Omer Saham 2040347222
Oren Sultan 201557972

1.

Init : Set $\pi(0,*,*,*) = 1$

Algorithm :

We will run the Viterbi algorithm with fixed values as emission i.e for the function pi

For k=1…n

For $t \in S_{k-2}, u \in S_{k-1}, v \in S_k \; for \; k \in \{1 \dots n\}$

$$\pi(k,t,u,v) = \max_{w \in S_{k-3}, x \in V}(\pi(k-1,w,t,u) \cdot q(v|w,t,u) \cdot e(x|v)$$

$$bp(k,t,u,v) = arg\max_{w \in S_{k-3}, x \in V}(\pi(k-1,w,t,u) \cdot q(v|w,t,u) \cdot e(x|v)$$

Set $(y_{n-2}, y_{n-1}, y_n) = argmax_{t,u,v}(\pi(n,t,u,v) \cdot q(STOP|t,u,v))$

For $k = (n-3) \dots . 1, y_k = bp(k+3, y_{k+1}, y_{k+2}, y_{k+3})$

$For \; k = 1 \dots . n \; , x_k = argmax_{x \in V}\big(e(y_k|x)\big)$

$Set \; y_{n+1} = STOP$

Return $x_1 \dots \dots x_n, y_1, \dots . , y_n, y_{n+1}$

2.

b.

1. Done. In the code
2. **Known words error rate:** 0.066
   Unknown words error rate: 0.789
   Total Error rate: 0.074

c.

1. Done. In the code
2. Done. In the code
3. **Error rate of viterbi(regular):** 0.2

Extra explanation: we added in our code a constant of `EPSILON = 0.000000000001` which is configurable, we added it to emission | transition | viterbi results when they were zeros. We added it because it changed our error rate from 0.8 to 0.2.
Also we tried to add it without condition of if they were zero(anyway) which is the (“Dirichlet prior” technique) but for some reason it was still bad, so we remained with our solution.
* We talked about it with the lecturer.

d.

1. Done. In the code
2. **Error rate of viterbi(add one smoothing):** 0.67

e.

1. Done in the code

2. **Error rate of viterbi(psuedo words):** 0.19
3. **Error rate of psuedo + add 1 smoothing:** 0.64
   Top 10 most dominant tags which differ from test set and psuedo + add 1 smoothing technique output (Confusion Matrix):
   [(('NN', 'JJT-HL'), 70), ((',', ',-HL'), 58), (('NP', 'NNS-TL-HL'), 47), (('IN', 'IN HL'), 45), (('NNS', 'JJT-HL'), 39), (('AT', 'AT-HL'), 30), (('JJ', 'JJT-HL'), 25), (('CC', 'CC-HL'), 24), (('VBD', 'JJT-HL'), 22), (('IN', 'IN-TL'), 21)]