NEELABHO CHAKRABORTY

AI ENGINEER · LOCAL-FIRST AI, OPEN-SOURCE & INTELLIGENT SYSTEMS

neelabhochakraborty20@gmail.com · GitHub · India · Open to relocation and remote work

SUMMARY

AI engineer building intelligent systems with a focus on local-first AI, privacy, and practical applications of LLMs. I turn ideas into working products, from on-device AI assistants to autonomous document-processing systems. Delivered two end-to-end products for paying clients in just two weeks each. Creator of **Nook**, an open-source native macOS application bringing local AI directly to the desktop.

Currently pursuing a dual degree in **Data Science & AI at BITS Pilani (9.34 CGPA)** and **Computer Applications at Techno India University (8.35 CGPA)**.

EXPERIENCE

FREELANCE AI ENGINEER — 2026

Remote · Paid client engagements, each scoped, built and delivered in two weeks

- **Private document intelligence:** architected and shipped a system that turns unstructured invoices, receipts and PDFs into structured, searchable Excel and CSV records using OCR, Qwen 3 8B and a Python pipeline, deployed with its LLM runtime on the client's own machine so sensitive financial data never leaves the premises.
- **Trust-but-verify LLM layer:** engineered guardrails so model output is never trusted blindly; deterministic checks recompute invoice arithmetic, confidence thresholds route weak extractions to fallback pipelines, and duplicate detection blocks re-processed documents, so hallucinated values cannot silently reach the database.
- **Automated business intelligence (second client):** built a pipeline that ingests multiple Excel sources, cleans and consolidates them, computes KPIs, renders charts and emails a formatted weekly PDF report, replacing a manual reporting workflow end to end.

PROJECTS

NOOK — v1.0.0, Oct 2026

Open-source macOS app with a private on-device AI assistant

`Swift 6, SwiftUI, AppKit, Ollama · github.com/Neelabho/nook · MIT`

- Designed and shipped a 10,800-line native app that turns the MacBook notch into a command surface for a file shelf, a local LLM assistant and Spotify control; App Sandbox, hardened runtime, no telemetry, no dependencies.
- **LLM client:** streaming chat over Ollama's NDJSON API with batched 20 Hz UI updates, cancellable generation, bounded context, reasoning-tag parsing across chunk boundaries and delimited user input to resist prompt injection; privacy enforced in code by accepting loopback endpoints only and refusing cloud models.
- **Crash-proof storage engine:** database-style import transactions with staged copies, streaming SHA-256 verification, atomic no-overwrite commits and a self-healing index rebuilt from sidecar metadata, so no crash or full disk can lose a file.
- **Systems design:** window behaviour modelled as a pure 8-state reducer that returns side effects as data; Spotify sign-in via OAuth 2.0 with PKCE over an RFC 8252 loopback listener and Keychain token storage.
- **Performance and security:** about 0% idle CPU, under 2 wake-ups/s and 14 MB memory; CI fails the build on unexpected sandbox entitlements, with automated secret scanning.

LLM PAPER REPLICATIONS — In progress, Nov 2026

Reimplementing five LLM papers from scratch, with one results table for all five

`Python, PyTorch (MPS), Hugging Face transformers, Ollama`

- **Papers:** nucleus sampling (Holtzman et al., 2019), prompt-format sensitivity, LLM-as-a-judge with error bars, ColPali visual document retrieval, and efficient guided generation (Willard and Louf, 2023).
- **What it builds:** a custom decoding loop on GPT-2 Large, a benchmark harness with variance and confidence intervals, a judge-agreement study with a local Qwen 3 8B judge, a late-interaction retrieval scorer evaluated by nDCG, and an FSM-constrained decoder for structured output.

URBAN FLOOD INTELLIGENCE — Live, Nov 2026

Geospatial AI platform for urban waterlogging, piloted in Kolkata

`Python, XGBoost, PyTorch, SHAP, GeoPandas, FastAPI, React, MapLibre · neelabho.github.io/urban-flood-intelligence`

- Combined 12 public datasets (Sentinel-1 SAR, Copernicus DEM, IMD, CHIRPS, ERA5, OpenStreetMap) with a source-cited waterlogging inventory to rank 250 m cells; seven models and three baselines under spatial block cross-validation, held-out ROC-AUC 0.86, with SHAP explanations and bootstrap uncertainty.
- Showed that rainfall flags waterlogging days (ROC-AUC 0.91) but not 24 h ahead, and that Sentinel-1 cannot label urban waterlogging; shipped as a ten-page public web app with live rainfall.

OPENLEDGER — In development

Company intelligence for Indian public companies

`Python, PDF parsing, MCP`

- A platform being built to turn the annual reports of Indian public companies into structured company intelligence.

OTHER BUILDS

- **SmartQueue AI:** (React, Supabase, Twilio; solo MVP): multilingual WhatsApp patient-queue platform with a clinic dashboard.
- **Presentation Maker:** (macOS, Windows 11): desktop app in active use by a working counsellor; builds formatted presentations from deterministic layout rules.

EDUCATION

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE (BITS), PILANI — 2025 – 2029 (expected)

BS, Data Science and Artificial Intelligence Development · CGPA 9.34/10

TECHNO INDIA UNIVERSITY — 2024 – 2028 (expected)

Bachelor of Computer Applications (BCA) · CGPA 8.35/10

UNIVERSITY OF MANITOBA, CANADA — 2021 – 2024

Coursework toward BSc Computer Science

TECHNICAL SKILLS

LANGUAGES	Python (pandas, scikit-learn), Swift, TypeScript/JavaScript, SQL, Java, C++, C
AI AND LLMS	Local LLM deployment (Ollama, Qwen 3 8B), streaming inference, structured extraction, LLM output validation and guardrails, prompt design, MCP, RAG, OCR, PDF parsing (MinerU), Excel and PDF report automation
ENGINEERING	SwiftUI, AppKit, React, Supabase, Twilio, OAuth 2.0 (PKCE), Git, GitHub Actions CI

ADDITIONAL

AWARD	Top 3, Web Design, Science, Technology and Innovation (STAI) competition, Oman, 2018–19
VOLUNTEERING	Statistics outreach, University of Manitoba: taught probability to school children with a Monte Carlo demo
OPEN SOURCE	River (online machine learning), merged PR #2086, Nov 2026: implemented the fix for additive seasonal initialisation in the Holt-Winters forecaster, with regression tests and a before-and-after benchmark (early forecast error on additive data fell from 5.98 to 2.23 MAE)
LANGUAGES	English, Hindi, Bengali (fluent); Arabic, French (understand)