

# DAFFA AHMAD PANGREKSA

6289687008503 | daffaahmadpangreksa@gmail.com | linkedin.com/in/daffaahmadpangreksa | https://dapp-eng.github.io

Driyorejo, Gresik, Jawa Timur

Mahasiswa aktif S1 Sains Data di Universitas Negeri Surabaya. Memiliki pengalaman praktis dalam machine learning, natural language processing, dan analisis data melalui proyek akademik, kompetisi, dan riset aktif di laboratorium big data. Terbiasa bekerja dengan Python dan berbagai framework data science dari tahap preprocessing hingga deployment model.

## Pendidikan

**Universitas Negeri Surabaya - Surabaya, Indonesia**

Jun 2024 - Jul 2028 (Expected)

*S1 Sains Data, 3.92/4.00*

- Mata Kuliah Relevan: Machine Learning/ AI, Data Mining, Data Wrangling, Exploratory Data Analysis, Pemrograman (Python), Struktur Data & Algoritma, Basis Data, Statistika.

## Pengalaman Profesional

**Data Science and AI Researcher**

Mar 2026 - Present

*Laboratorium Big Data Analysis, FMIPA, UNESA*

- Melakukan riset aktif di bidang Graph Neural Network (GNN) untuk pemodelan data relasional berskala besar menggunakan Neo4j dan NetworkX.
- Mengembangkan pipeline eksperimen berbasis Jupyter Notebook sebagai bagian dari luaran penelitian ilmiah yang sedang berlangsung.

**Asisten Dosen Struktur Data dan Algoritma**

Apr 2026 - Present

*Prodi Sains Data, FMIPA UNESA*

- Membimbing 36 mahasiswa dalam konsep algoritma searching, sorting, rekursi, dan analisis kompleksitas waktu serta ruang.
- Memandu sesi praktikum dan mengevaluasi tugas pemrograman untuk memastikan pemahaman mahasiswa terhadap implementasi struktur data.

**Asisten Dosen Pemrograman Dasar**

Sep 2025 - Dec 2025

*Prodi Sains Data, FMIPA UNESA*

- Membimbing 72 mahasiswa dari 2 kelas dalam penguasaan logika pemrograman Python.
- Memimpin sesi praktikum mingguan dan memberikan umpan balik terstruktur untuk meningkatkan efisiensi teknis kode mahasiswa.

**Staf Divisi Acara - ACTION Competition 2025 (Organisasi)**

Jun 2025 - Dec 2025

*HMP Sains Data (HIMASADA) Universitas Negeri Surabaya*

- Merancang dan mengelola alur kompetisi infografis tingkat nasional (150+ peserta) secara terstruktur dan lancar.
- Menjadi moderator utama, memfasilitasi komunikasi antara 3 juri, peserta, dan panitia.
- Mengelola dan merekap data akhir lomba secara akurat untuk memastikan transparansi penilaian.

**RANK 6 - Data Mining Competition (DISCO)**

May 2025 - Jun 2025

*Universitas Trunojoyo Madura*

- Membangun model prediksi laba UMKM menggunakan XGBRegressor dengan optimasi Optuna, menghasilkan  $R^2$  sebesar 0,91.
- Melakukan segmentasi bisnis berbasis persentil menjadi 3 tier profitabilitas dari 13.564 data UMKM lintas sektor industri.

*HMP Sains Data (HIMASADA) Universitas Negeri Surabaya*

- Mengelola kebutuhan logistik acara besar (300+ peserta) dengan kesiapan alat 100% tanpa kendala.
- Berkoordinasi dengan 5+ divisi dan vendor eksternal, mencapai efisiensi biaya  $\pm 15\%$  sesuai anggaran.
- Mengawasi distribusi, pemeliharaan, dan pengembalian perlengkapan dengan sistem inventaris yang tertata.

## **Pengalaman Proyek**

---

### **Analisis Sentimen Komparatif Program MBG via LDA dan Fine-Tuned IndoBERT**

Sep 2025 - Dec 2025

*Universitas Negeri Surabaya*

- Memproses 4.968 tweet dan 3.418 komentar YouTube melalui scraping dual- platform serta pipeline preprocessing berbasis regex, stopword removal, dan tokenisasi.
- Menerapkan LDA topic modeling dengan optimasi coherence score menghasilkan 10 topik optimal di X (skor 0,498) dan 3 topik di YouTube (skor 0,407).
- Mengklasifikasikan sentimen menggunakan fine-tuned IndoBERT (confidence rata-rata 0,89), mengungkap sentimen negatif lebih dominan di YouTube (49,1%) dibanding X (46,9%).

### **Prediksi Risiko Diabetes pada Dataset NHANES 2017-2020**

Sep 2025 - Dec 2025

*Universitas Negeri Surabaya*

- Mengintegrasikan 10 file XPT dari CDC API mencakup demografi, BMI, dan riwayat klinis, serta merekayasa fitur komposit seperti LAP Index, VAI Score, dan Metabolic Risk Score.
- Membandingkan XGBoost, Random Forest, dan MLP; Random Forest mencapai hasil terbaik pasca-koreksi leakage dengan  $F1 = 0,526$  dan  $AUC = 0,895$ .

### **Benzene Estimator: Analisis Kualitas Udara Multi-Sensor**

Dec 2025

*Universitas Negeri Surabaya*

- Membangun aplikasi web berbasis Streamlit untuk estimasi konsentrasi Benzene ( $C_6H_6$ ) secara real-time menggunakan data sensor array dari UCI ML Repository.
- Mengembangkan model OLS dengan statsmodels dan scikit-learn, dilengkapi visualisasi gauge chart Plotly interaktif serta modul evaluasi metrik  $R^2$ , MAE, dan RMSE.
- Menerapkan pipeline preprocessing otomatis berbasis IQR dan EDA interaktif dengan heatmap korelasi serta scatter plot 3D multivariat.

### **Klasifikasi Genre Musik dan Sistem Rekomendasi berbasis Digital Signal Processing (DSP) dan Machine Learning (ML)**

Oct 2025 - Nov 2025

*Universitas Negeri Surabaya*

- Merancang pipeline DSP menggunakan filter Butterworth orde 5 dan mengekstraksi 58 fitur akustik (ZCR, MFCC, Chroma STFT) dari 999 track audio dataset GTZAN.
- Melatih 4 model supervised learning; MLP meraih akurasi tertinggi 76,50% ( $F1 = 0,767$ ), dengan MFCC terkonfirmasi sebagai fitur paling diskriminatif (separation ratio 1,55x).
- Membangun mesin rekomendasi berbasis cosine similarity di ruang fitur 58 dimensi untuk identifikasi track akustik serupa lintas genre.

### **Indeks Keterjangkauan Pangan Regional Jawa Timur**

Nov 2025 - Dec 2025

*Universitas Negeri Surabaya*

- Mengintegrasikan 3 sumber data heterogen yaitu dokumen UMK, data harga Bank Indonesia, dan scraping Tokopedia untuk 11 komoditas di 9 kota Jawa Timur.
- Membangun Market Affordability Index (MAI) dan E- Commerce Access Index (EAI) ternormalisasi MinMaxScaler, mengungkap gap daya beli 2x antara Surabaya (MAI >120) dan Sumenep (MAI ~60).
- Mempublikasikan dataset final secara publik di Kaggle sebagai 'East Java Affordability Index'.

## Sertifikasi dan Keahlian

---

- **Sertifikasi Dicoding:** Data Science; Dasar Artificial Intelligence; Introduction to Financial Literacy
- **Sertifikasi Digital Talent Scholarship, Kemenkominfo:** Analisis Data
- **Bahasa Pemrograman:** Python (Pandas, NumPy, Scikit-learn, TensorFlow, Statsmodels, Librosa)
- **Machine Learning & AI :** Supervised Learning, NLP, Graph Neural Network, Deep Learning, Feature Engineering
- **Data & Visualisasi:** SQL, Streamlit, Plotly, Matplotlib, Seaborn
- **Tools & Platform:** Jupyter Notebook, Git, Neo4j, NetworkX, Figma, Tableau, PowerBI, Microsoft Office
- **Bahasa:** Bahasa: Indonesia (Native), Inggris (Menengah)
- **Soft Skills:** Komunikasi, Kerja Tim, Kepemimpinan, Manajemen Waktu, Detail-oriented