Vatsalya Anand

vatsalya.anand@gmail.com | (716) 247-3950 | linkedin.com/in/vatsalyaanand | github.com/vatsalyaanand

EDUCATION

University at Buffalo, The State University Of New York — **Buffalo, NY**
Master of Science in Data Science — May 2024

Birla Institute of Technology and Science (BITS), Pilani — **Pilani, India**
Bachelor of Engineering in Civil Engineering (Minor in Finance) — Jul 2021

SKILLS

Programming Languages: Python, R, MATLAB, SQL (PostgreSQL, Redshift, SQLite)
Application Software: NumPy, Pandas, Matplotlib, Plotly, Sklearn, Dplyr, Tidyr, ggplot2, Flask, Shiny, PyTorch, Rattle
Machine Learning: Regression, Classification, GAN, CNN, LSTM, Deep Learning, Transformers, Ensemble
Tools: Git, Docker, AWS Cloud, Azure, GCP, MS 365, PowerBI, Service Now, Hadoop, Spark

EXPERIENCE

Wipro — **Mumbai, India**
Data Scientist — Jul 2021 – Dec 2022

- Spearheaded implementation of predictive models, employing cutting-edge algorithms including decision trees and neural networks in R and Rattle. Achieved a cost reduction of 10% in cold calling operations.
- Pioneered development of automated visualization dashboards, seamlessly integrating BI tools such as Power BI and ServiceNow. This innovation streamlined workflows, eliminating manual interventions of 5 people.
- Innovatively engineered a Python-based web scraper to efficiently gather data from internal websites. Designed and executed a robust ETL process to seamlessly load obtained data into Microsoft SQL servers, enhancing data management and accessibility.

Cuemath — **Bengaluru, India**
Data Analyst Intern — Jul 2020 – Dec 2020

- Conducted in-depth analysis of Funnels, Uninstalls, and Retention Cohorts using a combination of Python and SQL (Redshift), deriving actionable insights to improve overall retention on app by 8%.
- Spearheaded integration of Big Query with CSV files stored in Buckets on the Google Cloud Platform, architecting new table schemas. Initiative revamped data storage and accessibility, contributing to streamlined data management processes and saving daily manual labor of approximately 1.5 hours.
- Applied advanced web scraping techniques using Selenium in Python to extract Daily Active Users (DAU) and Monthly Active Users (MAU) data from AppsFlyer. Leveraged pandas data frames to generate comprehensive reports, providing valuable metrics for performance evaluation and strategic planning.

PROJECTS

End to End MLOps | Python, Docker, AWS

- Designed and performed a comprehensive MLOps project, employing Python, Docker, and Flask to create a modular architecture for seamless data processing, model training, and deployment.
- Orchestrated CI/CD pipelines with GitHub Actions, deploying project onto AWS ECR and managing its execution on AWS EC2 instances, ensuring continuous integration and deployment in a production environment.

Fake News Detection | Python, Flask

- Engineered an effective fake news detection model using NLP and deep learning (CNNs, Bi-LSTMs), showcasing strong skills in Python programming and data preprocessing.
- Optimized and fine-tuned model for superior performance, culminating the integration into a user-friendly web application with Flask. This implementation allowed users to assess authenticity of news.

Analysis of Traffic Stop Receipts | Python, SQL, Flask

- Applied advanced SQL operations in PostgreSQL to analyze traffic stop data, revealing key patterns and trends through statistical analysis, enhancing data-driven decision-making in law enforcement.
- Leveraged Flask and Python to construct a dynamic web app seamlessly integrating with PostgreSQL. Interface allows stakeholders to efficiently access and visualize traffic stop data.