SYED ALI

basisvectors | basisvectors | syed.ali165@gmail.com | +91 91089 93774

WORK EXPERIENCE

Machine Learning Researcher, Beatoven.ai — Jan 2024 – Present

Keywords: PyTorch, Diffusion Transformers, LoRA, Distributed Training, MLOps, Data Orchestration — *Bangalore, India*

Specializing in generative audio models, large-scale data infrastructure, and foundation model training and optimization

- Designed Beatoven.ai's *Composer* architecture — an embedding-guided loop retrieval and arrangement model for controllable music generation from text and audio embeddings.
- Engineered low-level optimizations (GPU/CPU parallelization, multithreading, resumable batch processing) achieving a 7.5× *speedup* in training data preprocessing, reducing cleanup turnaround from multiple days to a few hours.
- Built large-scale data curation, preprocessing, and augmentation pipelines for *Maestro*, Beatoven's 2.5B parameter diffusion foundation model trained on several multi-million licensed music track datasets.
- Built fine-tuning pipelines for SFX, genre, and artist adaptation; optimized hyperparameters for *2× faster convergence* ($48 \rightarrow 24$ GPU hours) and improved audio quality.
- Built modular and reusable distributed training and inference pipelines for in-house embedding extraction models, improving maintainability across research teams and ensuring scalable, collaborative development.

AI Developer (Full-Stack & Data Systems), Beatoven.ai — July 2021 – Dec 2023

Keywords: Flask, React, MongoDB, Python, REST APIs, Microservices, NLP — *Bangalore, India*

Specializing in full-stack development, large-scale data ingestion, and metadata automation for music AI systems

- Built of Beatoven.ai's *artist portal and internal data ingestion system* — a full-stack Flask + React platform enabling large-scale artist submissions, metadata management, and human annotation workflows.
- Designed and deployed a lightweight *NLP model* for metadata extraction from unstructured text, reducing manual tagging overhead by $> 70\%$ and enabling high-quality dataset curation for early music generation models.
- Built scalable, asynchronous ingestion and cleaning pipelines handling *thousands of artist submissions*, integrated directly with Beatoven's music composition and embedding systems.
- Designed UX dashboards and automation workflows that streamlined team operations, reducing repetitive manual processing and forming the backbone of Beatoven's early data ingestion infrastructure.

Founder & Lead Developer, Tarab Instruments — June 2024 – Present

Keywords: Embedded Systems, Audio DSP, PCB Design, Analog Circuits — *Bangalore, India*

Designing and building custom MIDI controllers, Eurorack modules, and guitar pedals

- Founded Tarab Instruments, a boutique hardware studio specializing in custom audio hardware for musicians and producers.
- Teaching workshops on music electronics, synthesis, and DIY audio hardware

RELEVANT PROJECTS

CrateDig — AI Sample Pack and Music Library Manager — gh/CrateDig

Thesis Project, Dec 2023 - May 2024

Built an intelligent audio search app for music producers to organize and explore large sample libraries using natural language and reference audio queries. CrateDig analyzes audio, extracts embeddings, and enables fast similarity search via a local, offline desktop UI for Win/MacOS—streamlining sample discovery and retrieval from hours to seconds.

EDUCATION

Ramaiah University Of Applied Sciences — 2020 - 2024

Bachelor of Technology in Computer Science and Engineering — *Bangalore, India*

Relevant Coursework: Machine Learning, System Design, DSA, Automata Theory, Digital Signal Processing, Operating Systems, Database Management Systems