

Bibek Dhakal

Associate AI / Machine Learning Engineer

Kathmandu, Nepal | imbibek8366@gmail.com | linkedin.com/in/bibek-dhakal-771ba5334 | github.com/Bibek-Dhakal

SUMMARY

Entry-level ML engineer with about one year of cumulative software engineering experience (internships and contracts, 2023–2025), plus an ML internship and an ML apprenticeship in 2026. Focused on honest evaluation (baselines, leakage-safe validation) and on turning models into tested, containerized services. **Seeking:** a full-time Associate / Junior ML Engineer or AI Engineer role on a team where I can contribute to real systems and grow long term. Available immediately.

EXPERIENCE

AI / ML Engineering Intern, FlyRank AI Jul 2026 – Sep 2026

- Built a CTR opportunity-scoring model that ranks web pages for SEO metadata review; queried a ~78.8M-row Parquet warehouse with DuckDB and modelled on a 30k-row anonymized slice.
- Compared a hand-written rule, Logistic Regression and Random Forest on a client-holdout split: Random Forest reached Precision@50 of 0.64 vs. 0.26 for the rule (random chance 0.29).
- Showed a random split overstated performance (0.94 vs. 0.64 on unseen clients) and planted two deliberate leaks to confirm the validation caught them; published a research paper with limitations.

Data Science & ML Apprentice, Skill Shikshya Apr 2026 – Jun 2026

- Studied ML mathematics, classical ML and deep-learning fundamentals; built ML applications served with FastAPI and containerized with Docker; completed and defended a final project.

Software Engineering: Internships & Contracts (alongside BCA) Intermittent, 2023 – 2025

- Walkers Hive** (Intern, Oct–Dec 2025): designed and built the AcademiaOS MVP with a FastAPI and Celery backend, Next.js frontend, HTTP-only cookie auth and role-based access control, deployed with Docker.
- Nextwave Technology** (Contract, Apr–Jul 2025): shipped Flutter features and fixes and supported a Google Play Store launch; migrated corporate websites to Next.js.
- Walkers Hive** (Contract, Nov 2024–Apr 2025): built an e-commerce admin panel (React, MUI, Redux-Saga) and Next.js frontends integrated with PHP backends.
- CodSoft** (Android Development Intern, Dec 2023–Jan 2024): built Flutter apps with Firebase Authentication, local persistence and BLoC state management.

PROJECTS

Customer Churn Risk Intelligence | Python, scikit-learn, MLflow, FastAPI, Docker GitHub

- Pandera/Pydantic data contracts and MLflow tracking; Logistic Regression (5-fold CV ROC-AUC 0.85) beat Random Forest and LightGBM. Served via Dockerized FastAPI returning four risk tiers; CI/CD with GitHub Actions, Pytest, Ruff.

NanoTransformer & Forge-LM | PyTorch, ONNX, FastAPI, Docker GitHub

- GPT-2-style decoder (~28M parameters, TinyStories) using my own BPE tokenizer, LexiByte (on PyPI); exported to INT8 ONNX and served on CPU. Learning-scale project, not state-of-the-art.

Aegis Omnisearch Agent | Python, FAISS, ONNX Runtime, FastAPI GitHub

- Lightweight ReAct-style RAG agent (Gemini API) with local FAISS retrieval and INT8 ONNX CPU inference, designed for limited memory.

SKILLS

Languages: Python, SQL, TypeScript, Dart

ML & AI: PyTorch, scikit-learn, XGBoost, LightGBM, NumPy, Pandas, Hugging Face, ONNX Runtime, INT8 quantization, RAG (FAISS)

Backend & MLOps: FastAPI, Docker, MLflow, Pandera, Pydantic, GitHub Actions, Pytest, Celery, PostgreSQL, DuckDB

Frontend: React, Next.js, Flutter, Tailwind CSS

EDUCATION

Bachelor of Computer Application (BCA), Tribhuvan University, Nepal Completed Aug 2026

CERTIFICATIONS

FlyRank AI Machine Learning Internship | Skill Shikshya Data Science & ML Diploma | Anthropic Academy: Building with the Claude API, Claude Code in Action, MCP Advanced Topics | Kaggle: Intro & Intermediate ML