

Udit Arora

New York, NY
<https://uditarora.com>

uditarora@nyu.edu

EDUCATION

Courant Institute of Mathematical Sciences, New York University

Master of Science in Computer Science; GPA: 3.926/4.0

Courses: Machine Learning, Deep Learning, NLP, Distributed Systems, Security, Deep RL

New York, NY

Jan 2020 – Dec 2021

Netaji Subhas Institute of Technology, University of Delhi

Bachelor of Engineering in Computer Engineering

Merit Scholarship; Founding Member, NSIT Web Team

Delhi, India

Aug 2012 – May 2016

EXPERIENCE

Salesforce

Software Engineering Intern

San Francisco, CA

Jun 2021 – Aug 2021

- Added support for using past interactions as features in the Einstein Recommendation Builder modeling pipeline.
- Helped improve ERB adoption by supporting the CPQ (Configure, Price, Quote) Template GA.

Product Labs, IIIT Hyderabad

Research Engineer

Hyderabad, India

Jul 2019 – Dec 2019

- Productized commercially viable research output in collaboration with research labs at the institute.
- Implemented a pipeline for object detection and tracking in thermal images.
- Developed an end-to-end system for crowd-sourced speech data collection and transcription for 22 Indian languages.
- Architected a badminton player tracking tool used by Star Sports and a chatbot used by SBI (India's largest bank).

Laboratory for Computational Social Systems, IIIT Delhi

Research Assistant

Delhi, India

Aug 2018 – Jul 2019

- Researched on user, content (tweet) and action (retweet/quote) level detection of blackmarket-driven collusion for gaining inorganic appraisals on Twitter using different embedding methods and machine learning techniques.

Microsoft

Software Engineer

Hyderabad, India

Jun 2016 – Jul 2018

- **Excel:** Developed innovative ways of sharing Excel content, later productionized as a part of the Fluid framework. Enabled integration of Excel with a messaging app and published a patent with the USPTO for the same.
- **Kaizala:** Delivered a time-driven release of Kaizala messaging app's UWP application. Solved critical problems for workflows like heterogeneous message views and custom cards to ensure good performance.
- **Skype for Business:** Developed the Cloud Call Analytics feature for Skype for Business Server 2019, building a secure pipeline to upload call telemetry data online - among the most popular features of the 2019 release.
- **Hackathon:** Winner of the 2017 company-wide hackathon - created a utility to scan and analyze receipts in Excel.

PUBLICATIONS

- **Arora, U.**, Huang, W., He, H. (2021), *Types of Out-of-distribution texts and how to detect them*. Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP) (*Acceptance rate: ~20%*)
- Dutta, H., **Arora, U.**, Chakraborty, T. (2021), *ABOME: A Multi-platform Data Repository of Artificially Boosted Online Media Entities*. Proceedings of the International AAAI Conference on Web and Social Media (ICWSM)
- Khullar, A.*, **Arora, U.*** (2020), *MAST: Multimodal abstractive summarization with trimodal hierarchical attention*. Proceedings of the First EMNLP Workshop on NLP Beyond Text (* = equal contribution)
- **Arora, U.**, Dutta, H., Joshi, B., Chetan, A., Chakraborty, T. (2020), *Analyzing and detecting collusive users involved in blackmarket retweeting activities*. ACM Transactions on Intelligent Systems and Technology (*Impact Factor: 3.971*)
- **Arora, U.** et. al. (2019), *Multitask learning for blackmarket tweet detection*. ASONAM 2019 (*Acceptance rate: 15%*)

TECHNICAL SKILLS AND ACADEMIC SERVICE

- **Languages:** Python, C/C++, Javascript (JS), C#, Java
- **Other:** PyTorch, Keras, scikit-learn, Spark, MySQL, Git
- **Program Committee:** EMNLP Workshop on NLP Beyond Text, 2020

PROJECTS

- **Out-of-distribution Detection:** Collaborating with Prof. He He (NYU) on research to develop machine learning algorithms for detection of out-of-distribution text data using transformers and improve model robustness.
- **Multimodal Text Summarization:** Text summarization using different modalities of a video - designed and implemented a hierarchical attention based architecture that achieved SOTA results. ([link](#))
- **Covid News Analyzer:** A web-based tool for analysis of Covid-19 articles on different metrics using ML. ([link](#))
- **Fashion Product Classification:** Classification of fashion product images using convolutional neural networks with multitask learning, utilizing product metadata to improve accuracy by 3.27%. ([link](#))
- **Automated Cricket Umpire:** A tool to visualize ball trajectory and make cricket umpiring decisions from a single smartphone camera feed - used OpenCV for the ball tracking mechanism and Python for visualization. ([link](#))
- **Guidance System for Visually Impaired:** A Raspberry Pi based device to detect crosswalks/staircases using computer vision, and a companion android app to activate the device and provide haptic feedback to users. ([link](#))
- **Chain Reaction Single Player:** An Android game with bots of 10 difficulty levels that use a modified minimax algorithm with probability-based selection for faster computation - 10,000+ users and coverage by tech blogs. ([link](#))
- **NSIT Percentage Calculator:** A web-app for analyzing college grades - used by 20,000+ NSIT students. ([link](#))