

# Analisi dei Sentimenti sulla Maternità Surrogata come Reato Universale: Un Approccio Basato sui Dati di Reddit

Professor Delfina Malandrino

Student Giovanni Forlenza

## Abstract

I reati universali, come la tortura, il genocidio, la riduzione in schiavitù e i crimini contro l'umanità, sono riconosciuti come tali dalla comunità internazionale per la loro gravità assoluta. Recentemente, l'Italia ha approvato una legge che definisce **la maternità surrogata** (nota anche come gestazione per altri o "utero in affitto") un reato universale, rendendo perseguibili penalmente i cittadini italiani che ricorrono a questa pratica anche all'estero, indipendentemente dalla sua legalità nel paese di origine. Questo provvedimento ha sollevato un acceso dibattito sia a livello nazionale che internazionale.

Questo studio si propone di analizzare le reazioni delle persone rispetto a questa nuova legge attraverso un'analisi sistematica dei commenti su Reddit. Utilizzando le API della piattaforma, saranno raccolti commenti pertinenti, che verranno poi sottoposti a un modello di **sentiment analysis** pre-addestrato per identificare l'orientamento emotivo e le opinioni espresse dagli utenti. L'obiettivo è fornire una panoramica delle percezioni pubbliche su questo tema controverso, evidenziando i sentimenti predominanti e i principali argomenti di discussione.

## Metodologia

### Raccolta dati

Per identificare i thread pertinenti su Reddit, saranno utilizzate parole chiave specifiche che coprono sia i termini tecnici sia quelli più colloquiali legati al tema della maternità surrogata. Tra queste parole chiave figurano: **"maternità surrogata"**, **"utero in affitto"**, **"gestazione per altri"**, **"reato universale"** e **"legge maternità surrogata"**. Questi termini sono stati scelti per garantire una copertura ampia e rappresentativa delle discussioni sull'argomento.

Oltre alle parole chiave, saranno applicati alcuni filtri per selezionare i contenuti più rilevanti. Innanzitutto, verranno considerati solo i commenti scritti in **italiano**, utilizzando strumenti per il rilevamento della lingua che consentano di escludere quelli in altre lingue. Saranno inoltre prese in esame soltanto le risposte che abbiano una lunghezza minima di **20 caratteri**, al fine di escludere interventi troppo brevi o non significativi, come semplici espressioni di accordo o disaccordo. Per mantenere l'analisi focalizzata su discussioni recenti, saranno inclusi soltanto commenti pubblicati negli **ultimi 12 mesi**.

Infine, verranno utilizzate le API di Reddit per cercare post e commenti che contengano almeno una delle parole chiave definite. La ricerca sarà ulteriormente affinata concentrandosi su subreddit rilevanti, come **r/Italia** o **r/LegalAdviceItaly**, dove è più probabile trovare discussioni approfondite sul tema. Questi criteri garantiranno che i dati raccolti siano rilevanti e utili per l'analisi dei sentimenti.

### Pre-elaborazione dati

Per garantire la qualità dei dati raccolti, sarà effettuata una fase di pre-elaborazione del testo. In questa fase, i commenti saranno convertiti interamente in lettere minuscole per uniformare il formato e facilitare l'analisi. Successivamente, saranno rimosse le **stopwords**, ovvero le parole comuni e prive di rilevanza semantica (ad esempio, articoli, preposizioni e congiunzioni), insieme alla **punteggiatura** e ai contenuti non testuali, come URL o emoji.

Inoltre, sarà applicata una procedura di **normalizzazione del testo**, che include tecniche come la lemmatizzazione o lo stemming, per ridurre le parole alla loro forma base. Questo processo consente di migliorare la qualità dei dati e garantisce una rappresentazione più coerente delle informazioni testuali per l'analisi successiva.

## Analisi dei sentimenti

Per l'analisi dei sentimenti, sarà utilizzato un modello pre-addestrato, come **AIBERTO** o **UmBERTo**, integrato attraverso la libreria **spaCy**. Questi modelli, progettati per la lingua italiana, garantiscono una comprensione approfondita del testo e una classificazione accurata delle emozioni espresse nei commenti.

I commenti saranno classificati in tre categorie principali di sentimenti: **positivi**, **negativi** o **neutrali**. Questa suddivisione permetterà di analizzare l'orientamento emotivo prevalente e di identificare eventuali polarizzazioni nel dibattito pubblico sulla maternità surrogata.

## Analisi dei Risultati

Una volta classificati i commenti in base ai sentimenti, verrà effettuata un'analisi statistica per determinare la **distribuzione dei sentimenti** nel dataset. Questa analisi consentirà di quantificare la percentuale di commenti positivi, negativi e neutrali, fornendo una panoramica generale dell'orientamento emotivo degli utenti.

Inoltre, verrà effettuato un confronto tra le diverse pagine (o subreddit) per valutare se vi siano differenze significative nell'orientamento dei sentimenti. Questo permetterà di comprendere se alcune pagine mostrano una maggiore inclinazione verso sentimenti positivi, negativi o neutrali rispetto ad altre, offrendo ulteriori spunti sull'impatto della comunità di riferimento sul dibattito.

## Visualizzazione e interpretazione

Per facilitare la comprensione dei risultati, saranno creati **grafici** che rappresentano visivamente la distribuzione dei sentimenti nei commenti raccolti. Saranno utilizzati grafici comparativi per mostrare le differenze nell'orientamento dei sentimenti tra le diverse pagine o subreddit.

Infine, i risultati ottenuti saranno analizzati alla luce delle loro **implicazioni etiche e sociali**. La discussione si concentrerà su come le opinioni pubbliche, emerse dai dati, riflettano il dibattito contemporaneo sulla maternità surrogata e sul suo riconoscimento come reato universale.