

## DSHC : Assignment - 2

Q1. Illustrate the analysis of social media data in the detection & tracking of infectious disease outbreaks.

① Data Collection:

Choose the Social Media Platform:

Select a social media platform for data collection. Twitter is often a popular choice due to its real-time & public accessibility.

Data Gathering:

Use API to collect relevant data.

② Data Preprocessing:

text cleaning

Clean the collected text data to remove noise, including irrelevant tweets, URLs, special characters & duplicates.

Sentiment Analysis:

perform sentiment analysis to understand the overall sentiment (positive / negative / neutral) of the tweets related to the disease.

③ Keyword Extraction:

Identify disease-related keywords.

extract keywords & hashtags related to the disease from the cleaned text.

④ Geospatial Analysis:

extract location info:

determine the geographical location of users sharing disease-related tweets. This can be done by analysing user profiles, geotags or other location based info.



- ⑤ Temporal Analysis:  
time series Analysis.  
analyze the temporal aspect of the data to detect spikes or trends in disease related mentions over time.
- ⑥ Anomaly Detection:  
Statistical Analysis:  
employ statistical methods to detect unusual patterns in the data. Sudden spikes in the no. of disease-related tweets could be indicative of an outbreak.
- ⑦ Visualization:  
Create Visualizations:  
Develop visual representations of the data to communicate findings effectively.
- ⑧ Alerts & Reporting:  
Automated Alerts:  
Implement automated alert sys. that triggers notifications when significant changes or spikes in disease-related mentions are detected.
- ⑨ Verification:  
Cross-reference with Official Data:  
verify the social media findings with official health org. data.
- ⑩ Action.  
Public Health response:  
if an outbreak is confirmed, public health authorities can take appropriate measures.



Q2) Illustrate a case study on the use of social media analysis as a source of public health info.

⇒ ① title: twitter analysis of the COVID-19 Pandemic: A source of public health info.

② Methodology:

Data Collection -

- twitter data related to Covid-19 was collected using the twitter app. keywords such as Covid-19, coronavirus, pandemic. were used to filter tweets.

Data Preprocessing:

- the collected tweets were cleaned to remove retweets, duplicate content, & non english tweets.

Sentiment analysis was performed to categorize tweets as positive / negative / neutral.

Topic Modeling:

latent dirichlet allocation (LDA) was employed to identify prevalent topics within Covid-19 related tweets.

Sentiment Analysis:

sentiment analysis revealed the public's emotional response to the pandemic which can be indicative of stress, fear or hope.

③ Finding:

Emerging topics

Sentiment trends

Regional Insights.



- ④ Impact:
- Early Warning Syst:
  - Resource Allocation.
  - Public Engagement.

Q3) list popular online patient communities & then discuss the analysis of data from.

- ⇒ ① Popular Online Patient Communities:
- there are several online patient communities that provide platforms for patients to share their experiences, seek advice & find ~~spc~~ support.
- ② PatientLikeMe:
- this is a patient powered ~~resource~~ <sup>research</sup> n/w that aims to improve lives & contribute to medical research.
- ③ Smart Patients:
- this is an online community where patients & caregivers learn from each other about treatments, clinical trials, the latest science & how it all fits.
- ④ HealingWell:
- It is a thriving community where patients & caregivers can connect on ~~more~~ than 200 health communities.
- ⑤ Daily Strength:
- This online community provides support groups for people who are facing similar issues members can express their thoughts.



Q4) list popular online doctor communities & then discuss the analysis of data for one particular disease & drug-associated genes that are being tested from!

⇒

① Popular Online Doctor Communities:

Several online communities cater specifically to doctors & healthcare professionals. They offer platforms for sharing knowledge, discussing case studies & networking.

② Sermo:

this is a virtual doctors lounge where licensed physicians can anonymously talk real world medicine.

③ Doximity:

It's the largest comm of healthcare prof. in the country, with over 70% of US doctors as verified members.

④ Medscape:

Medscape provides a platform for physicians & healthcare professionals to access medical content, including the latest medical news & expert perspectives.

@

⑤ Quantamind:

this platform offers a collaborative approach to learning & practice growth. Doctors can share their insights & learn from each other.



Q5) list the various syndromic surveillance sys. based on social media & then discuss the analysis's from...

① Syndromic Surveillance sys based on social media.  
Several syndromic surveillance sys leverage social media data to track & predict disease outbreaks.

② HealthMap: <sup>Outbreaks</sup> Monitoring & real time surveillance of emerging public health threats.

③ ProMED-mail:

The program for monitoring emerging diseases (ProMED-mail) is a global electronic reporting sys.

④ Flu Near You:

this uses crowd sourced data to provide real time tracking of flu activity.

⑤ SickWeather:

This is a social health n/w & real time map of sickness & symptoms in your local area, which allows you to check for the spread of illness in your local area.

⑥ BioCaster:

this is a health-related news aggregator that uses ml to track the spread of disease outbreaks.