IndoBERT-NER Gold: SOTA Indonesian Named Entity Recognition (NER)
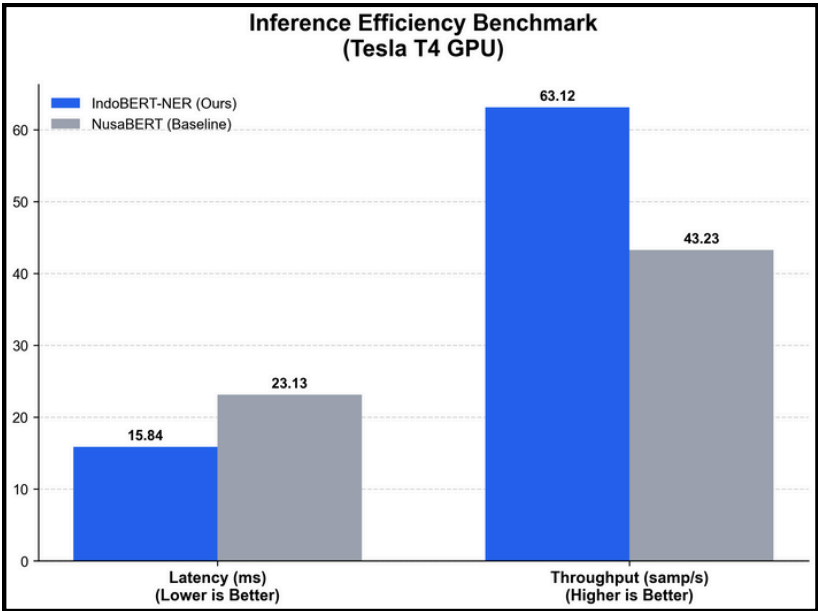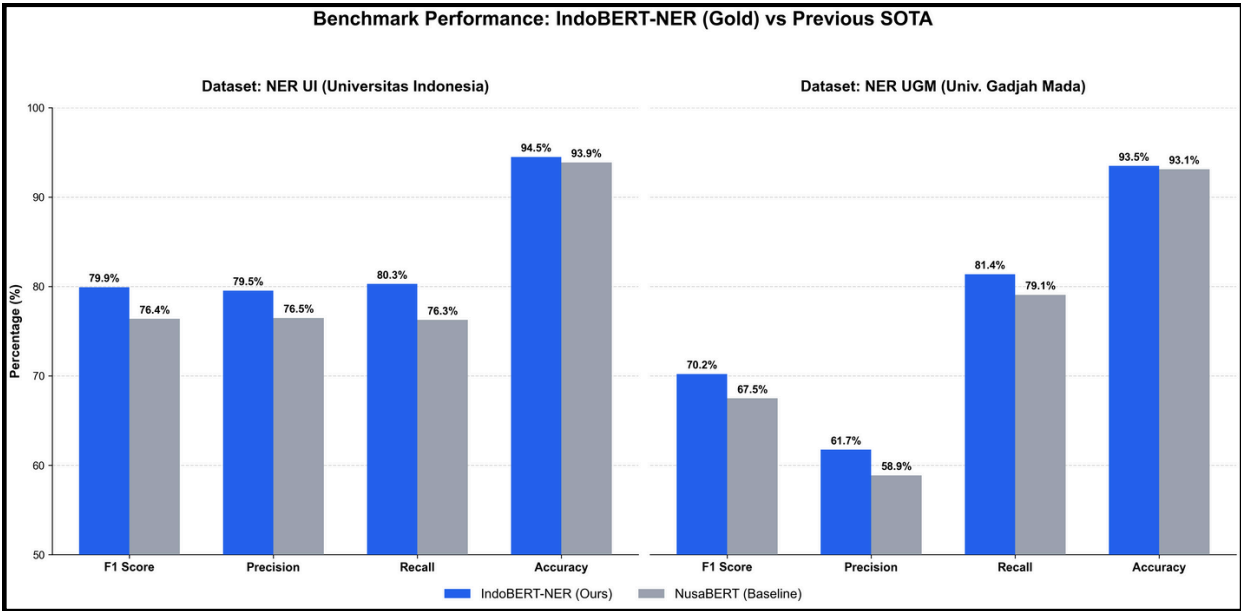
Year Accomplished
2025

Role/Position
NLP Research Engineer & Data Engineer

Publication Link
https://huggingface.co/treamyracle/indobert-ner-gold

A State-of-the-Art (SOTA) Large-scale Named Entity Recognition model for Indonesian, engineered using a novel Two-Stage Curriculum Learning strategy to outperform existing benchmarks.

- The Challenge: Existing Indonesian NER models (like NusaBERT) are often limited to 3 basic labels (PER, ORG, LOC) or trained on small datasets, resulting in low recall and an inability to capture rich semantic contexts like Events or Facilities.
- The Engineering Solution: Designed a "Silver-to-Gold" training pipeline. Phase 1 involved "warming up" the model on 130k synthetic sentences (Silver Data) to learn 19 diverse entity tags. Phase 2 focused on rigorous fine-tuning using the human-annotated NERGRIT corpus (Gold Data) to refine grammatical precision and boundary detection.
- The "Win-Win" Result: Achieved a 79.91% F1 Score on the standard NER_UI benchmark (+3.54% higher than the previous SOTA), proving that synthetic pre-training significantly boosts downstream performance.
- Hardware Optimization: Despite utilizing a "Large" architecture (334M params) for deeper understanding, the model achieves a throughput of 63 samples/sec (46% higher capacity than competitor Base models) and a latency of 15.84 ms, making it highly efficient for production environments.

Benchmark Performance: IndoBERT-NER (Gold) vs Previous SOTA

Dataset: NER UI (Universitas Indonesia) — F1 Score 79.9% / 76.4%; Precision 79.5% / 76.5%; Recall 80.3% / 76.3%; Accuracy 94.5% / 93.9%

Dataset: NER UGM (Univ. Gadjah Mada) — F1 Score 70.2% / 67.5%; Precision 61.7% / 58.9%; Recall 81.4% / 79.1%; Accuracy 93.5% / 93.1%

IndoBERT-NER (Ours) / NusaBERT (Baseline)

Inference Efficiency Benchmark (Tesla T4 GPU)

IndoBERT-NER (Ours) / NusaBERT (Baseline)

Latency (ms) (Lower is Better): 15.84 / 23.13
Throughput (samp/s) (Higher is Better): 63.12 / 43.23

Name
Arif Athaya Harahap

University
Student of University Brawijaya

Contact information
arifathayaharahap@gmail.com

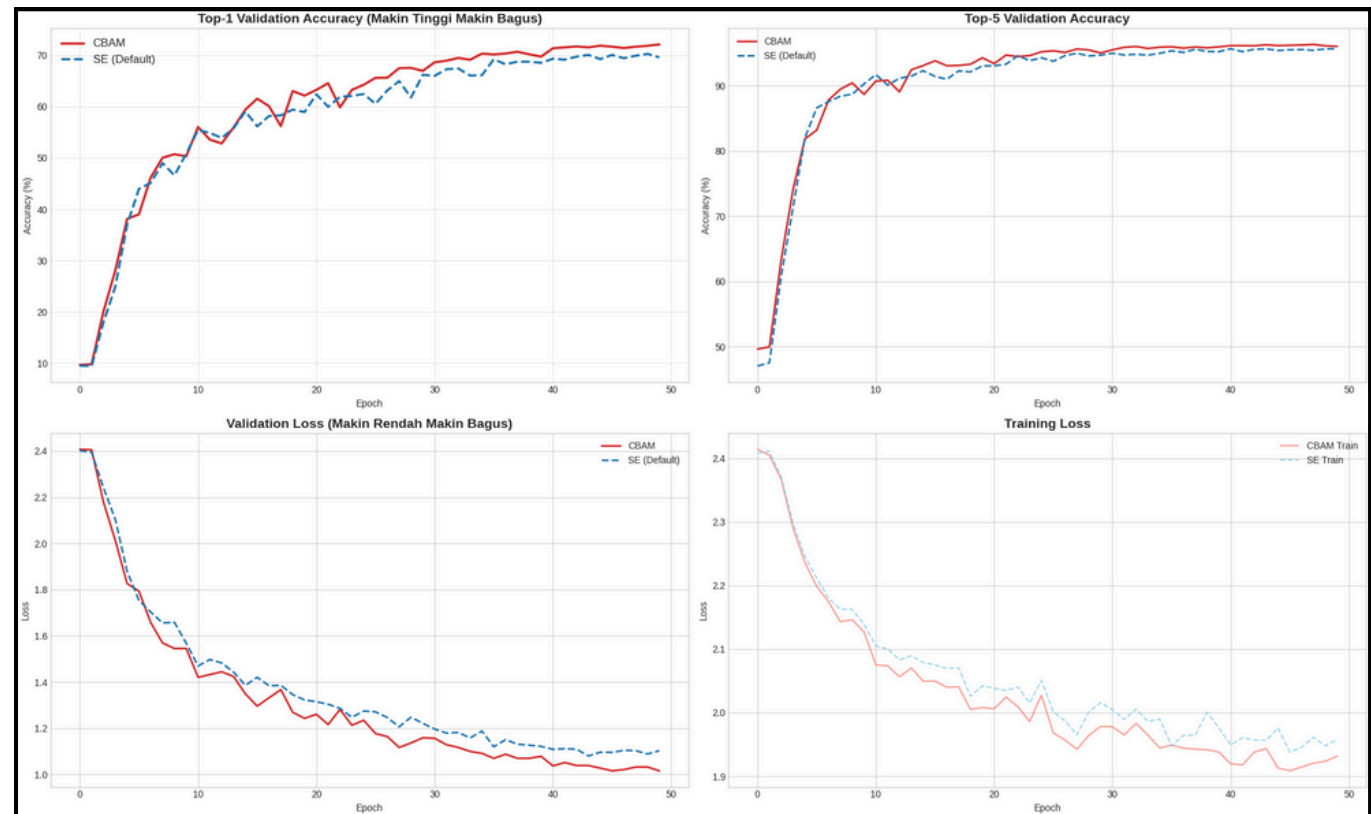# RepViT-CBAM: Hardware-Aware Mobile Vision Optimization

Year Accomplished
2025

Role/Position
Computer Vision Research Engineer

Publication Link
https://github.com/Tream yracle/RepViT_Enhanced

A custom-architected lightweight vision model designed to overcome the spatial limitations of standard RepViT architectures on edge devices.

- The Challenge: Identified that the default Squeeze-and-Excitation (SE) blocks in RepViT discarded critical spatial information ("where" an object is) via Global Average Pooling, limiting localization
- The Engineering Solution: Re-engineered the architecture by replacing SE blocks with Convolutional Block Attention Modules (CBAM) ($r=4$) to enforce both channel and spatial attention
- The "Win-Win" Result: Achieved a 1.83% increase in Top-1 Accuracy (70.27% to 72.10%) on Imagenette.
- Hardware Optimization: Contrary to theoretical complexity, the modified model reduced inference latency by ~0.3ms (15% speedup) on iPhone 15 (CoreML). This proved that Apple Neural Engine (ANE) is better optimized for CBAM's structural convolutions than SE's pooling operations5.

Benchmark Result
iPhone 15 Neural Engine

SE (Baseline)
2.101 ms
Rata-rata per gambar

CBAM (Ours)
1.792 ms
Rata-rata per gambar

Mulai Test

Selesai (Average of 100 runs)

Name
Arif Athaya Harahap

University
Student of University Brawijaya
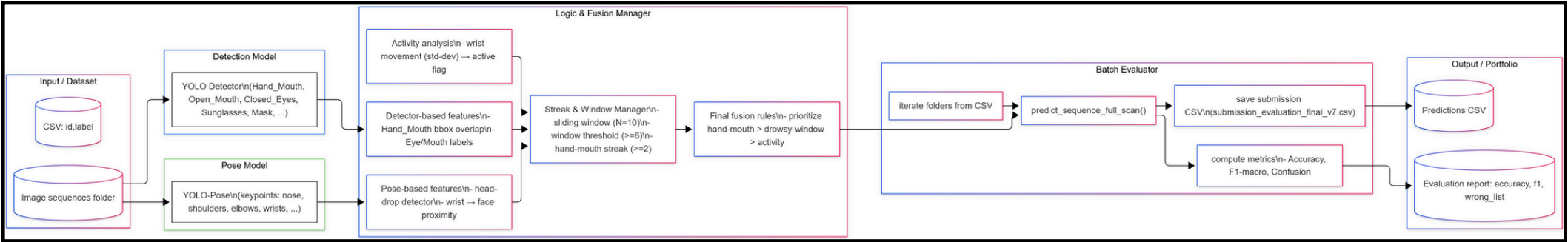
Contact information
arifathayaharahap@gmail.com

YOLO11: Advanced Drowsiness Detection Model

Year Accomplished
2025

Role/Position
AI Engineer & Data Engineer

Publication Link
https://huggingface.co/spaces/treamyracle/drowsiness-detector

I designed and implemented a computer-vision-based drowsiness detection pipeline that integrates a YOLO object detector with YOLO-Pose for keypoint extraction. The system includes dataset labeling, video sequence preprocessing, and a CSV-based evaluation format (id, label) for batch testing. I developed a multi-source fusion logic combining detector outputs, pose features, and activity analysis with a strict prioritization strategy. Hand-mouth confirmation uses both the Hand_Mouth bounding box and a minimum two-frame pose streak to significantly reduce false positives. Head-drop detection is derived from changes in nose–shoulder positioning, while activity analysis leverages wrist-movement standard deviation to filter out high-motion frames. A sliding window mechanism (N=10, threshold ≥ 6) ensures that only persistent drowsy events are classified (achieved 90% accuracy on a 700 dataset of 1 second video).

Name
Arif Athaya Harahap

University
Student of University Brawijaya

Contact information
arifathayaharahap@gmail.com

# Context-Aware NLP Grammatical Error Correction System

Year Accomplished
2025

Role/Position
NLP Backend Engineer

Publication Link
https://github.com/nlp-tugas-akhir/nlp-backend

A high-performance text correction API specifically engineered for the complexities of the Indonesian language (Bahasa Indonesia).

- The Architecture: Built a Hybrid NLP Pipeline that combines BERT-based Named Entity Recognition (NER) (to protect names/places from being "corrected") with SymSpell (for ultra-fast edit-distance spelling correction).
- The Linguistic Engine: Unlike generic spellcheckers, I implemented a custom Morphological Rule Engine from scratch to handle Indonesian phoneme melting rules (KPST) and complex suffix stacking (e.g., memp- > mem-, -kan/-i).
- System Reliability: Engineered a robust dictionary generation system that programmatically builds valid word variations from root words to reduce Out-Of-Vocabulary (OOV) errors.
- Full-Stack Integration: Wrapped in FastAPI with support for multi-format parsing (.pdf, .docx, .txt), making it a production-ready microservice.
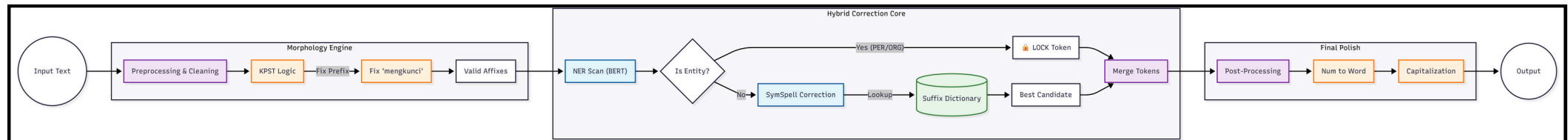
Name
Arif Athaya Harahap

University
Student of University Brawijaya

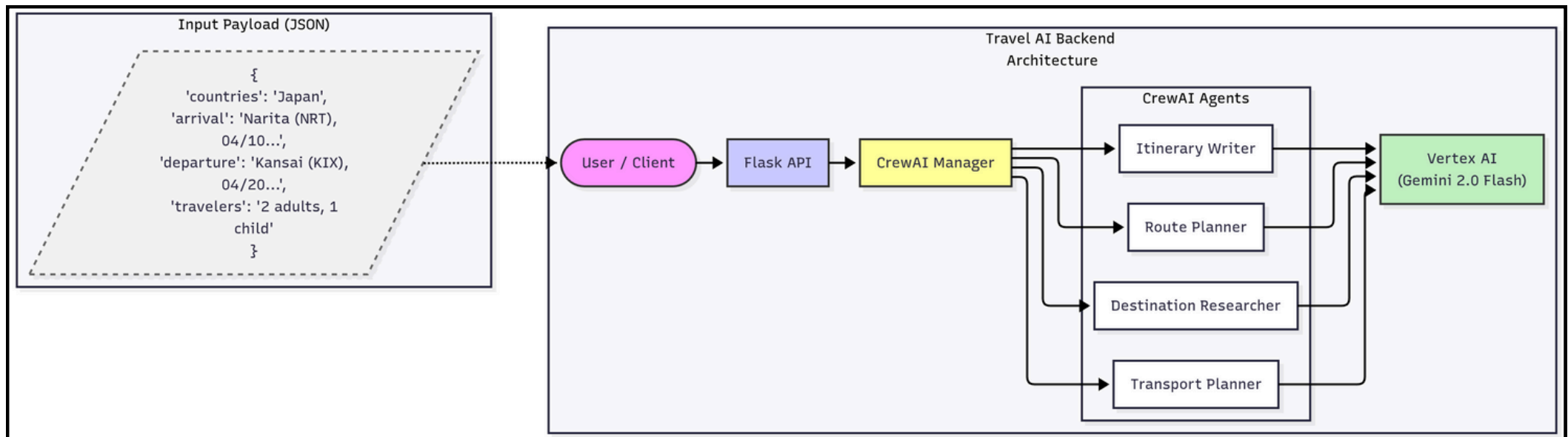**Contact information**
arifathayaharahap@gmail.com

Multi-Agent Gemini Itinerary Planner

Year Accomplished
2025

Role/Position
AI Engineer & Backend Developer

Publication Link
https://github.com/Treamyracle/GSC2025_SPS

A sophisticated backend API built with Flask and CrewAI that orchestrates an autonomous multi-agent system for travel planning. Leveraging Google Vertex AI (Gemini 2.0 Flash), the system coordinates specialized agents including Route Planners, Transport Researchers, and Itinerary Writers to autonomously generate comprehensive, structured travel itineraries from raw user inputs. The infrastructure is containerized and deployed on Google Cloud Run (no longer deployed i ran out of free credits) for serverless scalability.

Input Payload (JSON)

```
{
'countries': 'Japan',
'arrival': 'Narita (NRT),
04/10...',
'departure': 'Kansai (KIX),
04/20...',
'travelers': '2 adults, 1
child'
}
```

Travel AI Backend Architecture

User / Client → Flask API → CrewAI Manager

CrewAI Agents
- Itinerary Writer
- Route Planner
- Destination Researcher
- Transport Planner

→ Vertex AI (Gemini 2.0 Flash)

Name
Arif Athaya Harahap

University
Student of University Brawijaya

Contact information
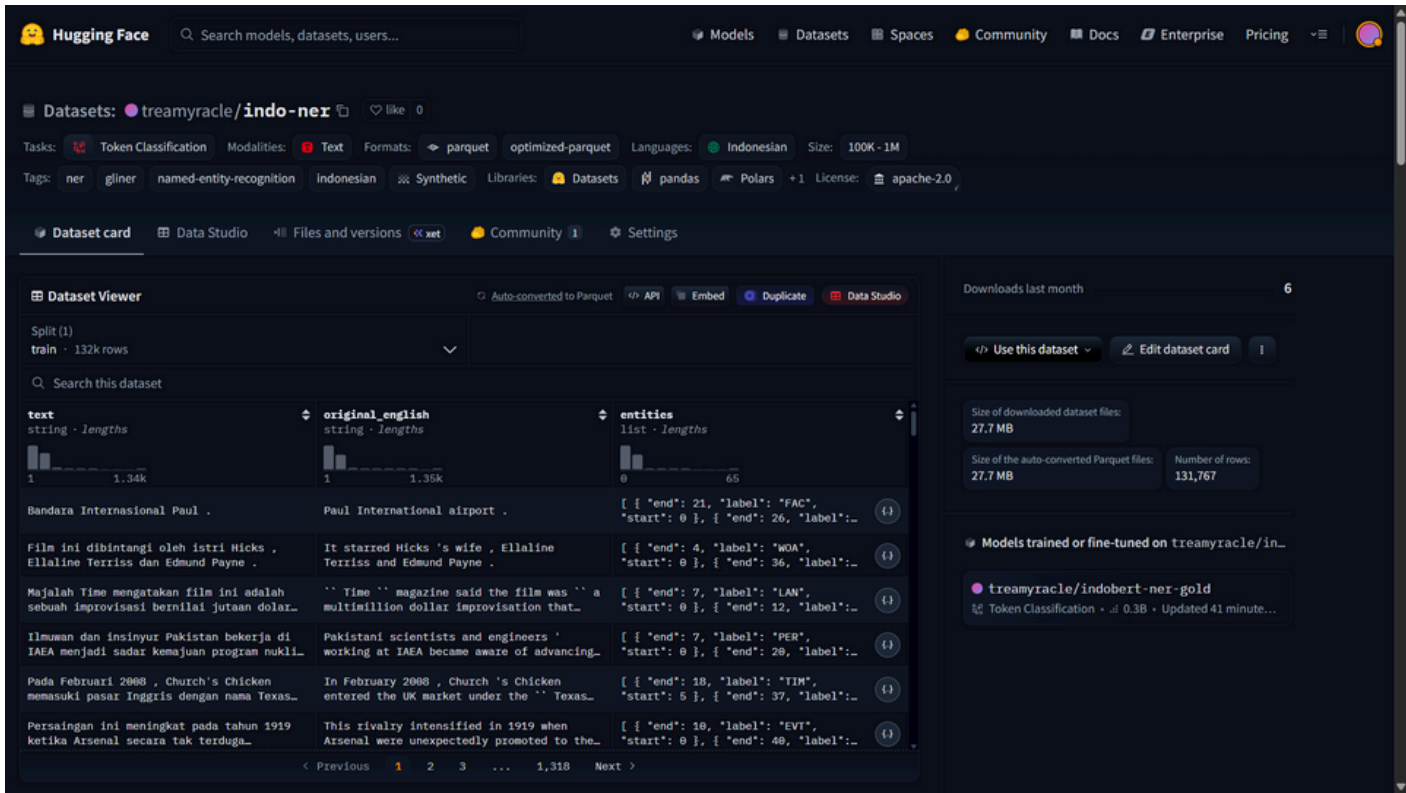arifathayaharahap@gmail.com

Indo-NER: Large-Scale Synthetic Indonesian NER Dataset

Year Accomplished
2025

Role/Position
AI Engineer & Backend Developer

Publication Link
https://huggingface.co/datasets/treamyracle/JudiOnline-Instruct-ID

Indo-NER is a massive synthetic dataset comprising over 130,000 sentences designed to address the scarcity of diverse Named Entity Recognition (NER) data in Bahasa Indonesia.

- The Methodology: Leveraged high-quality English NER datasets and applied a Neural Machine Translation pipeline, followed by an automated tagging process using GLiNER-Multi-v2.1.
- Rich Schema: Unlike standard datasets that only track 3 labels, this dataset captures 19 distinct entity types (including Events, Facilities, Laws, and Quantities), providing a richer semantic understanding for downstream models.
- The Impact: Served as the foundational "Silver Data" for pre-training the IndoBERT-NER (Gold) model, directly contributing to its State-of-the-Art performance by teaching the model general entity contexts before fine-tuning.

Name
Arif Athaya Harahap

University
Student of University Brawijaya
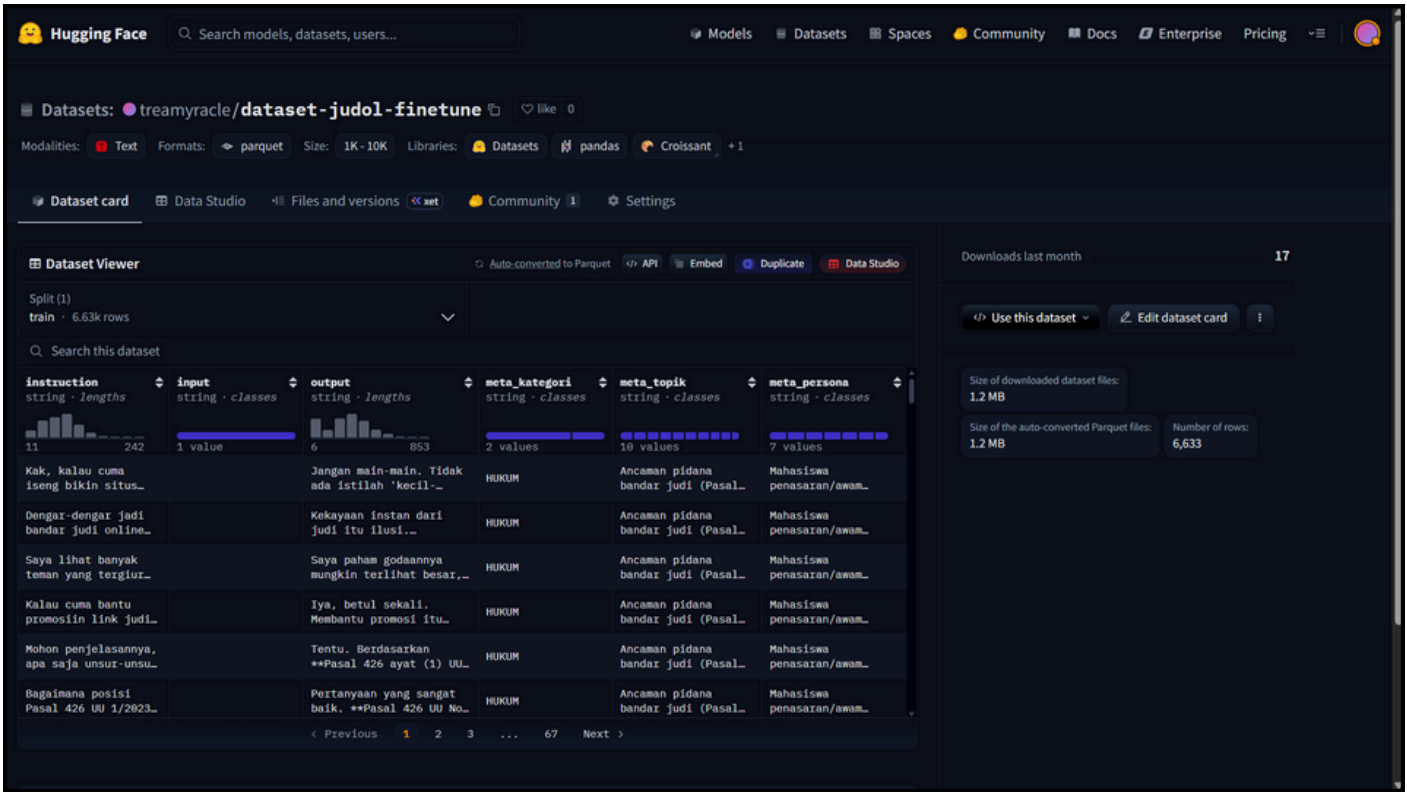
Contact information
arifathayaharahap@gmail.com

Instruct LLM Dataset: Online Gambling Awareness & Indonesian Law

Year Accomplished
2025

Role/Position
AI Researcher & Data Engineer

Publication Link
https://huggingface.co/datasets/treamyracle/JudiOnline-Instruct-ID

dataset-judol-finetune is a synthetic dataset generated using LLM API, leveraged with Google Search grounding tools to ensure high factual accuracy. The dataset focuses on the Indonesian online gambling landscape, covering general information, local slang definitions, and safety alignment (refusal mechanisms for gambling tips). When used to fine-tune the Qwen 2.5 7B model, this data achieved significantly higher domain-specific accuracy compared to the standard base instruct model (400% increase of accuracy measured by data test loss).

Name
Arif Athaya Harahap

University
Student of University Brawijaya

Contact information
arifathayaharahap@gmail.com

Instruct LLM Dataset: Indonesian Lawyer

Year Accomplished
2025

Role/Position
AI Researcher & Data Engineer

Publication Link
https://huggingface.co/datasets/treamyracle/Lawyer-Instruct-Translated-ID

Lawyer-Instruct-Translated-ID is a specialized dataset designed for instruct-tuning Large Language Models (LLMs) within the legal domain. I implemented a neural machine translation pipeline to curate high-quality training data, specifically engineered to fine-tune base models. The resulting dataset enables models to adopt a professional legal persona, allowing them to generate responses with the tone, structure, and terminology characteristic of a lawyer.
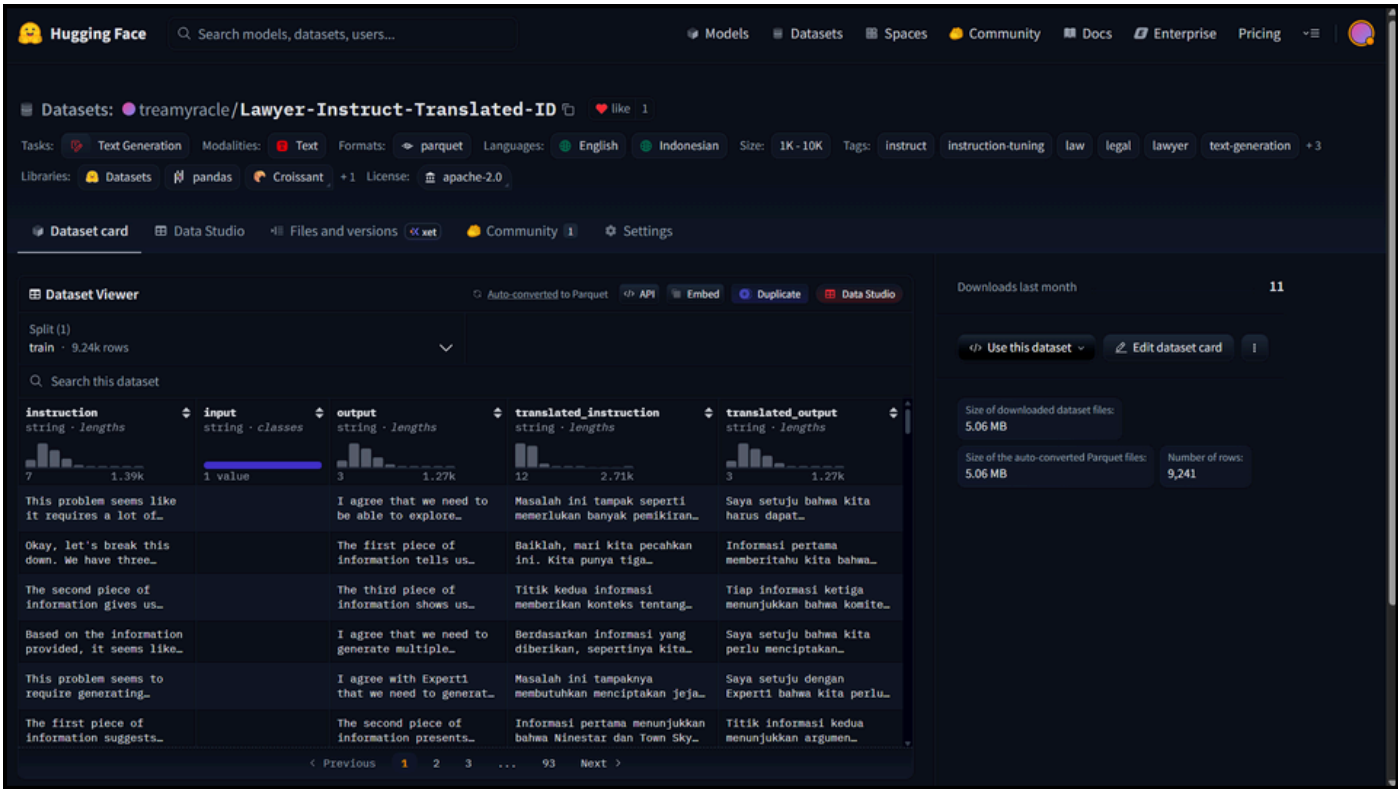
Name
Arif Athaya Harahap

University
Student of University Brawijaya
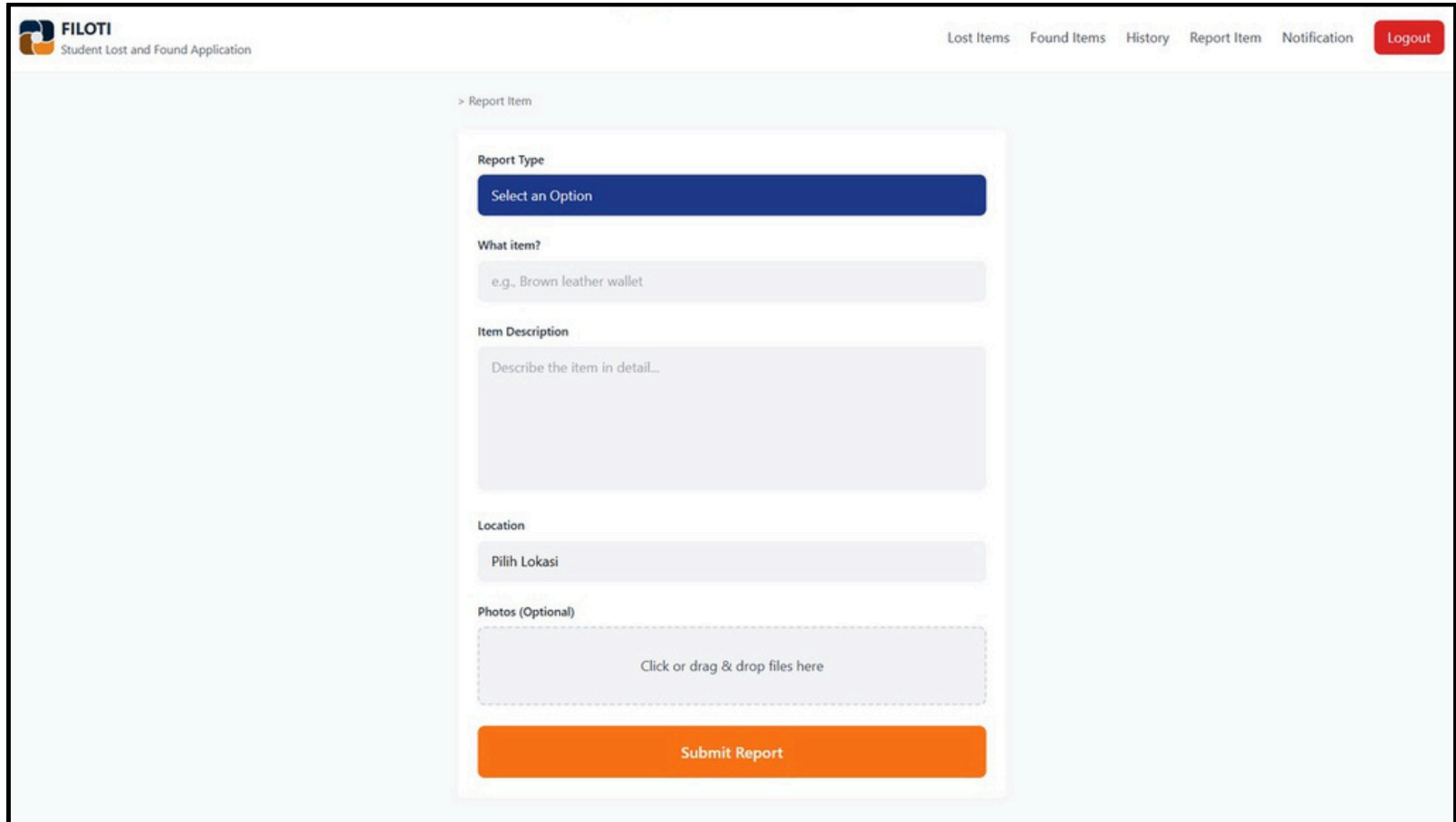
Contact information
arifathayaharahap@gmail.com

Website Application: FILKOMreserV

Year Accomplished
2024

Role/Position
Fullstack Developer

Publication Link
https://potong.in/FilkomR eserV

a centralized web application collaboratively built with a small team of 4 to solve the previously manual, fragmented, and error-prone room reservation system at the Faculty of Computer Science. The platform integrates the entire workflow, from requests to approvals, by implementing a real-time availability calendar and an automated admin system. Built with HTML & CSS, Golang, and Supabase, our team's solution successfully eliminates the risk of double bookings, reduces administrative workload, and streamlines the reservation process for hundreds of users.
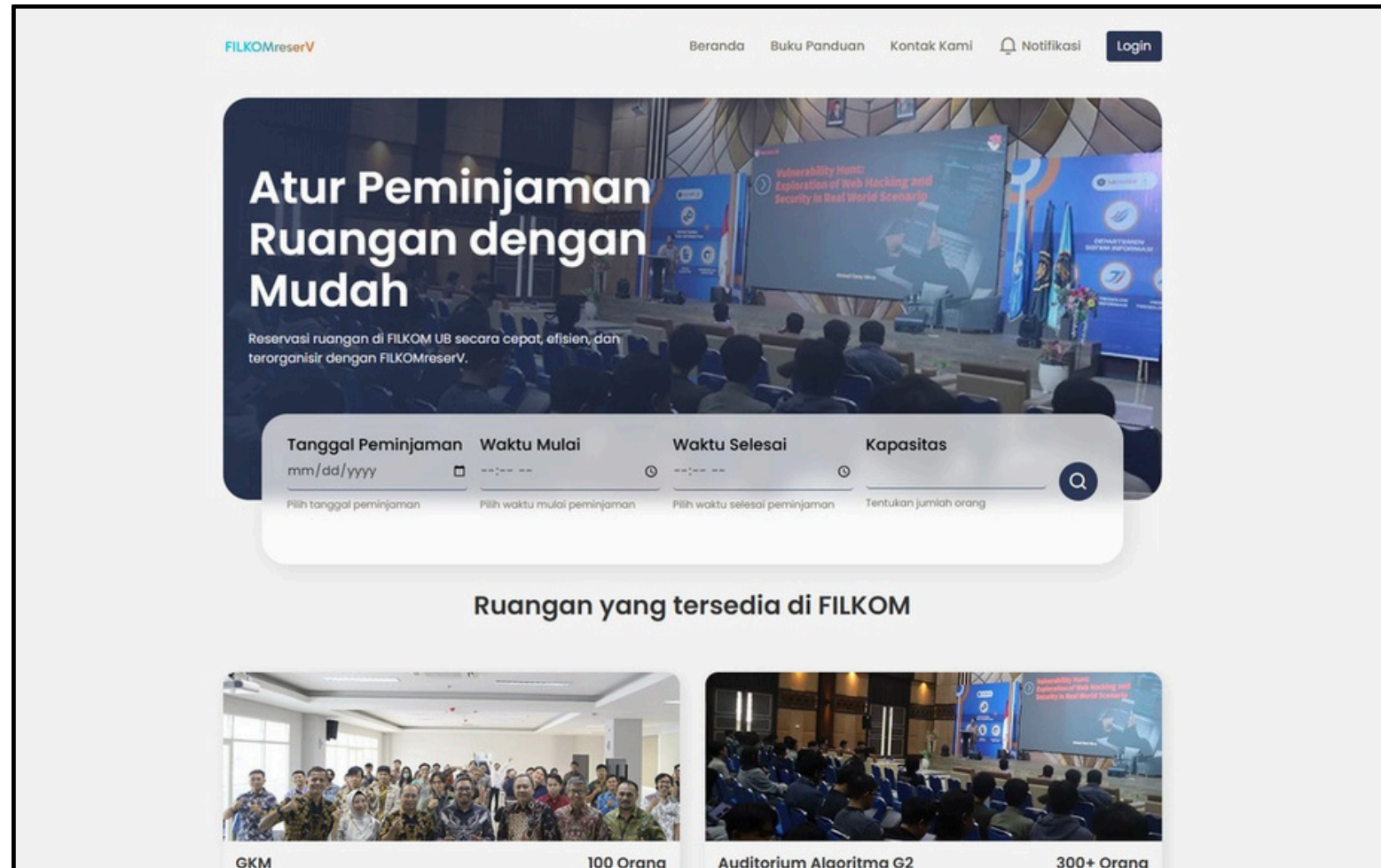
Name
Arif Athaya Harahap

University
Student of University Brawijaya

Contact information
arifathayaharahap@gmail.com

Website Application: FILOTI

Year Accomplished
2024

Role/Position
Fullstack Developer

Publication Link
https://potong.in/FILOTI

a centralized lost and found management system collaboratively built with a team to replace the faculty's manual, security-handled process and ineffective announcements via WhatsApp. The platform enables security staff to easily log found items with images, while allowing all faculty members to browse and search a structured database for their lost belongings. Built with React, Golang, and Supabase, our team's solution significantly improves item visibility and recovery rates, ensuring crucial announcements are no longer buried in irrelevant chat noise.

Name
Arif Athaya Harahap

University
Student of University Brawijaya
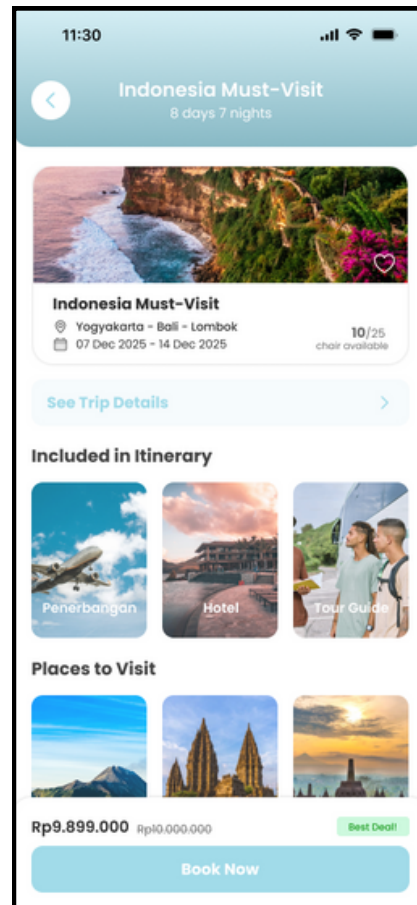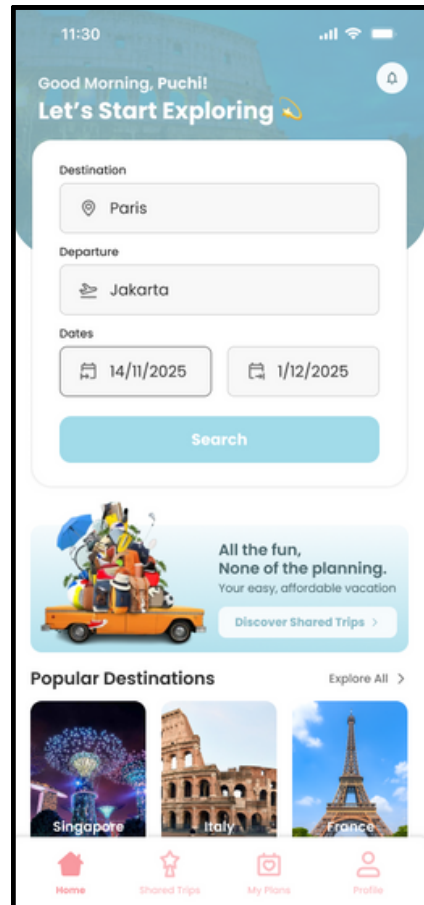
Contact information
arifathayaharahap@gmail.com

UI/UX Mobile Application: Viatrix

Year Accomplished
2025

Role/Position
Product Manager, UX Researcher, and UI Designer

Publication Link
https://potong.in/ViatrixPrototype

an innovative AI-based smart travel application collaboratively designed by team EL KAPITANO (team of 2 people) to replace the complex, time-consuming, and fragmented process of manual trip planning. The platform directly addresses the significant time burden an average of 17 hours and 42 minutes per trip and high stress that 96.8% of travelers experience during planning. The platform enables users to automatically generate a complete, personalized itinerary with integrated booking for flights and hotels, simply by entering their destination, dates, and number of passengers. Furthermore, it allows travelers to utilize a "Shared Travel" feature to find and join existing trips, making travel more affordable and social by splitting costs for guides and transportation. Designed with Figma, FigJam, and Google Forms , our team's solution significantly reduces planning time and cognitive load , ensuring a smarter, more inclusive, and cost-effective travel experience.

Name
Arif Athaya Harahap

University
Student of University Brawijaya

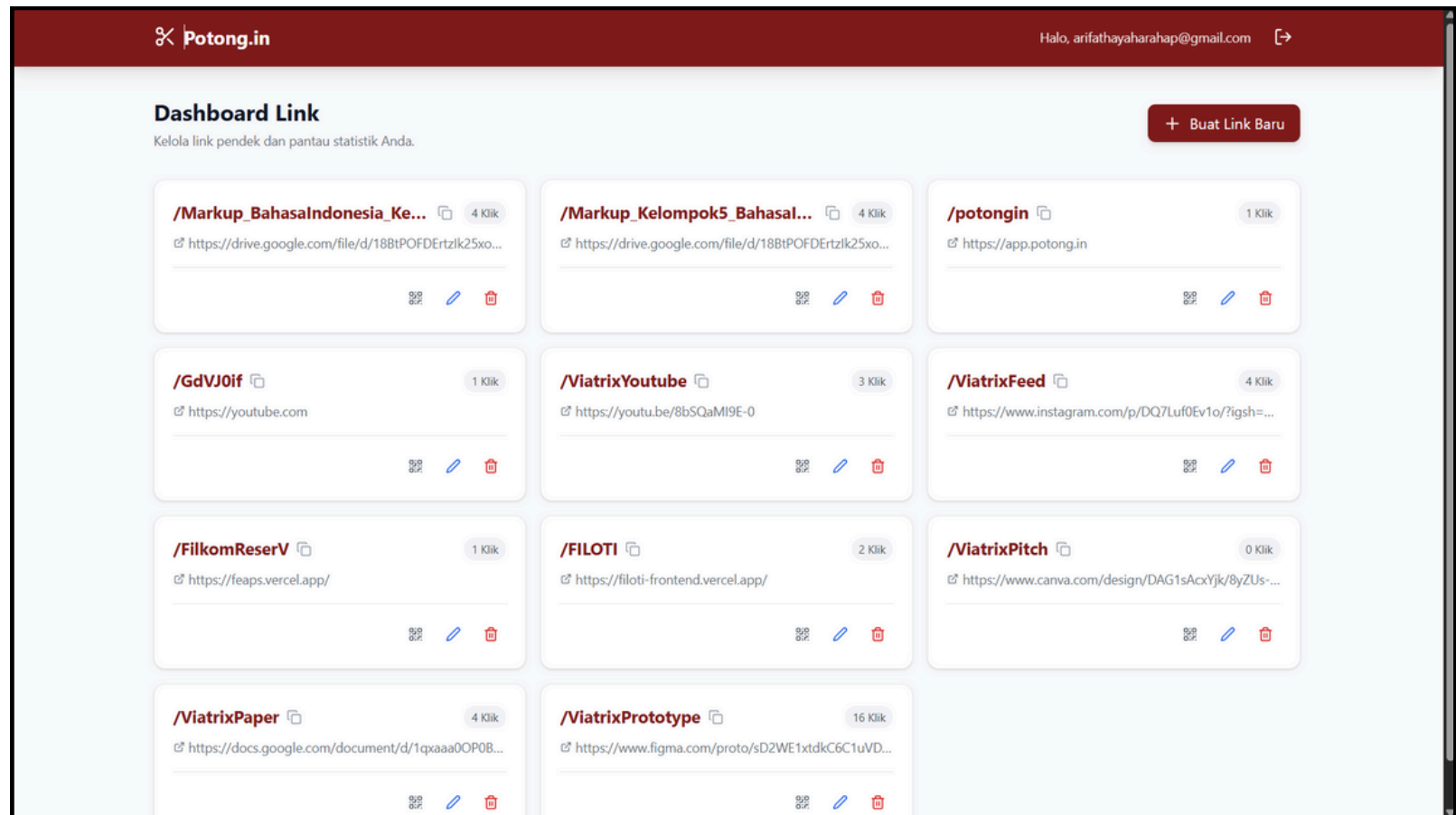**Contact information**
arifathayaharahap@gmail.com

Website Application: POTONGIN

Year Accomplished
2025

Role/Position
Product Manager & Fullstack Developer

Publication Link
https://app.potong.in/

potong.in is a full-stack URL shortener web application developed in collaboration with a partner. We engineered the backend API using Golang (Gin framework) to create a fast and reliable service, while the user interface was built with React.js to provide a seamless and interactive user experience. The project is live and deployed on Vercel.

Name
Arif Athaya Harahap

University
Student of University Brawijaya

Contact information
arifathayaharahap@gmail.com