

# Abinesh Mathivanan

[✉ abineshmathivanan31@gmail.com](mailto:abineshmathivanan31@gmail.com) [linkedIn linkedin.com/in/abineshmathivanan](https://linkedin.com/in/abineshmathivanan)  
[GitHub github.com/Abinesh-Mathivanan](https://github.com/Abinesh-Mathivanan) [Website abinesh-mathivanan.vercel.app](https://abinesh-mathivanan.vercel.app)

## Education

<b>KIT-Kalaignarkarunanidhi Institute of Technology</b> B.Tech in Artificial Intelligence and Data Science (CGPA: 8.74 / 10.00)	Expected June 2026 Coimbatore, TamilNadu
<ul style="list-style-type: none"><li><b>Relevant Coursework:</b> Data Structures and Algorithms (C), Data Science and Intro to ML, OOPs (Python), Discrete Mathematics, Probability and Queuing Theory, Database Management Systems, Networking &amp; TCP/IP</li></ul>	

## Experience

<b>RSystems</b> Agentic AI Engineer	June 2025 Noida, India
<ul style="list-style-type: none"><li>Built Agentic AI applications using CrewAI, Autogen, Langgraph and SmolAgents</li><li>Involved in writing inference kernels to deploy local LLMs using SGLang, LMCache, and Triton. Improved NER Pipeline to have process time &lt;2s per request</li><li>Built and trained foundational models such as LG-Exaone with RL for coding and mathematical tasks</li></ul>	
<b>PhobosQ</b> Software Engineer	May 2024 – Apr 2025 Coimbatore, India
<ul style="list-style-type: none"><li>Engineered the Luna project interface using TypeScript, React, and ShadCN UI, orchestrating AI frameworks like Langchain and Gemini APIs.</li><li>Implemented multi-threaded word document processing, achieving 200+ pages parsed in under 58 seconds and voice-enabled systems with &lt;3 ms latency.</li></ul>	

## Technical Skills

**Languages:** Python, C++, Go, Javascript, Typescript, Tile-lang, SQL

**Frameworks:** PyTorch, Triton, Jax, TensorFlow, Keras, Tailwind, Next.js, Astro.js

**Databases:** MySQL, MongoDB, Postgres, ChromaDB, Cassandra

**DevOps & Tools:** Git, Github, Docker, Azure ML, Firebase, Supabase, Node.js, Bun, ONNX Runtime, EC2 instance, Linux

**Focused on:** Inference Optimization, Foundational Model Engineering, Distributed Training, Building Frameworks, Engineering Networking architectures and XLA-Programming

## Projects

<b>DeepCode – LeetCode for ML</b>   Next.js, TypeScript, Python, Rust, Docker, Firebase, Google Bigtable, Cassandra	↗
<ul style="list-style-type: none"><li>Built a ML problem-solving platform with a Next.js and a scalable backend with Python and Rust for code execution, containerized with Docker.</li><li>Deployed using Firebase (backed by Google Bigtable) and Render, implementing distributed caching systems ensuring 99% availability for real-time inference. Achieved 1K+ API hit requests within the first 24 hours of launch.</li></ul>	
<b>Beens-MiniMax - LLM</b>   Python, PyTorch, HuggingFace, Triton, LaTeX	↗
<ul style="list-style-type: none"><li>A 103-million-parameter SLM with hybrid attention mechanism (softmax &amp; lightning) following MoE style architecture. Trained with Wikitext-103 for the base model and Ultrachat_200k for Instruction-SFT</li><li>Trained in Kaggle 2x T4 GPUs for 15 GPU hours &amp; Instruct-SFT trained for 6 GPU hours. Model weights and technical report was published in Github and Kaggle</li></ul>	
<b>go-torch</b>   Go, Python, OpenCL C, BLAS	↗
<ul style="list-style-type: none"><li>A simple torch-like deep learning library written in Go. Implemented go-routines and BLAS to the core, performing 1024 x 1024 in 118ms</li><li>Made the library 115x faster compared to the first version and wrote custom Intel kernels. Crossed 100+ stars in GitHub.</li></ul>	
<b>in-love.js – JavaScript Library</b>   JavaScript, TypeScript, Node.js, npm, Git, Tailwind	↗
<ul style="list-style-type: none"><li>Created an open-source JavaScript library for love-proposal sites, using node modules such as Babel.js, Webpack, and Puppeteer. Testing modules such as Mocha, Jest, and Chai were implemented for real-time testing across webs.</li><li>Published on the npm official platform and crossed 100+ downloads within the first week of launch.</li></ul>	