

MOCH. ARIL INDRA PERMANA

Phone: +6282118454945 ◊ Email: arilindra21@gmail.com

Portfolio◊ LinkedIn ◊ Github

EDUCATION

Padjadjaran University

Bachelor of Physics

GPA: 3.68/4.00

Aug 2013 - Jan 2017

PROFESSIONAL EXPERIENCES

PT Pakar Digital Global (Paper.id)

Senior Data Scientist

Jun 2023 - Present

Sunter, North Jakarta (Full-Time)

- Serve as a Senior Data Scientist, spearheading the development and deployment of cutting-edge machine learning models and advanced analytics solutions.
- Key responsibilities include optimizing script execution, enhancing classification models, and streamlining operational processes.
- **Business Specialties:** Data Science, Machine Learning, Process Automation
- **Tech Stacks:** Python, Airflow, MySQL, Metabase, Looker, ArangoDB, VertexAI, GCP, DocumentAI, BigQuery, FastAPI, Flask

List of Projects:

1. Script Legacy Improvement & Python Wrapper Development:

- Drove significant improvements in script execution speed, debugging numerous issues.
- Reduced execution time for error seeding detection by approximately 98.33% and optimized the master segment table process by about 80%.
- Developed the "askquinta" library, now indexed on PyPi, streamlining team workflows and simplifying script development.

2. Modeling and Deployment Dashboard for Item Classification & User Scoring:

- Boosted classification model accuracy from 76% to 83%.
- Successfully deployed a user-friendly interface (UI) for model accessibility.
- Developed a scoring system for users based on business type and product transaction types.

3. KYC and KYB NWOW (New Way of Working) Dashboard and Analysis:

- Optimized data flow/pipeline, facilitating easier access for the Ops team.
- Introduced a data change tracker, improving upon the previous last data updated system.
- Created a comprehensive analysis dashboard with Looker for monitoring and optimizing KYC/KYB processes, enabling Ops to identify and address user pain points effectively.

4. Finance Operation & Bank Statement Analysis Dashboard:

- Built a Flask-based dashboard for automated data processing related to disbursement and Journal Entry, achieving a process time reduction of over 98%.
- Developed auto reconciliation and Source of Fund selection for seamless automated disbursement processes.

- Engineered a Flask dashboard for comparing user transactions against bank statements using Gemini, significantly reducing checking time by over 98%.

5. **OCR DocumentAI & Gemini:**

- Designed a schema for OCR modeling.
- Annotated data for receipts, invoices, and bank statements.
- Successfully deployed an OCR API and innovated prompt engineering for OCR using the Gemini Model, enhancing performance and reducing costs.

6. **Invoice Validation & KYC/KYB Process Automation:**

- Implemented detection of signatures, stamps, and duty stamps on invoices.
- Classified signatures as wet or digital to ensure transaction authenticity via invoice validation.
- Developed ID card clarity scoring to ensure quality uploads.
- Created face similarity scoring between selfies and ID cards.
- Validated companies by matching Nomor Induk Berusaha (NIB), deed of company (Akta), Nomor Pokok Wajib Pajak (NPWP), and decree of Kemenkumham (SK).
- Ensured accuracy in user-entered company information by matching company photos and business types.

PT Bukalapak.com

Assistant Manager Audit Data Analytics

Jan 2022 - Jun 2023

South Jakarta (Full-Time)

- Functioned as the Assistant Manager for Audit Data Analytics, leading initiatives to enhance data-driven auditing processes and fraud detection mechanisms.
- **Business Specialties:** Audit Data Analytics, Fraud Detection, Process Optimization
- **Tech Stacks:** Python, Airflow, MySQL, Redash, Looker, BigQuery, FastAPI, Flask, Archer

List of Projects:

1. **Audit Dashboard Automation:**

- Developed automated dashboards for monitoring and auditing using Redash and Looker.
- Increased team efficiency by eliminating repetitive audit tasks, allowing for more focused and strategic auditing.

2. **Fraud Analysis & Value Audit:**

- Conducted comprehensive fraud investigations, both data-driven and field audits, to identify internal and external fraud activities.
- Created dashboards to monitor transaction anomalies, successfully identifying and proving fraudulent activities by suspicious users.
- Implemented network analysis to uncover "syndicates" and "account farming" operations, significantly enhancing fraud detection capabilities.
- Provided crucial input on the KYC process by identifying potential bugs and fraud vulnerabilities within the application.
- Discovered discrepancies in financial reporting within subsidiary finance teams, leading to corrective actions.
- Identified revenue enhancement opportunities in the Payment Gateway system due to miscalculated fees, leading to substantial financial improvements.

3. Data Analysis & Chatbot Implementation:

- Conducted Quadrant Analysis to identify optimal advertising partners for the Marketing team, enhancing targeted advertising strategies.
 - Identified system bugs that potentially reduced revenue, such as miscalculated system-based computations, and recommended corrective measures.
 - Assisted team members in extracting data via ChatBot, streamlining data retrieval processes.
 - Developed automated responses for common audit data queries and automated data calculations, significantly reducing manual intervention and increasing productivity.
-

PT. Telkom Indonesia

Data Scientist

Jan 2020 - Dec 2021

South Jakarta (Full-Time)

- Worked as a Data Scientist, focusing on developing advanced analytics solutions and leveraging machine learning models.
- **Business Specialties:** Legal Analytics, Social Media Analysis, NLP
- **Tech Stacks:** Python, Elasticsearch, MySQL, FastAPI, Flask, Docker, Gephi, Airflow, Spark

List of Projects:

1. Legal Analytics Dashboard:

- Scraped data from 60+ websites to obtain legal documentation in PDF format.
- Developed OCR models to extract data from legal documents, including judicial reviews in the Constitutional Court (MK) and Supreme Court (MA).
- Indexed OCR results in Elasticsearch and created the corresponding database schema.
- Built APIs to facilitate data interaction between Elasticsearch and the frontend team.
- Developed a search engine and word suggestion feature.
- Implemented a plagiarism checker for document comparison.
- Modeled NLP for the harmonization of legal regulations.

2. Social Media Analysis:

- Crawled social media data based on specific topics.
 - Analyzed engagement, virality, sentiment, and performed topic clustering and network analysis to provide insights.
 - Achieved over 80% user satisfaction in decision-making processes based on analysis results.
-

PT. Indonesia Indicator

Data Scientist

Mar 2018 - Dec 2019

South Jakarta (Full-Time)

- Served as a Data Scientist, focusing on the development and implementation of data-driven solutions for political and brand analysis.
- **Business Specialties:** Political Campaign Analysis, Brand Analysis, Social Scoring
- **Tech Stacks:** Python, Elasticsearch, MySQL, MongoDB, FastAPI, Flask, Gephi, Java, JavaScript

List of Projects:

1. Political Campaign Profiling & Brand Analysis:

- Conducted extensive profiling for political campaigns on social media, providing valuable insights into campaign strategies and public sentiment.

- Performed comprehensive brand analysis to assess brand performance and market positioning.
- Built an auto-reporting engine for daily reports, significantly reducing manual reporting efforts and ensuring timely delivery of insights.

2. Social Scoring System & Social Media Analysis:

- Developed a social scoring system for assessing credit risk, leveraging social media data to enhance traditional credit scoring methods.
- Conducted social media and network analysis for political issues, identifying key influencers and network dynamics to inform strategic decisions.

3. Bot-Integrated Reporting Engine:

- Implemented a bot-integrated reporting engine to automate the generation of analytical reports, increasing efficiency and accuracy.

LICENSES & CERTIFICATIONS

Data Science for Business Professional (DSBIZ)

Extract Text Data with Python and Regex

Introduction to Deep Learning

Computer Vision: Python OCR & Object Detection Quick Starter

Docker for Absolute Beginners

Deep Learning in Computer Vision

HONORS & AWARDS

Speaker in Webinar provided by Edspert: "The Awesome Microsoft Excel: Membangun Karir Impian sebagai Data Scientist" (Apr 2023)

Top 30 Telkomsel Tech Titans League Series 3 - Machine Learning (Oct 2021)

Top 30 Telkomsel Tech Titans League Series 1 - Data Science (Sep 2021)

Speaker in Webinar LAWebinar 2.0 "Data Driven Law Culture & The Future of Legal Technology" (Jun 2021)

Top 100 Telkomsel Tech Titans League Series 1 - Data Analytics (Oct 2020)