POS – Viterbi – Likelihood sampling

Sentence: Janet will back the bill

|  |  |  |
| --- | --- | --- |
| Tag | Description | Example |
| DT | Determiner | a, the |
| RB | Adverb | Quickly |
| NN | Sing or mass noun | Llama |
| JJ | Adjective | Yellow |
| VB | Verb base | eat |
| MD | Modal | Can, should |
| NNP | Proper noun, sing | IBM |

Step 1:

V1(1) = P(NNP|start) \* P(Janet|NNP) = 0.28\*0.000032 = 0.000009

V1(2) = p(MD|start) \* P(Janet|MD) = 0.0006\*0 = 0

V1(3) = p(VB|start) \* P(Janet|VB) = 0.0031\*0 = 0

V1(4) = p(JJ|start) \* P(Janet|JJ) = 0.045 \* 0 = 0

V1(5) to V1(7) = 0

Since only V1(1) output with a value that's greater of 0, we will prune V1(2) to V1(7)

So “Janet” is most likely to be a “proper noun”.

Step 2:

V2(1) = 0.000009\*P(NNP|NPP)\*P(will|NNP) = 0.000009\* 0.3777\*0 = 0

V2(2) = 0.000009 \*P(MD|NNP)\* P(will|MD) = 0.000009\*0.011\*0.308431 = 3.053e-8

V2(3) = 0.000009\*P(VB|NNP)\*P(will|VB) = 0.000009\*0.0009\*0.000028 = 2.268e-13

V2(4) = 0.000009\*P(JJ | NNP)\*P(will |JJ) = 0.000009\*0.0084\*0 = 0

V2(5) = 0.000009\*P(NN|NNP)\*P(will|NN) = 0.000009\*0.0584\*0.0002 = 1.051e-10

V2(6) = 0.000009 \*P(RB|NNP)\*P(will|RB) = 0.000009\*0.009\*0 = 0

V2(7) = 0.000009\* P(DT|NNP)\*P(will|DT) = 0.000009\*0.0025\*0= 0

V2(2) gives the greatest output, which is 3.053e-8

So “will” is most likely to be a “modal”.

Current path: NNP, MD

Step 3:

V3(1) = 3.053e-8\*P(NNP|MD)\*P(back|NNP) = 3.053e-8\*0.0008\*0 = 0

V3(2) = 3.053e-8\*P(MD|MD)\*P(back|MD) = 3.053e-8\*0.0002\*0 = 0

V3(3) =

“NNP, MD, VB”:

3.053e-8\*P(VB|MD)\*P(back|VB) = 3.053e-8\*0.7968\*0.000672 = 1.635e-11

“NNP, VB, VB”:

2.268e-13\*P(VB|VB)\*P(back|VB) = 2.268e-13\*0.005\*0.000672 = 7.62e-19

“NNP, NN, VB”:

1.051e-10\*P(VB|NN)\*P(back|VB) = 1.051e-10\*0.0014\*0.000672 = 9.888e-17

In which the path with NNP, MD, VB achieves the best result with 1.635e-11.

Current path: NNP, MD, VB

Therefore, the result is:

V3(3) = 3.053e-8\*P(VB|MD)\*P(back|VB) = 3.053e-8\*0.7968\*0.000672 = 1.635e-11