Eshaal Malik

Applied Machine Learning & Data Science | Tabular ML, rigorous validation, decision support

Karachi, Pakistan | +92 309 9878009 | eshaalmzaa@gmail.com

GitHub | LinkedIn | Portfolio

Summary

Applied machine learning and data science student focused on tabular ML, honest validation, and decision-support tools. At FlyRank AI, I built a leakage-aware workflow on 30,000 content pages and showed that grouped validation scores came out lower than a random split, avoiding an optimistic estimate. I also founded StrateUpward, a data science service and decision-intelligence prototype for businesses.

Experience

Machine Learning Intern, FlyRank AI | Remote | July 2026 – August 2026

- Analyzed 30,000+ historical content-page observations from 32 anonymized clients across 44 source columns to study search visibility, freshness, engagement, and discoverability signals.
- Turned a ranking problem into a reproducible ML workflow: leakage-aware feature construction (25 leakage-related columns excluded), benchmarking of 4 regression models on 26,612 modeling observations, and MAE / RMSE / R^2 evaluation.
- Validated with client-grouped held-out splits and found lower performance than a random split, documenting how optimistic estimates appear when related observations cross train/test boundaries.
- Built a transparent opportunity-scoring layer (demand, staleness, competition, lifecycle) to rank 27,532 eligible records for human editorial review, without claiming causal or guaranteed ranking effects.

Project write-up | GitHub

Additional internships (all remote, Python-based)

- **Machine Learning Intern, FutureXcel** (Feb 2026): Prepared datasets with outlier handling, normalization, and feature selection; built and evaluated ML models and documented findings for reproducible review.
- **Data Science Intern, SyntecxHub** (Jan 2026): Cleaned and validated 10,000+ records for two AI/data-science projects; built Power BI visuals from structured datasets.
- **Data Science Intern, Code Alpha** (Dec 2025): Built and evaluated three ML models across classification and regression tasks and compared their performance.

- **Data Science Intern, Arch Technologies** (Nov 2025 – Jan 2026): Completed 3+ classification tasks and built an end-to-end data pipeline to prepare structured data for modeling.

Projects

StrateUpward: Business Data Intelligence | Founder | Python, Pandas, Plotly, ML | Prototype

- Prototype platform that turns CSV and Excel datasets into evidence-based insights and decision-support reports for businesses.
- Workflows for data quality assessment, cleaning, EDA, statistical analysis, visualization, anomaly detection, and segmentation.
- Adding predictive modeling and AI-assisted interpretation, with transparent methods and data-grounded findings.

AQI-PAK Strategic Command: Air-Quality Forecasting | Time-series ML

- Built a forecasting system for major Pakistani cities using ensemble models (XGBoost, LightGBM, Random Forest), with reproducible preprocessing, feature engineering, and validation on 123,000+ hourly air-quality records, delivered as a Streamlit dashboard. **BW-to-RGB: Image Colorization** | Python, Computer Vision, Deep Learning | GitHub

- Built an image colorization project that converts black-and-white images into colorized RGB outputs using image-processing and ML techniques.

Mini GPT: Lightweight Text Generation | Python, NLP, Generative AI | GitHub

- Built a small GPT-style model to learn how language models predict context and generate text.

GenCode AI: AI Coding Assistant | Live demo

- Built an AI coding assistant for code generation, code explanation, and debugging suggestions.

Technical Skills

- **Core:** Python, Pandas, NumPy, Scikit-learn, XGBoost, LightGBM, Random Forest, feature engineering, model evaluation and validation, time-series forecasting **Visualization & tools:** Matplotlib, Seaborn, Plotly, Power BI, Streamlit, Git, GitHub
- **Familiar with:** PyTorch, Transformers, NLP, computer vision, generative AI and LLM applications, REST APIs, React (via Lovable)
- **AI-assisted development:** Lovable, Claude Code, Google Antigravity, ChatGPT. I treat AI output as a draft: I review and test every change, check results against the data, and rerun the analysis so findings stay reproducible.

Education

- **Intermediate, Computer Science**, Government Degree College (SRE Majeed), Karachi |Expected 2027
- **Matriculation, Computer Science**, Board of Secondary Education, Karachi | A+ (80.45%)

Achievements

- 1st place, Machine Learning Hackathon, Saylani Mass IT Training (S.M.I.T)
- 1st position, Python examination, S.M.I.T
- Data Science & Artificial Intelligence Diploma, S.M.I.T (Apr 2024 – Dec 2025)

Certifications

- AI & Data Science, Saylani Mass IT Training
- Data Science, Machine Learning & AI, eHunarMand Pakistan
- Python for Data Science, IBM Skills Network
- AI Prompt Engineering, IBM Cognitive Class
- Power BI Fundamentals, Cambridge International

Languages

English, Urdu