

Utkarsh Tyagi

☎ +1 240 501 3163 ✉ utkarsht@umd.edu 🌐 utkarsh4430.github.io 🔗 utkarsh4430 📄 utkarsh4430

EDUCATION

University of Maryland, College Park

08/2023 – 05/2025

M.S. in Computer Science, Advised by Prof. Dinesh Manocha - GPA 3.8/4.0

Maryland, USA

Delhi Technological University (Formerly Delhi College of Engineering)

2017 – 2021

Bachelor of Technology in Computer Science - GPA 8.7/10, Major GPA 9.05/10

Delhi, India

PUBLICATIONS

- [CompA: Addressing the Gap in Compositional Reasoning in Audio-Language Models](#)
ICLR 2024
- [AdVerb: Visually Guided Audio Dereverberation](#)
ICCV 2023
- [MMER: Multimodal Multi-task Learning for Speech Emotion Recognition](#)
InterSpeech 2023
- [ACLM: Selective-Denoising based Data Augmentation for Low-Resource Complex NER](#)
ACL 2023
- [BioAug: Conditional Generation based Data Augmentation for Low-Resource Biomedical NER](#)
SIGIR 2023
- [CoSyn: Detecting Implicit Hate Speech in Online Conversations Using a Context Synergized Hyperbolic Network](#)
EMNLP 2023
- [DALE: Generative Data Augmentation for Low-Resource Legal NLP](#)
EMNLP 2023
- [ABEX: Data Augmentation for Low-Resource NLU via Expanding Abstract Descriptions](#)
Under review at **ACL 2024**
- [ASPIRE: Language-Guided Data Augmentation for Improving Robustness Against Spurious Correlations](#)
Under review at **ACL 2024**
- [CoDa: Constrained Generation based Data Augmentation for Low-Resource NLP](#)
Under review at **NAACL 2024**
- [Do Vision-Language Models Understand Compound Nouns?](#)
Under review at **NAACL 2024**

RESEARCH EXPERIENCE

GAMMA Lab, University of Maryland

College Park, Maryland

Research Assistant

August 2022 – Present

- Currently researching multimodal learning; working on video representation learning for long-form video understanding
- My research focuses on low-resource (labeled data and compute) learning with applications in speech, NLP, or vision. In this area, I solve problems using self-supervised learning, synthetic data augmentation, etc.
- Published at **ICLR, ICCV, ACL, EMNLP, InterSpeech, SIGIR**
- Advised by [Prof. Dinesh Manocha](#)

Multimodal Digital Media Analysis Lab

IIITD, Delhi

Machine Learning Researcher

April 2022 – March 2023

- Worked under the supervision of **Dr. Rajiv Ratn Shah** in the areas of Speech and Language Processing.
 - Worked on multi-lingual automatic speech scoring systems for low-resource Indian languages.
 - Also explored novel architectures to detect implicit hate speech in online conversations. Paper accepted at **AAAI 2023**
- Defactify**

INDUSTRY EXPERIENCE

Atlassian India LLP

Bangalore, Karnataka

Software Development Engineer 2

July 2021 – August 2023

- Worked on the Atlassian's Issue Create Experience to deliver a reliable and performant interface which helped improve key business metrics and customer satisfaction ratings.
- Achieved **99.996% frontend reliability** and scaled the experience for over 250k customers with more than 2M issues created every day. Improved performance TTI from 3.9s to **1.2s** using modern technology stack.

Software Engineer Intern

Jan 2021 – June 2021

- Identified the limitations for improvements and bottlenecks in the existing legacy API implementation and proposed revamping. Utilized this performant API to develop the user interface.
- Reduced the **latency** of modal load by **95.57%** by avoiding redundant server-side database queries and scaled to 1M customers.

Samsung R&D Institute Bangalore

Bangalore, Karnataka

Machine Learning Intern

May 2020 – July 2020

- Worked with the **Voice Intelligence** team of Samsung to improve Bixby's wakeup word detection.
- Researched keyword spotting techniques and built an efficient on-device LSTM-CTC based, vocabulary-independent keyword spotter

PATENTS

- Apparatuses, methods, and computer program products for generating an abstractive context summary scheduling interface configured for scheduling and outputting abstractive context summaries for multi-party communication channels (Patent Pending)
U.S. Patent Application No. 17/936,695
- Apparatuses, methods, and computer program products for generating and selectively outputting abstractive context summaries for multi-party communication channels (Patent Pending)
U.S. Patent Application No. 17/936,705

TECHNICAL SKILLS

Languages: (Highly Proficient) Python, JavaScript, C++ (Moderate) Java, SQL

Frameworks: Pytorch, FastAI, Tensorflow, Scikit-Learn, ReactJS, Jest, Enzyme, NLTK

Certifications and Training:

- Machine Learning by Stanford University
- Competitive Programming by St Petersburg University
- Game Theory by Stanford University
- Data Analysis with Python by IBM

PROJECTS

Occluded Facial Expression Recognition | *Deep Learning, Image Processing*

- Developed a framework for recognizing facial expressions in occluded images using non-occluded images as privileged information
- The technique rendered an average gain of 3.90% over the baseline for 3 standard benchmarking datasets

ANN-GWO Intrusion Detection System | *Neural Networks, Swarm Algorithms*

- Developed a hybrid IDS by using Grey Wolf Algorithm instead of backpropagation with artificial neural networks
- Utilized MIT Darpa 1998 dataset and achieved SOTA results

ACHIEVEMENTS

- **EMNLP 2023** - Industry Track Reviewer
- **JEE Mains 2017: 99.49%** percentile, Pan India
- **People's choice award 2021 & 2022:** Atlassian's internal hackathon
- **Winner** of Digital Management Inc. India Hackathon, 2019
- **2nd position** in Student Hackday 2019 organised by Skillenza
- **Barclays India Hackathon 2019 Top 4**, All India
- Achieved **5 star rating** on Codechef platform for competitive programming