Elias Rosenberg
CS72
Coto-Solano
May 3, 2021

Exercise 5 responses

Part I:

(4) Report the results of the classification in a Word/LibreOffice/PDF document. What was the accuracy, precision and recall of the program?
The results were similar for both companies in terms of their accuracy, precision, recall, and feature terms. Here are the exact stats for both.
Google:
accuracy: 0.885
pos precision: 0.9230769230769231
pos recall: 0.84
neg precision: 0.8532110091743119
neg recall: 0.93

Amazon:
accuracy: 0.89
pos precision: 0.90625
pos recall: 0.87
neg precision: 0.875
neg recall: 0.91

(5) Report the results of the "most informative features" in a Word/LibreOffice/PDF document. You need to describe what is positive and what is negative about working at Amazon. You also need to describe what is positive and what is negative about working at Google. Finally, in addition to your written document, please submit your Python code and a JPEG/PNG file with a screen capture showing the results from your program.

Amazon: The three most positive features for Amazon were “great, nice, smart.” Two of the concerning negative features were (‘long’, ‘hours’) and (‘life, ‘balance’), which suggests that workers at Amazon might not have a great work-life balance. Other positives of working with Amazon possibly include good benefits, and opportunities for promotion and learning based on other features.

Google: The three most positive features for Google were “great, perks, free.” From a glance, there were more negative features in the Google output than the Amazon one. Some standouts include (‘hard’, ‘to’), ‘politics’, and ‘difficult.’ Unlike Amazon, where it seems that the people are the best part of the company, it might be the opposite at Google if the “politics” of the working

environment put off some employees. Based on the positives, however, maybe the benefits and ‘perks’ are worth it.

Part II:

(5) Report your results on the same Word/LibreOffice/PDF document where you reported the results from the first part. What are the results for each of the labels? (Verbs, pronouns, etc.). Did the program make any mistakes when tagging the example sentences? What were the mistakes, and how do they relate to the precision/recall/F1 results? What recommendations would you give to improve the program? Make sure you include a screen capture of your program’s output.

Here are the results of the inputs in the instructions and their outputs.

Input1: “kua qaere au ki rarotonga”
kua - tam
qaere - v
au - pron
ki - tam
rarotonga - n

Input2: “kua aere au ki rarotonga”
kua - <UNK>
aere - pron
au - tam
ki - prep
rarotonga - n

Input3: “ka tunu tere i taro”
ka - prep
tunu - v
tere - n
i - prep
taro - n

Input4: “kia orana kotou”
kia - tam
orana - v
kotou - <UNK>

This program makes loads of mistakes, but professor Rolando says this is to be expected. Inputs 1 and 3 are commonly the most successful inputs, with 1 yielding 100% accuracy every time, and 3 having 100% accuracy most of the time (not this time, unfortunately). Both inputs 2 and 4 are consistently wrong on their first and last terms respectively, labeling both ‘kua and

‘kotou’ as unknown even though they’re in the training set. For the second example, this throws the whole sentence off, except for the end. I imagine this is because the program can’t figure out what is supposed to come on either side of the verb, which it doesn’t recognize because it’s misspelled. The inaccuracy of this program, however, is to be expected, as the accuracy, recall, and precision scores are all kind of bad.

	precision	recall	f1-score	support
0	0.59	0.70	0.64	123
1	0.60	0.65	0.62	71
2	0.43	0.29	0.35	31
3	0.66	0.38	0.48	50
4	0.38	0.39	0.39	66
accuracy			0.55	341

To improve the output, some stemming of the input sentences would likely be a huge help, as I am doubtful the program recognizes all the conjugations very effectively. Also, more training data would greatly improve the accuracy of the program. There’s just not that much training data in Maori as there is in english unfortunately.