Namrata **Mukhija**

(+1) 201-920-7119 | nm3571@nyu.edu | Namrata96 | namratamukhija | namrata96.github.io

Education

Courant School of Mathematical Sciences, New York University — *New York, USA*

M.S. Computer Science — *Jan 2021 - Present*

GPA: 3.87/4. Expected graduation: May 2023.

Netaji Subhas Institute of Technology — *Delhi, India*

B.E. Information Technology — *2014 - 2018*

Percentage: 78.85%, CGPA: 8.635/10

Work Experience

AI & ML Research Intern (NLP) — *New York, USA*

J.P. Morgan Chase & Co. | Machine Learning Center of Excellence | Paper — *June 2022 - Aug 2022*

- Implemented paraphrase generation using large transformer-based language models (GPT-3 and Pegasus).
- Fine-tuned and assessed performance improvement for BERT-based Named Entity Recognition models using paraphrased data.
- Work published at NeurIPS 2022 Workshop on Synthetic Data for Empowering ML Research.

Natural Language Processing Research Intern — *Bangalore, India*

Microsoft Research | Advisors: Dr. Amit Sharma — *April 2022 - May 2022*

- Implemented controlled counterfactual generation using fine-tuned GPT-2. Surveyed literature on counterfactual generation and controlled generation from language models.

Natural Language Processing Research Intern — *Bangalore, India*

Microsoft Research | Advisors: Dr. Monojit Choudhury and Dr. Kalika Bali | Paper 1, Paper 2 — *June 2021 - Aug 2021*

- Formulated six conundrums that NLP technologists must resolve while developing and deploying language technologies (LT) for social good for any LT community. Formulated principled techniques for prioritization of NLP technologies and methods for inclusion of the end-user during the development cycle.

Software Engineer 2 — *Noida, India*

Microsoft India R&D — *July 2018 - Jan 2021*

- Successfully implemented a Distributed Data Structure to mitigate data inconsistency and data loss issues in Fluid Table. Developed Fluid Table de-serialization for enabling paste on Word, Excel, and PowerPoint apps.
- Implemented and shipped a re-architecture for PowerPoint Actions in the PowerPoint team with benefits of safe rollback and reliable undo support. Developed and shipped the USD Performance Analyzer tool and USD 4.1 with Single Sign-On and caching features.
- **Awarded across 2500+ employees** in Experiences and Devices engineering group India - **E+D India Leadership Award** for generating energy in the team, demonstrating "All for One, One for All" mindset, and championing diversity and inclusion. Promoted twice in 2.5 years in recognition of the contributions.

Natural Language Processing Research Collaborator — *Delhi, India*

MIDAS lab, Indraprastha Institute of Information Technology | Advisor: Dr. Rajiv Ratn Shah | Code, Website — *September 2018 - February 2020*

- Implemented topic segmentation and discourse analysis experiments to analyze the relationship between formality and virality of Reddit posts.
- Worked on abstractive text summarization. Implemented a LSTM-based encoder architecture for learning the representation of keywords extracted using the TextRank algorithm for news articles. Implemented an attention-based pointer-generator network to produce summaries.

Natural Language Processing Research Intern — *Gurgaon, India*

Aspiring Minds | Code — *Dec 2017 - Jan 2018*

- Developed a BiLSTM model to automate grammar scoring on essays.
- Implemented augmented C&W word embeddings which would take into account the usage information of a word to treat grammatical errors as informative. This model replaced a third-party solution, and thereby, estimated a cost saving of around **$12,000/year**.

Publications

(* indicates equal contribution)

- **Mukhija Namrata***, Sharma Saket*, Joshi Aviral*, et al. (2022), Systematic Review of Effect of Data Augmentation using Paraphrasing on Named Entity Recognition. NeurIPS 2022 Workshop on Synthetic Data for Empowering ML Research. https://openreview.net/forum?id=rc2h1h89aDi
- **Mukhija Namrata***, Bali Kalika*, Choudhury Monojit*, and Diddee Harshita*(2022), The Six Conundrums of Building and Deploying Language Technologies for Social Good. ACM SIGCAS/SIGCHI Conference on Computing and Sustainable Societies (COMPASS). https://doi.org/10.1145/3530190.3534792
- **Mukhija Namrata**, Choudhury Monojit, & Bali Kalika (2021), Designing Language Technologies for Social Good: The Road not Taken. arXiv preprint. https://arxiv.org/abs/2110.07444

- Chakraborty Pinaki, Arora Udit, **Mukhija Namrata** et al. (2019), OSAVA: An Android app to visualize and teach algorithms used in operating systems. JEET, doi:10.16920/jeet/2019/v32i4/145517

Skills

Deep Learning Frameworks Keras, PyTorch
Programming Languages Python, C, C++, C#, TypeScript, JavaScript, SQL
Other Git, OpenCV, NumPy, React, Node.js, pandas, Transformers, HuggingFace, Linux (Ubuntu), Windows

Projects

Multi-modal Image Captioning

New York University

- Working on improving multi-modal image captioning using CLIP-based models as a research assistant under Prof. He He and Dr. Chen Zhao.

Out-of-Distribution Detection

New York University

- Worked on out-of-distribution detection for summarization (transformer-based language models) under Prof. He He.

Traffic Sign Classification

Advisor: Dr. Deepika Kukreja | Code

- Developed and trained a Convolutional Neural Network model on LISA Traffic Sign Dataset in Keras for classifying 47 different types of road signs. **Accuracy: 97.75%**.

OSAVA

Advisor: Dr. Pinaki Chakraborty | Code

- A desktop/android app for visualizing various operating system algorithms - developed using Kivy framework and Python.

Leadership Experience

AI Ambassador — *Global and India*

Women in AI | Website | LinkedIn APAC — *Sep 2019 - August 2021*

- Head, WaiDATATHON APAC 2021 - Spearheading the initiative to spread awareness and envisage AI solution for domestic violence and mental health across the world. Responsible for managing end-to-end Program pipeline along with execution of rest of the pillars Sponsorship, Recruitment, and Communications.
- Program Chair, WaiSUMMIT 2020 - Organized a virtual global-scale Summit which saw participation of 40+ countries, 160+ women speakers, and 4000+ attendees to encourage and facilitate women participation in AI field. Responsible for Program design, management, and end-to-end execution.
- Head, WaiCOUNCIL India. Responsible for organizing multiple technical, career, networking and educational events for increasing participation of Indian women in AI. Forged partnerships with leading AI organizations in India. Led to a 300% growth in India Women in AI community within a year.
- Invited as a panelist by Swiss Embassy, Delhi to speak on Women in AI and technology in India and Switzerland.

NLP Research Lead — *Delhi, India*

NLP group, NSIT — *Aug 2017 - Jan 2018*

- Led a team of 2 undergraduates and implemented sarcasm detection achieving an f-score 0.75.

Core Team Member & NLP Speaker — *India*

Women in Machine Learning & DS, Delhi | LinkedIn — *Sep 2018 - January 2020*

- Organized a national-level machine learning conference to connect women in ML with industry ML projects, speakers from IBM Research and TCS Research. Speaker at 'Deep Dive into NLP' on training encoder-decoder architectures and beam search decoding. Slides here

Mentor — *Delhi, India*

Learn IT, Girl! | Blog — *Oct 2017 - Dec 2017*

- Mentored a student through the development of a Google Maps Chatbot developed using Python which she deployed on Facebook.

NSIT Developer Student Club Lead — *Delhi, India*

Google Developer Student Club, NSIT | Website, Code — *Jan 2018 - Jan 2019*

- Spearheaded Google's initiative to equip NSIT students with mobile development skills by developing solutions for startups and non-profits.
- Organized multiple Android hands-on workshops & multiple projects were developed by students applying learning from these workshops.

Scholarships

- 2022 Grace Hopper Conference (Orlando), 2021 CRA-WP Grad Cohort for Women, 2017 Grace Hopper Conference (India), and 2014 INSPIRE - Higher Education

Courses

- Machine Learning, Causal Inference, Natural Language Processing with Representation Learning, Machine Learning for Healthcare, Programming Languages, Honors Analysis of Algorithms, Operating Systems