AKSHAY VENKATESAN

(934)-451-9623 ◇ aksvenkatesa@cs.stonybrook.edu ◇ linkedin.com/akshayvenkatesan

EDUCATION

Stony Brook University, NY — Aug 2023 - Dec 2024
Master of Science in Computer Science — GPA: 4.0/4.0
***Relevant Coursework** : Distributed Systems, Databases, Machine Learning, Network Security*
***Teaching Assistant** : CSE351: Introduction to Data Science*

SRM Institute of Science and Technology, India — Jun 2016 - May 2020
Bachelor of Technology in Computer Science — GPA: 3.8/4.0
***Relevant Coursework** : Data Structures, Algorithms, Cloud Computing, Operating Systems*

EXPERIENCE

JP Morgan Chase & Co — Bangalore, India
Software Engineer — *Feb 2023 - Jul 2023*

- Led the migration of data infrastructure from ***Hadoop(Hive)*** to ***AWS Cloud (Redshift and Athena)***, optimizing data storage and resulting in a ***45% cost reduction*** and markedly improving model training time.
- Spearheaded ETL processes using ***Alteryx*** and architected data models to integrate API calls from LinkedIn Marketing, Marketo, and Google Marketing, increasing marketing efficacy through a **20% uplift in conversion**.
- Facilitated time critical business transaction and service data analysis using ***Airflow and Kafka***.
- Executed big data transformations with ***Spark*** and conducted textual data preprocessing using ***NLTK and spaCy*** for data science model readiness, handling up to 14,000 words and 300 sequences daily.
- Designed and deployed robust ML pipelines using ***AWS SageMaker and Lambda***, specifically for forecasting service request volumes, efficiently processing over **10 million records** and enhancing predictive accuracy.

Tata Consultancy Services — Bangalore, India
Software Engineer — *Nov 2020 - Feb 2023*

- Developed ***PySpark*** *and* ***Scala*** solutions on Azure Databricks, processing ***over 100 million records daily***.
- Optimized system level file integrity checks using ***UNIX shell scripts***, enhancing the effectiveness by 2X.
- Defined efficient ***T-SQL*** procedures for handling automated vendor payment settlement processing.
- Managed and transformed data across various formats (***JSON, CSV, Parquet, XML***) for diverse applications.
- Crafted log analytics pipeline handling 10,000 files daily, incorporating ***Flask API*** calls to automate data workflows and integrate with a central log monitoring application for streamlined operations.

PROJECTS

- **Adaptive Distributed System** | *Python, C++, Reinforcement Learning, Docker* :
  Leveraged ***multi-armed bandit*** Reinforcement Learning to develop a decentralized distributed transaction processing system. This system dynamically selects the best ***BFT*** protocol and data sharding stucture.
- **Key-Value Store** | *C++, RPC*:
  Designed and implemented a highly available, fault-tolerant distributed key-value storage system using the ***RAFT*** consensus protocol in C++, performing cross shard transactions using 2PC with concurrency control.
- **Jump Proxy for Enhanced TCP Service Security** | *Go, Crypto Library*:
  Developed a jump proxy system to add an encryption layer for ***TCP*** services, significantly improving security by using symmetric key decryption for traffic. Ensured robust protection against potential pre-auth vulnerabilities.
- **Energy Monitoring Platform** | *Node.js, React, Python, AWS EMR, SQL*:
  Deployed a real-time full stack **IoT-based energy monitoring system** utilizing Random Forest for forecasting, alongside AWS EMR and S3 for data processing, achieving a 15% reduction in monthly energy consumption.

TECHNICAL SKILLS

Languages: Python, C++, SQL, Scala, Java, Go, Shell, R Programming, Javascript
Technologies: Git, Docker, Spark, ETL, Hadoop, Kafka, MongoDB, Redis, Snowflake, Power BI, NLP, PostgreSQL
ML Libraries: pandas, NumPy, scikit-learn, plotly, PyTorch, TensorFlow, Keras, LDA, NLTK, spaCy, Prophet
Platforms: Databricks, Azure(Synapse, Data Lake, Data Factory), AWS (S3, EMR, EC2, Lambda, Redshift)