Application of Recurrent Neural Networks for Language Translation

Nguyen Do Nhat Huy, Tieu Nhat Thanh, Nguyen Thanh Nhan

Abstract—**This report investigates the application of Recurrent Neural Networks (RNNs) in the domain of language translation. The study focuses on exploring the architecture, implementation, and performance of RNN models specifically designed for English-Vietnamese translation tasks. Through the development and training of RNN models, including Gated Recurrent Units (GRUs) and Encoder-Decoder architectures, we evaluate their effectiveness in handling translation challenges. The results of the experiments are analyzed to assess model performance and provide recommendations for future improvements.**

I. Introduction

Language translation plays a crucial role in today's globalized world, where effective communication across different languages is essential. In the field of education, translation facilitates the sharing of academic materials and promotes student exchange programs, thereby broadening educational opportunities. Culturally, translation enables access to diverse media such as films, books, and music from various cultures, fostering greater cross-cultural understanding and appreciation. Politically, accurate translation supports international negotiations and the effective communication of national messages. The integration of artificial intelligence (AI) in language translation has the potential to significantly enhance the quality and efficiency of translations, offering valuable advancements in this critical field.

In this project, the team aims to:

- Investigate the architecture and operational mechanisms of Recurrent Neural Networks (RNNs).
- Develop an RNN model for language translation tasks.
- Expand the dataset and evaluate the performance of the model.

II. Methodology

A. Recurrent Neural Network (RNN)

Recurrent Neural Networks (RNNs)[1] are a type of artificial neural network designed to handle sequential data. Unlike traditional feedforward neural networks, RNNs feature connections that form directed cycles, allowing information to persist over time. This characteristic makes RNNs particularly suitable for tasks involving sequences, such as language translation.

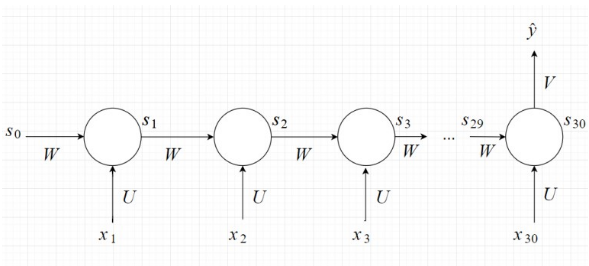

Fig. 1. RNN Workflow

B. Gated Recurrent Unit (GRU)

Gated Recurrent Units (GRUs)[2] are a variant of RNNs that introduce mechanisms to control the hidden state. The main difference between traditional RNNs and GRUs is the presence of gating mechanisms that determine when to update and when to reset the hidden state. This allows the GRU to retain important information from previous observations and discard irrelevant data, improving the model's ability to handle complex sequences.

C. Encoder-Decoder Architecture

The Encoder-Decoder architecture is a widely used model for sequence-to-sequence tasks, such as machine translation. The encoder component transforms the input sequence into a fixed-length context vector, which encapsulates the relevant information. This context vector is then fed into the decoder, which generates the output sequence. In the context of language

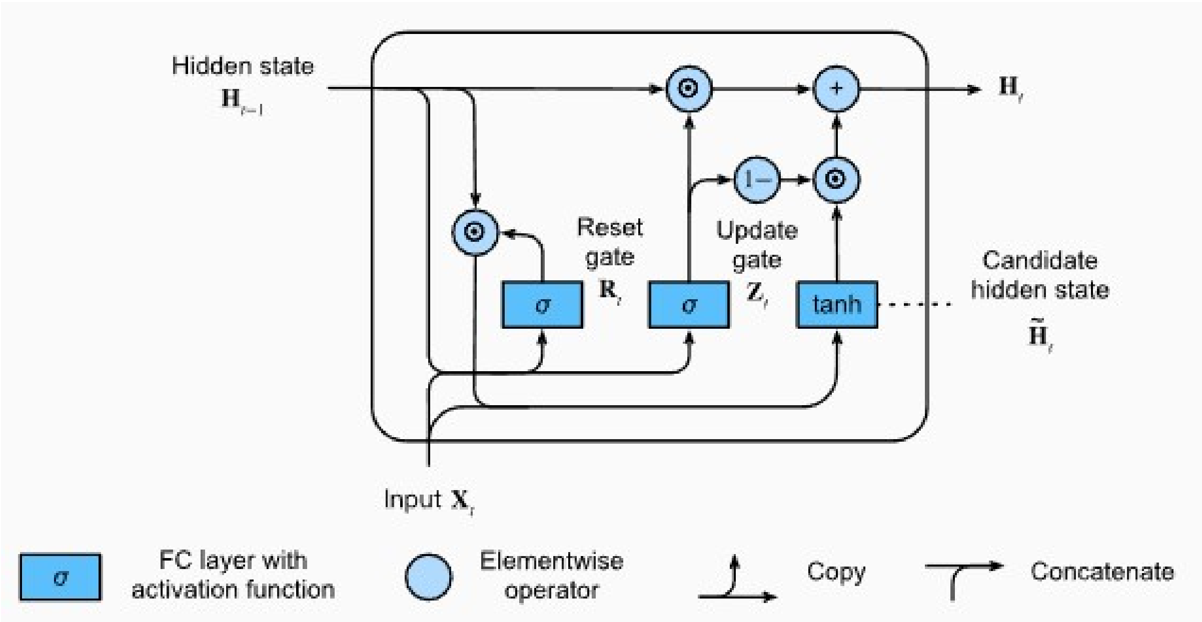

Fig. 2. GRU Architecture

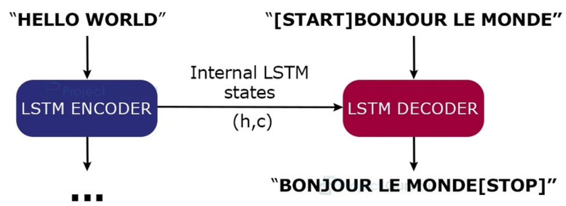

Fig. 3. Encoder-Decoder Visualization

translation, the encoder converts a source sentence into a context vector, and the decoder uses this vector to produce the translated sentence in the target language.

III. Experiment

A. Dataset Overview

The experiment utilizes several datasets:

- **vie.txt**: Contains English and Vietnamese pairs with 9429 rows.
- **Book1.xlsx**: Comprises Vietnamese and English pairs with 254090 rows.
- **Combined Dataset**: Includes Vietnamese sentences with 263482 rows.

B. System Specification

The experiments were conducted on a x64-based PC with the following specifications:

- Processor: Intel(R) Core(TM) i7-1065G7 CPU @ 1.30GHz, 4 Cores, 8 Logical Processors.
- Installed Physical Memory (RAM): 16 GB (15.8 GB usable), with 5.42 GB available.

C. Hyperparameters

The following hyperparameters were used for training the RNN model:

- batch_size: 64
- epochs: 10
- latent_dim: 256
- num_samples: 40000
- Number of unique input tokens: 167
- Number of unique output tokens: 94
- Max sequence length for inputs: 193
- Max sequence length for outputs: 192

D. Unexpected Problems

Several issues were encountered during the experiment:

- **Memory Error**: The system reported "Unable to allocate 14.2 GiB for an array with shape (263518, 263, 110) and data type float32." This was due to the large data size and hardware limitations. The solution involved reducing the data size to 227,546 samples and using a smaller data type (float16).
- **System Freeze**: The system froze during one-hot encoding due to excessive memory usage. To address this, the number of samples was limited to 40,000.

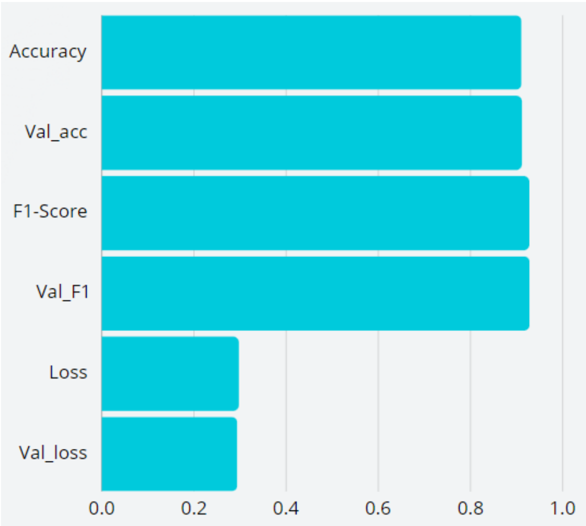

Fig. 4. Final Statistics

E. Results

The RNN model achieved the following performance metrics:

- Accuracy: 91.06%
- F1_score: 0.9277
- Loss: 0.2974
- Validation Accuracy: 91.16%
- Validation F1_score: 0.9283
- Validation Loss: 0.2940
- Training Time: 128 minutes 37 seconds

IV. Analysis

The model demonstrated significant improvement in accuracy, F1_score, and loss over 10 epochs. The high accuracy and F1_score on both training and validation sets indicate a well-balanced model in terms of precision and recall. The decreasing loss values suggest that the model effectively learned and minimized errors. However, increasing the dataset size only marginally improved effectiveness compared to previous training runs, indicating that further enhancements may be necessary.

V. Conclusion

In conclusion, this study has successfully explored and applied RNNs for language translation, with positive outcomes in English-Vietnamese translation tasks. The RNN model showed strong performance metrics, highlighting its capability to handle sequential data and maintain translation accuracy. Future work should focus on integrating more advanced techniques, such as Transformer models or hybrid approaches, to further enhance translation quality and address current limitations.

References

[1] R. M. Schmidt, "Recurrent neural networks (rnns): A gentle introduction and overview," *CoRR*, vol. abs/1912.05911, 2019.

[2] K. Cho, B. van Merrienboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using rnn encoder-decoder for statistical machine translation," 2014.