

## **Grading Criteria for Assignment 1**

### **Correctness: 75 points**

#### **Part A: 30 points**

Question	1	2	3	4	5
Points	5	10	5	5	5

#### **Part B: 45 points**

Question	1	2	3	4	5	6
Points	0	10	5	5	20	5

### **Grading criteria for correctness**

For each output file that reports probabilities, we will calculate the percent difference between your solutions and the correct values. For all other files, we will calculate the percent of tags or tokens that match the solutions. The percent correctness of a question maps to a percentage of points earned for that question as follows:

% Similarity	Question Points %
$\geq 95$	100
85-95	90
65-85	80
35-65	50
$\leq 35$	30

### **Design and Documentation: 25 points**

Running time	Part A	Part B
	5	5
Clarity of code	5	
Quality of report	10	

#### **Running time:**

Part A: 100% if running time  $\leq 5$  minutes, otherwise 0.

Part B: 100% if running time  $\leq 25$  minutes, otherwise 0.

#### **Clarity of code:**

Code will be examined for clarity of implementation. Sufficient comments will yield full credit.

#### **Report:**

All perplexity.py and pos.py values must be included. Analysis must demonstrate understanding of perplexity as a measure of a language model and POS% as a measure of a tagger.

#### **Late day policy**

10% off for each day. After three days, the assignment will be given a score of zero.