Sanchit Sinha

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

About Me:

PhD student in Computer Science with interest in local and global explainability, trustworthiness and robustness of DNNs

EDUCATION

University of Virginia Charlottesville, Virginia

Doctor of Philosophy (Ph.D.) *in Computer Science* 05/2021 - 05/2025 (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 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

Biocomplexity Institute and Initiative, UVA Charlottesville, Virginia

Reasearch Assistant 08/2019 – 05/2021

- Worked on the simulation and modeling of global commodity trade networks using spatio-temporal approaches
- Focused on food supply networks with use-cases in the form of tomato production in Senegal and Nepal. [5]

Western Digital Bangalore, India

SWE Intern, Client Storage Solutions - Validation 05/2018 – 08/2018

- Automate firmware installation on hundreds of SSDs in parallel - Awarded “Think Big (Automation Initiatives)” Award.
- Automation effort for automatic test curation using ML awarded the most “Out-of- Box Idea Award”.

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

PROJECT WORK

Understanding and Enhancing Robustness of Concept Models *Under Review at AAAI ’23* 01/2022 – 05/2022

Proposed and analyzed 3 novel attack algorithms and defense strategies for adversarial attacks on concept-based models

Self-Supervised Concept Extraction as priors for SENN *Under Review* 06/2022 – present

Propose a VAE-focussed robust concept extraction framework to initialize SENN models to learn generilizable robust concepts.

Perturbing Inputs for Fragile Interpretations in NLP *(Master’s Thesis)* 01/2020 – 05/2021

- Disrupt performances of DNN explainability methods like Integrated Gradients and LIME on deep NLP models.[1]

Automated Face Detection and Recognition in Primates *(Bachelor’s Thesis)* 01/2018 – 05/2019

- Designing an end to end biometric system to detect and recognize primate faces in the wild performing data cleaning, detection in the wild, face normalization, face alignment and facial recognition using a novel triplet loss algorithm[3,4]

Automated Video Summarization 01/2019 – 05/2019

- Global attention/LSTM based video summarization algorithm improving performance by 5% on standard benchmarks[5]

Unsupervised Image to Image Translation using GANs 01/2019 – 05/2019

- Add semi-supervision in unsupervised (CycleGAN) to obtain super-linear increase in performance wrt supervised methods

Publications - Best viewed in Google Scholar

- [1] Perturbing Inputs for Fragile Interpretations in Deep NLP EMNLP-Blackbox, 2021
- [2] Realistic Commodity Flow Networks to Assess Vulnerability of Food Systems Complex Networks, 2021
- [3] Triplet Transform Learning for primate face recognition IEEE ICIP, 2019
- [4]Exploring Bias in primate face detection and recognition ECCV-W, 2018
- [5]Video Summarization using Global Attention with Memory Networks and LSTM IEEE BigMM, 2019