

YASH JAIN



yashjain@gatech.edu



404.429.4813



jinga-lala



yash-jain



bitshots

EDUCATION

Georgia Institute of Technology
MS in Computer Science
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:

Machine Learning with Limited Supervision
Machine Learning
Introduction to Graduate Algorithms

Undergraduate Courses:

Natural Language Processing
Organization of Web Information
Artificial Intelligence
Advanced Methods in Satellite Image Processing
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, IIT Bombay CS, 2021

Summer Internship Fellowship, Aalto University, Finland (2021) & Carnegie Mellon University, USA (2019)

All India Rank 29, JEE Advanced 2017

All India Rank 8, KVPY Fellowship- Dept. of Science & Technology, Govt. of India, 2016

Gold & Silver Medal representing India- IJSO, Argentina 2014

INTERESTS

Ubiquitous Computing, Self-supervised learning, Natural Language Processing, Human Centric Machine Learning, Human Computer Interaction, Graphical Neural Nets

PUBLICATIONS

- Submitted* Group Supervised Learning for Human Activity Recognition
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
Group Supervised Learning: Extending SSL to Multi-Device Settings
Proposed a new problem setting, Time-synchronous multi-device systems (TSMDS). Formulated a novel framework, Group Supervised Learning (GSL), addressing the TSMDS problem, utilizing the principles of contrastive learning in a group setting. Outperformed supervised and semi-supervised approaches by 0.15 in F-1 score.
- May-July'20 Flipkart | Data Science Intern
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 all in real-time.
- May-Jul'19 Carnegie Mellon University, USA | Research Intern
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

- 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
- Jun-May'20 Answer-type Inference in QA systems
Prof. Soumen Chakrabarti
Outperformed an existing Answer-Type Inference RNN based manuscript system by almost 20% across 4 Question-Answering datasets by deploying BERT architecture
- 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