



# Distilling from Twitter: New Perspectives in Healthcare Organizations Using Association Rule Mining

Aditya Singhal  
Student ID: 1154832

Course: (2023W) COMP-5112-WA - Research Methodology Comp Sci

## CONTENTS

- What is Association Rule Mining?
- Critical Review
- Main Objectives
- Research Overview
- Data Set
- Example





Association rule mining is a rule-based machine learning method for discovering interesting if/then statements that help uncover relationships between variables in large databases.

# Literature Review



- *[1] Usage of Twitter for health literacy promotion*

- *[2] COVID-19 related health beliefs on Twitter*

- *[3] Frameworks for monitoring Social media via content analysis*

- *[4] Importance of public engagement on Twitter*

- [1] Zhou, J.; Liu, F.; and Zhou, H. 2018. Understanding health food messages on Twitter for health literacy promotion. *Perspectives in public health*, 138(3): 173–179.
- [2] Wang, H.; Li, Y.; Hutch, M.; Naidech, A.; Luo, Y.; et al. 2021. Using tweets to understand how COVID-19–Related health beliefs are affected in the age of social media: Twitter data analysis study. *Journal of medical Internet research*, 23(2): e26302.
- [3] Mendhe, C. H.; Henderson, N.; Srivastava, G.; and Mago, V. 2020. A scalable platform to collect, store, visualize, and analyze big data in real time. *IEEE Transactions on Computational Social Systems*, 8(1): 260–269.
- [4] Singhal, A.; Baxi, M. K.; and Mago, V. 2022. Synergy Between Public and Private Health Care Organizations During COVID-19 on Twitter: Sentiment and Engagement Analysis Using Forecasting Models. *JMIR medical informatics*, 10(8): e37829.

# Literature Review

## Association Rule Mining

- *[5] ARM to analyze human behaviour on social media*

- *[6] ARM for topic extraction on social media*

- *[7] ARM for making recommendations*

- [5] Raihan, M.; Islam, M. T.; Ghosh, P.; Hassan, M. M.; Angon, J. H.; and Kabiraj, S. 2020. Human behavior analysis using association rule mining techniques. In *2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, 1–5. IEEE.
- [6] Koukaras, P.; Tjortjis, C.; and Rousidis, D. 2022. Mining association rules from COVID-19 related twitter data to discover word patterns, topics and inferences. *Information Systems*, 109: 102054.
- [7] Orama, J. A.; Borra's, J.; and Moreno, A. 2021. Combining cluster-based profiling based on social media features and association rule mining for personalised recommendations of touristic activities. *Applied Sciences*, 11(14): 6512.

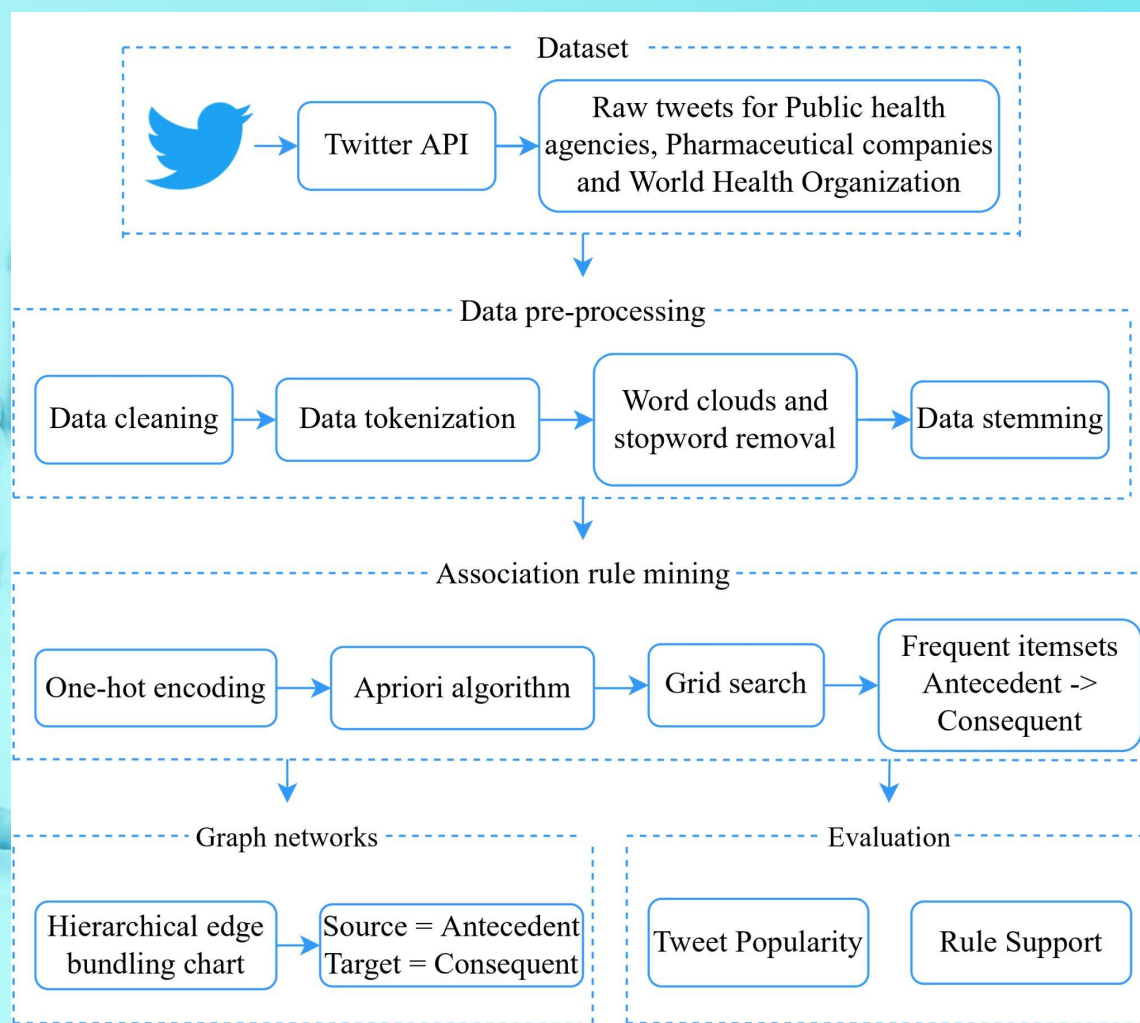




## Objectives:

- *Examine Twitter usage by US and Canadian health agencies and pharmaceutical companies*
- *Identify text patterns influencing the tweets' content.*





# Data Set



Name of organization (Twitter handle)	Total tweets, N
Public health agencies	
Centers for Disease Control and Prevention (CDCgov)	7,511
Indian Health Service (IHSgov)	1,632
Health Canada and PHAC (GovCanHealth)	52,695
Government of Canada for Indigenous (GCIndigenous)	3,725
<b>Total</b>	<b>65,563</b>
Pharmaceutical companies	
AstraZeneca (AstraZeneca)	1,284
Glaxo SmithKline (GSK)	2,359
Johnson & Johnson (JNKNews)	2,368
Novartis (Novartis)	715
Pfizer (pfizer)	2,474
<b>Total</b>	<b>9,200</b>
Non-governmental organization	
World Health Organization (WHO)	24,581

Table 1: Number of tweets for each organization.

### Tweet text

UPDATE: If you are fully vaccinated against #COVID19, you can resume activities without wearing a mask or staying 6 feet apart, except where required by federal, state, local, tribal or territorial laws, incl. local business and workplace guidance. More: <https://t.co/FJMon7WIFO>

### Tokenized text

['updat', 'fulli', 'vaccin', 'covid', 'resum', 'activ', 'without', 'wear', 'mask', 'stay', 'feet', 'apart', 'except', 'requir', 'feder', 'state', 'local', 'tribal', 'territori', 'law', 'incl', 'local', 'busi', 'workplac', 'guidanc']

Figure 2: Sample tweet from the Centers for Disease Control and Prevention (CDC).



***Thank you!***

***Any questions?***