Homework 3

Homework 3
CISE 4930/6930 Fall 2016
Due: September 19, 2016

Part I -- IN YOUR GROUP

Submit one writeup per group, as a PDF.

Include the names of the people in your group and include the names of people whose HW2 you are reviewing.

1. **Read the HW2 writeup that the other group (to whom you were assigned) submitted. Rate the following items:**

Item	Strongly Disagree	Somewhat Disagree	Somewhat Agree	Strongly Agree
a. The writeup describes the data and why the group chose to use that data.				
b. This writeup gives enough detail that I could follow it to replicate what the group did (in terms of what words they eliminated; whether they stemmed words, etc.)				
c. Overall, the writeup convinces me that the analysis the group conducted was technically sound.				
d. The writeup was well organized and readable, communicating clearly to the reader.				
e. The syntax and grammar of the writeup were clear and sound.				

2. For HW2 question 1 part b, indicate whether you arrived at the same or a different answer for items *i-vii*. If you differed, comment on the differences.

3. For HW2 question 4, read the analyses and interpretations of the group and discuss whether they are technically sound. If the group did not provide enough information or if you believe their interpretation was not technically sound, explain why. If the group described clearly and you believe their test and interpretation was technically sound, be specific about why you are convinced that it is correct. Do this for parts a, b, and c.

BONUS: Optionally, for 10 extra points on this homework (for a total possible score of 110) - compute all of the items in HW2 questions 2 and 3 on the other group's data set.

Part II -- INDIVIDUALLY

Submit one writeup per person, This must be your own individual work. Submit this as a separate PDF. You use the natural language toolkit of your choice. Python NLTK, or Stanford CoreNL are widely used but you may choose what you wish, or you may write your own code from scratch.

1. **Consider selected speeches by Abraham Lincoln and Nelson Mandela, as specified below. Answer the following questions.**
   a. Train a bigram language model on Lincoln's first inaugural address plus the Gettysburg address. (Treat those as the training corpus, call it LB-Train.) Call this language model *LB* (Lincoln bigram). https://www.nps.gov/liho/learn/historyculture/speeches.htm
   b. Compute the perplexity of your language model LB on Lincoln's second inaugural address. (Treat this as the testing data.) Call this testing data LB-Test
   c. Train a bigram language model on the combined corpus of Nelson Mandela's address to the court in the 1960's, "I am Prepared to Die" http://db.nelsonmandela.org/speeches/pub_view.asp?pg=item&ItemID=NMS010&txtstr=Dates:%201960%20-%201970
   And his speech upon receiving the International Freedom Award in 2000, http://db.nelsonmandela.org/speeches/pub_view.asp?pg=item&ItemID=NMS1389&txtstr=Dates:%201980%20-%202000. Call this language model MB (Mandela bigram) and call the training data MB-train
   d. Compute the perplexity of MB on MB-test, which is the following address by Nelson Mandela to an African National Congress rally: http://db.nelsonmandela.org/speeches/pub_view.asp?pg=item&ItemID=NMS1106&txtstr=Dates:%201990%20-
   e. Compare the perplexities of LB on LB-train and MB on MB-train. Are they drastically different? Discuss why or why not.
   f. Compute the perplexity of MB on LB-train and the perplexity of LB on MB-train. (You are using a language model on a testing set that you know is not from the same speaker as your training set.) Discuss your results.

2. **Build a first-order Markov model on the following sequence of tags. You may do this by hand or by writing a program to do it. Do not use a Markov modeling package. Do not use code shared with you by any other classmate. This work must be entirely your own. The goal is for you to fully understand how first-order Markov models are computed from the ground up. You will be asked to do this on the exam.**
   a. What are the state space and transition matrix for the model?
   b. Draw a state diagram (nodes and directed edges) to accompany this model.
   c. What is the highest transition probability of all? Interpret this finding.
   d. If you eliminate edges with less than 0.01 probability, how do you interpret the graph now? Explain why you would or would not want to do this.

Sequence:
M,Q,A,S,S,S,S,S,S,S,Q,A,S,Q,A,S,ACK,Q,A,Q,Q,S,S,ACK,S,S,ACK,U,S,ACK,Q,A,S,S,S,S,S,Q,F,ACK,Q,Q,S,D,Q,U,Q,ACK,Q,A,S,M,Q,S,M,S,Q,S,A,U,ACK,S,S,Q,D,D,Q,S,S,Q,A,Q,M,S,ACK,D,S,S,SU,Q,S,S,S,S,S,Q,A,Q,A,S,Q,A,Q,Q,A,Q,ACK,S,S,U,ACK,D,F,Q,A,S,U

Key to the meaning of these tags (from a corpus of collaborative dialogues in computer science) (credit: Fernando Rodríguez et al., research)

Tag	Description	Examples
S	Statement of information or explanation	*We need to create a program for kids to learn math.*
U	Opinion or indication of uncertainty	*unsure how to add strings together*
D	Explicit instruction	*wait put the if back*
SU	Polite or indirect instruction	*maybe we can do if user choice = +*
ACK	Acknowledgement	*oh ok gotcha*
M	Meta-comment or reflection	*hmmm*
QYN	Yes/no question	*can the answer be negative?*
QWH	Wh- question	*how do I take in their input?*
AYN	Answer to yes/no question	*yea*
AWH	Answer to wh-question	*the program should be able to generate erroneous questions*
FP	Positive task feedback	*oh nice*
FNON	Non-positive task feedback	*thats weird*
OEX	Irrelevant to the task	*wow its sweet in this room*