

Alma Mater Studiorum Università di Bologna
Archivio istituzionale della ricerca

L'OPINIONE ITALIANA SULLA DIDATTICA A DISTANZA ATTRAVERSO UNA 'SENTIMENT ANALYSIS' DEI TWEET NEL BIENNIO 2020-2021.

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

Published Version:

Stracqualursi Luisa (2022). L'OPINIONE ITALIANA SULLA DIDATTICA A DISTANZA ATTRAVERSO UNA 'SENTIMENT ANALYSIS' DEI TWEET NEL BIENNIO 2020-2021. *INDUZIONI*, Anno 2020(60-61), 47-61 [10.19272/202000902003].

Availability:

This version is available at: <https://hdl.handle.net/11585/862960> since: 2022-11-18

Published:

DOI: <http://doi.org/10.19272/202000902003>

Terms of use:

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>).
When citing, please refer to the published version.

(Article begins on next page)

L'OPINIONE ITALIANA
SULLA DIDATTICA A DISTANZA
ATTRAVERSO UNA 'SENTIMENT ANALYSIS'
DEI TWEET NEL BIENNIO 2020-2021

Luisa Stracqualursi*

Abstract · The italian opinion on distance learning through a 'sentiment analysis' of tweets in the two-year period 2020-2021 · In 2020, many efforts have been made to limit the diffusion of COVID-19 infection. Faced with this emergency, everyone's lives were changed through 'social isolation'. So, starting from March 2020, the Italian school system experienced a massive application of Distance Learning. These changes brought about to a certain degree of tension, which have also accelerated the use of innovative educational tools and led the way for a digitalisation of learning. Natural language processing (NLP) is a very powerful technique for social media data processing. Social media can provide the information to understand public opinions of different social phenomena. This paper implements a Sentiment Analysis over 25,000 Italian tweets talking about distance learning from March 2020 to November 2021. The analysis highlighted significant differences between Italian regions and the change in Sentiment due to the transition from an emergency Distance Learning to the Integrated Digital Learning.

Keywords · Distance learning, Sentiment Analysis, Twitter.

1. Introduzione

La Didattica a Distanza (DaD) è una forma di didattica che avviene senza la presenza degli insegnanti e degli studenti in aula, ma avvalendosi di strumenti informatici e di internet. In Italia la DaD è partita nel marzo 2020 (DPCM 11/3/2020) in forme diverse, coinvolgendo primo e secondo ciclo di istruzione e l'istruzione superiore.¹

* luisa.stracqualursi@unibo.it, Dipartimento di Scienze Statistiche, Università di Bologna, Italia.

¹ Fonte <https://www.miur.gov.it/sistema-educativo-di-istruzione-e-formazione>. Il sistema scolastico è organizzato: primo ciclo di istruzione (scuola primaria, scuola secondaria di primo grado); secondo ciclo di istruzione (scuola secondaria di secondo

Dall'anno scolastico 2020/2021, il sistema scolastico italiano ha cercato di avviare un processo di transizione dalla DaD alla Didattica Digitale Integrata (DDI), che avrebbe integrato momenti a distanza (fondati sul digitale) e in presenza. Purtroppo, la riapertura delle scuole è stata graduale, limitata ad alcune regioni e spesso solo temporanea. All'aumentare dei contagi da *Sars-Cov2*, molte regioni sono repentinamente tornate nel corso del 2021 alla didattica a distanza per alcuni ordini o gradi di scuola o per tutti, come è avvenuto nelle cosiddette "zone rosse".

La prima DaD sperimentata è stata quindi una didattica di emergenza, sostitutiva e non integrativa della didattica in presenza, che ha portato un certo disagio agli studenti e alle loro famiglie. Oggi finalmente stiamo sperimentando una didattica a distanza integrata con quella in presenza, non più emergenziale, ma in forma progettata e consapevole.

I dati dei social media possono fornire le informazioni necessarie per comprendere le percezioni del pubblico verso svariati fenomeni sociali (Barber e Liu, 2011). Gli utenti attivi sui social media sono costantemente in crescita e nell'ultimo rapporto Digital del 2020 erano ben 35 milioni in Italia (Kemp, 2020). Ecco perché ha senso ipotizzare che i commenti sui social media possano rivelare l'evoluzione nel sentimento percepito durante l'intera fase attuativa ed evolutiva della DaD.

Tra i vari social media si è scelto Twitter essendo il social che si è sempre contraddistinto per essere il più 'immediato'. I suoi tweet consentono di esprimere e diffondere, in modo veloce e conciso, opinioni, pensieri ma soprattutto emozioni.

La Sentiment Analysis è un sottocampo del Natural Language Processing (NLP) e dell'Intelligenza Artificiale (AI) che attraverso tecniche di machine learning, è in grado di analizzare il sentimento sottostante un messaggio di testo, quantificarlo ed etichettarlo come positivo, negativo o neutro.

In questo paper si applicherà la Sentiment Analysis a circa 25.000 tweet italiani riguardanti la didattica a distanza. Lo scopo è quello di evidenziare l'evoluzione nel sentimento percepito dagli italiani su

grado, istruzione e formazione professionale); istruzione superiore (Università, istituzioni di Alta Specializzazione, ...).

L'opinione italiana sulla didattica a distanza

questo tema durante l'intera fase attuativa ed evolutiva della DaD. L'analisi verrà condotta sia a livello nazionale che regionale, per cogliere possibili differenze dovute alle differenti modalità di attuazione della DaD nelle varie regioni.

2. I dati

I dati relativi ai tweet italiani sono stati forniti dalla piattaforma 'trackmyhashtag.com' ed estratti a partire da marzo 2020 fino a novembre 2021 usando come keyword per l'estrazione le parole 'didattica a distanza'.

Il dataset è composto da 25.180 tweet e contiene 19 variabili. Le Tabelle 1 e 2 riportano le caratteristiche del dataset. Vi sono dati mancanti solo sulle variabili 'Location' e 'User bio'. Questo perché non sempre gli utenti di Twitter inseriscono la loro posizione geografica e la breve biografia.

Tabella 1. Caratteristiche dei dati estratti – variabili di scala.*

<i>variabili di scala</i>	<i>Descrizione</i>	<i>mean</i>	<i>Std</i>	<i>min</i>	<i>25%</i>	<i>50%</i>	<i>75%</i>	<i>max</i>
Retweets Received	# re-tweet ricevuti	1,4	17,9	0	0	0	0	1.930
FavouritesReceived	# like ricevuti	6,9	108,3	0	0	0	1	10.310
User Followers	# followers dell'utente	66.900	386.598	0	243	976	4.338	24.027.220
User Following	# following dell'utente	1.161,9	3.102,7	0	157	468	1.202	185.869
FavouritesCount	# totale like dell'utente	37.558,4	84.214,7	0	308	6.632	36.979	1.305.615
StatusesCount	# totale tweet dell'utente	200.231	628.295	1.000.000	7.118	27.305	87.232	3.747.943

(*) *nessun valore mancante*

Tabella 2. Caratteristiche dei dati estratti – variabili categoriche.

<i>variabili categoriche</i>	<i>Descrizione</i>	<i># Nan</i>	<i>modalità</i>
Id	identificativo tweet	0	25180
URL	URL tweet	0	24892
Date	data del tweet	0	12965
Text	testo del tweet	0	12856
Type	tipologia di tweet (re-tweet, tweet, replay)	0	3
Client	fonte del tweet (Android, iPhone, web app,...)	0	335
Location	posizione geografica dell'utente	15611	20
UserId	identificativo utente	0	25180
Username	nome account utente	0	13851
UserBio	breve biografia utente	3866	11375
Verified or Non-Verified	utente verificato o non verificato	0	2
Profile URL	indirizzo del profilo utente	0	13851
User Account Creation Date	data creazione dell'account utente	0	13849

Le Figure 1 e 2 riportano i valori assoluti e le percentuali delle principali variabili categoriche. I dati contenuti sono prevalentemente tweet e re-tweet (Fig. 1a), solo il 5,94% è relativo a messaggi di ‘reply’.

La maggior parte degli utenti Twitter non sono verificati (Fig. 1b) e la fonte del tweet è generalmente lo smartphone e la web app di Twitter (Fig. 2a). Osservando invece i nomi degli utenti che twittano maggiormente appare evidente che sono testate giornalistiche, personaggi politici e notiziari di informazione (Fig. 2b) e che la proporzione maggiore di tweet in Italia proviene dal Lazio e dalla Lombardia (Fig. 2c).

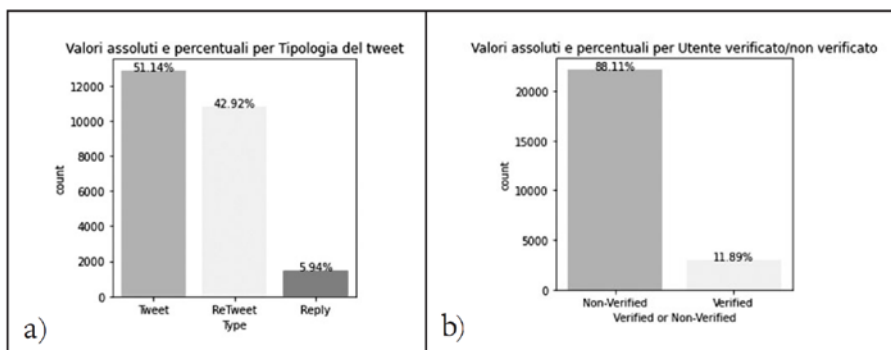


Fig. 1. Distribuzioni percentuali per ‘tipologia tweet’ e ‘verifica utente’.

L'opinione italiana sulla didattica a distanza

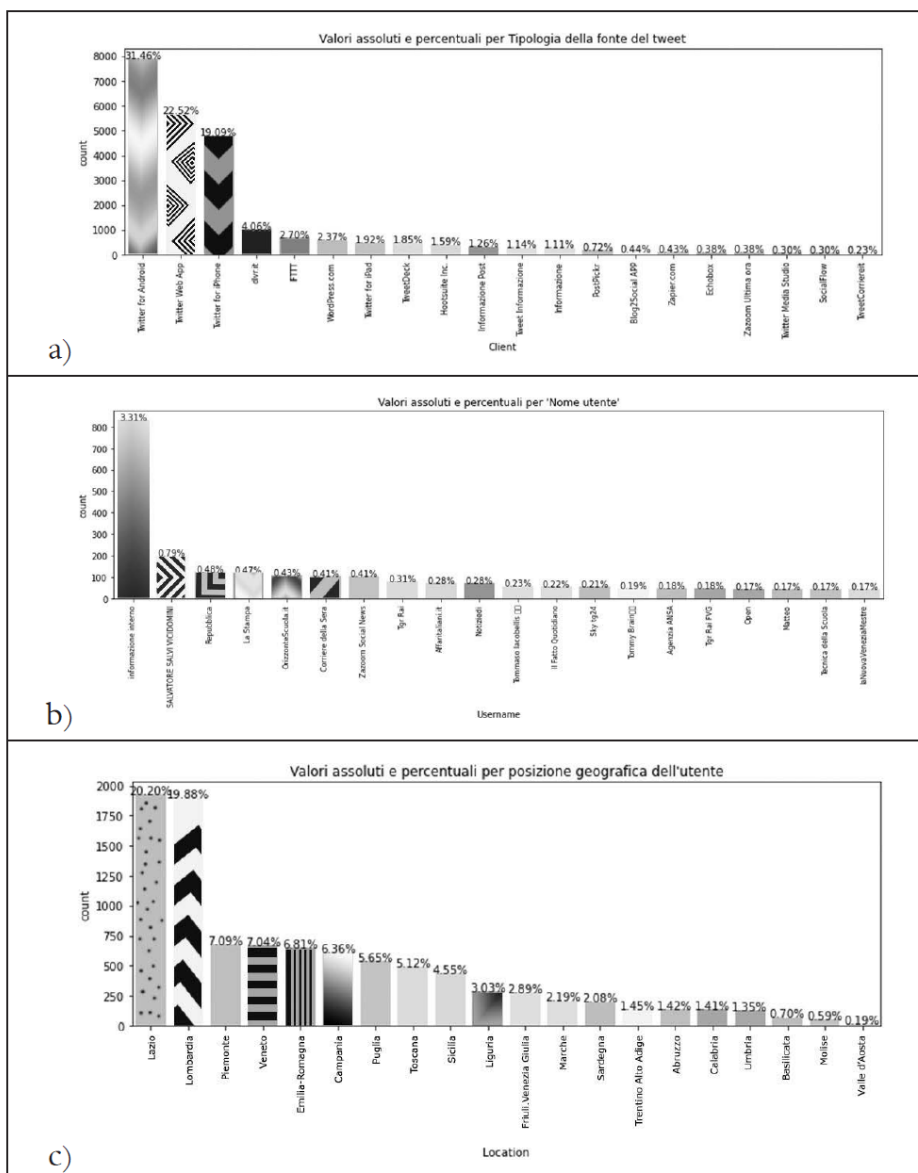


Fig. 2. Altre distribuzioni percentuali delle principali variabili categoriche.

2. 1. Analisi esplorativa dei dati

La Fig. 3 riporta il numero giornaliero di tweet nell'intero periodo di osservazione. Contrariamente a quanto ci si sarebbe aspettati, all'introduzione della prima DaD nel marzo 2020 gli italiani non hanno mostrato grande interesse nel twittare su questo tema. Il primo aumento significativo nel numero dei tweet giornalieri è avvenuto solo

con la seconda ondata Covid e il conseguente DPCM 3/11/2020. Tale decreto prevedeva la reintroduzione della DaD per le classi seconda e terza della scuola secondaria di i grado e per tutte le classi della secondaria di ii grado nelle zone rosse (quelle ad elevato rischio di contagio), mentre fissava la didattica digitale integrata per le zone arancioni e gialle (quelle con rischio moderato-basso). Tuttavia, solo con la terza ondata pandemica si ha il picco massimo di tweet, con oltre 500 tweet giornalieri, a seguito dell'approvazione del DPCM 2/3/2021 che ripristinava la misura restrittiva tanto temuta: la DaD totale con chiusura di tutte le scuole di qualsiasi ordine o grado nelle aree 'rosse'. Tale picco nel numero di tweet è stato seguito il mese successivo da un altro con più di 450 tweet in un giorno, in concomitanza con il successivo DL 22/4/2021, n. 52, noto anche come decreto delle 'riaperture'.

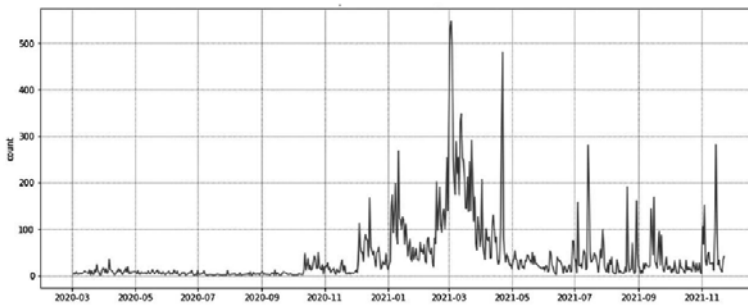


Fig. 3. Numero giornaliero di tweet.

Il grafico in Fig. 4b riporta il tasso di tweet (Ratio) ogni 1.000 abitanti per regione, ottenuto come:

$$Ratio_i = \frac{num. tweet_i}{Popol. residente_i} \cdot 1000 \quad \text{dove } i = 1, \dots, 20$$

Dal cartogramma si può notare che le regioni con un maggior numero di tweet sono state il Lazio e la Lombardia, ma non c'è da sorprendersi poiché sono anche le più popolate. Se invece si osserva la Fig. 4b, appare evidente che oltre ai laziali, hanno twittato di più i friulani rispetto ai lombardi e che anche i liguri sono stati dei buoni twittatori.

L'opinione italiana sulla didattica a distanza

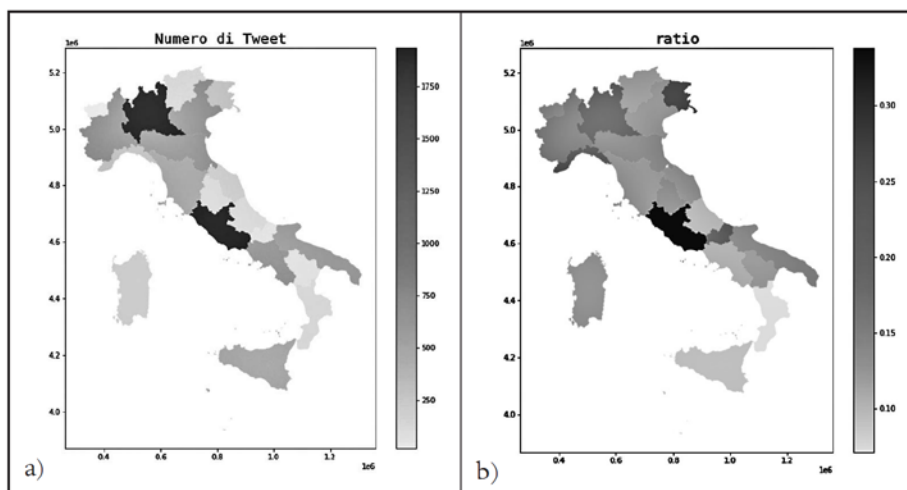


Fig. 4. Numero dei tweets per regione valori assoluti (a) e tasso di tweet (b).

L'analisi esplorativa dei dati testuali fa uso di una tecnica di visualizzazione detta 'nuvola di parole' o 'word cloud'. Per preparare i dati per la rappresentazione sono state rimosse le 'stop words', ossia quelle parole usate spesso all'interno di frasi che non contengono reali informazioni. La lista di stop words utilizzata è stata quella della libreria NLTK¹ di Python.

Dopo aver inizializzato l'oggetto WordCloud della libreria 'Python', è stata creata la nuvola con il metodo 'generate_from_text' per alimentarla direttamente con il testo dei vari tweet. La forma scelta per il 'cloud' è stata quella di un 'monitor', uno degli strumenti essenziali della DaD. Il criterio di rappresentazione adottato per identificare i termini più frequenti è stato il 'bi-gramm' ossia coppie di parole ripetute anziché parole singole, per enfatizzare il significato che scaturisce appunto dall'associazione dei termini.

La Fig. 5a mostra la nuvola sui dati nazionali nell'intero periodo di osservazione e contiene le parole più frequenti nei testi dei tweet. Sono parole di indignazione e rabbia se osservate sull'intero periodo, ma diventano meno dure e più propositive se restringiamo il periodo di osservazione agli ultimi quattro mesi dell'anno 2021 (Fig. 5b) ossia all'inizio del nuovo anno scolastico 2021-22.

¹ NLTK è una piattaforma per la creazione di programmi Python per il Natural Language Processing. Fornisce librerie di elaborazione del testo per classificazione, tokenizzazione, fare stemming, tagging, analisi e ragionamento semantico. <https://www.nltk.org/>.



Fig. 5. Word cloud tweet sulla DaD.

3. La metodologia

3.1. Preelaborazione del testo

Per preparare i dati per la successiva Sentiment Analysis sono state applicate le seguenti tecniche di Natural Language Processing (NLP) per la pulizia e la rimozione del testo non utile:

- rimozione di URL e indirizzi email
- rimozione di caratteri come virgolette isolate e simboli di testo a capo
- sostituzioni caratteri HTML con l'equivalente Unicode al fine di uniformare il testo
- rimozione di tag HTML (come “<div>”, “<p>”, ecc.)
- rimozione suffisso iniziale RT nei re-tweet
- eliminazione della punteggiatura, eccetto il punto esclamativo ‘!’ che porterà ad amplificazioni del ‘sentiment’ nel modello impiegato nell’analisi.

3.2. La Sentiment Analysis e il modello VADER

La *Sentiment Analysis* o *opinion mining*, è un sottocampo dell’elaborazione dell’NLP che cerca di identificare ed estrarre opinioni, sentimenti ed emozioni contenute all’interno di un testo (Pang, 2008; Liu,2012).

L'opinione italiana sulla didattica a distanza

Il modello VADER (Valence Aware Dictionary and sEntiment Reasoner) è uno strumento open source per la Sentiment Analysis basato su di un 'lessico' e regole predeterminate ed è particolarmente adatto per analizzare sentimenti ed emozioni espressi nei social media (Hutto *et al.*, 2014).

Un numero considerevole di approcci all'analisi del Sentiment si basa su di un lessico di sentimenti (o opinioni), ossia in un elenco di caratteristiche lessicali (ad es. parole) che sono generalmente etichettate in base al loro orientamento semantico come positive o negative (Liu, 2010).

Tuttavia, è sempre di maggior interesse non solo determinare la polarità binaria (positiva contro negativa), ma anche la forza del sentimento espresso nel testo, per capire, ad esempio, quanto si sentono favorevoli o sfavorevoli le persone verso un determinato tema o argomento. Nasce così la necessità di riconoscere e studiare le cause dei cambiamenti nell'intensità del sentimento nel tempo (Wilson *et al.*, 2004).

Il modello VADER è stato ottimizzato proprio per questo. Alle parole presenti nel lessico sono associati dei pesi detti 'punteggi di valenza' a seconda dell'intensità del sentimento che esprimono. Più in particolare il punteggio di ciascuna frase viene calcolato sommando i "punteggi di valenza" che ha ogni singola parola nel lessico di VADER e poi normalizzando tale punteggio, affinché vari tra 0 ed $|1|$, ossia tra -1 (sentimento negativo più estremo) e +1 (sentimento positivo più estremo). Si ottiene così una misura unidimensionale del sentimento per una determinata frase, con la denominazione di 'punteggio composito normalizzato ponderato' (compound), noto in letteratura anche come "Sentiment score" (S_s).

VADER ha molti vantaggi rispetto ai tradizionali metodi di analisi del Sentiment, tra cui il fatto che non richiede dati di training per il modello ed è veloce a livello computazionale. Nella determinazione del punteggio relativo al sentimento (S_s) si considerano alcuni punti chiave che possono variarne l'intensità, quali ad esempio:

- l'uso del punto esclamativo o le lettere maiuscole, che aumentano l'intensità del sentimento espresso;
 - l'uso dei modificatori di grado, ossia particolari "parole" che aumentano o diminuiscono l'intensità del sentimento espresso.
- Ad esempio: «La didattica è molto buona» ha maggior intensità

della frase «la didattica è buona», mentre se al contrario avessi usato nella prima frase ‘poco buona’ questo ne avrebbe ridotto l’intensità;

- l’uso di congiunzioni come ‘ma’ segnala uno spostamento della polarità del sentimento, con il sentimento che però segue la congiunzione dominante. «La spiegazione del docente è ottima, ma la DaD è terribile» ha sentimenti contrastanti, con la seconda metà che detta la valutazione complessiva;
- l’uso di emoji, slang ed emoticon che il modello riesce a riconoscere e valutare aumentando o diminuendo l’intensità del relativo sentimento espresso.

4. La Sentiment Analysis dei tweet sulla DaD

Sul dataset dei tweet italiani aventi come tema la DaD è stato applicato il modello VADER implementato tramite la libreria NLTK di Python. Poiché attualmente VADER supporta esclusivamente la lingua inglese, i tweet preelaborati in lingua italiana sono stati tradotti in inglese tramite l’API ‘googletrans’ prima di essere elaborati per l’analisi. Ai fini della determinazione del Sentiment score (Ss) sono stati considerati tutti i record, inclusi i ‘re-tweet’ di uno stesso messaggio da parte di utenti differenti. Tale scelta è stata motivata dal fatto che anche la condivisione di uno stesso testo da parte di utenti differenti rappresenta comunque un’opinione condivisa e pertanto concorre alla determinazione del sentimento globale.

Allo stesso tempo, è stato mappato il punteggio emotivo in tre categorie: “positivo”, “negativo” e “neutro” attraverso le medesime impostazioni dei parametri suggerite dagli autori (Hutto *et al.*, 2014), come indicato in Tabella 3.

L’analisi condotta dapprima a livello globale è stata poi effettuata a livello temporale per evidenziare la variazione del sentimento nei diversi tempi di attuazione della DaD ed a livello spaziale utilizzando come unità elementari le regioni italiane.

L'opinione italiana sulla didattica a distanza

Tabella 3. Sentiment Polarity (SP) e Sentiment score (S_s).

CRITERIO PER LA CLASSIFICAZIONE DEL SENTIMENT	
Condizione	Sentiment Polarity (SP)
$S_s \leq -0.05$	Negative
$S_s \geq 0.05$	Positive
$-0.05 \leq S_s \leq +0.05$	Neutral

4. 1. I risultati

L'analisi, condotta sull'intero periodo di osservazione, ha mostrato 8.876 tweets negativi, contro 8.184 positivi (Fig. 6). Poiché i tweet neutri riguardano prevalentemente notizie e informazioni twittate da testate giornalistiche o media, non sono stati inclusi nelle analisi successive.

La Fig. 7 mostra la distribuzione del 'Sentiment score' (S_s) per intervalli di ampiezza differenti 0.25 e 0.05. L'intervallo modale per il S_s risulta essere compreso tra 0 e + 0.05 e rientra nella polarità neutra di sentimento (vedi Tabella 3).

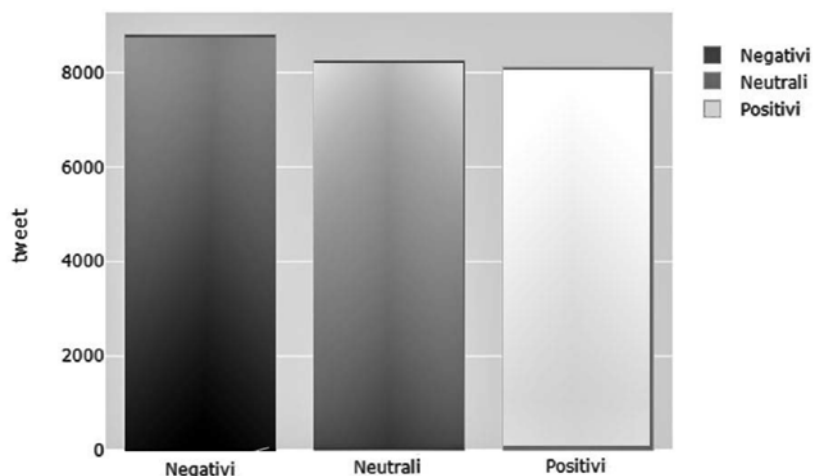


Fig. 6. *Sentiment Analysis*– intero periodo di osservazione.

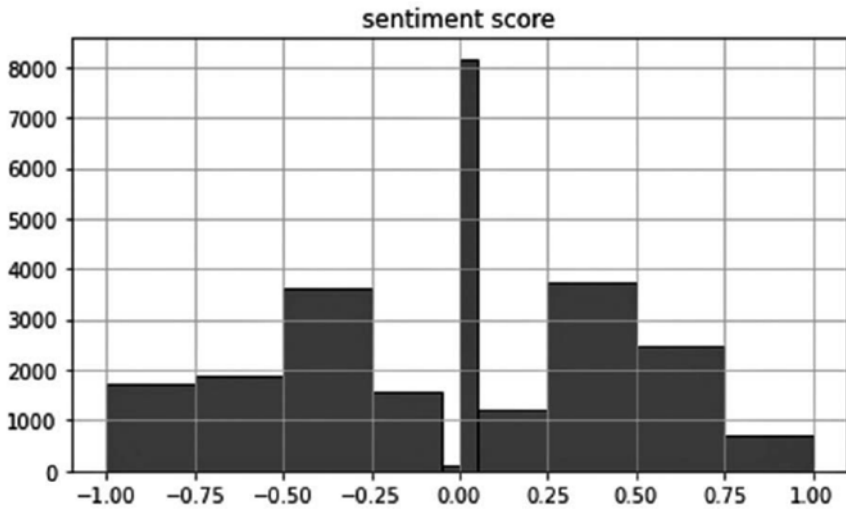


Fig. 7. Istogrammi *Sentiment score*.

L'esplorazione del 'word cloud' è stata condotta per l'intero periodo di osservazione nei due sottogruppi di tweet con polarità positiva e negativa. Tra le coppie di parole positive si leggono: "niente stressante", "molto facile", "chiara oggi", mentre tra quelle con sentimento negativo permangono termini di indignazione, con evidenziate le parole 'tasse' e 'sconti' che sottolineano il malcontento relativo alle tasse pagate e alla mancanza di un eventuale rimborso per il servizio scolastico ritenuto non ottimale durante la DaD di emergenza.



Fig. 8. Nuvole di parole per il 'sentiment' positivo e negativo.

L'opinione italiana sulla didattica a distanza

A livello temporale l'analisi del Sentiment ha evidenziato cambiamenti nell'atteggiamento degli italiani in concomitanza con i diversi decreti. In particolare, in Fig. 9 si nota un picco di sentimento negativo il 22 aprile 2021 contestualmente all'emanazione del DL 22/4/2021, n. 52, detto delle 'riaperture', che contrariamente alle attese degli italiani sanciva le riaperture solo in termini gradualisti ed a seconda del grado di rischio epidemico delle diverse aree.

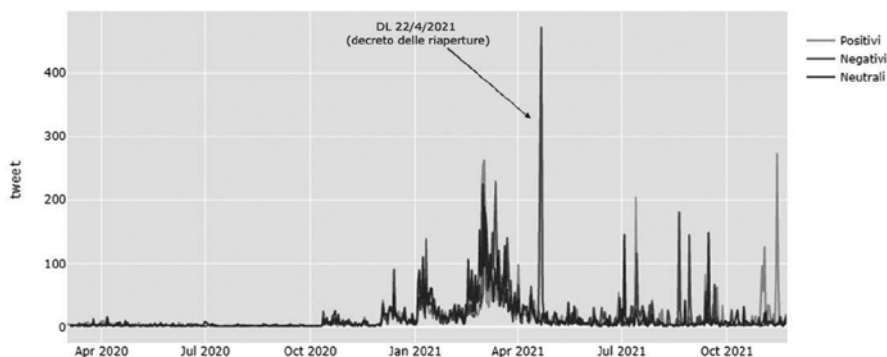


Fig. 9. Serie storica del Sentiment nei tweet italiani (periodo marzo 2020-novembre 2021).

Sempre nella serie storica del 'Sentiment' si può notare che i primi picchi di tweet con un sentimento positivo (linea grigio chiaro) sono iniziati a partire dall'anno scolastico 2021-2022. Più in particolare, il giorno 15 novembre ha registrato il picco positivo più alto della serie storica, scatenato dalla bozza del decreto fisco-lavoro apparsa sui media che prevedeva il rinnovo dei congedi straordinari per i genitori con i figli in DaD.¹

A livello regionale l'analisi del Sentiment score medio ha mostrato però forti differenze. La Fig. 10 riporta il S_s medio per regione, e mostra regioni con netta prevalenza di uno score medio negativo (in grigio scuro) come il Veneto e il Friuli, e regioni con score medio positivo (grigio con righe) come Umbria, Valle d'Aosta e Sardegna.

¹ Nella bozza il congedo è al 50% della retribuzione e vale per lavoratori con figli sotto i 14 anni, mentre la misura non è retribuita per i figli tra i 14 e i 16 anni.

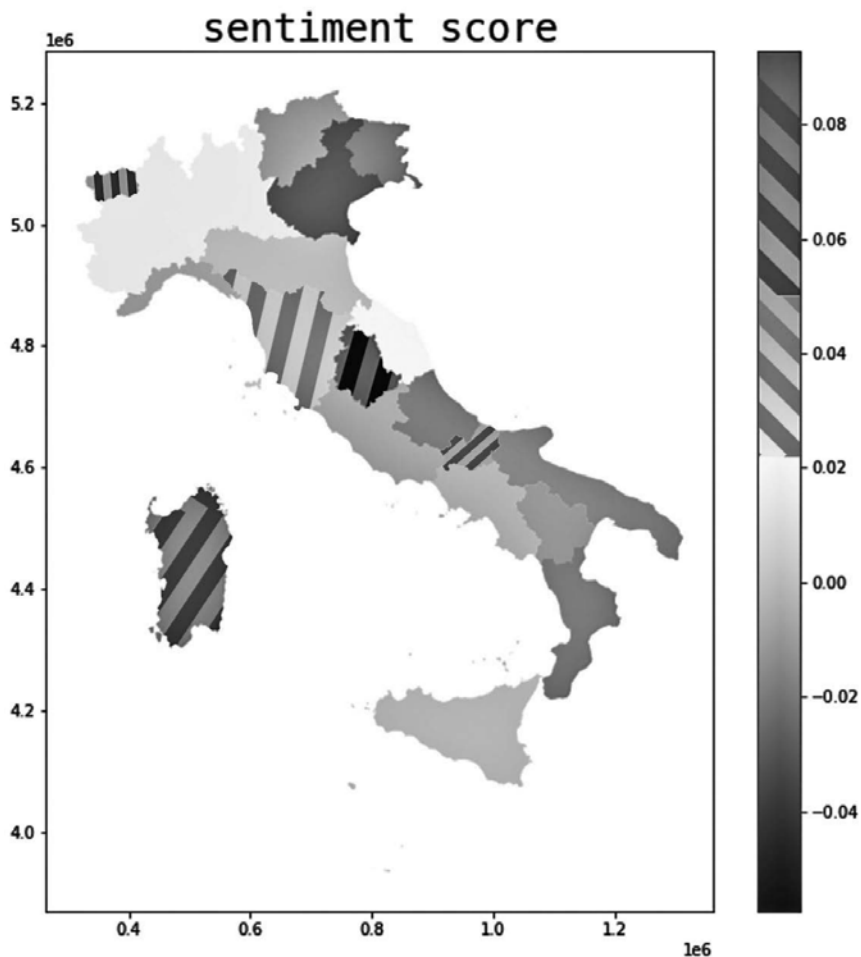


Fig. 10. *Sentiment score medio* per le diverse regioni italiane.

Conclusioni

Al fine di analizzare l'atteggiamento degli italiani nei confronti della Didattica a Distanza, sono state progettate ed eseguite una serie di analisi sul set di dati raccolti da Twitter che trattavano tale tema. La Sentiment Analysis applicata ai dati ha mostrato un'opinione prevalentemente negativa nei confronti della DaD fino ad agosto 2021. Il principale argomento di malcontento ha riguardato le tasse scolastiche pagate e la mancanza di un eventuale rimborso per il servizio scolastico ritenuto non ottimale durante la DaD di emergenza.

L'opinione italiana sulla didattica a distanza

L'analisi condotta nelle varie regioni ha poi mostrato il fluttuare del sentimento medio, con aree di forte malcontento in Veneto e in Friuli, frutto non tanto di situazioni scolastiche differenti quanto di notizie o provvedimenti restrittivi regionali. Dall'analisi della serie storica dei 'Sentiment' giornalieri appare evidente che il fluttuare dell'opinione degli italiani è da attribuire prevalentemente a notizie mediatiche e provvedimenti legislativi, mentre il trend nel lungo periodo ha mostrato il cambiamento nell'opinione degli italiani. L'inversione è avvenuta solo con l'avvento del nuovo anno scolastico (2021-2022) grazie ai progressi compiuti dal sistema scolastico italiano e all'applicazione della Didattica Digitale Integrata.

Bibliografia

- G. Barbier, H. Liu (2011), *Data Mining in Social Media*, in *Social Network Data Analytics*, New York, Springer, pp. 327-352.
- C. J. Hutto, E. E. Gilbert (2014), *VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text*, in *Eighth International Conference on Weblogs and Social Media (ICWSM-14)*, Ann Arbor, MI, pp. 216-225.
- S. Kemp (2020), *Digital 2020: Italy*, <https://datareportal.com/reports/digital-2020-italy?rq=italy>.
- B. Liu, (2010), *Sentiment Analysis and Subjectivity*, in *Handbook of Natural Language Processing*, 2nd Edition, New York, CRC Press.
- B. Liu (2012), *Sentiment Analysis and Opinion Mining*, San Rafael, CA, Morgan & Claypool.
- B. Pang, L. Lee (2008), *Opinion mining and sentiment analysis*, «Foundations & Trends in Information Retrieval», 2 (1), pp. 1-135.
- T. Wilson, J. Wiebe, R. Hwa (2004), *Just how mad are you? Finding strong and weak opinion clauses*, in *Proceedings of the National Conference on Artificial Intelligence*, NCAI-04s.