

TEJAS VEDAGIRI

Data Engineer

Led the successful completion of complex software projects, utilizing my expertise in Apache Spark, resulting in a 70% cost reduction and saving over 100 hours per week. Currently pursuing a master's at Florida State University. My career spans multiple programming languages, Databases and tools, specializing in transforming proof of concepts into enterprise software. Actively engage in machine learning and deep learning projects within the Big Data ecosystem.

tejasvedagiri1223@gmail.com 

(448)-200-6054 

Tallahassee, United States 

tejasvedagiri.com/ 

linkedin.com/in/tejas-vedagiri-66ab07174 

github.com/Tejasvedagiri 



SKILLS

Java Sacal Spark Python Hadoop HDFS Hive Cassandra CarbonData BigQuery Phoenix
HBase Solr MySQL MSSql Spark ML Machine Learning Deep Learning Air Flow Cloud Fusion
ADF Data Bricks Atlas PyTorch Tensorflow SpringBoot FastAPI Grafana Delta-Lake
Cloud Spanner



WORK EXPERIENCE

Student Worker

Center for Ocean-Atmospheric Prediction Studies

09/2022 - Present

Tallahassee, FL

Collaborated with JPL NASA to seamlessly integrate OIIP Viewer into COAPS ecosystem, enhancing data visualization capabilities and enabling advanced analysis, optimizing research outcomes.

Achievements/Tasks

- Revamped **OIIP Data Viewer** with **Podman Compose**, enabling researchers to deploy global data point visualization services. Making it a 0-click deployment and establishing a scalable, production-grade solution.
- Spearheaded the fine-tuning of **GeoServer**, **Solr**, **Postgres**, and the **OIIP Data Viewer**, resulting in a 30% increase in system performance and enriched user experience.
- Guided the transition to harvesting data from **ERDDAP** as the primary data source using **Python** and enacted **PostgreSQL** for storage, remarkably improving data handling capabilities for research projects and leading to refined accessibility and user-friendliness.
- Proficiently craft and conceptualize webpages within the **Joomla** (CRM) environment, skillfully implementing innovative features and functionalities.
- Enforced rigorous quality control measures to maintain web page integrity, resulting in a 80% decrease in broken links and a 20% increase in user satisfaction.

Contact : Michael McDonald - michael.mcdonald@fsu.edu



WORK EXPERIENCE

Data Engineer

Walmart Global Tech India

10/2020 - 08/2022

Bangalore, India

Spearheaded innovative data engineering initiatives at Walmart, collaborating closely with the Finance team to drive analytics, planning, and budgeting innovation; achieved a 70% decrease in cloud cost and saved 100+ hours per month through automation.

Achievements/Tasks

- Led the architectural design and successful implementation of 'Ensure', a data quality and reconciliation framework powered by **Apache Spark 3.2.1** with the capability to read from any **Big Data** source, ensuring 100% accurate and auditable data, leading to improved decision-making and enhanced customer satisfaction.
- Erected and carried through custom connectors on the **Apache Spark** platform, enabling seamless data consumption via DAX queries for **PowerBI**; revolutionized data insights for the organization, reducing data processing time by 70%.
- Successfully optimized both Code, Cloud infrastructure and **Big Data** storage, resulting in a remarkable 70% reduction in resource utilization costs. This not only streamlined operations but also significantly contributed to cost savings.
- Conducted a comprehensive evaluation of 5+ Data Engineering orchestration tools, including **ADF**, **Cloud Fusion**, and Internal solutions. This assessment aimed to identify the best-fit solution considering development efficiency, project timelines, and resource costs.
- Optimized **Google BigQuery SQL** queries and **Apache Spark** in Memory operations, resulting in 30% to 40% performance improvements. This optimization query response times effectively reduced the cost associated with data read operations, demonstrating a keen focus on efficiency and cost-effectiveness.
- Efficiently optimized and simplified operational workflows for corporate users, focusing on onboarding and access control procedures.

Contact : Sridhar Swarna - Sridhar.Swarna@walmart.com

Software Engineer

Software Engineer at Temenos India PVT LTD

04/2018 - 10/2020

Bangalore, India

Joined Htrunk, a startup later acquired by Temenos, into a dominant force in the banking software industry Developed and applied data-driven strategies that propelled client growth by 10x and positioned the company as a leader in innovation.

Achievements/Tasks

- Leveraged **Spark Kafka** to establish a robust data ingestion framework, enabling real-time data processing and seamless integration with Spark ML, resulting in an 80% reduction in latency and improved predictive capabilities.
- Architected and designed a seamless UI-based pipeline builder, leveraging **Apache Spark** and **JS**, to integrate Spark ML into the **Temenos Data Lake** Ecosystem.
- Enabled real-time loan predictions and empowered users with up-to-date insights. Engineered a robust text search feature using **Apache Solr**, integrating **Apache Spark** for data ingestion and UI development; accelerated text search performance by 40% and facilitated seamless user interaction.
- Streamlined metadata management and data lineage processes by putting into effect the use of the **Apache Atlas** framework, Optimizing data governance for a team of 200+ stakeholders.
- Created a framework to collect custom metrics from **Apache Spark** and executed a highly scalable monitoring module using the Executor framework, resulting in a 90% increase in job tracking and success rate monitoring.
- Formulated and enacted a customized Excel sheet processor utilizing **Python** to automate data updates, streamlining report distribution to clients via email; achieved a 85% reduction in processing time and enhanced data accuracy for timely decision-making.
- Utilized advanced technology, including **Spacy** and **TensorFlow Seq2Seq**, to seamlessly integrate natural language processing with **SQL** query generation.
- Crafted and applied a Single-Context Multiuser API service for **Apache Spark**, enhancing collaborative data processing capabilities. Reducing cost by 70%
- Constructed a tailored data consumption solution within **Apache Spark's** structured streaming framework, enabling efficient data ingestion through a custom sink.

Contact : Mallesha Mahadeva - mallesha.mahadeva@coindcx.com



ACHIEVEMENTS

Hackathon Victory at Walmart

Emerged as the victor in the Walmart Hackathon, outperforming more than 500 rival teams; demonstrated exceptional analytical ability and technical competence, earning recognition for exceptional leadership and achievement

Successful ETL Product Training

Efficiently delivered ETL product training to GM Bank Philippines in 2019.

Top 10 Placement in Temenos Hackathon

Ranked within the top 10 teams out of 200 participants in a highly competitive Temenos-hosted hackathon, demonstrating exceptional investigative and creative thinking abilities that led to industry recognition.

Designated as the Go-To Technical Expert for Data Lake and Spark

Consistently designated as the technical expert for Data Lake and Spark, serving the majority of Temenos clients.



PERSONAL PROJECTS

Stock Manager

- Planned and built a comprehensive enterprise application leveraging cutting-edge technologies, including **ReactJS**, **FastAPI**, **Kafka**, **MySQL**, **SpringBoot**, and **Apache Spark**. Key accomplishments include:
- Formulated and launched a **ReactJS**-powered dashboard presenting real-time portfolio value, dividend growth rate, recommended purchase amounts, and monthly dividends.
- Devised a **FastAPI**-based data ingestion system, executing data retrieval from Yahoo Finance through CRON scheduling and publishing it to **Kafka** for processing.
- Used **Apache Spark** for data processing by reading data from **Kafka** and seamlessly updating the **MySQL** database.
- Effectively employed **SpringBoot** with **OAuth2** for secure authentication and performed the necessary calculations to provide a user-friendly interface with essential financial metrics.

Loan Prediction with Apache Spark Machine Learning

- Developed and implemented **Apache Spark Machine Learning** models to predict loan approvals, resulting in a 20% increase in loan processing efficiency.
- Applied **Apache Spark** connectors to seamlessly integrate data from various databases (e.g., **MySQL**, **PostgreSQL**) into the **Apache Spark** environment, facilitating real-time data analysis.

Real-time Data Processing with Kafka and Apache Spark

- Employed **Kafka** and **Apache Spark** to ingest and process real-time data streams, enabling the team to make timely loan predictions.
- Constructed and executed custom sources and sinks for **Apache Spark** to address specific data ingestion and output requirements, enhancing data pipeline efficiency.
- Wrote custom source and sinks for **Apache Spark**.

Spark Metrics Monitoring

- Set up **Apache Spark** metrics monitoring using **Grafana**, **Graphite**, and **Carbon Cache** to provide real-time insights into **Apache Spark** job performance and resource usage.
- AWS-like **Apache spark** cluster management and deployment.
- Apache Spark** Session Management for performance optimization.

Automated/Self-Driving Car BOT (CNN & OpenAI)

- Masterminded and fabricated a self-driving car bot using **Convolutional Neural Networks (CNN)** within the GTA 5 gaming environment.
- Executed reinforcement learning algorithms using the **OpenAI GYM** framework.
- Effectuated a Doom Bot using **Deep-Q** Neural Networks for advanced gaming AI.

Face Recognition and Object Detection

- Designed a face recognition system using **Convolutional Neural Networks (CNN)**.
- Crafted an object detection system using the **TensorFlow Google API**.

Explainable AI for Loan Defaulters (Oracle Skater with TensorFlow)

- Built an explainable AI system for predicting loan defaults using **Oracle Skater** and **TensorFlow**.

Super-Resolution GANs for Enhancing Smartphone Camera Rolls (Master Project)

- Completed a master's project focused on implementing **SRGANs** to enhance photos taken with smartphone cameras.



SOFT SKILLS

Technical

Communication

Critical thinking

Multitasking

Teamwork

Creativity

Leadership

Time management

Listening

Public speaking



TECHNICAL SKILLS

Platforms	Linux (Centos, Ubuntu), Windows.	Languages	Java1.8 (proficient) > Scala 2.12 > Python > JS > Shell > C++ (intermediate)
Tools Used	Eclipse, Sonar Lint, Maven, Android studio, Spyder, Jupyter-Server/Notebook, and DataBricks	Databases	Hive, HBase + Phoenix, Cassandra, Mongo DB, RDBMS stack, Janus, Delta-Lake, BigQuery, Synapse, CloudSpanner.
Code Repository	Github, Azure Repos, BitBucker, SVN	Development	Agile Development (Kanban)



INTERESTS

Formula One

Automation

Techonology

Dota2

Deep Learning

asd



EDUCATION

Master's in computer science

Florida State University

08/2022 - Present

Tallahassee, FL - CGPA -3.77

Bachelor of Engineering in Computer Science

NITTE Meenakshi Institute of Technology

08/2014 - 06/2018

Bangalore, India - CGPA - 8.16