

0.d

The advantage of character level model stems from the fact that two words can be very similar, for example "walk" and "walked". With a word level model we don't leverage this similarity and ignore this information which may lead to poorer performance. Character level model can infer that the sequence "walk" and the sequence "walked" are similar. If we treat words with capital letters as different words, e.g. "Walk" and "walk" then it's even more confusing to word level model. The disadvantage of character level model is that it increases the length of the sequence. If we have a document of 100 words, and the average length of a word is for example 4, then with a word level model the sequence length is 100 and with a character level model the length is 400.

References