SUNNY DHAMNANI

@ dhamnanisunny.1402@gmail.com in linkedin.com/in/sdhamnani/

Education

GEORGIA INSTITUTE OF TECHNOLOGY

MS in Computer Science (specialization: Machine Learning)

Aug 2019 - Dec 2021 GPA - 4.0/4.0

- Teaching Assistant, CS6601 Artificial Intelligence, Georgia Tech.

INDIAN INSTITUTE OF TECHNOLOGY (IIT) KHARAGPUR

B. Tech. in Computer Science and Engineering (Honors)

July 2012 - June 2017 GPA - 9.50/10.0

Internship Experience

FACEBOOK

May – Aug 2021

Software Engineering Intern (ML) Seattle, WA

- Productionized multi-modal self-supervised learning framework for short-form video recommendation in Facebook NewsFeed. Implementation used distributed machine-learning stack at Facebook.
- Offline evaluation of the framework lead to 7% precision gains over current production model.

FACEBOOK

May – Aug 2020

Research Intern (Core Data Science) Menlo Park, CA

- Built an unsupervised anomaly detection framework in PyTorch for multivariate time-series data.
- Identified compromised accounts and detected page topic shifts.
- **The project was featured at a Facebook integrity event**.

CODENATION

May – July 2016

Software Engineering Intern Bengaluru, India

- Built static code analyzer for best software development practices in Java.
- Productionized correction modules to filter out false positives ($\sim 50\%$).
- **Awarded best intern. Received a full-time job offer.**

ADOBE RESEARCH

May – July 2015

Summer Research Intern Bengaluru, India

- Improvement factor of 3.15 in predicting the interests of a new visitor. Created a visualization tool using Django and JavaScript.
- **Published a research paper (ref: R4). Received a full-time job offer.**

Honors & Awards

- **Larry Heck Fellowship** at Georgia Institute of Technology, 2019.
- **Best Paper Award**, IEEE ICTS4eHealth, ISCC, 2018.
- India Rank 99, **ACM-ICPC** Asia Amritapuri Multisite Regionals Programming Contest, 2016.

Featured in the News

M1) **The Wall Street Journal** (link)

M2) **The Drum** (link)

M3) **Axios** (link)

Technical Skills

Proficient: C++, Python, PyTorch, SQL, Git

Comfortable: Java, Keras, TensorFlow, Spark, Django

Familiar: JavaScript, JQuery, PHP, OpenCV

Full-time Work Experience

ADOBE RESEARCH

June 2017 – Aug 2019

Research Scientist/Engineer Bengaluru, India

- **Published 4 research papers and filed 6 US patents.**
- Built bot-detection framework in PySpark using one-class learning. Improved classification accuracy by 15%. **(ref: R2, M1, M2, P1, P2)**
- Created a fast interpretable classification framework in Python and C++. Achieved 4-orders speed-up in training. **(ref: R4, M3, P3, P4, P5)**
- Devised a deep learning model in TensorFlow to estimate effect of marketing strategies. Outperformed causal inference baselines by 30%. **(ref: R3, P6)**
- Built a variant of GAN to produce synthetic data. Improved privacy-utility trade-off. **(ref: R1)**
- Head instructor of ML for over 200 engineers at Adobe. Mentored 8 summer interns (2018, 2019) at Adobe.

Research Papers

R1) HAR: Hardness Aware Reweighting for Imbalanced Datasets. IEEE Conference on Big Data (**BigData**), 2021. Publication-link

R2) Data-Sharing Economy: Value-Addition from Data meets Privacy. ACM International Conference on Web Search and Data Mining (**WSDM**), 2021. Publication-link

R3) Botcha: Detecting Malicious Non-Human Traffic in the Wild. In Workshop on Online Misinformation- and Harm-Aware Recommender Systems, (**OHARS@RecSys**), 2020. Publication-link

R4) Multiple Treatment Effect Estimation using Deep Generative Model with Task Embedding. The Web Conference (**WWW**), 2019. Publication-link

R5) RAPID: Rapid and Precise Interpretable Decision Sets. IEEE International Conference on Big Data (**BigData**), 2019. Publication-link

R6) **Best Paper Award.** Understanding Psycholinguistic Behavior of predominant drunk texters in Social Media. IEEE workshop on ICT solutions for ehealth, IEEE Symposium on Computers and Communications (**ISCC**), 2018. Publication-link

R7) Segmentation and Early Classification of Visitors Based on Reading Interests. ACM IKDD Conference on Data Science, (**CODS**), 2016. Publication-link

Patents

P1) Classification of website sessions using one-class labeling techniques. (15/793,001).

P2) Detecting robotic internet activity across domains utilizing one-class and domain adaptation machine-learning models. (15/982,393).

P3) Marketing simulations with perturbation based explanations. US Patent filed.

P4) Efficiently determining local machine learning model feature contributions. US Patent filed.

P5) Interpretable and fast multi-class classification. US Patent filed.

P6) What-if analysis for personalized targeting using a deep generative causal model. US Patent filed.