

Diksha Damahe

B.Tech Computer Science and Engineering (Artificial Intelligence and Machine Learning)

+91 97709 25370 | dikshadamahe1@gmail.com

LinkedIn | GitHub | ORCID ID | Scopus ID

Profile

B.Tech CSE (AI & ML) student at VIT Bhopal with strong foundations in **applied statistics**, **hypothesis testing**, **time-series analytics**, and **reproducible Python research infrastructure**. Experienced in building statistical modeling pipelines, backtesting discipline, feature engineering, and leakage-safe out-of-sample validation to prevent overfitting and look-ahead bias. Active Leader in the **FinTech Club**, passionate about bridging fundamental investment hypotheses with rigorous quantitative testing, factor modeling, and market data analytics.

Education

VIT Bhopal University, Bhopal, Madhya Pradesh

B.Tech in Computer Science and Engineering (Artificial Intelligence and Machine Learning), CGPA: **8.06**
2023 – 2027

Relevant Coursework: Applied Statistics, Econometrics Fundamentals, Probability Theory, Machine Learning, Data Structures & Algorithms, Database Management Systems.

Technical Skills

Quantitative & Statistics: Hypothesis Testing (t-tests, p-values, multiple testing corrections), Econometrics, Time-Series Analysis, Linear/Logistic Regression, Factor Modeling, Cross-Sectional Ranking, Out-of-Sample Validation, Signal Construction, Biases Prevention (Look-Ahead Bias, Survivorship Bias, Overfitting Control)

Python Research Stack: NumPy, Pandas, statsmodels, scikit-learn, SciPy, Matplotlib, Seaborn, Jupyter, Git, Reproducible Code Practices

Backtesting & Financial Concepts: Point-in-Time Data Handling, Transaction Cost & Slippage Modeling, Turnover Analysis, Risk Metrics (Sharpe Ratio, Max Drawdown), Market Microstructure Basics, NSE/BSE Equity Fundamentals

Machine Learning & Analytics: XGBoost, LightGBM, Isolation Forest (Anomaly Detection), Feature Selection, SHAP Explainability, Model Calibration

Databases & Core CS: SQL (PostgreSQL, Advanced Queries), Data Structures & Algorithms, Object-Oriented Programming (OOP), REST APIs

Experience

Hindustan Copper Limited (HCL), Malanjkhand, Madhya Pradesh

May 2026 – July 2026

Quantitative Data & AI Engineering Intern

- Engineered **RecovAI**, a real-time statistical and predictive process monitoring engine for industrial concentrator operations, analyzing 18+ months of time-series data to forecast recovery rates with an R^2 of **0.9720** and **MAE of 0.1042%**.
- Built data preprocessing and feature transformation pipelines incorporating time-lagged variables, rolling window stats, and robust outlier removal to handle non-stationary sensor feeds without look-ahead leakage.
- Deployed Isolation Forest anomaly detection to identify operational deviations and optimized multi-variable chemical dosing recommendations with clear interpretability (SHAP).

Projects

RecovAI — AI-Powered Copper Flotation Intelligence System

GitHub | Live Demo

Python, FastAPI, Streamlit, XGBoost, Isolation Forest, scikit-learn, Pandas, Matplotlib, SHAP

- Developed a industrial process analytics platform analyzing 18+ months of sensor data to forecast copper recovery with an R^2 of **0.9720** and **MAE of 0.1042%**.
- Implemented non-stationary time-series feature engineering, Isolation Forest anomaly detection, reagent dosing optimization, and SHAP explainability.

Quantitative Factor Research & Backtesting Engine (Python)

GitHub

Python, Pandas, NumPy, statsmodels, scikit-learn, Matplotlib

- Built a reproducible modular Python backtesting framework to test cross-sectional momentum and mean-reversion factor hypotheses on financial time-series data.
- Implemented point-in-time data handling, lag controls, dynamic transaction cost/slippage adjustments, and turnover calculation to ensure realistic strategy performance metrics.
- Conducted hypothesis testing on alpha signals, applying multiple testing corrections (Bonferroni / FDR) to guard against p-hacking, data snooping, and spurious correlations.
- Evaluated strategies via Sharpe ratio, Sortino ratio, max drawdown, and out-of-sample walk-forward validation to verify robustness against regime shifts.

NAFLD Risk Stratification Pipeline — Statistical ML Benchmark

GitHub

Python, scikit-learn, XGBoost, LightGBM, CatBoost, statsmodels, SHAP

- Designed and statistically benchmarked 24 classification models on NHANES multi-variable dataset, achieving top ROC-AUC of 0.9679 with rigorous cross-validation.
- Enforced strict leakage-safe validation using stratified k-fold splits and in-fold feature selection to prevent overfitting and optimistic evaluation bias.
- Conducted feature attribution using SHAP and statistical significance tests, making complex model outputs legible and actionable for domain specialists.

Chemical Equipment Parameter Analytics & Visualizer

GitHub | Live Demo

Django, React, PyQt5, Pandas, Matplotlib, ReportLab

- Developed a cross-platform data visualization tool for processing high-frequency sensor and parameter datasets.
- Built interactive trend charts, automated summary reporting, and structured dataset history endpoints for clear, transparent data communication.

Aegis Vault – Secure Data Management Dashboard

GitHub | Live Demo

Next.js, TypeScript, MongoDB, Prisma, Tailwind CSS

- Built an interactive data management application with real-time search, categorization filters, and encrypted data processing workflows.

Veil — Minimalist Android Streaming Aggregator

GitHub

Flutter, Dart, ExoPlayer, TMDB API, TypeScript

- Developed a cross-platform Android streaming media aggregator leveraging **TMDB API** for metadata and **ExoPlayer** for video playback.
- Integrated a self-hosted custom resolver (cinepro-org/core OMSS) to dynamically aggregate and resolve stream sources.

Additional Projects (Selected):

- **Shruti – Neumorphic Bilingual Audio-Text Syncer** (Flutter, Dart, Riverpod, Firebase, Firestore, Whisper AI): bilingual mobile app using Whisper AI to synchronize Osho discourses with real-time text. [GitHub](#)
- **IELTS Precision & Progress** (React, TypeScript, Vite, Tailwind CSS, Framer Motion): conversion-focused platform with metrics, testimonials, and micro-interactions. [GitHub](#) | [Live Demo](#)
- **VyomGarud – Counter-UAV Defense Website** (Next.js, React, Tailwind, Framer Motion): cinematic, data-led interface design with animated metrics and responsive visual storytelling. [GitHub](#) | [Live](#)
- **Hostel-Bites – Meal Management Platform** (JavaScript, Firebase, Firestore): real-time order

Publications

Secure Phishing Analysis and Robust Training Algorithm (SPARTA) for Phishing Website Detection

Diksha Damahe, Ajay Kumar Phulre

2026 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS), Bhopal, India

Detection rate: **99.15%** on PhishTank and UCI benchmarks; ensemble-based statistical detection framework.

DOI: 10.1109/SCEECS68810.2026.11430134 | Scopus ID: 60617196400 | ORCID ID: 0009-0002-0499-2572

Certifications

- **Applied Machine Learning in Python** — University of Michigan (Nov 2025)
- **Advanced SQL for Technical Interviews** — Bosscoder Academy (Feb 2025)
- Cloud Computing — Elite NPTEL Online Certification, IIT Kharagpur (Jan–Apr 2025)
- Introduction to Internet of Things — Elite NPTEL Online Certification, IIT Kharagpur (Jan–Apr 2026)
- AWS Technical Essentials & Certification — Amazon Web Services / Intellipaat (Sep 2025)
- Python 3 & Data Processing Essentials — Vityarthi / Lernx

Leadership & Quantitative Interest

Content & Research Lead, FinTech Club — VIT Bhopal University

- Lead research presentations and content discussions on quantitative finance topics, algorithmic trading concepts, factor investing, and financial market metrics.
- Organize workshops bridging computer science concepts with modern financial technology and data-driven investing strategies.

Volunteer, National Service Scheme (NSS) — VIT Bhopal University

- Participated in community outreach initiatives and selected in district-level Pre-RDC selection parade.