BBM 497: NLP LAB THE FOURTH ASSIGNMENT

A PREPRINT

Ufuk Baran Karakaya
ufuk.karakaya@hacettepe.edu.tr

June 2, 2019

1 Introduction

In this assignment, the code implements the Feed Forward Network to generate poems according to the dataset. Dynet Framework is used to implement model and train the model epoch by epoch. In this report includes the definition of problem, the approach, analysis of parameters and its effects.

2 Algorithm

The algorithm is based on main problem of the project which is creation of model. The code implements the model with using dynet framework and train the model epoch by epoch. First of all, the code reads the data and holds the poems, and unique words of the content. Count of unique words is used for implementation of parameters. Dynet model is based on one-hot vectors. The FNN predicts the next word according to first element of the bigram.

The code updates the weights of the model according to loss ratio. The main purpose is obtaining ideal weights for the prediction. The one hot vectors has the zeros and one 1 element which refers to the word. Thus, loss ratio is calculated from this index and cross-entropy shows the loss. Before the loss calculation, activation function (tanh) is used for balanced distribution and softmax function optimize the range.

After the train process the the code saves the model with optimized parameters and generation part load the model with essential parameters. However, this project cannot overcome the finding same index in the generation part. Some errors in the loss function or, low epoch counts may lead this problem.

2.1 Train the Model

The model based on this formula:

$$f(x) = U.tanh(W.x + b) + d \tag{1}$$

- W = hidden size, length of unique words
- U = length of unique words, hidden size
- b = hidden size
- d = length of unique words

2.2 Calculation of Loss

the model is trained with the formula and calculation of loss is based on cross- entropy. If we consider the one-hot vector elements, it is equal to negative loss function.

The code calculates the loss with using pick command of dynet library. The code picks the specific index is related to word. The model is updated with loss value and epoch by epoch the process repeats itself.

Table 1: Epoch - Loss- Hidden Size

Epoch	Loss	HiddenSize
1	28.904285	10
1	23.914406	20
1	11.66194	100
5	13.343988	10
5	10.785385	20
5	8.328138	100
20	11.95466	10
20	8.854932	20
20	3.048399	100

3 EFFECTS OF PARAMETERS

The effects of parameters such as dimension of hidden layer and, epoch number are crucial for loss ratio. The model is trained epoch by epoch and loss ratio should reduce in every epoch because the weights are updated according to loss.

On the otherhand, hidden dimension holds the details of the model, so many hidden neurons provide more robustness for solution. But, it takes much time than less neurons.

4 References

- bbm 497 4th problemset sheet
- bbm 495 lecture slides
- http://www.phontron.com/class/mtandseq2seq2018/assets/slides/mt-fall2018.chapter5.pdf