

Pytanie 1.

Wraz z zwiększającą się liczbą stanów ukrytych (użyłem następujących ilości stanów ukrytych: 5, 10, 15, 25, 35 oraz 45) oraz zwiększając ilość cykli treningowych (użyłem następującej ilości cykli treningowych: 50, 100, 150, 200), można wysnuć następujące wnioski:

- zwiększanie ilości stanów ukrytych z 5 na 10, 15, 25 oraz 35 powoduje zwiększenie logarytmów prawdopodobieństw wszystkich testowanych sekwencji, dla HMM trenowanego z sekwencją angielską logarytm prawdopodobieństwa angielskiej sekwencji testowych osiąga swoje maksimum przy 35 jednostkach ukrytych, jest ono równe -4586.91, analogicznie z sekwencją francuską, której osiąga maksimum przy 45 stanach ukrytych równe -4242.47,
- dla 45 stanów ukrytych logarytm prawdopodobieństwa sekwencji angielskiej oraz francuskiej maleje (przy trenowaniu HMM z sekwencją angielskich liter logarytm prawdopodobieństwa testowanej sekwencji angielskiej, analogicznie z sekwencją francuską), natomiast rosną logarytmy prawdopodobieństwa dla testowanych sekwencji innych niż zastosowanych do treningu – trenując HMM z sekwencją angielskich liter logarytmy prawdopodobieństwa testowanej sekwencji francuskiej oraz polskiej rosną,
- zwiększanie ilości cykli treningowych z 50 na 100, 150 oraz 200 powoduje zmniejszenie logarytmu prawdopodobieństwa, największe logarytmy prawdopodobieństw wszystkich sekwencji osiągnięte zostały dla ilości cykli treningowych równej 50,
- większe logarytmy prawdopodobieństwa dla polskiej sekwencji testowej zawsze otrzymujemy używając HMM wytrenowane z sekwencją angielskich liter niż wytrenowane z sekwencją francuskich liter,
- zwiększenie ilości cykli treningowych powoduje zmniejszanie się logarytmów prawdopodobieństwa dla polskiej sekwencji testowej używając HMM wytrenowane z sekwencją angielskich jak i francuskich liter,
- największy logarytm prawdopodobieństwa dla polskiej sekwencji testowej otrzymujemy dla HMM wytrenowane z sekwencją angielskich liter z ilością ukrytych stanów równą 45, przy ilości cykli treningowych równej 50, logarytm ten jest równy -6733.89

Przykładowa tablica z logarytmami prawdopodobieństw testowanych sekwencji angielskiej, francuskiej oraz polskiej przez obydwie HMM, z ilością ukrytych stanów równą 45 oraz ilością cykli treningowych równą 50

	ukrytych stanow=45		ilosc cykli treningowych=50	
	english	french	english	polish
english	-4841.971507	-5449.745273	-6733.890084	
french	-5452.546263	-4616.144401	-7887.969751	