



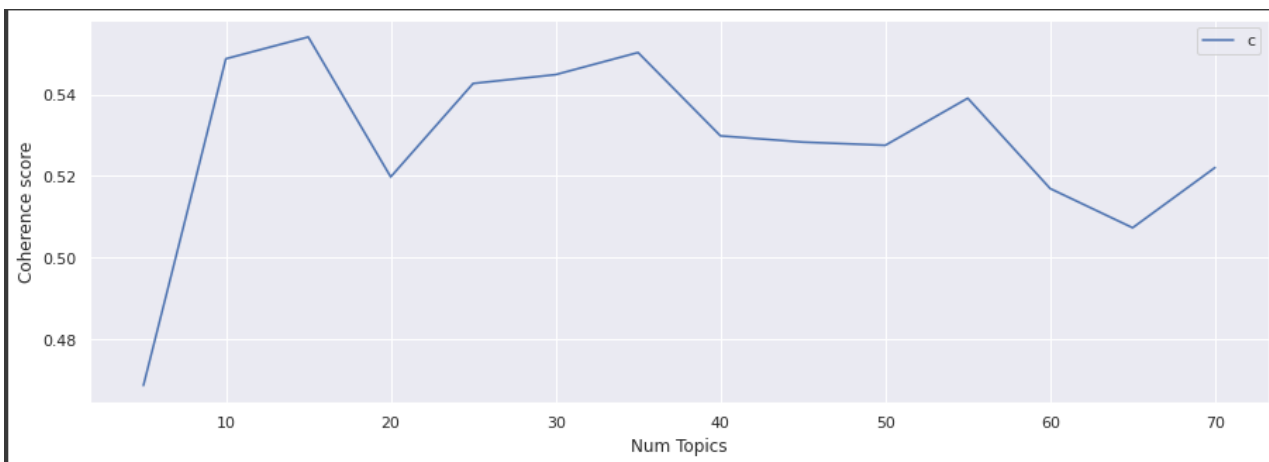
# PARAMETER TUNING FOR LDA

---



Iteration 1: - 34min

corpus=corpus, id2word=id2word, num\_topics= num\_topics,  
random\_state= 0, update\_every= 1, chunksize=1000,  
iterations= 50, passes=10, alpha='auto', per\_word\_topics=True



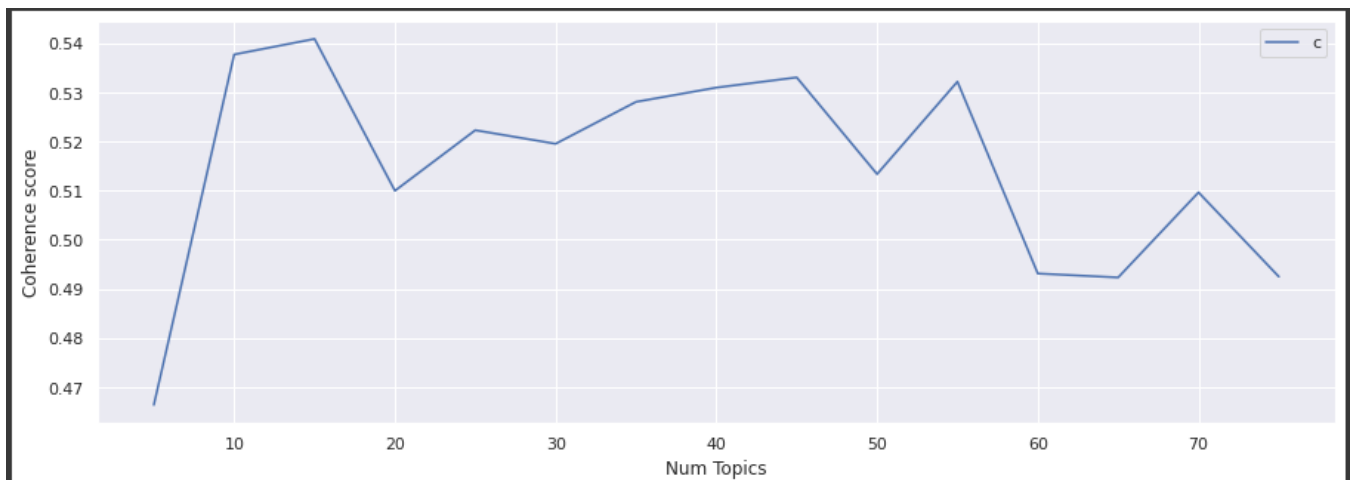
```
Num Topics = 5 has Coherence Value of 0.469
Num Topics = 10 has Coherence Value of 0.549
Num Topics = 15 has Coherence Value of 0.554
Num Topics = 20 has Coherence Value of 0.52
Num Topics = 25 has Coherence Value of 0.543
Num Topics = 30 has Coherence Value of 0.545
Num Topics = 35 has Coherence Value of 0.55
Num Topics = 40 has Coherence Value of 0.53
Num Topics = 45 has Coherence Value of 0.528
Num Topics = 50 has Coherence Value of 0.528
Num Topics = 55 has Coherence Value of 0.539
Num Topics = 60 has Coherence Value of 0.517
Num Topics = 65 has Coherence Value of 0.507
Num Topics = 70 has Coherence Value of 0.522
```

## PARAMETER TUNING FOR LDA

...

Iteration 2 – 38 min

`corpus=corpus, id2word=id2word, num_topics =  
num_topics, random_state=0, update_every=1, chunksize=1000,  
iterations=100, passes=10, alpha='auto', per_word_topics=True`



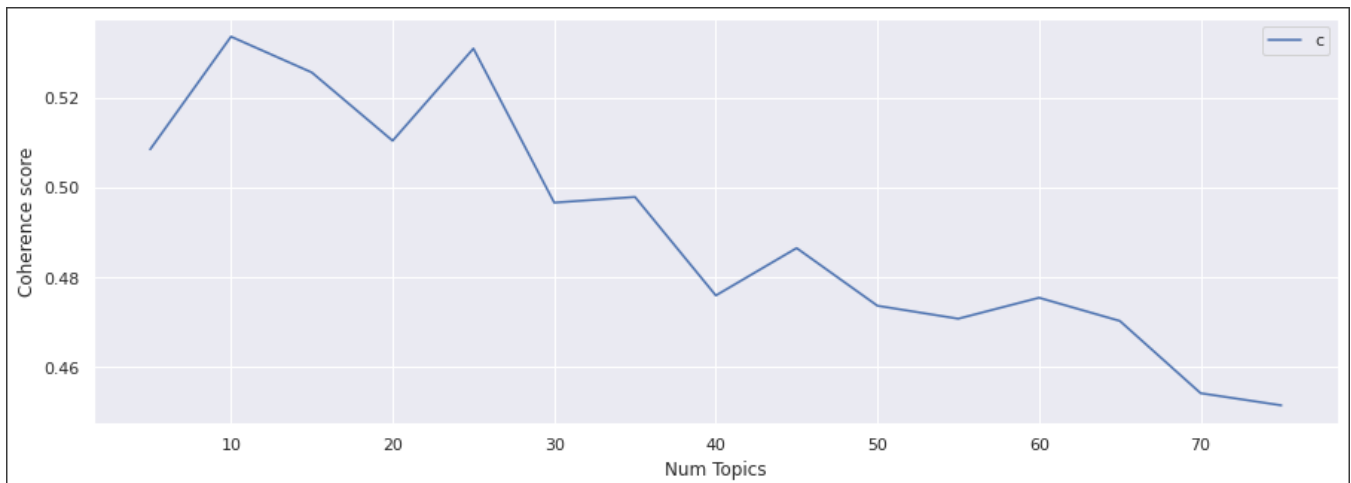
```
Num Topics = 5 has Coherence Value of 0.466  
Num Topics = 10 has Coherence Value of 0.538  
Num Topics = 15 has Coherence Value of 0.541  
Num Topics = 20 has Coherence Value of 0.510  
Num Topics = 25 has Coherence Value of 0.522  
Num Topics = 30 has Coherence Value of 0.520  
Num Topics = 35 has Coherence Value of 0.528  
Num Topics = 40 has Coherence Value of 0.531  
Num Topics = 45 has Coherence Value of 0.533  
Num Topics = 50 has Coherence Value of 0.513  
Num Topics = 55 has Coherence Value of 0.532  
Num Topics = 60 has Coherence Value of 0.493  
Num Topics = 65 has Coherence Value of 0.492  
Num Topics = 70 has Coherence Value of 0.510  
Num Topics = 75 has Coherence Value of 0.492
```

## PARAMETER TUNING FOR LDA



Iteration 3 – 1 hour 12 min

`corpus=corpus, id2word=id2word, num_topics=num_topics,  
random_state=0, update_every=1, chunksize=100, iterations=50,  
passes=10, alpha='auto', per_word_topics=True`



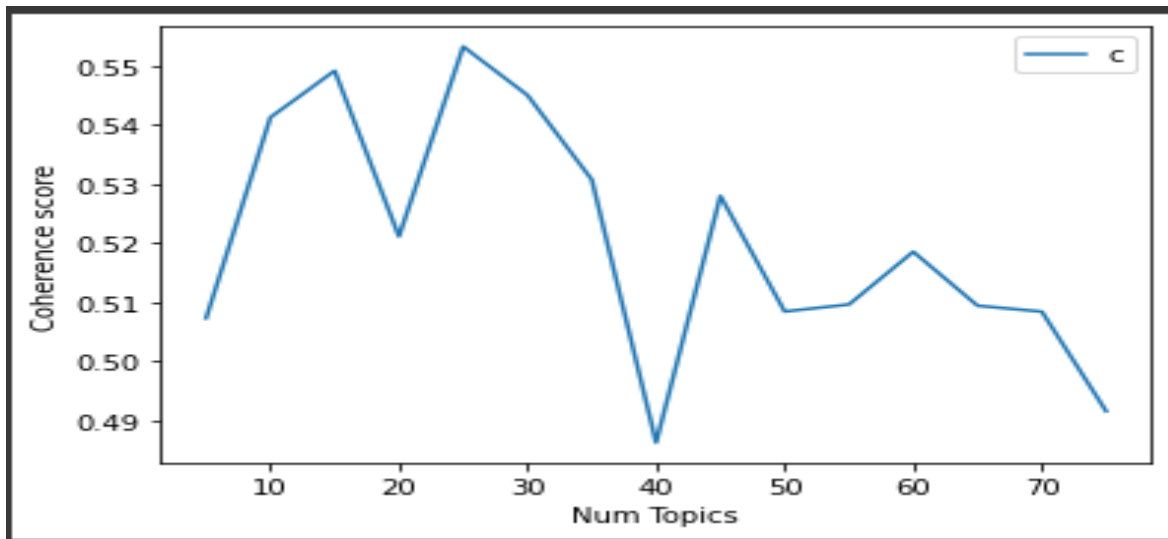
```
Num Topics = 5 has Coherence Value of 0.508  
Num Topics = 10 has Coherence Value of 0.534  
Num Topics = 15 has Coherence Value of 0.526  
Num Topics = 20 has Coherence Value of 0.51  
Num Topics = 25 has Coherence Value of 0.531  
Num Topics = 30 has Coherence Value of 0.497  
Num Topics = 35 has Coherence Value of 0.498  
Num Topics = 40 has Coherence Value of 0.476  
Num Topics = 45 has Coherence Value of 0.486  
Num Topics = 50 has Coherence Value of 0.474  
Num Topics = 55 has Coherence Value of 0.471  
Num Topics = 60 has Coherence Value of 0.475  
Num Topics = 65 has Coherence Value of 0.47  
Num Topics = 70 has Coherence Value of 0.454  
Num Topics = 75 has Coherence Value of 0.451
```

## PARAMETER TUNING FOR LDA

...

Iteration 4: 52 min

`corpus=corpus, id2word=id2word, num_topics=num_topics,  
random_state=42, update_every=1, chunksize=500, iterations=50,  
passes=20, alpha='auto', per_word_topics=True`



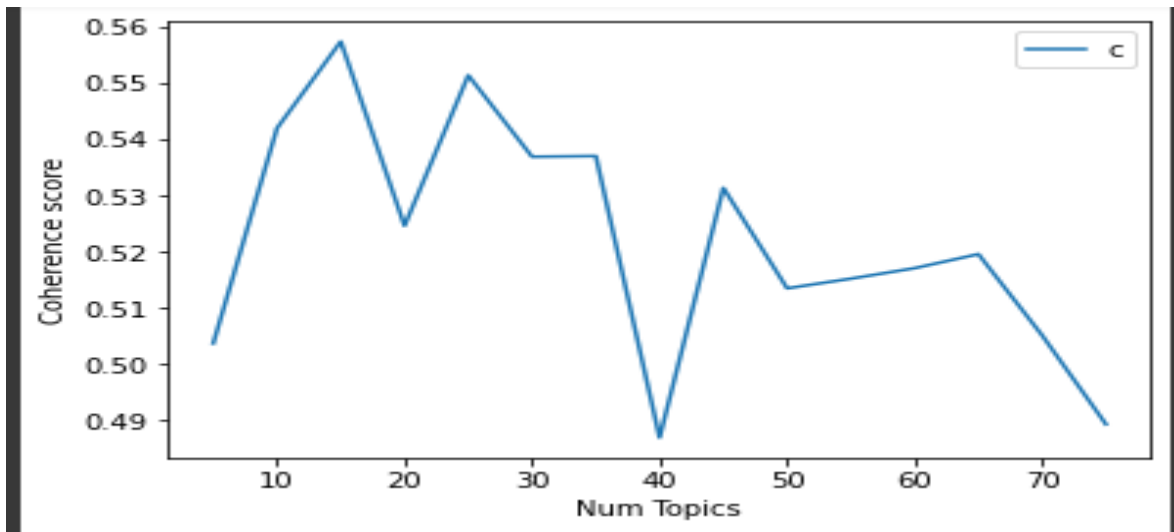
```
Num Topics = 5   has Coherence Value of 0.507
Num Topics = 10  has Coherence Value of 0.541
Num Topics = 15  has Coherence Value of 0.549
Num Topics = 20  has Coherence Value of 0.521
Num Topics = 25  has Coherence Value of 0.553
Num Topics = 30  has Coherence Value of 0.545
Num Topics = 35  has Coherence Value of 0.531
Num Topics = 40  has Coherence Value of 0.486
Num Topics = 45  has Coherence Value of 0.528
Num Topics = 50  has Coherence Value of 0.508
Num Topics = 55  has Coherence Value of 0.51
Num Topics = 60  has Coherence Value of 0.519
Num Topics = 65  has Coherence Value of 0.509
Num Topics = 70  has Coherence Value of 0.508
Num Topics = 75  has Coherence Value of 0.492
```

## PARAMETER TUNING FOR LDA

...

Iteration 5:

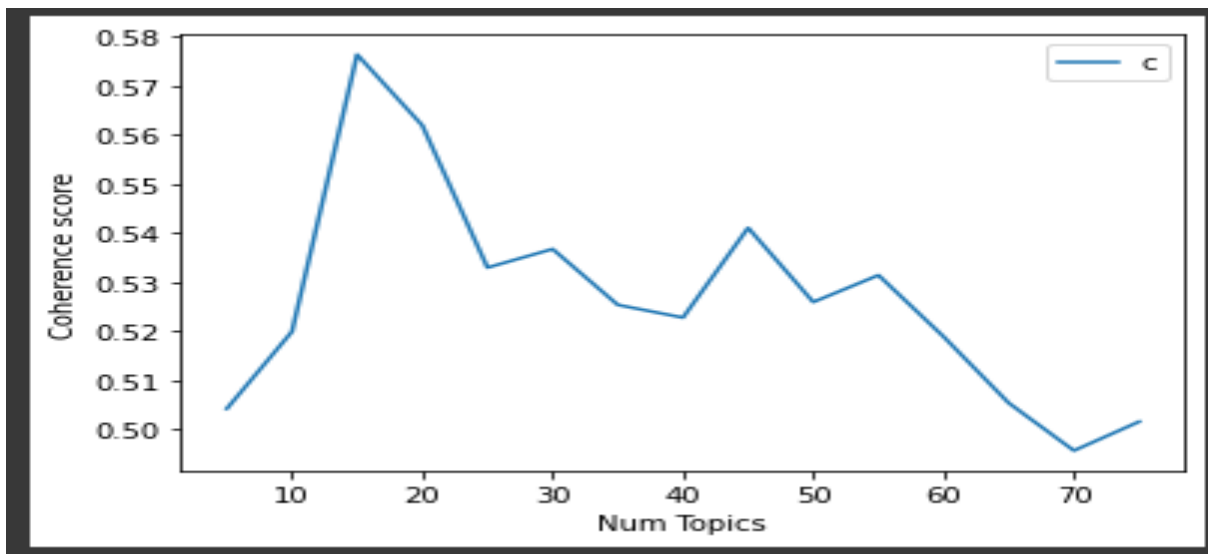
```
corpus=corpus, id2word=id2word, num_topics=num_topics,  
random_state=42, update_every=1, chunksize=500, iterations=50,  
passes=20, alpha='symmetric', decay = 0.6, per_word_topics=True
```



```
Num Topics = 5   has Coherence Value of 0.504  
Num Topics = 10  has Coherence Value of 0.542  
Num Topics = 15  has Coherence Value of 0.557  
Num Topics = 20  has Coherence Value of 0.525  
Num Topics = 25  has Coherence Value of 0.551  
Num Topics = 30  has Coherence Value of 0.537  
Num Topics = 35  has Coherence Value of 0.537  
Num Topics = 40  has Coherence Value of 0.487  
Num Topics = 45  has Coherence Value of 0.531  
Num Topics = 50  has Coherence Value of 0.513  
Num Topics = 55  has Coherence Value of 0.515  
Num Topics = 60  has Coherence Value of 0.517  
Num Topics = 65  has Coherence Value of 0.52  
Num Topics = 70  has Coherence Value of 0.505  
Num Topics = 75  has Coherence Value of 0.489
```

Iteration 6: 49 min

corpus=corpus, id2word=id2word, num\_topics=num\_topics,  
random\_state=42, update\_every=1, chunksize=1000, iterations=50,  
passes=10, alpha='symmetric', decay = 0.8, per\_word\_topics=True



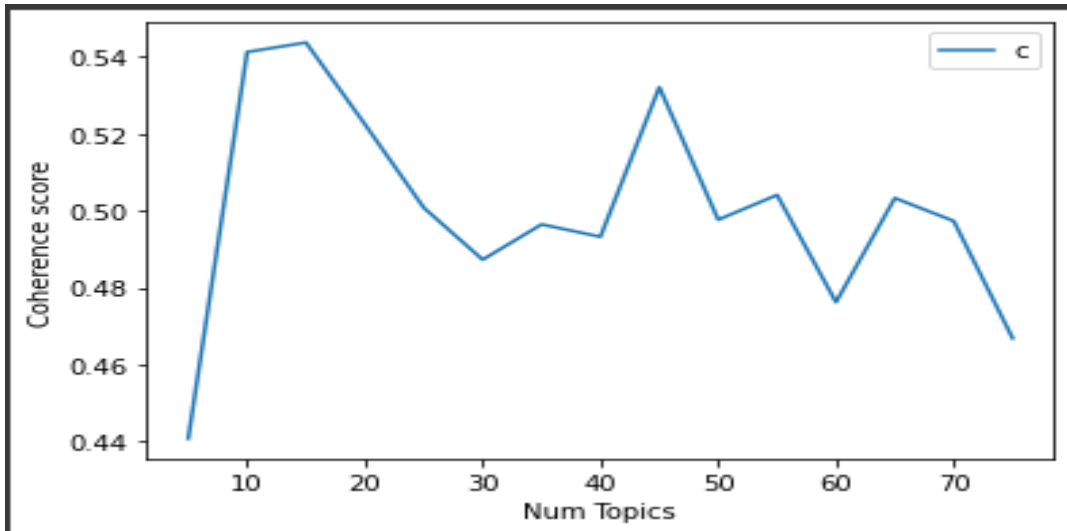
```
Num Topics = 5   has Coherence Value of 0.504
Num Topics = 10  has Coherence Value of 0.52
Num Topics = 15  has Coherence Value of 0.576
Num Topics = 20  has Coherence Value of 0.562
Num Topics = 25  has Coherence Value of 0.533
Num Topics = 30  has Coherence Value of 0.537
Num Topics = 35  has Coherence Value of 0.525
Num Topics = 40  has Coherence Value of 0.523
Num Topics = 45  has Coherence Value of 0.541
Num Topics = 50  has Coherence Value of 0.526
Num Topics = 55  has Coherence Value of 0.531
Num Topics = 60  has Coherence Value of 0.519
Num Topics = 65  has Coherence Value of 0.505
Num Topics = 70  has Coherence Value of 0.496
Num Topics = 75  has Coherence Value of 0.502
```

## PARAMETER TUNING FOR LDA

...

Iteration 7: 37 min

corpus=corpus, id2word=id2word, num\_topics=num\_topics,  
random\_state= 0, update\_every=1, chunksize=4000, iterations=50,  
passes=10, alpha='auto', decay = 0.6, per\_word\_topics=True

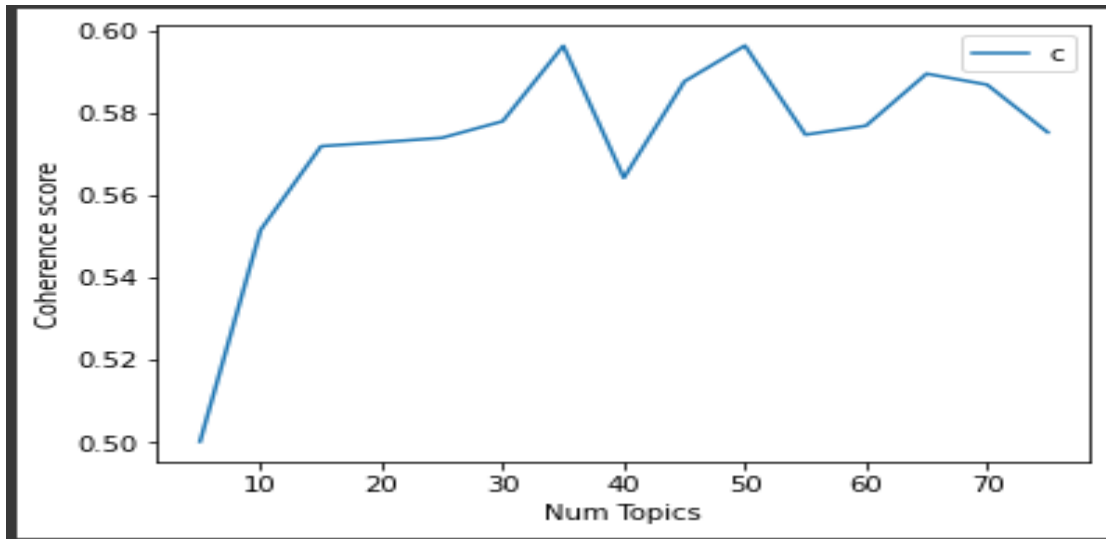


```
Num Topics = 5 has Coherence Value of 0.441
Num Topics = 10 has Coherence Value of 0.541
Num Topics = 15 has Coherence Value of 0.544
Num Topics = 20 has Coherence Value of 0.523
Num Topics = 25 has Coherence Value of 0.501
Num Topics = 30 has Coherence Value of 0.487
Num Topics = 35 has Coherence Value of 0.496
Num Topics = 40 has Coherence Value of 0.493
Num Topics = 45 has Coherence Value of 0.532
Num Topics = 50 has Coherence Value of 0.498
Num Topics = 55 has Coherence Value of 0.504
Num Topics = 60 has Coherence Value of 0.476
Num Topics = 65 has Coherence Value of 0.503
Num Topics = 70 has Coherence Value of 0.497
Num Topics = 75 has Coherence Value of 0.467
```



Iteration 8:

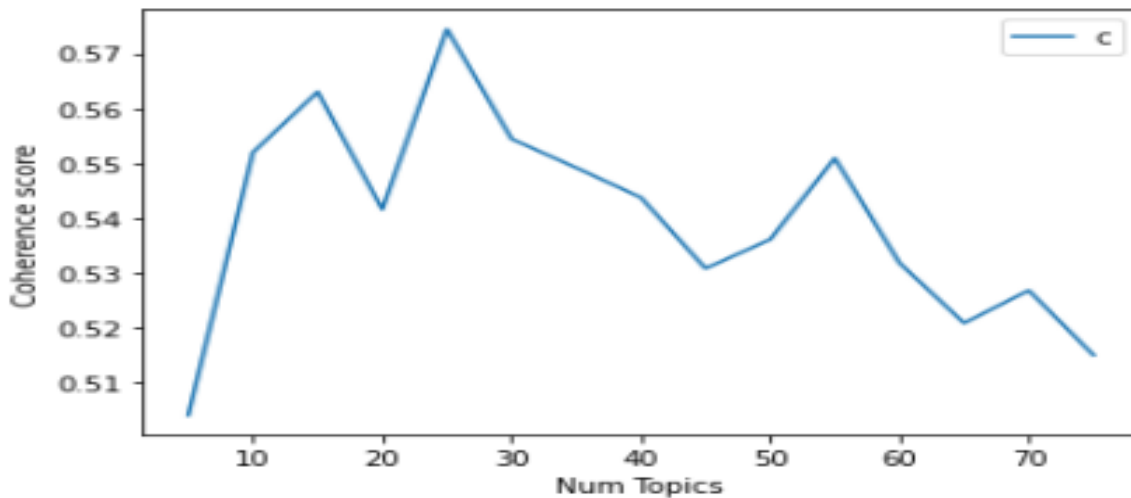
## LDA MALLET



```
Num Topics = 5 has Coherence Value of 0.5
Num Topics = 10 has Coherence Value of 0.552
Num Topics = 15 has Coherence Value of 0.572
Num Topics = 20 has Coherence Value of 0.573
Num Topics = 25 has Coherence Value of 0.574
Num Topics = 30 has Coherence Value of 0.578
Num Topics = 35 has Coherence Value of 0.596
Num Topics = 40 has Coherence Value of 0.564
Num Topics = 45 has Coherence Value of 0.588
Num Topics = 50 has Coherence Value of 0.596
Num Topics = 55 has Coherence Value of 0.575
Num Topics = 60 has Coherence Value of 0.577
Num Topics = 65 has Coherence Value of 0.59
Num Topics = 70 has Coherence Value of 0.587
Num Topics = 75 has Coherence Value of 0.575
```

Iteration 9: 25 min

**Corpus=corpus, id2word=id2word, num\_topics=num\_topics,  
random\_state=0, update\_every=1, chunksize=500, iterations=10,  
passes=10, alpha='symmetric', per\_word\_topics=True**



Num Topics = 5 has Coherence Value of 0.504

Num Topics = 10 has Coherence Value of 0.552

Num Topics = 15 has Coherence Value of 0.563

Num Topics = 20 has Coherence Value of 0.542

**Num Topics = 25 has Coherence Value of 0.574**

Num Topics = 30 has Coherence Value of 0.554

Num Topics = 35 has Coherence Value of 0.549

Num Topics = 40 has Coherence Value of 0.544

Num Topics = 45 has Coherence Value of 0.531

Num Topics = 50 has Coherence Value of 0.536

Num Topics = 55 has Coherence Value of 0.551

Num Topics = 60 has Coherence Value of 0.532

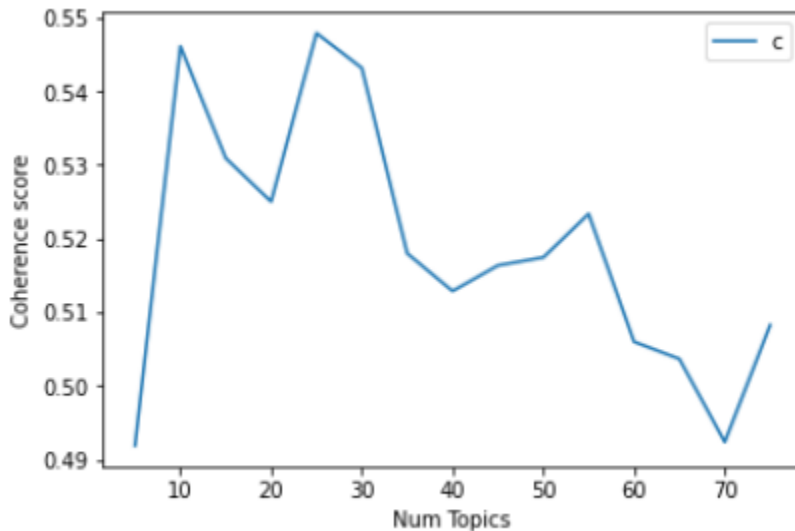
Num Topics = 65 has Coherence Value of 0.521

Num Topics = 70 has Coherence Value of 0.527

Num Topics = 75 has Coherence Value of 0.515

Iteration 10: - 32 min

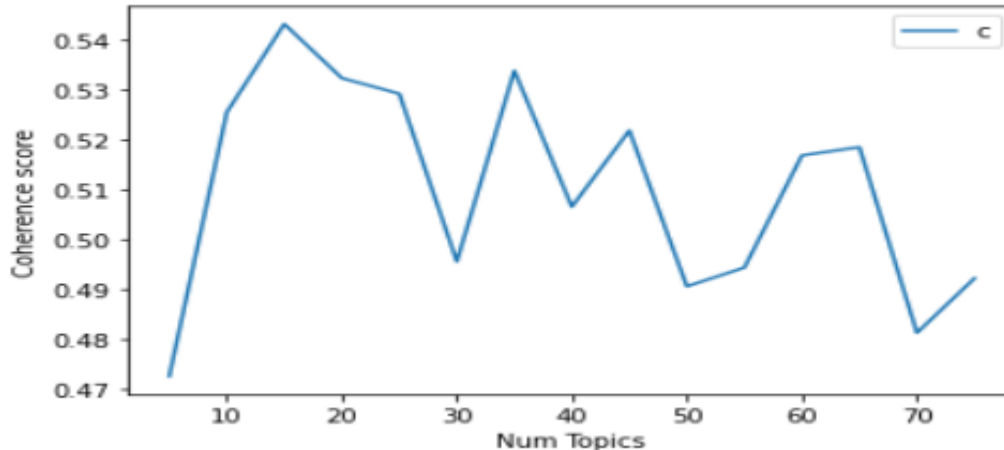
**Corpus=corpus, id2word=id2word, num\_topics=num\_topics,  
random\_state=0, update\_every=1, chunksize=500, iterations=50,  
passes=10, alpha='symmetric', per\_word\_topics=True**



Num Topics = 5 has Coherence Value of 0.492  
Num Topics = 10 has Coherence Value of 0.546  
Num Topics = 15 has Coherence Value of 0.531  
Num Topics = 20 has Coherence Value of 0.525  
**Num Topics = 25 has Coherence Value of 0.548**  
Num Topics = 30 has Coherence Value of 0.543  
Num Topics = 35 has Coherence Value of 0.518  
Num Topics = 40 has Coherence Value of 0.513  
Num Topics = 45 has Coherence Value of 0.516  
Num Topics = 50 has Coherence Value of 0.517  
Num Topics = 55 has Coherence Value of 0.523  
Num Topics = 60 has Coherence Value of 0.506  
Num Topics = 65 has Coherence Value of 0.504  
Num Topics = 70 has Coherence Value of 0.492  
Num Topics = 75 has Coherence Value of 0.508

Iteration 11: - 34 min

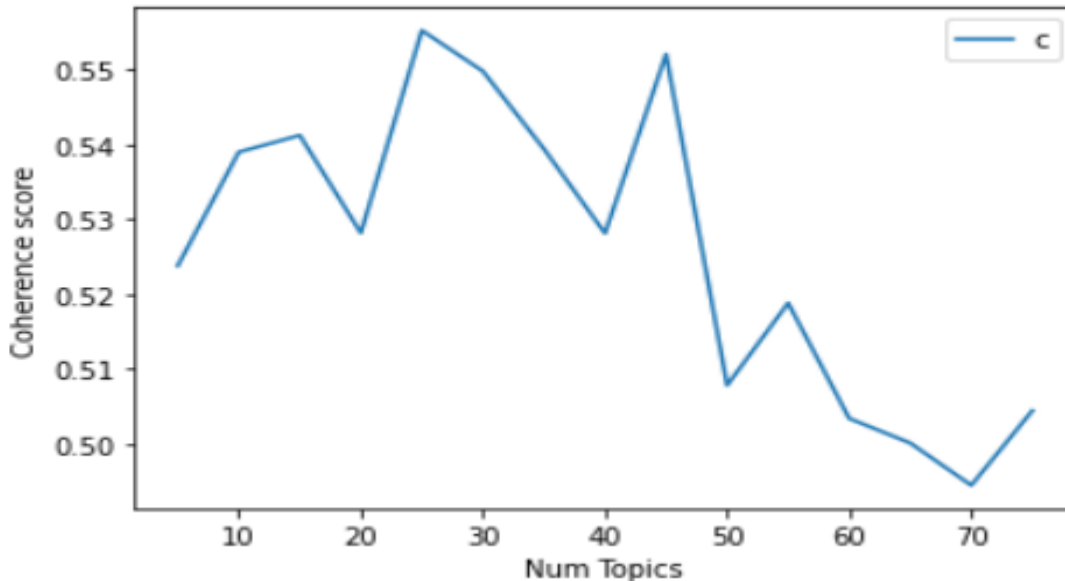
**Corpus=corpus, id2word=id2word, num\_topics=num\_topics,  
random\_state=0, update\_every=1, chunksize=1000, iterations=200,  
passes=10, alpha='symmetric', per\_word\_topics=True**



Num Topics = 5 has Coherence Value of 0.473  
Num Topics = 10 has Coherence Value of 0.525  
**Num Topics = 15 has Coherence Value of 0.543**  
Num Topics = 20 has Coherence Value of 0.532  
Num Topics = 25 has Coherence Value of 0.529  
Num Topics = 30 has Coherence Value of 0.496  
Num Topics = 35 has Coherence Value of 0.534  
Num Topics = 40 has Coherence Value of 0.507  
Num Topics = 45 has Coherence Value of 0.522  
Num Topics = 50 has Coherence Value of 0.491  
Num Topics = 55 has Coherence Value of 0.494  
Num Topics = 60 has Coherence Value of 0.517  
Num Topics = 65 has Coherence Value of 0.518  
Num Topics = 70 has Coherence Value of 0.481  
Num Topics = 75 has Coherence Value of 0.492

Iteration 12: - 60 min

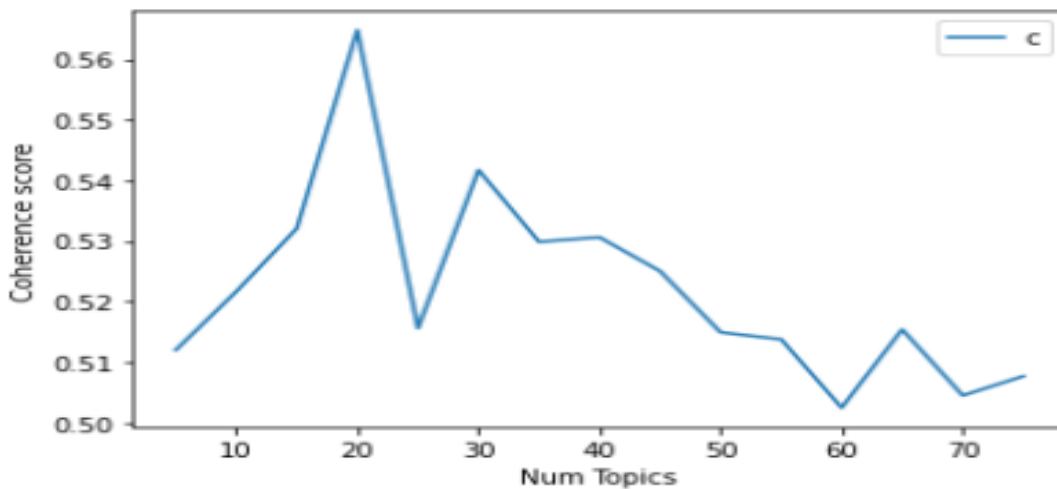
**Corpus=corpus, id2word=id2word, num\_topics=num\_topics,  
random\_state=42, update\_every=1, chunksize=1000, iterations=500,  
passes=20, alpha='auto', per\_word\_topics=True**



Num Topics = 5 has Coherence Value of 0.524  
Num Topics = 10 has Coherence Value of 0.539  
Num Topics = 15 has Coherence Value of 0.541  
Num Topics = 20 has Coherence Value of 0.528  
**Num Topics = 25 has Coherence Value of 0.555**  
Num Topics = 30 has Coherence Value of 0.55  
Num Topics = 35 has Coherence Value of 0.539  
Num Topics = 40 has Coherence Value of 0.528  
Num Topics = 45 has Coherence Value of 0.552  
Num Topics = 50 has Coherence Value of 0.508  
Num Topics = 55 has Coherence Value of 0.519  
Num Topics = 60 has Coherence Value of 0.503  
Num Topics = 65 has Coherence Value of 0.5  
Num Topics = 70 has Coherence Value of 0.494  
Num Topics = 75 has Coherence Value of 0.504

Iteration 13: - 56 min

**Corpus=corpus, id2word=id2word, num\_topics=num\_topics,  
random\_state=42, update\_every=1, chunksize=1000, iterations=200,  
passes=20, alpha='auto', per\_word\_topics=True**



Num Topics = 5 has Coherence Value of 0.512  
Num Topics = 10 has Coherence Value of 0.522  
Num Topics = 15 has Coherence Value of 0.532  
**Num Topics = 20 has Coherence Value of 0.565**  
Num Topics = 25 has Coherence Value of 0.516  
Num Topics = 30 has Coherence Value of 0.542  
Num Topics = 35 has Coherence Value of 0.53  
Num Topics = 40 has Coherence Value of 0.531  
Num Topics = 45 has Coherence Value of 0.525  
Num Topics = 50 has Coherence Value of 0.515  
Num Topics = 55 has Coherence Value of 0.514  
Num Topics = 60 has Coherence Value of 0.502  
Num Topics = 65 has Coherence Value of 0.515  
Num Topics = 70 has Coherence Value of 0.504  
Num Topics = 75 has Coherence Value of 0.508