

SAURABH SAMPATKUMAR LOYA

+1 (801)-680-5240 | s.loya21@gmail.com | [linkedin/saurabhloya](https://www.linkedin.com/in/saurabhloya) | saurabhloya.github.io

SKILLS AND CERTIFICATIONS

Languages: Python, Java, C++, JavaScript, TypeScript, SQL

Databases: MySQL, PostgreSQL, MongoDB

Frontend & Backend: Angular, React, Spring, Flask, Django, Apache Kafka, Apache Airflow, Apache Spark

ML/Data: NumPy, Pandas, PyTorch, TensorFlow, Scikit-Learn, Spacy, LangChain

Development & Visualization Tools: Git, Jira, Docker, Tableau, Power BI

Certification: AWS Certified Cloud Practitioner, Azure Fundamentals

EDUCATION

Master in Computer Science, The University of Utah (GPA: 3.6/4) | Salt Lake City, Utah Aug 2023 – May 2025

Bachelor in Computer Science and Engineering, MIT World Peace University (GPA: 3.8/4) | Pune, India Aug 2017 – May 2021

WORK EXPERIENCE

Data Scientist Intern, BMW Bank of NA, Salt Lake City, Utah Jan 2025 - May 2025

- Prepared a high-quality dataset by performing **data mining, data analysis, and cleaning** efforts using **Python and SQL**, enabling the development of a machine learning-driven loan optimization model impacting \$100M–\$300M in daily originations.
- Engineered **40 features from 688 raw attributes** by applying feature engineering techniques to high-dimensional, unstructured data using **Pandas** and **NumPy**, enhancing model accuracy and data usability.
- Built interactive **Tableau** dashboards to visualize loan performance trends, enabling the Credit and Risk team to make informed decisions on a \$9B+ financial portfolio.

Software Development Intern, H7 BioCapital, San Francisco, California June 2024 - Sept 2024

- Crafted a full-stack scheduling application using **Django, HTML/CSS**, and **JavaScript** to streamline investor–startup interactions, reducing email coordination time by ~60%.
- Integrated **Google Calendar & Zoom API** to automate meeting creation, reminders, and calendar between investors and startups.
- Translated business requirements into well-defined technical stories in **Jira**, supporting **agile** development and implementation.

Software Engineer, Volkswagen Group Technology Solution, Pune, India Aug 2021 - July 2023

- Engineered a global **SaaS survey platform**, adopted by 10+ Volkswagen entities to streamline survey creation, distribution, and analytics across departments and geographies.
- Developed a **Query Management System** with **RESTful APIs** to automate HR inquiry routing, reducing average resolution time from **5 days to 3 days** and achieving approximately 40% improvement in turnaround.
- Innovated a task automation solution using **Python, Selenium**, and **Apache Kafka** to replace manual RPA processes, reducing task time and cutting RPA license costs by over **\$15,000 annually**.
- Optimized an **OCR pipeline** with **Azure Cognitive Services**, achieving 8x faster document processing and 15% accuracy gain.
- Integrated **LLM and NLP** into an internal chatbot, enhancing responses quality and expanding its knowledge base by 30%.
- Designed and tested backend services using **JUnit** and **Mockito**, validating business logic and API–database interactions.

Founding Engineer, Venkatesh Agencies, Maharashtra, India May 2020 – May 2021

- Developed **full-stack web applications** streamlining inventory, billing, and customer workflows handling for a medical distributor.
- Managed **MySQL** database schema, indexing, and query optimization, reducing report generation time from 6 min to under 30 sec.
- Analyzed multi-source sales and inventory data, covering **200K+ SKUs** and 50+ pharmacy partners.
- Built interactive dashboards using **Tableau** to monitor real-time sales, forecast shortages, and visualize seasonal demand.

PROJECTS

- Distributed Key-Value Store Using Raft Consensus Algorithm:** Implemented a fault-tolerant, replicated **key-value store** in **Golang** using Raft, ensuring strong consistency, linearizability, and 4s leader failover with log replication and persistence.
- CITI Bike Rental Data Analytics and Forecasting:** Conducted **Exploratory Data Analysis** on 5M+ ride records and trained a predictive model using **Apache Spark** and Facebook Prophet, optimizing CITI Bike demand forecasts and inventory planning.
- LLM based Recommendation System:** Crafted a medicine recommendation system using **Llama 2, LangChain**, and **Chroma DB** with Retrieval-Augmented Generation (**RAG**), achieving an **F1 score of 0.87** in providing relevant medication suggestions.
- Android Malware Detection:** Compared various machine learning models from scratch including Neural Networks, Perceptron, and others, to detect malicious Android applications using system call frequency features, achieving a highest **F1 score of 0.89**.

HONORS AND RECOGNITION

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