Megha Thukral

mthukral3@gatech.edu www.linkedin.com/in/megha-thukral/ +1 (470)815-1982

Education

Georgia Institute of Technology Atlanta, USA
PhD Machine Learning - School of Interactive Computing *GPA 4/4* *Aug 2023-Present*

Coursework: Mathematical foundations of machine learning

Georgia Institute of Technology Atlanta, USA
Masters in Computer Science *GPA 4/4* *Aug 2021-May 2023*

Coursework: Deep Learning, Artificial Intelligence, Machine Learning, Computer Vision, Ubiquitous Computing

Research

Cross-Domain HAR: Few Shot Transfer Learning for Human Activity Recognition *-under submission at ACM TIST*

- Achieved significant few shot classification performance improvement over end-to-end training (by 10-15%), self-supervised learning (by 5-8%), and naive transfer (by 2-6%) across domains for activity recognition systems.
- Implemented a semi-supervised learning methodology combining self-training, consistency regularization with self-supervision loss to adapt representations learnt from source domain to a low resource domain.
- Evaluated and did extensive experimentation for Few Shot Classification in the target domains and analyzed target conditions for successful transfer.

How Much Unlabeled Data is Really Needed for Effective Self-Supervised Human Activity Recognition? *-accepted at ACM ISWC*

- Focused on the pre-training data efficiency of self-supervised methods to reduce computational costs and guide data collection practices.
- Investigated three established SSL methods (Autoencoders, Contrastive Predictive Coding, SimCLR) for HAR and three publicly available datasets (HHAR, PAMAP2, RealWorld).
- Found that Contrastive Predictive Coding (CPC) is the most data-efficient method, requiring as little as 15 minutes of sensor data to achieve competitive activity recognition performance.

Experience

Georgia Institute of Technology Atlanta, USA
Graduate Research Assistant *Aug 2021 -Present*

- Contributing to the NSF AI Caring project, I am developing foundational AI techniques for longitudinal behavioral tracking and change in routine detection in the elder population.

Bloomberg LP - AI Group New York, USA
Software Engineer(ML) Intern *May 2022 - July 2022*

- Built and deployed an FAQ Retrieval Model to handle repeated queries for ML Platform's chatbot. Used pretrained Sentence BERT model to implement Question-Query similarity algorithm.
- Tested, bench-marked quality of model and published the model by containerizing it using Bloomberg's internal MLOps platform built on top of a Kubernetes cluster.
- Automated the train-deploy-infer ML workflow using Argo Workflows which assisted other teams to build and onboard their bots with minimal effort.
- Tools: PyTorch, scikit-learn, KServe, Argo, Kubernetes, KServe, CI/CD Jenkins, Python unit testing

Projects

Transfer learning for Plankton Image Classification Atlanta, USA
Course Instructor: Dr. Danfei Xu *Aug 2022 - Dec 2022*

- Implemented a self-training-based transfer learning approach to create classification models for plankton images, effectively handling variations in imaging systems and limited labeled data resources.
- Achieved 5 way-5 shots performance surpassing few-shot supervised baseline by 10 to 20% and naive transfer by 1 to 2% for target imaging datasets

Where to put it? Best on-body IMU sensor placement Atlanta, USA
Course Instructor: Dr. Thomas Ploetz *Jan 2022 - May 2022*

- Conducted a semester-long project focused on optimizing the placement of on-body IMU sensors for enhanced human activity recognition.
- Performed comprehensive data collection (7 users for 6 activities), annotation, segmentation, and performance analysis, ultimately identifying the hip as the optimal sensor placement for the specified activities and compared our results with RealWorld dataset.

Visual Question Answering Atlanta, USA
Course Instructor: Dr. Mahdi Roozbahani *Sept 2021 - Dec 2021*

- Created a visual question answering system employing a feedforward neural network.
- Utilized a combination of visual features extracted from a pre-trained InceptionV3 backbone and text features from both bag-of-words and a BERT pretrained model. Employed k-means clustering to refine the dataset, retaining pertinent images, and visualized the outcomes with t-SNE plots

Scholastic Achievements

- Secured **Meritorious Student Scholarship** from undergrad institute Punjabi University, Patiala (2012 to 2015)
- Achieved **All India Rank of 62** among 100000 candidates in GATE, CS - National Level Graduate Entrance Exam (2015)

Skills

- **Programming/ML modelling:** Python, Pytests, Deep Learning, Scikit-learn, pandas,
- **Tools:** Pytorch, Keras, , Argo, VS Code, Git, Jupyter Scikit-learn, KServe, Kubernetes, SQL Developer, Anaconda