

Report - Assignment 4

Tuning Efforts

For the Hidden Markov Model (HMM) tagger, the parameters used to tune the tagger include the Lidstone, MLE, ELE, and LaPlace probability distribution estimators. Using these various probability distribution estimators can produce a more robust HMM tagger, as the performance of the probability distribution estimators is dependent on the training dataset. By comparing the performance of these estimators, we can determine the best estimator to choose for calculating the accuracy. The parameters that yield the best performance for the HMM tagger are the Lidstone probability distribution estimates, with a gamma value of 0.00001.

For the Brill tagger, the parameters used to tune the tagger include three template collections. These are composed of the nltkdemo18, fntbl37, and brill24 template collections. This influenced the accuracy of the tagger because different templates capture different linguistic nuances. The parameters that yield the best performance for the Brill tagger is the brill24 template collection, with a max rule count of 200.

Tagger Performance

| Tagger | Input (dataset used) | Accuracy of the Tagger |
|--------|----------------------|------------------------|
| hmm | data/test.txt | 81.05 |
| hmm | data/test_ood.txt | 78.68 |
| brill | data/test.txt | 72.45 |
| brill | data/test_ood.txt | 67.71 |

Tagger Errors

| HMM Misclassifications | Brill Misclassifications |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <ol style="list-style-type: none">1. Misclassifies "'s" as POS instead of VBZ. The possessive marker is highly context-dependent, and having contractions in datasets can be hard for taggers to discern.2. Misclassifies many out-of-vocabulary (OOV) words as NN. This is most likely due to the tagger recognizing that the NN tag appears many times throughout the dataset, assigning it a | <ol style="list-style-type: none">1. Misclassifies verbs as nouns (NN) in sentences where the tag of the verb is highly context-dependent, or the verb is at the start of a sentence. Other words are also misclassified as NN if the rules are not in the baseline rules. This led to overclassification of NN in the data.<ol style="list-style-type: none">a. Example: Lines 1, 15, and 102 in test_brill.txt are VB |

| | |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| higher transition probability. Because of the higher transition probability, OOV default to a NN tag. This can be a form of bias the HMM tagger has learned. | missclassified as NN. 2. Misclassifies "'s" as POS instead of VBZ. The possessive marker is highly context-dependent, and having contractions in datasets can be hard for taggers to discern. 3. All NNP are misclassified as NN. This is most likely due to NNP not being an initial rule, leading to misclassification. |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

Comparison of Tagger Errors

| <u>Similarities</u> Between HMM Tagger and Brill Tagger Misclassifications | <u>Differences</u> Between HMM Tagger and Brill Tagger Misclassifications |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <ul style="list-style-type: none">Both taggers misclassify the "'s" possessive marker as POS instead of VBZ. Because there is much contextual ambiguity surrounding possessive markers/contractions in terms of tagging, it is expected that both the HMM and Brill tagger would struggle to correctly classify these.Both taggers misclassify many out-of-vocabulary (OOV) words as NN, leading to overclassification of NN in all datasets. | <ul style="list-style-type: none">The difference between misclassifications for in domain vs out of domain data was higher for Brill tagging in comparison to HMM tagging. Brill tagging seems to struggle more with out of domain data, which can be due to HMM being able to make more contextual inferences based off of the sentence a word is located in. This is most likely why more NN misclassification is present in the out-of-domain Brill tagged dataset. |