MANASVI ALAM

+91-6374501358 • manasvialam03@gmail.com • LinkedIn/manasvi alam

EDUCATION

Sastra Deemed University, School of Computing — Tamil Nadu, India
B.Tech in Computer Science and Engineering(Specialization in AI and Data Science) — August 2020 - May 2024
– **Vice-Chairperson**, Entrepreneurship Cell of Sastra

PUBLICATIONS

1. Reducing Hepatitis C Diagnostic Disparities with Automated deep learning for HCV Antigen Detection, *Science Advances*
Link: *https://www.science.org/doi/10.1126/sciadv.adt3803* **(IF ~ 12.5, Q1)**
2. Deep Learning-Enabled ELISA Microfluidic Cartridge for Viral Pathogen Detection, *Advanced Materials Technologies*
Link: *https://doi.org/10.1002/admt.202500025* **(IF ~ 6.2, Q1)**
3. Enhancing IVF Success Prediction with AI, *Fertility and Sterility, ASRM (Poster)*
Link: *https://www.fertstert.org/article/S0015-0282%2824%2901161-0/fulltext* **(IF ~ 7.00, Q1)**
4. Deep Learning Models for Classification of Remote Sensed Images, *Applied Intelligence, SCIE Springer Journal (Under review)*

WORK EXPERIENCE

Exponential AI — Hyderabad, India
Machine Learning Engineer (Full-time) — November 2024 - Present

1. Claim Denial Prediction for CARC 4 & 252 codes, and Recommendation system
***Client**: U.S. private equity-backed national anesthesia and perioperative management firm with ~ $600 million in annual revenue*

- Formulated a **Proof of Concept (POC)** for Claim Denial Prediction for **Claim Adjustment Reason Codes(CARC) 4 and 252**, along with a recommendation system to minimise process delays and save providers' time and costs.
- Headed a **10-member team** to develop an **LLM-powered system targeting retrieval precision**, lower latency, and scalable design by fine-tuning domain-specific embedding, reranker models, and optimizing model pipelines after extracting data from EDI 837 & 835.
- Integrated **RAG pipelines** with payer-specific rules for context-aware recommendations, improving automation and acceptance rates, and achieving an accuracy of **98.2%**. Estimated financial savings **exceeding over $20M/year** in the initial months by saving the possible denials and recommendations-related operational costs.

2. Email-Intake Processing and Multi-format Attachment Parser
Client: U.S. based $247B-revenue global health insurer serving ~20 million medical customers

- Built and deployed a **scalable LLM-based data extraction pipeline** for structured and automated data collection by separately developing email summarisation, entity extraction, and intent/action revelation modules.
- Architected a **multi-format attachment parsing pipeline** comprising OCR, parsers, and content classifiers for suggesting context-aware actions to the user, with a speedup of over **89%**.
- Integrated the pipeline to the client's system using **API-based services**, increasing **operational efficiency** by **67%**.

3. Contract Page Classification and Extraction
Client: $ ~2B+ U.S. healthcare-IT company powering Medicaid and public-health systems across all 50 states

- Performed **classification** of contract pages on a 4200-page dataset using a fine-tuned Electra model, obtaining a **95.7% F1-score**.
- Implemented a pipeline with OpenCV, Tesseract OCR, and LayoutLMv3 for **layout parsing**, reducing PDF extraction latency whilst achieving **92.3%** accuracy in **data extraction** with a hybrid model involving a data parser and RoBERTa-based semantic validation, currently in the **production environment** in **client systems**.
- Improved email summarisation efficiency by **85%** token reduction for 60+ Excel attachments, benchmarking Claude and Gemini LLMs, and optimising PostgreSQL sync rate by designing structured tables and automating pipelines.

Harvard Medical School, Shafiee Lab — Boston, USA
AI Research Intern under Dr. Hadi Shafiee — February 2024 - July 2024

1. AI-driven Hepatitis C Diagnostic System

- Developed **SPyDERMAN** (Smartphone-based Pathogen Detection Resource Multiplier using Adversarial Networks), a **novel GAN-based domain adaptation framework** to **combat low viral load detection** sensitivity and dataset shortage for HCV detection.
- Designed an end-to-end smartphone Point-of-care model pipeline and user interface for HCV cAg image classification on the **Cherokee Nation dataset**, achieving **94.59%** overall accuracy. Reduced detection threshold to **574 IU/ml**, enabling early and error-free Hepatitis C diagnosis in resource-limited settings.

2. Adversarial Learning for Microfluidic Viral Detection (VISTA)

- Contributed to VISTA by developing an Adversarial Network model for a bubbling immunoassay-based system that detects **SARS-CoV-2 antigen** and **HCV core antigen** using minimal sample volumes. Integrated AI with microfluidics to create a cost and time-efficient alternative to traditional ELISA testing, optimizing models to handle noisy data from biological samples.
- Accomplished an accuracy of **95.45%** for HCV, **93.3%** for SARS-CoV-2 patient samples, and **100% sensitivity** across all isolates through training and tuning to handle all cases of biological samples.

3. Embryo & Patient Data–Driven IVF Outcome Prediction Model

- Analyzed the effectiveness of various deep learning models in **predicting live birth outcomes** for IVF patients by incorporating patient demographics, cycle data, and dynamic oocyte/embryo imaging data.
- This analysis highlights the critical need for embryonic development metrics, which led to a **13%** improvement in prediction accuracy.

Fidelity Investments Bengaluru, India
Full-stack intern June 2023 - August 2023

- **Led** a team of **15 interns** to create a **NextGen UI** that eases workflows and interactions of operations for the Brokerage Research and Innovations centre, **combining 4** different business platforms containing customer and beneficiary information into a **single interface**.
- Decreased document handling and transaction processing time from **8-10 minutes** to **2-3 minutes** within a span of **4 weeks**.
- Presented project results to the **business team and industry head**, highlighting a **90% increase** in employee efficiency.

Sonata Software Bengaluru, India
Web Development Intern February 2022 - May 2022
Website: ***https://www.intachblr.org/ulsoorsomeshwara/***

- Developed an **easy-to-navigate** website with visually appealing features like 360-degree site views for the **Indian National Trust for Art and Cultural Heritage**, as a part of the organization's goal to spread awareness and increase accessibility of **Indian Heritage**.
- Worked with a team of 8 members to develop a functional website using HTML, CSS, JavaScript, and Bootstrap.

PROJECTS

Kidney Tumor Detection, Deep Learning August 2024
Link - ***https://github.com/manasvialam/Kidney-Disease-Classification***

- Developed and fine-tuned a deep learning model for **tumor versus normal kidney classification**, achieving **98.3%** accuracy through the utilization of logging, exception handling, and utility modules that **increased process workflow stability**.
- Simplified model **version control** and **experiment tracking** by using MLflow and DVC Pipeline for data intake to the evaluation stage, deploying the model as a web service with React and Docker, which enabled **accessibility** and a user interface.

Remote Sensed Image Recognition, Deep Learning February 2023 – May 2023

- Conducted a **comparative analysis** of deep learning approaches to classify remote sensed images from the UCM dataset, evaluating CNNs, transfer learning, and fine-tuning techniques across **ResNet**, **EfficientNetB7**, and **MobileNet** architectures.
- This analysis achieved **95.62%** accuracy, improving by more than **8%** over the previous benchmark established in IEEE, and was deployed on the web using the Streamlit API to enable accessible image classification.

ACHIEVEMENTS

- **1st Place Winner**, Intelligent Farming Systems Hackathon, National Level Techno-Management Fest(DAKSH), SASTRA University; awarded Rs. 50,000 cash prize
- **Participant**, MIT GrandHack 2024(Selected from hundreds of global applicants) and AI Cures Conference, Boston, USA
- **Event host, Head & Vice-Chairperson**, Entrepreneurship Cell, SASTRA University; hosted and managed multiple events with 600-700 attendees and supported the establishment of 10+ startups across various domains

COMMUNITY INVOLVEMENT

- **Volunteer, Rashtriya Seva Samithi (RASS)**: Worked with disabled children to develop foundational skills for day-to-day activities, tailoring support to their individual learning rate and comfort level. Taught basic English and economics concepts, such as loans and entrepreneurship, to low-income women, enabling them to support themselves financially, and provided emotional support to senior citizens experiencing loneliness.
- **Volunteer, SASTRA Deemed University COVID-19 Outreach**: Delivered COVID-19 guidance and public health information as part of SAHO(SASTRA's Hope), regularly tracking the health of around 30 patients to provide them with critical information on vaccinations, oxygen supply, and bed availability. This tracking enabled me to serve as the point of contact between patients and medical authorities, facilitating immediate support in case of medical requirements or sudden health escalation.

ADDITIONAL INFORMATION

Technical Skills: Python, C++, R, SQL, TensorFlow, Keras, Scikit-learn, Pytorch, NumPy, Pandas, Seaborn, Matplotlib, NLTK, OpenCV, LangChain, ChromaDB, MLflow, DVC, Docker, Kubernetes, Git, Bitbucket, FastAPI, Flask, Hadoop, PostgreSQL, MongoDB, JavaScript, React, Angular, HTML, CSS, Bootstrap, Tableau, Insomnia

Coursework: Machine Learning Techniques, Time series & Sequential Data Analytics, Data Analytics for Healthcare Applications, Computer Vision & Pattern Recognition, Natural Language Processing, Predictive Analytics & Data Visualization, Data Structures, Data Warehousing & data mining, Fundamentals of AI & Data Science, Computer Networks, Forensics, Database Management Systems, Discrete Structures, Digital System Design, Computer Organisation

Certifications: Introduction to Statistics (Stanford University), Google Data Analytics Professional Certificate, Machine Learning A-Z (Python & R), Introduction to TensorFlow & Keras, Machine Learning Onramp, Machine Learning Pipelines with Azure ML Studio, Building custom regional reports with Google Analytics