Jhalak Surve

Boston, MA | (551) 344 8320 | surve.j@northeastern.edu | LinkedIn | Github

EDUCATION

Northeastern University *Boston, MA*
Master of Science in Information Systems *May 2024*
Relevant Coursework: AI Generative Modeling, Data Science Engineering Methods

Rajiv Gandhi Proudyogiki Vishwavidyalaya *Indore, India*
Bachelor of Technology in Computer Science and Engineering *June 2021*
Relevant Coursework: Data Warehousing, Machine Learning, Python Programming, Software Engineering

TECHNICAL SKILLS

Programming Languages:	Python, SQL, R, Java
AI and Machine Learning:	Generative Models (GAN, VAE, Transformer), Neural Networks, LLMs
Databases & Data Management:	Microsoft SQL Server, Google Bigquery, MySQL, MongoDB
Tools:	Git, Docker, AWS (EC2, Lambda, S3), Linux, Agile, Apache Kafka, Django

PROFESSIONAL EXPERIENCE

SOFTWARE ANALYST | Yardi Software India Pvt. Ltd. | Pune, India *July 2021 – August 2022*

- Optimized data storage and ETL processes by developing and managing a repository of 500+ SQL Server database objects and by leveraging SQL Server Integration Services (SSIS) to enhance data pipelines.
- Developed 15+ reports using SQL Server Reporting Services (SSRS) and utilized Google Bigquery to manage 300+ client customizations across 4 FinTech projects, leading to advanced data analysis and reporting.

DATA ANALYST INTERN | Gyrix Technolabs LLP | Indore, India *October 2020 – December 2020*

- Contributed to implementing a global data integration strategy using Apache Kafka and Python, optimizing cross-border collaboration, and reducing project development timelines by 10 days.
- Enhanced AI-driven customer support using Natural Language Processing (NLP) techniques, resulting in a 15-minute reduction in response time.

PROJECTS

Travel Chatbot using ChatGPT *February 2024 – March 2024*

- Developed a travel chatbot with Django and React for easy exploration of travel details.
- Implemented GPT-3.5 integration for processing natural language queries, converting them into SQL queries for database interaction facilitating dynamic data retrieval and response generation.

Aspect Based Sentiment Analysis *December 2023 – December 2023*

- Attained 93% accuracy in sentiment analysis of processed reviews using NLTK and TextBlob, advocated Parts of Speech (POS) tagging to extract key aspects from 50,000 reviews and inform targeted improvements.
- Engineered sentiment extraction using pattern recognition, categorizing customer reviews in 5 classes with 81% test accuracy using SVM and SemEval-2015 Dataset.

ML Model Evaluation using H2O AutoML *September 2023 – November 2023*

- Conducted thorough predictive modeling on a Kaggle dataset using three methods (Random Forest, Logistic Regression, and KNN Classifier), and employed advanced techniques like H2O.ai AutoML and Statsmodels for model evaluation.

Boston Blue Bikes Analysis *January 2023 - April 2023*

- Implemented an ArangoDB multi-model database, integrating graph and document models, utilizing live data collected from GFBS feed using a python script with a daily data update frequency of 500 records.
- Created optimized AQL queries and Power BI dashboards, reducing query execution time from 5 s to 3 s.

COMMUNITY INVOLVEMENT

- Crafted 2 Kaggle tutorials for understanding Neural Network Type Classification using CNN and Typeface MNIST dataset, and Dimensionality Reduction Techniques such as PCA and t-SNE.
- Working as an Application Processor at Northeastern University which involves reviewing the application materials for Graduate and PhD level applications across 9 colleges.