

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

Topic 61 companion sheet

G. Minello

C.R.M.A. Santagiustina

M. Warglien

This file contains the following supplementary information for Topic 61 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
data for the study	analysis->used (9.9%)	data->analysis->used (4.3%)	data->analysis->used->study (1.6%)

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	0.842469	0.007974	0.000122	0.000159	0.641858
rank	112	44	12	46	115

Based on the aforementioned measures, Topic 61 has been classified as a CROSS-CUTTING topic.

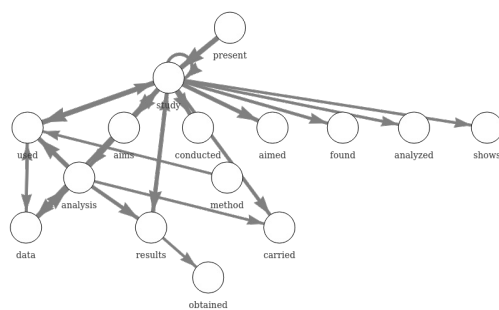


Figure 1: Network of top 25 bigrams (i.e., edges) by weight.

component
furthermore addition
aimed conducted
obtained
used present
results data
study
analysis
descriptive
main aims
analyze
analyzed
indonesia
quantitative
statistical
method comparative
findings aim
moreover

A word cloud of keywords related to the study. The most prominent word is 'study'. Other significant words include 'analysis', 'type', 'results', 'conducted', 'obtained', 'used', 'aims', 'present', 'various', 'moreover', 'related', 'addition', 'provided', 'T1', 'total', 'carried', 'found', 'different', 'aimed', 'according', 'investigated', 'applied', and 'analyzed'. The words are arranged in a circular pattern, with 'study' at the center.

3



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

