Vishruth T Venugopal

 Boston, MA   teluguvenugopal.v@northeastern.edu   857 544 9260   Vishruth008   in vishruthtv

Education

Northeastern University   *Sept 2023 – May 2025*
MS in Information Systems

- **Coursework:** Advances in Data Science, Designing Advanced Database Architectures, Prompt Engineering and Gen AI

Don Bosco Institute of Technology   *Sept 2019 – May 2023*
BS in Information Science and Engineering

- **Coursework:** Database Management, Big Data Analytics, Data Mining and Data Warehousing, Python Application Programming

Technical Skills

Languages: SQL, R, Python (Pandas, NumPy, Matplotlib, Seaborn, Scikit-learn, TensorFlow, SciPy)
Analytical Tools: PowerBI, Tableau, Looker, MS Excel, Airflow, Talend, Alteryx, GitHub
Databases: MySQL, MongoDB, PostgreSQL, Snowflake, BigQuery, Microsoft SQL Server
ML Frameworks: PyTorch, XGBoost, LightGBM, H2O.ai

Experience

Data Science Intern   *Bengaluru, India*
Rapsol Technologies   *Aug 2022 – Oct 2022*

- Streamlined **data preprocessing pipelines** with **SQL, Pandas, and NumPy** to process **50,000+** records, reducing inconsistencies by **25%** and preparing datasets for model training.
- Optimized machine learning models with **Scikit-learn** and **TensorFlow**, applying **hyperparameter tuning** and **cross-validation** to improve reliability by **20%** and analyze **10,000+** data entries
- Created interactive Tableau dashboards to visualize key metrics, including **transaction patterns** and **customer demographics**, enabling actionable insights for strategic planning

Projects

History of Massachusetts RAG System Q&A Bot

- Built an interactive Q&A bot for Massachusetts history using a **Retrieval-Augmented Generation** framework, integrating **LLMs** with a vector database to enable semantic search across **100,000+** historical records, enhancing accessibility for educators and history enthusiasts
- Designed a user-friendly interface with **Streamlit and Python**, leveraging **pandas and NLTK** for data preprocessing and integrating LLMs to deliver accurate, context-aware responses with features like **refined queries and bookmarking**

Malicious URL Detection System using Machine Learning and Deep Learning   *GitHub*

- Developed a malicious URL detection system using machine learning models, including **XGBoost and LightGBM**, achieving over **95%** detection accuracy while optimizing data preprocessing and visualization workflows with **pandas, matplotlib, and seaborn**
- Implemented a **robust evaluation pipeline** with metrics such as **confusion matrix, classification report, and accuracy score**, streamlining model validation and improving efficiency by **20%**

Swiftbank - Bank Management Database System   *GitHub*

- Architected a MySQL database managing **10,000+** records for accounts, loans, transactions, and branches, ensuring data integrity with **indexing and constraints**
- Automated processes with **SQL triggers**, **stored procedures**, and **views**, while using **AES encryption** and Tableau to enhance security and visualize transaction patterns, boosting efficiency by **40%**