



## Mathematics & Statistics Colloquium

---

Wednesday, December 6, 2023, 4:15pm-5:15pm

G5 Rolla Building and online via zoom<sup>1</sup>

Please join for refreshments and coffee (start at 4pm)  
before the talk!

---

### Subha Maity

Ph.D. Candidate  
Department of Statistics  
University of Michigan

---

## Investigations of algorithmic biases caused by underrepresentation of minority groups

**Abstract.** The problem of algorithmic bias, where machine learning algorithms reflect biases that are prevalent in their training datasets, is widely recognized as a major concern. In this talk, I will discuss two of my projects related to algorithmic biases that are caused by underrepresentation of minority groups. In the first project, we demonstrate that when learning representations from standard contrastive learning methods, the representations of minority groups merge with the representations of certain similar majority groups. We refer to this phenomenon as representation harm and demonstrate that it leads to allocation harms in downstream classification tasks. In the second project, we investigate whether enforcing group fairness is aligned with improving model performance. In light of the long-held belief that enforcing fairness comes at the cost of reduced model performance, we present an alternative perspective on the problem. In cases where the machine bias is due to the underrepresentation of minority groups, we show that enforcing fairness is often in line with improving model performance on a balanced test dataset. Furthermore, we derive necessary and sufficient conditions for such an alignment.

**Biographical Sketch.** Subha Maity is pursuing his Ph.D. in Statistics at the University of Michigan under the supervision of Prof. Moulinath Banerjee and Prof. Yuekai Sun. Prior to graduate study, he was a student at the Indian Statistical Institute in Kolkata, where he received his Bachelor and Master's degrees in Statistics. His primary research focuses on transfer learning and algorithmic fairness. More generally, he takes a keen interest in problems arising at the interface of statistical theory and machine learning.

<sup>1</sup>Zoom meeting ID: 94694762630  
password: mathstat