PARIBESH REGMI

Graduate Research Assistant — Rochester Institute of Technology

Rochester, NY, 14623

✉ : pr8537@rit.edu   ✳   🔗 : www.regmiparibesh.com.np

EDUCATION

Computing and Information Science   *Rochester Institute of Technology*
PhD Degree   *08/2021 - Present*
Advisor: Prof. Rui Li
Relevant courses: *Statistical Machine Learning, Deep Learning, Deep Learning Security, Foundations of Algorithms, Software Engineering*

Electronics and Communication Engineering   *IOE, Tribhuvan University*
Bachelor's degree   *2014 - 2018*
Thesis: *Nepali Speech Recognition Using RNN-CTC Model*

WORK EXPERIENCE

Lab of Use-Inspired Computational Intelligence (LUCI)   *2021 - Present*
Graduate Research Assistant   *Rochester, New York*

Amazon.com Inc.   *Jun - Sep 2025*
Applied Science Intern   *Santa Cruz, California*

LogPoint   *2018 - 2021*
Solutions Engineer   *Lalitpur, Nepal*

- *Solved system/software issues at the customer's end.*
- *Troubleshooted/maintained system and software associated to cybersecurity, networking, Linux, user and entity behavior analysis (UEBA).*

RESEARCH INTERESTS

Machine Learning, Deep Learning, Bayesian Methods, Generative Models - (VAE, Diffusion/Flow Models), Graph Neural Networks, Vision-Language Models (VLMs)

PAPERS/PUBLICATIONS

Shortcut Diffusion Training with Cumulative Consistency Loss: An Optimal Control View
Paribesh Regmi; Sandesh Ghimire; Rui Li
International Conference on Learning Representations (ICLR), 2026

Bayesian Neighborhood Adaptation for Graph Neural Networks
Paribesh Regmi; Rui Li; Kishan KC
Transactions of Machine Learning Research (TMLR), 2025

AdaVAE: Bayesian Structural Adaptation for Variational Autoencoders
Paribesh Regmi; Rui Li
Thirty-Seventh Conference on Neural Information Processing Systems (NeurIPS), 2023

Predicting Biomedical Interactions with Probabilistic Model Selection for Graph Neural Networks
Kishan KC; Rui Li; Paribesh Regmi; Anne Haake
arxiv.org

Nepali Speech Recognition Using RNN-CTC Model
Paribesh Regmi; Arjun Dahal; Basanta Joshi
International Journal of Computer Applications, 2019

RESEARCH/PROJECTS

Adaptive Token Pruning for Efficient Inference in Video-Language Models *Summer Internship 2025*
- *Developed an algorithm to efficiently process video inputs in vision-language models by reducing the number of visual tokens for LLM to process.*
- *The algorithm prunes out the redundant visual tokens in two stages: frame and token-level pruning. We also determine the number of tokens to prune based on the dynamism in the video content.*
- *This allows the deployment of VLMs in low-resource edge devices.*

Efficient Sampling in Diffusion Generative Models *2024 - Present*
- *Developing algorithm to improve sampling efficiency in diffusion-based generative models.*
- *The algorithm allows training for a few-step denoising along with the base diffusion model in a single training pass.*

Representation learning on graphs *2022 - 2024*
- *Applying Bayesian model selection to enhance graph representations by inferring appropriate neighborhood scope for message aggregation in a graph convolutional network (GCN).*
- *Using graph characteristics to infer the most plausible set of neighbors for message aggregation in a GCN.*
- *Expressivity analysis shows that our approach improves the expressivity of a GCN with larger layer depth.*

Bayesian model selection in VAE *2022 - 2023*
- *Developing a Bayesian model selection framework to infer an optimal model structure in variational autoencoders, guided by the data*
- *The framework eliminates the need to fine-tune network complexity for the encoding and decoding networks*
- *The framework is compatible with the state-of-the-art VAE regularization methods as well as various VAE variants, further improving their performance*

Leveraging deep learning in graphs for biomedical interaction prediction *2021 - Present*
- *Application of developed graph algorithms to real-world biomedical scenarios, like inferring the interactions in the datasets like PPI(Protein-Protein Interaction), DTI(Drug-Target Interaction), etc.*

Nepali Speech Recognition *2018 - 2019*
- *Application of deep learning to enhance the Nepali speech recognition system, transitioning from a limited vocabulary size to a large corpus. Connectionist Temporal Classification (CTC) loss aided in enabling end-to-end training of the recurrent neural network model.*
- *Defined a Nepali language character set of 67 characters.*

AWARDS

Fully funded Ph.D./ Research Assistantship at RIT *2021 - Present*
Full financial support for my Ph.D. from NSF grants

Fusemachine AI Fellowship Award *2017 - 2018*
Fellowship offered by Fusemachines (fusemachines.com) for AI and Machine Learning study

Full Scholarship for Bachelor's in Engineering *2014 - 2018*
Ranked 28th among 13,000 applicants in the engineering entrance examination to gain a full scholarship

SKILLS

Programming — **General:** Python, moderate expertise in Java and C++; **ML and DL:** pytorch, scikit-learn, numpy; **Visualization:** matplotlib

Troubleshooting — Solving system(Linux) and software related issues. Three years of work experience in troubleshooting.

Leadership — Former event manager at Nepalese Student Association, Rochester Institute of Technology (NSA-RIT)