

Lab5 Report

fmonteroperez1 - aca15fm

For this lab our task was to learn how to create natural language models using pytorch. We had to analyze some code from the tutorials and adapt it to use our toy training set.

1 Dense Embedding Model

This language model is based on the idea similar words share similar contexts so we want to model language like classification. The model separates the training data into trigrams and uses the first two elements as context for the third one (the target word).

The model has two layers, namely linear1 and linear2 which map the embedding vectors to a 128 space after they pass through a ReLU activation function. Then they map it back to the vocabulary size space. Then the output of the model turns the vectors of integers into probabilities from which we can later find the max to convert it into words.

2 Sanity Check

To test that the implementation is correct with a small data set. We use the five given sentences in order to train the model. Our model should predict the word "mathematician" when it sees the context "START The". The word mathematician is chosen over the word physicist because it is seen more times in that context than the latter.

The word "mathematician" is predicted more than five times in a row with learning rate 0.002 and 15 epochs. 15 epochs seems to be a good number of epochs due to a loss below 60 in comparison to epochs numbers smaller than 12 or greater than 17. Aside from the loss at 15 epochs, at this number we start to consistently predict the word mathematician for that context.

3 Testing The Model

The bigram language model would be a good fit for our toy training set because currently we were only taking into account the 2 previous words as context to decide what the target word is. The bigram model takes into account all the words in a sentence and returns the most likely combination so it would be able to return the correct word for the given sentence because it's similar to those sentences it

knows. But with this model we can see that the vectors for mathematician and physicist are closer together than that of philosopher and mathematician.