribadi atau arahan dari dosen sehingga standar kesantunan yang digunakan sering kali berbeda pada setiap individu (Astiti & Raharja, 2023). Kondisi ini membuka peluang pemanfaatan teknologi untuk membantu proses evaluasi bahasa secara otomatis.

Perkembangan teknologi di bidang *Natural Language Processing* memungkinkan sistem komputer untuk memahami, menganalisis, dan memproses bahasa manusia. Dalam beberapa tahun terakhir, metode berbasis *deep learning* dan model *transformer* mulai banyak digunakan untuk berbagai tugas pengolahan bahasa, seperti klasifikasi teks dan transformasi gaya bahasa. Salah satu model yang dapat digunakan adalah IndoBERT, yang mampu memahami konteks dalam bahasa Indonesia dan sering digunakan untuk mengklasifikasikan teks. Selain itu, model IndoT5 dapat digunakan untuk melakukan pengubahan atau transformasi teks sehingga menjadi lebih formal dan sesuai dengan konteks komunikasi akademik.

Meskipun saat ini telah tersedia teknologi kecerdasan buatan generatif seperti ChatGPT yang mampu membantu memperbaiki atau menghasilkan teks yang lebih formal, penggunaannya dalam konteks komunikasi akademik masih memiliki beberapa keterbatasan. ChatGPT merupakan model *general-purpose* yang tidak dirancang secara khusus untuk memahami standar kesantunan komunikasi akademik mahasiswa-dosen di Indonesia. Hasil transformasi yang diberikan dapat bersifat tidak konsisten karena sangat bergantung pada *prompt* yang diberikan pengguna serta belum memiliki mekanisme klasifikasi tingkat

kesantunan yang terstruktur. Selain itu, belum terdapat jaminan bahwa hasil transformasi benar-benar sesuai dengan konteks akademik maupun mempertahankan makna asli pesan secara optimal. Oleh karena itu, diperlukan sistem yang lebih terarah agar proses deteksi dan transformasi kesantunan bahasa dapat dilakukan secara lebih konsisten dan sesuai dengan konteks komunikasi akademik.

Selain itu, penelitian-penelitian sebelumnya umumnya masih memisahkan proses deteksi kesantunan dan transformasi bahasa menjadi dua bagian yang terpisah. Penelitian ini berfokus pada pengembangan sistem berbasis IndoBERT dan IndoT5 untuk mendeteksi serta meningkatkan kesantunan teks komunikasi akademik secara otomatis. Meskipun penelitian ini menggunakan istilah komunikasi mahasiswa-dosen, sistem yang dikembangkan difokuskan pada pemrosesan teks yang dikirim oleh mahasiswa kepada dosen sebagai bentuk komunikasi satu arah. Sistem yang dikembangkan diharapkan tidak hanya berfungsi sebagai alat bantu, tetapi juga sebagai media edukasi bagi mahasiswa dalam memahami dan menerapkan kesantunan berbahasa tulis dalam komunikasi akademik. Dengan demikian, mahasiswa dapat menyusun pesan akademik yang lebih tepat, santun, dan sesuai dengan kaidah komunikasi di lingkungan akademik.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dirumuskan sebelumnya, terdapat beberapa permasalahan utama yang menjadi fokus pada penelitian ini sebagai berikut:

- a. Bagaimana membangun model klasifikasi menggunakan metode *fine-tuning* IndoBERT yang mampu mengidentifikasi tingkat kesantunan teks *chat* mahasiswa-dosen yang dioperasikan ke dalam dua kategori, yaitu “sopan” dan “tidak sopan”?
- b. Bagaimana membangun model transformasi kalimat menggunakan metode *fine-tuning* IndoT5 yang mampu mengubah gaya bahasa teks yang teridentifikasi “tidak sopan” menjadi lebih santun tanpa mengubah makna dari kalimat tersebut?

1.3 Batasan Masalah

Batasan penelitian ditetapkan untuk membatasi ruang lingkup pembahasan agar tetap terarah sehingga tujuan dari penelitian tercapai. Dalam pelaksanaannya, lingkup penelitian dibatasi pada:

- a. Penelitian ini difokuskan pada konteks komunikasi akademik dari mahasiswa-dosen dengan menggunakan bahasa Indonesia. Ruang lingkup penelitian ini mencakup identifikasi tingkat kesantunan teks *chat* akademik, penggunaan kata kasar seperti makian tidak dijadikan indikator utama dalam proses klasifikasi.
- b. Data yang digunakan dalam penelitian ini berupa teks berbahasa Indonesia yang diperoleh dari hasil survei serta unggahan TikTok yang menampilkan ulang tangkapan layar percakapan dari platform X dan Facebook, serta data tambahan yang dihasilkan melalui ChatGPT, yang telah melalui proses kurasi manual untuk menyesuaikan dengan konteks komunikasi akademik.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut:

- a. Mengembangkan sistem berbasis *Natural Language Processing* untuk meningkatkan kesantunan teks komunikasi akademik mahasiswa-dosen.
- b. Menghasilkan *dataset* teks *chat* akademik berbahasa Indonesia yang representatif dan sesuai dengan konteks komunikasi akademik sebagai dasar dalam pengembangan model.
- c. Membangun dua model berbasis *fine-tuning*, yaitu model klasifikasi menggunakan IndoBERT untuk mengidentifikasi teks *chat* ke dalam kategori sopan dan tidak sopan sebagai representasi tingkat kesantunan, serta model transformasi teks menggunakan IndoT5 untuk mengubah kalimat tidak sopan menjadi lebih santun tanpa menghilangkan makna aslinya.
- d. Menilai kinerja model klasifikasi dan model transformasi teks dalam mendeteksi serta meningkatkan kesantunan teks komunikasi akademik secara efektif.

1.5 Manfaat Penelitian

Manfaat dari hasil penelitian ini diharapkan dapat berguna sebagai berikut:

- a. Memberikan kontribusi dalam menjaga budaya kesantunan berbahasa dalam komunikasi digital akademik di Indonesia, khususnya di lingkungan perguruan tinggi.
- b. Menjadi referensi dalam pengembangan data dan sistem pembelajaran bahasa yang santun dengan memanfaatkan pesan nyata mahasiswa melalui media digital yang telah dikurasi.
- c. Sebagai media edukasi dan alat bantu bagi mahasiswa dalam menyusun teks pesan yang lebih santun kepada dosen melalui sistem yang dikembangkan dalam penelitian ini.

1.6 Metodologi Penelitian

Tahapan penelitian yang dilakukan dalam pengembangan sistem ini secara umum meliputi beberapa langkah utama sebagai berikut:

- a. Pengumpulan Data
Data yang digunakan dalam penelitian ini berupa teks *chat* komunikasi akademik antara mahasiswa dan dosen dalam bahasa Indonesia. Data diperoleh dari beberapa sumber, yaitu Google Form, platform TikTok yang memuat ulang percakapan dari media sosial X dan Facebook, serta data tambahan yang dihasilkan menggunakan ChatGPT sebagai pelengkap variasi data.
- b. Pelabelan Data
Pelabelan data dilakukan secara manual ke dalam dua kategori, yaitu sopan dan tidak sopan. Proses ini bertujuan untuk memastikan kesesuaian label dengan konteks komunikasi akademik berdasarkan indikator kesantunan berbahasa.
- c. Persiapan dan *Preprocessing* Data
Tahapan ini dilakukan untuk menyiapkan data agar siap digunakan dalam proses pemodelan. Proses yang dilakukan meliputi pembersihan data, normalisasi teks, serta penghapusan elemen yang tidak relevan sehingga kualitas dataset tetap terjaga.
- d. Pembangunan Model
Penelitian ini menggunakan dua model utama, yaitu IndoBERT untuk melakukan klasifikasi tingkat kesantunan teks, serta IndoT5 untuk melakukan transformasi teks menjadi lebih santun. Model dikembangkan menggunakan pendekatan *fine-tuning* terhadap data yang telah diproses.

- e. **Evaluasi Model**
Evaluasi dilakukan untuk mengukur kinerja model klasifikasi dan transformasi teks menggunakan metrik evaluasi yang sesuai, sehingga dapat diketahui tingkat efektivitas sistem yang dikembangkan.
- f. **Integrasi Sistem Berbasis Web**
Tahap akhir penelitian adalah mengintegrasikan model yang telah dibangun ke dalam sebuah sistem berbasis web, sehingga pengguna dapat melakukan proses deteksi dan transformasi teks secara langsung

1.7 Sistematika Penulisan

Sistematika penulisan dalam penyusunan laporan tugas akhir ini terdiri dari lima bab, yaitu sebagai berikut:

a. BAB I PENDAHULUAN

Bab ini memuat latar belakang, rumusan masalah, batasan penelitian, tujuan penelitian, manfaat penelitian, metodologi penelitian, serta sistematika penulisan. Seluruh pembahasan pada bab ini berfokus pada permasalahan kesantunan teks *chat* mahasiswa dalam komunikasi akademik dengan dosen.

b. BAB II KAJIAN PUSTAKA

Bab ini berisi kajian pustaka yang meliputi tinjauan penelitian terdahulu yang relevan serta landasan teori yang mendukung penelitian terkait kesantunan teks *chat* mahasiswa dalam komunikasi akademik dengan dosen.

c. BAB III METODOLOGI

Bab ini menjelaskan tahapan penelitian yang dilakukan, meliputi pengumpulan data, pelabelan data, proses *preprocessing*, perancangan, dan pembangunan model menggunakan IndoBERT dan IndoT5, serta metode evaluasi yang digunakan dalam penelitian.

d. BAB IV HASIL DAN PEMBAHASAN

Bab ini menyajikan hasil dari setiap tahapan metodologi penelitian, meliputi pengumpulan data, pelabelan, *preprocessing*, pembangunan model, evaluasi, serta integrasi sistem berbasis web. Selain itu, disajikan pula hasil pengujian model dan analisis terhadap performa sistem dalam mengubah teks *chat* menjadi lebih santun dalam konteks akademik.

e. **BAB V KESIMPULAN DAN SARAN**

Bab ini berisi kesimpulan yang diperoleh dari hasil penelitian serta saran yang dapat digunakan untuk pengembangan penelitian selanjutnya.

BAB II

KAJIAN PUSTAKA

2.1 Komunikasi Akademik Mahasiswa-Dosen

Komunikasi di lingkungan akademik, terutama yang melibatkan mahasiswa dan dosen, memiliki peran penting dalam menciptakan suasana belajar yang kondusif. Komunikasi yang efektif tidak hanya berfungsi sebagai sarana penyampaian informasi, tetapi juga berkontribusi dalam membangun hubungan yang baik serta meningkatkan kepercayaan antar pihak yang terlibat (Sahabuddin et al., 2024). Penerapan komunikasi yang baik dapat membangun hubungan profesional dan sikap saling menghormati antara mahasiswa dan dosen (Poneni, 2023). Hubungan antara mahasiswa dan dosen memiliki perbedaan peran yang bersifat hierarkis, di mana dosen berperan sebagai pembimbing akademik dan mahasiswa sebagai pihak yang dibimbing. Perbedaan peran ini mempengaruhi cara berkomunikasi, terutama dalam pemilihan bahasa, struktur penyampaian pesan, serta kejelasan maksud. Oleh karena itu, komunikasi akademik menuntut penggunaan bahasa yang sesuai dengan konteks formal dan norma yang berlaku di lingkungan pendidikan tinggi.

Seiring dengan perkembangan teknologi digital, komunikasi akademik juga mengalami pergeseran ke media digital seperti aplikasi pesan instan dan email (Holidin et al., 2025). Meskipun memberikan kemudahan dalam berkomunikasi, penggunaan media digital juga menimbulkan tantangan dalam menjaga kualitas komunikasi, terutama karena keterbatasan konteks nonverbal dalam pesan berbasis teks. Dalam praktiknya, pesan yang disampaikan sering kali belum memenuhi kaidah komunikasi akademik yang baik, baik dari segi struktur maupun penggunaan bahasa (Eksan et al., 2021). Oleh karena itu, komunikasi akademik tidak hanya bergantung pada isi pesan, tetapi juga pada bagaimana pesan tersebut disusun dan disampaikan. Aspek seperti pemilihan kata, penggunaan bahasa baku, struktur pesan, serta kesesuaian dengan konteks relasi antara mahasiswa dan dosen menjadi faktor penting dalam menentukan kualitas komunikasi. Hal ini menunjukkan bahwa kesantunan berbahasa merupakan elemen yang tidak terpisahkan dalam komunikasi akademik. Dalam kajian ini, komunikasi akademik dipahami sebagai bentuk interaksi yang memiliki norma dan etika tertentu, sehingga dapat dianalisis secara sistematis.

2.2 Kesantunan Berbahasa dalam Komunikasi Akademik

Kesantunan berbahasa merupakan penggunaan bahasa yang memperhatikan norma sosial dan hubungan interpersonal untuk menjaga serta keharmonisan komunikasi (Purnomo, 2025). Dalam konteks komunikasi akademik, kesantunan berbahasa menjadi aspek penting karena adanya perbedaan peran antara mahasiswa dan dosen yang menuntut penggunaan bahasa yang sesuai dengan norma formal. Kesantunan dalam komunikasi akademik tidak hanya ditentukan oleh pilihan kata, tetapi juga oleh struktur penyampaian pesan serta kesesuaian dengan konteks komunikasi.

Kesantunan berbahasa tidak bersifat universal, melainkan dipengaruhi oleh latar belakang budaya, kebiasaan sosial, serta konteks komunikasi penutur. Dalam masyarakat multikultural seperti Indonesia, bentuk komunikasi yang dianggap santun pada suatu kelompok budaya belum tentu dipersepsikan sama oleh kelompok budaya lainnya. Penelitian mengenai pragmatik budaya menunjukkan bahwa kesantunan berkaitan erat dengan nilai sosial, hierarki, identitas budaya, serta hubungan antara penutur dan lawan tutur (Purnomo, 2025). Aspek tersebut berkaitan erat dengan kajian pragmatik dalam linguistik.

Dalam kajian linguistik, kesantunan berbahasa lebih dekat dengan pendekatan pragmatik dibandingkan semantik. Semantik berfokus pada makna literal suatu kata atau kalimat, sedangkan pragmatik memperhatikan makna berdasarkan konteks penggunaan bahasa, hubungan antara penutur dan lawan tutur, serta situasi komunikasi yang melatarbelakanginya. Dalam komunikasi mahasiswa-dosen, suatu tuturan dapat memiliki makna yang berbeda bergantung pada konteks sosial, tingkat formalitas, dan cara penyampaian pesan. Oleh karena itu, penilaian kesantunan dalam penelitian ini tidak hanya mempertimbangkan makna literal teks, tetapi juga memperhatikan konteks komunikasi akademik dan relasi antara mahasiswa dan dosen sebagai bagian dari aspek pragmatik bahasa.

Selain itu, perbedaan budaya dan latar belakang daerah juga dapat memengaruhi gaya komunikasi mahasiswa dalam berinteraksi dengan dosen maupun civitas akademika lainnya. Penelitian sebelumnya menunjukkan bahwa mahasiswa dari daerah yang berbeda membawa norma komunikasi dan kebiasaan berbahasa yang berbeda pula, sehingga dapat memengaruhi cara menyampaikan pendapat, penggunaan sapaan, tingkat formalitas bahasa, maupun ekspresi kesantunan dalam komunikasi akademik (Taslim et al., 2024). Perbedaan gaya komunikasi tersebut berpotensi menimbulkan perbedaan persepsi terhadap tingkat kesantunan apabila tidak dipahami sesuai konteks budaya masing-masing.

Selain itu, penggunaan unsur bahasa dan gaya komunikasi tertentu dalam lingkungan akademik juga dapat menjadi penanda penghormatan, identitas budaya, serta upaya menjaga keharmonisan komunikasi. Oleh karena itu, penelitian ini tidak berfokus pada standar kesantunan dari satu wilayah atau budaya tertentu saja, melainkan menggunakan indikator kesantunan umum dalam komunikasi akademik formal, seperti penggunaan salam pembuka, identitas diri, penyampaian maksud yang jelas, serta penggunaan bahasa yang sopan dan menghormati lawan tutur. Variasi sumber data yang digunakan dalam penelitian ini diharapkan dapat merepresentasikan beragam gaya komunikasi mahasiswa dengan latar belakang budaya dan kebiasaan berbahasa yang berbeda. Dengan demikian, dataset yang digunakan tidak hanya merepresentasikan pola komunikasi dari satu wilayah tertentu, tetapi juga mencerminkan keragaman komunikasi akademik mahasiswa dalam konteks bahasa Indonesia secara umum.

Bentuk kesantunan dapat ditunjukkan melalui penggunaan bahasa baku, penyampaian maksud secara jelas, penggunaan salam pembuka, serta sikap menghargai lawan tutur (Eksan et al., 2021). Sebaliknya, penggunaan bahasa informal, pesan yang terlalu singkat tanpa konteks, atau penggunaan bahasa campuran dan bahasa gaul dapat dianggap kurang sopan dalam konteks akademik (Reynaldi et al., 2024). Konsep tersebut juga dijelaskan dalam teori *politeness* dalam kajian pragmatik. Dalam kajian pragmatik, konsep kesantunan berbahasa berkaitan dengan upaya penutur dalam menjaga hubungan sosial dan menghormati lawan tutur melalui penggunaan bahasa yang sesuai dengan norma komunikasi. Brown dan Levinson (1987) menjelaskan bahwa kesantunan (*politeness*) berfungsi untuk meminimalkan potensi ancaman terhadap muka (*face*) lawan tutur dalam interaksi sosial. Dalam konteks komunikasi akademik mahasiswa-dosen, penggunaan bahasa yang santun menjadi penting karena terdapat relasi sosial dan hierarki akademik yang perlu diperhatikan (Mahas et al., 2025).

Secara operasional, pesan akademik yang santun dalam konteks komunikasi mahasiswa-dosen dapat diidentifikasi melalui beberapa indikator, yaitu: (1) adanya salam pembuka, (2) pengenalan identitas diri, (3) penyampaian maksud yang jelas dan runtut, serta (4) penggunaan bahasa yang baku dan sopan. Sebaliknya, pesan yang tidak santun cenderung tidak memiliki struktur yang jelas, menggunakan bahasa informal, serta tidak mempertimbangkan konteks relasi antara mahasiswa dan dosen. Indikator tersebut menjadi dasar dalam penelitian ini untuk mengklasifikasikan teks *chat* mahasiswa-dosen ke dalam kategori sopan dan tidak sopan.

Dalam penelitian ini, istilah “kesantunan” digunakan sebagai konsep teoritis yang merujuk pada kajian pragmatik mengenai penggunaan bahasa dalam konteks sosial. Sementara itu, istilah “sopan” dan “tidak sopan” digunakan sebagai representasi operasional dalam proses

pelabelan data untuk mempermudah proses klasifikasi teks. Penggunaan label *biner* tersebut dilakukan agar proses anotasi memiliki batasan yang lebih jelas dan konsisten dalam membedakan teks yang sesuai maupun tidak sesuai dengan norma komunikasi akademik. Dengan demikian, label “sopan” dan “tidak sopan” dalam penelitian ini tidak dimaksudkan sebagai penilaian absolut terhadap individu, melainkan sebagai kategori klasifikasi berdasarkan indikator kesantunan berbahasa yang telah ditentukan.

Dalam penelitian ini, pelabelan data dilakukan secara manual ke dalam dua kelas utama, yaitu sopan (0) dan tidak sopan (1). Penggunaan label numerik 0 dan 1 dilakukan sebagai representasi kategorikal dalam proses klasifikasi *machine learning* untuk mempermudah proses pelatihan dan evaluasi model. Penggunaan representasi numerik pada label kelas merupakan pendekatan yang umum digunakan dalam penelitian klasifikasi teks *berbasis Natural Language Processing (NLP)* (Kono et al., 2025). Dengan demikian, label numerik tersebut tidak memiliki makna linguistik tertentu, melainkan berfungsi sebagai penanda kelas pada data. Dengan demikian, label numerik tersebut tidak memiliki makna linguistik tertentu, melainkan berfungsi sebagai penanda kelas pada data dan digunakan secara konsisten pada seluruh tahapan penelitian.

2.3 Komunikasi Digital

Perkembangan teknologi informasi telah membawa perubahan signifikan dalam pola komunikasi di lingkungan pendidikan tinggi (Taqwa et al., 2025). Interaksi akademik yang sebelumnya dilakukan secara tatap muka kini banyak berlangsung melalui media komunikasi digital. Penggunaan media digital memberikan kemudahan dalam penyampaian informasi karena lebih fleksibel dari segi waktu dan tempat (Puspa et al., 2024). Namun, komunikasi berbasis teks juga menghadirkan tantangan terkait etika dan kesantunan berbahasa, karena keterbatasan konteks nonverbal dapat meningkatkan risiko kesalahpahaman. Beberapa mahasiswa masih menggunakan gaya komunikasi informal yang kurang sesuai dengan konteks akademik (Reynaldi et al., 2024). Oleh karena itu, pemanfaatan teknologi untuk membantu evaluasi dan transformasi komunikasi digital menjadi penting guna mendukung interaksi yang efektif antara mahasiswa dan dosen.

Dalam konteks penelitian ini, komunikasi digital difokuskan pada teks *chat* akademik antara mahasiswa-dosen. Berbeda dengan surat formal, teks *chat* cenderung bersifat lebih singkat, menggunakan bahasa yang lebih informal, serta terkadang bersifat multitafsir (Ratnasari & Fitriyanti, 2024). Karakteristik ini dapat menyebabkan pesan menjadi kurang jelas

atau tidak sesuai dengan norma komunikasi akademik. Selain itu, keterbatasan ruang dan kebiasaan komunikasi yang cepat pada media *chat* sering kali membuat pesan tidak disusun secara lengkap, seperti tidak adanya salam pembuka, pengenalan diri, maupun penyampaian maksud yang runtut. Kondisi ini menunjukkan bahwa teks *chat* akademik memiliki kompleksitas tersendiri yang berbeda dari teks formal, sehingga memerlukan perhatian khusus dalam analisis bahasa. Oleh karena itu, diperlukan suatu pendekatan yang mampu menganalisis dan mengolah teks secara otomatis berdasarkan konteks bahasa, seperti *Natural Language Processing* (NLP), untuk membantu mengidentifikasi serta meningkatkan kualitas komunikasi akademik berbasis teks.

2.4 *Natural Language Processing* untuk Analisis Teks

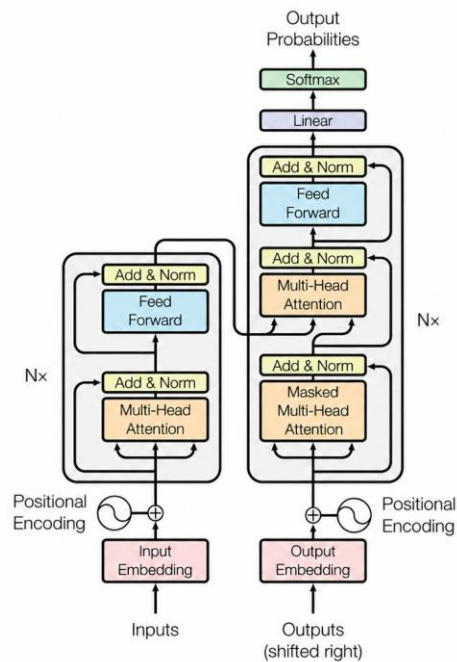
Natural Language Processing (NLP) merupakan salah satu bidang dalam ilmu komputer dan kecerdasan buatan yang berfokus pada interaksi antara komputer dan bahasa alami manusia. Bahasa alami merujuk pada bahasa yang digunakan dan dipahami manusia dalam komunikasi sehari-hari yang berfungsi sebagai media penyampaian informasi dari satu individu kepada individu lainnya, sehingga NLP dikembangkan untuk memungkinkan komputer memahami serta memproses informasi secara otomatis (Prasetyo et al., 2021). Dalam perkembangannya, NLP banyak digunakan untuk berbagai tugas pengolahan teks, seperti klasifikasi teks, analisis sentimen, penerjemahan bahasa, serta transformasi teks. Berbagai penelitian menunjukkan bahwa penerapan NLP juga menghadapi sejumlah tantangan dalam pemrosesan bahasa alami, seperti kesalahan ejaan, variasi gaya penulisan, serta kemampuan model dalam menangkap sentimen subjektif secara lebih akurat (Nurwanda et al., 2024).

Tantangan ini menjadi relevan dalam konteks komunikasi akademik berbasis teks *chat* yang cenderung singkat, informal, dan berpotensi multitafsir. Dalam penelitian ini, NLP dimanfaatkan untuk menganalisis tingkat kesantunan dalam teks *chat* mahasiswa-dosen melalui proses klasifikasi, serta melakukan transformasi teks untuk menghasilkan kalimat yang lebih sesuai dengan norma komunikasi akademik. Dengan demikian, NLP tidak hanya digunakan untuk memahami teks, tetapi juga untuk mendukung peningkatan kualitas komunikasi akademik secara otomatis. Oleh karena itu, pendekatan NLP dipilih dalam penelitian ini karena kemampuannya dalam memproses, memahami, dan mentransformasikan teks secara kontekstual sesuai dengan kebutuhan analisis kesantunan berbahasa.

2.5 Deep Learning dan Transformer

Deep Learning merupakan salah satu cabang kecerdasan buatan yang meniru cara manusia memperoleh jenis pengetahuan dengan mempelajari pengetahuan melalui data (Amriza & Supriyadi, 2021). Konsep ini diperkenalkan oleh Geoffrey Hinton pada tahun 2006 sebagai pengembangan dari metode *machine learning* konvensional, terutama dalam mengatasi keterbatasan *feature engineering*, karena mampu mempelajari representasi fitur secara otomatis dari data. Pendekatan ini menggunakan jaringan saraf tiruan berlapis untuk memahami pola data yang kompleks. Dalam konteks pengolahan bahasa alami, *deep learning* memiliki keunggulan dalam menangani berbagai jenis data serta memahami hubungan kontekstual dalam teks. Hal ini menjadi penting dalam analisis teks *chat* akademik yang memiliki karakteristik singkat, informal, dan berpotensi multitafsir. Oleh karena itu, *deep learning* banyak digunakan dalam berbagai tugas *Natural Language Processing* termasuk klasifikasi teks dan transformasi bahasa (Rachman & Santoso, 2021).

Salah satu arsitektur *deep learning* yang banyak digunakan dalam pemrosesan bahasa alami adalah *transformer*. Arsitektur ini memanfaatkan mekanisme *attention* untuk memahami hubungan antar kata dalam suatu kalimat secara kontekstual. Berbeda dengan model sekuensial seperti *Recurrent Neural Network* (RNN) dan *Long Short-Term Memory* (LSTM), *transformer* mampu memproses data secara paralel sehingga lebih efisien serta lebih efektif dalam menangkap konteks bahasa (Tan et al., 2022). Keunggulan *transformer* terletak pada kemampuannya dalam memahami konteks global dalam teks, termasuk hubungan antar kata yang memiliki jarak jauh (*long-distance dependency*). Kemampuan tersebut menjadikan arsitektur *transformer* sangat relevan dalam penelitian ini, khususnya dalam mendukung model IndoBERT untuk klasifikasi kesantunan dan IndoT5 untuk transformasi teks dalam konteks komunikasi akademik.

Gambar 2.1 Arsitektur *Transformer*

Sumber: (Vaswani et al., 2017)

Gambar 2.1, menunjukkan bahwa model *transformer* terdiri dari dua komponen utama, yaitu *encoder* dan *decoder*. Komponen utama *transformer* adalah *attention mechanism*, khususnya *self-attention*, yang memungkinkan model memahami hubungan antar kata dalam satu kalimat secara menyeluruh. *Encoder* berfungsi untuk membangun representasi konteks dari teks masukan, sedangkan *decoder* menghasilkan keluaran teks berdasarkan representasi tersebut. Mekanisme ini membuat *transformer* lebih efektif dalam menangkap ketergantungan jarak jauh (*long-distance dependency*) dibandingkan model sekuensial sebelumnya. Dalam penelitian ini, *deep learning* dengan *arsitektur transformer* digunakan karena kemampuannya dalam memahami konteks bahasa secara mendalam dan mempelajari pola secara otomatis, sehingga lebih efektif dalam melakukan klasifikasi tingkat kesantunan serta transformasi teks dibandingkan pendekatan tradisional.

2.6 IndoBERT untuk Klasifikasi Teks

IndoBERT merupakan model bahasa berbasis *deep learning* yang dikembangkan dari arsitektur *Bidirectional Encoder Representations from Transformer* (BERT) yang diperkenalkan oleh Google pada tahun 2018. Model BERT dirancang untuk memahami bahasa alami dengan memanfaatkan konteks dua arah dalam satu kalimat. Pendekatan *bidirectional*

memungkinkan model untuk mempertimbangkan hubungan kata sebelum dan sesudah token sehingga mampu menangkap makna kalimat secara lebih komprehensif. Dalam proses pelatihannya, BERT melalui dua tahapan yaitu, *pre-training* dan *fine-tuning*. Pada tahap *pre-training*, model mempelajari representasi bahasa secara umum menggunakan *Masked Language Modeling* (MLM), yaitu dengan menyembunyikan beberapa kata dalam melatih model untuk memprediksi kata yang hilang berdasarkan konteksnya (Ramdan et al., 2025).

IndoBERT merupakan adaptasi dari BERT yang dilatih menggunakan korpus teks berbahasa Indonesia, sehingga mampu memahami karakteristik bahasa Indonesia dengan lebih baik. Model IndoBERT dipilih karena model berbasis bahasa Indonesia yang telah dilatih untuk menyelesaikan berbagai tugas pemrosesan bahasa alami, seperti klasifikasi teks, analisis sentimen, dan pemahaman konteks bahasa. Keunggulan IndoBERT terletak pada kemampuannya dalam memahami konteks bahasa secara dua arah, sehingga efektif dalam mengidentifikasi makna implisit dalam teks, termasuk nuansa kesantunan yang tidak selalu dinyatakan secara eksplisit (Huzaifa et al., 2025). Hal ini menjadi penting dalam analisis teks chat akademik yang memiliki variasi bahasa dan konteks yang beragam.

Meskipun sebagian data penelitian berasal dari media sosial dan mengandung unsur bahasa informal, penelitian ini tetap menggunakan IndoBERT karena fokus utama penelitian berada pada konteks komunikasi akademik mahasiswa-dosen yang memiliki karakteristik semi-formal. IndoBERT dipilih karena telah dilatih pada korpus umum bahasa Indonesia sehingga mampu memahami konteks bahasa secara lebih luas dan stabil dalam domain akademik. Beberapa penelitian lain menggunakan IndoBERTweet untuk teks media sosial karena model tersebut lebih difokuskan pada karakteristik bahasa informal seperti *slang*, singkatan, dan gaya percakapan Twitter (Koto et al., 2021). Namun, dalam penelitian ini IndoBERT dinilai lebih sesuai untuk mendukung tugas klasifikasi kesantunan dalam konteks komunikasi akademik. Dalam penelitian ini, IndoBERT digunakan sebagai model klasifikasi untuk mengidentifikasi tingkat kesantunan teks komunikasi akademik mahasiswa-dosen ke dalam dua kategori, yaitu sopan dan tidak sopan, karena kemampuannya dalam menangkap konteks bahasa secara mendalam dan memahami makna implisit, sehingga lebih efektif dalam membedakan tingkat kesantunan dibandingkan pendekatan tradisional.

2.7 IndoT5 untuk Transformasi Teks

IndoT5 merupakan model bahasa berbasis *Text-to-Text Transfer Transformer* (T5) yang digunakan untuk menyelesaikan berbagai tugas pemrosesan bahasa alami melalui pendekatan teks-ke-teks (Fuadi et al., 2023). Model T5 merupakan varian arsitektur *Transformer* yang mengubah berbagai permasalahan berbasis teks dalam NLP ke dalam format input dan *output* teks. Pendekatan ini memungkinkan berbagai tugas seperti ringkasan teks, pelabelan frasa, dan generasi teks otomatis diproses dalam satu kerangka yang sama (Puspitasari et al., 2024). IndoT5 dikembangkan dari model multilingual mT5 dengan menyesuaikan model agar spesifik untuk bahasa Indonesia. Sebagai model *pre-trained transformer*, IndoT5 mampu memahami karakteristik linguistik bahasa Indonesia tanpa mengubah makna dari kalimat awal.

Keunggulan IndoT5 terletak pada kemampuannya dalam melakukan transformasi teks secara kontekstual, yaitu mengubah gaya bahasa suatu kalimat tanpa mengubah makna utamanya. Hal ini menjadi penting dalam konteks penelitian ini yang berfokus pada perubahan tingkat kesantunan dalam teks chat akademik. Dalam penelitian ini, IndoT5 digunakan untuk mentransformasikan teks chat mahasiswa yang teridentifikasi tidak sopan menjadi kalimat yang lebih santun sesuai dengan norma komunikasi akademik. IndoT5 dipilih karena menggunakan pendekatan *text-to-text transformer* yang memungkinkan model menghasilkan kalimat baru secara kontekstual dengan tetap mempertahankan maksud utama pesan. Kemampuan tersebut dinilai sesuai dengan tujuan penelitian, yaitu mengubah tingkat kesantunan teks tanpa menghilangkan informasi dan makna asli dari komunikasi akademik mahasiswa-dosen.

2.8 Preprocessing Teks

Preprocessing teks merupakan tahapan awal dalam pemrosesan bahasa alami yang bertujuan untuk membersihkan dan menyiapkan data sebelum digunakan dalam pelatihan model (Irmayanti & Hermanto, 2026). Tahapan ini dilakukan untuk mengurangi *noise* sehingga data yang digunakan memiliki kualitas yang lebih baik dan dapat diproses secara optimal oleh model. Secara umum, *preprocessing* mencakup proses normalisasi teks dan penghapusan karakter yang tidak relevan. Dalam penelitian ini, *preprocessing* berperan penting dalam menyiapkan data teks *chat* akademik agar dapat dianalisis secara lebih efektif oleh model. Data yang telah melalui tahap *preprocessing* kemudian digunakan sebagai masukan dalam proses klasifikasi dan transformasi (Kono et al., 2025).

2.9 Evaluasi

Pengujian model dilakukan untuk mengukur kinerja sistem dalam menyelesaikan tugas yang diberikan. Dalam konteks penelitian NLP ini, evaluasi dilakukan terhadap dua model utama, yaitu model klasifikasi kesantunan menggunakan IndoBERT dan model transformasi teks menggunakan IndoT5. Selain itu, dilakukan juga evaluasi manual sebagai bentuk validasi tambahan terhadap hasil luaran sistem. Dengan demikian, kombinasi evaluasi otomatis dan manual memberikan gambaran yang lebih komprehensif terhadap kinerja sistem yang dikembangkan.

2.9.1 Evaluasi Model Klasifikasi (IndoBERT)

Evaluasi model klasifikasi dilakukan untuk mengukur kemampuan model dalam mengidentifikasi tingkat kesantunan teks *chat* mahasiswa-dosen ke dalam kategori sopan dan tidak sopan. Pengukuran dilakukan menggunakan *confusion matrix*, kemudian dihitung beberapa metrik evaluasi, yaitu *accuracy*, *precision*, *recall*, dan *F1-Score*. Metrik-metrik tersebut digunakan karena mampu memberikan gambaran menyeluruh mengenai performa model dalam membedakan tingkat kesantunan secara akurat.

Confusion Matrix

Confusion matrix merupakan sebuah metode evaluasi yang digunakan untuk menganalisis seberapa baik model dalam menghasilkan prediksi terhadap data sebenarnya. Matriks ini memberikan gambaran akurasi jumlah prediksi yang benar maupun salah yang dihasilkan oleh model. *Confusion matrix* memiliki empat komponen utama, yaitu:

- a. TP (*True Positive*) sebagai prediksi model yang tidak sopan dan teksnya benar tidak sopan.
- b. TN (*True Negative*) sebagai prediksi model yang sopan dan teksnya benar sopan.
- c. FP (*False Positive*) sebagai prediksi model tidak sopan padahal teksnya sopan
- d. FN (*False Negative*) sebagai prediksi model sopan padahal teksnya tidak sopan.

Berdasarkan nilai-nilai ini berbagai metrik evaluasi seperti *accuracy*, *precision*, *recall*, dan *F1-Score* dapat dihitung untuk menilai performa model klasifikasi.

Accuracy

Accuracy merupakan metrik evaluasi yang digunakan untuk mengukur seberapa benar prediksi teks dari keseluruhan yang ditebak oleh model. Metrik ini dihitung dengan membandingkan jumlah prediksi yang benar terhadap total keseluruhan data yang diuji. Nilai

accuracy yang tinggi menunjukkan bahwa model mampu mengklasifikasikan data dengan ketepatan yang baik.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2.1)$$

Precision

Precision merupakan metrik evaluasi yang digunakan untuk mengukur tingkat seberapa tepat prediksi persentase model yang tidak sopan. *Precision* menunjukkan proporsi jumlah prediksi positif yang benar dibandingkan dengan seluruh prediksi positif yang dihasilkan oleh model. Secara matematis *precision* dapat dituliskan sebagai berikut:

$$Precision = \frac{TP}{TP + FP} \quad (2.2)$$

Recall

Recall merupakan metrik evaluasi yang digunakan untuk mengukur kemampuan model dalam mengidentifikasi seluruh data yang termasuk ke dalam kelas positif. Dalam penelitian ini, *recall* menunjukkan seberapa besar persentase pesan yang tidak sopan yang berhasil dikenali oleh model dibandingkan dengan seluruh pesan tidak sopan yang sebenarnya terdapat dalam *dataset*. Nilai *recall* yang tinggi menunjukkan bahwa model memiliki kemampuan yang baik dalam menangkap sebagian besar pesan yang tidak sopan tanpa banyak melewatkan data yang teridentifikasi. Rumus *recall* dapat dituliskan sebagai berikut:

$$Recall = \frac{TP}{TP + FN} \quad (2.3)$$

F1-Score

F1-Score merupakan metrik evaluasi keseimbangan model antara *precision* dan *recall* dalam satu nilai. Nilai *F1-Score* dihitung berdasarkan rata-rata harmonik dari *precision* dan *recall*. Rumus *F1-Score* dapat dituliskan sebagai berikut:

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (2.4)$$

2.9.2 Evaluasi Model Transformasi Teks (IndoT5)

Evaluasi transformasi teks dilakukan dengan membandingkan hasil keluaran model dengan teks referensi menggunakan metrik BLEU, ROUGE, dan METEOR untuk menilai kualitas kesesuaian dan makna kalimat. Metrik ini digunakan untuk mengukur kualitas kesesuaian dan makna kalimat yang dihasilkan oleh model IndoT5.

BLEU, ROUGE, dan METEOR

Penelitian ini juga melakukan evaluasi terhadap model transformasi teks yang dibangun menggunakan IndoT5. Evaluasi dilakukan menggunakan metrik *BLEU*, *ROUGE-1*, *ROUGE-2*, *ROUGE-L*, dan *METEOR* untuk mengukur kualitas hasil transformasi kalimat yang dihasilkan oleh model. *BLEU* (*Bilingual Evaluation Understudy*) digunakan untuk mengukur tingkat kesamaan antara teks generasi dengan teks referensi berdasarkan kesamaan *n-gram*. *ROUGE* (*Recall-Oriented Understudy for Gisting Evaluation*) merupakan metrik evaluasi yang digunakan untuk membandingkan kesamaan antara teks referensi dan teks hasil generasi. Dalam penelitian ini digunakan beberapa varian *ROUGE*, yaitu *ROUGE-1* yang mengukur kesamaan *unigram* atau kesamaan kata, *ROUGE-2* yang mengukur kesamaan *bigram* atau kesamaan pasangan kata, serta *ROUGE-L* yang mengukur kesamaan berdasarkan urutan kalimat. Sementara itu, *METEOR* (*Metric for Evaluation of Translation with Explicit Ordering*) digunakan untuk mengukur kesamaan makna dan sinonim pada kalimat. Penggunaan metrik-metrik tersebut memungkinkan evaluasi yang lebih komprehensif terhadap kualitas hasil transformasi teks yang dihasilkan oleh model.

2.9.3 Evaluasi Manual

Evaluasi manual merupakan metode penilaian yang dilakukan oleh manusia untuk menilai kualitas hasil pemrosesan teks, terutama dalam aspek yang sulit diukur secara otomatis, seperti kesesuaian makna dan tingkat kesantunan. Evaluasi ini dilakukan dengan melibatkan evaluator yang memiliki pemahaman terhadap konteks bahasa dan komunikasi. Dalam pengolahan bahasa alami, evaluasi manual sering digunakan sebagai pelengkap evaluasi otomatis untuk memastikan bahwa hasil yang dihasilkan sistem tidak hanya sesuai secara teknis, tetapi juga relevan secara kontekstual.

2.10 Penelitian Terdahulu

Pencarian literatur pada penelitian terdahulu dilakukan menggunakan Google Scholar, ResearchGate, ScienceDirect, dan IEEE Xplore. Rentang tahun yang digunakan dalam pencarian ini antara tahun 2021-2026. Pencarian literatur menggunakan kata kunci ‘*text classification*’, ‘*sentiment analysis*’, ‘*transformer model*’, ‘*IndoBERT*’, ‘*IndoT5*’, serta ‘*text transformation*’. Tujuan dari pencarian literatur ini adalah untuk mengidentifikasi perkembangan metode yang digunakan dalam klasifikasi teks dan transformasi teks pada penelitian sebelumnya. Hasil dari identifikasi kemudian digunakan sebagai dasar dalam menentukan pendekatan yang digunakan dalam penelitian ini. Analisis penelitian terdahulu disajikan pada Tabel 2.1 dan Tabel 2.2.

Tabel 2.1 Penelitian Terdahulu Klasifikasi Teks

Peneliti	Metode	Performa	Fokus	Jenis Model	Konteks
(Fikri & Muliono, 2025)	Naïve Bayes	79,75%	Klasifikasi bahasa kasar	<i>Machine Learning</i>	Media Sosial
(Akhfasy et al., 2024)	Support Vector Machine	97,74%	Klasifikasi deteksi ujaran kebencian	<i>Machine Learning</i>	Media Sosial
(Sari et al., 2024)	Logistic Regression	68,5%	Klasifikasi konten kasar	<i>Machine Learning</i>	Media Sosial
	Naïve Bayes	63,1%			
(Hakim et al., 2026)	Fine -Tuning IndoBERT	88,34%	Klasifikasi komentar kasar	<i>Transformer</i>	Media Sosial

Berdasarkan Tabel 2.1, penelitian terdahulu dianalisis menggunakan beberapa bingkai analisis, yaitu berdasarkan fokus penelitian, metode yang digunakan, performa model, bahasa dan konteks data, integrasi sistem, serta metode evaluasi. Dari segi fokus penelitian, seluruh penelitian pada Tabel 2.1 berfokus pada klasifikasi bahasa kasar, ujaran kebencian, dan komentar toksik pada media sosial berbahasa Indonesia. Penelitian oleh Fikri & Muliono (2025), Akhfasy et al. (2024), dan Sari et al. (2024) menggunakan metode *machine learning* klasik untuk melakukan klasifikasi teks kasar dan ujaran kebencian. Sementara itu, penelitian Hakim et al. (2026) menggunakan pendekatan *fine-tuning* IndoBERT untuk melakukan klasifikasi komentar toksik berbahasa Indonesia.

Ditinjau dari metode yang digunakan, penelitian terdahulu menunjukkan adanya perkembangan dari metode machine learning klasik seperti Naïve Bayes, Support Vector Machine (SVM), dan Logistic Regression menuju model berbasis *transformer* seperti IndoBERT. Metode *machine learning* klasik masih bergantung pada representasi fitur seperti *TF-IDF* dan *bag-of-words* sehingga pemahaman konteks bahasa menjadi terbatas. Sebaliknya, model *transformer* seperti IndoBERT mampu memahami hubungan konteks antar kata secara lebih mendalam melalui mekanisme *attention* serta proses *pre-training* pada korpus bahasa Indonesia dalam jumlah besar. Oleh karena itu, penelitian ini memilih pendekatan *fine-tuning* IndoBERT karena dinilai lebih efektif dalam memahami konteks penggunaan bahasa kasar dan tidak kasar dibandingkan metode *machine learning* klasik.

Dari segi performa model, penelitian Akhfasy et al. (2024) memperoleh performa tertinggi menggunakan metode Support Vector Machine dengan akurasi sebesar 97,74%, sedangkan penelitian Hakim et al. (2026) menggunakan *fine-tuning* IndoBERT dan memperoleh akurasi sebesar 88,34%. Meskipun beberapa metode *machine learning* klasik menunjukkan performa akurasi yang tinggi, model *transformer* seperti IndoBERT memiliki keunggulan dalam memahami konteks bahasa secara lebih mendalam dan kontekstual. Hal ini menjadikan IndoBERT lebih sesuai digunakan pada penelitian yang membutuhkan pemahaman konteks bahasa secara kontekstual.

Ditinjau dari segi bahasa dan konteks data, seluruh penelitian menggunakan data berbahasa Indonesia dengan konteks media sosial seperti Twitter, dan YouTube. Konteks tersebut masih didominasi oleh penggunaan bahasa informal dan komunikasi umum masyarakat. Penelitian pada konteks komunikasi formal, khususnya komunikasi akademik antara mahasiswa dan dosen, masih sangat terbatas. Padahal, komunikasi akademik memiliki karakteristik bahasa yang berbeda karena menuntut penggunaan bahasa yang lebih sopan, formal, dan kontekstual. Jika ditinjau dari aspek integrasi sistem, penelitian terdahulu masih berfokus pada satu tugas saja, yaitu klasifikasi bahasa kasar atau ujaran kebencian tanpa disertai proses transformasi teks menjadi bahasa yang lebih sopan. Selain itu, dari sisi evaluasi, penelitian terdahulu sebagian besar hanya menggunakan evaluasi otomatis seperti *accuracy* dan *F1-score* sehingga belum mempertimbangkan evaluasi manual terhadap tingkat kesopanan bahasa secara kontekstual.

Tabel 2.2 Penelitian Terdahulu Transformasi Teks

Peneliti	Metode	Performa	Fokus	Jenis Model	Konteks
(Christiant & Pratama, 2022)	Tag and Generate	<i>BLEU-1</i> : 53,48 <i>BLEU-2</i> : 40,30	Transformasi	<i>Deep Learning</i>	Umum
(Zaman et al., 2025)	Enhanced Fine-Tuned Sahabat-AI	<i>Joint Score</i> : 0,7295	<i>Detoxification</i> teks <i>toxic</i> menjadi <i>non-toxic</i>	<i>LLM Transformer</i>	Media Sosial
(Robet et al., 2025)	Fine-Tuning IndoT5	<i>BLEU</i> : 50,1 <i>ROUGE-L</i> : 61,7	Transformasi teks bahasa Indonesia	<i>Transformer</i>	Umum

Berdasarkan Tabel 2.2, penelitian terdahulu pada bidang transformasi teks dianalisis berdasarkan fokus penelitian, metode yang digunakan, performa model, bahasa dan konteks data, integrasi sistem, serta metode evaluasi. Dari segi fokus penelitian, penelitian transformasi teks umumnya berfokus pada proses *paraphrasing*, *style transfer*, dan *text detoxification* untuk menghasilkan teks yang lebih natural tanpa mengubah makna utama kalimat. Penelitian Christiant & Pratama (2022) menggunakan metode *Tag and Generate* untuk melakukan transformasi teks melalui proses *paraphrase*. Sementara itu, penelitian Zaman et al. (2025) berfokus pada *detoxification* teks *toxic* menjadi *non-toxic* menggunakan model berbasis *large language model* (LLM). Penelitian lainnya oleh Robet et al. (2025) menggunakan pendekatan *fine-tuning IndoT5* untuk melakukan transformasi atau *paraphrase* teks bahasa Indonesia secara otomatis.

Ditinjau dari metode yang digunakan, penelitian transformasi teks menunjukkan perkembangan dari pendekatan *deep learning* menuju model berbasis *transformer* dan *large language model*. Penelitian Christiant & Pratama (2022) masih menggunakan metode *Tag and Generate* berbasis *deep learning*, sedangkan penelitian Zaman et al. (2025) dan Robet et al. (2025) mulai memanfaatkan model *transformer* seperti Sahabat-AI dan IndoT5. Model *transformer* dinilai lebih efektif dalam tugas transformasi teks karena mampu memahami hubungan konteks antar kata serta menghasilkan teks yang lebih natural dan koheren dibandingkan metode tradisional. Oleh karena itu, penelitian ini memilih IndoT5 sebagai model transformasi teks karena memiliki kemampuan *text-to-text generation* dan *paraphrasing* yang baik dalam bahasa Indonesia.

Dari segi performa model, penelitian Christiant & Pratama (2022) memperoleh nilai BLEU-1 sebesar 53,48 dan BLEU-2 sebesar 40,30. Penelitian Zaman et al. (2025) memperoleh Joint Score sebesar 0,7295 pada tugas *detoxification teks toxic* menjadi *non-toxic*. Sementara itu, penelitian Robet et al. (2025) menggunakan fine-tuning IndoT5 dan memperoleh nilai BLEU sebesar 50,1 serta ROUGE-L sebesar 61,7. Hasil tersebut menunjukkan bahwa model *transformer* mampu menghasilkan teks transformasi yang tetap mempertahankan makna utama kalimat dengan kualitas yang cukup baik. Ditinjau dari segi bahasa dan konteks data, penelitian transformasi teks sebagian besar menggunakan bahasa Indonesia dengan konteks umum dan media sosial. Penelitian Zaman et al. (2025) secara khusus menggunakan konteks media sosial yang berkaitan dengan teks *toxic*, sedangkan penelitian Robet et al. (2025) masih berfokus pada transformasi teks umum. Penelitian pada konteks komunikasi formal, khususnya komunikasi akademik mahasiswa-dosen, masih sangat terbatas. Selain itu, penelitian terkait penggunaan IndoT5 untuk transformasi kalimat kasar menjadi lebih sopan juga masih jarang ditemukan karena sebagian besar penelitian IndoT5 masih berfokus pada *paraphrasing* umum.

Jika ditinjau dari aspek integrasi sistem, penelitian terdahulu pada bidang transformasi teks masih berfokus pada proses transformasi saja tanpa disertai proses klasifikasi bahasa kasar terlebih dahulu. Selain itu, dari sisi evaluasi, penelitian terdahulu sebagian besar menggunakan evaluasi otomatis seperti *BLEU*, *ROUGE*, dan *Joint Score* sehingga belum mempertimbangkan evaluasi manual terhadap tingkat kesopanan bahasa secara kontekstual.

Berdasarkan analisis tersebut, dapat diidentifikasi bahwa terdapat beberapa gap dalam penelitian terdahulu, yaitu: (1) belum adanya penelitian yang mengintegrasikan klasifikasi dan transformasi teks dalam satu sistem terpadu, (2) masih terbatasnya penelitian yang berfokus pada konteks komunikasi akademik mahasiswa-dosen, dan (3) belum optimalnya penggunaan evaluasi manual atau validasi ahli dalam menilai kualitas kesopanan hasil transformasi teks. Oleh karena itu, penelitian ini diusulkan untuk mengisi celah tersebut dengan mengembangkan sistem yang mengintegrasikan klasifikasi kesopanan menggunakan IndoBERT dan transformasi kalimat tidak santun menjadi lebih santun menggunakan IndoT5 dalam konteks komunikasi akademik mahasiswa-dosen.

BAB III

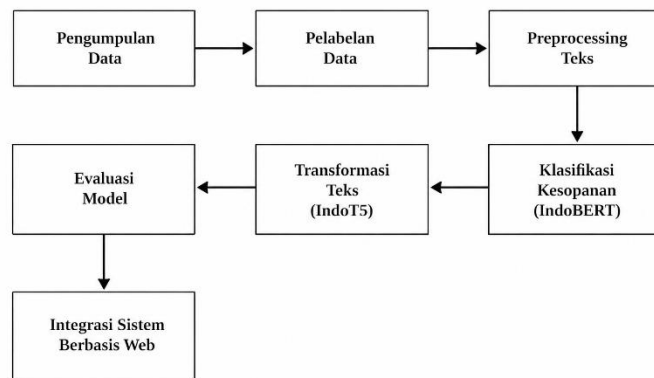
METODOLOGI PENELITIAN

3.1 Jenis dan Pendekatan Penelitian

Penelitian ini merupakan penelitian terapan yang berfokus pada pengembangan sistem berbasis kecerdasan buatan untuk meningkatkan kualitas komunikasi akademik dalam bentuk teks *chat* mahasiswa-dosen. Penelitian ini bertujuan untuk menghasilkan solusi yang tidak hanya mampu menganalisis tingkat kesantunan, tetapi juga mentransformasikan teks menjadi lebih santun secara otomatis. Pendekatan yang digunakan dalam penelitian ini adalah pendekatan eksperimen dalam bidang *Natural Language Processing* (NLP) dengan memanfaatkan metode *deep learning* berbasis arsitektur *transformer*. Model yang digunakan dalam penelitian ini adalah IndoBERT untuk tugas klasifikasi kesantunan dan IndoT5 untuk transformasi teks. Penelitian ini termasuk dalam kategori pengembangan sistem berbasis model *machine learning*, di mana sistem dirancang untuk memproses input berupa teks *chat* akademik, kemudian menghasilkan *output* berupa klasifikasi tingkat kesantunan serta transformasi teks menjadi lebih santun. Sistem yang dikembangkan berfokus pada teks yang dikirim oleh mahasiswa kepada dosen sebagai bagian dari konteks komunikasi akademik mahasiswa-dosen. Dengan demikian, penelitian ini tidak hanya berfokus pada analisis data, tetapi juga pada implementasi sistem berbasis web yang terintegrasi. Berdasarkan jenis dan pendekatan penelitian yang telah dijelaskan, tahapan penelitian disusun secara sistematis untuk mendukung pengembangan sistem dan dijelaskan pada subbab berikutnya.

3.2 Alur Penelitian

Alur penelitian ini disusun secara sistematis untuk menggambarkan tahapan yang dilakukan dalam pengembangan sistem yang diusulkan. Secara umum, penelitian ini terdiri dari beberapa tahapan utama, yaitu pengumpulan data, pelabelan data, *preprocessing*, proses klasifikasi, transformasi teks, evaluasi, dan integrasi sistem. Seluruh tahapan penelitian ini divisualisasikan dalam Gambar 3.1.



Gambar 3.1 Alur Penelitian

Pada Gambar 3.1, tahapan penelitian dimulai dari pengumpulan data berupa teks *chat* mahasiswa-dosen yang diperoleh dari berbagai sumber sesuai dengan kebutuhan penelitian. Selanjutnya, dilakukan proses pelabelan data secara manual untuk mengelompokkan teks ke dalam kategori sopan dan tidak sopan berdasarkan indikator kesantunan yang telah ditentukan. Selanjutnya, dilakukan tahapan *preprocessing*, yaitu proses pembersihan dan penyiapan data agar dapat digunakan dalam proses pemodelan. Tahapan berikutnya adalah proses klasifikasi menggunakan model IndoBERT untuk mengidentifikasi tingkat kesantunan teks. Teks yang terklasifikasi tidak sopan kemudian diproses menggunakan model IndoT5 untuk ditransformasikan menjadi kalimat yang lebih santun tanpa mengubah makna utama. Setelah itu, dilakukan evaluasi terhadap hasil klasifikasi dan transformasi teks menggunakan metrik yang telah ditentukan untuk mengukur performa sistem. Tahapan terakhir adalah integrasi sistem dalam bentuk aplikasi berbasis web yang memungkinkan pengguna untuk memasukkan teks dan memperoleh hasil klasifikasi serta transformasi secara langsung.

3.3 Pengumpulan Data

Data yang digunakan dalam penelitian ini berupa teks komunikasi akademik antara mahasiswa dan dosen dalam bahasa Indonesia. Pengumpulan data dilakukan dari beberapa sumber yaitu Google Form, media sosial berupa unggahan TikTok yang menampilkan ulang tangkapan layar percakapan akademik dari platform X dan Facebook, serta data tambahan yang dihasilkan menggunakan ChatGPT. Pengumpulan data melalui Google Form dilakukan dengan menyebarkan kuesioner kepada mahasiswa aktif yang memiliki pengalaman berkomunikasi dengan dosen melalui media *chat*. Responden diminta untuk menuliskan contoh pesan yang pernah digunakan dalam komunikasi akademik. Sementara itu, data dari media sosial diperoleh

secara manual dengan mengumpulkan unggahan atau tangkapan layar percakapan mahasiswa dan dosen yang relevan dengan konteks akademik. Penelitian ini tidak menggunakan data percakapan WhatsApp secara langsung karena mempertimbangkan aspek privasi dan kerahasiaan komunikasi personal antara mahasiswa dan dosen.

Data asli yang diperoleh terdiri atas 106 data hasil Google Form dan sekitar 90 data dari media sosial TikTok. Jumlah data tersebut dinilai masih terbatas untuk proses pelatihan model, sehingga dilakukan pengembangan variasi kalimat menggunakan teknik *prompting* dengan bantuan ChatGPT berdasarkan indikator kesantunan yang telah ditentukan sebelumnya. Data tambahan yang dihasilkan menggunakan ChatGPT digunakan sebagai pelengkap untuk memperkaya variasi struktur kalimat dalam *dataset*, khususnya dalam bentuk ekspresi bahasa yang mungkin jarang ditemukan pada data nyata. Proses pengembangan dilakukan dengan tetap mempertahankan konteks, maksud pesan, serta tingkat kesantunan dari data asli sehingga model diharapkan mampu mempelajari pola bahasa yang lebih beragam.

Namun demikian, penggunaan data sintetis dari ChatGPT memiliki beberapa keterbatasan. Data yang dihasilkan cenderung memiliki struktur bahasa yang lebih rapi dan formal dibandingkan komunikasi nyata, sehingga berpotensi mengurangi tingkat natural data. Selain itu, terdapat kemungkinan bias gaya bahasa yang dihasilkan oleh model, yang dapat memengaruhi distribusi data. Oleh karena itu, seluruh hasil pengembangan data kemudian melalui proses pemeriksaan dan kurasi manual untuk memastikan bahwa kalimat hasil generasi tetap sesuai dengan konteks komunikasi akademik, tidak menyimpang dari makna asli, serta sesuai dengan kategori kesantunan yang ditentukan sebelum dimasukkan ke dalam *dataset* penelitian.

Proses pengumpulan data dilakukan pada rentang waktu Oktober 2025 hingga Desember 2025. Untuk memastikan kualitas data, diterapkan beberapa kriteria dalam proses seleksi. Kriteria inklusi meliputi: (1) teks merupakan komunikasi akademik mahasiswa-dosen, (2) menggunakan bahasa Indonesia, dan (3) memiliki konteks akademik seperti konsultasi skripsi, bimbingan, permohonan izin, atau komunikasi administratif. Sebaliknya, kriteria eksklusi meliputi: (1) teks yang tidak memiliki konteks akademik, (2) penggunaan bahasa yang tidak jelas atau tidak lengkap, (3) teks yang terlalu pendek dan tidak merepresentasikan maksud komunikasi, serta (4) data yang terduplikasi dari sumber yang berbeda. Selanjutnya, dilakukan proses kurasi dengan menyeleksi teks yang sesuai dengan konteks komunikasi akademik serta menghapus data yang tidak memenuhi kriteria. Selain itu dilakukan pemeriksaan duplikasi untuk memastikan tidak terdapat data yang sama antar sumber. Seluruh data yang telah dikurasi

kemudian disimpan dalam format *Comma Separated Value* (CSV) untuk memudahkan proses pengolahan pada tahap selanjutnya. Total data yang diperoleh berjumlah 6.230 kalimat, yang terdiri atas 2.220 kalimat dari Google Form, 2.008 kalimat dari media sosial TikTok, dan 2.002 kalimat dari ChatGPT. Detail contoh data dapat dilihat pada Lampiran A, contoh pengembangan data menggunakan prompting ChatGPT untuk Google Form ditampilkan pada Lampiran B, contoh pengembangan data menggunakan *prompting* ChatGPT untuk TikTok ditampilkan pada Lampiran C dan *draft* kuesioner Google Form ditampilkan pada Lampiran D.

3.4 Pelabelan Data

Data yang telah dikumpulkan dan disimpan dalam format *Comma Separated Value* (CSV) terlebih dahulu diperiksa untuk memastikan tidak terdapat duplikasi data. Setelah data bersih dari duplikasi, tahap selanjutnya adalah proses pelabelan data yang bertujuan untuk memberikan kategori tertentu pada setiap data yang digunakan dalam proses pemodelan. Pelabelan data dilakukan secara manual ke dalam dua kelas, yaitu label 0 untuk kalimat sopan dan label 1 untuk kalimat tidak sopan. Pemilihan dua kelas ini dilakukan untuk menyederhanakan proses klasifikasi serta menghindari ambiguitas dalam penentuan label. Kategori tambahan seperti “netral” atau “kurang sopan” berpotensi menimbulkan interpretasi subjektif yang tinggi, sehingga dapat menurunkan konsistensi dalam proses pelabelan data. Dengan menggunakan dua kelas yang memiliki batasan yang lebih jelas, proses anotasi dapat dilakukan secara lebih konsisten berdasarkan indikator kesantunan yang telah ditentukan. Selain itu, pendekatan klasifikasi biner juga mempermudah model dalam mempelajari perbedaan karakteristik antara teks yang sesuai dan tidak sesuai dengan norma komunikasi akademik.

Proses pelabelan dilakukan dengan membaca serta menganalisis setiap teks *chat* mahasiswa-dosen berdasarkan konteks komunikasi akademik. Dalam proses pelabelan data, penilaian kesantunan dilakukan dengan mempertimbangkan konteks komunikasi akademik secara umum dan tidak berorientasi pada karakteristik budaya atau wilayah tertentu. Hal ini dilakukan karena gaya komunikasi dan persepsi terhadap kesantunan dapat dipengaruhi oleh latar belakang budaya, kebiasaan berbahasa, konteks sosial, serta relasi antara penutur dan lawan tutur. Oleh karena itu, pelabelan difokuskan pada indikator kesantunan formal yang umum digunakan dalam komunikasi mahasiswa-dosen, seperti penggunaan salam pembuka, identitas diri, penyampaian maksud yang jelas, serta penggunaan bahasa yang sopan dan

menghormati lawan tutur. Pelabelan dilakukan oleh peneliti sebagai anotator tunggal dengan mengacu pada indikator kesantunan berbahasa. Penggunaan anotator tunggal dalam penelitian ini dilakukan untuk menjaga konsistensi penilaian berdasarkan indikator kesantunan yang telah ditetapkan. Indikator kesantunan yang digunakan dalam penelitian ini mengacu pada penelitian sebelumnya Eksan et al. (2021) dan Mahas et al. (2025) yang kemudian diadaptasi ke dalam beberapa indikator operasional, yaitu: (1) adanya salam pembuka, (2) pengenalan identitas diri, (3) kejelasan tujuan pesan, (4) penggunaan bahasa baku dan sopan, (5) bentuk permintaan yang tidak memaksa, serta (6) tidak menggunakan singkatan atau bahasa informal. Indikator ini digunakan sebagai pedoman agar proses pelabelan dilakukan secara konsisten.

Selain itu, dilakukan proses pengecekan ulang terhadap hasil pelabelan untuk meminimalkan kesalahan dan bias. Dalam beberapa kasus, terdapat teks yang secara literal terlihat sopan, namun tidak sesuai dengan etika komunikasi akademik, misalnya penggunaan bahasa yang terlalu singkat tanpa konteks atau nada terkesan mendesak. Sebaliknya, terdapat juga teks yang tetap dianggap santun karena menggunakan struktur komunikasi yang sesuai. Oleh karena itu, pelabelan dilakukan dengan mempertimbangkan konteks pragmatik komunikasi secara menyeluruh, tidak hanya berdasarkan pada kata-kata yang digunakan. Contoh hasil pelabelan data ditunjukkan oleh Tabel 3.1.

Tabel 3.1 Contoh Hasil Pelabelan

Teks	Label	Bentuk Teks Sopan atau Tidak Sopannya
Bu, saya udah ngirim tugasnya, cepet diperiksa ya	1	Bu, saya sudah mengirimkan tugasnya. Mohon kesediannya untuk diperiksa jika berkenan. Terima kasih sebelumnya.
Pak saya telat soalnya macet	1	Mohon maaf, Pak. Saya datang sedikit terlambat karena kondisi jalan sedang macet.
Bu kapan nilai saya keluar? lama banget	1	Permisi, Bu. Saya ingin menanyakan apakah nilai mata kuliah ini sudah keluar? Terima kasih sebelumnya.
Mohon informasi dari Ibu terkait jadwal pengumpulan revisi terakhir.	0	Bu <i>deadline</i> revisi terakhir kapan ya?
Mohon arahan Ibu mengenai tata cara pengajuan sidang skripsi.	0	Bu, gimana sih cara daftar sidang skripsi?
Mohon informasi tambahan dari Ibu terkait poin penilaian yang belum terpenuhi.	0	Bu, yang kurang di nilai saya tuh bagian mana ya?

3.5 *Preprocessing Data*

Preprocessing data merupakan tahapan awal dalam pemrosesan bahasa alami yang bertujuan untuk membersihkan dan menyiapkan data sebelum digunakan dalam proses pelatihan model. Tahapan ini dilakukan untuk mengurangi *noise* sehingga model dapat mempelajari pola bahasa dengan lebih efektif. Pada penelitian ini, *preprocessing* dilakukan secara sistematis melalui beberapa tahapan sebagai berikut: (1) persiapan data, (2) *case folding*, (3) penghapusan URL, (4) penghapusan *mention*, (5) penghapusan *hashtag*, (6) normalisasi spasi, (7) penghapusan simbol dan emoji, serta (8) penghapusan teks dengan panjang tertentu. *Case folding* dilakukan dengan mengubah seluruh huruf pada teks menjadi huruf kecil untuk menyamakan bentuk penulisan kata. Selanjutnya, elemen yang tidak memiliki makna linguistik seperti URL, *mention* media sosial, simbol, dan emoji dihapus karena tidak memberikan kontribusi terhadap pemahaman makna teks (Kono et al., 2025).

Tahap berikutnya adalah normalisasi spasi untuk menghilangkan spasi berlebih agar struktur teks menjadi lebih rapi dan konsisten. Selain itu, dilakukan penghapusan teks dengan panjang kurang dari lima karakter karena dianggap tidak memiliki informasi yang cukup untuk merepresentasikan konteks komunikasi akademik, misalnya teks seperti “Pagi, Bu”, “ya”, atau “sip”. Batasan ini ditetapkan berdasarkan pengamatan terhadap data dalam penelitian ini. Data yang telah melalui tahapan *preprocessing* kemudian digunakan sebagai masukan dalam proses klasifikasi kesantunan menggunakan model IndoBERT serta pada tahap transformasi teks menggunakan IndoT5. Seluruh proses *preprocessing* diimplementasikan menggunakan bahasa pemrograman Python dengan dukungan pustaka *PyTorch* dan *Hugging Face Transformers*.

3.6 *Pembangunan Model Klasifikasi Kesopanan*

Model klasifikasi kesopanan dalam penelitian ini dibangun menggunakan model bahasa IndoBERT dengan *pre-trained* model *indobenchmark/indobert-base-p1*. Model ini kemudian dilakukan proses *fine-tuning* menggunakan *dataset* komunikasi akademik mahasiswa-dosen yang telah melalui tahap *preprocessing*. *Dataset* dibagi menjadi data latih, data validasi, dan data uji dengan rasio 70%:15%:15% menggunakan metode *stratified split* untuk menjaga keseimbangan distribusi kelas sopan dan tidak sopan pada setiap *dataset* (Tarwoto et al., 2025). Berdasarkan pembagian tersebut, diperoleh sebanyak 4.354 data latih, 933 data validasi, dan 934 data uji. Hal ini bertujuan agar model mampu mempelajari pola kesopanan bahasa secara optimal tanpa mengalami bias terhadap salah satu kelas.

Pada proses pelatihan, digunakan *tokenizer* bawaan IndoBERT dengan panjang maksimum input (*max sequence length*) sebesar 128 token. Model dilatih menggunakan *optimizer AdamW* dengan *learning rate* sebesar $2e-5$ serta *loss function cross-entropy* untuk tugas klasifikasi biner. Pelatihan dilakukan dengan *batch size* 8 untuk data latih dan 16 untuk data validasi serta pengujian. Pemilihan parameter model, khususnya jumlah *epoch* dan *batch size*, dilakukan melalui serangkaian eksperimen untuk memperoleh konfigurasi terbaik. Pemilihan *hyperparameter* pada penelitian ini dilakukan berdasarkan hasil eksperimen pelatihan model serta mempertimbangkan stabilitas performa selama proses *fine-tuning*. Penggunaan *learning rate* sebesar $2e-5$ mengacu pada praktik umum *fine-tuning* model *transformer* berbasis BERT yang banyak digunakan pada penelitian sebelumnya (Baharuddin & Naufal, 2023). Sementara itu, variasi *epoch* dan *batch size* diuji untuk melihat pengaruhnya terhadap performa model, khususnya nilai akurasi dan potensi *overfitting* selama proses pelatihan. Hasil uji coba parameter tersebut ditunjukkan pada Tabel 3.2.

Tabel 3.2 Hasil Uji Coba Parameter IndoBERT

Epoch	Batch Size	Akurasi
2	4	96,78%
3	4	96,56%
4	4	95,56%
2	8	95,67%
3	8	97,00%
4	8	96,11%

Berdasarkan Tabel 3.2, konfigurasi dengan 3 *epoch* dan *batch size* 8 menghasilkan akurasi tertinggi sebesar 97,00% dibandingkan dengan konfigurasi lainnya. Hasil ini menunjukkan bahwa peningkatan jumlah *epoch* tidak selalu meningkatkan performa model. Pada beberapa konfigurasi, seperti *epoch* 4, terjadi penurunan akurasi yang mengindikasikan potensi *overfitting*. Oleh karena itu, konfigurasi dengan 3 *epoch* dan *batch size* 8 dipilih sebagai parameter terbaik dalam penelitian ini. Seluruh proses pelatihan dan evaluasi model diimplementasikan menggunakan bahasa pemrograman Python dengan dukungan pustaka *PyTorch* serta *Hugging Face Transformers*. Proses komputasi dilakukan dengan memanfaatkan perangkat GPU (CUDA). Kemampuan IndoBERT dalam memahami konteks bahasa secara dua arah (*bidirectional context*) memungkinkan model untuk mengidentifikasi

tingkat kesantunan berdasarkan hubungan antar kata dalam kalimat. Nilai *confidence* pada hasil klasifikasi diperoleh dari probabilitas *output* model IndoBERT yang dihitung menggunakan fungsi *Softmax* pada *library* PyTorch (`torch.nn.functional.softmax`). Probabilitas tertinggi dari hasil prediksi kelas kemudian digunakan sebagai *confidence score* untuk menunjukkan tingkat keyakinan model terhadap hasil klasifikasi. *Output* dari tahapan ini berupa label klasifikasi yang menunjukkan apakah suatu teks termasuk dalam kategori sopan atau tidak sopan.

3.7 Pembangunan Model Transformasi Teks

Model transformasi teks dalam penelitian ini dibangun menggunakan model IndoT5 dengan pendekatan *sequence-to-sequence* untuk memperbaiki kalimat yang terdeteksi tidak sopan agar menjadi lebih santun tanpa mengurangi makna utama pada pesan teks. *Dataset* transformasi telah disusun dalam bentuk pasangan teks sejak tahap pelabelan, yaitu berupa teks tidak sopan dan versi perbaikannya yang lebih santun. Pasangan data ini digunakan sebagai data *input-output*, di mana teks tidak sopan berperan sebagai input dan teks santun sebagai target dalam proses pelatihan model. Untuk menjaga agar makna asli tetap dipertahankan, proses penyusunan teks santun dilakukan dengan mempertahankan inti pesan serta tujuan komunikasi dari kalimat awal, sementara perubahan hanya dilakukan pada struktur kalimat dan pilihan kata agar sesuai dengan norma komunikasi akademik.

Dataset dibagi menggunakan metode *stratified split* yang menggunakan rasio 80%:10%:10% untuk data latih, data validasi dan data uji (Kono et al., 2025). Berdasarkan pembagian tersebut, diperoleh sebanyak 2.491 data latih, 311 data validasi, dan 312 data uji. Perbedaan rasio pembagian data pada model klasifikasi dan model transformasi dilakukan karena karakteristik kedua tugas yang berbeda. Pada model klasifikasi menggunakan IndoBERT, pembagian data 70%:15%:15% digunakan untuk memberikan proporsi data validasi dan pengujian yang lebih besar sehingga evaluasi performa klasifikasi dapat dilakukan secara lebih stabil dan terukur. Sementara itu, pada model transformasi menggunakan IndoT5, rasio 80%:10%:10% dipilih karena tugas generasi teks memerlukan data pelatihan yang lebih banyak agar model mampu mempelajari variasi pola transformasi kalimat secara lebih optimal.

Format input ke dalam model menggunakan pendekatan *text-to-text*, yaitu teks tidak sopan sebagai input dan teks santun sebagai output tanpa menggunakan *prompt* tambahan khusus. Proses pelatihan dilakukan selama 3 *epoch* dengan *batch size* sebesar 4 untuk data latih dan 8 untuk data validasi. Pemilihan parameter ini didasarkan pada hasil eksperimen awal yang

menunjukkan bahwa konfigurasi tersebut memberikan hasil transformasi yang paling stabil. Model dilatih menggunakan *tokenizer* bawaan dengan pendekatan generatif berbasis *transformer*. Evaluasi dilakukan untuk mengukur kemampuan model dalam menghasilkan kalimat yang lebih santun dengan tetap mempertahankan kesesuaian makna dari teks asli.

3.8 Evaluasi Model

Evaluasi model dilakukan untuk mengetahui performa sistem yang telah dibangun, baik pada tahap klasifikasi kesopanan maupun transformasi teks. Evaluasi ini dilakukan dengan menggunakan berbagai metrik evaluasi serta penilaian manual untuk memperoleh gambaran kinerja sistem secara menyeluruh.

3.8.1 Evaluasi Model Klasifikasi

Evaluasi pada model klasifikasi kesopanan dilakukan menggunakan data uji yang telah dipisahkan pada tahap pembagian *dataset*. Model IndoBERT digunakan untuk memprediksi label pada data uji, kemudian hasil prediksi dibandingkan dengan label sebenarnya. Metrik evaluasi yang digunakan meliputi *confusion matrix*, *precision*, *recall*, dan *F1-score*. *Confusion matrix* dibentuk berdasarkan perbandingan antara label prediksi dan label sebenarnya untuk mengetahui jumlah prediksi benar dan salah pada masing-masing kelas. Nilai *precision*, *recall*, dan *F1-score* dihitung berdasarkan hasil *confusion matrix* tersebut. Metrik ini digunakan untuk menilai kemampuan model dalam mengklasifikasikan teks komunikasi akademik mahasiswa-dosen ke dalam kategori sopan dan tidak sopan secara akurat.

3.8.2 Evaluasi Model Transformasi

Evaluasi pada model transformasi teks dilakukan dengan membandingkan hasil keluaran model IndoT5 dengan teks referensi yang telah tersedia dalam pasangan data. Perbandingan ini bertujuan untuk mengukur sejauh mana hasil transformasi mendekati teks target yang diharapkan. Metrik yang digunakan meliputi *BLEU*, *ROUGE-1*, *ROUGE-2*, *ROUGE-L*, dan *METEOR*. Nilai metrik dihitung berdasarkan tingkat kesamaan antara teks hasil transformasi dengan teks referensi pada data uji. Evaluasi ini digunakan untuk menilai kualitas hasil transformasi dalam menghasilkan kalimat yang lebih santun sesuai dengan target kesantunan.

3.8.3 Evaluasi Manual

Evaluasi manual dilakukan untuk menilai kualitas hasil sistem secara kontekstual, khususnya dalam aspek kesantunan berbahasa dan kesesuaian makna yang tidak sepenuhnya dapat diukur melalui metrik otomatis. Proses evaluasi ini melibatkan seorang evaluator yang merupakan dosen di bidang psikologi. Sampel yang digunakan dalam evaluasi manual berjumlah 100 data uji yang diperoleh dari 10 responden. Setiap responden diminta untuk memberikan contoh teks komunikasi akademik, yang kemudian diproses oleh sistem untuk dilakukan klasifikasi dan transformasi. Hasil keluaran sistem selanjutnya digunakan sebagai bahan evaluasi manual.

Evaluator meninjau pasangan teks yang terdiri dari kalimat asli (input) dan hasil transformasi (output). Penilaian dilakukan secara deskriptif dengan mengacu pada indikator kesantunan berbahasa, yaitu: (1) penggunaan salam pembuka, (2) kejelasan penyampaian tujuan atau informasi, (3) penggunaan bahasa yang baku dan sopan, (4) penggunaan ungkapan permohonan yang santun, (5) kelancaran atau keterbacaan kalimat, serta (6) tidak menetapkan waktu atau agenda secara sepihak kepada dosen. Berdasarkan indikator tersebut, evaluator menilai apakah hasil transformasi telah meningkatkan tingkat kesantunan tanpa mengubah maksud utama pesan, dengan tetap mempertimbangkan konteks komunikasi akademik antara mahasiswa dan dosen. Hasil penilaian kemudian dikategorikan ke dalam dua label, yaitu sopan dan tidak sopan.

Selanjutnya, hasil evaluasi manual dibandingkan dengan label yang dihasilkan oleh sistem untuk mengukur tingkat kesesuaian antara keduanya. Berdasarkan perbandingan tersebut, setiap data diklasifikasikan ke dalam kategori *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN), yang selanjutnya digunakan untuk menghitung akurasi sistem. Dengan demikian, evaluasi manual dalam penelitian ini tidak hanya berfungsi sebagai pelengkap evaluasi otomatis, tetapi juga sebagai validasi kontekstual terhadap kualitas sistem dalam menghasilkan transformasi teks yang sesuai dengan norma komunikasi akademik.

3.9 Deteksi Konteks Akademik

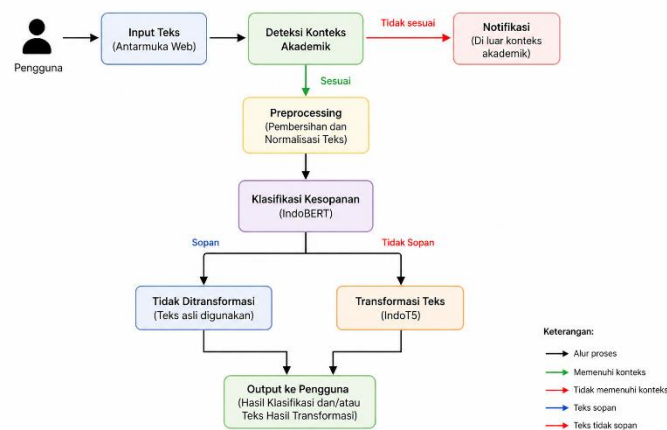
Deteksi konteks akademik dilakukan sebagai tahap awal dalam sistem untuk menyaring teks yang akan diproses. Pada tahap ini, sistem menggunakan pendekatan berbasis aturan *rule-based* dengan memanfaatkan kumpulan kata kunci yang berkaitan dengan konteks komunikasi akademik. Pendekatan *rule-based* dengan *keyword matching* digunakan dalam tahap deteksi

konteks akademik karena memiliki keunggulan dalam hal kesederhanaan implementasi dan efisiensi komputasi. Metode ini memungkinkan sistem untuk dengan cepat mengidentifikasi teks yang relevan berdasarkan keberadaan kata kunci tertentu tanpa memerlukan proses pelatihan model tambahan.

Namun demikian, pendekatan ini memiliki keterbatasan, yaitu tidak mampu memahami konteks secara implisit serta bergantung pada keberadaan kata kunci secara eksplisit dalam teks. Akibatnya, teks yang memiliki makna akademik tetapi tidak mengandung kata kunci tertentu berpotensi tidak terdeteksi. Oleh karena itu, metode ini digunakan sebagai tahap awal penyaringan dengan mempertimbangkan keseimbangan antara efisiensi dan akurasi. Kata kunci yang digunakan dikelompokkan ke dalam beberapa kategori seperti dosen (misalnya “Pak”, “Bu”, dan ”dosen”), mahasiswa (misalnya “mahasiswa”, “saya mahasiswa”), aktivitas akademik (seperti “tugas”, “ujian”, ”skripsi”, dan “kuliah”), serta kata yang berkaitan dengan interaksi akademik (seperti “konsultasi”, “izin”, dan “revisi”). Proses deteksi dilakukan dengan memeriksa apakah teks yang dimasukkan mengandung satu atau lebih kata kunci dari daftar tersebut. Jika teks mengandung kata kunci yang relevan, maka teks dianggap berada dalam konteks akademik. Sebaliknya, jika tidak ditemukan kata kunci yang sesuai maka teks dikategorikan sebagai di luar konteks akademik. Apabila teks tidak memenuhi kriteria konteks akademik, sistem akan menghentikan proses dan menampilkan notifikasi kepada pengguna.

3.10 Integrasi Sistem

Tahap akhir dalam penelitian ini adalah integrasi model klasifikasi dan model transformasi teks ke dalam sebuah sistem berbasis web. Sistem dikembangkan menggunakan bahasa pemrograman Python pada sisi *backend* dengan *framework* FastAPI, sedangkan pada sisi *frontend* digunakan HTML, CSS (Bootstrap), dan JavaScript. *Framework* FastAPI dipilih karena memiliki performa yang ringan, mendukung komunikasi berbasis API secara *asynchronous*, serta mudah diintegrasikan dengan model *machine learning* berbasis Python. Selain itu, FastAPI juga mempermudah proses pengiriman data antara *frontend* dan *backend* dalam format JSON sehingga proses inferensi model dapat dilakukan secara lebih efisien. Model IndoBERT dan IndoT5 dimuat pada server sehingga proses klasifikasi dan transformasi dapat dilakukan secara langsung. Alur kerja sistem disajikan pada Gambar 3.2.



Gambar 3.2 Alur Kerja Sistem

Alur kerja sistem dimulai dari pengguna yang memasukkan teks melalui antarmuka web. Teks kemudian dikirim ke *backend* dalam format JSON untuk diproses lebih lanjut. Tahap pertama yang dilakukan adalah deteksi konteks akademik menggunakan pendekatan berbasis aturan. Jika teks tidak termasuk ke dalam konteks akademik, maka sistem akan menghentikan proses dan menampilkan notifikasi kepada pengguna. Jika teks memenuhi konteks akademik, maka teks akan diproses melalui tahap *preprocessing* sebelum dilakukan klasifikasi menggunakan model IndoBERT. Pada tahap integrasi sistem, teks yang telah diproses melalui model klasifikasi IndoBERT akan ditentukan tingkat kesantunannya. Apabila teks diklasifikasikan sebagai “sopan”, maka sistem tidak akan melakukan proses transformasi dan langsung menampilkan hasil klasifikasi kepada pengguna. Sebaliknya, apabila teks diklasifikasikan sebagai “tidak sopan”, maka teks tersebut akan diproses lebih lanjut menggunakan model IndoT5 untuk ditransformasikan menjadi kalimat yang lebih santun. Pendekatan ini bertujuan untuk menjaga keaslian teks yang sudah sesuai dengan norma komunikasi akademik serta menghindari perubahan yang tidak diperlukan tanpa mengubah makna. Hasil akhir dari proses tersebut kemudian dikirim kembali ke antarmuka pengguna untuk ditampilkan sebagai output. Dengan alur ini, sistem mampu bekerja secara terintegrasi mulai dari deteksi konteks, klasifikasi hingga transformasi teks dalam satu pipeline sistem yang terintegrasi.

BAB IV

HASIL DAN PEMBAHASAN

4.1 Pengumpulan Data

Data yang digunakan dalam penelitian ini berjumlah 6.230 data teks komunikasi akademik antara mahasiswa-dosen dalam bahasa Indonesia. Pengumpulan data dilakukan dari tiga sumber utama, yaitu Google Form, TikTok, dan ChatGPT, dengan tujuan memperoleh variasi gaya bahasa dan konteks komunikasi akademik yang beragam. Distribusi data berdasarkan sumber pengumpulan data di tunjukkan pada Tabel 4.1.

Tabel 4.1 Distribusi Sumber Data

Sumber Data	Data Asli	Jumlah Data Setelah di Kembangkan
Google Form	106	2.220
TikTok	90	2.008
ChatGPT	2.002	2.002
Total	6.230	

Berdasarkan Tabel 4.1, data dari Google Form memiliki jumlah yang paling besar, yang menunjukkan bahwa sebagian *dataset* berasal dari komunikasi nyata mahasiswa kepada dosen sehingga mencerminkan penggunaan bahasa yang autentik dalam konteks akademik. Sementara itu, data dari TikTok dan ChatGPT memiliki jumlah yang relatif seimbang, yang berkontribusi dalam memperkaya variasi struktur kalimat dan gaya bahasa. Komposisi data yang relatif seimbang antar sumber memberikan keuntungan dalam meningkatkan keberagaman *dataset*. Namun demikian, perbedaan karakteristik bahasa dari masing-masing sumber berpotensi menimbulkan variasi gaya bahasa, di mana data dari media sosial cenderung lebih informal dibandingkan dengan data dari Google Form. Oleh karena itu, variasi ini dapat mempengaruhi proses pembelajaran model dalam mengenali pola kesantunan bahasa. Selain itu, sesuai dengan metodologi pada Bab 3, sebagian data dari Google Form dan TikTok telah dikembangkan menjadi beberapa variasi kalimat untuk meningkatkan jumlah dan keberagaman data, sehingga model dapat mempelajari pola bahasa yang lebih luas. Proses pengembangan data dilakukan dengan membuat beberapa variasi kalimat dari data asli tanpa mengubah konteks dan makna utama pesan. Variasi tersebut meliputi perubahan struktur kalimat, pilihan kata, serta bentuk penyampaian pesan yang tetap merepresentasikan konteks komunikasi akademik mahasiswa-dosen.

Pengembangan data dilakukan untuk meningkatkan jumlah dataset dan memperkaya variasi pola bahasa sehingga model dapat mempelajari bentuk komunikasi yang lebih beragam. Variasi sumber data yang digunakan diharapkan dapat merepresentasikan beragam gaya komunikasi mahasiswa dengan latar belakang budaya dan kebiasaan berbahasa yang berbeda, sehingga dataset tidak berfokus pada karakteristik komunikasi dari satu wilayah tertentu saja. Meskipun demikian, variasi budaya dan konteks komunikasi yang kompleks tetap berpotensi memengaruhi interpretasi kesantunan oleh model. Sementara itu, data dari ChatGPT tidak melalui proses pengembangan tambahan karena data yang dihasilkan telah memiliki variasi struktur kalimat yang cukup beragam. Meskipun demikian, pengembangan data tetap dilakukan secara terkontrol untuk menghindari duplikasi makna yang berlebihan. Selanjutnya, data yang dikumpulkan mencakup berbagai konteks komunikasi akademik, seperti konsultasi skripsi, bimbingan akademik, permohonan izin kuliah, serta komunikasi administratif lainnya. Distribusi data berdasarkan jenis komunikasi akademik ditunjukkan pada Tabel 4.2.

Tabel 4.2 Distribusi Jenis Komunikasi

Jenis Komunikasi	Jumlah Data
Bimbingan akademik	2.125
Konsultasi Skripsi	1.536
Administrasi / Prodi	1.288
Lainnya	600
Permohonan izin kuliah	482
Konfirmasi Asisten	199
Total	6.230

Berdasarkan Tabel 4.2, data dalam penelitian ini didominasi oleh kategori bimbingan akademik dan konsultasi skripsi. Hal ini menunjukkan bahwa sebagian besar komunikasi mahasiswa-dosen berkaitan dengan proses pembelajaran dan penyelesaian studi. Dominasi pada jenis komunikasi tertentu menunjukkan bahwa *dataset* memiliki fokus yang kuat pada konteks akademik utama, namun juga berpotensi menyebabkan ketidakseimbangan distribusi data. Ketimpangan ini dapat mempengaruhi kemampuan model dalam memahami konteks komunikasi yang jumlah datanya lebih sedikit, seperti konfirmasi asisten atau permohonan izin. Meskipun demikian, keberadaan berbagai jenis komunikasi menunjukkan bahwa *dataset*

memiliki cakupan konteks yang cukup beragam, sehingga tetap mampu merepresentasikan variasi komunikasi akademik mahasiswa-dosen secara umum.

4.2 Pelabelan Data

Setelah proses pengumpulan data selesai dilakukan, seluruh data kemudian melalui tahapan pelabelan untuk mengidentifikasi tingkat kesantunan pada setiap teks. Setiap teks diklasifikasikan ke dalam dua kategori, yaitu teks sopan (label 0) dan tidak sopan (label 1). *Dataset* hasil pelabelan kemudian digunakan sebagai dasar dalam proses pemodelan. Distribusi jumlah data pada masing-masing kategori kesantunan ditunjukkan pada Tabel 4.3.

Tabel 4.3 Distribusi *Dataset* Berdasarkan Kategori Kesantunan

Label	Kategori	Jumlah Data
0	Sopan	3.115
1	Tidak Sopan	3.115
Total		6.230

Berdasarkan Tabel 4.3, terlihat bahwa *dataset* memiliki distribusi yang seimbang antara kelas sopan dan tidak sopan. Komposisi yang seimbang ini menunjukkan bahwa *dataset* memiliki representasi yang baik terhadap kedua kelas, sehingga dapat membantu model dalam mempelajari pola kesantunan secara optimal tanpa cenderung bias terhadap salah satu kategori. Keseimbangan jumlah data antar kelas juga berpotensi meningkatkan performa model klasifikasi, khususnya dalam mengurangi risiko bias prediksi yang sering terjadi pada *dataset* yang tidak seimbang. Dengan distribusi yang seimbang, model memiliki peluang yang sama untuk mempelajari karakteristik masing-masing kelas. Namun demikian, meskipun jumlah data seimbang, perbedaan gaya bahasa antar sumber data tetap dapat mempengaruhi pola pembelajaran model. Data yang berasal dari media sosial cenderung memiliki gaya bahasa yang lebih informal dibandingkan dengan data dari Google Form, sehingga berpotensi mempengaruhi karakteristik teks dalam masing-masing kategori kesantunan. Contoh *dataset* dapat dilihat pada Tabel 4.4.

Tabel 4.4 Contoh *Dataset* Klasifikasi Kesopanan

Teks	Label
Assalamualaikum, Bu. Saya mau bimbingan. Bisa kapan ya?	Tidak Sopan
Pak, saya udah kirim bab 3, tolong dicek ya pak	Tidak Sopan
Bu, sistem KRS error terus, gimana ini, bisa ke kampus aja gak?	Tidak Sopan
Mohon izin Ibu, saya ingin memastikan kehadiran asisten laboratorium untuk sesi besok.	Sopan
Ibu, saya ingin menanyakan apakah nilai praktikum sudah diinput ke sisitem akademik.	Sopan

Berdasarkan Tabel 4.4, perbedaan utama antara kedua kategori terletak pada tingkat formalitas bahasa dan struktur penyampaian pesan. Teks yang termasuk dalam kategori sopan cenderung menggunakan kalimat yang lebih lengkap tersebut, bahasa yang baku, serta memperhatikan etika komunikasi akademik. Sebaliknya, teks tidak sopan cenderung menggunakan bahasa yang lebih informal, struktur kalimat yang lebih sederhana, serta penyampaian yang lebih langsung. Selain digunakan untuk klasifikasi, *dataset* dalam penelitian ini juga memuat pasangan kalimat untuk pelatihan model transformasi teks menggunakan IndoT5. Contoh pasangan data ditunjukkan pada Tabel 4.5.

Tabel 4.5 Contoh *Dataset* Pasangan Kalimat

Teks	Label	Target
Pak, tanda tangan presensi ujian harus ke kelas nggak sih, kan ujiannya take home?	1	Assalamualaikum, Pak. Mohon izin bertanya, apakah tanda tangan presensi ujian perlu dilakukan di kelas, mengingat ujiannya bersifat take home?
Bu, mau tanya apakah nilai ujiannya dapat dilihat? Kalau iya lihat di mana ya, Bu	1	Assalamualaikum, Bu. Mohon izin bertanya, apakah nilai ujian sudah dapat dilihat? Jika iya, di mana saya bisa melihatnya?
Pak, mohon izin, apakah Bapak dapat menandatangani surat pengantar magang saya hari ini	0	Pak, saya udah kirim surat magangnya, tolong ditandatangani ya.
Mohon izin, Bu. Saya sudah menyusun draft artikel, apakah Ibu berkenan meninjau dan memberi masukan?	0	Bu, saya udah kirim draft, coba dibaca dulu ya.

Pada Tabel 4.5, data dengan label tidak sopan, pasangan yang digunakan merupakan versi kalimat yang telah ditransformasikan menjadi lebih santun. Sementara itu, pada data dengan label sopan, pasangan yang digunakan merupakan versi kalimat yang telah

ditransformasikan menjadi lebih santun. Sementara itu, pada data dengan label sopan, pasangan yang digunakan merupakan variasi kalimat kurang sopan dengan makna yang serupa. Penyusunan pasangan ini bertujuan agar model mampu mempelajari pola perubahan bahasa dari tidak sopan menjadi lebih sopan tanpa mengubah makna utama dari pesan yang disampaikan. Dengan demikian, hasil pelabelan tidak hanya berfungsi sebagai dasar klasifikasi, tetapi juga sebagai fondasi dalam pembentukan pasangan data untuk proses transformasi teks.

4.3 Preprocessing Data

Pada tahap ini dilakukan *preprocessing* untuk membersihkan dan menyiapkan data teks sebelum digunakan dalam proses pemodelan. Data yang diperoleh dari berbagai sumber masih mengandung ketidakkonsistenan, seperti penggunaan huruf kapital yang beragam, keberadaan URL, serta penggunaan bahasa informal. *Preprocessing* dilakukan melalui beberapa tahapan, yaitu pembersihan data, normalisasi teks, serta penghapusan elemen yang tidak relevan. Proses ini bertujuan untuk mengurangi *noise* dan meningkatkan konsistensi data sehingga model dapat lebih mudah mengenali pola bahasa. Dampak dari proses *preprocessing* terhadap *dataset* ditunjukkan pada, yang memperlihatkan perubahan jumlah data sebelum dan sesudah *preprocessing*. Selain itu, contoh perubahan teks sebelum dan sesudah *preprocessing* di tunjukkan pada Tabel 4.13.

4.3.1 Persiapan Data

Tahap pertama dalam *preprocessing* adalah persiapan data atau pembersihan data, yaitu dengan menghapus data yang tidak valid atau tidak lengkap. Pada tahap ini, dilakukan pengecekan terhadap data yang memiliki nilai kosong (*missing value*), khususnya pada kolom teks dan label. Data yang memiliki nilai kosong pada salah satu kolom akan dihapus karena tidak dapat digunakan dalam proses pelatihan model dan berpotensi menyebabkan *error* saat proses *training*. Selain itu, dilakukan konversi tipe data pada kolom label menjadi bentuk *integer* untuk memastikan kesesuaian format data dengan kebutuhan model klasifikasi. Hasil dari proses ini menunjukkan bahwa *dataset* yang digunakan hanya terdiri dari data yang lengkap dan valid. Dengan demikian, kualitas data menjadi lebih baik serta mengurangi potensi gangguan (*noise*) dalam proses pembelajaran model. Detail implementasi pembersihan data dapat dilihat pada Lampiran E.

4.3.2 Case Folding

Setelah proses pembersihan data dilakukan, tahap selanjutnya adalah *case folding*. Tujuan dari proses ini adalah untuk menghindari perbedaan representasi kata yang disebabkan oleh penggunaan huruf kapital. Dengan adanya *case folding*, kata yang memiliki makna sama tetapi ditulis dengan variasi huruf besar dapat dianggap sebagai satu kesatuan oleh model. Misalnya, “Pak” dan “pak” akan diubah menjadi “pak”. Hasil dari proses ini menunjukkan bahwa seluruh teks menjadi konsisten dalam penulisan, sehingga membantu model dalam mengenali pola kata tanpa dipengaruhi oleh perbedaan huruf kapital. Detail implementasi *case folding* dapat dilihat pada Lampiran E. Hasil dari *case folding* dapat dilihat dari Tabel 4.6 berikut.

Tabel 4.6 Hasil *Case Folding*

Sebelum <i>Case Folding</i>	Sesudah <i>Case Folding</i>
Bu, saya udah ngirim tugasnya, cepet diperiksa ya	bu, saya udah ngirim tugasnya cepet diperiksa ya
Pak, saya lupa kirim tugas kemarin, baru inget sekarang	pak, saya lupa kirim tugas kemarin, baru inget sekarang

4.3.3 Penghapusan URL

Tahap berikutnya adalah penghapusan URL yang terdapat dalam teks. URL biasanya muncul dalam bentuk tautan yang mengarah ke halaman tertentu di internet, seperti *link* berbagi file, referensi materi, atau sumber informasi lainnya. Dalam konteks komunikasi akademik antara mahasiswa dan dosen, keberadaan URL umumnya tidak memberikan kontribusi langsung terhadap makna utama dari pesan yang disampaikan. Selain itu, keberadaan URL dianggap *noise* karena memiliki struktur karakter yang berbeda dengan teks biasa dan tidak relevan dalam proses analisis. Jika tidak dihapus, maka proses tokenisasi dapat terganggu dan berpotensi menurunkan kualitas representasi teks pada model. Oleh karena itu, pada tahap ini dilakukan penghapusan seluruh pola URL yang terdapat dalam teks menggunakan pendekatan *regular expression*. Hasil dari proses ini menunjukkan bahwa teks menjadi lebih bersih dan hanya mengandung informasi yang relevan, sehingga membantu model dalam memahami isi pesan secara optimal. Detail implementasi penghapusan URL dapat dilihat pada Lampiran E. Hasil dari penghapusan URL dapat dilihat di Tabel 4.7.

Tabel 4.7 Contoh Teks Hasil Penghapusan URL

Sebelum penghapusan URL	Sesudah Penghapusan URL
Pak. Ini linknya untuk pengumpulan tugas saya https://bit.ly/abc	pak, ini linknya untuk pengumpulan tugas saya

4.3.4 Penghapusan *Mention*

Tahap berikutnya adalah penghapusan *mention* yang biasanya ditandai dengan simbol “@” diikuti nama pengguna. *Mention* umumnya digunakan dalam media sosial untuk menandai atau menyebut pengguna tertentu, namun dalam konteks komunikasi akademik, elemen ini tidak memiliki pengaruh terhadap makna utama pesan. Keberadaan *mention* dalam teks dapat menambah variasi yang tidak diperlukan serta berpotensi mengganggu proses analisis teks oleh model. Oleh karena itu, *mention* dihapus untuk menjaga konsistensi dan kebersihan data. Hasil dari proses ini menunjukkan bahwa teks menjadi lebih bersih dan terfokus pada isi pesan utama, sehingga membantu model dalam memahami konteks komunikasi secara optimal. Detail implementasi penghapusan *mention* dapat dilihat pada Lampiran E. Hasil penghapusan *mention* dapat dilihat pada Tabel 4.8.

Tabel 4.8 Contoh Perbandingan Teks Sebelum dan Sesudah Penghapusan *Mention*

Sebelum Penghapusan <i>Mention</i>	Sesudah Penghapusan <i>Mention</i>
@Bambang Susanto Pak, saya sudah kirim tugas nya ya di <i>email</i>	pak, saya sudah kirim tugas nya ya di <i>email</i>

4.3.5 Penghapusan *Hashtag*

Hashtag merupakan kata atau frasa yang diawali dengan simbol “#” dan umumnya digunakan dalam media sosial untuk mengelompokkan topik tertentu. Dalam penelitian ini, *hashtag* tidak memiliki peranan penting dalam menentukan makna kesantunan dalam teks. Namun, berbeda dengan *mention* kata yang terdapat dalam *hashtag* masih memiliki makna yang relevan. Oleh karena itu, proses yang dilakukan dengan menghilangkan simbol “#” dan mempertahankan kata yang mengikutinya. Pendekatan ini bertujuan agar tetap dipertahankan tanpa menyertakan simbol yang diperlukan. Hasil dari proses ini menunjukkan bahwa teks menjadi lebih bersih tanpa kehilangan makna utama, sehingga model dalam memahami konteks secara lebih optimal. Detail implementasi penghapusan *hashtag* dapat dilihat pada Lampiran E. Hasil penghapusan *hashtag* dapat dilihat pada Tabel 4.9.

Tabel 4.9 Contoh Teks Perbandingan Sebelum dan Sesudah Penghapusan *Hashtag*

Sebelum penghapusan <i>Hashtag</i>	Sesudah Penghapusan <i>Hashtag</i>
#bimbingan Bu, saya mau bimbingan hari senin besok	bu, saya mau bimbingan hari senin besok

4.3.6 Penghapusan Spasi Berlebih

Tahap selanjutnya adalah penghapusan spasi berlebih yang terdapat dalam teks. Spasi yang tidak konsisten sering muncul akibat penulisan yang tidak tersruktur, seperti penggunaan spasi ganda dalam satu kalimat. Spasi berlebih dapat mempengaruhi proses tokenisasi serta menyebabkan ketidakkonsistenan dalam representasi teks. Oleh karena itu, dilakukan normalisasi spasi dengan mengubah semua spasi berlebih menjadi satu spasi serta menghapus spasi di awal dan di akhir teks. Hasil dari proses ini menunjukkan bahwa struktur teks menjadi lebih rapi dan konsisten, sehingga meningkatkan kualitas representasi data pada model. Detail implementasi penghapusan spasi berlebih dapat dilihat pada Lampiran E. Hasil penghapusan spasi berlebih dapat dilihat pada Tabel 4.10.

Tabel 4.10 Contoh Teks Sebelum dan Sesudah Penghapusan Spasi Berlebih

Sebelum Penghapusan Spasi Berlebih	Sesudah Penghapusan Spasi Berlebih
Bu, apakah saya bisa bimbingan besok	bu, apakah saya bisa bimbingan besok

4.3.7 Penghapusan Karakter Khusus

Tahap berikutnya adalah penghapusan karakter khusus yang tidak relevan hal ini berfungsi untuk mengurangi *noise* dalam data dengan hanya mempertahankan karakter yang dianggap penting, seperti huruf, angka, serta tanda baca dasar. Penghapusan karakter khusus meliputi simbol, emoji, atau tanda lain yang tidak termasuk dalam huruf maupun angka. Hasil dari proses ini menunjukkan bahwa teks menjadi lebih bersih dan terstandarisasi, sehingga membantu model dalam memahami pola bahasa secara lebih konsisten. Detail implementasi penghapusan karakter khusus berlebih dapat dilihat pada Lampiran E. Hasil penghapusan karakter khusus berlebih dapat dilihat pada Tabel 4.11

Tabel 4.11 Contoh Teks Sebelum dan Sesudah Penghapusan Karakter Khusus

Sebelum Penghapusan Karakter Khusus	Sesudah Penghapusan Karakter Khusus
Bu, maaf banget saya salah kirim tugasnya 🙏🙏	bu, maaf banget saya salah kirim tugasnya

4.3.8 Penghapusan Kalimat yang Kurang dari Lima Karakter

Tahapan terakhir dalam *preprocessing* adalah penghapusan kalimat yang memiliki panjang kurang dari lima karakter. Kalimat dengan panjang yang sangat pendek umumnya tidak memiliki informasi yang cukup untuk merepresentasikan makna dalam konteks komunikasi akademik. Keberadaan data yang seperti ini dapat dianggap *noise* dan berpotensi menurunkan performa model dalam memahami pola bahasa. Hasil dari proses ini menunjukkan bahwa data yang tidak informatif berhasil dihilangkan, sehingga *dataset* menjadi lebih berkualitas dan representatif terhadap konteks komunikasi akademik. Detail implementasi penghapusan kalimat yang kurang dari lima karakter dapat dilihat pada Lampiran E. Hasil penghapusan kalimat pendek dapat dilihat pada Tabel 4.12.

Tabel 4.12 Contoh Teks Sebelum dan Sesudah Penghapusan Kalimat Pendek

Sebelum Penghapusan Kalimat Pendek	Sesudah Penghapusan Kalimat Pendek
Bu, saya izin tidak masuk hari ini	bu, saya izin tidak masuk hari ini
Pak, saya gajadi bimbingan	(dihapus)

Berdasarkan keseluruhan proses *preprocessing* yang telah dilakukan, jumlah data mengalami penurunan dari 6.230 menjadi 6.221 data, sebagaimana ditunjukkan pada Tabel 4.13.

Tabel 4.13 Perbandingan *Dataset* Sesudah dan Sebelum *Preprocessing*

Tahapan	Total Data	Sopan	Tidak Sopan
Data Awal	6.230	3.115	3.115
Setelah Persiapan dan <i>Preprocessing</i>	6.221	3.107	3.114

Berdasarkan Tabel 4.13, penurunan jumlah data ini menunjukkan bahwa proses *preprocessing* berhasil menghapus data yang tidak valid dan informatif, seperti data kosong, simbol, serta kalimat yang terlalu pendek. Meskipun terjadi pengurangan jumlah data, distribusi antar kelas tetap relatif seimbang, sehingga tidak menimbulkan bias yang signifikan dalam proses pelatihan model. Contoh hasil *preprocessing* secara keseluruhan ditunjukkan pada Tabel 4.14.

Tabel 4.14 Contoh Teks Sebelum dan Sesudah *Preprocessing*

Kalimat Asli	Kalimat Setelah Preprocessing
Bu, bagaimana cara daftar sidang skripsi?	bu bagaimana cara daftar sidang skripsi
Pak, saya mau izin telat datang bimbingan dikarenakan jalanan macet banget 🚗	pak saya mau izin telat datang bimbingan dikarenakan jalanan macet banget

Berdasarkan Tabel 4.14, *preprocessing* berhasil menghasilkan teks yang lebih bersih, konsisten, dan mudah dipahami oleh model. Hal ini berpotensi meningkatkan kemampuan model dalam mengenali pola bahasa pada tahap klasifikasi dan transformasi teks.

4.4 Hasil Klasifikasi Kesopanan Menggunakan IndoBERT

Model IndoBERT merupakan model berbasis *transformer* yang telah dilatih sebelumnya pada korpus bahasa Indonesia, sehingga memiliki kemampuan dalam memahami konteks bahasa secara lebih mendalam. Pada penelitian ini, model IndoBERT digunakan untuk mengklasifikasikan teks komunikasi akademik mahasiswa-dosen ke dalam dua kategori, yaitu sopan dan tidak sopan. Penerapan model ini bertujuan untuk mengidentifikasi tingkat kesantunan secara otomatis, sehingga dapat membantu dalam memahami pola komunikasi akademik yang digunakan oleh mahasiswa.

4.4.1 Data Split

Sebelum proses pelatihan model dilakukan, dataset dibagi menjadi tiga bagian utama, yaitu data *training*, *validation*, dan *testing*. Pembagian ini bertujuan untuk memastikan bahwa model tidak hanya mampu mempelajari data latih, tetapi juga mampu melakukan generalisasi terhadap data baru yang belum pernah dilihat sebelumnya. Proses pembagian data dilakukan dalam dua tahap dengan pendekatan *stratified sampling* untuk menjaga keseimbangan distribusi kelas pada subset data. Pada tahap pertama, sebanyak 30% data dipisahkan sebagai data sementara (*temporary data*), sedangkan 70% sisanya digunakan sebagai data *training*. Selanjutnya, data sementara tersebut dibagi kembali menjadi data *validation* dan data *testing* dengan proporsi yang sama, sehingga diperoleh pembagian akhir sebesar 70% untuk data *training*, 15% untuk *validation*, dan 15% untuk *testing*. Pendekatan ini memastikan bahwa distribusi kelas sopan dan tidak sopan tetap seimbang pada setiap subset, sehingga dapat mengurangi potensi bias model selama proses pelatihan dan evaluasi. Kode program ditunjukkan pada Gambar 4.1.

```

train_texts, temp_texts, train_labels, temp_labels = train_test_split(
    df_valid['cleaned_text'].tolist(),
    df_valid['label'].tolist(),
    test_size=0.3,
    random_state=CONFIG['random_state'],
    stratify=df_valid['label']
)

val_texts, test_texts, val_labels, test_labels = train_test_split(
    temp_texts, temp_labels,
    test_size=0.5,
    random_state=CONFIG['random_state'],
    stratify=temp_labels
)

```

Gambar 4.1 Kode Program *Data Split*

Berikut Tabel 4.15, adalah distribusi dataset yang di bagi menjadi tiga bagian utama, yaitu data *training* sebanyak 4.354 data, data *validation* sebanyak 933 data, dan data *testing* sebanyak 934 data. Data *training* digunakan untuk melatih model dalam mempelajari pola kesopanan bahasa, sedangkan data *validation* digunakan untuk memantau performa model selama proses pelatihan serta mencegah terjadinya *overfitting*. Sementara itu, data *testing* digunakan sebagai data evaluasi akhir untuk mengukur performa model secara objektif.

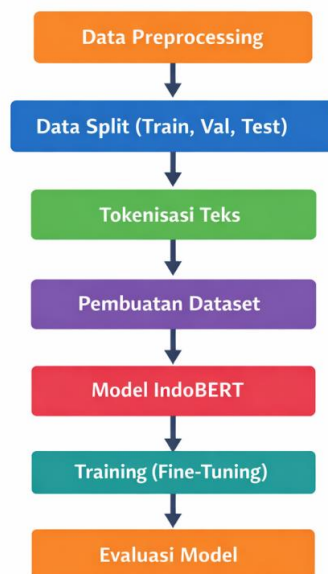
Tabel 4.15 Distribusi *Dataset*

Dataset	Jumlah Data
Training	4.354
Validation	933
Testing	934
Total	6.221

Berdasarkan Tabel 4.15, jumlah data *training* yang lebih besar memungkinkan model mempelajari pola kesantunan secara lebih optimal, sementara data *validation* dan *testing* cukup representatif untuk mengevaluasi performa dan *testing* tetap cukup representatif untuk mengevaluasi performansi model. Distribusi yang seimbang antar kelas pada setiap subset juga menunjukkan bahwa proses pembagian data telah dilakukan dengan baik dan tidak menimbulkan bias yang signifikan.

4.4.2 Proses Pelatihan Model IndoBERT

Pada tahap ini dilakukan proses pelatihan model IndoBERT untuk membangun model yang mampu memahami pola bahasa dalam komunikasi akademik serta mengklasifikasikan teks ke dalam kategori sopan dan tidak sopan secara otomatis.



Gambar 4.2 Alur Proses Pelatihan IndoBERT

Berdasarkan Gambar 4.2, alur proses pelatihan model dimulai dari data yang telah melalui tahapan *preprocessing*, kemudian dibagi menjadi data *training*, *validation*, dan *testing*. Selanjutnya, teks diubah menjadi token menggunakan *tokenizer* IndoBERT agar dapat diproses dalam bentuk numerik. Data yang telah ditokenisasi kemudian dikonversi menjadi format tensor sesuai kebutuhan model berbasis *deep learning*. Tahap berikutnya adalah inisialisasi model IndoBERT sebagai *pretrained model*, yang kemudian dilakukan proses *fine-tuning* menggunakan *dataset* yang telah disiapkan. Proses pelatihan dilakukan secara iteratif selama beberapa *epoch* hingga model mampu mengenali pola kesantunan dalam data. Kode program pada Gambar 4.3.

```

bert_model = BertForSequenceClassification.from_pretrained(
    CONFIG['bert']['model_name'],
    num_labels=2,
)
  
```

Gambar 4.3 Kode Program Inisialisasi Model

Model indoBERT digunakan sebagai model dasar yang telah memiliki pemahaman awal terhadap bahasa Indonesia. Parameter *num_labels=2* menunjukkan bahwa model akan melakukan klasifikasi ke dalam dua kategori, yaitu sopan dan tidak sopan.

```
training_args = TrainingArguments(  
    num_train_epochs=3,  
    learning_rate=2e-5,  
    per_device_train_batch_size=8  
)
```

Gambar 4.4 Kode Program Pengaturan *Training*

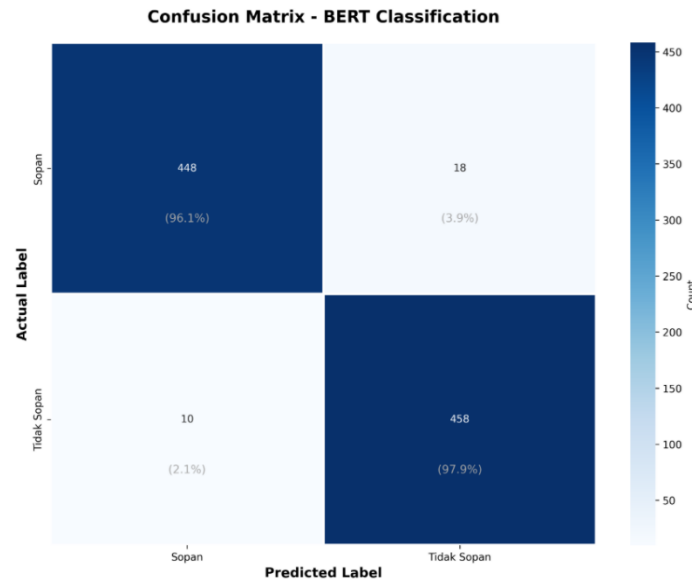
Pada Gambar 4.4, digunakan parameter pelatihan yang digunakan meliputi jumlah *epoch*, *learning rate*, dan *batch size*. Pemilihan *learning rate* bertujuan untuk menjaga kestabilan proses *fine-tuning*, sedangkan jumlah *epoch* yang terbatas membantu mengurangi risiko *overfitting*. Berdasarkan proses pelatihan dilakukan, model IndoBERT mampu mempelajari pola kesantunan dalam teks komunikasi akademik dengan baik. Hal ini didukung oleh penggunaan model *pretrained* yang telah memahami struktur bahasa Indonesia, sehingga proses *fine-tuning* menjadi lebih efektif dibandingkan pelatihan dari awal. Selain itu, pemilihan parameter pelatihan yang sesuai berkontribusi terhadap kestabilan proses pembelajaran model, sehingga menghasilkan performa yang tinggi pada tahap evaluasi yang ditunjukkan pada subbab berikutnya.

4.4.3 Hasil Evaluasi Model IndoBERT

Evaluasi dilakukan untuk mengukur performa dalam mengklasifikasikan teks kesopanan. Evaluasi ini dilakukan menggunakan data *testing* yang sebelumnya tidak digunakan dalam proses pelatihan, sehingga hasil yang diperoleh dapat mencerminkan kemampuan model dalam menggeneralisasi terhadap data baru.

Confusion Matrix

Performa model dianalisis menggunakan *confusion matrix* untuk melihat distribusi hasil prediksi secara lebih detail. Gambar *confusion matrix* disajikan pada Gambar 4.5.



Gambar 4.5 *Confusion Matrix*

Berdasarkan Gambar 4.5, *confusion matrix* digunakan untuk mengetahui jumlah prediksi yang benar maupun yang salah yang dilakukan oleh model. *True Positive* (TP) adalah jumlah data tidak sopan yang berhasil diklasifikasikan dengan benar yaitu berjumlah 458. *True Negative* (TN) adalah jumlah data sopan yang berhasil diklasifikasikan dengan benar yaitu berjumlah 448. *False Positive* (FP) adalah jumlah data sopan yang salah diklasifikasikan sebagai tidak sopan yaitu berjumlah 18. *False Negative* (FN) adalah jumlah data tidak sopan yang salah diklasifikasikan sebagai sopan yaitu berjumlah 10. Berdasarkan distribusi tersebut, terlihat bahwa jumlah kesalahan *False Positive* lebih tinggi dibandingkan *False Negative*. Hal ini menunjukkan bahwa model cenderung lebih sensitif dalam mendeteksi teks tidak sopan, sehingga beberapa teks yang sebenarnya sopan masih diklasifikasikan tidak sopan.

Hasil Evaluasi Model IndoBERT

Pada penelitian ini, evaluasi model dilakukan menggunakan beberapa metrik, yaitu *accuracy*, *precision*, *recall*, dan *F1-score*. Keempat metrik ini digunakan untuk memberikan gambaran yang lebih komprehensif terhadap performa model dalam melakukan klasifikasi. Berdasarkan *confusion matrix* tersebut, diperoleh nilai evaluasi model sebagai berikut.

Tabel 4.16 Performa Model IndoBERT

Metrik	Nilai
<i>Accuracy</i>	97,00 %
<i>Precision</i>	96,20 %
<i>Recall</i>	97,87 %
<i>F1-Score</i>	97,02 %

Pada Tabel 4.16, metrik accuracy sebesar 97,00% menunjukkan bahwa model mampu mengklasifikasikan sebagian besar data dengan benar. Dalam konteks penelitian ini, nilai accuracy yang tinggi menunjukkan bahwa model telah mampu memahami pola kesantunan dalam komunikasi akademik mahasiswa-dosen secara umum. Metrik precision sebesar 96,20% menunjukkan bahwa prediksi model terhadap kelas tidak sopan cukup akurat dan memiliki kesalahan yang rendah. Hal ini berarti ketika sistem mendeteksi teks sebagai teks tidak sopan, sebagian besar prediksi tersebut memang benar, sehingga sistem tidak memberikan perbaikan yang tidak diperlukan pada teks yang sebenarnya sudah sopan. Metrik *recall* sebesar 97,87% menunjukkan bahwa model mampu mendeteksi hampir seluruh data tidak sopan dengan baik. Hal ini berarti *recall* yang tinggi sangat penting karena memastikan bahwa sebagian besar teks yang tidak sopan dapat teridentifikasi dan diproses lebih lanjut pada tahap transformasi. Metrik *F1-Score* sebesar 97,02% menunjukkan keseimbangan yang baik antara *precision* dan *recall*. Hal ini mengindikasikan bahwa model tidak hanya akurat, tetapi juga konsisten dalam mendeteksi kedua kelas secara seimbang.

Tingginya nilai pada seluruh metrik tersebut menunjukkan bahwa model IndoBERT mampu mengenali pola kesantunan dalam teks dengan sangat baik. Hal ini dipengaruhi oleh beberapa faktor. Pertama, model IndoBERT merupakan model *pretrained* yang telah dilatih pada korpus bahasa Indonesia dalam jumlah besar, sehingga memiliki pemahaman yang baik terhadap struktur dan konteks bahasa. Kedua, penggunaan klasifikasi biner, yaitu “sopan” dan “tidak sopan”, membuat batas antar kelas menjadi lebih jelas sehingga memudahkan model dalam membedakan karakteristik masing-masing kategori. Selain itu, proses *preprocessing* yang dilakukan, seperti pembersihan teks dan penghapusan data yang tidak relevan, turut meningkatkan kualitas *dataset* sehingga membantu model dalam mempelajari pola yang lebih konsisten. Namun demikian, performa yang sangat tinggi ini juga perlu dianalisis secara kritis. Kemungkinan adanya pola yang relatif sederhana dalam dataset atau distribusi data yang tidak terlalu kompleks dapat membuat model mencapai performa tinggi, tetapi belum tentu

sepenuhnya mencerminkan kemampuan generalisasi pada data yang lebih beragam di dunia nyata. Dengan demikian, meskipun hasil yang diperoleh menunjukkan performa yang sangat baik, evaluasi lebih lanjut dengan data yang lebih variatif tetap diperlukan untuk memastikan robustitas model dalam berbagai konteks komunikasi akademik.

Contoh Hasil Prediksi Model

Untuk memberikan gambaran yang lebih jelas mengenai performa model, berikut diambil beberapa contoh hasil prediksi yang disajikan pada Tabel 4.17.

Tabel 4.17 Contoh Hasil Prediksi Model IndoBERT

Teks	Label Asli	Prediksi
Saya izin gak masuk kelas ya, pak. Sakit	Tidak Sopan	Tidak Sopan
Bu, saya sudah mengirimkan tugas. Mohon dicek	Sopan	Sopan
Pak, saya izin telat karena macet	Tidak Sopan	Sopan

Berdasarkan Tabel 4.17 di atas, terlihat bahwa model mampu mengklasifikasikan sebagian besar data dengan benar. Namun, terdapat beberapa kesalahan prediksi, terutama pada teks yang memiliki tingkat kesopanan ambigu. Contoh kesalahan prediksi dapat dilihat pada kalimat “Pak, saya izin telat karena macet” yang diklasifikasikan sebagai sopan, padahal termasuk kategori tidak sopan. Kesalahan ini terjadi karena kalimat tersebut tidak mengandung kata-kata yang secara eksplisit menunjukkan ketidaksopanan, sehingga model menanggapinya sebagai bentuk komunikasi yang masih dapat diterima. Selain itu, kesalahan *False Positive* umumnya terjadi pada kalimat yang menggunakan struktur sederhana atau bahasa yang kurang formal, sehingga model menganggapnya sebagai tidak sopan, meskipun dalam konteks tertentu masih dapat diterima dalam komunikasi akademik.

Hal ini menunjukkan bahwa model masih memiliki keterbatasan dalam memahami aspek pragmatik dan konteks sosial, khususnya dalam membedakan antara bahasa informal yang masih sopan dan bahasa yang benar-benar tidak sopan. Hasil evaluasi yang diperoleh, model IndoBERT menunjukkan performa yang sangat baik dalam mengklasifikasikan teks kesopanan. Namun demikian, performa yang tinggi tersebut belum sepenuhnya mampu mencerminkan kemampuan model dalam memahami konteks secara mendalam, terutama pada teks yang bersifat ambigu. Secara keseluruhan, model IndoBERT menunjukkan kemampuan yang kuat dalam mengenali pola kesantunan teks, tetapi masih memerlukan peningkatan dalam

aspek pemahaman konteks pragmatik agar dapat mengurangi kesalahan klasifikasi pada kasus-kasus tertentu.

4.5 Hasil Transformasi Kesopanan Menggunakan IndoT5

Model IndoT5 merupakan model berbasis *transformer* yang dirancang untuk tugas *text-to-text generation*, sehingga mampu melakukan tugas pengolahan bahasa termasuk mentransformasikan kalimat dan perbaikan gaya bahasa. Dalam penelitian ini, IndoT5 digunakan untuk mengubah teks yang teridentifikasi tidak sopan menjadi teks yang lebih santun dan sesuai dengan kaidah komunikasi akademik. Proses transformasi teks hanya dilakukan pada teks yang diklasifikasikan sebagai tidak sopan oleh model IndoBERT. Jika teks sudah termasuk kategori sopan, maka sistem tidak akan melakukan perubahan teks yang memerlukan perbaikan. Berikut adalah alur proses transformasi teks dalam penelitian ini pada Gambar 4.6.



Gambar 4.6 Alur Transformasi Teks Pada Model IndoT5

Berdasarkan diagram Gambar 4.6 tersebut, sistem bekerja secara berurutan dengan memanfaatkan dua model utama, yaitu IndoBERT untuk pengklasifikasian teks dan IndoT5 untuk transformasi teks. Pendekatan ini memungkinkan sistem tidak hanya mendeteksi tingkat kesantunan, tetapi juga memberikan solusi berupa transformasi teks secara otomatis.

4.5.1 Implementasi IndoT5

Model IndoT5 digunakan untuk melakukan transformasi teks dari bentuk tidak sopan menjadi lebih sopan. Model yang digunakan merupakan model *pretrained* berbasis *sequence-to-sequence* yang telah dilatih untuk parafrase bahasa Indonesia. Pendekatan ini memungkinkan model menghasilkan teks baru dengan struktur kalimat yang lebih baik serta meningkatkan tingkat kesantunan. Berikut adalah kode program inisialisasi IndoT5 pada Gambar 4.7.

```
t5_tokenizer = T5Tokenizer.from_pretrained(CONFIG['t5']['model_name'])
t5_model = T5ForConditionalGeneration.from_pretrained(
    CONFIG['t5']['model_name']
)
```

Gambar 4.7 Kode Program Inisialisasi IndoT5

Kode pada Gambar 4.7 di atas digunakan untuk memuat model IndoT5 dan *tokenizer* yang sesuai. Model yang digunakan adalah *IndoT5-base-paraphrase*, yang telah dilatih untuk melakukan parafrase teks dalam bahasa Indonesia. Pada gambar di atas terdapat fungsi untuk mempersiapkan data sebelum dimasukkan ke dalam model IndoT5 yaitu *input_text*, teks tidak sopan dan *target_text*, teks sopan hasil transformasi. Kedua teks diubah menjadi token, kemudian label diberikan agar model dapat belajar memetakan teks tidak sopan menjadi teks sopan.

```
t5_training_args = Seq2SeqTrainingArguments(
    num_train_epochs=3,
    learning_rate=3e-4,
    per_device_train_batch_size=4
)
t5_trainer = Seq2SeqTrainer(
    model=t5_model,
    args=t5_training_args,
    train_dataset=train_t5_dataset,
    eval_dataset=val_t5_dataset
)
```

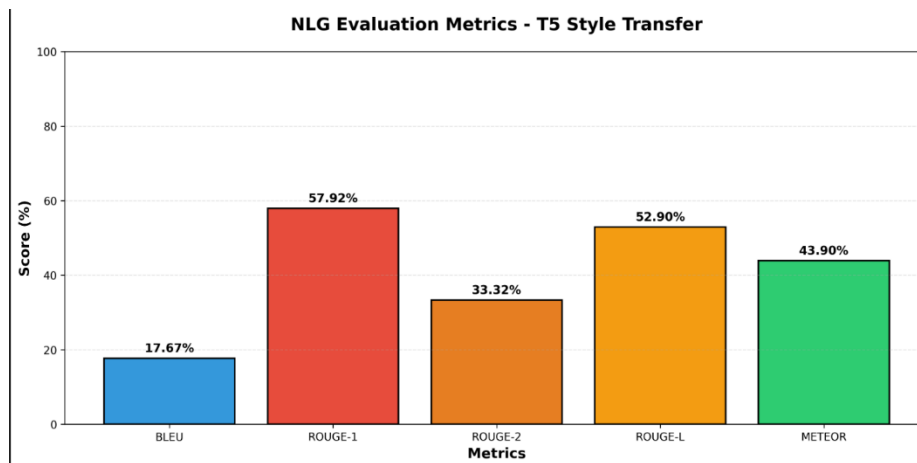
Gambar 4.8 Kode Program *Training* IndoT5

Pada Gambar 4.8, parameter pelatihan digunakan untuk mengatur proses pembelajaran model, seperti jumlah *epoch*, *learning rate*, dan *batch size*. Pengaturan parameter ini bertujuan untuk menjaga kestabilan proses pelatihan serta menghindari *overfitting*. Berdasarkan proses pelatihan yang dilakukan, IndoT5 mampu mempelajari hubungan antara teks tidak sopan dan versi teks yang lebih sopan melalui pendekatan pasangan data (*input-output*). Hal ini menunjukkan bahwa model tidak hanya melakukan perubahan kata, tetapi juga mampu memperbaiki struktur kalimat secara lebih menyeluruh. Penggunaan model *pretrained* memberikan keuntungan dalam memahami struktur bahasa Indonesia, sehingga proses pelatihan menjadi lebih efisien. Selain itu, pendekatan *sequence-to-sequence* memungkinkan model menghasilkan kalimat yang lebih natural sesuai dengan konteks komunikasi akademik. Meskipun demikian, kemampuan model dalam menjaga makna asli teks tetap menjadi tantangan, terutama pada kalimat yang ambigu atau memiliki konteks yang tidak eksplisit. Oleh

karena itu, evaluasi lebih lanjut diperlukan untuk menilai kualitas hasil transformasi secara menyeluruh.

4.5.2 Evaluasi Model IndoT5

Evaluasi model IndoT5 digunakan untuk mengukur kualitas hasil transformasi teks. Berbeda dengan IndoBERT yang dievaluasi menggunakan metrik klasifikasi, IndoT5 sebagai model *text generation* dievaluasi menggunakan metrik *Natural Language Generation* (NLG), yaitu *BLEU*, *ROUGE*, dan *METEOR*. Evaluasi dilakukan dengan membandingkan teks hasil prediksi model dengan teks referensi. Berikut grafik hasil evaluasi NLG IndoT5 pada Gambar 4.9.



Gambar 4.9 Grafik Hasil Evaluasi IndoT5

Tabel 4.18 Hasil Evaluasi Model IndoT5

Metrik	Nilai
<i>BLEU</i>	0,1767
<i>ROUGE-1</i>	0,5792
<i>ROUGE-2</i>	0,3332
<i>ROUGE-L</i>	0,5290
<i>METEOR</i>	0,4390

Pada Tabel 4.18, nilai metrik *BLEU* sebesar 0,1767 menunjukkan bahwa tingkat kesamaan kata secara langsung antara hasil prediksi dan teks referensi relatif rendah. Dalam penelitian ini, *BLEU* yang tidak terlalu tinggi justru menunjukkan bahwa model tidak hanya menyalin teks referensi, tetapi mampu melakukan parafrase dengan variasi kalimat yang berbeda. Hal ini sesuai dengan tujuan transformasi kesantunan yang tidak menuntut kesamaan

kata secara identik. Nilai *ROUGE-1* menghasilkan nilai 0,5792, mengukur kesamaan *unigram* (kata tunggal) menunjukkan bahwa sebagian besar kata dalam hasil prediksi relevan dengan referensi. Hal ini mengindikasikan bahwa model masih mampu mempertahankan informasi utama dalam kalimat meskipun terjadi perubahan gaya bahasa. Metrik *ROUGE-2* menghasilkan 0,3332, mengukur kesamaan *bigram* (dua kata berurutan) menunjukkan tingkat kesamaan struktur kalimat. Hal ini menunjukkan bahwa model sering melakukan perubahan struktur kalimat, yang merupakan karakteristik umum dalam tugas parafrase dan transformasi gaya bahasa. Metrik *ROUGE-L* menghasilkan nilai 0,5290, mengukur kesamaan urutan kata terpanjang yang menunjukkan kesesuaian struktur kalimat secara keseluruhan. Hal ini mengindikasikan bahwa struktur kalimat yang dihasilkan tetap koheren dan tidak menyimpang jauh dari makna asli. Metrik *METEOR* menghasilkan nilai 0,4390, mengukur kesamaan semantik antara teks prediksi dan referensi, termasuk sinonim dan variasi kata. Metrik ini menunjukkan bahwa meskipun terdapat perubahan kata dan struktur, makna utama dari kalimat masih dapat dipertahankan oleh model.

Secara keseluruhan, hasil evaluasi menunjukkan bahwa model IndoT5 memiliki kemampuan yang baik dalam melakukan transformasi teks kesantunan. Model mampu menghasilkan variasi kalimat yang lebih santun tanpa harus menyalin teks referensi secara langsung. Perbedaan nilai pada beberapa metrik evaluasi menunjukkan bahwa model cenderung menghasilkan variasi struktur dan pilihan kata yang berbeda dari teks referensi, namun makna utama kalimat masih dapat dipertahankan. Oleh karena itu, evaluasi manual tetap dilakukan untuk memastikan kesesuaian makna dan konteks komunikasi akademik pada hasil transformasi yang dihasilkan model.

4.5.3 Contoh Hasil Transformasi

Contoh hasil transformasi teks menggunakan model IndoT5 ditunjukkan pada Tabel 4.19, berdasarkan contoh dibawah ini, model IndoT5 mampu menghasilkan teks yang lebih sopan dan formal dibandingkan teks asli dan juga mampu menambahkan unsur kesopanan seperti “mohon maaf” yang merupakan ciri komunikasi akademik.

Tabel 4.19 Contoh Hasil Transformasi Teks IndoT5

Kalimat Awal	Hasil Transformasi
Bu, tugas ini dinilainya kapan ya? Saya belum lihat nilainya di Google Classroom.	Assalamu'alaikum Bu, saya ingin menanyakan hasil penilaian tugas ini di Google Classroom.
Pak, UTSnya kok jadwalnya beda ya sama yang di formulir?	Assalamualaikum, Pak. Mohon izin bertanya, apakah ada perbedaan jadwal untuk UTS dengan yang ada di formulir?
Misi pak saya mau bimbingan skripsi besok apa bisa?	Assalamu'alaikum, Pak. Mohon izin, saya ingin melakukan bimbingan skripsi besok. Apakah Bapak berkenan?

Berdasarkan contoh tersebut, model IndoT5 mampu meningkatkan tingkat kesantunan dengan menambahkan struktur kalimat yang lebih lengkap, penggunaan bahasa yang lebih baku, serta unsur kesantunan seperti salam pembuka dan ungkapan permohonan. Selain itu, makna utama dari kalimat masih dapat dipertahankan dengan baik. Namun demikian, model cenderung menghasilkan kalimat yang lebih panjang dan formal dibandingkan dengan teks asli. Dalam beberapa konteks komunikasi singkat, hasil transformasi tersebut dapat terasa terlalu formal atau kurang natural. Selain itu, terdapat kemungkinan bahwa beberapa kasus tertentu model dapat mengubah makna asli kalimat. Terutama teks awal bersifat ambigu atau tidak lengkap. Hal ini menunjukkan bahwa meskipun model mampu meningkatkan kesantunan, aspek pemahaman konteks dan makna masih menjadi tantangan. Dengan demikian, hasil transformasi IndoT5 secara umum menunjukkan peningkatan kesantunan yang signifikan, tetapi tetap memerlukan evaluasi lebih lanjut untuk memastikan kesesuaian makna dan konteks komunikasi akademik. Hasil Transformasi secara lengkap dapat dilihat pada Lampiran F.

4.6 Deteksi Konteks Akademik

Pada penelitian ini, sistem juga dilengkapi dengan deteksi konteks akademik. Tujuannya agar memastikan bahwa teks yang diproses oleh sistem benar-benar berada dalam konteks komunikasi akademik antara mahasiswa-dosen. Tahapan ini dilakukan sebelum proses klasifikasi kesantunan, sehingga sistem hanya memproses teks yang relevan dengan domain penelitian. Deteksi konteks akademik dilakukan menggunakan pendekatan berbasis *keyword matching*, yaitu dengan mencocokkan kata-kata dalam teks input dengan daftar *keyword* yang berkaitan dengan konteks akademik. *Keyword* dikelompokkan dalam beberapa kategori, seperti dosen, mahasiswa, tugas, ujian, kuliah, skripsi, dan lainnya. *Keyword matching* dapat dilihat pada Lampiran G. Jika teks yang dimasukkan mengandung satu atau lebih *keyword* yang

relevan, maka teks dianggap berada dalam konteks akademik dan akan diproses ke tahap selanjutnya. Sebaliknya, jika tidak ditemukan kata kunci yang sesuai, sistem akan menghentikan proses dan menampilkan notifikasi kepada pengguna bahwa teks di luar konteks akademik.

4.6.1 Implementasi Deteksi Konteks Akademik

Pada tahap implementasi, sistem melakukan pencocokan *keyword* terhadap teks input yang telah dinormalisasi. Semakin banyak kategori *keyword* yang terdeteksi dalam teks, maka akan semakin tinggi tingkat *confidence* bahwa teks tersebut termasuk dalam konteks akademik. Sebagai contoh, teks seperti “Pak”, “bimbingan”, dan “skripsi”. Hal ini menunjukkan bahwa metode ini efektif dalam mengenali teks yang secara eksplisit menggunakan istilah akademik. Namun, terdapat kasus dimana teks yang sebenarnya termasuk konteks akademik tidak terdeteksi. Misalnya, kalimat “Apakah saya bisa bertemu besok untuk membahas revisi?” dapat tidak terdeteksi apabila tidak mengandung kata kunci eksplisit, meskipun secara makna termasuk komunikasi akademik. Hal ini menunjukkan keterbatasan metode dalam memahami konteks implisit. Selain itu, terdapat kemungkinan terjadinya kesalahan deteksi (*False Positive*), yaitu ketika teks yang bukan konteks akademik terdeteksi sebagai akademik. Contohnya, kalimat “Bu, saya izin tidak masuk kerja hari ini” dapat terdeteksi sebagai konteks akademik karena mengandung kata “Bu” dan “izin”, meskipun tidak berkaitan dengan lingkungan perkuliahan.

4.6.2 Output Sistem Deteksi Konteks

Output sistem deteksi konteks berupa keputusan apakah teks termasuk dalam konteks akademik atau tidak, disertai dengan tingkat *confidence*. Jika teks terdeteksi sebagai konteks akademik, maka sistem akan melanjutkan ke tahap klasifikasi kesantunan. Sebaliknya, jika tidak terdeteksi, sistem akan menampilkan notifikasi kepada pengguna. Berdasarkan hasil implementasi, metode *keyword matching* mampu mendeteksi konteks akademik dengan cukup baik, terutama pada teks yang secara eksplisit mengandung istilah akademik. Penggunaan kategori *keyword* yang beragam juga membantu meningkatkan cakupan deteksi terhadap berbagai jenis komunikasi akademik. Pemilihan metode ini didasarkan pada kemudahan implementasi dan efisiensi dalam proses penyaringan awal, karena tidak memerlukan pelatihan model tambahan dan dapat bekerja dengan cepat. Namun demikian, metode ini memiliki keterbatasan dalam memahami makna secara kontekstual, terutama pada teks yang tidak

mengandung *keyword* secara langsung. Oleh karena itu, pendekatan ini lebih tepat digunakan sebagai tahap awal *filtering* sebelum proses klasifikasi menggunakan model lebih kompleks.

Dengan demikian, deteksi konteks akademik dalam penelitian ini berfungsi sebagai mekanisme penyaringan awal untuk memastikan bahwa hanya teks yang relevan yang diproses lebih lanjut, meskipun belum mampu memahami konteks secara mendalam. Pengujian dilakukan menggunakan contoh teks dengan konteks akademik dan non-akademik untuk menunjukkan kemampuan sistem dalam melakukan validasi konteks sebelum proses klasifikasi kesantunan dijalankan. Hal ini dilakukan karena sistem yang dikembangkan tidak hanya berfokus pada klasifikasi kesantunan, tetapi juga memastikan bahwa input yang diproses sesuai dengan ruang lingkup penelitian, yaitu komunikasi akademik antara mahasiswa dan dosen. Oleh karena itu, teks non-akademik digunakan sebagai pembanding untuk menunjukkan bahwa sistem mampu menolak input di luar konteks akademik.

4.7 Integrasi Sistem

Pada tahap ini, model klasifikasi kesopanan menggunakan IndoBERT dan model transformasi teks menggunakan IndoT5 diintegrasikan ke dalam sebuah sistem berbasis web. Integrasi ini bertujuan untuk mengimplementasikan model agar dapat digunakan secara langsung oleh pengguna dalam konteks nyata. Tampilan sistem ditunjukkan pada Gambar 4.10.



The screenshot displays the SIPAN web application interface. At the top, the logo 'SIPAN' is shown with a graduation cap icon. Below it, the text 'Asisten Cerdas untuk Mengecek dan Memperbaiki Kesantunan Chat ke Dosen' is displayed. The main section is titled 'Input Teks untuk Analisis' and contains a checkbox labeled 'Masukkan Teks (Konteks Akademik):'. Below the checkbox is a text input field with the example text 'Contoh: Pak, kapan ujian UTS dilaksanakan?'. A small note below the input field states '① Pastikan teks Anda berkaitan dengan konteks akademik (dosen, mahasiswa, kuliah, tugas, dll)'. At the bottom of the form, there are two buttons: 'Cek Kesopanan' and 'Reset'.

Gambar 4.10 Antarmuka Sistem Deteksi dan Transformasi Kesantunan Komunikasi Akademik Berbasis Web

Sistem ini dirancang dengan arsitektur *client-server*, dimana pengguna memasukkan teks melalui antarmuka web (*frontend*), kemudian teks tersebut dikirim ke server (*backend*) untuk diproses oleh model. Alur kerja sistem ini adalah sebagai berikut:

- a. Pengguna memasukkan teks ke dalam sistem
- b. Sistem melakukan deteksi konteks akademik
- c. Jika teks termasuk konteks akademik, maka dilakukan klasifikasi menggunakan IndoBERT namun jika tidak akan muncul pemberitahuan “Maaf teks bukan konteks akademik”
- d. Jika hasil klasifikasi adalah tidak sopan, maka teks diproses menggunakan IndoT5
- e. Hasil akhir ditampilkan kepada pengguna.

Integrasi ini memungkinkan sistem bekerja secara otomatis secara *end-to-end*, mulai dari deteksi hingga transformasi teks secara otomatis. Sehingga pengguna tidak perlu menjalankan dua model secara terpisah. Hal ini meningkatkan efisiensi penggunaan sistem serta memberikan pengalaman yang lebih praktis bagi pengguna. Dari sisi manfaat, sistem ini dapat membantu mahasiswa dalam menyusun pesan komunikasi akademik yang lebih santun, terutama dalam situasi di mana mahasiswa belum memahami secara penuh etika berbahasa kepada dosen. Dengan adanya fitur transformasi otomatis, sistem juga berperan sebagai media pembelajaran dalam memahami pola komunikasi akademik yang tepat. Namun demikian, integrasi sistem ini juga memiliki beberapa keterbatasan. Sistem sangat bergantung pada hasil klasifikasi IndoBERT, sehingga kesalahan pada tahap klasifikasi dapat mempengaruhi proses selanjutnya. Selain itu, hasil transformasi IndoT5 tidak selalu sempurna, terutama pada teks yang bersifat ambigu atau memiliki konteks yang tidak lengkap. Dalam menangani input yang tidak termasuk konteks akademik, sistem secara langsung menghentikan proses dan memberikan notifikasi kepada pengguna. Hal ini bertujuan untuk menjaga agar sistem tetap digunakan sesuai dengan domain penelitian. Namun, pendekatan ini juga dapat menyebabkan beberapa teks yang sebenarnya relevan tetapi tidak terdeteksi secara eksplisit menjadi tidak diproses. Dengan demikian, integrasi sistem dalam penelitian ini tidak hanya menggabungkan dua model, tetapi juga membentuk suatu pipeline yang mampu mendeteksi dan mentransformasikan kesantunan teks secara otomatis, meskipun masih memiliki keterbatasan dalam memahami konteks secara menyeluruh.

4.8 Evaluasi Manual

Evaluasi manual dilakukan untuk menilai performa sistem berdasarkan persepsi manusia terhadap kesantunan bahasa dalam konteks komunikasi akademik. Evaluasi ini melibatkan seorang evaluator yang memiliki latar belakang di bidang psikologi, yaitu Ibu Fani Eka Nurtjahjo, S. Psi., M. Psi., dosen psikologi di Universitas Islam Indonesia. Keterlibatan evaluator bertujuan untuk menilai aspek pragmatik yang tidak sepenuhnya dapat diukur melalui metrik otomatis, seperti kesantunan, kejelasan maksud, dan kesesuaian konteks komunikasi. Evaluasi manual dilakukan terhadap 100 data uji yang diperoleh dari 10 responden. Dari jumlah tersebut, seluruh data digunakan untuk mengevaluasi hasil klasifikasi model IndoBERT, sedangkan sebanyak 67 data yang teridentifikasi sebagai tidak sopan digunakan untuk mengevaluasi hasil transformasi model IndoT5. Pemilihan sampel ini bertujuan untuk merepresentasikan variasi data tanpa harus mengevaluasi seluruh dataset secara manual.

Hasil evaluasi menunjukkan bahwa secara umum sistem mampu mengklasifikasikan dan mentransformasikan teks menjadi lebih santun dengan tetap mempertahankan makna utama. Namun demikian, terdapat beberapa perbedaan antara hasil evaluasi manual dan evaluasi otomatis. Hal ini disebabkan karena metrik otomatis seperti *BLEU*, *ROUGE*, dan *METEOR* hanya mengukur kesamaan teks secara struktural, sedangkan evaluasi manual mempertimbangkan aspek konteks, kesantunan, serta kejelasan maksud dalam komunikasi akademik. Dengan demikian, evaluasi manual dalam penelitian ini berperan sebagai validasi kontekstual terhadap hasil sistem, sekaligus memberikan gambaran yang lebih komprehensif mengenai kualitas output yang dihasilkan.

4.8.1 Evaluasi Manual IndoBERT

Evaluasi manual pada tahap ini dilakukan dengan membandingkan hasil prediksi model IndoBERT dengan hasil penilaian oleh evaluator terhadap data yang sama. Setiap hasil kemudian dikategorikan ke dalam empat jenis, yaitu *True Positive* (TP), *True Negative* (TN), *False Positive* (FP) dan *False negative* (FN).

Tabel 4.20 *Confusion Matrix* Evaluasi Manual

Kategori	Jumlah
<i>True Positive</i> (TP)	60
<i>True Negative</i> (TN)	29
<i>False Positive</i> (FP)	6
False Negative (FN)	5
Total	100

Berdasarkan hasil Tabel 4.20 di atas, dapat dijelaskan bahwa sebanyak 60 data (TP) berhasil diklasifikasikan dengan benar sebagai tidak sopan oleh sistem dan sesuai dengan penilaian ahli psikologi. Sebanyak 29 data (TN) berhasil diklasifikasikan dengan benar sebagai kalimat sopan. Sebanyak 6 data (FP) merupakan kesalahan dimana sistem menanggapi teks tidak sopan, namun menurut ahli psikologi termasuk sopan. Dan sebanyak 5 data (FN) merupakan kesalahan di mana sistem menganggap teks sopan, namun sebenarnya tidak sopan menurut ahli psikologi. Berdasarkan *confusion matrix* tersebut, diperoleh nilai evaluasi model sebagai berikut pada Tabel 4.21.

Tabel 4.21 Evaluasi Manual Model IndoT5

Metrik	Nilai
<i>Accuracy</i>	89,00 %
<i>Precision</i>	90,91 %
<i>Recall</i>	92,31 %
<i>F1-Score</i>	91,60 %

Berdasarkan hasil Tabel 4.21, di atas diperoleh nilai akurasi sebesar 89,00%. Selain itu, nilai precision sebesar 90,91%, recall sebesar 92,31%, dan F1-score sebesar 91,60% menunjukkan bahwa model memiliki tingkat kesesuaian yang cukup tinggi terhadap penilaian evaluator. Namun demikian, nilai akurasi ini lebih rendah dibandingkan dengan evaluasi otomatis sebelumnya yang mencapai 97,00%. Perbedaan ini menunjukkan bahwa performa model yang tinggi secara metrik tidak selalu mencerminkan kesesuaian dengan penilaian manusia. Berdasarkan *confusion matrix*, kesalahan *false positive* terjadi ketika model mengklasifikasikan teks sebagai tidak sopan, tetapi menurut evaluator termasuk sopan. Hal ini umumnya terjadi pada teks dengan gaya bahasa informal namun masih dapat diterima dalam konteks komunikasi akademik. Sementara itu, kesalahan *false negative* terjadi ketika model

menganggap teks sebagai sopan, tetapi menurut evaluator termasuk tidak sopan. Kesalahan ini biasanya muncul pada teks yang secara struktur terlihat netral, tetapi secara pragmatik kurang santun. Perbedaan antara evaluasi otomatis dan manual ini menunjukkan bahwa kesantunan bahasa tidak hanya bergantung pada pola kata, tetapi juga pada aspek pragmatik seperti maksud, konteks, dan norma komunikasi. Model cenderung menilai berdasarkan pola linguistik, sedangkan evaluator mempertimbangkan konteks komunikasi secara lebih luas. Selain itu, perbedaan ini juga dipengaruhi oleh subjektivitas dalam penilaian manusia, di mana standar kesantunan dapat bervariasi tergantung pada perspektif dan pengalaman evaluator. Dengan demikian, hasil evaluasi manual menunjukkan bahwa meskipun model IndoBERT memiliki performa yang tinggi secara otomatis, evaluasi berbasis manusia tetap diperlukan untuk menilai kesesuaian kesantunan secara lebih kontekstual.

4.8.2 Evaluasi Manual IndoT5

Evaluasi manual pada model IndoT5 dilakukan terhadap 67 sampel data uji, yang merupakan sebagian dari keseluruhan dataset. Evaluasi ini bertujuan untuk menilai apakah hasil transformasi teks yang dihasilkan oleh model telah menjadi lebih santun dibandingkan teks asli berdasarkan penilaian evaluator. Evaluasi manual dalam penelitian ini melibatkan seorang evaluator yang memiliki latar belakang di bidang psikologi. Keterlibatan psikolog didasarkan pada pertimbangan bahwa kesantunan berbahasa tidak hanya berkaitan dengan struktur linguistik, tetapi juga dengan aspek komunikasi interpersonal dan konteks sosial. Evaluator berperan dalam menilai apakah hasil transformasi teks telah sesuai dengan norma komunikasi yang santun, mempertimbangkan kejelasan makna, serta kesesuaian konteks relasi antara mahasiswa dan dosen. Dengan demikian, evaluasi manual tidak hanya berfungsi sebagai validasi teknis, tetapi juga sebagai validasi kontekstual terhadap kualitas hasil transformasi yang dihasilkan oleh sistem. Berikut hasil evaluasi manual disajikan pada Tabel 4.22.

Tabel 4.22 Hasil Evaluasi Manual IndoT5

Keterangan	Jumlah
Hasil menjadi sopan	54
Total data uji	67

Berdasarkan keterangan tabel di atas dapat disimpulkan bahwa perhitungan akurasi IndoT5 adalah sebagai berikut:

$$Accuracy = \frac{54}{67} \times 100\% = 80,60\%$$

Hasil evaluasi diperoleh tingkat keberhasilan sebesar 80,60%. Nilai ini menunjukkan bahwa sebagian besar teks berhasil ditransformasikan menjadi lebih santun sesuai dengan penilaian evaluator. Namun demikian, terdapat 13 data yang belum sepenuhnya sesuai. Hal ini menunjukkan bahwa meskipun model mampu menghasilkan kalimat yang lebih formal, tidak semua hasil transformasi dianggap tepat secara pragmatik oleh evaluator. Jika dibandingkan dengan evaluasi otomatis sebelumnya yang menggunakan metrik NLG (*BLEU*, *ROUGE*, dan *METEOR*), hasil evaluasi manual ini menunjukkan adanya perbedaan. Model dapat memperoleh skor metrik yang cukup baik, namun belum tentu seluruh hasil transformasi dinilai santun oleh manusia. Beberapa penyebab perbedaan tersebut antara lain: (1) model IndoT5 berfokus pada parafrase, bukan secara khusus pada kesantunan, (2) model tidak sepenuhnya memahami konteks sosial dalam komunikasi mahasiswa–dosen, dan (3) serta adanya variasi bahasa yang dihasilkan model yang tidak selalu sesuai dengan preferensi evaluator. Selain itu, beberapa kesalahan yang ditemukan antara lain: (1) kalimat menjadi terlalu panjang atau bertele-tele, (2) penggunaan bahasa yang terlalu formal sehingga kurang natural, dan (3) serta dalam beberapa kasus terjadi perubahan makna dari teks asli. Hal ini menunjukkan bahwa dalam tugas transformasi kesantunan, kesamaan kata atau struktur kalimat saja tidak cukup untuk menilai kualitas hasil. Aspek pragmatik, konteks, dan persepsi manusia menjadi faktor penting yang tidak sepenuhnya dapat diukur dengan metrik otomatis. Dengan demikian, evaluasi manual memberikan gambaran bahwa model IndoT5 memiliki kemampuan yang cukup baik dalam meningkatkan kesantunan teks, tetapi masih memerlukan penyempurnaan agar hasil transformasi lebih sesuai dengan konteks dan penilaian manusia.

4.8.3 Analisis Keseluruhan Sistem

Secara keseluruhan, hasil evaluasi manual menunjukkan bahwa sistem yang dikembangkan memiliki kemampuan yang baik dalam melakukan deteksi dan transformasi kesantunan teks *chat* mahasiswa–dosen. Namun, performa sistem menunjukkan karakteristik yang berbeda pada setiap komponen. Model klasifikasi berbasis IndoBERT menunjukkan performa yang kuat dalam mengidentifikasi tingkat kesantunan teks. Hal ini terlihat dari hasil evaluasi otomatis yang tinggi serta didukung oleh evaluasi manual yang juga menunjukkan

tingkat kesesuaian cukup baik. Model mampu mengenali pola kesantunan berbasis struktur kalimat dan pilihan kata secara konsisten, terutama pada teks yang memiliki perbedaan yang jelas antara bahasa formal dan informal.

Di sisi lain, model transformasi berbasis IndoT5 menunjukkan kemampuan yang cukup baik dalam mengubah teks menjadi lebih santun, namun masih memiliki keterbatasan. Model cenderung berhasil pada teks dengan struktur sederhana, tetapi masih mengalami kesulitan pada teks yang ambigu, memiliki konteks implisit, atau memerlukan pemahaman hubungan sosial yang lebih dalam. Dalam beberapa kasus, hasil transformasi menjadi terlalu panjang, terlalu formal, atau mengalami sedikit perubahan makna.

Selain itu, komponen deteksi konteks akademik berbasis *keyword matching* terbukti efektif dalam menyaring teks yang secara eksplisit mengandung istilah akademik. Namun, metode ini masih lemah dalam mendeteksi konteks yang bersifat implisit dan berpotensi menghasilkan kesalahan deteksi pada teks yang tidak relevan tetapi mengandung kata kunci tertentu. Hasil evaluasi manual juga menunjukkan adanya perbedaan antara penilaian sistem dan penilaian manusia. Hal ini menegaskan bahwa kesantunan bahasa tidak hanya bergantung pada pola linguistik, tetapi juga pada aspek pragmatik, konteks sosial, dan interpretasi makna yang lebih kompleks. Oleh karena itu, meskipun sistem mampu memberikan hasil yang cukup baik secara otomatis, validasi manusia tetap diperlukan untuk memastikan kesesuaian dalam konteks komunikasi akademik.

Pemilihan model dalam penelitian ini difokuskan pada pendekatan berbasis *deep learning*, yaitu IndoBERT untuk klasifikasi dan IndoT5 untuk transformasi teks. Penelitian ini tidak melakukan perbandingan langsung dengan model *machine learning* tradisional seperti Naive Bayes atau Support Vector Machine (SVM), karena tujuan utama penelitian adalah mengembangkan sistem yang mampu melakukan deteksi sekaligus transformasi kesantunan secara terintegrasi, bukan membandingkan performa algoritma klasifikasi. Selain itu, model transformasi seperti IndoT5 tidak memiliki padanan langsung pada pendekatan *machine learning* tradisional, sehingga penggunaan model berbasis *transformer* menjadi lebih relevan dalam konteks penelitian ini. Pendekatan ini memungkinkan sistem untuk memahami konteks bahasa secara lebih mendalam serta menghasilkan kalimat yang lebih fleksibel dan natural. Meskipun demikian, penelitian selanjutnya dapat mempertimbangkan penggunaan model pembanding sebagai baseline untuk memberikan evaluasi yang lebih komprehensif.

Dari sisi praktis, sistem yang dikembangkan memiliki potensi sebagai alat bantu bagi mahasiswa dalam menyusun pesan akademik yang lebih santun. Sistem tidak hanya

memberikan hasil klasifikasi, tetapi juga menyediakan alternatif transformasi kalimat, sehingga dapat berfungsi sebagai media pembelajaran dalam memahami etika komunikasi akademik. Namun demikian, sistem ini masih memiliki beberapa keterbatasan, antara lain ketergantungan pada hasil klasifikasi awal, keterbatasan model dalam memahami konteks pragmatik secara mendalam, serta pendekatan deteksi konteks yang masih berbasis aturan sederhana. Oleh karena itu, pengembangan lebih lanjut diperlukan, khususnya dalam meningkatkan pemahaman konteks dan aspek sosial dalam bahasa. Dengan demikian, penelitian ini menunjukkan bahwa pendekatan berbasis model bahasa dapat digunakan untuk mendukung komunikasi akademik, namun belum sepenuhnya dapat menggantikan penilaian manusia dalam memahami kesantunan secara kontekstual.

4.8.4 Analisis Kasus Sulit

Berdasarkan hasil evaluasi, kesalahan pada sistem umumnya terjadi pada teks yang memiliki konteks yang kurang jelas atau tidak lengkap. Teks yang terlalu singkat atau tidak menyertakan informasi yang cukup, seperti tujuan atau identitas, menyebabkan model kesulitan dalam menentukan tingkat kesantunan secara akurat. Selain itu, pada model transformasi IndoT5, ditemukan beberapa kasus di mana hasil output mengalami perubahan informasi yang tidak diinginkan. Salah satu contohnya adalah pada teks berikut. Input: “Assalamualaikum, saya asisten Ibu di mata kuliah Kewirausahaan Syariah.” Output: “Assalamu’alaikum, Ibu. Saya asisten Ibu di mata kuliah nama mata kuliah.” Pada contoh tersebut, model menghasilkan frasa “nama mata kuliah” yang tidak sesuai dengan informasi pada teks asli. Kesalahan ini kemungkinan dipengaruhi oleh pola data pelatihan yang digunakan dalam proses *fine-tuning* model.

Dalam dataset, terdapat beberapa contoh teks transformasi yang menggunakan template umum, seperti “nama mata kuliah” atau “(sebutkan mata kuliah)” sebagai *placeholder*. Kondisi tersebut menunjukkan bahwa model masih dipengaruhi oleh pola tertentu dalam data pelatihan, sehingga pada beberapa kasus menghasilkan *output* yang mengandung *placeholder* atau informasi yang kurang spesifik. Hal ini menunjukkan bahwa kualitas dan konsistensi data pelatihan sangat berpengaruh terhadap hasil transformasi yang dihasilkan oleh model. Meskipun demikian, model tetap mampu menghasilkan transformasi yang baik pada sebagian besar teks dengan struktur sederhana. Namun, pada beberapa kasus tertentu yang mengandung konteks ambigu, makna implisit, atau relasi sosial yang kompleks, model masih memiliki keterbatasan dalam memahami konteks pragmatik secara menyeluruh. Temuan ini

menunjukkan bahwa meskipun model mampu meningkatkan kesantunan teks, masih terdapat keterbatasan dalam mempertahankan informasi spesifik dari input, terutama pada bagian yang berkaitan dengan entitas tertentu. Oleh karena itu, pada penelitian selanjutnya perlu dilakukan pembersihan dataset dari *placeholder* atau *template* agar model dapat menghasilkan *output* yang lebih akurat dan kontekstual.

BAB V

KESIMPULAN DAN SARAN

Bab ini menyajikan kesimpulan dari hasil penelitian yang telah dilakukan berdasarkan rumusan masalah yang telah ditetapkan. Selain itu, disampaikan pula saran yang dapat digunakan sebagai arahan pengembangan pada penelitian selanjutnya.

5.1 Kesimpulan

Berdasarkan hasil penelitian mengenai pengembangan sistem deteksi dan transformasi kesantunan teks *chat* mahasiswa-dosen menggunakan pendekatan *fine-tuning* IndoBERT dan IndoT5, dapat disimpulkan bahwa:

- a. Penelitian ini berhasil membangun model klasifikasi kesantunan teks chat mahasiswa-dosen menggunakan metode *fine-tuning* IndoBERT yang mampu mengidentifikasi tingkat kesantunan ke dalam dua kategori, yaitu sopan dan tidak sopan. Model klasifikasi yang dihasilkan menunjukkan performa yang sangat baik dengan akurasi mencapai 97%, sehingga mampu mendeteksi tingkat kesantunan teks komunikasi akademik secara efektif.
- b. Penelitian ini juga berhasil membangun model transformasi teks menggunakan metode *fine-tuning* IndoT5 yang mampu mengubah teks yang teridentifikasi tidak sopan menjadi lebih santun tanpa menghilangkan makna utama kalimat. Berdasarkan hasil evaluasi menggunakan BLEU, ROUGE, dan METEOR, model transformasi mampu menghasilkan kalimat yang lebih sesuai dengan kaidah komunikasi akademik dan mempertahankan konteks pesan secara umum.
- c. Penelitian ini berhasil mengintegrasikan proses klasifikasi dan transformasi teks ke dalam satu sistem berbasis web yang terintegrasi. Integrasi ini menjadi kontribusi penelitian karena penelitian terdahulu umumnya hanya berfokus pada salah satu tugas, yaitu klasifikasi atau transformasi teks secara terpisah. Dengan demikian, sistem yang dikembangkan mampu melakukan proses deteksi dan perbaikan kesantunan teks secara otomatis dalam satu alur proses.
- d. Berdasarkan hasil evaluasi manual oleh evaluator, sebagian besar hasil transformasi menunjukkan peningkatan tingkat kesantunan dan kesesuaian dengan konteks komunikasi akademik. Namun demikian, masih ditemukan beberapa kasus yang menunjukkan perbedaan penilaian, terutama pada teks yang mengandung makna implisit, konteks pragmatik, serta variasi latar belakang budaya dan gaya komunikasi penutur.

Secara umum, penelitian ini menunjukkan bahwa pendekatan berbasis model *transformer* memiliki potensi yang baik sebagai alat bantu dalam meningkatkan kualitas komunikasi akademik digital, khususnya dalam mendukung mahasiswa menyusun pesan yang lebih santun kepada dosen. Meskipun demikian, sistem yang dikembangkan masih memiliki keterbatasan dalam memahami aspek pragmatik, konteks sosial, serta keragaman budaya komunikasi, sehingga evaluasi dan pertimbangan manusia tetap diperlukan dalam penggunaan sistem secara nyata.

5.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, beberapa saran yang dapat diberikan adalah sebagai berikut:

- a. Penelitian selanjutnya disarankan menggunakan dataset yang lebih luas dan beragam agar model mampu mempelajari variasi komunikasi akademik secara lebih optimal serta meningkatkan kemampuan generalisasi model pada berbagai konteks bahasa.
- b. Berdasarkan hasil evaluasi, model masih mengalami keterbatasan dalam memahami konteks pragmatik, seperti makna implisit, ironi, sarkasme, serta relasi sosial pada beberapa kasus tertentu. Oleh karena itu, penelitian selanjutnya disarankan menggunakan pendekatan yang lebih mampu memahami konteks bahasa secara komprehensif agar hasil klasifikasi dan transformasi teks menjadi lebih natural dan kontekstual. Selain itu, aspek pragmatik dan keragaman budaya juga dapat dipertimbangkan lebih lanjut dalam proses pelabelan maupun evaluasi model, sehingga sistem mampu memahami variasi kesantunan bahasa dari berbagai latar belakang budaya dengan lebih baik.
- c. Metode deteksi konteks akademik yang digunakan pada penelitian ini masih berbasis *keyword matching* sehingga belum mampu memahami makna kalimat secara menyeluruh. Oleh karena itu, penelitian selanjutnya disarankan mengembangkan pendekatan deteksi konteks yang lebih adaptif dan kontekstual. Selain itu, sistem juga dapat dikembangkan ke platform yang lebih mudah diakses, seperti aplikasi *mobile* atau integrasi dengan media komunikasi digital, agar implementasi sistem dapat digunakan secara lebih luas.
- d. Penelitian ini masih memiliki keterbatasan pada proses evaluasi pelabelan data karena belum menerapkan pengukuran *inter-annotator agreement* seperti Cohen's Kappa. Oleh karena itu, penelitian selanjutnya disarankan untuk melibatkan lebih banyak annotator dan melakukan evaluasi *agreement* guna meningkatkan reliabilitas pelabelan data.

- e. Penelitian selanjutnya dapat melibatkan lebih banyak evaluator atau ahli bahasa dalam evaluasi manual hasil transformasi teks, serta melakukan pengujian langsung kepada lebih banyak pengguna sasaran, agar hasil sistem dapat divalidasi tidak hanya secara teknis tetapi juga secara pragmatik dan kontekstual.

Secara umum, penelitian ini masih memiliki ruang pengembangan pada aspek kualitas dataset, evaluasi pelabelan, serta kemampuan model dalam memahami konteks komunikasi akademik secara lebih mendalam, termasuk makna implisit, sarkasme, ironi, dan variasi bahasa dalam komunikasi digital. Oleh karena itu, pengembangan lebih lanjut diharapkan dapat meningkatkan kinerja sistem sehingga dapat digunakan secara lebih optimal sebagai alat bantu dalam menyusun teks chat mahasiswa-dosen yang santun. umum penelitian ini masih memiliki ruang pengembangan, khususnya dalam meningkatkan pemahaman konteks pragmatik dan variasi bahasa dalam komunikasi akademik.

DAFTAR PUSTAKA

- Abidin, K., & Wandu. (2023). Etika Komunikasi antara Mahasiswa dan Dosen dalam Interaksi Akademik melalui Media Digital. *Jurnal Ilmu Komunikasi*, 6(1).
- Aeni, E. S., Sahmini, M., & Nugraha, V. (2025). Penilaian karakter berdasarkan kesantunan berbahasa pada mahasiswa berorientasi penilaian sistem. *Sawerigading*, 31(1), 100–108. <https://doi.org/10.26499/sawer.v31i1.1539>
- Akhfasy, M. N., Ningsih, A. K., & Ilyas, R. (2024). Analisis sentimen pengguna Instagram terhadap komentar Karen's Diner menggunakan metode algoritma support vector machine dengan operator seleksi fitur information gain. *Jurnal Mahasiswa Teknik Informatika*, 8(3), 2699–2705. <https://doi.org/10.36040/jati.v8i3.9533>
- Amriza, R. N. S., & Supriyadi, D. (2021). Komparasi metode machine learning dan deep learning untuk deteksi emosi pada text di sosial media. *Jurnal JUPITER*, 13(2), 130–139.
- Anugrawati, N., & Syam, U. K. (2024). Faktor-faktor yang mempengaruhi strategi kesopanan dalam konteks akademik: kajian karakter di Universitas Muhammadiyah Makassar. *Jurnal Onoma: Pendidikan, Bahasa, Dan Sastra*, 10(4), 4305–4314. <https://doi.org/10.30605/onoma.v10i4.4666>
- Astiti, P., & Raharja, D. M. (2023). Peran komunikasi dalam pendidikan era digital. *Gandiwa: Jurnal Komunikasi*, 3(2), 41–48. <https://doi.org/10.30998/g.v3i2.2514>
- Baharuddin, F., & Naufal, M. F. (2023). Fine-tuning IndoBERT for Indonesian exam question classification based on Bloom's Taxonomy. *Journal of Information Systems Engineering and Business Intelligence*, 9(2), 253–263. <https://doi.org/10.20473/jisebi.9.2.253-263>
- Christiant, A., & Pratama, A. R. (2022). Implementasi Deep Learning untuk Mengubah Kalimat Tidak Sopan menjadi Sopan. *AUTOMATA*, 3(1).
- Eksan, E. R., Hafid, A., & Putra, T. Y. (2021). Kesantunan Berbahasa Mahasiswa terhadap Dosen di UNIMUDA Sorong: Tinjauan Pragmatik. *Jurnal Keilmuan, Bahasa, Sastra, Dan Pengajarannya*, 2(1), 16–23. <https://ejournal.unimudasorong.ac.id/index.php/jurnalfrasa/article/view/859>
- Fikri, R. A., & Muliono, R. (2025). Sentiment Towards Social Media Politeness Ambassadors: A Case Study Using the Naive Bayes Method. *Journal of Informatics*

- and Telecommunication Engineering*, 8(3Spc), 127–135.
<https://doi.org/10.31289/jite.v8i3Spc.14404>
- Fuadi, M., Wibawa, A. D., & Sumpeno, S. (2023). idT5: Indonesian Version of Multilingual T5 Transformer. *ArXiv Preprint ArXiv:2302.00856*, 1.
<https://doi.org/10.48550/arXiv.2302.00856>
- Hakim, L. N., Hana, F. M., & Wahyudin, W. C. (2026). Klasifikasi Komentar Toksik Berbahasa Indonesia di Media Sosial Berbasis Fine-Tuning IndoBERT. *Jurnal Riset Komputer*, 13(1), 202–209. <https://doi.org/10.30865/jurikom.v13i1.9449>
- Holidin, A., Haryanto, & Purnomo, A. C. (2025). Transformasi digital dalam pendidikan serta penggunaan teknologi komunikasi oleh mahasiswa. *Journal of Science, Technology, and Innovation*, 1(2), 268–277. <https://doi.org/10.65310/c1168c79>
- Huzaifa, M., Nalawati, R. E., Marcheta, N., Kurniawan, A., Ramadhanty, F. A., Rosnawati, Anggari, S. I., & Sagita, F. (2025). Model IndoBERT dalam pengenalan kebudayaan Jawa berbasis question answering. *Seminar Nasional Inovasi Vokasi*, 4, 416–423. <https://prosiding.pnj.ac.id/index.php/sniv/article/view/3882>
- Irmayanti, D., & Hermanto, T. I. (2026). Development of Hybrid K-Means DBSCAN Algorithm for Optimization of Landslide-Prone Area Clusters based on Web-GIS. *Sistemasi: Jurnal Sistem Informasi*, 15(1), 109–120.
<https://doi.org/10.32520/stmsi.v15i1.5671>
- Kono, M. F., Fajri, I. N., & Pristyanto, Y. (2025). Public Sentiment Analysis on Corruption Issues in Indonesia Using IndoBERT Fine-Tuning, Logistic Regression, and Linear SVM. *Journal of Applied Informatics and Computing (JAIC)*, 9(5), 2616–2628. <https://doi.org/10.30871/jaic.v9i5.10537>
- Koto, F., Lau, J. H., & Baldwin, T. (2021). INDOBERTWEET: A Pretrained Language Model for Indonesian Twitter with Effective Domain-Specific Vocabulary Initialization. *Association for Computational Linguistics*, 10660–10668.
<https://doi.org/10.18653/v1/2021.emnlp-main.833>
- Kurniadini, S. Z., & Putranto, T. D. (2025). Pola komunikasi interpersonal mahasiswa Indonesia (review of literature). *Jurnal Ilmu Komunikasi*, 15(2), 84–102.
<https://doi.org/10.36989/didaktik.v10i3.4130>
- Mahas, N. H., Muhsyanur, & Verlin, S. (2025). Pola Kesantunan Berbahasa Indonesia dalam Komunikasi Akademik Mahasiswa Universitas Islam As'adiyah Sengkang:

- Analisis Sosiopragmatik. *Aksentuasi: Jurnal Ilmiah Pendidikan Bahasa Dan Sastra Indonesia*, 6(1), 138–149.
- Nurwanda, Suarna, N., & Prihartono, W. (2024). Penerapan NLP (natural language processing) dalam analisis sentimen pengguna Telegram di PlayStore. *Jurnal Mahasiswa Teknik Informatika*, 8(2), 1841–1846. <https://doi.org/10.36040/jati.v8i2.8469>
- Poneni, I. S. (2023). An analysis of linguistic politeness strategies in online academic communication. *International Journal of Economic, Technology and Social Sciences*, 6(1), 358–365. <https://doi.org/10.53695/injects.v6i1.1541>
- Prasetyo, V. R., Benarkah, N., & Chrisintha, V. J. (2021). Implementasi natural language processing dalam pembuatan chatbot pada program information technology Universitas Surabaya. *Teknika*, 10(2), 114–121. <https://doi.org/10.34148/teknika.v10i2.370>
- Purnomo, B. (2025). Cultural Pragmatics of Politeness: The Role of Javanese Krama Inggil in Indonesian Academic Communication. *Journal of Linguistics, Culture and Communication*, 3(2), 248–263. <https://doi.org/10.61320/jolcc.v3i2.248-263>
- Puspa, E. S., Safarudin, M. S., Kurnia, H., Irawan, Mustofa, & Faridah, L. (2024). Pemanfaatan Media Sosial untuk Meningkatkan Komunikasi Akademik di Perguruan Tinggi. *Journal of Human and Education*, 4(6), 1367–1373. <https://doi.org/10.31004/jh.v4i6.1808>
- Puspitasari, W., Ramdani, D., & Maulana, A. M. (2024). IndoT5 (Text-to-Text Transfer Transformer) Algorithm for Paraphrasing Indonesian Language Islamic Sermon Manuscripts. *Khazanah Journal of Religion and Technology*, 2(2), 63–73. <https://doi.org/10.15575/kjrt.v2i2.1093>
- Rachman, F. P., & Santoso, H. (2021). Perbandingan model deep learning untuk klasifikasi sentiment analysis dengan teknik natural language processing. *Jurnal Teknologi Dan Manajemen Informatika*, 7(2), 103–112. <https://doi.org/10.26905/jtmi.v7i2.6506>
- Ramdan, M., Rakhmat Umbara, F., & Ilyas, R. (2025). Analisis sentimen pengguna aplikasi Sambara (E-Samsat) Jawa Barat menggunakan metode IndoBERT. *Jurnal Global Ilmiah*, 2(12).

- Ratnasari, D., & Fitriyanti, N. (2024). Pengaruh media sosial terhadap gaya komunikasi mahasiswa di era digital. *Central Publisher*, 2(9), 2616–2623. <https://doi.org/10.60145/jcp.v2i9.529>
- Reynaldi, Latief, D., & Tajibu, K. (2024). Etika Komunikasi Mahasiswa pada Dosen di Media Sosial WhatsApp. *Jurnal Peurawi:Media Kajian Komunikasi Islam*, 7(2), 131–154.
- Robet, Kohsasih, K. L., & Darwin, J. (2025). Penerapan Algoritma Transformer dalam Aplikasi Parafrase Teks Otomatis. *TAMIKA*, 5(1), 103–109. <https://doi.org/10.46880/tamika.Vol5No1.pp103-109>
- Sahabuddin, R., Arif, H. M., Rezky, M. A., Pratama, F. Y., Dirafathi, M., & Umama, N. (2024). Analisis pengaruh komunikasi internal terhadap prestasi akademik mahasiswa: peran kepercayaan sebagai variabel mediasi. *Maximal Journal : Jurnal Ilmiah Bidang Sosial, Ekonomi, Budaya Dan Pendidikan*, 2(1), 38–44. <https://malaqbipublisher.com/index.php/MAKSI>
- Sari, A. K., Irsyad, A., Aini, D. N., Islamiyah, & Ginting, S. E. (2024). Analisis sentimen Twitter menggunakan machine learning untuk identifikasi konten negatif. *Adopsi Teknologi Dan Sistem Informasi (ATASI)*, 3(1), 64–73. <https://doi.org/10.30872/atasi.v3i1.1373>
- Tan, K. L., Lee, C. P., Anbananthen, K. S. M., & Lim, K. M. (2022). RoBERTa-LSTM: A Hybrid Model for Sentiment Analysis With Transformer and Recurrent Neural Network. *IEEE Access*, 10, 21517–21525. <https://doi.org/10.1109/ACCESS.2022.3152828>
- Taqwa, R. H., Waldan, R., & Sari, P. A. (2025). Interpersonal communication and digital technology in the academic practices of Islamic higher education students. *Mawa Izh: Jurnal Dakwah Dan Pengembangan Sosial Kemanusiaan*, 16(01), 88–110. <https://doi.org/10.32923/maw.v16i01.5641>
- Tarwoto, Nugroho, R., Azka, N., & Sayudha Rendra Graha, W. (2025). Analisis sentimen ulasan aplikasi mobile JKN di Google PlayStore menggunakan IndoBERT. *Jurnal JTIK (Jurnal Teknologi Informasi Dan Komunikasi)*, 9(2), 495–505. <https://doi.org/10.35870/jtik.v9i2.3340>
- Taslim, M., Suherli, & Rusdin, F. R. (2024). Etika komunikasi di lingkungan akademik: evaluasi praktik dan tantangan di Universitas Almarisah Madani. *Jurnal Ilmiah Wahana Pendidikan*, 10(13), 1053–1063.

- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention Is All You Need. *Advances in Neural Information Processing Systems*, 30.
- Zaman, B., Humam, N., & Raharjana, I. K. (2025). Large Language Model-Based Detoxification for Bahasa Indonesia. *Cybernetics and Information Technologies*, 25(3), 3–21. <https://doi.org/10.2478/cait-2025-0019>

LAMPIRAN

Lampiran A. Contoh Data dari Berbagai Sumber

Contoh Teks	Sumber
Bu, saya udah ngirim tugasnya, cepet diperiksa ya	Chat GPT
Pak, saya telat soalnya macet	Chat GPT
Bu, kapan nilai saya keluar? lama banget	Chat GPT
Mohon informasi dari Ibu terkait jadwal pengumpulan revisi terakhir.	Chat GPT
Mohon izin Ibu, saya terlambat hadir karena ada kegiatan kampus yang bersamaan.	Chat GPT
Bu saya minta izin tidak ikut bimbingan kali ini	Goolge Form
Asaalamualaikum Bu, mau tanya buat ujian boleh open note selemba a4?	Goolge Form
Saya mau bimbingan, bisa minta waktu bapak sebentar?	Goolge Form
Pak tolong acc sekawan plis	Goolge Form
Assalamualaikum bu, sebelumnya izin bertanya bu apakah besok ibu dapat kebersamai kami untuk perkuliahan besok..? 🙏🌸	Goolge Form
Assalamualaikum, Bapak. Mohon maaf telah mengganggu waktunya hari ini. Izin, Pak, saya ingin menindaklanjuti revisi yang sudah saya kirim dua minggu lalu, tetapi belum ada tanggapan dari Bapak. Jika berkenan, apakah saya dapat melakukan bimbingan secara luring besok? Apabila Bapak berkenan, kira-kira pada pukul berapa saya dapat menemui Bapak? Terima kasih, Pak.	Tiktok
Assalamualaikum, selamat sore Pak. Mohon maaf mengganggu waktunya. [nama dan NIM]. GCR saya error Pak, belum bisa mengerjakan tugas dari pagi Pak	Tiktok
Bu, saya mau tanya kalau untuk proses pengambilan data untuk skripsi saya enak nya gimana ya Bu?	Tiktok
Ibu saya mau tanya kalo konversi sks untuk program pertukaran pelajar gitu di acc ga ya dari prodi?	Tiktok
Bu, mau tanya apakah nilai ujiannya dapat dilihat? Kalau iya lihat di mana ya, Bu	Tiktok

Lampiran B. Bentuk Pengembangan Kalimat *Prompting* ChatGPT untuk Kalimat Google Form

Contoh Teks	Prompt	Bentuk Pengembangan Kalimat
Asaalamualaikum Bu, mau tanya buat ujian boleh open note selembat A4?	Buatkan variasi kalimat komunikasi akademik mahasiswa kepada dosen yang lebih santun dengan mempertahankan makna asli. Gunakan indikator kesantunan berupa penggunaan sapaan formal, permohonan izin, dan bahasa baku formal.	Assalamualaikum, Ibu. Mohon izin bertanya, apakah ujian nanti diperbolehkan open note satu lembar A4?
		Assalamualaikum, Ibu. Saya ingin menanyakan, apakah saat ujian boleh membawa catatan satu lembar ukuran A4?
		Permisi, Ibu. Mohon izin bertanya, apakah ujian diperbolehkan open note satu lembar A4, Bu?
		Selamat pagi, Ibu. Mohon maaf mengganggu, saya ingin memastikan apakah ujian nanti boleh open note satu lembar A4.
		Assalamualaikum, Ibu. Apakah diperkenankan membawa catatan satu lembar A4 saat ujian nanti, Bu?
		Assalamualaikum, Ibu. Mohon izin, apakah pada saat ujian diperbolehkan open note satu lembar A4? Terima kasih, Ibu.
		Selamat sore, Ibu. Mohon izin bertanya, apakah ujian nanti boleh membawa catatan satu lembar A4, Bu?
		Selamat sore, Ibu. Mohon izin bertanya, apakah ujian nanti boleh open note satu lembar A4, Bu? Terima kasih sebelumnya.

Lampiran C. Bentuk Pengembangan Kalimat *Prompting* ChatGPT untuk Kalimat TikTok



Contoh Teks	Prompt	Bentuk Pengembangan Kalimat
Mohon maaf mengganggu waktunya pak/bu, saya mau meminta tanda tangan guna keperluan lomba.	Buatkan variasi kalimat komunikasi akademik mahasiswa kepada dosen yang lebih santun dengan mempertahankan makna asli. Gunakan indikator kesantunan berupa sapaan formal, permohonan izin, dan bahasa baku formal.	Mohon maaf mengganggu waktunya, Pak/Bu. Saya izin ingin meminta tanda tangan yang akan digunakan untuk keperluan lomba.
		Permisi, Pak/Bu. Mohon maaf sebelumnya, saya ingin meminta tanda tangan untuk keperluan lomba.
		Mohon izin, Pak/ bu. Saya hendak meminta tanda tangan sebagai syarat administrasi lomba.
		Maaf mengganggu waktunya, Pak/Bu. Saya ingin meminta tanda tangan Bapak/Ibu untuk keperluan berkas lomba yang sedang saya ikuti.
		Assalamualaikum, Pak/Bu. Mohon maaf, saya ingin meminta tanda tangan untuk melengkapi berkas lomba.
		Selamat siang, Pak/Bu. Mohon maaf menggagu waktunya, saya ingin meminta tanda tangan untuk kebutuhan lomba.
		Maaf mengganggu, Pak/Bu. Saya ingin memohon tanda tangan untuk keperluan mengikuti lomba. Terima kasih sebelumnya.

Lampiran D. Daftar Pertanyaan Google Form

The screenshot shows a Google Form with a yellow header. The header contains the logo of Universitas Islam Indonesia (UII) and the text 'PROGRAM STUDI INFORMATIKA PROGRAM SARJANA'. The main title of the form is 'Kuesioner Penelitian: Sistem Pengubah Teks Chat Mahasiswa kepada Dosen Menjadi Lebih Santun Menggunakan Fine-Tuning IndoBERT dan IndoT5'. Below the title, there is a greeting 'Assalamu'alaikum Wr. Wb.' and an introduction from the researcher, Belina Eka Ariesty (NIM 22523061), a student of the Informatics Program at UII Yogyakarta. The introduction states that the researcher is conducting a final assignment research with the title 'Sistem Pengubah Teks Chat kepada Dosen Menjadi Lebih Santun Menggunakan Fine-Tuning IndoBERT dan IndoT5'. The form is currently in edit mode, as indicated by the pencil icon in the bottom right corner.

A. Identitas Responden

No	Data Responden
1.	Nama
2.	Usia
3.	Pendidikan

B. Pertanyaan Utama

No	Pertanyaan
1.	Seberapa sering anda menghubungi dosen via chat?
2.	Tuliskan contoh chat yang menurut anda kurang sopan ketika menghubungi dosen.
3.	Tuliskan versi yang lebih sopan dari chat tersebut.
4.	Apa konteks komunikasi pada chat tersebut?
5.	Apakah anda merasa kesulitan menyusun teks chat yang sopan ke dosen?

Lampiran E. Implementasi Pembersihan Data

```
df_valid = df.dropna(subset=['text', 'label']).copy()
df_valid['label'] = df_valid['label'].astype(int)
print(f"\n Valid data: {len(df_valid)}")
print(f"    - Sopan (0): {len(df_valid[df_valid['label']==0])}")
print(f"    - Tidak Sopan (1): {len(df_valid[df_valid['label']==1])}")
def preprocess_text(text):
    text = str(text).lower()
    text = re.sub(r'http\S+|www\S+|https\S+', '', text)
    text = re.sub(r'@\w+', '', text)
    text = re.sub(r'#(\w+)', r'\1', text)
    text = re.sub(r'\s+', ' ', text).strip()
    text = re.sub(r'^a-zA-Z0-9\s.,!?', '', text)
    return text
df_valid['cleaned_text'] = df_valid['text'].apply(preprocess_text)
df_valid = df_valid[df_valid['cleaned_text'].str.len() >
5].reset_index(drop=True)
```

Lampiran F. Contoh Hasil Transformasi

Input	Output	Kategori Hasil
Selamat pagi, Pak. Saya kemarin sudah mengikuti kelas pengganti melalui Zoom meeting. Tapi, sampai saat ini, presensinya masih belum direkap. Mohon bantuannya ya, pak.	Selamat pagi, Pak. Saya kemarin sudah mengikuti kelas pengganti melalui Zoom meeting. Tapi, sampai saat ini, presensinya masih belum direkap. Mohon bantuannya ya, pak.	Berhasil (Santun)
Selamat Pagi, Bu. Apakah tugas matematika lanjut dapat dibuka kembali? Saya lupa submit	Selamat Pagi, Bu. Apakah tugas matematika lanjut dapat dibuka kembali? Saya lupa submit	Berhasil (Santun)
P info bimbingan minggu ini.	Assalamu'alaikum, Pak. Mohon izin, saya ingin menyampaikan informasi terkait jadwal bimbingan minggu ini.	Gagal
Bapak, teman saya berlaku curang saat ujian	Selamat siang Bapak, saya ingin menanyakan apakah ada praktik curang saat ujian berlangsung.	Gagal
Pak aku besok bimbingan ya	Assalamualaikum, Pak. Mohon izin, apakah besok saya dapat melakukan bimbingan dengan Bapak?	Berhasil (Tidak Santun)
Ibu saya izin gak masuk kelas karena sakit	Assalamualaikum, Ibu. Mohon izin tidak bisa hadir di kelas hari ini karena sedang sakit.	Berhasil (Tidak Santun)

Lampiran G. *Keyword Maching*

```
'academic_keywords': {
  'dosen': ['dosen', 'pak', 'bu', 'bapak', 'ibu', 'profesor', 'prof', 'dr',
'guru besar'],
  'mahasiswa': ['mahasiswa', 'mahasiswi', 'mhs', 'kami mahasiswa', 'saya
mahasiswa'],
  'tugas': ['tugas', 'assignment', 'pekerjaan rumah', 'pr', 'kerjakan',
'mengerjakan', 'dikerjakan'],
  'nilai': ['nilai', 'grade', 'skor', 'penilaian', 'hasil ujian'],
  'ujian': ['ujian', 'uts', 'uas', 'quiz', 'kuis', 'test', 'tes', 'exam'],
  'kuliah': ['kuliah', 'perkuliahan', 'kelas', 'mata kuliah', 'matkul',
'course', 'lecture'],
  'skripsi': ['skripsi', 'thesis', 'tesis', 'penelitian', 'riset',
'jurnal', 'paper'],
  'kampus': ['kampus', 'universitas', 'kampus', 'fakultas', 'jurusan',
'prodi', 'program studi'],
  'akademik': ['akademik', 'akademis', 'perkuliahan', 'semester', 'sks',
'transkrip', 'ipk', 'ip'],
  'jadwal': ['jadwal', 'schedule', 'jam kuliah', 'waktu kuliah'],
  'revisi': ['revisi', 'perbaikan', 'koreksi', 'feedback'],
  'konsultasi': ['konsultasi', 'bimbingan', 'tanya', 'bertanya',
'bertemu'],
  'izin': ['izin', 'permohonan', 'mohon', 'maaf', 'permisi'],
}
}
```