SAURABH BHAUSAHEB ZINJAD

📞 +1 480 913 5544 ✉ saurabhzinjad@asu.edu github.com/Ztrimus in linkedin.com/in/saurabhzinjad

EDUCATION

Arizona State University, Tempe, USA   August 2023 - May 2025
Masters of Science in Computer Science (GPA: 4/4)

Relevant Courses: *Social Media Mining, Knowledge Representation and Reasoning Algorithms, Statistical Machine Learning,*

Pune Institute of Computer Technology(PICT), Savitribai Phule Pune University, India   July 2015 - June 2019
Bachelor of Engineering (GPA: 8.53/10)

Relevant Courses: *Data Structure and Algorithms, Object Oriented Programming, Operating System, System Programming, Computer Networks, Information Theory, Artificial Intelligence, Machine learning, Digital Video and Image Processing,*

TECHNICAL SKILLS

Programming Languages	Python, Java, SQL, Unix / Linux Scripting
Cloud Platforms & Databases	GCP, AWS, Big Query, Cassandra, Postgres, DynamoDB, MySQL
Data Engineering	Spark, Kafka, SnowFlake, Airflow, Tableau, ETL, Pandas, D3.js
DevOps / SRE	CI/CD, Git, Jenkins, Docker, Kubernetes, EC2, ECR, S3
Certified Google Cloud Platform	Associate Cloud Engineer.

WORK EXPERIENCE

BigCommerce   Austin, Texas
Data Science Intern   June 2023 - August 2023

· Constructed a logistic regression predictive classifier machine learning model for 3 million opportunities, utilizing **SQL**, **ETL**, and **data modeling techniques**, to predict customer retention post-7-day trial with increased accuracy of 12%.
· Executed insightful **Exploratory Data Analysis (EDA)**, implementing effective data strategies to elevate data quality and lay a robust foundation for predictive modeling.
· Presented findings and new ideas for Hyper-personalization using AI to the director of technology, resulting in enterprise-scale adoption for driving business growth.

Publicis Sapient   Bangalore, India
Data Engineer - II   June 2019 - July 2022

· Engineered complex **ETL** pipelines using **Apache Spark**, optimizing data **extraction, transformation,** and **loading** from diverse sources, such as **Redshift**, **S3**, **Kinesis Streams**, and **Kafka**.
· Utilized distributed computing frameworks such as **Apache Spark** to manage **large-scale data processing** tasks, enhancing performance and resource utilization by 15%.
· Ensured data accuracy, integrity, and consistency through robust data quality checks and automated testing frameworks.
· Designed and maintained **real-time data streaming** solutions with **Apache Spark,** and **GCP Cloud Run** staying current with the latest candidate activities.
· Utilized **AWS Lambda, and SNS** for scheduled data transformations, cutting pipeline run-time by 2 hours weekly.

PROJECTS

Search Engine for All file types - Opportunity Hackathon

· Spearheaded **Elasticsearch** implementation for blazing-fast search responses, with **millisecond response times**.
· Converted and stored every file type data as **vector embeddings**, ensuring **low-latency search** capabilities.
· Led **Python FAST API development**, providing efficient data access and retrieval.
· Used Machine Learning techniques such as BERT, OCR, ResNet50, and Image Captioning to parse Image features.
· **Collaborated** effectively with team members, optimizing task distribution for streamlined project completion.

Scalable Data Processing Pipeline - Neo4J, Docker, Kafka and Minikube

· Designed and implemented a highly scalable and available data processing pipeline using **Kubernetes**, Kafka, Docker, **Neo4j**.
· Orchestrated the setup of Kafka and Apache Zookeeper using Minikube, a lightweight Kubernetes implementation.
· Streamlined data ingestion, and processing into Neo4j, applying PageRank and BFS for graph-based data exploration.

Email Automation Marketing Tool

· Initiated and completed the development of a robust email automation project, streamlining job application outreach and networking efforts with individuals of similar interests. employed REST API integration to acquire and manage a comprehensive contact database from CRM tools, demonstrating prowess in **full-stack development** and **API design**.