

## **Part1 - challenge**

1. The two languages cannot be distinguished using bag-of-words approach, because the representation of good and bad words will be identical, because the order between the letters is the only thing that's different.
2. Bigrams would not work, because there isn't any informative bigram – every bigram can be in either word. trigram would work, only in words that contain informative trigram, that is: 'a[1-9]b' 'b[1-9]c' 'c[1-9]d' for good words, and 'a[1-9]c' 'c[1-9]b' 'b[1-9]d' for bad words.
3. Convolutional neural network may work, if the filter used would be big enough (at least 3 letters). Because like with bigrams and trigrams, only then we would get informative windows.