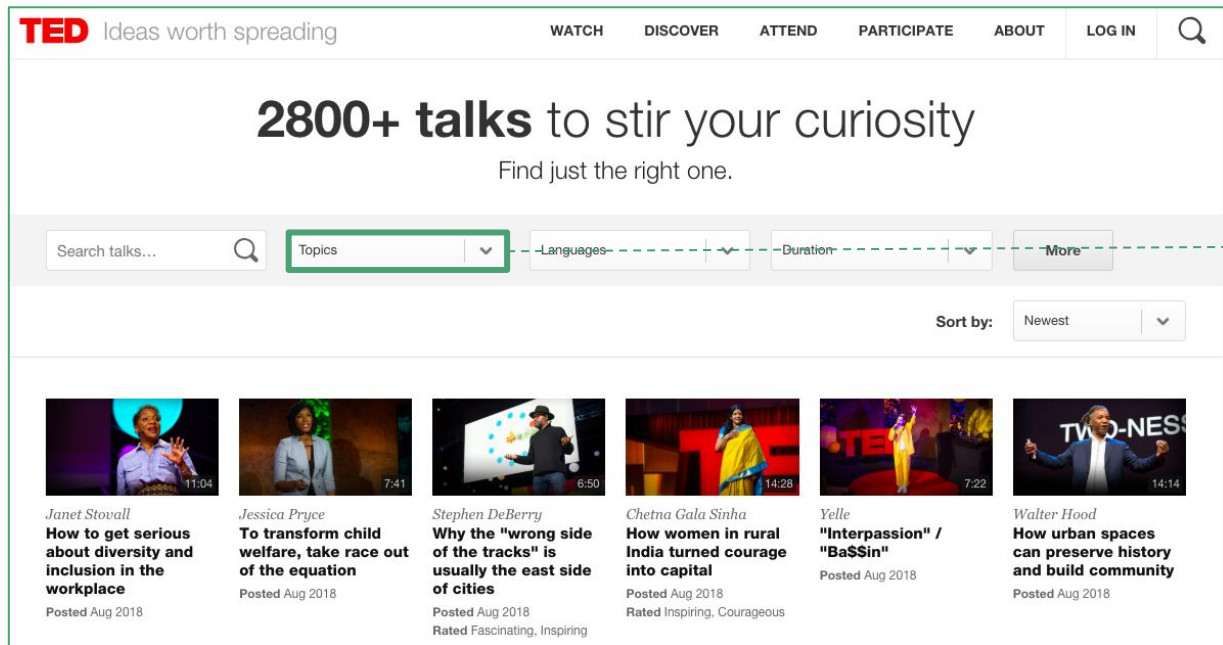


Predicting TED Talk Topics

Laura Borton
Project Fletcher

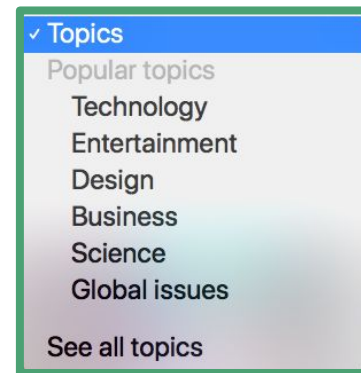
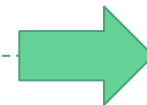


What is the Purpose?



The screenshot shows the TED website's search interface. At the top, the TED logo is followed by the tagline "Ideas worth spreading". Navigation links include WATCH, DISCOVER, ATTEND, PARTICIPATE, ABOUT, and LOG IN. A search bar contains the text "2800+ talks to stir your curiosity" and "Find just the right one." Below the search bar, there are filters for "Topics", "Languages", "Duration", and "More". The "Topics" filter is highlighted with a green box. To the right of the filters, there is a "Sort by:" dropdown menu set to "Newest". Below the filters, a list of talks is displayed, each with a thumbnail image, title, speaker name, and duration. The talks are:

- Janet Stovall**: How to get serious about diversity and inclusion in the workplace (11:04)
- Jessica Pryce**: To transform child welfare, take race out of the equation (7:41)
- Stephen DeBerry**: Why the "wrong side of the tracks" is usually the east side of cities (6:50)
- Chetna Gala Sinha**: How women in rural India turned courage into capital (14:28)
- Yelle**: "Interpassion" / "Ba\$\$in" (7:22)
- Walter Hood**: How urban spaces can preserve history and build community (14:14)




A list of popular topics from the TED website, enclosed in a green border. The list is titled "✓ Topics" and includes the following items:

- ✓ Topics
- Popular topics
- Technology
- Entertainment
- Design
- Business
- Science
- Global issues
- See all topics

Group talks into 6 popular topics

Data



The image shows a YouTube video player interface. The video is titled "Teach every child about food" by Jamie Oliver | TED2010. The video player shows a man (Jamie Oliver) speaking on stage, holding a piece of paper. The video has a play button in the center. The video player includes a progress bar, volume control, and a settings icon. The video is 21:47 long. The video player includes a share button, an add to list button, a like button, and a rate button. The video player includes a details section with tabs for "Details", "Transcript", and "Comments (1146)". The "Details" tab is selected. The details section includes a description of the talk and a list of related tags.

Jamie Oliver | TED2010

Teach every child about food

21:47

Details About the talk

Transcript 50 languages

Comments (1146) Join the conversation

Sharing powerful stories from his anti-obesity project in Huntington, West Virginia -- and a shocking image of the sugar we eat -- TED Prize winner Jamie Oliver makes the case for an all-out assault on our ignorance of food.

Related tags

- [Business](#)
- [Education](#)
- [Food](#)
- [Global Issues](#)
- [Health](#)
- [Obesity](#)

Years:

June 2006 - March 2018

Talks:

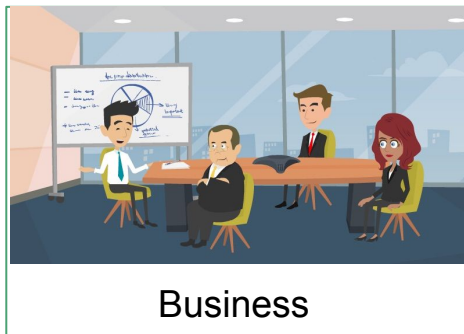
Over 2500

Relevant Features:

- Title
- Summary
- Transcript
- Tags

How well can a model predict?

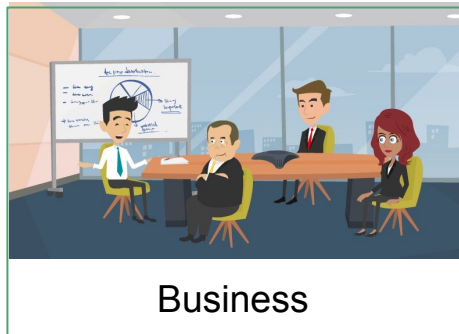
Predicted Talk Topic



?

=

True Talk Topic



Model

?

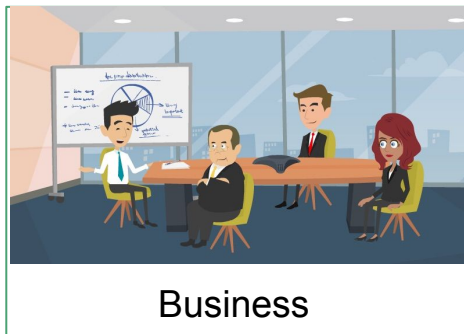
>

Monkey Throwing Darts



How well can a model predict?

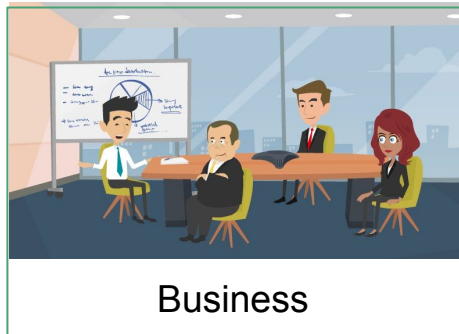
Predicted Talk Topic



55%

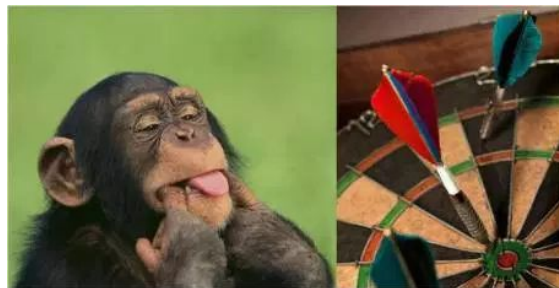


True Talk Topic



Linear SVM:
55% Accuracy

>



17% Accuracy

Can the talks be classified differently?

Ideal number of clusters (20 - 25) gives:

1. Topics already identified (but not the most popular)

Cluster 0:
Business /
Economy



Cluster 10:
Social Change /
Activism

Can the talks be classified differently?

Ideal number of clusters (20 - 25) gives:

1. Topics already identified (but not the most popular)

Cluster 0:
Business /
Economy



Cluster 10:
Social Change /
Activism

2. Nonsense: Same word, not topic

Warning: This talk might contain much more than you'd ever want to know about the way the world poops. But as sanitation activist (and TED Fellow) Francis de los Reyes asks ... doesn't everyone deserve a safe place to go?

Let's admit it: aerial photo drones and UAVs are a little creepy, and they come with big regulatory and safety problems. But aerial photos can be a powerful way of telling the truth about the world: the size of a protest, the spread of an oil spill, the wildlife hidden in a delta. Sergei Lupashin demos Fotokite, a nifty new way to see the world from on high, safely and under control.

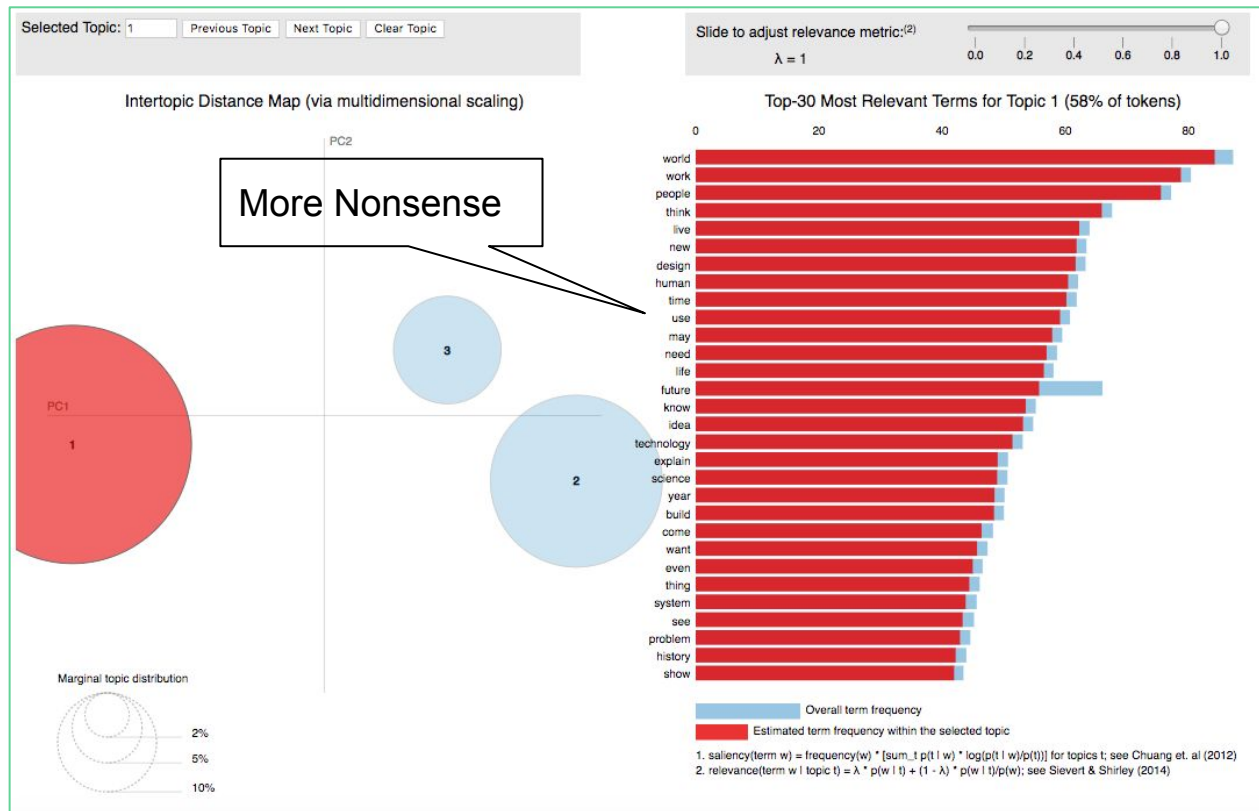
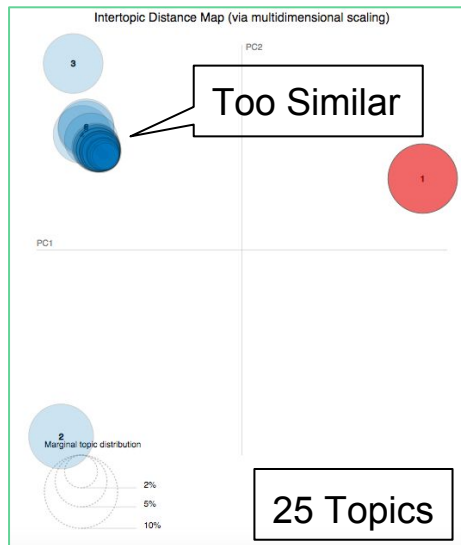
An epidemic of bad, inefficient, overcrowded meetings is plaguing the world's businesses ... and making workers miserable. David Grady has some ideas on how to stop it.



Cluster 4:
World



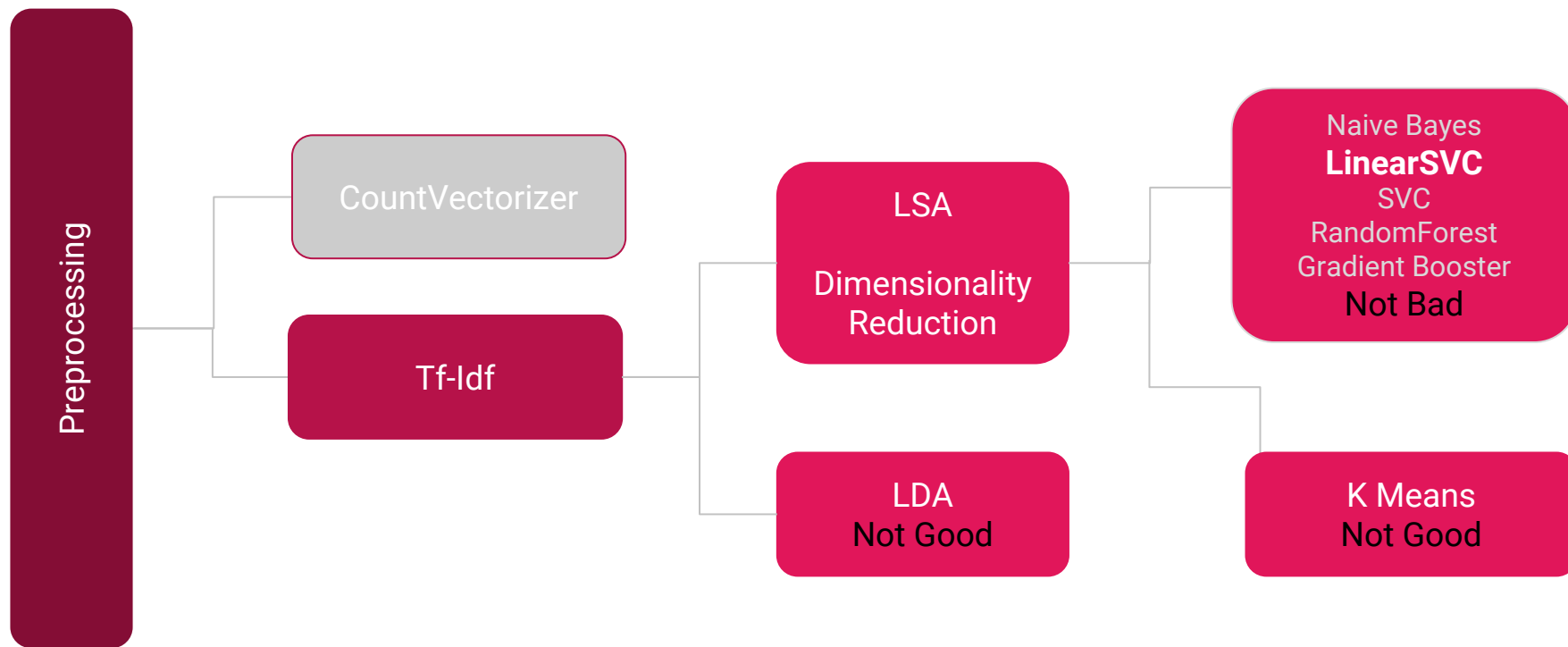
Can the talks be classified differently?



Going forward...

- ✚ Use full talk transcripts for better category prediction

Summary



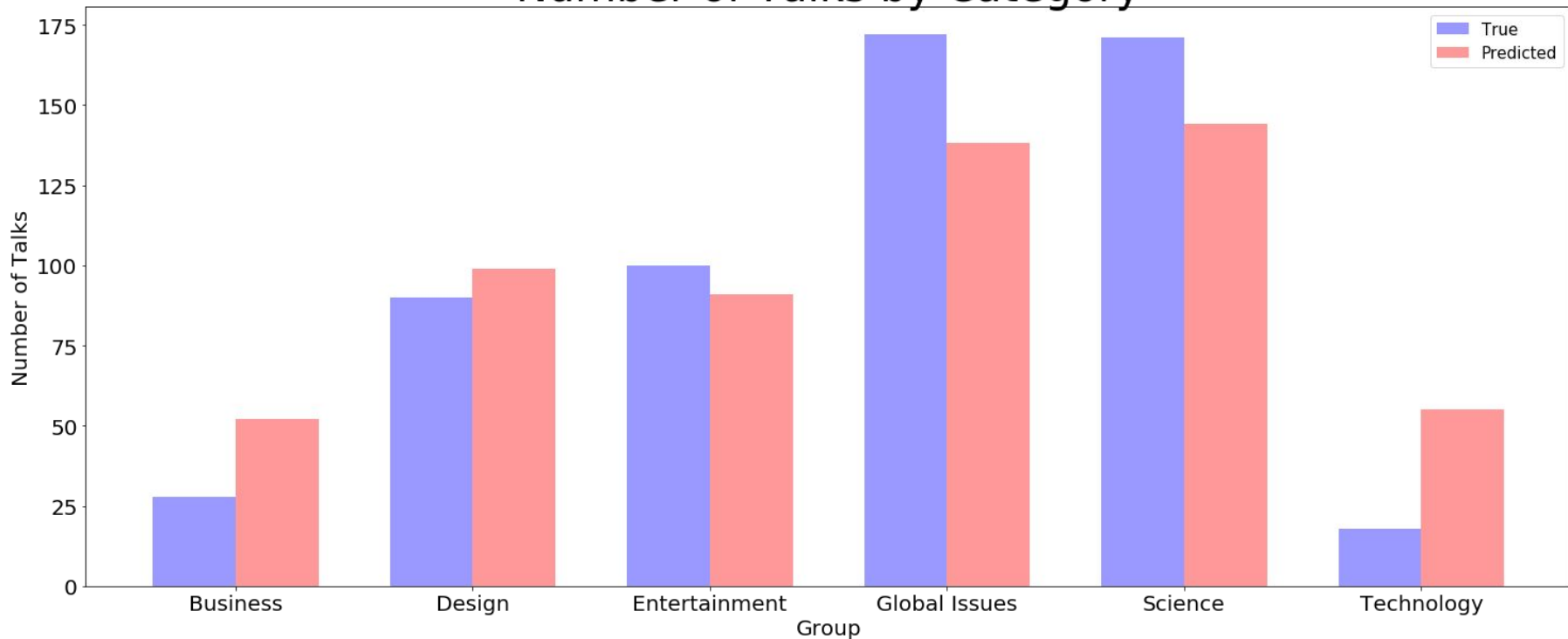
Appendix

Supervised Learning

Algorithm	Accuracy
Naive Bayes (Gaussian) (CountVectorizer)	0.45
LinearSVC	0.55
SVC - linear	0.54
SVC - rbf	0.53
Random Forest	0.46
Gradient Boosting	0.49

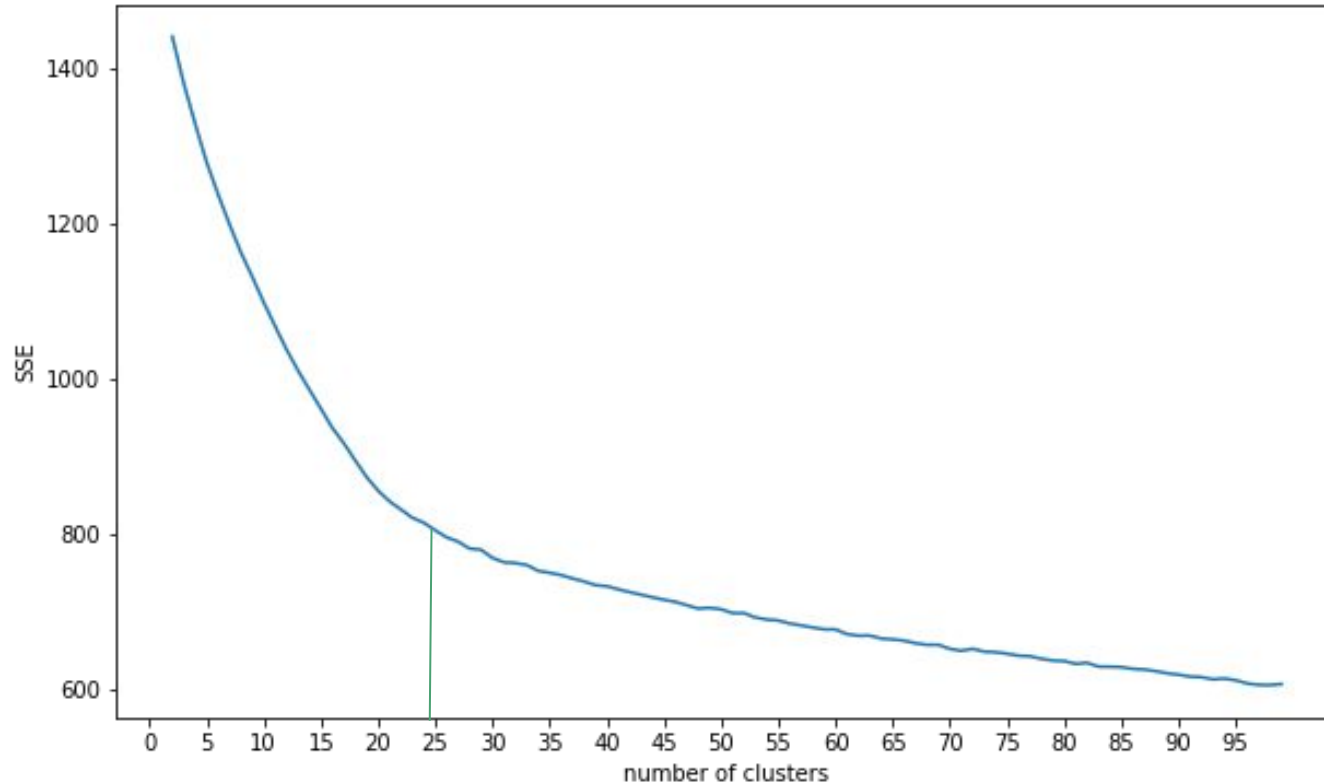
Tags and Topics (Test Data Set)

Number of Talks by Category



Talks without one of the six popular topic tags were removed (2200 → 1950 talks)

The maximum number of clusters: 25



Tags and Topics

Talk Topics (Jamie Oliver Example)



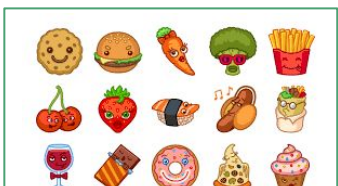
Business



Education



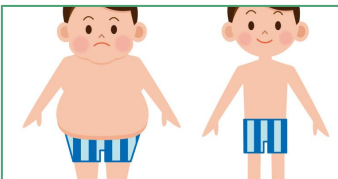
Global Issues



Food



Health

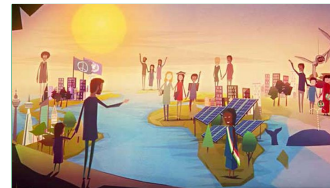


Obesity

Popular Topics



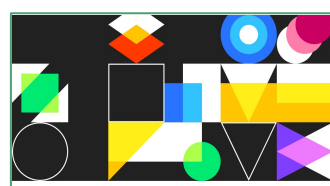
Business



Global Issues



Entertainment



Design



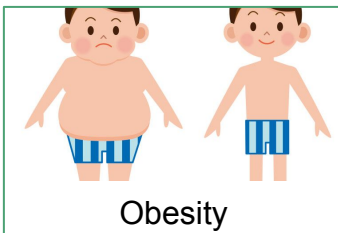
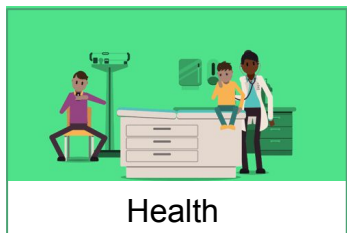
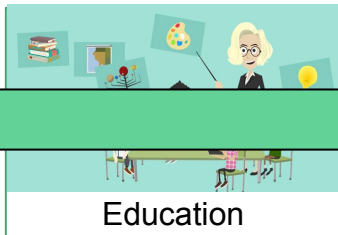
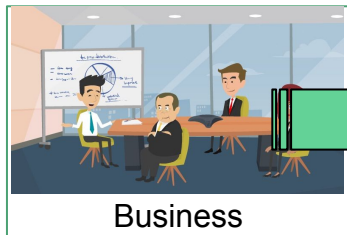
Technology



Science

Tags and Topics

Talk Topics (Jamie Oliver Example)



Popular Topics

