

# Paul J. Yoon

+1 (713) 320-1032 | [pauljy@stanford.edu](mailto:pauljy@stanford.edu) | [linkedin.com/in/pauljinyoon](https://www.linkedin.com/in/pauljinyoon)

## EDUCATION

### STANFORD UNIVERSITY

*Bachelor of Science, Mathematics. Minor in Music*

Expected Graduation: Jun 2027

GPA: 3.8/4.0

**Relevant Coursework:** Linear Algebra, Differential, and Integral Calculus of Several Variables, Computer Organization and Systems, Real Analysis, Probability Theory for Computer Scientists, Machine Learning, Math for Machine Learning

## PROFESSIONAL EXPERIENCE

### STANFORD DEPARTMENT OF MATHEMATICS

*Student Researcher*

Stanford, CA

Jun 2025– Aug 2025

- Blueprinted the full proof of Wigner’s Semicircle law into a 20+ page LaTeX blueprint bridging probability theory, combinatorics, and measure theory
- Implemented and machine-verified dozens of core lemmas and constructions in Lean, preparing contributions to Lean’s mathlib library
- Collaborated with a three person research team using Git version control workflows to maintain a shared codebase

### SUNDIAL

*Data Science Intern*

Palo Alto, CA

Jul 2024 – Sep 2024

- Sundial is a Series A startup building an automated data science and product analytics tool founded by two early Meta executives, one of whom went to Sequoia Capital before founding Sundial
- Developed classifier identifying fraudulent users using behavior-based thresholds, reducing false positives compared to existing process by 50% and boosting overall detection accuracy by 75%
- Created, trained, and tuned a time series seasonality model, outperforming existing model by 120% as measured by mean absolute percentage error (MAPE)

## PROJECTS

### Timestamping Video Game Eliminations with Computer Vision

*Python, LaTeX*

Apr 2025 – Jun 2025

- Curated a custom detection dataset (165 events, ~ 1,650 frames) with color-jitter augmentation strategies
- Fine-tuned YOLOv8-nano (2.5M params) on <200 images, achieving 0.61 F1 and 0.49s mean temporal error, doubling precision compared to a baseline template matcher
- Engineered a lightweight inference pipeline (OpenCV + ffmpeg) that processes a 14-minute VOD in 1.5 minutes on CPU, generating highlight clips with sub-frame accuracy

### An Exploratory Analysis of Feature Representation in Music Source Separation

*Python, LaTeX*

Jan 2025 – Mar 2025

- Implemented featurization approaches (STFT, Mel-spectrogram) within a Band-Split RNN to isolate vocal tracks
- Integrated HiFi-GAN for Mel-spectrogram inversion, improving audio reconstruction quality by reducing artifacts
- Optimized training via PyTorch AMP, hyperparameter tuning, and data augmentation on the MUSDB18 dataset

### Explicit/Implicit Heap Allocator

*Unix, C*

May 2024 – Jun 2024

- Implemented the “malloc”, “realloc”, and “free” functions optimizing for request throughput and memory utilization
- Incorporated an explicit list of nodes to assign optimal locations for new memory requests and lower memory fragmentation
- Achieved 91% memory utilization via testing on heap activity memory requests from Emacs, Cmake, and Firefox

## TECHNICAL SKILLS

**Languages:** Python, LaTeX, TypeScript, JavaScript, SQL, C++, C, HTML/CSS

**Frameworks/Libraries:** Pandas, NumPy, Matplotlib, scikit-learn, React, React Native, Next.js

**Developer Tools:** Git, Unix, Vim, VS Code, Apache Spark, Snowflake, Jupyter Notebook, Qt Creator