John Patrick Capocyan

johncapocyan@gmail.com | johncapocyan.com | linkedin.com/in/jpcai | github.com/J7P16

EDUCATION

University of Texas at Austin, B.S. in Computer Science, GPA: 3.9/4.0 May 2028

- **Relevant Courses:** Data Structures, Computer Organization & Architecture, Intro to Computer Science Research, Discrete Math, Linear Algebra, Probability & Statistics
- **Member:** Texas Luminescence, Machine Learning & Data Science Club, DiRP, UT Chess Club (A Team)
- **Awards:** Undergraduate Research Fellowship ($1000), Faculty Research Grants Scholarship ($500)

EXPERIENCES

Cenik Lab at UT Austin, AI Research Assistant - Austin, TX Jan 2026 – Present

- Trained convolutional and transformer-based architectures to predict RNA efficiency from ribosome profiling.
- Scaled DNA context window from 13k to 500k base pairs (38x increase), improving biological signal capture.
- Utilized Linux and SSH to access TACC's A100 GPUs, accelerating large-scale training for deep learning models.

Baker Group at UT Austin, Research Assistant - Austin, TX Oct 2025 – Dec 2025

- Processsed 10k+ quantum circuit samples using Pandas and NumPy, identifying 50+ critical circuit encodings.
- Generated 100+ Matplotlib visualizations to analyze relationships between circuit fidelity and fault tolerance.

iAnswer, Software Engineering Intern - Katy, TX May 2024 – Aug 2024

- Implemented RAG to fine-tune a LLM on 1000+ anonymized clinical PDFs, improving text querying by 20%.
- Built a PyTorch-based embedding pipeline producing dense vectors (768-dim) optimized for semantic retrieval.
- Connected React components with Node.js backend through REST APIs, improving response efficiency by 50%.

PROJECTS

Well Oil Production Model - Team Lead | Pandas, Jupyter Notebook, Sklearn Feb 2026

- Developed a gradient-boosted decision tree predicting 3-year cumulative oil output with a 0.106 error metric.
- Engineered physics-informed features using engineering equations, reducing 20+ raw variables into 10 inputs.
- Achieved 8th place out of 40+ teams (200+ competitors) at the 6th Annual Energy AI Hackathon (UT Austin).

Novel Vision Transformer Framework | Python, Tensorflow, OpenCV, Pandas Mar 2024 - Aug 2024

- Proposed a new ViT embedding pipeline, increasing classification performance by 27% in data-limited training.
- Reached 97.50% accuracy using only 1,200 labeled H&E images, outperforming CNN and baseline ViT models.
- Published sole-author paper in 2024 MIT Undergraduate Research Technology Conference (200+ participants).

CNN Lung Cancer Detection | Python, Tensorflow, OpenCV, Pandas, Sklearn Aug 2023 - Jan 2024

- Trained a DenseNet201 architecture on 30,000+ H&E tissue images, achieving 94.07% diagnostic accuracy.
- Attained 99.81% accuracy on LC25000 (5 classes) without stain normalization, showing strong generalization.
- Published sole-author paper in the 2024 Integrated STEM Education Conference (ISEC) (250+ participants).

LEADERSHIP AND ACTIVITIES

Texas Luminescence - Officer Board Member Jan 2025 - Present

- Architected a new recruitment process for 100+ applicants, assessing their technical and behavioral background.

Texas Luminescence - Project Team Lead Aug 2025 - Dec 2025

- Led development of HealthLens, a computer vision + RAG app, automating dermatology screening workflows.
- Coordinated an 11 member team across frontend, backend, and ML engineering, maintaining weekly milestones.

TECHNICAL SKILLS

Programming Languages: Python, Java, JavaScript, React, SQL
Tools: TensorFlow, PyTorch, Scikit-learn, OpenCV, RAG, Hugging Face, Pandas, NumPy, Matplotlib, Seaborn
Technologies: Deep Learning, Computer Vision, NLP, Jupyter Notebook, Git, VS Code, API Integration
Interests: Golf, Tennis, Chess, Coffee Brewing, Photography, Hiking, Gym, Reading, Journaling