CS 534 HW3

Ming-hung Shih

1.

```
epoch   1, updates 68, features 1494, train_err 2.54%, dev_err 5.26%
epoch   2, updates 80, features 1494, train_err 2.97%, dev_err 5.85%
epoch   3, updates 70, features 1494, train_err 2.67%, dev_err 6.14%
epoch   4, updates 76, features 1494, train_err 2.84%, dev_err 5.85%
epoch   5, updates 75, features 1494, train_err 2.86%, dev_err 6.14%
epoch   6, updates 75, features 1494, train_err 2.81%, dev_err 6.14%
epoch   7, updates 74, features 1494, train_err 2.73%, dev_err 5.85%
epoch   8, updates 70, features 1494, train_err 2.73%, dev_err 5.26%
epoch   9, updates 74, features 1494, train_err 2.78%, dev_err 6.14%
epoch  10, updates 74, features 1494, train_err 2.81%, dev_err 5.26%
best error rate 5.26%
```

2.

```
epoch   1, updates 68, features 1494, train_err 2.54%, dev_err 5.26%, avg_dev_err 4.97%
epoch   2, updates 80, features 1494, train_err 2.97%, dev_err 5.85%, avg_dev_err 4.97%
epoch   3, updates 70, features 1494, train_err 2.67%, dev_err 6.14%, avg_dev_err 4.97%
epoch   4, updates 76, features 1494, train_err 2.84%, dev_err 5.85%, avg_dev_err 5.26%
epoch   5, updates 75, features 1494, train_err 2.86%, dev_err 6.14%, avg_dev_err 5.26%
epoch   6, updates 75, features 1494, train_err 2.81%, dev_err 6.14%, avg_dev_err 4.68%
epoch   7, updates 74, features 1494, train_err 2.73%, dev_err 5.85%, avg_dev_err 4.68%
epoch   8, updates 70, features 1494, train_err 2.73%, dev_err 5.26%, avg_dev_err 4.68%
epoch   9, updates 74, features 1494, train_err 2.78%, dev_err 6.14%, avg_dev_err 4.68%
epoch  10, updates 74, features 1494, train_err 2.81%, dev_err 5.26%, avg_dev_err 4.68%
best error rate 5.26%
best error rate average 4.68%
```

3. unavg perceptron time 17.64 sec

   avg perceptron time 20.80 sec

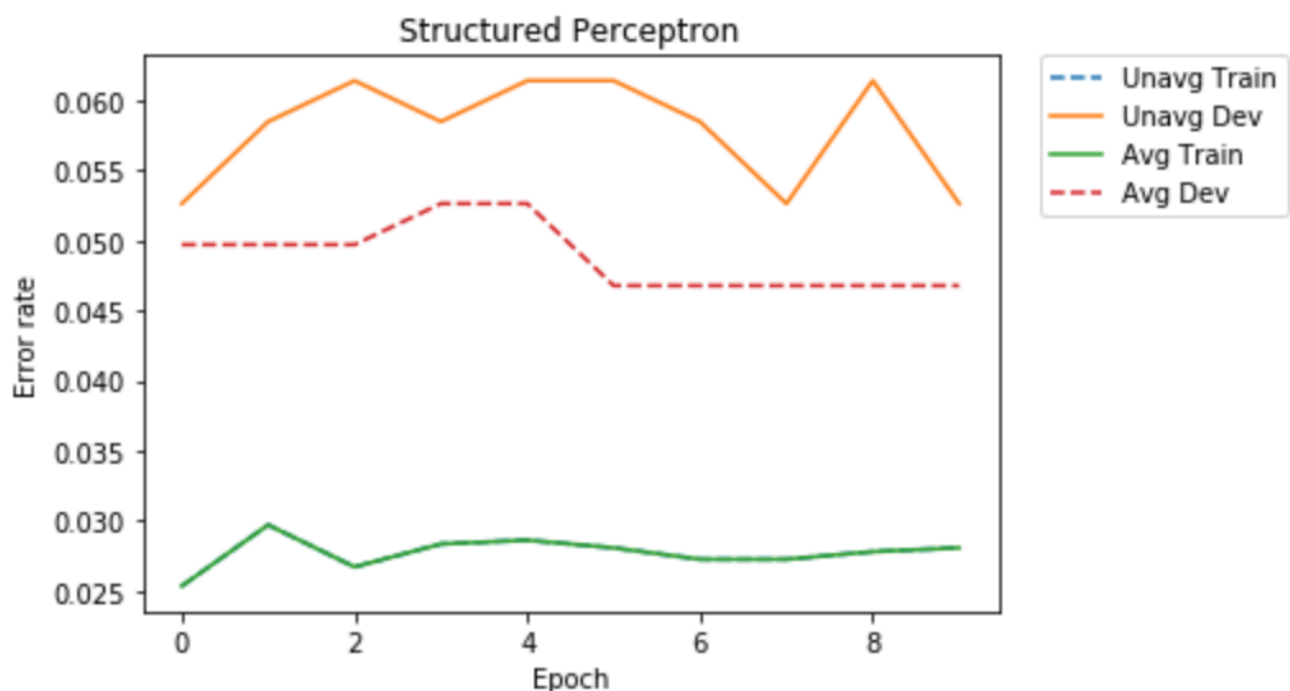avg perceptron is slower than unavg perceptron about

20.8/17.64 = 1.17913832 -> 117% time.

```
epoch  1, updates 68, features 1494, train_err 2.54%, dev_err 5.26%
epoch  2, updates 80, features 1494, train_err 2.97%, dev_err 5.85%
epoch  3, updates 70, features 1494, train_err 2.67%, dev_err 6.14%
epoch  4, updates 76, features 1494, train_err 2.84%, dev_err 5.85%
epoch  5, updates 75, features 1494, train_err 2.86%, dev_err 6.14%
epoch  6, updates 75, features 1494, train_err 2.81%, dev_err 6.14%
epoch  7, updates 74, features 1494, train_err 2.73%, dev_err 5.85%
epoch  8, updates 70, features 1494, train_err 2.73%, dev_err 5.26%
epoch  9, updates 74, features 1494, train_err 2.78%, dev_err 6.14%
epoch 10, updates 74, features 1494, train_err 2.81%, dev_err 5.26%
best error rate 5.26%
unavg perceptron time 17.64 sec
epoch  1, updates 68, features 1494, train_err 2.54%, dev_err 5.26%, avg_dev_err 4.97%
epoch  2, updates 80, features 1494, train_err 2.97%, dev_err 5.85%, avg_dev_err 4.97%
epoch  3, updates 70, features 1494, train_err 2.67%, dev_err 6.14%, avg_dev_err 4.97%
epoch  4, updates 76, features 1494, train_err 2.84%, dev_err 5.85%, avg_dev_err 5.26%
epoch  5, updates 75, features 1494, train_err 2.86%, dev_err 6.14%, avg_dev_err 5.26%
epoch  6, updates 75, features 1494, train_err 2.81%, dev_err 6.14%, avg_dev_err 4.68%
epoch  7, updates 74, features 1494, train_err 2.73%, dev_err 5.85%, avg_dev_err 4.68%
epoch  8, updates 70, features 1494, train_err 2.73%, dev_err 5.26%, avg_dev_err 4.68%
epoch  9, updates 74, features 1494, train_err 2.78%, dev_err 6.14%, avg_dev_err 4.68%
epoch 10, updates 74, features 1494, train_err 2.81%, dev_err 5.26%, avg_dev_err 4.68%
best error rate average 4.68%
avg perceptron time 20.80 sec
```

4.



5. By replacing <unk>, it makes model to construct the model easily. It reduces the weight of the (word, tag) because <unk> only appear one time. If we mark it as many new weight vector, the dimension will become very large and reduce the reliability.

## 2 Fearture Engineering

### Test 1 Prefix:

I tried to take the prefix of the word and combine with tag to test the result. For example, Taipei and Tainan has the same prefix so it can reduce the model with the same tag.

```
epoch  1, updates 68, features 1790, train_err 2.56%, dev_err 6.73%
epoch  2, updates 78, features 1790, train_err 2.86%, dev_err 6.14%
epoch  3, updates 72, features 1790, train_err 2.62%, dev_err 5.56%
epoch  4, updates 79, features 1790, train_err 2.92%, dev_err 5.85%
epoch  5, updates 78, features 1790, train_err 2.84%, dev_err 5.56%
epoch  6, updates 82, features 1790, train_err 3.05%, dev_err 7.02%
epoch  7, updates 77, features 1790, train_err 2.73%, dev_err 6.43%
epoch  8, updates 76, features 1790, train_err 2.75%, dev_err 6.43%
epoch  9, updates 84, features 1790, train_err 2.94%, dev_err 6.43%
epoch 10, updates 82, features 1790, train_err 2.86%, dev_err 5.85%
best error rate 5.56%
unavg perceptron time 23.33 sec
epoch  1, updates 68, features 1790, train_err 2.56%, dev_err 6.73%, avg_dev_err 6.14%
epoch  2, updates 78, features 1790, train_err 2.86%, dev_err 6.14%, avg_dev_err 5.85%
epoch  3, updates 72, features 1790, train_err 2.62%, dev_err 5.56%, avg_dev_err 5.85%
epoch  4, updates 79, features 1790, train_err 2.92%, dev_err 5.85%, avg_dev_err 5.26%
epoch  5, updates 78, features 1790, train_err 2.84%, dev_err 5.56%, avg_dev_err 5.26%
epoch  6, updates 82, features 1790, train_err 3.05%, dev_err 7.02%, avg_dev_err 5.26%
epoch  7, updates 77, features 1790, train_err 2.73%, dev_err 6.43%, avg_dev_err 5.56%
epoch  8, updates 76, features 1790, train_err 2.75%, dev_err 6.43%, avg_dev_err 5.56%
epoch  9, updates 84, features 1790, train_err 2.94%, dev_err 6.43%, avg_dev_err 5.56%
epoch 10, updates 82, features 1790, train_err 2.86%, dev_err 5.85%, avg_dev_err 5.56%
best error rate average 5.26%
avg perceptron time 23.72 sec
```

However, it doesn't work in this dataset because error rate is as same as MLE but spend more time.

### Test 2 Suffix:

I tried to take the Suffix of the word and combine with tag to test the result. For example, Singing and Playing has the same Suffix so it can reduce the model with the same tag.

```
train_err 2.73%
dev_err 5.26%
epoch  1, updates 67, features 1764, train_err 2.56%, dev_err 6.73%
epoch  2, updates 71, features 1764, train_err 2.70%, dev_err 6.43%
epoch  3, updates 72, features 1764, train_err 2.64%, dev_err 5.26%
epoch  4, updates 80, features 1764, train_err 2.97%, dev_err 6.73%
epoch  5, updates 79, features 1764, train_err 2.84%, dev_err 7.31%
epoch  6, updates 76, features 1764, train_err 2.89%, dev_err 5.56%
epoch  7, updates 80, features 1764, train_err 2.89%, dev_err 7.31%
epoch  8, updates 75, features 1764, train_err 2.59%, dev_err 5.85%
epoch  9, updates 69, features 1764, train_err 2.37%, dev_err 6.73%
epoch 10, updates 77, features 1764, train_err 2.81%, dev_err 6.14%
best error rate 5.26%
unavg perceptron time 19.33 sec
epoch  1, updates 67, features 1764, train_err 2.56%, dev_err 6.73%, avg_dev_err 6.14%
epoch  2, updates 71, features 1764, train_err 2.70%, dev_err 6.43%, avg_dev_err 5.85%
epoch  3, updates 72, features 1764, train_err 2.64%, dev_err 5.26%, avg_dev_err 5.56%
epoch  4, updates 80, features 1764, train_err 2.97%, dev_err 6.73%, avg_dev_err 5.56%
epoch  5, updates 79, features 1764, train_err 2.84%, dev_err 7.31%, avg_dev_err 4.97%
epoch  6, updates 76, features 1764, train_err 2.89%, dev_err 5.56%, avg_dev_err 5.56%
epoch  7, updates 80, features 1764, train_err 2.89%, dev_err 7.31%, avg_dev_err 5.26%
epoch  8, updates 75, features 1764, train_err 2.59%, dev_err 5.85%, avg_dev_err 5.26%
epoch  9, updates 69, features 1764, train_err 2.37%, dev_err 6.73%, avg_dev_err 5.26%
epoch 10, updates 77, features 1764, train_err 2.81%, dev_err 6.14%, avg_dev_err 5.26%
best error rate average 4.97%
avg perceptron time 26.85 sec
```

Fortunately, it works because it got better error rate in epoch 5.


Test 3 Suffix+Bigram

Combining with previous tag, current tag, and Suffix to get better error rate.

```
epoch  1, updates 68, features 3804, train_err 2.67%, dev_err 6.14%
epoch  2, updates 71, features 4063, train_err 2.56%, dev_err 4.97%
epoch  3, updates 55, features 4063, train_err 1.94%, dev_err 5.56%
epoch  4, updates 42, features 4063, train_err 1.39%, dev_err 6.14%
epoch  5, updates 50, features 4063, train_err 1.72%, dev_err 5.26%
epoch  6, updates 43, features 4063, train_err 1.50%, dev_err 4.68%
epoch  7, updates 37, features 4063, train_err 1.17%, dev_err 6.14%
epoch  8, updates 35, features 4063, train_err 1.12%, dev_err 6.14%
epoch  9, updates 36, features 4063, train_err 1.23%, dev_err 4.97%
epoch 10, updates 33, features 4063, train_err 1.09%, dev_err 4.97%
best error rate 4.68%
unavg perceptron time 19.21 sec
epoch  1, updates 68, features 3804, train_err 2.67%, dev_err 6.14%, avg_dev_err 5.85%
epoch  2, updates 71, features 4063, train_err 2.56%, dev_err 4.97%, avg_dev_err 4.97%
epoch  3, updates 55, features 4063, train_err 1.94%, dev_err 5.56%, avg_dev_err 4.68%
epoch  4, updates 42, features 4063, train_err 1.39%, dev_err 6.14%, avg_dev_err 4.68%
epoch  5, updates 50, features 4063, train_err 1.72%, dev_err 5.26%, avg_dev_err 4.97%
epoch  6, updates 43, features 4063, train_err 1.50%, dev_err 4.68%, avg_dev_err 4.68%
epoch  7, updates 37, features 4063, train_err 1.17%, dev_err 6.14%, avg_dev_err 4.68%
epoch  8, updates 35, features 4063, train_err 1.12%, dev_err 6.14%, avg_dev_err 4.68%
epoch  9, updates 36, features 4063, train_err 1.23%, dev_err 4.97%, avg_dev_err 4.68%
epoch 10, updates 33, features 4063, train_err 1.09%, dev_err 4.97%, avg_dev_err 4.97%
best error rate average 4.68%
avg perceptron time 23.99 sec
```