

AARYAN NAGPAL

aaryannagpal@gmail.com | +91-9599799022
linkedin.com/in/aaryannagpal | github.com/aaryannagpal | aaryannagpal.github.io

Technical Proficiency: Python, C, C++, Time Series Foundation Models, Synthetic Data, Pandas, TensorFlow, LLMs, RAG, BERT, PostgreSQL

EDUCATION

P.G. Diploma in Adv. Studies and Research	Ashoka University	Comp. Science; Math Minor	2025	3.69/4
B.Sc. (Hons.)	Ashoka University	Comp. Science; Physics Minor	2024	3.65/4

PROFESSIONAL EXPERIENCE

Birla AI Labs, Aditya Birla Group | AI Engineer Jun 2025 – Present

- Build synthetic data infrastructure for an in-house Time Series Foundation Model using mixup, transforms, and generator augmentation.
- Drive synthetic data research by probing model-derived data, testing generator strategies, and evaluating real-synthetic training mixtures.
- Develop a Python evaluation pipeline to benchmark TSFM performance across datasets, metrics, benchmarks, and experimental data regimes.

HP India | Software Engineer Intern Jun 2022 – Nov 2024

- Built tree- and regression-based ML models with Pandas to predict printer support issues, improving prediction accuracy from 50% to 80%.
- Automated manual call-audit review using a BERT-based model using TensorFlow, improving review consistency across 100+ call transcripts.
- Built a Python-Flask retrieval system for JSON manufacturing documents as part of a broader workflow to automate documentation processing.

RESEARCH EXPERIENCE

FoodComputing Lab | FKG.in Project May 2024 – Aug 2024

- Implemented a RAG-based query system over GraphDB for Indian food knowledge graphs supporting health- & finance-aware food decisions.
- Reduced repeated LLM calls by 50% using SentenceTransformers caching and integrated Translate and Bhashini APIs for multilingual access.

Ashoka University | Folk Computing Project Mar 2023 – Aug 2023

- Developed a multilingual health chatbot with acoustic interfaces, enabling local-language healthcare communication for underserved users.
- Researched modeling and inference techniques under Dr. Jia Li and Prof. Ramesh C. Jain, collaborating with Stanford, HealthUnity, and UCL.

PROJECTS

Clinical NL2SQL for MIMIC-IV | NLP and Healthcare Jan 2025 – May 2025

- Designed a dual-model pipeline using Qwen 2.5 and Phi-4 to translate natural-language clinical questions into executable PostgreSQL queries.
- Benchmarked 9 LLMs across prompting strategies on EHRSQL 2024, analyzing query correctness and healthcare-specific SQL performance.

ICU Outcome Prediction in Sepsis | ML and Healthcare Jul 2024 – Dec 2024

- Analyzed MIMIC-IV sepsis cohorts to identify relevant clinical and administrative features for ICU mortality and length-of-stay prediction.
- Used selected clinical features to benchmark ICU outcome models and built a visualization tool for predictions, feature effects, and patient risk.

nanoC Compiler | Compilers and Programming Languages Aug 2023 – Dec 2023

- Built a nanoC compiler using Flex, Bison, and C, implementing lexical analysis, parsing, symbol tables, TAC, and core semantic translation.
- Extended the backend with activation-record handling and assembly generation from quads for functions, control flow, globals, and expressions.

PUBLICATIONS

Mix, Don’t Pick: Why Synthetic Corpus Composition Matters for TSFM Pretraining | FMSD Workshop @ ICML 2026 May 2026

- Framed TSFM pretraining as a corpus composition problem, showing synthetic data benefits are architecture-dependent across model families.
- Showed equal-weight synthetic mixtures match or outperform generators, while real-synthetic mixtures yield most robust pretraining corpora.
- Also accepted to GlobalSouth ML and Forecasting as a New Frontier of Intelligence Workshops @ ICML 2026 under a related title/version.

time2time: Simulating Rare Events in TSFMs | BERT²S Workshop @ NeurIPS 2025 Sep 2025

- Introduced activation transplantation to causally intervene in TSFM hidden states and simulate rare-event forecasts such as market crashes.
- Demonstrated evidence of a latent concept space governing TSFM predictions, enabling semantic what-if analysis across Toto and Chronos.

TEACHING EXPERIENCE

Ashoka University | Teaching Assistant, CS 1102: Introduction to Computer Science Spring 2025

- Assisted Profs. Debayan Gupta and Aalok Thakkar for 115+ students, covering OCaml and C++, algorithmic thinking, and cryptography.

Ashoka University | Teaching Assistant, CS 1204: Data Structures and Algorithms Monsoon 2024

- Assisted Prof. Anirban Sen for 85+ students across C++ programming, arrays, linked lists, trees, graphs, recurrences, and correctness proofs.

Ashoka University | Teaching Assistant, CS 1206: Design and Analysis of Algorithms Spring 2024

- Guided 90+ students under Prof. Anirban Sen through randomized and greedy algorithms, dynamic programming, and NP-completeness.

AWARDS AND ACCOMPLISHMENTS

Cum Laude | Ashoka University 2024, 2025

Earned institutional honors for academic achievement across undergraduate and postgraduate study.

Teaching Assistantship Excellence Award | Ashoka University 2024

Recognized for outstanding mentorship and teaching assistance; selected from 30 TAs with a 5/5 student evaluation score.

Dean’s List | Ashoka University 2022, 2023, 2024

Recognized for strong academic performance across multiple semesters.