



UNIVERSITÀ DI PISA

Corso di Laurea in Informatica Umanistica

LAUREA TRIENNALE

Twitter come rete sismica

Candidato: *Serena Tardelli*

Relatore: *Prof. Maurizio Tesconi*

Correlatore: *Prof. Stefano Cresci*

Prof. Mariantonietta Noemi La Polla

Anno Accademico 2013-2014

Indice

1. Introduzione	1
1.1. Web 2.0 e social media	1
1.2. Obiettivi.....	2
2. Social media ed emergenze	4
2.1. Social Media analysis nelle emergenze	7
3. Progetto SoS - Social Sensing	11
3.1. Estensione del progetto	12
4. Strumento di annotazione manuale: #tweeTag	13
4.1. Descrizione	15
4.2. Annotazione globale.....	15
4.3. Annotazione testuale	18
4.4. Annotazione timeline.....	19
4.5. Inter Annotator Agreement.....	20
4.6. Implementazione.....	23
4.7. Esempi di utilizzo del tool	25
4.7.1. Considerazioni.....	29
5. Sistema per la rilevazione dei terremoti tramite tweets	30
5.1. Acquisizione dei dati	30
5.2. Campagna di annotazione dei dati acquisiti mediante tool #tweeTag	31
5.3. Addestramento del Classificatore	33
5.4. Simulazione delle attività degli utenti Twitter	35
6. Rappresentazione grafica USA dei risultati incrociando diversi dataset	37
6.1. Elenco dataset.....	37
6.1.1. Dati USGS - posizione, profondità, magnitudo terremoti	37
6.1.2. Terremoti rilevati dal sistema	39
6.1.3. Situazione demografica degli Stati Uniti	41
6.2. Creazione mappa	42
6.3. Considerazioni.....	44
7. Conclusioni e progetti futuri	47

8. Bibliografia 49

1. Introduzione

1.1. Web 2.0 e social media

Al giorno d'oggi le persone utilizzano sempre di più il web per condividere informazioni, aggiornarsi sulle news e scambiarsi opinioni su argomenti di qualsiasi natura. Il consistente ricorso a questa nuova tecnologia è stato favorito dall'evoluzione del web nel Web 2.0, un nuovo paradigma di comunicazione caratterizzato dalla dinamicità dei contenuti e da un alto livello di interazione. Ciò è stato reso possibile grazie all'affermarsi di servizi web che consentono lo scambio in tempo reale di informazioni e contenuti generati dagli utenti stessi: i social media. In questi luoghi virtuali, a differenza di ciò che avviene nei media tradizionali come televisione e giornali, l'informazione si sposta in molteplici direzioni, consentendo alle persone di interagire in maniera multilaterale.

Esistono diversi tipi di social media che si differenziano sulla base delle modalità e del grado di interazione che offrono ai loro utenti. Tra i più famosi troviamo forum (LiveJournal, Gutefrage), Wiki (Wikipedia, Wikitravel), Chat/texting (Skype, Whatsapp) e, in particolare, Social Network (Facebook, Twitter, Google+).

I Social Network sono piattaforme di condivisione che permettono alle persone di creare un profilo (semi) pubblico in un dominio, in modo da potersi connettere con altri utenti all'interno della stessa rete. L'accesso a tali strumenti sta diventando sempre più alla portata di tutti e, ad oggi, quelli più famosi, quali Facebook e Twitter, sono diventati veri e propri mezzi di comunicazione adottati giornalmente per interazioni online e condivisione di osservazioni, contenuti, sentimenti e opinioni.

Ogni giorno milioni di persone nel mondo postano messaggi su questi social per condividere pensieri, idee, strategie, risorse. Ciò comporta la generazione di una massiccia e imponente quantità di dati detta Big Data che può essere accessibile dalla rete e, se estratta a dovere, permette di ottenere informazioni utili e interessanti per qualunque settore.

Le informazioni condivise però possono essere false o inaccurate: è dunque necessario ricorrere a tecniche che siano in grado di filtrare, categorizzare, classificare e analizzare tali contenuti. Questo è un obiettivo difficile da raggiungere, in quanto i dati presenti sul web sono caratterizzati da tre principali problemi di gestione:

1. *Dimensione*: i dati da analizzare hanno grandi volumi;
2. *Rumore*: è difficile individuare oggetti d'interesse tra i dati a causa della loro natura non strutturata e delle loro provenienza da fonti diverse;
3. *Dinamismo*: i dati vengono prodotti in tempo reale e cambiano spesso.

Il Data Mining è una disciplina che ha come obiettivo quello di estrarre informazioni da enormi dataset mediante una vasta gamma di tecniche, dalla creazione di modelli statistici all'apprendimento automatico, per rilevare conoscenze utili e recuperare informazioni.

Persone, celebrità, aziende, organizzazioni e governi comprendono l'importanza di monitorare le attività sui Social Network al fine di ottenere informazioni di varia natura, come sapere qual è l'opinione pubblica riguardo a certe elezioni o come la loro immagine venga percepita, individuare l'andamento del mercato, captare il trend del momento, ecc.

Un aumento della comunicazione, e di conseguenza dei dati, nei social media può essere innescato da cause che possono essere suddivise in endogene ed esogene [9]. Le cause endogene si riferiscono ai fenomeni in cui un'idea o "meme" guadagna popolarità mediante un processo di contagio virale e si diffonde rapidamente attraverso la rete. Le cause esogene si riferiscono a eventi su larga scala che accadono nel mondo fisico e che sono di grande interesse per gli utenti dei social media (disastri o situazioni di emergenza).

1.2. Obiettivi

L'obiettivo di questo lavoro è stato quello di sfruttare i dati provenienti dai social media, in particolare Twitter, per studiare il comportamento degli utenti durante situazioni di emergenza, come, ad esempio, i terremoti.

Twitter è un servizio di microblogging che permette ai suoi utenti di trasmettere brevi messaggi testuali, i tweet, a un audience di follower. La scelta di utilizzare Twitter in questo lavoro è data dalla sua popolarità (vengono postati circa 500 milioni di tweet al giorno¹) e dalla rapidità di comunicazione che permette di concentrare le informazioni rilevanti di un certo fenomeno in 140 caratteri. Infatti una caratteristica importante di Twitter è la sua natura real time: gli utenti postano frequentemente quello che stanno facendo o pensando. Il tempo di vita di un tweet è brevissimo, perciò vengono generati numerosi aggiornamenti di stato dai quali è possibile ricavare informazioni utili relative a eventi reali in tutto il mondo, come ad esempio elezioni, olimpiadi, attentati o disastri naturali.

¹ Twitter, Inc Common Stock, 9 June 2014

Lo scopo di questo lavoro è stato quello di stabilire se fosse possibile determinare in tempo quasi-reale l'avvenimento di un terremoto intercettando i tweet di utenti che ne stanno parlando in quel momento.

Il lavoro effettuato si può riassumere in 6 fasi:

1. Raccolta di un campione di tweets in lingua inglese potenzialmente inerenti ad un terremoto;
2. Creazione di uno strumento web di annotazione manuale apposito per i dati provenienti da Twitter;
3. Annotazione manuale, usando il nuovo strumento, del campione di tweet raccolti, al fine di determinare quali di questi tweet segnalano effettivamente la presenza di un terremoto in corso e quali sono rumore da scartare;
4. Utilizzo del dataset annotato per addestrare un modello di apprendimento supervisionato;
5. Simulazione di situazioni di allerta terremoto su Twitter per testare il modello e cercare di rilevare attraverso i tweet i terremoti realmente avvenuti;
6. Verifica della bontà dei dati ottenuti attraverso una rappresentazione grafica su di una mappa.

I risultati ottenuti sono promettenti, a sostegno dell'idea dell'importanza di sfruttare tali dati a vantaggio della comunità. A partire da questi risultati si potrebbe pensare, in futuro, di tracciare l'evoluzione di un terremoto, valutare i danni prodotti e inviare i soccorsi. Ancora, l'ideale sarebbe riuscire ad applicare lo stesso paradigma per individuare eventi di tipo diverso dai terremoti, ad esempio alluvioni, incendi, o addirittura prevedere considerevoli eventi di interesse nazionale, come disordini pubblici.

2. Social media ed emergenze

I social media hanno il vantaggio di facilitare e accrescere la consapevolezza delle persone circa situazioni di qualsiasi natura. La “consapevolezza della situazione” viene definita come "la percezione di elementi nell'ambiente entro un volume di tempo e di spazio, la comprensione del loro significato, e la proiezione del loro stato nel prossimo futuro." [12].

Sempre più spesso, durante eventi importanti o disastri, la popolazione anziché aspettare gli avvisi ufficiali utilizza i social media come fonte di informazione per:

1. Ricercare informazioni relative all'evento;
2. Esprimere opinioni in merito all'accaduto;
3. Favorire l'azione (partecipando, donando o coordinando i soccorsi);
4. Esprimere sostegno, solidarietà, ecc.

Organizzazioni e governi negli ultimi anni hanno scoperto l'importanza di monitorare questi canali per identificare problemi, capire l'ambiente circostante e rispondere ai problemi dei cittadini. Il recupero e la valorizzazione di tali informazioni può avere un ruolo cruciale nella scelta di decisioni tempestive al fine di minimizzare l'impatto dei disastri sulla popolazione e sulle infrastrutture, ma anche per un migliore consapevolezza della situazione o verificare le informazioni ottenute attraverso altri canali.

Già in passato i social media hanno avuto un ruolo fondamentale nelle comunità per avvertire di aree o situazioni non sicure, informare amici e famiglie che qualcuno era al sicuro e raccogliere fondi per risollevere i disastri.

Uno dei primi casi di utilizzo dei social media in caso di emergenza si è registrato il 16 aprile del 2007 [44], quando uno studente della Virginia Tech University sparò a decine di studenti e professori uccidendo 32 persone: dopo mezz'ora dalla tragedia gli studenti iniziarono a postare messaggi su Facebook chiedendo ai loro amici se stessero bene e dopo soli 90 minuti dall'accaduto fu pubblicata la prima pagina di Wikipedia sulla tragedia. Lo studio condotto in [29] e [46] ha scoperto che la comunità online fu in grado di compilare accuratamente i nomi delle vittime prima del comunicato ufficiale.

Un altro esempio di uso dei social media come mezzo di comunicazione in casi di emergenza è stato durante il terremoto del Sichuan in Cina nel 2008 [39]: furono i messaggi su Twitter a diffondere la notizia a livello mondiale, prima ancora dei media tradizionali o dei comunicati stampa ufficiali. Inoltre a causa di danni alle infrastrutture di telecomunicazione, i

servizi di rete fissa e telefonia cellulare si interruppero, per cui molti utilizzarono Internet per chiedere aiuti e informazioni. In particolare fu molto usato il servizio web Tianya Club², uno dei forum più popolari in Cina che vanta oltre 20 milioni di account registrati. Entro 1 minuto dal terremoto i messaggi relativi all'evento iniziarono a comparire sul forum [32].

Twitter è stato usato anche durante il terremoto di magnitudo 7.0 che colpì Haiti il 13 gennaio 2010 [45]. I cittadini colpiti usavano i loro telefoni cellulari per scattare foto della loro situazione e postarli via Twitter [30]. Per coloro che avevano perso i telefoni fissi Facebook diventò l'unico modo per comunicare con i loro cari e conoscere la sorte di altri [28]. Oltre 2,3 milioni di tweet furono inviati tra il 12 e il 14 gennaio [41] e oltre 1.500 messaggi di stato di Facebook per minuto contenevano la parola "Haiti" [17]. In risposta alla tragedia, Twitter aggiornò il suo blog ufficiale per annunciare la campagna di raccolta fondi della Croce Rossa: entro 48 ore oltre 3 milioni di dollari erano stati raccolti, grazie in gran parte alla sua diffusione virale tramite Twitter. [40]. La Croce Rossa chiese anche l'aiuto di personaggi popolari su Twitter, tra cui la first lady Michelle Obama, per retweetare messaggi importanti per conto dell'organizzazione [27]. Inoltre i social in quel periodo furono usati per cercare le persone disperse postando su internet le foto e per dare aiuti umanitari. Anche Google in quel periodo mise su il servizio web *Google Person Finder*³ per trovare le persone scomparse [47].

L'11 marzo 2011 un terremoto di magnitudo 9 colpì il Giappone. Il terremoto provocò poi uno tsunami. A seguito di questo disastro le linee telefoniche si intasarono e le persone per comunicare scelsero Twitter, che registrò picchi di 1.200 tweet al minuto solo su Tokio [43]. Addirittura la compagnia elettrica di Tokio fece un account ufficiale sul social network per comunicare ai cittadini tutti i disservizi. Twitter in quel periodo divenne un mezzo di comunicazione quasi unico e ufficiale e su Facebook ci furono tantissimi status update che contenevano la parola "earthquake" o "tsunami" nelle 24 ore successive all'evento [10].

Questi esempi servono a far capire la mobilitazione che si manifesta nel web e la quantità di dati che vengono generati a seguito di un evento. La partecipazione della popolazione nelle situazioni di emergenza non è di panico e caos [26]. Le vittime del disastro non perdono il controllo, ma prendono decisioni rapide sulla base delle informazioni a loro disposizione al momento, che spesso permettono loro di salvare la propria vita e aiutare coloro che li circondano prestando i primi soccorsi [36].

La possibilità di inviare messaggi a un vasto pubblico da ovunque usando dispositivi mobili, o rete wireless se disponibile, sta rendendo piattaforme come Twitter sempre più usate

² www.tianya.cn

³ www.google.org/personfinder

per questi scopi e spesso finiscono con l'essere legittimate dagli utenti come fonte autorevole del ciclo di diffusione delle informazioni. I vantaggi di tali canali si possono identificare in tre principali opportunità:

1. In primo luogo, i messaggi possono raggiungere tempestivamente e facilmente un vasto pubblico.
2. In secondo luogo, tali strumenti possono essere utilizzati per fornire le direttive necessarie a gestire una situazione di emergenza.
3. Infine, i social media permettono un flusso di informazioni rapido e aggiornato, che è un aspetto essenziale nelle segnalazioni di emergenza.

Tuttavia è necessario prendere in considerazione anche i fattori di rischio legati all'utilizzo dei social media come principale fonte di informazione. Tra questi rientrano:

1. *Accuratezza delle informazioni*: le informazioni raccolte sui social media sono generalmente veritiere, ma esiste la possibilità di incappare in informazioni inesatte o imprecise che possono essere facilmente duplicate e inoltrate con il rischio di diffondere disinformazione. Le informazioni imprecise possono diventare un problema rilevante, soprattutto durante le catastrofi: alcuni studi [10] hanno rivelato che le informazioni obsolete e inesatte (rapporti falsi, una erronea posizione del pericolo) sono state riportate durante il terremoto del 2011 in Giappone quando alcuni tweet di richiesta di assistenza venivano "retweettati" anche dopo che le vittime erano state salvate, ostacolando e confondendo i soccorsi [43]. Un'altra preoccupazione è che alcuni individui o organizzazioni potrebbero intenzionalmente fornire informazioni inesatte per confondere gli interventi di soccorso. Per mitigare tali rischi, c'è bisogno di considerare in che modo trattare le richieste di assistenza fatte attraverso piattaforme di social media e sviluppare un sistema per verificare l'autenticità e rispondere in modo tempestivo e coordinato;
2. *Limitazioni tecnologiche*: anche se i social media possano migliorare alcuni aspetti della risposta alle emergenze e ai disastri, l'eccessivo affidamento sulla tecnologia potrebbe essere problematico in caso di interruzioni di corrente prolungate. Ad esempio, dopo l'uragano Irene nel 2011 molti residenti subirono interruzioni di corrente della durata di oltre 48 ore che, data la durata media delle batterie di smartphone e tablet, ha impedito di fatto il ricorso ai social media come unica fonte di informazione [7];

3. *Costi*: il pubblico non ha molti costi nell'utilizzo le nuove tecnologie per avere più informazioni, ma le organizzazioni che vogliono sfruttare tali dati per essere più vicini ai cittadini devono considerare il numero del personale necessario per monitorare l'enorme volume di messaggi in arrivo da parte del pubblico durante un disastro [22] o il costo che comporta la creazione di una piattaforma automatizzata;
4. *Privacy*: Infine, l'uso delle piattaforme sociali come mezzo di informazione da parte delle organizzazioni potrebbe sollevare delle questioni legate alla privacy, in quanto anche se i dati sono di pubblico dominio, i cittadini potrebbero sentirsi osservati dalle autorità [6].

2.1. Social Media analysis nelle emergenze

L'analisi dei dati sul web si distingue in statica e dinamica:

- L'analisi *statica* consiste nell'analizzare dati generati da tempo e conservati da qualche parte, ed è più facile da portare avanti rispetto ai dati delle reti generati in tempo reale;
- L'analisi *dinamica*, al contrario, analizza i flussi che vengono prodotti in streaming dai social network e sono più difficili da gestire. I dati in questo scenario sono generati rapidamente e in grande quantità.

Analizzare statisticamente i social media e usarli per disseminare informazioni e ricevere riscontri dagli utenti tramite messaggi, post e sondaggi è una procedura consolidata. Un trend più recente è quello di sfruttare l'intelligenza collettiva in maniera dinamica per identificare, processare e comprendere elementi critici di un incidente o di una situazione, e quindi fare scelte migliori nell'invio di soccorsi, aiuti e operazioni di salvataggio.

Alcune organizzazioni hanno già incluso il monitoraggio e la comunicazione dei social media nelle emergenze di massa, ad esempio la croce rossa americana⁴ ha aperto nel 2012 il Social Media Digital Operation Center for Humanitarian Relief, una base operativa che monitora i social media per ricevere e fornire "ulteriori informazioni agli utenti circa le zone colpite

⁴ www.redcross.org

durante le emergenze, anticipare le esigenze del pubblico e collegare le persone con le risorse di cui hanno bisogno, come il cibo, l'acqua, riparo e sostegno anche emotivo.”⁵

Monitorare le attività degli utenti in tempo reale durante le emergenze è però un task difficile a causa degli ostacoli descritti nel paragrafo 2.1.3, soprattutto per quanto riguarda i costi amministrativi, in quanto il numero del personale necessario per monitorare manualmente l'enorme flusso di informazioni che viene prodotto da più fonti è incerto [8]. Quando la computazione umana è limitata sono necessari metodi automatici di elaborazione dei messaggi.

Negli ultimi anni gli studi si sono spinti nel cercare di creare strumenti in grado di integrare questo tipo di analisi dei dati dei social media all'interno di sistemi automatizzati. Esistono diversi approcci computazionali per l'elaborazione automatica dei contenuti dei social media in scenari di emergenza.

La maggior parte dei social media fornisce l'accesso ai suoi contenuti attraverso le Application Programming Interface (API) che possono essere di due tipi: quelle che consentono di interrogare un archivio di messaggi passati (search API) e quelle che permettono di accedere al feed dei dati in tempo reale (streaming API) per un certo periodo di tempo, per coordinate GPS, o per parole chiave.

La maggior parte degli studi e dei sistemi realizzati utilizza i dati provenienti da Twitter, oltre che per la sua popolarità, perché le streaming API⁶ fornite dalla piattaforma sono più “aperte” rispetto a quelle degli altri social network. Ad esempio Facebook consente l'accesso ai post più recenti solo a partner selezionati⁷, mentre Google plus⁸ e Tumblr⁹ non espongono streaming API ma forniscono solo i post più recenti di un utente specifico (Google plus)¹⁰ o con un particolare tag (Tumblr)¹¹. Il seguente stato dell'arte è focalizzato, per queste ragioni, prevalentemente sull'utilizzo dei dati provenienti da Twitter.

Prima di eseguire l'analisi vera e propria del fenomeno in corso sui social è necessaria una pre-elaborazione che filtri i dati. La scelta del metodo di pre-elaborazione dei dati dipende dal tipo di dati e dagli obiettivi dell'analisi. Si può scegliere di ridurre la quantità dei dati da elaborare rimuovendo i messaggi quasi duplicati (De-duplicazione), oppure filtrando il rumore dei messaggi raccolti attraverso un'etichettatura a mano, mediante parole-chiave, o in base all'area geografica, sia attraverso coordinate GPS che mediante Geo-tagging, cioè trovando

⁵ www.redcross.org/news/press-release/The-American-Red-Cross-and-Dell-Launch-First-Of-Its-Kind-Social-Media-Digital-Operations-Center-for-Humanitarian-Relief

⁶ April 2014, www.dev.twitter.com/docs/api/streaming

⁷ April 2014, [www.developers.facebook.com/docs/public feed](http://www.developers.facebook.com/docs/public-feed)

⁸ www.plus.google.com

⁹ www.tumblr.com

¹⁰ April 2014, www.developers.google.com/+api/latest/

¹¹ April 2014, www.tumblr.com/docs/en/api/v

riferimenti geografici nel testo e collegandoli ai luoghi corrispondenti. Altri metodi di pre-elaborazione includono l'applicazione di tecniche di NLP che lavorano al livello sintattico e hanno il compito di individuare le unità sintattiche del testo (tokenizzazione, la part-of-speech tagging) e l'estrazione di features, ovvero delle caratteristiche di ogni dato sotto forma di vettore numerico.

Per la rilevazione di un evento esistono diversi approcci computazionali, tra cui il metodo della burst detection [11], che analizza la frequenza di una parola chiave legata a un particolare evento in una finestra temporale fissa. Se le frequenze osservate sono molto superiori rispetto a quelle registrate in passato allora viene rilevato un evento. Nell'applicazione di questo metodo tuttavia, va tenuto conto dei problemi legati a parole particolari (o hashtag nel caso di Twitter) che ricorrono maggiormente con cadenza variabile o fissa: ad esempio l'hashtag “#tbt” (“Throwback Thursday”) viene usato nei social il giovedì per condividere vecchie foto o ricordi [5].

Un altro metodo per rilevare eventi sui social media è quello di integrare i dati provenienti da fonti media tradizionali: un esempio è lo studio presentato in [42] che si concentra nel trovare articoli che descrivono eventi e utilizzare le parole chiave nel titolo e nel primo paragrafo per ottenere i tweet legati all'evento.

Una volta individuato l'evento, i dati ad esso collegati possono essere analizzati mediante diversi metodi computazionali. Esistono approcci che classificano gli elementi nel loro insieme, come la classificazione supervisionata e non supervisionata, e tecniche che estraggono informazioni utili direttamente dal contenuto di uno o più dati (Information Extraction).

La classificazione supervisionata consiste nella realizzazione sulla base di dati etichettati a mano, di un algoritmo a scopo predittivo capace di etichettare nuovi elementi in maniera automatica; metodi di apprendimento automatico non supervisionato includono il clustering, che cerca di individuare e spiegare dei modelli nascosti nei dati non etichettati.

Esempi di progetti e studi che mirano a migliorare la conoscenza delle situazioni di emergenza sono *Crisis Mapping* [25] che impiega volontari per raccogliere, classificare e localizzare i messaggi e usa i dati elaborati per addestrare metodi che eseguano tali operazioni automaticamente, e *SensePlace2* [23], un sistema che filtra ed estrae le informazioni geografiche, temporali e tematiche di tweets al fine di presentarli in una mappa. *TwitterMonitor* [24], raccoglie i tweet dal flusso di Twitter e rileva gli argomenti di tendenza (notizie, eventi) in tempo reale ed estrae le informazioni dei tweet per analizzare e descrivere il trend.

Altri studi [1] analizzano l'uso di Twitter durante i terremoti in Giappone e osservano che i tweets provenienti dalle aree colpite indicano richieste di aiuto e avvisi, mentre i tweets

provenienti dalle aree più lontane dall'epicentro includono altri tipi di informazioni, come preoccupazione o condoglianze.

TEDAS [35] rileva, analizza e identifica eventi collegati a disastri e alla criminalità raccogliendo i tweets sulla base di alcune parole chiave e hashtag, li classifica utilizzando metodi di apprendimento supervisionato in base al contenuto e alle loro caratteristiche (cioè, URL, hashtags, menzioni), e infine estrae informazioni sulla posizione tramite GPS o posizione presenti nel testo del tweet.

Un approccio simile è stato utilizzato per rilevare i terremoti su Twitter in [33]. Questo sistema controlla in tempo reale il flusso delle parole chiave "terremoto" e "#eqnz" ("earthquake Nuova Zelanda") e applica un metodo di burst detection per analizzare la frequenza di tali parole in una finestra temporale fissa. Se le frequenze osservate sono molto superiori rispetto a quelle registrate in passato allora viene rilevato l'evento.

Rilevare un evento è utile in fase di *Early Warning*, ovvero per allertare tempestivamente un determinato sito di interesse prima che l'evento potenzialmente pericoloso lo raggiunga. Lo studio condotto in [38] ha sviluppato un sistema di rilevamento terremoti in Giappone che fa uso di un classificatore per filtrare i tweet ed è stato in grado di individuare l'89.7% (70 su 78) dei terremoti avvenuti nell'arco di due mesi. Il sistema spedisce inoltre un mail di avvertimento ai luoghi più distanti dall'epicentro prima che vengano colpiti dall'onda del terremoto.

Un'architettura per applicazioni di Early Warning in un contesto di Social Sensing è stata sviluppata anche al CNR di Pisa [4]. Questo sistema di rilevamento terremoti monitora il flusso dei tweet in lingua italiana con delle keyword specifiche ("terremoto" e "scossa") e filtra il rumore mediante un classificatore. Dei dati raccolti solo il 12% supera la fase di filtraggi; a questi viene poi applicato il metodo della burst detection. Il sistema mostra ottime performance nel rilevare i terremoti con magnitudo superiore a 3.5 ed è in fase di ampliamento per aggiungere delle componenti che riescano a fare una stima dei danni a seguito dell'evento (Damage Assessment) [3].

3. Progetto SoS - Social Sensing

Il lavoro sviluppato per questa tesi si inquadra nell'ambito delle attività del progetto SoS - Social Sensing¹², sviluppato presso l'area della ricerca di Pisa, dall'Istituto di Informatica e Telematica del CNR.

Il progetto SoS – Social Sensing è portato avanti dal gruppo di ricerca Web Applications for the Future Internet (WAFI) che si occupa di social media analysis e ha come scopo quello di monitorare ciò che viene condiviso sui social media, in particolare su Twitter, al fine di identificare gli eventi che possono destare “allarme sociale”, ad esempio terremoti, inondazioni, alluvioni o altre situazioni di emergenza civile.

Uno degli obiettivi principali del progetto è realizzare un prototipo di piattaforma per il supporto decisionale nella gestione dell'emergenza, che analizzi in tempo reale i contenuti condivisi dagli utenti per rintracciare situazioni di pericolo nel momento stesso in cui avvengono nel mondo reale, anticipando i canali di informazione tradizionali.

Gli end users potenzialmente interessati a tali informazioni sono:

- I cittadini, che possono accedere a uno strumento di informazione tempestivo in caso di emergenze;
- La protezione civile;
- Le forze dell'ordine.

Il paradigma sfruttato da SoS – Social Sensing è quello dell'Opportunistic Crowdsensing, ossia degli utenti come dei sensore sociali: quando avviene un grande evento in una zona questi sensori si attivano e forniscono informazioni che possono essere raccolte e analizzate per fornire feedback istantanei.

In questo progetto sono stati presi in considerazione i terremoti come base di partenza, perché nella creazione di prototipi di rilevamento automatico di eventi tramite tweet, i terremoti sono facili da riscontrare grazie a fonti ufficiali come INGV¹³ e USGS¹⁴. Ciò è molto importante per avere una valenza scientifica e un riscontro senza il quale altrimenti non sarebbe possibile misurare l'accuratezza del sistema.

Un secondo obiettivo del progetto Social Sensing è quello di cercare di fare in tempo reale delle stime per il Damage Assessment, cioè per la valutazione dei danni a seguito di un

¹² www.socialsensing.it

¹³ “Istituto Nazionale di Geofisica e Vulcanologia”. www.ingv.it

¹⁴ “United States Geological Survey”. www.usgs.gov

terremoto, attraverso l'analisi delle parole più usate nei tweet o di immagini postate da cui è possibile risalire alle zone particolarmente colpite.

Il principio è quello di monitorare le parole chiavi usate nei tweet da tanti utenti in un lasso di tempo molto breve per individuare degli eventi in corso. Ad esempio parole chiavi come “scossa” e “terremoto” sono molto efficaci per intercettare tutti i messaggi che si riferiscono a quell'evento, e ogni evento in generale ha parole più ricorrenti che vengono utilizzate per descriverlo. In questo modo si può raccogliere tutta la mole di informazioni relativa a certi tipi di calamità, analizzarla e cercare di:

- Capire che c'è stato un evento (rilevamento/detection);
- Descriverlo;
- Fornire delle informazioni potenzialmente utili a prendere delle decisioni a seguito dell'evento.

Il volume dei tweet raccolti dipende dalla grandezza dell'evento ma anche da quante persone ci sono nella zona, poiché i “sensori” non sono distribuiti in maniera equivalente nel territorio.

Per l'Italia questo progetto ha già portato alla realizzazione di un prototipo per rilevare i terremoti attraverso l'analisi dei tweet.

3.1. Estensione del progetto

Questa tesi rappresenta l'ampliamento della piattaforma di SoS - Social Sensing: in particolare la metodologia usata per monitorare il territorio italiano è stata estesa al fine di coprire le zone di lingua inglese.

In particolare l'aggiunta di questa funzionalità si è articolata nelle seguenti fasi:

- Recupero dati da Twitter il cui testo contiene la parola “quake” o “shake”;
- Creazione di un web-tool per l'annotazione manuale di tweet;
- Annotazione manuale dei tweet raccolti per creare un training set con cui addestrare un modello supervisionato;
- Addestramento modello supervisionato in grado di riconoscere se un tweet riguarda un terremoto appena avvertito;
- Simulazione di situazioni di allerta terremoto su Twitter per testare il modello;
- Valutazione della bontà dei dati.

4. Strumento di annotazione manuale: #tweeTag

L'idea dello sviluppo di uno strumento per annotare manualmente i dati provenienti da Twitter è nata proprio dall'esigenza di ricavare in tempo reale, attraverso i tweet, informazioni su un dato evento critico in corso in maniera automatica. Per realizzare algoritmi di apprendimento e tecniche di machine learning basati su tali dati è infatti necessaria una fase iniziale di addestramento con dati annotati manualmente da umani che riescono a capire il contenuto testuale. Gli strumenti convenzionali e consolidati di annotazione sono addestrati su documenti lunghi in cui il testo è ampio.

Poiché i tweet sono composti da un massimo di 140 caratteri e hanno caratteristiche peculiari come la presenza di link¹⁵, abbreviazioni, hashtag¹⁶ e menzioni¹⁷, devono essere trattati in maniera diversa da come i sistemi convenzionali analizzano i testi ordinari.

I sistemi che annotano correttamente circa il 90% delle sequenze da un corpus giornalistico possono avere un calo di prestazioni fino al 50% su testi più informali [31].

Il problema è noto in letteratura ed esistono esempi di tool per l'annotazione manuale di social media data, ma non adattabili a più tipi di task grazie alla struttura del database di appoggio come #tweeTag.

Lo studio portato avanti in [20] ha evidenziato che le caratteristiche di microblogging sono più utili nella classificazione di tipo Sentiment Analysis grazie alla presenza di emoticon. In questo campo altri studi come [37] hanno condotto un lavoro di Sentiment Analysis¹⁸ usando un training set di tweet annotati manualmente, e per semplificare e accelerare gli sforzi di annotazione hanno costruito Tweenator4, anch'esso uno strumento di annotazione online che consente agli utenti di assegnare facilmente un'etichetta che descrive il sentimento espresso dal messaggio. Tuttavia Tweenator4¹⁹ è in grado di annotare tweet al solo scopo di assegnare un'etichetta negativa, positiva o neutra al sentimento espresso in un Tweet. Con #tweeTag è possibile fare oltre a questo tipo di annotazione anche altre campagne con scopi totalmente differenti personalizzando le etichette da attribuire.

¹⁵ “Collegamenti”

¹⁶ Parole o combinazioni di parole precedute dal simbolo cancelletto (“#”)

¹⁷ Citazione di un utente Twitter

¹⁸ L'analisi del sentimento è l'identificazione e l'estrazione delle valutazioni soggettive da opinioni, frasi, messaggi espressi in linguaggio naturale

¹⁹ www.tweenator.com

Un altro strumento presentato in [21] valuta la credibilità dei tweet. In questo lavoro, i tweet di un evento sono stati annotati da tre annotatori umani per valutarne la credibilità con uno specifico tool. Le etichette utilizzate, “Credibile”, “Non credibile” e “Non Rilevante”, sono state definite staticamente. Per elaborazioni successive sono stati considerati i soli tweet che almeno uno dei tre annotatori ha considerato credibili, senza calcolare il grado di accordo, come invece è possibile stabilire con #tweeTag. Inoltre l'interfaccia con cui si annotano i tweet è molto intuitiva, ma il tweet non ha l'embedded view, che rimane una prerogativa di #tweeTag.

Uno studio sull'annotazione del contenuto testuale del tweet è stato invece condotto in [13], in cui viene utilizzata la piattaforma di Amazon MTurk²⁰ per raccogliere semplici annotazioni di entità nominate dal testo dei tweet. Sfruttando i principi del crowdsourcing²¹, in cambio di un compenso di 5 cent per ogni tre annotazioni effettuate, ogni annotatore doveva contrassegnare le parole dei tweet proposti dal sistema usando una tra le seguenti etichette: persona (PER), organizzazione (ORG), posizione (LOC), "nessuna delle precedenti" (NONE) o incertezza. Il lavoro è molto simile a quello presentato in #tweeTag ma con interfaccia più ridondante a causa della ripetizione delle etichette per ogni parola. Sebbene un prodotto come MTurk sia ormai affidabile e ben collaudato, non permette la completa gestione dei dataset. Se è richiesto un lavoro più specifico nell'annotazione dei tweet potrebbe non essere lo strumento ideale da prendere in considerazione, in quanto non completamente adattabile e poco economico.

A causa di tutte queste mancanze non è stato possibile utilizzare per questo lavoro i tool già esistenti e si è resa necessaria la realizzazione di #tweeTag, un web-tool pensato per un approccio specifico nell'annotazione dei tweets. L'idea iniziale era di fare uno strumento per ottenere un output di dataset di tweet annotati riguardanti situazioni di emergenza e pericolo (catastrofi naturali, terremoti, alluvioni, ecc.), ma in una seconda fase il tool è stato esteso per permettere la realizzazione di più campagne di annotazione in base alle diverse necessità (ad esempio per la Sentiment Analysis).

Per “campagna di annotazione” si intende l'insieme delle operazioni che servono a preparare e annotare i tweet per uno specifico obiettivo. Ogni campagna ha dei tag, ovvero delle classi da assegnare ai tweet a seconda del task che la campagna si propone, e un tag “Skip” che dà la possibilità di non catalogare il tweet e saltarlo nel caso di ambiguità o dubbi.

#tweeTag è in grado di predisporre campagne di annotazione in tre modalità differenti: la prima per esaltare le macro caratteristiche del tweet nel suo complesso in funzione di un dato

²⁰ “Amazon Mechanical Turk”: piattaforma internet di crowdsourcing. www.mturk.com

²¹ Crowdsourcing è il processo attraverso cui si ottengono servizi, idee, risposte o contenuti da un ampio gruppo di persone, generalmente una community online, invece che dai fornitori tradizionali, con evidenti benefici di costo e di varietà di contributi.

contesto (se è inerente o meno ad un certo fenomeno, se è positivo o negativo, ecc.); la seconda per annotare il testo del tweet in modo da identificare informazioni riguardanti l'argomento o il contenuto (se vengono descritti vari tipi di danno a cose e/o persone, ecc.); la terza per valutare uno specifico utente attraverso l'analisi della sua timeline. L'applicazione è stata sviluppata in modo da eseguire campagne di annotazione personalizzate, al fine di soddisfare ogni tipo di esigenza.

4.1. Descrizione

#tweeTag è un'applicazione web accessibile mediante un qualsiasi browser. La scelta di un tool online consente un potenziale utilizzo futuro per un'interazione di tipo crowdsourcing. L'interfaccia sviluppata è innovativa, intuitiva e user-friendly, in modo da rendere il lavoro di annotazione manuale più gradevole e comodo.

Il tool consente di effettuare il login e accedere a una o più campagne di annotazione. L'applicazione è strutturata in tre pagine principali: la pagina di accesso, quella di annotazione e il profilo personale. Una volta effettuato il login l'utente viene reindirizzato alla pagina iniziale dove può selezionare la campagna di annotazione tra quelle che gli sono state assegnate; effettuata la scelta della campagna inizia (o prosegue) l'annotazione nell'apposita pagina.

L'apposita pagina, in cui si sottomettono le annotazioni, e la pagine del profilo si presentano con layout differenti a seconda del tipo di annotazione che richiede la campagna. I tipi di annotazione possono essere tre: globale, testuale e timeline.

4.2. Annotazione globale

Il primo tipo di annotazione è l'annotazione del tweet nel suo complesso, in modo da inquadrarlo all'interno di una particolare circostanza. L'annotazione globale presenta una vista embedded del tweet che permette un'analisi oltre che del testo, anche del contesto, ovvero delle informazioni relative al tweet, quali data e ora di pubblicazione, geolocalizzazione, numero di retweet e replies ecc., grazie alle quali l'annotatore è in grado di fornire un'annotazione più accurata e precisa in base al task che gli viene richiesto.

Questo primo tipo di annotazione è ideale per le campagne il cui obiettivo è quello di ricavare l'argomento di discussione del tweet o altre generalizzazioni a livello globale, ad

esempio se i tweet sono inerenti o meno ad un certo fenomeno preso in considerazione, se esprimono positività o negatività, e così via.



Figura 1. Esempio di tweet con vista embedded

Nella schermata principale vengono visualizzati la domanda da porsi durante l'annotazione, i primi 5 tweet della campagna da annotare. I tweet sono mostrati in ordine cronologico e c'è la possibilità di scegliere la data e l'ora dalla quale far partire i tweet spostandosi lungo l'arco temporale grazie alla presenza di un calendario realizzato con Datepicker e una casella apposita per l'orario. Una volta completata l'annotazione per un singolo tweet questo scomparirà e uno nuovo verrà accodato infondo alla lista rimasta e così via fino alla completa annotazione del dataset assegnato. L'utente può inoltre scegliere di annotare prima dei tweet che hanno una parola chiave, grazie a una pratica textbox in cui inserire del testo, e scegliere, grazie a un pratico calendario, di etichettare per primi i tweet della campagna postati in una certa data.



Figura 2. Pagina annotazione globale

Per visualizzare tutte le annotazioni effettuate fino a un certo momento è presente una pagina profilo. La pagina del profilo permette anche di revisionare, modificare e correggere le varie annotazioni compiute per una data campagna e fornisce una piccola statistica percentuale dei tweet ancora da annotare e di quelli già annotati.



Figura 3. Pagina profilo

4.3. Annotazione testuale

Il secondo tipo di annotazione è pensato per analizzare in maniera più approfondita il testo del tweet, in modo da sviscerare le informazioni sostanziali del contenuto. Nell'annotazione di dataset di tweet riguardanti catastrofi ad esempio, un task fondamentale è quello di definire per ogni tweet che parla di un danno, qual è il danno stesso (ad esempio crolli, dispersi, feriti), quali sono le entità coinvolte o qual è il luogo dell'avvenimento. Una annotazione di questo genere è importante per il Damage Assesment.

L'annotazione testuale prende in considerazione il solo testo del tweet su cui l'annotatore può selezionare una o più parole contigue etichettandole con diverse categorie. Vengono recuperati con una query uno alla volta i tweet che l'utente non ha ancora annotato per la campagna selezionata. Il testo del tweet viene scomposto a livello server nelle singole frasi e, a loro volta, nei singoli token che le compongono grazie allo strumento di analisi linguistica sviluppato dall'Italian NLP Lab²² del CNR. Il risultato viene mostrato nell'apposita pagina dove è stata inserita lato client l'opzione di selezionare i singoli token o gruppi di token adiacenti fino a ricoprire l'intera frase e annotarli con le etichette appropriate mostrate lateralmente. Ogni annotazione scelta va a riempire una tabella riepilogativa a fondo pagina con tutte le annotazioni effettuate sulle frasi fino a quel momento e una volta terminate è possibile inviare tutto in una volta sola al database.

Infine anche per questo tipo di annotazione c'è un profilo riepilogativo per rivedere le annotazioni e modificarle, tramite un link che rimanda alla pagina principale.

²² www.linguistic-annotation-tool.italianlp.it

Tweet totali: 1632 di cui 1632 annotati

frase:
 ITALIAN NEWS TERREMOTO ABRUZZO / I morti salgono a 150

frase:
 Riaperti i collegamenti con Roma : Secondo ... <http://tinyurl.com/dx3ore>

Clear frase | Ritorna al profilo senza salvare

Elementi selezionati:

Selezione	num frase	Token_in	Token_out	Value	Commento	Invia e torna al profilo
ABRUZZO	frase 1	4	4	GPE	<input type="text"/>	<input type="button" value="X"/>
Roma	frase 2	5	5	GPE	<input type="text"/>	<input type="button" value="X"/>
morti	frase 1	7	10	persone danneggiate	<input type="text"/>	<input type="button" value="X"/>

Figura 4. Pagina annotazione testuale

4.4. Annotazione timeline

Il terzo tipo di annotazione permette di valutare gli utenti Twitter attraverso l'annotazione della loro timeline. Ciò serve a individuare tendenze, consuetudini o anomalie nell'utilizzo di Twitter da parte dell'utente o identificare la presenza di profili falsi o inattivi.

L'annotazione dell'utente prende in considerazione gli ultimi 20 tweet della timeline mostrandoli attraverso una vista embedded.

La pagina di annotazione della timeline si presenta divisa in due colonne principali sovrastate dalla domanda a cui rispondere nell'annotazione. Nella colonna di sinistra si trova la lista di tutti gli screen name degli utenti che devono essere annotati. Nella colonna di destra viene visualizzata la timeline dell'utente selezionato. Al momento della selezione di un utente, insieme alla timeline, appaiono, vicino allo screen name dell'utente, anche i bottoni con le opzioni di annotazione e una textarea per eventuali commenti. Una volta annotato l'utente questo scomparirà dalla lista insieme alla sua timeline. Nel caso in cui l'annotatore pensi che la timeline di un dato utente non dia le informazioni necessarie per annotarlo, può cliccare sullo screen name dell'utente, che dotato di un link diretto alla pagina Twitter dell'utente, permetterà un'analisi più approfondita. Le annotazioni effettuate sono riepilogate nella pagina del profilo dell'annotatore con anche la possibilità di modificarle.

Nel profilo dell'annotazione timeline vengono mostrati gli screen name degli utenti annotati con la relativa annotazione e commento e con la possibilità di modifica.

Screen Name	Bot	Altro
@BiaManasse	<input type="checkbox"/>	<input type="checkbox"/>
@ContessaAlda	<input type="checkbox"/>	<input type="checkbox"/>
@_THESIGN_	<input type="checkbox"/>	<input type="checkbox"/>
@carloglietti	<input type="checkbox"/>	<input type="checkbox"/>
@maseratwit	<input type="checkbox"/>	<input type="checkbox"/>
@citazionime	<input type="checkbox"/>	<input type="checkbox"/>
@ilPetardo	<input type="checkbox"/>	<input type="checkbox"/>
@CarlaPoretti	<input type="checkbox"/>	<input type="checkbox"/>
@MickyMaruca	<input type="checkbox"/>	<input type="checkbox"/>

Figura 5. Pagina annotazione timeline

4.5. Inter Annotator Agreement

Una campagna di annotazione può essere assegnata a uno o più annotatori per avere più di una interpretazione dei dati e garantire che le annotazioni siano il più veritiere possibile.

Al termine di una campagna del tweet globale, per valutare il grado di accordo tra due o più annotatori è stato inserito il calcolo dell'Inter Annotator Agreement (K di Cohen).

#tweetTag
The annotation tool for Twitter data

Scegli la campagna Annotation Profilo

Calcolo Inter Annotator Agreement

L'**Inter Annotator Agreement** e' il grado di accordo tra gli annotatori nel classificare i dati in base alle categorie proposte. La misura dell'agreement permette di valutare la bontà dell'attività di annotazione svolta sul dataset. I dati raccolti possono essere considerati affidabili a seconda di quanto gli annotatori concordino nel classificare i tweet nelle classi proposte. Il coefficiente con cui è stato calcolato l'agreement è il k di Cohen, che esprime la concordanza reale, escludendo cioè quella dovuta al caso.

• Campagna scelta: **Campagna filtraggio tweets per terremoti inglesi**

• Annotatori che hanno partecipato alla Campagna (scegli gli annotatori su cui calcolare il grado di accordo):

meletti
 serena
 salvatore

[Calcola Kappa di Cohen](#)

K = 0.65834939997116 [Mostra i tweet su cui sono in disaccordo](#)

Figura 6. Pagina del calcolo dell'Agreement

La misura dell'Agreement permette di valutare la bontà dell'attività di annotazione svolta sul dataset. I dati raccolti possano essere considerati affidabili a seconda di quanto gli annotatori concordino nel classificare i tweet nelle classi proposte.

Il tool prevede due tipi di utenze:

- *Annotatori*: utenti che possono partecipare alle campagne che sono state assegnate loro e consultare solo le proprie annotazioni;
- *Super Annotatori*: utenti che hanno la possibilità di partecipare alle campagne assegnate e in più controllare l'Agreement delle campagne terminate assegnate a più di un annotatore. Inoltre possono entrare ancora di più nel dettaglio per vedere le annotazioni dei tweet su cui gli annotatori sono in disaccordo ed eventualmente disambiguare i casi andando a modificare le annotazioni di uno degli annotatori.

The screenshot shows the '#tweetTag' web application interface. At the top, there is a navigation bar with three tabs: 'Scegli la campagna', 'Annotation', and 'Profilo'. Below the navigation bar, the page title is 'Pagina del Super Annotatore'. A sub-header explains the tool's purpose: 'Da Super Annotatore oltre a controllare l' agreement puoi vedere ancor più nel dettaglio le annotazioni dei tweet su cui gli annotatori sono in disaccordo. Puoi disambiguare i casi controversi scegliendo un annotatore e modificando le sue annotazioni.'

Below the explanation, there is a section 'Modifica le annotazioni di:' with three radio button options: 'meletti', 'serena', and 'salvatore'. A link 'Oppure Torna al calcolo dell' agreement' is also present. The main content area is titled 'Tweet annotati' and includes a pagination control 'Pagina: 1 >>'. The content is organized into a table with three columns: 'Tweet', 'Annotazioni discordanti', and 'Tag Super Annotatore'.

Tweet	Annotazioni discordanti	Tag Super Annotatore
	<ul style="list-style-type: none"> • salvatore -> Inerente • serena -> Non Inerente 	<input type="button" value="Inerente"/> <input type="button" value="Non Inerente"/> <input type="button" value="Skip"/>
	<ul style="list-style-type: none"> • salvatore -> Skip • serena -> Inerente 	<input type="button" value="Inerente"/> <input type="button" value="Non Inerente"/> <input type="button" value="Skip"/>

Figura 7. Pagina modifica discordanze

Alla fine del processo di annotazione si ha un dataset annotato e ripulito da qualsiasi discordanza tra gli annotatori.

La presenza di una forte discordanza nelle annotazioni potrebbe essere causata sia da un errore da parte di un annotatore, sia da una cattiva interpretazione del task preso in esame. Può anche accadere che le linee guida di una campagna non siano spiegate correttamente e che quindi tutti gli annotatori commettano lo stesso errore producendo ugualmente un alto livello di Agreement o che si verifichi un alto grado di accordo tra gli annotatori per puro caso, ad esempio se le categorie sono poche è più probabile che essi si trovino d'accordo, in quanto ci sono meno opzioni tra cui scegliere. Quindi è bene tenere presente che un buon Agreement non garantisce sempre la validità delle annotazioni finali.

Attualmente la verifica del grado di accordo è disponibile solo per l'annotazione globale. Il coefficiente con cui è stato calcolato l'Agreement è il k di Cohen [8], che esprime la concordanza reale, escludendo cioè quella dovuta al caso. Per calcolarlo è necessario prima misurare la concordanza osservata (A_o), ossia il rapporto in percentuale delle valutazioni su cui concordano gli utenti che annotato gli stessi dati in maniera indipendente l'uno dall'altro. La concordanza osservata (A_o) si ottiene calcolando la media degli elementi su cui concordano gli annotatori. La concordanza osservata deve essere in seguito corretta tenendo conto

dell'Agreement casuale (A_e), ovvero della probabilità che gli annotatori attribuiscano per puro caso la stessa categoria a un elemento qualunque. La probabilità dovuta al caso (A_e) che gli annotatori concordino su di una qualsiasi categoria è il prodotto della probabilità che entrambi classifichino un elemento come appartenente a quella categoria.

L'equazione finale restituisce una stima della concordanza tra due annotatori:

$$\kappa = \frac{A_o - A_e}{1 - A_e}$$

Dove " $A_o - A_e$ " è la quantità di Agreement dovuta al caso e " $1 - A_e$ " è la stima di Agreement non dovuta al caso.

Il κ di Cohen può assumere valori compresi nell'intervallo [0,1]:

- $\kappa = 0$ concordanza interamente casuale;
- $\kappa = 1$ concordanza perfetta.

Per l'interpretazione dei valori di κ è stata usata la scala di Landis e Koch²³:

K	CONCORDANZA
≤ 0	Nulla
0.01 - 0.20	Scarsa
0.21 - 0.40	Modesta
0.41 - 0.60	Moderata
0.61 - 0.80	Sostanziale
0.81 - 1.00	Eccellente

Tabella I. Scala di Landis e Koch

4.6. Implementazione

Il punto di partenza dell'implementazione di #tweeTag è la progettazione del database di appoggio MySQL²⁴, indispensabile per memorizzare i dati riguardanti gli utenti, e le attività di annotazione svolte.

²³ J. R. Landis, G. G. Koch. Biometrics. 1977.

²⁴ Sistema per la gestione di basi di dati relazionali. www.mysql.com

Il database è organizzato in quattro tabelle:

- *Annotations*: contiene le annotazioni effettuate dai singoli annotatori;
- *Annotators*: contiene le informazioni relative ai profili utenti;
- *Assignments*: permette di assegnare le campagne a certi annotatori;
- *Campaigns*: contiene le campagne di annotazione da effettuare su dataset salvati in database esterni;
- *Tags*: contiene le etichette da assegnare.

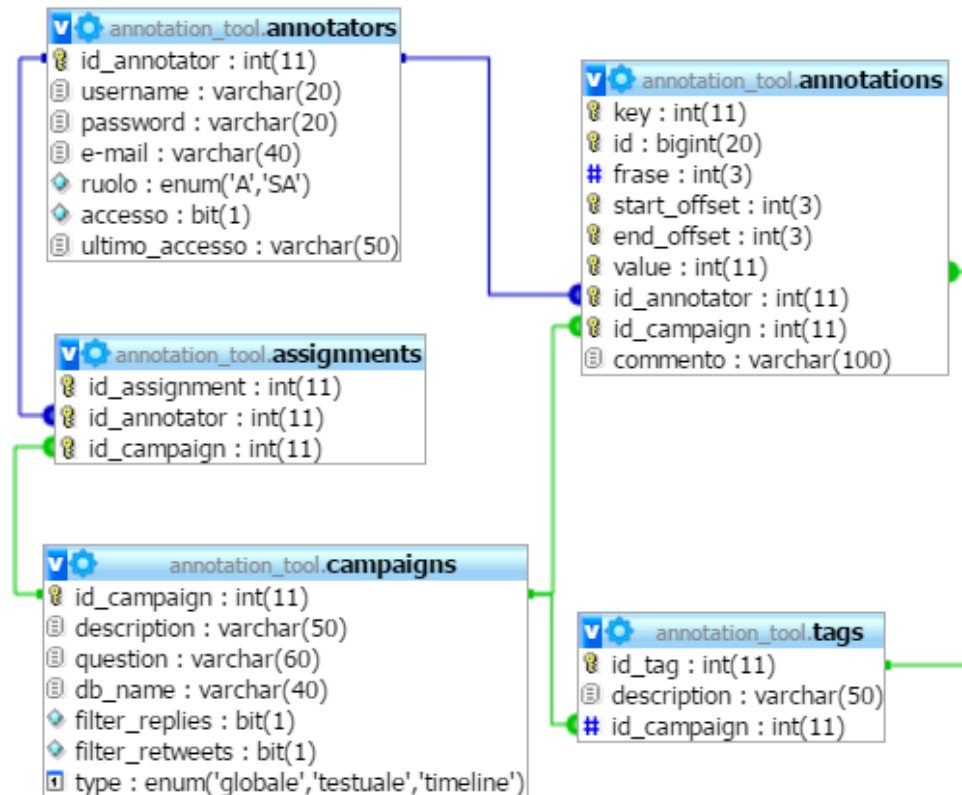


Figura 8. Schema ER del Database

Attualmente le credenziali delle utenze, le campagne, i tags e le assegnazioni sono gestite direttamente dall'amministratore.

L'applicazione è stata sviluppata su sistema operativo Windows 7, adottando come browser di riferimento Mozilla Firefox. Per creare un ambiente di sviluppo facile da gestire è stato installato XAMPP²⁵, una piattaforma software gratuita costituita dal server HTTP

²⁵ www.apachefriends.org

Apache²⁶, il database MySQL e tutti gli strumenti necessari per utilizzare i linguaggi di programmazione PHP.

Per sviluppare l'applicazione dinamica è stato usato PHP²⁷ 5.3.8, un linguaggio di scripting lato server, interpretato, con licenza open source e indipendente dalla piattaforma o dal sistema operativo. Le pagine HTML vengono create nel momento stesso in cui queste vengono richieste, in modo da aggiornarne i dati in tempo reale. Per la gestione dinamica delle componenti HTML si è fatto ricorso alla libreria Javascript jQuery²⁸ 1.8.2. Inoltre è stata adottata jQueryUI²⁹, libreria open source di plugin basata su jQuery che fornisce interazioni ed animazioni e widget. In particolare è stato integrato il widget Datepicker³⁰.

E' stata utilizzata anche AJAX³¹, una tecnica di sviluppo software per la realizzazione di applicazioni web interattive. Lo sviluppo di applicazioni HTML con AJAX si basa su uno scambio di dati in background fra web browser e server, consentendo l'aggiornamento dinamico di una pagina web senza esplicito ricaricamento da parte dell'utente e quindi senza interferire con il comportamento della pagina esistente.

I dati scambiati in background sono in formato JSON³², adatto per lo scambio dei dati in applicazioni client-server e soprattutto in applicazioni web realizzate con tecnologia AJAX. La classificazione dei tweet può avvenire senza mai ricaricare la pagina grazie all'utilizzo della tecnologia AJAX.

4.7. Esempi di utilizzo del tool

Per testare il funzionamento dell'applicazione sono state organizzate diverse campagne di annotazione.

La sezione dell'annotazione globale è stata testata con sei diverse campagne di annotazione, l'annotazione testuale con una e quella della timeline con due campagne. Le prime tre campagne di tipo globale hanno analizzato tre dataset di tweet riguardanti calamità naturali avvenute in Italia, in particolare il terremoto in Emilia Romagna del 2012, l'alluvione in Sardegna del 2013 e l'alluvione di Senigallia del 2014, per un totale di 4.595 tweet e l'annotazione richiedeva per ogni tweet di rispondere alla domanda "Il tweet si riferisce ad un

²⁶ www.apache.org

²⁷ "PHP: Hypertext Preprocessor". www.php.net;

²⁸ www.jquery.com

²⁹ www.jqueryui.com/

³⁰ www.jqueryui.com/datepicker

³¹ "Asynchronous JavaScript and XML". www.w3schools.com/ajax/

³² "JavaScript Object Notation". www.json.org

evento? Se sì, ci sono danni?”. In base a questa domanda i tweet venivano suddivisi in tre categorie:

- “*Non inerente*”, per annotare i tweet che non si riferiscono ad un particolare evento (es. “Mio figlio è un terremoto”);
- “*Senza danno*”, per annotare i tweet che si riferiscono ad un evento ma che non danno alcuna informazione sui danni provocati (es. Esempio: “Abito al 4 piano, se crolla tutto... Non riesco neanche a pensarlo... #terremoto”);
- “*Con danno*”, per annotare i tweet inerenti ad un dato evento e che portano informazione sui danni che sono stati provocati in modo che questi possano essere quantificabili (es. “A #Carpi solo calcinacci e qualche crepa. Menomale! #terremoto”).

Dei 4.595 tweet i tre annotatori hanno annotato circa 1.531 tweet ciascuno. I tweet risultano così ripartiti: 593 non inerenti a un evento, 1.320 tweet che riportano danno, 2.393 tweet inerenti all’evento ma che non conducono a un danno e 289 annotazioni saltate a causa di ambiguità e tweet con restrizione.

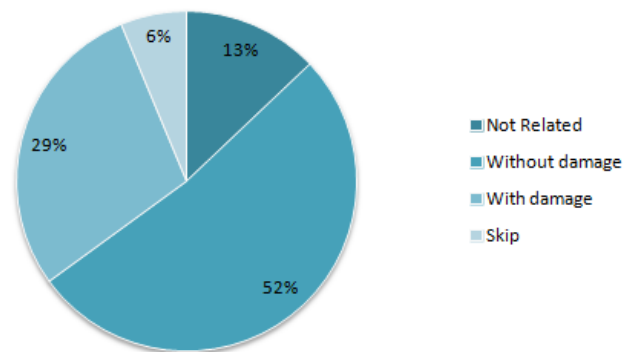


Grafico 1. Risultati della prima campagna di annotazione globale

Le successive campagne di annotazione globali con cui è stato testato #tweeTag prevedevano tipologie di analisi che non erano state contemplate nell’idea iniziale d’uso del tool. In queste campagne si è cercato di annotare in una le emozioni espresse nei tweet (Sentiment Analysis), e nell’altra la tempestività del messaggio rispetto ad un evento. Ciò è stato possibile modificando i tag a livello di database in base al task richiesto.

Nella campagna di Sentiment Analysis all’annotatore veniva chiesto di riconoscere se i tweet si riferissero ad un evento e di rispondere alla domanda “il tweet comunica emozioni di paura?”: se sì dare un valore al grado di paura espresso nel tweet. Le opzioni di scelta possibile per annotare il tweet erano:

- “*Non inerente*”, per annotare i tweet che non si riferiscono ad un particolare evento (es. "come stai?");
- “*Non comunica paura*”, per i tweet che si riferiscono ad un evento ma che non esprimono alcun sentimento di paura (es. “questa mattina ho avvertito terremoto”);
- “*Lieve paura*”, riferito ad un determinato tweet che rende appena percepibile un sentimento (es. “nottata passata in piedi causa terremoto”);
- “*Moderata paura*”, per annotare i tweet che trasmettono un sentimento di moderata paura (es. “scossa di terremoto forte, temo sia successo qualcosa, speriamo niente di grave”);
- “*Forte paura*”, per i tweet che esprimono un sentimento forte (es. “oh mio Dio il terremoto, muoio dalla paura!!!”).

Di 1.500 tweet sono risultati: 4 non annotati, 64 saltati, 289 tweet con forte paura, 212 con moderata paura, 131 con lieve paura, 788 tweet non comunicano paura e 11 non inerenti.

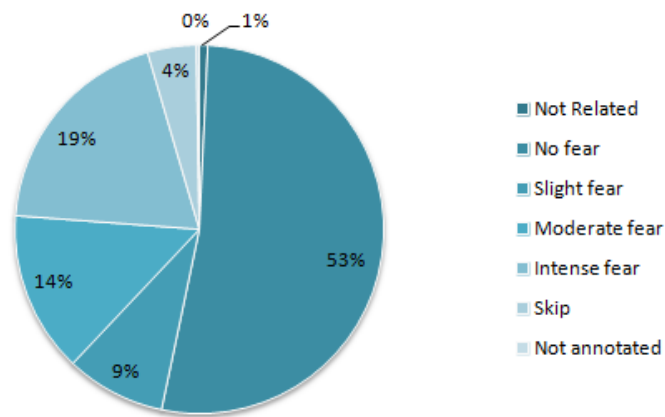


Grafico 2. Risultati della seconda campagna di annotazione globale

L’ultima campagna di annotazione globale è stata portata avanti da tre annotatori e analizzava un dataset di 2.414 tweet in lingua inglese postati durante il terremoto avvenuto il 15 gennaio 2014 in California. Veniva richiesto all’annotatore di rispondere alla domanda “il tweet segnala un terremoto in corso?”. In base a questa domanda le categorie presenti per l’annotazione erano:

- “*Inerente*”, per indicare tweet che si riferiscono al terremoto in corso preso in considerazione (es. “questo terremoto è stato lieve, speriamo sia finito qui”);
- “*Non inerente*”, per annotare tweet che si riferiscono ad un terremoto passato (es. “la scossa più forte di terremoto che ho mai sentito risale al 1997”).

Di 2.414 tweet annotati: 1.034 sono risultati inerenti, 860 non inerenti e 520 saltati. Con questi valori di dataset annotati sarebbe già possibile iniziare ad addestrare qualche semplice algoritmo di apprendimento supervisionato.

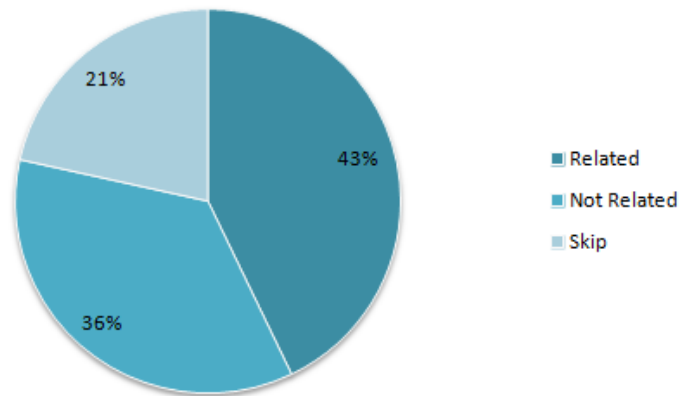


Grafico 3. Risultati della terza campagna di annotazione globale

Esperimenti di annotazioni testuali sono stati effettuati su una sola campagna il cui dataset era composto dai tweet annotati come “con danno” durante le prime tre campagne di annotazione globale. Nell’annotazione testuale l’annotatore si trova davanti il solo testo del tweet e gli viene richiesto di annotare diversi tipi di entità utili a capire se nel testo sono presenti informazioni riguardanti dei danni a cose e/o persone. Le macro categorie prese in considerazione per l’annotazione testuale in questione sono entità nominate, entità di dominio ed eventi. Per quanto riguarda le entità nominate le possibili scelte dell’annotatore sono tra:

- “*Nome di persona*”, utilizzato per indicare i nomi di persone coinvolte nell’evento (es. Marco, Mario Rossi);
- “*Nome di luogo*”, usato per indicare i nomi di monti, laghi, ecc. (Es. Il Cimone, L’appennino apuano, Lago di Garda);
- “*Nome di Geo Political Entities*”, usato per indicare entità geografiche con limiti amministrativi (es. Milano, Toscana, Provincia di Palermo);
- “*Organizzazioni*” (es. croce rossa, RAI, Nutella).

All’interno della categoria entità di dominio si trovano le etichette:

- “*Persona danneggiata*”, che indicano gli individui che hanno subito un danno (morto, disperso, ferito);
- “*Entità danneggiata*”, utilizzata per indicare le entità inanimate che hanno subito un danno (es. calcinacci, strada dissestata, ponte crollato, crepe).

Per quanto riguarda la categoria degli eventi troviamo una sola etichetta:

- “*Verbo evento*” che viene utilizzato per indicare l’azione avvenuta (es. è crollato, è annegato).

Le due campagne di annotazione delle timeline invece hanno analizzato un complessivo di 2.477 utenti, chiedendo agli annotatori di indicare se questi fossero utenti reali o fake. La prima campagna analizzava un dataset di utenti ottenuti random da Twitter, mentre la seconda si concentrava su di un campione di follower di Obama.

4.7.1. Considerazioni

I test effettuati sono stati utili al fine di valutare il grado di adeguatezza dell’applicazione verso il task proposto. I risultati ottenuti hanno evidenziato uno strumento funzionale, usabile e efficace secondo gli obiettivi iniziali. I dataset annotati possono essere riutilizzati per un’analisi statistica o per addestrare algoritmi di machine learning supervisionato.

I risultati dei test mostrano un metodo intuitivo per presentare le annotazioni e un grande potenziale per focalizzare l’attenzione sui tweet e inquadrarli al meglio.

La particolarità che caratterizza #tweeTag è la possibilità di concentrare in un unico strumento tre approcci diversi all’annotazione di dati provenienti da Twitter e grazie alla struttura del database di appoggio e alla capacità di adattarsi a ogni tipo di task, possibili progetti futuri possono prevedere l’ampliamento dell’applicazione come strumento di annotazione di riferimento per i tweet, con una configurazione di tipo crowdsourcing e una piattaforma open source, permettendo una registrazione sia per gli annotatori che per persone esterne (requester) che desiderano creare mediante interfaccia una campagna di annotazione, in modo da adeguarsi a ogni tipo di esperienza.

5. Sistema per la rilevazione dei terremoti tramite tweets

La fase successiva di questo lavoro si è concentrata nella realizzazione di un sistema in grado di rilevare la presenza di terremoti in nazioni di lingua inglese, monitorando l'attività su Twitter. Il sistema analizza il testo dei tweet per capire se ciò che cerca di trasmettere l'utente è una situazione di pericolo; individua di un burst di tweet, cioè un certo numero di tweet simili in un breve intervallo di tempo; e associa ad essi un evento in corso, in questo caso un terremoto.

Una fase essenziale per il raggiungimento di tale obiettivo è stato lo sviluppo di un modello di classificazione supervisionato che, attraverso esempi di tweet annotati manualmente, riesca a stabilire (assegnando appunto la classe di appartenenza) se i tweet analizzati si riferiscano o meno a un terremoto.

Un classificatore basato su apprendimento supervisionato fa uso di un insieme di addestramento per apprendere un certo tipo di comportamento: il modello di classificazione; esistono anche algoritmi di apprendimento non supervisionato, che non fanno uso di un insieme di addestramento. Il classificatore creato in questo ambito è supervisionato e il dataset utilizzato per addestrarlo è composto da un campione di tweet già etichettati e valutati inerenti o meno a un terremoto in corso mediante il tool #tweeTag.

5.1. Acquisizione dei dati

Innanzitutto sono stati acquisiti i dati da Twitter. Per questo lavoro la campagna di acquisizione è stata realizzata prendendo da Twitter i tweet e le informazioni sui relativi utenti postati nell'arco di 98 giorni, a partire dal 30 maggio 2014 fino al 4 settembre 2014.

I tweet sono stati raccolti:

- In lingua inglese;
- Contenenti almeno una keyword tra “quake” o “shake”, o le loro derivazioni.

In totale sono stati raccolti circa 500 mila tweet memorizzati in un database MySQL.

Per avere una lista dei terremoti avvenuti in quel periodo sono stati scaricate da USGS tutte le informazioni dei terremoti registrati tra il 30 maggio 2014 e il 4 settembre 2014 nelle nazioni in cui si parla prevalentemente inglese (USA, Antigua and Barbuda, Australia, The Bahamas, Barbados, Belize, Canada, Grenada, Guyana, Irlanda, Jamaica, New Zealand, St Kitts and Nevis, St Lucia, St Vincent and the Grenadines, Trinidad and Tobago, Regno Unito).

In totale sono stati recuperati 1.488 terremoti con intensità di magnitudo tra 3 e 7.9.

5.2. Campagna di annotazione dei dati acquisiti mediante tool #tweeTag

Del dataset di tweet acquisito è stato estratto un campione di 5.590 tweet da annotare manualmente. Il campione è stato estratto con due script:

- 3.234 scelti in maniera randomica;
- 2.356 scelti tra i tweet postati entro i 5 minuti dopo l'avvenimento di un terremoto indicato da USGS, per essere sicuri di annotare un campione inerente.

L'annotazione di questo dataset è stata effettuata grazie allo strumento di annotazione manuale web #tweeTag da tre annotatori. La domanda che veniva loro posta è stata “*Il tweet si riferisce a un terremoto appena accaduto?*”. La classe da assegnare a un tweet era una a scelta tra tre classi disponibili:

- “*Inerente*”, il tweet parla di un terremoto appena avvertito;
- “*Non inerente*”, il tweet non si riferisce a un terremoto appena avvenuto (es. riguarda terremoti passati, giornali che riportano una scossa, ecc.);
- “*Skip*”, tweet ambiguo e non annotabile.

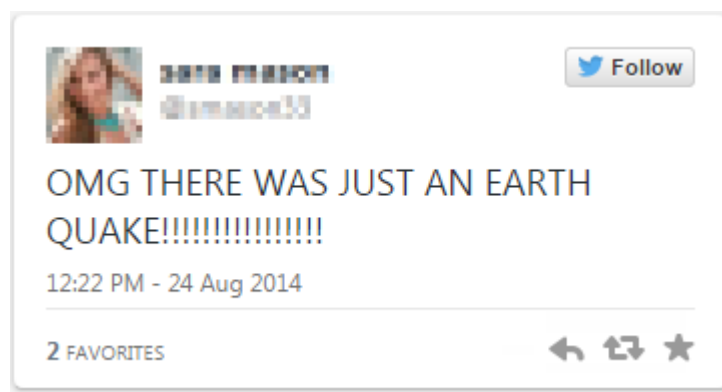


Figura 9. Esempio di tweet inerente

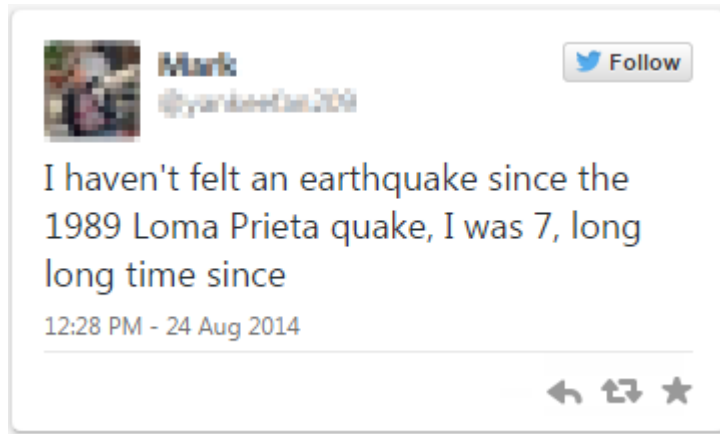


Figura 10. Esempio di rumore

Terminata la campagna di annotazione l'accordo tra gli annotatori si è rivelato sostanziale e un super annotatore si è occupato di scegliere tra le discordanze degli annotatori. Alla fine il dataset è risultato così ripartito:

- 1.698 tweet inerenti;
- 3.771 tweet non inerenti;
- 85 tweet “skippati”.

Eliminati i tweet non annotati (“Skip”) si è ottenuto il dataset di addestramento composto da 5.469 tweets.

Come si può notare dal numero di tweet non inerenti, il rumore è molto ricorrente in questo tipo di dati. Durante lo studio dei tweet è infatti emerso che i tweet contenenti la parola shake apportano una notevole ambiguità. In lingua inglese infatti la parola “shake” può significare diverse cose: scossa, fare cenno di no con il capo, affaticamento, stretta di mano e molto altro ancora. Anche i retweet³³ e i replies³⁴ sono informazione di rimbalzo perché chi la inoltra non è più il testimone oculare dell'evento. Inoltre utenti e bot³⁵ potrebbero cavalcare il trend del momento e utilizzare le keyword per fare spam. Quindi riuscire a filtrare il rumore è un task essenziale al fine di rilevare l'evento tra la marea di informazioni che attraversano Twitter ogni minuto.

I tweet annotati come “Inerenti” sono accomunati da alcune caratteristiche: il testo è molto corto, spesso in lettere maiuscole e senza punteggiatura. Inoltre contengono emoticon e alcune

³³ Condivisioni di tweet originali da parte di altri utenti

³⁴ Risposte ai tweet postati

³⁵ Contrazione di robot. Programma che si comporta come una persona.

espressioni di sorpresa (es. “omg”). Le persone che hanno appena avvertito un terremoto utilizzano nei loro post un linguaggio informale, diretto, che esprime spavento e con una struttura grammaticale semplice.

5.3. Addestramento del Classificatore

A partire dai tweet annotati sono stati individuati una serie di attributi chiamati features, ovvero caratteristiche rilevanti del tweet espresse in formato numerico, dove uno di questi attributi è la classe, che in questo caso è una delle categorie di annotazione (“Inerente”, “Non inerente”). In questo modo il classificatore è riuscito a imparare a riconoscere la classe di appartenenza di un tweet in funzione dei valori degli altri attributi.

Nel dettaglio l’attributo classe a cui un tweet può appartenere in questo ambito può assumere valore:

- 0: se il tweet è rumore;
- 1: se il tweet si riferisce a un terremoto in corso.

Le features estratte per ogni tweet dal DB sono state:

- Numero di caratteri del testo del tweet;
- Numero delle parole nel testo del tweet;
- Presenza di punteggiatura all’interno del testo;
- Numero di punti esclamativi nel testo del tweet;
- Numero di URLs nel testo;
- Numero di menzioni nel testo;
- Numero di caratteri maiuscoli rispetto al numero totale di caratteri;
- Presenza di numeri il cui formato rimanda alla magnitudo di un terremoto;
- Classe del tweet.

Lo strumento utilizzato per la classificazione è Weka³⁶. Weka è un software open source per l’apprendimento automatico sviluppato nell’università di Waikato in Nuova Zelanda. Weka è un ambiente software interamente scritto in Java che consente di applicare dei metodi di

³⁶ “Waikato Environment for Knowledge Analysis”. www.cs.waikato.ac.nz/ml/weka

apprendimento automatici (learning methods) ad un set di dati (dataset) e analizzarne il risultato. È possibile attraverso questi metodi avere una previsione del comportamento di nuovi dati.

L'algoritmo di apprendimento utilizzato su Weka è l'albero di decisione J48 con 10-cross validation, che suddivide il dataset in dieci sottoinsiemi, allena il sistema su nove di questi e lo testa sul sottoinsieme restante; ripete il meccanismo dieci volte cambiando ogni volta il sottoinsieme su cui effettuare il test e calcola la media dei risultati. Il modello generato è stato in grado di classificare correttamente l'81.642% (4465 su 5469) dei tweet costituenti il training set, un risultato molto soddisfacente. L'output del modello contiene inoltre una serie di indici che forniscono delle indicazioni sui risultati del processo di classificazione:

- *TP Rate*: indica la percentuale di campioni effettivamente positivi rilevati tali dal classificatore;
- *FP Rate*: complementare al TP Rate, indica la percentuale di campioni effettivamente negativi rilevati tali dal classificatore;
- *Precision*: misura di correttezza che corrisponde alla percentuale degli elementi correttamente rilevati rispetto al totale degli elementi rilevati presenti nel sistema ($Precision = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}$);
- *Recall*: misura di completezza che corrisponde alla percentuale degli elementi correttamente rilevati rispetto al totale degli elementi presenti nel sistema ($Recall = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}$);
- *F-Measure*: media armonica pesata ($2 * \frac{Recall * Precision}{Recall + Precision}$).

Le prestazioni del classificatore addestrato sono quelle riportate in tabella:

<i>TP Rate</i>	<i>FP Rate</i>	<i>Precision</i>	<i>Recall</i>	<i>F-Measure</i>	<i>Classe</i>
0.726%	0.143%	0.696%	0.726%	0.71%	1
0.857%	0.274%	0.874%	0.857%	0.866%	0

Tabella II. Dettagli dell'accuratezza del classificatore per classe

Weka consente di salvare il modello generato e di riutilizzarlo in seguito per predire la classe di appartenenza di nuovi tweet mai visti prima. Se il modello è soddisfacente potrà essere impiegato in applicazioni real time per eseguire la predizione e il filtraggio di tweet ricavati direttamente dallo stream di Twitter, grazie alla sua velocità di decisione.

5.4. Simulazione delle attività degli utenti Twitter

Per questo lavoro il modello non è stato utilizzato in applicazioni real time ma è stato testato con un simulatore messo a disposizione dal CNR per ricreare lo stream di Twitter recuperando i tweet già raccolti dal database come se venissero appena postati.

I tweet recuperati vengono dati in pasto al modello che scarta quelli inutili. Quando viene intercettato un burst di tweet inerenti in una breve finestra temporale, ovvero si assiste al verificarsi di un grosso numero di tweet che segnala l'evento appena avvertito dagli utenti, all'interno di una finestra temporale ristretta, avviene la rilevazione dell'evento. Questa rilevazione viene confrontata con i terremoti ufficiali segnalati da USGS e se c'è un match viene registrato nel DB.

Il risultato della simulazione produce tre rilevazioni distinte:

- True positive (TP): i terremoti riconosciuti dal sistema e confermati da USGS;
- False Positive (FP): i terremoti individuati ma non confermati da USGS;
- False Negative (FN): i terremoti non riconosciuti dal sistema ma realmente avvenuti.

Dalla simulazione dello stream di tweet sono stati rilevati, attraverso i messaggi degli utenti, 30 terremoti. Tra questi gli eventi Veri Positivi (TP), cioè i terremoti che risultano realmente avvenuti e confermati dalla fonte ufficiale USGS, sono 12. Andando ad analizzare i tweet che hanno contribuito alla rilevazione dei Veri Positivi è stato possibile accertare che essi si riferissero realmente ai terremoti che avevano una corrispondenza con i dati USGS, ad esempio il testo dei tweet conteneva il luogo in cui stava avvenendo il terremoto, oppure il luogo di provenienza degli utenti che hanno postato i tweet corrispondeva con il luogo del terremoto rilevato. Gli altri 18 terremoti risultano Falsi Positivi, ovvero non confermati. I terremoti rilevati sono molto inferiori rispetto al totale dei terremoti ufficiali raccolti, ma bisogna tenere in considerazione il fatto che la maggior parte di essi sono situati nell'oceano, in zone poco popolate come l'Alaska o in profondità tali da essere non percepibili. Infatti, come verrà mostrato nel capitolo successivo, le rilevazioni coincidano con le zone demograficamente più popolate, ovvero quelle in cui sono presenti più sensori in grado di condividere in massa l'avvenimento sui social in tempo reale.

E' sicuro che i risultati evidenziati dal sistema hanno un potenziale margine di miglioramento essendo questo un lavoro in fase sperimentale. Sarà possibile utilizzare il modello anche in applicazioni real time una volta completati ulteriori test.

6. Rappresentazione grafica USA dei risultati incrociando diversi dataset

La fase finale del lavoro di questa tesi è stata quella di incrociare i risultati ottenuti nelle fasi precedenti per ottenere una rappresentazione grafica e semplificata dell'esito finale.

Quello che si è cercato di fare è una rappresentazione su mappa della situazione degli Stati Uniti durante il periodo preso in considerazione, dal 30 maggio 2014 al 4 settembre 2014, mostrando la situazione demografica, i terremoti effettivamente avvenuti e i terremoti rilevati dal sistema attraverso il monitoraggio dell'attività degli utenti Twitter.

L'obiettivo finale è stato dimostrare come il sistema sia maggiormente predisposto a rilevare eventi di allerta in zone maggiormente popolate, in quanto presumibilmente più gente avrà la possibilità di twittare al riguardo. Inoltre si è cercato di stabilire che tipo di relazione c'è tra i terremoti rilevati e la loro magnitudo e profondità. Si supponeva che i terremoti rilevati fossero quelli con magnitudo più alta e profondità minore.

La scelta di mostrare solo gli Stati Uniti nella mappa è data dal fatto che i dati raccolti sono avvenuti prevalentemente sul territorio statunitense. La maggior parte dei terremoti avvenuti nelle altre nazioni si trovano principalmente a largo dell'oceano e quindi non sono percepibili.

Per la realizzazione della mappa è stato necessario raccogliere in maniera opportuna tutti i dati necessari da fonti ufficiali.

6.1. Elenco dataset

6.1.1. Dati USGS - posizione, profondità, magnitudo terremoti

La lista dei terremoti avvenuti tra il 30 maggio 2014 e il 4 settembre 2014 è stata estratta dal sito ufficiale di USGS, che conserva le informazioni di tutti i terremoti avvenuti nel mondo.

Sono stati recuperati e inseriti in un DB i terremoti avvenuti in tale periodo nelle nazioni in cui si parla prevalentemente inglese (USA, Antigua and Barbuda, Australia, The Bahamas,

Barbados, Belize, Canada, Grenada, Guyana, Irlanda, Jamaica, New Zealand, St Kitts and Nevis, St Lucia, St Vincent and the Grenadines, Trinidad and Tobago, Regno Unito).

In totale sono stati recuperati 1.488 terremoti con intensità di magnitudo tra 3 e 7.9. I risultati sono stato salvati in un database insieme ad altre informazioni per ogni terremoto, tra cui data e ora di avvenimento, geolocalizzazione, nome del luogo, epicentro, magnitudo, profondità ecc.

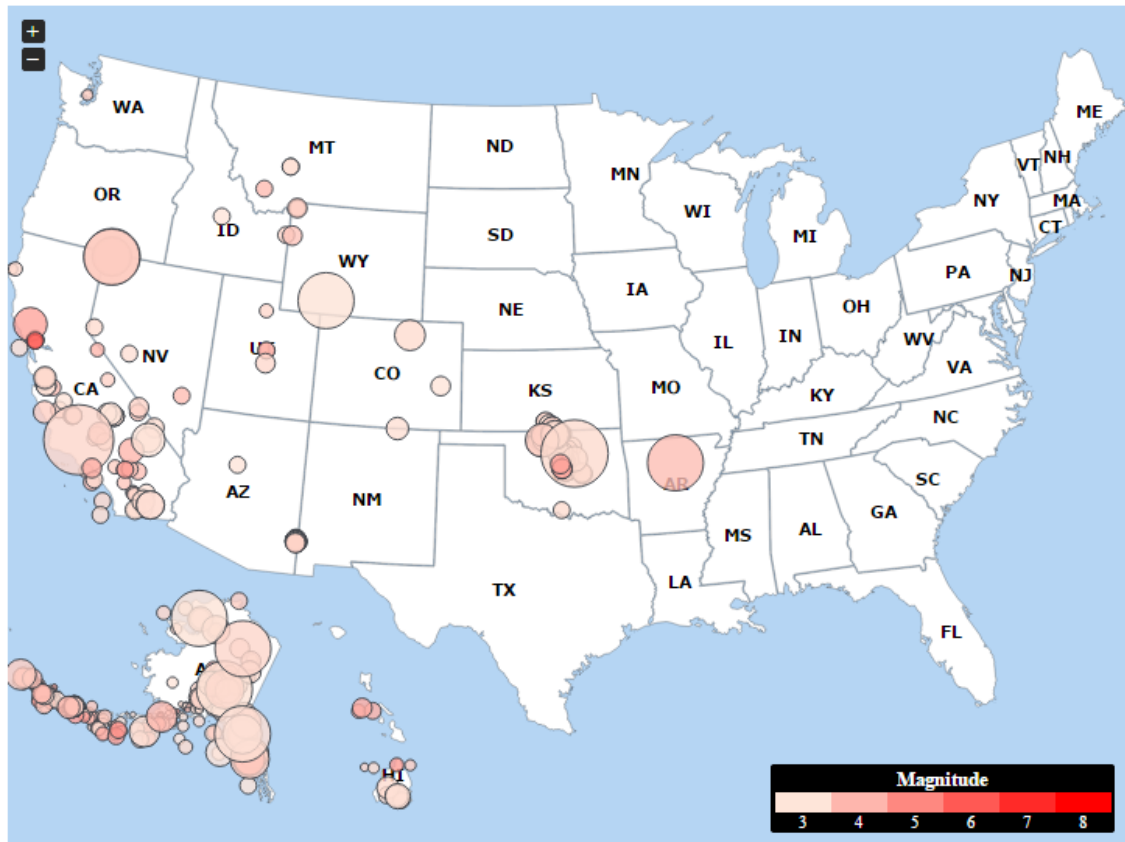


Figura 11. Terremoti avvenuti in USA dal 30 maggio al 4 settembre 2014 (USGS.gov)

La maggior parte dei terremoti recuperati sono avvenuti nell'oceano o in profondità tali da non essere percepiti.

USGS fornisce tra le informazioni aggiuntive anche un campo "felt" che indica il numero totale degli utenti che hanno garantito di aver sentito il terremoto attraverso sondaggi. Affidandoci quindi alle informazioni di USGS se si escludono i terremoti con campo "felt" uguale a zero si ottengono 227 terremoti così ripartiti negli stati Stati Uniti:

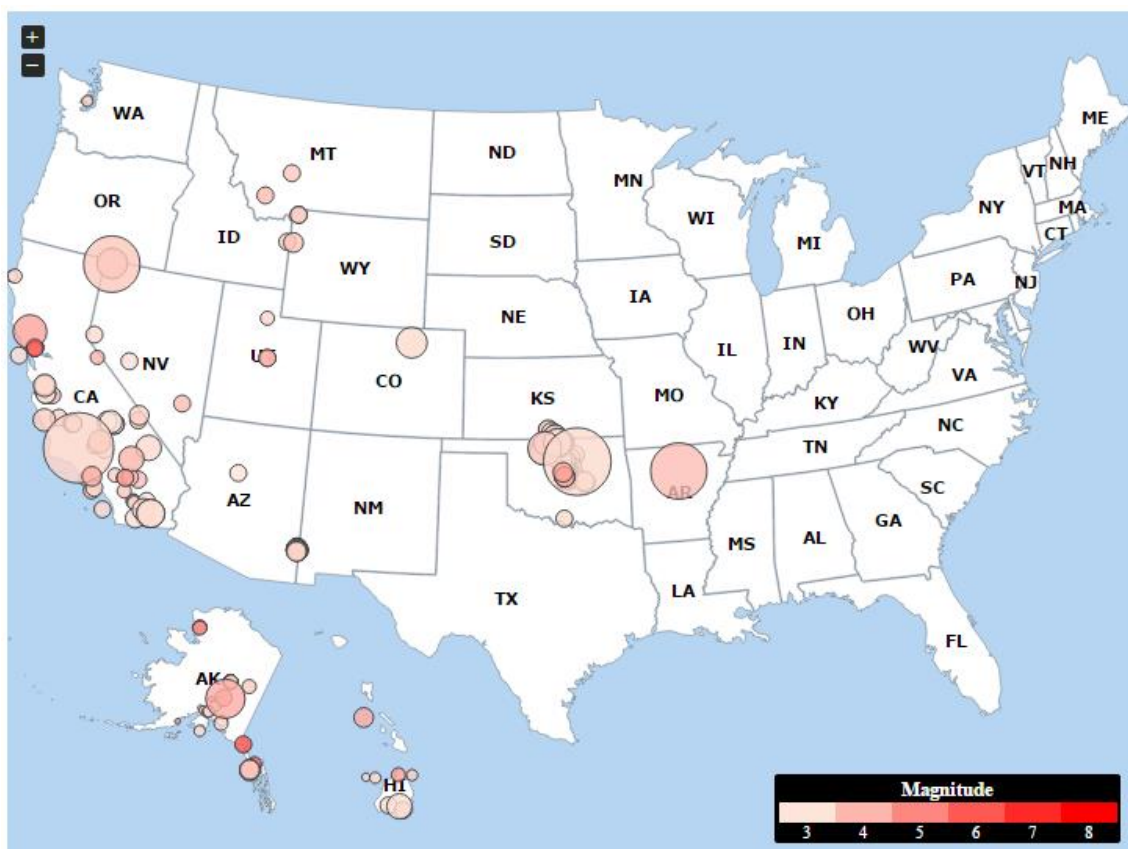


Figura 12. Terremoti avvenuti in USA nel periodo considerato escludendo quelli non rilevati nei sondaggi (USGS.gov)

Questa riduzione è utile per avere un quadro più semplificato della situazione senza perdere i terremoti concreti, in quanto i terremoti eliminati sono quelli avvenuti in oceano o nelle coste più isolate dell'Alaska, che il sistema non rileva a causa della scarsa presenza di sensori sociali.

6.1.2. Terremoti rilevati dal sistema

Come secondo obiettivo si è cercato di creare una mappa che mostrasse i terremoti che il sistema ha rilevato grazie alla presenza di un cospicuo numero di utenti che ha tweettato le proprie sensazioni poco dopo averlo avvertito.

I risultati del sistema hanno indicato 30 terremoti rilevati grazie alle informazioni ricavate dai tweet, di cui 12 terremoti confermati con la fonte ufficiale USGS, 9 avvenuti negli Stati Uniti e 3 in Nuova Zelanda. La rappresentazione su mappa USA mostra che i terremoti rilevati dal

sistema si trovano principalmente nella zona occidentale del paese, in particolare in California, nota zona sismica molto popolata.

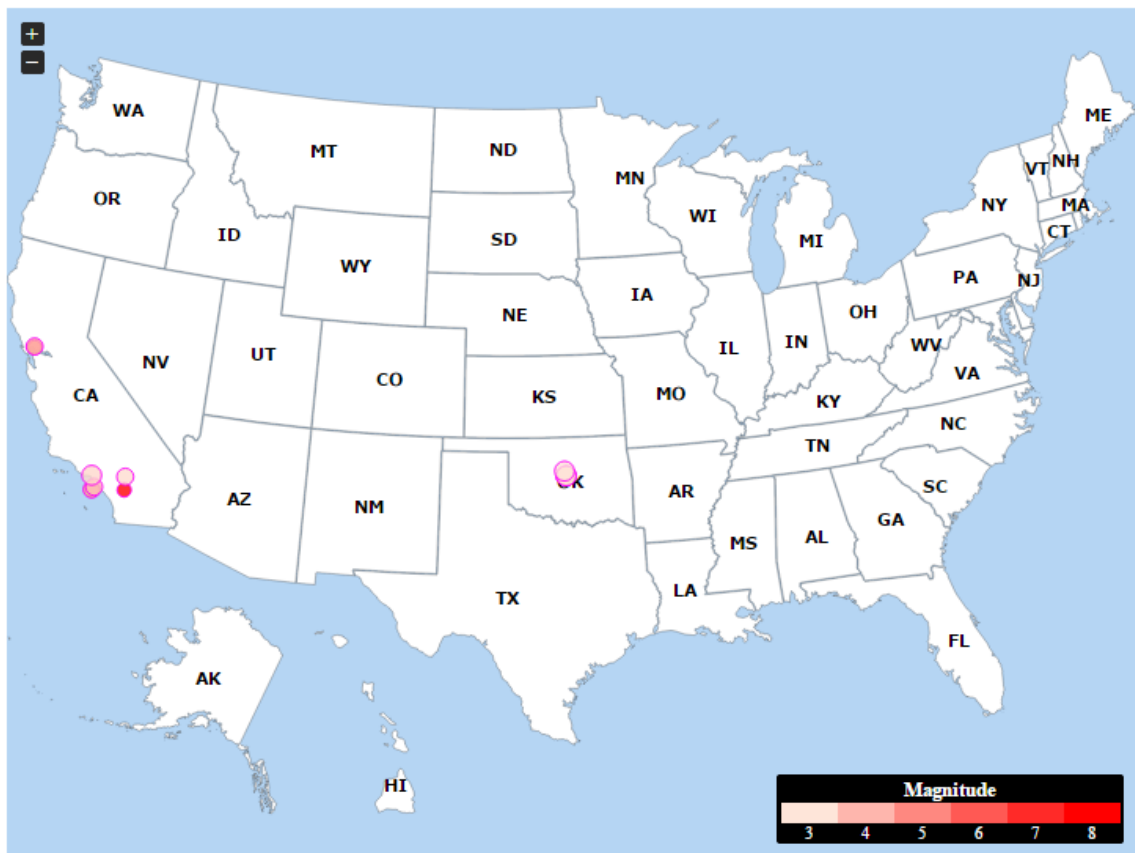


Figura 13. Terremoti rilevati dal sistema analizzando i tweet del periodo considerato

Oltre alle informazioni su tutti terremoti del mondo, USGS mette anche a disposizione una lista con i terremoti che ritiene più significativi dato un certo periodo. Andando a verificare il periodo preso in considerazione, USGS segnala 28 terremoti più significativi tra le migliaia che detiene. Di questi 28 terremoti 13 sono avvenuti in nazioni in cui si parla la lingua inglese, cioè 10 negli USA e 3 in Nuova Zelanda. Considerando gli Stati Uniti, dei 10 terremoti avvenuti negli Stati Uniti che USGS ritiene come i più importanti, 6 di questi rientrano tra quelli individuati dal sistema attraverso le informazioni condivise dai sensori sociali. Questo a sostegno della validità del sistema, che nella sua fase iniziale riesce a individuare pochi terremoti ma quelli più significativi.

6.1.3. Situazione demografica degli Stati Uniti

Per rappresentare la situazione demografica statunitense sono stati scaricati i dati dal sito ufficiale Census.gov³⁷. I dati raccolti sono le stime della popolazione del 2014 per gli stati, rilasciate il 23 dicembre 2014, e le stime della popolazione a livello di contea del 2013, in quanto per il 2014 non sono ancora disponibili. La scelta di usare anche i dati riferiti alle contee permette di avere un quadro più chiaro e dettagliato della popolazione, accentuando ancora di più eventuali discrepanze tra le zone maggiormente popolate, in cui presumibilmente c'è maggiore attività di networking, e quelle più isolate dove i sensori sociali sono scarsi.

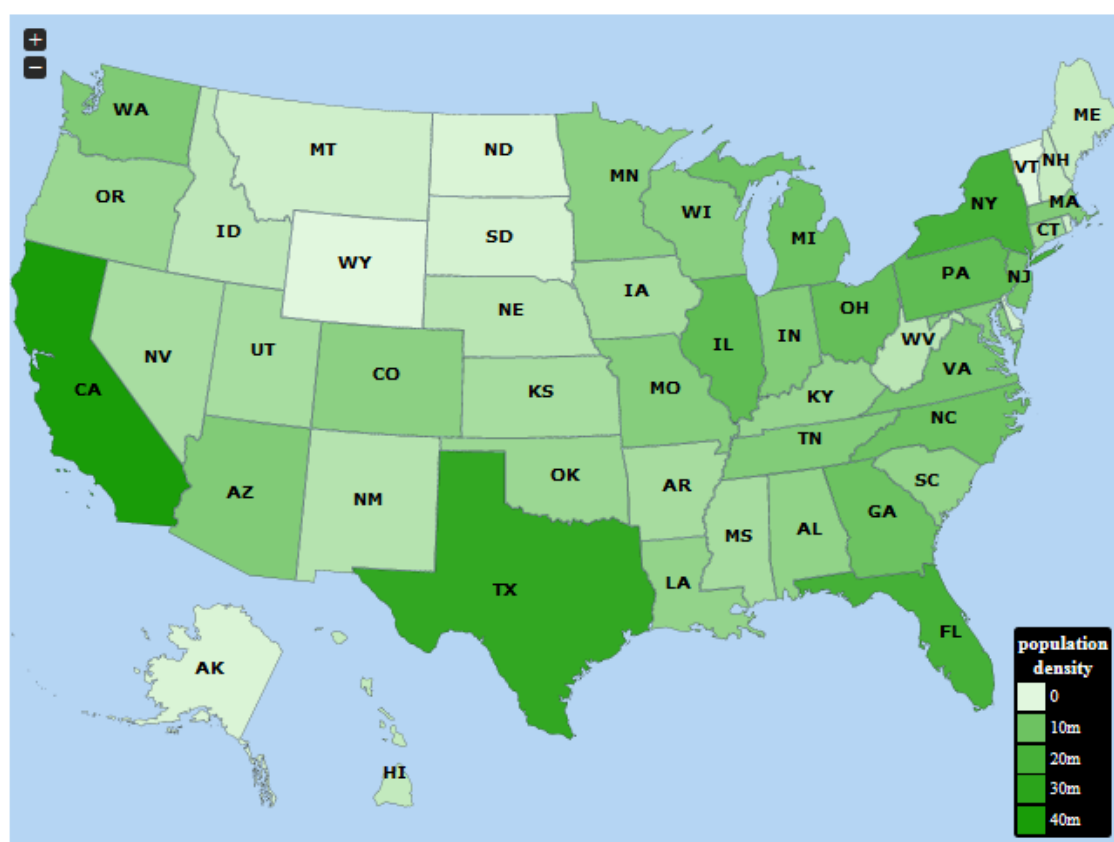


Figura 14. Popolazione stimata degli USA per l'anno 2014 (census.gov)

³⁷ www.census.gov

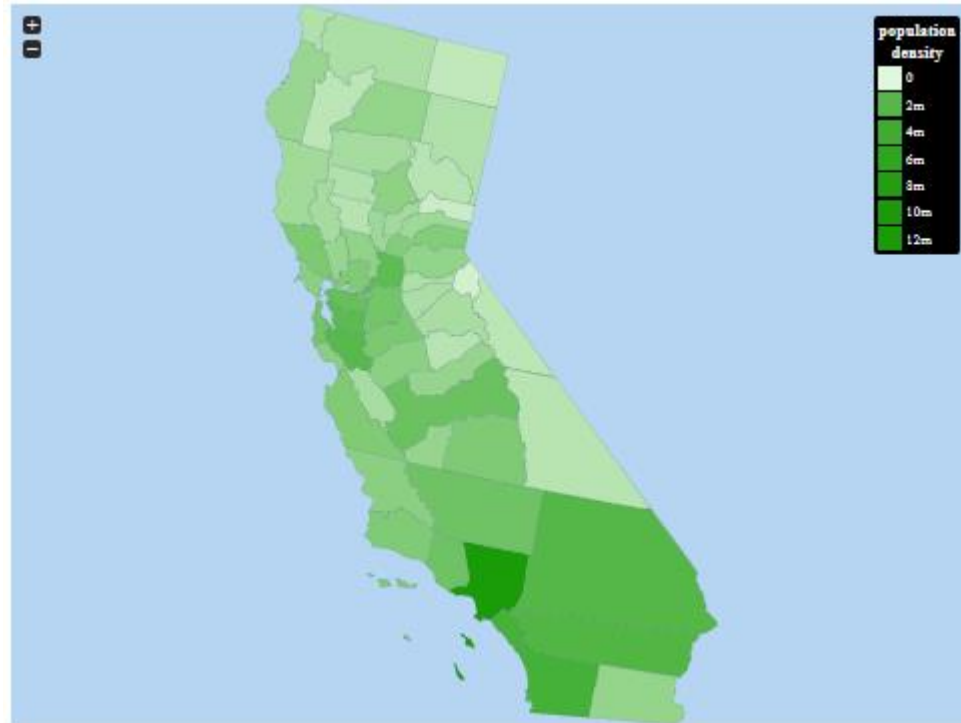


Figura 15. California. Popolazione stimata per contea anno 2013 (census.gov)

6.2. Creazione mappa

Per valutare la bontà dei risultati è stato pensato di mostrare su un'unica mappa i dataset raccolti. Sono state valutate diverse librerie per la realizzazione della mappa. In un primo momento è stato usato il servizi messi a disposizione da Google Maps³⁸ tramite API che consente la ricerca e la visualizzazione di mappe geografiche della Terra e la possibilità modificarle, creando marker personalizzati sulle mappe, geocodificare di indirizzi, ecc. Sebbene questo servizio sia tra i più consolidati per la creazione di certi tipi di applicazioni dinamiche, tuttavia il risultato era poco soddisfacente: infatti non permette di mostrare i terremoti e la popolazione nella stessa mappa.

La seconda scelta è stata la libreria Highcharts³⁹, che permette di creare mappe e grafici interattivi tramite API. Questa ha portato alla creazione di una iniziale mappa che mostrava la popolazione anche a livello delle contee, consentendo una visione dettagliata della situazione demografica. Tuttavia si è scoperto in seguito che Highcharts non supporta l'inserimento delle

³⁸ www.developers.google.com/maps

³⁹ www.highcharts.com

coordinate geolocalizzate, per cui non è stato possibile aggiungere i terremoti alla mappa demografica e la scelta è stata abbandonata.

Infine è stata trovata la libreria open source jVectorMap⁴⁰. Questa è stata la scelta finale in quanto ha permesso di inserire in un'unica mappa sia la situazione demografica che i terremoti geolocalizzati.

Per realizzare la mappa è stato adottato anche qui il linguaggio PHP 5.3.8, la libreria Javascript jQuery1.8.2 e la tecnica AJAX per la gestione dinamica del grafico.

Sono stati recuperati con una query i dati demografici salvati sul DB e mostrati. Allo stesso modo sono state estratte le coordinate latitudine e longitudine dei terremoti segnalati da USGS e mostrati sulla mappa sotto forma di cerchio. Più la magnitudo è alta, più il cerchio è di colore rosso; la grandezza del raggio del cerchio è inversamente proporzionale alla sua profondità: tanto più è piccolo il cerchio, tanto più è profondo il terremoto.

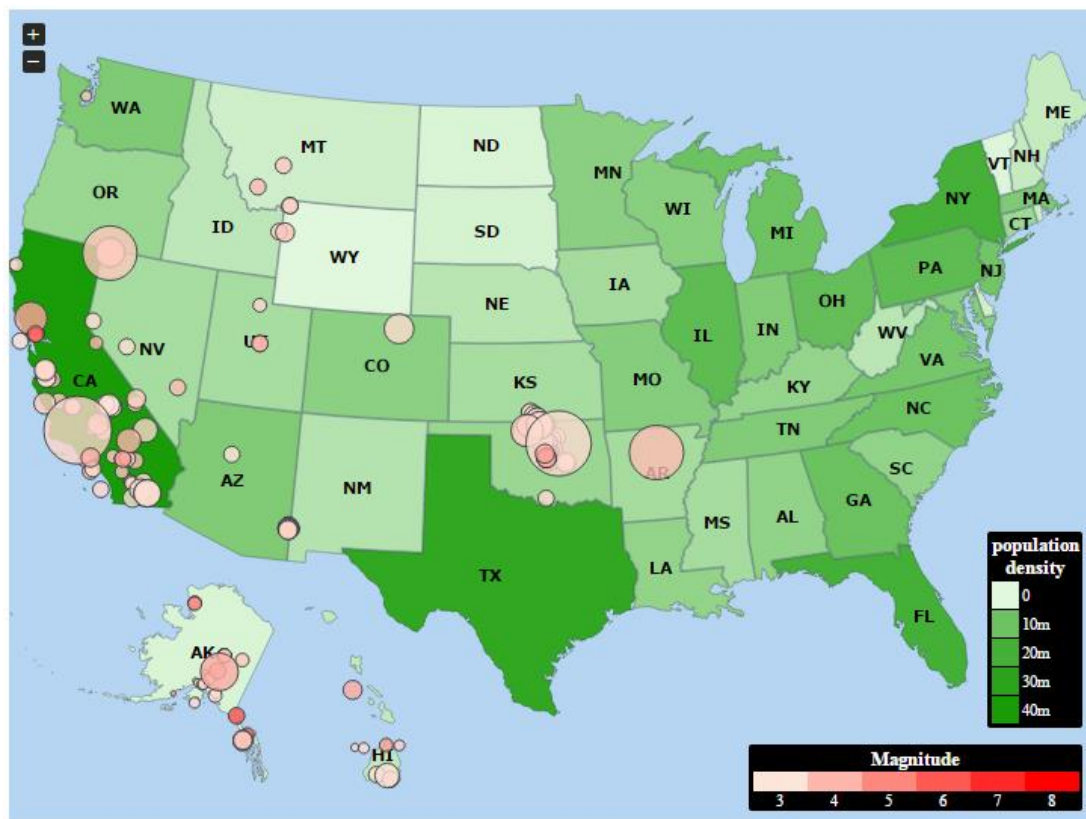


Figura 16. Terremoti forniti da USGS e popolazione sulla stessa mappa

⁴⁰ www.jvectormap.com

Infine sono stati cercati di inserire nella solita mappa i terremoti rilevati dal sistema:

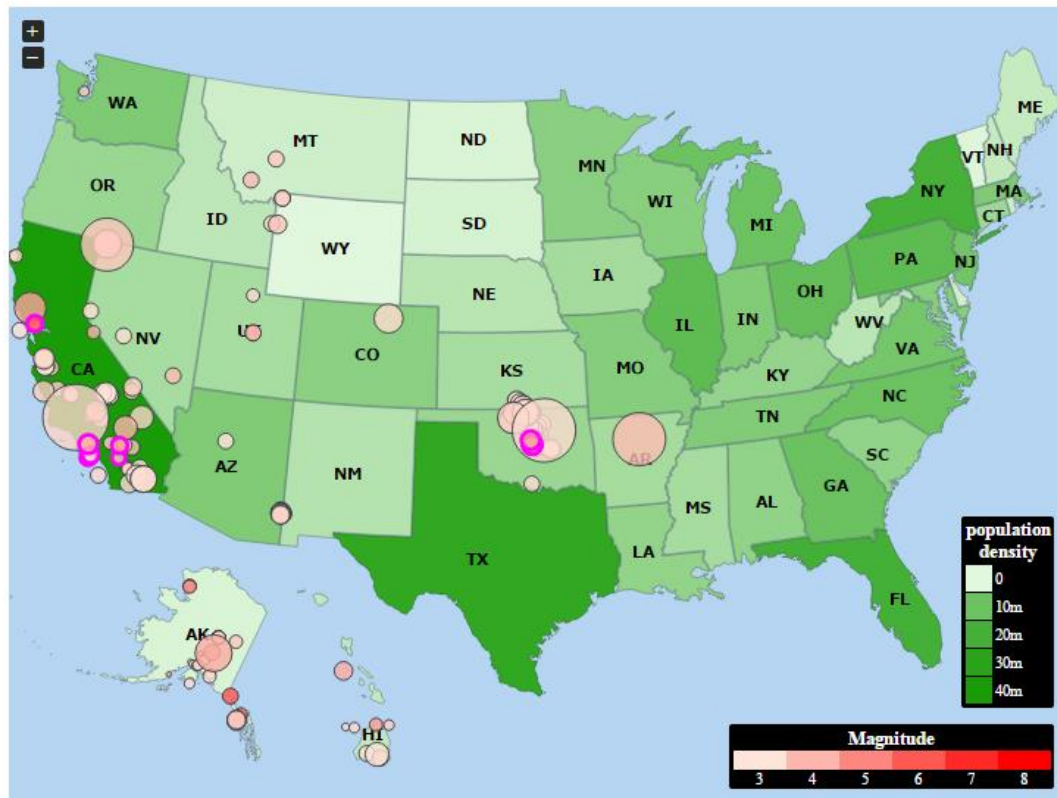


Figura 17. Terremoti rilevati dal sistema (bordo rosa), terremoti ufficiali e popolazione sulla stessa mappa

6.3. Considerazioni

Com'è possibile osservare dalla mappa la California è una zona soggetta a movimenti sismici. Il motivo per cui la maggior parte dei terremoti rilevati dal sistema si trova in questo stato è che è una zona molto più popolata rispetto alle altre come l'Alaska o le Hawaii. Nelle zone popolate i sensori sociali sono un numero molto alto e più persone hanno la possibilità di twittare un evento contemporaneamente e "attirare l'attenzione". Andando a osservare nel dettaglio lo stato della California è possibile notare come la maggior parte dei terremoti rilevati siano nelle contee più popolate dello stato, soprattutto la contea di Los Angeles. Il terremoto più a nord invece ha magnitudo 6.02, segno che una tale scossa desta l'attenzione di tanti sensori sociali, indipendentemente dal livello demografico del luogo.

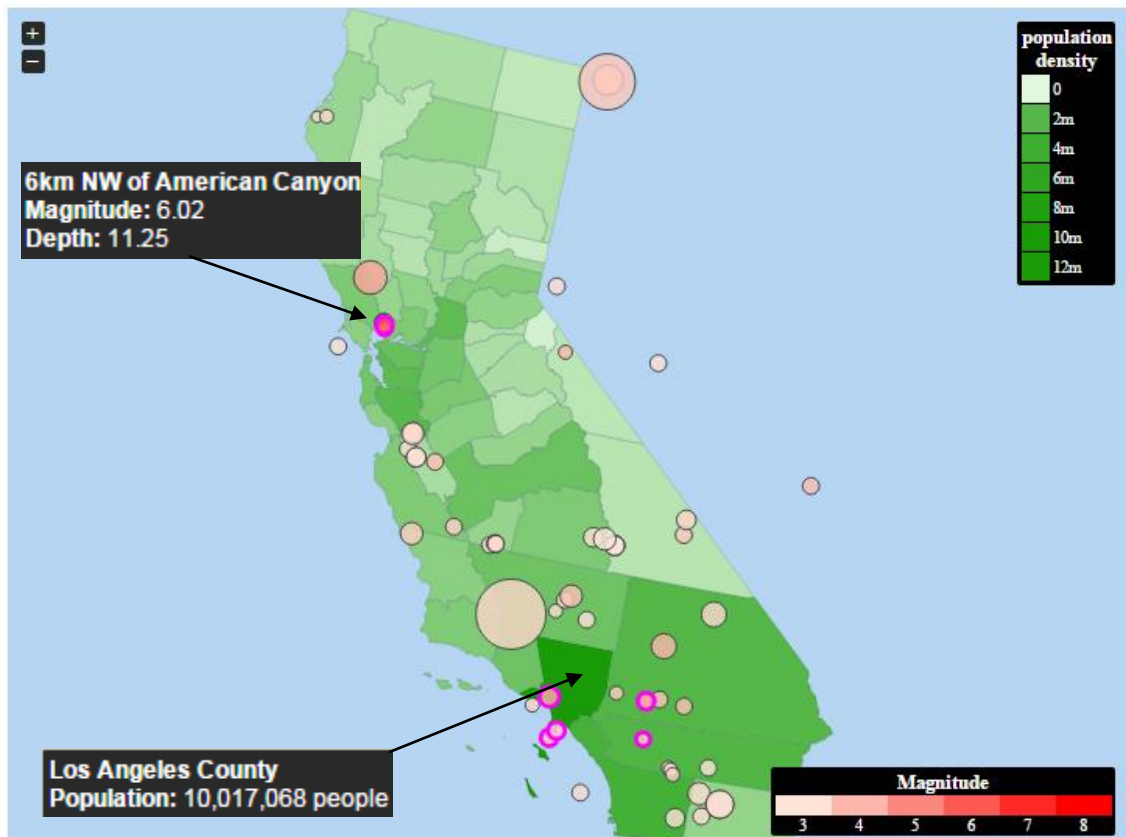


Figura 18. Terremoti rilevati in California

Gli altri terremoti rilevati tra quelli ufficiali hanno epicentro nello stato dell'Oklahoma. Nonostante la grande presenza di scossa il sistema ha rilevato dall'attività degli utenti Twitter solo due di queste. Tuttavia è possibile anche in questo caso osservare come le rilevazioni coincidano con le zone demograficamente avvantaggiate, a sostegno dell'importanza del ruolo che assumono le persone come "sensori sociali" in uno studio come questo. Inoltre il sistema ha rilevato i due terremoti con magnitudo maggiore rispetto agli altri, rispettivamente 4.3 e 4.4 magnitudo.

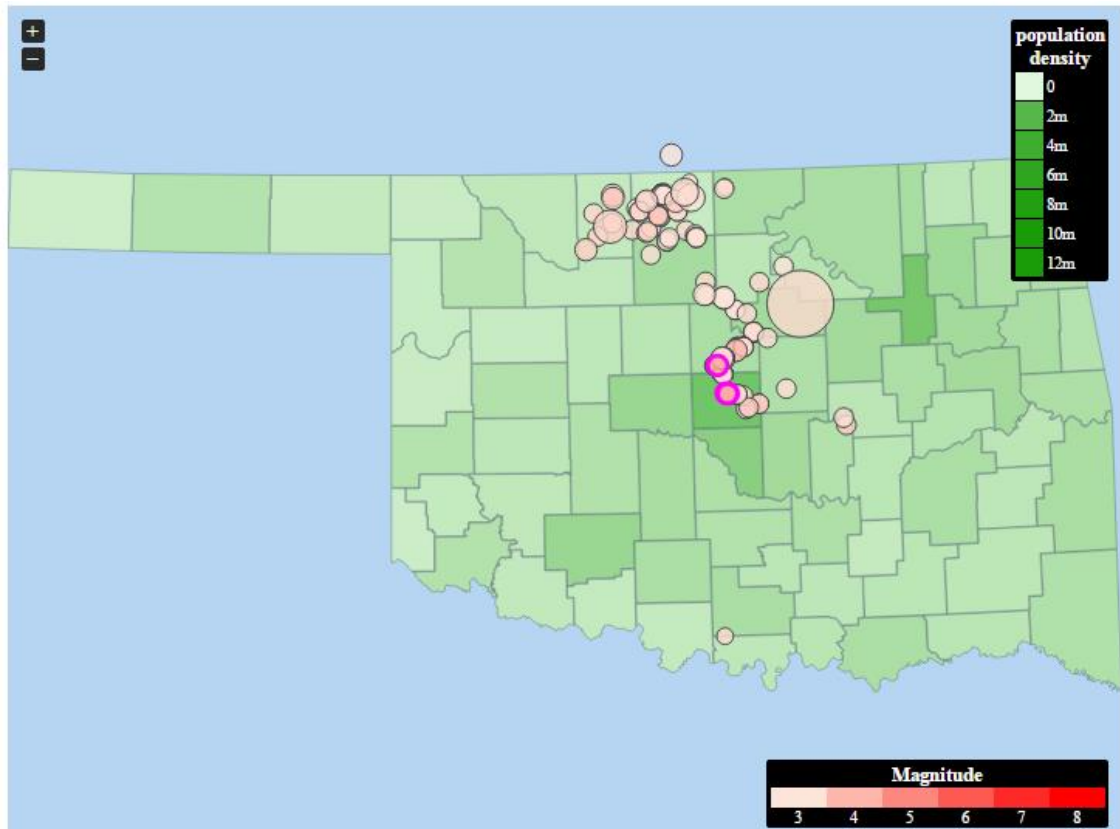


Figura 19. Terremoti rilevati in Oklahoma

Il tutto dimostra che il sistema è propenso a rilevare terremoti con magnitudo superiore a 4 e in zone particolarmente popolate. La profondità dei terremoti rilevati non supera i 15km, quindi è possibile dire che anche essa incide sulle prestazioni del sistema.

La bontà dei dati garantisce che i risultati ottenuti sono in linea con la realtà. Il sistema può ancora essere migliorato ma i risultati sono molto promettenti per il futuro, soprattutto in tempi di crisi dove se un certo tipo di comunicazione viene interrotta, le persone potranno comunque usare gli altri mezzi per rimanere in contatto tramite social media e ricevere aiuti.

7. Conclusioni e progetti futuri

La Rete ha mille potenzialità, tra queste la tendenza corrente è quella di utilizzare i dati provenienti dai Social Network a vantaggio della comunità, sfruttando la reazione di massa che si scatena sulle piattaforme sociali durante o subito dopo le catastrofi naturali.

Realizzare e perfezionare un sistema come quello proposto è possibile e può essere utile non solo al fine di individuare terremoti nelle zone in cui la connessione internet è avanzata e la presenza dei sismografi molto bassa, ma anche livello generale applicandolo ad altri ambiti settoriali, come l'individuazione di incendi, inondazioni, problemi di traffico stradale ed incidenti, ecc. E' importante tuttavia tenere presente che i sensori umani non sono distribuiti uniformemente sul territorio e la loro sensibilità risulta generalmente influenzata da elementi che non influiscono sul funzionamento dei sensori tradizionali (ora del giorno, ecc.). Eventuali progetti futuri devono quindi tenere in considerazione tutti questi fattori e potrebbero concentrarsi su contesti diversi dagli eventi sismici. A livello teorico si potrebbero individuare sui Social Network delle discussioni particolarmente calde relative a certi eventi (disordini allo stadio, ecc.) e, di conseguenza, prevenirli o gestirli.

Attualmente questo progetto è ancora in fase di avanzamento. Uno dei suoi sviluppi più interessanti in corso è quello di contattare in modalità automatica gli utenti coinvolti in un terremoto appena dopo la rilevazione da parte del sistema, chiedendo di condividere ulteriori informazioni.

Dal lato umanistico un tipo di progetto come questo si rivela utile per la società in quanto si propone di collaborare con le forze dell'ordine come sistema di supporto nelle situazioni di emergenza. Esso potrebbe aiutare a salvare vite umane, ad esempio spedendo i soccorsi nelle zone maggiormente colpite indicate nei contenuti postati sui social dalle persone coinvolte oppure avvertendo i luoghi in cui il disastro non è ancora.

Alla base di tale lavoro non può mancare uno studio del comportamento dell'essere umano nell'utilizzo di questi nuovi mezzi di comunicazione:

- Dal punto di vista *linguistico* è stato già sottolineato come durante momenti di panico i messaggi postati siano molto corti e serrati, con la presenza di tutte le caratteristiche già menzionate, tra cui parolacce e presenza di punti esclamativi. Un'altra interessante analisi nell'uso della lingua riguarda l'uso di determinati

termini per individuare discussioni relative a certi eventi: nell'italiano monitorare le parole "scossa" e "terremoto" bastano per raccogliere tutte le informazioni necessarie; mentre nella lingua inglese i rispettivi "shake" e "quake" funzionano in maniera leggermente diversa, in quanto "shake" fatica a disambiguare l'evento;

- Dal punto di vista *psicologico e sociologico* le persone anche in situazioni di emergenza rimangono connesse, aggiornando e controllando il proprio profilo. Prima di sfuggire al pericolo preferiscono postare due parole sui Social Network, molto spesso non per chiedere aiuto ma semplicemente per guadagnarsi quei famosi "15 minuti di notorietà" di cui diceva Andy Warhol. L'idea è quella di sfruttare questo nuovo disagio moderno che porta all'attaccamento perenne alla tecnologia e alla dipendenza da Social Network.



Figura 20. Cartello di avviso inventato per promuovere il progetto SoS. Social Sensing

Questo è un filone di ricerca in continua espansione alimentato dall'evoluzione delle piattaforme sociali. Sempre più persone sono coinvolte in tale progresso tecnologico ed è necessario studiare metodi interessanti al fine di sfruttare l'intelligenza collettiva del pubblico a vantaggio della comunità.

8. Bibliografia

1. Acar, A., & Muraki, Y. (2011). Twitter for crisis communication: lessons learned from Japan's tsunami disaster. *International Journal of Web Based Communities*, 7(3), 392-402.
2. Adedoyin-Olowe, M., Gaber, M. M., & Stahl, F. (2013). A Survey of Data Mining Techniques for social media Analysis. *arXiv preprint arXiv:1312.4617*.
3. Avvenuti, M., Cresci, S., La Polla, M. N., Marchetti, A., & Tesconi, M. (2014, March). Earthquake emergency management by social sensing. In *Pervasive Computing and Communications Workshops (PERCOM Workshops), 2014 IEEE International Conference on* (pp. 587-592). IEEE.
4. Avvenuti, M., Cresci, S., Marchetti, A., Meletti, C., & Tesconi, M. (2014, August). EARS (earthquake alert and report system): a real time decision support system for earthquake crisis management. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 1749-1758). ACM.
5. Becker, H., Naaman, M., & Gravano, L. (2011). Beyond Trending Topics: Real-World Event Identification on Twitter. *ICWSM, 11*, 438-441.
6. Burns, A., & Eltham, B., (2009). Twitter Free Iran: An Evaluation of Twitter's Role in Public Diplomacy and Information Operations in Iran's 2009 Election Crisis. *Paper presented at Communications Policy & Research Forum 2009*.
7. Chew, C., & Eysenbach, G. (2010). Pandemics in the Age of Twitter: Content Analysis of Tweets during the 2009 H1N1 Outbreak. *PloS one*, 5(11), e14118.
8. Cohen, J. (1968). Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. *Psychological bulletin*, 70(4), 213.
9. Crane, R., & Sornette, D. (2008). Robust dynamic classes revealed by measuring the response function of a social system. *Proceedings of the National Academy of Sciences*, 105(41), 15649-15653.
10. Doan, S., Vo, B. K. H., & Collier, N. (2012). An analysis of Twitter messages in the 2011 Tohoku Earthquake. In *Electronic Healthcare* (pp. 58-66). Springer Berlin Heidelberg.
11. Ebina, R., Nakamura, K., & Oyanagi, S. (2011, November). A real-time burst detection method. In *Tools with Artificial Intelligence (ICTAI), 2011 23rd IEEE International Conference on* (pp. 1040-1046). IEEE.
12. Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 37(1), 32-64.

13. Finin, T., Murnane, W., Karandikar, A., Keller, N., Martineau, J., & Dredze, M. Annotating Named Entities in Twitter Data with Crowdsourcing. In *Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk* (pp. 80-88).
14. Giroux, J., Roth, F., & Herzog, M. (2013). Using ICT & social media in Disasters: Opportunities & Risks for Government.
15. Hughes, A. L. (2012). Supporting the Social Media Needs of Emergency Public Information Officers with Human-Centered Design and Development (*Doctoral dissertation, UNIVERSITY OF COLORADO AT BOULDER*).
16. Imran, M., Castillo, C., Diaz, F., & Vieweg, S. (2014). Processing Social Media Messages in Mass Emergency: A Survey. *arXiv preprint arXiv:1407.7071*.
17. Keen, J. (2010). Facebook, Twitter 2-way 'lifeline' for Haiti. USA Today. Retrieved from <http://www.webcitation.org/5qbzC2Ur2>
18. Kongthon, A., Haruechaiyasak, C., Pailai, J., & Kongyoung, S. (2012, July). The role of Twitter during a natural disaster: Case study of 2011 Thai Flood. In *Technology Management for Emerging Technologies (PICMET), 2012 Proceedings of PICMET'12:* (pp. 2227-2232). IEEE.
19. Kouloumpis, E., Wilson, T., & Moore, J. (2011). Twitter sentiment analysis: The good the bad and the omg!. *ICWSM, 11*, 538-541.
20. Kumar, S. Ranking Assessment of Event Tweets for Credibility.
21. Lesperance, A. M., Godinez, M. A., & Olson, J. R. (2010). Social Networking for Emergency Management and Public Safety.
22. Lindsay, B. R. (2011). Social Media and Disasters: Current Uses, Future Options, and Policy Considerations.
23. MacEachren, A. M., Jaiswal, A., Robinson, A. C., Pezanowski, S., Savelyev, A., Mitra, P. & Blanford, J. (2011, October). Senseplace2: Geotwitter analytics support for situational awareness. In *Visual Analytics Science and Technology (VAST), 2011 IEEE Conference on* (pp. 181-190). IEEE.
24. Mathioudakis, M., & Koudas, N. (2010, June). Twittermonitor: trend detection over the twitter stream. In *Proceedings of the 2010 ACM SIGMOD International Conference on Management of data* (pp. 1155-1158). ACM.
25. Meier, P. (2011). New information technologies and their impact on the humanitarian sector. *International review of the Red Cross*, 93(884), 1239-1263.
26. Mitchell, J. T., Thomas, D. S., Hill, A. A., & Cutter, S. L. (2000). Catastrophe in reel life versus real life: perpetuating disaster myth through Hollywood films. *International Journal of Mass Emergencies and Disasters*, 18(3), 383-402.

27. Morgan, J. (2010). Twitter and Facebook users respond to Haiti crisis. BBC. Retrieved from <http://www.webcitation.org/5qc0IvXpw>

28. Nielsen. (2010). Social media and mobile texting a major source of info and aid for earthquake in Haiti. Nielsen Wire. Retrieved from <http://www.webcitation.org/5qbzJq9rR>

29. Palen, L., Vieweg, S., Liu, S., & Hughes, A. (October 2007). Crisis informatics: Studying crisis in a networked world. In *Proceedings of the 3rd International Conference on e-Social Science, Ann Arbor, Michigan, USA*.

30. Parr, B. (2010). Haiti earthquake: Twitter pictures sweep across the web. Mashable. Retrieved from <http://www.webcitation.org/5qbz66pMp>

31. Poibeau, T., & Kosseim, L. (2001). Proper name extraction from non-journalistic texts. *Language and computers*, 37(1), 144-157.

32. Qu, Y., Wu, P., & Wang, X. (January 2009). Online community response to major disaster: a study of Tianya forum in the 2008 Sichuan earthquake. In *Proceedings of the 42nd Hawaii International Conference on System Sciences, Waikoloa, Big Island, Hawaii, USA*.

33. Robinson, B., Power, R., & Cameron, M. (2013, May). A sensitive twitter earthquake detector. In *Proceedings of the 22nd international conference on World Wide Web companion* (pp. 999-1002). International World Wide Web Conferences Steering Committee.

34. Rogstadius, J., Kostakos, V., Laredo, J., & Vukovic, M. (2011). Towards real-time emergency response using crowd supported analysis of social media. In *Proceedings of CHI workshop on crowdsourcing and human computation, systems, studies and platforms*.

35. Rui Li, Kin Hou Lei, Ravi Khadiwala, and KC-C Chang. (2012). Tedas: A twitter-based event detection and analysis system. In *Proc. of ICDE. IEEE*, 1273–1276.

36. Russell R. Dynes. 1970. *Organized Behavior in Disaster*. Heath LexingtonBooks.

37. Saif, H., He, Y., & Alani, H. (2012). Alleviating Data Sparsity for Twitter Sentiment Analysis. *Making Sense of Microposts (#MSM2012)*.

38. Sakaki, T., Okazaki, M., & Matsuo, Y. (2010, April). Earthquake shakes Twitter users: real-time event detection by social sensors. In *Proceedings of the 19th international conference on World Wide Web* (pp. 851-860). ACM.

39. Scoble, R. (2008). Twittering the earthquake in China. Scobleizer. Retrieved from <http://scobleizer.com/2008/05/12/quake-in-china/>

40. Siegel, J. (2010). Wyclef Jean's Haiti relief Tweets raise \$1 million. ABC News. Retrieved from <http://www.webcitation.org/5qc0qd5ZW>

41. Sysomos. (2010). Haitian earthquake dominates Twitter. Sysomos Blog. Retrieved from <http://www.webcitation.org/5qbzQx8S4>
42. Tanev, H., Ehrmann, M., Piskorski, J., & Zavarella, V. (2012, May). Enhancing Event Descriptions through Twitter Mining. In *ICWSM*.
43. Taylor, M., Wells, G., Howell, G., & Raphael, B. (2012). The role of social media as psychological first aid as a support to community resilience building. *Australian Journal of Emergency Management, The*, 27(1), 20.
44. Times, N. Y. (2007). Virginia Tech shooting leaves 33 dead. New York Times. Retrieved from <http://www.webcitation.org/5q11aE6pK>
45. US Department of Homeland Security (2010). Privacy impact assessment for the Office of Operations Coordination and Planning Haiti social media disaster monitoring initiative. www.dhs.gov/xlibrary/assets/privacy/privacy_pia_ops_haiti.pdf
46. Vieweg, S., Palen, L., Liu, S., Hughes, A., & Sutton, J. (2008, May). Collective intelligence in disaster: Examination of the phenomenon in the aftermath of the 2007 Virginia Tech shooting. In *Proceedings of the 5th International ISCRAM Conference, Washington DC, USA*.
47. Yates, D., & Paquette, S. (2011). Emergency knowledge management and social media technologies: A case study of the 2010 Haitian earthquake. *International Journal of Information Management*, 31(1), 6-13.
48. Yin, J., Lampert, A., Cameron, M., Robinson, B., & Power, R. (2012). Using social media to Enhance Emergency Situation Awareness. *Intelligent Systems, IEEE*, (27), 52-59.

Ringraziamenti

Desidero ringraziare le persone che mi hanno sostenuto durante questo lungo percorso con i loro consigli o semplicemente con la loro presenza.

In particolare uno ringraziamento speciale va ai miei genitori e a mia sorella, che continuano a sostenermi ogni giorno e mi aiutano nei momenti difficili: senza di voi non sarei ciò che sono.

Un indispensabile grazie ai professori che hanno reso possibile questo lavoro, il mio relatore Maurizio Tesconi e i miei correlatori Stefano Cresci e Mariantonietta Noemi La Polla: grazie per la vostra disponibilità e la pazienza che avete dimostrato negli ultimi mesi, sia durante la stesura della tesi che durante il tirocinio al CNR, è stata un'esperienza difficile, ma fantastica!

Grazie anche alla mia amica e collega Jessica con cui ho condiviso questo percorso negli ultimi anni: grazie mille per il supporto, gli appunti, i consigli e soprattutto le risate!

Ringrazio parenti e amici che a modo loro hanno contribuito a rassicurarmi e sostenermi nell'ultimo periodo.

Infine grazie anche a tutte le persone con cui ho avuto il piacere di entrare in contatto al CNR.

Senza tutti voi non avrei potuto raggiungere questo traguardo. Grazie.