

Assignment 3 - RNN Acceptors and BiRNN Transducers

Itay Etelis, 209041474,
Matan Avitan, 318287455

June 2023

1 Part 1: RNN Acceptor

1.1 Understanding the Challenge

Q1: Can the two languages be distinguished using a bag-of-words approach? Explain why.

A1: The simple answer to this is no, a BOW approach is not sufficient to differentiate these two languages. The BOW methodology can count the frequency of words in a sequence, but it cannot account for the sequence's order. In this particular challenge, sequence order is a critical factor, given that the same words can appear in different orders to form different sentences. To illustrate this formally, let's consider two sequences: Seq_1 belongs to the category of "good sequences" while Seq_2 is classified as a "bad sequence". Let's assume that $Seq_1 = 1a1b1c1d1$ and $Seq_2 = 1a1c1d1b1$. If we utilize the BOW method, both sequences will be represented by identical feature vectors. As a result, any model relying on the BOW approach would yield the same output for both, symbolically represented as $M(Seq_1) = M(Seq_2)$. However, as outlined in the assignment, a model that accurately distinguishes the two should provide different predictions for each. Hence, the BOW approach is insufficient in this case.

Q2: Can the two languages be distinguished using a bigram or trigram based approach? Explain why.

A2: In this case the answer is also no, both bigram and trigram approaches are also insufficient for distinguishing these two languages. Unlike the bag-of-words method, the n-gram approach does consider the order of elements in a sequence to a certain extent. However, in this scenario, this feature does not prove beneficial. The languages under consideration allow for multiple random length letters to be placed between key letters (a, b, c, d). Because of this variability, the n-gram methodology cannot maintain the relation across a long sequence, thus proving ineffective for this particular task.

Q3: Can the two languages be distinguished using a CNN model? Explain why.

A3: The languages presented in the question cannot be effectively distinguished using a CNN model. Although CNNs excel at capturing local patterns within data, the languages in question require a global understanding of the sequence order. Since the filter sizes in CNNs are fixed, setting them to a sufficiently large value to capture the necessary global dependencies would make the training process more challenging and less robust. This would likely lead to lower performance in distinguishing between the languages.