Andrew Nicola 801136465 HW4

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.

```
Epoch 0, Loss: 3.8442407544041086
Epoch 5, Loss: 3.0207676385265354
Epoch 10, Loss: 2.6042925533769647
Epoch 15, Loss: 2.1649717358406533
Epoch 20, Loss: 1.8346153032949253
Epoch 25, Loss: 1.3240529088796085
Epoch 30, Loss: 0.8900329556275436
Epoch 35, Loss: 0.4836146333867115
Epoch 40, Loss: 0.26111635948008943
Epoch 45, Loss: 0.1752822861058507
Input: I am cold, Target: J'ai froid, Predicted: J'ai froid
Input: The sun sets in the evening, Target: Le soleil se couche le soir, Predicted: Le soleil se couche le soir
Input: We love music, Target: Nous aimons la musique, Predicted: Nous aimons la musique
Input: She dreams of flying, Target: Elle rêve de voler, Predicted: Elle rêve de voler
Input: They visit the Eiffel Tower, Target: Ils visitent la tour Eiffel, Predicted: Ils visitent la tour Eiffel
Input: The cat is sleeping, Target: Le chat dort, Predicted: Le chat dort
Input: He sleeps deeply, Target: Il dort profondément, Predicted: Il dort profondément
Input: She wears a red dress, Target: Elle porte une robe rouge, Predicted: Elle porte une robe rouge
Evaluation Loss: 0.10351606466596432, Accuracy: 0.974025974025974
```

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.

```
Epoch 0, Loss: 3.7347967089242338
Epoch 5, Loss: 3.040594063643646
Epoch 10, Loss: 2.576128551690167
Epoch 15, Loss: 2.171311986630687
Epoch 20, Loss: 1.6594764834948361
Epoch 25, Loss: 1.1862098950635427
Epoch 30, Loss: 0.6757440529067942
Epoch 35, Loss: 0.3689108398554251
Epoch 40, Loss: 0.21861835086030515
Epoch 45, Loss: 0.12960668354182112
Input: She studies hard for exams, Target: Elle étudie dur pour les examens, Predicted: Elle étudie dur pour les examens
Input: He sings beautifully, Target: Il chante magnifiquement, Predicted: Il chante magnifiquement
Input: The wind blows gently, Target: Le vent souffle doucement, Predicted: Le vent souffle doucement
Input: They listen to the radio, Target: Ils écoutent la radio, Predicted: Ils écoutent la radio
Input: She speaks French fluently, Target: Elle parle français couramment, Predicted: Elle parle français couramment
Input: The children play in the park, Target: Les enfants jouent dans le parc, Predicted: Les enfants jouent dans le parc
Input: We are friends, Target: Nous sommes amis, Predicted: Nous sommes amis
Input: The sun sets in the evening, Target: Le soleil se couche le soir, Predicted: Le soleil se couche le soir
Evaluation Loss: 0.07297424324620162. Accuracv: 0.987012987012987
```

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.

```
Epoch 0, Loss: 4.01476361393265
Epoch 5, Loss: 3.247063229595649
Epoch 10, Loss: 2.703021876164698
Epoch 15, Loss: 2.107596046657715
Epoch 20, Loss: 1.772784441433059
Epoch 25, Loss: 1.2986708844747528
Epoch 30, Loss: 0.7382359566810918
Epoch 35, Loss: 0.3674306252044002
Epoch 40, Loss: 0.20514154019975778
Epoch 45, Loss: 0.11993708974769916
Input: Elle enseigne l'anglais à l'école, Target: She teaches English at school, Predicted: She teaches English at school
Input: Tu es fatigué, Target: You are tired, Predicted: You are tired
Input: Nous sommes amis, Target: We are friends, Predicted: We are friends
Input: Les fleurs fleurissent au printemps, Target: The flowers bloom in spring, Predicted: The flowers bloom in spring
Input: Elle nage dans l'océan, Target: She swims in the ocean, Predicted: She swims in the ocean
Input: Nous regardons un film ensemble, Target: We watch a movie together, Predicted: We watch a movie together
Input: Les oiseaux gazouillent le matin, Target: The birds chirp in the morning, Predicted: The birds chirp in the morning
Input: Elle danse avec grâce, Target: She dances gracefully, Predicted: She dances gracefully
Evaluation Loss: 0.07725600125838297, Accuracy: 1.0
```

Andrew Nicola 801136465 HW4

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.

```
Epoch 0, Loss: 4.033506503603526
Epoch 5, Loss: 3.042147725632549
Epoch 10, Loss: 2.655589724825575
Epoch 15, Loss: 2.0213546907599507
Epoch 20, Loss: 1.4257666318300117
Epoch 25, Loss: 0.832981334797158
Epoch 30, Loss: 0.39717954489170354
Epoch 35, Loss: 0.19716492170121244
Epoch 40, Loss: 0.1111489831843749
Epoch 45, Loss: 0.07492291003087824
Input: Nous apprenons quelque chose de nouveau chaque jour, Target: We learn something new every day, Predicted: We learn something new every day
Input: Il attend le bus, Target: He waits for the bus, Predicted: He waits for the bus
Input: La pluie tombe doucement, Target: The rain falls gently, Predicted: The rain falls gently
Input: Le restaurant sert une délicieuse cuisine, Target: The restaurant serves delicious food, Predicted: The restaurant serves delicious food
Input: Il chante dans le chœur, Target: He sings in the choir, Predicted: He sings in the choir
Input: Elle danse avec grâce, Target: She dances gracefully, Predicted: She dances gracefully
Input: Elle nage dans l'océan, Target: She swims in the ocean, Predicted: She swims in the ocean
Input: Il écrit une lettre, Target: He writes a letter, Predicted: He writes a letter
Evaluation Loss: 0.05187436859840475, Accuracy: 1.0
```

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