

LDA TASK_VARYING_K

Samingun Handoyo - 0880816

Present key word in distribution of frequency

	Key word	Frequency
0	activity	1
1	addition	1
2	application	1
3	asymptotic	1
4	availability	1

	Key word	Frequency
63	version	1
64	weighted	1
65	withour	1
66	without	1
67	xed	1

LDA topic modeling with k=5

```
[(0,
  '0.018*"model" + 0.009*"network" + 0.008*"learning" + 0.007*"data" + '
  '0.007*"risk" + 0.006*"inflation" + 0.006*"financial" + 0.006*"deep" + '
  '0.005*"factor" + 0.005*"research"'),
 (1,
  '0.017*"model" + 0.014*"risk" + 0.012*"market" + 0.010*"volatility" + '
  '0.007*"time" + 0.006*"portfolio" + 0.006*"cryptocurrencies" + 0.006*"data" '
  '+ 0.005*"tail" + 0.005*"price"'),
 (2,
  '0.020*"model" + 0.011*"method" + 0.011*"data" + 0.011*"variable" + '
  '0.007*"estimator" + 0.006*"parameter" + 0.005*"nonparametric" + '
  '0.005*"adaptive" + 0.005*"change" + 0.005*"test"'),
 (3,
  '0.016*"model" + 0.011*"risk" + 0.008*"distribution" + 0.008*"time" + '
  '0.007*"data" + 0.006*"method" + 0.005*"factor" + 0.005*"gaussian" + '
  '0.005*"asset" + 0.005*"high"'),
 (4,
  '0.016*"market" + 0.015*"risk" + 0.009*"index" + 0.007*"network" + '
  '0.007*"financial" + 0.007*"model" + 0.006*"based" + 0.005*"currency" + '
  '0.005*"tail" + 0.005*"crix"')]
```

Perplexity: -7.63496186687069

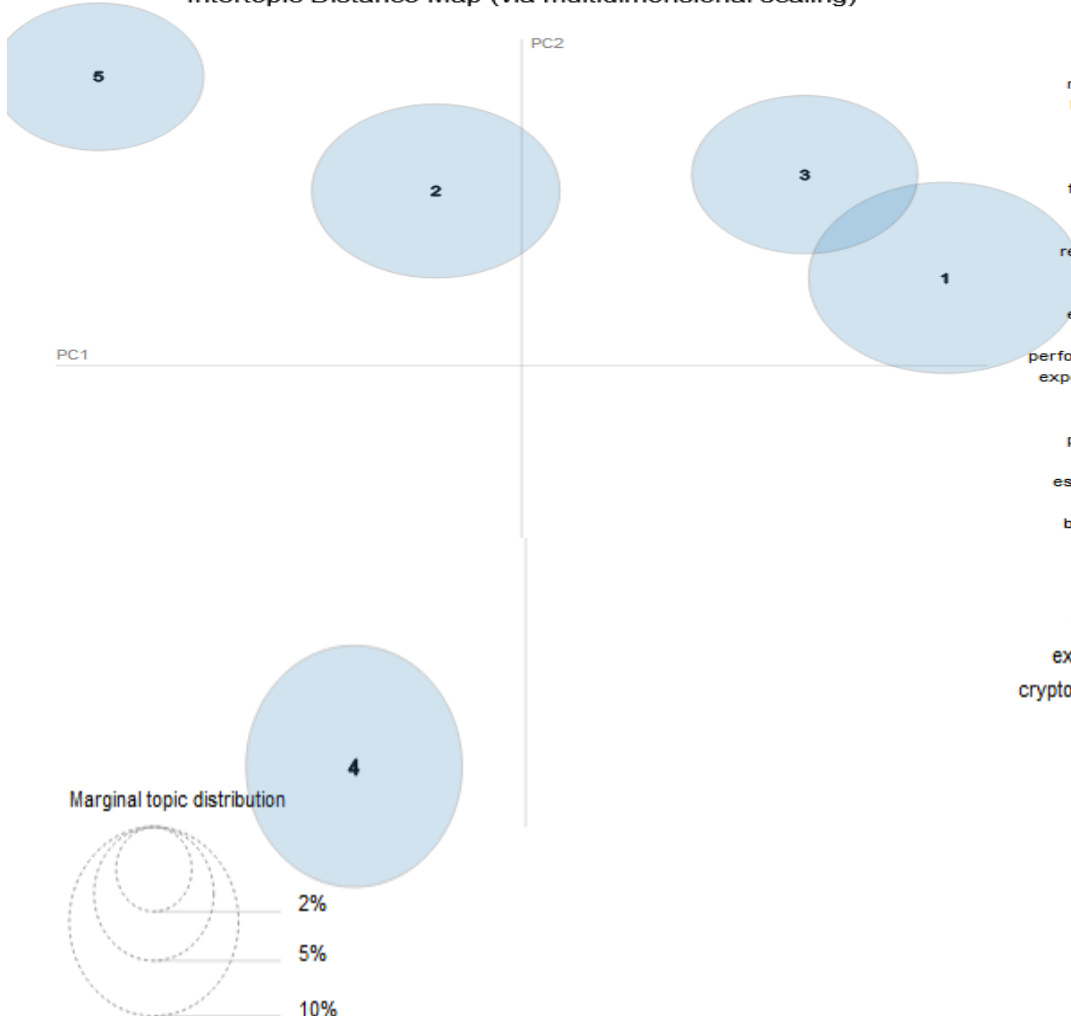
Coherence Score: 0.29801775800298713

The LDA visualization with k=5

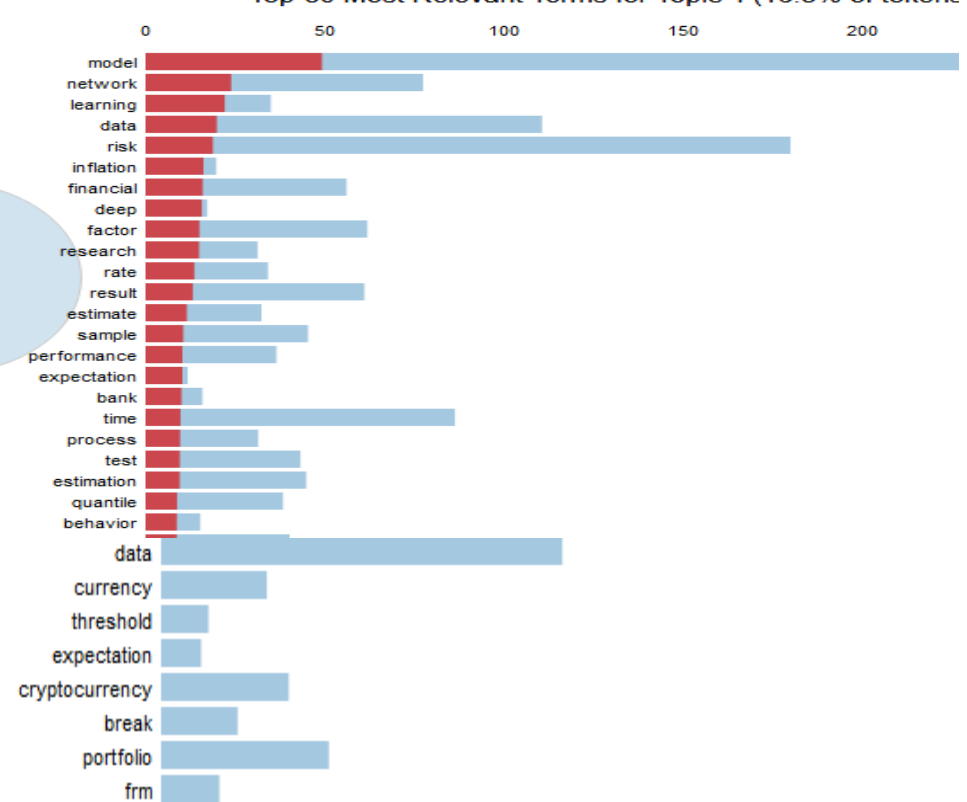
Selected Topic:

Slide to adjust relevance metric:(2)
 $\lambda = 1$ 0.0 0.2 0.4 0.6 0

Intertopic Distance Map (via multidimensional scaling)



Top-30 Most Relevant Terms for Topic 4 (16.3% of tokens)



Overall term frequency

Estimated term frequency within the selected topic

1. $saliency(term\ w) = frequency(w) * [\sum_t p(t|w) * \log(p(t|w)/p(t))]$ for topics t ; see Chuang et al. (2012)

2. $relevance(term\ w|topic\ t) = \lambda * p(w|t) + (1 - \lambda) * p(w|t)/p(w)$; see Sievert & Shirley (2014)

The LDA visualization with k=10

Selected Topic: 3

Previous Topic

Next Topic

Clear Topic

Slide to adjust relevance metric:(2)

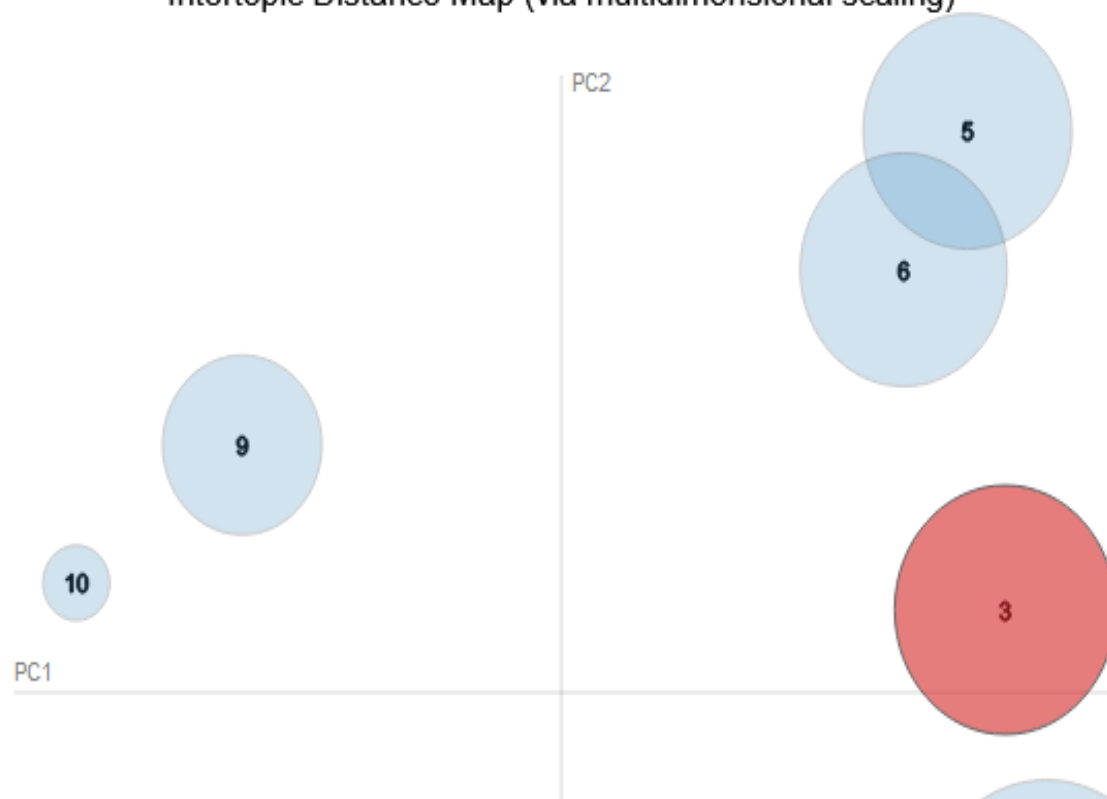
$\lambda = 1$

0.0

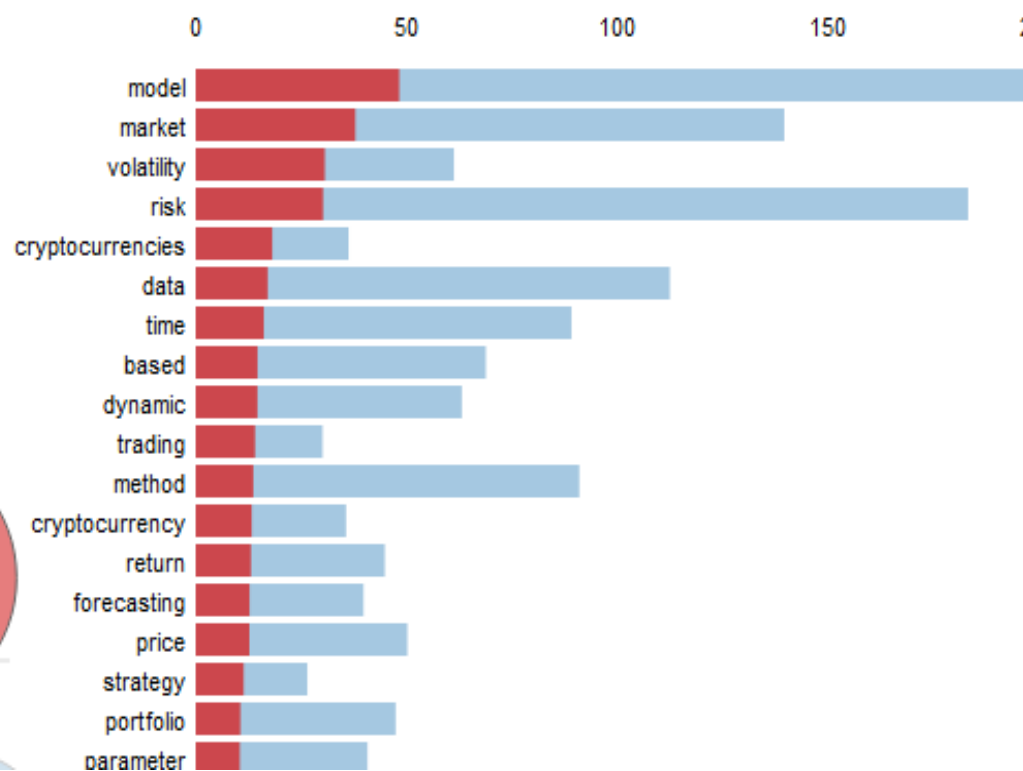
0.2

0.4

Intertopic Distance Map (via multidimensional scaling)



Top-30 Most Relevant Terms for Topic 3 (12.8%)



The LDA visualization with k=10

