# Sommario

[Sommario 1](#_Toc32227094)

[Premessa 3](#_Toc32227095)

[Capitolo 1 4](#_Toc32227096)

[Introduzione 4](#_Toc32227097)

[Il progetto CANTICO 4](#_Toc32227098)

[Capitolo 2 10](#_Toc32227099)

[Introduzione 10](#_Toc32227100)

[NLP – Natural Language Processing 10](#_Toc32227101)

[Sentiment Analysis 13](#_Toc32227102)

[Emotion Mining 16](#_Toc32227103)

[Tools 17](#_Toc32227104)

[NLP – Natural Language Processing 17](#_Toc32227105)

[Sentiment Analysis 20](#_Toc32227106)

[Emotion mining 26](#_Toc32227107)

[Visual Emotion Mining 29](#_Toc32227108)

[Capitolo 3 31](#_Toc32227109)

[Introduzione 31](#_Toc32227110)

[Identificazione dei tool 31](#_Toc32227111)

[Ambiente di sviluppo 34](#_Toc32227112)

[Architettura dell’analizzatore 35](#_Toc32227113)

[Capitolo 4 39](#_Toc32227114)

[Introduzione 39](#_Toc32227115)

[Costruzione del dataset 39](#_Toc32227116)

[Definizione dell’oracolo 40](#_Toc32227117)

[Elaborazione dati sperimentali 44](#_Toc32227118)

[Analisi dei dati sperimentali 45](#_Toc32227119)

[Analisi della polaritá 45](#_Toc32227120)

[Analisi delle emozioni 47](#_Toc32227121)

[Conclusioni 49](#_Toc32227122)

[Indice delle figure e delle tabelle 51](#_Toc32227123)

[Bibliografia 52](#_Toc32227124)

# Premessa

L’elaborato ha come scopo la realizzazione di uno degli obbiettivi presentati nell’ambito di un progetto di ricerca nazionale volto a migliorare l’esperienza degli utenti nella fruizione degli spettacoli teatrali. In questo contesto si è cercato di fare un’analisi dello stato dell’arte alla ricerca di possibili tool utilizzabili ai fini del progetto.

All’interno del Capitolo 1 vi è una introduzione che consente di inquadrare l’elaborato all’interno del progetto globale. Individuato il contesto del progetto si cerca, sempre all’interno del Capitolo 1, di contestualizzare le tecniche di NLP (Natural Language Processing), necessarie alla realizzazione del progetto, sia dal punto di vista dell’evoluzione degli approcci teorici, sia dal punto di vista della realizzazione concreta. Si sono, infatti, ricercati diversi tool adatti allo scopo richiesto e se ne è fatta un’analisi in termini di utilizzabilità. Successivamente nel Capitolo 2 si mostra come sono stati effetti i test per verificare il livello di efficienza dei tool individuati all’intero del contesto d’interesse. Contestualmente vengono presentate le modalità di creazione del dataset utilizzato per testare i tools individuati.

## Capitolo 1

## Introduzione

In questo capitolo si vuole analizzare il contesto entro cui l’elaborato si sviluppa, descrivendo il progetto a cui si cerca di contribuire.

## Il progetto CANTICO

L’arte italiana, ed in particolare l’opera lirica e il teatro di prosa, non sono soltanto la punta di diamante del nostro Paese, ma una risorsa economica strategica, un patrimonio da valorizzare, con produzioni e un’offerta formativa d’eccellenza riconosciute a livello internazionale. Tuttavia, in questi ultimi anni la marginalità di questo settore del Made in Italy è in costante diminuzione. Infatti, i teatri d’opera, o meglio le Fondazioni lirico-sinfoniche che li gestiscono, attraversano da anni una profonda crisi economica che rischia di minarne alla radice non solo la qualità ma finanche la sostenibilità. Secondo la relazione della Corte dei Conti sulle Fondazioni Lirico-Sinfoniche per l’esercizio 2014 i ricavi derivanti da vendite e prestazioni, ossia i ricavi da mercato, rappresentino una frazione relativamente bassa dei valori di produzione, con una riduzione dell’8,5% nel 2014 rispetto all’anno precedente. La stessa relazione mette in evidenza un saldo negativo fra valore della produzione e relativi costi, dovuto in larga misura alla contrazione dei contributi in conto esercizio. Il rischio è che tale situazione di crisi porti ad una diminuzione della qualità delle produzioni, peraltro già in qualche modo annunciata dalla forte caduta dei costi di produzione registrata a partire dal 2008. Analogo discorso vale per il teatro di prosa. L’Italia ospita oltre 25 mila luoghi di spettacolo, che hanno prodotto, solo 2014, 131.000 rappresentazioni artistiche di cui il 63% sono spettacoli in prosa. Gli spettacoli

in prosa teatrale risultano infatti i più seguiti in Italia, con un incasso superiore a 200 milioni di euro, con una variazione del -3,9% nel triennio 2012-2014. Lo stato attuale evidenzia la necessità di migliorare le capacità di autofinanziamento dei teatri lirici e di prosa, sia attraverso l’incremento di pubblico sia attraverso l’arricchimento dei servizi che essi erogano. In tale ottica, si inserisce il progetto **CANTICO** che sta per *“PLATFORM TO ATTRA****C****T MORE* ***A****UDIE****N****CE AND SPREAD OPERA AND PROSE* ***T****HEATER BY USING* ***I****MMERSIVE TE****C****HNOL****O****GIES”* che intende sviluppare e sperimentare metodi e strumenti ICT per la creazione di servizi mirati a valorizzare il patrimonio delle produzioni liriche e dei teatri di prosa italiani. La piattaforma sviluppata nell’ambito del progetto CANTICO, seppure focalizzata sul teatro di prosa e l’opera lirica, mira ad implementare servizi general purpose, quali ad esempio ticketing, marketing e fidelizzazione, ed a offre le risorse necessarie alla realizzazione di ulteriori verticalizzazioni in altri settori di eccellenza e tradizione del made in Italy, come ad esempio gli spettacoli musicali e di danza.

Quattro sono le direzioni verso cui l’intero progetto intende operare:

1. Favorire la crescita della base di pubblico che frequenta i teatri d’opera e di prosa, in particolare mediante l’uso di nuove tecnologie per attrarre i più giovani;
2. Estendere la fruizione delle produzioni oltre il tradizionale canale della rappresentazione dal vivo, attraverso strumenti di fruizione remota innovativi, sia in tempo reale sia on-demand, e attraverso l’adozione di tecniche e metodologie per favorire l’accessibilità ai disabili;
3. Accompagnare la crescita di un pubblico appassionato e consapevole e formare gli artisti del domani, attraverso azioni di sensibilizzazione e formazione rivolte a studenti di ogni ordine e grado basate su tecnologie innovative, e attraverso l’adozione di soluzioni in grado di favorire una didattica inclusiva;
4. Migliorare la capacità di programmazione dei teatri lirici e di prosa attraverso strumenti innovativi di raccolta ed analisi dei dati.

L’idea alla base del progetto è di dare vita ad una piattaforma multicanale capace di comunicare in maniera chiara, efficace e coinvolgente. La piattaforma dovrà supportare i servizi tradizionali e dovrà essere aperta all’integrazione con servizi di terza-parte, in una logica di sussidiarietà e cooperazione finalizzati alla creazione di una esperienza user-centered. Si vuole inoltre sviluppare una piattaforma che dovrà consentire la creazione di una community in cui condividere informazioni e impressioni su un evento, fino a creare dei veri e propri “percorsi esperienziali”.

L’intero progetto è suddiviso in nove obbiettivi programmatici, che vanno da uno studio preliminare degli strumenti disponibili per effettuare una personalizzazione dell’esperienza da parte dell’utente, fino all’obbiettivo finale che corrisponde ad una sperimentazione della piattaforma. Il progetto viene sviluppato da diversi enti che contribuiscono in maniera diversa durante tutti le varie fasi descritte dai nove obbiettivi programmatici. Tali soggetti sono:

* Innovaway Spa, in qualità di soggetto capofila
* Universitá degli studi del Sannio, in qualitá di Organismo di ricerca
* Il Conservatorio “N.Sala”, in qualitá di ente pubblico
* Meditel, Officina Rambaldi, Step, Digitalcomoedia e E@l Software, in qualitá di PMI

In particolare questo elaborato si inserisce nell’ambito del primo punto programmatico. Questo ha l’obiettivo di studiare i metodi e gli strumenti che consentono di personalizzare l’esperienza d’uso da parte degli utentei, attraverso attività di ricerca applicata in diversi settori, fra cui tecniche avanzate di profiling, data mining, affective computing, analytics context aware, sentiment analysis e social network analysis.Per raggiungere questi risultati si fará riferimento alle conoscenze già disponibili in Innovaway e Unisannio. Saranno usate le competenze tecnico-scientifiche di Unisannio sulle tematiche di user profiling, data mining, Affective Computing e Sentiment Analysis, in parte sviluppate in progetti precedenti. Fra gli altri si ricorda MyOpenGov, dove si è utilizzata una profilazione avanzata e tecniche per l’analisi di social networking per raccomandare i servizi ed i mush-up più adatti ad uno specifico utente. Un secondo esempio è OpenGoal, che ha utilizzato tecniche di social engineering per favorire l’elicitazione dei bisogni dei cittadini, oppure TEMOTEC, che ha sviluppato approcci di crowdsourcing per la raccolta di informazioni sullo stato di conservazione dei beni culturali, e infine Collective Knowledge, che ha sviluppato modelli e tecnologie abilitanti la collective knowledge in reti produttive e sociali. Saranno poi utilizzate le conoscenze di Innovaway sulle tematiche di Crowdsourcing Analysis e Social Network Analysis acquisite dal team impiegato per il progetto SNECS, acronimo di *“Social Network delle Entità dei Centri Storici”.* Nell’ambito del progetto SNECS, infatti, sono state svolte attività di ricerca di base su tali problematiche tecnico-scientifiche la cui conoscenza verrà sfruttata da Innovaway per lo sviluppo di soluzioni pratiche e specifiche nell’ambito del progetto.

L’obiettivo è suddiviso in diverse attivitá. In particolare la prima ha l’obiettivo di studiare e analizzare metodi e strumenti per la personalizzazione dell’esperienza dell’utente. In particolare, sarà analizzato lo stato dell’arte e sarà condotta una ricerca applicata su tecniche avanzate di User Profiling, Affective Computing e Sentiment Analysis. L’applicazione di tali tecniche per l’estrazione di informazioni implicite, precedentemente sconosciute, da dati anche non strutturati, e la rilevazione di informazioni emotive e di interazione sociale per l’acquisizione dati sul comportamento e/o lo stato fisico dell’utente, consentiranno di identificare gli interessi di un utente, ovvero gli interessi che ha in comune con altri utenti, anche in un evento specifico, e potrà aiutare gli organizzatori dell’evento ad attrarre più visitatori ad eventi simili in futuro e a migliorare i servizi e le esperienze future.

La seconda attività, delegata a Innovaway, è incentrata sullo studio e l’analisi critica di metodi, strumenti e tecniche di condivisione e caratterizzazione della community e del pubblico interessato. In particolare, saranno studiate tecniche di crowdsourcing, attraverso le quali coinvolgere in modo partecipativo il pubblico dei teatri, in particolare i giovani e i giovanissimi. Saranno, inoltre, studiate tecniche di analytics context aware incentrate sul concetto che un sistema contestualmente consapevole anticipa le esigenze dell'utente e propone in modo proattivo il contenuto o il servizio più adatti e personalizzati. Infine, verranno studiate tecniche di Social Network Analysis per la definizione di flussi di informazione e per comprendere in che modo avvengono le interazioni tra ruoli e persone differenti all’interno di un medesimo contesto organizzativo, nella fattispecie quello del pubblico e degli appassionati di lirica e di prosa teatrale.

La terza attivitá, avrà l’obiettivo di definire e caratterizzare la componente di personalizzazione dell’esperienza utente, attraverso la sintesi delle attività di ricerca svolte nell’ambito della prima attività e la prototipizzazione in ambiente di laboratorio. La personalizzazione andrà oltre le tradizionali capacità di adattamento della interfaccia di fruizione dei servizi, spingendosi verso la realizzazione di sistemi di adattamento e raccomandazione dei contenuti. La componente dovrà essere in grado di profilare gli utenti rispetto allo specifico dominio del teatro d’opera e di prosa, sfruttando sia dati espliciti sia dati inferiti, formulare raccomandazioni, anche a partire dall’estrazione e dal mining dei dati relativi alle reti sociali in cui l’utente opera, valutare la reazione dell’utente ai suggerimenti ed alle personalizzazioni proposte, ed usare tali reazioni quale input per le successive attività di personalizzazione. La componente realizzerà strategie di mining delle associazioni tra profili di utenti e classi servizi utilizzati in passato in grado di prevedere l'utilizzo futuro dei servizi da parte di nuovi utenti (con profili simili) o di evolverne l'associazione contestualmente all'evoluzione dei profili stessi, al fine di meglio adattare l'erogazione dei servizi della piattaforma CANTICO all'utenza.

Infine l’ultima attività dell’obbiettivo, realizzata da Innovaway, definirà la componente di personalizzazione dell’esperienza utente, finalizzando le attività di ricerca svolte nell’ambito della seconda attività.

Il seguente elaborato rappresenta l’inizio della realizzazione dell’obiettivo appena descritto, cercando di dare il via al completamento della prima attività delegata all’Università degli studi del Sannio. Nell’elaborato si è avvita la fase di individuazione e studio di metodi e strumenti di personalizzazione e condivisione dell'esperienza utente. In particolare si vuole completare un’attività di ricerca analizzando le tecniche di sentiment analysis e di emotion mining, individuando approcci e tools atti a estrarre informazioni implicite, che consentiranno di identificare gli interessi di un utente. A partire da queste sarà poi possibile ricavare gli interessi che questi ha in comune con altri utenti, anche in un evento specifico, e ciò potrà aiutare gli organizzatori degli eventi, nel caso specifico del progetto CANTICO quelli legati al mondo dello spettacolo, ad attrarre più visitatori ad eventi simili in futuro e a migliorare i servizi e le esperienze future.

A seguito di questo capitolo introduttivo, il successivo sarà utilizzato per analizzare lo stato dell’arte e per presentare i principi alla base del Natural Language Processing in generale e come le tecniche si sono evolute nel tempo. Verranno poi presentate la Sentiment Analysis e l’Emotion Mining per comprendere come queste possano essere sfruttate all’interno del progetto considerato. Si presenteranno, infine, alcuni dei tool che potrebbero essere utilizzati per completare questo tipo d’analisi.

# Capitolo 2

Analisi automatica del testo

## Introduzione

In questo capitolo si presenterà uno studio sullo stato dell’arte su ció che riguarda il Natural Language Processing, la Sentiment Analysis e l’Emotion Mining. Si cercherà di dare una rapida panoramica su ció di cui si occupano teoricamente questi settori per poi procedere ad una analisi piú precisa e puntuale dei tool attualmente disponibili. Si andranno infine ad analizzare due strumenti che tentano di riconoscere i sentimenti da un punto di vista dell’analisi dell’espressione facciale, e che potrebbero essere utilizzabili in una fase diversa del progetto CANTICO.

## NLP – Natural Language Processing

“ Il Natural Language Processing è il processo di trattamento automatico mediante un [calcolatore elettronico](https://it.wikipedia.org/wiki/Calcolatore_elettronico) delle informazioni scritte o parlate in una [lingua naturale](https://it.wikipedia.org/wiki/Lingua_(linguistica)) ”

questa la definizione che da Wikipedia di NLP. Questa semplice definizione non è in grado di sintetizzare la complessità del campo in cui l’NLP opera. Si tratta di una ricerca interdisciplinare che abbraccia informatica, intelligenza artificiale e linguistica, il cui scopo è quello di sviluppare algoritmi in grado di analizzare, rappresentare e quindi “comprendere” il linguaggio naturale, scritto o parlato, in maniera similare o addirittura più performante rispetto agli esseri umani. Tale “comprensione” è determinata dal capire, e quindi essere poi in grado di usare, il linguaggio a varie granularità, dalle parole, in relazione al loro significato ed alla appropriatezza d’uso rispetto ad un contesto, fino alla grammatica ed alle regole di strutturazione sia delle frasi a partire dalle parole sia dei paragrafi e delle pagine a partire dalle frasi. Tra gli ambiti in cui le tecniche si sono maggiormente mostrate al pubblico troviamo i sistemi conversazionali intelligenti. Questi sistemi, tra cui vanno annoverati i più famosi:

* Alexa di Amazon
* Siri di Apple
* Google Assistant di Google

sono in grado non solo di imitare il dialogo umano, ma anche di rispondere a domande su argomenti di diversa natura (dalle ultime news ai film in TV) e di eseguire task complessi (come pianificare un viaggio).

Volendo capire nel dettaglio di cosa si occupa NLP possiamo dire che, in primo luogo, questa fornisce soluzioni per analizzare la struttura sintattica del testo, associando alle singole parole le rispettive categorie morfologiche (ad es. nome, verbo, aggettivo), identificando entità e classificandole in categorie predefinite (ad es. persona, data, luogo), estraendo dipendenze sintattiche (ad es. soggetti e complementi) e relazioni semantiche (ad es. iperonimia, meronimia). In secondo luogo, consente di comprendere la semantica del testo, identificando il significato delle parole, anche relazionato al contesto e alle modalità di utilizzo (ad es. ironia, sarcasmo, sentimento, umore), classificandolo in categorie predefinite (ad es. sport, geografia, medicina) o sintetizzandone il contenuto.

Volendo invece fare una piccola ricapitolazione dello sviluppo storico dell’NLP bisogna partire dagli anni settanta, durante i quali la maggior parte degli approcci di NLP hanno provato a formalizzare nei computer i vocabolari e le regole del linguaggio naturale, con risultati però non sempre soddisfacenti. La rigidità e non scalabilità delle regole scritte a mano non hanno, infatti, reso possibile una gestione estensiva dell’estrema variabilità ed ambiguità, funzionale, lessicale e sintattica, del linguaggio naturale. Agli inizi degli anni ottanta e, in maniera più estesa negli anni novanta, sono apparsi i primi approcci di NLP statistico o basato su corpora, ovvero una collezione di testi selezionati e organizzati per facilitare le analisi linguistiche. In particolare, essi hanno cercato di risolvere le problematiche degli approcci a regole, imparando autonomamente le preferenze lessicali e strutturali da corpora di grandi dimensioni, senza richiederne una codifica manuale preliminare. La maggior parte degli approcci di NLP statistico, ancora oggi largamente utilizzati in diversi prodotti commerciali, elabora il set di parole presenti in un testo (bag-of-words), a volte arricchito con le differenti categorie morfologiche, consentendo una migliore gestione dei differenti tipi di ambiguità del linguaggio e, al contempo, garantendo robustezza, capacità di generalizzazione e resilienza in presenza di errori.

Nonostante i risultati convincenti ottenuti in differenti applicazioni, ad esempio coi motori di ricerca e nell’estrazione della conoscenza, l’esigenza di migliorare ulteriormente le capacità di comprensione automatica di contenuti in linguaggio naturale, raggiungendo livelli similari a quelli dell’essere umano. Ed è per questo motivo che si è avviata la ricerca di nuove tecniche per migliorare l’efficacia di questi sistemi. Grandi progressi compiuti e importanti i risultati si sono ottenuti, soprattutto grazie allo sviluppo delle tecniche di intelligenza artificiale in particolare quelle basate sull’applicazione di algoritmi di deep learning.

Intorno al 2011, per la prima volta un semplice algoritmo basato sul deep learning è stato applicato a differenti problemi di NLP, tra cui l’identificazione di entità e l’assegnazione di categorie morfologiche a parole, mostrando prestazioni sensibilmente migliori rispetto ad altri approcci rappresentativi dello stato dell’arte.

Questa primo utilizzo degli algoritmi di intelligenza artificiale applicato ai problemi affrontati dall’NLP rappresenta la svolta nell’ambito della ricerca. Da allora, sono stati realizzati algoritmi sempre più complessi basati sul deep learning per affrontare problemi di NLP ancora non risolti o trattati in passato ma con risultati non soddisfacenti.

Il deep learning, in estrema sintesi, si basa sul concetto di rete neurale artificiale, ossia un modello matematico ispirato, dal punto di vista funzionale, ai sistemi neurali biologici del cervello umano. Una prima caratteristica fondamentale di tali reti è che sono in grado di apprendere, in maniera autonoma e contestuale, sia una rappresentazione di tipo gerarchico delle migliori caratteristiche descrittive dei dati di ingresso (non necessariamente intellegibili dall’essere umano) sia le modalità con le quali combinare al meglio tali informazioni per la risoluzione di un compito specifico. Una seconda caratteristica di rilievo è che tali reti, in maniera similare al cervello umano, sono in grado di imparare dalle loro esperienze, ossia di migliorare le proprie prestazioni nella risoluzione di un problema complesso in funzione della quantità di esempi con cui sono addestrati.

Tali sistemi, si vedrà, vengono utilizzati nell’ambito dello sviluppo degli algoritmi di NLP, ed in particolare sono state prese in considerazione per quanto riguarda la Sentimenti Analysis e l’Emotion Mining.

## Sentiment Analysis

“L'analisi del sentiment o sentiment analysis (nota anche come opinion mining) è un campo dell'elaborazione del linguaggio naturale che si occupa di costruire sistemi per l'identificazione ed estrazione di opinioni dal testo.”

Questa è la definizione che Wkipedia [1] da della Sentiment Analysis.Si basa sui principali metodi di linguistica computazionale e di analisi testuale. L'analisi del sentiment è utilizzata in molteplici settori: dalla politica ai mercati azionari, dal marketing alla comunicazione, dall'ambito sportivo a quello delle scienze mediche e naturali, dall'analisi dei social media alla valutazione delle preferenze del consumatore. La sentiment analysis viene utlizzata per determinare il l’emozione che si nasconde dietro le parole e per comprende attitudini, opinioni ed emozioni espresse all’interno di un testo. L’obbiettivo finale è comprendere se un testo risulta esprimere un sentimento positivo, negativo o neutro. Tale processo per quanto banale agli umani risulta estremamente complesso per una macchina. Risulta, però, uno strumento estremamente potente per ricavare informazioni analizzabili a partire da una enorme quantità di dati grezzi.

Questo tipo di tecnologia trova il suo naturale sviluppo nell’ambito dei social network dove è possibile ottenere in pochissimo tempo tantissimi dati grezzi da processare e che possono essere utilizzati per ricavare importanti informazioni circa l’opinione pubblica. La capacità di comprendere in maniera rapida e veritiera l’attitudine dei consumatori risulta ormai uno strumento imprescindibile di qualsiasi società. Ciò le consente di avere un feedback “diretto” e immediato rispetto alle proprie iniziative e anche rispetto a quelle dei competitors. Uno degli esempi forse più emblematici dell’importanza di tale approccio risulta essere che già nel 2012 il presidente Obama utilizzava strumenti di Sentiment Analysis per analizzare le opinioni degli americani in vista delle elezioni.

Nel corso del tempo la sentiment analysis ha assunto sempre maggiore importanza e per questo ha visto una notevole evoluzione nel tempo. Sostanzialmente è possibile menzionare le seguenti tecniche ad oggi:

* Rule-based: sono sistemi che ricavano la polarizzazione del testo attraverso un set di regole create manualmente.
* Automatic: sono i sistemi che sfruttano il machine learning per ricavare autonomamente le regole a partire dai dati
* Hybrid: sono i sistemi che combinano i due approcci.

La sentiment analysis basata su regole risulta molto complessa da gestire perché deriva direttamente dagli approcci NLP maggiormente interessati alla semplice analisi del testo e per questo spesso poco efficaci nell’analisi di come le parole sono messe in sequenza. In questo caso gli algoritmi esplodono molto velocemente verso una complessità molto elevata che li rende molto complessi da manutenere e ancor di più da evolvere nel tempo.

Gli approcci automatici si basano sulla capacità delle intelligenze artificiali moderne di apprendere autonomamente. Questo sgrava l’uomo dal compito di ricavare e poi immettere manualmente delle regole che gli algoritmi devono seguire. Queste tecniche basate su ció che definiamo machine learning, nell’ambito della sentiment analysis, possono essere assimilate ad un problema di classificazione. Un testo fornito in esame deve essere classificato cioè come positivo, negativo o neutrale. Per fare ciò gli algoritmi devono attraversare due fasi:

* Training, ovvero l’addestramento
* Prediction, ovvero il processo di analisi del testo

Durante la fase di addestramento si procede ad “insegnare” al modello ad associare un particolare input (ovvero un particolare testo) al corrispondente output (una delle possibili polarizzazioni positivo, negativo o neutrale). Lo si fa utilizzando un tipo di input precedentemente taggato. Si procede allora a trasferire il contenuto informativo del testo in un feature vector che identifica il testo. A questo punto il modello è in grado di associare una serie di feature vectors ai possibili tag.

La fase di predizione invece parte da un input generico che viene a sua volta trasformato in un feature vector. Il modello a questo punto analizza il feature vector del testo appena creato e lo associa con una certa percentuale a quelli con i quali è stato addestrato, riuscendo quindi a generare un tag da associare a quel testo generico.

Come si evince da questa generica presentazione del processo contenuto all’interno degli strumenti degli step che segue un algoritmo di sentiment analysis è l’estrazione del feature vector. Tale meccanismo altro non è che una semplice trasformazione di un testo in una rappresentazione numerica. Tale processo può seguire diversi meccanismi e tra i più classici si ricordano la bag-of-words oppure la bag-of-ngrams. Accanto a queste più recentemente si è utilizzata la tecnica del word vectors che consente di avere un rappresentazione simile per parole che hanno un significato simile, consentendo un miglioramento dei classificatori.

La parte di classificazione interviene solo a questo punto e sfrutta tipicamente dei modelli statistici quali ad esempio Naïve Bayes, Logistic Regression, Support Vector Machines, or Neural Networks (deep learning). Questa fase come si vede arriva alla fine e non opera sul testo ma su una rappresentazione numerica del testo cosa che li rende molto performanti.

Mostrando il processo su cui si basano gli algoritmi di sentiment analysis si è voluto rendere evidente come il processo di addestramento risulti il tassello fondamentale per la buona riuscita della comprensione del testo. Con ciò si intende che, un modello addestrato su un determinato tipo di input, riesce a classificare con un certo livello di precisione input dello stesso tipo, mentre perde di efficacia a mano a mano che ci si allontana dal contesto con il quale è stato addestrato. In questa ottica la cosa più immediata da pensare è quella di riaddestrare un modello con il nuovo tipo di input. Tale approccio è stato seguito da un team di ricercatori rivelando risultati inaspettati e mostrando come risulti poco efficace riaddestrare un modello già funzionante su di un nuovo contesto, sia dal punto di vista delle performance che da quello della precisione. Secondo Lin e al [2] il riaddestramento dei modelli su un nuovo contesto per effettuare la sentiment analysis è necessario ma non rappresenta uno strumento per garantire un grande miglioramento dell’accuratezza durante la predizione.

## Emotion Mining

Una sorta d’evoluzione della sentiment analysis è rappresentata dall’emotion mining. Questa visone semplice non rappresenta in realtà nel dettaglio ciò che è in realtà l’emotion mining ma è un buon punto di partenza per comprendere come l’emotion mining operi e perché viene oggi sempre più utilizzata.

È stata definita evoluzione perché il processo utilizzato per creare algoritmi di emotion mining è lo stesso di quello presentato nel paragrafo precedente. Si utilizzano le stesse tecniche basate sul machine learning e sulla costruzione del modello, la differenza importante tra i due approcci è costituita dai tipi di tag. Nel caso della sentiment analysis tipicamente si intende classificare un testo come positivo, negativo o neutrale, nell’ambito dell’emotion mining si fa riferimento ad un numero maggiore di tag che descrivano lo stato emozionale di una persona. I tipici tag sono quelli della gioia, della paura, della rabbia e così via. In questo caso ci troviamo di fronte ad un numero molto maggiore di possibili emozioni e soprattutto ad una definizione non universalmente chiara di quali sono quelle da rappresentare. Come si vedrà successivamente i diversi tool hanno utilizzato classificazioni diverse sebbene si vadano a ricercare spesso lo stesso tipo di emozioni all’interno del testo. Questa maggiore variabilità nella possibilità di classificare le emozioni, rende gli strumenti di emotion mining più sensibili e difficili da sviluppare. La fase di classificazione infatti diventa molto più complessa e molto spesso non univoca, un testo infatti può esprimere contemporaneamente paura e rabbia magari con percentuali diverse. Per questo motivo molto spesso si tende a vedere la distribuzione dei sentimenti piuttosto che andare ad identificarne uno solo, rendendo quindi l’analisi più veritiera e utile.

Sulla base di questa descrizione si evince il concetto di evoluzione considerato in senso lato. Risulta evidente, infatti, il differente scopo che fa evolvere la sentiment analysis e l’emotion minning. Il primo risulta infatti estremamente immediato nella sua analisi, e per questo ampiamente utilizzato nei campi presentati precedentemente. L’emotion mining richiede una analisi più fine e anche dei dati elaborati, per riuscire a capire la risposta del consumatore, ad esempio, rispetto ad una particolare campagna pubblicitaria. Una programmazione accurata infatti potrebbe cercare di stimolare alcune determinate emozioni nelle persone, rendendo lo questo strumento molto più efficace e versatile rispetto alla semplice definizione di polarità che offre la sentiment analysis. Questa necessaria spiegazione vuole sottolineare come, quindi, i due strumenti vadano ad estrarre informazioni diverse da un testo, e per questo motivo possono essere utilizzate insieme per operare in maniera complementare e riuscire ad ottenere maggiori dati per classificare l’opera oggetto dell’indagine. In questo caso, facendo riferimento al progetto presentato durante la fase introduttiva, è possibile ricavare dagli utenti una migliore descrizione di cosa potrebbe suscitare in loro l’intera esperienza teatrale al fine non solo di migliorare il tipo di spettacolo presentato ma anche l’infrastruttura che lo supporta.

## Tools

Si vogliono ora presentare una serie di esempi di tool che sono stati individuati nella letteratura che potrebbero essere utilizzati ai fini della ricerca.

### NLP – Natural Language Processing

Alcuni dei tool individuati potrebbero essere una base per costruire degli strumenti di sentiment analysis e di emotion mining. Tra gli altri si sono individuati tool come la Stanford CoreNLP che presa singolarmente si limita ad effettuare un’analisi base del testo. La libreria fornita funziona come un jar esterno e viene fornita un’api molto precisa e di facile interpretazione. Ne segue che il suo utilizzo e la sua installazione risultano molto semplici e immediati.

Questa libreria è stata sviluppata per funzionare con la lingua inglese e per questo motivo potrebbe non essere indicata per gli scopi del progetto ma, si è individuato un team di sviluppo, composto da Palmeto e Moretti [3], dell’istituto Bruno Kessler Center in Information and Communication Technology, che sta portando avanti il progetto TINT (The Italian NLP Tool), sempre basata sulla Stanford CoreNLP ma in grado di operare sulla lingua italiana. Questo team ha provveduto modificare la Sfanford CoreNLP per adattarla all’italiano, una delle lingue non supportate nativamente dalla Stanford CoreNLP.

Tabella 1 Tools per l'analisi testuale

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Tool | Lingue supportate | Italiano | Documentazione | Tecnica | Utilizzato |
| Stanford CoreNLP | Inglese, arabo, cinese, francese, tedesco, spagnolo | No | Disponibile una Api completa | Rete neurale | [2], [4] |
| TINT | Italiano | Si | Sono disponibili diversi tutorial | Stanford CoreNLP | [3], [5] |

La tabella riporta un semplice confronto tra le due librerie analizzate. Le funzionalità di questi strumenti sono potenti e risultano perfetti per completare egregiamente delle analisi del testo. Sia la versione inglese che quella adattata all’italiano consentono di effettuare delle precise analisi del testo. [3]

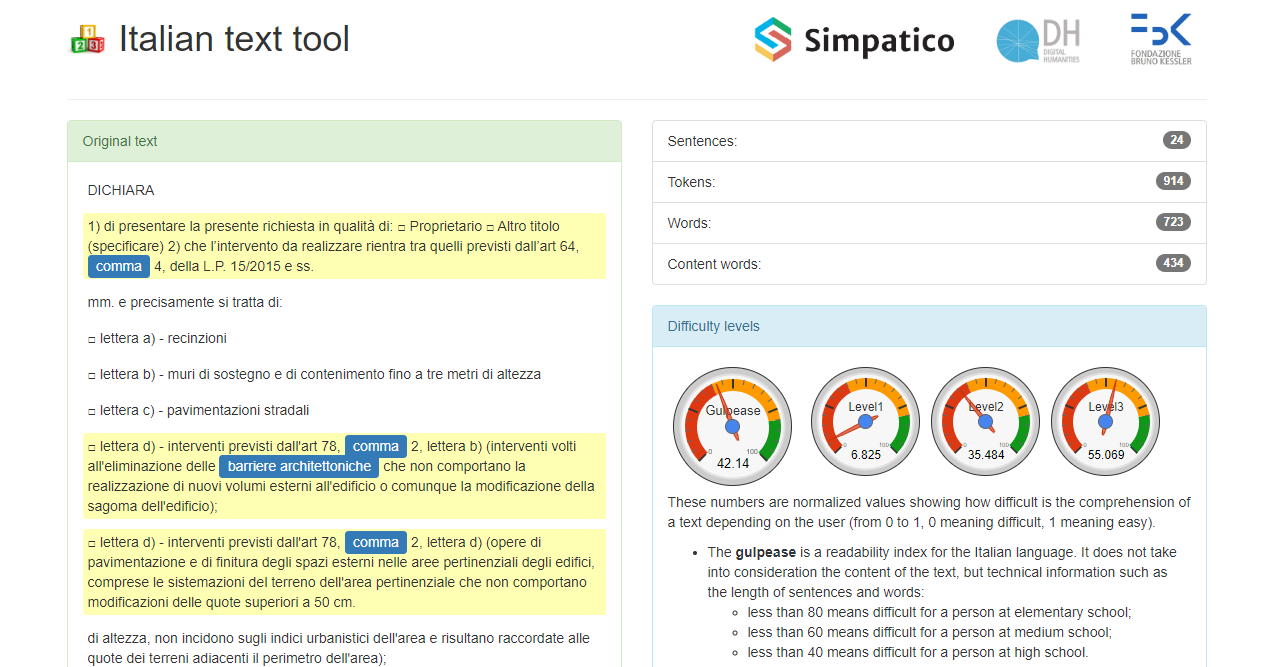


Figura Esempio analisi TINT [5]

In Figura 1 [5] si mostra un esempio di analisi fornita da TINT e si evince la precisione con cui viene effettuata l’analisi. Si possono ricavare importanti informazioni circa la composizione del testo, il numero delle frasi, il numero delle parole e anche il numero di parole significative. Viene anche fornita un’analisi del tipo di linguaggio utilizzato e della difficoltà di comprensione. Per maggiori informazioni si rimanda a Aprosio e Moretti [3] che hanno sviluppato il tool per degli scopi specifici che non sono sovrapponibili a quelli di questo elaborato.

Questo esempio chiarisce perché si è allargata la ricerca ad altri tool di analisi del testo. I due presentati, infatti, posseggono tutte le caratteristiche di efficienza e semplicitá d’uso per fungere da base si sviluppo di quelli che sono i tool necessari a perseguire gli scopi del progetto. Per come si è presentato nella premessa, infatti, non si necessita di tool atti alla semplice analisi testuale, quanto di tool che riescano ad estrarre dal testo informazioni relative all’emozioni espresse dal testo. Per questo motivo tool di NLP generici, per quanto efficienti, sono stati scartati rapidamente e si è provveduto a ricercare soluzioni piú specifiche all’ambito di ricerca.

### Sentiment Analysis

Ci sono numerose soluzioni commerciali che integrano numeroso funzionalitá aggiuntive a quella della semplice sentiment analysis. Diverse soluzioni commerciali sono necessariamente rivolte principalmente ad un mercato business e tentano di offrire un servizio piú completo come ad esempio WatsonNaturalLanguageUnderstanding di IBM in grado di fornire uno strumento per effettuare analisi del testo e sentiment analysis in diverse lingue a partire dai dati estratti da social network, sondaggi e blog pubblici. Dal punto di vista aziendale queste soluzioni rappresentano il metodo piú veloce per ottener un’analisi rapida e costante della loro percezione presso il pubblico. Questo perché vengono presentate dashboard con delle analisi dei dati giá fatte che possono essere giá utilizzate dall’azienda cosí come sono, tale approccio elimina il compito della ricerca da parte dell’azienda cliente, le informazioni vengono automaticamente reperite on line, tramite i social network ad esempio, e direttamente analizzate. Accanto a questo tipo d’approccio tali societá ed altre come MeanigCloud e GetSentiment utilizzano una metodologia diversa, esponendo un’api che puó essere interrogata con qualsiasi tipo di testo. E tipicamente questi servizi forniscono analisi molto approfondite. La Figura 2, estratta dalla demo di Menaningcloud [6], consente di vedere il livello di dettaglio della sentiment analysis che non si limita ad effettuare analisi solo sulle singole frasi ma estrae dei concettie e poi analizza questi. Tale livello di dettaglio puó essere raggiunto anche in altri aspetti dell’analisi, ad esempio per quanto riguarda l’estrazione di concetti dal testo e la correlazione di questi ultimi.

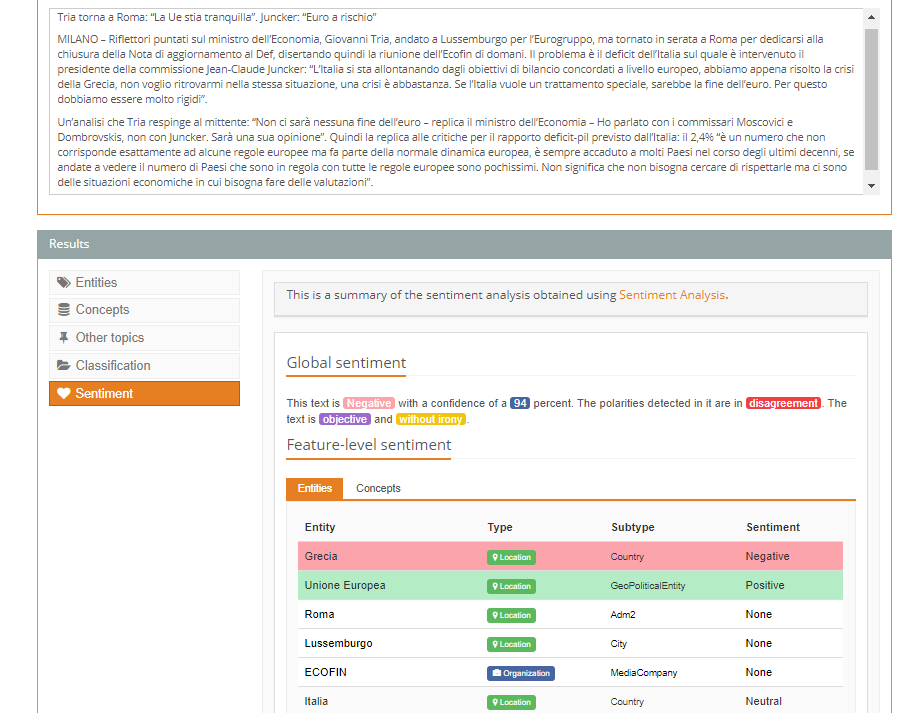


Figura Esempio analisi Meaningcloud [6]

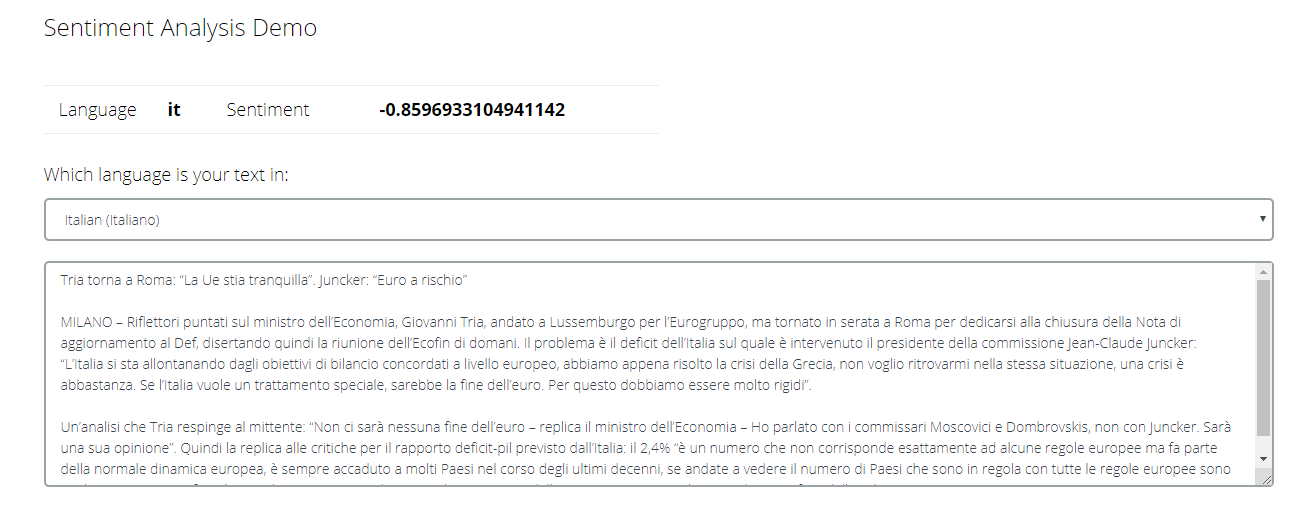


Figura esempio analisi Repustate [7]

Lo stesso tipo di servizio è offerto ad esempio da Repustate. In Figura 3 si vede come il servizio messo a disposizione da Repustate [7] permette di ottenere la risposta e sostanzialmente consente di ricavare il tipo di sentiment espresso nel testo testo, funzionando con diverse lingue.

Oltre a queste sono state individuate una serie di alternative open source da utilizzare. Tutti sono in grado di effettuare analisi sulla lingua inglese e, se non espressamente riportato, ci si riferirá a questa nei successivi paragrafi.

Uno dei tool piú famosi è sicuramente SentiStrength. Questo strumento è stato inizialmente addestrato con dei commenti di MySpace, e si basa su una sentiment word strenght list, ossia una collezione di termini classificati come positivi o negativi e associati ad un dato valore che ne indica la forza. SentiStrength assegna un valore ad ogni parola analizzata e ottiene il valore della frase effettuando una somma. Questo approccio semplice lo rende facilmente adattabile a nuovi contesti, semplicemente arricchendo e aggiungendo liste di parole nuove al suo core. SentiStrength è disponibile in una duplice versione: una per Windows ridotta e con funzionalitá leggermente ridotte, una seconda che opera su java che puó funzionare su qualsiasi sistema operativo. Accanto a queste due principali, sono stati sviluppati una serie di pacchetti per operare su altri linguaggi e piattaforme: Python, Ruby, Weka e Gate. Sono disponibili in beta versione di SentiStrenght in grado di operare su altre lingue, ma a seguito di un rapido test attraverso il demo offerto online si nota rapidamente come non venga fornito un servizio particolarmente accurato.

Una seconda alternativa è rappresentata da NLTK, anch’esso basato su regole e che è esplicitamente rivolto a testi estratti dai social media. Questo tool scritto in Python viene fornito come libreria. Offre servizi di analisi del testo e anche di sentiment Analisys. È possibile lavorare su qualsiasi sistema operativo, e viene fornita una API consultabile on line per capire velocemente come funziona. L’installazione risulta semplice ed intuitiva e anche il funzionamento non è complesso. I limiti imposti da questa libreria sono legati al fatto che supporta esclusivamente la lingua inglese.

Uno strumento presentato prima come tool generico, in realtá, presenta al suo interno una parte dedicata alla sentiment analysis. La Stanford CoreNLP presenta al suo interno un SentimentAnnotator, questa in particolare a differenza delle prime due soluzioni presentate sfrutta un approccio basato su Recursive Neural Network ed è quindi capace di analizzare la polarizzazione di una frase a partire da come le parole sono composte per definire il significato della frase, e non “semplicemente” ricavando il risultato da una somma. Questa è stata addestrata sulle recensioni di film. Come anticipato precedentemente la Stanford CoreNLP puó essere utilizzata come una libreria esterna, essendo fornita sottoforma di Jar integrabile in un progetto java. L’api correlata di esempi è disponibile online accanto ad uno java doc che consente di operare facilmente sulla libreria. Nuovamente il limite di questa libreria sono le lingue supportate. Come sempre il tool è stato sviluppato per operare sull’inglese poi nel tempo sono state supportate arabo, cinese, francese, spagnolo e tedesco, ma purtroppo non l’italiano.

Altro tool preso in esame è Emotion Mining Toolkit (EMTk), ossia un tool sviluppato dal Politecnico di Bari e che è composto da due parti: un analizzatore di polaritá, pySenti4SD e uno strumento di emotion mining, EmoTxt. Entrambi gli strumenti si basano sulla combinazione di un approccio n-gram, un’analisi lessicale e i risultati ottenuti da SentiStrength al fine di dare un risultato migliore. Questo tool viene fornito in diverse forme. In particolare è possibile configurare la propria macchina in modo da lavorare con il tool completo, oppure è possibile ottenere un container docker che ha giá tutto il necessario per avviare il tool. La parte di analisi di polaritá, pySenti4SD, lavora con python ed è semplice da utilizzare nella versione docker velocissima da configurare e avviare, l’api fornita invece risulta non completa essendo un progetto ancora in fase di sviluppo. Il problema principale di questo tool resta comunque il funzionamento esclusivo con la lingua inglese. La seconda componente, EmoTxt verrá analizzato nel paragrafo successivo.

Accanto a questi progetti piú grandi in rete è possibile reperire numerosi progetti sviluppati da singoli utenti. Ad esempio il Sentiment Analysis on Tweets [8], è un tool che consente di confrontare diversi approcci alla sentiment analysis. L’analisi in questo caso è rivolta specificamente a dei tweets ed il progetto confronta l’efficacia di vari metodi di classificazione (Naive Bayes, Foresta randomica, Massima Entropia ed altre), verificando come i metodi basati sulle reti neurali ottengono i risultati migliori nella classificazione. Questo tool si basa su python ma necessita anche di java per funzionare, è semplice da utilizzare e tutte le possibilitá d’uso sono presenti nella documentazione fornita. È possibile riaddestrare il tool oppure utilizzare il modello presente per analizzare il testo. La lingua supportata attualmente è solo quella inglese. Risulta uno strumento molto utile per confrontare i vari metodi di classificazione.

Un altro esempio sempre disponibile on line [9] utilizza NLTK come core e usa una Word Sense Disambiguation e le statistiche sull’occorrenza delle parole. I metodi utilizzati per classificare sono Naive Bayes e la Massima Entropia ed anche in questo caso si sono analizzati i tweet. Risulta estremamente semplice da installare, il setup provvede a recuperare tutte le dipendenze, e anche da utilizzare. Il progetto presenta una serie di esempi chiari sulle principali funzionalitá ma non è presente una documentazione completa attualmente. Allo stesso tempo non risulta piú disponibile un demo online, e inoltre la sola lingua supportata è quella inglese.

L’ultimo tool preso in considerazione, e che poi è stato scelto per effettuare i test, è Sentita [10]. Questo strumento sviluppato in Python si basa sul deep learning in particolare su una LSTM-CNN bidirezionale (Long Short-Term Memory Network) che opera a livello delle parole. Il modello riceve in input una parola che contiene la rappresentazione delle singole parole, ed emette due segnali compresi tra 0 e 1, uno per il rilevamento del sentimento positivo ed uno per il rilevamento del sentimento negativo. I due segnali potrebbero essere triggerati dalla stessa frase in input se questa contiene contemporaneamente sentimenti positivi e negativi. Questo tool, per quanto simile ad altri nelle modalitá con cui opera, è stato scelto perché in grado di operare sulla lingua italiana, cosa fondamentale per il tipo di attivitá che si intende perseguire nell’ambito del progetto. Il tool si basa su python e fa ampio utilizzo di librerie ampiamente note come Tensorflow. L’installazione risulta essere semplice e il setup provvede ad ottenere tutte le necessarie dipendenze.

Nella Table 2 sono elencati i tool open source presi in considerazione. Il tool scelto, ovvero SentIta, oltre ad essere semplice da utilizzare e da installare, ha l’enorme vantaggio di riuscire a funzionare direttamente con l’italiano che, considerato il progetto a cui si fa riferimento, è un vantaggio enorme. E sulla base dei test condotti riesce anche a predire in maniera molto efficiente la polaritá del testo sottoposto ad analisi.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Tool | Lingue | Italiano | Tecnica | Utilizzato |
| SentiStrength | Molte, tra cui: Inglese, tedesco, spagnolo, italiano, russo, turco, arabo, cinese, | Si | Basata su regole | [2] |
| NLTK | Inglese | No | Basata su regole | [2] |
| Stanford CoreNLP (SentimentAnnotator) | Inglese, arabo, cinese, francese, tedesco, spagnolo | No | Rete neurale | [2], [4] |
| EMTk (pySenti4SD) | Inglese | No | Caratteristiche lessicali | [2] |
| Sentiment Analysis on Tweets | Inglese | No | Varie Reti neurali | [8] |
| Sentiment Classifier | Inglese | No | NLTK | [9] |
| SentIta | Italiano | Si | Rete neurale CNN | [10] |

Tabella 2 Tools per la sentiment analysis

### Emotion mining

Anche nell’ambito dei tool di emotion mining sono stati individuati alcuni esempi di strumenti forniti dalle aziende a pagamento ed in grado di offrire un’esperienza completa all’utente. Alcuni esempi possono essere Twinword che offre una vastitá di strumenti, compresa una parte di sentiment analysis, e che è in grado di individuare la percentuale di emozioni (gioia, tristezza, sorpresa, disgusto, rabbia e paura) presenti in un testo. La demo online [11] consente di lavorare solo con la lingua inglese ma fornisce una serie di demo con cui è possibile valutare la loro accuratezza. Nell’esempio mostrato nella Figura 4 si vede come il tool sia in grado di mostrare in valore percentuale le emozioni mostrate nel testo fornito. L’api proposta dal tool viene offerta sulla base dei testi inviati analizzati al mese, in base al numero disponibile aumenta il prezzo dell’abbonameto.



Figura esempio analisi twinword [11]



Figura esempio paralleldots [12]

Un secondo esempio di tool a pagamento analizzato è quello sviluppato da ParallelDots, visibile in Figura 5 [12]. Questa soluzione fornisce anche strumenti di sentiment analysis accanto a quella di emotion mining ed è inoltre in grado di operare con numerose lingue, tra cui l’italiano. Anche in questo caso vengono mostrate le percentuali di emozioni (gioia, rabbia, eccitazione, tristezza, paura, noia) presenti all’interno di un testo, e anche in questo caso viene fornita una demo [12] dove è possibile testare il loro servizio con un testo libero. L’abbonamento offerto, in questo caso, non solo comprende un numero variabile di richieste di analisi, ma anche un sovrapprezzo per il supporto ad altre lingue oltre l’Inglese.

In entrambe le soluzioni è fornita una documentazione molto dettagliata che permette di integrare i servizi nelle proprie applicazioni in maniera abbastanza semplice.

Accanto a queste soluzioni a pagamento si sono ricercate, come prima, alcune soluzioni open source che peró sono risultate molto piú difficili da individuare. Uno dei progetti analizzati è quello dell’EmoTxt presentato precedentemente, nell’ambito del tool EMTk. La componente che si occupa dell’analisi delle emozioni (in questo caso

gioia, rabbia, tristezza, amore, sorpresa e paura) è in grado di individuare per ogni testo solo l’emozione desiderata, ossia va ricercata la singola emozione in ogni testo dato come input, quindi vanno fatte interrogazioni successive per ricercare tutte le emozioni. In particolare fornisce in output un file che indica se il testo in input contiene l’emozione con cui si è avviata la classificazione. Il tool analizzato purtroppo, ha dimostrato numerosi problemi durante la fase d’installazione per mancanza di dipendenze. Alcune delle librerie su cui si basa, infatti, non sono piú disponibili generando numerosi problemi in fase d’installazione. Questo si ripercuote purtroppo sui risultati ottenuti, che in questo caso risultano essere negativi. Questo modulo dell’EMTk, infatti, risulta non funzionante. Quando viene avviato il processo di analisi il risultato è un’operazione inconcludente che genera dei file intermedi ma non quelli previsti dagli sviluppatori. La mancanza di documentazione dettagliata delle varie fasi, e un progetto intrinsecamente complesso, concorrono ad aumentare la difficoltá nell’individuare il problema e quindi a porvi rimedio. La stessa situazione si ripete all’interno del container docker fornito dagli sviluppatori che non richiede alcuna installazione e configurazione della macchina ospite. I problemi incontrati fino ad ora, e la necessitá di effettuare numerose interrogazioni per ricavare un’analisi completa di un testo considerato hanno spinto alla ricerca di un tool alternativo per effettuare le indagini nell’ambito dell’emotion mining.

Il secondo tool open source preso in considerazione, e ancora in fase di sviluppo, è EmoNet [13]. Il tool sviluppato dal gruppo di ricerca NLP della University of British

Columbia è uno strumento molto avanzato che funziona attraverso un modello costruito su una GRNN (Gated Recurrent Neural Networks). Il modello è in grado di estrarre numerose emozioni da un testo, seguendo la ruota delle emozioni di Plutchik. Gli otto settori voglio indicare che esistono otto dimensioni primarie delle emozioni che come si vede sono opposte nella ruota (gioia – tristezza). Il tool è in grado di ricavare la percentuale delle otto emozioni primarie (gioia, fiducia, paura, sorpresa, tristezza, disgusto, rabbia, anticipazione). Queste sono rappresentate dagli autori [13] negli otto settori della ruota delle emozioni di Plutchik mostrata in Figura 5. Il tool lavora con python e sfrutta alcune famose e facili da reperire librerie python, garantendo la massima compatibilitá e facilitá d’uso.



Figura Ruota delle emozioni [13]

A differenza di quanto fatto per il tool di sentiment analysis, in questo caso, non è stato possibile individuare un tool open source che supportasse la lingua italiana. Per questo motivo si è cercata una soluzione alternativa che sará presentata nel capitolo successivo che riguarda la spiegazione della fase di test. In questo caso si è scelto di procedere con il tool EmoNet, in grado di fornire per ogni interrogazione un quadro emozionale completo del testo considerato.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Tool** | **Lingue supportate** | **Documentazione** | **Tecnica** | **Utilizzato** |
| EmoTxt (EMTk) | Inglese | Scarsa | Caratteristiche lessicali | [2], |
| EmoNet | Inglese | Scarsa | Rete neurale ricorrente | [13] |

Tabella 3 Tool per l'emotion mining

In tabella vengono confrontate le caratteristiche principali dei due tool open source presi in considerazione per quanto concerne l’emotion mining.

## Visual Emotion Mining

L’analisi delle emozioni testuale potrebbe essere affiancata da alcuni strumenti in grado di fare un’analisi delle emozioni a partire da alcune immagini. Necessariamente questo tipo di analisi dovrebbe essere eseguita in maniera diversa, ad esempio all’uscita del teatro per riuscire ad analizzare la reazione degli utenti al termine dello spettacolo. Le analisi visuali, in questo caso, potrebbero comunque risentire dell’influenza di fattori esterni che potrebbero inficiare la veridicità dei dati ottenuti. In questo senso, quindi, potrebbe essere piú utile effettuare un’analisi delle reazioni direttamente durante la proiezione per cercare di capire quali sono le reazioni a caldo degli spettatori. In ognuna di queste ipotesi le sfide tecnologiche da affrontare sarebbero notevoli, anche solo per riuscire ad accumulare i dati da analizzare, cosa molto piú semplice da fare nel caso dell’analisi testuale. Per questi motivi si è riservata una parte marginale a questo tipo di approccio analitico.

Si sono, comunque, individuati due progetti open source rappresentativi di questo tipo di analisi:

* Il primo sviluppato da Omar Ayman [14], è in grado di riconoscere le emozioni (in questo caso rabbia, disgusto, paura, gioia, tristezza, sorpresa, neutra) trasmesse dalle espressioni facciali. Anche in questo caso vengono mostrate in percentuale le emozioni trasmesse da ogni espressione. L’autore dichiara per il suo progetto un’accuratezza del 66% che rappresenta un buon traguardo.
* Il secondo è in corso di sviluppo da Juan Carlos Garfias Tovar [15], si concentra sull’individuazione delle emozioni espresse dal un essere umano. Si basa su una cascata di classificatori per identificare l’espressione facciale, e anche eventuali altri gesti che potrebbero avvalorare un’emozione ad esempio il movimento di una mano.

Questo tipo di approccio risolverebbe uno dei problemi principali nell’analisi testuale, quello della lingua. Come evidenziato nei paragrafi precedenti infatti la lingua rappresenta un grosso ostacolo allo sviluppo degli strumenti di questo tipo, che ha favorito alcuni dei tool presi in esame rispetto ad altri. Nel caso dell’analisi visuale questi problemi verrebbero superati facendo riferimento ad un tipo di comunicazione non verbale che è di natura universale.

# Capitolo 3

## Introduzione

In questo capitolo si vuole presentare l’analizzatore costruito. Si partirá dalla descrizione del processo decisionale rispetto ai tool, identificando i motivi principali che hanno fatto scegliere alcuni dei tool precedentemente descritti. Seguiranno poi una breve analisi dell’ambiente utilizzato e di quali sono i prerequisiti necessari all’esecuzione dei tool e dell’analizzatore. Viene infine presentata l’architettura realizzata e il processo che consente di estrarre le informazioni da un semplice commento in lingua italiana.

## Identificazione dei tool

A partire dall’analisi dello stato dell’arte effettuate nel capitolo precedente, si sono individuati i tool piú adeguati ad effettuare le analisi. Per effettuare la scelta si è tenuto in considerazione il particolare contesto in cui ci si sarebbe trovati ad operare. Si ricorda, infatti, che il progetto di riferimento in cui si attesta questo elaborato è CANTICO. Questi rappresenta un progetto interno del ministero dell’istruzione italiano, e mira esplicitamente a migliorare l’esperienza del pubblico presso i teatri italiani, cosí come descritto nel capitolo 1. Per questi motivi si è data particolare attenzione alle lingue supportate dai vari tool individuati. Il progetto, infatti, sará sviluppato con il fine di operare con la lingua italiane e per queste motivazioni, si è posta particolare attenzione a quali sono le lingue supportate dai vari tool. Dall’analisi effettuata nel capitolo precedente, si evince facilmente come la maggior parte dei tool a pagamento valutati fornisca supporti in numerose lingue, da cui si evince come non sia impossibile costruire e progettare un tool in una data lingua. Allo stesso tempo si nota, peró, come per la totalitá dei tool open source individuati sia la lingua inglese quella supportata e quella con cui il tool si è tipicamente sviluppato. Questa scelta degli sviluppatori deriva, ovviamente, dall’ampia diffusione di questo idioma in tutto il mondo e dalla facilitá con cui ormai è possibile tradurre in maniera precisa espressioni da qualsiasi lingua all’inglese. Per questa ragione è stata avviata una ricerca approfondita per individuare dei tool che supportassero nativamente l’italiano. Nell’ambito della sentiment analysis si è riuscito ad individuare SentIta, un tool che sebbene sia ancora in una fase di sviluppo iniziale, è stato ideato con il preciso scopo di operare sull’italiano, proprio a causa della mancanza di uno strumento simile anche secondo gli sviluppatori [10]. SentIta rappresenta il solo tool open source disponibile, attualmente, in grado di effettuare una sentiment analysis direttamente sulla lingua italiana. Per questa ragione si è deciso di selezionare proprio quest’ultimo come tool per effettuare le analisi, nonostante la limitazione presente in questa versione ancora in fase di sviluppo legata alla scarsa efficacia del tool per testi superiori alle 35 parole.

Nell’ambito, invece, dell’emotion mining non è stato possibile individuare un tool open source che supportasse l’italiano. Per questo motivo si è cercato di approcciare a questo problema in maniera diversa. Durate le fasi iniziali dello sviluppo di questo elaborato, si è riflettuto, infatti, sull’ipotesi di modificare un tool basato sulle reti neurali e addestrato sulla lingua inglese facendo sì che questo potesse operare sull’italiano. Tale ipotesi rendeva necessario un processo di riaddestramento che avrebbe dovuto permettere al tool di riconoscere una lingua totalmente nuova e, a partire da questa, completare il processo di sentiment analysis. Per quanto affascinate e apparentemente fattibile, questo modus operandi si rivela secondo Lin e altri [2] ampiamente deficitario in termini di performance e accuratezza, sebbene il loro lavoro avesse degli obbiettivi diversi. La loro analisi, infatti, prevedeva un riaddestramento atto a modificare esclusivamente il contesto operativo dei tool presi in considerazione, mantenendo la lingua con cui questi erano stati sviluppati. Con ció non si vuole affermare che tale procedimento non sia una reale ipotesi, perfettamente perseguibile. Il progetto TINT, infatti, rappresenta l’esempio perfetto di come sia possibile derivare da un tool preesistente, in questo caso Stanford CoreNLP, un nuovo strumento perfettamente funzionante con una nuova lingua. In quel caso come, descritto da Palmero e Moretti [3], si è proceduto a riaddestrare una parte della Stanford CoreNLP, costruendo una serie di moduli a partire da questi che siano in grado di analizzare correttamente del testo in italiano. I vantaggi di strumenti in grado di operare direttamente sull’italiano sono sicuramente notevoli, e infatti, quasi la totalitá dei servizi offerti da tool a pagamento offre un supporto a numerose lingue. Proprio con questo spirito sono nati, in ambito open source, progetti quali TINT e SentIta sviluppati da team direttamente interessati ad avere uno strumento che fosse in grado di effettuare analisi direttamente sul testo originale. Sarebbe infatti possibile tradurre di volta in volta il testo dalla lingua originale all’inglese, tipicamente, per operare poi con tool che supportano solo quell’idioma. Tale approccio è stato per molto tempo sconsigliato perché i servizi di traduzione risultavano poco efficaci nel trasmettere in maniera precisa il contenuto di una frase. Ogni lingua, infatti, ha i propri costrutti e le proprie peculiarità e risulta complesso convertire determinati concetti da una lingua all’altra, di conseguenza effettuare determinate analisi su un testo tradotto porta ad una fallace interpretazione del testo di partenza. Fortunatamente, nel corso del tempo, gli strumenti di traduzione sono diventati sempre piú efficaci e sempre piú in grado di trasferire in maniera efficace le sfaccettature che si cerca di dare ad una determinata frase. Per questo motivo si è cercato, nell’ambito di questo elaborato, di verificare la reale efficacia di un approccio di questo tipo.

Tornando quindi all’analisi dei tool di emotion mining, si è deciso di optare per il tool EmoNet, piú recente e semplice da utilizzare rispetto a EmoTxt, che invece si è rivelato arduo da configurare e da installare. La particolare struttura di EmoNet, inoltre, consente di avere un quadro completo di un testo, descrivendolo sempre come la somma delle emozioni principali della sfera umana.

A partire da queste scelte si è derivata un’architettura che consentisse, a partire da un commento in italiano, di ricavare il valore di polaritá capace di descriverlo e un valore che indichi la presenza di emozioni all’interno del testo stesso. Avendo a disposizione solo strumenti di emotion mining in grado di operare sull’inglese, si è cercato di individuare uno strumento che operasse una traduzione del testo in italiano fornito. Si è individuato in google translate il tool piú indicato per rispondere a questo scopo. Questi, infatti, non solo nel corso del tempo è riuscito a raggiungere un livello di accuratezza nella traduzione estremamente alto, soprattutto verso la lingua inglese ma, inoltre, mette a disposizione una API semplice e completa che puó essere facilmente utilizzata. Per questi motivi si è cercato di integrare tale strumento nell’architettura finale. A tale scopo si è individuato Googletrans, una libreria python gratuita che implementa le funzionalitá della Google Translate API. Tale libreria è compatibile con tutte le versioni di Python (2.7+ e la 3.4+) e riesce a costruire delle chiamate http verso Google Translate e a gestire la risposta in modo da ottenere esattamente la funzionalitá ricercata. In questo modo l’architettura è completamente trasparente all’operazione di traduzione che viene fatta direttamente online e riesce ad ottenere rapidamente una versione tradotta della frase. Il tool è semplice e rapido da installare, consente una serie di personalizzazioni anche sulle modalitá di comunicazione con la Google Translate Api che non sono state prese in considerazione perché poco rilevanti ai fini dell’elaborato.

Avendo in ogni caso la necessitá di tradurre il testo dall’italiano all’inglese per poter operare sui tool di emotion mining, si è deciso di sfruttare questa necessitá per effettuare un ulteriore controllo per quanto concerne il calcolo della polaritá. Per Questa ragione si é stati costretti ad individuare un tool aggiuntivo per completare l’analisi. In virtú del contesto utilizzato si è deciso di optare per la Stanford CoreNLP, in particolare, ovviamente, per il modulo che si occupa della sentiment analysis. La scelta è stata motivata principalmente dai risultati ottenuti da Lin e altri [2], dove si evidenzia come far operare analizzatori basati sul machine learning in contesti diversi da quelli in cui sono stati addestrati ne riduce di molto le prestazioni. La libreria Stanford CoreNLP è stata addestrata sulle recensioni cinematografiche degli utenti, e rappresenta, di conseguenza l’alternativa migliore rispetto agli scopi di questo elaborato. In questo caso, infatti, si analizzeranno recensioni di spettacoli teatrali, che sono perfettamente assimilabili a quelle rilasciate sulle opere cinematografiche su cui il tool è stato addestrato.

## Ambiente di sviluppo

Durante l’analisi dei tool ci si è reso conto come questi facciano ampio utilizzo del linguaggio Python e di librerie legate al machine learning e al deep learning che proprio con questo linguaggio sono ampiamente utilizzate e diffuse. In particolare si è notato come sia SentIta che EmoNet siano stati sviluppati direttamente in python e sono nati per essere utilizzati con quel linguaggio. Per questo motivo si è poi optato per un tool quale Googletrans che rappresenta un semplice wrapper python di una api remota. Prima di procedere con una descrizione piú puntuale dell’architettura del tool costruito, si vuole presentare l’ambiente d’esecuzione su cui sono stati effettuati i test e come, quest’ultimo, è stato configurato per poter operare. Si è deciso di utilizzare un’ambiente Linux in particolare la versione 19.10 di Ubuntu, per semplicitá e facilitá di configurazione e utilizzo. Sul sistema operativo è stata installata la versione 3.7 di Python, necessario per tutti i tool utilizzati. Accanto a questi si è provveduto ad installare le diverse dipendenze dei tool, in particolare Keras per SentIta, oltre a PyTorch, transformers, numpy e pandas per EmoNet. Infine si è installato Tensorflow, la nota libreria utilizzata per l’apprendimento automatico, che viene utilizzata da entrambi i tool selezionati. Per quanto concerne, invece, la parte di sentiment analysis effettuata sul testo in inglese si è reso necessario reperire la libreria Stanford CoreNLP. La libreria è scritta in Java ed è semplice da reperire e non richiede alcuna installazione. Sono stati inoltre sviluppati nel tempo numerosi wrapper che consentono di interfacciare la libreria stessa con molti linguaggi. Nel caso specifico si è utilizzato python e per questo si è fatto riferimento al wrapper Pycorenlp che ha consentito in poco tempo di accedere alle funzionalitá della libreria. La Stanford CoreNLP, in queste particolari configurazioni, viene avviata come server locale e puó essere acceduta per sfruttarne i servizi esposti. Il funzionamento è semplice e immediato e la documentazione reperibile online facilita il lavoro. Altre librerie secondarie che vengono gestite direttamente dai tool in fase di installazione e per questo non si sono prese in esame, in questo senso si vuole notare come l’ausilio Pip (Python Packet Index) faciliti di molto il lavoro degli sviluppatori consentendo di recuperare molti delle dipendenze utilizzate dai tool in maniera automatica.

In aggiunta alle dipendenze strettamente legate ai tool d’analisi utilizzati, si rende, ovviamente, necessaria una connessione alla rete per tutta la fase d’analisi per consentire al Googletrans di procedere all’invocazione del servizio remoto. Allo stesso modo si sono utilizzate altre librerie tipiche di python per gestire i file csv, per poter interagire con il sistema operativo, e per operare sul formato xml attraverso cui si deciso di interagire con la Stanford CoreNLP.

## Architettura dell’analizzatore

Conclusa la doverosa premessa sull’ambiente operativo si vuole ora presentare la struttura dell’analizzatore costruito. L’architettura è stata costruita intorno ai tool sopra descritti, al fine di farli comunicare e operare al meglio. Per questa ragione si è deciso di operare su dei file di testo piatti, quali i csv, che contenessero i dati di partenza, ovvero le recensioni delle opere teatrali analizzati, verrá poi descritto nel capitolo successivo come è stato costruito il dataset utilizzato. L’utilizzo dei file di tipo csv ha consentito una piú veloce integrazione con i fogli di calcolo che sono stati utilizzati per taggare il dataset manualmente e per effettuare le analisi desiderate.

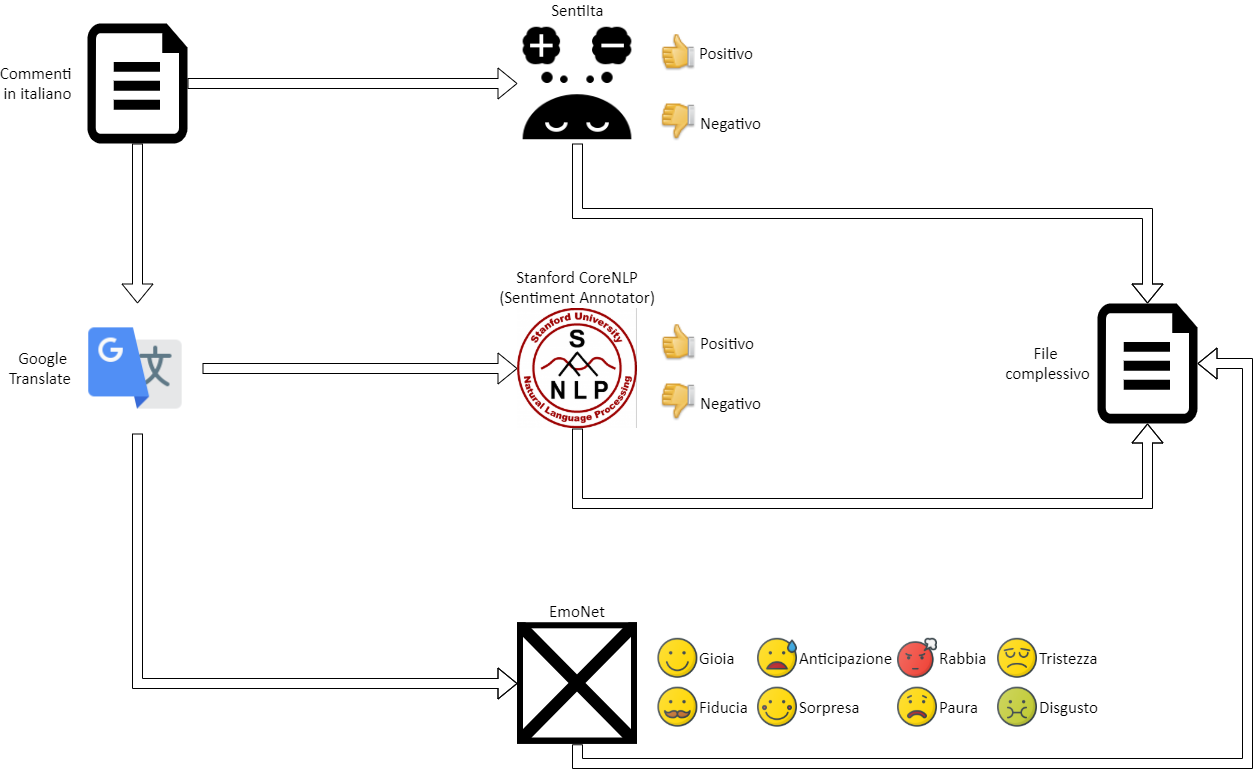


Figura Architettura del tool realizzato

In Figura 8 si è deciso di rappresentare l’architettura del tool costruito. In particolare si vogliono evidenziare i tool utilizzati e le loro interazioni. A partire da un semplice file di testo, vengono estratti i commenti in italiano. Questi vengono dati in input al tool SentIta che si occupa di definirne la polaritá, e contemporaneamente viene interpellato il servizio di traduzione fornito da Google Translate. L’output del servizio di traduzione, ovvero i commenti in lingua inglese, vengono poi elaborati dal servizio offerto di Stanford CoreNLP, o meglio dal modulo Sentiment Annotator della libreria, il quale effettua un secondo calcolo sulla polaritá. I risultati ottenuti dai due tool utilizzati per la sentiment analysis saranno confrontati nel capitolo successivo, si tra di loro che rispetto all’oracolo utilizzato. Oltra alla Stanford CoreNLP, anche EmoNet utilizza l’output generato da Google Translate per verificare la presenza di emozioni all’interno del testo in inglese. Tutti i tool poi sintetizzano le informazioni calcolate in un unico file di testo, che è diventata la base per le successive fasi d’analisi dei dati.

Accanto alla struttura generale che si è utilizzata, si mostra in Figura 8 l’activity diagram del tool costruito con i vari passaggi che il commento iniziale segue per poter essere analizzato. Si vuole cercare in questo modo di facilitare la comprensione e dettagliare meglio l’intero progetto. I commenti letti dal file csv vengono inseriti in una struttura e per ognuno viene avviato il processo d’analisi. Il primo passo è la chiamata al modulo SentIta che procede a calcolare la polaritá direttamente sul commento in italiano. Segue la comunicazione con il modulo Googletrans che procede ad invocare, tramite chiamata http, il servizio offerto da google per ottenere la traduzione. Il testo tradotto diventa la base per completare la fase di analisi. Viene prima utilizzato dal componente che si occupa del calcolo di polaritá, pycorenlp. I risultati vengono poi utilizzati per aggiornare il commento e si passa poi alla fase successiva. Questa consiste nell’analisi del testo inglese effettuata dal modulo EmoNet che procede a verificare la presenza di emozioni nel testo inglese. Completata quest’ultima fase l’oggetto commento, viene aggiornato nuovamente e risulta completo, si procede, quindi, con il commento successivo. Una volta terminato il loop, quindi una volta analizzati tutti i dati, si procede alla creazione di un nuovo file csv dove sono inserite tutte le nuove informazioni ottenute. Queste poi verranno confrontate con le analisi effettuate manualmente da diversi operatori per verificare il grado di affidabilità dei tool considerati, nel contesto di riferimento. Il processo di confronto appena descritto sará dettagliano nel capitolo successivo, previa descrizione del dataset utilizzato e della sua costruzione.

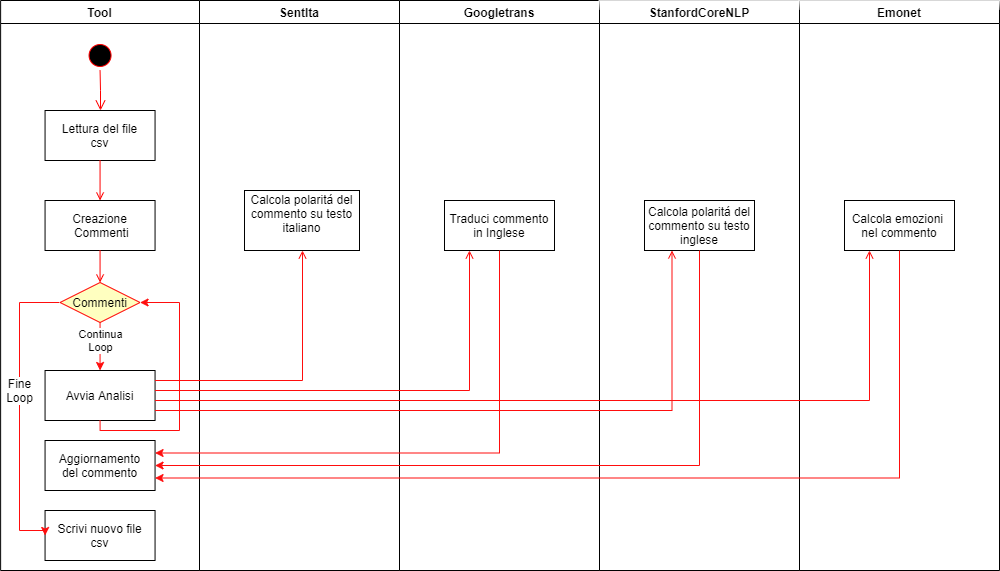


Figura 8 Activity diagram del tool realizzato

# Capitolo 4

## Introduzione

In questo capitolo si procederá a presentare il dataset utilizzato per poi concentrarsi sull’analisi dei dati ottenuti.

## Costruzione del dataset

Il progetto CANTICO, come presentato nel capitolo 1, mira a migliorare l’esperienza di un pubblico interessato ad opere teatrali. Per questa ragione sin da subito si è rivelato chiaro come fosse necessario utilizzare un dataset di commenti che riguardassero opere e spettacoli teatrali. Il progetto, inoltre, opera espressamente in italia e questo rende necessario l’utilizzo proprio di questa lingua per i commenti. Tali necessità si sono, purtroppo, ben presto scontrate con l’impossibilitá di trovare un dataset con queste caratteristiche, e per questa ragione, si è deciso di procedere alla costruzione di un dataset piú rappresentativo. La ricerca di commenti e recensioni di spettacoli teatrali si è rivelata particolarmente ardua per una serie di ragioni. Prima fra tutte la scarsitá di recensioni disponibile in rete su questo tipo di argomento. Purtroppo il teatro e gli spettacoli teatrali ricevono poco spazio sulla rete, ed è anche per questo che si sta cercando di sviluppare un progetto come CANTICO. Accanto a questa scarsa presenza, si aggiunge la difficoltá nel reperire commenti realmente utilizzabili. Ció che, infatti, risulta essere maggiormente disponibile su internet e sui siti specializzati in opere teatrali sono le lunghe recensioni di critici teatrali. Queste ultime sono ricche di informazioni e in grado, per un appassionato, di sintetizzare bene le qualitá e i difetti delle performance degli artisti di questo settore, ma non risultano utili ai fini di questo elaborato. Tali recensioni vengono regolarmente redatte dai critici e poi analizzate dagli artisti e dai direttori dei teatri da sempre. Questo tipo d’analisi riguarda un aspetto ampiamente coperto dalla produzione teatrale e che per questo viene, nel contesto dell’elaborato, tralasciato. Ció che si cerca di fare, in questo caso infatti, è quello di ricavare maggiori informazioni da ció che è direttamente il pubblico ad esprimere e a comunicare sull’opera. Per questo tipo di analisi quindi sono maggiormente indicati i classici commenti di poche righe che gli utenti lasciano al termine di esperienze organizzate. L’esempio piú classico di questo tipo di approccio è dato da qualsiasi sito di e-commerce che consente di postare una recensione di un prodotto acquistato, che sia una summa del prodotto in sé e dell’intera esperienza d’acquisto anche rispetto al venditore o alle terze parti intervenute. Parimenti si possono citare come esempio le recensioni di locali e ristoranti che negli ultimi anni sono minuziosamente descritti dalla maggior parte dei loro clienti. Questo processo di recensione si basa su testi e analisi molto piú sintetiche e dirette delle classiche recensioni di un critico del settore. In questo caso è proprio il loro enorme numero e la loro varietá che consente, poi, agli utenti di effettuare una scelta rispetto all’acquisto o meno. In questa ottica questo tipo di recensioni è molto difficile da reperire nell’ambito degli spettacoli teatrali. A quanto pare, attualmente, il pubblico che assiste abitualmente agli spettacoli teatrali è poco incline a rilasciare commenti e recensioni in rete. Questa attitudine ha causato problemi nella costruzione del dataset, che lo ha portato ad essere costituito da un numero non statisticamente significativo di elementi, ma utile ai fini dell’analisi qui effettuata. In particolare il dataset è stato costruito con le recensioni reperite sul sito tripadvisor.it, limitando la ricerca agli spettacoli e alle opere teatrali. Tale limitazione ha generato un numero basso di attrazioni utilizzabili, che peró avevano un numero discreto di recensioni. A partire da queste recensioni si è estratto il testo e si è costruito un dataset che poi è stato analizzato. Il dataset finale è costituito da 268 commenti in lingua italiana di lunghezza variabile che rappresentano una buona base di partenza per effettuare le analisi richieste.

## Definizione dell’oracolo

Prima di procedere con l’analisi dei dati ottenuti, si è proceduto a costruire l’oracolo con cui verranno poi confrontati i dati ottenuti dagli strumenti automatici. A tal fine si è reso necessario identificare manualmente le informazioni contenute nel dataset utilizzato. La procedura di validazione è stata eseguita da due persone, che hanno provveduto prima a taggare manualmente il dataset, per poi concordare un oracolo definitivo. Il dataset, costituito da una semplice sequenza di frasi, è stato prima classificato singolarmente da due persone costruendo, ognuno, una struttura come mostrata in Figura 9. Si è deciso di procedere classificando la polaritá del commento e individuando la presenza o l’assenza delle 8 emozioni principali di Plutchik. Per facilitare l’analisi successiva la polaritá è stata valutata sulla base di una scala a 5 livelli, dal piú basso Negativo al piú alto Positivo. Le emozioni invece sono state indicate con la sola presenza di un flag “Yes” o “No”. Questo processo iniziale ha portato alla definizione di due dataset taggati individualmente. A partire da questi ultimi è si è ricavato l’oracolo definitivo.

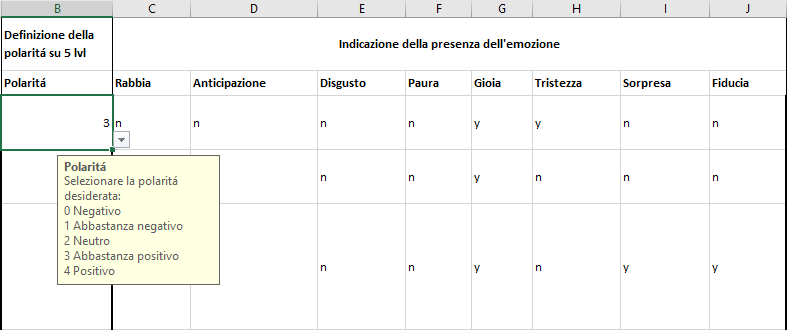


Figura Esempio dataset taggato

Inizialmente infatti il tagging individuale ha portato ad una discrepanza tra i due dataset ottenuti, si discuterá a breve di questa ulteriore difficoltá strettamente legata all’uomo. In particolare si è notata una discrepanza di circa il 13% per quanto concerne la definizione della polaritá dei commenti, mentre si è notata una discrepanza di circa il 5% in media per quanto concerne l’individuazione delle emozioni, come evidenziato con maggiori dettagli nella Figura 10. Sono mostrati anche i risultati di un ulteriore analisi sulla polaritá, che è stata effettuata collassando i valori positivi e negativi della scala a 5 livelli, generando un notevole miglioramento nella percentuale di agreement tra i due soggetti. Giá da questa prima analisi si nota come tra sole due persone vi sia una differenza di interpretazione del medesimo testo. Per questa ragione, la stesura dell’oracolo è passata attraverso una seconda revisione dell’intero dataset ad opera dei due utenti che hanno provveduto a taggarlo inizialmente. Questa revisione, effettuata in contemporanea ha portato a ció che poi è stato considerato come oracolo. In questo processo si sono cercati di minimizzare le discrepanze d’interpretazione tra i due utenti e si è provveduto ad analizzare in maniera quanto piú oggettiva possibile il testo.

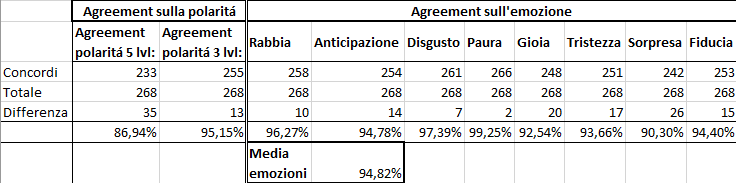


Figura Agreement definizione oracolo

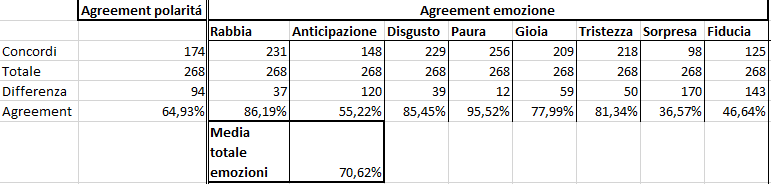


Figura Analisi percezione tra 3 persone

Ad avvalorale la tesi secondo cui vi sono notevoli difficoltá di agreement giá tra esseri umani, si è esteso il confronto ad una terza persona, che non ha peró partecipato alla definizione dell’oracolo, ma che é stata utilizzata per analizzare la variabilitá della percezione delle emozioni semplicemente tra persone diverse. Anche in questo caso è stato seguito il medesimo processo di classificazione in modo da avere dei dati confrontabili. Attraverso questo processo si è visto come sia davvero complicato avere la stessa percezione, e come ognuno percepisca le informazioni in maniera del tutto personale. Il dato piú significativo da questo punto di vista è sicuramente quello per cui solo su 10 elementi del dataset di partenza vi è stata una classificazione unanime tra tutti i partecipanti, che corrisponde a circa il 4% del dataset. Questo elemento è sicuramente il piú significativo per descrivere la varbilitá dell’argomento che si sta trattando in questo elaborato. Prendendo in esame, invece, solo le emozioni si nota come l’unanimitá si ottenga in circa il 6% del dataset, mentre considerando solo la polaritá classificata su 5 valori si raggiungono risultati nell’ordine del 64%. I dati mostrati in Figura 11, confermano come sia molto basso il livello di agreement tra i soggetti che hanno proceduto al tagging, e si puó immaginare che aumentando il numero di persone le percentuali di agreement possano solo peggiorare. Nella stessa immagine si vede come le emozioni, analizzate singolarmente abbiano risultati variabili e probabilmente legati alla percezione delle singole persone e a come queste interpretano una determinata emozione. Per rendere piú chiaro questo concetto si consideri uno dei commenti che fanno parte del dataset utilizzato: “Abbiamo apprezzato molto questo concerto d'opera di 60 minuti in una bella chiesa di Lucca. Eravamo in città solo per una sera, ma saremmo andati tutte le sere se fossimo rimasti qualche giorno in più. Siamo rimasti colpiti dal talento dei due cantanti e dalla maestria del pianista.”. Questo commento, nell’oracolo è stato classificato come totalmente positivo, 4 sulla scala di valori utilizzata, e si riscontra la presenza di Gioia, Tristezza e Sorpresa. La gioia è legata all’apprezzamento mostrato, mentre la sorpresa si evince dal fatto che i protagonisti sono rimasti colpiti dal talento dei due cantanti. La tristezza invece si potrebbe ricavare dal rammarico legato al fatto che coloro i quali hanno assistito allo spettacolo hanno trascorso una sola notte in cittá. Questa tristezza, peró, viene estratta dal testo attraverso una deduzione del contesto e per questo, non solo potrebbe risultare difficile da rilevare per un tool automatico ma, è anche oggetto di dibatto tra le persone che devono classificarla. I tre soggetti che hanno proceduto alla classificazione manuale, infatti, hanno dato 3 dato tre interpretazioni diverse di questa stessa frase. Due su tre hanno assegnato un valore di polaritá totalmente positivo, mentre un terzo ha assegnato un valore in meno sempre ritenendolo un commento positivo. Per quanto concerne le emozioni, invece, tutti e tre i tester hanno rilevato la presenza di gioia e sorpresa all’interno del commento, mentre uno solo ha indicato come presente la tristezza, ed uno solo ha indicato la presenza di fiducia. Questa divergenza d’opinione evidenzia come in alcune situazioni sia davvero complesso valutare cosa sia espresso nel testo e come anche i tool automatici possano essere poco oggettivi in senso assoluto, dal momento che non esiste una interpretazione univoca.

Terminata questa giusta disamina per la definizione del dataset, analizziamo brevemente come è caratterizzato l’oracolo. Questi è composto da 268 commenti in italiano che riguardano una serie di spettacoli, che spaziano dalle opere teatrali classiche fino ad alcuni spettacoli piú moderni legati a proiezioni in piú dimensioni e a percorsi culturali alla scoperta dei monumenti storici. Di questi, circa il 67% è stato identificato come totalmente positivo, circa il 18% come positivo, circa il 4% come neutro, un ulteriore 4% è stato identificato come negativo e, infine circa un 8% come totalmente negativo. Per quanto concerne le emozioni si è visto che la gioia è l’emozione piú presente con un riscontro nel 74% del dataset, seguono con valori molto vicini la sorpresa con il 39%, la fiducia con il 34% e la tristezza con il 26%. Meno presenti sono la rabbia con circa il 9% di riscontri, l’anticipazione con il 6% e il disgusto con l’1%. Il sentimento della paura, invece, non è stato mai riscontrato nella definizione dell’oracolo.

## Elaborazione dati sperimentali

Terminata la definizione dell’oracolo si sono utilizzati i tool selezionati e si è provveduto ad analizzare il dataset. Una volta terminata l’analisi si è reso necessario una ulteriore elaborazione dei dati al fine di renderli confrontabili con quelli presenti nell’oracolo. I tool, infatti, forniscono in output una serie di dati che sono piú ad un livello piú fine di dettaglio, ma che a volte puó risultare in una informazione ricca di rumore e per questo si è reso necessario effettuare delle semplificazioni. Si partirá analizzando, prima di tutto la polaritá ottenuta. Quest’ultima viene espressa in maniera diversa da SentIta e dalla Stanford CoreNLP. Il tool che opera sul testo inglese, Stanford CoreNLP, esprime il risultato di un commento dividendolo nelle frasi che lo compongono e per ognuno fornisce un valore di polaritá su una scala a 5 livelli, come quella utilizzata nell’oracolo. A partire da questo livello di dettaglio si è deciso di sommare i valori delle singole frasi e di farne la media, in modo da avere un valore unico che fosse rappresentativo di ogni commento. SentIta, invece, espone i propri risultati definendo un valore tra 0 e 1 di positivitá del testo fornito in input, e un valore tra 0 e 1 per definire la negativitá, entrambi i valori sono stati poi ricondotti su base 100 per migliorare la leggibilitá. A partire da questi valori si sono testate diverse soglie per ricondurre l’analisi alla medesima scala a 5 valori utilizzata precedentemente. In particolare si è tentato un approccio che verificasse la distanza tra il valore della positivitá e il valore della negativitá, riconducendo questi ai diversi livelli della scala. Definito l’approccio si sono utilizzate diverse soglie di conversione:

* Prima si è tentato di identificare i commenti neutri dati da una differenza tra positivitá e negativitá. La soglia sotto la quale si è identificato il commento neutro è stata posta una volta pari a 20% ed una volta pari al 30%.
* Poi si è cercato di individuare i commenti estremamente negativi o positivi, verificando che la differenza tra i due valori fosse superiore nel primo caso al 50% nel secondo al 60%.
* Qual’ora tutte queste verifiche non fossero risultate vere, si è provveduto a verificare se uno solo dei due valori, positivo o negativo, superasse la soglia del 50%, in tal caso veniva classificato in quel modo.
* L’ultima verifica, che veniva effettuata solo se tutte le precedenti non si fossero verificate, controllava se entrambi i valori fossero risultati superiori o inferiori alla soglia del 50% e nel caso classificava il commento come neutro.

Nel paragrafo successivo saranno discussi i risultati ottenuti.

Per quanto concerne, invece, l’emotion mining, si ricorda che si è utilizzato il tool EmoNet. Questi fornisce in output, una classificazione dell’input, individuando la percentuale di presenza delle 8 emozioni principali nel testo fornito. In particolare la somma delle 8 emozioni riconosciute deve sempre essere uguale al 100%. Ne segue che l’output del tool necessita di una conversione alla piú semplice rappresentazione presente nell’oracolo. Il modo piú semplice era convertire tutti i valori superiori allo 0 nel flag che indica la presenza di quell’emozione e tutti i valori pari allo zero nel flag opposto. Tale approccio peró non sarebbe risultato efficace. La modalitá di funzionamento del tool, infatti, lo porta quasi sempre all’individuazione di tutte le 8 emozioni principali, alcune, peró, sono caratterizzate da percentuali irrisorie. Per queste ragioni si sono individuate una serie di soglie minime che rappresentano consentono la conversione al flag che indica la presenza di quella specifica emozione nel commento. Le soglie utilizzate sono: il 50%, il 40%, il 30%, il 20%, il 10% e il 5%. I risultati ottenuti con le varie soglie saranno discussi nel paragrafo successivo.

## Analisi dei dati sperimentali

### Analisi della polaritá

L’analisi della polaritá è partita da un confronto dei risultati dei tool utilizzati. Questi, infatti, hanno operato su lingue diverse e quindi si è verificato quanto i risultati dell’uno combaciassero con quelli dell’altro. Il confronto, come anticipato precedentemente, è stato possibile solo attraverso una fase di conversione dei risultati di entrambi i tool ad un unico valore su una scala a 5 livelli che fosse rappresentativo dell’intero commento. Durante questa fase si sono utilizzati diversi approcci che hanno generato conversioni diverse, che sono state tutte utilizzate per il confronto. In Figura 12 si mostrano le percentuali di agreement tra i tool utilizzati, sfruttando i meccanismi di conversione precedentemente elencati. Si nota come i due tool non mostrino dei risultati decisamente distanti. Con la prima configurazione utilizzata nella fase di conversione dei risultati ottenuti con SentIta, si vede come la percentuale di agreement con la Stanford CoreNLP sia decisamente bassa, prossima al 34%. Tale percentuale migliora leggermente sfruttando la seconda configurazione di conversione, ma non raggiunge mai risultati elevati. Solo collassando i valori positivi e negativi su una scala a 3 livelli per entrambi i dati forniti dai tool, si riesce ad ottenere un agreement che supera di poco il 50%. Da sola quest’analisi risulta poco utile ma, in realtá, si rivela importante quando si confrontano i risultati ottenuti con l’oracolo. Tale confronto, infatti, permetterá di definire quale tool sia piú preciso nell’individuazione della corretta polaritá. Il confronto delle varie interpretazioni dei risultati del tool è mostrato in Figura 13. Da questa si evince chiaramente come la prima versione della conversione dal tool SentIta sia quella piú vicina ai risultati veri contenuti nell’oracolo. Tale risultato risulta comunque essere molto poco significativo, dal momento che il tool recepisce in maniera corretta la polarizzazione della frase, solo nel 40% dei casi. Questo risultato risulta essere davvero deludente, in quanto lascia un margine d’errore troppo ampio per poter sfruttare il tool al pieno delle proprie funzionalitá. Purtroppo questo risulta essere il dato migliore anche rispetto alla Stanford CoreNLP, la quale non riesce ad andare oltre una precisione del 15%. Questi risultati deludenti consentono di affermare che un’architettura basata sulla traduzione del testo dall’italiano all’inglese al fine di rilevarne la polaritá, risulta essere molto imprecisa, mentre un tool sviluppato per riconoscere la lingua italiana gode di una precisione notevolmente superiore. Questi risultati, ottenuti con una versione primordiale del tool SentIta, lasciano ben sperare per il futuro. Soprattutto quando il tool avrá risolto i problemi nell’analisi dei testi piú lunghi, questi potrebbe rappresentare un ottimo strumento per il calcolo della polaritá. Un possibile miglioramento dei risultati potrebbe essere ottenuto analizzando il testo nelle varie frasi che lo compongono e poi effettuare il calcolo del valore medio, in maniera simile al procedimento che la Stanford CoreNLP segue in maniera nativa.

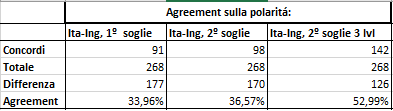


Figura Confronto tool Ita - Ing

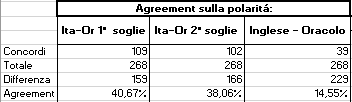


Figura Confronto tool-oracolo

### Analisi delle emozioni

Il tool utilizzato per l’analisi di emotion mining, ovvero EmoNet, si è rivelato abbastanza efficiente. La scelta delle soglie è stata dettata dai risultati ottenuti durante le analisi. Come si evince dalla Figura 14, in tutti i casi presi in esame il tool riesce a riscontrare la presenza di un’emozione almeno nel 50% dei casi. Questa condizione risulta sempre piú valida all’aumentare della soglia di scarto utilizzata per filtrare i dati in input. Tale andamento ha portato ad aumentare la soglia minima di riconoscimento del tool per attivare il flag “Yes” sulla particolare emozione. Inizialmente, infatti, si erano utilizzate soglie dal 5% al 30%, riscontrando un notevole aumento della precisione del tool. Tale miglioramento si vede soprattutto analizzando la media totole delle emozioni, nel grafico in Figura 14 le ultime colonne. Si evidenzia come ci sia un miglioramento di hit di 10 punti percentuale tra la soglia imposta al 5% e quella al 30%. In particolare si passa da un’hit rate pari al 69% nel caso della soglia al 5% ad un’hit rate dell’81% con la soglia del 30%. Tale andamento ha portato ad aumentare ancora la soglia al fine di verificare fino a quando questo andamento sarebbe stato rispettato. Utilizzando le soglie al 40% e al 50% si e riscontrato lieve ulteriore miglioramento, che rischia peró di inficiare i dati e portare a delle letture sporche. I dati ottenuti, analizzati sul dataset di riferimento, mostrano come il tool sia particolarmente efficiente nell’ individuare le emozioni non presenti nel testo. Tale affermazione è giustificata dagli altissimi hit rate presenti sulle emozioni di paura, disgusto e anticipazione, che rappresentano, secondo l’oracolo, i sentimenti meno presenti nel dataset. Questo dato non inficia la validitá dei dati utilizzati, piuttosto avvalora la tesi sulla difficoltá nel percepire in maniera oggettiva le emozioni nel testo da parte degli esseri umani. Il sentimento della gioia, di gran lunga il piú presente nel dataset, è l’unica emozione che all’aumentare della soglia utilizzata fa registrare un’hit rate leggermente piú basso, si abbassa di circa il 4%. Questo dato potrebbe indicare che il rumore dell’analisi viene distribuito in maniera uniforme alla percentuale di emozione rilevata. Dal momento che la gioia rappresenta l’emozione con la percentuale di rilevamento in media piú alta, si attesta intorno al 67% mentre tutte le altre sono ampiamente al di sotto del 10%, la gioia subisce maggiormente gli effetti del cambio di soglia. Una seconda ipotesi è legata ad un semplice ragionamento numerico. La gioia rappresenta l’emozione piú presente nel dataset sia secondo l’oracolo sia secondo EmoNet in qualsiasi condizione di soglia considerata. Visti i numeri piú elevati di presenza la gioia è piú soggetta ad errori. L’emozione che, peró, ha fatto registrare i risultati peggiori è la sorpresa. Questa si è rivelata molto ardua da individuare da parte degli uomini, che ne hanno indicato la presenza molto piú spesso del tool automatico. Questo dato è molto interessante, perché probabilmente legato ad una interpretazione errata del concetto di sorpresa da parte dei tester evidenza che conferma, ancora una volta, quanto sia complesso cercare di definire in maniera univoca ed oggettiva questi dati.



Figura Analisi emotion mining

# Conclusioni

L’elaborato si prometteva di verificare lo stato dell’arte nell’ambito della sentiment analiysis e dell’emotion mining al fine di adattare questi strumenti al progetto CANTICO. Oltre ai risultati strettamente legati allo sviluppo e al software, si è potuto verificare con mano la difficoltà che vive in questo periodo storico il teatro e l’opera. La mancanza di un dataset dedicato a questo settore non è, di per sé, un qualcosa di sorprendente. La vera criticitá è consistita nella semplice ricerca di commenti, opinioni e recensioni dei fruitori di tali spettacoli. Questi dati grezzi, sono presenti sulla rete in maniera massiva su una enorme quantitá di argomenti, dalla semplice descrizione di oggetti acquistati on line, fino alla descrizione delle piú svariate esperienze, dai ristoranti ai film passando dai libri. Questa condizione, si ritiene, possa essere estremamente significativo di quanto il teatro sia lontano dal mondo digitale e di quanto il progetto CANTICO possa rappresentare una svolta per avvicinare questi due mondi.

Questa analisi preliminare è stata seguita dalla fase di definizione dello stato dell’arte rispetto agli obbiettivi prefissati. In questa fase si è notato come l’intero Naturale Language Processing abbia fatto, negli anni, notevoli progressi. Tali risultati si sono visti anche nelle branche dedicate alla sentiment analysis e all’emotion mining, con l’intelligenza artificiale che ha significato, in molti ambiti tra cui questo, cambiare radicalmente approccio ed essere in grado di ottenere dei buoni risultati e delle ottime performance. In quest’ambito le difficoltá nella definizione dello stato dell’arte sono state legate principalmente ad aspetti linguistici. L’intenzione era, infatti, quella di individuare tool che fossero in grado di operare con la lingua italiana ma, purtroppo, non è stato possibile realizzare pienamente questo scopo, é stato possibile individuare un solo tool che operasse nativamente sull’italiano. In particolare il tool è SentIta, dedicato alla sentiment analysis, e accanto a quest’ultimo si sono individuati una serie di tool in grado di operare sulla lingua inglese. In questa condizione i due approcci possibili erano quelli legati alla creazione di un tool totalmente nuovo in grado di operare sull’italiano nell’ambito dell’emotion mining, oppure tentare prima un approccio piú conservativo, che tentasse di sfruttare strumenti che sono molto complessi da sviluppare. In quest’ottica si è deciso di sfruttare i tool basati sulla lingua inglese, e di verificarne l’efficienza sulla lingua italiana, integrando un tool di traduzione che possiamo ormai definire stabile ed efficiente quale google translate. Per questa ragione si sono selezionati altri due tool, il sentiment Annotator presente nella Stanford CoreNLP ed EmoNet. Il primo è un tool di sentiment analysis i cui risultati sono stati confrontati con SentIta, mentre il secondo è il tool dedicato all’emotion mining.

Le particolari condizioni operative in cui il progetto CANTICO si troverá ad operare, ha fatto sì che la creazione ad hoc di un dataset fosse una scelta obbligata. Per questa ragione si è costruito un dataset di recensioni in italiano di opere teatrali e spettacoli, con la conseguente creazione manuale di un oracolo che potesse essere sfruttato per la fase di verifica. Durante questo processo ci si è resi conto di quanto questo tipo di analisi sia complessa e molto soggetta all’opinione personale. Analizzando i risultati ottenuti da diverse persone nell’ambito della definizione dell’oracolo, si è visto come giá tra queste non vi fosse accordo e, anzi, all’aumentare degli individui interpellati la variabilità dei risultati aumentasse inesorabilmente. Queste informazioni diventano estremamente preziose quando si analizzano i risultati ottenuti dai tool. Questi, infatti, hanno rivelato un’hit rate variabile tra il 15% ed il 40% nell’ambito della sentiment analysis, mostrando come un tool che opera direttamente sull’italiano sia molto piú efficiente. Il risultato ottenuto risulta comunque essere abbastanza basso e potrebbe essere soggetto a miglioramenti operando delle modifiche nell’approccio utilizzato in questo elaborato. Per quanto concerne, invece, l’emotion mining, l’approccio si è rivelato efficiente mostrando un’hit rate tra il 70% e l’80% al variare delle condizioni d’analisi, un valore molto alto se si considera la variabilitá intrinsecamente legata alla definizione delle emozioni. Proprio in questo senso i risultati ottenuti vanno interpretati e va sottolineato come strumenti di analisi emozionale e polaritá potrebbero, forse, essere una soluzione per approcciare a questo tipo di problemi in maniera oggettiva. Questo obbiettivo è ancora lontano ma, lo stato dell’arte mostra come ci si stia avvicinando sempre piú al raggiungerlo. Probabilmente peró, saranno sempre e comunque gli esseri umani a non riuscire a trovare un accordo definitivo nell’ambito della sfera emozionale, e questa condizione non potrá essere gestita in maniera oggettiva e univoca dalle macchine che cercheranno di imparare da noi come interpretare i sentimenti.

# Indice delle figure e delle tabelle

[Figura 1 Esempio analisi TINT [5] 18](file:///C:\Users\39393\Documents\Git\TheaterAnalysis\Tesi\Tesi.docx#_Toc32227052)

[Figura 2 Esempio analisi Meaningcloud [6] 20](file:///C:\Users\39393\Documents\Git\TheaterAnalysis\Tesi\Tesi.docx#_Toc32227053)

[Figura 3 esempio analisi Repustate [7] 21](file:///C:\Users\39393\Documents\Git\TheaterAnalysis\Tesi\Tesi.docx#_Toc32227054)

[Figura 4 esempio analisi twinword [11] 26](file:///C:\Users\39393\Documents\Git\TheaterAnalysis\Tesi\Tesi.docx#_Toc32227055)

[Figura 5 esempio paralleldots [12] 26](file:///C:\Users\39393\Documents\Git\TheaterAnalysis\Tesi\Tesi.docx#_Toc32227056)

[Figura 6 Ruota delle emozioni [13] 28](file:///C:\Users\39393\Documents\Git\TheaterAnalysis\Tesi\Tesi.docx#_Toc32227057)

[Figura 7 Architettura del tool realizzato 36](file:///C:\Users\39393\Documents\Git\TheaterAnalysis\Tesi\Tesi.docx#_Toc32227058)

[Figura 8 Activity diagram del tool realizzato 37](file:///C:\Users\39393\Documents\Git\TheaterAnalysis\Tesi\Tesi.docx#_Toc32227059)

[Figura 9 Esempio dataset taggato 41](file:///C:\Users\39393\Documents\Git\TheaterAnalysis\Tesi\Tesi.docx#_Toc32227060)

[Figura 10 Agreement definizione oracolo 41](file:///C:\Users\39393\Documents\Git\TheaterAnalysis\Tesi\Tesi.docx#_Toc32227061)

[Figura 11 Analisi percezione tra 3 persone 42](file:///C:\Users\39393\Documents\Git\TheaterAnalysis\Tesi\Tesi.docx#_Toc32227062)

[Figura 12 Confronto tool Ita - Ing 46](file:///C:\Users\39393\Documents\Git\TheaterAnalysis\Tesi\Tesi.docx#_Toc32227063)

[Figura 13 Confronto tool-oracolo 46](file:///C:\Users\39393\Documents\Git\TheaterAnalysis\Tesi\Tesi.docx#_Toc32227064)

[Figura 14 Analisi emotion mining 47](file:///C:\Users\39393\Documents\Git\TheaterAnalysis\Tesi\Tesi.docx#_Toc32227065)

[Tabella 1 Tools per l'analisi testuale 18](#_Toc32227066)

[Tabella 2 Tools per la sentiment analysis 25](#_Toc32227067)

[Tabella 3 Tool per l'emotion mining 29](#_Toc32227068)

# Bibliografia

|  |  |
| --- | --- |
| [1] | Wikipedia, «Wikipedia,» [Online]. [Consultato il giorno 30 Gennaio 2020]. |
| [2] | B. Lin, F. Zampetti, G. Bavota, D. P. Massimiliano, M. Lanza e R. Oliveto, «Sentiment Analysis for Software Engineering: How far can we go?,» *Proceedings of the 40th International Conference on Software Engineering,* pp. 94--104, 2018. |
| [3] | A. Palmero e G. Moretti, «Italy goes to Stanford: a collection of CoreNLP modules for Italian,» *ArXiv e-prints,* 2016. |
| [4] | M. Masudur Rahman, C. K Roy e I. Keivanloo, «Recommending insightful comments for source code using crowdsourced knowledge,» *Proceedings of SCAM 2015 (15th International Working Conference on Source,* vol. IEEE, pp. 81-90, 2015. |
| [5] | A. Palmero e G. Moretti, «TINT Demo,» [Online]. Available: http://simpatico.fbk.eu/demo/. [Consultato il giorno 30 Gennaio 2020]. |
| [6] | Meaning Cloud, «Demo,» [Online]. Available: https://www.meaningcloud.com/demos/text-analytics-demo#. [Consultato il giorno 30 Gennaio 2020]. |
| [7] | Repustate, «Demo Sentiment analysis,» [Online]. Available: https://www.repustate.com/sentiment-analysis-api-demo/. [Consultato il giorno 30 Gennaio 2020]. |
| [8] | A. Ansari, A. Seenivasan, A. Anandan e R. Lakshmanan, «Twitter Sentiment Analysis,» 12 Novembre 2017. [Online]. Available: https://github.com/abdulfatir/twitter-sentiment-analysis. [Consultato il giorno 30 Gennaio 2020]. |
| [9] | T. Cagan, S. L. Frank e R. Tsarfaty, «Generating subjective responses to opinionated articles in social media: an agenda-driven architecture and a turing-like test,» *Proceedings of the Joint Workshop on Social Dynamics and Personal Attributes in Social Media,* pp. 58-67, 2014. |
| [10] | G. Nicola, «SentIta,» 2019. [Online]. Available: https://nicgian.github.io/Sentita/. [Consultato il giorno 30 Gennaio 2020]. |
| [11] | Twinword, «Demo Emotion Mining,» [Online]. Available: https://www.twinword.com/api/emotion-analysis.php. [Consultato il giorno 30 Gennaio 2020]. |
| [12] | Paralleldots, «Demo Emotion Mining,» [Online]. Available: https://www.paralleldots.com/emotion-analysis. [Consultato il giorno 30 Gennaio 2020]. |
| [13] | M. Abdul-Mageed e L. Ungar, «EmoNet: Fine-Grained Emotion Detection with Gated Recurrent Neural,» *Proceedings of the 55th annual meeting of the association for computational linguistics,* vol. 1, pp. 718--728, 2017. |
| [14] | O. Ayman, «GitHub,» 2018. [Online]. Available: https://github.com/omar178/Emotion-recognition#p4. [Consultato il giorno 30 Gennaio 2020]. |
| [15] | J. C. Garfias Tovar, «Github,» [Online]. Available: https://github.com/SeaWar741/Emotion-detection. [Consultato il giorno 30 Gennaio 2020]. |