Arun Nemani

✉ neman014@gmail   ⌨ arun-nemani   🎓

🌐 U.S. Citizen   💬 arunnemani.com   in

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 - TC | 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 Electrocardiograms • Echocardiograms Phenotyping • cardiovascular disease

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 🔗
Aug 2019 - present
- Founded Draycon Labs, a consultancy focused on implementing and licensing cloud-based, bespoke ML/AI workflows with applications in medical imaging, laboratory workflow optimizations, and preclinical studies.
- Service clients including a Fortune 500 medical device firm with **100%** license renewal rate for cloud-based deep learning ML software.

Staff Machine Learning Scientist | Tempus Labs 🔗
Aug 2019 - present
- Co-invented the Tempus Atrial Fibrillation (AF) prediction algorithm aimed to robustly predict AF risk within one year. Model is deployed in **three** hospital systems for a clinical trial and currently **undergoing FDA 510k validation**.
- Invented and patented cloud-based Tempus ECG training platform, powering all multi-modal research and production deep learning workflows including **FDA Software as Medical Device (SaMD)** builds.
- Served as sole data engineer for Tempus Cardio with responsibilities ranging from cloud migration, EHR + ECG dataset ingress, database standup, and establishing coding standards to scale technical team to **10+** 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.

Sr. Data Scientist | Food Genius (acquired by US Foods) 🔗
Jul 2018 - Aug 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
- Developed and validated machine learning models using brain imaging data to classify surgical motor skill proficiency.
- Thesis work validated models to robustly (ROC - AUC = **0.94**) predict motor skill levels with a **113%** higher accuracy than current US Surgery Board Certification methods.
- Led three multi-institutional NIH clinical studies with **$2M** of funding in collaboration with Massachusetts General Hospital, Harvard Medical School, Yale Medical School, and University at Buffalo.

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
- 17/026,092 | 17/829,356 | 17/829,357 | 17/814,229 | 18/09,320