

Shriya Dikshith

LinkedIn | [Portfolio](#) | Boston, MA | (617) 390-4991 | shriya5dxt@gmail.com

Education

Northeastern University, *Master's in information systems*, Boston, USA

Coursework: Data Management and Database Design, UI/UX, Web Design, Data Science Engineering Methods, Data Structures

Technical Skills

Programming languages	: Python, SQL, R, Java, HTML5, CSS, JavaScript, MATLAB, ETL/ELT
Databases	: MS SQL, Vertica, Snowflake, Oracle, MongoDB, NoSQL, Big query, PostgreSQL, GCP
Libraries and Packages	: Keras, TensorFlow, PyTorch, Flask, PPTx, NumPy, SciKit Learn, lambda functions, PySpark
Visualizations	: Tableau, Power BI, Excel, Informatica, Alteryx, Looker, ggplot, Plotly
Other	: GIT, dbt, Dimensional data modeling, ETL, BI tools, DAX, OLAP, Data Governance, Agile Methodologies, SQL Server, AWS, User Experience, Databricks, VBA, LLMs, Google Cloud, Azure Databricks

Work Experience

Amazon, Boston, MA, USA, *ML Data Associate*

Jul. 2024 – Present

- Optimized LLM data processing workflows, contributing to a 40% increase in model accuracy for Alexa smart home models by refining data inputs and enhancing feature engineering processes
- Enhanced the performance of Alexa smart home models by conducting detailed dataset reviews, identifying patterns, and optimizing workflows, leading to a 20% increase in model accuracy, and a 25% improvement in team efficiency and data integrity

Mimecast, Lexington, MA, USA, *Data Engineer*

Jan. 2023 – Aug. 2023

- Partnered with data engineers and product owners to manage including the development and optimization of dbt models, ingesting data to Snowflake, reducing ELT runtime from 12 hours to under 20 minutes
- Enhanced data quality and governance** by leading the migration of Tableau workbooks and flows from Tableau Server to Tableau Cloud, achieving a 40% reduction in migration time and resolving systemic data quality issues
- Conducted comprehensive data analyses and presented KPIs to key stakeholders, including the identification of 10 high-impact workbooks, leading to a better data usage and strategic decision-making

Northeastern University, Boston, USA, *Data Analyst*

Aug. 2022 – Jan. 2023

- Automated Snowflake ETL/ELT processes using Alteryx Designer, dbt and Informatica Cloud, achieving an 80% efficiency increase, showcasing skills in database scalability and performance enhancement
- Created a real-time Tableau dashboard that enabled finance teams to identify cost-saving opportunities, reducing operational expenses by 15% and saving 20+ man-hours per week through data-driven insights

HP Inc., Bangalore, India, *Data Scientist (Customer Experience)*

Jan. 2020 – Jan. 2022

- Engineered a custom BERT model using NLP and LLMs for hierarchical issue detection, processing 30M customer requests across 50+ languages; sentiment analysis on social media text with PyTorch, driving actionable insights from customer reviews
- Directed a CX team in developing a **Python pipeline** and Tableau dashboard for HP's global printers, cutting resolution time by 20%, saving \$1.5M annually, and reducing Databricks GPU cluster costs by \$10,000/month through strategic error analysis
- Pioneered the creation of a Python **ETL data pipeline** in **Azure Databricks**, achieving a 90% automation rate for weekly large data loads, exemplifying leadership in big data processing and database management excellence
- Delivered end-to-end analytics, from problem statement to model deployment, and created a **Tableau dashboard**, reducing customer support issues by 25% and earning recognition for strategic insights in the Weak Signals project

Academic Projects

Customer Complaint Management Database System | *SQL, Tableau, Schema design, Oracle – PL/SQL* 🐍

- Designed and implemented a customer complaints management database in Oracle, incorporating ER diagrams, normalized schema, and ACID transactions to efficiently track, analyze, and resolve product, account, payment, and service-related issues within warranty periods, leading to improved customer satisfaction

Building a Robust API | *JSON, Elastic Search, Redis, Springboot, Postman, MongoDB, VSCode*

- Implemented a REST API for structured JSON data with CRUD, merge, and delete functionalities, ensuring data validation and providing a detailed JSON schema for the use case

Amazon Product Review Text Mining | *Python, NLP, Pandas, Machine learning algorithms, NumPy* 🐍

- Employed a python based verbatim analysis with PyATE framework to identify and address common electronic product issues, leading to improved product quality assessment and increased customer satisfaction