



POLITECNICO
MILANO 1863

Fake news e social network

Su Nature Scientific Reports una metodologia per individuarle e capire come si diffondono

Milano, 24 marzo 2020 - Negli ultimi anni le informazioni ingannevoli e dannose, le cosiddette fake news, hanno avuto una crescita esplosiva sui social network con un'influenza sull'opinione pubblica che ha spesso causato gravi reazioni sociali e politiche.

Francesco Pierri, dottorando del corso interdipartimentale in Data Analytics and Decision Sciences del Politecnico di Milano, in collaborazione coi Professori Carlo Piccardi e Stefano Ceri, ha recentemente pubblicato su **Nature Scientific Reports** un lavoro in cui mostra una **metodologia di machine learning che permette di distinguere accuratamente le notizie attendibili da quelle fuorvianti osservando il modo in cui si propagano sui social media**.

Lo studio, partendo da casi noti secondo cui le notizie false si diffondono più velocemente, più in profondità e in modo più ampio rispetto alle informazioni fattuali per effetto di social bots, camere di risonanza e pregiudizi algoritmici e umani, ha evidenziato che le comunità di utenti che diffondono notizie pubblicate su fonti affidabili interagiscono diversamente con esse rispetto agli utenti che condividono notizie da fonti inattendibili.

Oggi i ricercatori sono al lavoro per creare una piattaforma in grado di distinguere in tempo reale le notizie sulla base della loro attendibilità e di individuare in maniera efficace i contenuti dannosi che circolano sui social network. La piattaforma si baserà sulla metodologia sviluppata, combinata con altri tipi di analisi già disponibili.

La metodologia

Il metodo di classificazione automatica dell'attendibilità della notizia è basato sul machine learning e sfrutta le discrepanze nella diffusione tra notizie più o meno veritiere su Twitter, considerando le reti di diffusione originatesi dalle numerose interazioni tra utenti che condividono notizie on-line. Questa metodologia permette di bypassare la complessità di un approccio basato sul contenuto in quanto è indipendente dal testo della notizia, e potrebbe essere combinata con altri tipi di analisi già disponibili per individuare efficacemente i contenuti dannosi che circolano sui social network.

Dato l'ampio dibattito nella comunità scientifica su una definizione unica di cattiva informazione, i ricercatori hanno focalizzato la loro attenzione ad una vasta gamma di notizie fuorvianti e potenzialmente dannose che includono ma non si limitano alle "fake news".

Hanno quindi raccolto e monitorato la diffusione di migliaia di news pubblicate nel 2019 su decine di giornali on-line statunitensi, sia quelli cosiddetti di “disinformazione” che quelli tradizionali e più attendibili. Nell’arco di tre settimane sono stati raccolti più di 3 milioni di tweets condivisi da circa mezzo milione di utenti unici, e sono state costruite circa 12mila reti di diffusione associate ad altrettante news.

Rappresentando ogni rete con un insieme di features topologiche (e.g. numero di componenti connesse, diametro della rete) ed utilizzando degli algoritmi classici di machine learning per fare classificazione (e.g. regressione logistica e k-Nearest Neighbors), sono stati in grado di ottenere un’accuratezza fino al 94% sulle reti più grandi.

Nonostante il diverso orientamento politico della fonte, è stato osservato che le notizie tradizionali e attendibili si diffondono diversamente da quelle meno attendibili e di contro-informazione, e che il classificatore è in grado di distinguerle accuratamente.

Complessivamente, i risultati mostrano che le comunità di utenti che condividono news poco attendibili tendono ad essere più connesse tra loro e le notizie meno attendibili si diffondono più in profondità ma con un’audience inferiore rispetto alle notizie tradizionali.

In un altro lavoro, attualmente in peer-review, la stessa metodologia è stata applicata al contesto italiano con risultati simili.

L’articolo “Topology comparison of Twitter diffusion networks effectively reveals misleading information” è disponibile online all’indirizzo www.nature.com/articles/s41598-020-58166-5

Francesco Pierrì

Francesco Pierrì è un dottorando al 2 anno con una borsa ministeriale del programma multi-disciplinare “Data Analytics and Decision Sciences” (Corso di Dottorato Interdipartimentale del Dipartimento di Elettronica, Informazione, Bioingegneria, Dipartimento di Ingegneria Gestionale e Dipartimento di Matematica). La sua tesi di dottorato, sviluppata coi Professori Stefano Ceri e Fabio Pammolli, sta investigando la diffusione di informazione fuorviante e dannosa sui social network.

Ha pubblicato 5 articoli scientifici sul tema, tre su riviste internazionali e 2 in conferenze internazionali, e 2 articoli sono in fase di peer-review.