

Deteksi Penipuan Suara Berbahasa Indonesia: Studi Komparatif Antara Arsitektur LALM dan ASR+LLM

M. Maulana Firdaus R.
Institut Teknologi Bandung
Bandung, Indonesia
maulanafirdaus63@gmail.com

Atqiya Haydar Luqman
Institut Teknologi Bandung
Bandung, Indonesia
atqiyahaydar15@gmail.com

M. Fauzan Azhim
Institut Teknologi Bandung
Bandung, Indonesia
mfauzan.az23@gmail.com

Abstract

Penipuan melalui percakapan telepon merupakan ancaman serius di Indonesia, di mana metode deteksi konvensional berbasis teks seringkali gagal menangkap isyarat vokal krusial. Penelitian ini melakukan studi komparatif pertama untuk mengevaluasi efektivitas pendekatan *Large Audio Language Models* (LALM) dibandingkan dengan pipeline ASR+LLM yang dioptimalkan untuk deteksi penipuan suara dalam konteks Bahasa Indonesia. Dua sistem dikembangkan dan di-*fine-tune* menggunakan QLoRA pada dataset sintesis berisi 1.034 percakapan: satu berbasis LALM (Qwen2-Audio-7B) dan satu lagi sebagai baseline (ASR-Whisper + LLM Sailor2-8B). Berlawanan dengan hipotesis awal, hasil pengujian menunjukkan bahwa sistem baseline ASR+LLM mencapai kinerja yang nyaris sempurna dengan F1-Score sebesar **99.03%**, secara signifikan mengungguli sistem LALM yang memperoleh **95.19%**. Studi ini menyimpulkan bahwa pipeline ASR+LLM, jika didukung oleh komponen SOTA dan LLM yang memiliki spesialisasi pada bahasa target, merupakan solusi yang lebih andal dan efisien untuk kasus ini. Temuan ini menyoroti pentingnya spesialisasi pre-training model dan kualitas sinyal tekstual, serta memberikan panduan berbasis bukti untuk pengembangan sistem deteksi penipuan di Indonesia.

1 Pendahuluan

1.1 Latar Belakang

Tindak penipuan telekomunikasi terus berkembang menjadi semakin kompleks dan canggih, menimbulkan ancaman serius bagi keamanan dan stabilitas ekonomi global. Kerugian finansial di seluruh dunia telah mencapai angka yang sangat besar, menggarisbawahi urgensi pengembangan metode deteksi yang lebih efektif [1].

Di Indonesia, situasi ini tidak jauh berbeda dan bahkan menunjukkan tren yang semakin mengkhawatirkan. Masyarakat terus dihadapkan pada berbagai modus penipuan siber yang semakin beragam, mulai dari rekayasa sosial melalui panggilan telepon hingga serangan teknis. Salah satu modus yang sangat meresahkan adalah penipuan melalui pengiriman file aplikasi berbahaya (APK) yang disamarkan sebagai undangan pernikahan, resi pengiriman paket, atau surat tilang [2]. Skala masalah ini tercermin dari data Kementerian Komunikasi dan Informatika (Kominfo) yang mencatat adanya 130.000 korban penipuan online pada tahun 2022 [3], serta data Kepolisian RI (Polri) yang menunjukkan bahwa penipuan online merupakan salah satu jenis kejahatan siber yang paling banyak dilaporkan [4].

Salah satu insiden yang menyoroti kerentanan terhadap penipuan berbasis data pribadi adalah kasus yang menimpa mahasiswa dan alumni Institut Teknologi Bandung (ITB). Sejumlah mahasiswa dan

054 alumni menerima panggilan telepon mencurigakan dari nomor tidak dikenal. Pihak ITB secara
055 resmi menanggapi isu ini, mengonfirmasi bahwa panggilan tersebut merupakan bagian dari modus
056 penipuan (*scam*) dan mengakui adanya dugaan kebocoran data yang kemungkinan dimanfaatkan
057 oleh pelaku [5–7]. Kasus ini menggarisbawahi bagaimana kebocoran data pribadi dapat menjadi
058 bahan bakar bagi pelaku penipuan untuk meluncurkan aksinya dengan lebih meyakinkan.

059 Kemajuan dalam *Large Language Models* (LLM) menawarkan solusi baru, namun pendekatan
060 berbasis teks murni (ASR+LLM) berisiko menghilangkan informasi paralinguistik. Perkembangan
061 *Large Audio Language Models* (LALM) menawarkan terobosan penting untuk mengatasi kelemahan ini [1, 8]. Oleh karena itu, penelitian ini berfokus pada pemanfaatan LALM untuk membangun sebuah **sistem pertahanan proaktif**. Tujuannya bukan hanya untuk mendeteksi, tetapi untuk **mencegah keberhasilan penipuan** dengan mengidentifikasi indikator risiko secara *real-time*.

066 1.2 Rumusan Masalah

067 Penelitian ini dirancang untuk menjawab pertanyaan-pertanyaan berikut:

- 069 1. Bagaimana perbandingan kinerja antara sistem deteksi berbasis LALM dengan sistem baseline ASR+LLM pada dataset audio sintesis yang bersih?
- 070 2. Faktor-faktor teknis apa (spesialisasi pre-training, konfigurasi pelatihan, modalitas data) yang memengaruhi perbedaan kinerja antara kedua arsitektur tersebut?
- 071 3. Bagaimana kinerja kedua arsitektur berubah saat dihadapkan pada data audio dunia nyata yang mengandung derau dan ketidaksempurnaan, dan apa implikasinya terhadap pemilihan arsitektur untuk implementasi riil?

078 1.3 Tujuan Penelitian

079 Tujuan dari penelitian ini adalah:

- 081 1. Mengimplementasikan dan melatih dua sistem deteksi penipuan: satu berbasis LALM (Qwen2-Audio) dan satu lagi sebagai baseline berbasis ASR+LLM (Whisper+Sailor2).
- 082 2. Melakukan evaluasi komparatif yang objektif terhadap kedua sistem pada data sintesis dan melakukan studi kasus kualitatif pada data riil.
- 083 3. Menganalisis secara mendalam pertukaran (*trade-off*) antara kedua arsitektur untuk memberikan rekomendasi berbasis bukti.

089 1.4 Manfaat Penelitian

090 Penelitian ini diharapkan memberikan manfaat sebagai berikut:

- 092 1. Menyediakan **analisis komparatif berbasis bukti pertama** antara pendekatan LALM dan ASR+LLM untuk deteksi penipuan dalam Bahasa Indonesia, memberikan panduan yang jelas bagi industri dan peneliti selanjutnya.
- 093 2. Mengidentifikasi bahwa **pipeline ASR+LLM yang dioptimalkan dengan model spesialis bahasa** merupakan solusi yang sangat kuat, efisien, dan dalam kasus ini, lebih unggul dari pendekatan multimodal yang lebih kompleks.
- 094 3. Menghasilkan wawasan tentang **tantangan nyata dalam penerapan LALM**, menyoroti pentingnya kualitas data audio dan spesialisasi pre-training model sebagai faktor krusial.
- 095 4. Memberikan **rekomendasi arsitektur yang paling andal (ASR+LLM)** untuk diimplementasikan sebagai sistem pertahanan proaktif, yang dapat membantu memitigasi risiko finansial dan operasional.

106 1.5 Batasan Penelitian

107 Penelitian ini memiliki beberapa batasan yang perlu diperhatikan:

1. **Validasi pada Data Sintetis:** Evaluasi kuantitatif sepenuhnya dilakukan pada set pengujian yang juga berasal dari data sintetis. Terdapat kemungkinan adanya *domain gap*, di mana kinerja model bisa berbeda saat dihadapkan pada data percakapan penipuan di dunia nyata.
2. **Generalisasi Skenario:** Kemampuan model terbatas pada skenario-skenario penipuan yang telah didefinisikan. Model kemungkinan akan kesulitan mengenali modus penipuan yang sepenuhnya baru.
3. **Fokus pada Kelayakan, Bukan Efisiensi:** Penelitian ini lebih menekankan pada pembuktian konsep. Aspek efisiensi komputasi untuk implementasi secara *real-time* tidak menjadi fokus utama.

2 Kajian Teori

2.1 Deteksi Penipuan Telekomunikasi

Deteksi penipuan telekomunikasi adalah proses identifikasi aktivitas penipuan yang terjadi melalui kanal komunikasi suara. Seiring meningkatnya kompleksitas modus penipuan, metode deteksi pun turut berevolusi. Metode deteksi tradisional yang mengandalkan verifikasi manual dan pencocokan pola berbasis aturan (*rule-based*) terbukti tidak lagi memadai karena sifatnya yang kaku dan sulit beradaptasi dengan taktik penipuan yang cepat berubah [1].

Pendekatan modern mulai memanfaatkan kecerdasan buatan, khususnya *Large Language Models* (LLM). Arsitektur yang umum digunakan adalah mengubah data audio menjadi teks menggunakan *Automatic Speech Recognition* (ASR), kemudian menganalisis teks tersebut dengan LLM [9]. Namun, pendekatan ini memiliki kelemahan fundamental, yaitu hilangnya informasi paralinguistik selama proses konversi ASR. Isyarat vokal yang krusial seperti nada suara, emosi, keraguan yang terdengar dari jeda bicara, atau penekanan pada kata-kata tertentu akan hilang, padahal informasi tersebut sering kali menjadi petunjuk utama adanya niat penipuan [1].

2.2 Large Audio Language Models (LALM)

Untuk mengatasi keterbatasan pada sistem berbasis teks, *Large Audio Language Models* (LALM) muncul sebagai sebuah terobosan. LALM adalah model skala besar yang mampu memproses sinyal audio secara langsung (*end-to-end*), memungkinkannya untuk menjembatani kesenjangan modalitas antara suara dan teks [1]. Dengan memproses audio mentah, LALM dapat menganalisis tidak hanya konten linguistik (apa yang dikatakan), tetapi juga konteks vokal (bagaimana sesuatu dikatakan).

Beberapa contoh LALM terkemuka yang menunjukkan kapabilitas canggih ini antara lain **Qwen2-Audio** yang dapat menerima beragam jenis masukan audio untuk dianalisis, dan **GLM-4-Voice** yang dirancang sebagai sistem dialog suara interaktif yang mampu menyesuaikan emosi dan nada suara [1, 8]. Kemampuan ini sangat relevan untuk tugas deteksi penipuan karena model dapat secara holistik mengevaluasi kewajaran sebuah percakapan.

3 Penelitian Terkait

Bab ini meninjau beberapa karya penelitian sebelumnya yang relevan di bidang deteksi penipuan menggunakan model bahasa. Tinjauan ini bertujuan untuk memetakan lanskap penelitian yang sudah ada, memahami pendekatan yang telah dieksplorasi, dan pada akhirnya memperjelas posisi serta kebaruan dari penelitian yang diusulkan.

Beberapa penelitian telah mengeksplorasi penggunaan LLM untuk deteksi penipuan berbasis teks. Singh et al. [10] mengusulkan mekanisme deteksi *real-time* menggunakan LLM yang diperkuat dengan *Retrieval-Augmented Generation* (RAG) untuk memverifikasi legitimasi percakapan. Sementara itu, Shen et al. [9] mengembangkan kerangka kerja berbasis LLM yang dapat memberikan peringatan dini kepada pengguna, dengan fokus pada analisis konten percakapan yang juga telah diubah menjadi teks.

Karya yang paling mendekati pemanfaatan LALM secara spesifik adalah yang dilakukan oleh Ma et al. [1]. Mereka memperkenalkan TeleAntiFraud-28k, dataset audio-teks pertama untuk deteksi

penipuan dan menunjukkan bahwa model LALM yang di-*fine-tune* pada data audio-teks menunjukkan peningkatan kinerja yang sangat signifikan dibandingkan model yang tidak dilatih secara khusus.

Meskipun penelitian-penelitian tersebut sangat foundational, sebuah kesenjangan riset yang jelas masih ada: hampir seluruh penelitian yang ada berfokus pada bahasa Inggris atau Mandarin. Hingga saat ini, belum ada studi yang secara khusus mengeksplorasi pengembangan dan adaptasi LALM untuk deteksi penipuan percakapan telepon dalam konteks Bahasa Indonesia. Kekosongan inilah yang menjadi motivasi utama dan mengisi kebaruan dari penelitian yang diusulkan ini.

4 Metodologi Penelitian

Alur kerja penelitian, seperti yang diilustrasikan pada Gambar 1, dibagi menjadi lima tahap utama: persiapan data sintesis, desain sistem dan pemilihan model, pelatihan model dengan QLoRA, protokol dan metrik evaluasi, serta implementasi antarmuka pengujian.

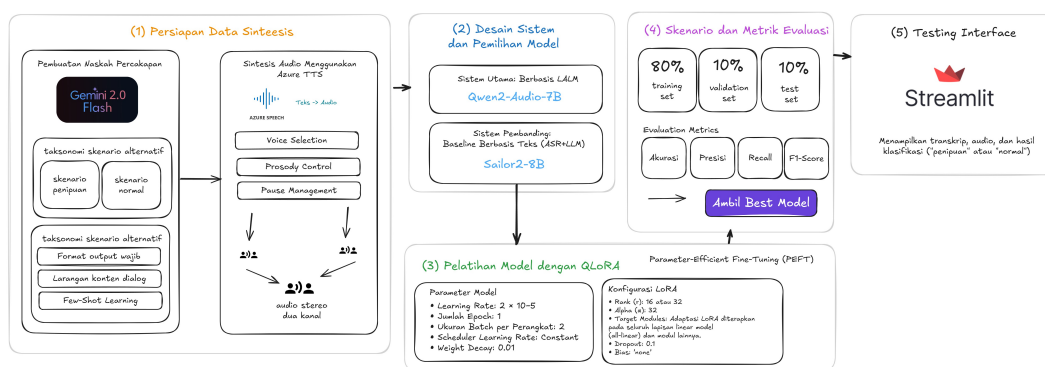


Figure 1: Diagram Alur Kerja Metodologi Penelitian

4.1 Tahap 1: Persiapan Data Sintesis

Mengingat kelangkaan dataset publik untuk domain ini dalam Bahasa Indonesia, langkah fundamental pertama adalah pembuatan dataset sintesis yang realistis dan beragam.

4.1.1 Pembuatan Naskah Percakapan dengan Gemini Flash

Naskah percakapan yang menjadi dasar dataset ini dibuat secara terprogram menggunakan model bahasa **Gemini 2.0 Flash**. Proses ini melibatkan:

- **Pengembangan Taksonomi Skenario:** Sebuah taksonomi detail untuk skenario penipuan (misal: Penipuan APK, Penipuan Perbankan) dan normal (misal: Layanan Janji Temu, Interaksi Keluarga) yang spesifik untuk konteks Indonesia telah dirancang.
- **Rekayasa Prompt Tingkat Lanjut:** Digunakan teknik *few-shot learning* di mana LLM diberikan peran sebagai "Mesin Transkripsi" dengan aturan keluaran yang sangat ketat untuk memastikan format dialog yang konsisten (Nama Peran: Ucapan) dan bebas dari artefak narasi.
- **Pembersihan Paska-Generasi:** Naskah yang dihasilkan oleh LLM kemudian melewati tahap pembersihan minor untuk memastikan tidak ada format atau teks yang tidak diinginkan sebelum disimpan ke dalam format CSV.

4.1.2 Sintesis Audio Menggunakan Azure TTS

Naskah teks kemudian dikonversi menjadi audio menggunakan layanan **Azure Cognitive Services for Speech**. SSML (*Speech Synthesis Markup Language*) dimanfaatkan secara ekstensif untuk:

- **Variasi Suara:** Menetapkan suara neural yang berbeda (id-ID-ArdiNeural dan id-ID-GadisNeural) untuk peran penipu dan korban.
- **Realisme Prosodi:** Mengontrol nada, kecepatan, dan volume bicara secara dinamis untuk menyimulasikan emosi seperti urgensi, kebingungan, atau keramahan yang manipulatif.

Hasil akhir dari tahap ini adalah kumpulan file audio stereo dua-kanal dengan format WAV, 16-bit, dan laju sampel 16kHz.

4.2 Tahap 2: Desain Sistem dan Pemilihan Model

Penelitian ini mengadopsi desain komparatif untuk mengukur secara objektif keunggulan pendekatan LALM. Dua sistem dibangun menggunakan model dengan skala parameter setara (7 Miliar).

4.2.1 Sistem Utama: Berbasis LALM

Sistem utama menggunakan model *open-source* **Qwen2-Audio-7B** [8]. Model ini dipilih karena kemampuannya sebagai LALM canggih yang dapat memproses input multimodal (audio dan teks) secara end-to-end.

4.2.2 Sistem Pembanding: Baseline Berbasis Teks (ASR+LLM)

Sebagai pembanding, dibangun sebuah arsitektur ASR+LLM. Penting untuk dicatat bahwa komponen ASR hanya digunakan pada **tahap pengujian (inferensi)**. Pada tahap pelatihan, model LLM dilatih langsung pada naskah teks asli untuk performa maksimal. Arsitektur ini terdiri dari:

- **Komponen ASR (Hanya untuk Pengujian):** Untuk mentranskripsi audio dari test set menjadi teks, digunakan model **Whisper** dari OpenAI.
- **Komponen LLM:** Model yang digunakan adalah **Sailor2-8B**. Model ini di-*fine-tune* langsung pada data naskah teks bersih yang dihasilkan pada Tahap 1.

4.3 Tahap 3: Pelatihan Model dengan QLoRA

Untuk efisiensi komputasi, kedua model diadaptasi menggunakan metode **QLoRA** (*Quantized Low-Rank Adaptation*). Implementasi teknisnya disesuaikan untuk setiap model.

4.3.1 Pelatihan Sistem LALM (Qwen2-Audio-7B)

Pelatihan dilakukan menggunakan library Hugging Face Transformers. Konfigurasi utamanya meliputi kuantisasi model dasar ke 4-bit ('nf4'), LoRA dengan rank (r) 32 dan alpha (α) 32 yang menargetkan semua lapisan linear, serta penggunaan *data collator* kustom untuk memformat input multimodal dan memfokuskan proses belajar pada prediksi token klasifikasi.

4.3.2 Pelatihan Sistem Baseline (Sailor2-8B)

Pelatihan diakselerasi menggunakan library **Unsloth**. Prosesnya melibatkan pemformatan teks dengan templat prompt spesifik, diikuti dengan adaptasi LoRA ($r=16$, $\alpha=32$) dan penggunaan *data collator* yang memastikan fungsi kerugian hanya dihitung pada token klasifikasi akhir.

Hyperparameter Pelatihan Utama Untuk menjaga konsistensi, hyperparameter pelatihan utama seperti *learning rate* (2×10^{-5}), jumlah *epoch* (1), dan *scheduler* ('constant') diterapkan secara setara pada kedua proses pelatihan.

4.4 Lingkungan Komputasi

Seluruh proses eksperimen, mulai dari pelatihan model hingga evaluasi, dilakukan menggunakan platform **Kaggle Notebooks** dengan spesifikasi perangkat keras dan lunak sebagai berikut:

- **CPU:** Intel(R) Xeon(R) CPU @ 2.20GHz, 2 cores, 4 threads

- **RAM:** 32 GB
- **GPU:** NVIDIA Tesla P100, 16 GB VRAM
- **Platform:** Kaggle Notebook (Kernel berbasis Ubuntu 20.04)
- **Library Utama:** Python 3.11, PyTorch 2.6, Transformers, PEFT, Unsloth, Librosa, Streamlit

Pemilihan lingkungan ini, khususnya dengan ketersediaan GPU NVIDIA Tesla P100, memungkinkan dilakukannya proses *fine-tuning* model skala 7B secara efisien menggunakan teknik QLoRA.

4.5 Tahap 4: Protokol dan Metrik Evaluasi

Evaluasi dilakukan secara ketat untuk mengukur dan membandingkan kinerja kedua sistem.

4.5.1 Protokol Pengujian dan Pembagian Data

Total dataset sintetis yang berhasil dibuat adalah sebanyak **1.034 sampel**. Dataset ini kemudian dibagi menjadi tiga set data yang independen dengan rasio **80% untuk pelatihan, 10% untuk validasi, dan 10% untuk pengujian**. Rincian distribusi untuk setiap set disajikan pada Tabel 1.

Table 1: Distribusi Sampel pada Set Pelatihan, Validasi, dan Pengujian

Jenis Set	Jumlah Sampel	Persentase	Kelas Normal	Kelas Penipuan
Training	827	80%	413	414
Validation	103	10%	52	51
Test	104	10%	52	52
Total	1034	100%	517	517

Setelah kedua model selesai di-*fine-tune*, performa final mereka dievaluasi hanya pada **Test Set** (104 sampel) yang datanya belum pernah dilihat sebelumnya oleh model. Distribusi kelas yang seimbang (*balanced*) dipertahankan di setiap set untuk memastikan proses evaluasi yang adil dan tidak bias terhadap kelas tertentu.

Pengujian Tambahan pada Data Riil. Selain pengujian kuantitatif pada data sintetis, sebuah studi kasus kualitatif dilakukan pada 1 sampel audio penipuan riil untuk menguji ketahanan (*robustness*) model terhadap kondisi audio dunia nyata.

4.5.2 Metrik Kuantitatif

Kinerja pada tugas determinasi penipuan diukur menggunakan metrik klasifikasi standar:

- **Akurasi:** Persentase prediksi yang benar secara keseluruhan.
- **Presisi:** Dari semua yang diprediksi sebagai "Penipuan", seberapa banyak yang benar-benar penipuan. Metrik ini penting untuk meminimalkan *false positive*.
- **Recall (Sensitivitas):** Dari semua kasus "Penipuan" yang asli, seberapa banyak yang berhasil dideteksi oleh model. Metrik ini krusial untuk meminimalkan *false negative*.
- **F1-Score:** Rata-rata harmonik dari Presisi dan Recall, memberikan gambaran performa yang seimbang.

4.5.3 Metrik Efisiensi

- **Waktu Inferensi:** Diukur sebagai waktu rata-rata (dalam detik) yang dibutuhkan setiap model untuk memproses satu sampel audio dari *test set*.

4.6 Tahap 5: Implementasi Antarmuka Pengujian

Sebagai bukti konsep dan untuk demonstrasi fungsional, prototipe sistem diimplementasikan dalam sebuah antarmuka web interaktif menggunakan **Streamlit**. Aplikasi ini memungkinkan pengguna untuk mengunggah file audio, kemudian sistem akan menjalankan inferensi pada kedua model (LALM dan baseline) secara paralel dan menampilkan hasil klasifikasi keduanya secara berdampingan untuk perbandingan langsung.

5 Hasil Eksperimen dan Pengujian

5.1 Perbandingan Kinerja Kuantitatif

Hasil perbandingan performa kedua sistem disajikan secara ringkas pada Tabel 2. Metrik Presisi, Recall, dan F1-Score yang ditampilkan adalah nilai rata-rata terbobot (*weighted average*) dari kedua kelas untuk perbandingan yang adil.

Table 2: Perbandingan Kinerja dan Efisiensi Model(Inferensi belum termasuk whisper)

Metrik	Baseline (ASR+Sailor2)	LALM (Qwen2-Audio)
Akurasi	99.03%	95.19%
Presisi	99.03%	95.54%
Recall	99.03%	95.19%
F1-Score	99.03%	95.19%
Waktu Inferensi (detik/sampel)	0.45s	5.3605

Untuk analisis yang lebih granular, Tabel 3 menyajikan laporan klasifikasi lengkap yang merinci performa pada setiap kelas ("Biasa/Normal" dan "Penipuan").

Table 3: Laporan Klasifikasi Detail untuk Kedua Model

Model	Kelas	Presisi	Recall	F1-Score
Baseline (ASR+Sailor2)	Normal	1.00	0.98	0.99
	Penipuan	0.98	1.00	0.99
LALM (Qwen2-Audio)	Normal	0.91	1.00	0.95
	Penipuan	1.00	0.90	0.95

5.2 Analisis Matriks Kebingungan

Untuk memberikan gambaran visual mengenai jenis kesalahan yang dibuat oleh masing-masing model, disajikan matriks kebingungan (*confusion matrix*) pada Gambar 2.



Figure 2: Matriks Kebingungan Kedua Model pada Test Set

Dari matriks tersebut, terlihat bahwa:

- **Sistem Baseline (ASR+Sailor2)** membuat total 1 kesalahan, yaitu salah mengklasifikasi 1 sampel "Biasa" sebagai "Penipuan" (*False Positive*). Sistem ini tidak membuat kesalahan dalam mendeteksi kasus penipuan (*zero False Negatives*).

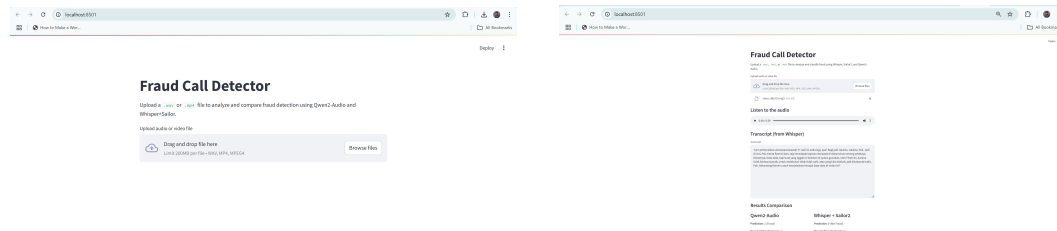
- **Sistem LALM (Qwen2-Audio)** membuat total 5 kesalahan, yaitu gagal mendeteksi 5 sampel "Penipuan" (*False Negatives*). Sistem ini sempurna dalam mengidentifikasi sampel "Normal" (*zero False Positives*).

Secara keseluruhan, data menunjukkan hasil yang berlawanan dengan hipotesis awal. Sistem baseline (ASR+Sailor2) secara tak terduga menunjukkan performa yang lebih tinggi dan lebih seimbang pada semua metrik utama dibandingkan dengan sistem LALM (Qwen2-Audio). Analisis mendalam mengenai temuan ini akan dibahas pada bab berikutnya.

5.3 Implementasi dan Verifikasi Fungsional Prototipe Streamlit

Sebagai bukti konsep (*proof-of-concept*), prototipe antarmuka web yang dirancang pada Tahap 5 Metodologi telah berhasil diimplementasikan menggunakan framework **Streamlit**. Aplikasi ini menyediakan platform untuk demonstrasi dan analisis interaktif.

Pengujian fungsional menunjukkan bahwa semua komponen bekerja sesuai rencana. Fungsionalitas utama aplikasi diilustrasikan pada Gambar 3. Gambar 3a menunjukkan tampilan awal di mana pengguna dapat mengunggah file audio, sementara Gambar 3b menampilkan antarmuka setelah audio diproses, yang menyajikan hasil klasifikasi dari kedua model secara berdampingan.



(a) Tampilan awal antarmuka aplikasi.

(b) Tampilan hasil setelah audio diproses.

Figure 3: Tampilan Antarmuka Prototipe Aplikasi Pengujian Berbasis Streamlit

Alur kerja aplikasi yang telah terverifikasi adalah sebagai berikut:

- Aplikasi mampu menerima dan memproses unggahan file audio dari pengguna.
- Kedua model yang telah di-*fine-tune*—sistem LALM dan sistem baseline—berhasil dimuat dan dapat menjalankan proses inferensi.
- Hasil prediksi dari kedua model, beserta transkrip teks dari komponen ASR, berhasil ditampilkan pada antarmuka pengguna untuk dianalisis.

Keberhasilan implementasi ini berfungsi sebagai validasi fungsional dari keseluruhan alur kerja penelitian, dari data hingga aplikasi interaktif, dan telah digunakan secara efektif sebagai alat bantu untuk melakukan analisis kualitatif.

5.4 Studi Kasus Tunggal pada Data Riil

Selain pengujian kuantitatif, sebuah studi kasus mendalam dilakukan pada **satu sampel audio penipuan riil** untuk menguji kemampuan generalisasi model pada kondisi dunia nyata yang mengandung derau (*noise*) dan cara bicara yang tidak ideal.

Hasil dari studi kasus tunggal ini menyoroti kelemahan krusial pada pipeline ASR+LLM:

- **Sistem Baseline (ASR+Sailor2):** Gagal mendeteksi penipuan (Prediksi: Normal).
- **Sistem LALM (Qwen2-Audio):** Berhasil mendeteksi penipuan (Prediksi: Penipuan).

Kegagalan sistem baseline berakar pada kualitas transkripsi yang dihasilkan oleh ASR Whisper. Kinerjanya pada Bahasa Indonesia untuk percakapan non-standar terbukti suboptimal, menghasilkan transkrip yang secara semantik tidak koheren.

Transkrip Aktual Hasil ASR (Whisper) pada Data Riil:

turn perkenalkan-amaawawawawah !!! Jadi ini anda lagi, apa? Bagi jadi Jakarta. Jakarta, Pak. Jadi di sini, Pak, Kanta Ramini baru saja mendapat laporan daripada di dokamanan omong selabaya, Dekaitnya, Data-data, tapi buat yang nggak ini kelahan di syalan gunakan, oleh Tihah No, karena tidak bertawa jawab, untuk melakukan tidak tidak naik, atau yang kita silaikh, jadi dikakanda hadir, Pak, Nekanteng Ramini untuk menjelaskan kenapa Data-data di randa ini?

Transkrip yang dihasilkan di atas dipenuhi dengan kata-kata yang tidak relevan dan struktur kalimat yang rusak. Teks ini kemudian diumpankan ke model Sailor2, yang tentu saja tidak dapat menemukan pola penipuan dari input yang secara semantik sudah hancur. Sebaliknya, keberhasilan LALM menunjukkan kemampuannya bekerja tanpa bergantung pada kesempurnaan transkripsi.

6 Analisis Hasil

6.1 Interpretasi Metrik Kinerja Detail

Analisis mendalam pada setiap metrik memberikan gambaran yang lebih jernih tentang perilaku dan karakteristik masing-masing model.

6.1.1 Interpretasi Laporan Klasifikasi

Laporan klasifikasi (Tabel 3) mengungkapkan "kepribadian" dari kedua model:

- **Sistem Baseline (ASR+Sailor2):** Model ini menunjukkan kinerja yang sangat seimbang dan nyaris sempurna. Dengan **Recall 1.00** untuk kelas "Penipuan", artinya model ini berhasil **mendeteksi semua kasus penipuan** yang ada di test set tanpa ada yang terlewat (*zero False Negatives*). Nilai **Presisi 0.98** untuk kelas yang sama menunjukkan bahwa model ini sangat andal, meskipun ada satu kasus di mana ia salah menandai percakapan normal sebagai penipuan. Ini adalah karakteristik model yang "agresif melindungi".
- **Sistem LALM (Qwen2-Audio):** Model ini menunjukkan karakteristik yang berbeda. Dengan **Presisi 1.00** untuk kelas "Penipuan", artinya setiap kali model ini menandai sebuah panggilan sebagai penipuan, prediksinya selalu benar (*zero False Positives*). Namun, kelemahannya terletak pada **Recall 0.90**, yang berarti ia **gagal mendeteksi 5 dari 52 kasus penipuan**. Ini adalah karakteristik model yang "konservatif", yang hanya akan bertindak jika bukti penipuannya sangat kuat, sehingga menghindari alarm palsu namun dengan risiko melewatkan beberapa kasus.

6.1.2 Interpretasi Matriks Kebingungan

Matriks kebingungan (Gambar 2) memvisualisasikan jenis kesalahan ini secara langsung.

- **Kesalahan Sailor2 (1 FP, 0 FN):** Satu-satunya kesalahan model ini adalah *False Positive*. Dari perspektif bisnis dan keamanan pengguna, kesalahan ini cenderung lebih bisa ditoleransi. Sebuah alarm palsu mungkin sedikit mengganggu pengguna, namun jauh lebih baik daripada membiarkan satu pun kasus penipuan lolos tanpa terdeteksi.
- **Kesalahan Qwen2 (0 FP, 5 FN):** Kelima kesalahan model ini adalah *False Negative*. Ini adalah jenis kesalahan yang paling kritis dalam sistem deteksi penipuan. Artinya, ada 5 pengguna dalam skenario pengujian ini yang akan menjadi korban karena sistem gagal memberikan peringatan. Meskipun model ini tidak pernah mengganggu pengguna dengan alarm palsu, keagalannya dalam mendeteksi penipuan yang nyata merupakan kelemahan yang signifikan.

6.2 Interpretasi Waktu Inferensi

Data waktu inferensi menunjukkan adanya pertukaran (*trade-off*) yang jelas antara performa dan efisiensi.

- Waktu inferensi LALM (**5.36 detik**) terbukti secara signifikan lebih lambat dibandingkan komponen LLM pada sistem baseline (**0.45 detik**).
- Waktu inferensi Sailor2 memiliki waktu yang lebih cepat yaitu 0.45s. Yang mana model ini mampu mengklasifikasikan suatu percakapan dengan cepat dan tepat dibanding LALM.

6.3 Analisis Kinerja Komparatif

Berlawanan dengan hipotesis awal, hasil eksperimen dengan jelas menunjukkan bahwa sistem base-line berbasis teks (ASR+Sailor2) secara signifikan **mengungguli** sistem LALM (Qwen2-Audio). Dengan F1-Score mencapai **99.03%** dibandingkan **95.19%** pada LALM, sistem yang secara arsitektur lebih sederhana ini justru menunjukkan kinerja yang mendekati sempurna.

Temuan ini mengindikasikan bahwa untuk dataset sintesis yang dihasilkan, **sinyal tekstual lebih dominan dan informatif** daripada sinyal audio. Kemungkinan besar, naskah penipuan yang dibuat memiliki pola linguistik yang sangat jelas (frasa kunci, struktur kalimat) yang dapat dengan mudah ditangkap oleh model bahasa yang kuat. Sementara itu, sinyal audio dari TTS, meskipun natural, mungkin tidak memiliki nuansa paralinguistik yang cukup kaya atau bahkan berpotensi menjadi "distraksi" bagi model LALM.

6.4 Analisis Faktor Teknis yang Mempengaruhi Hasil

Perbedaan performa yang signifikan ini dapat diatribusikan pada beberapa faktor teknis yang berbeda dalam proses pelatihan kedua model.

6.4.1 Spesialisasi Model Dasar (Pre-training)

Faktor paling fundamental adalah latar belakang pre-training kedua model.

- **Sailor2-8B**: Model ini secara spesifik telah melalui proses pre-training pada korpus data yang besar dari kawasan **Asia Tenggara (SEA)**, termasuk Bahasa Indonesia. Ini memberikannya "keunggulan awal" (*head start*) dengan pemahaman yang mendalam terhadap tata bahasa, kosakata, dan konteks kultural Bahasa Indonesia.
- **Qwen2-Audio-7B**: Meskipun merupakan model LALM yang sangat kuat, basis pengetahuannya lebih bersifat umum dan global, tanpa spesialisasi regional yang sama dengan Sailor2.

Karena Sailor2 sudah "familiar" dengan bahasa target, proses *fine-tuning* menjadi lebih efektif karena hanya perlu beradaptasi pada tugas spesifik klasifikasi penipuan, bukan belajar bahasa dari dasar.

6.4.2 Perbedaan Konfigurasi LoRA dan Parameter yang Dilatih

Terdapat perbedaan signifikan dalam konfigurasi adaptasi LoRA untuk kedua model:

- **Sailor2-8B**: Menggunakan rank (r) yang lebih rendah yaitu **16**, dengan total parameter yang dilatih hanya sekitar **46 juta (0.57%)**.
- **Qwen2-Audio-7B**: Menggunakan rank (r) yang lebih tinggi yaitu **32**, dengan jumlah parameter yang dilatih lebih dari dua kali lipat, yaitu sekitar **108 juta (1.28%)**.

Meskipun secara intuisi parameter yang lebih banyak seharusnya memberikan kapasitas belajar yang lebih besar, hasil yang terjadi justru sebaliknya. Ini memperkuat argumen spesialisasi pre-training: Sailor2 tidak memerlukan adaptasi yang masif (cukup dengan $r=16$) karena fondasinya sudah kuat. Sebaliknya, Qwen2 memerlukan adaptasi yang lebih besar ($r=32$) untuk mengejar pemahaman konteks, namun tetap tidak dapat mengalahkan keunggulan pre-training dari Sailor2.

6.4.3 Perbedaan Ukuran Batch Pelatihan

Ukuran batch juga menjadi faktor pembeda yang penting.

- **Sailor2-8B**: Dilatih dengan *batch size* sebesar 8.
- **Qwen2-Audio-7B**: Dilatih dengan *batch size* yang jauh lebih kecil, yaitu 2, kemungkinan besar karena keterbatasan memori VRAM saat memproses data audio yang kompleks.

Ukuran batch yang lebih besar pada pelatihan Sailor2 memungkinkan estimasi gradien yang lebih stabil dan akurat pada setiap langkah pembaruan, yang secara umum dapat mengarah pada konvergensi model yang lebih baik dan hasil akhir yang lebih robust.

6.4.4 Perbedaan Skala Model (8B vs 7B)

Faktor lain yang perlu dipertimbangkan adalah perbedaan ukuran antara kedua model. **Sailor2 (8B)** memiliki sekitar 1 miliar parameter lebih banyak daripada **Qwen2-Audio (7B)**. Meskipun perbedaannya tidak masif, kapasitas model yang lebih besar secara teoretis memberikan Sailor2 keunggulan dalam hal kemampuan representasi dan penghafalan pola. Keunggulan kapasitas ini, ketika dikombinasikan dengan spesialisasi pre-training dan kondisi pelatihan yang ideal, kemungkinan besar turut berkontribusi pada skor kinerjanya yang lebih tinggi.

6.5 Analisis Studi Kasus Riil: Keterbatasan ASR pada Bahasa Lokal

Studi kasus pada satu sampel data riil mengungkap narasi yang berbeda dan menyoroti kelemahan fundamental dari pipeline ASR+LLM saat berhadapan dengan bahasa non-Inggris dalam konteks non-formal.

- **ASR sebagai Titik Kegagalan Utama**: Kegagalan sistem baseline bukan disebabkan oleh derau audio semata, melainkan oleh **keterbatasan inheren model ASR Whisper dalam mentranskripsi Bahasa Indonesia secara akurat**. Seperti yang ditunjukkan pada hasil transkripsi, output yang dihasilkan bukanlah teks yang tidak lengkap, melainkan teks yang secara semantik rusak dan penuh dengan "halusinasi" kata. Teks yang tidak koheren ini membuat model Sailor2 mustahil untuk melakukan penalaran dan menemukan pola penipuan yang sebenarnya ada.
- **Keunggulan Arsitektur End-to-End LALM**: Sebaliknya, keberhasilan LALM pada kasus yang sama ini secara kuat mengindikasikan keunggulan arsitektur *end-to-end*. Dengan memproses sinyal audio secara langsung, LALM tidak terpengaruh oleh kegagalan pada tahap transkripsi. Model ini kemungkinan besar berhasil menangkap isyarat kecurigaan langsung dari fitur audio seperti intonasi, ritme bicara, atau jeda yang manipulatif, sebuah informasi yang hilang dan digantikan oleh "sampah semantik" pada pipeline ASR+LLM.

Studi kasus ini secara nyata mengilustrasikan bahwa secanggih apapun komponen LLM (seperti Sailor2), kinerjanya akan selalu dibatasi oleh kualitas input yang diterimanya. Dalam hal ini, komponen ASR menjadi "mata rantai terlemah" yang membuat seluruh pipeline berbasis teks menjadi rapuh.

6.6 Implikasi Hasil Penelitian

Temuan-temuan ini memiliki implikasi penting yang melampaui sekadar perbandingan dua model:

1. **Pentingnya Model Berbasis Regional**: Penelitian ini menunjukkan bahwa untuk tugas yang sangat bergantung pada nuansa bahasa lokal, model yang memiliki spesialisasi pre-training pada data regional (seperti Sailor2 untuk Bahasa Indonesia) dapat mengungguli model generalis yang lebih besar atau bahkan multimodal.
2. **Pipeline ASR+LLM Masih Sangat Relevan**: Jangan meremehkan arsitektur ASR+LLM. Jika didukung oleh komponen ASR SOTA (seperti Whisper) dan LLM spesialis bahasa, pipeline ini terbukti sangat kompetitif, efisien, dan bahkan bisa lebih unggul untuk kasus di mana sinyal tekstual sangat kuat.
3. **Arah Riset LALM untuk Bahasa Non-Inggris**: Untuk memaksimalkan potensi LALM pada bahasa seperti Bahasa Indonesia, proses *fine-tuning* sederhana mungkin tidak cukup. Diperlukan riset lebih lanjut pada teknik adaptasi domain yang lebih canggih, atau bahkan kelanjutan pre-training LALM pada korpus audio-teks Bahasa Indonesia yang besar untuk membangun pemahaman fundamental yang setara dengan model teks spesialis.

7 Kesimpulan dan Saran

7.1 Kesimpulan

Penelitian ini telah berhasil merancang, melatih, dan mengevaluasi dua prototipe sistem deteksi penipuan telekomunikasi. Kesimpulan utama dari studi ini bersifat dualistik dan menyoroti adanya pertukaran (*trade-off*) krusial antara kinerja di lingkungan ideal dan ketahanan di lingkungan dunia nyata.

1. **Pada Data Sintetis, Pipeline ASR+LLM Unggul Mutlak:** Dalam lingkungan pengujian yang terkontrol menggunakan data audio sintetis yang bersih, sistem baseline **ASR+LLM (Whisper+Sailor2)** menunjukkan kinerja yang nyaris sempurna (**F1-Score 99.03%**). Keunggulan ini diatribusikan pada kombinasi kuat antara spesialisasi pre-training model Sailor2 pada Bahasa Indonesia dan kondisi pelatihannya yang ideal pada naskah teks asli yang bersih.
2. **Pada Data Riil, LALM Menunjukkan Potensi Ketahanan yang Lebih Baik:** Narasi berbalik pada studi kasus tunggal menggunakan data penipuan riil. Sistem baseline gagal total karena komponen **ASR (Whisper)** tidak mampu menghasilkan transkrip yang akurat dari audio yang mengandung ketidaksempurnaan khas dunia nyata. Kegagalan ini melumpuhkan model Sailor2 yang sangat bergantung pada input teks berkualitas. Sebaliknya, sistem **LALM (Qwen2-Audio)** berhasil mendeteksi penipuan pada kasus yang sama, mengindikasikan bahwa arsitektur *end-to-end* memiliki potensi ketahanan (*robustness*) yang lebih superior karena tidak memiliki titik kegagalan tunggal pada tahap transkripsi.

7.2 Saran untuk Penelitian Selanjutnya

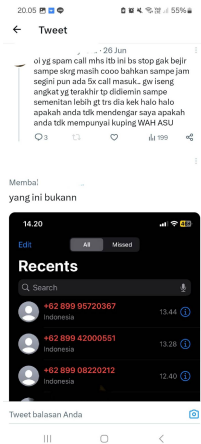
Berdasarkan temuan ini, berikut adalah beberapa arah penelitian strategis yang direkomendasikan:

1. **Validasi Skala Besar pada Data Riil:** Langkah paling krusial adalah melakukan pengujian ekstensif pada dataset riil yang lebih besar dan beragam. Hal ini diperlukan untuk memvalidasi secara kuantitatif temuan awal dari studi kasus bahwa LALM memang lebih tangguh, dan untuk memahami pada titik mana penurunan kualitas ASR mulai secara signifikan merusak kinerja pipeline baseline.
2. **Pengembangan Arsitektur Hibrida:** Mengingat setiap arsitektur memiliki keunggulan-nya masing-masing, merancang sistem hibrida adalah arah yang sangat menjanjikan. Misalnya, sebuah sistem dapat menggunakan ASR+LLM yang cepat untuk penyaringan awal. Jika kualitas transkripsi terdeteksi rendah (misalnya, dengan metrik kepercayaan ASR) atau hasilnya ambigu, barulah sampel tersebut diteruskan ke LALM untuk analisis audio yang lebih mendalam.
3. **Meningkatkan Komponen "Mata Rantai Terlemah":** Daripada meninggalkan pipeline ASR+LLM, penelitian dapat difokuskan untuk memperkuat komponen ASR-nya. Melakukan *fine-tuning* model Whisper atau ASR lainnya secara khusus pada dataset audio telepon yang bising dan beragam dalam Bahasa Indonesia dapat secara signifikan meningkatkan kinerja keseluruhan dari pipeline tersebut.
4. **Peningkatan Realisme Data Sintetis untuk LALM:** Untuk meningkatkan kinerja LALM pada data bersih, penelitian selanjutnya dapat fokus pada teknik sintesis audio yang lebih canggih, yang mampu mereplikasi tidak hanya derau, tetapi juga nuansa emosi manipulatif, keraguan, dan disfluensi yang khas dalam percakapan penipuan riil.

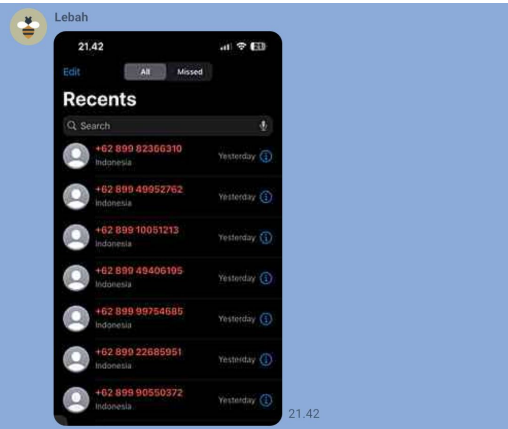
A Dokumentasi Panggilan Spam yang Diterima

Berikut adalah beberapa dokumentasi bukti panggilan tidak dikenal (spam) yang diterima oleh penulis dan rekan-rekan selama periode penelitian. Panggilan-panggilan ini menjadi salah satu motivasi studi kasus dan menunjukkan pola serangan yang nyata. Terlihat bahwa mayoritas panggilan spam berasal dari nomor dengan prefiks yang serupa, terutama '+62 899'.

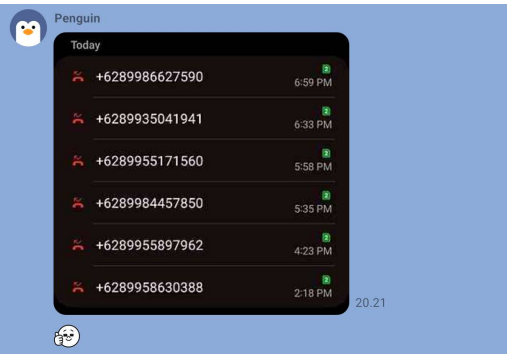
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701



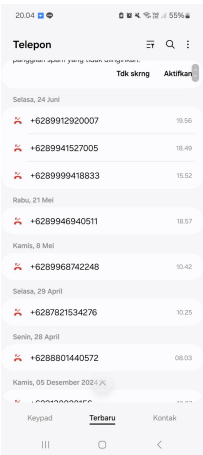
(a) Keluhan di media sosial disertai riwayat panggilan spam yang diterima.



(b) Contoh riwayat panggilan tak terjawab dari nomor-nomor tidak dikenal.



(c) Contoh lain dari serangkaian panggilan spam dalam satu hari.



(d) Riwayat panggilan spam dari berbagai tanggal dengan prefiks nomor yang berbeda.

Figure 4: Kumpulan bukti dokumentasi panggilan spam yang diterima penulis dan rekan.

B Contoh Sampel Dataset Sintetis

Berikut adalah dua contoh sampel data yang dihasilkan dari proses persiapan data, masing-masing merepresentasikan percakapan normal dan percakapan penipuan.

B.1 Contoh 1: Percakapan Normal

ID : 1001
Kategori Skenario : Layanan Pesan Antar Makanan
Detail Skenario : Kurir mengonfirmasi alamat pengantaran.
Label : 0 (Normal)

Transkrip Dialog:

Penelepon: Selamat siang, dengan Bapak Andi?
Penerima: Iya, betul.
Penelepon: Saya kurir GoFood, Pak. Mau konfirmasi alamatnya

di Jalan Cendrawasih nomor 15, betul?
 Penerima: Betul, Mas. Nanti rumahnya yang pagarnya warna hitam ya.
 Penelepon: Oh, siap Pak. Ini saya sudah dekat. Terima kasih.
 Penerima: Oke, sama-sama.

B.2 Contoh 2: Percakapan Penipuan

ID : 6
Kategori Skenario : Penipuan Layanan Pelanggan
Detail Skenario : Staf e-commerce palsu menginformasikan masalah pesanan.
Label : 1 (Penipuan)

Transkrip Dialog:

Penipu: Selamat pagi, Bu Aini. Saya Rina dari XYZShop.
 Korban: Halo, ya?
 Penipu: Ini mengenai pesanan Anda untuk tanggal 15 Oktober dengan nomor REF-XYZ45678. Kami mendeteksi masalah kecil - karton pengiriman rusak saat tiba di gudang kami.
 Korban: Oh tidak, bagaimana sekarang?
 Penipu: Jangan khawatir, Bu Aini. Untuk mempercepat proses perbaikan dan mengganti produk Anda, silakan akses tautan berikut: t.co/DummyLinkInstall.
 Korban: Apa... aplikasi? Mengapa bukan langsung diperbaiki saja?
 Penipu: Prosedur baru kami memerlukan akses ini, Bu Aini, demi keamanan data Anda. Harap instal dalam waktu 15 menit untuk verifikasi dan tindakan selanjutnya.
 Korban: Baiklah, saya unduh. Tetapi tolong pastikan ini aman.
 Penipu: Terima kasih, Bu Aini! Setelah instalasi selesai, Anda akan menerima konfirmasi dari tim kami dalam beberapa menit. Jika ada masalah, silakan hubungi 08123456789.
 Selamat hari!

Korban: Oke, terima kasih... semoga cepat selesai.

References

- [1] Z. Ma, P. Wang, K. Wu, M. Huang, J. Wang, Y. Yang, X. Lv, W. Tang, Y. Pang, and Y. Kang, "Teleantifraud-28k: An audio-text slow-thinking dataset for telecom fraud detection," 2025.
- [2] Desa Limbangan Kendal, "Awas hati-hati penipuan kiriman undangan dalam bentuk file apk," Jan 2023, [Online; diakses 7 Juni 2025]. [Online]. Available: <http://limbangan-kendal.desa.id/kabardetail/QmhTVStyQ0YrS1QwdTR2YVZkZ3dXZz09/awa-s-hati-hati-penipuan-kiriman-undangan-dalam-bentuk-file-apk.html>
- [3] InfoPublik.id, "Catatan kominfo: Korban penipuan online capai 130 ribu pada 2022," Feb 2023, [Online; diakses 7 Juni 2025]. [Online]. Available: <https://infopublik.id/kategori/nasional-sosial-budaya/715547/catatan-kominfo-korban-penipuan-online-capai-130-ribu-pada-2022>
- [4] PatroliSiber.id, "Statistik kejahatan siber," 2025, [Online; diakses 7 Juni 2025]. [Online]. Available: <https://patroliSiber.id/statistic/>
- [5] Tempo.co, "Ramai mahasiswa dan alumni ditelepon nomor berawalan 0888*, itb telusuri dugaan kebocoran data," Apr 2025, [Online; diakses 7 Juni 2025]. [Online]. Available: <https://www.tempo.co/digital/ramai-mahasiswa-dan-alumni-ditelepon-nomor-berawalan-0888-itb-telusuri-dugaan-kebocoran-data-1304645>
- [6] CNN Indonesia, "Itb sebut data mahasiswa dan alumni bocor diteror penipuan," Apr 2025, [Online; diakses 7 Juni 2025]. [Online]. Available: <https://www.cnnindonesia.com/nasional/20250430235534-20-1224504/itb-sebut-data-mahasiswa-dan-alumni-bocor-diteror-penipuan>

756 [7] Institut Teknologi Bandung, “Tanggapan itb terhadap dugaan kebocoran data dan panggilan
757 spam kepada mahasiswa dan alumni,” Apr 2025, [Online; diakses 7 Juni 2025]. [Online].
758 Available: [https://itb.ac.id/berita/tanggapan-itb-terhadap-dugaan-kebocoran-data-dan-pang-](https://itb.ac.id/berita/tanggapan-itb-terhadap-dugaan-kebocoran-data-dan-panggilan-spam-kepada-mahasiswa-dan-alumni/62290)
759 [gilan-spam-kepada-mahasiswa-dan-alumni/62290](https://itb.ac.id/berita/tanggapan-itb-terhadap-dugaan-kebocoran-data-dan-panggilan-spam-kepada-mahasiswa-dan-alumni/62290)
760 [8] Y. Chu, J. Xu, Q. Yang, H. Wei, X. Wei, Z. Guo, Y. Leng, Y. Lv, J. He, J. Lin *et al.*, “Qwen2-
761 audio technical report,” 2024.
762 [9] Z. Shen, S. Yan, Y. Zhang, X. Luo, G. Ngai, and E. Y. Fu, “‘it warned me just at the right
763 moment’: Exploring llm-based real-time detection of phone scams,” 2025.
764 [10] G. Singh, P. Singh, and M. Singh, “Advanced real-time fraud detection using rag-based llms,”
765 2025.
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809