

Surya Narayanan Hari
suryanarayananhari@gmail.com
650.304.8805

Education

- **Stanford University** *2021*
 - M.S. MS&E (Data Science & Machine Learning Focus)
 - B.A. Economics, Stanford University (2019)
 - B.S. Mathematics, Stanford University (2019)

First-author publications

1. Using Distributional Robustness to mitigate the reliance on spurious correlates in histopathology - preprint
 - Led to two abstracts accepted to the Workshop on Machine Learning in Computational Biology
 - Led to an abstract accepted to the NeurIPS 2021 workshop on Distributional Shifts - Poster Presentation
2. Senior Thesis published as a chapter in a book from the Harvard lab for Entrepreneurship and Development. - Link
 - Paper was also accepted to be published in Stanford Undergraduate Research Journal and undergraduate journals at Cornell Economics and Brown University.

Work Experience

- **Computational Biologist, Dana Farber Cancer Institute** *October 2020 - Present*
 - Developed an understanding of biomarkers used to make inferences from histopathology scans in lung cancer
 - Found novel Computer Vision algorithms that offer distributional robustness
- **Machine Learning Intern, Seventh.ai** *June - September 2020*
 - Using Data Science and NLP to construct a metric to measure IP Activity
 - Used Doc2Vec, BERT and classical ML algorithms such as XGBoost
- **Research Intern, Fujitsu Labs America** *January - September 2020*
 - Building an efficient sequencing model using MIP, MILP and QUBO formulations
 - Developed expertise in using solvers such as Gurobi and Google ORTools
 - Worked with team of scientists to deliver solutions to the client, patent pending

Research Experience

- **Research Assistant, Stanford MS&E:** Implemented Top Trading Cycles (TTC), Serial Dictatorship, Deferred Acceptance algorithms and formulations using MIP, MILP using CVXpy as a programming language and Gurobi / SCS solvers to solve combinatorial optimization problems. Modeled non-convex programs in Gurobi.
- **Research Assistant, Harvard Econ.:** Performed Data Analysis and implemented Deep Learning Algorithms for binary classification task. Improved accuracy on highly imbalanced dataset from 50 % using class weights to 98.5 % using SMOTE.
- **Research Assistant, Stanford GSB:** Research Assistant on Machine Learning for Optical Character Recognition (OCR) and text classification algorithms. Implemented an API that implements TF-IDF, Hash Vectorization, Count-Vectorization and Doc2Vec as Data Access Layers and XG-Boost, Logistic Regression and SVM as Prediction Layers. Implemented deep neural network using Ernie & 1D-CNNs in Keras, used Semi-supervised learning, measured precision and recall.

- **Research Assistant, Stanford SIEPR advised by Nobel Laureate Bill Sharpe:** Automated extraction of Key Financial Metrics from comptroller reports using PyPDF2, PDFTools API & RegEx. Automated work reduced manual efforts by 70-75%
- **Research Assistant, Stanford Law School:** Involved in Qualitative Research under Professor Emeritus Paul Brest. Under his tutelage, studied implementations of behavioral nudges & presented research findings to fellow research cohort.

Projects

- **Planted Clique Model:** Understanding the effect of network structures on success of seeding and vaccination strategies. Examining failure of degree centrality, random seeding on a dumbbell graph. Proposed solution of max-cut for effective vaccination. Proposed theoretical results. [Link](#)
- **Causal Inference:** Literature covers a positive correlation between temperature and crime and makes generalizations to account for the effect of confounders. This work offers contradictory evidence to popular literature using an agnostic regression and sensitivity analysis. [Link](#)
- **Meta Learning:** Built few-shot model from SNAIL, ProtoNets, MAML and Transfer Learning on an NIH dataset for detecting Malignant Tumors on CT scans. [Link](#)
- **Computer Vision:** Built Deep Learning model for detecting Malignant Tumors on CT scans. Trained Resnet, VGG networks for Transfer Learning on an NIH dataset. Final algorithm implemented Region Proposal Network with IOU bounding box regression. [Link](#)
- **Natural Language Processing:** Developed two BiDAF-BERT Ensembles for Question Answering using 1) a rule based and 2) deep learning based classifier to reduce the energy in processing queries by bigger models. Saved between 58% and 75% energy at deployment time. [Link](#)

Skills

- **Programming** Python, C, C++, R, SQL, Keras, PyTorch, Tensorflow
- **Natural Language Processing** Transformer models, Attention, Seq2Seq, MT, SMT, BERT
- **Machine Learning** GMMs, Kernel methods, SVMs, Naive Bayes, PCA, CNN, RNNs, LSTM, Adam, Bias/Variance Reduction including Dropout, Early Stop, Regularization, BatchNorm, Data Augmentation using SMOTE
- **Convex Optimization:** LP Solvers Gurobi, Google ORTools, Interfaces CVXPY, LP, MILPs, Integer Programming, Combinatorial Programming, Convex Optimization
- **Databases** Understanding of Database system architecture, storage, query optimization, transaction management, fault recovery, and parallel processing of data-intensive systems

Other Experiences

- **CRO, Kudla Fund** *June 2019 - Jan 2020*
 - Entrusted with the responsibility as a Chief Risk Officer (CRO) of the Kudla Fund with AUM of \$1 Million, in ensuring Portfolio maintenance within laid framework
 - Recommend Portfolio exits based on Mark to Market Gains / Losses.
 - Created Python script to produce Portfolio Net Present Value, Sharpe Ratio, distribution of Daily Returns relative to benchmark using Trading-Calendars and Openpyxl packages
- **Summer Scholar, Quality Council of India** *June - August 2018*
 - The main objective of QCI is to encourage the implementation of quality standards in production methods.

- During one rotation, I reported on the responsibility of Quality Control in Transactions involving Coal that lead to Purchasers paying spot-prices
- In another, I evaluated the implementation of Swacch, a project to make India ODF (Open Defecation Free) and wrote a cost-benefit analysis to clean the river Ganges.

- **Analyst Intern, Omidyar Technology Ventures**

June - August 2015

- Omidyar Technology Ventures (OTV) is a Venture Capital firm investing in early stage technology companies.
- My tasks is to scope high-growth technology verticals and identifying early stage companies therein.

- **Analyst Intern at Edelweiss Financial Services**

June - August 2014

- Edelweiss, founded by a first-generation entrepreneur, is a well-diversified financial services firm.
- Attached to the Sales & Trading division, I supported the lead research analyst with required data points on the travel and tourism sector.

Interests

- **Awards**

- Represented India at the International Olympiad in Linguistics in 2012, winning an Honorable Mention.

- **Service**

- Perform Stand Up Comedy shows to Cancer Patients and their families waiting for treatment at Tata Memorial Cancer Hospital in Mumbai, India.