Name: Suhaila Ahmed Hassan
Track: AI
Branch: Alexandria

Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling

Objective:

The paper evaluates the performance of two advanced recurrent neural network units, Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU), in comparison with the traditional tanh RNN on sequence modeling tasks, including polyphonic music and speech signal modeling.

Problem with RNNs

RNNs struggle with long-term dependencies due to vanishing/exploding gradients.
LSTM and GRU are gating mechanisms designed to overcome this limitation by enabling better gradient flow and memory retention.

Architectures:

LSTM:

- Contains separate memory cells.
- Uses input, forget, and output gates.
- Controls information exposure through an output gate.

GRU:

- Merges memory and hidden state.
- Uses update and reset gates.
- Exposes a full hidden state directly without output control.

LSTM & GRU:

- Both use additive updates rather than overwriting states, which helps preserve information and gradients.

Experiments:

Datasets:

- Polyphonic Music: Nottingham, JSB Chorales, MuseData, Piano-midi (binary vectors representing music notes).
- Speech: Proprietary Ubisoft datasets (predicting future audio samples based on past ones).

Method:

- Models are trained with the same number of parameters to ensure a fair comparison.

Goal:

- Maximizing sequence log-likelihood.

Results:

- LSTM and GRU outperform traditional tanh RNNs on both music and speech tasks.
- GRUs and LSTMs performed comparably and it is difficult to conclude which is better.
- On some datasets, GRU slightly outperforms LSTM in many cases, where GRUs have faster convergence in training time and better generalization.