Sanjay Kariyappa

Last Updated on 14th November 2024

https://sanjaykariyappa.github.io
sanjaykariyappa@gmail.com | 678.650.5017 | Mountain View, CA

SUMMARY

Sr. AI Research Scientist at Nvidia, working on AI Security and Privacy

EDUCATION

GEORGIA TECH
PhD in ECE
Dec 2022 | Atlanta, GA
GPA: 4.0 / 4.0

GEORGIA TECH
MS in ECE
Dec 2014 | Atlanta, GA
GPA: 4.0 / 4.0

SRI JAYACHAMARAJENDRA COLLEGE OF ENGINEERING
BS in ECE
June 2013 | Mysore, India
GPA: 3.78 / 4.0

LINKS

Github:// **sanjaykariyappa**
LinkedIn:// **sanjay-kariyappa**
Twitter:// **@sanjayatwork**

RESEARCH INTERESTS

Machine learning, deep learning, privacy, security, federated learning, explainable AI, uncertainty estimation, semi-supervised learning, computer architecture, ML accelerators

COURSEWORK

Statistical Machine Learning
Digital Image Processing
Advanced Computer Architecture
ML Hardware Acceleration
Advanced Memory Systems

SKILLS

Programming Languages:
• Python • C • C++
• Matlab • Latex
Software Libraries:
Pytorch • Tensorflow • Keras
• Pandas • Numpy

PUBLICATIONS

Progressive Inference: Explaining Decoder-Only Sequence Classification Models Using Intermediate Predictions
[ICML 2024] S Kariyappa, F. Lécué, S. Mishra, C. Pond, D. Magazzeni, M. Veloso

SHAP@k: Efficient and PAC Identification of Top-k Features
[AAAI 2024 - Oral] S. Kariyappa, L. Tsepenekas, F. Lécué, D. Magazzeni

Cocktail Party Attack: Breaking Aggregation-Based Privacy in Federated Learning using Independent Component Analysis
[ICML 2023] S. Kariyappa, C. Guo, K. Maeng, W. Xiong, Ed Suh, M. K. Qureshi, H. S. Lee

ExPLoit: Extracting Private Labels in Split Learning
[SaTML 2023] S. Kariyappa, M. K. Qureshi

Bounding the Invertibility of Privacy-preserving Instance Encoding using Fisher Information [NeurIPS 2023] K. Maeng, C. Guo, S. Kariyappa, Ed Suh

MAZE: Data-Free Model Stealing Attack Using Zeroth-Order Gradient Estimation
[CVPR 2021] S. Kariyappa, A. Prakash, M. K. Qureshi

Tolerating Noise in PCM-Based AI Accelerators via Noise-Aware Training
[IEEE Transactions on Electron Devices 2021] S. Kariyappa et al.

Protecting DNNs from Theft using an Ensemble of Diverse Models
[ICLR 2021] S. Kariyappa, A. Prakash, M. K. Qureshi

Defending Against Model Stealing Attacks with Adaptive Misinformation
[CVPR 2020] S. Kariyappa, M. K. Qureshi

PrivRecourse: Generating Realistic and Privacy-Preserving Recourse Paths
[XAI-FIN 2023] S. Pentyala, S. Sharma, S. Kariyappa, F. Lécué, D. Magazzeni

Improving Adversarial Robustness of Ensembles with Diversity Training
S. Kariyappa, M. K Qureshi

WORK EXPERIENCE

NVIDIA | Sr. AI Research Scientist
Nov 2024 – present | Santa Clara, CA

JP MORGAN CHASE | Sr. AI Research Associate
Feb 2023 – Nov 2024 | Palo Alto, CA

META | AI Research Intern (FAIR)
May 2022 - Aug 2022 | Boston, MA

- Developed a novel attack on federated learning to break aggregation based privacy using independent component analysis. **(paper)**

FACEBOOK | Software Engineering Intern
May 2021 - Aug 2021, May 2020 - Aug 2020 | Menlo Park, CA

- Explored the use of semi-supervised learning techniques to improve conversion prediction models for online advertising.

IBM | Research Intern
May 2019 – Aug 2019 | San Jose, CA

- Developed Noise-Resilient DNNs that are robust against hardware noise for PCM-based analog AI hardware. **(paper)**