Report

Writeup:

1. Describe how you handled unknown words in hmm1.
   Answer: the words with count less than 3 in the train set was changed to **UNK** with its tag also it was checked it the word matches on of pattern. If yes, we add it to the counter of the pattern. In test phase if the word is unknown, we searched if there is a pattern which matches the word if not we took tags from **UNK** words

2. Describe your pruning strategy in the Viterbi hmm.
   Answer: Each sentence we start only with two tags: "START", "START", after each word we update the first tag by set it to the second tag options and the second tag set we update with the best tag of the current word for Greedy but for Viterbi we take only tags where this word was associated with. We also choosing the best tag from the set if we have the word in dictionary or from pattern set or if there is no pattern we choose all tags

3. Report your test scores when running each tagger (hmm-greedy, hmm-viterbi, maxent-greedy, memm-viterbi) on each dataset. For the NER dataset, report token accuracy accuracy, as well as span precision, recall and F1.

Accuracy on POS:

HMM greedy 0.9405918292648591

HMM viterbi 0.9592589464200242

MEMM greedy 0.9584424360165554

MEMM viterbi 0.9642706309654532

Accuracy on NER:

HMM Greedy

```
Accuracy: 0.9449183760518051

All-types   Prec:     0.7919586868314276 Rec:        0.722652305621003
      ORG   Prec:     0.6345875542691751 Rec:     0.6539895600298284   F-Score:     0.6441424899008446
      LOC   Prec:      0.849197247706422 Rec:     0.8062057702776265   F-Score:     0.8271432560737224
     MISC   Prec:     0.7687861271676301 Rec:     0.7212581344902386   F-Score:     0.7442641298265249
      PER   Prec:     0.8881900768693222 Rec:     0.6900108577633007   F-Score:     0.7766575007638252
```

HMM Viterbi

```
Accuracy: 0.9566481833339796

All-types   Prec:         0.8497678737233055 Rec:        0.7701110737125547
    MISC  Prec:         0.8416075650118203 Rec:        0.7722342733188721    F-Score:        0.8054298642533938
    LOC   Prec:         0.9057057057057057 Rec:        0.8209036472509527    F-Score:        0.8612221587664193
    ORG   Prec:         0.7156644394951744 Rec:        0.7188665175242357    F-Score:        0.7172619047619049
    PER   Prec:         0.9115913555992141 Rec:         0.755700325732899    F-Score:        0.8263579697239538
```

MEMM Greedy

```
Accuracy: 0.962425840474621

All-types   Prec:         0.868260144524736 Rec:        0.7886233591383373
    MISC  Prec:         0.867948717948718 Rec:        0.7342733188720173    F-Score:        0.7955346650998825
    ORG   Prec:         0.8175618073316283 Rec:         0.715137956748695    F-Score:        0.7629276054097057
    PER   Prec:         0.854006586169045 Rec:        0.8447339847991314    F-Score:        0.8493449781659389
    LOC   Prec:         0.9210850801479655 Rec:        0.8132825258573761    F-Score:        0.8638334778837815
```

MEMM Viterbi

```
Accuracy: 0.9728178680832913

All-types   Prec:         0.8776290630975143 Rec:        0.8497139010434197
    PER   Prec:         0.8914316125598722 Rec:        0.9093376764386536    F-Score:        0.9002956194571351
    LOC   Prec:         0.9300578034682081 Rec:        0.8758845944474687    F-Score:        0.9021586767591814
    MISC  Prec:         0.879372738238842 Rec:        0.7906724511930586    F-Score:        0.8326670474014849
    ORG   Prec:         0.7878326996197719 Rec:        0.7725577926920209    F-Score:        0.7801204819277109
```

4. Is there a difference in behavior between the hmm and maxent taggers? discuss.
   Answer: MEMM is normalized globally and HMM is normalized locally, MEMM using Logistic Regression model which slow down the performance.

5. Is there a difference in behavior between the datasets? discuss.
   Answer: Yes, there is a huge number of tags with name "O' and less other tags with location, organization etc, but in POS there is a lot of examples for each tag. In addition, the amount of tags in POS is much bigger than in NER.

6. What will you change in the hmm tagger to improve accuracy on the named entities data?
   Answer: We have got a good accuracy on hmm, but we can improve it with more examples of MISC,LOC,PER and ORG or probably add other tags which can be used as probability to NER.

7. What will you change in the memm tagger to improve accuracy on the named entities data, on top of what you already did?
   Answer: We change the word pattern and remove patterns which have all options, because it doesn't give us nothing

8. Why are span scores lower than accuracy scores?
   Answer: Because number of 'O" general occurrences is much bigger, and we get better accuracy because we right on them instead of span.