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

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]*

Under the Surface: Tracking the Artifactuality of LLM-Generated Data   *Under Review @ ACL 2024*

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   *[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

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*

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 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

[To be Presented]

Work Experience

Minnesota NLP Lab

Research Assistant   *Jun. 2021 - Present*

- 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 - Present*

- 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.

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

Teaching Experience

Teaching Assistant

The Practice of DB Systems   *Fall 2018 - Present*

- Designing and grading assignments, holding office hours and lecturing.
- **Terms Served :** - Fall 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 lecturing.

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

Mentorship Experience

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

Awards & Accomplishments

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

Academic Service

Reviewer

NLP and Advertising Conferences *Fall 2022 - Present*
- IC2S2 2024
- In2Writing Workshop 2024
- ACL 2024
- LREC-COLING 2024
- NLP+CSS Workshop @ EMNLP 2023
- 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.