

Sequence-to-Sequence Learning with LSTM

Deep Learning

Objective

The goal of this assignment is to implement a sequence-to-sequence (seq2seq) model for machine translation from English to Spanish. You will explore two architectures:

1. LSTM Encoder-Decoder without Attention

2. LSTM Encoder-Decoder with Attention

- Bahdanau (Additive) Attention
- Luong (Multiplicative) Attention

You will evaluate the performance of each model on a English-Spanish dataset.

Dataset

You are provided with a text file containing English-Spanish sentence pairs, one pair per line, separated by a tab:

```
Hello. Hola.  
How are you? ¿Cómo estás?  
I am fine. Estoy bien.  
...
```

Note:

If the dataset is too large to train in reasonable time on your system, you may sample a subset (e.g., 10,000 sentence pairs), and split it into:

- 80% training
- 10% validation
- 10% testing

Ensure all preprocessing (tokenization, lowercasing, padding, etc.) is properly handled.

Part 1: LSTM Encoder-Decoder without Attention

- Implement a vanilla seq2seq model using LSTM layers for both the encoder and decoder.
- Use teacher forcing during training.
- Use word embeddings (e.g., learned embeddings or pretrained like GloVe if desired).
- Evaluate the model using BLEU score.

Part 2: LSTM Encoder-Decoder with Attention

- Extend the model by adding attention mechanisms:
 1. Bahdanau Attention
 2. Luong Attention
- Compare the performance of each attention mechanism with the (without attention).
- Visualize attention weights for a few example translations.
- Evaluate the model using BLEU score.

Bonus (Optional)

You may also deploy your trained model using a simple web app or API (e.g., using Flask, Streamlit) that takes an English sentence as input and returns the translated Spanish output. **Please note that they will not contribute to your final grade.**