

PGM Assignment 3

Ajita Shree
2015MCS2328

November 15, 2016

Model Description

Table 1:

1	OCR
2	OCR + Transition
3	OCR + Transition + Skip Factors
4	OCR + Transition + Skip Factors + Pair-Skip Factors

Gibbs Sampling

Table 2: TestCase 1 : data-tree.dat

		Gibbs Sampling		
Model	CharAcc	wordAcc	LL	time
1	58.445040	8.333333	-6.986933	0.360586s
2	67.828418	16.666667	-6.749648	0.362549s
3	67.828418	16.666667	-6.749648	0.424486s
4	68.096515	17.857143	-6.603487	0.421195s
		Loopy BP		
1	60.053619	4.761905	-6.988995	0.000310s
2	67.560322	15.476190	-6.746949	0.000259s
3	67.560322	15.476190	-6.746949	0.000306s
4	67.560322	16.666667	-6.606466	0.000245s

Table 3: TestCase 2 : data-loops.dat

		Gibbs Sampling		
Model	CharAcc	wordAcc	LL	time
1	51.079137	3.571429	-7.977534	0.138204s
2	56.834532	3.571429	-7.804875	0.147725s
3	56.834532	3.571429	-7.804875	0.152705s
4	57.553957	10.714286	-7.485028	0.144800s
		Loopy BP		
1	53.956835	3.571429	-7.996443	0.000073s
2	56.115108	7.142857	-7.817972	0.000090s
3	56.115108	7.142857	-7.817972	0.000090s
4	56.834532	7.142857	-7.546444	0.000095s

Table 4: TestCase 3 : data-treeWS.dat

		Gibbs Sampling		
Model	CharAcc	wordAcc	LL	time
1	57.505519	8.152174	-7.778342	0.893264s
2	64.128035	12.500000	-7.509077	1.008133s
3	64.017660	14.130435	-7.342652	1.015450s
4	63.686534	15.217391	-7.223134	0.919674s
		Loopy BP		
1	57.836645	8.152174	-7.778992	0.000536s
2	65.894040	15.217391	-7.517586	0.000587s
3	66.114790	15.217391	-7.344032	0.000662s
4	66.556291	15.760870	-7.213167	0.000646s

Table 5: TestCase 4 : data-loopsWS.dat

		Gibbs Sampling		
Model	CharAcc	wordAcc	LL	time
1	55.709877	9.230769	-7.846686	0.662152s
2	66.049383	11.538462	-7.524869	0.005747s
3	66.203704	11.538462	-7.331848	0.058373s
4	67.592593	13.846154	-6.919762	0.127113s
		Loopy BP		
1	60.030864	8.461538	-7.816212	0.000319s
2	64.197531	11.538462	-7.548425	0.744294s
3	66.049383	13.076923	-7.346568	0.735013s
4	66.512346	15.384615	-6.944808	0.651548s

Parameter Learning in Probabilistic Graphical Models

Table 6: Accuracies

	Bayesian Network	
DataSet	AvgAcc	AvgLL
Andes	77.900004	-70.377605
Hepan 2	78.552177	-16.185782
Insurance	82.375771	-5.811494
	Markov Network	
DataSet	AvgAcc	AvgLL
Andes	78.774385	-48.813523
Hepar2	80.056651	-16.521046
Insurance	80.077212	-7.937592

From the above table, it is evident that the accuracies corresponding to Markov network model has increased corresponding every file.

Table 7: Andes File

C-Value (x-axis)	Accuracy	Log-likelihood(y-axis)
1	78.774385	-48.813523
10	78.859131	-48.815645
100	78.367441	-49.832606
1000	79.557311	-48.931437
10000	76.548906	-49.642471
100000	77.033707	-53.692681

Table 8: Hepar2 File

C-Value (x-axis)	Accuracy	Log-likelihood(y-axis)
1	80.056651	-16.521046
10	80.056651	-16.520643
100	79.431651	-16.548352
1000	80.670204	-16.502720
10000	78.891463	-17.471397
100000	76.802064	-18.415796

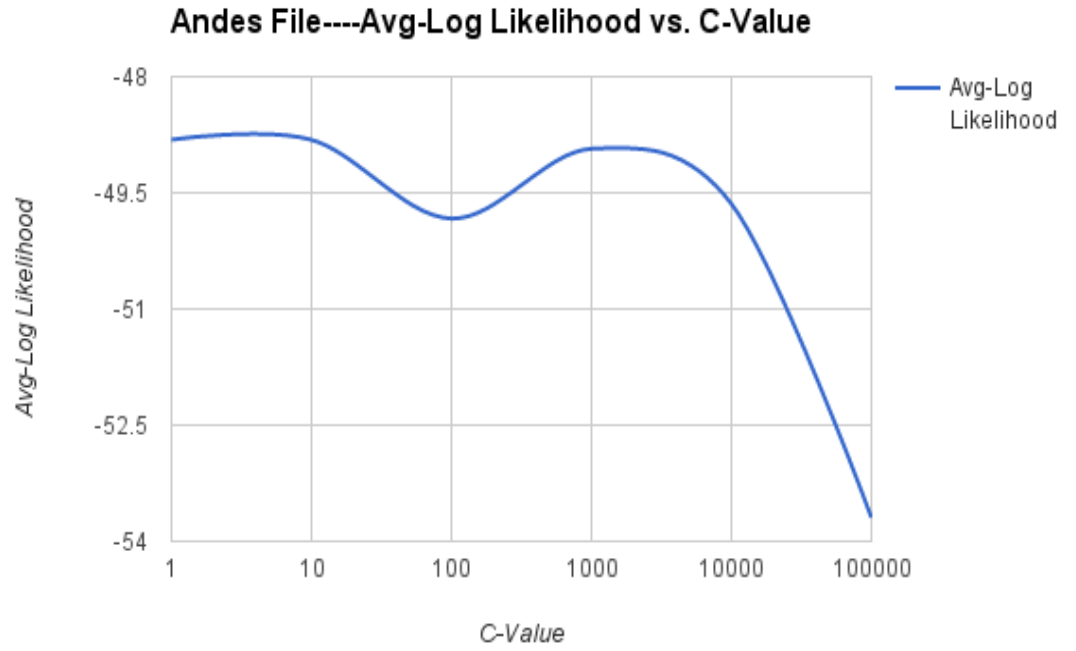


Figure 1: Andes File

Table 9: Insurance File

C-Value (x-axis)	Accuracy	Log-likelihood(y-axis)
1	80.077212	-7.937592
10	81.489806	-7.561210
100	80.077212	-7.926377
1000	81.400521	-7.653545
10000	81.659413	-8.202600
100000	76.455417	-10.674843

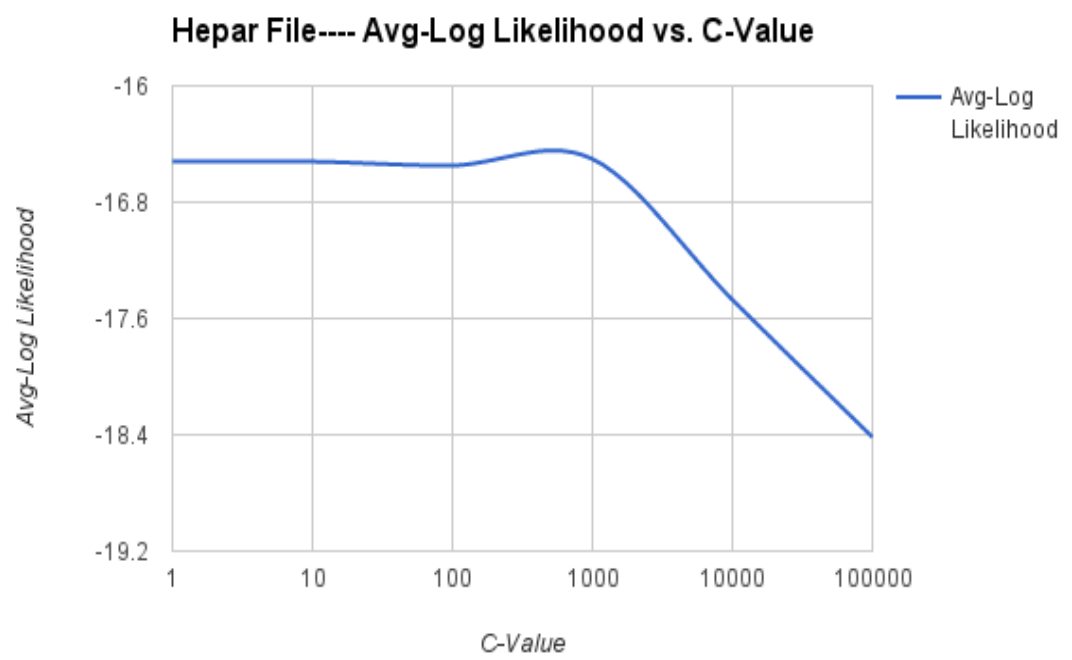


Figure 2: Hepar2 File

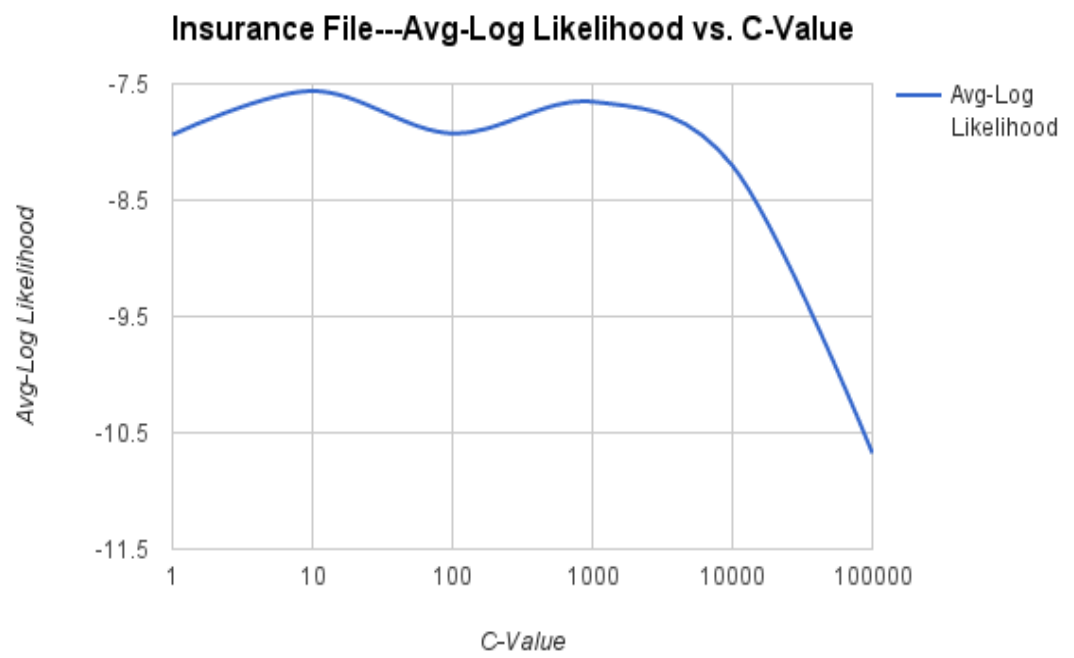


Figure 3: Insurance File