

NIKHIL ADVANI

New York, USA | +1 (716) 939-6884 | nadvani@buffalo.edu | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

SUMMARY

Results-driven Data Scientist with 1 year of experience in machine learning and data analytics. Proficient in Python, R, SQL, Power BI, and Tableau. Currently pursuing a master's in Data Science from the University at Buffalo, with expected graduation in December 2024. Actively seeking full-time opportunities to leverage my skills and contribute to impactful projects.

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: 3.9

Relevant Courses: Statistical Data Mining, Machine Learning, Database Management Systems, Probability & Statistics

SKILLS & TOOLS

- **Programming Languages:** Python, R, SQL, SAS, DAX
- **Tools:** Microsoft (Power BI, Office, Excel, Word, PowerPoint), Tableau, AWS (S3, Redshift, Glue, DMS), MongoDB, Oracle, Groq, Streamlit, Vector Database (Chroma, FAISS, Qdrant, Milvus, Weaviate)
- **Data Science & Analytics:** Python (Pandas, NumPy, Matplotlib, SciPy, Scikit-learn), ETL, Data Science Pipeline (cleansing, wrangling, visualization, modeling, interpretation), Statistics, LLM (OpenAI GPT, Llama, Mistral, Claude, Bard, Gemini)
- **Machine Learning:** Regression, Classification, Clustering

EXPERIENCE

Data Scientist/Analyst - EPL Limited

May 2022 - Apr 2023

- Developed a **logistic regression model** to offer relevant product features to at-risk clients, leading to a boost in monthly customer retention by **0.92%**
- Implemented time series forecasting models using **ARIMA & SARIMAX** for sales forecasting, enabling targeted marketing and boosting sales efforts by **22%**
- Designed **Power BI** dashboards to identify areas of improvement for sales and marketing teams, presented analysis to stakeholders, leading to an **YoY revenue gain of \$500k**
- Monitored data pipelines with AWS Redshift, S3, Glue, DMS for real-time integration, reducing data access delays by **32%**
- Mentored a team of 3 interns, guiding them through building **sophisticated reports** and end-to-end projects

Software Development Intern - Numeric Infosystem Pvt. Ltd.

Aug 2021 - Nov 2021

- Created an Inventory Management website integrating **Django, python, SQL, javascript**, utilized by 5 clients, streamlining operations and enhancing efficiency by **20%**
- Leveraged Python, SQL and Excel to clean and prepare client data, enabling data scientists to build predictive models that **lifted ROI by 7.3% annually**

RESEARCH EXPERIENCE

Research Assistant – UB School of Pharmacy | NY, USA

Sep 2024 – Present

- Developing a Retrieval Augmented Generation (**RAG**) System using **LLM's** to streamline drug development research, focusing on compliance regulations and disease recognition by integrating data from clinical trials and immunogenicity studies

Research Assistant – X-Lab@UB | NY, USA

Jun 2024 – Jul 2024

- Initiated a role-based access control system within a **RAG** application using Streamlit, securing access for 100+ pathologists and 500 clients, while integrating 1000+ research papers and documents. Leveraged **Langchain tools and agents** to streamline access, **reducing retrieval time by 40%**.

ACADEMIC PROJECTS

News Research Chatbot: Python, LLM, OpenAI, Generative AI ([Link](#))

May 2024

- Built an LLM powered research chatbot using **Langchain, OpenAI, FAISS** and **Streamlit** 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

Real Estate Price Prediction: Python, ML ([Link](#))

Feb 2024

- Engineered models for predicting house prices through **Exploratory Data Analysis** and **Feature Engineering**.
- Applied preprocessing techniques, 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: Tableau ([Link](#))

Jan 2024

- Constructed and implemented **KPIs** for a concise overview of car sales performance
- Translated business requirements into actionable insights for a **customized dashboard** with interactive features