Antoine Gaton

November 3, 2024

CS379

Unit 4: Alice in Wonderland LSTM

* Description: This code implements a Long Short-Term Memory (LSTM) neural network to generate haiku poetry based on patterns learned from "Alice in Wonderland".
  + The model:
    - Processes text data from Alice in Wonderland for creative input
    - Uses character-level tokenization and syllable counting for haiku structure (5-7-5)
      * I wanted to get creative with the code but need further tweaking.
    - Implements a deep LSTM architecture with embedding layer for text understanding
    - Trains on sequences of 50 characters to predict the next character
    - Generates haiku poetry using temperature-based sampling (0.7) for creativity
      * I would like to see output with temperature-based sampling of 2.0
    - Supports continuous learning through model checkpointing
    - Features rich console output for monitoring training and generation progress
* Why LSTM?
  + LSTMs are ideal for sequential data like text due to their ability to maintain long-term dependencies
  + They can remember important patterns while forgetting irrelevant information through their gate mechanisms
  + Unlike simple RNNs, LSTMs avoid the vanishing gradient problem, making them better at learning long sequences
  + The model can capture the writing style and vocabulary patterns from the source text
* Distributed Implementation Considerations:
  + Data Parallelism: Training can be distributed across multiple GPUs/machines using tf.distribute
  + Model Parallelism: Large models can split layers across different devices
  + Parameter Server Architecture: Central servers can store model parameters while workers handle computation
  + Asynchronous Training: Multiple workers can train independently and periodically sync parameters
  + Distributed TensorFlow strategies can be implemented for scalability
* Implementation Choices:
  + Character-level tokenization provides finer granularity for creative text generation
  + Embedding layer reduces dimensionality and captures character relationships
  + Two-layer LSTM architecture balances complexity and performance
  + Temperature sampling adds controlled randomness for more creative outputs
  + Syllable counting ensures proper haiku structure while maintaining creative freedom

A screenshot of a computer program

Description automatically generated

A screenshot of a computer

Description automatically generated

