Thong Pham

ThongPhamBusiness@gmail.com | Linkedin.com | (714) 487 3430 | Thornton, CO | U.S Citizen

EDUCATION

University of Colorado, Boulder   **Boulder, CO**
Master of Science: Computer Science (**Specialization: Machine Learning, GPA 4.0 / 4.0**)   *June 2024 - June 2026*
University of California, Irvine   **Irvine, CA**
Bachelor of Arts: Business Economics   *September 2019 - December 2022*

WORK EXPERIENCE

Kashmir World Foundation   **Richmond, VA**
Machine Learning Engineer Intern   *September 2024 - Present*

- Utilized **XGBoost**, **CatBoost**, and **SVM** on **large-scale bioacoustic data** (over 200K audio clips) to achieve **91.6% classification accuracy**, demonstrating end-to-end data ingestion, hyperparameter tuning, and deployment.
- Implemented **dimensionality reduction** techniques (PCA, LDA) and optimized feature engineering strategies, reducing computational overhead by **20%** without compromising accuracy—enabling near real-time species identification in conservation projects.
- Performed **exploratory data analysis (EDA)** and **feature selection** (random forests, correlation analysis) on complex signal-based datasets to enhance model interpretability for domain experts studying endangered species.
- Created interactive dashboards in **Power BI** and **Tableau**, translating intricate model outcomes into actionable insights for **non-technical stakeholders** within wildlife conservation groups.

Scale AI   **Remote, Online**
Prompt Engineer   *November 2024 - February 2025*

- Developed customized **evaluation metrics** (BLEU, perplexity, word embeddings) to benchmark large language models across multiple text corpora, improving fine-tuning strategies for high-stakes language applications.
- Built **TF-IDF, NER**, and **embedding-based** workflows to enrich textual data, increasing downstream accuracy for tasks such as sentiment analysis and entity recognition by **15%**.
- Analyzed usage patterns and token consumption in production LLMs, reducing overall **token costs by 16%** through improved prompt design and model inference optimization.
- Established **A/B testing** and drift detection frameworks to monitor real-time performance, ensuring consistent **response accuracy** for applications in chatbot, search, and content generation.

True Path Financial   **Irvine, CA**
Associate Data Analyst   *February 2023 to April 2024*

- **Applied Statistical Modeling:** Built **machine learning models** (K-Nearest Neighbors, Isolation Forest) on transactional datasets exceeding **500K records** for financial anomaly detection, increasing **fraud detection accuracy by 11.3%**.
- Transitioned key financial data from SQL to **MongoDB**, improving **query performance by 30%** and enabling **faster analytics** on semi-structured customer data.
- Leveraged **Power BI**, **Tableau**, and **Seaborn** to create interactive dashboards, guiding executive decisions on investment strategies and operational improvements.
- Performed **A/B testing**, chi-square tests, and **statistical modeling** on new financial products, reducing **operational risk** by identifying unprofitable segments early.

PROJECTS

AI Pong with NEAT | *Python, NEAT (NeuroEvolution of Augmenting Topologies)*   *December 2024*

- Developed a self-learning AI agent that uses NEAT to evolve neural network topologies and master the game of Pong from scratch. Over the course of **300+ generations**, the AI achieved near-human performance, demonstrating **robust improvements** in reflex-based decision-making. By optimizing reward structures and population parameters, **total training time was reduced by 20%**.

Retro Kong AI (PPO) | *Python, Proximal Policy Optimization (PPO)*   *December 2024*

- Developed an AI agent to play the 1980s arcade game "**Kong**," with barrels rolling toward the player's character. **Using Proximal Policy Optimization** in a custom environment mirroring original game physics, the agent refined its platform navigation and dodging strategies. **By tuning learning rate and clip range**, it exhibited near-human reflexes over multiple training epochs, illustrating PPO's effectiveness in mastering classic arcade challenges.

TECHNICAL SKILLS

Programming Languages: Python (TensorFlow, Scikit-learn, Keras, PyTorch, Seaborn, OpenCV), R, Java, C/C++, Swift, SQL.
Tools & Frameworks: Git (CI/CD), Power BI, Tableau, Excel, PowerPoint, Word, Vantage API, Microsoft Outlook.
Data Engineering & Big Data: MongoDB, AWS (S3, EC2), Google Cloud (GCP).