Xianglong Xu

Personal Website: sean652039.github.io

412-452-5278 xix110@pitt.com linkedin.com/in/xianglongxu github.com/Sean652039

Education

University of Pittsburgh **August 2023 – May 2025**

Master of Science in Information Science *GPA: 3.95/4.0*

North China Electric Power University **September 2019 – June 2023**

Bachelor of Science in Computer Science - Specialization in AI *GPA: 3.59/4.0*

Experience

Vosyn Inc. **May 2024 – August 2024**

AI Software Developer Intern *Chicago, Illinois(Remote)*

- Led end-to-end development of a speaker diarization system with significantly improved performance:
  - Achieved 10x faster execution compared to pyannote clustering by implementing a novel Chinese Restaurant Process-based approach.
  - Integrated an end-to-end Speaker Change Detection model based on research by Hervé Bredin and Antoine Laurent.
  - Developed custom evaluation scripts using AMI Meeting Corpus to calculate Diarization Error Rate (DER).
- Enhanced the diarization system robustness and efficiency through multiple optimizations:
  - Implemented WebRTC's VAD with ring buffer for improved speech detection accuracy.
  - Optimized memory usage by processing audio in-memory using io.BytesIO, eliminating disk I/O overhead.
  - Engineered robust audio processing pipeline to handle non-standard audio files.
- Conducted comprehensive research and documentation for speech technology advancement:
  - Gained expertise in xTTS model architecture and principles.
  - Collected and analyzed research papers to keep team updated on latest methodologies.
  - Created and maintained detailed technical documentation covering API usage, system design, and performance metrics.

Yunnan Gold Mining Group Co. LTD. **February 2022, August 2022**

Information Engineer Intern *Kunming, Yunnan(Onsite)*

- Supported digital mine management initiatives by assisting with data visualization for safety monitoring, maintaining digital maintenance records, and participating in digital manufacturing reform meetings.
- Assisted HR operations by helping maintain employee database, conducting basic HR data analysis, and supporting departmental communication and training programs.

Projects

Music Generation with Deep Learning **September 2024 – December 2024**

- Engineered a dual-model music generation system integrating Stable Diffusion for roll plot generation from MIDI files and RBM for refining and smoothing musical outputs. The combined approach significantly improved performance, achieving a CLAP score of 0.9863, the highest among tested models, and reducing FAD to 13.3207, compared to the RNN baseline's CLAP score of 0.9723 and FAD of 6.6732.
- Developed an advanced preprocessing pipeline for roll plots, incorporating temporal truncation and pitch range adjustments to facilitate efficient processing of musical data in a 512x768 pixel format, optimizing data compatibility with model architectures.
- Created custom visual encoding strategies leveraging RGB and grayscale mappings to optimize information representation, ensuring seamless integration with Stable Diffusion and enhancing interpretability.

AI Snake **March 2024 – May 2024**

- Developed and implemented a deep reinforcement learning AI agent using MLP & CNN architectures for the "Snake" game, resulting in a 5% improvement in average game performance.
- Engineered a sophisticated Markov chain model to analyze game states, incorporating multiple parameters including spatial positioning, snake metrics, directional vectors, and body configuration.
- Innovated a novel scoring mechanism inspired by Go game principles, developing a recursive territory calculation algorithm to optimize snake movement patterns.
- Designed and implemented a dynamic penalty system based on space utilization analysis, successfully promoting more efficient snake navigation strategies.

LSTM-based text classification **December 2023 – January 2024**

- Engineered a high-performance text classification system with jieba segmentation, implementing a custom neural network by extending nn.Module and leveraging Tencent AI Lab's pre-trained embeddings, achieving 95%+ accuracy.
- Developed a robust vocabulary generation pipeline using statistical word frequency analysis and designed efficient preprocessing workflows for efficient sequence conversion and text normalization.