

ANAS NORANI

BS Data Science, NUST SEECS | AI & Data Science Research | Machine Learning

Islamabad, Pakistan | anorani.bsds24seecs@seecs.edu.pk | +92 320 806 8311 | linkedin.com/in/anas-norani | github.com/anasnorani1

RESEARCH PROFILE

Third-year Data Science undergraduate at NUST SEECS with research experience in machine learning, time-series classification, statistical modelling, anomaly detection, NLP, and academic manuscript writing. Co-authored an IEEE-format research paper involving literature review, experimental design, model benchmarking, cross-validation, interpretability analysis, and results reporting. Experienced in Python, Excel reporting, API-driven data validation, and remote analytical work. Seeking a paid remote Research Intern or Research Assistant role in AI & Data Science, with strong interest in literature review, research experimentation, technical writing, and deadline-driven academic research.

EDUCATION

B.S. Data Science

2024 – 2028

National University of Sciences & Technology, School of Electrical Engineering & Computer Science

- Relevant coursework: Machine Learning, Probability & Statistics, Advanced Statistics, Data Structures & Algorithms, Introduction to Data Science, NLP, Data Visualization.
- Academic focus: applied machine learning, interpretable AI, time-series modelling, data-driven decision systems, and research-oriented AI applications.

RESEARCH EXPERIENCE

Research Paper: F1 Active-Aerodynamic Proxy Forecasting

Jan 2026 – Jun 2026

IEEE-format Manuscript; Machine Learning, Time-Series Classification, and Interpretable AI

Tools: Python, PyTorch, Scikit-Learn, XGBoost, FastF1, Integrated Gradients, Bootstrap Testing

- Co-authored a complete IEEE-format research manuscript benchmarking machine learning and deep learning models for one-second-ahead Formula 1 telemetry-based aerodynamic proxy forecasting.
- Conducted a structured literature review covering active aerodynamics, motorsport telemetry, time-series classification, anomaly detection, model generalisation, and interpretable ML.
- Designed a two-phase ML study using 159,712 telemetry frames from four architecturally different F1 circuits: Monaco, Monza, Silverstone, and Suzuka.
- Identified a methodological weakness in the original framing and helped redesign the study into a stronger anticipatory time-series classification benchmark.
- Implemented and compared nine models, including Logistic Regression, Random Forest, XGBoost, CausalLSTM, CausalGRU, TCN, CNN1D, and CausalTransformer under Leave-One-Circuit-Out evaluation.
- Applied Integrated Gradients for interpretability and diagnosed circuit-specific speed memorisation as a key generalisation limitation in recurrent models.
- Found that a compact Logistic Regression model achieved strong performance with $AUC = 0.963 \pm 0.021$, outperforming several deep learning architectures.
- Contributed to methodology writing, related-work synthesis, results interpretation, robustness checks, figures, tables, and final academic formatting.

WORK EXPERIENCE

Data Research Analyst

Mar 2026 – Present

BisViews

- Conducted commercial and business registry research across Angola, Colombia, Chile, Vietnam, Dominican Republic, Italy, and the United Kingdom.
- Built structured validation workflows for business names, addresses, categories, franchise status, geographic consistency, and country-level data quality.
- Used Google Places API and Google Geocoding API to verify business identity, address quality, coordinates, and foreign-location leakage.
- Prepared Excel-based audit reports highlighting duplicates, closed businesses, inconsistent addresses, taxonomy issues, franchise/solo mismatches, and records requiring manual verification.
- Automated parts of the data quality workflow using Python, pandas, and openpyxl, improving consistency and reducing manual review time.
- Worked under strict validation requirements, strengthening attention to detail, documentation quality, independent problem-solving, and deadline-based delivery.

Data Analytics Intern

Jun 2026 – Present

Hackney Digital, UK – Remote

- Analyzed Shopify and ecommerce performance data, including sessions, conversion rate, average order value, completed checkouts, traffic sources, customer type, and revenue trends.
- Prepared analytical reports explaining sales changes, traffic quality, customer behaviour, checkout performance, and channel-level performance shifts.
- Investigated conversion-rate and revenue decline scenarios by comparing current performance against previous periods and

identifying traffic-source-level drivers.

- Communicated findings in clear business language for non-technical stakeholders, including whether revenue changes were caused by traffic volume, checkout drop-off, AOV, or customer mix.

Data Analytics Job Simulation

Jun 2026

Deloitte Australia – Forage

- Completed a Deloitte Australia job simulation involving data analysis, forensic technology, Tableau dashboarding, and Excel-based business analysis.
- Created an interactive Tableau dashboard to analyze operational data, identify business trends, and communicate key insights clearly.
- Used Excel to classify equality scores, segment data into meaningful categories, and draw business conclusions from structured datasets.
- Strengthened practical skills in dashboard development, spreadsheet-based analysis, data interpretation, and consulting-style insight communication.

...

SELECTED RESEARCH & DATA SCIENCE PROJECTS

Climate Variability & Drought Risk Analysis in Pakistan

Python, SciPy, Statsmodels, Plotly, Pandas, ERA5-Land Data

- Conducted multi-station statistical analysis across 15 Pakistani cities using climate observations from 1990–2023.
- Computed SPI drought indicators at 3-, 6-, and 12-month timescales and compared probability distributions using maximum likelihood estimation and AIC.
- Applied ARIMA modelling, Mann-Kendall trend testing, two-way ANOVA, Tukey HSD, and regression analysis to study drought and climate variability patterns.

Patient Readmission Risk Prediction

Python, Scikit-Learn, XGBoost, Pandas, Cross-Validation

- Compared Logistic Regression, Random Forest, and XGBoost for 30-day hospital readmission risk prediction using cross-validation and feature importance analysis.
- Used feature engineering, missing-value handling, and model comparison to identify key clinical risk factors for healthcare decision support.

Real-Time Reddit Sentiment Analysis Dashboard

Python, PRAW, RoBERTa, VADER, Streamlit, SQLite

- Built an NLP pipeline for collecting Reddit posts and analyzing sentiment using both VADER and transformer-based RoBERTa models.
- Developed a Streamlit dashboard with background ingestion, SQLite storage, model switching, and real-time sentiment monitoring.

TECHNICAL SKILLS

Programming	Python, R, SQL, Java, C++
Machine and Deep Learning	Scikit-Learn, XGBoost, PyTorch, TensorFlow, Logistic Regression, Random Forest, Time-Series Classification
Data Analysis	Pandas, NumPy, SciPy, Statsmodels, Excel, Data Cleaning, EDA, Statistical Testing
NLP	HuggingFace Transformers, RoBERTa, VADER, NLTK, Text Classification, Sentiment Analysis
Research Methods	Literature Review, Paper Writing, Experimental Design, Cross-Validation, Bootstrap Confidence Intervals, Model Interpretability
Visualization	Matplotlib, Seaborn, Plotly, Power BI, Tableau
Tools	Git, GitHub, Jupyter Notebook, Google APIs, SQLite, Streamlit, LaTeX, Overleaf

CERTIFICATIONS & LEADERSHIP

- **Programming in Python** - Meta, Coursera.
- **AI Developer** - IBM, Coursera.
- **General Secretary, NUST Entrepreneurs Club:** contributed to student innovation, event coordination, entrepreneurship-focused campus activities, and communication management.
- **SEECs Open House 2026 Coordination:** co-led coordination involving students, projects, and departmental participation across CS, SE, and ECE.
- **Campus Ambassador, National Incubation Center:** promoted innovation-driven startup initiatives and connected student cohorts with entrepreneurship resources.