Vedant Shirpurkar

Buffalo, NY, 14214 | (716) 787-0704 | [Gmail](#)| [LinkedIn](#) | [GitHub](#)

EDUCATION

University at Buffalo,The State University of New York   Aug 2023 - Dec 2024

Master of Professional Studies, Data Science, **GPA 4.0/4.0**

Coursework: Data Mining, Machine Learning, Data Visualization,Python, Time Series Analysis, Applied Mathematics, Data Analysis in R , Database Management System

TECHNICAL SKILLS

Programming : Python, R Programming, Java, SQL, C++

Packages   : Pandas, NumPy, Matplotlib, Seaborn, Scikit-Learn, transformers, R

Shiny Tools   : Git, AWS,MLFlow, PySpark, Databricks, Apache Airflow, Tableau, AWS Lambda, AWS Glue, Docker, Snowflake, SQL, Power BI, LangChain, Groq, ChromaDB, FAISS, MySQL, Astronomer Airflow

Domain Skills : Machine Learning, Exploratory Data Analysis, Scalable ETL pipelines, Hypothesis testing, ANOVA

Certifications : **Google Advanced Data Analytics Certificate**

PROFESSIONAL EXPERIENCE

Oracle Database Administrator, Cognizant Technology Solutions Pvt. Ltd, India   Dec 2022 - Jul 2023
- Conducted systematic audits and updates on user access controls, strengthening data protection mechanisms by   20% to safeguard sensitive information against unauthorized breaches within the organization.
- Identified root causes of database-related issues, including Failed Backups, and delivered timely resolutions to minimize downtime, resulting in a 15% increase in system uptime.

Major Incident Manager, Cognizant Technology Solutions Pvt. Ltd, India   Sep 2021 - Dec 2022
- Reviewed and analyzed over 100 incidents and issues, identifying root causes and implementing streamlined workflows for resolution, resulting in a 20% decrease in incident resolution time.
- Constructed an advanced ServiceNow Dashboard that seamlessly integrated essential incident data, streamlining the monitoring process for critical issues and contributing to an 8% improvement in urgent response times impacting service delivery.

PROJECTS

Health Insurance Analysis and Classification :
- Predicted insurance purchases using models like Logistic Regression, tree-based models, boosting methods, and ANN,achieving a top F1 score of 0.88 on 380k customer records.
- Engineered features with frequency encoding for high-cardinality categories and resolved class imbalance to boost model performance.

Customer Segmentation and Recommendation System :
- Designed a customer segmentation model using RFM analysis and PCA, applying K-Means, Hierarchical Clustering, and DBSCAN for accurate clustering. Enhanced predictive accuracy with a fine-tuned Random Forest classifier.
- Deployed the model as a web app, enabling real-time, personalized product recommendations that supported more effective targeted marketing strategies.

Real-Time Data Streaming, Transformation and Analytics :
- Developed real-time data pipeline using Apache NiFi on Amazon EC2 (Docker) to ingest data into Amazon S3.
- Implemented Snowflake Snowpipe for continuous data loading from Amazon S3 to Snowflake staging table.
- Engineered automated data movement through the implementation of Snowflake Streams and Tasks, enabling real-time updates to target tables while processing 10,000 rows simultaneously for optimized reporting efficiency.

Conversational RAG System with Chat History :
- Implemented Hugging Face Embeddings in conjunction with Chroma for document retrieval processes; achieved rapid response times averaging under two seconds per query from users accessing the system in high-traffic scenarios.
- Integrated the Groq API and Llama3-70B model into a Streamlit web app, enabling users to securely upload PDFs and ask context-aware questions based on document content for robust, accurate reformulation.

LEADERSHIP AND VOLUNTEER EXPERIENCE
- Led the college club for 8 major college events with a focused team of 20 members.
- Trained and mentored junior team members while addressing concerns, resulting in an impressive increase in student participation and engagement metrics by approximately one-third.