Analisi del comportamento degli utenti attraverso lo studio delle recensioni su TripAdvisor

1. **Introduzione**

Con l’avvento di Internet, i ristoratori hanno iniziato a pubblicizzare la propria attività tramite siti web e social network. Questa tendenza è risultata utile per due motivazioni:

* Maggiore visibilità per le attività di ristorazione, con conseguente aumento della clientela.
* Ricezione di feedback riguardanti il servizio fornito e quindi di effettuare modifiche volte a migliorare le loro offerte.

Le recensioni sono quindi diventate fondamentali per i ristoratori, in quanto sono uno degli strumenti utilizzati dai clienti per selezionare i ristoranti che più rispondono alle proprie esigenze e gusti. Tramite le recensioni è possibile intuire più precisamente quali servizi vengono maggiormente apprezzati e quali invece dovrebbero essere migliorati secondo l’utenza, ad esempio un cliente potrebbe apprezzare la cucina offerta da un ristorante, ma potrebbe esprimersi negativamente riguardo il personale.

Una delle piattaforme web più utilizzate nel campo delle recensioni online di ristoranti è TripAdvisor.

1. **TripAdvisor**

Tripadvisor, Inc. è un'azienda statunitense di viaggi e ristoranti che pubblica recensioni di alberghi, bed and breakfast e ristoranti, prenotazioni di alloggi e altri contenuti relativi ai viaggi. Include anche forum di viaggi interattivi ed è diffusa in tutto il mondo.

Il sito vanta un largo bacino di utenza tra cui: ristoratori e albergatori, i quali pubblicizzano le loro attività, e clienti, i quali utilizzano il sito sia come strumento di selezione, sia come intermediario per le loro prenotazioni oppure per lasciare recensioni.

Per la piattaforma cruciale è l’attività di recensione, infatti ogni anno vengono assegnati premi alle attività meglio recensite che ne attestano la qualità. Questi premi sono molto ambiti dalle attività poiché permettono di ottenere una maggiore visibilità al pubblico.

**2.1 Recensioni**

Le recensioni sono lo strumento con cui un cliente previa consumazione nell’attività di riferimento propone una critica nei confronti del locale. Qualsiasi persona dotata di un account registrato sulla piattaforma può pubblicare una propria recensione, indicando solitamente un titolo a scopo esplicativo, un corpo della recensione ed una votazione che va da 1 a 5 stelline, più è alto il valore delle stelline più è stata positiva l’esperienza del cliente.

Più recensioni positive ottiene un locale, più questo acquisisce visibilità all’interno della piattaforma, tale meccanismo è molto simile alla presentazione delle pagine sui browser, i siti più ricercati sono tra le prime pagine, mentre quelle meno cercate sono “nascoste” nelle pagine successive.

Una recensione può riguardare diversi aspetti del locale recensito, ad esempio una recensione per un ristorante può parlare di cibo, location, costi e servizio offerto. Tali informazioni risultano indispensabili nello studio affrontato nell’elaborato proposto.

1. **Obiettivo del lavoro svolto**

Nel seguente elaborato ci siamo posto l’obiettivo di studiare ed analizzare il comportamento dei recensori all’interno della piattaforma, in particolare per le recensioni delle attività di ristorazione.

I ristoranti in questo caso di studio sono stati selezionati in base alla loro regione di appartenenza, principalmente abbiamo raccolto i ristoranti più famosi dei capoluoghi delle regioni italiane, più specificamente sono stati selezionati approssimativamente lo stesso numero di ristoranti per il Nord, Centro e Sud dell’Italia.

I capoluoghi selezionati sono (Da Nord a Sud):

* Ancona
* Bologna
* Milano
* Trieste
* Genova
* Venezia
* Firenze
* L’Aquila
* Roma
* Perugia
* Potenza
* Napoli
* Palermo
* Cagliari
* Catanzaro

Per ciascuna delle città selezionate, sono stati considerati i primi 13 ristoranti che comparivano inserendo la città come input di ricerca nel sito. Per ciascuno dei 13 ristoranti sono state considerate all’incirca 15 recensioni, per un totale di 2925 recensioni. I valori elencati hanno inoltre rappresentato il workload dello studio condotto.

Lo scopo del lavoro svolto è quello di individuare attraverso l’analisi delle recensioni, la quantità di recensioni che sono positive o comunque costruttive, la quantità di recensioni negative e quelle invece che si frappongono tra le due o più semplicemente quelle neutrali.

Questo tipo di analisi oltre ad essere fatto sul totale di tutte le recensioni per tutte le città selezionate, è stata eseguita anche su una suddivisione per zone della penisola, ovvero nord, sud e centro.

Una simile analisi, ma più capillare è stata fatta sugli aspetti delle recensioni, cioè fra tutte le recensioni rilevate, ne abbiamo estrapolato il contesto e ne abbiamo analizzato il sentimento, per comprendere se quest’ultimo è stato positivo, negativo o neutrale. Gli aspetti analizzati riguardano essenzialmente le caratteristiche principali di un’attività di ristorazione quali il cibo, la location, il costo e il servizio offerto. Tale analisi è utile quindi a capire più precisamente quali aspetti del ristorante sono stati apprezzati maggiormente e quali invece sono stati criticati.

* 1. **Tecnologie e metodologie utilizzate**

L’estrazione dei dati è stata costruita attraverso l’impiego di un tool di scraping pubblicato da Maximilian Copelli sul sito [www.apify.com](http://www.apify.com). Il tool consente di prelevare le recensioni e i ristoranti in base ad un input predefinito, in particolare la città, restituendo le informazioni richieste in un file JSON. Le informazioni contenute all’interno del file restituito comprendono:

* Nome del ristorante
* Città in cui è ubicato il ristorante
* Il tipo di cucina del ristorante, ad esempio: sushi, orientale, pizzeria, etc…
* Le recensioni

Per la lettura dei dati e per le elaborazioni core del progetto è stato utilizzato Java 8, che ci ha concesso di prelevare i dati dai file ed organizzare i risultati ottenuti dalla fase di sentiment analysis in file di output.

La sentiment analysis e l’opinion mining delle recensioni sono stati realizzati con i cognitive services offerti da Microsoft Azure. In particolare l’operazione di sentiment analysis ci ha permesso di valutare il sentiment di una particolare recensione, mentre l’opinion mining ci ha permesso di estrapolare le parole chiave all’interno della recensione e di fornire anche in questo caso un sentiment correlato.

Per utilizzare i cognitive services si necessita di un account iscritto alla piattaforma di Azure e di effettuare una subscription a tali servizi. La subscription fornisce delle chiavi di accesso collegate all’account che consentono di autenticare un client che verrà utilizzato per richiamare i metodi offerti dal servizio. I metodi sono chiamati attraverso semplici istruzioni Java e restituiscono risposte in formato JSON.

L’autenticazione è stata effettuata come segue:

public static void inizialize(String key, String endpoint) {

*client* = TextAnalyticsSamples.*authenticateClient*(key, endpoint);

*client* = TextAnalyticsSamples.*authenticateClient*(key, endpoint);

}

Dove key è la chiave di sottoscrizione, ed endpoint è l’indirizzo della risorsa in cloud.

Le operazioni di sentiment analysis, invece sono state effettuate come segue:

JSONObject sentimentInfo = new JSONObject();

sentimentInfo = TextAnalyticsSamples.sentimentAnalysisWithOpinionMining(client,textOfReview);

Un esempio dei dati prelevati dal metodo e poi formattati nel formato utile al nostro scopo è il seguente:

{

​​​​"RISTORANTI": [

{

​​​​"RECENSIONI": [

{

​​​​...

}​​​​,

{

​​​​"RECENSIONE": "Moglie vegana, ogni tanto tocca anche a me Mia moglie è vegana, quindi ogni tanto lo devo essere anche io. Non nascondendo un certo disinteresse per la filosofia dietro a questa dieta, devo anche ammettere che però questo ristorante mi ha parzialmente fatto ricredere sui miei ideali da carnivoro. Mi hanno servito un’ ottima bowl piena di cose molto buone, e per una sera non ho sentito la mancanza della ciccia!",

"RATE RECENSIONE": 4,

"SENTIMENT RECENSIONE": {

​​​​"SCORE": [0.51, 0, 0.4],

"SENTIMENT": "mixed",

"SENTENCES": [

{

​​​​"SENTENCE SENTIMENT": "neutral",

"SCORE": [0.01, 0.98, 0.01],

}​​​​,

{

​​​​"SENTENCE SENTIMENT": "negative",

"SCORE": [0.01, 0, 0.99],

}​​​​,

{

​​​​"SENTENCE SENTIMENT": "positive",

"SCORE": [1, 0, 0],

"ASPECTS": [

{

​​​​"OPINION MINING": [

{

​​​​"OPINION SENTIMENT": "positive",

"OPINION": "great"

}​​​​

],

"ASPECT": "bowl",

"SENTIMENT ASPECT": "positive"

}​​​​,

{

​​​​"OPINION MINING": [

{

​​​​"OPINION SENTIMENT": "positive",

"OPINION": "good"

}​​​​

],

"ASPECT": "things",

"SENTIMENT ASPECT": "positive"

}​​​​

]

}​​​​

]

}​​​​

}​​​​

],

"CITTA'": "ancona",

"NOME": "Zucchero a Velo",

"RANKING": 4.5

}​​​​,

...

]

}​​​​

**3.2 Limitazioni sullo studio**

Le limitazioni riscontrate riguardano soprattutto l’impiego di tecnologie, ad esempio per Apify non è stato possibile prelevare le informazioni delle recensioni per alcune città, in particolare per la città di Bari, l’output non era veritiero, infatti i ristoranti rilevati erano ubicati in Argentina.

Non è stato possibile, inoltre prelevare più di 13 recensioni per ristorante, poiché i cognitive services di Azure fornivano un numero di transazioni gratuite limitate.

Per utilizzare nello specifico il servizio di opinion mining è stato necessario tradurre le recensioni in lingua inglese, poiché Azure mette a disposizione tale funzionalità solo per testi in lingua inglese. Di conseguenza, prima di dare in input la recensione al servizio, si è implementato uno script su Google Apps Script, per effettuarne la traduzione.

Per via di questo inconveniente alcuni degli aspetti rilevati potrebbero essere poco consistenti con la realtà, poiché l’operazione di traduzione non è efficiente al 100%.

**3.3 Rappresentazione grafica del workflow**

