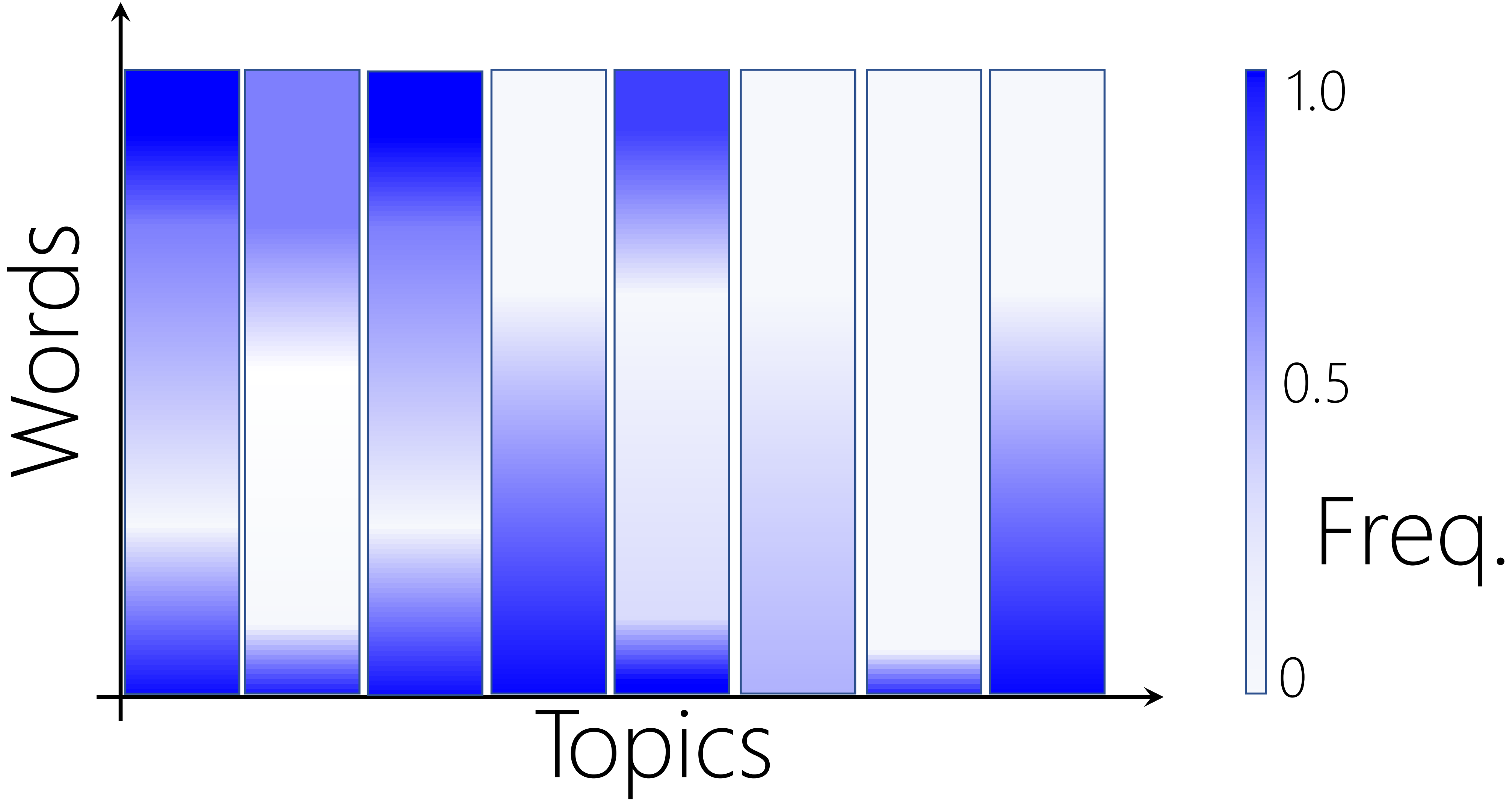
Probabilistic Topic Models

Probabilistic Topic Models

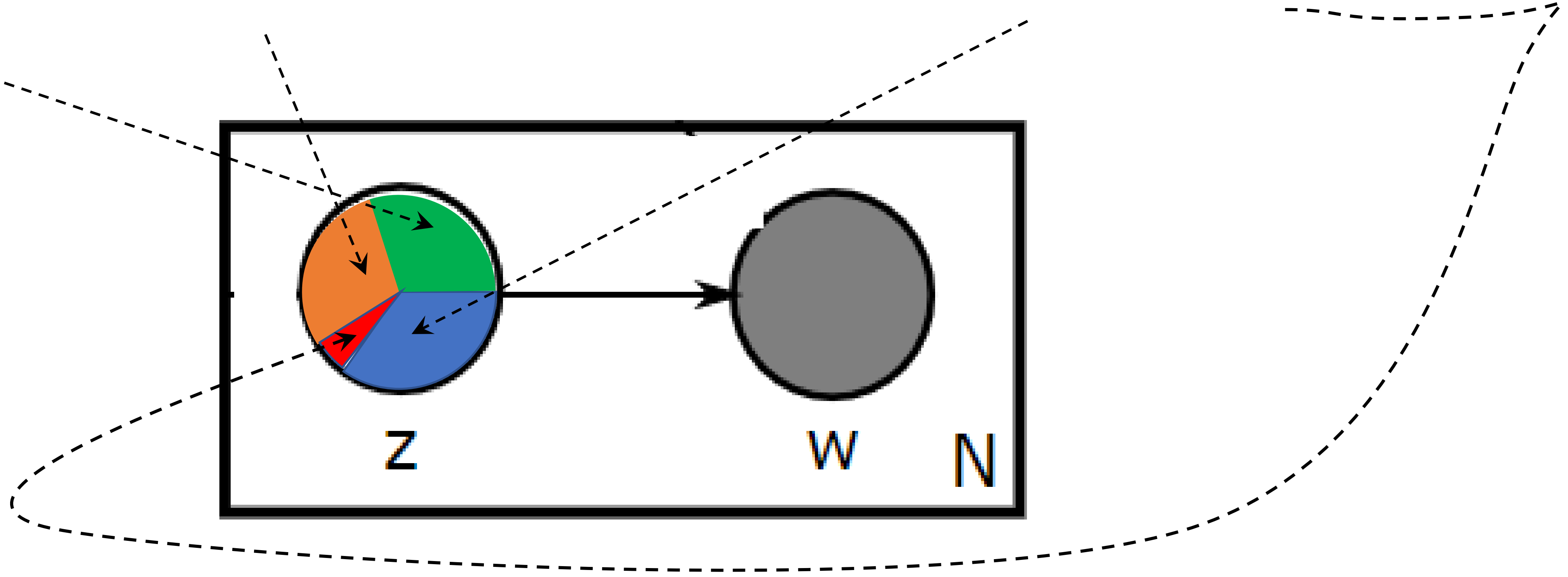


Probabilistic Topic Models

Let's write a document?

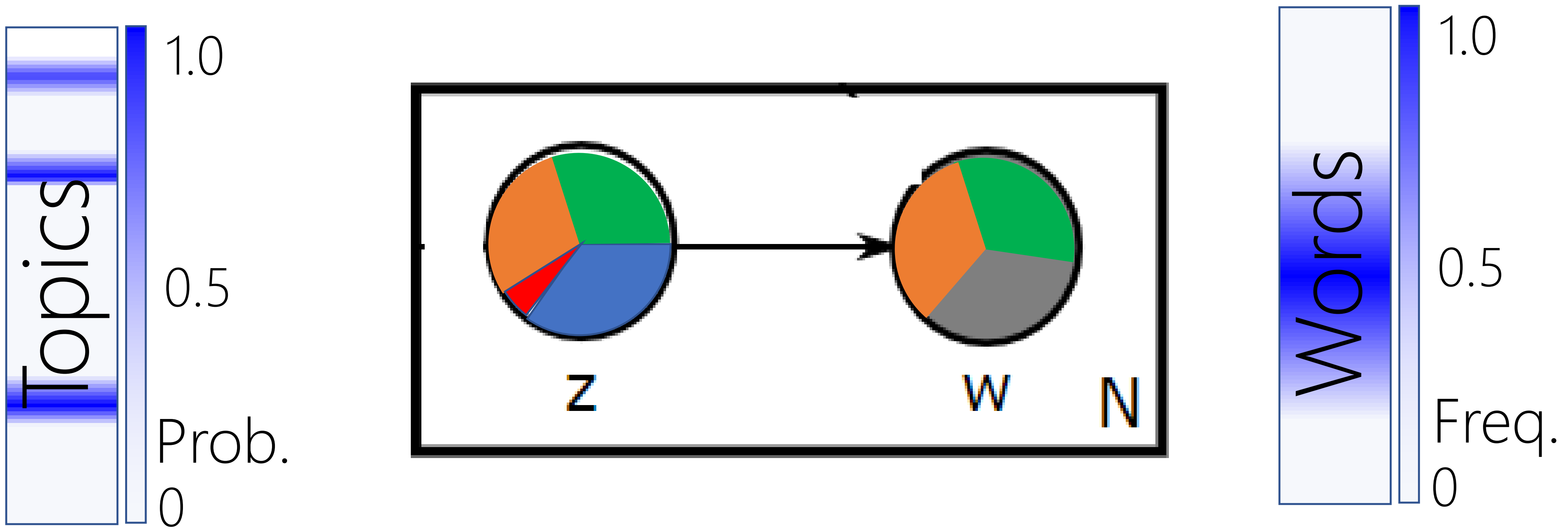
About what? *topic*?

Concert in soccer stadium for Nowruz and else.



Probabilistic Topic Models

- A document is about *all* topics but different distributions over topics
- A topic have *all* words but with different distributions over words



Journal of Machine Learning Research 3 (2003) 993-1022 Submitted 2/02; Published 1/03

Latent Dirichlet Allocation

David M. Blei BLEI@CS.BERKELEY.EDU
Computer Science Division
University of California
Berkeley, CA 94720, USA

Andrew Y. Ng ANG@CS.STANFORD.EDU
Computer Science Department
Stanford University
Stanford, CA 94305, USA

Michael I. Jordan JORDAN@CS.BERKELEY.EDU
Computer Science Division and Department of Statistics
University of California
Berkeley, CA 94720, USA

Editor: John Lafferty

Probabilistic Topic Models

Assumptions:

1) #topics = $\boldsymbol{k}$ → Z = {topics $\boldsymbol{z_i}$}, |Z| = $\boldsymbol{k}$, for the whole corpus.

Probabilistic Topic Models

Assumptions:

1) #topics = $\boldsymbol{k}$ → Z = {topics $\boldsymbol{z}_i$}, |Z| = $\boldsymbol{k}$, for the whole corpus.

2) Each doc is about all $\boldsymbol{k}$ topics but with different distributions

C={docs $\boldsymbol{d}_i$| $\boldsymbol{d}_i$=[$\boldsymbol{d}_{i1}, \boldsymbol{d}_{i2}, \ldots, \boldsymbol{d}_{ik}$]}$_{i=1}^{M}$ =[Matrix]$_{M\times k}$

Probabilistic Topic Models

Assumptions:

1) #topics = $\boldsymbol{k}$ → Z = {topics $\boldsymbol{z}_i$}, |Z| = $\boldsymbol{k}$, for the whole corpus.

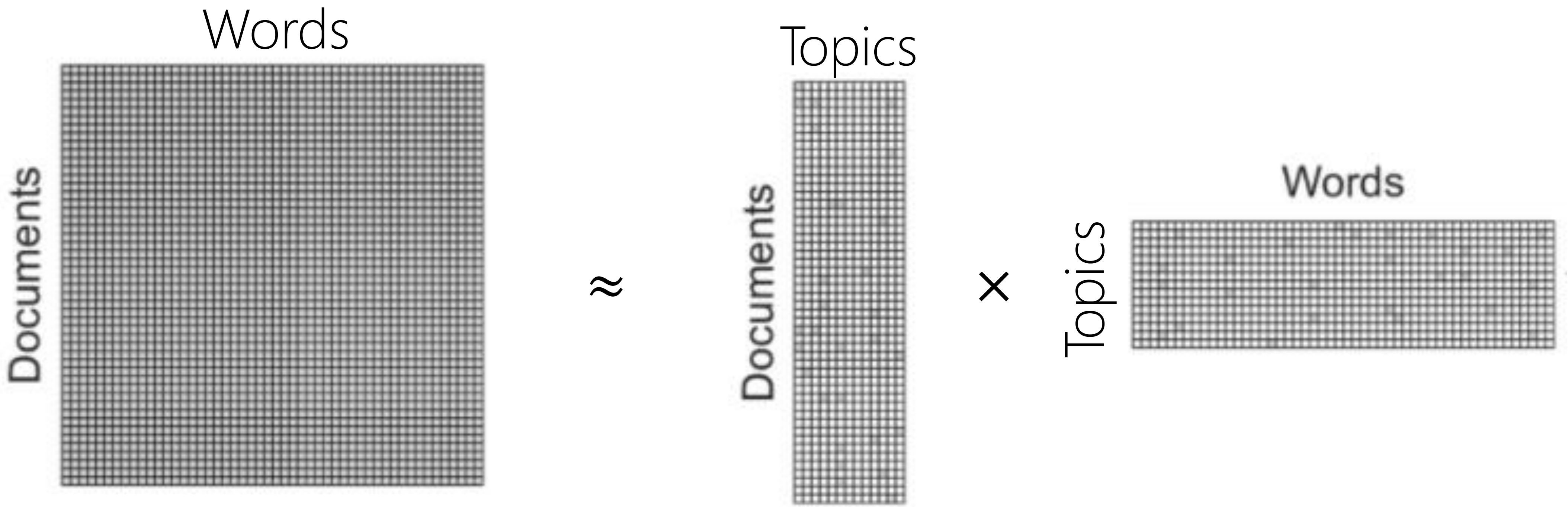
2) Each doc is about all $\boldsymbol{k}$ topics but with different distributions

C={docs $\boldsymbol{d_i}$| $\boldsymbol{d_i}$=$[\boldsymbol{d_{i1}}, \boldsymbol{d_{i2}}, \ldots, \boldsymbol{d_{ik}}]\}_{i=1}^{M}$ =$[\text{Matrix}]_{M\times k}$

3) Each topic $\boldsymbol{z}_i$ $_{1 \le i \le k}$, has all words but with different distributions

Z={topics $\boldsymbol{z}_i$ | $\boldsymbol{z_i}$=$[\boldsymbol{z_{i1}}, \boldsymbol{z_{i2}}, \ldots, \boldsymbol{z_{iV}}]\}_{\boldsymbol{i=1}}^{\boldsymbol{k}}$ =$[\text{Matrix}]_{k\times V}$

Probabilistic Topic Models



LDA can be seen as non-Negative Matrix Factorization

But learning algorithm is probabilistic!

Probabilistic Topic Models

Initializations:

1) For d_i in C={docs d_i| d_i=$[d_{i1}, d_{i2}, \ldots, d_{ik}]\}_{i=1}^{M}$ =$[\text{Matrix}]_{M\times k}$
   - $d_i = [d_{i1}, d_{i2}, \ldots, d_{ik}] \sim Dirichlet(\alpha = [\alpha_1, \alpha_2, \ldots, \alpha_k])$

$$f(x_1, \ldots, x_K; \alpha_1, \ldots, \alpha_K) = \frac{1}{\mathrm{B}(\boldsymbol{\alpha})} \prod_{i=1}^{K} x_i^{\alpha_i - 1}$$

where $\{x_k\}_{k=1}^{k=K}$ belong to the standard $K-1$ simplex, or in other words:

$$\boxed{\sum_{i=1}^{K} x_i = 1 \text{ and } x_i \geq 0 \text{ for all } i \in \{1, \ldots, K\}}$$

We've seen this before. where?

Probabilistic Topic Models

Initializations:

1) For d_i in C={docs d_i| d_i=$[d_{i1}, d_{i2}, \ldots, d_{ik}]\}_{i=1}^{M}$ =$[\text{Matrix}]_{M\times k}$
   - $d_i=[d_{i1}, d_{i2}, \ldots, d_{ik}] \sim Dirichlet(\alpha = [\alpha_1, \alpha_2, \ldots, \alpha_k])$

$$f(x_1, \ldots, x_K; \alpha_1, \ldots, \alpha_K) = \frac{1}{\mathrm{B}(\boldsymbol{\alpha})} \prod_{i=1}^{K} x_i^{\alpha_i - 1}$$

where $\{x_k\}_{k=1}^{k=K}$ belong to the standard $K-1$ simplex, or in other words:

$$\boxed{\sum_{i=1}^{K} x_i = 1 \text{ and } x_i \geq 0 \text{ for all } i \in \{1, \ldots, K\}}$$

Softmax! Transforming to vector of probs. But here we generate it!

Pólya's urn [edit source]

Consider an urn containing balls of K different colors. Initially, the urn contains α_1 balls of color 1, α_2 balls of color 2, and so on. Now perform N draws from the urn, where after each draw, the ball is placed back into the urn with an additional ball of the same color. In the limit as N approaches infinity, the proportions of different colored balls in the urn will be distributed as $\mathrm{Dir}(\alpha_1,\ldots,\alpha_K)$.[20]

For a formal proof, note that the proportions of the different colored balls form a bounded $[0,1]^K$-valued martingale, hence by the martingale convergence theorem, these proportions converge almost surely and in mean to a limiting random vector. To see that this limiting vector has the above Dirichlet distribution, check that all mixed moments agree.

Each draw from the urn modifies the probability of drawing a ball of any one color from the urn in the future. This modification diminishes with the number of draws, since the relative effect of adding a new ball to the urn diminishes as the urn accumulates increasing numbers of balls.

Probabilistic Topic Models

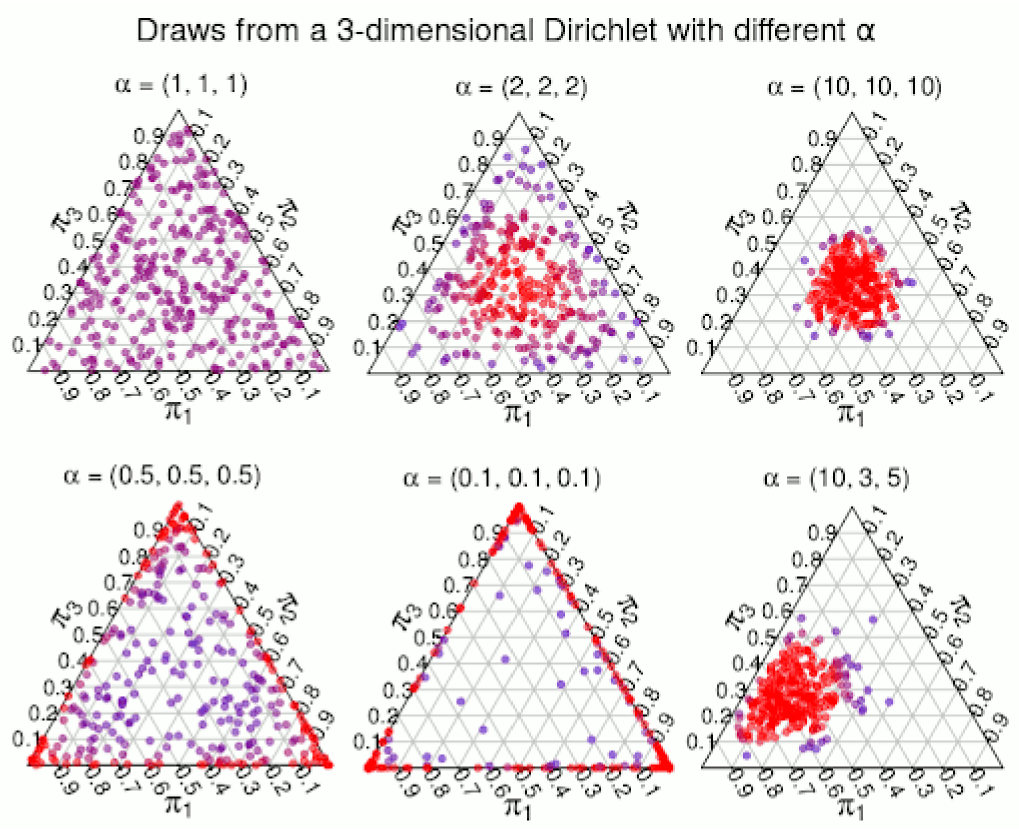
Initializations:

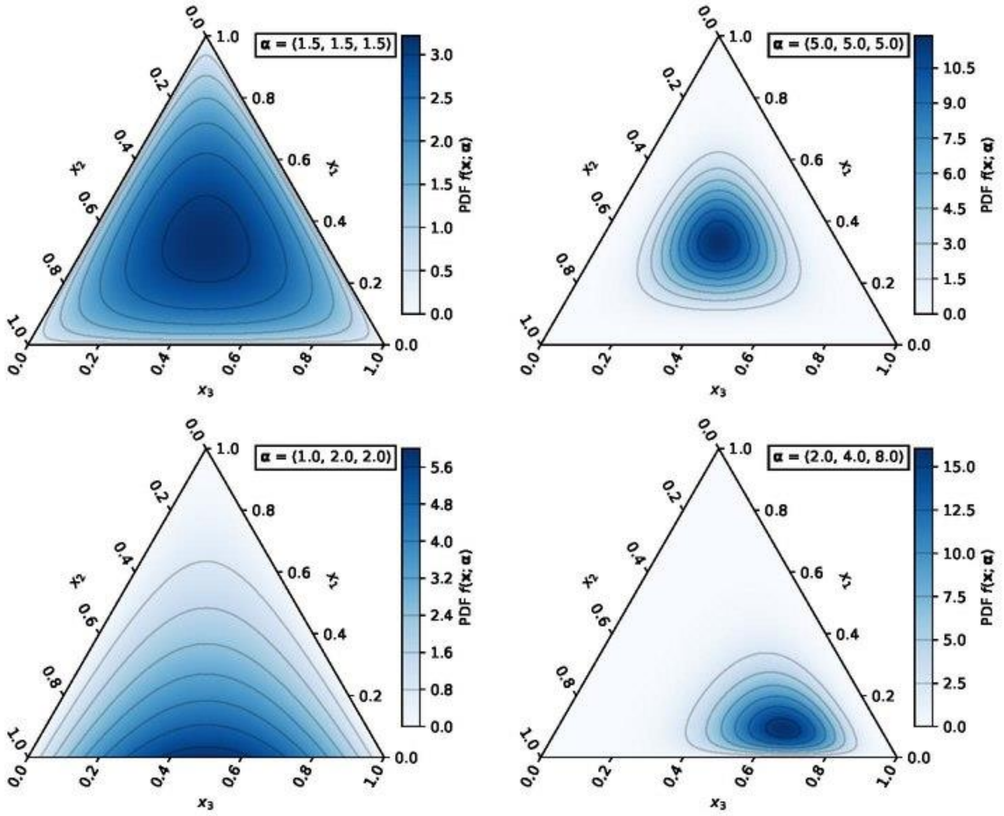
1) For d_i in C={docs d_i| d_i=[$d_{i1}, d_{i2}, \ldots, d_{ik}$]$\}_{i=1}^{M}$ =[Matrix]$_{M\times k}$
   - $d_i = [d_{i1}, d_{i2}, \ldots, d_{ik}] \sim Dirichlet(\alpha = [\alpha_1, \alpha_2, \ldots, \alpha_k])$

$$f(x_1, \ldots, x_K; \alpha_1, \ldots, \alpha_K) = \frac{1}{\mathrm{B}(\boldsymbol{\alpha})} \boxed{\prod_{i=1}^{K} x_i^{\alpha_i - 1}}$$

We've seen this before. where?

Probabilistic Topic Models



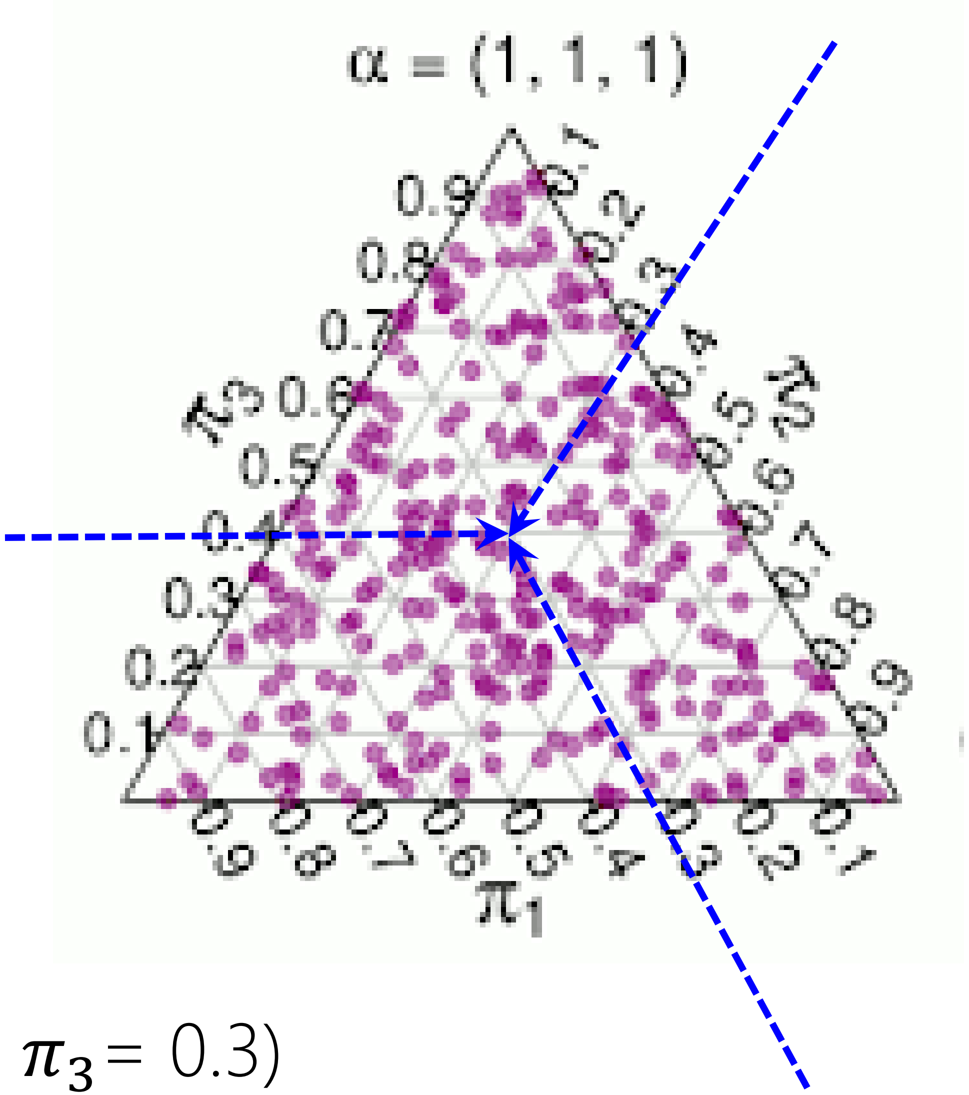


http://www.sumsar.net/blog/2015/04/the-non-parametric-bootstrap-as-a-bayesian-model/

https://en.wikipedia.org/wiki/Dirichlet_distribution

Probabilistic Topic Models

Random chance to π_1, π_2, π_3



(π_1 = 0.3, π_2 = 0.4, π_3 = 0.3)

Simplex

From Wikipedia, the free encyclopedia

For other uses, see Simplex (disambiguation).

The four simplexes which can be fully represented in 3D space.

In geometry, a **simplex** (plural: **simplexes** or **simplices**) is a generalization of the notion of a triangle or tetrahedron to arbitrary dimensions. The simplex is so-named because it represents the simplest possible polytope in any given space.

For example,

- a 0-simplex is a point,
- a 1-simplex is a line segment,
- a 2-simplex is a triangle,
- a 3-simplex is a tetrahedron,
- a 4-simplex is a 5-cell.

Specifically, a ***k*-simplex** is a k-dimensional polytope which is the convex hull of its $k + 1$ vertices. More formally, suppose the $k + 1$ points $u_0, \dots, u_k \in \mathbb{R}^k$ are affinely independent, which means $u_1 - u_0, \dots, u_k - u_0$ are linearly independent. Then, the simplex determined by them is the set of points

$$C = \left\{ \theta_0 u_0 + \cdots + \theta_k u_k \;\middle|\; \sum_{i=0}^{k} \theta_i = 1 \text{ and } \theta_i \geq 0 \text{ for } i = 0, \dots, k \right\}.$$

A **regular simplex**[1] is a simplex that is also a regular polytope. A regular k-simplex may be constructed from a regular $(k - 1)$-simplex by connecting a new vertex to all original vertices by the common edge length.

The **standard simplex** or **probability simplex** [2] is the simplex whose vertices are the k standard unit vectors and the origin, or

$$\{x \in \mathbb{R}^k : x_0 + \cdots + x_{k-1} = 1, x_i \geq 0 \text{ for } i = 0, \dots, k-1\}.$$

Probabilistic Topic Models

Initializations:

1) For d_i in C={docs d_i| d_i=[$d_{i1}, d_{i2}, \dots, d_{ik}$]}$_{i=1}^{M}$ =[Matrix]$_{M\times k}$
   - d_i=[$d_{i1}, d_{i2}, \dots, d_{ik}$] $\sim$ $Dirichlet(\alpha = [\alpha_1, \alpha_2, \dots, \alpha_k])$

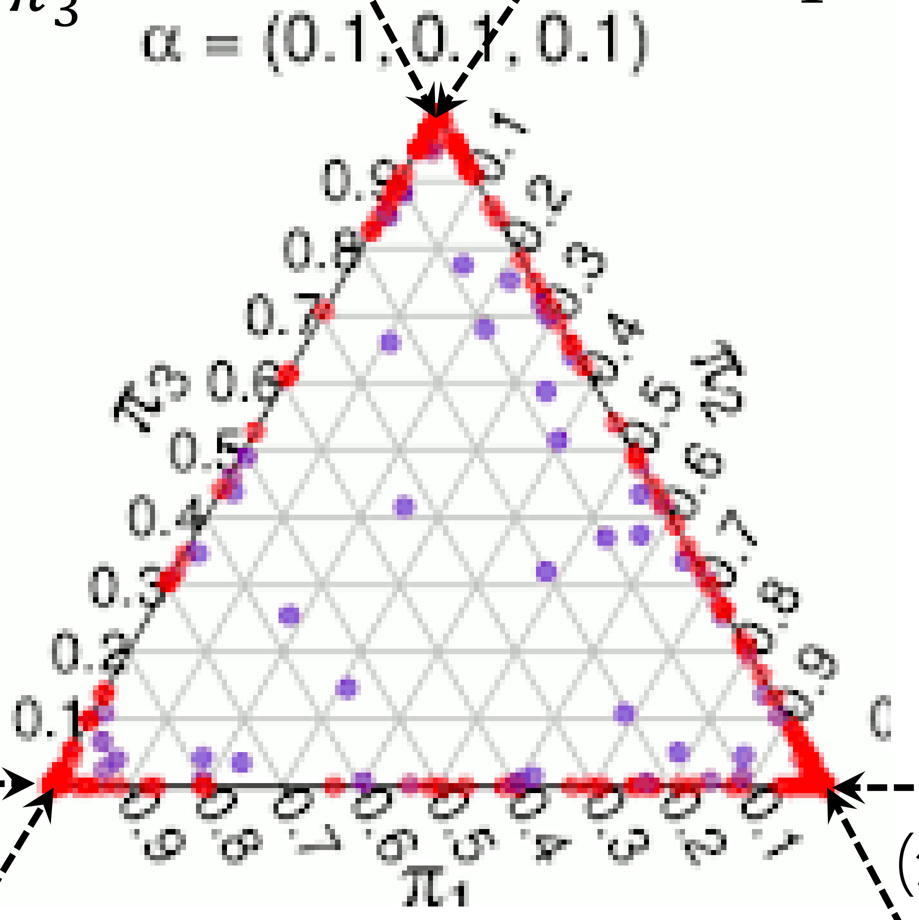
$$f(x_1, \ldots, x_K; \alpha_1, \ldots, \alpha_K) = \frac{1}{\mathrm{B}(\boldsymbol{\alpha})} \prod_{i=1}^{K} x_i^{\boxed{\alpha_i}-1}$$

Prior weight of topic k in a document
- Initially the same for all topics
- Normally $\alpha_{1\leq j\leq k} < 1$, e.g. 0.1,
- Prefer sparse topic distributions, i.e., few topics per document

Probabilistic Topic Models

Sparse chance to π_1, π_2, π_3

(π_1= 0.0, π_2= 0.0, π_3= 1.0)



(π_1= 1.0, π_2= 0.0, π_3= 0.0)

(π_1= 0.0, π_2= 1.0, π_3= 0.0)

Probabilistic Topic Models

Initializations:

2) For z_i in Z={topics z_i| z_i=[$z_{i1}, z_{i2}, \dots, z_{iV}$]}$_{i=1}^{k}$ =[Matrix]$_{k\times V}$
- $z_i=[z_{i1}, z_{i2}, \dots, z_{iV}] \sim Dirichlet(\beta=[\beta_1, \beta_2, \dots, \beta_V])$

$$f(x_1, \dots, x_V; \beta_1, \dots, \beta_V) = \frac{1}{\mathrm{B}(\beta)} \prod_{i=1}^{K} x_i^{\boxed{\beta_i}-1}$$

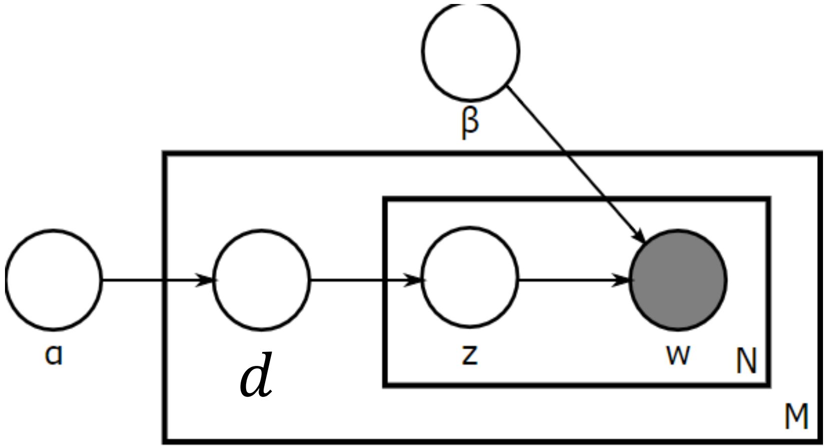
Prior weight of topic k in a document
- Initially the same for all topics
- Normally $\beta_{1\leq j\leq V}$ << 1, e.g. 0.001,
- Very sparse word distributions, i.e., few words per topics

Probabilistic Topic Models

Generative Process:

$$p(\theta, \mathbf{z}, \mathbf{w} \mid \alpha, \beta) = p(\theta \mid \alpha) \prod_{n=1}^{N} p(z_n \mid \theta) p(w_n \mid z_n, \beta),$$

Plate Notation
Graphical Model

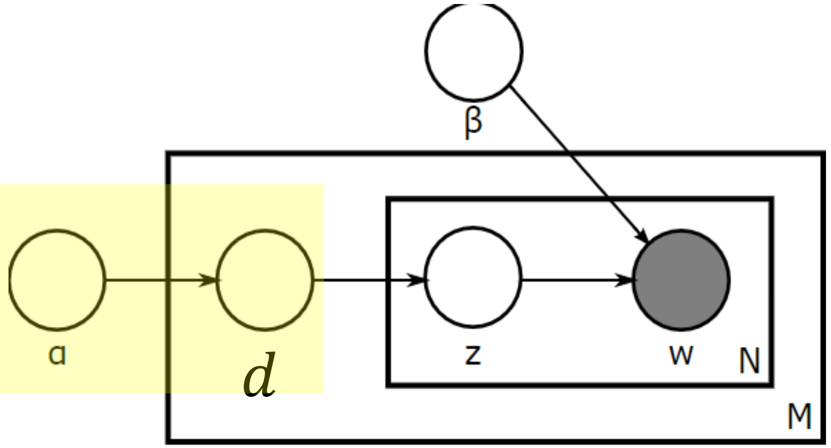


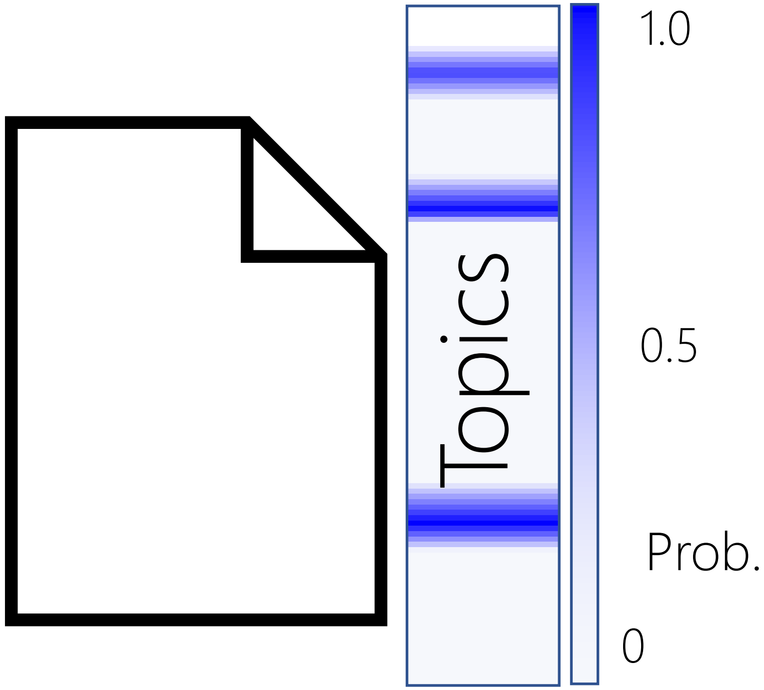
Probabilistic Topic Models

Generative Process:

$$p(\theta, \mathbf{z}, \mathbf{w} \mid \alpha, \beta) = p(\theta \mid \alpha) \prod_{n=1}^{N} p(z_n \mid \theta) p(w_n \mid z_n, \beta),$$

Select the topic distribution for the doc



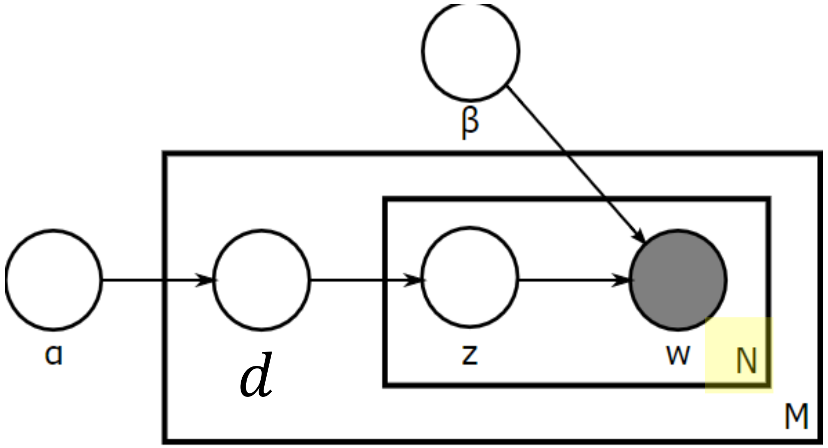


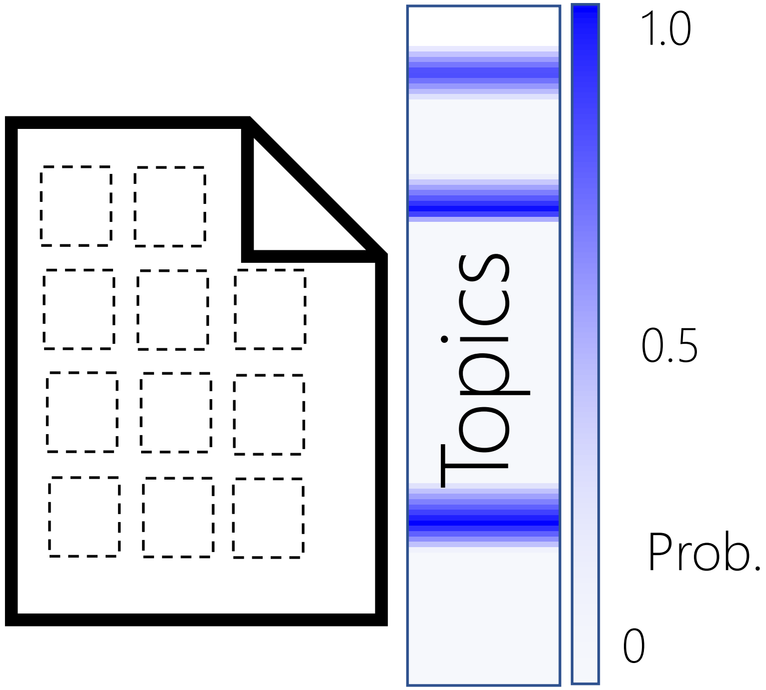
Probabilistic Topic Models

Generative Process:

$$p(\theta, \mathbf{z}, \mathbf{w} \mid \alpha, \beta) = p(\theta \mid \alpha) \prod_{n=1}^{N} p(z_n \mid \theta) p(w_n \mid z_n, \beta),$$

For each word in the doc



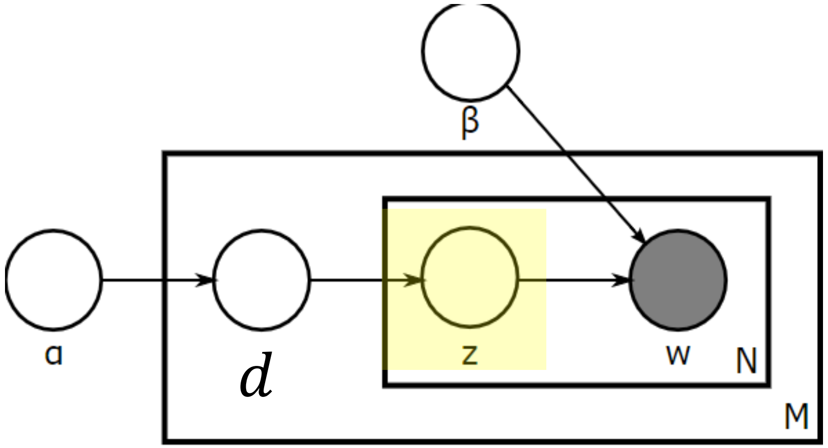


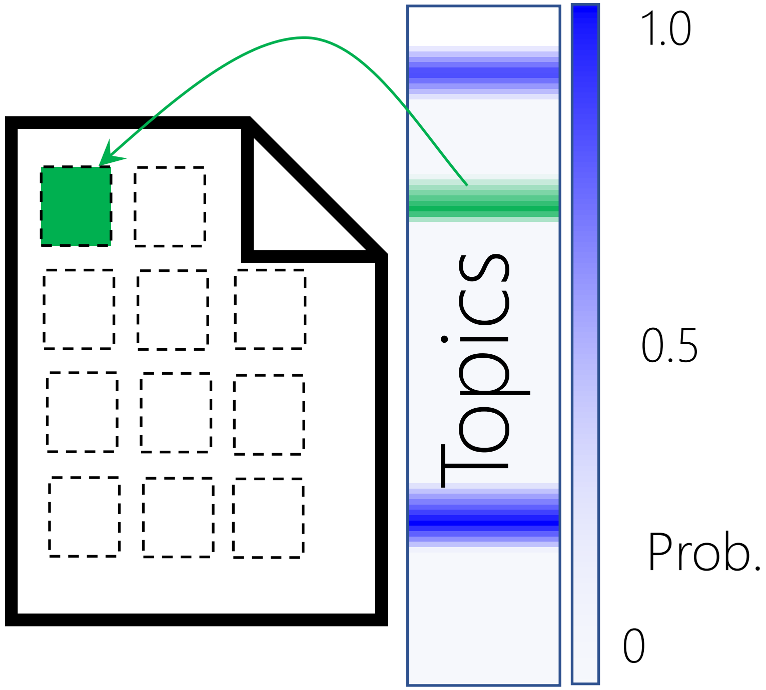
Probabilistic Topic Models

Generative Process:

$$p(\theta, \mathbf{z}, \mathbf{w} \mid \alpha, \beta) = p(\theta \mid \alpha) \prod_{n=1}^{N} p(z_n \mid \theta) p(w_n \mid z_n, \beta),$$

Select a topic from the topic distribution of the doc



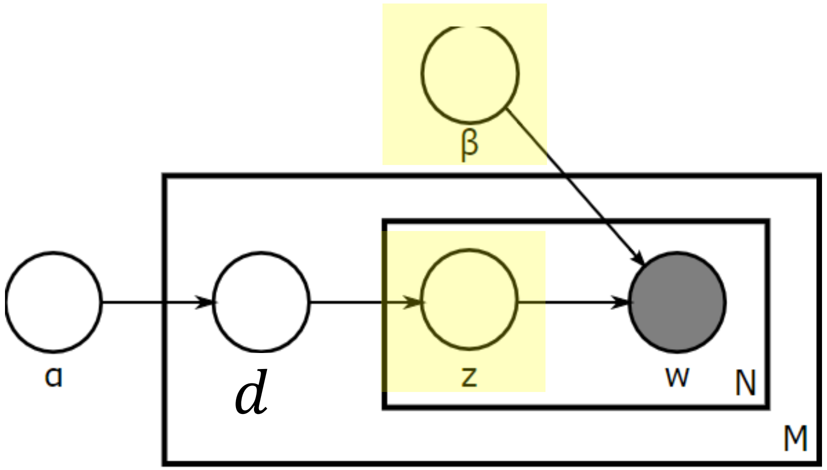


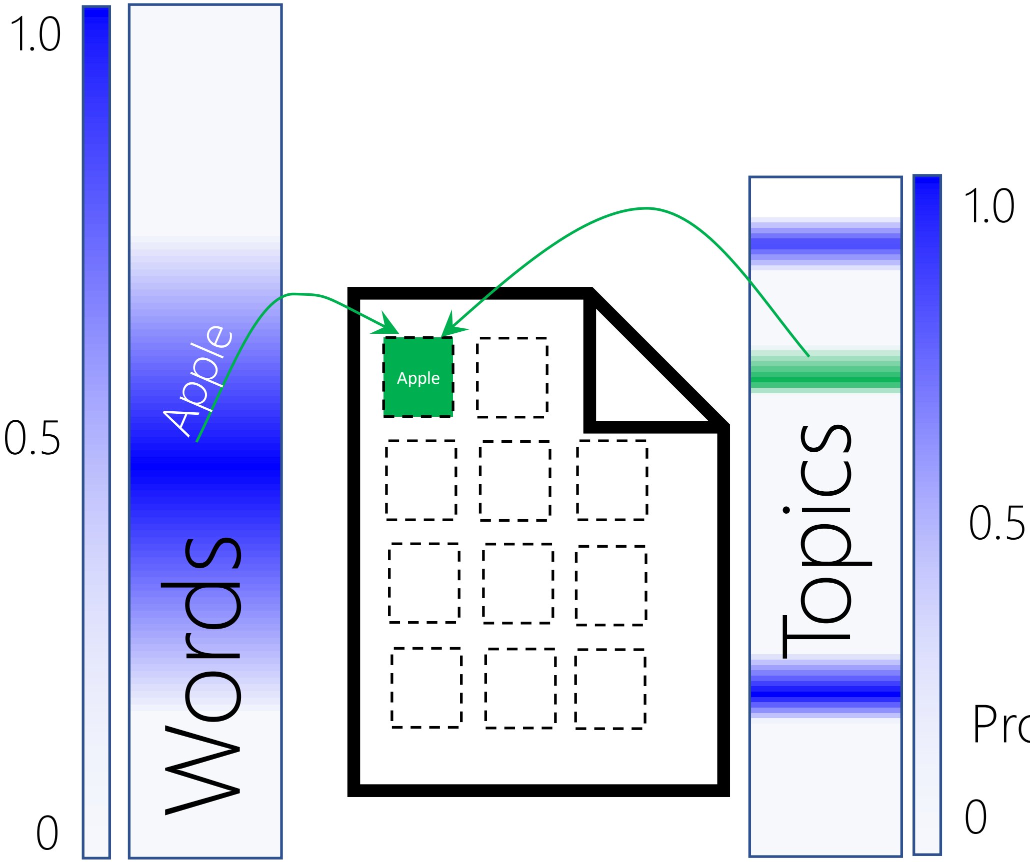
Probabilistic Topic Models

Generative Process:

$$p(\theta, \mathbf{z}, \mathbf{w} \mid \alpha, \beta) = p(\theta \mid \alpha) \prod_{n=1}^{N} p(z_n \mid \theta) p(w_n \mid z_n, \beta),$$

Select the word from the topic distribution





Probabilistic Topic Models

Generative Process:

For $\boldsymbol{d_i}$ in C

For $\boldsymbol{j = 0{:}\ N}$

$\boldsymbol{z}$= Choose a topic $\boldsymbol{z_i}$ in Z based on $[\boldsymbol{d_{i1}}, \boldsymbol{d_{i2}}, \ldots, \boldsymbol{d_{ik}}]$

$\boldsymbol{w}$ = Choose a word $\boldsymbol{w_j}$ in V based on $\boldsymbol{z}$=$[\boldsymbol{z_1}, \boldsymbol{z_2}, \ldots, \boldsymbol{z_V}]$

Probabilistic Topic Models

Generative Process:

For d_i in C

For $j = 0: N$

z= Choose a topic z_i in Z based on $[d_{i1}, d_{i2}, \ldots, d_{ik}]$

w = Choose a word w_j in V based on $z=[z_1, z_2, \ldots, z_V]$

$$p(\theta, \mathbf{z}, \mathbf{w} \mid \alpha, \beta) = p(\theta \mid \alpha) \prod_{n=1}^{N} p(z_n \mid \theta) p(w_n \mid z_n, \beta),$$

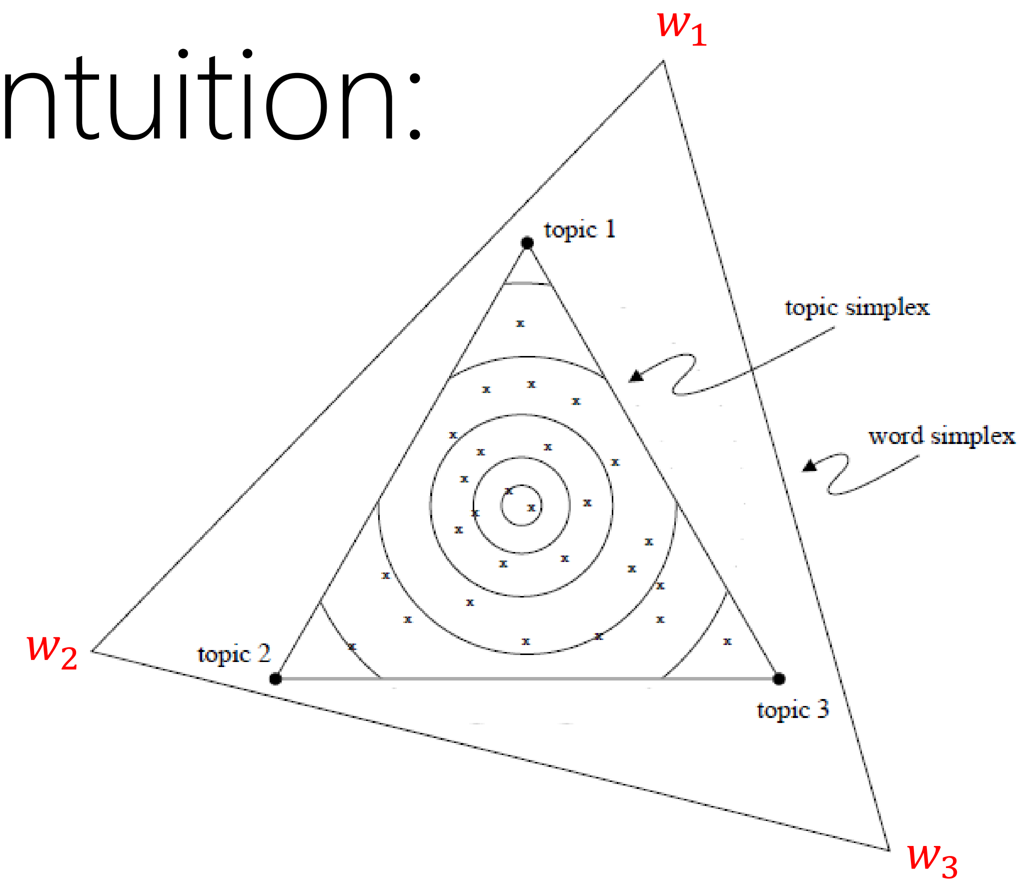
- Bernoulli distribution:

1 out of two mutually exclusive options = success (p) or failure ($q = 1 - p$)

- Generalized Bernoulli (Categorical) distribution = 1 out of N mutually exclusive options p_i

Probabilistic Topic Models

Geometric Intuition:



For d_i in C

For $j = 0{:}\ N$

z= Choose a topic z_i in Z based on $[d_{i1}, d_{i2}, \ldots, d_{ik}]$

w = Choose a word w_j in V based on z=$[z_1, z_2, \ldots, z_V]$

Probabilistic Topic Models

Optimization

Generated Docs ↔ Observed Docs

$$\ell(\alpha,\beta) = \sum_{d=1}^{M} \log p(\mathbf{w}_d \,|\, \alpha, \beta).$$

For d_i in C

For $j = 0 : N$

z = Choose a topic z_i in Z based on $\boxed{d_i = [d_{i1}, d_{i2}, \ldots, d_{ik}]}$

w = Choose a word w_j in V based on $\boxed{z = [z_1, z_2, \ldots, z_V]}$

Probabilistic Topic Models

Question: what's the difference?

For d_i in C

z= Choose a topic z_i in Z based on $[d_{i1}, d_{i2}, \ldots, d_{ik}]$

For $j = \mathbf{0}: N$

w = Choose a word w_j in V based on z=$[z_1, z_2, \ldots, z_V]$

Probabilistic Topic Models

Question: what's the difference?

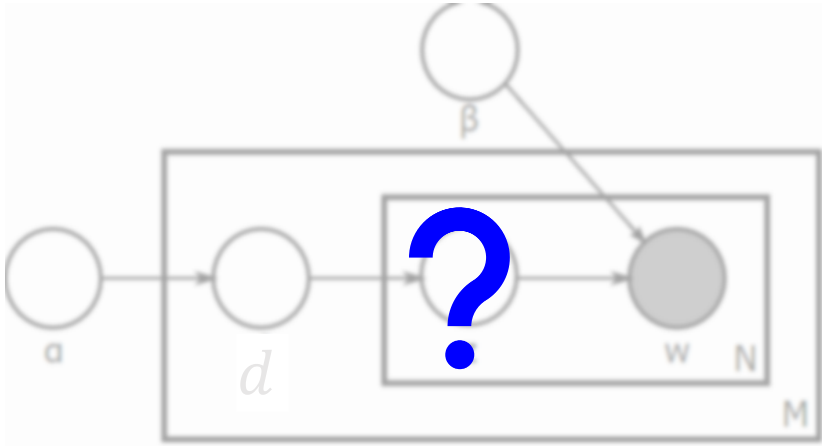
For d_i in C

z= Choose a topic z_i in Z based on $[d_{i1}, d_{i2}, \ldots, d_{ik}]$

For $j = 0{:}\ N$

w = Choose a word w_j in V based on z=$[z_1, z_2, \ldots, z_V]$

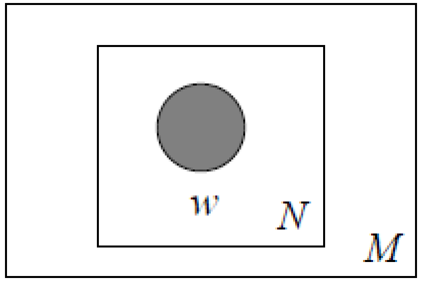
$$p(\theta, \mathbf{z}, \mathbf{w} \mid \alpha, \beta) = p(\theta \mid \alpha) p(z_n \mid \theta) \prod_{n=1}^{N} p(w_n \mid z_n, \beta)$$



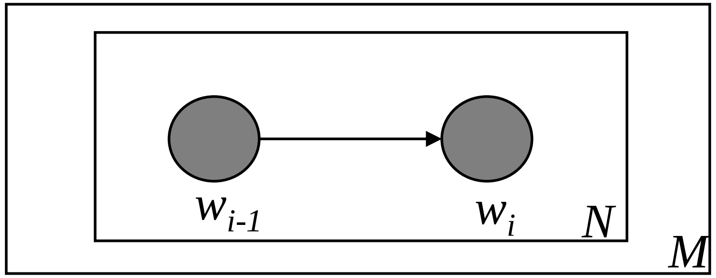
Probabilistic Topic Models

LDA vs. or as a n-gram LM

$$p(\mathbf{w}) = \prod_{n=1}^{N} p(w_n).$$



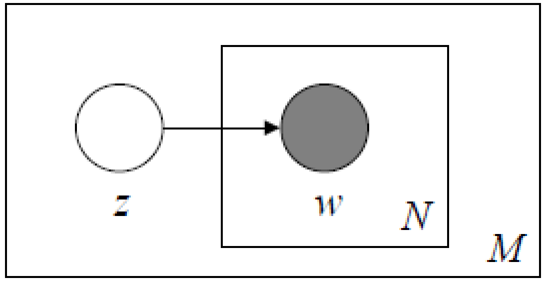
(a) unigram



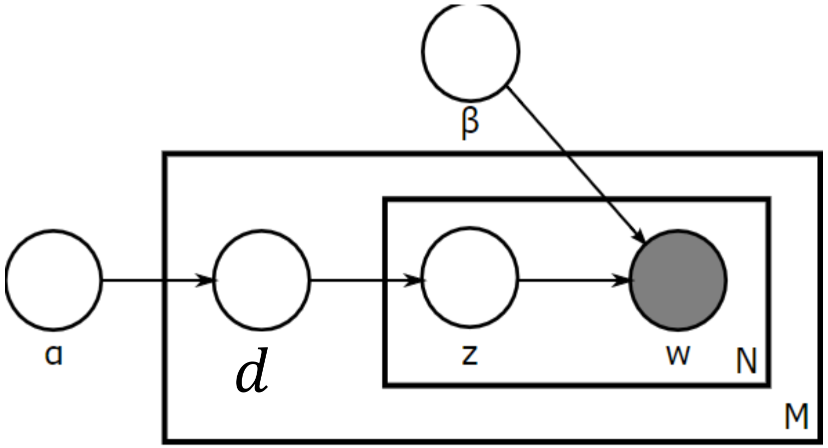
(b) bigram

$p(\mathbf{w})$ = ?

$$p(\mathbf{w}) = \sum_z p(z) \prod_{n=1}^{N} p(w_n \mid z).$$



(c) mixture of unigrams



Probabilistic Topic Models: API

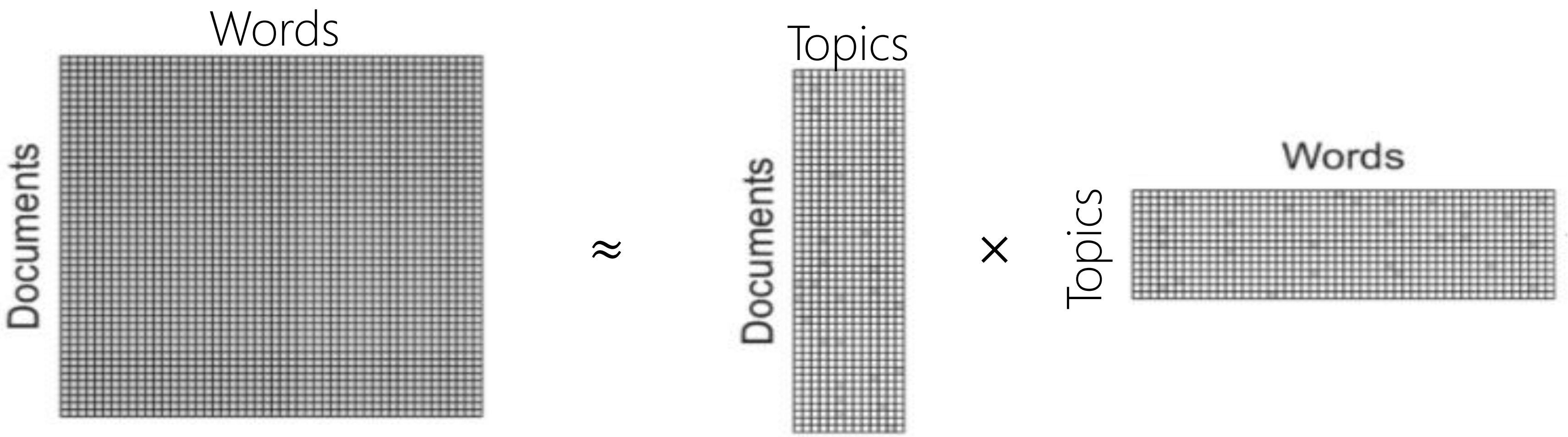
- MALLET (http://mallet.cs.umass.edu/topics.php)
- Gensim (https://radimrehurek.com/gensim/)
  - Gensim wrapper for MALLET

Probabilistic Topic Models: Applications

- Corpus Summarization
- Document Clustering / Classification
- Document Summarization
- User Modeling

Probabilistic Topic Models: Applications

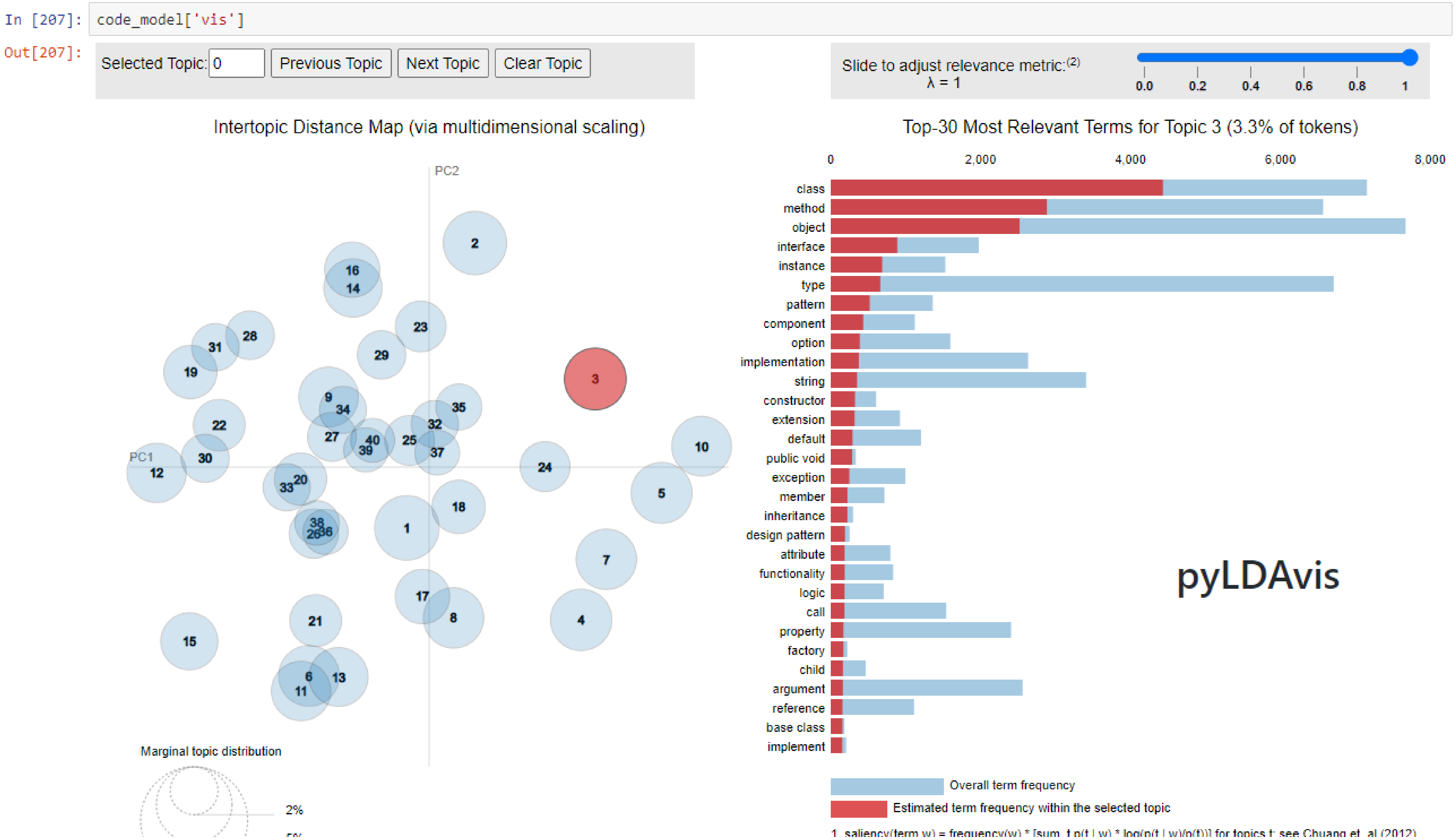
- Corpus Summarization
  - Corpus → Topics → Top-k Words



Probabilistic Topic Models: Applications

1	Week	NASA	Space	Science	Scientist	.com	Phys.org	Planet	Community	Time
	0.085	0.054	0.051	0.048	0.032	0.031	0.026	0.02	0.019	0.018
2	Message	Email	Day	England	Time	France	Manchester	Assassination	Venezuela	Game
	0.041	0.038	0.037	0.035	0.026	0.024	0.024	0.024	0.02	0.02
3	Quebec	Pará	Latin	Delaware	System	Mexico	Sweden	Travel	Como	Louisiana
	0.079	0.043	0.036	0.027	0.026	0.026	0.021	0.021	0.021	0.017
4	DONT	Ur	Fuck	Internatio nal	Music	RT!	God	Human	Airport	Feces
	0.039	0.033	0.026	0.019	0.018	0.016	0.015	0.015	0.014	0.014
5	Hootsuite	Party	Ll	Sea	Light-year	.co	Life	People	Shit	Human
	0.077	0.067	0.043	0.043	0.036	0.024	0.022	0.017	0.015	0.015
6	Street	China	Business	Reuters	Stock	Ireland	Bank	Reut	Economics	0
	0.127	0.042	0.038	0.031	0.027	0.027	0.016	0.016	0.015	0.014
7	Mail	Sunday	Time	People	Love	Person	Hope	Thought	Nice	Christmas
	0.057	0.05	0.048	0.042	0.029	0.026	0.023	0.023	0.018	0.014
8	MSNBC	News	White	Committe e	World	Fox	Hootsuite	President	Party	PBS
	0.079	0.055	0.036	0.033	0.029	0.029	0.026	0.024	0.023	0.019
9	BBC	News	World	Sky	Ireland	Death	TGR	Murder	.uk	PH
	0.317	0.261	0.172	0.026	0.006	0.006	0.006	0.006	0.005	0.005
10	Foursqua re	Wales	Facebook	Engagem ent	Time	Wedding	International	Family	Shanghai	CNNMon ey
	0.077	0.043	0.036	0.035	0.033	0.031	0.021	0.019	0.017	0.016

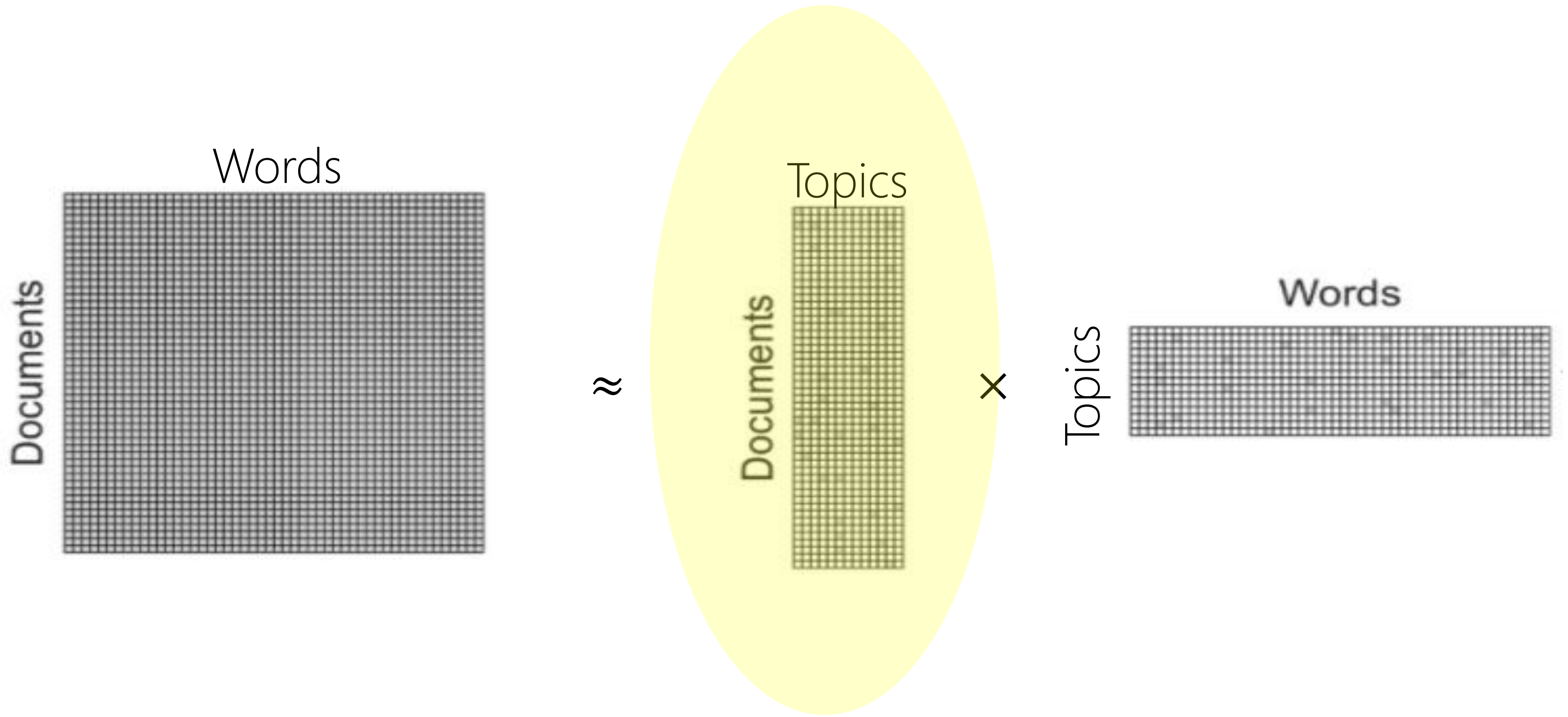
Configuration:	
# of Topics	40
Dataset	1/14 zabel
Representation	USER
Topic Detector	MALLET
Preprocessing	TagME
# of Users	1619
# of Tweets	167572
Filter	Yes
TagME threshold	0.01



https://nbviewer.jupyter.org/github/bmabey/hacker_news_topic_modelling/blob/master/HN%20Topic%20Model%20Talk.ipynb

Probabilistic Topic Models: Applications

- Document Clustering / Classification



Probabilistic Topic Models: Applications

- Document Clustering / Classification
  - Each document is a feature vector of topics
  - Similar documents have similar vectors/topics


```
0 MyNewsData\10085.txt 4 0.140485312899106 5 0.08556832694763729 48 0.05491698595146871 25 0.033205619412515965 37 0.02937420178799489 38 0.0280970625798212 31 0.022988505747126436 42 0.021711366538952746 0 0.020434227330779056
28 0.019157088122605363 27 0.019157088122605363 26 0.017879948914431672 15 0.017879948914431672 30 0.016602809706257982 12 0.016602809706257982 3 0.016602809706257982 1 0.016602809706257982 49 0.01532567049808429 46
0.01532567049808429 35 0.01532567049808429 22 0.01532567049808429 16 0.01532567049808429 40 0.0140485312899106 29 0.0140485312899106 19 0.0140485312899106 17 0.0140485312899106 7 0.0140485312899106 44 0.01277139208173691 39
0.01277139208173691 36 0.01277139208173691 33 0.01277139208173691 24 0.01277139208173691 21 0.01277139208173691 20 0.01277139208173691 18 0.01277139208173691 10 0.01277139208173691 47 0.011494252873563218 45 0.011494252873563218
43 0.011494252873563218 41 0.011494252873563218 34 0.011494252873563218 32 0.011494252873563218 23 0.011494252873563218 14 0.011494252873563218 13 0.011494252873563218 11 0.011494252873563218 9 0.011494252873563218 8
0.011494252873563218 6 0.011494252873563218 2 0.011494252873563218
1 MyNewsData\10711.txt 48 0.15815295815295816 4 0.13015873015873017 5 0.12005772005772006 28 0.10851370851370852 15 0.09437229437229437 11 0.05656565656565657 2 0.043001443001443 38 0.03261183261183261 17 0.030014430014430014 42
0.013564213564213565 7 0.012121212121212121 25 0.011255411255411256 21 0.008369408369408369 46 0.00808080808080808 19 0.00808080808080808 49 0.007792207792207792 16 0.006349206349206349 12 0.006349206349206349 37
0.006060606060606061 30 0.006060606060606061 3 0.006060606060606061 31 0.005772005772005772 10 0.005772005772005772 0 0.005772005772005772 45 0.005483405483405483 36 0.005483405483405483 47 0.005194805194805195 44
0.005194805194805195 39 0.005194805194805195 22 0.005194805194805195 20 0.005194805194805195 33 0.004906204906204906 32 0.004906204906204906 43 0.004329004329004329 34 0.004329004329004329 26 0.004329004329004329 6
0.004329004329004329 41 0.00404040404040404 35 0.00404040404040404 24 0.00404040404040404 1 0.003751803751803752 29 0.003463203463203463 27 0.003463203463203463 23 0.003463203463203463 14 0.003463203463203463 40
0.0031746031746031746 13 0.0031746031746031746 8 0.0031746031746031746 18 0.002886002886002886 9 0.002886002886002886
2 MyNewsData\10856.txt 15 0.31922398589065254 5 0.2177150695669214 38 0.06094454242602391 4 0.0536939055457574 28 0.049578679208308836 48 0.036449147560258674 46 0.028610621203213796 14 0.017244757985498727 26
0.016852831667646482 30 0.016068979031941995 6 0.016068979031941995 11 0.013325494806976288 37 0.009994121105232217 17 0.00725063688026651 36 0.006858710562414266 8 0.006466784244562022 35 0.006270821085635900 5 3
0.006270821085635900 5 49 0.005682931608857535 39 0.005486968449931412 5 16 0.005291005291005291 19 0.004899078973153047 44 0.004507152655300804 23 0.004507152655300804 42 0.004311189496374682 45 0.004115226337448560 9
0.004115226337448560 29 0.003919263178522437 7 0.003919263178522437 25 0.003723300019596316 34 0.003527336860670194 31 0.003527336860670194 12 0.003527336860670194 10 0.003527336860670194 43 0.003331373701744072 18
0.003331373701744072 24 0.003135410542817950 2 22 0.003135410542817950 2 13 0.002939447383891828 5 40 0.002743484224965706 2 33 0.002743484224965706 2 0 0.002743484224965706 2 41 0.002547521066039584 5 21 0.002547521066039584 5 20
0.002547521066039584 5 2 0.002351557907113462 7 47 0.002155594748187341 32 0.002155594748187341 1 0.002155594748187341 27 0.001959631589261218 7
```

Probabilistic Topic Models: Applications

- Document Summary

NEW	MILLION	CHILDREN	SCHOOL
FILM	TAX	WOMEN	STUDENTS
SHOW	PROGRAM	PEOPLE	SCHOOLS
MUSIC	BUDGET	CHILD	EDUCATION
MOVIE	BILLION	YEARS	TEACHERS
PLAY	FEDERAL	FAMILIES	HIGH
MUSICAL	YEAR	WORK	PUBLIC
BEST	SPENDING	PARENTS	TEACHER
ACTOR	NEW	SAYS	BENNETT
FIRST	STATE	FAMILY	MANIGAT
YORK	PLAN	WELFARE	NAMPHY
OPERA	MONEY	MEN	STATE
THEATER	PROGRAMS	PERCENT	PRESIDENT
ACTRESS	GOVERNMENT	CARE	ELEMENTARY
LOVE	CONGRESS	LIFE	HAITI

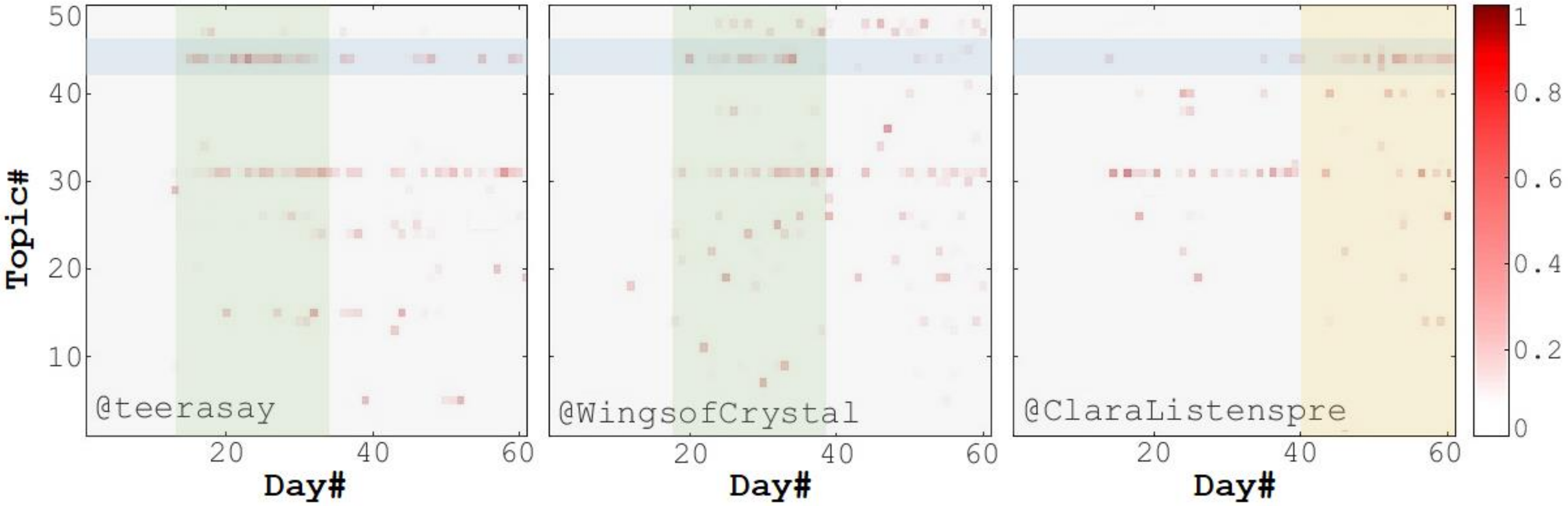
NEW	MILLION
FILM	TAX
SHOW	PROGRAM
MUSIC	BUDGET
MOVIE	BILLION
PLAY	FEDERAL

The William Randolph Hearst Foundation will give $1.25 million to Lincoln Center, Metropolitan Opera Co., New York Philharmonic and Juilliard School. "Our board felt that we had a real opportunity to make a mark on the future of the performing arts with these grants an act every bit as important as our traditional areas of support in health, medical research, education and the social services," Hearst Foundation President Randolph A. Hearst said Monday in announcing the grants. Lincoln Center's share will be $200,000 for its new building, which will house young artists and provide new public facilities. The Metropolitan Opera Co. and New York Philharmonic will receive $400,000 each. The Juilliard School, where music and the performing arts are taught, will get $250,000. The Hearst Foundation, a leading supporter of the Lincoln Center Consolidated Corporate Fund, will make its usual annual $100,000 donation, too.

The William Randolph Hearst Foundation will give $1.25 million to Lincoln Center, Metropolitan Opera Co., New York Philharmonic and Juilliard School. "Our board felt that we had a real opportunity to make a mark on the future of the performing arts with these grants an act every bit as important as our traditional areas of support in health, medical research, education and the social services," Hearst Foundation President Randolph A. Hearst said Monday in announcing the grants. Lincoln Center's share will be $200,000 for its new building, which will house young artists and provide new public facilities. The Metropolitan Opera Co. and New York Philharmonic will receive $400,000 each. The Juilliard School, where music and the performing arts are taught, will get $250,000. The Hearst Foundation, a leading supporter of the Lincoln Center Consolidated Corporate Fund, will make its usual annual $100,000 donation, too.

Probabilistic Topic Models: Applications

- User Modeling (will be discussed more …)



Distributed Memory Model of Paragraph Vectors (PV-DM)

aka Doc2Vec