

Sukumar Govindaraj

Arlington, TX | +1 (682) 405-4122 | govindaraj.sukumar@outlook.com | [GitHub](#) | [LinkedIn](#)

PROFESSIONAL SUMMARY

Software Engineer with 4 years' experience in data engineering and distributed systems. Skilled in building secure, scalable pipelines and APIs using Python, React, and Spark across AWS and Azure. Strong focus on software engineering, data modeling, audit-compliant design, and production-grade MLOps.

SKILLS

Technical Skills: Anomaly Detection System, Recommendation System, Predictive Modeling, Software Design Patterns, Database Schema Design
Programming Languages: Python, TypeScript, JavaScript, SQL, HTML/CSS, C++, Shell scripting
Frameworks and Tools: NumPy, PyTorch, PySpark, Scikit-learn, MLflow, DVC, Node.js, Pinecone, SSMS, React, Nodejs, Flask, PyCharm, VS Code, Linux
Cloud Platforms: AWS (EC2, SageMaker, S3, Lambda, Redshift, CloudWatch), Azure (ML, Synapse, Blob, Monitor)
Professional Skills: Flexible, Attention to Detail, Committed, Communication, Leadership, Problem-solving skills, Innovation, Critical Thinking

PROFESSIONAL EXPERIENCE

Shimadzu Research Center, UT Arlington

11/ 2023 – Present

Machine Learning Research Assistant | Anomaly Detection System

- **Monitoring** - Collaborated with research team in building a system to monitor lab equipment's health in real time using machine learning to help the team to detect issues early and reduce unexpected tool failures.
- **MLOps Automation** - Automated model training and updating to ensure predictions stay accurate without manual checks.
- **Visualization and Communication** - Created easy to understand dashboards and alert systems to support lab staff and shared weekly reports with the leadership, decision-making and reduced unplanned equipment downtime by **over 20%** through early anomaly detection.
- **Skills:** Predictive modelling, Experimentation, PyTorch, MLflow, Azure DevOps, PostgreSQL, Docker, CI/CD, Microsoft Excel, Power Point.

Victoria's Secret & Co, India

04/ 2022 – 07/ 2023

Data Engineer | Supply Chain Management - Inventory Forecasting

- **Requirement Gathering** - Collaborated with a planning team to organize sales and inventory data to help predict inventory in stores and warehouses.
- **ETL Pipelining** - Created scheduled workflows that kept data up to date by fetching, cleaning and delivering accurate data to reporting dashboards, helping teams make faster and better decisions.
- **Collaboration** - Worked closely with business users to improve report quality and reduced report delivery time by **70%**.
- **Skills:** ETL, Azure Data Factory, Synapse, Delta Lake, Git, Jira, Ad hoc, Delta Batching, data analytics, Agile.

Web Wings Global Solutions, India

01/ 2021 – 03/ 2022

Data Engineer | ETL & Automation for User Analytics

- **ETL** - Designed pipelines to move and organize large data for reporting and predictive analytics and improved the speed and quality of business analysis.
- **Data Modeling** - Improved and organized structured data layers to make dashboards faster, more reliable, and easier for business users to interpret.
- **Automation** - Automated daily tasks like data refreshes and cleanups to cut manual workload by **80%** across multiple workflows.
- **Skills:** ETL, Terraform, Azure Databricks, PySpark, SSIS, Version Control, Docker, Azure, KPI's, Relational Database, Quantitative analysis.

PROJECTS

Healthcare – Revenue Cycle Management

02/ 2025 – 03/ 2025

- **ETL** – Developed a revenue management pipeline to process billing and insurance data from hospitals, including patient claims and diagnostic codes.
- **Validation** - Added checks to catch unusual billing patterns and payment delays, helping healthcare teams make quicker decisions.
- **Skills:** AWS Glue, Redshift, PySpark.

Coffee Shop Chat Application-GenAI

08/ 2024 – 11/ 2024

- **Product Design** – Built a conversational AI assistant to help customers place orders, ask questions, and get recommendations in a coffee shop app.
- **User Experience** – Enhanced user interaction by guiding customers through a step-by-step ordering process with helpful responses and suggestion.
- **Skills:** React Native, Python, NLP, LLMs, RAG, Pinecone, Firebase, Hugging Face, Prompt Engineering, Vector Databases, ML lifecycle automation.

Credit Risk Modeling – Probability of Default & Loss Given Default Estimation

03/ 2024 – 04/ 2024

- **ML** - Developed classification models to predict Probability of Default (PD) and Loss Given Default (LGD) using credit bureau and financial history data.
- **LLMs** - Engineered features and handled categorical variables; applied model selection and parameter tuning for improved R^2 and error reduction.
- **Skills:** Financial Analysis, Data Visualization, Logistic Regression (for PD), Linear Regression / Decision Trees (for LGD), Quantitative Modeling, Testing.

EDUCATION

- Masters of Science in Applied Statistics and Data Science
- Bachelors of Engineering in Computer Science

University of Texas at Arlington, TX — GPA: 4.0/4.0
Sahyadri College of Engineering, India — GPA: 3.5/4.0