Samar Dikshit

Cambridge, MA | 929-217-0015 | samard05@gmail.com | LinkedIn | Website

Technical Skills

Programming Languages	Python, SQL, R, Cypher
Data Science Technologies	pandas, NumPy, scikit-learn, Matplotlib, Seaborn, Jupyter, PyTorch, pytest, NetworkX, dbt, Snowflake, git, Tableau, Dataiku, Elasticsearch, tidyverse, caret, VSCode, conda, Neo4j, REST APIs

Professional Experience

Sonos, Inc. – Boston, MA **January 2022 - Present**
Senior Data Scientist - Product Data (October 2023 - Present)

- Led a cross-functional team with hardware and user interface designers focused on understanding user behavior, control patterns, and feature discoverability across speaker categories to drive decisions for new products and features e.g. headphone design, optimal button placement, Bluetooth availability, power/battery saver mode etc.
- Supported the development of Gaussian Process models in Python to automatically set speaker volumes, reducing volume change interactions for users by over 50%.
- Built a feature generation framework with dbt and Snowflake that creates 7-10 million rows of product usage-based feature values per day for use in downstream machine learning tasks.
- Contributed to the development of a Python-based machine learning pipeline to cluster and segment over 10 million users using k-means with over 80 features.

Data Scientist - Product Data (January 2022 - September 2023)

- Developed SQL-based statistical and probabilistic methods to discover and predict usage routines for over 8 million users, achieving R^2 values between 0.6 - 0.85, and productionized them using dbt.
- Built a proof-of-concept collaborative recommendation system using Python for Sonos Radio, a radio/music service with an average weekly user base of 1 million.
- Collaborated in designing and conducting user research studies to understand the effect of new software automations on user behavior, leading to the development of improved volume control and speech enhancement features.

DTonomy Inc. – Cambridge, MA **January 2021 - August 2021**
Analytics Engineering Intern

- Analyzed cyberattack data in Python based on the MITRE ATT&CK database, and developed patterns to identify Defense Evasion and Lateral Movement attack patterns for the SOAR platform.
- Developed Node-RED automations to connect Elastic Security with the SOAR platform.

Center for Complex Network Research – Boston, MA **June 2020 - January 2021**
Research Assistant

- Collected, processed, and explored data with over 1.5 million samples related to philanthropies, non-profits, and universities from the LittleSis API using Python.
- Constructed and analyzed a network with more than 97,000 relationships between socio-political entities using NetworkX and pandas, identifying key factors that influence grants and donations.
- The created network data was on display at the Postmasters Gallery in New York City from Sept. 2022 - Oct. 2022.

Northeastern University – Boston, MA **May 2020 - December 2021**
Teaching Assistant

- Advanced Programming with Data (Fall 21), Programming with Data (Fall 20), Algorithms and Data (Summer 20)

Pepper Cloud – Bangalore, India **January 2019 - July 2019**
Software Engineering Intern

- Worked on developing a customer-facing chatbot for a CRM platform using Node.js and Dialogflow.

Education

Northeastern University, Boston, MA **December 2021**
Khoury College of Computer Sciences
Master of Science in Data Science GPA: 3.79 / 4
Relevant Courses: Supervised & Unsupervised Machine Learning, Data Processing, Information Retrieval, NLP, Complex Networks

Manipal Institute of Technology, Manipal, India **July 2019**
Department of Information and Communication Technology
Bachelor of Technology in Computer and Communication Engineering (Data Analytics minor) GPA: 8.22 / 10
Relevant Courses: Big Data, Data Mining & Predictive Analytics, Database Systems, Pattern Recognition, Advanced Programming