

Topic Modelling Resources

Keefellow

2022-03-13

Contents

1	About this collection	5
2	Textbooks	7
2.1	Mastering Text Mining with R	7
2.2	Practical Text Mining and Statistical Analysis for Non-structured Text Data Applications	7
2.3	Text Mining with R	7
2.4	Natural Language Processing with Python and spaCy	8
3	Website Resources	9
3.1	Apache OpenNLP library is a machine learning based toolkit for the processing of natural language text.	9
3.2	The Stanford Natural Language Processing Group	9
3.3	The Summer Institutes in Computational Social Science	9
3.4	Computational Social Science (Soc 690)	10
4	Journal Articles	11
4.1	A review of topic modeling methods	11
4.2	Using Topic Modeling Methods for Short-Text Data: A Compar- ative Analysis	11
4.3	Topic Modeling: A Comprehensive Review	11
5	YouTube Resources	13
5.1	SICSS 2018 - Topic Models/Structural Topic Models (Day 3. June 20, 2018)	13
5.2	Structure Topic Modelling	13

6	Structural Topic Modelling	15
6.1	stm: R Package for Structural Topic Models	15
6.2	Tutorials	15
7	Publications	17
7.1	The Structural Topic Model and Applied Social Science	17
7.2	Structural Topic Models for Open-Ended Survey Responses	17
7.3	Mining quality determinants of product-service systems from user-generated contents	17
7.4	COVID-19 Twitter Communication of Major Societal Stakehold- ers: Health Institutions, the Government, and the News Media	17
7.5	Topic Modeling and User Network Analysis on Twitter during World Lupus Awareness Day	18
7.6	Toward understanding the topical structure of hospitality litera- ture: Applying machine learning and traditional statistics.	18
7.7	Topic Modeling as a Strategy of Inquiry in Organizational Re- search: A Tutorial With an Application Example on Organiza- tional Culture	18
8	Yet to classify	19
8.1	Apache OpenNLP library is a machine learning based toolkit for the processing of natural language text.	19

Chapter 1

About this collection

There are plenty of resources for learning topic modelling, and this holds some of them.

Chapter 2

Textbooks

2.1 Mastering Text Mining with R

Chapter 5. Text Summarization and Clustering - Topic modeling - Latent semantic analysis - Machine learning-based text summarization - Text clustering - Sentence completion <https://learning.oreilly.com/library/view/mastering-text-mining/9781783551811/ch05.html>

2.2 Practical Text Mining and Statistical Analysis for Non-structured Text Data Applications

An Exploratory Literature Survey Using Text Mining and Beyond : A practical example <https://learning.oreilly.com/library/view/practical-text-mining/9780123869791/xhtml/CHP005.html#CESECTITLE0100>

2.3 Text Mining with R

Chapter on Topic Modelling <https://learning.oreilly.com/library/view/text-mining-with/9781491981641/ch06.html>

2.4 Natural Language Processing with Python and spaCy

Linguistic Primer <https://learning.oreilly.com/library/view/natural-language-processing/9781098122652/xhtml/app01.xhtml#lev174>

Extracting and using linguistic features <https://learning.oreilly.com/library/view/natural-language-processing/9781098122652/xhtml/ch04.xhtml#lev45>

Chapter 3

Website Resources

##Topic Models Learning and R Resources https://github.com/trinker/topicmodels_learning

3.1 Apache OpenNLP library is a machine learning based toolkit for the processing of natural language text.

<https://opennlp.apache.org/>

3.2 The Stanford Natural Language Processing Group

<https://nlp.stanford.edu/software/>

3.3 The Summer Institutes in Computational Social Science

Open source teaching and learning resources for computational social science <https://sicss.io/overview> Previous editions <https://github.com/compsocialscience/summer-institute>

3.4 Computational Social Science (Soc 690)

Course materials from Duke University https://cbail.github.io/comp_soc_grad/Home.html#An_Introduction_to_Text_Analysis

Chapter 4

Journal Articles

4.1 A review of topic modeling methods

<https://doi.org/10.1016/j.is.2020.101582> See Fig 8 for Decision tree for selecting a topic modeling method. <https://ars-els-cdn-com.libproxy.nie.edu.sg/content/image/1-s2.0-S0306437920300703-gr8.jpg>

4.2 Using Topic Modeling Methods for Short-Text Data: A Comparative Analysis

<https://www.frontiersin.org/articles/10.3389/frai.2020.00042/full>

4.3 Topic Modeling: A Comprehensive Review

<https://eprints.eudl.eu/id/eprint/682/1/eai.13-7-2018.159623.pdf>

Chapter 5

YouTube Resources

5.1 SICSS 2018 - Topic Models/Structural Topic Models (Day 3. June 20, 2018)

<https://www.youtube.com/watch?v=k-zkxdxyjTk>

5.2 Structure Topic Modelling

<https://www.youtube.com/watch?v=3kcjbzy4UPM>

Chapter 6

Structural Topic Modelling

6.1 stm: R Package for Structural Topic Models

<https://cran.r-project.org/web/packages/stm/vignettes/stmVignette.pdf>

6.2 Tutorials

http://thomaselliott.me/pdfs/earl/topic_modeling.html https://cbail.github.io/SICSS_Topic_Modeling.html
<https://me.eui.eu/andrea-de-angelis/blog/structural-topic-models-to-study-political-text-an-application-to-the-five-star-movements-blog/>

Chapter 7

Publications

7.1 The Structural Topic Model and Applied Social Science

<https://scholar.princeton.edu/files/bstewart/files/stmnips2013.pdf>

7.2 Structural Topic Models for Open-Ended Survey Responses

<https://scholar.harvard.edu/files/dtingley/files/topicmodelsopenendedexperiments.pdf>

7.3 Mining quality determinants of product-service systems from user-generated contents

<https://doi.org/10.1080/08982112.2021.1877305>

7.4 COVID-19 Twitter Communication of Major Societal Stakeholders: Health Institutions, the Government, and the News Media

<https://ijoc.org/index.php/ijoc/article/view/17147/3575>

7.5 Topic Modeling and User Network Analysis on Twitter during World Lupus Awareness Day

<https://dx.doi.org/10.3390%2Fijerph17155440>

7.6 Toward understanding the topical structure of hospitality literature: Applying machine learning and traditional statistics.

<https://psycnet.apa.org/doi/10.1108/IJCHM-11-2017-0714>

7.7 Topic Modeling as a Strategy of Inquiry in Organizational Research: A Tutorial With an Application Example on Organizational Culture

<https://doi.org/10.1177/1094428118773858>

Chapter 8

Yet to classify

8.1 Apache OpenNLP library is a machine learning based toolkit for the processing of natural language text.

<https://opennlp.apache.org/>

<https://learning.oreilly.com/library/view/mastering-text-mining/9781783551811/ch05.html>

<https://learning.oreilly.com/library/view/practical-text-mining/9780123869791/xhtml/CHP005.html#CESECTITLE0100>

<https://learning.oreilly.com/library/view/text-mining-with/9781491981641/ch06.html>

<https://learning.oreilly.com/library/view/natural-language-processing/9781098122652/xhtml/app01.xhtml#lev174>

<https://learning.oreilly.com/library/view/natural-language-processing/9781098122652/xhtml/ch04.xhtml#lev45> <https://www.youtube.com/watch?v=3kcjbzy4UPM>

Be familiar with document term matrix

<https://www.youtube.com/watch?v=k-zkxdxyjTk> download this video

<https://sicss.io/overview> https://cbail.github.io/comp_soc_grad/Home.html#An_Introduction_to_Text_Analysis

<https://github.com/compsocialscience/summer-institute>

<https://sicss.io/locations#2017>

<https://doi-org.libproxy.nie.edu.sg/10.1016/j.is.2020.101582> <https://ars-els-cdn-com.libproxy.nie.edu.sg/content/image/1-s2.0-S0306437920300703-gr8.jpg>

<https://www.frontiersin.org/articles/10.3389/frai.2020.00042/full>

<https://eprints.eudl.eu/id/eprint/682/1/eai.13-7-2018.159623.pdf>

A review of topic modeling methods <https://doi.org/10.1016/j.trc.2017.12.018>