



Abhishek Jani

IT Department

Roll No - 50

Experiment No.7
Medical Reviews Analysis from social media data
Date of Performance: 16/10/2023
Date of Submission:29/10/2023

Aim: To use social media data to analyse medical reviews

Objective: Given a dataset with social media data analyse, and interpret information and feedback about medical products, services, treatments, or healthcare facilities as expressed by individuals on social media platforms.

Theory:

Medical reviews analysis from social media data refers to the process of collecting, analyzing, and interpreting information and feedback about medical products, services, treatments, or healthcare facilities as expressed by individuals on social media platforms. This analysis is often conducted to gain insights into the effectiveness, safety, patient experiences, and public perceptions related to various aspects of healthcare. Here are the key components of medical reviews analysis from social media data:

Data Collection: Data is collected from various social media platforms such as Twitter, Facebook, Reddit, online forums, and healthcare-specific websites. This data can include text, images, videos, and other forms of user-generated content.

Text Mining and Natural Language Processing (NLP): Text mining and NLP techniques are used to extract and process the textual information within the collected data. This involves sentiment analysis, named entity recognition, and other linguistic and semantic analyses to understand the content and context of medical reviews.



Sentiment Analysis: Sentiment analysis is a critical component of this analysis. It helps categorize the sentiment of reviews as positive, negative, or neutral. Sentiment analysis can reveal the overall satisfaction of patients and the areas where improvements are needed.

Topic Modeling: Topic modeling techniques are used to identify and categorize common themes or topics discussed in the medical reviews. This can help healthcare organizations understand what aspects of their services or products are frequently mentioned and which areas require attention.

Adverse Event Detection: In the case of medical products or treatments, adverse event detection is crucial. By analyzing social media data, it's possible to identify potential adverse events associated with certain drugs or medical interventions.

Patient Experience Analysis: Analysis of social media data can provide insights into the patient experience, including factors like wait times, staff behavior, facility cleanliness, and the overall quality of care provided by healthcare facilities.

Public Perceptions and Trends: Social media data can reveal public perceptions about healthcare issues, emerging trends in medical treatments, and the popularity of different healthcare providers or services.

Insights for Healthcare Providers and Regulators: The findings from social media data analysis can be used to improve the quality of healthcare services and products, identify areas for improvement, enhance patient satisfaction, and ensure compliance with regulatory standards.

Data Privacy and Ethics: It's important to handle social media data with sensitivity to privacy and ethical considerations. Anonymization



and consent are crucial aspects when collecting and analyzing such data.

Reporting and Decision-Making: The results of the analysis are typically reported to healthcare providers, pharmaceutical companies, regulatory agencies, and other stakeholders to inform decision-making and improve the healthcare system.

Overall, medical reviews analysis from social media data is a valuable tool for gaining insights into the real-world experiences and perceptions of patients, consumers, and the general public regarding various aspects of healthcare, which can ultimately drive improvements in healthcare quality and patient care.

Code: -

```
AIHC-EXPT7.ipynb
File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

[5] import requests
    from bs4 import BeautifulSoup
    import nltk
    from nltk.sentiment import SentimentIntensityAnalyzer
    from nltk.tokenize import word_tokenize, sent_tokenize
    from nltk.corpus import stopwords

    # Ensure necessary NLTK datasets/libraries are downloaded
    nltk.download('vader_lexicon')
    nltk.download('stopwords')

    # Define the URL of the webpage containing medical reviews
    url = 'https://www.who.int/health-topics/coronavirus'

    # Set the user-agent header
    headers = {'User-Agent': 'Mozilla/5.0 (Windows NT 10.0; WOW64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/118.0.0.0 Safari/537.36'}

    # Send an HTTP GET request to the URL with the custom user-agent header
    response = requests.get(url, headers=headers)

    # Parse the HTML content using BeautifulSoup
    soup = BeautifulSoup(response.content, 'html.parser')
    reviews = soup.find_all('div', class_='review-container') # Adjust based on your webpage's structure

    # Initialize NLTK's sentiment analyzer
    sia = SentimentIntensityAnalyzer()

    for review in reviews:
        # Extract the review text
        review_text = review.find('p', class_='review-text').text # Adjust based on your webpage's structure
```



```
# Remove stopwords
stop_words = set(stopwords.words('english'))
filtered_tokens = [word for word in tokens if word.lower() not in stop_words]

# Calculate sentiment scores
sentiment_scores = sia.polarity_scores(review_text)

print("Review Text:")
print(review_text)
print("\nTokens:")
print(filtered_tokens)
print("\nSentiment Scores:")
print(sentiment_scores)
print("\n---\n")
```

[nltk_data] Downloading package vader_lexicon to /root/nltk_data...
[nltk_data] Package vader_lexicon is already up-to-date!
[nltk_data] Downloading package stopwords to /root/nltk_data...
[nltk_data] Package stopwords is already up-to-date!

Google Collaboratory Link: -

EXPT7.ipynb

Conclusion: -

Thus we have successfully performed Medical Reviews Analysis from social media data.

