

L'intelligenza artificiale contro la violenza di genere

Claudia Villante, Gabriele Paone, Maria Giuseppina Muratore

Istat

Abstract. Obiettivo dell'attività di diffusione dell'Istituto Nazionale di Statistica è garantire l'accesso ai risultati delle rilevazioni statistiche utilizzando tecnologie aggiornate. Allo stato attuale infatti l'ISTAT ha dedicato un sito web al tema della violenza di genere (<https://www.istat.it/it/violenza-sulle-donne>), dove sono riportati dati e informazioni sul fenomeno che sono in continuo aggiornamento. Il sistema informativo "violenza sulle donne" è stato realizzato in collaborazione con il Dipartimento per le Pari Opportunità della Presidenza del Consiglio dei Ministri e risponde all'esigenza di monitorare il fenomeno nei suoi vari aspetti, in linea con le indicazioni della Convenzione di Istanbul. Il sistema produce inoltre dati e indicatori necessari a verificare il raggiungimento degli obiettivi previsti dall'Agenda 2030 – Goal 5 "Raggiungere l'uguaglianza di genere ed emancipare tutte le donne e le ragazze"

Keywords. Gender based Violence, gender stereotype, big data, sentiment analysis, emotion analysis

Introduzione

L'obiettivo del progetto è quello di arricchire il patrimonio informativo statistico oggi disponibile con indagini sperimentali come la sentiment e l'emotion analysis avvalendosi di tecnologie all'avanguardia come il machine learning.

Allo stesso tempo lo scopo è quello di rafforzare l'efficacia e la qualità dei dati prodotti, predisponendo analisi, raccolte e bonifica di dati agli organi di governo e a tutti i soggetti pubblici e privati coinvolti nelle azioni di contrasto alla violenza di genere.

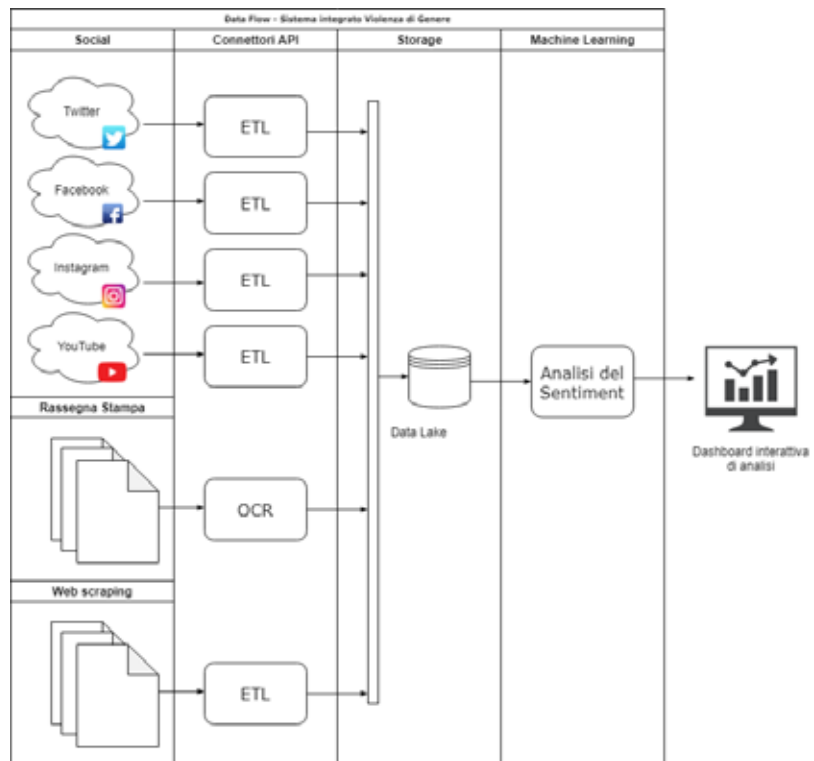
1. Approccio architetturale

Il primo passo è dunque quello di integrare le varie fonti. Il servizio si avvale di una soluzione che, a partire dai requisiti espressi dagli statistici dell'Istituto, abilita la raccolta delle informazioni sui diversi canali individuati e la definizione di viste-ambiti tematici (insights) di interesse, la selezione ed elaborazione di KPI specifici, l'addestramento degli algoritmi di calcolo e la produzione della reportistica per la fruizione dei risultati da parte degli utenti per una completa e certificata Web and Social Analysis.

È stata realizzata una dashboard con quattro quadri di analisi (insight) e una psicografica, basata su prodotti open source, per la visualizzazione e la navigazione delle informazioni. Il flusso completo alla base dell'erogazione del servizio viene rappresentato sinteticamente nella figura seguente.

In risposta alla sicurezza ed alla privacy dei dati (Regolamento UE – GDPR), i principali Social Network (Twitter, Facebook, Instagram, YouTube) stanno revisionando le proprie policy. La concreta compatibilità d'uso andrà quindi valutata in funzione delle specifiche

Fig. 1
Flusso dati



esigenze dell'Istituto ed in relazione alle policy una volta consolidate.

La configurazione dei canali considera un'attività di analisi preventiva finalizzata a:

- selezionare i canali più idonei (ovvero dove andare a raccogliere i dati)
- definire gli ambiti, temi, soggetti (ovvero cosa raccogliere)

L'accuratezza della configurazione dei canali, sotto il profilo qualitativo, risulta essenziale in quanto permette raccolte più mirate, evita acquisizioni di moli di dati scarsamente, o per nulla, significativi per le finalità dell'Istituto, e abilita l'ottenimento di risultati di maggior qualità nell'erogazione del servizio.

2. Gli Insights

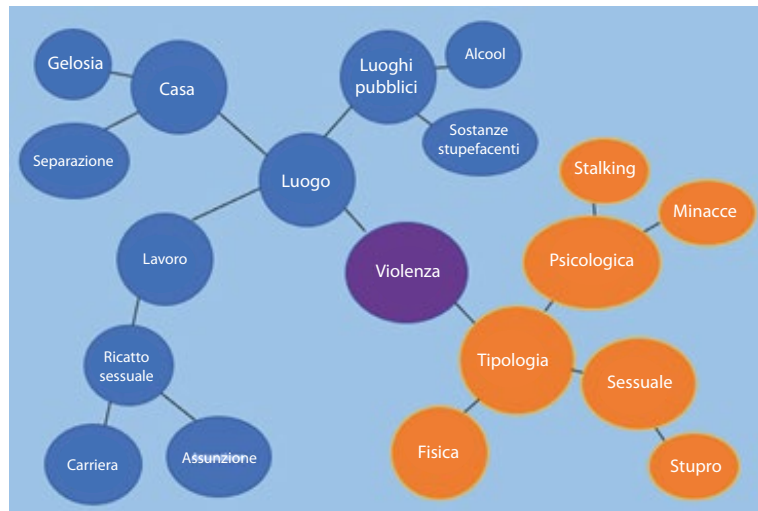
I contenuti acquisiti vengono analizzati, classificati ed organizzati in macro-ambiti. A puro titolo di esempio, la figura seguente riporta alcune rappresentazioni possibili.

3. La Dashboard interattiva

I principi architetturali precedentemente riassunti sono stati declinati nella realizzazione di una Dashboard interattiva che permette la configurazione e l'utilizzo di KPI standard e personalizzabili secondo i quali è possibile customizzare dashboard analitiche (es. analisi dei trend, andamento KPI di sintesi, analisi integrata dei contenuti).

A titolo di esempio viene di seguito fornito un elenco dei principali KPI implementabili per la fonte di tipo Social: - termini ricorrenti - hashtag frequenti - utenti più attivi - utenti più menzionati - distribuzione linguistica - distribuzione like (Instagram) - distribuzione

Fig. 2
Gli Insights



reply - distribuzione commenti – sentiment - trend temporali e picchi. Interessante il fatto che sia in fase di raccolta sia in fase di visualizzazione, sarà possibile analizzare i temi utilizzando dizionari personalizzati dall'utente.

4. La metodologia per la misurazione del Sentiment e dell'Emotion

Il servizio di sentiment analysis utilizza metodologie che si fondano su architetture di Deep Learning (Artificial Intelligence). La precisione dei risultati in termini di sentiment e di emotion è assicurata, oltre che dalla qualità intrinseca della metodologia, proprio dall'attività condotta in fase di addestramento, utilizzando tecniche di Active learning.

A partire da un ambito-tema o dal soggetto di interesse per l'Istituto, l'addestramento del motore (Machine Learning) si basa sulla disponibilità di un campione di testi certificati in termini di risposta sentiment (valutazioni positive, negative, neutre) ed emotion (gioia, sorpresa, amore, tristezza, rabbia, neutro); l'accuratezza delle classificazioni viene arricchita e perfezionata tramite l'impiego di opportuni cataloghi creati ad-hoc, secondo le specifiche esigenze del contesto di interesse, in questo caso la violenza di genere e gli stereotipi di genere.

L'addestramento è propedeutico per la successiva definizione del calcolo del sentiment e dell'emotion e, dunque, dell'erogazione del servizio. L'analisi del sentiment e dell'emotion risulta tanto più accurata quanto più ampio è il campione utilizzato in fase di addestramento (casistiche di classificazione). Al fine di perfezionare e rendere più affidabile i risultati delle analisi si prevedono fasi successive di tuning.

Conclusioni

L'analisi è iniziata a marzo 2020 e da questa data la piattaforma è costantemente aggiornata in tempo reale. Sulla base della proposta metodologica si prevede la produzione di una statistica sperimentale con cadenza trimestrale volta a restituire informazioni e colmare un gap informativo importante, offrendo informazioni inedite sul clima/mood che si crea intorno ai principali eventi riconducibili al tema della violenza contro le donne e de-

gli stereotipi di genere. L'obiettivo quindi è quello di leggere attraverso indici di positività o negatività come vengono veicolati i messaggi social sul fenomeno oggetto di analisi, ma anche di osservare come lo stesso fenomeno della violenza di genere, al pari di molti altri fenomeni sociali (come ad esempio il cyberbullismo) si espande nelle sue diverse forme digitali. Nuove estrinsecazioni del fenomeno richiedono, infatti, l'elaborazione di nuovi strumenti di analisi che siano in grado di coglierlo. Con una uscita periodica di reports inoltre si intende perseguire un ulteriore obiettivo di tipo culturale che attiene alla volontà di tenere alta l'attenzione dell'opinione pubblica sugli stereotipi di genere e l'immagine sociale della violenza, in linea con le raccomandazioni internazionali e con il raggiungimento degli obiettivi dell'Agenda 2030.

Riferimenti bibliografici

Bergvall, V. (1999). Toward a comprehensive theory of language and gender. *Language in Society*, 28(2), 273– 293. Retrieved from www.jstor.org/stable/4168929.

Berzal, Fernando & Matín, Nicolfás. (2002). Data mining: concepts and techniques by Jiawei Han and Micheline Kamber. *ACM SIGMOD Record*. 31. 66-68. 10.1145/565117.565130.

Bing Liu and Ian Lane. 2016. Attention-based recurrent neural network models for joint intent detection and slot filling. In *Interspeech 2016*, pages 685–689.

Bolasco S., 1999, "Analisi multidimensionale dei dati", Carocci Editore.

Bonadiman D., Castellucci G., Favalli A., Romagnoli R. and Moschitti A., (2017) *Neural Sentiment Analysis for a Real-World Application*, DOI: 10.4000/books.aaccademia.2357, Accademia University Press, Torino

Chakraborty, Goutam. (2014). *Analysis of Unstructured Data: Applications of Text Analytics and Sentiment Mining*.

Charles J. Fillmore. 1985. Frames and the semantics of understanding. *Quaderni di Semantica*, 6(2):222– 254.

Autori

Claudia Villante claudia.villante@istat.it

Dottore di ricerca in Sistemi sociali e analisi delle politiche pubbliche, dal 2018 e' primo ricercatore ISTAT, presso la Direzione Centrale Statistiche Sociali e Censimento della popolazione. E' stata policy officer presso ANPAL e ricercatore presso INAPP (ex-Isfol) occupandosi del monitoraggio e valutazione delle politiche pubbliche, in chiave di genere. Attualmente e' esperta nazionale distaccata presso EIGE (European Institute for Gender Equality).

Gabriele Paone gpaone@istat.it

Nasce nel 1975, è un ingegnere informatico e attualmente lavora come IT Senior Architect presso la Direzione informatica dell'ISTAT, inoltre è responsabile tecnico degli applicativi d'area "Salute, assistenza, previdenza, giustizia, lavoro, istruzione e formazione" all'interno dell'iniziativa "Sistemi Informativi per le statistiche sociali e il censimento della popolazione".

Fa parte della Commissione informatica dell'Ordine degli Ingegneri di Roma.

Maria Giuseppina Muratore muratore@istat.it

dottore di ricerca in Metodologia della Ricerca Sociale, primo ricercatore Istat dal 2004, è responsabile del Sistema Informativo sulla violenza di genere, nato dall'Accordo tra l'Istat e il Dipartimento per le Pari Opportunità presso la Presidenza del Consiglio nel 2017, delle rilevazioni sui Centri antiviolenza e le Case rifugio, dell'indagine sulla sicurezza delle donne, dell'indagine sulla sicurezza dei cittadini. Esperta su crime, crime classification e gender based violence è stata parte di molte Task Force e gruppi di lavoro presso le Nazioni Unite (UNECE, UNODC, UNWOMEN, UNSD), il Centre of Excellence on Crime, la Commissione Europea, EUROSTAT, il FRA (European Fundamental Rights Agency) e EIGE.