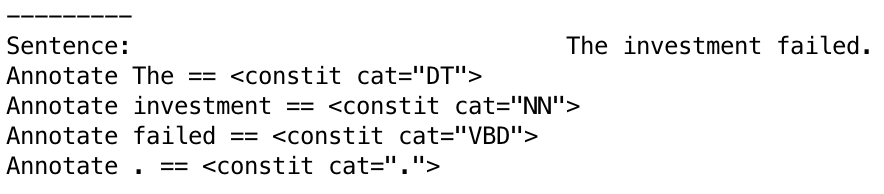
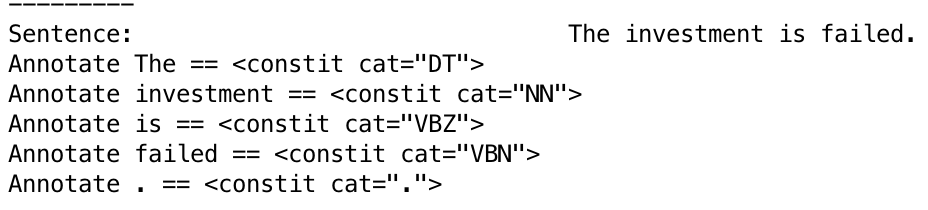
Q2. JET HMM Tagger:

a) The sentence “The investment failed” is correctly tagged.



b) In sentence “The investment is failed”, every word is tagged correctly, except “failed”. It is supposed to be in the category ‘JJ’.



**The investment is failed.**

**DT NN VBZ VBN .**

STATE start  47396  
ARC TO DT  9748  
STATE DT 98334

ARC TO NN 46646

STATE NN 159394

ARC TO VBZ 6967

STATE VBZ 25735

ARC TO VBN 3817

STATE VBN 24107

EMIT failed 53

ARC TO . 1202

STATE . 47396

ARC TO end 47396

State end

The computation of relative probability:

E(‘failed’ |VBN) \* P(DT|start)\* P(NN|DT) \* P(VBZ|NN) \* P(VBN|VBZ) \*P(.|VBN)=

(53|24107) \* (9748|47396) \* (46646|98334) \* (6967|159394) \* (3817|25735) \* (1202|24107)

**The investment is failed.**

**DT NN VBZ JJ .**

STATE start  47396  
ARC TO DT  9748  
STATE DT 98334

ARC TO NN 46646

STATE NN 159394

ARC TO VBZ 6967

STATE VBZ 25735

ARC TO JJ 1903

STATE JJ 73230

EMIT failed 6

ARC TO . 1780

STATE . 47396

ARC TO end 47396

State end

The computation of relative probability:

E(‘failed’ |JJ) \*P(DT|start)\* P(NN|DT) \* P(VBZ|NN) \* P(JJ|VBZ) \*P(.|JJ) =

(6|73230) \* (9748|47396) \* (46646|98334) \* (6967|159394) \* (1903|25735) \* (1780|73230)

From STATE start to STATE VBZ are common factor in two results, I just ignore their calculation. Thus, in short, I have:

E(‘failed’ |VBN) \* P(VBN|VBZ) \*P(.|VBN)

= (53|24107) \* (3817|25735) \* (1202|24107)

is greater than

E(‘failed’ |JJ) \* P(JJ|VBZ) \*P(.|JJ)

= (6|73230) \* (1903|25735) \* (1780|73230)

The relative probability of the right tagger(‘JJ’) is much lower than the wrong tagger(‘VBN’).

This leads to the wrong tagged result.