Natural Language Processing

AI Research Papers

RobFaltisco@gmail.com
San Francisco, CA
faltisco

The Data

30 Years of AI Research Papers at the Neural Information Processing Systems Conference (NeurIPS)

The largest AI Conference in the world (Wikipedia)

The LDA Model

In NLP, a Latent Dirichlet Allocation model finds hidden topics. The LDA model was used to analyze each of four decades of AI research papers: 1980s, 1990s, 2000s, and 2010s.

Motivation

- Artificial Intelligence topics have evolved from 1980s until today
- Can LDA topic modeling discover an evolution of topics?
- Can this pattern over these four decades be better understood using such LDA?
- Which visualization technique and/or model output assists best in understanding topic evolution?
  - Word Clouds?
  - Numeric Coefficients?
  - LDAvis interactive display?

Model Results

One way of viewing the results is looking at the key terms and the weights each of those terms has in "informing" the topic of interest. In the LDA topics found, one example output was as follows:

```
[(0,
  '0.043*"model" + 0.024*"response" + 0.022*"subject" + 0.022*"human" + '
  '0.019*"functional" + 0.016*"brain" + 0.016*"trial" + 0.014*"stimulus" + '
  '0.014*"movement" + 0.014*"experiment"'),
 (1,
  '0.104*"tree" + 0.081*"word" + 0.066*"topic" + 0.031*"language" + '
  '0.030*"structure" + 0.028*"model" + 0.025*"node" + 0.019*"discover" + '
  '0.018*"modeling" + 0.016*"context"'),
 (2,
  '0.047*"signal" + 0.046*"sparse" + 0.039*"code" + 0.032*"rate" + '
  '0.023*"neural" + 0.022*"response" + 0.020*"population" + '
  '0.019*"correlation" + 0.016*"threshold" + 0.015*"potential"'),
 (3,
  '0.060*"learn" + 0.049*"feature" + 0.029*"training" + 0.019*"classification
  '+ 0.019*"task" + 0.018*"performance" + 0.018*"learning" + 0.018*"train" +
  '0.017*"label" + 0.017*"class"'),
```

The top four (4) topics are shown, each with the ten (10) most important terms within that topic. For example, *Topic 0* is most informed by the term *model* and its coefficient is *0.043*. A human meaning was applied to these NLP results because the model doesn't truly understand meanings.

Text Cleaning

An important step in making the text ready for LDA models is preprocessing. The *7,241* papers were divided into each of the four decades. Then, the text was tokenized, then stop words removed. Next, the text was lemmatized and bigrams replaced (a bigram is a phrase like "neural network", that was turned into "neural_network.")

Dictionary and Corpus

A dictionary and corpus were built across all the terms in the body of papers. This enables model to build numerical analysis.

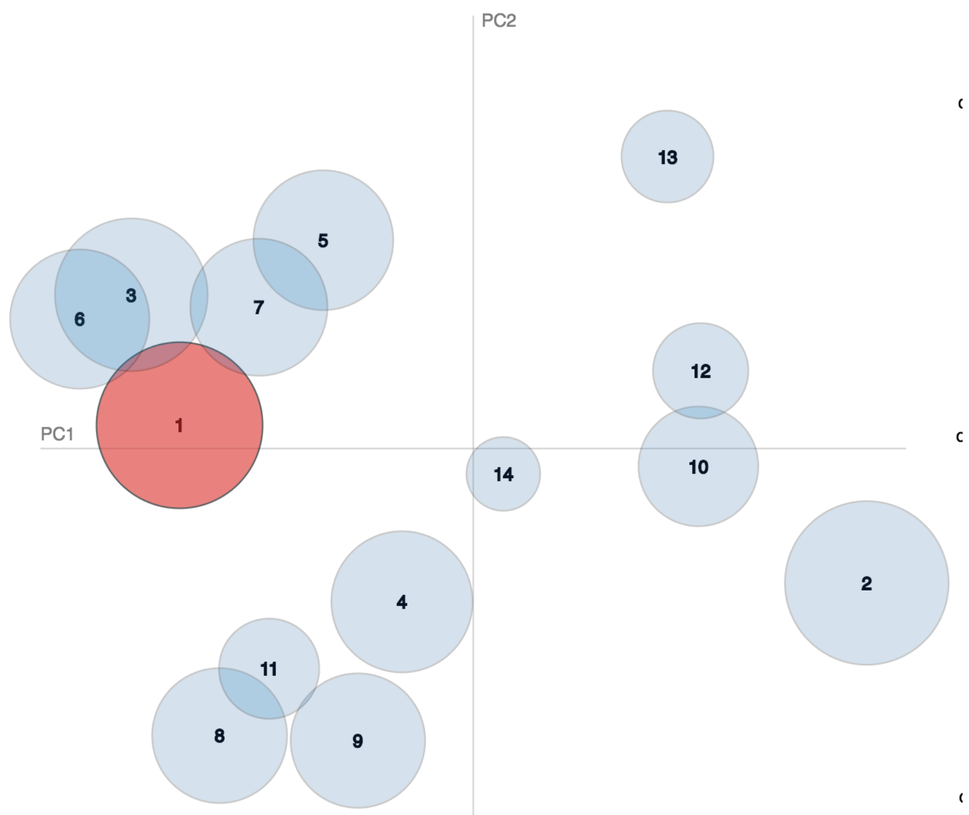
dict_1980s:

```
Dictionary(3342 unique tokens:
['accelerate', 'acceptable', 'accordingly',
'achieve', 'actual']...)
```

corp_1980s:

```
[[(0, 1), (1, 2), (2, 6), (3, 2),
(4, 3), (5, 2), (6, 1), (7, 1),
(8, 1), (9, 2), (10, 1),...]
```

Intertopic Distance Map (via multidimensional scaling)



Marginal topic distribution

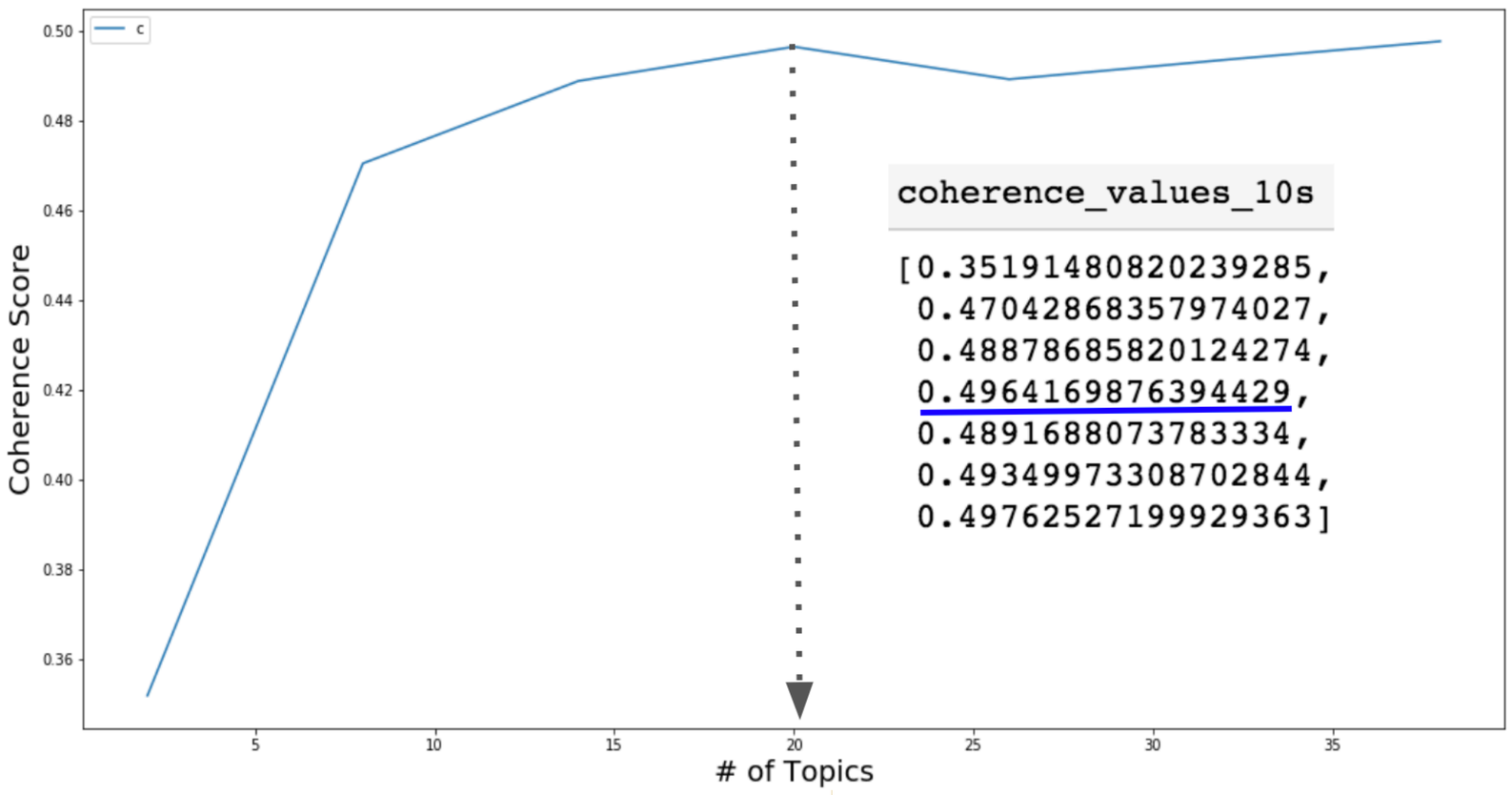
Output from the LDAvis visualization tool

*note LDAvis changes Topic 0 to 1

Coherence Analysis

An important "quality" measure of the LDA model is calculating the coherence value, which measures how well the topics fit together and how distinct each found topic is from the others.

2010s Optimal Number of Topics




```
coherence_values_10s

[0.35191480820239285,
 0.47042868357974027,
 0.48878685820124274,
 0.4964169876394429,
 0.4891688073783334,
 0.49349973308702844,
 0.49762527199929363]
```

By altering the desired number of topics and then building an LDA model for each number, and then plotting the coherence values, the optimal number of topics was chosen for each decade of papers. In the example above, for the 2010s, the num_topics parameter was ultimately set to 20. For other decades, 14.

Conclusions

It's important to note that much of the model analysis included using the visualization tool *LDAvis* (a partial screenshot is shown above). *LDAvis* has interactivity and graphical display functions that enable quicker, easier analysis. The final analysis of each decade's *Topic 1* using LDAvis revealed interesting results shown below. As an added bonus result, a 2016-17 models' results are included below. Here are the found topics, with human named titles:

1980s - "Modeling the Brain"
1990s - "Vision/Image Models"
2000s - "Probability Analysis Applied to Text Data"
2010s - "Computational Power and Linear Algebra"
2016-17s -"Neural Networks: Deep Learning (Images)"