PUSHP KHARAT

AI Researcher & Systems Engineer

kharatpushp16@outlook.com — +91 98696 05981 — Mumbai, India

github.com/Pushp-Kharat1 — linkedin.com/in/pushp-kharat — pushp-kharat1.github.io

RESEARCH PROFILE

AI researcher specializing in neural-symbolic reasoning and high-performance machine learning systems. Independently reproduced Google DeepMind's AlphaProof mathematical reasoning approach and authored a novel gradient boosting algorithm outperforming XGBoost by 18% on extreme class imbalance. Published researcher with production ML systems deployed at Fortune 500 companies achieving 100,000+ queries/second.

Core expertise: ML algorithm design, Monte Carlo Tree Search, Rust systems programming, SIMD optimization, formal verification, and production ML deployment.

RESEARCH & PUBLICATIONS

LEMMA: Neural-Symbolic Mathematical Reasoning Engine **Mar 2025 – Present** *Independent*
Research — PKBoost AI Labs *Rust, 29k LOC*

- Reproduced Google DeepMind's 2024 AlphaProof approach: hybrid neural-symbolic system combining Monte Carlo Tree Search with a Transformer policy network for automated mathematical reasoning
- Built 450+ formally verified transformation rules spanning IMO-level mathematics: algebra, calculus, trigonometry, number theory, inequalities, combinatorics, and polynomial manipulation
- Achieved 95.2% accuracy on single-step problems and 100% on multi-step problems across a 31-test benchmark suite (algebraic identities, derivatives, equation solving, trigonometry, multi-variable)
- Designed an end-to-end training pipeline: auto-generated 17,000 synthetic mathematical problems and trained a Transformer for 50 epochs to guide symbolic rule selection in proof search
- Key innovation: Provides formally verified proof traces with complete step-by-step justification, eliminating hallucinated intermediate reasoning steps common in LLM-based systems
- Technical architecture: Custom AST parser, MCTS with UCB selection, Transformer policy network (Candle), integrated numerical and symbolic verifier
- **Impact:** Open-source research prototype demonstrating scalable hybrid reasoning for automated theorem proving
- **Code:** github.com/Pushp-Kharat1/LEMMA — **License:** Mozilla Public License 2.0

PKBoost: Shannon-Entropy Gradient Boosting for Extreme Imbalance **Jun 2025 – Present**
Published Research — PKBoost AI Labs *Rust + PyO3, 13K LOC*

- Proposed a novel gradient boosting algorithm fusing Shannon entropy with Newton–Raphson optimization, outperforming XGBoost by 17.9% PR-AUC and LightGBM by 10.4% on credit card fraud detection (0.2% minority class, 284K samples)
- Demonstrated extreme drift resilience: 1.8% degradation under covariate shift vs. XGBoost (31.8%) and LightGBM (42.5%), enabling reliable learning under evolving data distributions
- Implemented systems-level optimizations: zero-copy architecture (31.7 MB training overhead), cache-aware data structures (64-byte alignment), 8x loop unrolling for SIMD auto-vectorization, ¡5ms histogram construction
- Achieved 45-second training on 170K samples; supports binary classification, multi-class (One-vs-Rest with softmax), and regression
- Designed an auto-tuning system that profiles dataset characteristics and derives optimal hyperparameters, eliminating manual tuning while matching or exceeding tuned XGBoost/LightGBM
- **Published:** Zenodo DOI 10.5281/zenodo.17541137
- **Adoption:** 3,200+ PyPI downloads — 62 GitHub stars — Featured on Kaggle
- **Code:** github.com/Pushp-Kharat1/PkBoost — **PyPI:** `pip install pkboost`

PRODUCTION SYSTEMS

Enterprise RAG System for HR Knowledge Management Dec 2025 – Present *PKBoost AI Labs — Value Score Business Solutions* *Rust + React, 6K LOC*

- Role: Lead systems engineer responsible for end-to-end architecture, performance optimization, and production deployment
- Built an ultra-high-performance document Q&A system achieving 100,000+ queries/second, ¡5ms vector search latency, and 300ms end-to-end response time including LLM inference
- Delivered a 10–100x performance improvement over database-backed baselines (USearch in-memory HNSW: 5ms vs. PostgreSQL pgvector: 50ms for 10K-vector search)
- Implemented production-grade security: JWT authentication, Argon2 password hashing, token-bucket rate limiting, SQL injection protection, CORS enforcement, and graceful shutdown with signal handling
- Developed multi-format ingestion pipeline (PDF, Excel, Word, text) with optional Tesseract OCR and semantic chunking using all-MiniLM-L6-v2 embeddings (384-dim)
- Designed fully async architecture using Tokio to handle 1,000+ concurrent connections with connection pooling and single-binary deployment (50MB for 10K vectors)
- Real deployment: Deployed at a Fortune 500 company (Under NDA) supporting 1,000+ employees with ¡5ms semantic search across 10,000+ document chunks
- **Tech Stack:** Rust (Axum), Tokio, USearch, FastEmbed-rs, PostgreSQL, React + Vite, Groq API (Llama 3.3)
- **Code:** github.com/PKBoost-AI-Labs/RAG

PROFESSIONAL EXPERIENCE

Founder & Lead Research Engineer Dec 2025 - Present

PKBoost AI Labs — Mumbai, India

- Founded independent AI/ML research lab focused on high-performance tabular ML, neural-symbolic reasoning systems, and production ML infrastructure
- Research priorities: Concept drift adaptation, formal mathematical reasoning, SIMD-optimized inference, interpretable gradient boosting
- Built and maintained 3 major open-source projects (PKBoost, LEMMA, RAG) with 25K+ lines of production Rust code and active user communities

Technical Intern Jun 2025 - Present

Value Score Business Solutions LLP — Mumbai, India

- Architected and deployed agentic RAG workflows using n8n automation and open-source LLMs for document-based question answering
- Built production Rust RAG agent with USearch vector search—demoed to a Fortune 500 for employee HR assistance (1,000+ user capacity)
- Developed LLM-powered email personalization system with Groq/Grok validation and quality checks
- Evaluated Zoho Catalyst platform for ML model deployment and CRM integration

Network Engineering Trainee Feb 2025 - Apr 2025

Artech Communications — Mumbai, India

- Configured high-availability hospital LAN with redundancy and failover
- Administered Linux/Windows servers with security hardening and validation
- Performed penetration testing and network security audits

TECHNICAL EXPERTISE

Languages	Rust, Python, C++, JavaScript/TypeScript, SQL
ML Frameworks	Custom implementations (GBDT, MCTS), Candle, PyO3, FastEmbed
Systems	SIMD optimization, cache-aware algorithms, zero-copy design, async I/O (Tokio), memory safety, performance profiling
Algorithms	Monte Carlo Tree Search, Newton-Raphson optimization, Shannon entropy, gradient boosting, approximate nearest neighbors (HNSW)
Mathematics	Information theory, numerical optimization, statistical learning, linear algebra, calculus, formal verification
ML Domains	Concept drift detection, extreme class imbalance, tabular ML, neural-symbolic reasoning, retrieval-augmented generation
Infrastructure	Docker, PostgreSQL, Linux, Git, CI/CD, systemd, Nginx
Tools	USearch (vector DB), SQLx, Axum, n8n automation, Pandas, NumPy

EDUCATION & CREDENTIALS

Diploma in Computer Technology 2022 - 2025*K.V.M Institute of Technology — Mumbai, India CGPA: 8.1/10*

Independent Research & Advanced Study (Self-Directed):

- Reproduced cutting-edge AI research (Google DeepMind's AlphaProof, AlphaZero)
- Published novel ML algorithm with formal benchmarking and evaluation
- Mentored by Ash Vardanian (Founder, Unum Cloud; Creator of USearch, SimSIMD)

Other coursework:

- Advanced Machine Learning: Gradient boosting internals, MCTS, Transformers
- Systems Programming: Cache optimization, SIMD vectorization, Rust concurrency
- Mathematics: Information theory, numerical optimization, linear algebra
- Formal Methods: Symbolic verification, proof systems, type theory

RECOGNITION & IMPACT

Research Impact

- Published researcher: Zenodo DOI 10.5281/zenodo.17541137
- 3,200+ production users of PKBoost library across fraud detection, medical diagnosis, and anomaly detection applications
- Featured on Kaggle with working notebooks demonstrating 86.56% PR-AUC on credit card fraud detection (0.173% fraud rate)

Open Source Contributions

- 100+ GitHub stars across projects (PKBoost, LEMMA, RAG, Etc.)
- Active maintenance with continuous updates and community engagement
- Mozilla Public License 2.0 and Apache 2.0 licensing for research reuse

Mentorship & Collaboration

- Mentored by Ash Vardanian, industry expert in high-performance vector search and SIMD optimization (USearch used by Anthropic, Cohere, major AI companies)

Community Leadership

- Founded PKBoost AI Labs as platform for open ML research
- Regular technical blog posts and documentation for reproducible research

Athletic Achievements

- MMA District Gold Medalist 2022 (5-1 record)
- TRCAC Chess Gold Medalist 2024