NIKHIL ADVANI

New York, USA

+1 (716) 939-6884 | nadvani@buffalo.edu | LinkedIn | GitHub

EDUCATION

Master of Professional Studies: Data Science and Applications **Aug 2023 - Expected Dec 2024**

University at Buffalo, The State University of New York | Buffalo, NY, USA **CGPA: 3.8**

Relevant Courses: Statistical Data Mining, Machine Learning, Database Management Systems, Python Programming

SKILLS & TOOLS

- **Programming Languages**: Python, R, SQL, DAX, C, C++, JavaScript
- **Tools**: Microsoft (Power BI, Office, Excel, Word, PowerPoint), Tableau, AWS S3, AWS Redshift, AWS Glue, MongoDB, Oracle, Groq, Streamlit
- **Data Science & Analytics**: Python (Pandas, NumPy, Matplotlib, SciPy, Scikit-learn), ETL, Data Science Pipeline (cleansing, wrangling, visualization, modeling, interpretation), Statistics, LLM
- **Machine Learning**: Regression, Classification, Clustering

EXPERIENCE

Volunteer Research Assistant – X-Lab@UB | NY, USA **Jun 2024 - Present**

- Initiated a role-based access control system within a **RAG** application using streamlit, enhancing security and privacy for over **100** pathologists and **500** clients
- Incorporated **1000+** research papers & numerous documents, providing comprehensive resources for pathologists & clients
- Employed Langchain tools and agents to streamline access control mechanisms, reducing document retrieval time by **40%**
- Enhanced the security framework by integrating **Qdrant** vector databased and **JSON Web Tokens (JWT)**

Data Scientist/Analyst - EPL Limited **May 2022 - Apr 2023**

- Managed end to end projects integrating **Power BI**, **AWS Cloud, Python,** and **SQL**
- Integrated data from sources such as **ODATA – SAP**, **MongoDB**, and **Postgres**, reducing retrieval time by **18%**
- Implemented ML models to predict sales, inventory, and production, achieving high accuracy rates
- Analyzed real time data to showcase important business metrics/visualizations and identified data anomalies
- Developed dashboards and written reports for **12** internal teams, enabling data-driven recommendations
- Supervised a team of 3 interns, guiding them through daily tasks and end-to-end projects

Software Development Intern - Numeric Infosystem Pvt. Ltd. **Aug 2021 - Nov 2021**

- Created an Inventory Management website using **Django** Framework, utilized by 5 companies, streamlining operations and enhancing efficiency by **20%**
- Drafted several **python** and **javascript** codes, increasing website interactivity by **25%**
- Collaborated, designed, and implemented HTML templates to elevate website's visual appeal
- Formulated **SQL** queries to retrieve and manipulate employee data within the database

ACADEMIC PROJECTS

Chatbot – Groq Inference Engine: Python, LLM, Llama, Generative AI (Link) **June 2024**

- Structured a **RAG** chatbot using **Groq** and opensource **Llama-3** model, reducing average response time by **50%**
- Combined multiple tools like Wikipedia, Pdf Search, and Arxiv, leading to intelligent information retrieval
- Set up **Chroma DB** vector database for embeddings storage and retrieval, boosting retrieval speed by **30%**
- Instituted **langchain agent** setup for accurate reasoning in much lesser time

News Research Chatbot: Python, LLM, OpenAI, Generative AI (Link) **May 2024**

- Built an LLM powered research chatbot using **Langchain**, **OpenAI**, and **FAISS**, reducing research time by 60%
- Established a data processing pipeline that loads and splits articles, achieving efficient chunking of data
- Integrated **OpenAI embeddings** and **FAISS indexing**, enabling storage and retrieval of relevant content, leading to a **30%** reduction in API call costs
- Streamlined user interface with **Streamlit**, facilitating user input and interaction

Real Estate Price Prediction: Python, ML (Link) **Feb 2024**

- Engineered models for predicting house prices through data analysis and **feature engineering**
- Applied preprocessing, including outlier removal and one-hot encoding, reducing data size by **27%**
- Predicted property prices using **Linear Regression**, **Decision Tree**, **Lasso,** and **Random Forest**
- Utilized **Grid Search CV** to determine optimal model and achieved **81.9%** accuracy

Car Sales Dashboard: (Link) **Jan 2024**

- Constructed and implemented **KPIs** for a concise overview of car sales performance
- Reviewed business requirements, translated them into actionable insights for a customized dashboard
- Displayed dynamic charts with interactive features for stakeholder insight into complex sales data