Academic Year: 2025-26

Subject: AI&ML in Healthcare

Name:	Jay Kore
Roll No & Branch:	58 - COMPS
Class/Sem:	BE/VII
Experiment No.:	07
Title:	Medical Reviews Analysis from social media data.
Date of Performance:	26-09-25
Date of Submission:	03-10-25
Marks:	
Sign of Faculty:	

Aim: To study Medical Reviews Analysis from social media data.

Objective: To apply text mining and natural language processing (NLP) for analyzing social media health reviews and to identify patient sentiments (positive, negative, neutral) regarding healthcare services and medicines. To understand trends, concerns, and experiences shared by patients online and to demonstrate how social media data can enhance healthcare research and decision-making.

Theory: With the rise of digital platforms, social media has become a powerful source of healthcare-related discussions. Patients often share their experiences about treatments, medicines, hospitals, and doctors. This user-generated data can be analyzed using **Natural Language Processing (NLP)** and **Machine Learning** techniques to gain valuable insights.

Steps in Analysis:

1. **Data Collection:** Extract reviews, tweets, or posts from platforms like Twitter, Facebook, or health forums.

2. **Preprocessing:** Clean text (remove noise, stop words, punctuation, emojis).

3. **Sentiment Analysis:** Apply AI models to classify opinions (positive/negative/neutral).

4. **Topic Modeling:** Identify recurring themes such as side effects, satisfaction, cost, or service quality.

5. **Visualization:** Use charts/word clouds to represent the findings.

Applications:

- Evaluating public opinion on new drugs or vaccines.

- Identifying side effects reported by patients.

- Measuring satisfaction with hospital services.

- Supporting healthcare policy decisions based on real-world feedback.

Benefits:

- Real-time monitoring of patient experiences.

- Cost-effective and large-scale feedback collection.

- Helps improve healthcare services and patient trust.

Challenges:

- Privacy and ethical concerns in using personal social media data.

- Fake reviews or misinformation.

- Language diversity and slang may complicate analysis.

Program and output

(Illustration – a simple Python program can fetch dummy review data and perform sentiment analysis using NLP libraries like NLTK/TextBlob.)

```python
import pandas as pd

from textblob import TextBlob

# Dummy Data Collection (simulating social media reviews)

data = {

    'review_id': [1, 2, 3, 4, 5, 6, 7, 8, 9, 10],

    'review_text': [

        "This hospital provided excellent care, I'm very satisfied.",

        "The new medication has too many side effects, feeling unwell.",

        "Doctor was professional and the appointment was on time. Neutral experience.",

        "Terrible service, long waiting times and rude staff.",

        "Highly recommend this clinic, friendly staff and effective treatment.",

        "Vaccine caused mild fever, but otherwise fine. A bit concerning.",

        "The staff at the emergency room were incredibly helpful.",

        "Pharmacy was out of stock, very inconvenient.",

        "My recovery has been smooth thanks to the therapy.",

        "Unsure about the treatment effectiveness, need more time to tell."
```

```
                                                                              ]

                                                                          }

                                                                    df =
pd.DataFrame(data)

# Preprocessing (TextBlob handles some normalization internally, but a simple clean-up can be added)

def clean_text(text):

    text = text.lower()

    return text

df['cleaned_review_text'] = df['review_text'].apply(clean_text)

# Sentiment Analysis

def get_sentiment(text):

    analysis = TextBlob(text)

    if analysis.sentiment.polarity > 0:

        return 'Positive'

    elif analysis.sentiment.polarity < 0:

        return 'Negative'

else:

        return 'Neutral'

df['sentiment'] = df['cleaned_review_text'].apply(get_sentiment)

# Topic Modeling (Simplified - identifying keywords)

keywords = {

    'side effects': ['side effects', 'unwell', 'fever', 'concerning'],

    'satisfaction': ['excellent care', 'satisfied', 'highly recommend', 'friendly', 'helpful', 'smooth recovery'],
```

'service quality': ['professional', 'on time', 'terrible service', 'long waiting times', 'rude staff', 'inconvenient'],

    'treatment effectiveness': ['effective treatment', 'vaccine caused', 'recovery has been smooth', 'unsure about treatment effectiveness']

}

```python
def identify_topics(text):

    found_topics = []

    for topic, kws in keywords.items():

        if any(kw in text for kw in kws):

            found_topics.append(topic)

    return ", ".join(found_topics) if found_topics else "General"

df['topics'] = df['cleaned_review_text'].apply(identify_topics)

print("Program Output:")

print(df[['review_text', 'sentiment', 'topics']].to_string(index=False))
```

Program Output:

```
                              review_text sentiment    topics

 This hospital provided excellent care, I'm very satisfied.  Positive             satisfaction

    The new medication has too many side effects, feeling unwell.  Negative       side effects

   Doctor was professional and the appointment was on time. Neutral experience.   Neutral service quality

         Terrible service, long waiting times and rude staff.  Negative service quality

   Highly recommend this clinic, friendly staff and effective treatment.  Positive satisfaction, treatment effectiveness
```

Vaccine caused mild fever, but otherwise fine. A bit concerning.   Neutral   side effects

The staff at the emergency room were incredibly helpful. Positive   satisfaction

Pharmacy was out of stock, very inconvenient. Negative service quality

My recovery has been smooth thanks to the therapy. Positive satisfaction, treatment effectiveness

Unsure about the treatment effectiveness, need more time to tell. Neutral treatment effectiveness

Conclusion:

The analysis of medical reviews from social media data provides valuable insights into patient perspectives, treatment outcomes, and healthcare quality. By applying data science and NLP, healthcare providers and researchers can identify trends, detect emerging issues, and improve decision-making. While challenges like privacy and misinformation exist, social media review analysis remains a powerful tool for enhancing healthcare delivery and patient satisfaction.