Muhammad Septian Nugroho

Salatiga • +62 82325279005 • muhammadsian00@gmail.com • LinkedIn • Portfolio

Motivated and detail-oriented IT graduate with a strong foundation in data science. Proficient in Python, Machine Learning, and AI. Eager to apply technical expertise and problem-solving skills to drive innovative solutions in a dynamic IT environment.

EDUCATION

Bachelor of Computer in Informatic Engineering ·
Satya Wacana Christian University
GPA: 3.77/4.00

- Relevant Coursework: Artificial Intelligence, Machine Learning, Database Systems, Statistic and Probability
- Thesis/Capstone Project: Web-Based Information System for Tingkir Lor-Salatiga Tourism Village (Laravel)

WORK EXPERIENCE

PT Zona Akselerasi Pendidikan
Data Analyst

- Performed feature engineering—handled missing values, encoded categorical variables, and scaled numerical features to improve model inputs.
- Built and optimized classification models (Logistic Regression, Random Forest, XGBoost), achieving strong performance as assessed by AUC and accuracy metrics.
- Led and collaborated with team to achieve a model accuracy of 85%

PROJECTS

Home Credit Risk (Company Hands-On)

- Built loan default risk prediction models with **ROC AUC ≈ 0.91**, strengthening credit risk assessment.
- Designed ML pipelines with preprocessing, feature engineering, and ensemble modeling for robust classification.
- Delivered insights through feature importance analysis and dashboards, improving transparency in financial decision-making.

Project Portfolio: E-commerce Recommendation Engine (Amazon Data Case Study)

- **Engineered an e-commerce recommendation system** on Amazon review dataset by implementing **collaborative filtering (user–item matrix factorization, cosine similarity)** and **content-based filtering (TF-IDF vectorization on product metadata)**.
- **Preprocessed and analyzed >100k+ product–review interactions**, including data cleaning, handling sparsity in user–item matrix, and applying dimensionality reduction to optimize computation.
- **Developed end-to-end pipeline in Python** using **Pandas, Scikit-learn, and Surprise library**, evaluated models with **RMSE and Precision@K metrics**, and compared performance of baseline vs hybrid recommendation approaches.

Project Portfolio: Ecommerce Sentiment Analysis & Chatbot

- Developed an end-to-end sentiment analysis solution for ecommerce product reviews using Natural Language Processing (NLP).
- Achieved an **81% precision** and **77% recall** by training a **Naive Bayes classifier**, outperforming more complex models like SVM and Random Forest.
- Successfully deployed the model into an interactive web application using **Streamlit**, demonstrating the ability to turn a machine learning model into a user-facing product.

Project Portfolio: E-commerce Purchase Prediction

- Designed and implemented a **machine learning pipeline** for purchase prediction, achieving **98%+ accuracy, precision, recall, F1, and ROC-AUC**.
- Engineered a **synthetic dataset of 100K+ records** to simulate real-world e-commerce user behavior.
- Built a **stacking ensemble (XGBoost, LightGBM, CatBoost, RandomForest)** optimized with Optuna, delivering **production-ready performance** and key business insight

Project Portfolio: Large-Scale Balanced Churn Predictor
- Built a **large-scale churn prediction system** achieving **97%+ accuracy, precision, recall, and F1**.
- Implemented **balancing strategies (SMOTE, undersampling, oversampling)** to handle severe class imbalance.
- Optimized a **stacking ensemble** of XGBoost, LightGBM, and CatBoost with Optuna for maximum performance.

SKILLS
Programming Languages: HTML/CSS, Javascript, PHP, Python, SQL
Databases: MySQL
Tools & Platforms: Adobe XD, ETL, Excel, Figma, Power BI, SPSS, VBA, Looker Studio
Frameworks & Libraries: CatBoost, Ensemble methods, Gradient Boosting, Keras / TensorFlow, LightGBM, Logistic Regression, Matplotlib, NumPy, Pandas, Random Forest (via Scikit-learn), Scikit-learn, Seaborn, XGBoost, Streamlit.
Other: Data Visualizations & Storytelling, ETL Pipelines, Feature Engineering, Predictive Modeling & Machine Learning (Logistic Regression, Random Forest, XGBoost), Model Deployment (Streamlit), NLP.

CERTIFICATIONS
Data Classification and Summarization Using IBM Granite
IBM
Artificial Intelligence Fundamentals
IBM
Database Programming wih PL SQL
Oracle Academy
Database Design and Programming with SQL
Oracle Academy