

SAHANA GIRISH

USA | (217) 979-5487 | sgirish2@illinois.edu | [LinkedIn](#) | [Website](#)

EDUCATION

University of Illinois Urbana-Champaign

Master of Science in Information Management (MSIM)

August 2023 - May 2025

4.00/4.00 GPA

Visvesvaraya Technological University

Bachelor of Engineering in Electronics and Communication Engineering

August 2017 - July 2021

3.64/4.00 GPA

SKILLS

Languages: Python, SQL, R, MATLAB

Libraries: Flask, Pandas, TensorFlow, Scikit-Learn, PyTorch, Keras, OpenCV, Requests, NumPy, Matplotlib, SciPy, Seaborn

Technologies: AWS (SageMaker, Lambda, Data Lake, ECR, S3, Glue), Azure (ML, Databricks, Data Factory), GCP, MLflow

Others: Git, Docker, Tableau, Power BI, Qlik, Jupyter, Spark, PyCharm, Hadoop, Evernote, Airflow, Agile, Jira, MS Suite

WORK EXPERIENCE

Graduate Associate – Grainger College of Engineering (UIUC)

August 2023 - May 2024, January 2025 - Present

- Optimize data processing and track supply chain issues by overseeing research order placement, tracking, and management
- Manage the department's facilities, HR, and front office databases and dashboards ensuring improved operational efficiency

Data Science Intern - Werfen (R&D)

August 2024 - December 2024

- Applied statistical curve fitting techniques to digitize 3B+ instrument time series data to support product feature development
- Formulated a scalable flagging strategy and threshold analysis to handle reported false positives, thereby reducing it by 96%
- Partnered with product managers in analyzing sales trends, UX and correlations, facilitating data-informed decisions
- Delivered customized datasets and enhanced data pipelines through complex SQL queries, resulting in 62.85% faster runtime

Data Science Intern - OSF Healthcare (Advanced Analytics)

May 2024 - August 2024

- Forecasted cardiac abnormalities with 88% sensitivity by implementing dual-detector LSTM and CNN time-series models
- Designed automated testing, validation and feedback loops using BERT and RLHF GenAI systems to monitor the chatbot
- Conducted A/B testing and control group analysis to validate positive reinforcement and multivariate testing, measuring a 20% improvement in its response performance
- Collaborated with cross-functional teams in sprint planning workshops to develop streamlined project roadmap flowcharts

Senior Data Scientist - Comviva

September 2021 - July 2023

- Led the design, development, Docker containerization, AWS SaaS deployment of an end-to-end automated AI/ML product
- Integrated REST APIs, model performance visualization tools and deep learning modules based on customer requirements
- Streamlined the cleaning and processing of large datasets to construct demand forecasting and upselling/cross-selling machine learning models, yielding an 8% to 10% reduction in customer churn for telecom clients
- Deployed Qlik Cloud and Monitor, obtaining a 30% run-time reduction in the dashboard generation process
- Accelerated SDLC by implementing sprint iterations, CRM skills for SaaS, boosting Product Life Cycle velocity by 60%
- Managed a team of five, overseeing the big data ETL processes by using Apache NiFi and PySpark

Machine Learning Intern - Compssoft Technologies

July 2021 - September 2021

- Performed market analytics and sentiment analysis on social media/website reviews using NBSVM, achieving 84% accuracy
- Increased accuracy to 86.7% through NLP enhancements such as cleaning, lemmatization, stemming, and tf-idf techniques
- Refined interactive Power BI dashboards with query folding and data modeling to effectively present performance metrics

PROJECTS

Business consultancy for a motion capture company to enter the US retail market

April 2024

- Provided a data-driven B2B strategy with in-store analytics, BI, and warehouse management system use cases by conducting detailed market research, and cohort analysis to examine customer lifetime value and shopping behaviors
- Recommended a framework for targeted marketing, supply chain optimization, and client acquisition to drive lead generation

Diabetes Prediction using Machine Learning Classifiers: Random Forest, Naïve Bayes & XGBoost

January 2022

- Processed the Pima dataset using 5-fold cross validation and maximized the metrics through hyperparameter tuning
- XGBoost outperformed Naïve Bayes with sensitivity, specificity and AUC scores of 81.2%, 94.5% and 2.02% respectively

Customer Segmentation and Next Best Offer Recommendation for retail e-commerce dataset

December 2023

- Segmented high-value customers using K-Means clustering and RFM analysis for marketing strategies to improve retention
- Built sales prediction models to determine the propensity score of customers accepting offers, attaining a recall of 86.1%