Part of Speech Tagging

General Requirements

CRITERIA	MEETS SPECIFICATIONS	STUDENT COMMENTS

Submission includes all files required for grading	• Includes `HMM Tagger.ipynb` displaying output for all executed cells
	
	• Includes `HMM Tagger.html`, which is an HTML copy of the notebook showing the output from executing all cells
		• The following files are present in the zip file:
		
		    • HMM_Tagger.ipynb
		
		    • HMM_Tagger.html
		
		    • HMM_Tagger.pdf

Submitted files are complete and do not include any disallowed changes	Submitted notebook has made no changes to test case assertions	No changes are made to any of the test scenarios/test case assertions.

Baseline Tagger Implementation

CRITERIA	MEETS SPECIFICATIONS	STUDENT COMMENTS
Student correctly implements the `pair_counts()` function	Emission count test case assertions all pass. • The emission counts dictionary has 12 keys, one for each of the tags in the universal tag set • "time" is the most common word tagged as a NOUN	• pair_counts() is implemented in cell 186. the approach is the count the number of occurrences for each tag and word. • As tested by the assert statement, time is the most common word tagged as noun.
Correct baseline MFC tagger implementation	Baseline MFC tagger passes all test case assertions and produces the expected accuracy using the universal tagset. • >95.5% accuracy on the training sentences • 93% accuracy the test sentences	• MFCTagger class is implemented in cell 187 and the results are as expected >95.5% for training and >93% for testing.

Calculating Tag Counts

CRITERIA	MEETS SPECIFICATIONS	STUDENT COMMENTS
Correct `unigram_counts()` implementation	All unigram test case assertions pass	unigram_counts() is implemented in cell 192 and all the tests pass
Correct `bigram_counts()` implementation	All bigram test case assertions pass	bigram_counts() is implemented in cell 193 and all the tests pass
Correct `start_counts()` and `end_counts()` implementation	All start and end count test case assertions pass	start_counts() is implemented in cell 194 and end_counts is implemented in cell 195 and all the tests pass

Basic HMM Tagger Implementation

CRITERIA	MEETS SPECIFICATIONS	STUDENT COMMENTS
Correct HMM network construction	All model topology test case assertions pass	The model is implemented in cell 196 and the assertions pass.
Correct basic HMM tagger implementation	Basic HMM tagger passes all assertion test cases and produces the expected accuracy using the universal tagset. • >97% accuracy on the training sentences • >95.5% accuracy the test sentences	The result of the accuracy are presented in cell 197