

Assignment 3 - Understanding The Challenge

324827328, 213713522

May 2025

1. Can the two languages be distinguished using a bag-of-words approach? Explain why.

No, a CBOW approach will not work here, because a CBOW approach does not give meaning to order, and in the described language, the meaning of words is critical. For example, in a CBOW approach, the following sequence: "1a2b3c4d5", which is part of the language, will be equivalent to "5d4c3b2a1" (The inverse sequence), which is **not** part of the language.

2. Can the two languages be distinguished using a bigram or trigram based approach? Explain why.

No, the two languages can not be distinguished using a bigram or trigram based approach, because in a bigram and a trigram approach, we only take into account the last 1 and 2 characters. For example, if will look at the sequence "12a34b56c78d90", when the model will look at the letter "c", it will only see the characters "6" (in a bigram), and "56" (in a trigram). The model will not see "b", so it will not know what came before, b or c, making it impossible for this approach to distinguish the languages. The issue in general for n-gram, is that distinguishing these languages requires capturing long-distance dependencies. For example, to know if a sequence is valid, you need to check if 'b' comes before 'c' - but they could be separated by many digits, making the dependency too distant for any reasonable n-gram model to capture.

3. Can the two languages be distinguished using a convolutional neural network? Explain why.

It might be possible to distinguish the languages using a convolutional neural network. In a convolutional neural network, the model has access to the entire sequence, making it possible for the model to know what came before, the character "b" or "c". For example, in the following convolution network(see figure 1), that we learned in class, we can see that even though in the first layer the network doesn't have access to the entire sequence, it can learn filters that effect how later layers see the entire sequence.

Dilated Convolutions

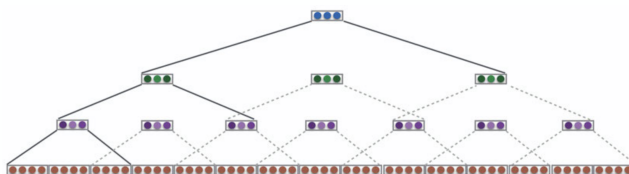


Figure 1: Convolutional Network Learned In Class