


KRITHIKA RAMESH

(+91) 9049986993  kramesh.tlw@gmail.com

 Semantic Scholar  GitHub  LinkedIn  Website

EDUCATION

Manipal Institute of Technology, Manipal

B.Tech in Information Technology, (Minor in Computational Intelligence)

2018 – 2022

CGPA: 8.12/10

EXPERIENCE

Research Intern

Microsoft Research India

August 2022 – Present

Supervisor: Dr. Sunayana Sitaram

- Domain: Fairness in Language Models
- Analyzed and documented the gaps and challenges for both monolingual (inclusive of languages beyond English) and multilingual NLP. Studied the effects of model compression methods on fairness in text classification. Work under review at EACL '23.
- Ongoing research includes devising model compression methods that minimize the bias in the compressed model, and a study to analyze bias in multilingual models.

Research Intern

Microsoft Research India

January 2022 – July 2022

Supervisor: Dr. Rahul Sharma

- Domain: Privacy-Preserving Machine Learning
- Implemented deep learning models such as the GCN-LSTM and the ConvLSTM architectures using the [EzPC](#) framework to predict air pollution levels using remote sensing data, whilst optimizing for a performance-efficiency tradeoff due to the additional costs of secure multi-party computation.
- Used the EMP toolkit to perform secure inference operations over the outputs. Work under review at PETS '23.

Research Intern

Indian School of Business, Hyderabad

May 2021 – Present

Supervisors: Dr. Sumeet Kumar, Dr. Ashiqur KhudaBukhsh

- Analyzed the inappropriate content generated by ASR systems in childrens' videos, and proposed a method using masked language modelling and phonetic embeddings to rectify these errors. The [work](#) was accepted to **AAAI '22** (AISI).
- Position paper on revisiting hate speech/abusive language lexicons with regard to the presence of words related to the queer community.
- Currently working on the multilingual analysis of lexicons and their inconsistencies, and the development of a more robust method to fix incorrectly transcribed abusive language.

RESEARCH EXPERIENCE AND PUBLICATIONS

'Beach' to 'Bitch': Inadvertent Unsafe Transcription of Kids' Content on YouTube

[Paper](#)

K. Ramesh, AR KhudaBukhsh, S. Kumar, Proceedings of the Thirty-Sixth AAAI Conference on Artificial Intelligence, AI for Social Impact Track

Revisiting Queer Minorities in Lexicons [Paper](#)

K. Ramesh, S. Kumar, AR KhudaBukhsh, Proceedings of the 6th Workshop on Online Abuse and Harms, NAACL, 2022

Towards Quantifying the Carbon Emissions of Differentially Private Machine Learning

[Paper](#)

R. Naidu*, H. Diddee*, A. Mulay*, A. Vardhan*, K. Ramesh*, A. Zamzam, Socially Responsible Machine Learning Workshop, ICML, 2021

Evaluation of Gender Bias in Hindi-English Machine Translation [Paper](#)

G.Gupta*, K.Ramesh*, S.Singh, Proceedings of the Third Workshop on Gender Bias in Natural Language Processing, ACL-IJCNLP, 2021

ADDITIONAL PROJECTS

Detecting gender bias with post-hoc interpretable methods

- Used post-hoc explanations from various algorithms such as LIME and SHAP to measure gender bias.
- Modified the Equity Evaluation Corpus (EEC), and extracted explanations from LIME and SHAP to find the bias associated with masculine, feminine and gender-neutral terms, and to identify the terms that contribute the most to this bias. Preliminary work accepted at WiNLP, EMNLP 2021.
- Ongoing work deals with using self-explaining neural network architectures and other interpretability methods to detect and mitigate demographic biases.

TECHNICAL STRENGTHS

Programming Languages	Python, C++, C, Java
AI Libraries/Frameworks	PyTorch, Tensorflow, keras, sk-learn, OpenCV, nltk, spaCy, etc
Other Software	NetLogo, GAMA

EXTRACURRICULARS

- **Queer in AI** talk at **ACL '21** on **Queer Representation in Indian Academia**
- **Volunteer** at **ICML '21**, **NAACL '22**
- **ACM and ACM-W, Manipal**: Core committee member.
- **Volunteer** at **Teach Code for Good**: An NGO in Manipal that aims to educate underprivileged children between the ages of 11-15 about the basics of coding in Python and HTML/CSS.
- **The Research Society-MIT**
Co-President and Co-Founder
 - Founded the Research Society in 2020, a student organization in the Manipal Institute of Technology, with the core aim of providing a platform for undergraduate research. With over a **100 members**, it spans **10 domains** some of which include biotechnology, robotics, and artificial intelligence.
 - **25+ papers accepted** to journal and conference proceedings around the world, with 40+ ongoing projects. Organized over 20 webinars where we invited proficient researchers in a variety of fields to discuss their work.
 - Assisted in mentoring and conducting the training phase for over 20 students in the **Artificial Intelligence and Cybersecurity** domains.