Sri Hari Sivashanmugam

Chicago, IL, USA

sivashanmugamsrihari@gmail.com | +1-312-478-9332 | [LinkedIn](#) | [Portfolio](#) | [GitHub](#)

PROFESSIONAL EXPERIENCE

Data Scientist

Chicago Department of Public Health   *April 2024 – Present | Chicago, IL*
- As part of the Data Modernization Initiative, developed a longitudinal data model for 3.2 M+ citizens, standardizing medical records from diverse sources & formats (HL7, FHIR, CCDA), enabling cross-agency data interoperability.
- Designed and implemented a scalable data lakehouse using Medallion Architecture in Azure.
- Engineered Spark-based ETL jobs to seamlessly integrate & process large volumes of data, reducing ingestion time by 28%.
- Built machine learning models with ~97% accuracy using SDOH data, supporting outbreak predictions & resource planning.
- Designed and deployed 40+ Tableau dashboards to monitor community health metrics, served as a data steward.
- Established Active Directory-based data governance, developed data catalogs and data inventories for transparency.
- Enhanced cross-functional data flows using Salesforce, Rhapsody, and Azure tools, improving operational efficiency.
- Collaborated with the Office of Emergency Management and Communications (OEMC) to enhance CHI311 reporting capabilities, improving city agency responsiveness and service delivery to citizens.
- Streamlined reporting workflows by integrating Azure Synapse Analytics, enabling faster data querying & actionable insights.
- Conducted advanced geospatial analyses identifying areas requiring targeted interventions, leveraging Python and Tableau.
- Automated data quality checks using Python, reducing manual validation time by 35% & ensuring high data accuracy.
- **Technologies:** Python, Spark, Azure Data Factory, Azure Blob Storage, Azure SQL Database, Azure Active Directory, Databricks, Tableau, Machine Learning, Salesforce, Rhapsody, Azure Data Lake, Azure Synapse Analytics, Azure Purview.

Data Scientist

Intellihot   *June 2023 – April 2024 | Chicago, IL*
- Developed asynchronous APIs and data pipelines, enhancing data accessibility by 97% and operational efficiency.
- Conducted time-series analyses for predictive maintenance, reducing system downtime by 15% and improving reliability.
- Designed PowerBI dashboards for real-time KPI monitoring, reducing reporting time by 95% improving decision-making.
- Integrated machine learning models into workflows, achieving a 12% increase in system reliability and process optimization.
- Collaborated with cross-functional teams to implement edge solutions, enhancing predictive analytics capabilities.
- Enabled seamless connectivity between disparate systems using CData connectors for streamlined data integration.
- **Technologies:** Python, SQL, Power BI, Streamlit, Scikit-learn, Azure, Thingworx, CData, Microsoft Excel.

Graduate Teaching Assistant

Illinois Institute of Technology   *Sep 2021 – May 2023 | Chicago, IL*
- Mentored 200+ students in Big Data Technologies, taught Hadoop, Spark, Kafka, emphasizing scalable data solutions.
- Conducted workshops on data engineering best practices, improving hands-on skills in large-scale data processing.
- Guided students on implementing machine learning techniques using cloud platforms, focusing on real-world scalability.
- **Technologies:** Python, R, SQL, Hadoop, Spark, Hive, Kafka, AWS, Azure, Tableau.

Data Scientist

Caterpillar   *Jan 2021 – Jul 2021 | Chennai, India*
- Boosted sales by 8% & repeat purchases by 50K/month through an optimized marketing attribution model (Shapley).
- Reduced features in a propensity model by over 50%, improving ROC-AUC scores by 15% & enhancing prediction reliability.
- Designed an optimized web page for 45K+ daily users, increasing user satisfaction by 30% & engagement rates.
- **Technologies:** Python, SQL, GCP, Power BI, Data Studio, Machine Learning.

TECHNICAL SKILLS
- **Programming & Tools:** Python, R, SQL, Tableau, Power BI, Data Studio, Streamlit, Git, Docker, Microsoft Excel.
- **Machine Learning:** Scikit-learn, Keras, tensorflow.
- **Big Data:** Spark, Kafka, Hadoop.
- **Cloud Platforms:** Azure, AWS, Databricks, Salesforce, Snowflake.
- **Advanced:** Data Engineering, Feature Engineering, A/B Testing

EDUCATION

Master of Data Science   *Illinois Institute of Technology* | May 2023 | Chicago, IL
- GPA: 3.7/4.0
- Relevant Coursework: Big Data Technologies, Time Series Analysis, Neural Networks, Data Mining, Statistical Learning.

Bachelor of Computer Science & Engineering   *Amrita Vishwa Vidyapeetham* | May 2021 | Coimbatore, India
- GPA: 8.95/10.0
- Relevant Coursework: Neural Network & Deep Learning, Machine Learning, Cloud Computing, Database Management.

PROJECTS
- **Workplace Safety Monitoring System:** Designed a solution using MobileNet and Blazepose for real-time workplace safety compliance, achieving 99% mask detection accuracy and 93% pose detection, reducing violations by 15%.
- **Google Ads Optimization:** Developed predictive models using CatBoost to optimize ad targeting, increasing impressions by 20% and boosting website traffic by 10%, while improving ROI by 15% through enhanced keyword strategies.
- **Custom Database Manager:** Engineered a custom DBMS in Python with a B+-tree index for efficient query execution, reducing response times by 30% and enabling seamless data management for large-scale applications.
- **Customer Churn Prediction:** Developed a machine learning model with 92% accuracy to predict customer churn, enabling targeted retention strategies and increasing customer loyalty by 10%.
- **Fraud Detection System:** Implemented an anomaly detection system using Scikit-learn to identify fraudulent transactions in real-time, reducing false positives by 20% and preventing potential losses.
- **Sales Forecasting with Time Series Analysis:** Designed a forecasting model using ARIMA and LSTM to predict monthly sales with 95% accuracy, assisting in inventory management and supply chain optimization.