**Homework #4**

GitHub: <https://github.com/pballou/ECGR_4106/tree/master/Homework>

1. GRU-based encoder-decoder architecture for English to French Translation
   1. Train loss: 0.0106
   2. Validation loss: 0.0102
   3. Validation accuracy: 100%
2. GRU-based encoder-decoder architecture with attention for English to French Translation
   1. Train loss: 0.0088
   2. Validation loss: 0.0082
   3. Validation accuracy: 100%
3. Problems 1 and 2, but for French to English Translation
   1. Without attention
      1. Train loss: 0.0098
      2. Validation loss: 0.0094
      3. Validation accuracy: 100%
   2. With attention
      1. Train loss: 0.0089
      2. Validation loss: 0.0083
      3. Validation accuracy: 100%

Conclusions:

* The dataset is pretty small which makes the problem relatively easy for a model of this complexity.
* My parameters are as follows:
  + Epochs = 51
  + Learning rate = 0.01
  + Hidden size = 1028
* I found that a larger learning rate of .01 actually helped the model learn faster without overfitting, which is ideal. In a larger, more complex dataset this would likely not work.
* Epochs at 50 and a hidden size of 1028 also helped the model learn better and are a bit overkill for this specific problem due to its simplicity, but it’s fun to see how good it can be, especially since the training was quick.
* Adding attention helped the training and validation be more accurate.
* French to English translation was actually slightly more accurate, though the attention models were basically the same.