MAWABA PASCAL DAO

ML Engineer

(321) 355-0510 pdao2015@my.fit.edu https://www.linkedin.com/in/mawaba-dao/ https://github.com/PascalPolygon
Melbourne

SUMMARY

Machine-Learning Engineer with four years of experience delivering real-time systems for defense and cybersecurity. I design and optimize low-latency inference and streaming pipelines, taking models from research to compliant production deployment.

LANGUAGES

English Native ●●●●●
French Native ●●●●●

SKILLS

Agile C/C++ Azure Git Java
JavaScript Mojo 🔥 Kafka
Keras/TensorFlow Machine Learning
Neural Networks NodeJS Python
PyTorch C# .NET TypeScript
TensorFlow Reinforcement Learning
Model-Based RL AWS React
MongoDB Data Analysis
Publication Writing

RELEVANT PUBLICATIONS

Boosting MCTS with Free Energy Minimization

Accepted at Neural Computation (MIT Press), Publication in November, 2025

Mawaba Pascal Dao, Adrian M. Peter

06/2025
https://arxiv.org/pdf/2501.13083

Hybrid MCTS-CEM planner with free-energy bonuses beats standalone CEM and MCTS on continuous-control tasks.

LEVIOSA: Natural Language-Based Uncrewed Aerial Vehicle Trajectory Generation

Electronics

Godwyll Aikins, Mawaba Pascal Dao, Koboyo Josias Moukpe, Thomas C. Eskridge, Kim-Doang Nguyen

11/2024
https://doi.org/10.3390/electronics13224508

LEVIOSA uses multimodal LLMs to turn text or speech commands into UAV-swarm flight paths for search-and-rescue, agriculture, and infrastructure inspection.

EXPERIENCE

ML Engineering Intern

Onlykit — 03/2025 - 08/2025 — San Fransisco, USA (Remote)

- **AI social-engineering training platform:** Delivered a real-time system that simulates vishing and other emerging cyber-attacks, enabling organizations to train employees and monitor potential insider threats.
- **Self-hosted, low-latency ML stack:** Built a fully self-hosted speech-to-speech pipeline that achieves less than 500 ms end-to-end on a single L40 GPU; no external APIs, enabling offline deployments.
- **Model design & training:** Curated and fine-tuned Kit-Llama-3_2-1B on 14 million JSONL rows (about 70 million prompt-answer pairs) drawn from Army, Air Force, Space Force, and law enforcement corpora; delivered BLEU 58.2— approximately 4.8 times the baseline Llama-3.2-1B— while producing 48% shorter, tactically concise replies.
- **Inference optimization of Sesame's CSM:** Integrated vLLM for production serving and streamed each newly generated RVQ token directly into Mimi codec (chunk-wise audio generation) to reduce per-turn latency by 35%.
- **Benchmarking & evaluation:** Compared Sesame, Orpheus, and F5 voice-cloning models on synthesis speed and MOS quality; selected the best-of-breed hybrid for live calls.
- **Real-time vishing simulation platform:** Integrated Twilio bidirectional audio with the transformer pipeline and architected a hierarchical decision engine: strategic campaign planner plus in-session chatbot.
- **Regulatory readiness:** Implemented TCPA-compliant SMS consent workflow.

Lab Manager/Graduate Research Assistant

Center for Advanced Data Analytics and Systems — 05/2021 - Present — Melbourne, FL

- **Created a real-time acoustic-event inference system:** Optimized lightweight edge models deployable on Raspberry Pi boards and Samsung Android phones for on-device detection.
- **Own day-to-day delivery on a $2.4 million AFRL contract:** Coordinate a five-member research team, run weekly stand-ups, escalate blockers to university resources, and submit comprehensive monthly progress reports.
- **Leading R&D on large-scale audio labeling:** developed an embedding-based label-propagation pipeline to auto-label 30 million audio clips; triplet-loss embedder attains 99% validation accuracy on ESC-50.
- **Built an interactive robustness-analysis platform (**React/TypeScript UI + multi-threaded Flask/PyTorch backend + Keycloak authentication on AWS EC2) that benchmarks acoustic-classification models on accuracy, precision, recall, F1, ROC, and PR curves.
- **Designed a 91 Mbps multi-group Android MANET** and an MQTT-based gunshot-sound collection network spanning 26 devices and six microphone types for D3SOE field tests..
- **Mentoring and enabling two undergraduates and two PhD candidates** on experimental design, software integration, and publication writing.

Software Engineer

Computer Task Group (CTG) — 03/2022 - 01/2025 — Melbourne, FL

- Part-Time
- Designed and implemented **event streaming software over Kafka message bus.**
- **Wrote, tested and containerized libraries** and services in C#, Python, Java & C++.
- **Implemented and Executed Azure pipelines,** deployed software with Ansible.
- Worked in a scrum team following agile methodology.

EDUCATION

Doctor of Philosophy (Ph.D.) Candidate in Computer Engineering

Florida Institute of Technology — 01/2021 - Present — Melbourne, FL

- Expected graduation May '25
- Doctor of Philosophy (Ph.D.) in Computer Engineering Candidate.
- Florida Institute of Technology, Melbourne, FL GPA 3.66.
- Researching effective tree representations for Monte Carlo Tree Search in model-based reinforcement learning, focusing on efficient planning in high-dimensional continuous state-action spaces and Reinforcement Fine Tuning of Large Language Models.

Bachelor of Science (B.S.) in Mechanical Engineering, Cum Laude

Florida Institute of Technology — 08/2015 - 05/2019 — Melbourne, FL

- Graduated May'19

INTERESTS

💎 🤖 Reinforcement Learning
💎 👍🏾 Jujitsu