

Unlabeled test set

Gold Accuracy = 67%

