

UTKARSH SARAOGI

857-426-9744 | saraogi.u@northeastern.edu | <https://www.linkedin.com/in/us1610> | <https://github.com/utkarsh1610-s>

EDUCATION

Northeastern University, Boston, MA

Exp. May 2027

Master of Science in Computer Science

- Coursework: System Design, Database Management Systems, Algorithms, Machine Learning, NLP

Manipal Institute of Technology, Manipal, KA

May 2022

BTech in Electronics and Communication

- Coursework: Data Structures & Algorithms, Data Science, Python, Cloud Computing

TECHNICAL SKILLS

Data Engineering: Airflow, dbt, SnowFlake, Apache Spark, DuckDB, ETL/ELT Pipelines, Data Warehousing

Database Systems: MySQL, MongoDB, Star Schema Design, Medallion Architecture, Parquet/Partitioning

Cloud & DevOps: AWS (S3, Lambda, SageMaker), Docker, Git, Jenkins CI/CD, Distributed Systems, Microservices

Programming & Core CS: Python, SQL, Java, Data Structures & Algorithms, System Design, MATLAB

Data Science & ML: A/B Testing, Causal Inference, Statistical Analysis, Scikit-learn, LLM Interpretability

PROFESSIONAL WORK EXPERIENCE

HEWLETT PACKARD ENTERPRISE

Aug. 2022 – Aug. 2025

Cloud Developer I

Bangalore, India

- **Data Pipeline Observability & DataOps:** Monitored distributed data flows across 15+ microservices using real-time logging; implemented data quality validation and root cause analysis, **reducing mean time to resolution by 40%** through systematic observability practices.
- **Regression Optimization:** Automated regression testing across 1,200+ test cases with parallel execution; **reduced test-cycle times by 55%** through optimized pipeline orchestration and cross-functional collaboration with DevOps teams.
- **ML-Powered Triage System:** Led triage team operations while contributing to the training of an AI assistant to automate failure classification by sourcing and labeling data through **4,000+ failure scenarios** (test, feature, server), achieving **65% improvement in classification accuracy** and reducing manual triage time.
- **Production Release Management:** Owned end-to-end testing lifecycle using Jenkins CI/CD; validated production deployments through comprehensive smoke, integration, and regression test suites.

HEWLETT PACKARD ENTERPRISE

Jan. 2022 – Jul. 2022

Graduate Intern

Bangalore, India

- **Test Automation & QA:** Executed daily automation suites, identified software defects, and contributed fixes to **Maintain test framework health** for the Switch Data Monitoring platform.
- **Code Quality Automation:** Created **automated PyLint validation scripts** integrated into the CI pipeline, eliminating production rejections due to lint errors and reducing pre-release debugging time.

PERSONAL PROJECTS

SoundStream Analytics | ELT, dbt, Airflow, duckDB, Pyspark, SQL

- **Data Warehousing & Pipeline Orchestration:** Built production-grade ELT pipeline processing 46M+ music records using medallion architecture (**Bronze → Silver → Gold**); designed star schema dimensional model for scalable analytics.
- **Data Quality & Performance Optimization:** Automated data validation, **removed 5,043 duplicates**, **handled schema drift**; achieved **60% storage compression** through columnar Parquet format with intelligent partitioning strategy.

Credit Risk Strategy Engine: Revenue Recovery | A/B Testing, Scikit-Learn

- **Statistical Modeling & Experimentation:** Developed fraud-resistant risk engine using **HistGradientBoosting** and **Counterfactual A/B Testing**; engineered compensating factor features from verifiable data, **recovering \$7.8M in potential revenue** and **reducing false rejections by 35%**.

Stock Portfolio Management System | Python, MySQL, REST-APIs, ETL

- **Data Pipeline & API Development:** **Built end-to-end data pipeline** extracting real-time market data from Finnhub API; **implemented SQL stored procedures** for data transformation and served portfolio analytics through REST API.

RESEARCH ACTIVITIES

- Researching LLM sycophancy at Northeastern—analyzing model responses across 3 models using hidden state probes to detect when models abandon correct answers under social pressure; exploring guardrail mechanisms to prevent flipping.
- Mentored 3 interns at HPE through comprehensive documentation and live knowledge transfer sessions reducing onboarding time by 2 months.