Arun Nemani

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

Education

PhD | Biomedical Engineering
RPI | 2013-2017

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

MS | Biomedical Engineering
RPI | 2010-2012

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

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

Skills

Languages
Expert
Python • SQL • bash

Proficient
Terraform • CSS • Matlab • C++ • R

Machine Learning
Open source frameworks
Tensorflow • Keras • PyTorch • scikit-learn • pandas • numpy • scipy catboost • lightgbm • XGBoost dask • shap • matplotlib • lifelines

Algorithms
deep learning • CNN • transformers gradient boosting • SVM • clustering regression • classification • survival PCA / LDA • decision trees • NLP model selection • model validation

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

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

Experience

Founder and Chief Scientist | Draycon Labs
Jul 2019 - present

- Founder of a consultancy focused on implementing ML/AI solutions for medical imaging, laboratory workflow optimizations, and pre-clinical studies.
- Build, validated, and deployed a cloud-based deep learning segmentation platform for end-end histology and automated imaging analysis workflows for a Fortune 500 medical device firm with **100%** license renewal rate for **4 years**.

Senior Staff Machine Learning Scientist | Tempus Labs
Aug 2019 - present

- Co-invented the Tempus Atrial Fibrillation (AF) algorithm aimed to predict first time AF risk within one year. Deep learning model is trained on **3.5M ECGs across 2M patients** (ROC - AUC = **0.84**) and is deployed in **three** hospital systems for clinical trials and currently undergoing **FDA 510k validation**.
- Inventor and maintainer of Tempus ECG training platform, a cloud-based distributed GPU modeling framework powering all multi-modal research and production models including **FDA Software as Medical Device (SaMD)** builds.
- 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 **477M** clinical notes from **2.1M** patients within **30** hours yielding **98%** sensitivity.
- Managed three data scientists to execute on cross-functional objectives, establish best practices, and mentorship.

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
- Managed four data scientists to execute on cross-functional business initiatives.

Research Scientist | Rensselaer Polytechnic Institute
Sep 2010 - Dec 2017

- Invented and validated a brain imaging based machine learning model to predict surgical motor skill (ROC - AUC = **0.92**) with a **113%** higher accuracy than current US Surgery Board Certification methods. Thesis work resulted in **two R01 grants**.
- Led three multi-institutional NIH clinical studies with **$2M** of funding.

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

- US20230245782A1, US11657921B2, US20230028783A1, US20210076960A1