Prem Gladstone Kanaparthi

ADDRESS: 3925 N ELM STREET. #1231, DENTON, TEXAS, 76207
MOBILE: +19408431006
MAIL: premgladstone1@gmail.com
Portfolio: https://prem060.github.io/portfolio/
Linkedin: https://www.linkedin.com/in/prem-gladstone-36593b1a5

PROFESSIONAL SUMMARY:

Detail oriented and results driven computer science professional with over two years of industry experience specializing in data engineering analytics and machine learning system development. Demonstrated expertise in architecting scalable data pipelines transforming raw data into meaningful insights and deploying end to end ML workflows across cloud platforms. Strong foundation in data modeling database design and distributed processing using tools such as Snowflake PostgreSQL Airflow and DBT. Proficient in Python and SQL for data transformation analytics automation and ML integration. Adept at leveraging cloud platforms like AWS for deploying robust scalable systems with CI CD and monitoring frameworks. Committed to solving complex data challenges by simplifying workflows enabling data driven decisions and contributing to enterprise wide data strategy in fast paced FinTech and analytics environments.

TECHNICAL SKILLS:

Programming Languages : Python, SQL, Java, C, C++, JavaScript, React, Bash, R
Data Engineering and Analytics Stack : Snowflake, PostgreSQL, Airflow, DBT, GitLab
ML Frameworks : PyTorch, TensorFlow, Keras, Scikit learn, Hugging Face Transformers
Transformer Models : XLNet, T5, ELECTRA, BERT
Model Optimization Tools : ONNX, TorchScript, quantization, pruning, batch inference
Model Lifecycle and MLOps : MLflow, SageMaker, GitHub Actions, GitLab CI CD, Jenkins
Web and API Frameworks : Flask, Django, FastAPI, RESTful API, Swagger, Postman
Cloud and Infrastructure : AWS, EC2, S3, RDS, Lambda, Kubernetes, Docker, GCP, Redshift
Monitoring and Benchmarking : TensorBoard, Seaborn, Matplotlib, Plotly, Jupyter
Databases : Snowflake, PostgreSQL, MySQL, MongoDB, SQLite, MariaDB, DB2
Softwares and Tools : Microsoft Office, Tableau, Jira, Confluence
Soft Skills : Problem solving, analytical thinking, strategic planning, communication

PROFESSIONAL EXPERIENCE:

Software Developer
L&T Mindtree, India
July 2021 – May 2022

- Built robust data pipelines using SQL and Python on PostgreSQL and Snowflake for real time analytics
- Collaborated with cross functional teams to translate analytical use cases into scalable data solutions

- Implemented Airflow DAGs for automated scheduling of ETL workflows across cloud environments
- Supported CI CD integrations via GitLab to deploy containerized applications and model APIs
- Conducted data profiling and schema design for analytical datasets and feature stores
- Optimized inference logic for transformer based models using ONNX and deployed on AWS Lambda
- Developed monitoring dashboards using Plotly and Seaborn for latency and throughput analysis

INTERNSHIPS AND PROJECTS:

Machine Learning Intern
GNV IT Solutions
Summer 2023

- Developed Snowflake based ELT pipelines transforming large scale financial data using DBT
- Engineered batch prediction workflows using Airflow and Python for automated scoring of classification models
- Tuned transformer models such as ELECTRA and BERT to enable text classification for regulatory tagging
- Integrated MLflow for model version tracking and lifecycle logging with SageMaker endpoints for deployment
- Benchmarked model performance by latency throughput and prediction quality under production load
- Delivered containerized inference APIs on Flask using Docker and AWS Gateway for secure real time scoring

Liver Disease Prediction System

- Used Snowflake and DBT to design and populate curated feature stores, enabling scalable and reliable access to dimensional health data for model development
- Engineered complex data transformation workflows using SQL and DBT to clean, normalize, and structure large clinical datasets
- Built and compared multiple classification models including Random Forest, SVM, and Logistic Regression to predict liver disease with explainable insights
- Automated the end to end ML workflow including model training, parameter tuning, and cross validation benchmarking with performance metrics such as accuracy, F1 score, and ROC AUC
- Integrated model scoring pipeline with monitoring to track changes in input distributions and model performance over time

Fake Profile Detection Tool

- Designed pipeline using DBT and SQL to curate behavioral metrics from raw user activity logs
- Used these features in training fraud detection classifiers with XGBoost and SVM
- Deployed and served models using FastAPI integrated into a React interface

CI CD Enabled RESTful Data API

- Developed modular APIs using Flask to expose ML model endpoints with dynamic routing

- Set up GitHub Actions to automate model training testing containerization and deployment to AWS EC2

Skills & Technologies Used
Machine Learning Libraries:

- Scikit-learn: Model development, training, testing, and evaluation
- Pandas: Data manipulation, transformation, and cleaning
- Matplotlib & Seaborn: Data visualization for model insights and results
- **Machine Learning Algorithms:**
- **Classification:** Logistic Regression, Random Forest, SVM, XGBoost
- **Regression:** Linear Regression, Decision Trees
- **Clustering:** K-Means, DBSCAN
- **Data Preprocessing:**
- Handling missing data
- Feature scaling
- Outlier detection
- Categorical encoding

EDUCATION:

University of North Texas
Master's in Computer Science
GPA: 3.63/4.00
2023 - 2025

GNA University
Bachelor's in Computer Science
GPA: 7.96/10.00
Graduated 2021

CERTIFICATIONS

- **AWS Certified Cloud Practitioner**
- **AWS Certified Developer – Associate**
- **Cybersecurity Fundamentals** – Palo Alto Networks
- **Machine Learning Certificate** – CEMS