COL776 Assignment1 Report

Shubham Rawat
2013CS10258

September 2016

1   OCR using Graphical Models

(a) Following accuracies were obtained for each of the datasets, for all three models i.e OCR mode, transition mode, combined model.

Table 1: Character accuracies

	OCR	Transition	Combined
small.dat	53.92%	66.27%	71.17%
allimages1.dat	58.39%	68.04%	70.84%
allimages2.dat	57.26%	67.69%	70.72%
allimages3.dat	57.25%	67.87%	70.63%
allimages4.dat	57.58%	68.25%	70.77%
allimages5.dat	58.53%	68.46%	71.07%

Table 2: Word accuracies

	OCR	Transition	Combined
small.dat	8.65%	25.96%	35.58%
allimages1.dat	11.19%	24.04%	31.49%
allimages2.dat	10.01%	24.18%	31.81%
allimages3.dat	9.92%	24.68%	31.95%
allimages4.dat	11.47%	24.68%	31.85%
allimages5.dat	11.56%	26.69%	33.32%

(b) Following log likelihoods were obtained for each of the datasets, for all three models.

Table 3: Average Log Likelihood (base 2)

	OCR	Transition	Combined
small.dat	-11.265	-10.238	-9.059
allimages1.dat	-11.363	-10.352	-9.048
allimages2.dat	-11.359	-10.350	-9.047
allimages3.dat	-11.347	-10.339	-9.038
allimages4.dat	-11.354	-10.345	-9.043
allimages5.dat	-11.336	-10.327	-9.028