

Assignment 3 result and readme

	Stemming	Position Weighting	Local Collocation Modeling	ACCURACY		
				Tank	Plant	Pers/place
1	unstemmed	#0-uniform	#1-bag-of-words	89.00	89.25	78.00
2	stemmed	#1-expndecay	#1-bag-of-words	90.00	89.50	68.75
3	unstemmed	#1-expndecay	#1-bag-of-words	87.75	89.75	76.50
4	unstemmed	#1-expndecay	#2-adjacent-separate-LR	86.50	90.25	76.75
5	unstemmed	#2-stepped	#1-bag-of-words	86.25	89.25	76.25
6	unstemmed	#3-yours	#2- adjacent-separate-LR	84.25	89.75	74.50
7	unstemmed	#2-stepped	#1-bag-of-words	95.50	98.25	91.25

All data files are included in the current directory of the project.

In hw3_uniform.py

Please change the doc_path = "" to match the different tests for "Tank", "Plant" and "Perplace". The output line with "*" is incorrect.

In hw3_expndecay.py

Please change the doc_path = "" to match the different tests for "Tank", "Plant" and "Perplace". And change every code sentence using the string ".x-" to ".X-" to match the different stemming methods. The output line with "*" is incorrect.

In hw3_stepped.py

Please change the doc_path = "" to match the different tests for "Tank", "Plant" and "Perplace". The output line with "*" is incorrect.

In hw3_yours.py

I use adjacent-separate-LR to mark the left or right word of the ".x-". And the weighting method is similar to "stepped" but I set the weight of words 4 away from the ".x-" to 2.

In hw3_Bayesian

I implement the simple Bayesian Classifier.