Apurv Dube PhD22021

Credit statement:

I attempted Task 2, parts 1, 2 and 3. The code and results are attached.

Corpus information

Count of sentences in corpus: 2400

No of words in corpus: 46532

No of unique words in corpus : 5429

1. **Top 5 bigrams:**
   a. With no smoothing:
   {('agreed', 'to'): 1.0, ('spritzer', 'out'): 1.0, ('cubicle', 'pretty'): 1.0, ('unattainable', 'because'): 1.0, ('strangers', 'of'): 1.0}
   b. With laplace smoothing:
   {('<s>', 'i'): 0.2693142638232665, ('i', 'feel'): 0.110412147505423, ('feel', 'like'): 0.03508771929824561, ('i', 'am'): 0.03188720173535792, ('<s>', 'im'): 0.027199591367641426}
   c. Kneser-Ney smoothing:
   {('href', 'http'): 0.9735794861393684, ('don', 't'): 0.9718722811177164, ('didn', 't'): 0.963330563767104, ('sort', 'of'): 0.9616955415104728, ('supposed', 'to'): 0.9281158442671915}

2. **Details of implementing the emotion component:**
   This approach basically checks top few potential bigrams by evaluating them with the emotion_scores() function. Using the score against the particular label supplied during generation, the probabilities are modified. Then the sampling proceeds as regular.

   The adjusted probabilities are computed by multiplying the original probabilities of the top k candidates by a factor that incorporates their alignment with the specified emotion, as determined by the emotion_scores() function. This adjustment biases the selection process towards words that are more likely to convey the desired emotion. The probabilities are then normalized by dividing each adjusted probability by the sum of all adjusted probabilities. Varying k affects performance by determining how many of the top candidates are considered for emotional alignment. A larger k increases the computational cost due to more calls to emotion_scores(), but it may lead to sentences that align more closely with the target emotion by considering a wider range of options.

3. **Generated samples:**
   a. No emotion component, max length 10, no smoothing:
      i. i can t irritate your life
      ii. i remember feeling fabulous and numb

   b. No emotion component, max length 10, Laplace smoothing:
      i. greg was years callous reactions recognize optimism unstable flow

ii. woke private serious event needing pakistan eight racing expert

c. No emotion component, max length 10, Kneser-Ney smoothing:
i like a pleasant
i just a little too popular

d. With emotion component, max length 10, Laplace smoothing:
Label = joy
journalism cousin emotionally enjoy imagined hometown milk deeply gracious
lets fighter sh note settle concerned rave create handed

Label = love
i href vision breath uncoiling purchase process drastic outraged
performances joke animated peers depressed sinks present fighting loki

Label = sadness
judge things paul enter christian thankful curious spain horny
bitter attacking beatles planner corners yu wp uninteresting portrayals

Label = surprise
considerate ica fearful agitated wave invigorated clearing o position
thereby mammogram attitudes miracles brds hp rightly resigned director

Label = fear
i feel mixed rid joining woman makeup lash main
i unfit coworker entire gym runs appreciate benjamin sophist

Label = anger
shouldn dwelling isnt foggy ants animals prom creamy off
countless balks exhibit box slice grabbing reflect episode burst

We sampled top 20 words for generating these sentences, we could have used more to potentially improve the emotion alignment.

e. With emotion component, max length 10, Kneser-Ney smoothing:
Label = joy
i simply does not know its not sure how
i am feeling so people can t want to

Label = love
i feel like some of feeling and he offers
i like getting no identity

Label = sadness
i have ever written although thats why wont you
i evolve to feel irritated but the job

Label = surprise

i feel your musicianship especially the only a boy
i don t feeling very apprehensive putting the reasons

Label = fear
i really need to make you have had enough
im sure how to know how to in between

Label = anger
i should be interesting n n n n n
i d have for all of death spiral