Yash Jain

jinga-lala.github.io | yash.jain3599@gmail.com | bitshots.github.io

EDUCATION

Indian Institute of Technology Bombay   Mumbai, India
Bachelor of Technology in Computer Science with Honors   *July 2017 – August 2021*
Cum. GPA: 8.91/ 10.0

Relevant Courses: Artificial Intelligence & Machine Learning, Organisation of Web Information, Speech, Web & Natural Language Processing, Bachelors Thesis Project-1&2, ML for Remote Sensing-II

Advisors: Prof. Soumen Chakraborty and Prof. Abir De

PUBLICATION

[1] Y. Jain, C. I. Tang, C. Min, F. Kawsar, and A. Mathur. Group Supervised Learning: Extending Self-Supervised Learning to Multi-Device Settings. *Workshop ICML: Self-Supervised Learning for Reasoning and Perception*, 2021.

[2] J. Wang, C. Pan, H. Jin, V. Singh, Y. Jain, J. I. Hong, C. Majidi, and S. Kumar. RFID Tattoo: A Wireless Platform for Speech Recognition. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 3(4), Dec. 2019.

EXPERIENCE

Group Supervised Learning: Extending SSL to Multi-Device Settings [paper] [poster]   Summer 2021
Guide: Akhil Mathur   *Nokia Bell Labs, UK*
Workshop Poster presentation   *ICML, 2021*
- Proposed a new problem setting, Time-synchronous multi-device systems (TSMDS), which exists in many domains in the real world but has not been thoroughly explored yet
- Formulated a novel framework, Group Supervised Learning (GSL), addressing the TSMDS problem, utilizing the principles of contrastive learning in a group setting
- Early results demonstrate that GSL outperforms supervised and semi-supervised training baselines proposed in the HAR literature by as high as 0.15 in F-1 score.

Automated E-commerce Question-Answering system   Summer 2020
Guide: Nikesh Garera & Nithish Pai   *Flipkart, India*
- 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
- System might be implemented to the company's website, subject to quality check approval

RFID Tattoo: A wireless platform for speech recognition [paper]   Summer 2019
Guide: Prof. Swarun Kumar   *Carnegie Mellon University, USA*
Best long paper award *(1 out of 121 papers)* | **U.S. Patent pending**   *UbiComp, 2020*
- Speech recognition platform for voice impairments through wafer-thin, battery-free and stretchable RFID Tattoos
- 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

Integrating Transductive And Inductive Embeddings Improves Link Prediction Accuracy   Spring 2021
Guide: Prof. Soumen Chakrabarti & Prof. Abir De | Bachelor's Thesis Project   *InfoLab, IIT Bombay*
Paper under double-blind review
- Combined transductive Node2Vec and inductive GNNs embeddings in a GNN to surpass SOTA in link prediction
- Provided alternative to node features in Online Social Network (OSN) graphs using transductive embeddings, protecting user privacy while improving link prediction performance

Deep Neural Matching Models for Graph Retrieval from Product Graph   Autumn 2020
Guide: Prof. Abir De & Prof. Soumen Chakrabarti | Bachelor's Thesis Project-2   *InfoLab, IIT Bombay*
- Developing a novel deep neural network approach for searching a subgraph from a corpus of sample graphs
- Working on decreasing the query time to sub-linear time by employing LSH hashing of corpus graphs

Answer-type Inference in QA systems [report] [slides]   Spring 2020
Guide: Prof. Soumen Chakrabarti | Bachelor's Thesis Project-1   *InfoLab, IIT Bombay*
- Outperformed an existing Answer-Type Inference RNN based manuscript system by almost 20% across 4 Question-Answering datasets by deploying BERT architecture

Awards

- Nominated for **Undergraduate Research Award** for exceptional work in Bachelors Thesis at IITB   *2021*
- Selected in **Aalto Science Institute internship programme** at Aalto University, Finland   *2021*
- Selected in **Summer Undergraduate Internship Program** at Carnegie Mellon University   *2019*
- **Gold Medalist** in theory and **Silver Medalist** in practicals representing India at 11th International Junior Science Olympiad held in Mendoza, Argentina   *2014*
- **All India Rank 8** in **KVPY Fellowship** by Dept. of Science & Technology, Govt. of India   *2016*
- **All India Rank 29** in **JEE-Advanced** out of 220,000 shortlisted candidates from 1.2 million students   *2017*

Key course Projects

Meta Self-learning with Noisy Student   Autumn 2020
Guide: Prof. Biplab Banerjee | Machine Learning for Remote Sensing-II course   *IIT Bombay*
- Combined the benefits of self-training with noisy student and MAML for few-shot image classification
- 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

Adversarial Reprogramming of Neural Networks   Autumn 2019
Guide: Prof. Ajit Rajwade & Prof. Suyash Awate | Digital Image Processing course   *IIT Bombay*
- Reprogrammed ImageNet classification model - ResNet50v2 on MNIST dataset to perform an adversarial task by finding a single constant learnable weight matrix, added to all test inputs; achieved an accuracy of 80%
- Illustrated the vulnerability in deep neural networks that perform a task chosen by the adversary despite not being trained to do this task originally

Comparing Extractive and Abstractive Text Summarisation   Autumn 2019
Guide: Prof. Ganesh Ramakrishnan | Aritificial Intelligence & Machine Learning course   *IIT Bombay*
- Implemented an extractive text summarizer using the page rank algorithm, achieved a ROGUE score of f-32%
- Comparatively, the abstractive summarizer built using seq2seq model with attention mechanism, achieved a ROUGE scores of only f-13%
- Latent Semantic analysis of the generated summary and the ground truth revealed an overlap of 79%

Quarter to Sixth Sense   Summer 2018
Institute Technical Summer Project | Electronics and Robotics Club   *IIT Bombay*
- Remodeled a wearable gesture interface by programming color tracking algorithms that perform functions in correspondence with hand gestures by recognizing the color markers on fingers

Position of Responsibility

Technical Head | *NGO We Listen, India* 🔗   Summer 2020 - Summer 2021
- Leading a team of interns in the development of a professional website for increasing mental health awareness
- Engineering a browser extension to automatically detect hate speech content on social media platforms while synchronously hiding it before the purview of the user, creating a virtual hate-free space

Undergraduate Teaching Assistant | *IIT Bombay*   Spring 2021, 2020 & Autumn 2018
Introduction to Machine Learning   *Spring 2021*
- Created teaching material and supervised exams during the online semester for 200+ students

Computer Programming & Utilization   *Spring 2020*
- Responsible for setting algorithmic coding questions for graded and ungraded labs for 500+ freshmen

Quantum Physics & its Applications   *Autumn 2018*
- Conducted weekly course tutorials for a batch of 50 freshmen and evaluated their term papers

Academic Resource Facilitator | *International Junior Science Olympiad cell (IJSO), HBCSE*   Summer 2018
- Student facilitator for mentoring top 35 students from India in IJSO Selection Camp 2018

Technical Skills

Strong: Python (with Tensorflow and PyTorch), C/C++, Bash, MATLAB
Familiar: Java, SQL (Postgres), JavaScript, HTML/CSS, Android, Arduino
Tools: Git, Google Cloud Platform, LaTeX, SolidWorks