Sujan Dutta

Rochester, NY | (585) 694-1217 | sd2516@rit.edu | github.com/Suji04 | YouTube | Google Scholar | sujandutta.com

EDUCATION

Rochester Institute of Technology — Rochester, NY
Ph.D. — Aug 2022 - Present
Computing and Information Sciences

Rochester Institute of Technology — Rochester, NY
Master of Science (MS) — Aug 2021 - Aug 2022
Data Science

Kalyani Government Engineering College — Kalyani, WB
Bachelor of Technology (B.Tech) — Aug 2017 - Jul 2021
Computer Science

WORK EXPERIENCE

Apple — Seattle, WA
Research Intern PhD — May 2023 – Aug 2023
- Proposed and implemented a finetuning framework using reinforcement learning with AI feedback (RLAIF) to improve the planning, task decomposition, and code generation abilities of smaller LLMs. This work will appear at ACL 2024 NLRSE.
- Successfully reduced language model sizes by 80%, maintaining equivalent performance levels and thereby optimizing efficiency.

Apple — Seattle, WA
Research Intern — May 2022 – Aug 2022
- Designed a large-scale differentially private (DP) personalized recommendation system for more than 100,000 users.
- Presented the project results at an internal company conference, highlighting the system's design and impact.

ASAW — Remote
Data Science Intern — Mar 2021 – Jul 2021
- Developed a natural language understanding tool for automated knowledge extraction and ranking documents.
- Integrated multiprocessing in an existing framework achieving a 4x speed-up in document parsing.

Kesowa Infinite Ventures — Remote
Machine Learning Intern — Mar 2020 – Jan 2021
- Implemented real-time (30 fps) computer vision systems using retina neural network for various tasks, including tree counting, crowd counting, and people tracking.

National Institute of Technology — Durgapur, WB
Research Intern — Jun 2019 – Jul 2019
- Developed machine learning powered data processing and imputation algorithms for spatiotemporal air pollution data, improving accuracy by 10%.

SKILLS

- Machine Learning (ML), Artificial Intelligence (AI), Natural Language Processing (NLP), Large Language Model (LLM), RLHF, Generative AI, Multimodal Models, AI Ethics, AI Safety, Natural Language Understanding, Computational Linguistics, Deep Learning, Conversational AI, Data Science, Research, Statistics, Leadership, Public Speaking.
- Python, C/C++, MATLAB, PyTorch, TensorFlow, NumPy, Pandas, NLTK, Spacy, HuggingFace.

CAREER HIGHLIGHTS

Publications: Author of 5+ publications in esteemed AI conferences including IJCAI and EMNLP.
YouTube Educator: Creator of the educational channel Normalized Nerd on Machine Learning, with over 96,000 subscribers.
Language Science Student Excellence Award: Received for the research in estimating disparate news media stances using NLP.
Conference Talk: Gave a tutorial talk on the evolution of NLP at the UPSTAT 2023.
ML Hackathon Finalist: Competed in the final round (top 20) of the ML hackathon at the IIT Madras.

PUBLICATIONS

- **S. Dutta**, B. Li, D. S. Nagin, and A. R. KhudaBukhsh, "A Murder and Protests, the Capitol Riot, and the Chauvin Trial: Estimating Disparate News Media Stance," in Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI-22, L. D. Raedt, Ed., AI for Good, International Joint Conferences on Artificial Intelligence Organization, pp. 5059–5065. DOI: 10.24963/ijcai.2022/702
- T. Weerasooriya, **S. Dutta**, T. Ranasinghe, M. Zampieri, C. Homan, and A. KhudaBukhsh, "Vicarious Offense and Noise Audit of Offensive Speech Classifiers: Unifying Human and Machine Disagreement on What is Offensive," in Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing, H. Bouamor, J. Pino, and K. Bali, Eds., Singapore: Association for Computational Linguistics, pp. 11648–11668. DOI: 10.18653/v1/2023.emnlp-main.713
- **S. Dutta**, P. Srivastava, V. Solunke, S. Nath, and A. R. KhudaBukhsh, "Disentangling Societal Inequality from Model Biases: Gender Inequality in Divorce Court Proceedings," in Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence, IJCAI-23, E. Elkind, Ed., AI for Good, International Joint Conferences on Artificial Intelligence Organization, pp. 5959–5967. DOI: 10.24963/ijcai.2023/661
- A. Khorramrouz, **S. Dutta**, and A. R. KhudaBukhsh, "For Women, Life, Freedom: A Participatory AI-Based Social Web Analysis of a Watershed Moment in Iran's Gender Struggles," in Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence, IJCAI-23, E. Elkind, Ed., AI for Good, International Joint Conferences on Artificial Intelligence Organization, pp. 6013–6021. DOI: 10.24963/ijcai.2023/667
- A. Dutta, A. Khorramrouz, **S. Dutta**, and A. R. KhudaBukhsh, "Down the Toxicity Rabbit Hole: A Novel Framework to Bias Audit Large Language Models," 2023. URL: https://arxiv.org/abs/2309.06415
- **S. Dutta** and K. Dasgupta, "A Shallow Approach to Gradient Boosting (XGBoosts) for Prediction of the Box Office Revenue of a Movie," in Proceedings of International Conference on Innovations in Software Architecture and Computational Systems: ISACS 2021, Springer, 2021, pp. 207–219. DOI: 10.1007/978-981-16-4301-9_16