Query Focused Abstractive Summarization Analysis on CNN-Daily Mail Dataset

Kailash Karthik S

March 2020

Rouge Scores

1 Batch 1

1.1 Generic Summary

```
-----------------------------------------------
1 ROUGE-1 Average_R: 0.28961 (95%-conf.int. 0.28755 - 0.29166)
1 ROUGE-1 Average_P: 0.49333 (95%-conf.int. 0.49029 - 0.49618)
1 ROUGE-1 Average_F: 0.35055 (95%-conf.int. 0.34853 - 0.35253)
-----------------------------------------------
1 ROUGE-2 Average_R: 0.12858 (95%-conf.int. 0.12684 - 0.13035)
1 ROUGE-2 Average_P: 0.22274 (95%-conf.int. 0.22003 - 0.22558)
1 ROUGE-2 Average_F: 0.15629 (95%-conf.int. 0.15441 - 0.15831)
-----------------------------------------------
1 ROUGE-L Average_R: 0.27058 (95%-conf.int. 0.26856 - 0.27256)
1 ROUGE-L Average_P: 0.46170 (95%-conf.int. 0.45870 - 0.46458)
1 ROUGE-L Average_F: 0.32773 (95%-conf.int. 0.32567 - 0.32975)
```

1.2 Query Focused Summary - Multiple Query Training

```
-----------------------------------------------
1 ROUGE-1 Average_R: 0.46633 (95%-conf.int. 0.46426 - 0.46853)
1 ROUGE-1 Average_P: 0.40088 (95%-conf.int. 0.39868 - 0.40312)
1 ROUGE-1 Average_F: 0.41887 (95%-conf.int. 0.41697 - 0.42069)
-----------------------------------------------
1 ROUGE-2 Average_R: 0.20640 (95%-conf.int. 0.20421 - 0.20846)
1 ROUGE-2 Average_P: 0.17731 (95%-conf.int. 0.17539 - 0.17928)
1 ROUGE-2 Average_F: 0.18509 (95%-conf.int. 0.18322 - 0.18688)
-----------------------------------------------
1 ROUGE-L Average_R: 0.43924 (95%-conf.int. 0.43712 - 0.44136)
1 ROUGE-L Average_P: 0.37789 (95%-conf.int. 0.37577 - 0.38009)
1 ROUGE-L Average_F: 0.39471 (95%-conf.int. 0.39284 - 0.39651)
```

2 Batch 2

2.1 Generic Summary

```
---------------------------------------------
1 ROUGE-1 Average_R: 0.35907 (95%-conf.int. 0.35539 - 0.36262)
1 ROUGE-1 Average_P: 0.25485 (95%-conf.int. 0.25220 - 0.25757)
1 ROUGE-1 Average_F: 0.27596 (95%-conf.int. 0.27330 - 0.27855)
---------------------------------------------
1 ROUGE-2 Average_R: 0.15762 (95%-conf.int. 0.15424 - 0.16064)
1 ROUGE-2 Average_P: 0.10882 (95%-conf.int. 0.10665 - 0.11094)
1 ROUGE-2 Average_F: 0.11843 (95%-conf.int. 0.11608 - 0.12063)
---------------------------------------------
1 ROUGE-L Average_R: 0.32891 (95%-conf.int. 0.32527 - 0.33218)
1 ROUGE-L Average_P: 0.23362 (95%-conf.int. 0.23104 - 0.23620)
1 ROUGE-L Average_F: 0.25281 (95%-conf.int. 0.25027 - 0.25519)
```

2.2 Query Focused Summary - Multiple Query Training

```
---------------------------------------------
1 ROUGE-1 Average_R: 0.53116 (95%-conf.int. 0.52747 - 0.53470)
1 ROUGE-1 Average_P: 0.19223 (95%-conf.int. 0.19023 - 0.19418)
1 ROUGE-1 Average_F: 0.26553 (95%-conf.int. 0.26341 - 0.26774)
---------------------------------------------
1 ROUGE-2 Average_R: 0.24041 (95%-conf.int. 0.23682 - 0.24407)
1 ROUGE-2 Average_P: 0.08304 (95%-conf.int. 0.08168 - 0.08440)
1 ROUGE-2 Average_F: 0.11554 (95%-conf.int. 0.11387 - 0.11721)
---------------------------------------------
1 ROUGE-L Average_R: 0.49443 (95%-conf.int. 0.49093 - 0.49796)
1 ROUGE-L Average_P: 0.17883 (95%-conf.int. 0.17693 - 0.18067)
1 ROUGE-L Average_F: 0.24700 (95%-conf.int. 0.24493 - 0.24911)
```

3 Batch 3

3.1 Generic Summary

```
---------------------------------------------
1 ROUGE-1 Average_R: 0.28961 (95%-conf.int. 0.28755 - 0.29166)
1 ROUGE-1 Average_P: 0.49333 (95%-conf.int. 0.49029 - 0.49618)
1 ROUGE-1 Average_F: 0.35055 (95%-conf.int. 0.34853 - 0.35253)
---------------------------------------------
1 ROUGE-2 Average_R: 0.12858 (95%-conf.int. 0.12684 - 0.13035)
1 ROUGE-2 Average_P: 0.22274 (95%-conf.int. 0.22003 - 0.22558)
1 ROUGE-2 Average_F: 0.15629 (95%-conf.int. 0.15441 - 0.15831)
---------------------------------------------
1 ROUGE-L Average_R: 0.27058 (95%-conf.int. 0.26856 - 0.27256)
1 ROUGE-L Average_P: 0.46170 (95%-conf.int. 0.45870 - 0.46458)
1 ROUGE-L Average_F: 0.32773 (95%-conf.int. 0.32567 - 0.32975)
```

3.2 Query Focused Summary - Multiple Query Training

```
---------------------------------------------
1 ROUGE-1 Average_R: 0.46685 (95%-conf.int. 0.46483 - 0.46900)
1 ROUGE-1 Average_P: 0.40065 (95%-conf.int. 0.39842 - 0.40288)
1 ROUGE-1 Average_F: 0.41911 (95%-conf.int. 0.41722 - 0.42095)
---------------------------------------------
1 ROUGE-2 Average_R: 0.20676 (95%-conf.int. 0.20462 - 0.20886)
1 ROUGE-2 Average_P: 0.17723 (95%-conf.int. 0.17536 - 0.17918)
1 ROUGE-2 Average_F: 0.18525 (95%-conf.int. 0.18340 - 0.18701)
---------------------------------------------
1 ROUGE-L Average_R: 0.43970 (95%-conf.int. 0.43757 - 0.44178)
1 ROUGE-L Average_P: 0.37759 (95%-conf.int. 0.37543 - 0.37977)
1 ROUGE-L Average_F: 0.39487 (95%-conf.int. 0.39305 - 0.39673)
```

Personal Note:

Batch 1
Source Dir : augmented_files_query
Target Dir : multiple_queries_again

Batch 2
Source Dir : augmented_files
Target Dir : query_1

Batch 3
Source Dir : augmented_files_test_query
Target Dir : query_2