

# Calcolare Opinioni e Controversialità

Valerio Basile

basile@di.unito.it

<http://valeriobasile.github.io/>



# > whoami

## *Post hoc vs Post-Doc*

### *The Post hoc Fallacy*

*To incorrectly assume "A" is the cause of "B" just because "A" preceded "B".*

*e.g. "All Professors have Ph.D.'s, therefore getting a Ph.D. means you'll get a Professor job (right?)"*



### *The Post-Doc Fallacy*

*To incorrectly assume you'll have a job just because you have a Ph.D.*

*e.g.  
"Now what??"*

JORGE CHAM © 2009

WWW.PHDCOMICS.COM

Content-Centered Computing group  
<http://hatespeech.di.unito.it/>

AI ↔ NLP ↔ Semantica ↔ Sentiment,  
Hate Speech

# Sentiment Analysis

a.k.a. Opinion Mining

“Given a set of evaluative text documents  $D$  that contain opinions (or sentiments) about an object, opinion mining aims to extract attributes and components of the object that have been commented on in each document  $d \in D$  and to determine whether the comments are positive, negative or neutral.”

Bing Liu

Grande richiesta dall'industria  
Particolare attenzione ai **social media**

# Sentiment Analysis



★★★FRANCO★★★  
@francodeandrea

Follow



Replying to @SanremoRai @ManuelaZeroOff

@lazerofficial bravissima, voce stupenda.  
Avanti così... #spingere!



Ramon 🇮🇹  
@RamonGrun

Follow



#PiazzaPulita tanto non ci commuovete.  
Siete complici dei trafficanti

# Sentiment Analysis



★★★FRANCO★★★  
@francodeandrea

Follow



Replying to @SanremoRai @ManuelaZeroOff

@lazerofficial bravissima, voce stupenda.  
Avanti così... #spingere!



Ramon 🇮🇹  
@RamonGrun

Follow



#PiazzaPulita tanto non ci commuovete.  
Siete complici dei trafficanti



# Lessici affettivi

paradisiaco	a	166146	0.875	0	1	0.875
piacente	a	166146	0.875	0	1	0.875
portentoso	a	166146	0.875	0	1	0.875
stupendo	a	166146	0.875	0	1	0.875
suggestivo	a	166146	0.875	0	1	0.875
udibile	a	173764	0.375	0	1	0.375
nefasto	a	176838	0.111	0.889	-0.8418425791	0.8959028965
fasto	a	177186	0.5	0	1	0.5
fausto	a	177186	0.5	0	1	0.5
favorevole	a	177186	0.5	0	1	0.5
propizio	a	177186	0.5	0	1	0.5
infausto	a	177963	0	0.625	-1	0.625
di_malaugurio	a	178126	0.25	0.75	-0.5903344706	0.790569415
abusivo	a	179486	0	0.875	-1	0.875
anticostituzionale	a	180211	0	0.75	-1	0.75
incostituzionale	a	180211	0	0.75	-1	0.75
disponibile	a	183053	0.375	0	1	0.375

Basile and Nissim (2013): Sentiment Analysis on Italian Tweets, WASSA 2013

# Lessici affettivi



**Spinoza**

@spinozait

Follow



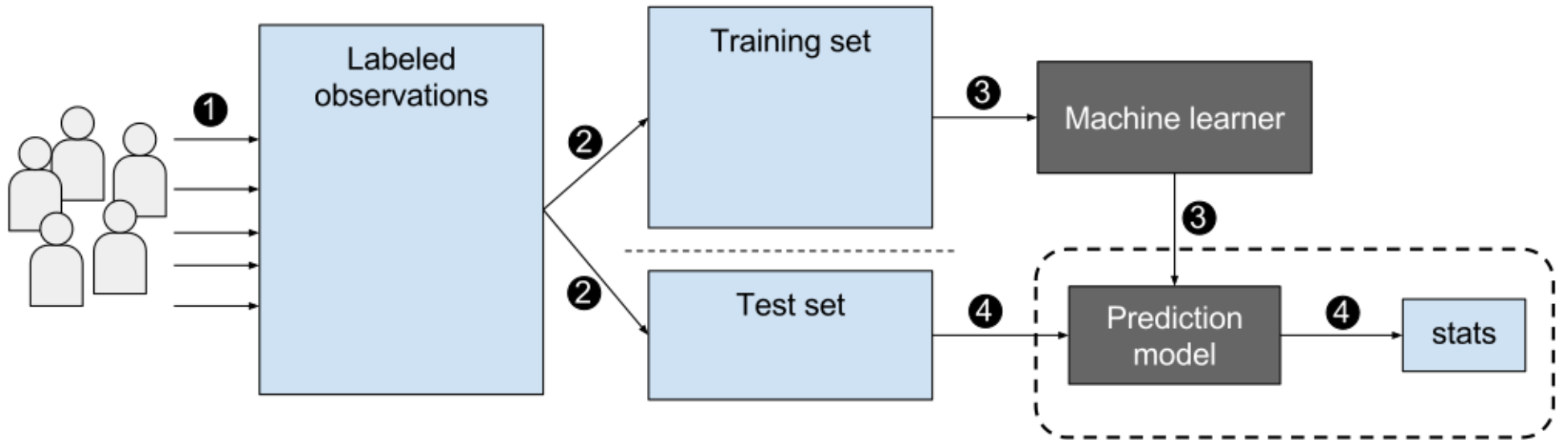
Trattativa Stato-mafia, Napolitano ha  
chiamato Mancino congratulandosi per  
l'assoluzione. Poi ha ingoiato la SIM.  
[@rostokkio]



?

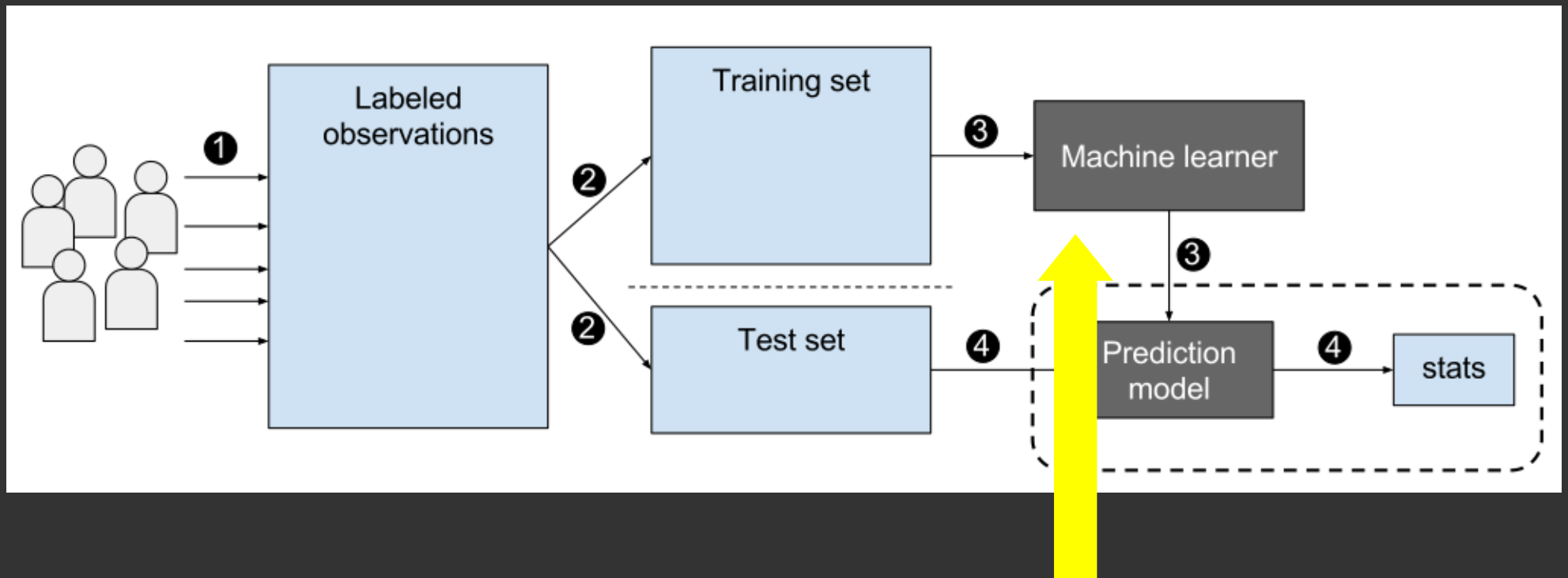


# Supervised Learning





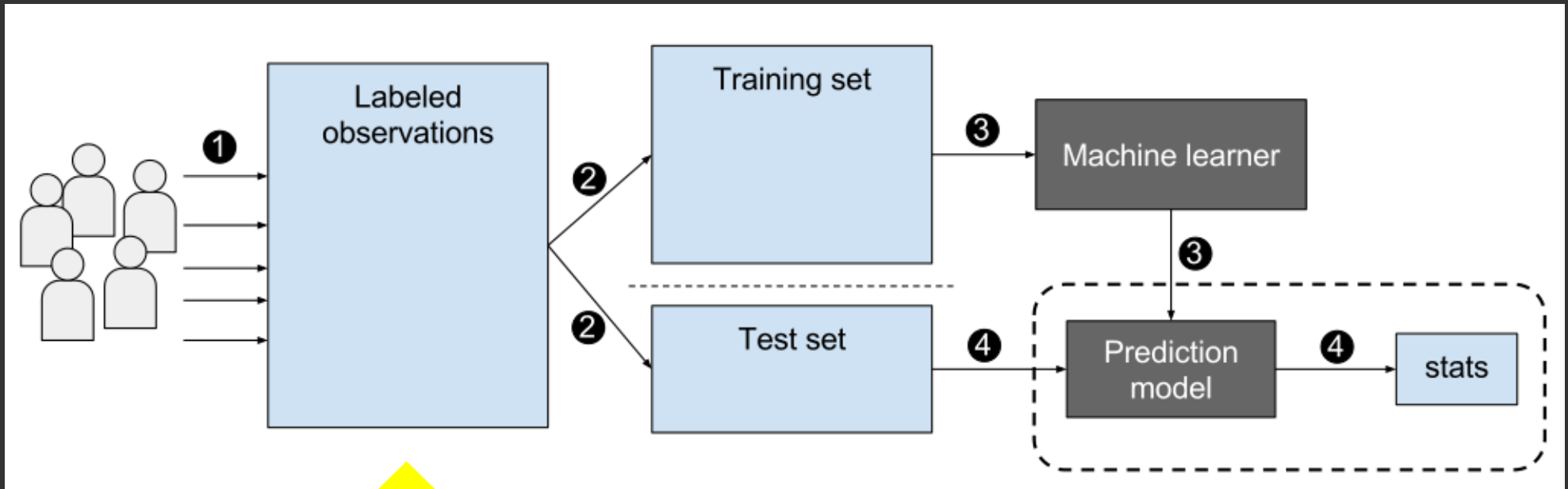
# Supervised Learning



Support Vector Machine,  
Recurrent Neural Network,

...

# Supervised Learning



**Spinoza**  
@spinozait

Follow

Trattativa Stato-mafia, Napolitano  
chiamato Mancino congratulando  
l'assoluzione. Poi ha ingoiato la S  
[[@rostokkio](#)]

Replied to [@SanremoRai](#) [@ManuelaZeroOff](#)

**FRANCO**  
@francodeandrea

Follow

[@lazeroofficial](#) bravissima, voce stupenda.  
Avanti così... [#spingere!](#)

**Ramon**  
@RamonGrun

[#PiazzaPulita](#) tanto non ci commuovete.  
Siete complici dei trafficanti

# L'annotazione manuale

- Esperti
- Crowdsourcing
- Distant Supervision
- Gamification
- ...



# L'annotazione manuale

text	HS	TR	AG
Detectan a miles de refugiados que se van de vacaciones a sus países con dinero de las ayudas sociales <a href="https://t.co/83HAI">https://t.co/83HAI</a>	0	0	0
El Gobierno fomentará la enseñanza de árabe y cultura marroquí en los colegios <a href="https://t.co/ETnKSfRVww">https://t.co/ETnKSfRVww</a>	0	0	0
#QuintanaRoo// 'TEIBOLERAS' SIN PAPELES <a href="https://t.co/MTvzWJG1bV">https://t.co/MTvzWJG1bV</a> <a href="https://t.co/XclGMf5Kha">https://t.co/XclGMf5Kha</a>	0	0	0
Yo lo que quiero es que desaparezcáis como club, maldita escoria gabacha de mierda @PSG_inside <a href="https://t.co/mwP11mg">https://t.co/mwP11mg</a>	0	0	0
Los árabes y los venezolanos jajaja <a href="https://t.co/4UPxNGicsX">https://t.co/4UPxNGicsX</a>	0	0	0
Unicef dará a Guyana 30 mil dólares en ayuda para refugiados venezolanos <a href="https://t.co/QFKsevk5jh">https://t.co/QFKsevk5jh</a>	0	0	0
Contaré que el fin un amigo me dijo algo que me partió el alma, el pasó por una glorieta donde había unos indocumentados	0	0	0
Intereses mutuos impulsan la creciente relación entre Emiratos Árabes y Costa Rica <a href="https://t.co/zX34lO3dGI">https://t.co/zX34lO3dGI</a>	0	0	0
En noviembre se cumplen 30 años desde que apareció el primer cadáver de un inmigrante en la costa de Tarifa. ¿No ha ha	0	0	0
El asesor para Oriente Próximo de la Casa Blanca ha puesto en marcha una iniciativa para acabar con la #UNRWA , con ot	0	0	0
Decenas de víctimas en ataques de la coalición árabe contra Al Hudaida <a href="https://t.co/iiacLOqfmw">https://t.co/iiacLOqfmw</a>	0	0	0
▼ ACNUR mantiene la llamada internacional para ayudar a los refugiados sirios #Sociedad <a href="https://t.co/dJWDVNLIGC">https://t.co/dJWDVNLIGC</a>	0	0	0
La izquierda española pide que se derribe la valla de Ceuta y Melilla para que entren los inmigrantes   El Municipio <a href="https://t.co/">https://t.co/</a>	0	0	0

Groningen Meaning Bank

<http://gmb.let.rug.nl>

Sentipolc 2014-2016

<http://www.di.unito.it/~tutreeb/sentipolc-evalita16/index.html>

ABSITA 2018

<http://sag.art.uniroma2.it/absita/>

HatEval 2019

<https://competitions.codalab.org/competitions/19935>



# L'annotazione manuale

**Tweet:**

RT @mariannethleme: In NL 14,2 miljoen varkens en 505 miljoen kippen geslacht na kort en ellendig leven. Bijdrage BNP 0,3% He-le-maal ni ...

—bramvanliere 2012-09-02 08:19:44

[view conversation](#)

Type of tweet	Twittering behavior	Topic of Tweet	Criticized party	Personal	Topic of personal
Normal post	Campaign Trail	N/A	N/A	Personal	Not applicable
@Reply	Campaign Promotion	Animal Rights	CDA	Mixed	Children
Retweet	Campaign Action	Civil/Human Rights	PvdA	Political	Pets
Retweet with comment	Call to Vote	Crime/Judicial Proceedings	SP		Marital Life/Relationships
	Political Report	Business/Economy	VVD		Family time
	Other Report	Education	PVV		Hobbies
	Own/Party Stance	Environment	GL		TV/Film
	Critique	EU	CU		Music
	Requesting Public Input	Government/Democracy	D66		Literature
	Advice/Helping	Social Welfare/Benefits	PvdD		Sports
	Acknowledgement	Health	SGP		Friends/Chatter
	Personal	Immigration/Integration	50Plus		Health/Well-being
	Other	Military/Defense	DPK		Religion/Spirituality
		Religion	Other Party		Fashion/Beauty
		Science/Technology			Food/Drink
		War/Conflicts			Tech
		World Affairs/Events			Past Life/Upbringing
		National Events/Heritage			Places/Travel
		Infrastructure			Other
		Campaign/Party affairs			
		Politicians/Personalities			
		Norms/Values			
		Media			

Reaction: Leader debates	Reaction: News coverage
Yes	Yes
No	No

# Inter-rated agreement

(o *reliability*)

- % agreement
- Cohen's Kappa
- Fleiss' Kappa
- Krippendorff's alpha
- ...

Annotation

→ Harmonization

→ Agreement

→ **Gold standard** data

# Crowdsourcing

## Monitor

Finalized rows are ready to be converted to test questions

[Convert to Test Questions](#)

[Dashboard](#)

[Alerts](#)

[Advanced Analytics](#)

[Contributors](#)

100%

Complete [i](#)

\$57

Cost [i](#)

42

Active Test Questions [i](#)

3,589

Rows [i](#)

0

Judgments By Hour [i](#)

10,834

Trusted Judgments [i](#)

2,509

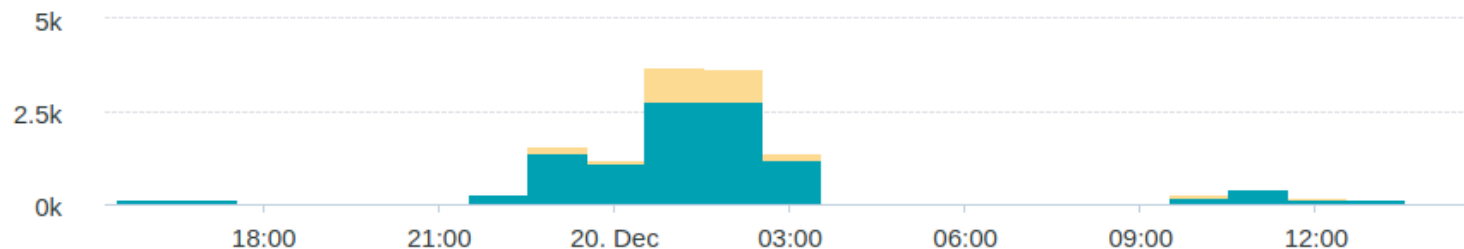
Untrusted Judgments [i](#)

0

Pending Judgments [i](#)

### Judgments By Hour

Estimated Completion: Dec 20, 2:10 pm



# Crowdsourcing

@La12twittera En los últimos 30 años ganaron 2 sudaca y una soronga, los sin barrios tienen en los últimos 20 años más chapa copera q estos amargos jajajaj

¿Este tweet expresa odio contra los inmigrantes? (required)

☐ No

☒ Sí

¿Cual es el tipo de objetivo? (required)

☐ Grupo

☐ Persona

¿Este tweet es agresivo? (required)

☐ No

☐ Sí

<https://www.figure-eight.com/>



# Gamification

?

wordrobe

play what you mean



 Senses Questions left until drawer is completed: 5 

The Who is currently touring in support of Endless Wire, its first **album** since 1982.

- ☒ one or more recordings issued together – originally released on 12-inch phonograph records (usually with attractive record covers) and later on cassette audiotape and compact disc
- ☐ a book of blank pages with pockets or envelopes – for organizing photographs or stamp collections etc

Place your bet: low  high

<https://www.wordrobe.org>

**Siamo d'accordo?**

# Fenomeni Altamente Soggettivi

- parte del discorso (**sintassi**)
- sensi, entità (**semantica**)
- sentimenti, emozioni
- ironia, sarcasmo
- razzismo, misoginia, omofobia

**pragmatica**

soggettività



# Fenomeni Altamente Soggettivi

**sintassi**

Valerio **sale** le scale

nome o verbo?

# Fenomeni Altamente Soggettivi

polarità



**Domenico**  
@Domyph

Follow

Siamo a 50 [#followers](#) grazie a tutti

 Translate Tweet

2:56 PM - 31 Jan 2013 from [Calderara di Reno, Emilia Romagna](#)



# Fenomeni Altamente Soggettivi

**ironia**

Follow



Pubblicata la classifica mondiale della libertà di stampa. Non possiamo dirvi altro.

# Fenomeni Altamente Soggettivi

## hate speech



**giandomenico** 

@giandomenic45

Follow



Replying to @matteosalvinimi


Ma se son pieni di musulmani che picchiano tutti in casa.!Voglio vedere quanti ne puniscono.!

# Omofobia online

Periodo: 11/12/2018 - 10/01/2019 Filtra


Tweet totali: 8159


	positivo	negativo	neutro	dubbio	Totale
Sentiment:	0	0	0	0	0
Non pertinenti: 0			Classificati: 0 (0.00%)		

 **Daniele Mont**  
@dmdarpizio

Concorsi per sole donne e taglio dei finanziamenti a università ed enti di ricerca con un basso indice di "Gender Equality". Nell'università inclusiva del futuro non basterà essere in gamba nella propria materia. [ilbolive.unipd.it/it/news/univer...](http://ilbolive.unipd.it/it/news/univer...)

11:48 AM - Dec 11, 2018

 **Università e donne: l'impegno irlandese per ...**  
Il governo irlandese ha ufficializzato l'entrata in vigore del Gender action plan, un piano triennale rivolto alle università e istituti di ricerca con [ilbolive.unipd.it](http://ilbolive.unipd.it)

 See Daniele Mont's other Tweets

✓ ✕ - ? 🗑



# Omofobia online

annotator	positive	negative	neutral	uncertain
7	2	0	0	0
5	7	2	0	3
13	41	20	3	11
10	63	37	2	3
15	53	40	2	4
2	3	4	0	0
8	3	4	0	0
9	34	47	7	10
16	10	26	13	54
12	24	78	1	0

# Controversialità

Una misura che:

- **massimizza** l'accordo all'interno dei gruppi
- **minimizza** l'accordo tra i gruppi

$$g_1 = \{1,2,3\}; g_2 = \{4,5,6\}$$

$$(\text{agreement}(g_1) + \text{agreement}(g_2)) (1 - \text{agreement}(g_1 \cup g_2))$$

Calcolata **per ogni istanza**

# Controversialità

$$(\text{agreement}(g_1) + \text{agreement}(g_2)) (1 - \text{agreement}(g_1 \cup g_2))$$

Es.:

$$(1, 1, 1) (1, 1, 1) \rightarrow 0.0$$

$$(0, 0, 0) (1, 0, 0) \rightarrow 0.27$$

$$(1, 0, 1) (0, 0, 0) \rightarrow 0.55$$

$$(1, 1, 1) (0, 0, 0) \rightarrow 1.0$$

# Controversialità

## Teoria gender Nature vs. nurture Scuole

1.000 @Nicsilvano No, non nasci omosessuale. Ci diventi ed è un fattore legato a traumi soprattutto. Che ti piaccia o meno.  
0.857 @antoniooguterres @UN invece di pensare allecazzate del Gender-clima, pensi a questo, alla sovrappopolazione africana, ... <https://t.co/oYo0B5cLyE>  
0.857 @alfredoferrante @MissItalia @andreavenuto @MichelaMirani @fishonlus @vinc\_falabella @isabillride @MyEDF... <https://t.co/j3RyQzXuix>  
0.800 UNAR, Ufficio Nazionale Antidiscriminazione Razziale sovvenziona il circolo omosessuale... <https://t.co/FdY0r94Ybo>  
0.800 Silvana De Mari a Otto e Mezzo: 'dittatura delle minoranze, omofobia e islamofobia psichiatrici di un governo... <https://t.co/Wogb7Rj7sV>  
0.800 Secondo #Provita Onlus, per avere figli eterosessuali servono «giochi sessiti, omofobi e stereotipati». La maggior... <https://t.co/5QYBNkjx03>  
0.800 @RVenieri @LaKlaretta Magra consolazione quando si usa la diffamazione e l'omofobia  
0.800 @rob\_milano @iaia90\_ @manginobrioches Sono solo due righe di un libro, siamo all'omofobia...  
0.800 Quelli che cianciano di "lobby gay" hanno creato una vera lobby omofoba, di omofobi professionisti. <https://t.co/pGp6AykJeD>  
0.800 Nasce "Famiglia e Vita": gli omofobi si riuniscono in un intergruppo parlamentare <https://t.co/UrLI6EiVa>  
0.800 La polizia russa si rifiuta di proteggere gli omosessuali dagli attacchi omofobi <https://t.co/5m5...07e> <https://t.co/F47uFu8Xxo>  
0.800 io: le multinazionali cavalcano l'onda del pride per guadagnare ma non fanno nulla di concreto per bidimensionare l... <https://t.co/K790wwJpV0>  
0.800 in fase di preparazione... il Calendario Meridionalista l'unico con i Santi Gender... contro ogni sessismo...  
0.800 @eroilroscio @cingackuk L'omofobia e' nel tuo cervello.  
0.800 Dopo il Gender è allarme streghe: la battaglia del neosenatore leghista nelle scuole <https://t.co/wIWkLitrfM>  
0.800 C'è ancora un giudice a Berlino!! La procura di Roma condanna le due paladine #LGBT, Appendino e Raggi, per aver c... <https://t.co/X5quysZw1z>  
0.800 @biggestinterist @FabDellaValle @Gazzetta\_it Quindi voi uomini afflitti dal complesso di inferiorità date dell'omos... <https://t.co/y6qVXH74NU>  
0.800 a occhio, vivete sulla luna. Non fatevi rimbambire dai gay pride e da tutto il resto. I vostri figli hanno bisogno... <https://t.co/k28x2gtASW>  
0.800 @Amos8125 @FartFromAmerica 2/2 Il libro chiave è Billy Budd di Melville, incomprensibile senza il gender Parla di o... <https://t.co/GcrqvDA0Dh>  
0.800 @AlessiaMorani @lauraleghista Alessia che brutta combinazione... PD, avvocatessa, tifosa del Milan... Ti manca essere lesbica e sei apposto...  
0.766 Ma così la vuole nascondere la sua omosessualità?! Il leghista Pillon: «Obbligheremo le donne a partorire. Vorrei... <https://t.co/ayjb2ziUjE>  
0.766 @Fontana3Lorenzo Non posso pensare che una persona intelligente come lei sia omofobo, certo, maschio e femmina form... <https://t.co/Z03EYmR6jR>  
0.750 Che poi tra tutte le parole che gli omosessuali potevano usare per descriversi hanno scelto proprio "gay" (gaio, fe... <https://t.co/vCwzKQggYi>  
0.694 Wow non sapevo che i gay e le lesbiche avessero poteri magici talmente tanto forti da riuscire a far crollare un po... <https://t.co/LJHdVJwaUR>  
0.694 @simondritt Una curiosità, ma se sentissi un omosessuale (entrambi i sessi) dire ad un etero che schifo quando bac... <https://t.co/PjRYHDZYJ1>  
0.694 @semprekimara @hotel\_solemar Quale dovrebbe essere il problema che tu sia una transgender? Conosco moltissime perso... <https://t.co/EIHj2WUrTr>  
0.694 @Red74Mike @Morbidosamente concordo su tutto, basta che non si esageri arrivando al gender x sui neonati perché a tutto c'è un limite  
0.694 @puffettainlov3 forse sei te che devi curare l'omofobia, sarai una persona migliore ☺  
0.694 Non sono omofobo, MA <https://t.co/LAQxnQyb0h>  
0.694 Non c'è nessuna emergenza razzismo. E neanche omofobia. <https://t.co/wiFH2f0W0b>  
0.694 @matteosalvinimi Quindi in Italia voi leghisti non siete omofobi, no... Legga queste scritte: è il suo NOME!! <https://t.co/9PJpBPFnxa>  
0.694 La7 dopo aver sdoganato fascisti, razzisti e omofobi oggi dà spazio a Silvana de Mari. Peccato che Goebbels e Himml... <https://t.co/5k0cLs0ub5>  
0.694 @jovino Perchè, qualcuno crede davvero che sia omofobo e razzista????  
0.694 @InfoMalgioglio ma pensa che questa tua ostentazione della omosessualità sia di orgoglio ai gay . Quelli seri ? Ma... <https://t.co/VYIYXQVSpl>  
0.694 il DDL #Pillon per l'alienazione genitoriale si ispira a Richard Gardner. Uno psichiatra omofobo che invitava le mo... <https://t.co/dgD00co7U0>  
0.694 @hidekivelasco La polizia è dalla mia parte i poliziotti sono omofobi come me  
0.694 Gay e lesbiche che sostenete Salvini, ciucciate! Ve lo dicono i vostri amichetti fasci. Godetevela, questa sindrome... <https://t.co/9hiiVj4cMT>  
0.694 Diritti gay Diritti lesbo ma del Diritto di ogni donna di passeggiare x strada senza subire violenze di sorta fino... <https://t.co/y7D0Y4wLT4>  
0.694 Collega Omofobo arriva dal nulla e fa "SIETE NA MASNADA DE FROCI". Penso sia la sua parola preferita dopo "Inter".  
0.694 @Chiara92210799 Ma che te frega a te! Mo pure omofobia sei. Saranno fatti suoi con chi dorme. Fatto sta che il vero... <https://t.co/CxyKLTG7AY>  
0.694 @ceg62 @Pontifex\_it lo schifo sono gli omofobi  
0.694 @byoblu Come far estinguere un po' di persone? Trovato....inventiamoci il gender! ☺  
0.667 @umbertogardini @lauraboldrini Salvini è un cafone razzista, maschilista, misogino, omofobo e pure ignorante come u... <https://t.co/lT6gid8e33>  
0.667 Tossica, con amici pusherini, e pure lesbica. Ciao papi ti penso 🍷

# Controversialità

## Tweet su Brexit

text	Fabio	Valerio	Komal	Sohail	Sumaia	Rizwan		Disagreement
illuminati own EU, provide for, and apply, muslim insurgency, as mean to global genocide; so muslims for EU, for now <a href="https://t.co/oeShKeINVb">https://t.co/oeShKeINVb</a>	0	0	0	1	1	1		100%
saw you on the deport list.good luck. muzzies love jews so pick country.	0	0	0	1	1	1		100%
First chance to celebrate leaving EU! Time out for dancing & fun.. ;) Best time in ages. Hope someone gets those Paki Landlords. x	0	1	0	1	1	1		89%
Good for you BRITS...You saved your countrv. Now seal vourselves from the	0	0	1	1	1	1		89%
More on how some migrant men, posing as children, behave i public	0	0	1	1	0	1		78%

PhD Sohail Akhtar

# Controversialità

- Controversialità come feature  
(à la *CrowdTruth*)
- E se non sappiamo niente degli annotatori?  
(PhD Komal Florio)
- Possiamo **predire** la controversialità?