Maria Nikitha Suresh

marianikitha.suresh@gmail.com | +1 520-333-0012 | www.linkedin.com/in/maria-nikitha-588994242 | **Github** | **Portfolio**

PROFESSIONAL EXPERIENCE

Machine learning Developer - ASTEC Lab, University of Arizona **Aug 2024 - Dec 2024**

- Developed a deep learning-based suture classification model (YOLOv8) to identify sutures as **Good, Tight, or Loose** with **95% accuracy**.
- Conducted **exploratory data analysis (EDA)** and **feature engineering** to enhance model interpretability.
- Implemented **precision, recall and F1-score** for model evaluation, optimizing for medical imaging challenges.
- Automated image preprocessing pipelines using Python (**OpenCV, NumPy**) to improve detection consistency and presented findings to research teams, showcasing **data-driven insights** for improving surgical training.

Student Research Worker - University of Arizona **Jul 2024 - Aug 2024**

- Worked on neural network-based video processing for **foreground-background** separation using **Bayesian Tensor Factorization**.
- Applied **probabilistic modeling and regularization techniques** to enhance video quality and feature extraction.
- Conducted performance tuning and model evaluation to improve output accuracy and robustness.

Data Analyst & Software Development Engineer - Protecto (OneDPO Private Limited) **Jun 2022 - Jul 2023**

- Built **interactive Tableau & Power BI dashboards** for **privacy risk analysis**, enabling data-driven decision-making.
- Designed and implemented frontend features using ReactJS for internal privacy risk analysis tools, improving user experience and engagement.
- Utilized **statistical modeling** to detect **anomalies and privacy risks**, ensuring compliance with industry standards.
- Collaborated cross-functionally to implement data-driven BI solutions for enterprise data workflows.

EDUCATION

University of Arizona, United States **Aug 2023 - May 2025**
Masters of Science : Data Science | GPA : 3.78

Christ University, India **June 2018 - May 2022**
Bachelor of Technology: Information Technology | GPA : 3.79

SKILLS: Deep Learning, Machine Learning, Data Engineering, Generative AI (GenAI), NLP, Data Analysis & Visualization.

TOOLS: Python, TensorFlow, PyTorch, Keras, Scikit-learn, OpenCV, NumPy, Pandas, SQL, R, C++, Git, AWS (Lambda, S3, Glue, Athena), Azure (Data Factory, Databricks, Data Lake Storage), Apache Spark, Transformers, GPT, BERT, NLTK, MySQL, MariaDB, MongoDB, R, Java, Bash, PHP.

PROJECT EXPERIENCE

Chikankari Design Generator using Generative AI | Capstone Project **Link**

- Built a generative pipeline using **Stable Diffusion v1.5** fine-tuned with **Advanced LoRA** on 4000+ Chikankari embroidery images to synthesize authentic pattern designs.
- Developed training workflows with **Hugging Face Diffusers**, including data augmentation, checkpointing, and LoRA adaptation.
- Evaluated model performance using **FID** and **CLIPScore** to assess realism and prompt alignment.
- Designed and implemented a **user interface** with Streamlit that allows users to input custom prompts and generate embroidery patterns, making the tool accessible for non-technical users.

Early Readmission Prediction using Random Forest, SMOTE, and AWS-Based ETL **Link**

- Developed a machine learning model to predict early hospital readmissions using a **Random Forest Classifier**, achieving a training accuracy of **96%** and a test accuracy of **88.99%.**
- Implemented an **ETL pipeline** on **AWS** to automate data extraction, transformation, and loading processes. Utilized AWS Glue for data cataloging and ETL operations.
- Performed data preprocessing and oversampling techniques, specifically using **SMOTE**, to address class imbalance in the dataset, enhancing model performance.

Transforming E-Commerce Data with Azure: End-to-End Pipeline Implementation **Link**

- Designed and implemented an end-to-end data pipeline using **Azure Data Lake** Storage, **Azure Data Factory,** and **Azure Databricks.**
- Ingested E-commerce data from multiple sources into Azure Data Lake and developed **ETL** workflows in Azure Data Factory for data extraction, transformation, and loading.
- Utilized **Apache Spark** on Azure Databricks for processing large-scale data, following Bronze, Silver, and Gold architecture.
- Leveraged **Delta Lake** for optimized storage, data versioning, and reliability. Implemented data quality checks and performance optimizations to enhance efficiency.