Saksham Singh Kushwaha

Website: https://sakshamsingh1.github.io/

Email : sxk230060@utdallas.edu

Mobile : +1-917-443-3706

EDUCATION

The University of Texas at Dallas
Ph.D. - Computer Science (Computer Vision and Audio)
May 2027(expected)
GPA: -/4.0

New York University, Courant
M.S. - Computer Science
May 2023
GPA: 3.95/4.0

Indian Institute of Technology, Delhi
B.Tech - Mathematics and Computing
May 2018
GPA: 7.97/10.0

EXPERIENCE

Nvidia | Deep Learning Intern
May 2022 - Aug 2022
Worked in Product Security team to efficiently detect anomalous user-behavior in AWS accounts
Remote, USA

- Developed and implemented a **multi-task autoencoder** that replaced up to 10 production models.
- Improved existing intrusion detection system by 65% (F-score) and reduced false positives by 50%

Sharechat | Data Scientist II
May 2021 - Aug 2021
*Built a scalable **Facial Recognition System** to handle the cold start problem*
Bangalore, India

- Implemented pre-trained **Arcface** & **Retinaface** model-based pipeline to efficiently utilize 6M faces/day
- Trained CNN based model to detect a face's gender which improved baseline accuracy by 5%

Zomato | Machine Learning Engineer II
July 2018 - April 2021
Part of Search, User personalization, and Logistics teams
Gurgaon, India

- Improved auto-suggestion search by 10% avg. rank, 4% CTR & 2% OTR using **point-wise ML** model
- Created DQN **RL** based rider dispatch service, improving next order time(3min) & order probability(9%)
- Developed similar restaurants service by creating restaurants' embedding using modified **Word2Vec**

PUBLICATIONS

- A **multimodal prototypical** approach for **unsupervised** sound classification (paper)
  ***Saksham Singh Kushwaha**, Magdalena Feuentes [INTERSPEECH 2023]*
  - *Developed an unsupervised classification approach leveraging local audio-text embedding relationships.*
  - *Outperformed text-to-audio zero-shot SOTA models (**AudioClip, CLAP, WavClip**) by 12%.*

- ***Sound source distance estimation** in diverse and dynamic acoustic conditions*
  ***Saksham Singh Kushwaha**, Iran Roman, Magdalena Feuentes, Juan Pablo Bello [WASPAA 2023]*
  - *Proposed a **CRNN** for estimating sound source distance for acoustically diverse annotated datasets.*
  - *Improved the SOTA approach for LOCATA by ∼0.1 meter using inverse-distance loss function.*

- *Analyzing the effect of equal-angle spatial discretization on **sound event localization & detection** (paper)*
  ***Saksham Singh Kushwaha**, Iran R. Roman, Juan Pablo Bello [DCASE Workshop 2022]*
  - *Empirically showed that equal-angle targets results in non-uniform localization error(LE) along elevation*
  - *Mitigated the biasness and improved localization using Fibonacci targets & multi-task angular error loss*

ACADEMIC EXPERIENCE

- **Reviewer**: MLSP 2023
- **Research Assistant**: Prof. Magdalena Fuentes (multimodal deep learning) [Aug'22-May'23], Prof. Raveesh Mayya (ML for digital policy change) [Aug'21-July'22]
- **Teaching Assistant**: Intro to python programming (Prof. Junpei Komiyama) [Jan'22-May'22], Discrete Mathematics (Prof. Simeon Ntafos) [May'23-Aug'23], Computer Science I (Prof. Scott Dollinger) [May'23-Aug'23]