SAHANA GIRISH

(217) 979-5487 | sgirish2@illinois.edu | LinkedIn | Website

EDUCATION

University of Illinois Urbana-Champaign **August 2023 – May 2025**
Master of Science in Information Management (MSIM) 4.00/4.00 GPA

Visvesvaraya Technological University **August 2017 – July 2021**
Bachelor of Engineering in Electronics and Communication Engineering 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), Azure (Data Factory, Databricks, Stream Analytics), MLflow
Others: Git, Docker, Tableau, Power BI, Qlik, Jupyter, Spark, PyCharm, Hadoop, Statistics, Airflow, Agile, Jira, MS Suite

WORK EXPERIENCE

***Data Science Intern,* Werfen – R&D** **August 2024 – Present**
- Applied statistical curve fitting techniques to digitize the instrument time series data to aid new product feature development
- Formulated a robust 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, extracted new features and optimized the existing pipeline through complex SQL queries leading to 62.85% faster runtime while adhering to GMP for quality, traceability, and regulatory compliance

***Data Science Intern,* OSF Healthcare – Advanced Analytics** **May 2024 – August 2024**
- Forecasted cardiac abnormalities by implementing a dual-detector LSTM and CNN time series models with 88% accuracy
- Designed insightful dashboards and assessed operational KPIs using statistical process control to support the mergers & acquisitions team in due diligence and market strategy analysis thereby enabling real-time monitoring of process variation
- Employed A/B testing and control group analysis to identify operational strategies and reduced potential risks by 20%
- Collaborated with cross functional teams in sprint planning workshops to develop efficient project roadmap flowcharts

***Senior Data Scientist,* Comviva** **September 2021 – July 2023**
- Led the design, development, containerization, AWS SaaS deployment of an automated AI/ML product - Mobilytix AIx
- Expanded Mobilytix AIx by incorporating REST API, model performance visualization tools and deep learning modules
- Streamlined the cleaning and processing of large datasets to construct demand forecasting machine learning models, yielding an 8% to 10% reduction in customer churn
- Accelerated SDLC by implementing sprint iterations, CRM skills for SaaS to improve Product Life Cycle velocity by 60%
- Supervised new joiners in meticulously extracting, transforming and loading big data by using Apache NiFi and Spark

***Machine Learning Intern*, Compsoft Technologies** **July 2021 – September 2021**
- Performed market analytics and sentiment analysis on social media comments using NBSVM, achieving 84% accuracy
- Augmented accuracy to 86.7% through NLP enhancements such as cleaning, lemmatization, stemming, and tf-idf techniques
- Designed data pipelines with AWS S3 and Glue, optimized SQL scripts to enable efficient ETL processes and data storage
- Optimized interactive Power BI dashboards through query folding and data modeling for presenting performance metrics

PROJECTS

Business consultancy for a motion capture company to enter the retail domain **April 2024**
- Provided a data-informed B2B strategy to enter the US retail market with in-store analytics, BI and warehouse management system use cases by carrying out detailed market research, product and data analysis, and visualizations
- Performed cohort analysis to examine customer lifetime value and shopping behaviors, providing actionable recommendations for targeted marketing efforts, supply chain optimization and customer retention plans

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 hyper parameter tuning
- XGBoost outperformed Naïve Bayes with a sensitivity, specificity and AUC 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 by K-Means clustering and RFM analysis for marketing strategies to improve retention
- Built predictive models to determine the propensity score of customers accepting offers and achieved a recall of 86.1%