SINDHUJA MAHESWARAN

smaheswa@syr.edu| **+1 3154509688** | www.linkedin.com/in/sindhuja-maheswaran| https://github.com/SindhujaMaheswaran

EDUCATION

Master Of Applied Data Science, Syracuse University, New York, USA	**12/2024**
Bachelor in Engineering, Computer Science, Anna University, India	**05/2018**

TECHNICAL SKILLS

Skills:

- **Data Analysis Software & Programming:** Python, R, SQL, Tableau, Power BI, MS Excel- Advanced Formulas, Macros, VLOOKUP, Pivot Tables and Pivot Charts, Conditional Formatting, PySpark, Linux, Apache Spark.
- **Database & Data Management:** Microsoft SQL Server, ETL, AWS, Azure, Docker, Data Warehouse, GIT, RDBMS, SSMS, Snowflake, Visual Studio, SSMS.
- **Statistical & Visualization Libraries**: Pandas, Scikit-Learn, TensorFlow, ggplot, Numpy, Keras, NLTK, Plotly, Matplotlib FLASK, BART, Plotly, Adam, Data Storytelling, Microsoft Office Suite, VGG16.

Certifications: Data Science hands-on work experience, Developing SQL Databases, Python Programming, Business Intelligence using Power BI, Hands-on experience on Data Warehouse, Google AI essentials, and Selinium Software Testing.

EXPERIENCE

Data Science Researcher – Nexis Student Lab Technology, Syracuse University, USA　　　　**09/2024- Present**

- Leading ongoing research and data engineering project for a scalable recommender system with AWS SageMaker, MXNet, and Gluon, tracking milestones to achieve a 20% improvement in user engagement.
- Building secure data pipelines with IAM and CloudWatch, enhancing system reliability and operational efficiency by 15% through detailed technical communication of data definitions and business rules.
- Applying advanced analytics to optimize data models, using data-driven decision making and model-driven solutions to target a 30% increase in recommendation accuracy.

Data Analyst – SK Multispecialty Hospital, India　　　　**08/2018 – 05/2022**

- Automated KPI-tracked reporting with data validation processes, enhancing data quality by 100% and resolving data integrity issues across 1,000+ weekly stakeholder presentations on public health data, boosting customer satisfaction.
- Performed healthcare analytics with SQL for data retrieval and competitive analysis, delivering actionable insights to improve data-driven decisions by 40% for non-technical stakeholders.
- Collaborated with cross-functional teams to define requirements, addressing business challenges and reducing report discrepancies by 25% through a collaborative mindset and ensuring strategic alignment.
- Developed ad-hoc analyses using root cause analysis to enhance data accuracy and support decision-making, aligning findings with stakeholder presentation needs.

HANDS-ON WORK EXPERIENCE

Data Science and Machine learning Intern- Greedge, India　　　　**05/2023 – 8/2023**

- Collaborated with tech leads and analysts in 40+ modeling innovation meetings, automated data processing usingPython and R, gathered business requirements, and implemented statistics and quantitative analysis.

ASL- Image Classification System | ML, Deep Learning, Image Processor, CNN, Adam, OS, OpenCV, SciKit-Learn

- Led the development of a CNN model with 686K+ parameters, achieving 92% model performance for ASL recognition, while applying data interpretation and quality controls to minimize misclassification
- Utilized image segmentation and problem-solving skills to enhance ASL detection accuracy by reducing error rates and demonstrating strong project management and interpersonal collaboration.

PERSONAL PROJECT

Text Summarization-API using Large Language Models | Flask, BART, PyTorch, Ngrok , NLP　　　　**10/2024**

- Developed a Flask-based ETL API using Hugging Face's BART model, to enhance summarization capabilities.
- Optimized prompt design techniques to generate accurate summaries, resulting in a 40% reduction in content length while retaining 90% of key information.

Apple Retail Sales Analysis| Advance SQL, Data Segmentation　　　　**04/2024**

- Implemented data gathering, reporting, and maintaining database structures using advanced SQL joins and aggregations on 1M+ sales records to enhance financial performance analysis and product insights.
- Analyzed P&L metrics and price pattern identification, supporting market growth predictions and improving strategic decision-making.

Crime Classification based on Geographical Area |Data Mining, Python, Sklearn, Seaborn, TensorFlow　　　　**11/2023**

- Enhanced anomaly detection accuracy by 70% through feature engineering, and statistical analysis on 1M+ records, identifying high-risk areas to support predictive models for crime trends.
- Applied machine learning model development, including CNN, gradient boosting, and logistic regression to improve crime trend classification and regression techniques, generating insights for project management.