

For a quick reference:

i) previous hidden unit, previous character and category

ii) previous hidden unit and previous character

iii) previous hidden unit and category

iv) previous hidden unit.

As you can see from the graph we can see that (i) performs the best, followed by (ii), (iii), and (iv) (Numbered according to the requirements in the question). This makes sense as since the previous character and category is being supplied in each step, it helps the RNN to learn better and faster. Then between the character and category, the character contains more information.This helps the RNN with the supplied previous character perform better. This is followed by the category and then at last we have the RNN which is not supplied the previous character and category.