IMPLEMENT VITERBI ALOGRITHM FOR DNA SEQUENCE

The program takes initial probabilities, emission probabilities, transitional probabilities and sequence as input

Below is how the user has to enter the input:

Step -1: Enter the number of the states in the sequence and their names

```
E:\Fall_2015\machine_learning\HW_3>homework.py
Enter the number of states in the Hidden Markov Model   : 2

Enter the name of the hidden states

State1 : S1

State2 : S2
```

Step-2 : Enter the number of features and their names

```
Enter the number of features : 4

Enter the name of the features

Feature1 : A

Feature2 : C

Feature3 : G

Feature4 : T
```

Step -3 : Initial probabilities are given as input from state s1 to sn

```
Enter the initial probabilities for all the states

Enter the value : 0.5

Enter the value : 0.5
```

Step -4: Transition probabilities are entered as below sequence

P(S1|S1) , P(S2|S1)….P(S2|S2)

```
Enter the transition probabilities as s1s1,s1s2...snsn-1,snsn
Enter the value : 0.8
Enter the value : 0.2
Enter the value : 0.2
Enter the value : 0.8
```

Step-5: Emission probabilities are entered as below sequence

P(A|S1) , P(C|S1)….P(G|S2),P(T|S2)

```
Enter the emission probabilities as p(A|s1),p(C|s1),p(G|s1)...p(C|s2),p(G|s2),p(
T|s2)
Enter the value : 0.3
Enter the value : 0.2
Enter the value : 0.3
Enter the value : 0.2
Enter the value : 0.1
Enter the value : 0.4
Enter the value : 0.1
Enter the value : 0.4
```

Step-6: Enter the sequence

```
Enter the sequence : CGTCAG
```

Step -7 : The program print Viterbi matrix and most likely hidden state sequence

```
Viterbi matrix :

PS1(C,1) : 0.100000
PS2(C,1) : 0.200000
PS1(G,2) : 0.024000
PS2(G,2) : 0.016000
PS1(T,3) : 0.003840
PS2(T,3) : 0.005120
PS1(C,4) : 0.000614
PS2(C,4) : 0.001638
PS1(A,5) : 0.000147
PS2(A,5) : 0.000131
PS1(G,6) : 0.000035
PS2(G,6) : 0.000010
Most likely hidden states at each step are :
Step - 1 : S1
Step - 2 : S1
Step - 3 : S1
Step - 4 : S1
Step - 5 : S1
Step - 6 : S1
```