Andrew Tirto Kusumo

Email: andrewtirtokusumo@gmail.com | Mobile: +62 813 1421 4039 |
linkedin: https://www.linkedin.com/in/andrew-tirto-kusumo/

A Data Plumber with a big enthusiast making my team works faster and better.
Also love analyzing data and wrangling with them.

Experience

Paper.id (paper.id)
Senior Data Engineer   **05/2024 – Present**
Tools: Airflow, DBT, Google BigQuery, ArangoDB, Datastream, Pub/Sub, Python (Pandas, FastAPI, google-cloud), Docker
√ Creating a streaming pipeline to ingest data from App DB to Google BigQuery from scratch using Google Datastream, Pub/Sub, and Dataflow
√ Fixed existing DBT ELT inefficiencies that can improve the development time by around 100%
√ Reducing cost in BigQuery by identifying the problem, helping to reduce cost around 20% per month
√ Responsible for guiding other Data Engineer in their respective projects
√ Created a cost management dashboard that tracks the cost of each projects each day

Flip.id (flip.id)
Data Engineer Manager   **12/2022 – 05/2024**
Senior Data Engineer   **09/2021 – 12/2022**
Tools: DBT, Google BigQuery, Apache Beam / Dataflow, Datastream, Pub/Sub, Python (Pandas, FastAPI, google-cloud), Docker, GitLab CI
√ Creating a streaming pipeline to ingest data from App DB to Google BigQuery from scratch using Google Datastream, Pub/Sub, and Dataflow
√ Create an end to end ELT pipeline with DBT from scratch, implementing test and query dependencies
√ Manage the cost in BigQuery by applying strict Partitioning and Clustering, helping to reduce cost around 20% per month
√ Managing Data Warehouse structure to scale with microservices infrastructure in Flip
√ Responsible for hiring a new Data Engineer team from zero
√ Creating job description, career framework, entry test, and interview questions for the Data Engineer team
√ Currently managing a team of 3 Data Engineers that reports directly to me
√ Make a pipeline that consume Facebook, Google, and Appsflyer API daily to automatically store it to DB
√ Make a pipeline to store Finance Journaling data to Netsuite
√ Creating a credit scoring POC for Flip's new lending product with Docker and FastAPI
√ Provisioning Redash and Looker Studio Analytics Dashboard for Analysts and end user
√ Responsible for our weekly releases

JULO (julo.co.id)
Senior Data Engineer   **1/2021 – 9/2021**
Data Engineer   **8/2018 – 1/2021**
Tools : Airflow, Ansible, AWS (All Data Tools), Alibaba, GCP (GCE, BigQuery, CloudSQL, GCS), Docker, Spark, Postgres, CircleCI
√ Making sure that our Data Pipeline (Airflow) is running alright 24/7 , fixing it if there's a problem
√ Responsible for making sure that the features my team generated is perfect for Data Scientist to use
√ Creating a replica to stream to the master DB for analytics purposes
√ Responsible for making daily shell scripting and bug fixing
√ Model deployment using Docker, H2O and feature implementation in Django
√ Make a pipeline that consume Facebook, Google, and Appsflyer API daily to automatically store it to DB
√ Designing and implementing a PG 10 partition using range for existing big tables
√ Designing action log data archiver that automatically partition a PG DB with almost billion rows of data

√ Integrating CircleCI in the repo for automatically testing and deploy
√ Responsible for weekly releases in our Data Platform
√ Maintaining and Reducing cost in AWS and GCP Platform

PHI-Integration (www.phi-integration.com)
Data Engineer and Analyst   **6/2017 – 5/2018**
√ Responsible for modeling in Azure Machine Learning using R to predict price of a goods
using linear model and SARIMA (Seasonal Autoregressive Integrated Moving Average)
√ Cleanse raw data to standardize data
√ Location data enrichment and visualization with Google Maps (Geocoding) for Marathon Event
√ Creating eBook for R basics and machine learning
√ Using Pentaho and R for data cleansing and mining
√ Crawling all news medias in Indonesia using R

Education

Multimedia Nusantara University, Jakarta, Indonesia
Bachelor of Computer Science | **GPA : 3.78 / 4.0**   **2014 - 2018**
Major: Computer Engineering, Concentration: Oracle Database Administrator
Thesis: Speech Recognition Analysis using Convolutional Neural Network for Indonesian Language
English Certification: TOEIC (950/990)
Japanese Certification: JLPT N3, N5

Honor & Awards

Academic Scholarship from University
√ Given to only top 5% (158 from 4500+) students per Year in Multimedia Nusantara University

Research & Projects

Flip Level Up Event - Speaker   **2023**
√ A project together with Algoritma and DANA.
I was invited to be one of the speakers in the event, speaking about Data Protection in Fintech.
(https://algorit.ma/data-science-fintech/)

DQLAB - Feature Engineering with Python   **2020**
√ A project together with DQlab where I create teaching material for them.
Feature Engineering with Python using Kaggle Titanic Dataset.
(https://dqlab.id/ & https://www.dqlab.id/belajar-data-engineer-dirumah-bersama-dqlab)

KUACI   **2019**
√ Create an Open Source KYC for KTP based on the data from Government
It translate people's KTP number into Location, Gender, and DOB
Contribute to Github Arctic Code Vault
(https://github.com/tenapril/kuaci)

Skills

Technical: Data Engineering, Feature Generation, Feature Monitoring, Model Monitoring, Model Deployment, CI/CD, Data Analyst, Instance Management, Predictive Machine Learning, Internet of Things, Linux OS, Arduino, NodeMCU, Android App, Web App, QA
Programming Languages: Python, SQL, PHP, HTML, Assembly, MATLAB, R, C, C++, C#, Java, ReactJS
Others: AWS, GCP, Alibaba, Azure Machine Learning, Github, Team Management, Wordpress, SEO
Language: English (full proficiency), Indonesian (native), Japanese (N3)