Bereket Eshete Kebede
UID - U00827234

Graduate Student
The University of Memphis, Fall 2021
EECE 7740 Neural Networks

Recurrent Neural Networks (RNN) for Machine Translation

Instructor: Prof. Zahangir Alom
Assignment #4 Report

Introduction

In this report, we implement Recurrent Neural Networks (RNN) models for machine translation where the trained model will translate sentences from English to French. RNNs are a class of neural networks that allow previous outputs to be used as inputs while having hidden states. RNN models are mostly used in the fields of natural language processing (NLP) and speech recognition. RNNs are preferred for this task because it takes computation historical information into consideration [2].

Our main task is to perform NLP using stacked traditional RNN, LSTM (Long short-term memory), and GRU (Gated recurrent units) as well as compare and contrast the performance of these models. LSTM is a type of RNN meaning unlike feedforward networks, it has a feedback connection. This feature enables it to process sequences of data, such as video, speech, or language. In contrast with traditional RNN, LSTM deals with vanishing gradient problems better than traditional RNN [5]. The other type of RNN we will use for our machine translation task is GRU. GRU is similar to LSTM but with a forget gate because it lacks an output gate, it has fewer parameters than LSTM, but has similar performance in NLP as LSTM [6]. Code for this report is at, https://github.com/bereketeshete/Natural-Language-Processing

Methodology

The methodology of this report follows implementing the following networks listed below and reporting the results.

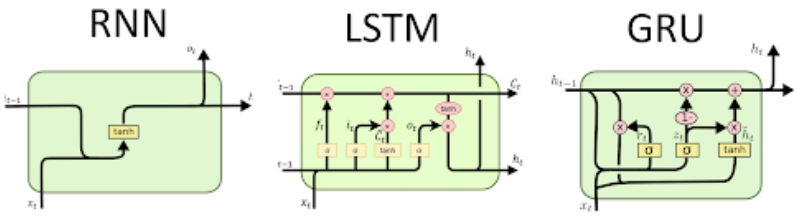
- Implementing stacked (2-layer) RNN
- Implementing stacked (2-layer) LSTM
- Implementing stacked (2-layer) GRU

In this report, we will perform machine translations from English to French and compare the performance among them, and provide at least ten translations from English to French for each model in the report. To perform the following training we will follow the following pre-processing since computers don't understand a language but only understand numbers.

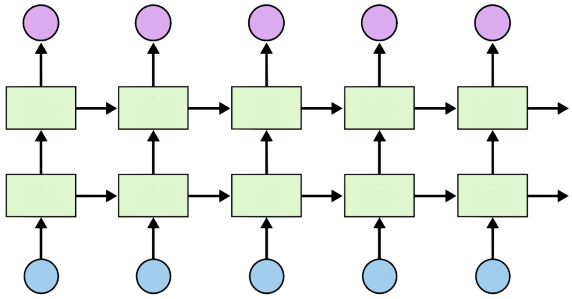
- Row data reading
- Encoding
- Tokenization
- Padding

Deep Learning Architecture

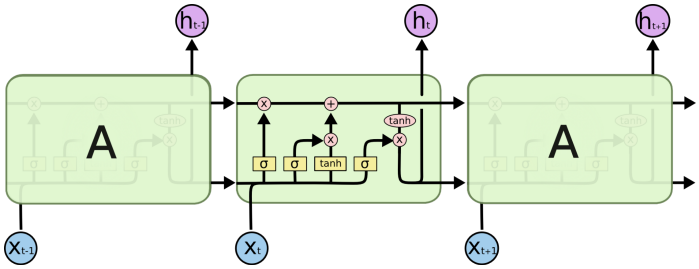
The following basic neural structures in the figure below will be implemented for natural language processing for our experiment after stacking their basic structure.



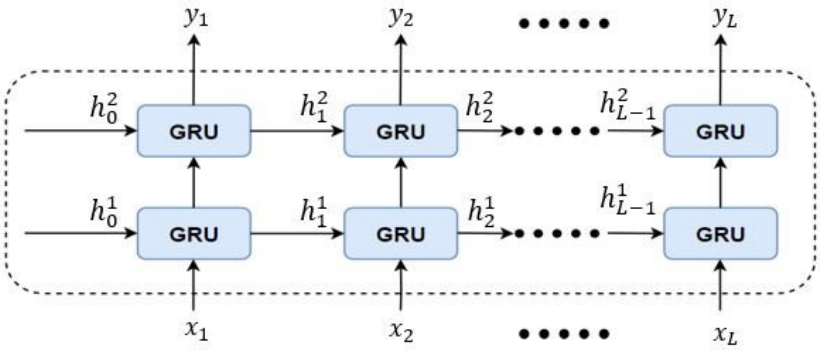
Stacked RNN



Stacked LSTM



Stacked GRU [4]



Experiment and Results

The results from the experiment of stacking RNN's for a machine translation is presented in the following section.

(a) Training and testing logs

Stacked RNN

Training accuracy: 74.31%
Testing accuracy: 74.60%

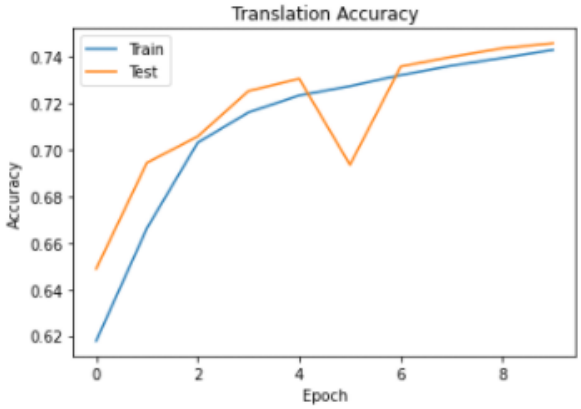


Figure 1. Training and testing accuracy plot for English to French machine translation using stacked RNN neural network.

Stacked LSTM

Training accuracy: 77.07%
Testing accuracy: 77.31%

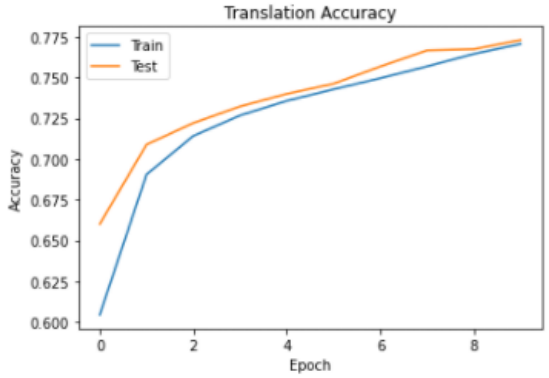


Figure 2. Training and testing accuracy plot for English to French machine translation using stacked LSTM neural network.

Stacked GRU

Training accuracy: 76.47%

Testing accuracy: 76.21%

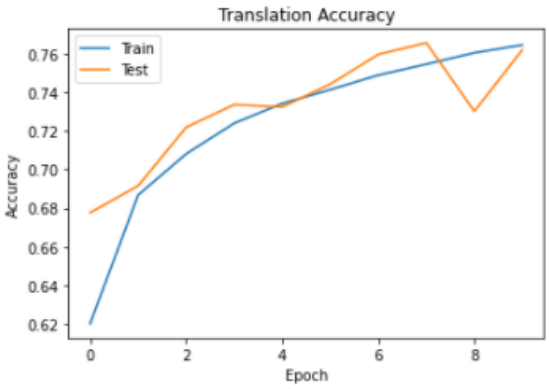


Figure 3. Training and testing accuracy plot for English to French machine translation using stacked GRU neural network.

(b) Discussion and comparison

The following table compares and contrasts the accuracy performance of the different machine translation RNN models we discussed in this report.

Model	Translation Accuracy
Stacked RNN	74.60%
Stacked LSTM	77.31%
Stacked GRU	76.21%

Table 1. Comparison of testing accuracy for the different neural networks used in this report.

Conclusion

In summary, we demonstrate RNN's are suitable for machine translation tasks because they use historical information for training tasks. By referring to table 1, in our experiment stacked LSTM performed the best in comparison to the traditional stacked RNN and the stacked GRU model. The LSTM RNN performed a 77.31% of English to French translation accuracy.

References

[1] **https://github.com/zahangircse/COMP_EECE_7or8740_NNs/blob/main/Lecture_14_RNN.ipynb**
[2] **https://predictivehacks.com/example-of-machine-translation-in-python-and-tensorflow/**
[3]https://github.com/zahangircse/COMP_EECE_7or8740_NNs/blob/main/Assignment_4_ref_codes.ipynb
[4]https://www.researchgate.net/figure/The-structure-of-stacked-GRU-The-two-layered-stacked-network-is-used-in-the-proposed_fig5_341841037
[5] https://en.wikipedia.org/wiki/Long_short-term_memory
[6] https://en.wikipedia.org/wiki/Gated_recurrent_unit