Vishwakarma Institute of Information Technology
Department of CSE-AI

Melody Generation using LSTM and GRU-based RNNs

Name	Roll No.
Arnav Gilankar	22311644
Parth Kairamkonda	22311585
Janhvi Saste	22311825
Rudra sawant	22420151

Guided by - Dr. Anuradha Yenkikar Ma’am

Melody Generation using LSTM and GRU-based RNNs

Welcome to this presentation on AI-Powered Music Composition. We will explore how deep learning techniques are revolutionizing creative expression through melody generation. Our focus is on using advanced Recurrent Neural Network architectures, specifically LSTM and GRU models, to create compelling, algorithmic music that captures the structure and emotion of human compositions. This journey combines technology and art to push the boundaries of automated music creation.

Presentation Overview

Project Goals and Objectives

A clear outline of what we aim to achieve with melody generation models.

Background on RNNs, LSTMs, and GRUs

Discussion of the neural architectures used for learning musical sequences.

Dataset and Preprocessing Steps

Details on the musical data and how it is prepared for training.

Model Architecture and Training Details

Insights into the design and training of our deep learning models.

Results and Evaluation Metrics

Evaluation strategies and examples of generated music quality.

Challenges and Future Directions

Current limitations and promising avenues for research.

Conclusion

Summarizing key takeaways and project impact.

Problem Statement: Algorithmic Music Generation

Music has a complex temporal structure that poses challenges for algorithmic generation. Capturing the dependencies between notes over time is critical to creating melodies that sound coherent and expressive. Traditional approaches like Markov models fall short in understanding long-term context, resulting in music that lacks depth and emotional nuance.

This project addresses these challenges by leveraging neural networks capable of modeling long-range temporal patterns, aiming to produce more sophisticated and emotionally resonant music beyond rule-based methods.

- Challenge: Capturing temporal dependencies in music
- Limitations of traditional Markov models
- Need for context-aware music creation
- Improving emotional depth in generated music

Objective: Deep Learning-Based Melody Generation

1. **Develop Melody Generation System**

   Create a system that can autonomously compose melodies using deep learning methods.

2. **Utilize LSTM and GRU Networks**

   Harness strengths of both Long Short-Term Memory and Gated Recurrent Unit models for sequence learning.

3. **Train on Diverse MIDI Dataset**

   Use a rich collection of classical and popular melodies to build a versatile model.

4. **Generate Coherent, Creative Output**

   Ensure melodies produced are musically coherent and exhibit creativity.

Model Architecture: LSTM and GRU Networks

Recurrent Neural Networks (RNNs)

RNNs are designed to handle sequential data, making them suitable for music. However, standard RNNs struggle with long-term dependencies due to vanishing gradients.

Long Short-Term Memory (LSTM)

- Overcomes vanishing gradient problem
- Employs memory cells and gates for controlled information flow
- Capable of learning long-range dependencies

Gated Recurrent Units (GRUs)

- Simplified variant of LSTM with fewer parameters
- Faster training times
- Often matches LSTM performance in sequence tasks

Dataset and Preprocessing

Dataset Composition

Utilized MIDI files of classical compositions from Bach and Mozart, alongside popular melodies from The Beatles and Queen. This combination ensures stylistic diversity and broad learning scope.

Preprocessing Steps

1. Convert MIDI files into numerical representations
1. Normalize note values and create sequential input data

1. Split data into training and validation sets (80/20)
1. Apply one-hot encoding for effective classification

Results and Evaluation

Generated Sample Melodies

Both LSTM and GRU models produced sample melodies exhibiting coherence and musical structure, demonstrating the models' ability to capture temporal patterns.

Evaluation Metrics

- Perplexity - Assesses model confidence and prediction accuracy
- Musicality Scores - Human evaluators rate melody quality and emotional impact
- Objective Metrics - Note range, interval distribution, and rhythmic complexity

A/B Testing

Comparative tests reveal nuanced differences between LSTM and GRU outputs, with each having strengths in different musical aspects.

Conclusion and Future Directions

Project Summary

The project successfully developed an AI-based melody generation system employing LSTM and GRU architectures, capable of producing creative, coherent musical sequences.

Limitations and Challenges

Current models face challenges with emotional expressiveness and real-time generation constraints. Data diversity and stylistic depth are additional factors to improve.

Future Research Avenues

- Incorporate attention mechanisms to better capture long-range dependencies
- Explore generation in diverse musical styles beyond classical and pop
- Enable real-time, interactive music creation applications
- Add complexities like chords, harmonies, and multi-instrument arrangements