GOWTHAMI SOMEPALLI

College Park, Maryland, USA
voice: (+1) 475-228-4318 · *email:* gowthami@umd.edu
web: https://somepago.github.io

RESEARCH INTERESTS

My research is focused on understanding the failure modes of deep learning models, both discriminative and generative. Lately, my efforts have primarily centered on understanding multimodal generative models, particularly Text-to-Image Diffusion models and Video Language Models.

EDUCATION

University of Maryland, College Park Aug. 2019 - Present
M.S. + Ph.D. in Computer Science (4.0/4.0)
Kulkarni fellow, Amazon Internship Fellowship
Advisor: Tom Goldstein

Indian Institute of Technology Madras Aug. 2006 - Dec. 2011
Bachelors + Masters in Mechanical Engineering
Advisor: Raghu Prakash

SELECTED PUBLICATIONS

Visit my Google Scholar T2ezBDsAAAAJ for complete list

- Understanding and Mitigating Copying in Diffusion Models NeurIPS, 2023
  G. Somepalli, V. Singla, M. Goldblum, J. Geiping, T. Goldstein
- Baseline defenses for adversarial attacks against aligned language models web, 2023
  N. Jain, Y. Wen, **G. Somepalli**, J. Kirchenbauer, P.Y. Chiang, M. Goldblum, A. Saha, J. Geiping, T. Goldstein
- Diffusion Art or Digital Forgery? Investigating Data Replication in Diffusion Models CVPR, 2023
  G. Somepalli, V. Singla, M. Goldblum, J. Geiping, T. Goldstein *TechCrunch coverage - link*
- How much Data is Augmentation Worth? web, ICLR 2023
  J. Geiping, **G. Somepalli**, R. Shwartz-Ziv, A. G. Wilson, T. Goldstein, M. Goldblum
- Investigating Reproducibility and Double Descent from the DB Perspective web, CVPR (Oral) 2022
  G. Somepalli, L. Fowl, A. Bansal, P. Yeh-Chiang, Y. Dar, R. Baraniuk, M.GoldBlum, T. Goldstein
- SAINT: Improved NNs for Tabular Data via Row Attention and Contrastive PT web, NeurIPS TRLW 2022
  G. Somepalli, M. Goldblum, A. Schwarzschild, C.B. Bruss, T. Goldstein *300+ stars on Github*
- PatchGame: Learning to Signal Mid-level Patches in Referential Games web, NeurIPS 2021
  K. Gupta, **G. Somepalli**, Anubhav, V. Jayasundara, M. Zwicker, A. Shrivastava
- Unsupervised Anomaly Detection with Adversarial Mirrored AutoEncoders web, UAI (Oral) 2021
  G. Somepalli, Y. Wu, Y. Balaji, B. Vizumuri, S. Feizi
- FUGUE : Characterizing functional genes across human tissues web, PLOS CompBio 2021
  G. Somepalli, S. Sahoo, S. Hannenhalli
- Adversarial Training against Poisons and Backdoors (pre-print) web, 2021
  J. Geiping, L. Fowl, **G. Somepalli**, M. Goldblum, M. Moeller, T. Goldstein

WORK EXPERIENCE

Meta AI 2023 – Present
Research Internship *New York, NY*
- Long video understanding with LLMs.

Amazon AWS AI 2022 – 2022
Summer Internship *Pasadena, CA*
- Improving the reproducibility and consistency of deep learning models for object detection and classification tasks.

University of Maryland and National Cancer Institute 2019 – 2020
Student Researcher *College Park, MD*
- Worked with Prof. Sridhar Hannenhalli on understanding the tissue-level function of genes, synthetic lethality in the context of selective cancer therapy and cancer development using single-cell data.
- Developed a supervised machine learning model to rank genes in a specific tissue in terms of functionality. The model uses tissue-specific expression-derived and network-derived features.

- Worked on a model to predict trigger mutations in cancer development using single-cell data.

Flipkart 2017 - 2018
Business Manager - Sell side *Bengaluru, India*

- Led the demand planning and forecasting for the TV category; revamped the legacy data collection and modeling techniques. Reduced the overstocking of TVs by 15% and reduced the lost potential sales due to out-of-stock issues by 10% within my 6 months of joining.
- Won Annual **Business Excellence Award** for the most business growth via product innovation.

Poolka Technologies 2015-2017
Cofounder *Bengaluru, India*

Poolka provides scalable visual recognition APIs to developers & businesses.

- Built Fairi, a fashion assistant chatbot that provides fashion recommendations based on social media trends and users' existing wardrobe. Iterated & improved the product with 2000+ beta users.
- Selected for **Microsoft Bizspark** and **IBM Global Entrepreneur** Program. (>$20000 cloud credits per year)

General Electric, Oil & Gas 2013 - 2015
Engineer/ Synergy Leader *Bengaluru, India*

As a technologist in GE, I worked primarily on following two projects -

- Upgrading reciprocating compressors (RC) for Petrochemical plants - Led team that designed and validated various RC components. Awarded **Project of the Quarter** (Oct 2014- Dec 2014) and the automation work on Torsional Vibration Analysis won **Kaizen Award** (Mar 2015).
- GE-Cameron Synergy - Was single point of contact in Bangalore during the GE acquisition of Cameron. Performed Root Cause Analysis (RCA) for multiple Cameron products. Was accoladed by Senior management – GE Oil & Gas, Italy for my role in the synergy.

INVITED TALKS

- ICML GenLaw workshop - **Understanding Data Replication in Diffusion Models** July 2023
- Google US, ML Collective, TikTok US - **Understanding Data Replication in Diffusion Models** 2023
- WiML workshop @NeurIPS - **Investigating Reproducibility from the DB Perspective** Nov 2022
- G-Research, UK - **SAINT: Transformers for tabular data.** Nov 2022
- Google Research, India - **Can Neural Nets Learn the Same Model Twice?** Apr 2022
- TOPML Workshop - **Investigating Double Descent from the DB Perspective.** Mar 2022
- ML Seminar Series, UMD - **Transformers for Vision - a mini survey**. Oct 2021
- Headstart, Bengaluru - **Learnings from running a B2B tech startup in India**. 2017

SHORT PROJECTS

Predicting side-effects of drugs and drug combinations web, Spring 2020

Supervisor: Hector C. Bravo, (Prof. in Department of Computer Science, UMD)

- Built a self-attention based approach to predict Adverse Side Reactions (ADRs). The neural network model exploits know drug protein reactions and protein protein reactions and can be used to predict ADRs in the early stages of drug development. Improved SOTA approaches by 5% AUPRC.

SKILLS

Languages	Python, C, R, JavaScript, LaTeX
Libraries/Frameworks	PyTorch/Tensorflow, Docker
Platforms	MacOS/Linux

MISCELLANEOUS

- Kulkarni fellowship for summer research (web). 2021
- Reviewer - NeurIPS, ICLR, CVPR 2022-Present
- **Founder of @MLSummaries** twitter and medium accounts. 4000+ followers. 2021-Present
- Initiated student-led Machine Learning Seminar Series, UMD. 2021-Present
- **Mentor**: Tech + Research program in Technica. 2020, 2023
- Master's research Scholarship by Ministry of Human Resources and Development, Govt. of India. May 2010
- National Talent Search Exam Scholarship by Govt. of India. 2005-11
- Prathibha Award by Andhra Pradesh State Govt. 2003