林建甫 r07944013

Environment

CSIE workstation(linux)

How to execute

- training
  ./train [number of iterations] [initial HMM model][training sequences file] [output model name]

- testing
  ./test [modellist] [testing data] [result file name]

Result

Accuracy: 0.8704
Iterations: 900
Predicted by Viterbi Algorithm

Analysis

Number of iterations

Iterations	**500**	**700**	**900**	**1100**	**1300**	**1500**	**1700**
Accuracy	0.8560	0.8656	**0.8704**	**0.8704**	0.8700	0.8700	0.8700
Iterations	**1900**	**2100**	**2300**	**2500**	**2700**	**2900**	**3100**
Accuracy	0.8696	0.8688	0.8680	0.8680	0.8676	0.8676	0.8676

$P(O|\lambda)$ and $P(O|\lambda, q)$

By directly using $P(O|\lambda)$ for prediction rather than $P(O|\lambda, q)$ (Viterbi), we can get slightly better accuracy which is **0.8812** (900 iterations).

Misprediction among models

Model Name	#Errors	#samples in testing_data1	Error Rate
Model1	0	498	0.0000
Model2	4	531	0.0075
Model3	1	486	0.0020
Model4	141	507	0.2781
Model5	178	478	0.3724

Using the model set trained 900 iterations.

We can find out that majority of errors come from misprediction of model4&5.