SAURABH SAMPATKUMAR LOYA

Salt Lake City, Utah (**Open to Relocation**) | +1 (801)-680-5240 | s.loya21@gmail.com | linkedin/saurabhloya | saurabhloya.github.io

SUMMARY

Data-driven Data Scientist with a Master's in Computer Science and 2+ years of work experience, specializing in Machine Learning, Natural Language Processing, and LLM. Skilled in problem-solving and innovation, with a customer-centric approach to building scalable solutions. Adept at leading cross-functional teams and leveraging data-driven insights for decision-making.

SKILLS AND CERTIFICATIONS

Languages: Python, Java, R, HTML, CSS, JavaScript, SQL

Framework & Libraries: Pandas, NumPy, Matplotlib, Scikit-Learn, Apache Spark, Apache Airflow, Apache Kafka, PyTorch, Tensorflow, Spacy, Flask, Django, Selenium, LangChain

Visualization Tools & Development: Tableau, Power BI, D3.js, Git, Jira, Docker, Microsoft Office, Microsoft Excel

Databases: MySQL, PostgreSQL, MongoDB, ChormaDB, Pinecone

Certification: AWS Certified Cloud Practitioner, Azure Fundamentals

EDUCATION

The University of Utah — Salt Lake City, Utah

Master of Science in Computer Science | **Graduate Teaching Assistant** | **GPA: 3.5/4** — *Aug 2023 – May 2025*

Coursework: Visualization for Data Science, Natural Language Processing (NLP), Deep Learning, Machine Learning

MIT World Peace University — Pune, India

Bachelor of Technology in Computer Science and Engineering | **Merit List Scholarship** | **GPA: 3.8/4** — *Aug 2017 – May 2021*

WORK EXPERIENCE

BMW Financial Services NA — Salt Lake City, Utah

Data Scientist Intern — *Jan 2025 – May 2025*

- Developing **machine learning-driven** loan selection optimization model**,** ensuring the most profitable loans are approved from a pool of booked loans, effectively maximizing profit and reducing risk for daily loan originations of **$100M–$300M**.
- Executing intricate feature engineering on a **high-dimensional dataset** to identify key predictors to optimize loan selection.
- Utilizing **SQL** and **ETL** processes to extract, transform, and load data for comprehensive analysis and reporting.
- Building **Tableau dashboards** to track loan performance, supporting data-driven decisions for a **$9B+ asset portfolio.**

Volkswagen Group Technology Solution — Pune, India

Software Engineer (focused on Data Science) — *Aug 2021 - July 2023*

- Engineered a global survey platform for Volkswagen Group as a scalable **Software as a Service (SaaS)** solution, enabling comprehensive participant engagement monitoring and analysis using **Tableau dashboards.**
- Implemented and optimized an **Optical Character Recognition (OCR)** pipeline using **Azure Cognitive Services** for invoice scanning, achieving an **8x speedup** and **15% accuracy** improvement in extracting invoice details.
- Elevated the intelligence of the chatbot by integrating **Large Language Model (LLM)** and advanced **NLP** technologies, driving to a **30% improvement** in its knowledge base through comprehensive analysis of user interactions and diverse data sources.
- Innovated a task automation solution with **Python, Selenium**, and **Apache Kafka** achieving a **60% reduction** in task completion time, eliminating the need for RPA, leading to significant operational cost savings.
- Resolved **50+ production issues**, including data quality and performance and led the refactoring of a **2 complex projects.**
- Supported the Volkswagen **AI/ML Center of Excellence (COE**) team by collaborating on **3 proofs of concepts (POC).**

PROJECTS

- **CITI Bike Rental Data Analytics and Forecasting**: Conducted **EDA** on 5M+ ride records and developed a **predictive model** using **Apache Spark** and **Facebook Prophet**, optimizing CITI Bike demand forecasts and inventory.
- **Scheduling & Analytics Platform for VCs:** Expanded a meeting scheduling system integrated with analytics, where VCs can view engagement trends and optimize meeting times and resources for maximum efficiency.
- **Android Malware Detection System:** Trained a malware detection model by comparing ML algorithms including **Perceptron**, **Logistic Regression, SVM, and Neural Networks,** achieving a **92% F1** score in classifying malicious apps based on system calls.
- **LLM based Recommendation System:** Crafted a medicine recommendation system using Llama 2, LangChain, and Chroma DB with Retrieval-Augmented Generation (RAG), achieving **90% accuracy** in providing medication.

HONORS AND RECOGNITION

- **Innovation Award** in recognition of creative thinking in automating training pipeline of chatbot using LLM.
- **Achievers Award** for successful project deliverable with Customer Satisfaction Survey rated above 9.