|  |  |
| --- | --- |
| Comprensione del comportamento scorretto nella discussione di COVID19 sui social network e sui media online | |
|  | |
| Studente/i | Relatore |
| Nogara Gianluca | Giordano Cremonese Silvia |
| Correlatore |
| Luceri Luca |
| Committente |
| - |
| Corso di laurea | Modulo |
| Ingegneria Informatica TP | C10344 Progetto di diploma |
| Anno |  |
| 2020/2021 |  |

|  |  |
| --- | --- |
| Data |  |
| - |  |

**Indice**

[Abstract 6](#_Toc81581261)

[Progetto assegnato 8](#_Toc81581262)

[1 Analisi 9](#_Toc81581263)

[1.1 Contesto 9](#_Toc81581264)

[1.2 Fruitori 9](#_Toc81581265)

[1.3 Requisiti 9](#_Toc81581266)

[1.3.1 Contenuti 9](#_Toc81581267)

[1.4.3 Rispetto della privacy 10](#_Toc81581268)

[1.5 Obiettivi 11](#_Toc81581269)

[2 Stato dell’arte 12](#_Toc81581270)

[2.1 Analisi dei dati 12](#_Toc81581271)

[2.2 Linguaggi di programmazione e ambiente di sviluppo 13](#_Toc81581272)

[2.3 Librerie 13](#_Toc81581273)

[2.4 Strumenti utilizzati 14](#_Toc81581274)

[3 Implementazione 14](#_Toc81581275)

[3.1.1 Limitazioni hardware 15](#_Toc81581276)

[4 Operazioni sui dati 15](#_Toc81581277)

[5 Piani di lavoro 15](#_Toc81581278)

[5.1 GitHub 15](#_Toc81581279)

[6 Conclusioni 15](#_Toc81581280)

[6.1 Problemi riscontrati 15](#_Toc81581281)

[6.2 Implementazioni future 15](#_Toc81581282)

[7 Fonti 16](#_Toc81581283)

**Indice delle figure**

Abstract

I social network e i media online (OSNEM) sono tecnologie basate su computer che consentono agli utenti di creare contenuti e intrattenere relazioni sociali.

Le OSNEM sono rapidamente cresciute dall’essere semplicemente un canale di aggregazione all’essere un fenomeno globale che ha suscitato un cambio di paradigma della nostra società, trasformando il modo in cui le persone accedono alle notizie, condividono opinioni, fanno affari e fanno politica.

In uno scenario del genere, l'accuratezza, la veridicità e l'autenticità del contenuto condiviso sono ingredienti necessari per mantenere una sana discussione online.

Tuttavia, negli ultimi tempi, le OSNEM hanno dovuto affrontare una notevole crescita di account e attività dannose, che minano intenzionalmente l'integrità delle conversazioni online condividendo, nelle piattaforme OSNEM, informazioni false e provocatorie per influenzare l'opinione pubblica e creare conflitti su questioni sociali o politiche.

Ciò è stato, e continua ad essere, drammaticamente vero per la discussione sul COVID19, che è stata minata da campagne di manipolazione e disinformazione a livello globale. È quindi necessario studiare i comportamenti peculiari delle entità dannose nel dibattito COVID19 per fornire un cambiamento sociale verso una migliore comprensione (alfabetizzazione) delle reti sociali.

Questo include il rilevamento di malintenzionati, l'identificazione di informazioni di scarsa credibilità e notizie false, la consapevolezza degli utenti e il controllo dei comportamenti scorretti.

Social networks and online media (OSNEMs) are computer-based technologies that allow users to create content and engage in social relationships.

OSNEMs have rapidly grown from being simply a channel of aggregation to being a global phenomenon that has sparked a paradigm shift in our society, transforming the way people access news, share opinions, do business, and make policy.

In such a scenario, accuracy, truthfulness and authenticity of shared content are necessary ingredients to maintain a healthy online discussion.

However, in recent times, OSNEMs have faced a significant growth of malicious accounts and activities, which intentionally undermine the integrity of online conversations by sharing, in OSNEM platforms, false and provocative information to influence public opinion and create conflict on social or political issues.

This has been, and continues to be, dramatically true of the COVID19 discussion, which has been undermined by global campaigns of manipulation and misinformation. Therefore, it is necessary to study the peculiar behaviors of malicious entities in the COVID19 debate to provide social change towards a better understanding (literacy) of social networks.

This includes detection of malicious actors, identification of information of low credibility and fake news, user awareness, and control of misbehavior.

Progetto assegnato

Lo scopo di questo progetto è quello di analizzare entità malevoli presenti in un dataset pubblico su COVID19. L’ OSNEM di riferimento è Twitter e l’attività di analisi e classificazione delle entità malevoli (come bot e trolls) viene fatta tramite diversi algoritmi esistenti.

Il progetto in questione consente quindi di analizzare il comportamento all'interno della discussione e l’impatto che hanno tali entità malevole nella discussione globale.

**COMPITI**

• Identificazione degli strumenti usati per la propagazione di notizie

• Identificazione delle entità nella discussione

• Analisi comportamentale delle entità malevoli e confronto con entità competenti

**OBIETTIVI**

L'obiettivo principale di questa tesi è quello di arrivare a distinguere tra entità/azioni corrette e quelle malevoli nel dibattito COVID19.

Identificare quindi tramite quali mezzi e comportamenti le varie entità malevoli influsicono nelle discussioni, analizzando hashtags, link condivisi e attività, facendo una comparazione tra chi ha l’obiettivo di informare correttamente

**TECNOLOGIE**

• Python

• Complex networks

# 1 Analisi

## 1.1 Contesto

Il progetto si inserisce nel contesto del Dipartimento Tecnologie Innovative della SUPSI e in particolare nel corso di Ingegneria Informatica. Si tratta di un progetto interno alla SUPSI, anche se di grande importanza sociale dal momento che le informazioni raccolte possono essere utilizzate per limitare la propagazione di entità malevoli sui social network e sensibilizare su un tema tanto delicato e attuale.

## 1.2 Fruitori

Il progetto in questione è destinanto ad un pubblico intenzionato a informarsi sulle OSNEM nel dibattito da COVID19, e di come le varie entità interagiscono sul tema.

La finalità scientifica del progetto lo rende pubblicabile ed espandibile per nuovi studi di vario genere e operazioni di aggiornamento sui dati.

## 1.3 Requisiti

### 1.3.1 Contenuti

Il principale requisito a livello contenutisco è quello di avere dei costanti feedback visivi dell’analisi effettuata sui dati, in modo da effettuare operazioni di ricerca sempre più dettagliata in funzione di risultati.

Una volta identificate le strategie di condivisione delle informazioni è quindi possibile andare a stabilire di quali tipologie di utenti si tratta e definire quindi nuove metriche di valutazione per la credibilità di tali utenti.

Identificando utenti malevoli è inoltre possibile effettuare studi sulle loro interazioni attive e passive, definendo delle community potenzialmente utili per ricavere ulteriori inormazioni

### 1.4.3 Rispetto della privacy

Per lo svolgimento del progetto sono state utilizzate grosse quantità di dati pubblici ottenuti dalla Twitter API.

Queste informazioni sono relative sia ai profili che ai singoli tweet pubblicati dagli utenti.

Per quanto riguarda i tweets tali informazioni sono ottenute da tweet considerati “Tweet pubblici”, visibili quindi a chiunque (anche senza Twitter).

Immagine che contiene testo

Descrizione generata automaticamente

Figura 1: Esempio della struttura di tweets filtrata su poche colonne

Dal momento che non sono informazioni confidenziali è possibile pubblicarle senza restrizioni.

Per quanto riguarda le informazioni sui profili, esse possono risultare più personali dal momento che contengono dati che possono risultare talvolta sensibili, come nome, posizione e followers.

Immagine che contiene testo

Descrizione generata automaticamente

Figura 2: Esempio della struttura di profili filtrata su poche colonne

Tuttavia si tratta di informazioni pubbliche che, nel caso non si volesse rendere visibili, potrebbero essere rimosse o protette modificando la visibilità del profilo.

Pertanto tutte le informazioni riportate e utilizzate non violano nessuna normativa sulla privacy.

## 1.5 Obiettivi

Gli obiettivi di questo progetto sono i seguenti:

* Investigazione sulle strategie di condivisione (tweet, retweet, risposte)
* Classificazione delle enità malevoli
* Investigazioni sulle attività delle attività malevoli
* Studio degli utenti che interagiscono con tali entità
* Studio della credibilità dei domini
* Comparazione delle attività tra entità malevoli e attendibili

# 2 Stato dell’arte

## 2.1 Analisi dei dati

Come situazione di partenza è stato messo a disposizione un piccolo dataset di 50,000 tweets con i relativi utenti per prendere confidenza con dati eterogenei.

I dataset in questione riguardano i profili utente e i singoli tweet, la struttura dei due dataset è completamente differente, in particolare quello dedicato agli utenti è molto ricco di colonne (ben 46), mentre quello dedicato agli utenti ne ha poco meno della metà (18).

Questa differenza è data dal fatto che, per quanto riguarda i tweet, abbiamo a disposizione molte più informazioni rispetto a dei semplici profili utente.

Di seguito sono mostratie le colonne più utilizzate dei due dataset:

**Tweets**

|  |  |  |
| --- | --- | --- |
| Campo | Tipologia | Descrizione |
| id | int64 | Id univoco del tweet |
| created\_at | object | Data di creazione del tweet |
| user\_id | int64 | Id univoco del profilo |
| user\_screen\_name | object | Nome utente del profilo |
| text | object | Contenuto del tweet |
| retweet\_count | int64 | Numero di retweet del tweet |
| in\_reply\_to\_user\_id | float64 | Id dell’utente che riceve la risposta |
| in\_reply\_to\_screen\_name | object | Username dell’utente che riceve la risposta |
| rt\_user\_id | float64 | Id dell’utente che riceve il retweet |
| rt\_user\_screen\_name | object | Username dell’utente che riceve il retweet |
| hashtags | object | Hashtag usati nel tweet |
| urls | object | Link condivisi nel tweet |

**Profili**

|  |  |  |
| --- | --- | --- |
| Campo | Tipologia | Descrizione |
| id | int64 | Id univoco dell’utente |
| Screen\_name | object | Username dell’utente |
| location | object | Località dell’utente |
| follower\_count | int64 | Numero di follower dell’utente |
| verified | bool | Booleano che indica se il profilo è verificato |
| geo\_enabled | bool | Booleano che indica se il profilo ha una location |

I dati da analizzare risultano eterogenei, sono presenti infatti valori booleani, valori interi, oggetti e valori in virgola mobile, è quindi importante effettuare un preprocessamento per un corretto studio.

## 2.2 Linguaggi di programmazione e ambiente di sviluppo

Dal momento che i dati su cui lavorare sono file con formato .csv di grossa dimensione la scelta del linguaggio è ricaduta su Python, questo anche perché fornisce una grande varietà di librerie e tools per data analysi e complex networks.

Anaconda ha svolto un compito molto importante per quanto riguarda la programmazione in Python e la realizzazione di feedback visivi, infatti dal momento che per un’analisi corretta si ha bisogno di generare report grafici, Anconda contiene il software Jupyter Notebook.

Quest’ultimo è un'applicazione Web open source che rappresenta lo strumento perfetto per effettuare operazioni di data analysis grazie alle celle che consentono di salvare rappresentazioni grafiche e testo senza dover rieseguire necessariamente tutto il codice.

## 2.3 Librerie

Durante lo sviluppo del progetto si è fatto uso di una serie di librerie che hanno semplificato lo svolgimento del progetto, le più importanti hanno dato un contributo significativo e sono state utilizzate con una certa frequenza

### 2.3.1 Pandas

Pandas è tra le librerie più importanti in Python dal momento che è stato sviluppato per la manipolazione e l'analisi dei dati.

In particolare, offre strutture dati e operazioni per manipolare tabelle numeriche e serie temporali.

L’utilizzo di pandas ha consentito lo sfruttamento dei DataFrame per la lettura dei file .csv e la gestione dei relativi dati contenuti.

I DataFrame hanno rappresentato il cuore pulsante del lavoro svolto dal momento che sono altamente gestibili e semplici nel filtraggio.

Lo sfruttamento della libreria ha permesso inoltre l’utilizzo delle Series, ndarray monodimensionali che espongono un’infità di metodi per la manipolazione dei dati. Questa tipologia di strutture sono state largamente utilizzate per effettuare classificazioni basate sul numero di occorrenze.

### 2.3.2 Matplotlib e Plotly

Matplotlib è una libreria per la creazione di grafici ed è stata largmente utilizzata per la rappresentazione grafica di diverse tipologie di classificazioni, quali istogrammi, barplot e distribuzioni varie.

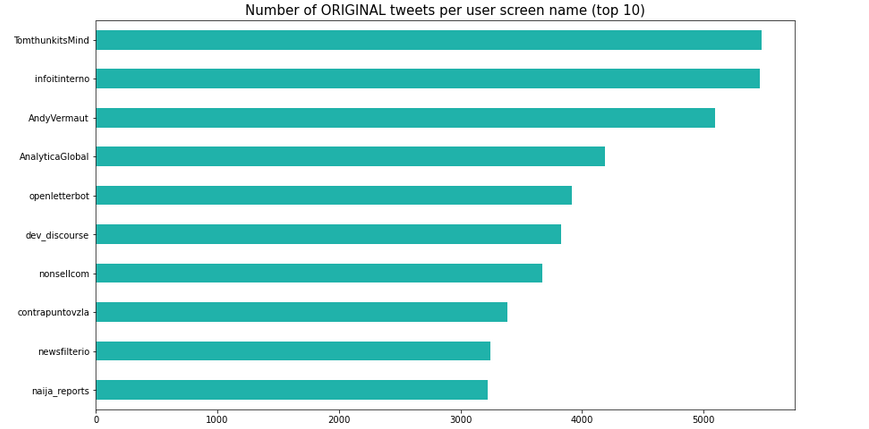
****

Figura 3: Esempio di plot con matplotlib (10 utenti che effettuano più tweet)

Dal momento che i plot generati sono statici essi risultano molto veloci da generare e leggeri da salvare, anche per quantità molto grosse di dati.

In contrapposizione vi è Plotly, libreria di grafici che viene creare grafici interattivi browser-based, che consente quindi di generare report dinamici ma che, in presenza di una quantità di dati significativa, risulta poco efficiente e pesante.

Immagine che contiene mappa

Descrizione generata automaticamente

Figura 4: Esempio di mappa interattiva generata con Plotly

### 2.3.3 NetworkX

Si tratta di una libreria realizzata per semplificare la generazione e lo studio di grafi e reti, è stata largamente utilizzata per classificare gli utenti in community e capire quindi i legami che hanno l’un l’altro.

**Esempio di network**

NetworkX è una libreria adatta per il funzionamento su dati molto consistenti: ad esempio, grafici con più di 10 milioni di nodi e 100 milioni di archi, dal momento che è realizzata in puro Python, infatti è ragionevolmente efficiente, molto scalabile e altamente portabile per l'analisi di reti e social network.

## 2.4 Strumenti utilizzati

Durante lo svolgimento del progetto sono vi è stata la necessità di utilizzare strumenti realizzati da terzi per poter dare una validità ai dati generati o ottenere determinate informazioni.

### 2.4.1 Botometer

Si tratta di un progetto del gruppo OsoMe (Observatory on Social Media) sviluppato all'Indiana University.

OSoMe è una collaborazione tra il Network Science Institute (IUNI), il Center for Complex Networks and Systems Research (CNetS) e la Media School dell'Indiana University.

Botometer è facilmente fruibile tramite l’apposito sito web, raggiungibile al link <https://botometer.osome.iu.edu/>, ma espone anche un’API pubblica.

Tutte le informazioni per l’integrazione di Botometer sono disponibili sulla pagina GitHub del progetto (al link <https://github.com/IUNetSci/botometer-python>).

Per poter utilizzare l’API è necessario passare per RapidAPI e da Twitter.

RapidAPI è l'hub di API più grande al mondo, registrandosi gratuitamente è possibile accedere a Botometer Pro e sottoscrivere il piano adatto per il lavoro.

Immagine che contiene testo, monitor, screenshot, schermo

Descrizione generata automaticamente

Figura 5: Piani offerti da Botometer Pro su RapidAPI

Una volta sottoscritto il piano si ottengono due key, rispettivamente una key\_rapidapi e una consumer\_key, per poter accedere appunto all’API.

Una volta ottenuti i dati da RapidAPI è necessario accedere alla Twitter API e ottenere i permessi da sviluppatore e realizzare quindi una Twitter app.

Anche in questo caso si otterranno delle keys: consumer\_secret, bearer\_token, access\_token e access\_token\_secret.

Tutte queste keys sono essenziali dal momento che vengono utilizzate da Botometer per poter effettuare lo score sugli utenti passati.

È possibile utilizzare Botometer tramite appositi metodi:



Per quanto riguarda lo svolgimento del progetto si è utilizzato uno script già esistente con una serie di credenziali, il quale è stato realizzato in modo tale da effettuare un cambio di credenziali nel momento in cui il numero di richieste disponibili termini, in modo da effettuare quante più richieste possibili giornalmente.

Per l’esecuzione, lo script, richiede che vengs passato un file .csv formato da campi id\_utente,nome\_utente ritornando un file .json contente gli score di Botometer.

- mediabiasfactcheck

# 3 Implementazione

Le prime analisi condotte, seppur molto superfciali, hanno avuto un certo riscontro con analisi precedentemente condotte da studentesse del Master, dal momento che il sample è un subset di un DataFrame molto più grande in utilizzo.

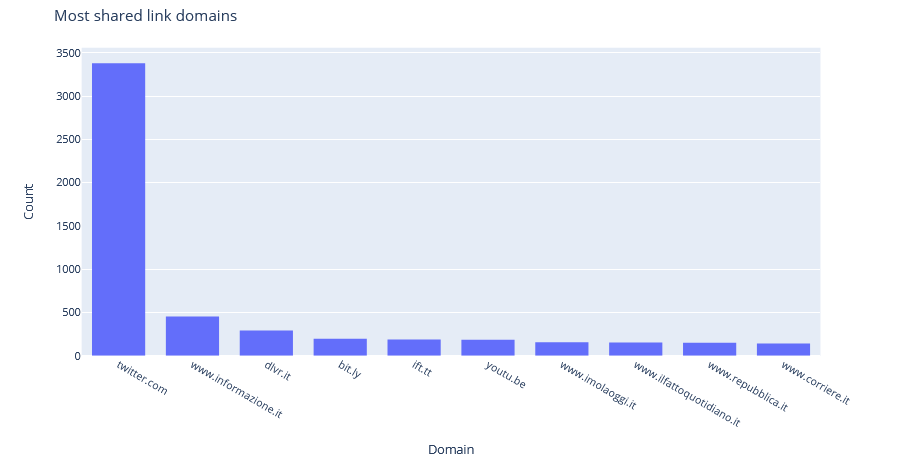


Figura 6: Esempio di analisi sui dati condotta (Domini più condivisi)

Dopo aver condotto studi sul contenuto dei vari campi (come url condivisi e hashtag) si è andato poi a verificare tramite una timeline

### 3.1.1 Limitazioni hardware

# 4 Operazioni sui dati

# 5 Piani di lavoro

## 5.1 GitHub

Il repository è stato uno strumento cruciale per quanto riguarda la sincronizzazione del gruppo e il mantenimento del codice aggiornato e ordinato.

Il lavoro effettuato tramite Git è iniziato il 06/04/2021, con l’iterazione-05, questo perché le prime cinque iterazioni riguardavano la ricerca di informazioni, l’installazione di LL e la familiarizzazione con il software in questione.

Le attività hanno avuto una cadenza settimanale e bisettimanale in funzione del lavoro svolto e delle Issues aperte non ancora svolte.

# 6 Conclusioni

## 6.1 Problemi riscontrati

## 6.2 Implementazioni future

# 7 Fonti

**Python:**

<https://botometer.osome.iu.edu/api>

<https://pypi.org/project/botometer/#description>

<https://cnets.indiana.edu/blog/2020/09/01/botometer-v4/>

<https://blog.quantinsti.com/detecting-bots-twitter-botometer/>

<https://rapidapi.com/developer/dashboard>

<https://networkx.org/documentation/stable/index.html>

**Python:**

https://help.twitter.com/it/safety-and-security/public-and-protected-tweets

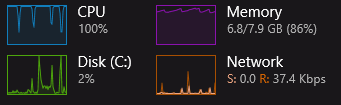
Problemi:

Memoria (parallelizzazione)

bottleneck (lettura in chunk)

Tipi di dati (parsing in date)

ParserError: Error tokenizing data. C error: out of memory



Confronto dei tempi

Da 1000+ secondi a 390 + secondi (users)



Tweets

E users