

# Namrata Deka

PHD STUDENT · MACHINE LEARNING

✉ [ndeka@cs.cmu.edu](mailto:ndeka@cs.cmu.edu) | 🏠 [namratadeka.github.io](https://namratadeka.github.io) | 📷 [namratadeka](https://www.instagram.com/namratadeka) | 📺 [namrata-deka](https://www.youtube.com/channel/UCv33333333333333333333)

## Summary

Ph.D. student in Machine Learning at Carnegie Mellon University developing robust representation learning methods for controllable multi-modal planning, generation and cross-modal alignment. Strong in Python/PyTorch/TensorFlow and end-to-end research prototyping; seeking a Summer 2026 research internship / PhD mentorship in multimodal modeling and generative AI.

## Education

### Carnegie Mellon University

Pittsburgh, USA

PH.D., MACHINE LEARNING

2023 - Present

- Advisors: Dr. Jeff Schneider, Dr. Kun Zhang

### University of British Columbia

Vancouver, Canada

M.SC., COMPUTER SCIENCE

2020 - 2023

- Advisor: Dr. Danica J. Sutherland
- Thesis: Kernel Methods for Invariant Representation Learning: Enforcing Fairness and Conditional Independence

### Indraprastha Institute of Information Technology

New Delhi, India

B.TECH., COMPUTER SCIENCE AND ENGINEERING

2013 - 2017

- Honors thesis/undergrad research advisors: Dr. Saket Anand, Dr. Sanjit K. Kaul

## Publications

### PEER-REVIEWED CONFERENCES

- Y. Shen, P. Zhu, Z. Li, S. Xie, **N. Deka**, Z. Liu, Z. Tang, G. Chen, K. Zhang. *Controllable Video Generation with Provable Disentanglement*, in the 14th International Conference on Learning Representations (**ICLR**), 2026.
- Z. He, R. Pogodin, Y. Li, **N. Deka**, A. Gretton, D. Sutherland. *On the Hardness of Conditional Independence Testing in Practice*, in the 39th Annual Conference on Neural Information Processing Systems (**NeurIPS**), 2025 (**Spotlight**).
- R. Pogodin\*, **N. Deka**\*, Y. Li\*, D.J. Sutherland, V. Veitch, A. Gretton. *Efficient Conditionally Invariant Representation Learning*, in the 11th International Conference on Learning Representations (**ICLR**), 2023 (**Oral/Top 5%**). \*Equal Contribution.
- N. Deka**, D.J. Sutherland. *MMD-B-Fair: Learning Fair Representations with Statistical Testing*, in the 26th International Conference on Artificial Intelligence and Statistics (**AISTATS**), 2023.

### CURRENTLY UNDER REVIEW

- K. Zhang, J. Sun, Y. Li, **N. Deka**, S. Xie, I. Ng. *SEDGE: Structural Extrapolated Data Generation*, 2026.
- P. Zhu, S. Xie, Z. Li, Y. Shen, **N. Deka**, H. Shrivastava, G. Chen, K. Zhang. *MoVA: Learning Asymmetric Dual Projections for Modular Long Video-Text Alignment*, 2026.

### WORKSHOPS & TECHNICAL REPORTS

- Y. Khandelwal, M. Arvind, S. Kumar, A. Gupta, S.K. Danisetty, P. Bagad, A. Madan, M. Lunayach, A. Annavajjala, A. Maiti, S. Jain, A. Dalmia, **N. Deka**, J. White, J. Doshi, A. Kanazawa, R. Panicker, A. Raval, S. Rana, M. Tapaswi. *NurtureNet: A Multi-task Video-based Approach for Newborn Anthropometry*, in the 7th **CVPR** Workshop on Computer Vision for Physiological Measurements (CVPM), 2024 (**Best Paper**).
- N. Deka**, D. Sutherland. *Learning Privacy-Preserving Deep Kernels with Known Demographics*, in the 36th **AAAI** Conference on Artificial Intelligence. Workshop on Privacy-Preserving Artificial Intelligence (PPAI), 2022.
- D. Sutherland, **N. Deka**. *Unbiased estimators for the variance of MMD estimators*. Technical report 2022 (arXiv:1906.02104).

## Academic Research Experience

---

### Carnegie Mellon University - Dept. of Machine Learning

ADVISORS: DR. JEFF SCHNEIDER, DR. KUN ZHANG

- Developing methods to learn structured representations from high-dimensional modalities (timeseries, text) to improve robustness and controllability in generation.
- Building multimodal concept-alignment pipelines for long timeseries-text understanding with an emphasis on evaluation under distribution shift.
- Prototyping LLM-based multimodal scientific assistants for experiment planning in data-driven domains (fusion/tokamak control), focusing on grounding and reliability.

### University of British Columbia - Dept. of Computer Science

ADVISOR: DR. DANICA J. SUTHERLAND

- Developed a conditionally invariant representation learning model using kernel-based conditional independence measures; published ICLR 2023 (Oral/Top 5%).
- Developed a fair representation learning model with statistical testing guarantees; published AISTATS 2023.

## Professional Experience

---

### Borealis AI

Vancouver, Canada

ML RESEARCH INTERN

Sep. 2022 - Jan. 2023

- Audited ML systems for systematic performance gaps via slice discovery; designed evaluation and diagnostic tools to surface under-represented group errors and guide mitigation.

### École Polytechnique Fédérale de Lausanne (EPFL)

Lausanne, Switzerland

COMPUTER VISION RESEARCH INTERN, SUMMER@EPFL

Jun. 2021 - Aug. 2021

- Researched articulated object modeling from video using 3D optical flow + depth; implemented training/inference pipeline and ablations.

### Wadhvani Institute for AI

Mumbai, India

AI RESEARCH FELLOW

Aug. 2018 - Aug. 2020

- Built a 3D vision system to screen low birth-weight babies with smartphone cameras for public health sectors in rural India.
- Developed a novel method to reconstruct 3D infant meshes at metric scale.

### Microsoft Research

Bengaluru, India

APPLIED ML INTERN

Jan. 2018 - Jul. 2018

- Built GDPR-oriented PII tagger for rare-sensitive categories; improved detection reliability through targeted data curation and error analysis.

## Voluntary Service

---

### PEER REVIEW

- Conference on Causal Learning and Reasoning (**CLear**): 2026
- Conference on Neural Information Processing Systems (**NeurIPS**): 2024
- Conference on Artificial Intelligence and Statistics (**AISTATS**): 2023

## Teaching Experience

---

Fall 2025 **Machine Learning in Practice**, Teaching Assistant, CMU

Fall 2022 **Intelligent Systems**, Teaching Assistant, UBC

Spring 2021 **Intelligent Systems**, Teaching Assistant, UBC

Spring 2017 **Computer Vision**, Teaching Assistant, IIIT-D

Fall 2016 **Advanced Programming**, Teaching Assistant, IIIT-D