

Arnav Rastogi

arnavrastogi.543@gmail.com | arnav.rastogi.net.in | linkedin.com/in/arnav-rastogi-6b868a22a | github.com/ARC345

Research Interests

DeepRL, Transformer Architectures, TRM, HRM, SSMs, Subquadratic Architectures, Linear-Time Sequence Models, Hardware-Aware Selective Scan, Structured State Space Duality (SSD), Hybrid Attention-SSM Models, Mechanistic Interpretability.

Education

Indian Institute of Technology, Jodhpur, BS in Applied AI and Data Science Jan 2025 – Jan 2029

- Core focus on Deep Learning, Linear Algebra, and Statistical Inference.

MITx, MicroMasters in Statistics and Data Science Sept 2024 – Sept 2026

- Advanced training in Probabilistic Modeling, Machine Learning foundations, and Time Series Analysis.

Projects

LLM Lab Jan 2025 – present

- Engineered a from-scratch character-level GPT implementation using PyTorch to study the internal dynamics of **Multi-Head Attention** and **Transformer blocks**.
- Implemented residual connections and layer normalization to optimize gradient flow and training stability.

Tailor (AI Framework) 2025 – present

- Built a modular AI framework using **Rust (Tauri)** and **Python**, enabling users to stitch together models and tools into custom assistants.
- Developed a highly customizable chat interface designed to be "tailored to you", adapting to individual user workflows and preferences.
- Implemented a bi-directional event system and plugin architecture, allowing seamless integration of custom tools and extensions.

ARC-Engine 2023

- Developed a high-performance OpenGL-based rendering engine featuring a custom **Entity Component System (ECS)** architecture for optimized memory layout and cache locality.
- Designed a real-time visualization pipeline using kernels, demonstrating expertise in low-level systems and compute-heavy architectures.

Lenia & Conway's Game of Life 2024

- Simulated artificial life forms using **Gaussian kernel evolution** and non-linear partial differential equations.
- Optimized grid update performance using OpenCV and kernel-based filtering, achieving real-time high-resolution simulations.

Sudoku Solver (Algorithms) 2023

- Investigated the efficiency of **Backtracking vs. Bitmasking** approaches for solving large-scale combinatorial optimization problems.
- Optimized search space using heuristics, reducing computation time for complex puzzles by 80%.

Technical Writing

Can Mamba Learn, Unlearn, and Retain Noise? (link): Feb 2026

Multi-Hop Reasoning in Transformers (link): Jan 2026

ReLU vs GELU (link): Jan 2026

Experience

Research Intern, Abhay Travels – Feb 2025 – Mar 2025

- Researched demand forecasting models for time-series data, engineering 15+ features to capture seasonal and non-linear trends.
- Built a predictive pipeline using **CatBoost**, outperforming baseline ARIMA models in RMSE on 5 years of historical booking data.

Data Analyst Intern, AMR India – Nov 2024 – Dec 2024

- Conducted statistical analysis on 4 years of international port data to forecast article-specific demand cycles using HSN codes.

Skills

Mathematical Foundations: Linear Algebra, Probability & Statistics, Optimization, Calculus

AI & Systems: PyTorch, JAX (learning), Transformers, CNNs, Reinforcement Learning

Engineering: C++, Python, SQL, OpenGL, CUDA (basics), Git