Gevan Putra Yudhistira

putra.gevan00@gmail.com +62857-6306-4263 Kota Malang https://mochaheree.github.io/

linkedin.com/in/gevan-py

SUMMARY

Data Analyst with Computer Engineering background, currently building data pipelines and dashboards for customer behavior analysis. Experience transforming raw chat data into structured insights using Python, SQL, and Looker Studio. Focused on translating data into actionable business decisions, particularly in reservation and conversion tracking systems.

PROJECTS

WhatsApp Reservation Analytics Pipeline for Customer Behavior and Conversion Tracking 04/2026 – Present
XOAN Social Hub, Malang, East Java
- Raw WhatsApp chat data was **transformed into a structured reservation funnel** using a custom data pipeline: JSON Chat Data → Cleaning & Parsing → Interaction Modeling → PostgreSQL Database → Looker Studio Dashboard & internal web app dashboard.
- Reservation demand existed, but businesses lacked visibility into how customer conversations translated into bookings, causing inconsistent reservation performance.

PORTOFOLIO

Cohort Retention & Churn Analytics for SaaS Subscription, *Personal Project*
- Applies cohort retention analysis to a SaaS subscription dataset to measure customer retention trends and identify churn risk factors.
- Insights are used to support data-driven strategies for improving subscription growth and reducing churn.

Customer Segmentation & Retention Strategy for E-commerce, *Personal Project*
- Analyzes an e-commerce customer dataset to understand customer behavior, retention patterns, and revenue distribution.
- The objective is to generate actionable business insights and propose strategic recommendations to improve customer retention and sustainable revenue growth.

EDUCATION & COURSES

Machine Learning and Data Science, *Online course by Zero to Mastery* 03/2026
- Applied data analysis and visualization techniques on a heart disease dataset using Pandas, Matplotlib, and Seaborn to extract insights. One key finding showed that stress level provided the clearest categorical signal despite exhibiting a non-linear pattern, with medium stress having the highest prevalence (~21.4%), followed by high (~19.8%) and low (~18.7%).
- Utilized a car sales dataset and mapped it to the Scikit-Learn machine learning workflow, identifying its suitability for regression tasks, particularly using Ridge Regression, Support Vector Regressor (SVR), and Random Forest Regressor. Conducted model comparison and evaluated Ridge Regression performance using regression metrics such as R^2, MAE, and MSE.
- Developed a deep learning model for multi-class image classification (dog breed classification) using TensorFlow and Keras, leveraging transfer learning with pretrained CNN architectures. Optimized model performance through data augmentation and fine-tuning, and performed comprehensive evaluation using accuracy and loss to ensure generalization on unseen data.
- Performed data cleaning, preprocessing, and feature engineering to support robust and effective model development.

Data Science Foundation, *Online course by IYKRA* 07/2025
- Applied fundamental data science concepts (data analysis, statistics, and data-driven decision making) through case studies on household furniture sales, customer profiling, and office supply distribution
- Applied data visualization techniques to communicate insights effectively
- Utilized Python, SQL, and UNIX for data processing and analysis across business case scenarios involving sales and customer data
- Gained understanding of data governance, ethics, and regulations
- Strengthened non-technical skills such as problem-solving and analytical thinking

Universitas Brawijaya, *Bachelor in Computer Engineering* 2018 – 2025 | Malang, East Java
- Final project: "Hand Gesture Recognition System Based on ESP32 and Machine Learning for Power Point Slide Control"
- Developed an end-to-end machine learning system that transforms real-time sensor data into actionable decisions. The project involved data collection using MPU6050 sensors, followed by data cleaning, preprocessing, and feature engineering (18 extracted features) to build a Random Forest classification model. The model achieved 86.7% accuracy and was deployed on an embedded system (ESP32), enabling efficient real-time prediction under hardware constraints and translating model outputs into automated actions.

SKILLS AND TOOLS

Skills
Python, Scikit-learn, SQL, Data cleaning & preprocessing, Data Analytic, Data Visualization, Basic ETL Concepts

Tools
Google colab, Google sheet, Looker Studio, PostgreSQL