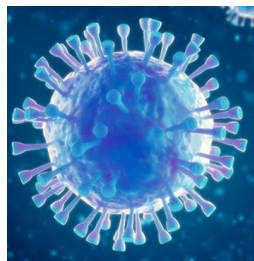


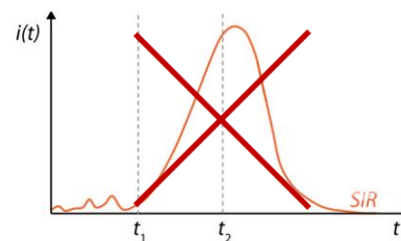
1. Overview of the problem
2. Dataset
3. Sentiment Analysis
4. Choropleth map of Italy
5. Results of the first analysis
6. Network
7. Topology
8. Users' Projection
9. Hashtags' Projection
10. Limitation
11. Future improvements



$$12 < R_t < 18$$



Whole life immunity



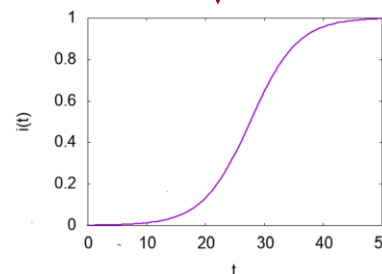
VACCINE HESITANCY



Not vaccinated children



- Fear about side effects
- Belief that children's immune system too immature
- Fake-news about vaccine-induced autism



2017

5000 measles cases

Lorenzin decree



TWITTER

- From May to December 2017
- 2 times per month ☒ 200 tweets per month
- Queries with Italian words/hashtags
- 28 combinations of words/hashtags:
vaccine-related + measles-related

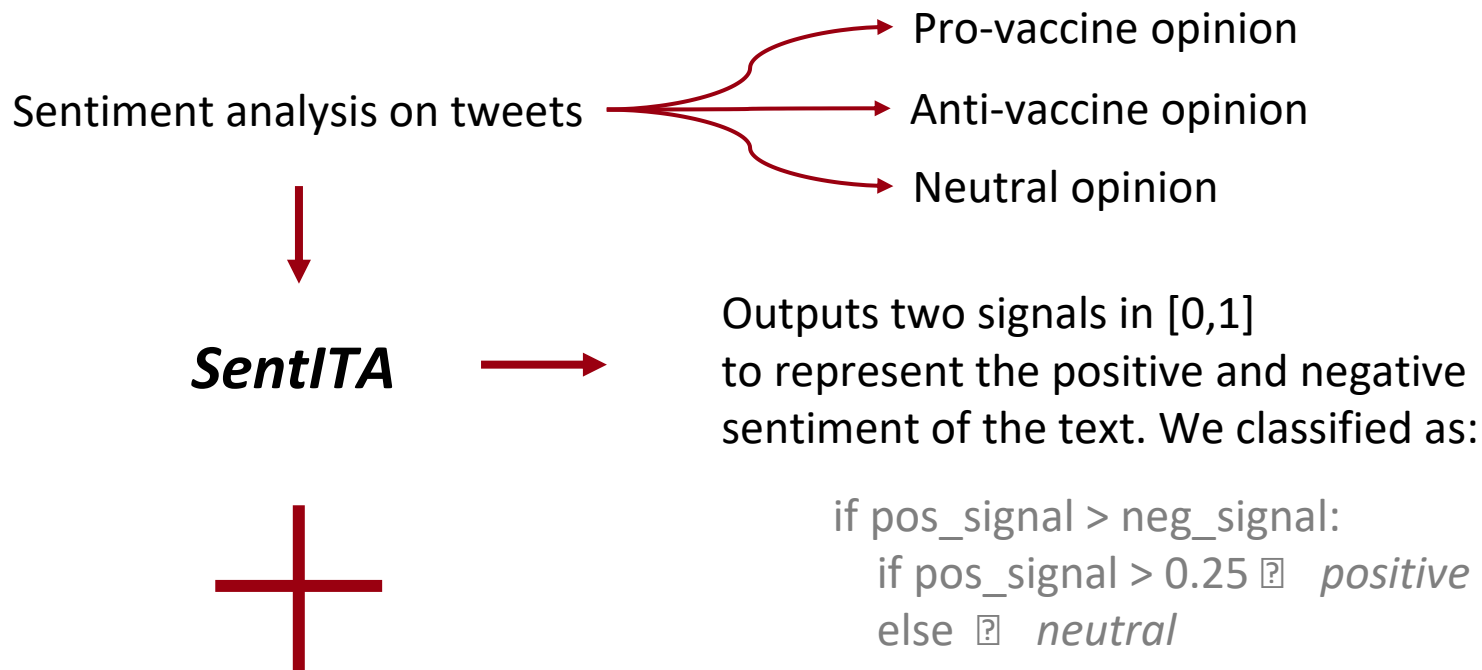
10340 tweets

- Tweet Id
- Day and Time
- Username
- Location
- Coordinates
- User_mentions
- Retweet count
- Full text
- Hashtags

6279 geolocalized tweets

tweet_id	created_at	username	location	coordinates	user_mentions	retweetcount	full_text	hashtags
0 863838208798412803	Sun May 14 19:27:32 +0000 2017	724648551561728000	Valdagno, Veneto	NaN	['mcocucci']	0	RT @mcocucci: Vaccinazioni e morbillo: oltre 2...	[]

3. Sentiment Analysis



Opinion-related words
during the queries to
Twitter API

PRO-WORDS	NEG-WORDS
Provax	Antivax
Freevax	Novax
IoVaccino	noobbligovaccinale
	libertàdiscelta

4. Choropleth map

Map cities' name into the corresponding regions' name
using a dictionary with 100 cities



3420 tweets

```
mapping = {
    'roma' : 'lazio',
    'rome' : 'lazio',
    'romeitaly' : 'lazio',
    'milano' : 'lombardia',
    'milan' : 'lombardia',
    'lecco' : 'lombardia',
    'rho' : 'lombardia',
    'monza' : 'lombardia',
```

The overall opinion of the each region
was computed as:

$$\frac{\#pro - \#con}{\#pro + \#con + \#neutral}$$



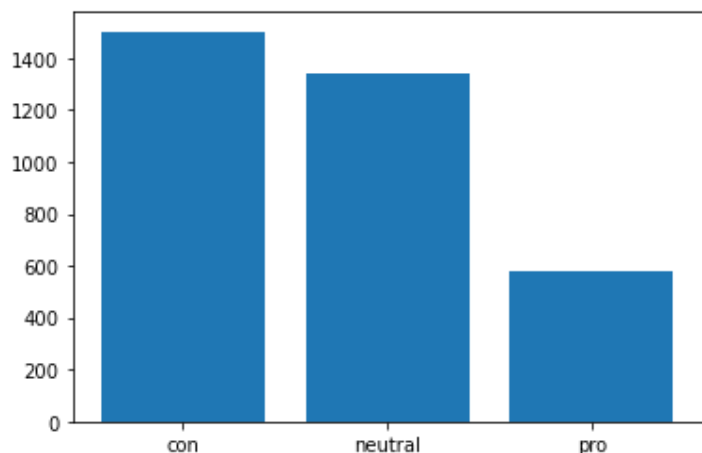
Python library to assign coordinates
to each region

location	count	lat	lon	pro	con	neutral	Opinion
lazio	724	44.525696	11.039437	48	327	349	-0.385359
lombardia	523	37.587794	14.155048	48	242	233	-0.370937
toscana	514	37.587794	14.155048	61	261	192	-0.389105
sicilia	441	43.458654	11.138920	54	191	196	-0.310658

...

abruzzo	23	37.587794	14.155048	1	10	12	-0.391304
valled'aosta	7	41.980804	12.766231	0	3	4	-0.428571
basilicata	2	40.860672	14.843984	0	1	1	-0.500000
molise	1	45.647666	11.866525	0	0	1	0.000000

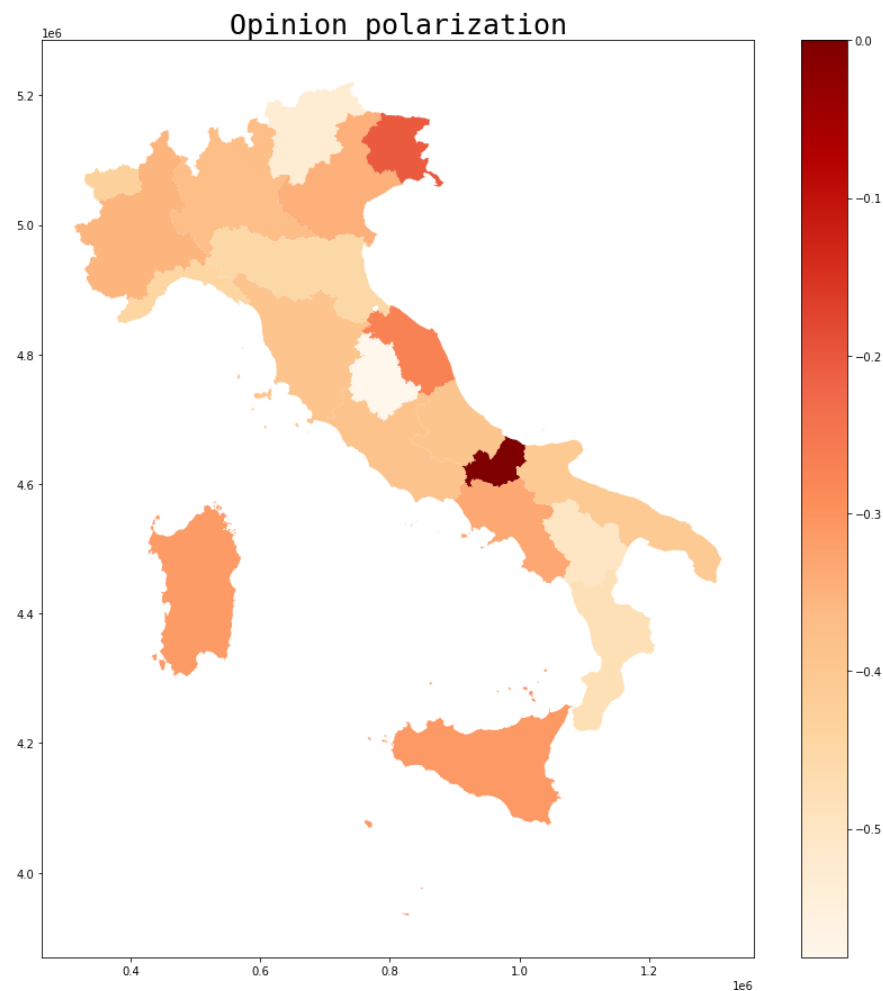
5. Results of the first analysis



44 % con 39 % neutral 17 % pro



- Affected from the tool used for the sentiment analysis
- Depend on the threshold (set to 0.25)



The opinion in each region is polarized towards a negative feeling about measles vaccine

username	location	usermentions	retweet_count	full_text	hashtags	opinion
----------	----------	--------------	---------------	-----------	----------	---------

140327940	emiliaromagna	["DarioBallini"]	0	chiss se autismo calato in modo proporzionale...	[autismo, morbillo, vaccini]	0
-----------	---------------	------------------	---	--	---------------------------------	---

556350111	sicilia	["l_gad48"]	0	vaccini morbillo autismo graphicnovel forse co...	[vaccini, morbillo, autismo, graphicnovel]	-1
-----------	---------	-------------	---	--	--	----

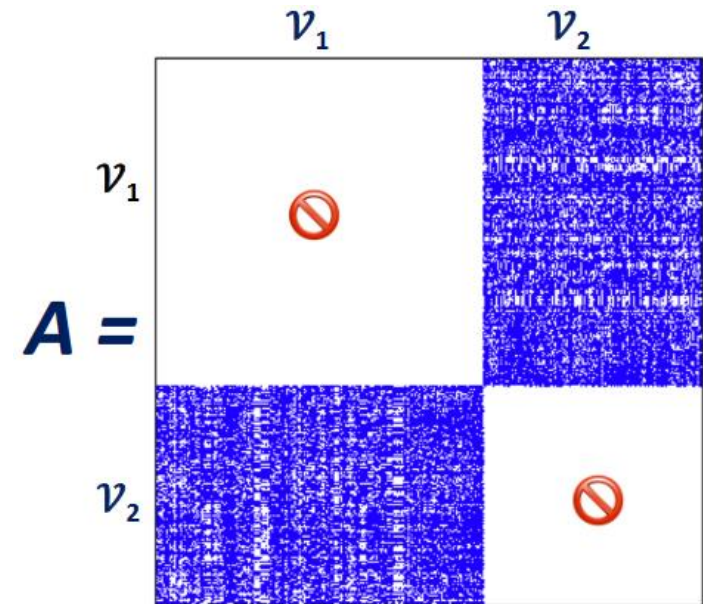
1364930161	sicilia	[]	1	vaccini morbillo autismo graphicnovel forse co...	[vaccini, morbillo, autismo, graphicnovel]	-1
------------	---------	----	---	--	--	----



$$\frac{\#pro - \#con}{\#pro + \#con + \#neutral}$$

	hashtags	pro	neutral	con	opinion
0	13luglio	0	1	0	0.000000
1	13ottobre	0	1	0	0.000000
2	16giugno	3	0	0	1.000000
3	18luglio	18	0	0	1.000000
4	18novembre	2	1	0	0.666667
...

3420 tweets



We decided to create a bipartite with users and hashtags as nodes sets

BIPARTITE: 1528 nodes , avg. degree 4.5

USER : 1049 nodes, avg. degree 575.2

HASHTAG: 479 nodes, avg. degree 65.2

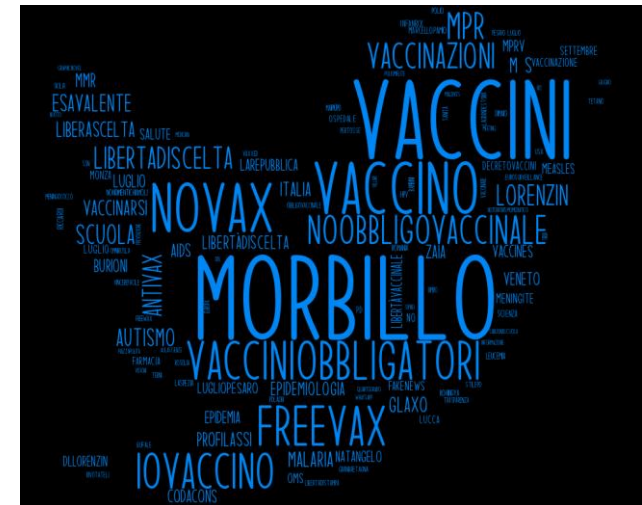
filtering:

users with weights < 2

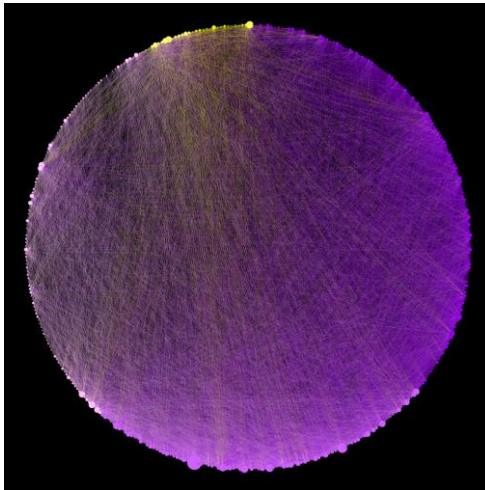
hashtags with weights < 1

USER: 397 nodes, avg. degree 53.2

HASHTAG: 206 nodes, avg. degree 22.4



	density	connected	diameter	avg. path length	gamma
user net.	0.12	no	4	2.05	7.6
hashtag net.	0.11	no	4	1.94	2.75

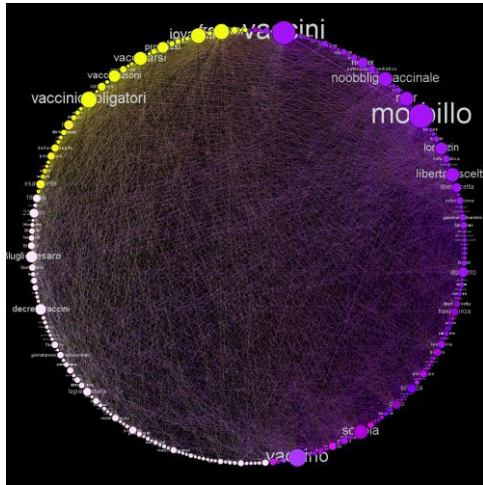


- 397 nodes
- 387 belonging to the giant component
- 54 average degree
- Only 5.53 % positive opinion
- Hubs present mainly in negative opinion group
- Average path length 2.04 $\longrightarrow \ln(\ln(N)) = 1.8$
- 6 communities

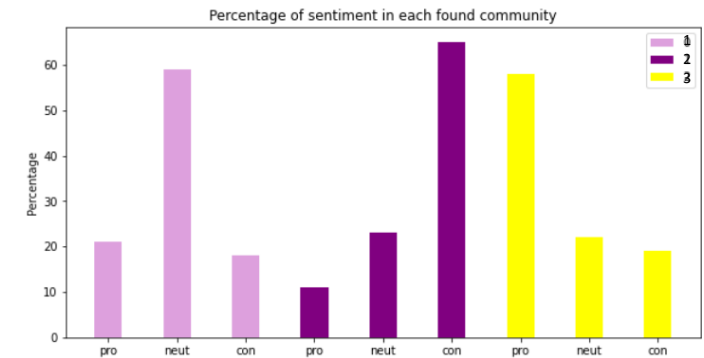
Legend:

- Users with a positive opinion
- Users with a negative opinion
- Users with a neutral opinion

HASHTAGS PROJECTION



- 206 nodes
- 204 belonging to the giant component
- 22 average degree
- 18 % positive
- 3 communities



Communication?

	To comm 1	To comm 2	To comm 3
comm 1	79.0	8.0	13.0
comm 2	38.0	46.0	16.0
comm 3	37.0	19.0	44.0

HUBS:

- Vaccini
- Morbillo
- Vaccino
- Vaccini Obbligatori

NEGATIVE

- Limited dataset
- Few geolocalized tweet → **6279**
- Sentiment analysis
- Amount of tweets per regions
- Network with few nodes



	location	pro	con	neutral
0	piemonte	33	112	79
1	valled'aosta	0	3	4
2	lombardia	48	242	233

- Bigger and well-proportioned dataset
- Training a network in a customized dataset
- Extract more information on opinions between and with regions
- Compare different periods
- Compare opinion with medical data

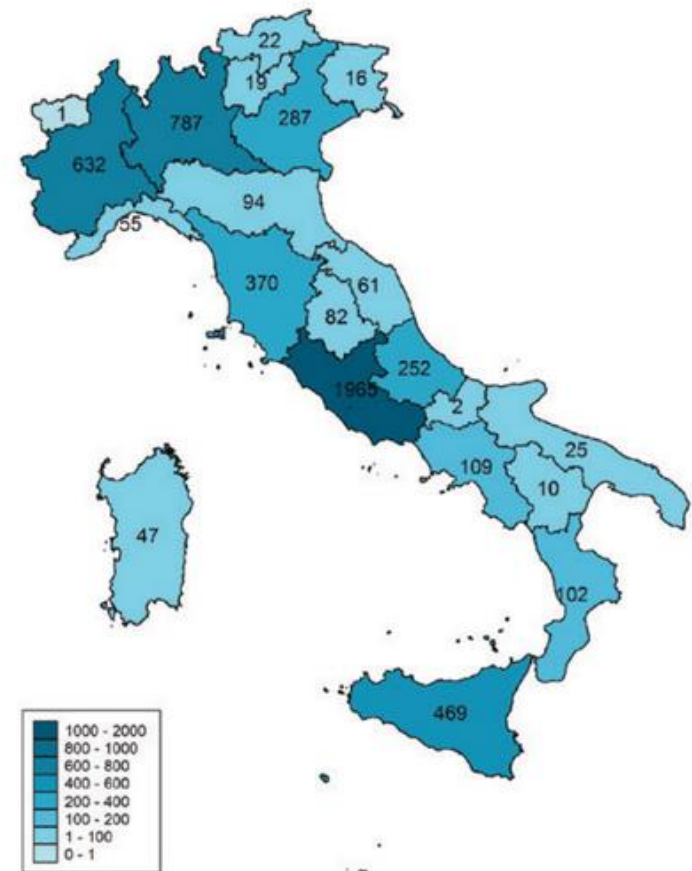


Figura 2. Distribuzione dei casi di morbillo segnalati per Regione. Italia 2017 (N = 5.407).

Taken from: L'epidemia di morbillo in Italia nel 2017.

Author: Antonietta Filia, Antonino Bella, Martina Del Manso, Maria Cristina Rota