

SURYA VISWANATH

+91 8608919869 (WhatsApp Available) | +1(316) 226-5276 | suryaviswanath1998@gmail.com | [LinkedIn](#) | Open to relocation

PROFESSIONAL SUMMARY

Data Engineer with experience in **Building Real-Time Data Pipelines**, **Developing Scalable ML Solutions**, and **Implementing Generative AI Applications** using **AWS Services** and **Open-Source LLMs**. Proven ability to **Deploy ETL Workflows**, **Manage Distributed Systems**, and **Deliver Business Insights** through **Interactive Dashboards**.

EXPERIENCE

Data Engineer | Sunlight Technologies (Client: Cogwise) | Suwanee, GA, USA | June 2023 – Present

- **Designed and deployed scalable, real-time data pipelines** using AWS services including **Kinesis, Kafka, Lambda, S3, Glue, Redshift**, and **RDS**, processing millions of events per day for analytics and alerting.
- Developed and optimized **ETL workflows with PySpark and AWS Glue**, transforming raw structured/semi-structured data for downstream consumption in **Snowflake** and **Redshift**.
- Built **real-time analytics dashboards** by indexing processed data into **OpenSearch**, and monitored pipeline performance using **CloudWatch** and **EventBridge**.
- Utilized **SQL (Snowflake SQL and Redshift SQL)** extensively to perform data transformations, join large datasets, implement business logic, and validate ingestion accuracy.
- Developed **Generative AI tools** using **AWS Bedrock (Claude, Titan)** and **LLMs (LLaMA)** with **prompt engineering**, supporting internal use cases like document summarization and query resolution.
- Implemented **CI/CD pipelines** using **GitHub** for data pipeline deployment, with schema validation & data integrity tests.
- Built **classification models** to automate entity tagging in telemetry datasets, reducing processing time and manual review effort.
- Participated in **Agile Scrum teams**, contributing to sprint planning, story estimation, and regular stand-ups.

Business Data Analyst | Nagarjuna Fertilizers and Chemicals Limited | India | June 2019 – May 2020

- Extracted, cleaned, and modeled datasets from **SAP and SQL-based systems**, working with **PostgreSQL** and **MySQL** to access and prepare data from financial, sales, and logistics modules.
- Designed **interactive Tableau dashboards** for C-level leadership, enabling real-time visibility into KPIs like revenue, product demand, and regional distribution.
- Developed **forecasting models** using **regression and time series techniques (ARIMA)** to optimize production and inventory planning cycles.
- Wrote complex **SQL queries** to support BI dashboards, perform **Trend Analysis**, and generate ad-hoc reports, improving data accuracy and analyst productivity.
- Automated **data preprocessing workflows** with Python (pandas, NumPy) for ML pipeline readiness, enhancing data quality and consistency.
- Collaborated with business stakeholders to identify key data metrics and implemented **data validation checks** to ensure accuracy in cross-functional reports.

PROJECTS

BCG GenAI Job Simulation on Forage– June 2025

CERTIFICATE

- Completed a job simulation focused on developing an AI-powered financial **Chatbot** for **BCG's GenAI Consulting team**.
- Extracted and analyzed financial data from **SEC 10-K filings** for **Microsoft, Tesla, and Apple**.
- Built a rule-based chatbot using **Python** that responds to predefined **financial** queries using structured JSON/CSV datasets.
- Applied data preprocessing, financial trend analysis (using pandas), and simplified financial reporting for non-technical users.

British Airways Data Science Virtual Experience Program – June 2025

CERTIFICATE

- Modeled lounge access demand at Heathrow Terminal 3 by grouping flights and estimating tier-wise eligibility across customer segments.
- Designed a reusable lounge eligibility lookup table and documented assumptions for future planning scenarios.
- Developed a **Random Forest Predictive Model** in Python to forecast customer holiday bookings using a booking dataset.
- Evaluated model performance with cross-validation and feature importance analysis, summarized findings in a presentation.

Tata Group Data Analytics Job Simulation on Forage - June 2025

CERTIFICATE

- Completed a job simulation with the **Financial Services team** at **Tata iQ**, focusing on **AI-powered data analytics and strategic decision-making**.
- Performed **exploratory data analysis (EDA)** using GenAI tools to evaluate data quality, detect early warning signs of **customer delinquency**, and structure insights for downstream predictive modeling.

- Proposed a **no-code predictive modeling framework** to assess risk, leveraging **Generative AI** to auto-generate model logic, variables, and interpretability guidelines for business users.
- Designed an **AI-driven collections strategy** using **agentic AI**, with a focus on **ethical AI governance**, **regulatory compliance**, and scalable deployment using automated decision trees and action pipelines.
- Demonstrated the use of GenAI in a **financial services context**, balancing technical feasibility with operational scalability and responsible AI practices.

Real-Time Vehicle Telemetry Processing System | May 2024

- Engineered a **real-time data pipeline** using **AWS Kinesis** to ingest vehicle telemetry streams at scale.
- Integrated **Apache Kafka** for resilient data distribution across microservices and downstream consumers.
- Delivered a production-ready system that **enabled live fleet monitoring** and **predictive maintenance** insights via Snowflake dashboards.

GenAI ChatBot for PDFs Repository | Feb 2024

- Built a **PDF-interactive chatbot** using **GPT-3.5-turbo** to extract answers from user-uploaded documents.
- Used **PyPDF2** for file parsing, and **Streamlit** for a clean front-end experience.
- Designed a user-friendly interface allowing question-based interaction with technical and non-technical PDFs.
- Applied **LLM prompt engineering** to deliver coherent and context-aware responses.

CAPSTONE - "Risk Reduction for Meritrust" | May 2023

- Led a **team of 5** in designing ML-based solutions to reduce financial risks in loan approvals.
- Achieved **70% accuracy** using ensemble models like **XGBoost** and **Random Forest**, trained on anonymized bank data.
- Collaborated with mentors to **translate business goals into ML models** and present insights to industry stakeholders.

Recognizing Bird Species by Audio Using Data Augmentation Techniques | May 2022

- Developed a **bird audio recognition model** using **augmented spectrogram data** from the **Xeno-Canto dataset**.
- Applied **transfer learning** with pre-trained models like **DenseNet201** and **VGG16**, tuned via Keras.
- Proposed an innovative augmentation pipeline to improve generalization across rare bird species.

TECHNICAL SKILLS

- **Programming:** Python, SQL, PostgreSQL, Spark, Sci-Kit Learn, Tensorflow, Keras, Numpy, Pandas, Matplotlib, Seaborn, XGBoost, Django
- **Databases:** MySQL, AWS Redshift-S3, Snowflake, SAP S4 Hana
- **Visualization Tools:** Tableau, Power BI, Advanced Excel, AWS Quicksight, SPSS, Microsoft Office, Streamlit
- **Machine Learning:** Classification, Regression, Clustering, Time Series Analysis, Neural Networks, CNN, RNN, NLP
- **Cloud Platforms:** Amazon Web Services (AWS), Microsoft Azure
- **ETL Tools:** Matillion, Fivetran, AWS Glue

EDUCATION

M.S. in Business Analytics with Specialization in Data Science | Wichita State University | Wichita

- **Coursework:** *Machine Learning and Deep Learning, Data Mining, Statistics, Time Series Forecasting, Data Visualization, ERP, Cloud Computing*

B.Tech in Computer Science | Sathyabama Institute of Science and Technology | Chennai

- **Coursework:** *Data Structures and algorithms, Operating Systems, Database Management, Object Oriented Analysis and design, Cloud Computing, DBMS*

PUBLICATION

Artificial Intelligent Approach for Prediction Analysis of Engineering Fault Detection | March 2020

- (Ref: <http://annalsofrscb.ro/index.php/journal/article/view/2429/2036>)