

Questions

“Training”

1. What is the `nn.GRU` module? How is this module similar to the LSTM? How is it different?
2. What is the “temperature”?
3. When generating text using a trained RNN model, do we obtain cleaner samples using a *high* temperature or a *low* temperature?
4. What loss function do we use when training an RNN model to generate text? Why?

“Generalization”

1. When generating text using a trained RNN model, why do we “prime” the RNN hidden state?
2. How does the “temperature” affect RNN training?
3. We did not set aside a “test set” for our generative RNN model. Why?
4. Is text generation a supervised or unsupervised learning problem?
5. How does the chunk length (the length of each training sequence) affect the training of RNN models to generate text?
6. During training, we use the ground-truth sequence as input to the RNN. However, during test time, we used the generated sequences as input. Can the disconnect between what is done during training and what is done be a problem? Why or why not?