Prisha Priyadarshini

609-751-6691 | prishapriya31@gmail.com | LinkedIn | Github | Personal Website

Education

Rutgers University – New Brunswick — New Brunswick, NJ

Bachelor of Science in Computer Science, Minor in Mathematics GPA: 3.65/4.00 — *Expected May 2027*

Dean's List – Spring 2026

Research Experience

Lead ML Researcher / Rutgers MBS Exchange — Jan. 2026 – Present

The Center for Discrete Mathematics & Theoretical Computer Science (DIMACS) — *New Brunswick, NJ*

- Led an empirical study under Dr. Linda Ness comparing interpretable decision tree models (**SPLIT, LicketySPLIT, LicketyRESPLIT, GOSDT**) with SOTA boosting methods (**XGBoost, CatBoost, LightGBM**) across 6 datasets.
- Demonstrated that preprocessing with **ThresholdGuessBinarizer (TGB)** eliminates performance gaps between boosting models and interpretable decision trees, achieving comparable accuracy while preserving interpretability.
- Analyzed the impact of **SMOTE** on **Rashomon set size**, showing that class balancing significantly expands model equivalence classes.
- Quantified **performance–complexity tradeoffs** using accuracy, recall, runtime, and Rashomon set size, finding that increased model complexity yields marginal performance gains but substantially larger Rashomon sets.

AI Researcher (Multi-Agent Deliberation & Consensus Dynamics) — Sept. 2025 – Present

Algoverse AI — *Remote*

- Built a multi-agent framework to study consensus formation, showing that **model-to-model deference drives convergence over independent reasoning**.
- Ran large-scale 20-round deliberation experiments across **GPT-4.1, Mistral, LLaMA, and Phi** models on subjective and objective benchmarks.
- Introduced quantitative metrics (Inter-agent disagreement rate, model deference rate) to analyze interaction dynamics, identifying **hierarchical influence patterns (up to 80% small to large model deference).**
- Developed a rotation-based framework to separate model identity from response quality, revealing that **large-to-small model deference rates increase relative to baseline**, slightly weakening hierarchical influence patterns.
- Showed that prompting strategies can control multi-agent behavior, **increasing disagreement and reducing deference, with implications for system reliability.**
- **Accepted to the ICML 2026 Pluralistic Alignment workshop** and currently under review at the ICML 2026 AI4GOOD workshop.

AI Researcher (Multimodal Machine Learning) — June 2025 – Dec. 2025

Algoverse AI — *Remote*

- Contributed to a hierarchical scene-captioning pipeline **integrating dynamic stride window selection and multimodal chain-of-thought (MMCoT) reasoning** for temporally coherent caption generation.
- Designed and implemented evaluation frameworks (BLEU-4, METEOR, CIDEr, BERTScore) and conducted ablation studies on frame sampling **(5, 20, 40)** and model aggregators **(Phi, Mistral, Qwen3).**
- Ran large-scale YouCook2 experiments on 30 RTX A6000 GPUs and achieved **17% higher CIDEr than GPT-4o** and **14% over VLLaMA-3**.
- Accepted for an **oral presentation at NeurIPS 2025 (7HVU) and accepted to AAAI 2026 (AI4EDU).**

Algorithms Research Shadow — Jan. 2025 – May 2025

The College of New Jersey — *Ewing, NJ*

- Researched classical and modern folding algorithms under Dr. Dimitris Papamichail (Nussinov, Zuker, LinearFold, MCTS).
- Implemented sparse dynamic programming to prune redundant structures and reduce runtime.
- Conducted SLURM-based large-scale experiments using ViennaRNA across thousands of RNA sequences.

Publications

- **Directional Influence and Consensus Formation in Multi-Agent Systems**
  Prisha Priyadarshini*, Aryan Shrivastava*
  Accepted to ICML 2026 Pluralistic Alignment Workshop
- **DynaStride: Dynamic Stride Windowing with MMCoT for Instructional Multi-Scene Captioning**
  Eddision Pham, ***Prisha Priyadarshini****, Adrian Maliackel, Kanishk Bandi, Cristian Meo, Kevin Zhu*
  Accepted to NeurIPS 2025 7HVU Workshop (Oral), AAAI 2026 AI4EDU Workshop

Industry Experience

Incoming Machine Learning Intern May 2026
Regeneron *Troy, NY*

- Incoming Summer 2026

Open Source Contributor (DeepChem) May 2026 – Present
Google Summer of Code *Remote*

- Selected as **1 of 5 out of 200+ applicants** for **Google Summer of Code 2026** under **DeepChem** to extend support for large-scale molecular language models and generative chemistry workflows.
- Extending DeepChem's **HuggingFaceModel** infrastructure with **generation, batching, causal LM, and distributed training (FSDP) support** for large decoder-only models including **OLMo 7B**.
- Developing a **multi-step molecular optimization pipeline** using **MoLFormer/ChemBERTa**-style models with iterative masking, constrained candidate generation, **QED optimization**, and **Tanimoto similarity**-based trajectory search for molecular generation tasks.
- Implementing and testing scalable molecular generation workflows in DeepChem, including batched inference, RDKit-based molecular validation, optimization trajectory tracking, and benchmarking preparation with **GuacaMol**.

AI Ignite Fellow May 2026
AI4ALL *Remote*

- Selected for a competitive fellowship with the AI4ALL Ignite Summer 2026 Accelerator.
- Working on a real-world AI/ML portfolio project to present to industry experts at the Ignite Symposium.

AI Associate Developer Oct. 2025 – Jan. 2026
Insurity – SpatialKey Team *Remote*

- Built **geospatial preprocessing pipelines over 6M+ records**, integrating climate and peril data via time alignment and spatial filtering.
- Engineered features using Pandas and GeoPandas to support ML models.
- Built and evaluated LightGBM-based multi-class peril classifiers; **applied SMOTE class balancing, increasing accuracy and F1 score by 10%.**
- Experimented with a Temporal Fusion Transformer (TFT) model, modeling seasonal dependencies via cyclical week encoding and trained GPU-accelerated PyTorch models in a managed Anaconda environment on NVIDIA RTX hardware.
- Additionally, developed a geospatial change-detection system for before/after wildfire satellite imagery using OpenCV and U-Net++ for high-resolution segmentation.

Projects

Knowledge Distillation for Reasoning in MoE and Dense LLMs | *Python, PyTorch, HuggingFace* April. 2026

- Developed an offline knowledge distillation pipeline using **LoRA (< 1% trainable parameters)** to transfer chain-of-thought reasoning and token-level distributions from large LLMs to a 3B student via **sequence and logit distillation (KL divergence + temperature scaling)**, achieving **4.5% accuracy gains** and **96%+ BERTScore** on **SciBench and TheoremQA.**

GenreBlender | *Python, PyTorch, Streamlit* Feb. 2026

- Engineered a Generative AI system combining Meta's MusicGen and **a 4-layer PyTorch Multilayer Perceptron (92% validation accuracy)** to create and quantitatively evaluate controllable music genre blends via a weighted probability framework in a Streamlit app.

PocketRAG | *Python, Flask, FAISS, Gemini, AWS, Docker, RAG* Sept. 2025

- Engineered a production RAG pipeline with document chunking, **FAISS vector retrieval**, and Gemini inference; containerized and deployed on **AWS EC2 (Docker + Gunicorn)** achieving **sub-second latency.**

Professional Development & Extracurriculars

Workshop Reviewer ICML 2026 Pluralistic Alignment Workshop
May 2026
- Reviewed and evaluated research submissions for the ICML 2026 Pluralistic Alignment Workshop through the OpenReview peer review process.

Mathematics Grader Rutgers University – New Brunswick
Dec. 2025 – Present
- Grade homework and exams for Multivariable Calculus across 3 sections.

Peer Tutor The College of New Jersey
Sept. 2024 – May 2025
- Provided one-on-one tutoring for Data Structures, Algorithms, Linear Algebra, and Statistics.

Recruitment Chair Kappa Theta Pi
Dec. 2024 – May 2025
- Led recruitment efforts and onboarded 10 new members for Spring 2025.

Student Athlete (NCAA Division III Tennis) The College of New Jersey
Aug. 2023 – May 2024
- Balanced 20+ hrs/week training with academics; earned All-NJAC recognition.

Technical Skills

AI & ML: PyTorch, Hugging Face, OpenCV, Transformers, scikit-learn
Research Domains: Multi-agent Systems, Interpretable ML, Multimodal ML
Programming Languages: Python, C, C++, JavaScript, Java
Libraries: NumPy, Pandas, GeoPandas, Matplotlib, Seaborn
Cloud, Systems, & MLOps: GitHub, AWS, Azure, Docker, CI/CD, pytest, Jira, Anaconda, Vercel
Certifications: AWS Certified Machine Learning — Specialty, AWS Certified Cloud Practitioner