

Università degli Studi del Sannio

Dipartimento di Ingegneria

Corso di Laurea Magistrale in Ingegneria Informatica

**Validazione empirica e applicazione di tecniche di Sentiment
Analysis e Emotion Mining su testo in Italiano**

Relatore

Prof. Massimiliano Di Penta

Correlatore

Dott.ssa Fiorella Zampetti

Candidato

Lucio De Luca

Matricola

399/168

Anno accademico 2018/2019

Sommario

Sommario	1
Abstract	3
1 Introduzione	4
2 Contesto.....	5
2.1 Il progetto CANTICO.....	5
3 Stato dell'arte	8
3.1 Introduzione	8
3.2 NLP – Natural Language Processing.....	8
3.3 Sentiment Analysis.....	11
3.4 Emotion Mining	14
3.5 Tools	15
3.5.1 NLP – Natural Language Processing	15
3.5.2 Sentiment Analysis.....	18
3.5.3 Emotion mining	24
3.6 Visual Emotion Mining	27
4 Analizzatore.....	29
4.1 Introduzione	29
4.2 Identificazione dei tool	29
4.3 Ambiente di sviluppo	32
4.4 Architettura dell'analizzatore.....	33
5 Analisi dei dati	37
5.1 Introduzione	37

5.2	Costruzione del dataset	37
5.3	Definizione dell'oracolo	38
5.4	Metriche d'analisi	42
5.5	Elaborazione dati sperimentali.....	43
5.6	Analisi dei dati sperimentali	44
5.6.1	Analisi della polarità.....	44
5.6.2	Analisi delle emozioni	48
6	Conclusioni	55
7	Indice delle figure e delle tabelle	58
8	Bibliografia	59
9	Ringraziamenti.....	62

Abstract

L'elaborato si sviluppa all'interno di un progetto più ampio che ha lo scopo di migliorare l'esperienza teatrale agli spettatori. Il progetto in questione è CANTICO e si articola in diversi obiettivi programmatici, questo lavoro ambisce a rispondere alla prima parte del primo obiettivo. Si è cercato, prima di tutto, di descrivere il contesto operativo entro cui si opera. Si è partiti dalla descrizione del progetto a cui si fa riferimento, le fasi operative, al fine di sottolineare il ruolo di questo elaborato nel progetto complessivo. Si è poi passati all'analisi dello stato dell'arte a proposito dei problemi legati al Natural Language Processing, con particolare attenzione alla Sentiment Analysis e all'Emotion Mining. Una volta definiti i tool da utilizzare e le motivazioni che hanno portato a sceglierli, si è presentato l'ambiente operativo in cui si è operato, sottolineando tutte le dipendenze utilizzate. Si presenterà, infine, l'architettura del tool costruito, seguita da una sostanziosa analisi dei dati su cui si è operato, a partire dalla costruzione del dataset, fino alle metriche utilizzate per l'analisi passando per le modalità di definizione dell'oracolo. Si concluderà presentando i risultati ottenuti individuando, poi, gli sviluppi futuri legati ad un possibile miglioramento dell'approccio utilizzato.

1 Introduzione

L'elaborato si propone come primo passo nella realizzazione del progetto CANTICO, presentato nel Capitolo 1. Il progetto punta a fornire una migliore esperienza agli spettatori dei teatri italiani e, per questa ragione, si vuole cercare di estrarre una serie di informazioni insite nel linguaggio naturale, rappresentato dai commenti di questi ultimi reperibili online. In questa ottica si vuole ottenere un'analisi dello stato dell'arte nell'ambito del Natural Language Processing, con una particolare attenzione alla Sentiment Analysis e all'Emotion Mining. Oltre ad una visione dello stato dell'arte, si vogliono individuare dei tool immediatamente utilizzabili e se ne vogliono testare le performance nell'ambito del contesto considerato, ossia quello teatrale. Tale contesto rappresenta un campo abbastanza particolare e richiede la creazione di un dataset che rifletta le sue peculiarità. I tool presi in considerazione sono stati integrati tra di loro in modo da avere un unico analizzatore che avviasse l'analisi complessiva del testo inserito come input.

La tesi si articola in 4 capitoli, di cui il primo è destinato alla descrizione del progetto CANTICO e come l'elaborato prevede di contribuire alla sua realizzazione. Il Capitolo 2 si occupa della presentazione dello stato dell'arte per quanto concerne la sentiment analysis e l'emotion mining, con una particolare attenzione alla descrizione dei tool attualmente disponibili. Segue una descrizione del processo che ha portato all'identificazione dei tool utilizzati in questo elaborato, cercando di sottolineare le motivazioni che hanno portato ad una determinata scelta. Nello stesso capitolo viene presentata l'architettura utilizzata dall'analizzatore costruito, definendo le modalità con cui i vari tool sono stati integrati tra di loro. L'ultimo capitolo, il quarto, è dedicato ai dati. In particolare si presenta il processo che ha portato alla costruzione del dataset, con la necessaria creazione di un oracolo da utilizzare per testare il l'analizzatore descritto nel capitolo 3. Vengono infine presentati i risultati ottenuti e le performance raggiunte dall'analizzatore rispetto a quelli dichiarati dai singoli tool. Seguono le conclusioni all'interno delle quali sono anche specificati i possibili sviluppi futuri del lavoro svolto.

2 Contesto

2.1 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 anche 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 nel 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 ATTRACT MORE AUDIENCE AND SPREAD OPERA AND PROSE THEATER BY USING IMMERSIVE TECHNOLOGIES"* 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 offrire 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 obiettivi programmatici, che vanno da uno studio preliminare degli strumenti disponibili per effettuare una personalizzazione dell'esperienza da parte dell'utente, fino all'obiettivo finale che corrisponde ad una sperimentazione della piattaforma

In particolare questo elaborato si inserisce nell'ambito del primo punto programmatico, che è suddiviso in diverse attività. 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. L'elaborato si concentra sull'applicazione di tali tecniche alle analisi del testo, con un semplice accenno ai possibili sviluppi delle stesse tecniche applicate all'analisi visuale degli utenti, ad esempio identificazione delle emozioni direttamente dall'immagine del viso di uno spettatore.

3 Stato dell'arte

3.1 Introduzione

In questo capitolo si presenterà uno studio sullo stato dell'arte su ciò che riguarda il Natural Language Processing, la Sentiment Analysis e l'Emotion Mining. Si cercherà di dare una rapida panoramica su ciò di cui si occupano teoricamente questi settori per poi procedere ad una analisi più 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.

3.2 NLP – Natural Language Processing

Il Natural Language Processing è un'area della ricerca che esplora come il linguaggio naturale immesso in un calcolatore può essere manipolato e trasformato in una forma che sia più utilizzabile per successive elaborazioni [1]. Questa definizione di NLP data nel 2010 non è più 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 simile 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.

Deng e Liu [2] hanno sottolineato, appunto, come lo sviluppo storico nel campo dell’NLP abbia seguito 3 ondate nelle ultime decadi. Si è partiti da un approccio basato sul razionalismo per poi passare ad un’analisi empirica del testo. Queste metodologie hanno ottenuto risultati convincenti in differenti applicazioni, ad esempio coi motori di ricerca e nell’estrazione della conoscenza, si è intensificata l’esigenza di migliorare ulteriormente le capacità di comprensione automatica di contenuti in linguaggio naturale, raggiungendo livelli simili 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. Questa secondo gli autori [2] rappresenta la terza ondata di sviluppo nel settore dell’NLP. La loro conclusione è, infatti, che la tecnologia del deep learning applicata all’NLP sia una rivoluzione dal punto di vista concettuale. I grandi progressi compiuti e gli importanti risultati raggiunti sarebbero stati inimmaginabili prima dell’avvento del deep learning.

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 simile 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.

3.3 Sentiment Analysis

La prima volta che la è stata definita la sentiment analysis è probabilmente per mano di Nasukawa e Yi [3] nel 2003 che indicano come problema principale di questa branca della ricerca l'individuazione della polarità positiva o negativa riguardo un determinato soggetto. Questa definizione è stata poi integrata da Dave e altri [4] che hanno indicato lo stesso concetto definendolo come opinion mining. Nel corso del tempo le due definizioni sono state utilizzate indifferentemente, come ci suggerisce Liu [5], ma in questo elaborato si farà riferimento all'analisi di polarità come sentiment analysis. Questa 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 utilizzata 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 mantenere 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 ciò 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 una 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 utilizzare il modello 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 utilizzare 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 [6] il riaddestramento dei modelli su un nuovo contesto per effettuare la sentiment analysis è necessario ma non rappresenta, in ogni caso, uno

strumento per garantire un grande miglioramento dell'accuratezza durante la predizione.

3.4 Emotion Mining

Liu [5] identifica in [7] il primo approccio alla sentiment analysis con l'aggiunta di una polarità neutra, e con l'aggiunta di una classificazione delle emozioni a partire da queste polarità. In questo articolo [7] non vi è una definizione di emotion mining quanto piuttosto di emotion classification ma risulta evidente come questa branca si sia sviluppata sempre nell'ambito della sentiment analysis e per questo la si potrebbe considerare come una sua evoluzione, che ha però uno scopo diverso. Questo punto di vista non rappresenta dettagliatamente 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 mining. 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.

3.5 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.

3.5.1 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

Tool	Lingue supportate	Italiano	Documentazione	Tecnica	Utilizzato
Stanford CoreNLP	Inglese, arabo, cinese, francese, tedesco, spagnolo	No	Disponibile una Api completa	Rete neurale	[6], [9]
TINT	Italiano	Si	Sono disponibili diversi tutorial	Stanford CoreNLP	[8], [10]

TABELLA 1 TOOL PER L'ANALISI TESTUALE

team di sviluppo, composto da Palmeto e Moretti [8], 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 Stanford CoreNLP per adattarla all'italiano, una delle lingue non supportate dalla Stanford CoreNLP.

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 [8].

In Figura 1 [10] 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



FIGURA 1 ESEMPIO ANALISI TINT [10]

Moretti [8] 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 semplicità d'uso per fungere da base di 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 più specifiche all'ambito di ricerca.

3.5.2 Sentiment Analysis

Ci sono numerose soluzioni commerciali che integrano numerose funzionalità aggiuntive a quella della semplice sentiment analysis. Diverse soluzioni commerciali sono necessariamente rivolte principalmente ad un mercato business e tentano di offrire un servizio più 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 più veloce per ottenere un'analisi rapida e costante della loro percezione presso il pubblico. Questo perché vengono presentate dashboard con delle analisi dei dati già fatte che possono essere già utilizzate dall'azienda così 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 società ed altre come MeanigCloud e GetSentiment utilizzano una metodologia diversa, esponendo un'api che può essere interrogata con qualsiasi tipo di testo. E tipicamente questi servizi forniscono analisi molto approfondite. La Figura 2, estratta dalla demo di

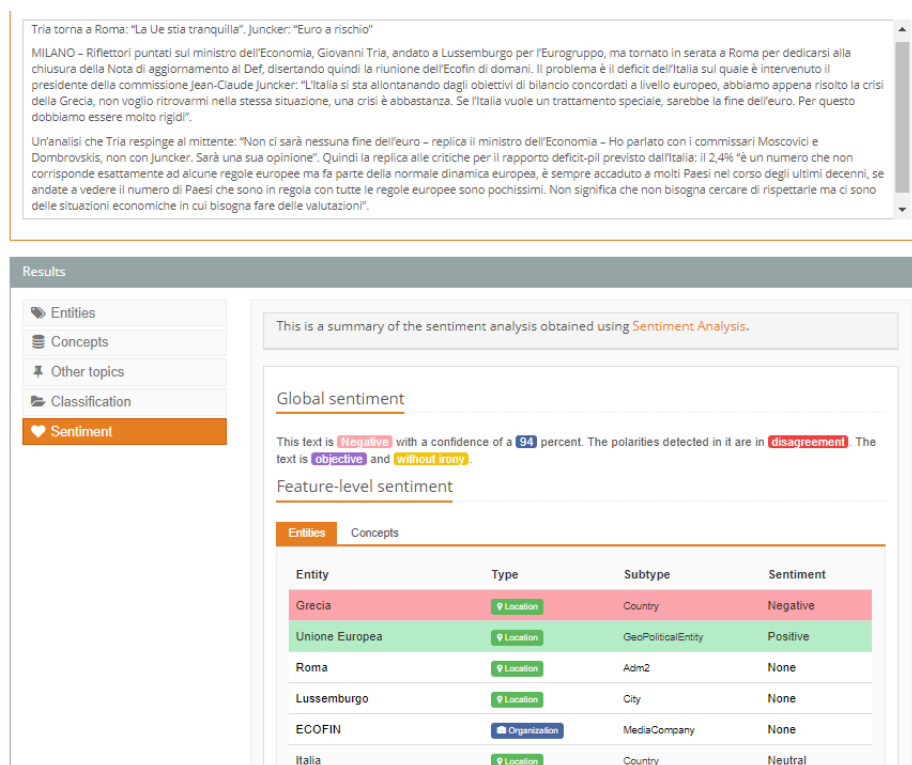


FIGURA 2 ESEMPIO ANALISI MEANINGCLOUD [11]

Sentiment Analysis Demo

Language	it	Sentiment	-0.8596933104941142
----------	----	-----------	---------------------

Which language is your text in:

Italian (Italiano)

Tria torna a Roma: "La Ue stia tranquilla". Juncker: "Euro a rischio"

MILANO – Riflettori puntati sul ministro dell'Economia, Giovanni Tria, andato a Lussemburgo per l'Eurogruppo, ma tornato in serata a Roma per dedicarsi alla chiusura della Nota di aggiornamento al Def, disertando quindi la riunione dell'Ecofin di domani. Il problema è il deficit dell'Italia sul quale è intervenuto il presidente della commissione Jean-Claude Juncker: "L'Italia si sta allontanando dagli obiettivi di bilancio concordati a livello europeo, abbiamo appena risolto la crisi della Grecia, non voglio ritrovarmi nella stessa situazione, una crisi è abbastanza. Se l'Italia vuole un trattamento speciale, sarebbe la fine dell'euro. Per questo dobbiamo essere molto rigidi".

Un'analisi che Tria respinge al mittente: "Non ci sarà nessuna fine dell'euro – replica il ministro dell'Economia – Ho parlato con i commissari Moscovici e Dombrovskis, non con Juncker. Sarà una sua opinione". Quindi la replica alle critiche per il rapporto deficit-pil previsto dall'Italia: il 2,4% "è un numero che non corrisponde esattamente ad alcune regole europee ma fa parte della normale dinamica europea, è sempre accaduto a molti Paesi nel corso degli ultimi decenni, se andate a vedere il numero di Paesi che sono in regola con tutte le regole europee sono

FIGURA 3 ESEMPIO ANALISI REPUSTATE [12]

Menaningcloud [11], consente di vedere il livello di dettaglio della sentiment analysis che non si limita ad effettuare analisi solo sulle singole frasi ma estrae dei concetti e poi analizza questi. Tale livello di dettaglio può 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.

Lo stesso tipo di servizio è offerto ad esempio da Repustate. In Figura 3 si vede come il servizio messo a disposizione da Repustate [12] permette di ottenere la risposta e sostanzialmente consente di ricavare il tipo di sentiment espresso nel 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 riferirà a questa nei successivi paragrafi.

Uno dei tool più 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 funzionalità leggermente

ridotte, una seconda che opera su java che può 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 Analysis. È possibile lavorare su qualsiasi sistema operativo, e viene fornita una API consultabile online per capire velocemente come funziona. L'installazione risulta semplice e 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 realtà, 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 può essere utilizzata come una libreria esterna, essendo fornita sotto forma di Jar integrabile in un progetto java. La 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 polarità, 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 già tutto il necessario per avviare il tool. La parte di analisi di polarità, pySenti4SD, lavora con python ed è semplice da utilizzare nella versione docker velocissima da configurare e avviare, la 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 verrà analizzato nel paragrafo successivo.

Accanto a questi progetti più grandi in rete è possibile reperire numerosi progetti sviluppati da singoli utenti. Ad esempio il Sentiment Analysis on Tweets [13], 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 possibilità 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 [14] 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 funzionalità ma non è presente una documentazione completa attualmente. Allo stesso tempo non risulta più 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 [15]. 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 modalità con cui opera, è stato scelto perché in grado di operare sulla lingua italiana, cosa fondamentale per il tipo di attività 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 Tabella 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 polarità 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	[6]
NLTK	Inglese	No	Basata su regole	[6]
Stanford CoreNLP (SentimentAnnotation)	Inglese, arabo, cinese, francese, tedesco, spagnolo	No	Rete neurale	[6], [9]
EMTK (pySenti4SD)	Inglese	No	Caratteristiche lessicali	[6]
Sentiment Analysis on Tweets	Inglese	No	Varie Reti neurali	[13]
Sentiment Classifier	Inglese	No	NLTK	[14]
SentIta	Italiano	Si	Rete neurale CNN	[15]

TABELLA 2 TOOLS PER LA SENTIMENT ANALYSIS

3.5.3 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 vastità 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 [16] 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. La api proposta dal tool viene offerta sulla base dei



FIGURA 4 ESEMPIO ANALISI TWINWORD [16]

Try our free demo now by typing a sentence or choose from the options in the drop-down menu.

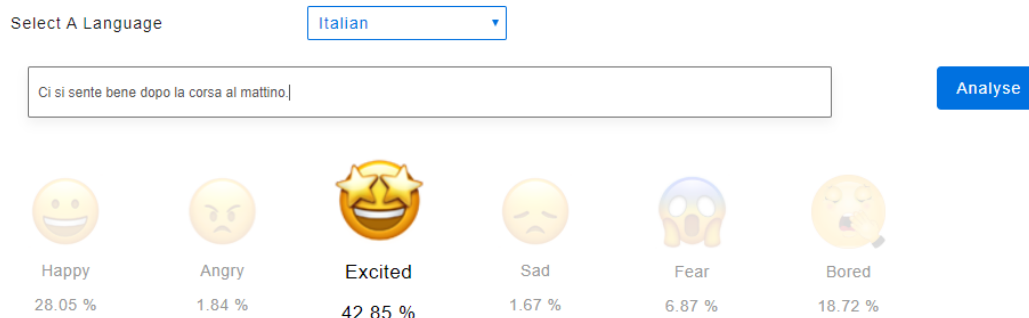


FIGURA 5 ESEMPIO PARALLELDOTS [17]

testi inviati analizzati al mese, in base al numero disponibile aumenta il prezzo dell'abbonamento.

Un secondo esempio di tool a pagamento analizzato è quello sviluppato da ParallelDots, visibile in Figura 5 [17]. 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 [17] 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 però sono risultate molto più 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 più 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 difficoltà 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 necessità 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 [18]. 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 [18] negli otto

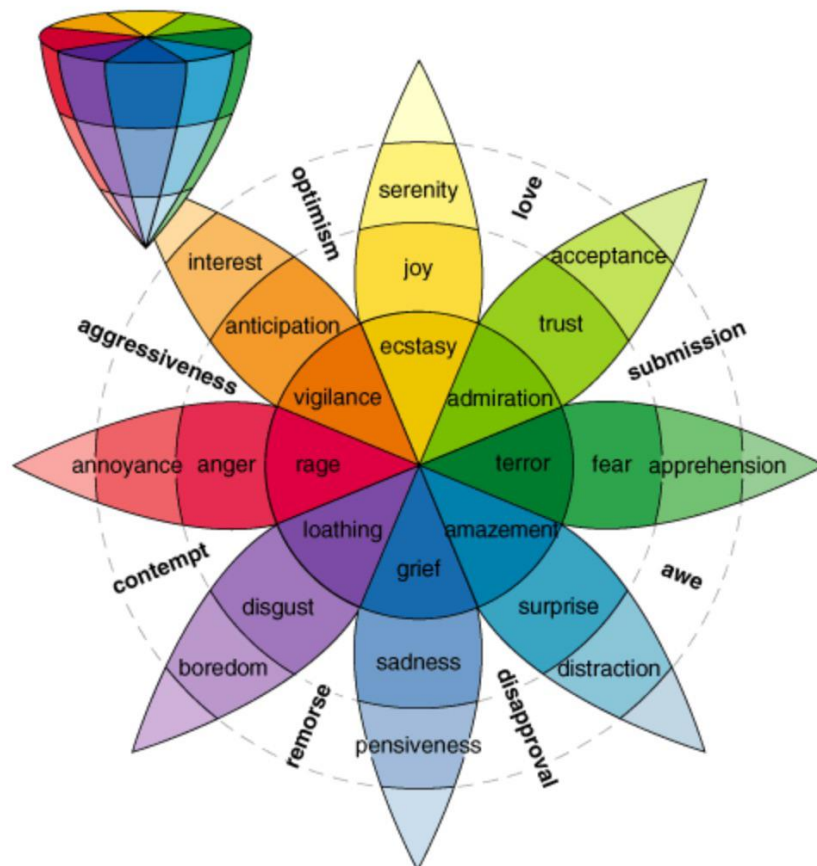


FIGURA 6 RUOTA DELLE EMOZIONI [18]

Tool	Lingue supportate	Documentazione	Tecnica	Utilizzato
EmoTxt (EMTk)	Inglese	Scarsa	Caratteristiche lessicali	[6],
EmoNet	Inglese	Scarsa	Rete neurale ricorrente	[18]

TABELLA 3 TOOL PER L'EMOTION MINING

settori della ruota delle emozioni di Plutchik mostrata in Figura 6. Il tool lavora con python e sfrutta alcune famose e facili da reperire librerie python, garantendo la massima compatibilità e facilità d'uso.

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 sarà 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.

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

3.6 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 più 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 più 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 [19], 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 [20], 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.

4 Analizzatore

4.1 Introduzione

In questo capitolo si vuole presentare l'analizzatore costruito. Si partirà 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.

4.2 Identificazione dei tool

A partire dall'analisi dello stato dell'arte effettuate nel capitolo precedente, si sono individuati i tool più 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. Questo rappresenta un progetto interno del ministero dell'istruzione italiano, e mira esplicitamente a migliorare l'esperienza del pubblico presso i teatri italiani, così come descritto nel capitolo 1. Per questi motivi si è data particolare attenzione alle lingue supportate dai vari tool individuati. Il progetto, infatti, sarà sviluppato con il fine di operare con la lingua italiana 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, però, come per la totalità 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 facilità 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 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 [15]. 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. Durante 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 affascinante e apparentemente fattibile, questo modus operandi si rivela secondo Lin e altri [6] ampiamente deficitario in termini di performance e accuratezza, sebbene il loro lavoro avesse degli obiettivi 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 ciò 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 [8], 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 totalità 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 più efficaci e sempre più 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, più 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 polarità 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 più 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 può 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 funzionalità 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 funzionalità

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 modalità 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 necessità di tradurre il testo dall'italiano all'inglese per poter operare sui tool di emotion mining, si è deciso di sfruttare questa necessità per effettuare un ulteriore controllo per quanto concerne il calcolo della polarità. Per Questa ragione si è stati costretti ad individuare un tool aggiuntivo per completare l'analisi. In virtù 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 [6], 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.

4.3 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 più 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 semplicità e facilità 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 funzionalità della libreria. La Stanford CoreNLP, in queste particolari configurazioni, viene avviata come server locale e può 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.

4.4 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

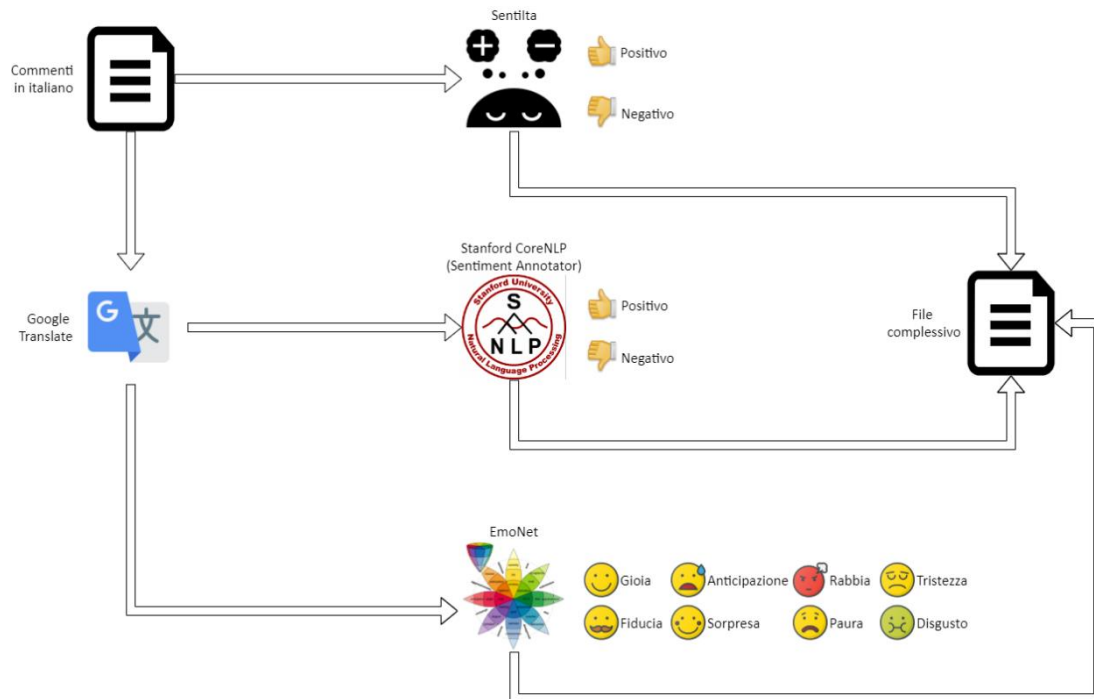


FIGURA 7 ARCHITETTURA DEL TOOL REALIZZATO

le recensioni delle opere teatrali analizzati, verrà poi descritto nel capitolo successivo come è stato costruito il dataset utilizzato. L'utilizzo dei file di tipo csv ha consentito una più veloce integrazione con i fogli di calcolo che sono stati utilizzati per taggare il dataset manualmente e per effettuare le analisi desiderate.

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 polarità, 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 polarità. 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. Oltre 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.

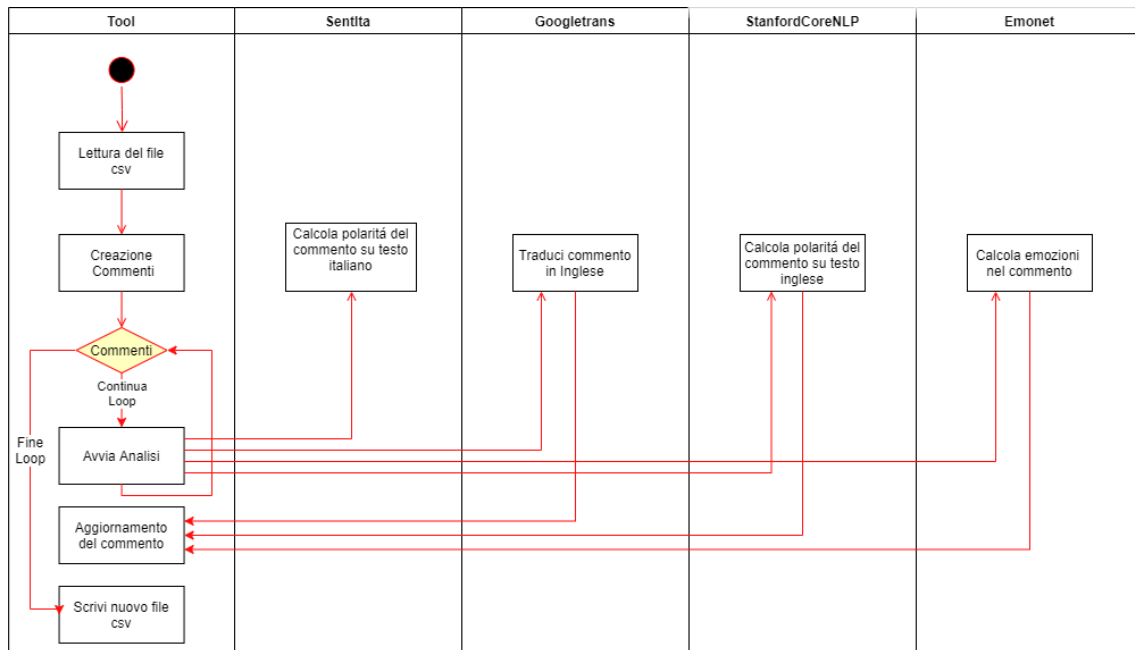


FIGURA 8 ACTIVITY DIAGRAM DEL TOOL REALIZZATO

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 polarità 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 polarità, 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 sarà dettagliano nel capitolo successivo, previa descrizione del dataset utilizzato e della sua costruzione.

5 Analisi dei dati

5.1 Introduzione

In questo capitolo si procederà a presentare il dataset utilizzato, descrivendone il processo di creazione. Saranno poi discussi i risultati ottenuti facendo analizzare il dataset dai tool considerati.

5.2 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'impossibilità di trovare un dataset con queste caratteristiche, e per questa ragione, si è deciso di procedere alla costruzione di un dataset più rappresentativo. La ricerca di commenti e recensioni di spettacoli teatrali si è rivelata particolarmente ardua per una serie di ragioni. Prima fra tutte la scarsità 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 difficoltà nel reperire commenti realmente utilizzabili. Ciò 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 qualità 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. Ciò che si cerca di fare, in questo caso infatti, è quello di ricavare maggiori informazioni da ciò 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 più 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 più sintetiche e dirette delle classiche recensioni di un critico del settore. In questo caso è proprio il loro enorme numero e la loro varietà 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 però 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.

5.3 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

Definizione della polarità su 5 lvl		Indicazione della presenza dell'emozione							
Polarità		Rabbia	Anticipazione	Disgusto	Paura	Gioia	Tristezza	Sorpresa	Fiducia
3	n	n	n	n	y	y	n	n	n
				n	n	y	n	n	n
				n	n	y	n	y	y

Polarità

Selezionare la polarità desiderata:

0 Negativo

1 Abbastanza negativo

2 Neutro

3 Abbastanza positivo

4 Positivo

FIGURA 9 ESEMPIO DATASET TAGGATO

di procedere classificando la polarità del commento e individuando la presenza o l'assenza delle 8 emozioni principali di Plutchik. Per facilitare l'analisi successiva la polarità è stata valutata sulla base di una scala a 5 livelli, dal più basso Negativo al più 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.

Inizialmente infatti il tagging individuale ha portato ad una discrepanza tra i due dataset ottenuti, si discuterà a breve di questa ulteriore difficoltà strettamente legata all'uomo. In particolare si è notata una discrepanza di circa il 13% per quanto concerne la definizione della polarità 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 polarità, che è stata effettuata collassando i valori positivi e negativi della scala a 5 livelli, generando un notevole miglioramento nella percentuale d'intesa tra i due

	Polarità		Emozioni							
	Intesa 5 lvl:	Intesa 3 lvl:	Rabbia	Anticipazione	Disgusto	Paura	Gioia	Tristezza	Sorpresa	Fiducia
Concordi	233	255	258	254	261	266	248	251	242	253
Totale	268	268	268	268	268	268	268	268	268	268
Differenza	35	13	10	14	7	2	20	17	26	15
	86,94%	95,15%	96,27%	94,78%	97,39%	99,25%	92,54%	93,66%	90,30%	94,40%
			Media emozioni	94,82%						

FIGURA 10 INTESA NELLA DEFINIZIONE DELL'ORACOLO

	Intesa polarità	Intesa emozioni							
		Rabbia	Anticipazione	Disgusto	Paura	Gioia	Tristezza	Sorpresa	Fiducia
Concordi	174	231	148	229	256	209	218	98	125
Totale	268	268	268	268	268	268	268	268	268
Differenza	94	37	120	39	12	59	50	170	143
Agreement	64,93%	86,19%	55,22%	85,45%	95,52%	77,99%	81,34%	36,57%	46,64%
		Media totale emozioni	70,62%						

FIGURA 11 ANALISI PERCEZIONE TRA 3 PERSONE

soggetti. Già 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 ciò 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 più oggettiva possibile il testo.

Ad avvalorare la tesi secondo cui vi sono notevoli difficoltà d'intesa già tra esseri umani, si è esteso il confronto ad una terza persona, che non ha però partecipato alla definizione dell'oracolo, ma che è stata utilizzata per analizzare la variabilità 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 più 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 più significativo per descrivere la viabilità dell'argomento che si sta trattando in questo elaborato. Prendendo in esame, invece, solo le emozioni si nota come l'unanimità si ottenga in circa il 6% del dataset, mentre considerando solo la polarità classificata su 5 valori si raggiungono risultati nell'ordine del 64%. I dati mostrati in Figura 11, confermano come sia molto basso il livello d'intesa tra i soggetti che hanno proceduto al tagging, e si può immaginare che aumentando il numero di persone le percentuali d'intesa 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 più 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 città. Questa tristezza, però, 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 dibattito 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 polarità 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

	Polarità su 5 livelli					Polarità su 3 livelli		
	P0	P1	P2	P3	P4	P1	P2	P3
N	20	11	10	48	179	31	10	227
%	7,46%	4,1%	3,73%	17,91%	66,79%	11,57%	3,73%	84,70%

TABELLA 4 POLARITÀ DEL DATASET UTILIZZATO

	Rabbia	Anticipazione	Disgusto	Paura	Gioia	Tristezza	Sorpresa	Fiducia
N	23	16	3	0	200	71	105	91
%	8,58%	5,97%	1,12%	0%	74,63%	26,49%	39,1%	33,96%

TABELLA 5 DEFINIZIONE DELLE EMOZIONI DEL DATASET

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 doverosa disamina per la definizione del dataset, analizziamo brevemente come è caratterizzato l'oracolo. Nella Tabella 4 si mostrano tutte le informazioni relative alla polarizzazione del dataset secondo l'oracolo, si noti la doppia rappresentazione sia su una scala a 5 livelli, sia su una scala a 3 livelli che è stata poi utilizzata per il confronto con i tool. In Tabella 5, invece, si riportano tutti i dati relativi alle emozioni riscontrate nel dataset. Si noti come il dataset non sia, purtroppo, uniformemente distribuito rispetto alla totalità delle emozioni. Questo è da imputare in parte alla mancanza di dati su cui costruire il dataset stesso, come discusso precedentemente. Allo stesso tempo l'argomento stesso in esame, le rappresentazioni teatrali, poco si contraddistingue per la presenza di emozioni quali la paura o il disgusto.

5.4 Metriche d'analisi

Al fine di valutare l'efficienza dell'architettura costruite si è deciso di estendere l'analisi rispetto al semplice d'intesa tra tool e oracolo, cosa che invece è stata considerata fino ad ora per confrontare i dati tra persone diverse. In particolare si vuole cercare di migliorare l'analisi utilizzando un approccio statistico. In particolare si utilizzeranno le metriche di precisione e recupero, o richiamo, a cui si farà riferimento con i corrispettivi inglesi di precision e recall. Queste sono due comuni classificazioni statistiche, utilizzate

in diversi ambiti del sapere, come ad esempio l'information retrieval. La precision può essere vista come una misura di esattezza o fedeltà, mentre la recall è una misura di completezza. In particolare la precision risponde alla domanda: “quanti elementi individuati sono rilevanti?”, la recall, invece, risponde alla domanda: “quanti elementi rilevanti sono stati selezionati?”. Per poter ottenere queste informazioni sono state costruite delle matrici di confusione che consentissero di calcolarle. Ogni colonna della matrice rappresenta i valori reali mentre ogni riga rappresenta i valori predetti dal tool, andando a costruire una matrice con 4 valori, i veri positivi, i veri negativi, i falsi positivi e i falsi negativi. La combinazione di questi valori consente di calcolare i valori di precision e recall. La terza metrica che si è presa in considerazione mette in correlazione la precision e la recall ed è la F-measure o F-score. Questa va a misurare l'accuratezza di un test e viene calcolata attraverso la media armonica di precision e recall, e può assumere valori che vanno da 0 a 1, più il valore è prossimo ad 1 più il test si può definire accurato.

5.5 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 più ad un livello più fine di dettaglio, ma che a volte può risultare in una informazione ricca di rumore e per questo si è reso necessario effettuare delle semplificazioni. Si partirà analizzando, prima di tutto la polarità 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 polarità su una scala a 5 livelli, come quella utilizzata nell'oracolo. A partire da questo livello di dettaglio si è dovuta operare una scelta al fine di ridurre il risultato dell'analisi ad un numero confrontabile. Per questa ragione si è deciso di valutare la maggioranza, tra positivi e negativi, tra i valori forniti dalla Stanford CoreNLP per tutte le frasi del commento. Tale scelta è stata, inoltre, integrata da un collasso della scala a 5 livelli a favore di una scala con soli 3 livelli: positivo, neutro e negativo. Tale scelta è motivata dalla necessità di avere un confronto meno soggetto a bias. SentIta,

invece, espone i propri risultati definendo un valore tra 0 e 1 di positività del testo fornito in input, e un valore tra 0 e 1 per definire la negatività, entrambi i valori sono stati poi ricondotti su base 100 per migliorare la leggibilità. A partire da questi valori si è calcolata la differenza tra polarità positiva e negativa in modo da avere una distribuzione di dati su cui si è fatta un'analisi statistica ottenendo il valore del secondo quartile. A partire da quel valore e dalla differenza tra polarità positiva e negativa, si è ricavato il corrispettivo valore su una scala a 3 livelli, esattamente come fatto per la Stanford CoreNLP. A partire da questi dati è stato fatto un primo confronto semplicemente confrontando quanto i due tool considerati dessero la medesima valutazione. Sul dataset composto da 268 commenti i due tool sono risultati in accordo su 135 commenti, risultato in una intesa pari al 50,37%. Questo dato di per sé è poco significativo e per questo non verrà incluso nell'analisi statistica del prossimo paragrafo.

Per quanto concerne, invece, l'emotion mining, si ricorda che si è utilizzato il tool EmoNet. Questo 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 più semplice rappresentazione presente nell'oracolo. Si è tentato di approssimare il problema nella maniera più semplice imponendo una soglia minima entro cui considerare presente una tale emozione all'interno dello specifico commento. Tale approccio, inizialmente, è stato implementato attraverso l'applicazione di una soglia comune a tutte le emozioni, ottenendo risultati poco soddisfacenti per via della distribuzione non uniforme del dataset. Per questo motivo si è condotta un'analisi statistica sulla distribuzione dei dati ottenuti dal tool sulle singole emozioni. A partire da quest'analisi si sono definite delle soglie variabili per ogni emozione. Si è deciso di testare i risultati per 3 soglie in particolare corrispondenti ai valori del primo, del secondo e del terzo quartile.

5.6 Analisi dei dati sperimentali

5.6.1 Analisi della polarità

L'analisi della polarità 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 3 livelli che fosse rappresentativo dell'intero commento. Il confronto ha mostrato come i due tool concordino circa nel 50% dei casi. Per quanto riguarda invece il confronto con i dati contenuti nell'oracolo, si è provveduto alla costruzione delle matrici di confusione per due polarità, positiva e negativa, per entrambi i tool. A partire da queste ultime si sono calcolate precision, recall e f-measure per ogni matrice ottenendo i risultati mostrati nelle figure dalla 12 alla 17. Si può notare come sia per entrambi i tool presi in considerazione si riescano ad ottenere dati positivi, con un f-measure che per entrambi i tool supera il valore di 0,9 nella capacità di riconoscere la polarità positiva. Vi è una lieve flessione per quanto riguarda l'analisi della polarità negativa. Questi dati confermano come l'analisi di polarità sia efficiente anche attraverso l'utilizzo di strumenti di traduzione come Google Translate. Si nota un valore leggermente inferiore per quanto concerne la recall calcolata sul tool Stanford CoreNLP nella valutazione della polarità positiva. I dati risultano quindi molto soddisfacenti e consentono di confermare che i tool utilizzati possono essere applicati al contesto di riferimento.

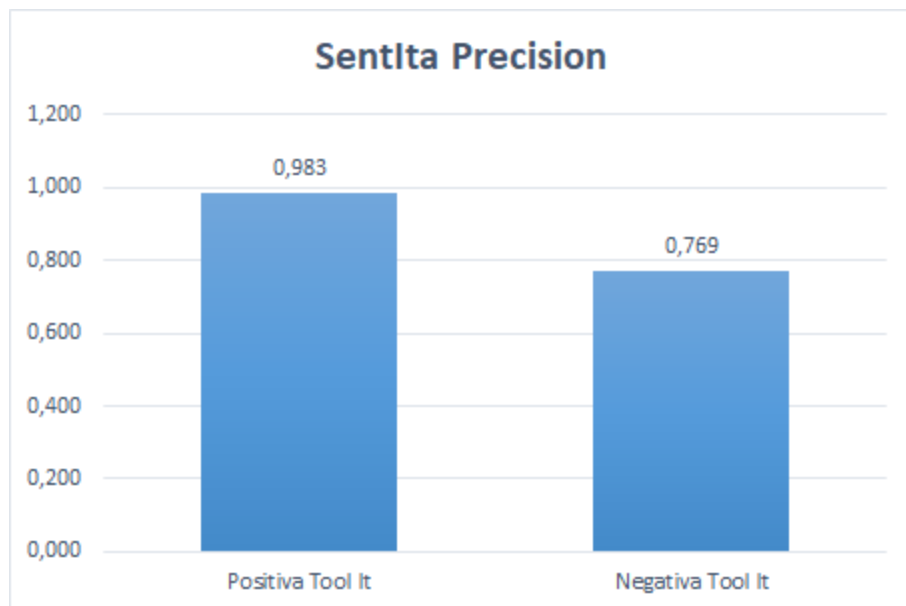


FIGURA 12 PRECISION SENTITA

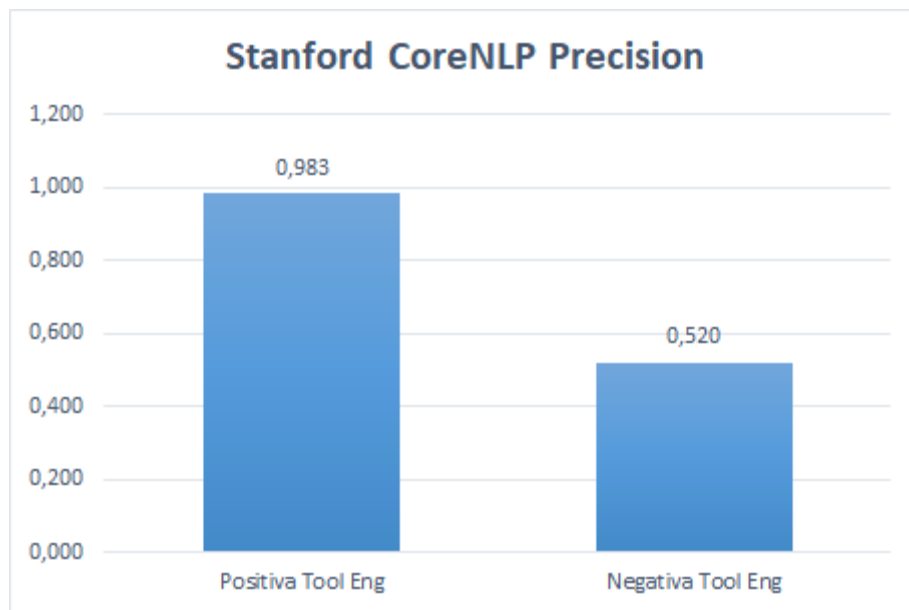


FIGURA 13 PRECISION STANFORD CORENLP

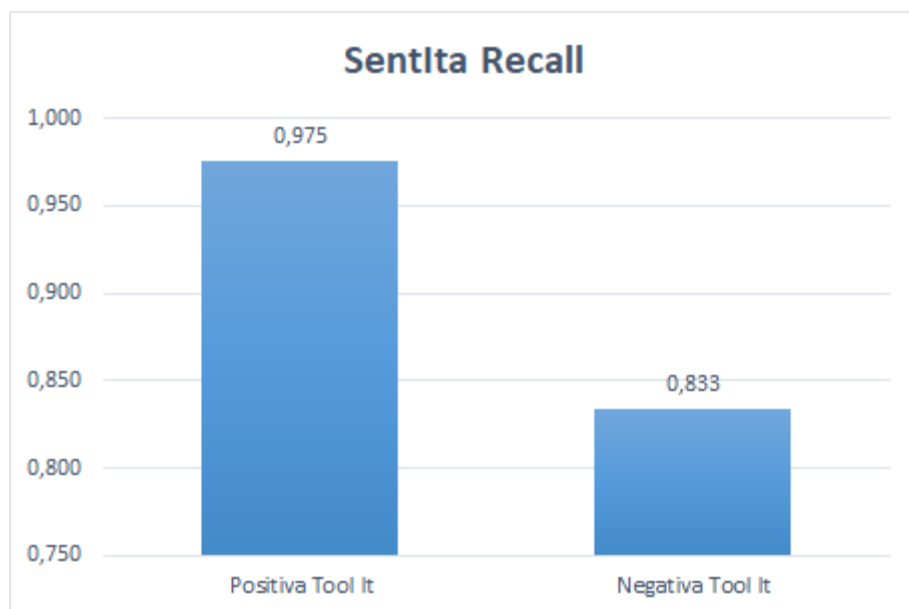


FIGURA 14 RECALL SENTITA

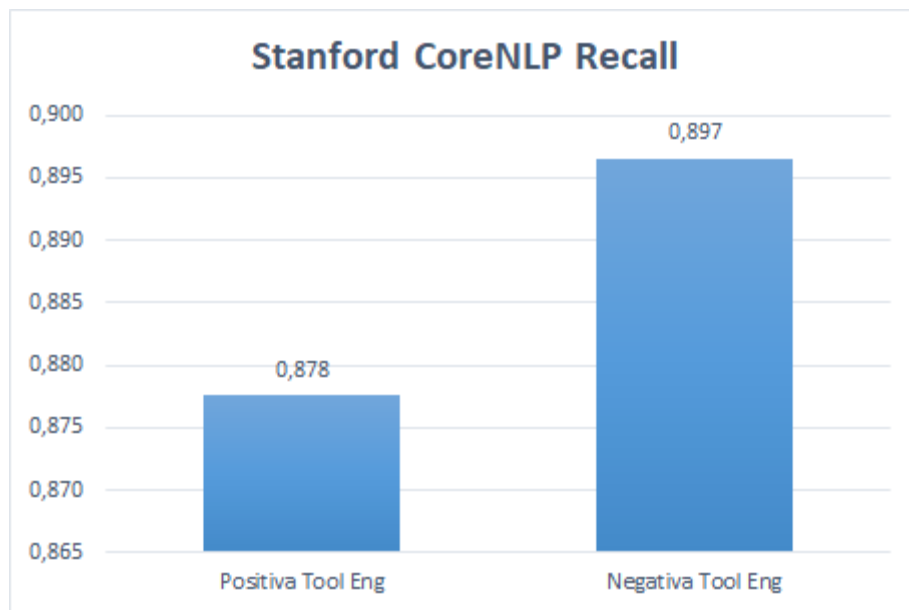


FIGURA 15 RECALL STANFORD CORENLP

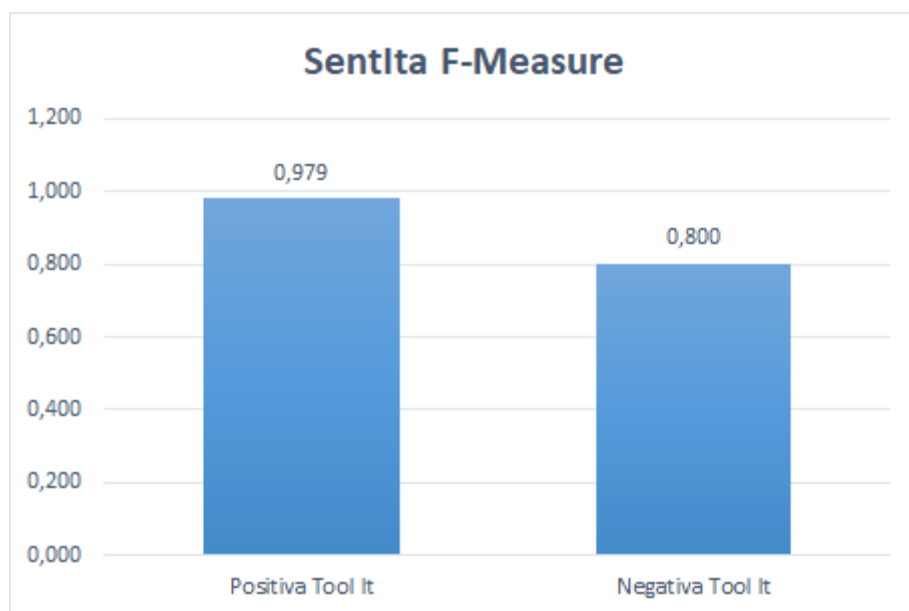


FIGURA 16 F-MEASURE SENTITA

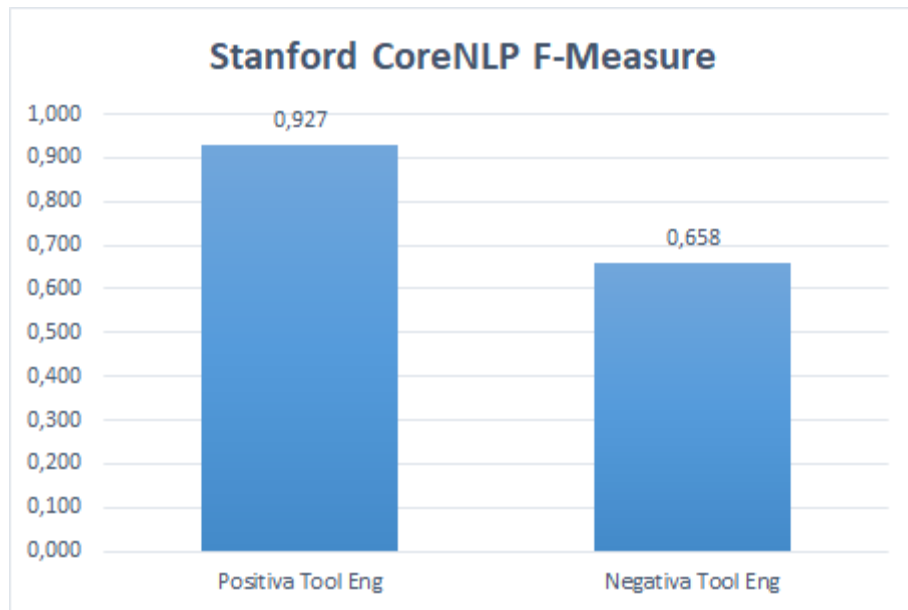


FIGURA 17 F-MEASURE STANFORD CORENLP

5.6.2 Analisi delle emozioni

Il tool utilizzato per l'analisi di emotion mining, ovvero EmoNet, si è scontrato con una serie di difficoltà che hanno sicuramente influito sulle sue performance, ma prima di esaminarle nel dettaglio si vogliono presentare i risultati ottenuti. Le analisi sono state presentate per ogni emozione, dal momento che il tool di riferimento fornisce sempre un dato su tutte le 8 emozioni principali prese in esame. Nella Figura 18 si mostrano i

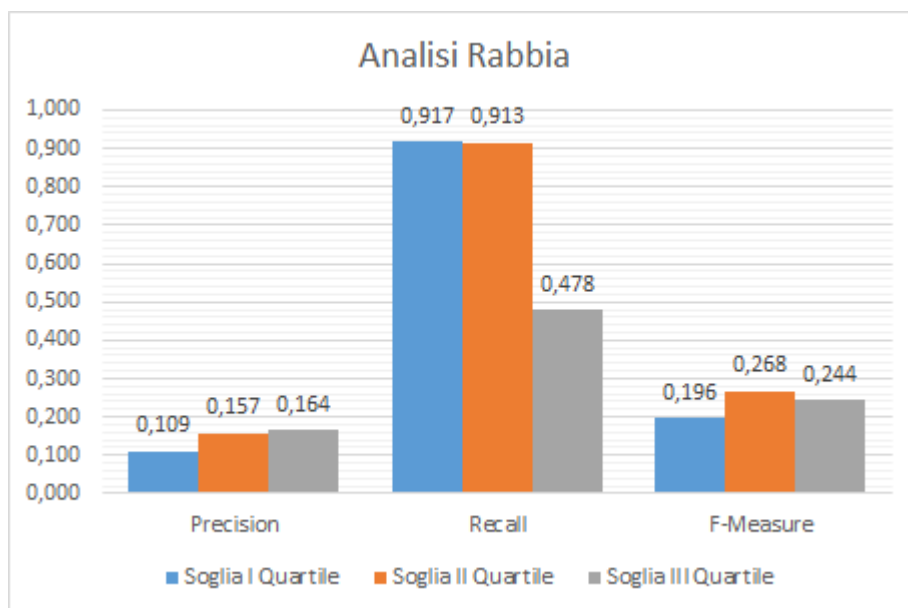


FIGURA 18 ANALISI RABBIA PER QUARTILE

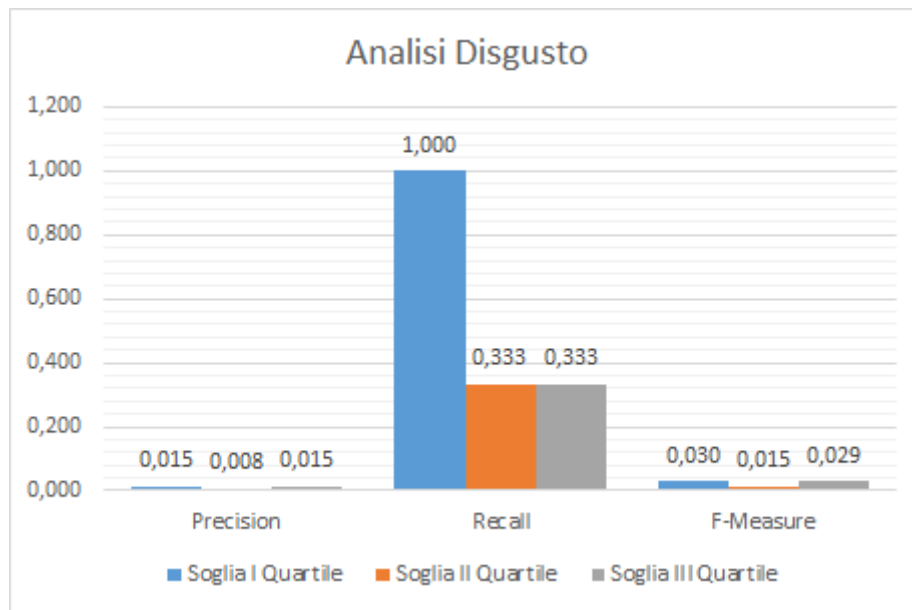


FIGURA 19 ANALISI DISGUSTO PER QUARTILE

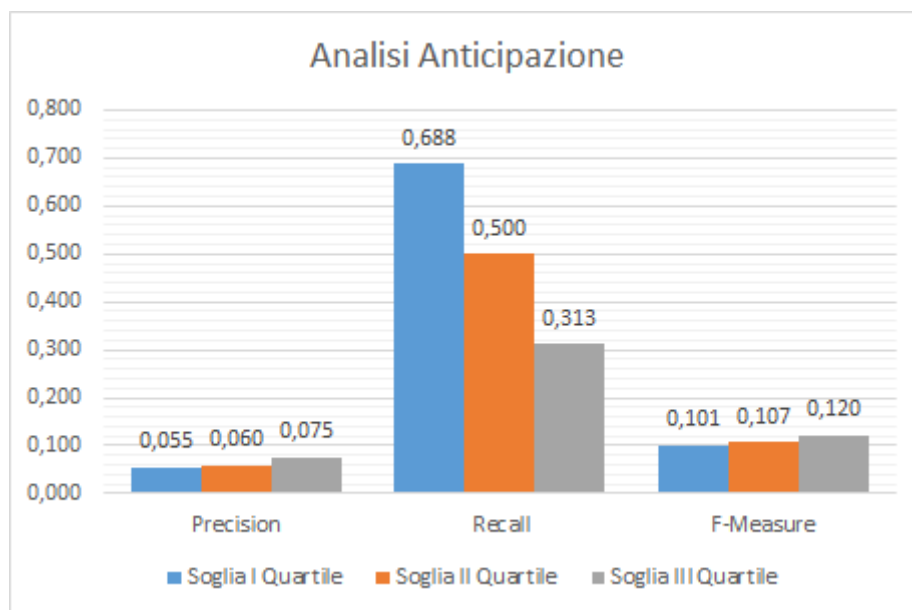


FIGURA 20 ANALISI ANTICIPAZIONE PER QUARTILE

risultati delle analisi effettuate sulla rabbia, mentre la 19 e la 20 presentano i risultati di disgusto e anticipazione. Il valore più significativo è quello dell’F-Measure pari, nel migliore dei casi, a 0,26. Tale risultato è raggiunto dalla rabbia, mentre il disgusto e l’anticipazione ottengono risultati inferiori. Per quanto concerne la paura non sono state prese in esame delle misure dal momento che il dataset utilizzato per il test non conteneva alcuna istanza di questo sentimento.

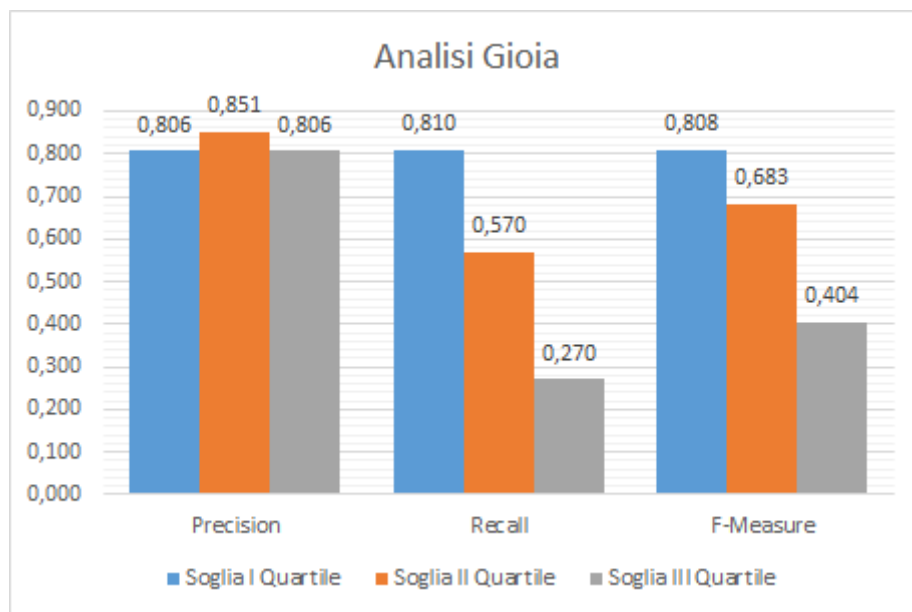


FIGURA 21 ANALISI GIOIA PER QUARTILE

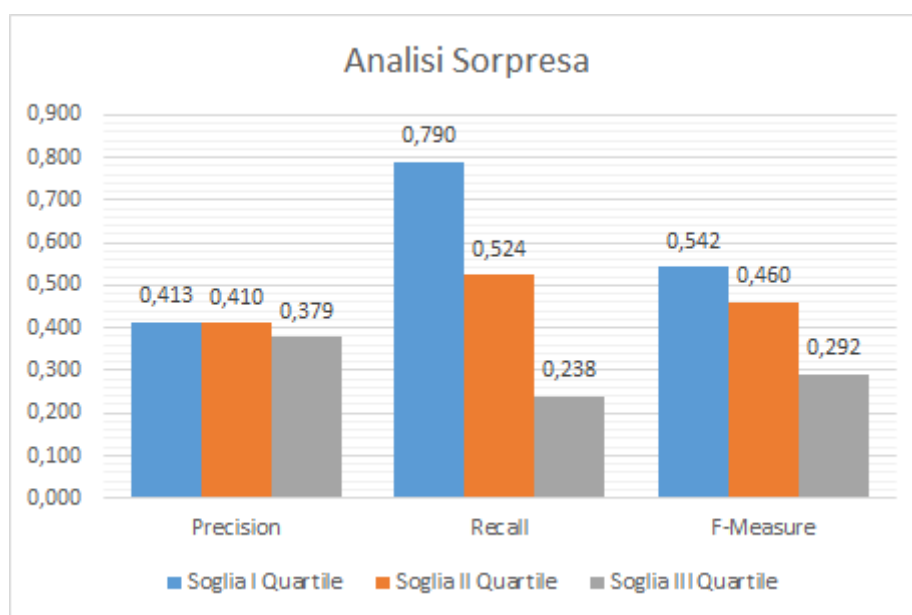


FIGURA 22 ANALISI SORPRESA PER QUARTILE

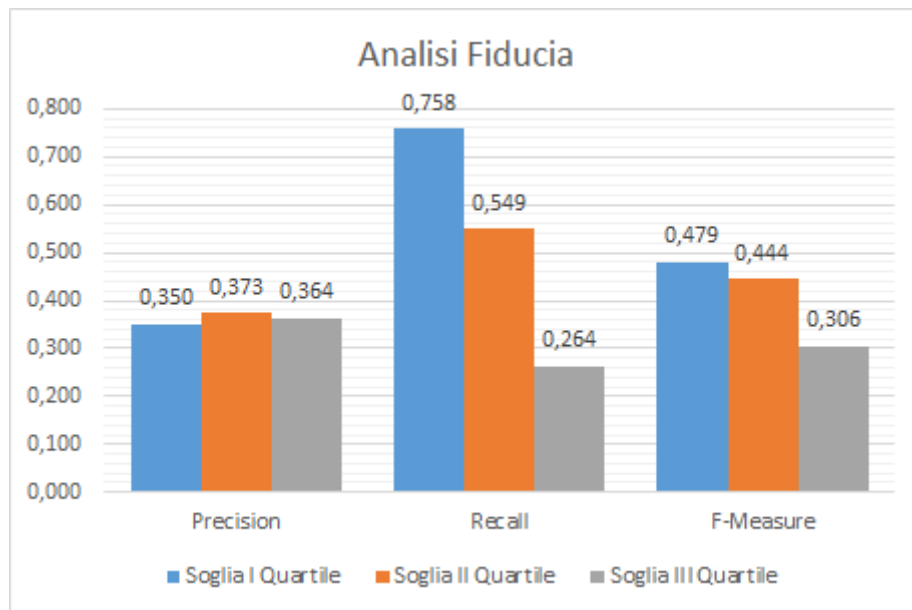


FIGURA 23 ANALISI FIDUCIA PER QUARTILE

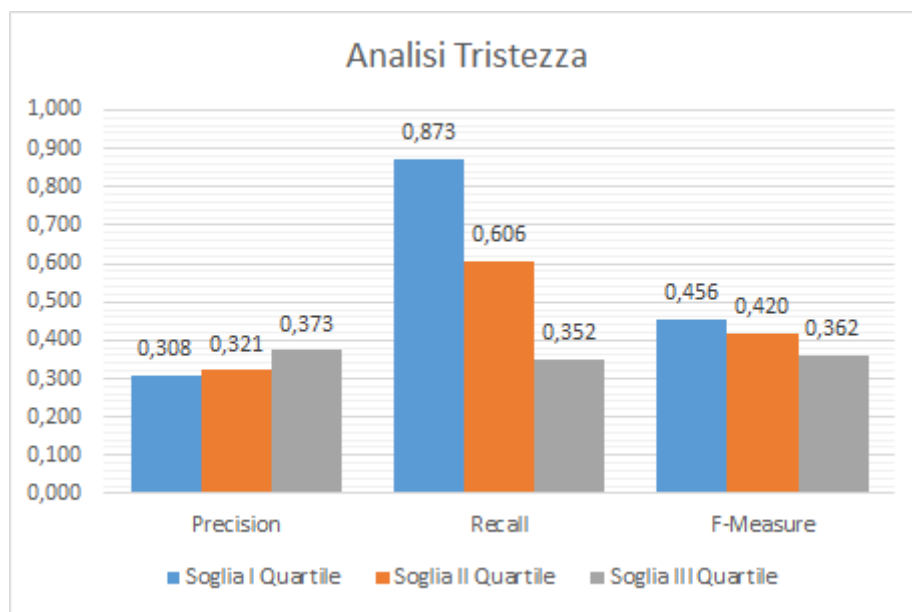


FIGURA 24 ANALISI TRISTEZZA PER QUARTILE

Nelle figure dalla 21 alla 24 si mostrano i risultati ottenuti da gioia, sorpresa, fiducia e tristezza che riescono a raggiungere delle buone performance, facendo registrare un valore di F-Measure tra 0,80 e 0,45. Questi dati denotano delle performance del tool abbastanza scarse, soprattutto se paragonate ai risultati ottenuti dal team di EmoNet [18]. Il tool infatti si presenta con un F-Measure superiore allo 0,9 per ogni emozione considerata, ma i valori così bassi riscontrati in questo elaborato possono essere legati ad una serie di fattori. Prima di tutto il contesto in si sta operando, come presentato

infatti da Lin e altri [6] utilizzare dei tool basati sul machine learning su contesti diversi da quelli su cui si sono addestrati ne peggiora notevolmente le performance. EmoNet è stato addestrato su un dataset molto vasto ma che era basato su tweet generici che quindi poco condividevano con il contesto teatrale su cui si è basato questo elaborato. Il secondo importante fattore da considerare è legato alla lunghezza stessa dei tweet che hanno costituito il dataset con cui EmoNet è stato addestrato. Questi infatti sono stati selezionati per essere più lunghi di 5 parole almeno, ma mediamente risultano essere molto più brevi delle recensioni di cui è composto il dataset realizzato per il progetto. Questo fattore ha sicuramente pregiudicato notevolmente le performance del tool nel nuovo contesto operativo. La fase di training, inoltre, potrebbe essere stata completata in maniera più esaustiva rispetto al riconoscimento di alcune emozioni rispetto ad altre. Esaminandolo, infatti, si nota come la presenza di tweet taggati con l'emozione di gioia, costituisca circa il 41% del dataset. Questo può aver fornito una capacità del tool di riconoscere meglio alcune emozioni, la gioia in particolare, rispetto alle altre. L'ultimo, ma forse più importante, fattore che ha contribuito alla perdita di performance del tool è il fatto di avere utilizzato un tool che lavora in lingua inglese e quindi essere stati costretti ad effettuare una traduzione del commento originale. Se questo fattore ha influito in maniera poco incisiva nell'ambito dell'analisi di polarità, in questo caso l'impatto è stato sicuramente più significativo. Questo risultato non deve sorprendere, si sta infatti facendo un'analisi più complessa in questo caso, di conseguenza la traduzione può far perdere, almeno in parte, il contenuto emozionale che l'utente aveva inserito nella frase. In aggiunta a questi fattori oggettivi va ricordato la difficoltà degli esseri umani nel definire in maniera univoca le emozioni. EmoNet è infatti stato addestrato su emozioni ricavate dall'analisi degli hashtag che caratterizzavano ciascun tweet. Vi è stata una validazione manuale di tutto il procedimento, ma comunque il tool non si basa su un oracolo definito direttamente da esseri umani e questo può averlo reso meno efficiente nell'analisi di un oracolo totalmente definito dall'uomo. Questi risultati avvalorano, ulteriormente, la tesi sulla difficoltà nel percepire in maniera oggettiva le emozioni nel testo, sia per un tool automatico, sia per le persone stesse. Per quanto concerne il confronto tra le varie soglie si è visto come il valore che garantisce le performance migliori in media su ogni emozione è quello legato al primo quartile. Tale valore consente di massimizzare le

performance del tool sebbene, in ogni caso, non ci sia un miglioramento così significativo tra un valore e l'altro.

Per completare l'analisi si è deciso di prendere in esame alcune recensioni che sono state classificate come false positive e false negative secondo alcune emozioni per cercare di capire le differenze tra oracolo e tool. Il primo commento da esaminare è lo stesso considerato nel paragrafo "Definizione dell'oracolo". Lo si riporta per una maggiore chiarezza:

"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."

Si ricorda che il commento era stato oggetto di dibattito durante la definizione dell'oracolo per la presenza dell'emozione della tristezza, estrapolata dal contesto entro cui il commento si inserisce. A differenza di come previsto il tool è riuscito a percepire la tristezza, che si è rivelata essere un vero positivo per tutte le soglie considerate, mentre ha avuto problemi a riconoscere i sentimenti della gioia e della sorpresa che si sono rivelate essere dei falsi negativi per le soglie al secondo e al terzo quartile. Allo stesso tempo il commento è risultato un falso positivo per emozioni quali paura e fiducia che non sono state affatto rilevate dagli utenti che hanno provveduto a compilare l'oracolo.

Il secondo commento che si vuole prendere come esempio è il seguente:

"buoni cantanti e pianista. Bel mix di momenti salienti dell'opera. L'acustica del teatro è purtroppo triste. Trambusto accogliente nel mezzo di Taormina. Ancora bel tempo il 17 settembre 2019."

Il commento è positivo, ma non trasmette gioia, quanto piuttosto la tristezza legata alla conformazione del teatro che rovina l'acustica. Questa l'analisi che è stata fatta durante la compilazione manuale dell'oracolo, che per questo non ha riportato né la gioia né la fiducia come emozioni presenti. Il commento è risultato essere un falso positivo per tutte le emozioni, tranne la tristezza, se si considera la prima soglia considerata. Questo

risultato migliora all'aumentare della soglia considerata, ma permette di comprendere come il tool, in questo caso, abbia assegnato quasi una certa percentuale ad ogni emozione sebbene questa rappresenti una frase molto descrittiva e priva di emozioni significative. Esso inoltre è risultato essere un falso negativo per quanto concerne la tristezza per le soglie al secondo e al terzo quartile. Questo commento rappresenta bene le difficoltà con cui ci si imbatte in questo tipo d'analisi. Il tool, probabilmente, percependo la gioia trasmessa nel descrivere un'esperienza assolutamente positiva classifica il commento come rappresentante di quel sentimento, mettendo in secondo piano analisi che gli esseri umani hanno ritenuto più valide.

6 Conclusioni

L'elaborato ha verificato lo stato dell'arte nell'ambito della sentiment analysis 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 un qualcosa di sorprendente. La vera criticità 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 quantità di argomenti, dalla semplice descrizione di oggetti acquistati on line, fino alla descrizione delle più 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 obiettivi prefissati. In questa fase si è notato come l'intero Natural 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 per essere in grado di ottenere dei buoni risultati e delle ottime performance. In quest'ambito le difficoltà nell'analisi 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, è stato possibile individuare un solo tool che operasse 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 più 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 troverà 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 già 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, nell'ambito della sentiment analysis, hanno ottenuto risultati soddisfacenti, con dei valori di F-Measure prossimi allo 0,9. Questi risultati hanno mostrato prima di tutto l'efficienza dei tool individuati, oltre a dimostrare che il processo di traduzione per questo tipo di analisi sia poco impattante.

Per quanto concerne, invece, l'emotion mining le performance del tool preso in considerazione si sono rivelate più deludenti. Le motivazioni che spiegano il perché del peggioramento dei risultati del tool rispetto a quelli ottenuti dagli sviluppatori [18] sono diverse. Tra le principali citiamo prima di tutto il fatto che si sta operando con un tool addestrato a funzionare in un contesto totalmente diverso, ed allo stesso tempo si sta operando su delle frasi mediamente molto più lunghe rispetto a quelle sulle quali il tool stesso è stato addestrato. Già da soli questi fattori hanno sicuramente impattato notevolmente sulle performance, che sono state ulteriormente peggiorate dall'aver operato su delle traduzioni che, probabilmente, hanno fatto perdere una parte del contenuto emozionale che gli utenti hanno cercato di infondere nel testo. In conclusione non ci si può ritenere soddisfatti dai risultati del tool che opera sull'emotion mining che,

per poter funzionare al meglio, andrebbe riaddestrato per operare nelle nuove condizioni imposte dal progetto CANTICO. Questo sviluppo futuro si scontrerà con la mancanza di un dataset abbastanza vasto per funzionare da training set e test set, che sia incentrato sulle recensioni di opere teatrali.

7 Indice delle figure e delle tabelle

FIGURA 1 ESEMPIO ANALISI TINT [10]	17
FIGURA 2 ESEMPIO ANALISI MEANINGCLOUD [11]	18
FIGURA 3 ESEMPIO ANALISI REPUSTATE [12]	19
FIGURA 4 ESEMPIO ANALISI TWINWORD [16]	24
FIGURA 5 ESEMPIO PARALLELDOTS [17]	24
FIGURA 6 RUOTA DELLE EMOZIONI [18]	26
FIGURA 7 ARCHITETTURA DEL TOOL REALIZZATO	34
FIGURA 8 ACTIVITY DIAGRAM DEL TOOL REALIZZATO	35
FIGURA 9 ESEMPIO DATASET TAGGATO	39
FIGURA 10 INTESA NELLA DEFINIZIONE DELL'ORACOLO	39
FIGURA 11 ANALISI PERCEZIONE TRA 3 PERSONE	40
FIGURA 12 PRECISION SENTITA	45
FIGURA 13 PRECISION STANFORD CORENLP	46
FIGURA 14 RECALL SENTITA	46
FIGURA 15 RECALL STANFORD CORENLP	47
FIGURA 16 F-MEASURE SENTITA	47
FIGURA 17 F-MEASURE STANFORD CORENLP	48
FIGURA 18 ANALISI RABBIA PER QUARTILE	48
FIGURA 19 ANALISI DISGUSTO PER QUARTILE	49
FIGURA 20 ANALISI ANTICIPAZIONE PER QUARTILE	49
FIGURA 21 ANALISI GIOIA PER QUARTILE	50
FIGURA 22 ANALISI SORPRESA PER QUARTILE	50
FIGURA 23 ANALISI FIDUCIA PER QUARTILE	51
FIGURA 24 ANALISI TRISTEZZA PER QUARTILE	51
TABELLA 1 TOOL PER L'ANALISI TESTUALE.....	16
TABELLA 2 TOOLS PER LA SENTIMENT ANALYSIS	23
TABELLA 3 TOOL PER L'EMOTION MINING	27
TABELLA 4 POLARITÀ DEL DATASET UTILIZZATO.....	41
TABELLA 5 DEFINIZIONE DELLE EMOZIONI DEL DATASET	42

8 Bibliografia

- [1] G. G. Chowdhury, *Introduction to Modern Information Retrieval*, Facet Publishing, 2010.
- [2] L. Deng e Y. Liu, *Deep Learning in Natural Language Processing*, Springer, 2018.
- [3] T. Nasukawa e J. Yi, «Sentiment analysis: Capturing favorability using natural language processing.,» *Proceedings of the 2nd international conference on Knowledge capture.*, pp. 70-77, 2003.
- [4] K. Dave, S. Lawrence e D. M. Pennock, «Mining the peanut gallery: Opinion extraction and semantic classification of product reviews.,» *Proceedings of the 12th international conference on World Wide Web.*, pp. 519-528, 2003.
- [5] B. Liu, *Sentiment Analysis and Opinion Mining*, Morgan & Claypool Publishers, 2012.
- [6] 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.
- [7] R. K. I. a. Y. M. Tokuhsa, «Emotion classification using massive examples extracted from the web.,» *Proceedings of the 22nd International Conference on Computational Linguistics. Association for Computational Linguistics*, vol. Volume 1, pp. 881-888, 2008.
- [8] A. Palmero e G. Moretti, «Italy goes to Stanford: a collection of CoreNLP modules for Italian,» *ArXiv e-prints*, 2016.

- [9] 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.
- [10] A. Palmero e G. Moretti, «TINT Demo,» [Online]. Available: <http://simpatico.fbk.eu/demo/>. [Consultato il giorno 30 Gennaio 2020].
- [11] Meaning Cloud, «Demo,» [Online]. Available: <https://www.meaningcloud.com/demos/text-analytics-demo#>. [Consultato il giorno 30 Gennaio 2020].
- [12] Repustate, «Demo Sentiment analysis,» [Online]. Available: <https://www.repustate.com/sentiment-analysis-api-demo/>. [Consultato il giorno 30 Gennaio 2020].
- [13] A. Ansari, A. Seenivasan, A. Anandan e R. Lakshmanan, «Twitter Sentiment Analysis,» 12 Novembre 2017. [Online]. Available: <https://github.com/abdufatir/twitter-sentiment-analysis>. [Consultato il giorno 30 Gennaio 2020].
- [14] 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.
- [15] G. Nicola, «SentIta,» 2019. [Online]. Available: <https://nicgian.github.io/Sentita/>. [Consultato il giorno 30 Gennaio 2020].
- [16] Twinword, «Demo Emotion Mining,» [Online]. Available: <https://www.twinword.com/api/emotion-analysis.php>. [Consultato il giorno 30 Gennaio 2020].

- [17] Paralleldots, «Demo Emotion Mining,» [Online]. Available: <https://www.paralleldots.com/emotion-analysis>. [Consultato il giorno 30 Gennaio 2020].
- [18] 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.
- [19] O. Ayman, «GitHub,» 2018. [Online]. Available: <https://github.com/omar178/Emotion-recognition#p4>. [Consultato il giorno 30 Gennaio 2020].
- [20] J. C. Garfias Tovar, «Github,» [Online]. Available: <https://github.com/SeaWar741/Emotion-detection>. [Consultato il giorno 30 Gennaio 2020].

9 Ringraziamenti

Vorrei dedicare questo spazio a chi, con dedizione e pazienza, ha contribuito al raggiungimento di questo obiettivo.

In primis, un ringraziamento speciale va al mio relatore Massimiliano Di Penta che mi ha seguito, con la sua infinita disponibilità, in ogni step della realizzazione dell'elaborato, fin dalla scelta dell'argomento.

Una menzione particolare va al mio Correlatore Fiorella Zampetti, per la sua immensa pazienza, per i suoi indispensabili consigli, e per le conoscenze trasmesse durante tutto il percorso di stesura dell'elaborato.

Ringrazio infinitamente mia madre e mio padre, senza i loro insegnamenti e senza il loro supporto questo lavoro di tesi, che corona il mio percorso universitario, non esisterebbe nemmeno. Grazie per avermi fatto sbagliare, l'unico modo per imparare e migliorare, supportando le mie scelte e soprattutto grazie per essere rimasti lì a rimediare ai miei errori.

Un grazie va poi a mia sorella che mi è sempre rimasta accanto, supportandomi e aiutandomi in tutti i modi possibili.

Grazie ai miei amici, tutti indistintamente, per essere stati sempre presenti. Grazie per aver ascoltato i miei sfoghi, per tutti i "progetti" che abbiamo condiviso, per tutte le ansie e soprattutto grazie per tutti i momenti di spensieratezza.

Grazie a Francesco Tedesco che più di tutti ha condiviso con me la lunga strada che ci ha portato ad oggi, e con cui posso festeggiare un traguardo raggiunto insieme.

Ringrazio la mia fidanzata per avermi trasmesso il suo immenso coraggio e la sua gioia di vivere. Grazie per tutto il tempo che mi hai dedicato. Grazie per tutti i ti amo e per tutte le volte che mi hai dato la forza durante i momenti difficili. Grazie per condividere questo giorno speciale, per averlo progettato e desiderato almeno quanto me. Grazie perché adesso, insieme, daremo il via a qualcosa di nuovo. Grazie perché ci sarai sempre.