Sachin Goyal

Ph.D. Student, Machine Learning Department
Carnegie Mellon University
Advisor: **Prof. Zico Kolter**

sachingo@andrew.cmu.edu | sach.goyalsachin@gmail.com
Webpage : https://saching007.github.io/
Github : www.github.com/SachinG007

EDUCATION

Carnegie Mellon University   *Pittsburgh, US*
Ph.D. in Machine Learning   *August '21 - Current*
Advisor: Prof. Zico Kolter

Indian Institute of Technology Bombay   *India*
B.Tech in Electrical Engineering with Minors in Computer Science and Engineering   *July '15 - August '19*
GPA: 9.11/10
Advisor: Prof. Subhasis Chaudhuri

WORK EXPERIENCE

Microsoft Research, India   *Bangalore*
Research Fellow in Machine Learning and Optimization Group   *July '19 - June '21*
Advisors: Dr. Prateek Jain and Dr. Harsha Vardhan Simhadri
Worked on resource-efficient and large-scale machine learning resulting in top-tier publication and Microsoft product impact.

RESEARCH INTERESTS

Domain Adaptation, Robust Fine-Tuning of Pretrained Models

PREPRINTS

Finetune like you pretrain: Improved finetuning of zero-shot vision models
Sachin Goyal, Ananya Kumar, Sankalp Garg, Zico Kolter, Aditi Raghunathan.
Under Review

PUBLICATIONS

MET : Masked Encoding for Tabular Data
Kushal Majmundar, **Sachin Goyal**, Praneeth Netrapalli and Prateek Jain.
Tabular Representation Learning Workshop @ **NeurIPS**, *2022* [Paper]

Test Time Adaptation via Conjugate Pseudo-Labels
Sachin Goyal*, Mingjie Sun*, Aditi Raghunathan and Zico Kolter.
*Advances in Neural Information Processing Systems (***NeurIPS***), 2022.* [Paper]

PAL: Pretext-based Active Learning.
Shubhang Bhatnagar, **Sachin Goyal**, Darshan Tank, Amit Sethi.
*British Machine Vision Conference (***BMVC), 2021***).* [Paper]

DROCC: Deep Robust One-Class Classification.
Sachin Goyal, Aditi Raghunathan, Moksh Jain, Harsha Simhadri and Prateek Jain.
*In International Conference on Machine Learning (***ICML), 2020***.* [Paper]

Indoor Distance Estimation using LSTMs over WLAN Network
Pranav Sankhe, Saqib Azim, **Sachin Goyal** , Tanya Choudhary, Kumar Appaiah , Sukumar Srikant
India Patent Application 201821047043, filed Dec' 2018. Patent Pending.
In IEEE Workshop on Positioning, Navigation and Communications (WPNC, 2019). [Paper]

Improving Self Super Resolution in Magnetic Resonance Images.
Sachin Goyal, Can Zhao, Amod Jog, Jerry L. Prince, Aaron Carass.
In **SPIE** *Conference on Medical Imaging and Biomedical Applications,* **2018**. [Paper]

Selected Projects

Robust fine-tuning of Image-Language Models

Advisor: Prof. Aditi Raghunathan and Prof. Zico Kolter   *June '22 - Ongoing,* **CMU**

Finetuning Image-Language models like CLIP give state-of-the-art accuracies on a variety of benchmarks. However, subtle differences in fine-tuning procedure lead to surprisingly large differences in in-distribution (ID) and out-of-distribution (OOD) accuracies. In our work, we show that simply using the same loss as pretraining, for fine-tuning the model on downstream labeled dataset outperforms alternative and complex fine-tuning approaches. Our work highlights that across various settings, using similar loss functions as pre-training leads to better adaptation to distribution shifts. ***(Under Review)***

Test Time Domain Adaptation

Advisor: Prof. Zico Kolter and Prof. Aditi Raghunathan   *Sept '21 - May '22,* **CMU**

Distribution shifts between training and test data is a pretty ubiquitous phenomenon. Most of the approaches for adapting to distribution shifts, require a prior access to the unlabeled data from the target domain, which is not always feasible. In our work, we develop a general class of unsupervised loss functions, which one can optimize at test-time, over the test samples, to adapt to the test distribution in an online fashion i.e. adapt while making predictions at the same time. ***(Published at NeurIPS '22)***

Low False Positive Rate (FPR) Classifiers for WakeWord Detection   [ICML 2020]

Advisors: Dr. Prateek Jain & Dr. Harsha Simhadri   *Sept '19 - May '20,* **Microsoft Research**

Wake-word systems (like Alexa and Cortana), activate when a particular "wake-word" is spoken. A low False Positive Rate (FPR) is a critical requirement in these systems. This is typically done by training a neural network on a corpus of both "wake-word" and "non-wake word" datasets, which is quite expensive. We proposed an effficient approach based on learning the "manifold" of wake-word samples, predicting inputs more than a threshold away from the manifold as a negative i.e. not a wake-word. This removes the requirement of training on a "non-wake word" corpus, significantly improving the efficiency. Our proposed approach, gave classifiers with upto 10% better recall at a fixed FPR across various data modalities like time-series data, tabular and image data.***(Published at ICML '20)***

EdgeML : Phonemes based KeyWord Spotting (KWS) on Resource Constrained Devices   [Project]

Advisors: Dr. Prateek Jain & Dr. Harsha Simhadri   *June '20 - May '21,* **Microsoft Research**

Edge Machine Learning aims to develop ML algorithms deployable on severely-resource constrained devices like microcontrollers, where the RAM and Flash are as low as **2MB**. We developed state-of-the-art **1mB** size phoneme prediction models, which were then used for various downstream tasks like keyword spotting on an ARM Cortex M7 processor. ***(EdgeML Official Microsoft Repository Link)***

Student Mentorship

Sankalp Garg : MS Student @ CMU MLD → Applied Scientist @ Amazon
Robust fine-tuning of image-language models.

Kushal Majmundar : Pre-doc Researcher @ Google Research India
Representation learning for unstrcutured tabular data.

Shubhang Bhatnagar : UG Student @ IIT Bombay → PhD Student @ UIUC
Active learning and efficient training of neural networks.

Miscellaneous

Service
- Conference Reviewing: PODS @ ICML' 22, SafeML @ NeurIPS' 22, AISTATS' 23
- Undergraduate Teaching Assistant : Biology 101   2017
- Hostel System Administrator , IIT Bombay   2019
- National Cadet Corps (NCC), Indian Air Force   '15-'17