CSCI 544 HW6
Qifan Wang
9369314001

4.

I would say it depends which metrics is better.

However, overall I would choose the conlleval metrics. The

conlleval metrics is more intuitive and with the changing of

feature, but it only contain one type, however the original

metrics contain B and I type of scores, which makes the

result more well-rounded. And I found that the original

metrics tend to have higher accuracy then the conlleval.

As I mentioned before, while I modify some test cases, such

as long word, the conlleval doesn't change at all, but the

original metrics shows the difference by less than 0.1%, but it

changed anyway. So from some perspective, I would say

that the original metrics is better.