

Lecture Note: Large Language Models – Bài Giảng: Mô Hình Ngôn Ngữ Lớn

Nguyễn Quân Bá Hồng*

Ngày 15 tháng 5 năm 2025

Tóm tắt nội dung

This text is a part of the series *Some Topics in Advanced STEM & Beyond*:

URL: https://nqbh.github.io/advanced_STEM/.

Latest version:

- *Lecture Note: Large Language Models – Bài Giảng: Mô Hình Ngôn Ngữ Lớn.*

PDF: URL: [.pdf](#).

TeX: URL: [.tex](#).

Mục lục

1 Basic LLMs	1
2 Miscellaneous	2
Tài liệu	2

1 Basic LLMs

Resources – Tài nguyên.

1. *Attention Is All You Need*.
2. [Ras24]. SEBASTIAN RASCHKA. *Build A Large Language Model (From Scratch)*.

Question 1 (LLMs vs. Music Generation). *Draw similarities between LLMs & Music Generation in Computer Music.*

– *Vẽ ra điểm tương đồng giữa LLM và ngành Tạo nhạc trong Âm nhạc máy tính.*

Grok3's DeepSearch. Large Language Models (LLMs) & music generation in computer music share several conceptual and technical similarities, as both involve generative processes that create structured outputs from complex data. Below, I outline key parallels:

1. Sequential Data Processing.

- *LLMs*: Operate on sequences of tokens (words, subwords, or characters) to predict the next token based on context, generating coherent text.
- *Music Generation*: Processes sequences of musical elements (notes, chords, or time steps) to predict subsequent elements, creating melodies, harmonies, or rhythms. Models like MIDI-based systems or audio spectrogram processors treat music as a time-ordered sequence, akin to text.

2. Pattern Learning and Generalization:

- *LLMs*. Learn linguistic patterns, grammar, and semantic relationships from vast text corpora, enabling them to generate contextually relevant sentences.
- *Music Generation*: Learn musical patterns, such as chord progressions, scales, or stylistic motifs, from datasets of compositions or audio. Models like MuseNet or Magenta capture genre-specific structures to generate music that aligns with learned styles.

3. Transformer Architecture.

- *LLMs*: Rely heavily on transformer models (e.g., GPT architectures) to model long-range dependencies in text, using attention mechanisms to weigh the importance of prior tokens.

*A scientist- & creative artist wannabe, a mathematics & computer science lecturer of Department of Artificial Intelligence & Data Science (AIDS), School of Technology (SOT), UMT Trường Đại học Quản lý & Công nghệ TP.HCM, Hồ Chí Minh City, Việt Nam.
E-mail: nguyenquanbahong@gmail.com & hong.nguyenquanba@umt.edu.vn. Website: <https://nqbh.github.io/>. GitHub: <https://github.com/NQBH>.

- *Music Generation*: Many modern music generation systems, such as Jukebox or Music Transformer, also use transformers to model long-range dependencies in musical sequences, ensuring coherence over extended passages.

4. *Training Objective (Next-Token Prediction)*:

- *LLMs*: Trained to minimize the error in predicting the next token in a sequence, optimizing for fluency and coherence.
- *Music Generation*: Similarly trained to predict the next note, chord, or audio frame, ensuring the output adheres to musical logic. For example, symbolic music models predict note sequences, while audio-based models like WaveNet predict waveform samples.

5. *Creative Output with Variability*:

- *LLMs*: Generate diverse text outputs by sampling from probability distributions (e.g., top-k or nucleus sampling), allowing for creative variations in responses.
- *Music Generation*: Use similar sampling techniques to introduce variability in generated music, enabling models to produce novel melodies or improvisations while adhering to learned structures.

6. *Contextual Awareness*:

- *LLMs*: Use context windows to maintain coherence, drawing on prior tokens to shape the narrative or argument.
- *Music Generation*: Maintain musical context, ensuring that generated notes or phrases align with the preceding musical structure, such as staying in key or following a rhythmic pattern.

7. *Fine-Tuning for Style*:

- *LLMs*: Can be fine-tuned on specific datasets to adopt particular writing styles, tones, or domains (e.g., legal or poetic text).
- *Music Generation*: Fine-tuned to emulate specific genres, composers, or instruments (e.g., Bach-style chorales or jazz improvisations), tailoring outputs to desired musical aesthetics.

8. *Evaluation Challenges*:

- *LLMs*: Difficult to evaluate due to subjective quality of text; metrics like BLEU are limited, and human judgment is often required.
- *Music Generation*: Similarly hard to assess, as musical quality is subjective. Metrics like pitch entropy or harmonic consistency are used, but human listening tests remain critical.

9. *Ethical and Bias Concerns*:

- *LLMs*: Can perpetuate biases present in training data, generating biased or harmful text if not carefully managed.
- *Music Generation*: May replicate biases in musical datasets, such as over-representing certain genres or cultural styles, potentially marginalizing others.

10. *Interdisciplinary Applications*:

- *LLMs*: Applied beyond text generation to tasks like code generation, translation, or dialogue systems.
- *Music Generation*: Used in interactive composition tools, film scoring, or therapeutic applications, extending beyond standalone music creation.

In summary, both LLMs and music generation leverage probabilistic, sequence-based modeling to create structured, creative outputs. Their reliance on similar architectures (e.g., transformers) and training paradigms highlights a deep connection between natural language and music as forms of human expression processed by AI. □

2 Miscellaneous

Tài liệu

[Ras24] Sebastian Raschka. *Build A Large Language Model (From Scratch)*. 1st edition. Manning Publishing, 2024, p. 343.