Alekhya Madanu

contact@alekhya.dev |linkedin.com/in/alekhya-madanu | +91 9121002980

Education

Indian Institute of Technology (IIT), Hyderabad

B.Tech in Computer Science and Engineering Science | Aug 2017 - Apr 2021

Courses: Data Structures | Algorithms | DBMS | Computer Networks | Wireless Networks & Security | OS | Algorithmic Techniques for Massive Data | Intro to Image Processing | Representation Learning | Data Mining | Foundations of Machine Learning | Fraud Analytics

Professional Experience

GOLDMAN SACHS | Hyderabad

Associate - Search Engineering Team, ML Division | Jan 2024– Present

Worked on our GenAI platform which provides search services on internal and public domain financial data in a conversational setting. I was tasked with building resiliency and configurability into our platform across all services and as part of this I have -

- Built a custom gRPC ratelimit service that fits into an **EnvoyProxy** service mesh. This service employs throttling based on customizable user configs and service agreements, enhancing the platform's robustness.
- Implemented a service discovery based blue green deployment technique which involved building an automated service discovery plane and dynamic routing through Envoy. This reduced our downtime and optimized release processes.
- Streamlined logging infrastructure by utilizing **BigQuery** and **Fluentd**. This initiative significantly improved platform observability and diagnostic capabilities.

Analyst - Search Engineering Team, ML Division | July 2021 – Jan 2024

Worked on GS's internal search engine, which aims to make multiple formats of structured and unstructured data searchable - such as e-communications, trades, tickets, news and more based on the requirements of internal clients.

- **Spark Trade Pipeline (Spark, Hive, Presto)**: Pioneered team's first Spark pipeline for trade processing, optimizing Hive table ingestion. Slashed processing time by nearly **50%**, **doubled** data throughput through custom logic to insert data into bucketed and partitioned Hive tables.
- **Search Result Refinement (MapReduce, Clustering)**: Devised MapReduce job for precise result pages via **document clustering**, optimized for **petabyte** scale data. Enhanced user experience by cleaning up broken links and irrelevant content worth almost **1 million** links.
- **Voice Pipeline Resilience (Flink, Kafka)**: Strengthened voice pipeline with Flink data streams, Kafka integration. Increased success rate by **15%**. Implemented retry strategies, **Monadic Exception Handling** for robust performance.
- **Near Real-Time Ingestion (Java, Spring Boot)**: Engineered real-time data pipeline with multi-threaded Java programs. Seamlessly integrated multiple data sources like Confluence and Elasticsearch using **REST APIs**.

HONEYWELL | Bengaluru

Data Science Intern | May 2020 – July 2020

- The project was based on **NLP** and **Information Retrieval** techniques. I tested out topic modeling and cluster analysis to group a database of jira error reports, using various word embeddings. Based on observations, built a predictive model to suggest viable solutions for a given error.
- Worked with **ML** frameworks such as **Pandas, Numpy, Scikit-learn, NLTK** and **Keras** to build the model.

Projects

CONTINUAL LEARNING FOR NETWORK DATA | Guide: Dr.Bheemarjuna Reddy. Final year research project where we implemented different **Continual Learning** algorithms on network flow data. Most of the papers published in continual learning are focused on **Computer Vision** and we tuned these algorithms to fit into our use case of network data. One notable algorithm implemented was **Class-Balancing Reservoir Sampling**.

ML BASED NETWORK TRAFFIC CLASSIFIER | Built a **Deep Neural Network** to classify network traffic based on flow statistics from wireshark, which can be used to prioritize bandwidth allocation and improve performance of applications. Explored **Reinforcement Learning** strategies using **Markov Decision Process** as a mathematical framework.

Skills

Languages: Java | Python | C++ | C | Shell | Javascript

Tech Stack: MapReduce | Spark | Flink | Kafka | Spring Boot | Hbase | Hive | Presto | HDFS | React.js | Vue.js