Arun Nemani

neman014@gmail | arun-nemani
U.S. Citizen | arunnemani.com

Education

PhD | Biomedical Engineering
Rensselaer Polytechnic Institute
2012-2017

- Research: machine learning, image processing, brain imaging, surgical skill assessment
- Advisors: Suvranu De, Xavier Intes

MS | Biomedical Engineering
Rensselaer Polytechnic Institute
2010-2012

- Research: image processing, virtual simulators, surgical skill assessment

BS Biomedical Engineering
University of Minnesota - Twin Cities
2005-2009

Skills

Languages
Python • SQL • bash
Terraform • Matlab • C++ • R

Machine Learning
Open source frameworks
PyTorch • Tensorflow • hugging-face • scikit-learn • pandas • numpy • scipy catboost • airflow • XGBoost dask • shap • sentence-transformers • lifelines

Algorithms
deep learning • CNN • transformers gradient boosting • contrastive-learning clustering • regression • classification survival • PCA / LDA • bayesian methods model selection • model validation

Healthcare
Real world data (RWD) • EHR • ICD-9/10 Electrocardiograms • Echocardiograms Medical imaging • Cardiology • Oncology

General
Databricks • AWS • GCP • Azure • (yup, all three)
LaTeX • linux • UNIX • git • docker spark • DBT • CUDA • OpenCV Kubernetes • CI/CD

Experience

Senior Staff Machine Learning Engineer | Altana AI
Nov 2014 - Present

- Designed, validated, and deployed a narcotics screening classifier trained on 200M transactions (PR - AUC = **0.22 @ 0.00001** prevalence), running at **50ms p50 latency** and handling **10M inferences/week** for a Fortune 50 firm. **Patent pending.**
- Designed, validated, and deployed a Bayesian regressor using **4B** global supply chain data to estimate declared value and flag anomalies for anti-laundering, miss-valuation, and tariff evasion; serves **3B** inferences/week for multiple U.S. federal agencies. **Patent pending.**

Senior Staff Machine Learning Scientist | Tempus AI
Aug 2019 - 2024

- Co-invented the **FDA 510(k) approved** Tempus Atrial Fibrillation (AF) algorithm aimed to predict first time AF risk within one year. Deep learning model is trained on **3.5M ECGs from 2M patients** (ROC - AUC = **0.84**) and is deployed **four+** hospital systems.
- Inventor and owner of the Tempus ECG training platform, a distributed cloud-based GPU modeling framework that powers all multi-modal research and cardio ECG production models, generating **$10M+** in revenue.
- Served as lead data engineer for Tempus Cardio with responsibilities ranging from creating databases for EHR and imaging data and successful integration of four healthcare partners consisting of **2.5M+ patients and 10TB+ worth of data** enabling data access to **10+** team members.
- Validated and implemented a de-identification pipeline to scalably de-identify **1 billion** clinical notes from **2.5M** patients within **40** hours yielding **98%** sensitivity.
- Led org-wide documentation protocols to establish best practices on coding standards, cloud computing fundamentals, and clinical validation methods.

Founder and Chief Scientist | Draycon Labs
Jul 2019 - Feb 2025

- Build, validated, and deployed a cloud-based deep learning segmentation platform for end-end histology and automated imaging analysis workflows for Fortune 500 medical device firms with **100%** license renewal rate for **6 years**.

Sr. Data Scientist | Food Genius (acquired by US Foods)
Jul 2018 - Jul 2019

- Designed, built, and deployed a full-stack, machine-learning based web app that predicts supply chain service levels enterprise wide, with **36%** higher balanced accuracy than food industry standards

Key Publications and Patents

Articles

- Prediction of mortality from 12-lead electrocardiogram voltage data using a deep neural network *Nature Medicine*
- Deep Neural Networks Can Predict New-Onset Atrial Fibrillation From the 12-Lead ECG and Help Identify Those at Risk of Atrial Fibrillation–Related Stroke *Circulation*
- Assessing bimanual motor skills with optical neuroimaging *Science Advances*

Patents

- US11869668B2, US20230245782A1, US11657921B2, US20230028783A1, US20210076960A1