To run the code

Python my_viterbi.py probs.txt sents.txt

Where, probs.txt- probability file, sents.txt- sentence file

Analysis of Results

- Viterbi algorithm calculates POS tags based on transition and emission probability
- Model considered here is bigram. The accuracy obtained on tried sentences is still very good. The reason being the probabilities given are good and works well for the tried test cases. However, the accuracy might fall for new sentences.

Results

Following are the obtained results in required format-

PROCESSING SENTENCE:['mark', 'has', 'fish']

FINAL VITERBI NETWORK
P(mark=noun) = 0.072
P(mark=verb) = 0.006
P(mark=inf) = 1e-08
P(mark=prep) = 1e-08
P(has=noun) = 4.62e-07
P(has=verb) = 0.001404
P(has=inf) = 1.32e-07
P(has=prep) = 2.16e-06
P(fish=noun) = 8.64864e-05
P(fish=verb) = 2.1021e-08
P(fish=inf) = 3.0888e-08
P(fish=prep) = 3.51e-08

FINAL BACKPTR NETWORK
Backptr(has=noun) = verb
Backptr(has=verb) = noun
Backptr(has=inf) = verb
Backptr(has=prep) = verb

Backptr(fish=noun) = verb
Backptr(fish=verb) = noun
Backptr(fish=inf) = verb
Backptr(fish=prep) = verb

BEST TAG SEQUENCE HAS PROBABILITY =4.32432e-05
fish->noun
has->verb
mark->noun

FORWARD ALGORITHM RESULTS
P(mark=noun) = 0.072
P(mark=verb) = 0.006
P(mark=inf) = 1e-08
P(mark=prep) = 1e-08
P(has=noun) = 4.627208501e-07
P(has=verb) = 0.00140401822503
P(has=inf) = 1.327200002e-07
P(has=prep) = 2.3100000002e-06
P(fish=noun) = 8.6633284752e-05
P(fish=verb) = 3.779412e-08
P(fish=inf) = 3.088802754e-08
P(fish=prep) = 3.511388292e-08

PROCESSING SENTENCE:['mark', 'bears', 'fish']

FINAL VITERBI NETWORK
P(mark=noun) = 0.072
P(mark=verb) = 0.006
P(mark=inf) = 1e-08
P(mark=prep) = 1e-08
P(bears=noun) = 9.24e-05
P(bears=verb) = 0.000936
P(bears=inf) = 1.32e-07
P(bears=prep) = 2.16e-06
P(fish=noun) = 5.76576e-05
P(fish=verb) = 4.2042e-06
P(fish=inf) = 2.0592e-08
P(fish=prep) = 2.34e-08

FINAL BACKPTR NETWORK
Backptr(bears=noun) = verb
Backptr(bears=verb) = noun
Backptr(bears=inf) = verb
Backptr(bears=prep) = verb

Backptr(fish=noun) = verb
Backptr(fish=verb) = noun
Backptr(fish=inf) = verb
Backptr(fish=prep) = verb

BEST TAG SEQUENCE HAS PROBABILITY =2.88288e-05
fish->noun
bears->verb
mark->noun

FORWARD ALGORITHM RESULTS
P(mark=noun) = 0.072
P(mark=verb) = 0.006
P(mark=inf) = 1e-08
P(mark=prep) = 1e-08
P(bears=noun) = 9.254417002e-05
P(bears=verb) = 0.00093601215002
P(bears=inf) = 1.327200002e-07
P(bears=prep) = 2.3100000002e-06
P(fish=noun) = 5.7805220256e-05
P(fish=verb) = 4.21769712e-06
P(fish=inf) = 2.059294692e-08
P(fish=prep) = 2.617202292e-08

PROCESSING SENTENCE:['mark', 'likes', 'to', 'fish', 'for', 'fish']

FINAL VITERBI NETWORK
P(mark=noun) = 0.072
P(mark=verb) = 0.006
P(mark=inf) = 1e-08
P(mark=prep) = 1e-08
P(likes=noun) = 4.62e-07
P(likes=verb) = 4.68e-06
P(likes=inf) = 1.32e-07
P(likes=prep) = 2.16e-06
P(to=noun) = 3.6036e-10
P(to=verb) = 3.003e-11
P(to=inf) = 1.019304e-06
P(to=prep) = 3.861e-07
P(fish=noun) = 2.62548e-08
P(fish=verb) = 5.351346e-08
P(fish=inf) = 1.019304e-14
P(fish=prep) = 1.08108e-14
P(for=noun) = 4.12053642e-12
P(for=verb) = 1.706562e-12

P(for=inf) = 1.17729612e-12
P(for=prep) = 3.07702395e-09
P(fish=noun) = 2.092376286e-10
P(fish=verb) = 1.8748440711e-13
P(fish=inf) = 3.7544364e-17
P(fish=prep) = 1.236160926e-16

FINAL BACKPTR NETWORK
Backptr(likes=noun) = prep
Backptr(likes=verb) = noun
Backptr(likes=inf) = verb
Backptr(likes=prep) = noun
Backptr(to=noun) = prep
Backptr(to=verb) = noun
Backptr(to=inf) = verb
Backptr(to=prep) = noun
Backptr(fish=noun) = prep
Backptr(fish=verb) = noun
Backptr(fish=inf) = verb
Backptr(fish=prep) = noun
Backptr(for=noun) = prep
Backptr(for=verb) = noun
Backptr(for=inf) = verb
Backptr(for=prep) = noun
Backptr(fish=noun) = prep
Backptr(fish=verb) = noun
Backptr(fish=inf) = verb
Backptr(fish=prep) = noun

BEST TAG SEQUENCE HAS PROBABILITY =1.046188143e-10
fish->noun
for->prep
fish->verb
to->inf
likes->verb
mark->noun

FORWARD ALGORITHM RESULTS
P(mark=noun) = 0.072
P(mark=verb) = 0.006
P(mark=inf) = 1e-08
P(mark=prep) = 1e-08
P(likes=noun) = 4.627208501e-07
P(likes=verb) = 4.6800607501e-06
P(likes=inf) = 1.327200002e-07

P(likes=prep) = 2.3100000002e-06
P(to=noun) = 5.4396594e-10
P(to=verb) = 3.99984e-11
P(to=inf) = 1.019576646e-06
P(to=prep) = 4.31913636e-07
P(fish=noun) = 2.62648071629e-08
P(fish=verb) = 5.35325592902e-08
P(fish=inf) = 1.47183036e-14
P(fish=prep) = 2.561559e-14
P(for=noun) = 4.12079988702e-12
P(for=verb) = 1.70709789919e-12
P(for=inf) = 1.17755866821e-12
P(for=prep) = 4.88860515048e-09
P(fish=noun) = 2.09342795202e-10
P(fish=verb) = 2.70843566994e-13
P(fish=inf) = 6.83675818254e-17
P(fish=prep) = 1.97062155061e-16
