Problem 1)

Develop a GRU-based encoder-decoder architecture for English to French Translation. Train the model on the entire dataset and evaluate it. Report training loss, validation loss, and validation accuracy. Also, try some qualitative validation, asking the network to generate French translations for some English sentences.

For problem 1, English to French translation was performed. This learning was done with a hidden size of 256, a learning rate of 0.0085 and in 50 epochs. The accuracy was 97.4% and the evaluation loss was 0.1035. The training loss over 45 epochs as well as some predictions can be seen below.

A screenshot of a computer

AI-generated content may be incorrect.

Problem 2:

Repeat problem 1, this time extend the network with attention. Train the model on the entire dataset and evaluate it on the entire dataset. Report training loss, validation loss, and validation accuracy. Also, try some qualitative validation as well, asking the network to generate French translations for some English sentences. Also, compare the results against problem 1.

For problem 2, English to French translation was performed. This learning was done with a hidden size of 256, a learning rate of 0.0085 and in 50 epochs. The accuracy was 98.7% and the evaluation loss was 0.0729. The training loss over 45 epochs as well as some predictions can be seen below. Comparing this to problem one, the model seemed to perform better for every measured metric.

A close-up of a text

AI-generated content may be incorrect.

Problem 3:

Repeat problems 1 and 2, this time try to translate from French to English. Train the model on the entire dataset and evaluate it on the entire dataset. Report training loss, validation loss, and validation accuracy. Also, try some qualitative validation as well, asking the network to generate English translations for some French sentences. Which one seems to be more effective, French-to-English or English-to-French?

For problem 3 part 1, French to English translation was performed. This learning was done with a hidden size of 256, a learning rate of 0.0085 and in 50 epochs. The accuracy was 100% and the evaluation loss was 0.07726. The training loss over 45 epochs as well as some predictions can be seen below.

A close up of a text

AI-generated content may be incorrect.

For problem 3 part 2, French to English translation was performed. This learning was done with a hidden size of 256, a learning rate of 0.0085 and in 50 epochs. The accuracy was 100% and the evaluation loss was 0.05187. The training loss over 45 epochs as well as some predictions can be seen below. Once again with attention added the model performs better. Overall, French to English translation was more accurate and had less loss than English to French.

A close-up of a text

AI-generated content may be incorrect.

Github: <https://github.com/Eskdagoat/4106/blob/main/NicolaAndrew_801136465_HW4.ipynb>