Aryan Thakur

aryannnthakurrr@gmail.com | +91 7065220661
linkedin.com/in/aryannnthakurrr | aryanthakur.vercel.app | github.com/Aryannnthakurrr

Education

University School of Information and Communication Technology (USICT) New Delhi, India
B.Tech in Computer Science (Artificial Intelligence) CGPA: 9.0/10.0 | 2024 - 2028
Titiksha Public School (CBSE)
Senior Secondary (Class 12): 84.2% | Secondary (Class 10): 94.4%

Experience

AI Engineer Intern | PepperLabs AI Solutions, Gurgaon JUN 2026 – AUG 2026

- Built the AI core and backend of BookBuddy, a live RAG tutor for CUET-UG that cites NCERT; parsed 126 chapters into 2,221 heading-aware chunks, linking 3,875 tables, formulas and figures to their source.
- Indexed chunks with LLM-generated student questions for FAISS + BM25 retrieval with confidence-gated re-retrieval; cut token use by about 60% in simulated tests via lighter models and rule-based routing.
- Developed a provider-agnostic LLM layer with retries and token tracking, and tier-priority scheduling with per-user concurrency caps, auth and daily quotas per paid tier.

Software Intern | Defence Research and Development Organisation (DRDO) JAN 2026 – FEB 2026

- Engineered a voice processing pipeline with speech recognition, RAG and a FastAPI backend.
- Earned a Letter of Recommendation from DRDO for technical excellence and on-time delivery.

Projects

Sorted - Household Admin Assistant | GitHub OCT 2026 – Present

- Built an assistant for bills, subscriptions and appointments that learns each message type with an LLM, then compiles it into a Python parser tested on held-out cases and run in an offline sandbox.
- Ran parsers in shadow mode until they matched the LLM, with fallback and regeneration when a format changes; compiled code rose from 0% to 62% of the workload over six simulated months.

LLM-Augmented Hybrid Retrieval for Informal Queries | GitHub APR 2026 – NOV 2026

- Compared three pipelines on a shared hybrid backend (BM25 + BGE-large in FAISS): plain hybrid, an LLM-enriched index, and LLM query rewriting over that index, for informal queries in a niche marketplace.
- Offline enrichment (6 buyer-style queries per listing) raised mean relevance by 9.8% and MRR by 9.9% over plain hybrid in pooled LLM-as-a-judge tests, with highly relevant results up from 37% to 47%.
- Rewriting added only 1% more: it helped vague queries but hurt 20 of 50 by adding constraints the user never stated, which supports enriching by default and rewriting only ambiguous queries.

Technical Skills

- **Languages:** Python, TypeScript, JavaScript, C++, SQL
- **Technologies:** FastAPI, MongoDB, PostgreSQL, FAISS, Sentence Transformers, Gemini API, Git
- **Core Concepts:** LLMs, RAG, Semantic Search, BM25 Ranking, Vector Embeddings, REST APIs, Agentic AI

Certifications & Coursework

Oracle Certified AI Vector Search Professional OCT 2025
Relevant Coursework: DBMS, Operating Systems, Data Structures & Algorithms, OOP

Leadership & Activities

Technical Team Leader | Hackathons (HackGenesis, HackCBS, CTRL ALT HACK, HackTU)

- Won 1st prize at HackGenesis (Christ University) and its AI quiz; finalist at CTRL ALT HACK and HackTU.
- Led teams of 3–4 at national-level and MLH-backed hackathons, owning architecture and final pitches.