

SARDER JUNAID AHMED

Data Analyst — Data Scientist — Machine Learning & Statistical Analytics

Dhaka, Bangladesh • +880 171-820-0525 • junaidahmedrupok@gmail.com

[Portfolio](#) • [GitHub](#) • [StatsPro Demo](#)

Data Analyst/Scientist with a track record of building end-to-end analytics platforms and production-ready ML systems. Turn raw data into automated insights and business recommendations. Published researcher with award-winning work. Experienced across classical ML, deep learning, NLP, and statistical analysis. Focused on delivering measurable impact, not just models.

EDUCATION

Bachelor of Science in Computer Science & Engineering Rajshahi University of Engineering & Technology (RUET), Bangladesh CGPA: 3.08/4.00 • Thesis Grade: A+	Feb 2025
Higher Secondary Certificate (HSC) Pabna Cadet College, Bangladesh • Perfect GPA: 5.00/5.00	2018
Secondary School Certificate (SSC) Pabna Cadet College, Bangladesh • Perfect GPA: 5.00/5.00	2016

SKILLS

Languages & Tools: Python (pandas, numpy, scikit-learn, XGBoost) — SQL — Git — Docker — Streamlit — Linux

Visualization: Matplotlib — Seaborn — Plotly — Power BI

Statistics: Hypothesis Testing — Regression — Clustering — PCA — Probability Theory — Bayesian Inference

Machine Learning: Classification — Ensemble Methods — Cross-Validation — Feature Engineering — SMOTE — XGBoost — NLP (TF-IDF) — Transformer Fine-Tuning (DistilBERT)

PORTFOLIO PROJECTS

StatsPro — Full-Stack Automated Statistical Analysis Platform Live Demo • GitHub - Built automated web app that transforms CSV → cleaned data → statistical tests → ML models → professional PDF/Excel reports in under 60 seconds; reduces analyst time by 95% Technical: 10 statistical tests, 15+ visualizations, 6 ML models with cross-validation, AI-powered insights; deployed on Streamlit Cloud	<i>2025 – Present</i>
IMDb Transformer Sentiment Analysis — DistilBERT Fine-Tuning GitHub - Fine-tuned pretrained DistilBERT on 10,000 IMDb reviews for binary sentiment classification; achieved 87.05% test accuracy with 99.5% inference confidence on strong positive inputs Technical: HuggingFace Transformers, PyTorch, AdamW + linear warmup scheduler, T4 GPU (Google Colab); training loss reduced from 0.4449 → 0.1615 over 3 epochs; compared against classical ML baseline	<i>2025</i>
IMDb Sentiment Analysis — NLP & Classical ML Comparison GitHub - Classified 50,000 movie reviews as positive/negative using TF-IDF vectorization; compared 3 models with Logistic Regression achieving 89.76% accuracy and 90.9% recall Technical: Unigram + bigram TF-IDF (10k features), stratified 80/20 split, Logistic Regression vs. LinearSVC vs. Naive Bayes; pipeline saved as production-ready <code>.pkl</code> artifacts	<i>2025</i>
Country Environmental Livability Analysis — Global Sustainability Benchmarking GitHub - Ranked and clustered 150+ countries by livability using 2019 global indicators (air quality, water access, sanitation, green space) Technical: K-Means and hierarchical clustering, PCA, correlation analysis, and geospatial visualizations; outputs interpretable country-level scores for sustainability research	<i>2024</i>

Economic Development Classification — ML for Global Policy Insights

2024

[GitHub](#)

- Built ML system achieving **98.3% accuracy** classifying 150+ countries by income level using World Bank indicators
- **Technical:** Random Forest, XGBoost, ensemble methods; missing data handling and multicollinearity control via VIF

Crime Pattern Analytics — Data-Driven Public Safety

2023 – 2024

[GitHub](#)

- Built ensemble classifier achieving **93% accuracy** predicting 6 crime types from spatial-temporal features; identifies high-risk zones for patrol allocation
- **Technical:** XGBoost, SMOTE for class imbalance, feature engineering from location/time data

Annual Enterprise Survey — Financial Performance EDA (New Zealand)

2026

[GitHub](#)

- Conducted comprehensive EDA on 55,620 enterprise records across 117 industries and 40 financial variables; uncovered concentration patterns showing manufacturing sectors dominate financial metrics
- **Technical:** One-way ANOVA across 5 categorical dimensions (all $p < 0.001$); right-skew and outlier analysis via log-scaled box plots; Pandas, NumPy, SciPy, Seaborn, Matplotlib

ANALYTICS EXPERIENCE

Researcher

Royal Scientific Publications, Bangladesh

- Conduct rigorous mathematical research and develop novel solutions addressing open problems in applied mathematics and computational science
- Collaborate on peer-reviewed publications synthesizing mathematical theory with practical applications
- Apply measure-theoretic and probabilistic frameworks to complex research problems

Data Analyst (Research Lab)

2022 – 2024

Young Learners' Research Lab, Rajshahi University of Engineering & Technology

- Led data analysis across cross-disciplinary projects; built reproducible pipelines and documentation standards adopted lab-wide
- Mentored 5+ junior analysts on Python and statistical methods; one mentee published their first paper as a direct outcome
- **Recognition:** Esteemed Alumni Award (2024); Certificate of Appreciation for Research Mentorship Excellence

Independent Data Analyst (Self-directed)

- *Annual Enterprise Survey (2026):* Analyzed 55,620 enterprise records across 117 industries; identified statistically significant patterns, built data cleaning pipeline and professional dashboards; [GitHub](#)

CERTIFICATIONS

- **Introduction to Data Science in Python** — DataCamp (2026)

TEACHING & MENTORING

- Tutored 50+ students in Statistics, Mathematics, and Programming; 5 achieved perfect GPAs on national exams; 4 admitted to top engineering universities
- Mentored 10+ university peers in Python, Machine Learning, and Statistical Analysis

ADDITIONAL

- **Communication:** Hosted weekly live radio show (2020–2021) — storytelling, audience engagement, live presentation under pressure
- **Awards:** Talentpool Scholarship (national merit), Principal's Prize for Outstanding Academic Performance