Geddy Dukes

AI/ML ENGINEER · SF BAY AREA · AVAILABLE IMMEDIATELY

geddydukes@gmail.com|(707) 799-1271|linkedin.com/in/geddy-dukes|github.com/geddydukes

✦ FEATURED ON TLDR AI

SUMMARY

ML Engineer who builds AI for high-stakes environments. Specialize in systems where hallucinations aren't acceptable and every decision needs an audit trail — from training **67M**-parameter models from scratch to shipping production underwriting platforms for **$30M** portfolios. M.S. Analytics candidate at Georgia Tech with deep expertise in neuro-symbolic architectures, evaluation infrastructure, and deploying reliable AI in regulated industries.

TECHNICAL PROJECTS

TinyLLM – Language Model from Scratch   `FEATURED · TLDR AI`
Python · PyTorch · WandB · NumPy · github.com/geddydukes/tinyllm

- Implemented GPT-style transformer from scratch (**67M params**, 512 hidden dim, 12 layers, 8 heads) with RoPE, RMSNorm, and SwiGLU — demonstrating first-principles understanding of architecture beyond API-level usage.
- Achieved **93.94%** exact-match accuracy on natural language → CLI command generation.
- Designed continual learning system with replay buffers, weight-preference regularization, and evaluation gating to prevent catastrophic forgetting.
- Trained entirely on consumer hardware (Mac Mini 24GB, ~13 hrs, ~204M tokens) with custom MPS optimizations.

Provenance – Agentic Credit Underwriting Platform   `IN DEVELOPMENT`
Python · FastAPI · PostgreSQL · Pydantic · LLM APIs · TypeScript · github.com/geddydukes/provenance

- Architected event-sourced "glass-box" AI platform where agents emit immutable artifacts and cannot mutate system state — guaranteeing 100% auditability and zero silent numerical drift.
- Designed claim-map narrative system that prevents hallucination by binding all LLM outputs to registry-validated metrics and deterministic calculations.
- Implemented deterministic replay with schema-pinned metrics and policy-gating for human-in-the-loop regulatory workflows — designed to cut loan analysis from weeks to days for CDFIs.
- Built Supervisor/state machine with snapshot governance ensuring all external data is immutable, hashed, and reproducible.

AgentFT – Agent Evaluation Framework   `OPEN SOURCE · NPM · PYPI`
Python · AsyncIO · PyTorch · Jinja2 · github.com/geddydukes/agentft

- Built evaluation harness treating AI agents as testable software artifacts — async execution, composite judging, trace-level logging, regression detection across versions.
- Implemented pluggable agent adapter interface with deterministic verification, fail-fast modes, and provider-based rate limiting for large-scale runs.
- Published on npm and PyPI. Everyone is building agents — almost no one knows how to test them.

Research Discovery Agent – Knowledge Graph   `LIVE DEMO`
TypeScript · Node.js · PostgreSQL · Gemini 2.5 · Zod · github.com/geddydukes/research-agent

- Architected neuro-symbolic system separating LLM extraction from deterministic validation, enforcing confidence thresholds and provenance tracking before graph insertion.
- Extracted **119** entities, **187** relationships with per-edge confidence scores (75–100%) and direct source citations — full audit trail on every claim.

EXPERIENCE

ML Research Engineer (Independent)   Nov 2025 – Present
San Francisco

- Developing Provenance: agentic credit underwriting platform with event-sourced architecture and neuro-symbolic validation for regulated finance.
- Expanding AgentFT evaluation framework and contributing to open-source AI reliability tooling.

Credit Analyst & Catalytic Capital Program Manager   Dec 2023 – Nov 2025
Community Vision Capital & Consulting · San Francisco

- Architected and deployed production financial analysis platform (TypeScript/Next.js/PostgreSQL/Supabase) with RBAC, audit logs, and real-time calculation pipelines — used by underwriting teams across millions in loan volume.
- Managed **$30M** catalytic capital program; personally underwrote deals up to $3M while ensuring full compliance with federal and state lending regulations.
- Encoded complex financial domain rules into auditable backend logic — became the deterministic foundation for the AI underwriting platform.

Catalytic Capital Program Analyst   Jun 2022 – Dec 2023
Community Vision · San Francisco

- Improved program operational efficiency ~20% through data modeling and analysis optimization.
- Coordinated budgeting and performance across ~$100M in programs; managed AHP/AHEAD compliance.

Finance Project Manager   May 2020 – Oct 2021
Dorjil Company · San Francisco

- Built Python/SQL tools securing $400K+ in subsidies; led multi-year audit recovering **$2M** in overdue subsidies through data-driven forensic accounting.
- Automated data pipelines achieving ~25% efficiency improvements in reporting and compliance.

Full-Stack Software Engineer (Contract)   Jun – Sep 2020
Blindly · San Francisco

- Implemented identity verification workflows (Node.js/React/Redux); shipped production authentication features in cross-functional team.

SKILLS

ML & AI SYSTEMS
Neuro-symbolic AI | LLM orchestration | Agent frameworks | PyTorch | Transformers | RAG | Vector search | Multi-agent systems | Eval infrastructure | Multimodal SSL | Continual learning | Apple Silicon MPS

ENGINEERING
Python | TypeScript | Node.js | React/Next.js | FastAPI | PostgreSQL | Redis | Docker | AWS | Vercel | Git/CI/CD | Event sourcing

DOMAIN
Regulated environments | Financial analysis | Lending compliance | Risk assessment | Audit trails | Underwriting

EDUCATION

MS Analytics (Computational Data)
Georgia Institute of Technology
Expected 2027

BS Business Administration
Western Governors University
2024

Software Engineering Immersive
General Assembly
2020