

YASH JAIN



yashjain@gatech.edu



404.429.4813



jinga-lala



yash-jain



bitshots

EDUCATION

Georgia Tech, ATL, USA
Masters of Science in Computer Science
GPA: 4.0 / 4.0 | Aug.2021-May.2023

IIT Bombay, India
B.Tech. Computer Science (With Honors)
GPA : 8.91 / 10 | Jul.2017-Jul.2021

KEY COURSEWORK

Graduate Courses:
ML with Limited Supervision
Machine Learning
Introduction to Graduate Algorithms

Undergraduate Courses:
Natural Language Processing
Organization of Web Information
Artificial Intelligence
Adv. Methods in Satellite Image Proc.
Machine Learning for Remote Sensing
Operating Systems
Data Structures and Algorithms

TECHNICAL SKILLS

Proficient-
• Python • C++ • Tensorflow • Pytorch
• Keras • MATLAB • SQL • GIT • \LaTeX

Familiar-
• Java • Bash • Spark • MapReduce

AWARDS & FELLOWSHIPS

- Research Excellence Award for B.Tech Thesis Project, IITB CS, 2021
- Recipient of Dhirubai Ambani Scholarship for supporting my Master's studies at GaTech, 2021
- Summer Internship Fellowship, Aalto University, Finland (2021)
- All India Rank 29, JEE Advanced 2017
- Gold & Silver Medal representing India- IJSO, Argentina 2014

PUBLICATIONS

- Under review ColloSSL: Collaborative Self-Supervised Learning for HAR
[Yash Jain*](#), Chi Ian Tang*, Chulhong Min, Fahim Kawsar, Akhil Mathur
- CIKM'21 Integrating Transductive And Inductive Embeddings
Improves Link Prediction Accuracy
Chitrang Gupta*, [Yash Jain*](#), Abir De, Soumen Chakrabarti
- ICML'21 Group Supervised Learning: Extending Self-Supervised Learning to Multi-Device Settings
[Yash Jain*](#), Chi Ian Tang*, Chulhong Min, Fahim Kawsar, Akhil Mathur
Workshop on Self-Supervised Learning for Reasoning and Perception
- UbiComp'20 RFID Tattoo: A wireless platform for speech recognition
Jingxian Wang, Chengfeng Pan, Haojian Jin, Vaibhav Singh, [Yash Jain](#), Jason I. Hong, Carmel Majidi, Swarun Kumar
UbiComp 2020 Best Paper Award, U.S. Patent Pending
IJCAI 2021 Sister Conferences Best Papers

WORK EXPERIENCE

- May-Aug'21 Nokia Bell Labs, UK | Research Intern | Dr. Akhil Mathur
Group Supervised Learning: Extending SSL to Multi-Device Settings
• Formulated a novel framework, Group Supervised Learning (GSL), which utilizes synchronous multi-device unsupervised data, extending the principles of contrastive learning to a group setting.
• Outperformed supervised and semi-supervised baselines by 0.15 in F-1 score in RealWorld dataset.
- May-July'20 Flipkart | Data Science Intern | Nikesh Garera & Nithish Pai
Automated E-commerce Question-Answering system
• Generated synthetic queries from a limited set of user query to increase the dataset size by more than 30%.
• Combined BERT and GPT-2 models for developing a target product-type classification system which would then prompt the text-generation model to answer user query in natural language
- May-Jul'19 Carnegie Mellon University | Research Intern | Prof. Swarun Kumar
RFID Tattoo: A wireless platform for speech recognition
• Speech recognition platform for voice impairments through wafer-thin, battery-free and stretchable RFID Tattoo.
• Collected own sensor data and implemented Random Forest model calibrated on the stretch of tags to achieve state of the art 86% accuracy on a vocabulary size of 100 most common English words

RESEARCH PROJECTS

- Present Generating virtual IMU data for effective activity recognition systems with a focus on eating Prof. Thomas Ploetz
Working on improving IMUTube for developing a DL based food-monitoring system using smartwatch data
- Present Group contrastive learning with noisy labels Prof. Judy Hoffman
Integrating the ideas of group contrastive learning with noisy labels for downstream object recognition tasks
- Dec-Apr'21 Integrating Transductive And Inductive Embeddings Improves Link Prediction Accuracy Prof. Soumen Chakrabarti & Prof. Abir De
Provided alternative to node features in Online Social Network (OSN) graphs using transductive embeddings, protecting user privacy while improving link prediction performance in GNNs
- Aug-Dec'20 Meta Self-learning with Noisy Student Prof. Biplab Banerjee
While training large networks using MAML is expensive, our proposed method allows for training of large student networks using few-shot pseudo labels which outperforms the teacher learnt using MAML in fewer epochs