

Swift-Transformer Project

A native implementation of Numpy Transformer Project

Arya Mirsepasi, Nov 2024

Features

- Rewritten from Python Numpy-Transformer to Swift
- Optimized for Apple Silicon
- Leverages MLX-Swift for array operations
- Utilizes native GPU for operations

36.75x Faster in Training

26.12x faster in Evaluation

Challenges

Challenges

- **Finding a NumPy Alternative in Swift:**
 - Tested libraries like Matft; lacked required functions
- **Custom Implementation Issues:**
 - Complexities with varying matrix dimensions
- **Memory Management Problems:**
 - Memory usage spiked to 50 GB, causing crashes after 11 epochs

Solutions

- **Adopted MLX-Swift:**
 - 99% of required functionality
 - Excellent documentation and support
- **File Format Transition:**
 - Switched from .pkl to .safetensors
 - Aligns with Swift and MLX-Swift standards

Memory Management Fixes

- **Autorelease Pools:**
 - Wrapped compute-intensive operations to manage memory
- **Cache Clearing:**
 - Used MLX to clear cache after each training epoch
- **Simplified Implementations:**
 - Implemented essential loss, activation, and optimizer classes

Memory usage stabilized under 6 GB after 20
epochs

Demo



Appendix

- **Github Repository:** <https://github.com/Aryamirsepasi/swift-transformer>
- **Based on:** Numpy-Transformer by AmritanshuV
 - Github Repository: <https://github.com/AmritanshuV/Numpy-Transformer>
- **Additional Resources:**
 - NOTES.md for notes, dependencies, acknowledgements
 - Performance data folder for detailed metrics
 - MLX Swift Github: <https://github.com/ml-explore/mlx-swift/tree/main>

The background of the slide features a dark, star-filled night sky. A vibrant green aurora borealis arches across the upper portion of the frame, its light glowing against the dark backdrop. In the lower half, the silhouettes of rugged mountain peaks are visible, their dark forms contrasting with the bright celestial lights.

Questions?

Thank you for your attention!