Kenneth Chu
Division de la Science des données
Data Science Division

CANDEV Data Challenge
Ottawa – January 18, 2020

Topic Modeling

Latent Dirichlet Allocation in R
Hierarchical Bayesian models

Objective (intuitively speaking)

www.clipartpanda.com

Given a "corpus" $\mathcal{C}$ of documents :

$$\mathcal{C} = \{D_1, \ldots, D_N\}$$

Want to find :

- the set $\mathcal{T} = \{\, T_k \,\}_{k=1}^{K}$ of "topics" that occur in $D_1, \ldots, D_N$,
- the **topic allocation vector**

$$\theta_i \;=\; (\theta_{i1}, \ldots, \theta_{iK}) \;\in\; \Delta^{K-1} \;\subset\; \mathbb{R}^K$$

of D_i, where θ_{ik} is the "proportion" of D_i attributable to T_k.

Latent Dirichlet Allocation

Journal of Machine Learning Research, 3 (2003), 993–1022

`http://www.jmlr.org/papers/v3/blei03a.html`

David M. Blei
U.C. Berkeley

Andrew Y. Ng
Stanford University

Michael I. Jordan
U.C. Berkeley

Jonathan K. Pritchard, Matthew Stephens and Peter Donnelly

Inference of Population Structure Using Multilocus Genotype Data

GENETICS June 1, 2000, vol. 155, no. 2, 945-959

https://www.genetics.org/content/155/2/945

Jonathan K. Pritchard
Stanford University

Matthew Stephens
University of Chicago

Peter Donnelly
University of Oxford

Features : document-term matrix

Matrix of word counts

	w_1	w_2	w_3	$\cdots$	w_V
D_1	3	0	0	$\cdots$	0
D_2	0	0	n_{23}	$\cdots$	9
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
D_N	0	1	0	$\cdots$	0

$\mathcal{V} = \{ w_j \}_{j=1}^{V}$ is

the vocabulary of words

that occur in

$\mathcal{C} = \{D_1, \ldots, D_N\}$.

Parameters : documents as mixtures of topics

	T_1	T_2	$\cdots$	T_K	$\mathcal{T}$
D_1	θ_{11}	θ_{12}	$\cdots$	θ_{1K}	θ_1
D_2	θ_{21}	θ_{22}	$\cdots$	θ_{2K}	θ_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
D_{N-1}	0.2	0.7	$0 \cdots 0$	0.1	θ_{N-1}
D_N	θ_{N1}	θ_{N2}	$\cdots$	θ_{NK}	θ_N

θ_i = topic allocation vector of D_i

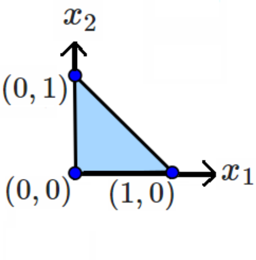
$\mathcal{T} = \{ T_k \}_{k=1}^{K}$ is the set of topics that occur in $\mathcal{C} = \{D_1, \ldots, D_N\}$.

Each D_i is regarded as a "mixture" (probability distribution) of topics :

$$\theta_i \in \Delta^{K-1} \iff \begin{cases} \theta_{ik} \geq 0, & \forall\ i, k \\ \sum_{k=1}^{K} \theta_{ik} = 1, & \forall\ i \end{cases}$$

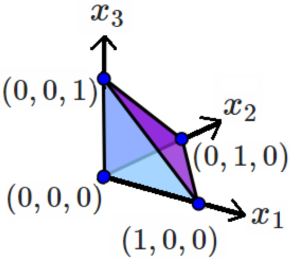
Probability simplexes

Parameter spaces of categorical distributions



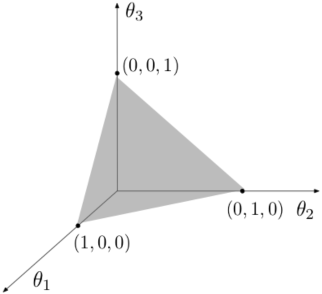
Δ^{2-1} = 1-simplex

Categorial(2)



Δ^{3-1} = 2-simplex

Categorial(3)



Δ^{3-1} = 2-simplex

Categorical(3)

Parameters : topics as mixtures of words

	w_1	w_2	w_3	$\cdots$	w_V	$\mathcal{V}$
T_1	φ_{11}	φ_{12}	φ_{13}	$\cdots$	φ_{1V}	φ_1
T_2	φ_{21}	φ_{22}	φ_{23}	$\cdots$	φ_{2V}	φ_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
T_K	φ_{K1}	φ_{K2}	φ_{K3}	$\cdots$	φ_{KV}	φ_K

φ_k = word allocation vector of T_k

$\varphi_k \in \Delta^{V-1} \subset \mathbb{R}^V$

Each T_k regarded as a "mixture" (probability distribution) of words :

$$\begin{cases} \varphi_{kj} \geq 0, & \forall\ k, j \\ \sum_{j=1}^{V} \varphi_{kj} = 1, & \forall\ k \end{cases}$$

(Postulated) data-generation mechanism

Parameters : vector φ_k of word probabilities of T_k

	w_1	w_2	w_3	$\cdots$	w_V	$\mathcal{V}$
T_1	φ_{11}	φ_{12}	φ_{13}	$\cdots$	φ_{1V}	φ_1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
T_K	φ_{K1}	φ_{K2}	φ_{K3}	$\cdots$	φ_{KV}	φ_K

Parameters : vector θ_i of topic probabilities of D_i

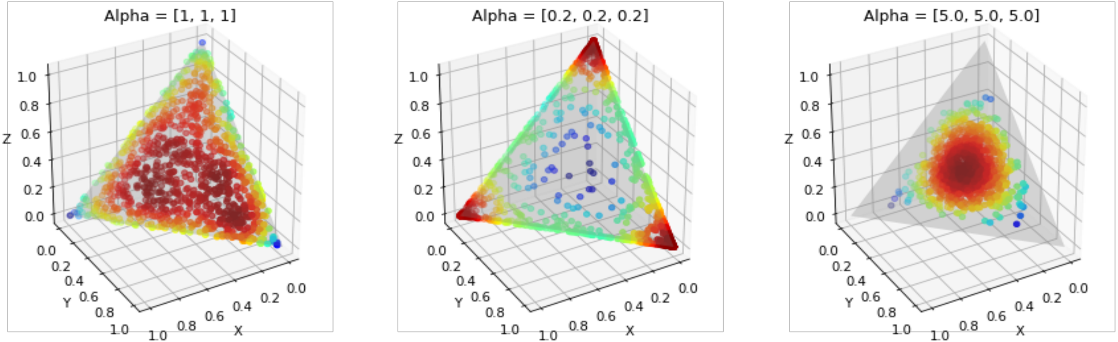
	T_1	T_2	$\cdots$	T_K	$\mathcal{T}$
D_i	θ_{i1}	θ_{i2}	$\cdots$	θ_{iK}	θ_i

Observed : vector n_i of word counts of D_i

	w_1	w_2	w_3	$\cdots$	w_V	$\mathcal{V}$
D_i	n_{i1}	n_{i2}	n_{i3}	$\cdots$	n_{iV}	n_i

- For each $k = 1, 2, \ldots, K$, choose
  $$\varphi_k \sim \text{Dirichlet}(\Delta^{V-1}; \beta)$$
- For each $i = 1, 2, \ldots, N$, choose
  $$\theta_i \sim \text{Dirichlet}(\Delta^{K-1}; \alpha)$$
- For each $j = 1, 2, \ldots, |D_i|$,
  - first choose topic
    $$T_{k(i,j)} \sim \text{Multinom}(\mathcal{T}; \theta_i)$$
  - then choose word
    $$W_{i,j} \sim \text{Multinom}(\mathcal{V}; \varphi_{k(i,j)})$$
- Let n_i be resulting vector of word counts for D_i.

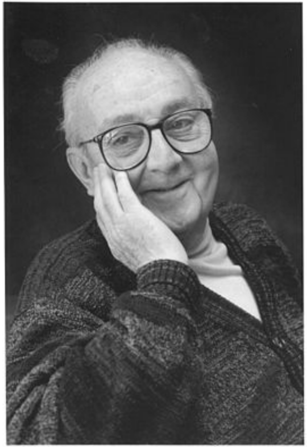
The Dirichlet distributions



https://towardsdatascience.com

George E. P. Box

"All models are wrong,
but
some are useful."



en.wikipedia.org/wiki/George_E._P._Box

Inference

- Hierarchical Bayes
- Input :
  - Hyperparameters[1] : α, β
  - Observed data : n_i
- Algorithm : WarpLDA

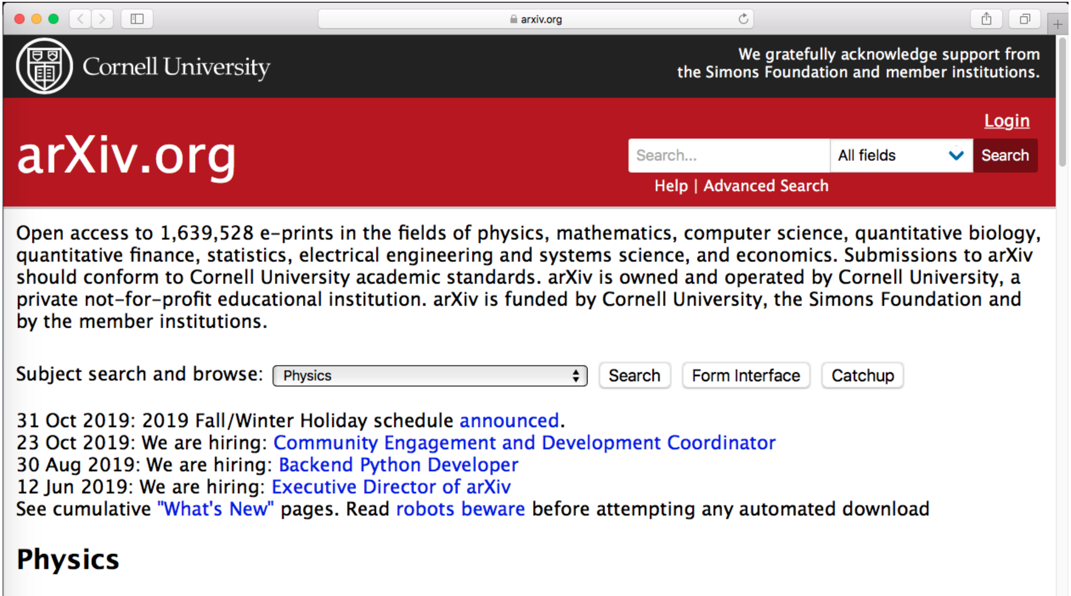
  `LatentDirichletAllocation`
  in R package `text2vec`

- Output : Maximum *a posteriori*[2] (MAP) estimates for

$$\left[\ \theta_i\ \right] \in (\Delta^{K-1})^N \subset \mathbb{R}^{N \times K}, \qquad \left[\ \varphi_k\ \right] \in (\Delta^{V-1})^K \subset \mathbb{R}^{K \times V}$$

1. in the sense of Bayesian statistics
2. mode of posterior distribution

Demo / Experiment


arxiv.org
Cornell University
We gratefully acknowledge support from the Simons Foundation and member institutions.
Login
arXiv.org
Search...
All fields
Search
Help | Advanced Search
Open access to 1,639,528 e-prints in the fields of physics, mathematics, computer science, quantitative biology, quantitative finance, statistics, electrical engineering and systems science, and economics. Submissions to arXiv should conform to Cornell University academic standards. arXiv is owned and operated by Cornell University, a private not-for-profit educational institution. arXiv is funded by Cornell University, the Simons Foundation and by the member institutions.
Subject search and browse:
Physics
Search
Form Interface
Catchup
31 Oct 2019: 2019 Fall/Winter Holiday schedule announced.
23 Oct 2019: We are hiring: Community Engagement and Development Coordinator
30 Aug 2019: We are hiring: Backend Python Developer
12 Jun 2019: We are hiring: Executive Director of arXiv
See cumulative "What's New" pages. Read robots beware before attempting any automated download
Physics

Computer Science > Machine Learning

AEGR: A simple approach to gradient reversal in autoencoders for network anomaly detection

Kasra Babaei, Zhi Yuan Chen, Tomas Maul

(Submitted on 21 Dec 2019)

Anomaly detection is referred to as a process in which the aim is to detect data points that follow a different pattern from the majority of data points. Anomaly detection methods suffer from several well-known challenges that hinder their performance such as high dimensionality. Autoencoders are unsupervised neural networks that have been used for the purpose of reducing dimensionality and also detecting network anomalies in large datasets. The performance of autoencoders debilitates when the training set contains noise and anomalies. In this paper, a new gradient-reversal method is proposed to overcome the influence of anomalies on the training phase for the purpose of detecting network anomalies. The method is different from other approaches as it does not require an anomaly-free training set and is based on reconstruction error. Once latent variables are extracted from the network, Local Outlier Factor is used to separate normal data points from anomalies. A simple pruning approach and data augmentation is also added to further improve performance. The experimental results show that the proposed model can outperform other well-know approaches.

Subjects: **Machine Learning (cs.LG)**; Machine Learning (stat.ML)

Cite as: arXiv:1912.13387 [cs.LG]

(or arXiv:1912.13387v1 [cs.LG] for this version)

Bibliographic data

[Enable Bibex(What is Bibex?)]

Submission history

From: Kasra Babaei [view email]

Download:

- PDF
- Other formats

(license)

Current browse context:

cs.LG

< prev | next >

new | recent | 1912

Change to browse by:

cs
stat
stat.ML

References & Citations

- NASA ADS

Export citation
Google Scholar

Bookmark

Science WISE

Demo / Experiment – data

- Downloaded **abstracts** of 1000 articles of each of the following six domains :

Domain	Description	count
cs-LG	machine learning (computer science)	1000
math-AG	algebraic geometry (mathematics)	999
physics-acc-ph	accelerator physics (physics)	1000
q-bio-GN	genomics (quantitative biology)	997
quant-ph	quantum physics (physics)	995
stat-ME	methodology (statistics)	986

- Removed (23) articles with 2 or more domain labels above.

Demo / Experiment – procedure

- Preprocessing
  convert to lower case, remove stop words
- Create / prune vocabulary
  - $\geq$ 5 occurences over all documents
  - 9076 words
- Generate document-term matrix $M \in \mathbb{Z}^{5977 \times 9076}$

- Perform LDA on M
  - # of topics : 10
  - $\alpha = 50/(\#\text{ of topics})$
  - $\beta = 1/(\#\text{ of topics})$

  which are defaults in `text2vec::LatentDirichletAllocation`
- Examine results

Topics as mixtures of words (probability distributions over vocabulary)

word	Topic1	Topic2	Topic3	Topic4	Topic5	Topic6	Topic7	Topic8	Topic9	Topic10
lep	0	0	0	0	0	0	0	0	8.569898E-05	2.283835E-05
bare	8.308581E-05	2.20085E-05	0	0	0	0	0	0	0	0
bracket	0	0	0	0	0	0	0	0.0001058627	0	0
whitening	0	0	0	0	0	0	3.909228E-05	0	6.427424E-05	0
bifurcation	4.15429E-05	0	6.939304E-05	0	0	0	0	0	0	0
transmitters	0	0	0	9.597482E-05	0	0	0	0	0	0
metarnaseq	0	0	0	0	0	0	0	0	0.0001071237	0
consisted	0	0	0	1.919496E-05	9.124296E-05	0	0	0	0	0
cushaw2	0	0	0	0	4.562148E-05	0	0	0	0	6.851505E-05
bigger	0	0	9.252406E-05	0	0	0	0	0	0	2.283835E-05
testbed	0	2.20085E-05	0	3.838993E-05	0	0	0	0	0	4.56767E-05
quantumness	0.0001038573	0	0	0	0	0	0	0	0	0
bmi	0	0	0	0	0	0	0	0	0.0001071237	0
reactor	0	0	0	9.597482E-05	0	0	0	0	0	0
composting	0	0	0	0	0	0	0	0	0	0.0001141918
fertilization	0	0	0	0	0.0001140537	0	0	0	0	0
5.5	0	0	0	9.597482E-05	0	0	0	0	0	0
dogma	0	0	0	0	6.843222E-05	0	0	0	0	4.56767E-05
informally	0	0	2.313101E-05	0	0	0	0	0	2.142475E-05	6.851505E-05
bcs	0	0	0	9.597482E-05	0	0	0	0	0	0
hardy's	0.0001038573	0	0	0	0	0	0	0	0	0
cohomologically	0	0	0	0	0	0	0	0.0001058627	0	0
achievement	0	0	0	5.758489E-05	0	0	0	0	4.284949E-05	0
ree	0.0001038573	0	0	0	0	0	0	0	0	0
multires	0	0	0	0	6.843222E-05	0	0	0	0	4.56767E-05
resort	0	4.401699E-05	0	0	0	0	0	0	2.142475E-05	4.56767E-05
griffiths	0	0	0	0	0	0	0	0.0001058627	0	0

Most likely words of each topic

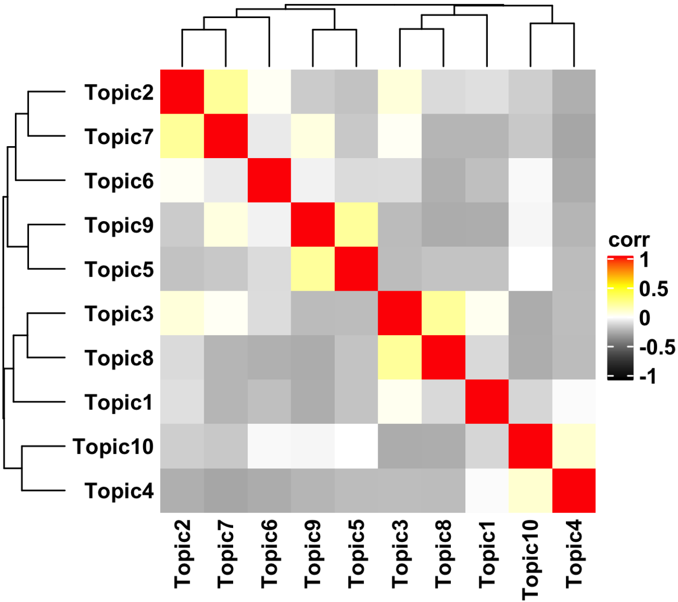
	Topic1	Topic2	Topic3	Topic4	Topic5	Topic6	Topic7	Topic8	Topic9	Topic10
1	quantum	algorithm	terms	beam	sequence	learning	distribution	prove	analysis	reads
2	states	problem	equations	electron	genome	network	inference	mathbb	expression	high
3	systems	matrix	structure	energy	dna	networks	estimation	group	genetic	read
4	system	optimal	series	laser	sequences	neural	regression	1	cancer	code
5	spin	optimization	functions	plasma	genomes	deep	model	2	studies	designed
6	entanglement	problems	properties	particle	protein	training	variables	give	individual	implementation
7	state	solution	surface	radiation	species	classification	bayesian	points	disease	lhc
8	dynamics	approximation	solutions	acceleration	regions	art	distributions	mathcal	wide	power
9	qubit	lower	investigate	beams	human	task	estimate	variety	types	long
10	coupling	bounds	kernel	accelerator	alignment	feature	likelihood	varieties	association	design
11	measurement	stochastic	matrices	bunch	genomic	domain	estimator	degree	seq	technology
12	entropy	number	general	rf	binding	tasks	variable	algebraic	data	fast
13	transition	convergence	examples	magnetic	sites	image	sampling	curves	variation	implemented
14	correlations	complexity	obtained	transverse	tree	input	estimators	smooth	variants	speed
15	coherence	sparse	point	10	mutations	prediction	causal	projective	analyses	quality
16	weak	derive	relations	ion	proteins	knowledge	estimates	theorem	population	project
17	qubits	convex	stability	electrons	diversity	learn	posterior	conjecture	identifying	parallel
18	operator	covariance	basic	proton	evolutionary	adversarial	conditional	surfaces	samples	technologies
19	entangled	theoretical	suitable	ray	regulatory	world	estimating	characteristic	clinical	correction
20	hamiltonian	adaptive	fact	pulse	genes	images	sample	0	patients	luminosity
21	atoms	exact	index	emittance	mrna	trained	random	polynomial	methylation	generation
22	coupled	inverse	definition	accelerators	transcription	machine	carlo	spaces	individuals	years
23	dynamical	minimum	space	operation	phylogenetic	architecture	assumptions	curve	differences	scientific
24	topological	upper	form	frequency	pattern	convolutional	simulation	rational	package	present
25	circuit	block	view	charge	coding	language	confidence	geometric	units	integrated
26	fidelity	linear	application	intensity	trees	generative	parametric	cohomology	snps	science
27	classical	algorithms	structures	muon	nucleotide	layer	monte	moduli	profiles	community

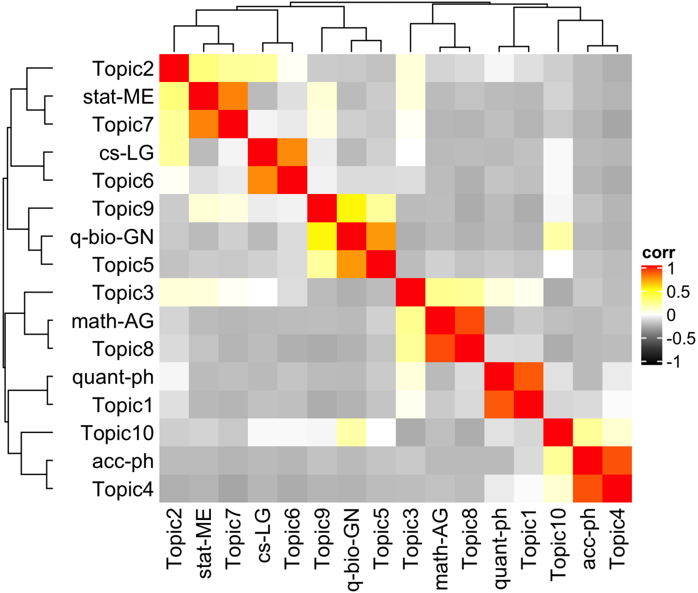
Documents as mixtures of topics (probability distributions over topics)

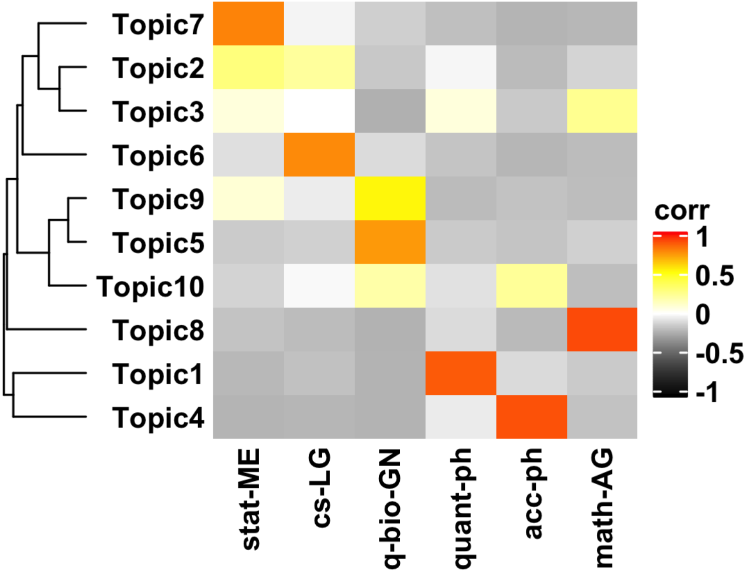
1	document	domain	Topic1	Topic2	Topic3	Topic4	Topic5	Topic6	Topic7	Topic8	Topic9	Topic10
4352	http://arxiv.org/abs/1804.05579v2	quant-ph	0.204	0.051	0.143	0.082	0.051	0.092	0.112	0.092	0.082	0.092
4353	http://arxiv.org/abs/1808.08530v1	quant-ph	0.383	0.060	0.081	0.094	0.054	0.060	0.067	0.087	0.054	0.060
4354	http://arxiv.org/abs/1808.08527v1	quant-ph	0.351	0.072	0.072	0.062	0.082	0.093	0.052	0.062	0.062	0.093
4355	http://arxiv.org/abs/1807.04483v2	quant-ph	0.469	0.063	0.063	0.050	0.056	0.038	0.038	0.088	0.038	0.100
4356	http://arxiv.org/abs/1804.08226v2	quant-ph	0.377	0.051	0.065	0.080	0.072	0.058	0.043	0.101	0.043	0.109
4357	http://arxiv.org/abs/1808.08515v1	quant-ph	0.257	0.073	0.101	0.165	0.055	0.046	0.046	0.101	0.083	0.073
4358	http://arxiv.org/abs/1808.08506v1	quant-ph	0.402	0.091	0.045	0.068	0.045	0.045	0.106	0.061	0.061	0.076
4359	http://arxiv.org/abs/1808.08505v1	quant-ph	0.402	0.063	0.080	0.063	0.054	0.071	0.054	0.098	0.063	0.054
4360	http://arxiv.org/abs/1808.01381v3	quant-ph	0.196	0.082	0.227	0.072	0.052	0.072	0.052	0.093	0.103	0.052
4361	http://arxiv.org/abs/1808.08471v1	quant-ph	0.179	0.060	0.164	0.052	0.067	0.119	0.127	0.097	0.045	0.090
4362	http://arxiv.org/abs/1808.10030v1	quant-ph	0.219	0.061	0.289	0.053	0.079	0.061	0.061	0.053	0.070	0.053
4363	http://arxiv.org/abs/1802.08804v3	quant-ph	0.387	0.070	0.049	0.077	0.063	0.035	0.099	0.092	0.056	0.070
4364	http://arxiv.org/abs/1802.10061v3	quant-ph	0.408	0.046	0.063	0.046	0.057	0.126	0.034	0.080	0.057	0.080
4365	http://arxiv.org/abs/1808.08429v1	quant-ph	0.256	0.093	0.093	0.047	0.085	0.054	0.062	0.078	0.054	0.178
4366	http://arxiv.org/abs/1808.05165v2	quant-ph	0.191	0.183	0.078	0.157	0.043	0.070	0.052	0.113	0.043	0.070
4367	http://arxiv.org/abs/1808.08386v1	quant-ph	0.296	0.038	0.081	0.204	0.048	0.032	0.038	0.167	0.043	0.054
4368	http://arxiv.org/abs/1808.08370v1	quant-ph	0.290	0.076	0.069	0.168	0.046	0.046	0.069	0.076	0.076	0.084
4369	http://arxiv.org/abs/1808.08343v1	quant-ph	0.301	0.068	0.087	0.068	0.107	0.078	0.068	0.078	0.097	0.049
4370	http://arxiv.org/abs/1808.08324v1	quant-ph	0.227	0.114	0.080	0.102	0.057	0.057	0.068	0.080	0.102	0.114
4371	http://arxiv.org/abs/1604.07517v3	quant-ph	0.269	0.084	0.084	0.050	0.042	0.109	0.151	0.084	0.042	0.084
4372	http://arxiv.org/abs/1808.06709v2	quant-ph	0.229	0.101	0.064	0.110	0.064	0.064	0.064	0.092	0.055	0.156
4373	http://arxiv.org/abs/1808.08305v1	quant-ph	0.284	0.125	0.091	0.057	0.057	0.057	0.091	0.091	0.080	0.068
4374	http://arxiv.org/abs/1705.09261v2	quant-ph	0.242	0.111	0.131	0.071	0.061	0.111	0.091	0.051	0.071	0.061
4375	http://arxiv.org/abs/1808.06009v2	quant-ph	0.246	0.038	0.038	0.269	0.062	0.054	0.054	0.054	0.085	0.100
4376	http://arxiv.org/abs/1803.07119v3	quant-ph	0.190	0.183	0.092	0.049	0.070	0.155	0.070	0.063	0.049	0.077
4377	http://arxiv.org/abs/1808.08261v1	quant-ph	0.125	0.125	0.113	0.163	0.075	0.063	0.075	0.100	0.075	0.088
4378	http://arxiv.org/abs/1808.08259v1	quant-ph	0.326	0.076	0.076	0.065	0.065	0.054	0.087	0.087	0.054	0.109
4379	http://arxiv.org/abs/1808.08246v1	quant-ph	0.351	0.099	0.072	0.045	0.099	0.054	0.063	0.072	0.072	0.072

Documents as mixtures of topics (probability distributions over topics)

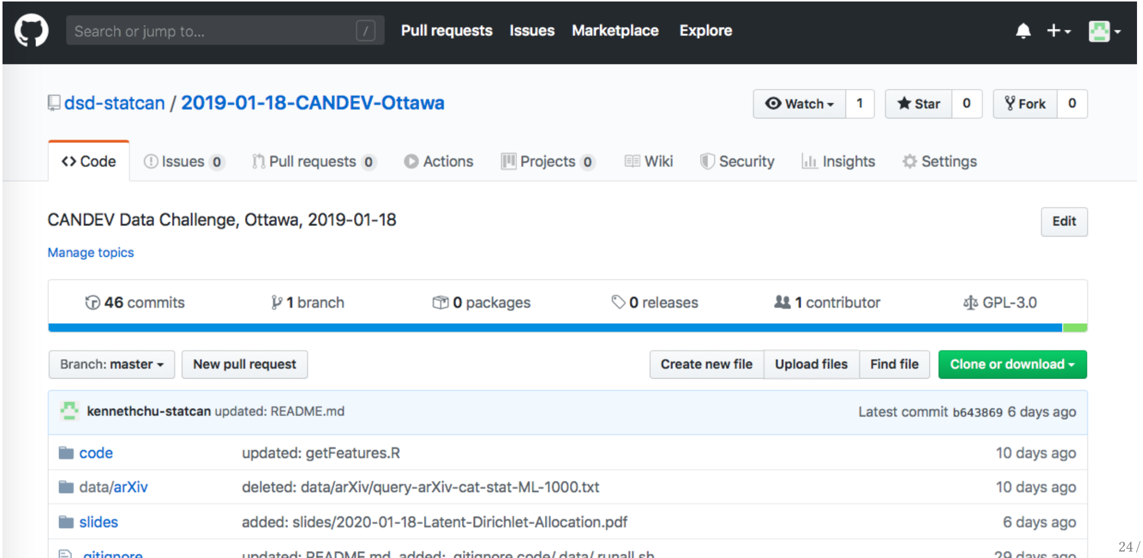
1	document	domain	Topic1	Topic2	Topic3	Topic4	Topic5	Topic6	Topic7	Topic8	Topic9	Topic10
873	http://arxiv.org/abs/1803.05407v2	cs-LG	0.041	0.179	0.049	0.089	0.057	0.293	0.073	0.057	0.081	0.081
874	http://arxiv.org/abs/1808.02651v1	cs-LG	0.061	0.052	0.096	0.052	0.070	0.374	0.052	0.043	0.061	0.139
875	http://arxiv.org/abs/1801.09720v3	cs-LG	0.204	0.290	0.142	0.043	0.049	0.037	0.105	0.037	0.037	0.056
876	http://arxiv.org/abs/1807.00442v2	cs-LG	0.051	0.140	0.059	0.051	0.074	0.243	0.125	0.044	0.096	0.118
877	http://arxiv.org/abs/1808.02610v1	cs-LG	0.061	0.130	0.122	0.061	0.046	0.214	0.107	0.069	0.145	0.046
878	http://arxiv.org/abs/1808.02602v1	cs-LG	0.054	0.125	0.125	0.054	0.071	0.188	0.080	0.054	0.143	0.107
879	http://arxiv.org/abs/1807.05490v2	cs-LG	0.052	0.078	0.078	0.052	0.043	0.319	0.129	0.052	0.121	0.078
880	http://arxiv.org/abs/1711.09535v3	cs-LG	0.061	0.079	0.085	0.037	0.067	0.311	0.159	0.104	0.037	0.061
881	http://arxiv.org/abs/1807.00374v3	cs-LG	0.024	0.053	0.034	0.039	0.058	0.522	0.058	0.077	0.068	0.068
882	http://arxiv.org/abs/1709.07308v2	cs-LG	0.041	0.147	0.162	0.041	0.071	0.183	0.051	0.147	0.086	0.071
883	http://arxiv.org/abs/1808.02480v1	cs-LG	0.079	0.055	0.094	0.055	0.087	0.394	0.087	0.055	0.055	0.039
884	http://arxiv.org/abs/1806.03972v3	cs-LG	0.071	0.080	0.062	0.044	0.062	0.319	0.071	0.062	0.115	0.115
885	http://arxiv.org/abs/1808.02546v1	cs-LG	0.058	0.282	0.087	0.068	0.058	0.097	0.058	0.058	0.107	0.126
886	http://arxiv.org/abs/1808.03147v1	cs-LG	0.080	0.232	0.145	0.043	0.036	0.116	0.123	0.036	0.058	0.130
887	http://arxiv.org/abs/1802.01894v2	cs-LG	0.131	0.208	0.082	0.049	0.038	0.219	0.060	0.137	0.033	0.044
888	http://arxiv.org/abs/1802.04434v3	cs-LG	0.048	0.306	0.068	0.054	0.061	0.088	0.041	0.088	0.075	0.170
889	http://arxiv.org/abs/1808.02513v1	cs-LG	0.047	0.071	0.087	0.087	0.063	0.220	0.071	0.055	0.063	0.236
890	http://arxiv.org/abs/1808.02510v1	cs-LG	0.051	0.154	0.120	0.043	0.060	0.282	0.103	0.043	0.043	0.103
891	http://arxiv.org/abs/1711.05136v5	cs-LG	0.040	0.105	0.056	0.048	0.040	0.387	0.065	0.040	0.081	0.137
892	http://arxiv.org/abs/1808.02474v1	cs-LG	0.041	0.074	0.088	0.047	0.047	0.480	0.054	0.061	0.061	0.047
893	http://arxiv.org/abs/1806.06063v2	cs-LG	0.122	0.082	0.112	0.071	0.071	0.224	0.133	0.061	0.051	0.071
894	http://arxiv.org/abs/1808.02458v1	cs-LG	0.069	0.169	0.200	0.038	0.046	0.100	0.131	0.085	0.115	0.046
895	http://arxiv.org/abs/1808.02435v1	cs-LG	0.062	0.196	0.144	0.062	0.072	0.175	0.082	0.072	0.062	0.072
896	http://arxiv.org/abs/1808.02433v1	cs-LG	0.056	0.176	0.083	0.065	0.046	0.315	0.074	0.065	0.046	0.074
897	http://arxiv.org/abs/1805.08809v2	cs-LG	0.078	0.174	0.087	0.043	0.061	0.304	0.096	0.043	0.061	0.052
898	http://arxiv.org/abs/1805.08273v2	cs-LG	0.064	0.100	0.055	0.055	0.045	0.073	0.364	0.045	0.136	0.064
899	http://arxiv.org/abs/1808.02394v1	cs-LG	0.072	0.063	0.090	0.081	0.045	0.288	0.099	0.063	0.063	0.135


Topic2
Topic7
Topic6
Topic9
Topic5
Topic3
Topic8
Topic1
Topic10
Topic4
corr
1
0.5
0
-0.5
-1
Topic2
Topic7
Topic6
Topic9
Topic5
Topic3
Topic8
Topic1
Topic10
Topic4


Topic2
stat-ME
Topic7
cs-LG
Topic6
Topic9
q-bio-GN
Topic5
Topic3
math-AG
Topic8
quant-ph
Topic1
Topic10
acc-ph
Topic4
corr
1
0.5
0
-0.5
-1


Topic7
Topic2
Topic3
Topic6
Topic9
Topic5
Topic10
Topic8
Topic1
Topic4
stat-ME
cs-LG
q-bio-GN
quant-ph
acc-ph
math-AG
corr
1
0.5
0
-0.5
-1

https://github.com/dsd-statcan/2019-01-18-CANDEV-Ottawa



Merci !!

Personne-ressource

Pour plus d'information,
veuillez contacter :

For more information,
please contact:

Kenneth Chu

kenneth.chu@canada.ca

613-852-7361