

LDA2Net: Digging under the surface of COVID-19 topics in literature

Topic 79 companion sheet

G. Minello

C.R.M.A. Santagiustina

M. Warglien

This file contains the following supplementary information for Topic 79 of the manuscript “*LDA2Net: Digging under the surface of COVID-19 topics in scientific literature*”:

- Human label and automatic n-gram label proposals (Table 1)
- Summary measures (Table 2)
- Network of top 25 bigrams (Figure 1)
- Wordclouds of top 25 words by node relevance measure (Figure 2)
- Wordclouds of top 25 bigrams by edge relevance measure (Figure 3)
- Filtered (0.99 percentile) topic network (Figure 4)

Table 1: Human and automatic label proposals. Automatic label candidate for largest word community of the topic. In parenthesis: absolute frequency of the walk out of a sample of size 1000.

Human label	2-gram label	3-gram label	4-gram label
impact on energy market	paper->shows (7.7%)	paper->shows->impact (7.7%)	paper->shows->impact->examine (3.2%)

Here follows the set of topic-specific measures that have been used to classify the topic and to analyse its structural properties (see manuscript for details):

Table 2: Summary measures

	JSD	Mean propensity	Variance propensity	Modularity	Barrat Clustering Coeff.
value	0.571981	0.007740	0.000368	0.127015	0.559598
rank	22	32	72	89	38

Based on the aforementioned measures, Topic 79 has been classified as a SPECIALIZED topic.

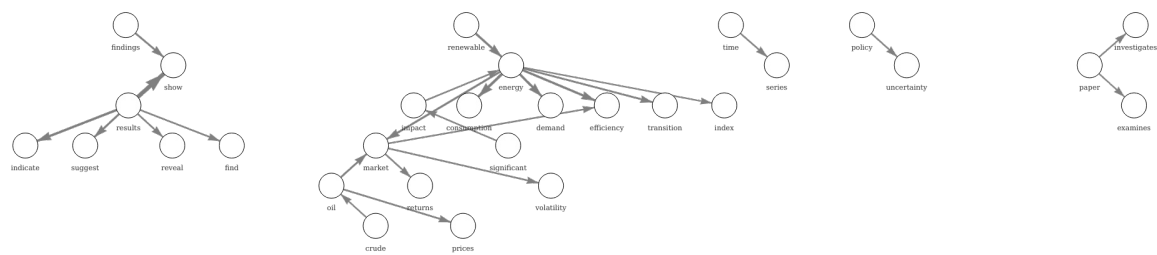


Figure 1: Network of top 25 bigrams (i.e., edges) by weight.

A word cloud of factors influencing oil prices. The words are arranged in a circular pattern, with 'show' and 'results' being the most prominent. Other words include 'prices', 'uncertainty', 'consumption', 'demand', 'oil', 'energy', 'market', 'impact', 'model', 'findings', 'significant', 'volatility', 'efficiency', 'index', 'indicate', 'suggest', 'paper', 'time', 'increase', 'effects', 'renewable', and 'returns'.

A word cloud featuring various terms such as "significant", "efficiency", "demand", "reveal", "paper", "oil", "energy", "results", "timefind", "returns", "consumption", "prices", "volatility", "uncertainty", "show", "market", "impact", "suggest", "index", "transition", and "increase". The words are arranged in a circular pattern around a central point, with varying font sizes indicating their relative frequency or importance.

[illegible]

A word cloud of terms related to the paper's content. The words are arranged in a circular pattern around the central word "paper". The words include: examines, returns, increase, market, analysis, provides, examine, impact, time, significant, volatility, show, results, shows, transition, demand, consumption, efficiency, period, and investigates. The word "paper" is the largest and most central, while "period" is the smallest.

[illegible]

Out-degree

Betweenness

PageRank

Figure 2: Top 25 unigrams (i.e., nodes) by measure.



Figure 3: Top 25 bigrams (i.e., edges) by measure.

