

Sai Manoj Jalam

☎* 352-709-2440

✉* jalam.saimanoj@ufl.edu

🌐* [linkedin](#)

🐙* [github](#)

Education

University of Florida

Master of Science in Computer Science: 3.83/4.0

August 2022 – May 2024

Gainesville, Florida, United States

Indian Institute of Technology(Indian School of Mines)

Bachelors of Technology in Computer Science: 7.8/10

June 2014 – May 2018

Dhanbad, Jharkhand, India

Technical Skills

Languages: Java, JavaScript, Python, C/C++, Swift, Scala, HTML, CSS

Big Data: Apache Hadoop, Hive, HBase, Spark, MapReduce, Apache Kafka, Airflow, Sqoop

Machine Learning/Data Science: Pandas, Numpy, Pytorch, XGBoost, Scikit-Learn

Cloud: Amazon (EC2, S3, CloudWatch, Lambda, ELB, DynamoDB, Cognito) , Azure Cloud Services

Tools, Frameworks and Databases: React.js, Swift, MySQL, PostgreSQL, Git, Docker, Kubernetes, Unix, CI/CD

Experience

Software Engineer/Data Scientist 📄 [PPT link] 📄 [Design link]

Oct 2022 – Present

University of Florida

Gainesville, Florida, United States

- Designed and developed an integrated ROAMM platform using **Swift, React**, and **AWS** to offer real-time health insights.
- Built a resilient **serverless backend** infrastructure on AWS using **EC2,ELB,API gateway** ,etc ensuring low latency and highly scalable performance for **15+ API's** and **10+ microservices** in Python and NodeJS.
- Implemented secure APIs leveraging **AWS Cognito** for authentication and **JWT for authorization**, facilitating real-time communication with the smartwatch app.
- Designed and deployed Database for **Optimized data storage** and retrieval mechanisms using **AWS RDS** and **DynamoDB**.
- Established robust monitoring and alerting mechanisms using AWS tools(**Cloudwatch** ,**SNS** , etc) to achieve and maintain **99.9% uptime** for the backend services on EC2.
- Built **data pipeline** to transfer data from databases(**RDS,DynamoDB**) to datalake (**AWS redshift**) and built data pipeline for complete **ETL** flow for **realtime analytics** .
- Managed and processed vast datasets to **analyze KPIs** from data collected and stored using **Pyspark,Numpy,Pandas,etc** and prepared reports.
- Deployed **CNNs and RNNs** on AWS, using **SageMaker** and **Lambda**, to analyze real-time sensor data for **predicting health events** like **fall and activity detection**.

Associate Software Engineer (Data Platform)

July 2021 – Aug 2022

Optum(United Health Group)

Hyderabad, India

- Designed and developed highly scalable claims **KPI data pipelines**, leveraging technologies such as **RDBMS, Big Data, and Cloud**, to streamline the collection, storage, modeling, and analysis of massive datasets from heterogeneous systems.
- Automated data import and export processes from **MySQL to HDFS using Sqoop jobs**, significantly reducing manual effort.
- Engineered real-time distributed **ETL data pipelines** to optimize claim data processing, resulting in a 50% reduction in processing time. Technologies utilized include **HDFS, Hive, PySpark, Kafka, and Airflow**.
- Developed an efficient injection pipeline to transfer data from **HDFS to Snowflake**, which dramatically improved dashboard response times for claims KPI application from 45 seconds to approximately 15 seconds.

Projects

Brain tumor Detection | Python, Pytorch, Image Processing

Dec 2022

- Developed a 95% accurate ML system for brain tumor recognition using custom CNNs, Pytorch, and transfer learning with pre-trained models(VGG-16,ResNet-50).

Stock market real time data analytics | Python, AWS, Kafka, Glue, Athena

February 2023

- Developed a Python-based Apache Kafka pipeline for real-time stock market data, utilizing AWS Glue, Athena, and SQL for data handling, and integrated with AWS S3 and RDS via Kafka Connect.

Spatial Hotspot Analysis using Apache Spark | Java, JavaFX Apache Spark, Hadoop, EC2, Cloud Watch

March 2022

- Conducted Spatial Hotspot Analysis on NYC yellow taxi trip records using Apache Spark and Hadoop File System in AWS EC2 cluster.

Spotify ETL pipeline | Python, Airflow, PostgreSQL

March 2023

- Used Spotify's API to read the data and perform some basic transformations and Data Quality checks finally loaded the retrieved data to PostgreSQL DB and then automated the entire process through airflow.

End-To-End From Web Scraping to Tableau | Python, Tableau, PostgreSQL

July 2023

- Executed end-to-end data projects, extracting web data using Python (Beautiful Soup, Scrapy). Cleaned and transformed data with Pandas, stored in PostgreSQL. Developed and shared insightful Tableau dashboards, emphasizing data-driven decision-making.