

6. (9 points) Suppose you have trained a Conditional Random Field model, but you realize one of the test sentences is unlabeled.

a. (3 points) Tag the following sentence using NER and the BIO tagset and the classes PERSON, LOCATION, ORGANIZATION, and OTHER:

It has been 53 years since NASA astronauts Neil Armstrong and Buzz Aldrin landed on the Moon

O O O O O O B-ORG O B-PER I-PER O B-PER I-PER O O O B-LOC

Now suppose you run the sentence from part a through the model, it returns the following prediction:

It has been 53 years since NASA astronauts Neil Armstrong and Buzz Aldrin landed on the Moon
O O O O O O O O O B-PER I-PER O O B-PER O O B-ORG

- b. (3 points) Calculate the **precision and recall** of named entity spans on this example. (This is the standard NER evaluation.)

precision: Neil Armstrong: 100% correct
Buzz Aldrin: 50% correct
Moon: 0% correct

$\frac{1+0.5+0}{3} = 50\%$ precision

Recall

Neil Armstrong: 1
Buzz Aldrin: 0.5
Moon: 0

$\frac{1.5}{4} = 37.5\%$ recall

- c. (3 points) Name **two** ways that you could modify your CRF model to increase **recall** (in general, not necessarily on this example). These could be changes to the model or to how it's trained.

- Increase sensitivity of our model; we may predict more false positives, but predicting more positives increases chance of recall increasing
- We could get wider training data set to increase chance that a tag has been seen before or is already in vocabulary.