

UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Dipartimento di Ingegneria
dell'Informazione

Corso di Laurea in
Ingegneria Informatica

Intelligenza Artificiale A.A. 2017/2018

FAKE NEWS DETECTION

A cura di:
Stefano Ivancich
Luca Lezzi
Luca Masiero

Docente:
Silvana Badaloni

INDICE

1. Il problema delle Fake News	1
1.1. Cosa sono le fake news e perché esistono	1
1.2. Cronaca recente: elezioni statunitensi del 2016	1
1.2.1. Influenza delle Ads russe	1
1.2.2. Cambridge Analytica	2
1.2.3. La strategia comunicativa di Trump	3
1.3. Le bolle Social	3
1.4. Non è un problema recente (Es. Cicerone 64 a.C.)	5
1.5. Tecniche Manuali: <i>fact cheking, debunkers</i> e il caso <i>Butac</i>	5
2. Principali tecniche automatiche ad oggi	7
2.1. Riconoscimento tramite i like messi ai post di Facebook	7
2.2. Analisi della coerenza tra titolo e testo di un articolo	9
2.3. Riconoscimento tramite l'analisi del testo	10
Bibliografia	11

1. Il problema delle Fake News

1.1. Cosa sono le fake news e perché esistono

Le *fake news*, secondo il New York Times, vengono definite come “*storie costruite con l'intenzione di ingannare, spesso per un fine economico o politico*” [1].

Da diverse ricerche nel campo della psicologia si viene a conoscenza che: “*la maggior parte delle decisioni non vengono prese basandosi sulla logica, ma basandosi sulle emozioni. La logica viene poi utilizzata per giustificare la scelta irrazionale che l'individuo ha preso*” [2].

Il Dizionario di Oxford, l'equivalente inglese dell'Accademia della Crusca, nel 2016 ha annunciato come parola descrittiva dell'anno: *Post Verità*, ovvero *la serie di vicende nelle quali i fatti oggettivi sono meno influenti sull'opinione pubblica rispetto alle emozioni che vengono suscitate dal modo di essere raccontate* [3].

Per il Collins, dizionario inglese più autorevole, la parola rappresentativa del 2017 è stata *Fake News* [4].

Le bufale sui Social Network si diffondono rapidamente, con un picco nelle prime 2 ore [5].

In questi due termini: *Post Verità* e *Fake News*, secondo alcuni, è riassunta la strategia comunicativa vincente della politica d'oggi [6].

1.2. Cronaca recente: elezioni statunitensi del 2016

1.2.1. Influenza delle Ads russe

Tra il giugno del 2015 e l'agosto del 2017, milioni di americani sono stati esposti agli annunci di Facebook e ai post generati da agenti russi che hanno cercato di influenzare il comportamento degli elettori e sfruttare le divisioni nella società americana su questioni scottanti. [7]

Hanno raggiunto 126 milioni di utenti su Facebook. [8]

Esempi:

If you were: Someone who “likes” patriotism



Heart of Texas
Sponsored · 🌐 Like Page

Border Patrol agents in South Texas arrested an illegal alien from Honduras that had previously been deported and convicted of Rape Second Degree. Thanks to Obama's and Hillary's policy, illegals come here because they wait for amnesty promised. The wrong course had been chosen by the American government, but all those politicians are too far from the border to see who actually sneaks through it illegally. Rapists, drug dealers, human traffickers, and others. The percent of innocent poor families searching for a better life is too small to become an argument for amnesty and Texas warm welcome.

DON'T MESS WITH TX BORDER PATROL

ALWAYS GUIDED BY GOD

3.1K Reactions 89 Comments 1.2K Shares

Like Comment Share

Fig.1

If you were: A Bernie Sanders supporter



Born Liberal
Sponsored · 🌐 Like Page

"If you ask me about the Clinton Foundation, do I have a problem when a sitting secretary of state and a foundation run by her husband collects many millions of dollars from foreign governments, many governments which are dictatorships... yeah, I do," Sanders said in an interview with Jake Tapper on CNN. So, Bernie has the point. I appreciate the way he maintains his independent opinion in spite of all the pressure that is on him as the "second Democratic option". Because that ... See More

**BERNIE SANDERS:
CLINTON FOUNDATION
IS A "PROBLEM"**

124 Reactions 5 Comments 21 Shares

Like Comment Share

Fig. 2

Fig. 1

If you were: Someone who supports veterans



The image shows a Facebook post from the page "Heart of Texas". The post is sponsored and features a headline: "HILLARY CLINTON HAS A 69 PERCENT DISAPPROVAL RATE AMONG ALL VETERANS". The text of the post reads: "Hillary Clinton has a 69 percent disapproval rate among all veterans. Indeed, there are many reasons for it. First of all, Benghazi: four people died on her watch and she did not send help. Secondly, Hillary refused to apologize to all veterans, when she has made several remarks about veterans 'embellishing' the situation at the VA. Finally, Hillary is the only one politician (except Barack Obama) who is despised by the overwhelming majority of American veterans. If Hillary becomes the President of the US, the American army should be withdrawn from Hillary's control according to the amendments to the Constitution." Below the text is a photograph of a soldier in full combat gear kneeling in a field. At the bottom of the post, it shows 800 Reactions, 84 Comments, and 308 Shares, along with Like, Comment, and Share buttons.

Fig. 3

1.2.2. Cambridge Analytica

Cambridge Analytica è stata fondata nel 2013 da Robert Mercer, uno dei finanziatori del sito d'informazione di estrema destra *Breitbart News*, diretto da Steve Bannon (consigliere e stratega di Trump durante la campagna elettorale e poi alla Casa Bianca). *Cambridge Analytica* è specializzata nel raccogliere dai Social Network un'enorme quantità di dati sui loro utenti: quanti "Mi piace" mettono e su quali post, dove lasciano il maggior numero di commenti, il luogo da cui condividono i loro contenuti e così via. Queste informazioni sono poi elaborate da modelli e algoritmi per creare profili di ogni singolo utente, con un approccio simile a quello della "psicometria", il campo della psicologia che si occupa di misurare abilità, comportamenti e più in generale le caratteristiche della personalità. Più "Mi piace", commenti, *tweet* e altri contenuti sono analizzati, più è preciso il profilo psicometrico di ogni utente.

Cambridge Analytica dice di avere sviluppato un sistema di "microtargeting comportamentale", che tradotto significa: pubblicità altamente personalizzata su ogni singola persona. I suoi responsabili sostengono di riuscire a far leva non solo sui gusti, come fanno già altri sistemi analoghi per il marketing, ma sulle emozioni degli utenti.

Il modello è studiato per prevedere e anticipare le risposte degli individui. Kosinski sostiene che siano sufficienti informazioni su 70 "Mi piace" messi su Facebook per sapere più cose sulla personalità di un soggetto rispetto ai suoi amici, 150 per saperne di più dei genitori del soggetto e 300 per superare le conoscenze del suo partner. Con una quantità ancora maggiore di "Mi piace" è possibile conoscere più cose sulla personalità rispetto a quante ne conosca il soggetto [9].

Cambridge analitica ha creato profili psicometrici di circa 87 milioni di utenti di Facebook prevalentemente americani [10].

Nell'estate del 2016, il comitato di Trump affidò a Cambridge Analytica la gestione della raccolta dati per la campagna elettorale.

1.2.3. La strategia comunicativa di Trump

Il giornale americano *Washington Post* ha contato tutte le affermazioni false e ingannevoli che Trump ha annunciato pubblicamente dal giorno in cui si è insediato alla Casa Bianca, in un anno sono già oltre 2000, con una media di quasi 6 affermazioni false al giorno [11]. Ad esempio, per ben 55 volte ha affermato che la sua riforma fiscale è il più grande taglio delle tasse nella storia degli stati uniti, quando in realtà è solo all'ottavo posto, addirittura 2 posti sotto l'amministrazione Obama [12].

La notizia più condivisa su Facebook durante la campagna elettorale americana è stata "Il Papa appoggia Donald Trump", notizia falsa che però è stata condivisa quasi 1 milione di volte [13].

Pope Francis Shocks World, Endorses Donald Trump for President, Releases Statement

TOPICS: Pope Francis Endorses Donald Trump



Fig. 4

1.3. Le bolle Social

Una "bolla di filtraggio" è il risultato del sistema di personalizzazione dei risultati di ricerche su siti che registrano la storia del comportamento dell'utente. Questi siti sono in grado di utilizzare informazioni sull'utente (come la posizione, i click precedenti, le ricerche passate) per scegliere selettivamente tra tutte le risposte quelle che vorrà vedere l'utente stesso. L'effetto è di isolare l'utente da informazioni che sono in contrasto con il suo punto di vista, effettivamente isolandolo nella sua bolla culturale o ideologica. Esempi importanti sono la ricerca personalizzata di Google e le notizie personalizzate di Facebook.

Circa il 67% degli Americani si informa sui social network [14], il 35% degli italiani usa Facebook come fonte di informazioni [15], leggendo quasi esclusivamente titoli [16].

Questi utenti non approfondiscono le tematiche, non verificano le fonti e prendono ogni affermazione per vera.

Queste bolle hanno fatto sì che nascessero delle vere e proprie comunità di utenti che disdegnano la scienza, credono nel *terrapiattismo* (con gruppi di anche 100.000 individui [17]), teorie del complotto come le *scie chimiche* (ovvero che le scie di condensazione visibili nell'atmosfera terrestre create dagli aerei non siano formate da vapore acqueo ma composte da agenti chimici o biologici, spruzzati in volo per mezzo di ipotetiche apparecchiature montate sui velivoli, per varie finalità).

Queste comunità sono tutte accumulate dall'odio verso le autorità, verso il metodo scientifico e verso la medicina ufficiale.

Alcuni esempi di bufale in Italia:



CONDIVIDI SE SEI INDIGNATO!!!!!!

PUTIN DA UNA LEZIONE DI GIUSTIZIA AL MONDO
 L'uomo della foto a sinistra si chiama AMIN KAMYA SALOUINI profugo, immigrato in Russia dal WAKANDA.
 PUTIN lo ha condannato a **70 anni di lavori forzati in Siberia** (a spese dello Stato di provenienza), per aver derubato un pensionato Russo.
"SE VIENI IN RUSSIA DEVI RISPETTARE CHI TI PORGE LA MANO, ALTRIMENTI LA PENA SARA' DOPPIA!!! SPECIE SE IL REATO E' CONSUMATO AI DANNI DI UNA PERSONA ANZIANA.
TRA L'ALTRO NON CI RISULTA CHE NEL WAKANDA CI SIA UNA GUERRA, PER CUI ESPELLEREMO DALLA RUSSIA TUTTI I FINTI PROFUGHI PROVENIENTI DA QUEL PAESE! E SE IL LORO PRESIDENTE NON E' D'ACCORDO CI SARANNO SPIACEVOLI CONSEGUENZE. QUINDI STIANO IN CAMPANA!"
CONDIVIDI IL GESTO DI PUTIN!!!!

Fig. 5: Esempio di bufala che incita al razzismo e al dissenso verso lo stato. È stata condivisa quasi 30.000 volte, che con una media di 1.000 amici a persona, questa bufala è stata vista potenzialmente da **30 milioni** di utenti italiani.

SOLO IN ITALIA



Mustafa
 36 anni, congolese
 pregiudicato
 vive in un bilocale
 messo a disposizione
 dal comune in
 un'area...

Leonardo
 78 anni, piemontese
 incensurato
 da due mesi vive in
 un parco perché lo
 Stato gli ha pignorato
 la casa...

Claudio [redacted]
 😂😂😂 Luciano, quello a sinistra è un giocatore di pallacanestro, Shaquille O'Neil e l'altro è Leonardo di Caprio, nel film REVENANT. Ancora credi in queste bufale? 😂

17 h Mi piace Rispondi

Luciano [redacted]
 CLAUDIO anche se sono bufale nella realtà e' così. Non conosco questi personaggi non conosco questo film ma in italia non siamo lontani da queste realtà'

14 h Mi piace Rispondi

Fig. 6: Bufala che incita al dissenso verso lo stato, con commento di un utente che dice "anche se è una bufala in Italia non siamo lontani da questa realtà".

Questo è Luca Boldrini ed è il figlio del fratello della presidente della camera Laura Boldrini. Pur avendo solo il diploma, è stato assunto

[redacted]

Boh, io l'ho vista e condivisa, non mi sono soffermata molto a controllare le fonti. In ogni caso, la cosa non mi stupirebbe. È noto che ci siano casi del genere in Italia. Magari la bufala va presa cmq come testimonianza di una realtà ESISTENTE e che si perpetua da sempre.

Fig. 7: Esempio di commento di un utente sotto un post di una bufala svelata, in cui dice "non mi sono soffermata a controllare le fonti. E' comunque noto che ci siano questi casi in Italia".

PosteMobile 61% 00:43

Vegana chiede aiuto per eliminare pidocchi senza ucciderli

Salve a tutti. Ho urgentissimo bisogno di un consiglio da parte dei più "esperti".
 Sono vegana soltanto da un anno e mezzo e mi trovo in una situazione mai affrontata prima, per cui ho bisogno di un consiglio immediato. Mia figlia di 9 anni è tornata da scuola con i pidocchi. Il giorno dopo, l'ho riportata a scuola, ma l'insegnante mi ha detto che non la poteva ammettere in classe con i pidocchi e poi mi ha consigliato uno shampoo specifico da acquistare in farmacia per eliminare gli insetti e le uova. Però, guardando su internet gli inci degli shampoo, ho capito che avrebbero UCCISO i pidocchi. Non so davvero come fare! E' da 3 giorni che tengo la bambina a casa e lei, giustamente, si lamenta, ma io non me la sento di ucciderli... non so proprio come affrontare la situazione. Nel frattempo lei sta perdendo giorni di scuola e questo non va affatto bene! Non ci sarebbero dei metodi "naturali" per allontanare i pidocchi? L'unica alternativa che mi viene in mente è di rasarle i capelli, ma lei dice che non vuole.
 Grazie a tutti,
 vi auguro una buona giornata! : D

Fig. 8: Madre che chiede aiuto in un forum su come scacciare i pidocchi della figlia senza consultare il medico perché non crede nella medicina ufficiale.

“Non c’è più il tempo di approfondire le questioni, ma attraverso i Social si cerca di farsi subito un’idea. Un’idea che sarà più forte quanto più la potremmo scagliare come una pietra contro qualcuno, l’importante è attaccare qualcuno. Nel tempo si è svuotato il dibattito politico di contenuti e lo si è riempito di masse di *haters*, di odiatori che non entrano mai nel merito, che si affrontano tra di loro per slogan.

Per attirare a sé l’elettore, questi slogan lo devono colpire alla pancia, sui sentimenti. “

Roberto Saviano [18]

1.4. Non è un problema recente (Es. Cicerone 64 a.C.)

Queste tecniche di conquista del consenso tramite la diffusione di notizie false non sono state inventate dai politici degli ultimi anni, ma ci sono sempre state.

Ad esempio, nel 64 a.C. Marco Tullio Cicerone era candidato alla carica di Console della Repubblica Romana. Suo fratello Quinto, per aiutarlo a vincere le elezioni, scrisse il *Commentariolum Petitionis*, manuale per la campagna elettorale. Più che di consigli politici, esso tratta di strategie di comunicazione, come quella di rispondere sempre “sì” alle richieste della gente anche se si sapeva già in partenza di non riuscire a soddisfarle. Tra questi consigli si legge anche: “*Abbi cura che sorga nei confronti dei tuoi avversari un sospetto, appropriato al loro comportamento o di colpa, o di lusso, o di sperpero*”. Non una verità quindi, ma un’accusa verosimile, un'accusa che rovinasse la reputazione dell'avversario. In altre parole, il fratello di Cicerone stava dicendo che le *Fake News* diffamatorie in campagna elettorale sono ammesse, anzi sono prassi utile. [19]

Già allora non importava smontare le argomentazioni del rivale, ma gettare il sospetto su di loro tramite notizie false.

1.5. Tecniche Manuali: *fact cheking*, *debunkers* e il caso *Butac*

Nel giornalismo la verifica dei fatti (anche verifica delle fonti [20], spesso arbitrariamente indicato con l'anglicismo *fact checking*) è il *lavoro di accertamento degli avvenimenti citati e dei dati usati in un testo o in un discorso* [21]. Questa pratica si applica in particolare alle informazioni date dai politici e, anche come autoverifica, alle notizie diffuse dai mezzi di comunicazione.

Negli ultimi anni stanno emergendo diversi siti di *debunking* come www.snopes.com e l'italiano www.butac.it.

Un *debunker* (in italiano: demistificatore o disingannatore) è *un individuo che mette in dubbio o smaschera ciarlatanerie, bufale, affermazioni false, esagerate, antiscientifiche, dubbie o pretenziose*.

Purtroppo, questi siti sono gestiti da persone che scrivono articoli non per lavoro ma per hobby e passione personale nel tempo libero, inoltre essendo siti web indipendenti e non ufficiali o regolamentati dallo stato sono soggetti alla legge.

Ad esempio, nella mattina di venerdì 6 aprile 2018 il sito www.butac.it era stato sottoposto a sequestro preventivo da parte della procura di Bologna, in seguito a una querela per diffamazione per un articolo pubblicato due anni e mezzo prima, nel quale erano segnalate dichiarazioni di un oncologo che promuoveva la “*medicina olistica*”, pratica senza fondamento scientifico. Invece di sequestrare il solo articolo è stato reso irraggiungibile tutto l'archivio [22].

In sostanza basta fare una semplice querela per “spegnerne” questi siti per giorni o mesi prima di poterli rimettere online.

Tutti questi fatti messi assieme, ovvero che:

- le *Fake News* sui Social Network si diffondono rapidamente, con un picco nelle prime 2 ore.
- C'è un forte interesse economico e politico alla loro diffusione.
- Le tecniche manuali come i siti di *debunking* e *fact checking* sono lente, non applicabili in larga scala e facilmente bloccabili.
- Le bolle Social fanno sì che le persone si isolino nelle loro convinzioni e pregiudizi che sono facilmente manipolabili.
- La persona media non ha tempo, voglia, risorse o capacità intellettuali per verificare le fonti e approfondire un tema, quindi prende ogni affermazione per vera.

evidenziano la necessità di un sistema automatico per il rilevamento delle *Fake News* e bufale.

2. Principali tecniche automatiche ad oggi

Presentiamo ora le principali tecniche automatiche che sono tutt'oggi utilizzate per il rilevamento di *Fake News*.

2.1. Riconoscimento tramite i like messi ai post di Facebook

Introduzione [23]

Le bufale possono essere identificate con grande precisione sulla base degli utenti che interagiscono con loro: in particolare, concentrando in particolare la nostra attenzione su Facebook, rispondiamo alla seguente domanda di ricerca: *si può identificare una bufala in base agli utenti che hanno messo "Mi piace" ad essa?*

L'obiettivo è quindi quello di classificare i post come bufale o non-bufale.

Vengono presentate due tecniche di classificazione.

Una consiste nell'applicare la *regressione logistica*, considerando l'interazione dell'utente con i post. L'altra tecnica consiste in un nuovo adattamento degli algoritmi booleani di *crowdsourcing (BLC)*, ovvero gli utenti forniscono etichette *True/False* per i post, indicando se viola le linee guida della community, senza però fare un presupposto sulla assoluta affidabilità degli utenti che segnalano è stata ottenuta una precisione di classificazione superiore al 99% anche quando il set di allenamento contiene meno dell'1% dei post.

Dataset

Consiste in tutti i post pubblici e "Mi piace" di un elenco di pagine Facebook selezionate durante il secondo semestre del 2016: dal 1° luglio 2016 al 31 dicembre 2016.

Le pagine di Facebook sono suddivise in due categorie: fonti di notizie scientifiche e fonti di notizie di cospirazione.

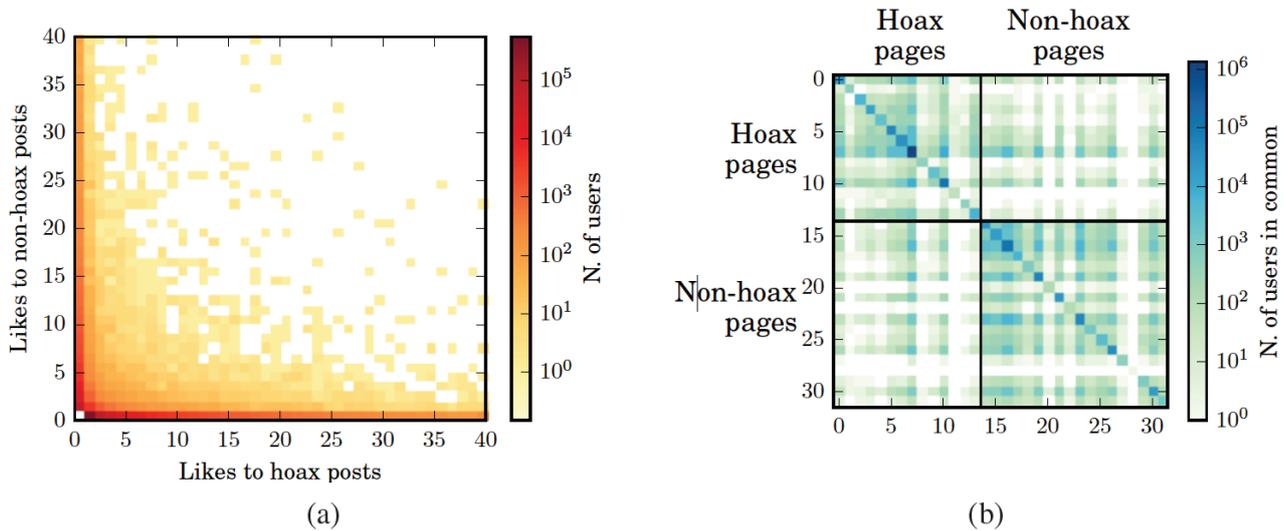
2 Dataset:

- *Dataset completo*: è composto da 15.500 post di 32 pagine (14 cospirative e 18 scientifiche), con oltre 2.300,00 "Mi piace" di oltre 900.000 utenti. Tra i post, 8.923 (57,6%) sono bufale e 6.577 (42,4%) non-bufale.

Gli utenti possono essere divisi in tre categorie in base a cosa hanno messo *like*: i) solo quelli a cui piacciono i post bufale, ii) solo quelli a cui piacciono i post non-bufale, e iii) quelli a cui piacciono almeno un post appartenente a una pagina di bufale e uno appartenente a una pagina non-bufale.

- *Dataset di intersezione*: comprende gli utenti a cui sono piaciuti sia i post bufala che quelli non-bufala ed i suddetti post.

Data ogni coppia di pagine, si studiano quanti utenti hanno messo *like* almeno un post di una pagina e un post dell'altra pagina. Ne risulta che le pagine di bufale hanno più utenti in comune con altre pagine di bufale che con pagine di non-bufale.



Tecnica 1: Classificazione tramite regressione logistica

È un *apprendimento supervisionato*, un *problema di classificazione binaria* in cui si classificano i post sulla base delle loro caratteristiche, cioè sulla base delle preferenze degli utenti.

Il modello di regressione logistica apprende un peso w_u per ogni utente $u \in U$.

La probabilità p_i che un post i non sia una bufala è data da $p_i = \frac{1}{1+e^{-y_i}}$, dove $y_i = \sum_{u \in U} x_{iu} w_u$.

$w_u > 0$ indica che l'utente u ha messo per la maggior parte *like* a non-bufale.

La regressione logistica si adatta bene a problemi con un insieme di caratteristiche molto ampio e uniforme, infatti in questo caso l'applicazione reale coinvolgerebbe fino a centinaia di milioni di utenti.

Tecnica 2: Classificazione tramite crowdsourcing dell'etichetta armonica booleana

Proponiamo algoritmi derivati dal *crowdsourcing*, precisamente dal problema del *crowdsourcing* (BLC) dell'etichetta booleana.

Nel problema BLC, gli utenti forniscono etichette *True/False* sui post, indicando se viola le linee guida della community.

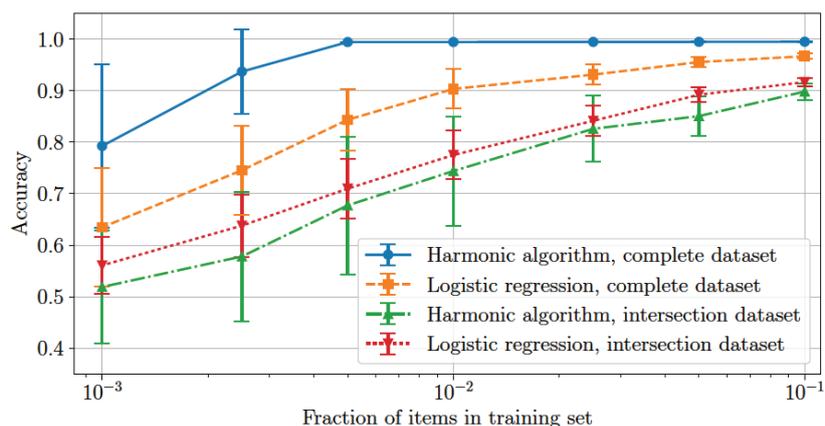
Consideriamo il "*Mi piace*" come mettere TRUE al post.

Il BLC standard presuppone che le persone siano più inclini a dire la verità che a mentire.

Nel nostro contesto, non possiamo presumere che gli utenti siano più propensi a dire la verità, cioè che preferiscono i post non-bufala.

I post bufala potrebbero avere più "*Mi piace*" di quelli non-bufale.

Faremo affidamento su una serie di post di training per i quali è nota la verità di base.



2.2. Analisi della coerenza tra titolo e testo di un articolo

Introduzione [24]

L'obiettivo è stimare la posizione di un articolo rispetto a un titolo; in particolare, il titolo può essere: d'accordo, in disaccordo, può mettere in discussione o essere del tutto estraneo al testo del corpo dell'articolo.

Questo modello è in grado di classificare correttamente le coppie titolo-corpo come correlate o non correlate con una precisione del 97%, è inoltre in grado di identificare correttamente la posizione (concorda, non concorda, discute, ecc...) con il 93% di precisione.

Dataset utilizzato

Il dataset consiste in 49.972 coppie di titoli e testo del corpo, ciascuno con un'etichetta: non correlati, d'accordo, in disaccordo o discussione.

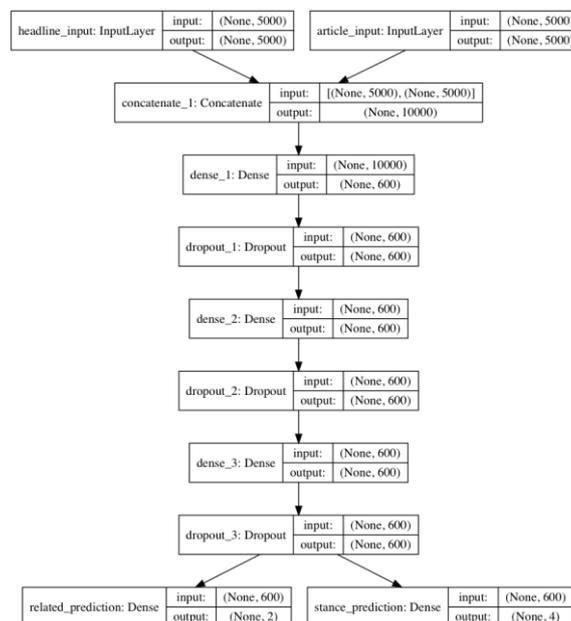
In totale si hanno 1683 articoli e 49972 titoli.

Architettura del modello

Il modello *multilayer perceptron* (con *multilayer perceptron* si intende una classe di *feedforward artificial neural network* nella quale le connessioni fra le varie unità non costituiscono cicli, come invece avviene nelle *recurrent neural networks*) *bag-of-words* (il *bag-of-words*, o *modello della borsa di parole*, è un metodo utilizzato nell'*Information Retrieval* ignorando l'ordine delle parole. In questo modello, ogni documento è considerato in quanto contiene parole, analogamente a una borsa; ciò consente una gestione di queste basata su liste, dove ogni borsa contiene determinate parole di una lista) utilizza un vocabolario di dimensioni fisse dei *token* più comuni. Questi token sono ridotti a un vettore di parole (BoW, per l'appunto) (la cui dimensione è la dimensione del vocabolario), che indica la distribuzione di token senza conservare alcuna informazione sull'ordinamento o sulla struttura.

Il vettore *bag-of-words* viene quindi passato attraverso diversi livelli nascosti con *dropout*, cioè con l'eliminazione graduale di unità attraverso livelli sia visibili che non, della rete.

Infine, si usano *layer* separati di una *softmax* (funzione esponenziale normalizzata) per fare due classificazioni: una decisione binaria sulla correlazione tra articolo e titolo, e una classificazione del titolo e dell'articolo come in accordo, in disaccordo, in discussione (nessuna posizione) o completamente estranei.



2.3. Riconoscimento tramite l'analisi del testo

Introduzione [25]

Si tratta di un classificatore in grado di prevedere se una notizia è falsa basandosi solamente sul suo contenuto. Possiede una precisione del 79%.

Dataset

Le etichette binarie 0/1 sono state assegnate rispettivamente agli articoli di notizie reali/fasulle. La collezione è costituita da circa 63.000 articoli.

Modello

La *RNN* (una *Recurrent Neural Network* è una particolare classe di *artificial neural network* nella quale le connessioni tra le unità formano un grafo diretto/orientato) è stata inizialmente implementata per processare ogni articolo nella sua interezza. Tuttavia, quando si è arrivati ad utilizzare un training sample da 26,970 parole, il costo computazionale troppo elevato ha reso impossibile il completamento della fase di training della rete. Di conseguenza è stato imposto un vincolo di 200 step temporali (consideriamoli come 200 passi da eseguire in sequenza), l'introduzione di un *iperparametro* che è un parametro di una distribuzione di probabilità a priori, nonché l'introduzione di sample per l'allenamento più corti costituiti da vettori di zeri all'inizio: in questo modo la rete sarà in grado di esaminare le parole mentre si sposta lungo i vari step temporali fino a raggiungere la fine dell'articolo.

Sebbene le RNN possano teoricamente catturare dipendenze a lungo termine, sono molto difficili da addestrare. Le *Gated Recurrent Units (GRUs)* sono progettate in modo da avere una memoria più persistente rendendo così più facile per le RNN acquisire dipendenze a lungo termine.

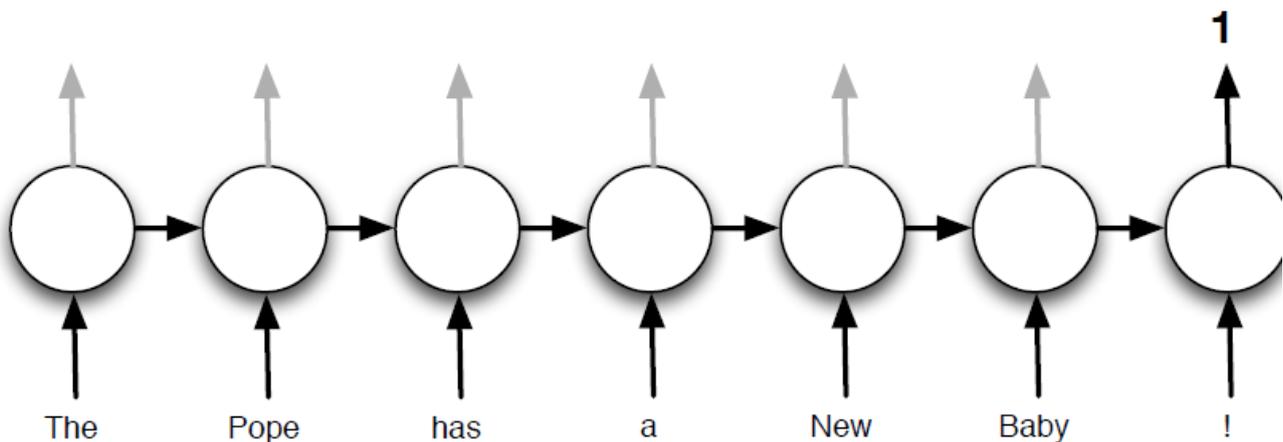


Figure 7: RNN Architecture for the Classification task

Bibliografia

- [1] Tavernisen, S.: As fake news spreads lies, more readers shrug at the truth. *New York Times*, 6 December 2016.
<http://nyti.ms/2lw56HN>
- [2] Decisions Are Emotional, Not Logical: The Neuroscience behind Decision Making. 2017.
<http://bigthink.com/experts-corner/decisions-are-emotional-not-logical-the-neuroscience-behind-decision-making>
Emotion and Decision Making. May 2014
https://scholar.harvard.edu/files/jenniferlerner/files/annual_review_manuscript_june_16_final.final_.pdf
- [3] Oxford dictionary. Word of the year 2016.
<https://en.oxforddictionaries.com/word-of-the-year/word-of-the-year-2016>
- [4] Collins. The Collins word of the year 2017.
<https://www.collinsdictionary.com/it/woty>
- [5] Michela Del Vicario, Alessandro Bessi, Fabiana Zollo, Fabio Petroni, Antonio Scala, Guido Caldarelli, H. Eugene Stanley, and Walter Quattrociocchi. The spreading of misinformation online. *Proceedings of the National Academy of Sciences*, 113(3):554–559, January 2016.
- [6] Alex Law. Post-truth and fake news. 2017.
<https://rke.abertay.ac.uk/ws/portalfiles/portal/8789997>
- [7] The Washington Post. The Facebook ads Russians targeted at different groups. November 2018.
https://www.washingtonpost.com/graphics/2017/business/russian-ads-facebook-targeting/?utm_term=.ff6b090495ae
- [8] The New York Times. Russian Influence Reached 126 Million Through Facebook Alone. October 2017.
<https://www.nytimes.com/2017/10/30/technology/facebook-google-russia.html>
- [9] Il Post. Il caso Cambridge Analytica, spiegato bene. March 2018.
<https://www.ilpost.it/2018/03/19/facebook-cambridge-analytica/>
- [10] Il Sole 24 ore. Facebook: i profili ceduti a Cambridge Analytica sono 87 milioni, 214mila italiani. April 2018.
<http://www.ilsole24ore.com/art/tecnologie/2018-04-04/facebook-profilii-ceduti-cambridge-analytica-sono-87-milioni-212600.shtml>
- [11] The Washington Post. President Trump has made more than 2,000 false or misleading claims over 355 days. 2018.
https://www.washingtonpost.com/news/fact-checker/wp/2018/01/10/president-trump-has-made-more-than-2000-false-or-misleading-claims-over-355-days/?noredirect=on&utm_term=.beea2c2d90b5
- [12] The Washington Post. President Trump's tax cut: Not 'the biggest' in U.S. history. 2017.
https://www.washingtonpost.com/news/fact-checker/wp/2017/11/01/president-trumps-tax-cut-not-the-biggest-in-u-s-history/?utm_term=.60e317b08af1

- [13] **CNBC. Biggest fake news stories of 2016. 2016.**
<https://www.cnn.com/2016/12/30/read-all-about-it-the-biggest-fake-news-stories-of-2016.html>
- [14] **Pew Research Center. News Use Across Social Media Platforms 2017. 2017.**
<http://www.journalism.org/2017/09/07/news-use-across-social-media-platforms-2017/>
- [15] **Censis: il 35,5% degli italiani si informa su Facebook. 2016.**
<http://www.fnsi.it/censis-il-355-degli-italiani-si-informa-su-facebook-tengono-radio-e-tv-ancora-in-calco-i-lettori-di-quotidiani>
- [16] **Forbes. 59 Percent Of You Will Share This Article Without Even Reading It. 2016.**
<https://www.forbes.com/sites/jaysondemers/2016/08/08/59-percent-of-you-will-share-this-article-without-even-reading-it/#12ddd7432a64>
- [17] **Gruppi facebook di persone che credono che la terra è piatta**
<https://www.facebook.com/groups/OfficialFlatEarthAndGlobeDiscussion/about/>
<https://www.facebook.com/groups/1538327313126456/>
- [18] **Roberto Saviano. Intervento Roberto Saviano fake news e post verità. January 2018.**
<https://www.raiplay.it/video/2018/01/Intervento-di-Roberto-Saviano-fake-news-e-postverita224---18012018-118988ae-2cb1-496b-9587-d277344d47b4.html>
- [19] **Il manuale per la campagna elettorale di Marco Tullio Cicerone (Commentariolum petitionis)**
- [20] **European Journalism Observatory: Verifica delle fonti, una guida per il digitale. .**
<http://it.ejo.ch/digitale/verification-handbook-ejc-verifica-ugc>
- [21] **Varesenews: "La cultura della verifica dei fatti è rivoluzionaria". 2012.**
<http://www3.varesenews.it/comuni/varese/articolo.php?id=248186>
- [22] **Il Post. Il sito BufaleUnTantoAlChilo è sotto sequestro. April 2018.**
<https://www.ilpost.it/2018/04/06/butac-sequestro-preventivo/>
- [23] **Eugenio Tacchini, Gabriele Ballarin, Marco L. Della Vedova, Stefano Moret, Luca de Alfaro. Some Like it Hoax: Automated Fake News Detection in Social Networks. April 2017.**
<https://arxiv.org/abs/1704.07506>
- [24] **Richard Davis, Chris Proctor. Fake News, Real Consequences: Recruiting Neural Networks for the Fight Against Fake News.**
<https://web.stanford.edu/class/cs224n/reports/2761239.pdf>
- [25] **Samir Bajaj. "The Pope Has a New Baby!" Fake News Detection Using Deep Learning. Winter 2017.**
<https://web.stanford.edu/class/cs224n/reports/2710385.pdf>