Bayuzen Ahmad

bayuzen19@gmail.com | 0852 5417 2133 | https://github.com/bayuzen19
https://www.linkedin.com/in/bayuzenahmad/

Summary

Data Scientist and ML/AI Engineer with 5+ years of experience delivering analytics, forecasting, computer vision, and RAG-based applications. Strong across the end-to-end lifecycle: data pipelines, model development, deployment, and stakeholder enablement. Hands-on with Python, SQL, PySpark, Azure (Cognitive Services, Azure OpenAI, AI Search), GCP, and BI tools (Power BI, Tableau, Looker Studio).

Experience

METRODATA   **NOV 2024 – PRESENT**

Data & AI

- Developed and implemented an end-to-end pipeline integrating OpenAI, OCR, and Cosmos DB to address specific customer business requirements, enabling automated document processing and data storage solutions.

- Delivered comprehensive training programs on Power BI and Microsoft Fabric, empowering customers to leverage advanced data analytics and business intelligence capabilities.

- Architected and implemented Azure Cognitive Services (Document Intelligence) integration for automated invoice text extraction, streamlining document processing workflows and reducing manual data entry.

- Successfully integrated LangChain and Azure OpenAI to develop custom chatbot solutions, enhancing customer service capabilities through advanced natural language processing.

- Designed and implemented Azure AI Search pipelines for RAG (Retrieval-Augmented Generation) chatbots, enabling context-aware responses and improved information retrieval from enterprise knowledge bases.

- Created bot translation documents using Azure AI services and App Services, and integrated the solution with Microsoft Teams.

ISMAYA GROUP (FREELANCE)   **MAR 2023 – NOV 2024**

Business Intelligence

- Developed an end-to-end web scraping solution using Selenium to extract data from Google Maps, including user information, ratings, and reviews. Designed a pipeline to store this data on Google Cloud Platform and created a customer review dashboard for two Ismaya concepts (restaurant and lifestyle) and one subsidiary (Legit Group) to monitor performance ratings and customer reviews.

- Developed and implemented an end-to-end OCR pipeline utilizing Microsoft's Phi-3 model to efficiently extract text from images and PDFs.

- Identifying transaction anomalies utilizing the Isolation Forest algorithm based on transaction occurrence patterns, total transaction frequency, received coin amounts, and variations in different name changes. Proficient in leveraging Isolation Forest for effective anomaly detection and ensuring data security and integrity.

- Creating customer universes using the RFM (Recency, Frequency, and Monetary) method to facilitate effective campaign processes and measure customer purchasing power based on transaction frequency, expenditure, and the last purchase date. Implemented RFM methodology resulted in a 2.4x increase in retention and a 1.2x improvement in Average Order Value (AOV).

- Develop a machine learning model for predicting which customers will make purchases next month using a classification model. Extract insights from the machine learning model to identify the significant factors that influence whether a customer will make a purchase in the upcoming month.
- Proficient in handling project product recommendation system, based on customer segmentation and time categories (morning, afternoon, evening, night, and midnight), to provide tailored recommendations that align with customer.
- Skilled in web scraping data from Grab Food platform using Beautiful Soup and Selenium for the purpose of pricing analytics.

FARMBYTE (FREELANCE)   AUG 2024 – DEC 2024
Machine Learning Engineer

- Perform web scraping on the DOSM (Department of Statistics Malaysia) website to collect agronomic price data.
- Developed a dashboard for credit scoring monitoring using Tableau. The project involved visualizing credit score trends, identifying risk patterns, and providing actionable insights for decision-making.
- Develop a forecasting model using multiple frameworks and perform model optimization to select the best model for predicting vegetable prices. Achieved a Root Mean Square Error (RMSE) between 0.1 and 1.2 with an average accuracy of 94.8%.

Altha Consulting (Project Based)   NOV 2024 – DEC 2024
AI Engineer

- Build a chatbot capable of interacting with PDF documents using the Vertex AI Gemini Pro model, integrated with the LangChain framework and a Retrieval-Augmented Generation (RAG) approach.

RUBY THALIB (FREELANCE)   NOV 2024 – FEB 2025
Computer Vision Consultant

- Evaluated WIMDa (AI weight-in-motion) classification outputs; identified extremely low confidence (~0.9%-1.45%) for "Uneven" predictions and class bias toward "Large" (~57%-58%); recommended threshold calibration and class-balanced retraining.
- Diagnosed likely drivers including camera angle, image quality/resolution, and detection-pipeline noise; proposed capture guidelines and targeted data collection to improve accuracy.
- Authored comprehensive documentation for a YOLOv8-based PPE compliance detection system for tunnel and jetty areas, covering camera configurations, observation zones, and PPE requirements by zone (helmet, safety vest, boots).
- Documented database schema (sicantik_apd_tunnel), SQLAlchemy connectivity, and best practices for configuring detection zones (polygon ROI; rectangles for PPE requirements).
- Produced deep-dive technical documentation for a Streamlit + LangChain document intelligence chatbot (RAG) using Vertex AI Gemini Pro and FAISS, including custom embeddings, file hashing, PDF text/table extraction, chunking, and Conversational Retrieval Chain setup.

Semudahitu.id (Project Based)   SEP 2024 – NOV 2024
AI Engineer

- Develop a chatbot with an SQL agent to perform automated data processing tasks, enabling seamless interaction with SQL databases for efficient query execution, data manipulation, and analysis.

INSIGNIA (FREELANCE)   APR 2024 – JUL 2024
Machine Learning Engineer

- Developed an end-to-end image generation pipeline using Stable Diffusion, ComfyUI, and OpenAPI to generate text for product descriptions. Deployed the model on Askdea.ai to assist SMEs in producing high-

quality images and compelling descriptions to enhance their sales.

ASTRA ISUZU MOTOR INDONESIA (PROJECT BASED)   NOV 2023 – APR 2024
Data Scientist

- Develop a time series model to predict monthly demands for over 21,000 materials with an accuracy exceeding 90%, utilizing an automated time series model selection process to identify the best models for each material.

PT SHARING VISION INDONESIA   MAR 2022 – NOV 2023
Data Scientist

- Developed a comprehensive dashboard for monitoring merchant performance, including the number of BRI and Non-BRI merchants across all regions in Indonesia, using Tableau and Looker Data Studio.Developed a dashboard for credit scoring monitoring using Tableau. The project involved visualizing credit score trends, identifying risk patterns, and providing actionable insights for decision-making.
- Developed an end-to-end automated machine learning pipeline using PySpark for classification tasks. The project encompassed data preprocessing, model training, and result monitoring, optimizing machine learning performance.
- Identify non-BRI customers who receive substantial money transfers from BRI (Bank Rakyat Indonesia) customers. The approach involves RFM (Recency, Frequency, Monetary) analysis and clustering to group customers based on their transaction patterns. The goal is to design an effective customer acquisition marketing campaign to attract potential customers to become BRI clients.
- Work with PT Bank Rakyat Indonesia (BRI) in the Digital Banking Development & Operations Division to support them in acquiring potential merchants by building analytics with vintage and cohort analysis, performing feature selection techniques and making model prediction.
- Use Beautiful Soup for data scraping and Selenium to automate web data scraping.
- Identify Bank Rakyat Indonesia savings customer patterns to determine whether they have a business or not by analyzing debit volume, credit volume, debit frequency, credit frequency, and analyzing customers based on geographical conditions.

SUPERAPP   JAN 2022 – FEB 2022
Business Intelligence Analyst

- Develop and manage business intelligence solutions.
- Provide reports through office applications to improve business processes.
- Analyze business requirements and processes to generate recommendations to assist business decision.
- Create and maintain documentation including designs, requirements, and user manual.

TSINGSHANG HOLDING GROUP   JAN 2020 – AUG 2021
Engineering and Analyst

- Supervise the final results of the AOD (Argon Oxygen Decarburization) process.
- Material checks, evaluates chemical content, and analyzes instances where chemical content exceeds internal standards.
- Utilize statistical methods, Minitab applications, and engineering analysis for this analysis..

Mentor Experience
AI INDONESIA   JAN 2024 - PRESENT
Machine Learning Mentor

- Create material and teaching machine learning Supervised and Unsupervised learning.

DIBIMBING.ID   **MAR 2023 - PRESENT**
Data Science Mentor

- Successfully led and nurtured a troop of +30 budding data scientists in mastering the fundamentals of
- Python, empowering them with the skills needed to harness the power of data.

PT SINAR TEKNOLOGI BANGSA (BANGALORE)   **AUG 2022 – JUN 2023**
Data Science Lecturer

- Handle a class with more than 10 students to teach them about Python for Data Science, such as list, dictionary, data wrangling, and numpy.

NARASIO DATA   AUG 2022 – NOV 2022
Data Science Mentor

- Handle more than 5 students in small group discussions to teach and strengthen their concepts in data preparation, statistics, data transformation, machine learning models, classification, and clustering

Skills & Abilities

- Technical: Mathematical Modeling, Classification, Regression, Clustering, Data Analysis and Visualization, CRISP-DM, Data Science, High Dimensionality Reduction, Feature Engineering, Feature Selection, Machine Learning Optimization.
- **Software & Tools**: Microsoft Excel, Tableau, R Studio, Minitab, Python, SQL, Power Bi, DAX, Google Data Studio, PySpark.
- **Interpersonal**: Team work, Leadership, Project Management, Problem Solving, Communication.

Education

SEPULUH NOPEMBER INSTITUTE OF TECHNOLOGY   **AUG 2015 - SEPT 2019**
Bachelor of Material Engineering, GPA: 3.58/4.00

Personal Projects

MARKETING CAMPAIGN OPTIMIZATION
Problem: Over the past three years, there has been a significant decrease in the conversion rate for customer acquisition. The company requires a machine learning solution for classifying customer potential to improve acquisition strategies.
Goals: Develop a machine learning model capable of predicting potential customers with a high chance of conversion.
Result: This project successfully increased customer acquisition by 4.3 times compared to the baseline model and reduced marketing campaign costs by 90%.

Portfolios Link: https://github.com/bayuzen19/marketing-Campaign

CUSTOMER SEGMENTATION
Problem: Develop customer segmentation to determine marketing strategies. The sample dataset summarizes the usage behavior of approximately 9000 active credit card holders over the past 6 months.
Goals: Create a credit card customer segmentation estimation model to assist the company in determining marketing strategies.
Result: Credit card customers are segmented into 3 clusters by using the K-Means algorithm with the Elbow and Silhouette Score as the evaluation methods.

Portfolios Link: https://github.com/bayuzen19/Clustering-Customer-Segmentation

LAPTOP PRICE PREDICTION
Problem: The company wants to understand laptop pricing across sales and purchase channels to help

distributors set market-appropriate prices.

Goals: Build a model that can predict laptop prices.

Result: The prediction has a 90.58% R-Square score by using the XGBoost algorithm.

Portfolios Link: https://github.com/bayuzen19/Laptop-Price-Prediction

DASHBOARD CHURN ANALYTICS

Problem: The problem is to design and develop a dashboard for monitoring churn.

Goals: Build a dashboard in Looker Studio.

Result: Link to dashboard.

Certificate & Training

CERTIFIED ASSOCIATE DATA SCIENTIST   NOV 2021

Badan Nasional Sertifikasi Profesi (BNSP) Indonesia

FULL STACK DATA SCIENCE   DEC 2022 – PRESENT

iNeuron

Related coursework: Python Basic to Advance, Django Framework, Machine Learning Supervised and Unsupervised, Deep Learning for Natural Language Processing and Computer Vision, Recommendation System with Deep Learning and Machine Learning, Tableau, Power BI, Marketing Mixing Model Analysis, Hive, Impala, Spark, Kafka.

FULL STACK DATA SCIENCE   FEB 2024 – MAR 2024

PT. Epam Digital Mandiri

Related coursework: Advance Machine Learning, Deep Learning, Streamlit, and Deployment.

ARTIFICIAL ENGINEERING : COMPUTER VISION   SEP 2023 – FEB 2024

AI Indonesia

Related coursework: Deep learning, Convulutional Neural Network, AlexNet, VGG, GoogleNet, Fater RCNN, YOLO, Object Detection,Self Driving Car.

DATA ENGINEERING   MAR 2023 – JUL 2023

Dibimbing.id

Related coursework: Object Oriented Programming, GIT, Advanced SQL and Optimization, Airflow, Kafka and Data Govermence.

BUSINESS INTELLIGENCE SPECIALIST   MAY 2022 – AUG 2022

Dibimbing.id

Related coursework: Business Intelligence Scenario, Exploratory Model Analysis & Model Explainability, customer segmentation using RFM and K-Means, Time Series Analytics and Forecasting, A/B Testing Uplift Modeling, Customer Lifetime Value / Marketing Channel Optimization, Casual Inference Statistics.

FULL STACK DATA SCIENCE BOOT CAMP   OCT 2021 – APR 2022

FullStack Data Bangalore Indonesia

Related coursework: Python, Data Wrangling, Data Preprocessing, Fundamental Statistics, Data Visualization, Machine Learning, Natural Languange Processing, Image Processing.

BUSINESS INTELLIGENCE ANALYST   DEC 2021 – FEB 2022

Binar Academy

Related coursework: Business Intelligence framework, Business Requirement Documents, Google Data Studio.

FULL STACK DATA SCIENCE BOOTCAMP   AUG 2021 – DEC 2021

Dibimbing.id

Related coursework: Python and R for Data Science, Statistics for Data Science, Data Cleaning, Data Preprocessing, Data Visualization, Supervised and Unsupervised Machine Learning, Google Data Studio.