Lab5 : Conditional VAE

0516069 翁英傑

1. Introduction

Seq2seq VAE is an common machine translation tool. In this lab we implement a conditional VAE where the condition is the tense of the verb.

2. Experiment setups

A. Dataloader:

In training setting, it gives a random tense of a verb with the tense class.

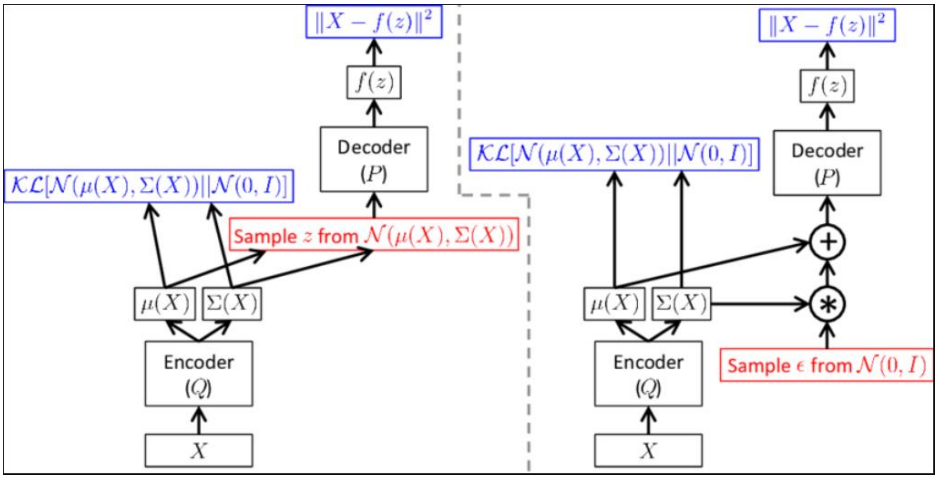
B. Encoder:

Add condition size in input layer, can be seen as increase input size.

C. Decoder:

Also add size in the latent layer, can be seen as increase feature. And take the ground truth as previous output if teacher force is true.

D. Reparameterization trick:



Shown as the graph, we sample a normal distribution sample_z and does the linear combination with the computed mean and variance so that the gradient can be compute.

E. RNN type: nn.GRU.

3. Hyper Parameters

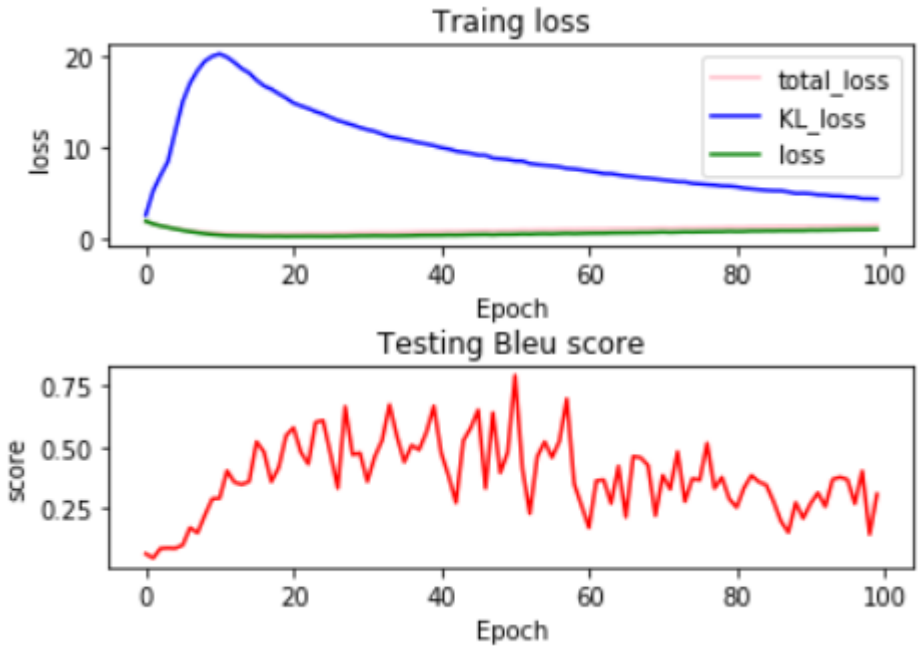
A. KL weight: (epoch%100)*0.001 (linear increase)

B. Learning rate: 1

C. Teacher forcing ratio: 1 - epoch*0.005

D. Epochs: 200

4. Results of testing



Teacher forcing is a trick when training RNN, since much more mistakes will be make in the early epochs. So taking the ground truth as the input of the next cell can help decrease the error and help converge. As the epoch increases, the teacher forcing rate can decrease.

While the KL-divergence might cause the encoder to ignore the input and thus we cycle the weight of the loss function, so that the problem could be solved.

The Bleu-4 score decreases might be the result of the influence of the KL-divergence, where I set 100 as a cycle that the weight of KL-divergence starts from 0. A delicate adjustment might fix this problem.