SABYASACHI SAHOO

iamsabyasachisahoo@gmail.com ◇ sabyasachis.github.io ◇ google-scholar ◇ ◇ ◇

Ph.D. candidate with industry experience working on reasoning, generalization, robustness & efficiency.

EDUCATION

Mila, Inria, Université Laval Aug 2021 - Present
Ph.D. in Machine Learning. Advisors: Christian Gagné & Frédéric Precioso. (GPA: 4/4) (Expected: Dec 2026)
Keywords: *reasoning*, *generalization*, *robustness*, *efficiency*, *LLMs*, *VLMs*, *finetuning*.

Indian Institute of Science (IISc) - Bangalore 2014 - 2016
Masters in Computational and Data Science. Advisor: Sathish Vadhiyar. (Top 5 in class)

National Institute of Technology (NIT) - Surat 2010 - 2014
Bachelors in Mechanical Engineering.

WORK EXPERIENCE

Amazon - Applied Scientist Intern Jun 2025 - Oct 2025

- **Reasoning Faithfulness**: Using chain-of-thought (CoT) reasoning faithfulness to improve LLM finetuning.

NVIDIA - System Engineer II Aug 2016 - May 2018

- **Self-Driving Car Display**: Implemented new features, reducing bugs (60%) and improving user experience.
- **Resource Management**: Designed modular device tree framework, cutting cross-team development time (30%).
- **Xavier Chip's Launch**: Implemented critical modules for successful launch & performance boost (25%).

Donut Research Labs - Deep Learning Engineer May 2018 - Feb 2019

- **Text Generation**: Improved Walmart's brand name extraction accuracy (40%) using a Seq2Seq model, pre-trained on product descriptions, fine-tuned on noisy annotations, and top-k sampling with beam search.
- **Extreme Text Classification**: Class-imbalance-aware LSTM improved accuracy (25%) & sped inference (10%).
- **Product Specs Extraction/OCR**: Single Shot Detector based information extraction, improving perf (35%).
- **Dataset Building**: for text generation & OCR projects (collection, annotation, cleaning).

ML Lab, IISc Bangalore - Research Associate Feb 2019 - Aug 2021

- **Image Generation** [4, 17, 25]: Enhanced LiDAR generation using an adversarial autoencoder, improved high-resolution image generation using Discrete Cosine Transform based Variational Autoencoders (VAEs), and made scene flow prediction adaptive by fusing learning-based and geometry-based approaches.
- **Multi-Task Generalization** [3, 24]: Improved domain adaptation and explainability for smartphone-transmitted X-rays using Multi-Task Learning and deployed it as a text-based reporting tool during the COVID-19 pandemic.
- **Model Explainability** [14, 15, 26]: Aligned concept extraction with human understanding, improved it using non-negative matrix factorization, and exposed adversarial vulnerabilities in popular explainability methods.
- **Robotics** [6, 9, 27]: Differentiable SLAM framework for LiDAR & cost-effective driver assistance system (ADAS).

ACADEMIC EXPERIENCE

Mila, Inria - Graduate Research Assistant Aug 2021 - Present

- **Test-time Finetuning** [1, 10]: Enhanced adaptation to distribution shifts using optimal layer selection for test-time adaptation and self-distillation of CLIP for zero-shot image classification.
- **Domain Generalization** [8, 13]: Improved generalization to distribution shifts using out-of-distribution methods and denoising incorrect predictions using diffusion models.
- **Adversarial Robustness** [2, 11, 12]: Enhanced adversarial defense by using softmax predictions to detect vulnerable samples cheaply and using test-time adaptation.
- **Out-of-Distribution (OOD)** [5, 16]: Improved OOD detection using gradients from OOD prototype and optimized microservice partitioning by reformulating it into a reinforcement learning (RL) problem.
- **Continual Learning** [7, 23]: Reduced forgetting in continual learning through hessian-aware low-rank approximation and showed that simple ensembling (bagging/dropout) can outperform complex strategies.

Computational and Data Science, IISc Bangalore - Graduate Research Assistant Aug 2015 - Aug 2016

- **Visual Semantic Search** [18]: Extracted hierarchical relationships between visually similar classes in CNNs.
- **Parallel Computing** [19, 21]: Scaled molecular dynamics on dragonfly supercomputer using a hierarchical graph partitioning and improved Traveling Salesman Problem using a hybrid CPU-GPU/CUDA implementation.

ONGOING WORKS

- S. Sahoo *et al.* **Reasoning Faithfulness improves LLM Finetuning**.
- S. Sahoo *et al.* **Adaptive Data Filtering for Improving Generalization**.
- S. Menon, S. Sahoo *et al.* **Learnable Layer Selection for Test-Time Finetuning**.
- M. Guerrier, S. Sahoo *et al.* **Safe and Robust Robot Navigation**.

SELECTED PUBLICATIONS

[1] S. Sahoo, M. ElAraby, J. Ngnawe, Y. Pequignot, F. Precioso, and C. Gagné. ***A Layer Selection Approach to Test Time Adaptation***. In: Association for the Advancement of Artificial Intelligence Conference. [link]. **AAAI**. **2025**.

[2] J. Ngnawe, S. Sahoo, Y. Pequignot, F. Precioso, and C. Gagné. ***Detecting Brittle Decisions for Free: Leveraging Margin Consistency in Deep Robust Classifiers***. In: Advances in Neural Information Processing Systems. [link]. **NeurIPS**. **2024**.

[3] M. Antony, S. T. Kakileti, R. Shah, S. Sahoo, C. Bhattacharyya, and G. Manjunath. ***Challenges of AI driven diagnosis of chest X-rays transmitted through smart phones: a case study in COVID-19***. In: Scientific Reports Journal. [link]. **Nature**. **2023**.

[4] S. Sahoo, P. Kumar, V. Shah, V. Kondameedi, A. Jain, A. Verma, C. Bhattacharyya, and V. Vishwanath. ***Dynamic to static lidar scan reconstruction using adversarially trained auto encoder***. In: Association for the Advancement of Artificial Intelligence Conference. [link]. **AAAI**. **2021**.

[5] M. ElAraby, S. Sahoo, Y. Pequignot, P. Novello, and L. Paull. ***GROOD: GRadient-aware Out-Of-Distribution detection in interpolated manifolds***. In: *Transactions on Machine Learning Research*. [link] **TMLR**. **2025**.

[6] P. Kumar, D. Vattikonda, V. B. S. Nadkarni, E. Dong, and S. Sahoo. ***Differentiable SLAM Helps Deep Learning-based LiDAR Perception Tasks***. In: British Machine Vision Conference. [link]. **BMVC**. **2023**.

[7] J. Li, R. Wang, Y. Lai, C. Shui, S. Sahoo, C. X. Ling, S. Yang, B. Wang, C. Gagné, and F. Zhou. ***Hessian Aware Low-Rank Weight Perturbation for Continual Learning***. In: Transactions on Knowledge and Data Engineering Journal. [link]. **TKDE**. **2023**.

[8] S. Sahoo, F. Zhou, Y. Pequignot, J. Ngnawe, F. Precioso, and C. Gagné. ***Domain Generalization by Minimizing Out-of-Distribution Detection***. In: Montreal AI Symposium. [link]. **MAIS**. **2022**.

[9] F. Aryan, D. Vattikonda, E. Dong, and S. Sahoo. ***Grad-lidar-SLAM: Fully differentiable global SLAM for lidar with pose-graph optimization***. In: IROS Workshop on Probabilistic Robotics in the Age of Deep Learning. [link]. **IROS Workshop**. **2022**.

[10] M. Sandhu, Y. Pequignot, S. Nashed, S. Sahoo, and L. Paull. ***CLIP-Enhance: Improving CLIP Zero-Shot Classification via von Mises-Fisher Clustering***. In: [link]. *Under review*. 2026.

[11] J. Ngnawe, S. Sahoo, Y. Pequignot, F. Precioso, and C. Gagné. ***Robust Fine-Tuning from Non-Robust Pretrained Models***. In: [link]. *Under review*. 2026.

PREPRINTS

[12] K. Samanta, S. Sahoo, and C. Gagné. ***Test Time Adaptation as an Adversarial Defense Strategy***. Internship Report. [link]. 2023.

[13] A. Verma, S. Sahoo, and C. Gagné. ***Diffusion based Pseudolabeling under Distribution Shifts***. Internship Report. [link]. 2023.

[14] D. Tiwari, R. Shah, S. Sahoo, and C. Bhattacharyya. ***Enhancing Explainability in Medical Images using Global Methods***. Masters Thesis. [link]. 2022.

[15] G. Parashar, S. Sahoo, and C. Bhattacharyya. ***Adversarial Robustness for Local Interpretable Methods***. Masters Thesis. [link]. 2021.

[16] S. Sahoo and K. Sellami. ***Automated Microservice Extraction using Reinforcement Learning***. [link]. 2021.

[17] D. Shanbag, S. Sahoo, C. Bhattacharyya, and V. V. ***An Approach For Accurate Sceneflow Prediction for LiDAR-based Sensors***. Masters Thesis. [link]. 2020.

[18] S. Sahoo and V. Kondameedi. ***Establishing Semantic relationships among Object Classes using Deep Networks for Image Classification***. [link]. 2015.

[19] S. Sahoo and V. Kondameedi. ***Hybrid Execution of Travelling Salesman Problem***. [link]. 2015.

THESES

[20] S. Sahoo, F. Precioso, and C. Gagné. "***Test-time Out-of-Distribution Generalization***". PhD Proposal. [link]. Mila/Université Laval, 2022.

[21] S. Sahoo and S. S. Vadhiyar. "***Hierarchical Task Mapping on Dragonfly topology for Scaling Molecular Dynamics***". Masters Thesis. [link]. IISc Bangalore, 2016.

[22] S. Sahoo, M. N. Yadav, V. Savalia, R. Soni, R. Agarwal, N. Lomash, and H. B. Naik. "***Thermoacoustic Energy Conversion Using Piezoelectric Diaphragm/Bi-Morph***". Bachelors Thesis. [link]. SVNIT Surat, 2014.

PROJECTS

[23] S. Sahoo, S. Karami, A. Safarnejadian, and A. Tupper. ***Deep Ensemble Methods for Vehicle Classification***. [link]. Université Laval. 2021.

[24] S. Sahoo, R. Shah, S. T. Kakileti, C. Bhattacharyya, and G. Manjunath. ***A new AI-driven platform will facilitate early-COVID interventions over Whatsapp***. [link]. Department of Science and Technology, Government of India. 2021.

[25] T. Varshney, S. Sahoo, V. Kondameedi, and C. Bhattacharyya. ***DCT-VAE: Capturing Low-level and High-level Features for Image Generation***. [link]. IISc Bangalore. 2021.

[26] S. Sahoo, A. Jain, R. Shah, and C. Bhattacharyya. ***Improving Automatic Concept Extraction for Global Model Explainability***. [link]. Niramai Health Analytix. 2021.

[27] V. Kondameedi, S. Shet, A. Verma, S. Sahoo, P. Kumar, C. Bhattacharyya, and S. Biswas. ***Frugal Advanced Driver Assistance System (ADAS) for Indian Roads***. [link]. TATA Motors. 2020.

[28] S. Sahoo, P. Kumar, C. Bhattacharyya, and V. V. ***Proximal Pose Search for Adapting SLAM in Dynamic Environments on Slow Moving UGVs***. [link]. Ati Motors. 2019.

TEACHING AND LEADERSHIP ROLES

- **Teaching Assistant**, Machine Learning course, Université Laval (2022/2023/2024): Involved with designing and grading quizzes/homeworks, and conducting tutorial sessions.
- **Student Mentor**, SHARE Research Labs (2020-21): Teaching and mentoring students for working towards a research paper for top-tier conferences.
- **Organizer** of various reading groups on topics like Machine Learning (2022-23), Out-of-Distribution (2022-25), and Autonomous Navigation (2019-20).
- **Placement Coordinator**, IISc (2015-16): Invited, organized, and coordinated on-campus placement for numerous industries and startups.

HONORS AND AWARDS

- Awarded IID Excellence scholarship 2022.
- Secured research funding from DEEL (2021-Present).
- Secured research funding from Ati Motors (2019-21), and ARTPARK (2021).
- Awarded distinction for my master's thesis.
- Won various competitions: 1st place in NeurIPS 2017 Challenge, top finalist in NVIDIA Reinforcement Learning Competition 2018, 3rd place in SO1 Product Recommendation Competition 2018.