Ankit Aich and Natalie Parde

Department of Computer Science   The University of Illinois at Chicago   {aaich2 , parde} @ uic.edu   ankitaich.com, natalieparde.com
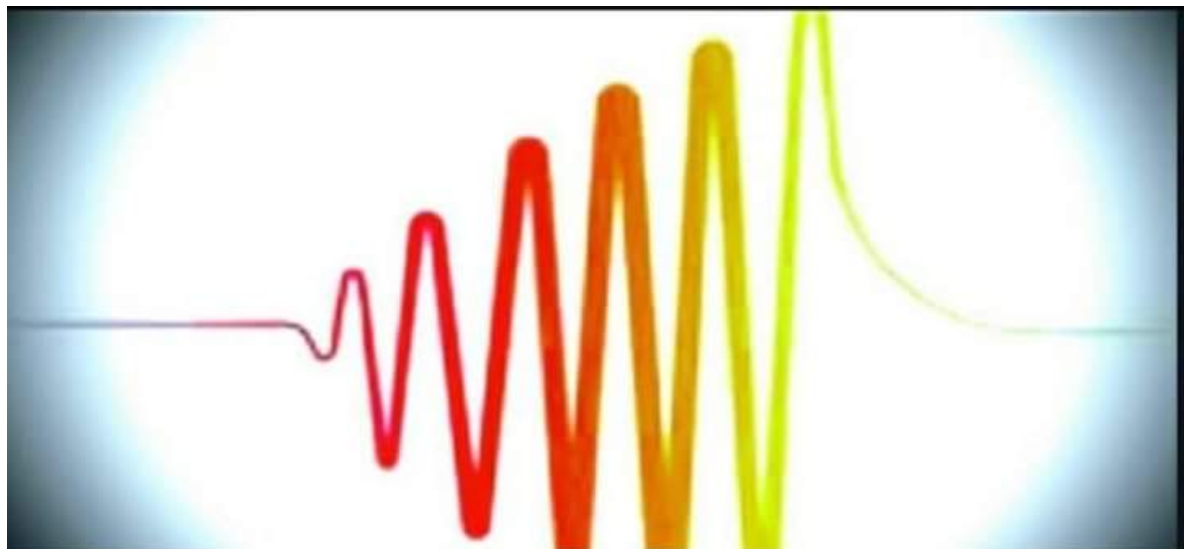
Case Studies

INDIA-Science behind 5th April 2020 lamp ritual

Scientific reason behind 5th April 2020 to light lamps , diyas candles at 9 pm for 9 minutes . Number 9 represents mars , number 5 mercury and 2 represents moon .
Number 9 represents planet Mars . At present mars is in makar rashi with planet Jupiter and Saturn . Moon is exactly situated at 180 degrees from these 3 planets .
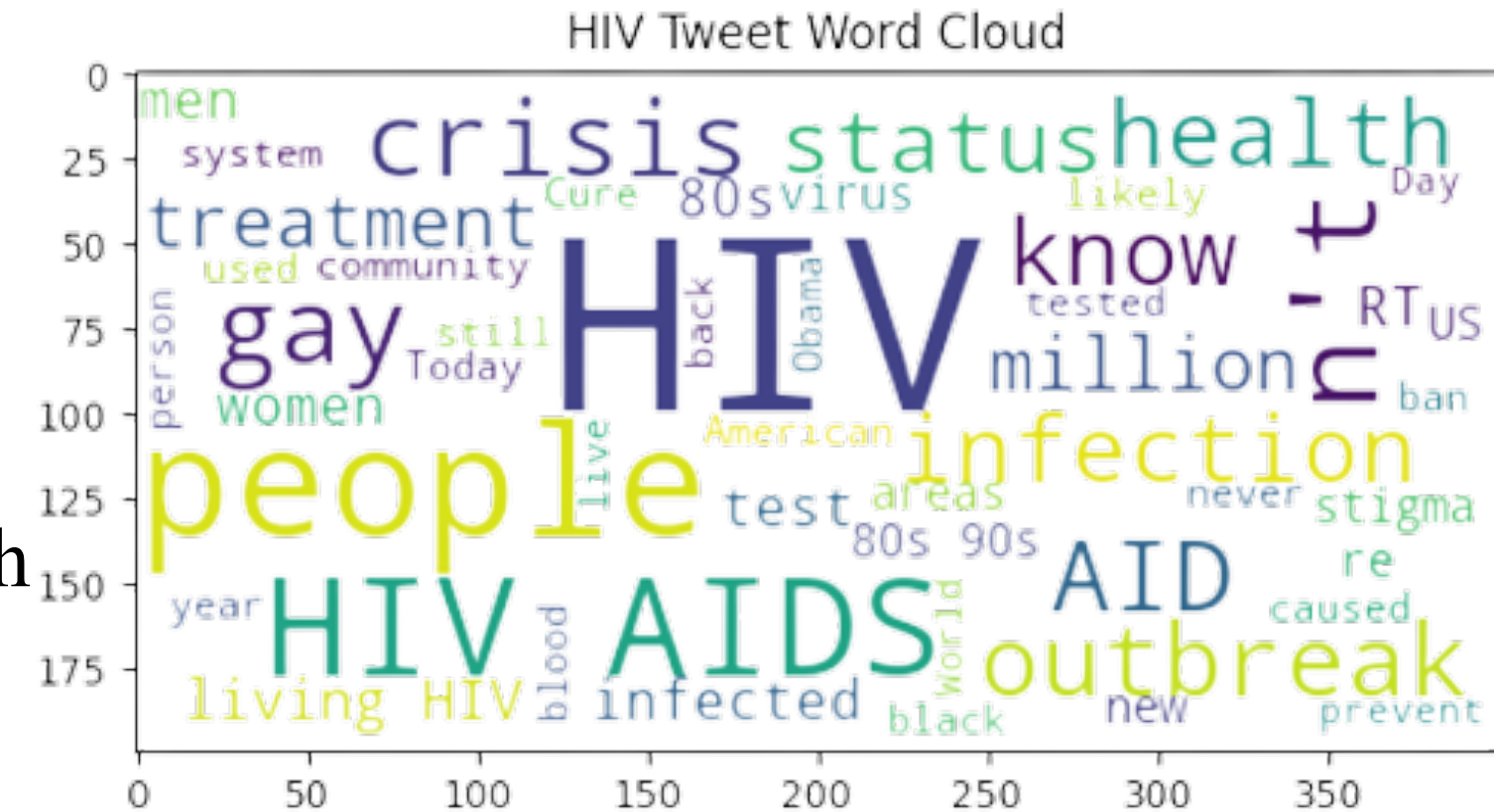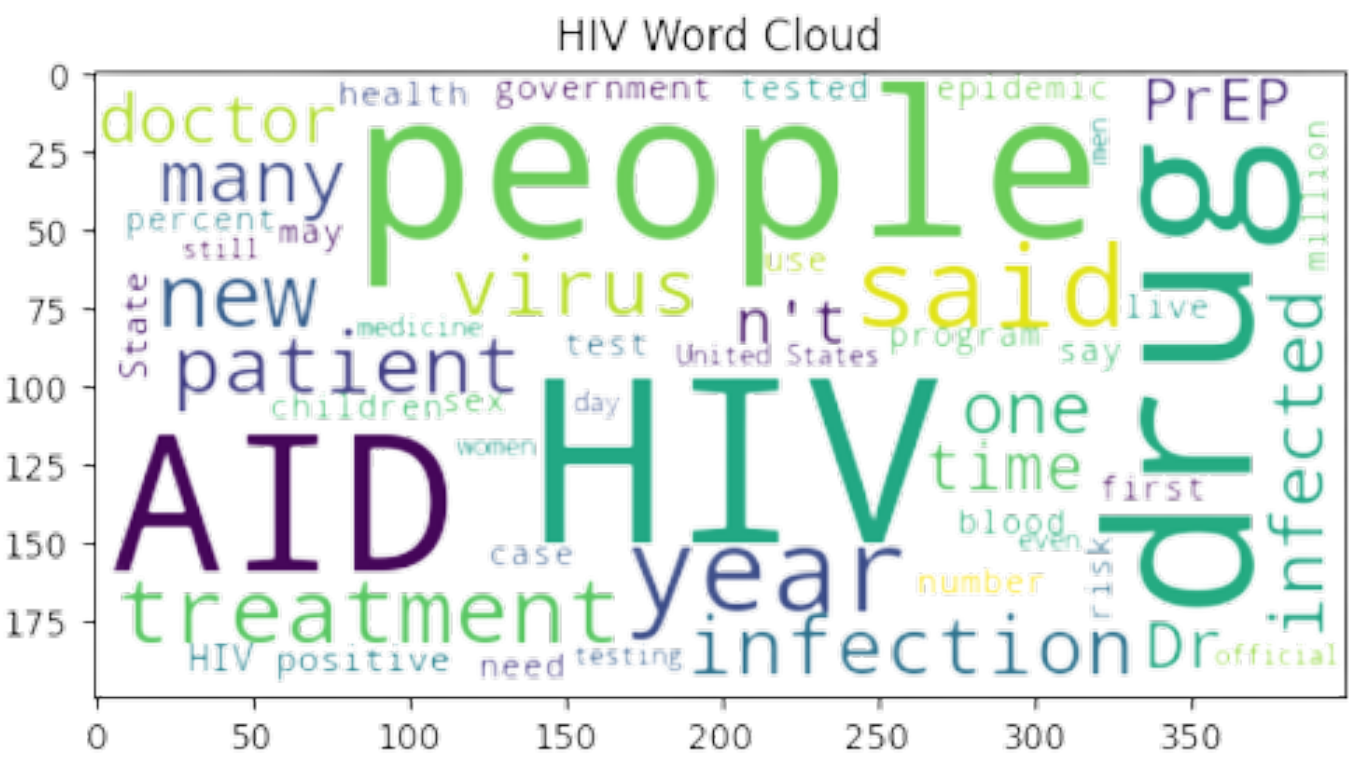
The conjunctions of moon and mars will create an atmospheric spark at 9.00 pm on 5th April 2020 that will increase the magnetic field of planets earth . Plus with lamps lightened and electric lights closed . It will increase the power of magnetic field plus moon will be at the shortest distance from earth . Moon is considered to control mind and water in human body .
It will increase the human immunity in the Indian subcontinent.

it will produce oxygen from lamps though in small quantities from each lamp but if we consider 50 percent of Indian population lighting lamps imagine the oxygen level in environment plus the magnetic field will use this oxygen reduce the impact of corona virus from atmosphere and help in respiratory immunity of human body and This combination of moon and mars help nature's fight against pandemic covid 19

Donald J. Trump
@realDonaldTrump

These are the things and events that happen when a sacred landslide election victory is so unceremoniously & viciously stripped away from great patriots who have been badly & unfairly treated for so long. Go home with love & in peace. Remember this day forever!

11:01 PM · Jan 6, 2021

Motivations

1. Widespread misinformation during the COVID-19 pandemic.
2. A lack of data from previous global health events.
3. Lack of articles from print media.

HIV Tweet Word Cloud
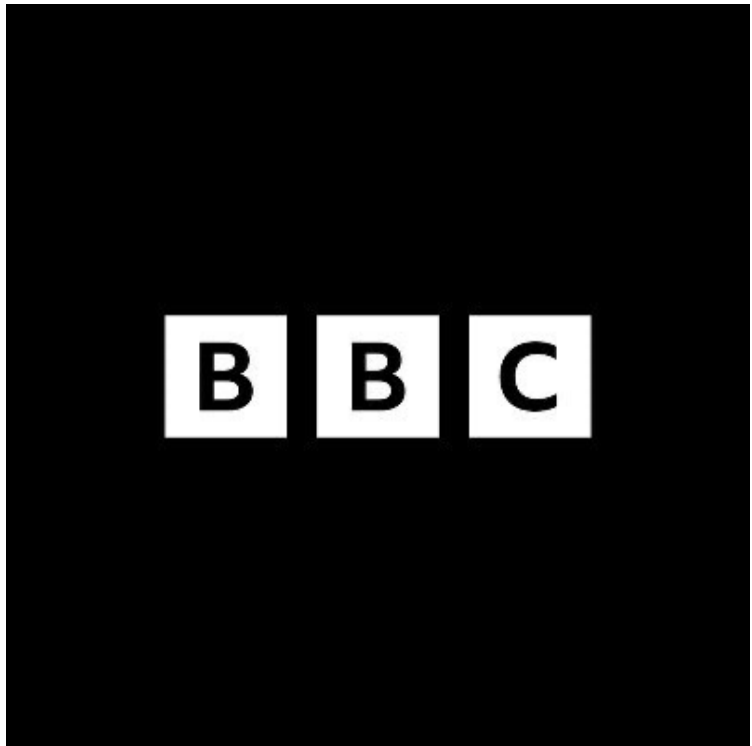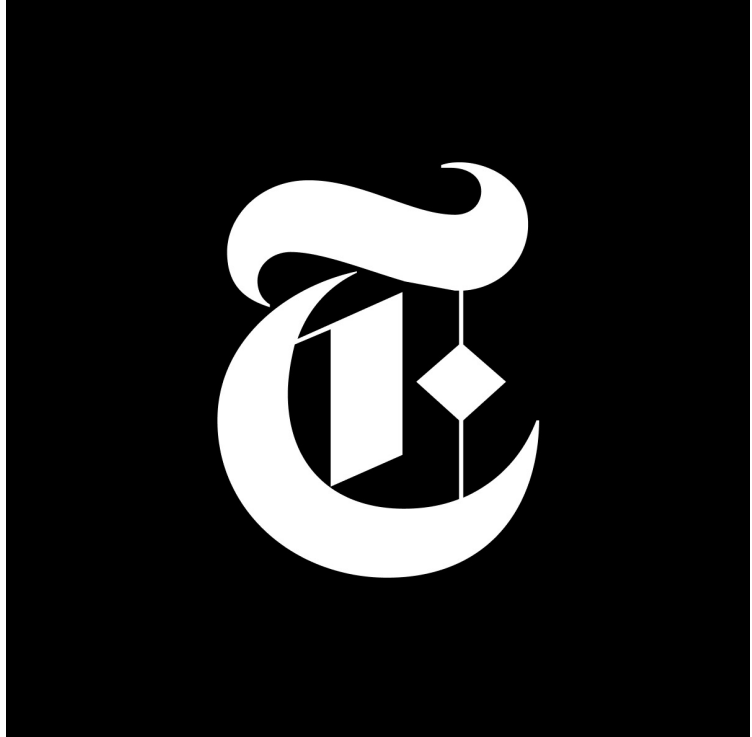

HIV Word Cloud

Linguistic Differences between information and misinformation

Misinformation can be an innocent mistake, but it's still dangerous.

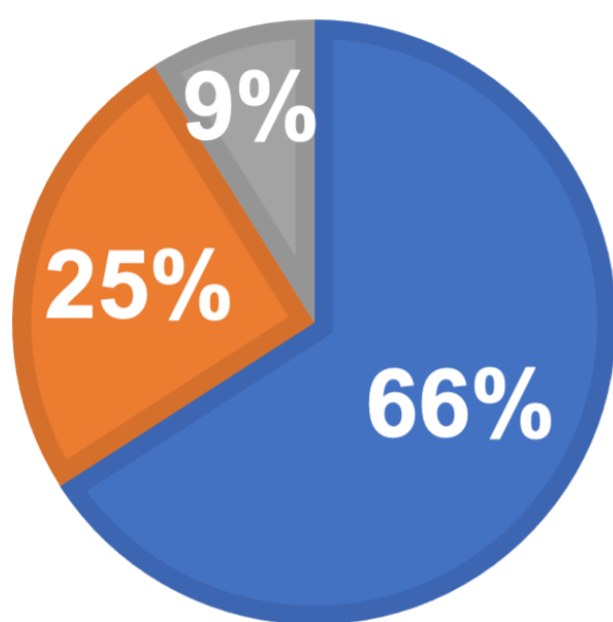Disinformation is dangerous, plus it serves someone else's agenda.

Misinformation -
Might be a lack of intent – *e.g., spread from social media for* concern. – still harmful. – can cause widespread effects.

Disinformation -
Has a presence of intent – caused due to someone else's agenda. Almost always has a harmful side effect(s)

1. 3 Million articles from 8 different sources
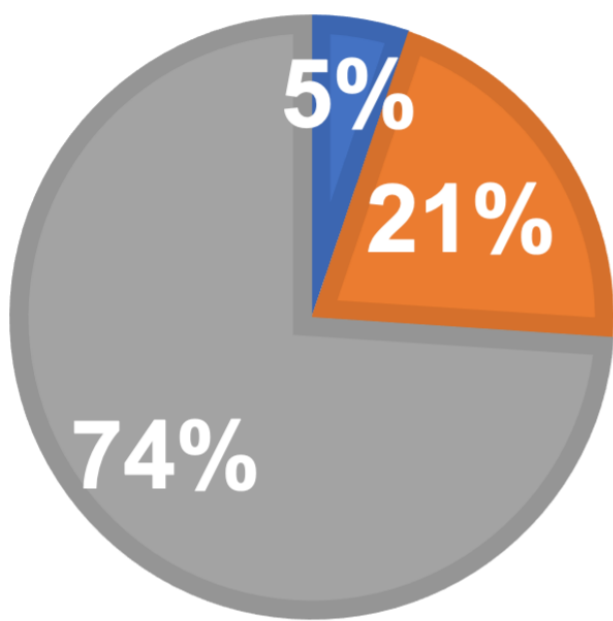2. Mix of print media, social media, online media and health centers.
3. Articles from 1900 to 2021
4. Average of 330 words and 6800 characters

SOCIAL MEDIA ANNOTATION DISTRIBUTION
■ Class 1 ■ Class 2 ■ Class 3

9% / 25% / 66%

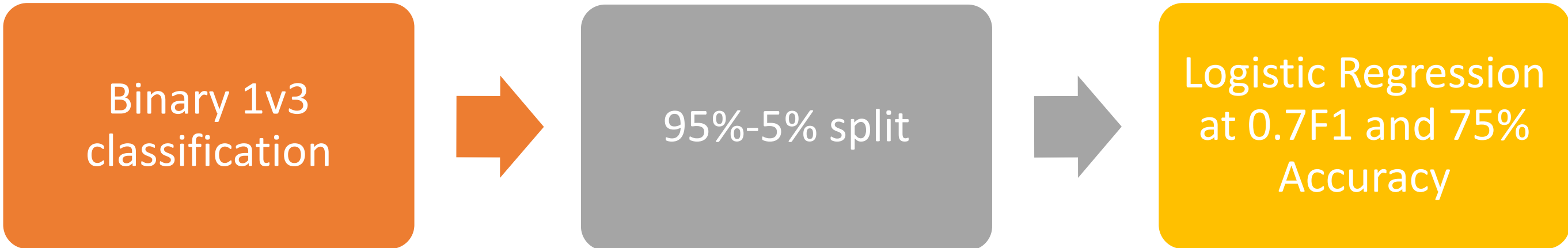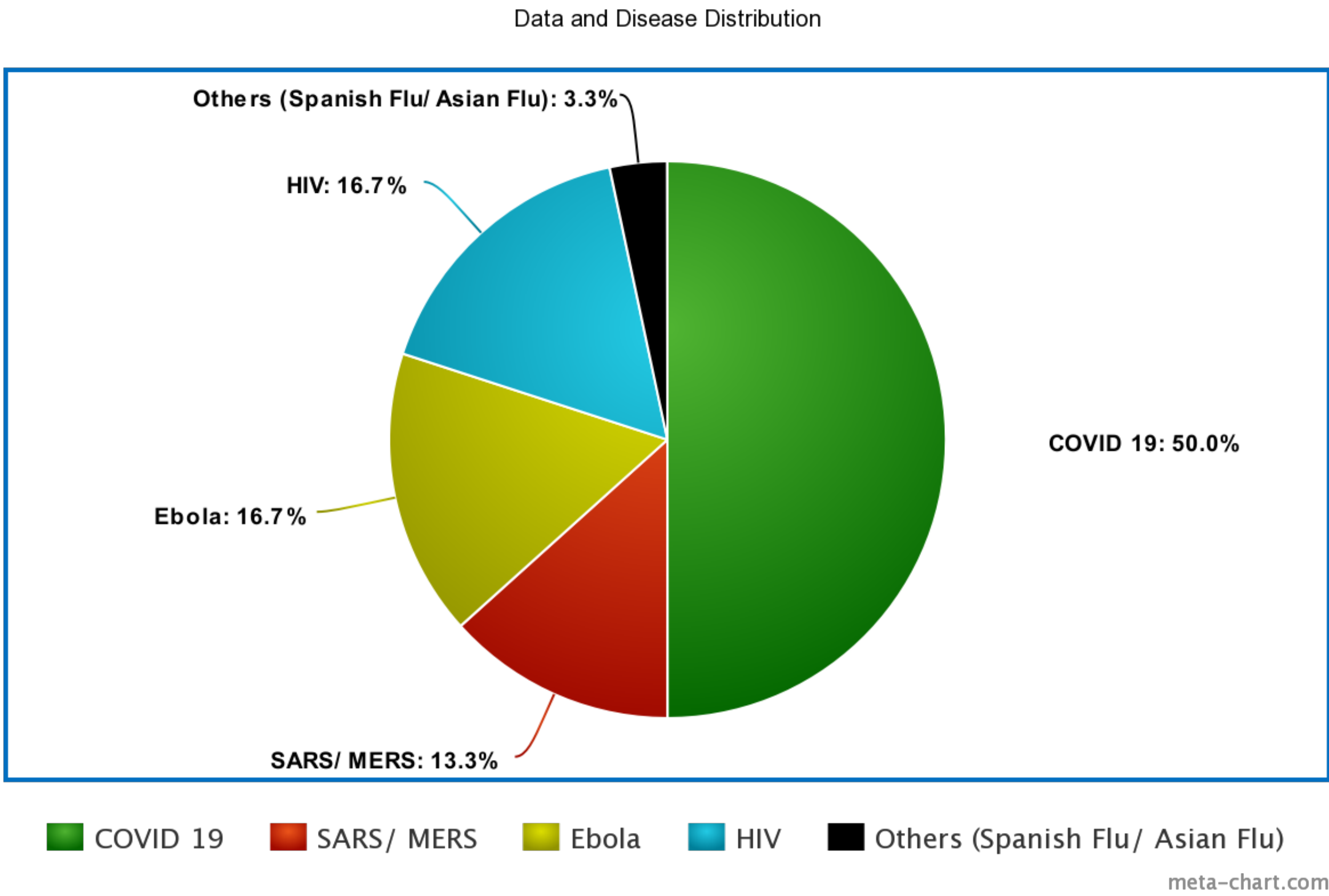| Description	Class
Personal opinions, known misinforma-tion, or information unverified during fact-checking	1
Information verified by < 5 sources	2
Known facts, or information supported by ≥ 5 sources	3

NEWS ANNOTATION DISTRIBUTION
■ Class 1 ■ Class 2 ■ Class 3

5% / 21% / 74%

Description	Class
Known misinformation, or information unverified during fact-checking	1
Information verified by < 4 sources	2
Known facts, or information supported by ≥ 4 sources	3

$$\kappa = 0.81$$

Our Data in Statistics

Data and Disease Distribution

Others (Spanish Flu/ Asian Flu): 3.3%
HIV: 16.7%
Ebola: 16.7%
COVID 19: 50.0%
SARS/ MERS: 13.3%

■ COVID 19 ■ SARS/ MERS ■ Ebola ■ HIV ■ Others (Spanish Flu/ Asian Flu)

meta-chart.com

Feature	Description
Trust Score	Number of Words in NRCLexicon under trust
Readability Index	$RI = 4.71\frac{c}{w} + 0.5\frac{w}{s} - 21.43$
Ease of Reading	$ER = 206.835 - 1.015\frac{w}{s} - 84.6\frac{l}{w}$
Difficult Words	$N(w) \; \forall \; w \in$ *TextStat Easy words* or w has ≥ 2 syllables
+NRC Scores	Number of Positive Words in NRC Word Emolex
-NRCScores	Number of Negative Words in NRC Emolex

C = Length of String, W = Number of Words, S = Number of Sentences in a corpus, l = Number of Syllables

Binary 1v3 classification → 95%-5% split → Logistic Regression at 0.7F1 and 75% Accuracy

1. **First work to establish misinformation dataset spanning all major health events.**
2. **In depth analysis shows recurring themes and features across misinformation and fact.**
3. **Can be used to train NLG systems for future fine grained misinformation detection systems.**

Future Directions