# Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling Summary

**Paper Link:** https://arxiv.org/pdf/1412.3555

The paper compares traditional tanh units with two more advanced gated architectures: the Long Short-Term Memory (LSTM) unit and the Gated Recurrent Unit (GRU). While LSTM has been widely adopted in tasks requiring the modeling of long-term dependencies, GRU is a more recent alternative designed to achieve similar benefits with a simpler structure. The study aims to empirically evaluate and compare the performance of these units.

The authors design experiments around two primary tasks: polyphonic music modeling and raw speech signal modeling. They use four well-known music datasets and two proprietary speech datasets. In each experiment, the RNN models were built to have approximately the same number of parameters to ensure a fair comparison. The models were trained using RMSProp with gradient clipping and weight noise, and performance was evaluated using negative log-likelihood on training and test sets.

The experimental results show that on the music datasets, all three models performed similarly, though GRUs slightly outperformed LSTMs in most cases. However, on the more challenging speech datasets, both GRU and LSTM significantly outperformed the traditional tanh-based RNNs. GRU achieved better performance on one of the speech datasets, while LSTM did better on the other. Additionally, GRUs often required fewer updates and less wall-clock time to converge, making them a computationally efficient alternative to LSTM.

The findings indicate that both GRU and LSTM offer substantial advantages over traditional RNNs with tanh units, particularly in tasks involving complex temporal dependencies. Although the results demonstrate the efficacy of gating mechanisms in improving sequence modeling, the comparison between GRU and LSTM is inconclusive. The authors conclude that the choice between GRU and LSTM may depend on the specific characteristics of the dataset and the task at hand.