Debarati **Das**

200 Union St SE, Minneapolis, MN 55455, USA

☐ (+1) 651-348-0431 | ✉ das00015@umn.edu | ⌂ https://debaratidas94.github.io/ | 🔗 debaratidas1994 | 🎓 Google Scholar

Summary

I am interested in the nuances of online social media interactions — understanding the role of user tastes, societal signals, and contextual factors in shaping user actions. Additionally, I am keen on delving into the stylistic examination of user-generated text. Merging insights from both computer science and social sciences, I utilize diverse techniques such as deep learning, graph learning, social network analysis, NLP/LLMs, and statistical approaches.

Education

University of Minnesota Twin Cities   *Minneapolis, MN*

Ph.D. Candidate in Computer Science - GPA: 4.00/4.00   *August 2018 - Present*

- **Research Interest:** Computational social science & Natural Language Processing
- **Advisors:** Dongyeop Kang and Jaideep Srivastava

University of Minnesota Twin Cities   *Minneapolis, MN*

M.S. in Computer Science - GPA: 3.70/4.00   *August 2017 - August 2019*

Selected Publications

Computer Science Papers

Prototypical Human-AI Collaboration Behaviors from LLM-Assisted Writing in the Wild   *EMNLP 2025*

S. Mysore, **D. Das**, H.Cao and B.Sarrafzadeh   *[Paper Link]*

LawFlow : Collecting and Simulating Lawyers' Thought Processes   *COLM 2025*

D. Das, K. Le*, R. Parkar*, K. Langis, B. Madson, C. Berryman, R. Willis, D. Moses, B. McDonnell, D. Schwarcz, and D. Kang   *[Paper Link]*

Under the Surface: Tracking the Artifactuality of LLM-Generated Data   *TACL*

D. Das, K. Langis, A. Martin, J. Kim, M. Lee, Z. Kim, S. Hayati, R. Owan, B. Hu, R. Parkar, R. Koo, J. Park, A. Tyagi, L. Ferland, S. Roy, V. Liu and D. Kang   *[Under Review]*

Which Modality should I use - Text, Motif, or Image? : Understanding Graphs with Large Language Models   *NAACL 2024*

D. Das, I. Gupta, J. Srivastava and D. Kang   *[Paper Link]*

Balancing Effect of Training Dataset Distribution of Multiple Styles for Multi-Style Text Transfer   *ACL 2023*

D. Das, D. Ma and D. Kang   *[Paper Link]*

AdBERT: An Effective Few Shot Learning Framework for Aligning Tweets to Superbowl Advertisements   *W-NUT @ COLING 2022, WiML @ ICML 2021*

D. Das, R. Chenchu, M. Abdollahi, J. Huh and J. Srivastava   *[Paper Link]*

Scalable clustering by aggregating representatives in hierarchical groups   *Journal of Pattern Recognition*

W. Xie, Z. Liu, **D. Das**, B. Che and J. Srivastava   *[Paper Link]*

A Computational Analysis of Mahabharata   *ICON 2016*

D. Das, B. Das and K. Mahesh   *[Paper Link]*

Modeling Memetics Using Edge Diversity   *Social Network Analysis and Mining*

Y. Gupta, A. Saxena, **D. Das** and S. Iyengar   *[Paper Link]*

Pseudo-Cores: The Terminus of an Intelligent Viral Meme's Trajectory   *Complex Networks VII*

Y. Gupta, **D. Das** and S. Iyengar   *[Paper Link]*

Application Papers

Thinking Like A Lawyer In The Age Of Generative AI: Cognitive Limits On AI Adoption Among Lawyers   *Minnesota Legal Studies - SSRN*

D. Schwartz, **D. Das**, D. Kang and B. McDonell   *[Paper Link]*

Publics' Perceptions of Legitimacy in Corporate Social Advocacy: A Computational Analysis of the Influence of Ideological Congruence

Public Relations Review

H. Xu, **D. Das**, J. Huh, H.Rim and J.Srivastava

[Paper Link]

Influence of Consumers' Temporary Affect on Ad Engagement: A Computational Research Approach

Journal of Advertising

X. Lu, **D. Das**, J. Huh and J. Srivastava

[Paper Link]

Rebuilding Social Connection and Enhancing Advertising Effects Through the Nostalgic Appeal during the Pandemic

Journal of Advertising Research

K. Kim, B. Chang, **D. Das**, S. Muthya, J. Huh, D. Kang and J. Srivastava

Under Review

Work In Progress: Reactions to Incivility: A Computational Approach to Measuring Affective Polarization on Twitter During the First 2020 U.S. Presidential Debate

ICA 2023

J. Gustin, C. Gansen, T. He, **D. Das**, J. Huh and J. Srivastava

[May 2023]

Work Experience

Microsoft, Redmond

Senior Applied Scientist

Aug 2025 - Present

- Understanding and steering the impact of Generative AI on human productivity.

Microsoft Research, Redmond

Researcher

Sep 2024 - July 2025

- Exploring contextual personalization in email drafting at the Office of applied research.

Microsoft Research, Redmond

Research Intern

Jun. 2024 - Sep 2024

- Developing a RAG-Assisted answer drafting tool to aid Enterprise QA Systems. Conducted user study-based analysis to understand user knowledge-seeking behavior and needs on the community QA forum.

Minnesota NLP Lab

Research Assistant

Jun. 2021 - Present

- Built a *human-expert dataset and benchmark* to understand legal document drafting and associated planning with AI assistance.
- Developed a fine-grained dataset of conflicting societal values and preferences for *Human-LLM alignment*.
- Conduct a rigorous analysis to identify *biases and artifacts from LLM-generated synthetic data*.
- Leveraging *LLMs like GPT4 for graph reasoning* tasks, primarily node classification, to enhance the accuracy and efficiency of the prediction.
- Explored the balancing effect of training data set distribution of multiple styles for *multi-style text transfer*. Developed a T5-based pipeline that can perform joint style analysis, transfer, and generation for different dimensions of style simultaneously using pseudo-parallel data sets.

Minnesota Computational Advertising Laboratory

Research Assistant

Jul. 2019 - Dec 2023

- Collected and processed large-scale Twitter data. Developed a BERT-based *political ideology classifier* incorporating a user's textual and behavioral features for a more comprehensive classification.
- Collected and processed large-scale Twitter data. Developed AdBERT, a *few-shot learning framework* to align tweets with the TV advertisements they refer to in the context of the Super Bowl.
- Implemented *latent attribute detection* for gender, location, and age and BERT-based classifiers for sentiment and emotion from social media text.

3M

NLP Research Intern

May 2020 - Aug. 2020

- Developed a pipeline to perform *Neural Extractive Summarization* from MIMIC-III ICU data using BERT. Evaluated and identified the most effective query embeddings for better summarization of clinical text.

3M

Research & Dev Intern

May 2019 - Aug. 2019

- Explored *Bayesian Neural Nets for large-scale time series* data. Implemented Variational Inference and Dropout-based Bayesian approximation methods as Bayesian Neural Layers for models to quantify the uncertainty of predictions.

Microsoft Research, India

UG Research & Dev Intern

Jan. 2017 - Jun. 2017

- Developed a *cloud-based pipeline* for analyzing data derived from smartphones and onboard automobile sensors to develop a real-time prediction system enabling driver safety and accident avoidance.

Indian Institute of Technology

UG Research Intern

May 2015 - Aug. 2015

- Developed decentralized search-based algorithms to inject *virality in a social network* by influencing the trajectory of meme propagation.
- Built a probabilistic ordering model where the probability of influence propagation *mimics a meme's flow* in a real-world social network.

Awards & Accomplishments

2024   **Recipient**, AnitaB.org Advancing Inclusion Scholarship
2019   **Recipient**, Computing Research Association-Women (CRA-W) scholarship
2016   **Recipient**, Grace Hopper Student Scholarship
2016   **Runner Up**, IET India Scholarship
2016   **Recipient**, National Network for Mathematical & Computational Biology Research Fellowship
2015   **Recipient**, Indian Academy of Sciences Research Fellowship

Presentations and Talks

Microsoft   *Redmond*
GUEST TALK   *March 2025*
Presented my research titled "LLMs as Sensors and Subjects: Expanding Human–AI Collaboration in Socio-Technical Systems"

Sparkway Learning PythonAI Workshop 2024   *Minneapolis*
GUEST TALK   *Dec 2024*
Gave a talk titled "When AI learns to Listen: Generative AI in the Social Sciences" to high school students.

Women in AI and Data Science (WiADS) 2024   *Minneapolis*
ORAL PRESENTATION   *Nov 2024*
Talk about LLM-generated artifacts and their impact on the AI ecosystem

10th International Conference on Computational Social Science   *Pennsylvania*
ORAL PRESENTATION   *July 2024*
Talk about our abstract titled, "Beyond Words: Incorporating Graph Data into LLMs for Comprehensive Social System Analysis"

Megagon Labs   *Virtual*
GUEST TALK   *May 2024*
Talk about my paper on stress testing LLM-generated data for artifacts

Hubbard School of Journalism (JOUR8500) at UMN   *Minneapolis*
GUEST LECTURE   *May 2023*
Talk about computational techniques used to analyze political ideologies and sentiments in social media data

Microsoft Research   *Bangalore*
GUEST TALK   *Feb 2017*
Talk about my paper on the computational analysis of the Mahabharata

Teaching Experience

Teaching Assistant
THE PRACTICE OF DB SYSTEMS   *Fall 2018 - Present*
- Designing and grading assignments, holding office hours, and conducting lectures.
- **Terms Served :** - Spring 2024, Summer 2023, Fall 2022, Summer 2022, Summer 2021, Spring 2020, Spring 2018

Teaching Assistant
NATURAL LANGUAGE PROCESSING   *Spring 2023*
- Mentoring group projects, designing tutorials, and organizing research poster presentations.
- Designing and grading assignments, holding office hours, and conducting lectures.

Instructor
CSE TALK PROGRAM   *Summer 2023*
- Designing course material and instructing prospective graduate students on how to serve as teaching assistants.

Teaching Assistant
INTRODUCTION TO ALGORITHMS AND DATA STRUCTURES USING JAVA   *Fall 2019*
- Grading assignments and holding office hours.

Teaching Assistant
INTRO TO ALGORITHMS AND PROGRAM DEVELOPMENT   *Fall 2018 - Fall 2023*
- Grading assignments and holding office hours.
- **Terms Served :** - Fall 2023, Spring 2019, Fall 2018

Academic Service

Reviewer

- **CS Conferences and Workshops:**
    - ACM SIGCHI 2026
    - NLP4PosImpact 2025 @ ACL 2025
    - COLING 2025
    - IC2S2 2024, 2025
    - In2Writing Workshop @ CHI 2024, 2025
    - Pluralistic-Alignment @ Neurips 2024
    - CustomNLP4U @ EMNLP 2024
    - ARR [2024 - Feb, June, August]
    - WiNLP Workshop @ EMNLP 2024
    - LREC-COLING 2024
    - NLP+CSS Workshop @ EMNLP 2023
- **Journals and other scholarly venues:**
    - Transactions on Knowledge Discovery from Data
    - Transactions on Visualizations and Computer Graphics
    - Humanities and Social Sciences Communications
    - Computational Communication Research Journal
    - Journal of Marketing Communications
    - American Academy of Advertising 2023

Co-organizer

Minnesota NLP Seminar   *Fall 2021 - Fall 2023*

- Organized seminars hosting speakers from different fields of Natural Language Processing

Officer

Computer Science Graduate Students Association   *Fall 2020 - Fall 2022*

- Advocate for CS graduate students
- Organize social and academic events for grad students.

Member

Diversity, Equity, and Advocacy Committee (CS-IDEA)   *Spring 2020 - Present*

- Work collaboratively to encourage students and staff to grow in their understanding of diversity.
- Promote equal opportunity for all.

Coordinator

Women in Computer Science   *Spring 2019- Spring 2022*

- Encourage and build a support system for female-identifying and non-binary graduate students in the department.
- Facilitate networking between academic/industry professionals and students.
- Organize events and outreach activities to create awareness about inclusion and diversity issues.

Mentorship Experience

[1]	**Jianing Wen**, Summer 2024 - Spring 2025	*UMN CS MS*
[2]	**Amrutha Shetty**, Spring 2024	*UMN CS MS*
[3]	**Sanjali Roy**, Fall 2023 - Spring 2024	*UMN CS UG*
[4]	**Ishaan Gupta**, Fall 2023	*UMN CS UG*
[5]	**David Ma**, Fall 2021 - Spring 2021	*UMN CS MS*
[6]	**Josh Spitzer-Resnick**, Fall 2021	*UMN CS MS*
[7]	**Aryan Srivastava**, Fall 2021	*High School*
[8]	**Haoyu Gong**, Spring 2021	*UMN CS PhD*
[9]	**Aadesh Salecha**, Spring 2021- Spring 2022	*UMN CS UG*
[10]	**Ahmed Hassan**, CSE Mentorship Program, Fall 2020	*UMN CS UG*
[11]	**Hao Zou**, Fall 2020	*UMN CS UG*
[12]	**Roopana Chenchu**, Spring 2020 - Spring 2021	*UMN CS MS*
[13]	**Mayura Nene**, Women in Science and Engineering (WISE) Mentorship Program, Spring 2020	*UMN CS MS*

Skills

Programming	Python, Java
ML/ NLP	Tensorflow, Pytorch, WEKA, OpenCV, NLTK, Spacy
Social Network Analysis	NetworkX, Gephi, GraphViz
DevOps	SQLite, MongoDB, AWS, Docker
OS	Linux, Windows, OSX