Sanchit Sinha

ss7mu@virginia.edu | github.com/sanchit97 | sanchitsinha.com |
scholar.google.com/citations?user=squ4_6IAAAAJ&hl=en

EDUCATION

University of Virginia   Charlottesville, Virginia
Doctor of Philosophy (Ph.D.) *in Computer Science*   05/2021 - 12/2025 (expected)
Advised by Dr. Aidong Zhang - improving interpretability, explainability, adversarial robustness, and concept extraction.
Master of Science (M.S) in Computer Science   GPA: 4.0/4.0   08/2019 - 05/2021
Elective Courses: Advanced Deep Learning, Machine Learning, Data Mining, NLP, Manifold Analysis, Graph Mining
IIIT-Delhi   New Delhi, India
Bachelor of Technology in Computer Science with Honors   GPA: 8.28/10   08/2015 - 05/2019
Elective Courses: Advanced ML, Artificial Intelligence, Parallel Programming, Advanced Algos, Collab Filtering, Biometrics

WORK EXPERIENCE

Amazon AGI   Cambridge, MA, USA
Applied Scientist Intern   05/2023 − 08/2023
- Improving warmup approache for improved in-context learning performance using second-order meta-learning approaches
- Beating standard meta-training approaches by a baseline minimum of 3%, a challenging feat not discussed before
- Seminal work on understanding dual optimization landscape of LLMs. Formalized insights on task selection.

Amazon Web Services (AWS), Amazon   Sunnyvale, CA, USA
Applied Scientist Intern, AWS Lex   05/2022 − 08/2022
- Implemented parameter efficient self-supervised accent domain adaptation on large speech models (HuBERT) using adapters
- Demonstrated improved performance on downstream speech tasks using general fine-tuning data by minimum 5%
- Improved generic accent information learned by large speech models without explicit labeling - reducing manual annotation

Unity Technologies (Unity 3D)   Seattle, WA, USA
ML-Computer Vision Intern, AI@Unity   05/2020 − 08/2020
- Implemented a real time video object tracking segmentation model with benchmark performance on public leaderboards
- Containerized deployment on GCP/AWS with ETL functionality, robust fine-tuning and scalable pipelining (Kubeflow)
- Designed multi-domain (including synthetic data) training algorithms (domain randomization) for better generalizability

FFmpeg - Google Summer of Code, 2017   Remote
Student Developer   05/2017 − 08/2017
- Nominated in a highly selective student open source developer program hosted by Google (code on Github profile)
- Designed/implemented audio processing decoder for Ambisonic AR-sound files to custom speaker array configuration

Publications - Best viewed in Google Scholar

- CoLiDR: Concept Learning using Aggregated Disentangled Representations   KDD, 2024
- [2]MAML-en-LLM: Model Agnostic Meta-Training of LLMs for Improved In-Context Learning   KDD, 2024
- [3]A Self-explaining Neural Architecture for Generalizable Concept Learning   IJCAI, 2024
- [4]Enhance Diffusion to Improve Robust Generalization   KDD, 2023
- [5]Don't stop self-supervision: Accent adaptation of speech representations via residual adapters   Interspeech, 2023
- [6]Understanding and Enhancing Robustness of Concept Models   AAAI, 2023
- [7] Perturbing Inputs for Fragile Interpretations in Deep NLP   EMNLP-Blackbox, 2021
- [8] Triplet Transform Learning for primate face recognition   IEEE ICIP, 2019
- [9]Exploring Bias in primate face detection and recognition   ECCV-W, 2018
- [10]Video Summarization using Global Attention with Memory Networks and LSTM   IEEE BigMM, 2019

PROJECT WORK

Improving Attention-based Concept Alignment with Scale-aware Representations *Under review, AAAI '24*
Proposing an improved architecture combining CNN and ViT's scale and patch representations for better concept alignment
Neurosymbolic Concept-based Reasoning with LLMs *Under review, AAAI '24*
Using LLM-Agents to extract, ground and compose concepts into neurosymbolic entities for better explainability and predictions. Seminal work linking grounding of concepts and neurosymbolic reasoning
Advancing Additive Models with Mixture of Experts (MoEs) *Under review, NeurIPS '24*
Utilizing Mixture of Experts as a tool to combine additive model features and model interactions improving performance
Unsupervised Image to Image Translation using GANs Add semi-supervision in unsupervised (CycleGAN) to obtain super-linear increase in performance wrt supervised methods

AWARDS

Student Travel Award - KDD 2024, AAAI 2023. **Amazon Conference Grant** - 2024
Reviewer - NeurIPS, ICML, ICLR, KDD, EMNLP (2022-present)
School of Engineering and Applied Science - PhD Fellowship 2021-22