

Misinformation Spread in Social Networks

Problem Statement



In the Information Age, social networking sites have become a notable agent for the spread of misinformation, fake news, and propaganda. Misinformation on social media spreads quickly in comparison to traditional media because of the lack of regulation and examination required before posting. These sites provide users with the capability to spread information quickly to other users without requiring the permission of a gatekeeper such as an editor who might otherwise require confirmation of the truth before allowing publication. A 2018 study of Twitter determined that, compared to accurate information, false information spread significantly faster, deeper, and more broadly.

In this study, we aim to better understand misinformation flows in social networks.

Related Work

The spread of misinformation has been an ongoing research topic for some time now but before understanding the spread of misinformation, we need a clear understanding of what constitutes misinformation. It usually refers to false and inaccurate information that may spread unintentionally or intentionally[2]. Substantial research indicates that social media is responsible for major web-based Misinformation [1], this is cause for major concern as more than 4.41 Billion people are projected to be social media users by 2025 [3]. To minimize the impact of misinformation on social media, many researchers are trying to understand how misinformation spreads, specifically its diffusion patterns. Based on data from Twitter, Oh et al [4] examined the working dynamics of rumors related to the 2010 Haiti disaster and discovered that informational uncertainty and anxiety are major components that affect the quick propagation of a rumor. Furthermore, they suggested that reliable information from reliable sources could lessen anxiety on Twitter, hence limiting the propagation of rumor. Domenico et al [5] investigated and provided a model for the dissemination of a scientific rumor concerning the Higgs boson. Individuals were more inclined to propagate the story if the majority of their friends tweeted it regularly, according to the researchers.

Other studies have analyzed how people distribute fake news in political elections using a combination of web browsing data and online survey data. According to the findings, the public's behavioral preferences have a major impact on misinformation trust, and the public is more ready to believe stories about politicians they support, even if the stories are untrue and full of improbable components [6].

Research Questions

Our study takes misinformation as its subject and examines the network structure of misinformation and factual information on social networks in order to understand how misinformation spreads and evolves on social media, as well as what characteristics make misinformation more successful. To achieve this goal, we propose the following research questions:

1. Is there any significant difference between spread patterns of information based on its veracity?
2. Do sentiments affect misinformation spread and if they do then what is the correlation between different sentiments and diffusion?
3. Does the presence of pictures, videos and weblinks contribute to how misinformation is received?

Methodology pipeline

We have considered tweets from twitter to perform our analysis.

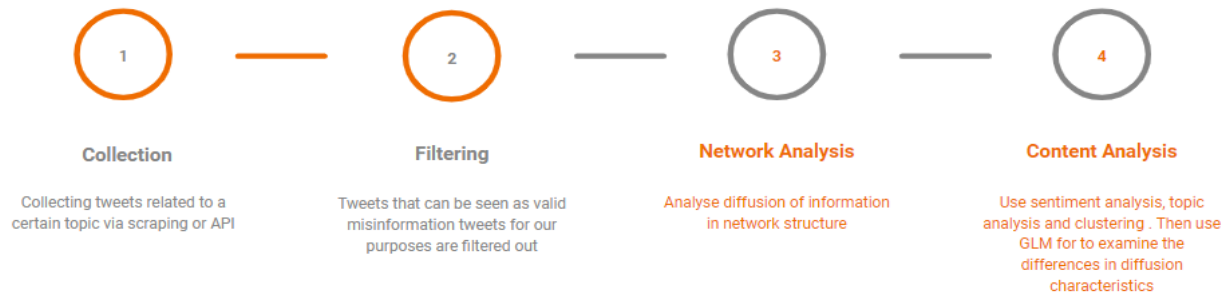


Figure 1. Methodology pipeline

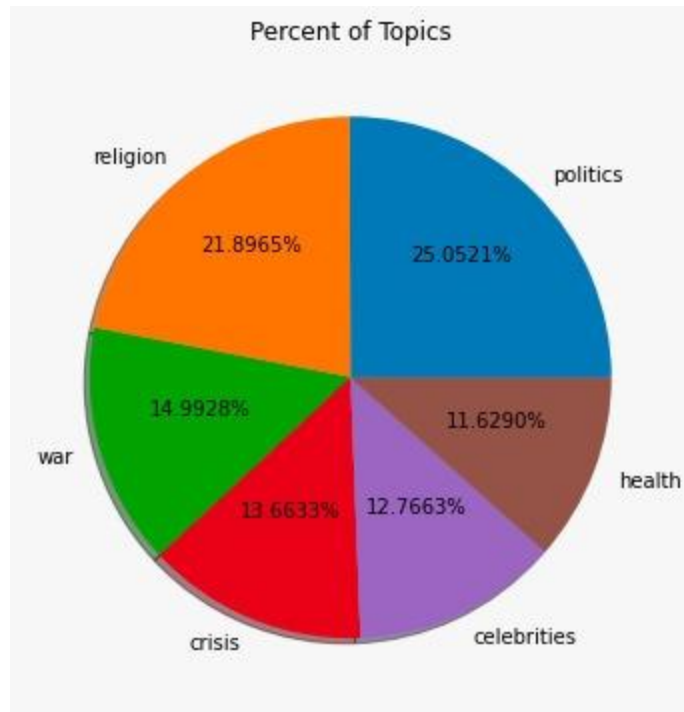
Data Collection

Survey

We conducted an online survey as a quantitative study to gather more depth on the opinions of people. We obtain a statistical representation of the responses of our participants when it comes to how they would classify a piece of information, tweet or picture, as true or false information and if they would further share that information. Our sampling method is non-probabilistic and we do convenience sampling.

Collecting Tweets

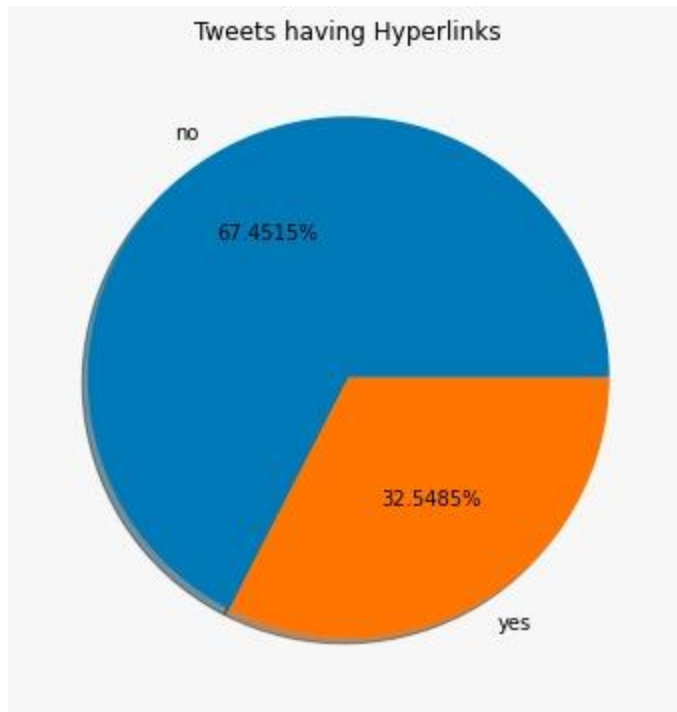
Twitter API is used to extract tweets from twitter which we will use in our analysis. During this extraction, no personal information is collected. Hence, the privacy of Twitter users is maintained. We collect tweets with hashtags like #Bollywood, #UkraineRussiaWar, #Covid-19, #health, #antivaxx, etc which are prone to misinformation.



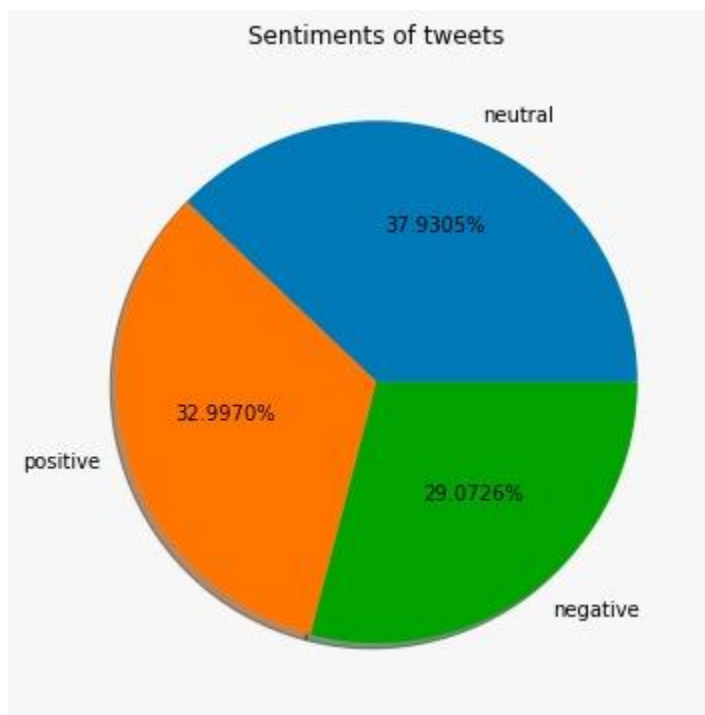
Classification

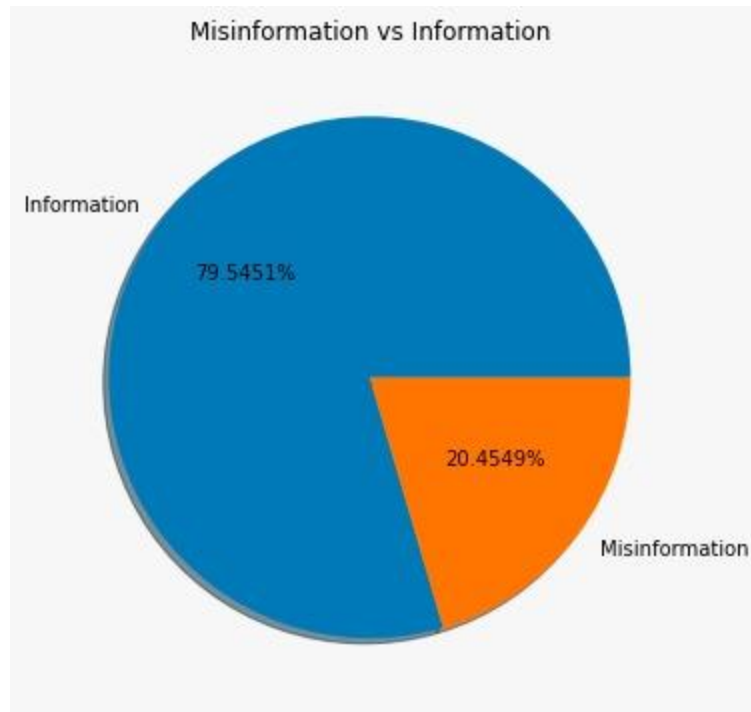
We use the “3 S” approach for classifying tweets as having information or misinformation

- 1. Subjectivity:** Tweets can be subjective or objective. Subjective tweets can be a person’s opinion, feelings, personal anecdotes, etc. Example: “Is it just me or does my dog look sad today?” There is no notion of information or misinformation for subjective tweets. Objective tweets, on the other hand, contain information. We first classify tweets as subjective or objective. Subjective tweets are discarded and objective tweets are further examined for information or misinformation.
- 2. Sources:** It is a common tactic to include links in tweets to establish credibility of the tweet content. However, all links are not equal. If a tweet has a trustworthy link (reputed sources), there is very less probability of the tweet content being misinformation. If a tweet has a trustworthy link, we classify it as information with high confidence.



3. **Sentiment Analysis:** For tweets which don't have any links, we run sentiment analysis. Tweets having misinformation usually use strong language to grab people's attention. They generally have a click-baiting vibe to them. So if an objective tweet without any trustworthy links is too positive or too negative, there's a very high probability that it contains misinformation





Network Analysis

Network Analysis helps us in understanding the structure of a relationship in social networks, a structure or process of change in natural phenomena. We perform our study and analysis on the data thus collected. We use a social network analysis method followed by cascading network analysis.

Diffusion network

Diffusion Network tries to capture the underlying mechanism of how events propagate through a complex network, whether the spreading of some social movement, some new fashion or innovation or it may be a marketing message through an online social network.

Content Analysis

T-Test:

A t-test is an inferential statistic that is used to see if there is a significant difference between the means of two groups that are similar in some ways. In statistics, the t-test is one of several tests used to test hypotheses. The T-Test or the Student's test is used to reject a null-hypothesis or

not reject it. Three crucial data values are needed to perform a t-test. They contain the mean difference, each group's standard deviation, and the number of data values in each group. We will be using the independent T-Test to determine the diffusion characteristics and draw a comparison, showing that there is a difference in the tweets, likes or comments between the misinformation and the true information. The t-value is found using formula:

$$t = (\text{Difference between group means}) / (\text{Variability of groups})$$

Correlation:

We make a correlation graph depicting the emotions being factors of how likely that influences the spread of misinformation. We use the Pearson's coefficient as a metric of correlation. It uses covariance and gives the relationship between the variables.

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

General Linear Model

The General Linear Model (GLM) is a useful framework for comparing how several variables affect different continuous variables. It is mainly used for model specification. It helps us to arrive at the exact equation that will most accurately summarize the data, thereby allowing us to summarize the research outcomes.

Ablations

The previous work doesn't account for the star topology that arises in the network due to forwarding relationship limitations of the twitter API. Therefore we don't get an accurate representation of the network structure. To solve this, we use a custom softmax function based on the influence of the user and the time difference of retweeting. We use a cascading network where the information cascade V of size n is defined as a series of messages v_i sent by user u_i at time t_i , i.e. $V = \{v_i = (u_i; t_i)\}$. We model our network as a cascade of tweets so that our graph does not suffer from the problem of not knowing the duration of a tweet: we now know how a tweet dissipates in a network and we prevent the star topology from being too prominent.

$$p_{ij} = \frac{m_i e^{-r(t_j - t_i)}}{\sum_{k=1}^{j-1} m_k e^{-r(t_j - t_k)}}$$

Results and Discussions

After collection of data, we performed some tests to analyze the spread of misinformation in our network based on certain diffusion characteristics .

This diffusion is observed by using a **Fruchterman-Reingold layout** as the graph representation of the tweets to obtain further characteristics of the diffusion network. The results from the survey show that people tend to share more information than misinformation.

T-Test

According to our null hypothesis, there is no statistically significant difference between the number of favorites, comments and retweets for the true information and misinformation. On performing the T-Test on our dataset, we get t-values 1.52, 1.92 and 1.67 for retweets, favorites and comments. We get 8 degrees of freedom, and a significance level of 5% as a one-tailed test ($p=0.05$). Our critical value is 1.86. Thus, our t-value is greater than our critical value in the number of favorites, so we can reject our null hypothesis for favorites and accept it for the other two cases.

This shows that the **diffusion of misinformation in favorites is more widespread** than the other categories.

	Attribute	Retweet true info	Retweet misinfo	Favourite true info	Favourite misinfo	Comments true info	Comments misinfo
0	Mean	61832.200000	10178.400000	11364.000000	1918.800000	210267.800000	37539.600000
1	Max	206891.000000	27585.000000	11364.000000	1918.800000	210267.800000	37539.600000
2	Min	4886.000000	1647.000000	1047.000000	89.000000	16810.000000	4171.000000
3	Std	75435.728679	9047.437617	10920.036465	1472.532838	228218.530014	36378.919465
4	Skew	1.843974	1.812885	1.370968	0.374266	1.607976	1.932215
5	Kurtosis	-0.130264	-0.079894	-0.511332	-1.449218	-0.309785	0.028726

Figure 2. Table with attributes

- From the above table, Kurtosis value is negative for all fields except comments misinfo, thus there is more **extreme data** diffused through comments.
- Skew values are all positive, this indicates that our extreme data is **spread more to the right** of our bell curve. It shows the level of distortion in our data.
- The number of true information is more than misinformation, thus the true information spreads more than misinformation.

- The number of comments under misinformation is larger because people tend to comment more under posts.
- The mean value with the standard deviation gives the spread of the data, assuming we have a normal distribution of data.
- Divergence of data is observed by this graph. We see that the **maximum divergence occurs because of the difference in the means of true and false information**, especially in the number of comments.
- So the diffusion index is higher in retweets and comments when compared to favorites.
- We see that overall, **true information diffuses more than mis-information**.

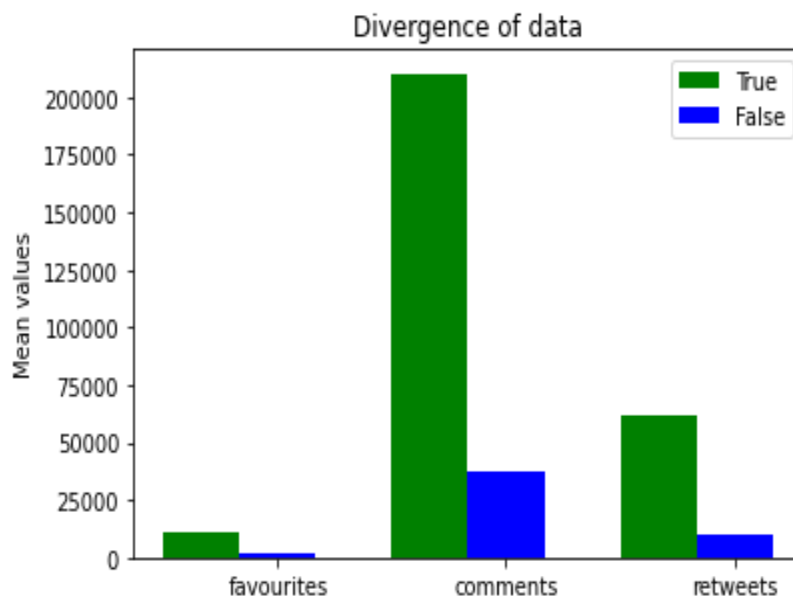


Figure 3. Divergence of data

Correlation

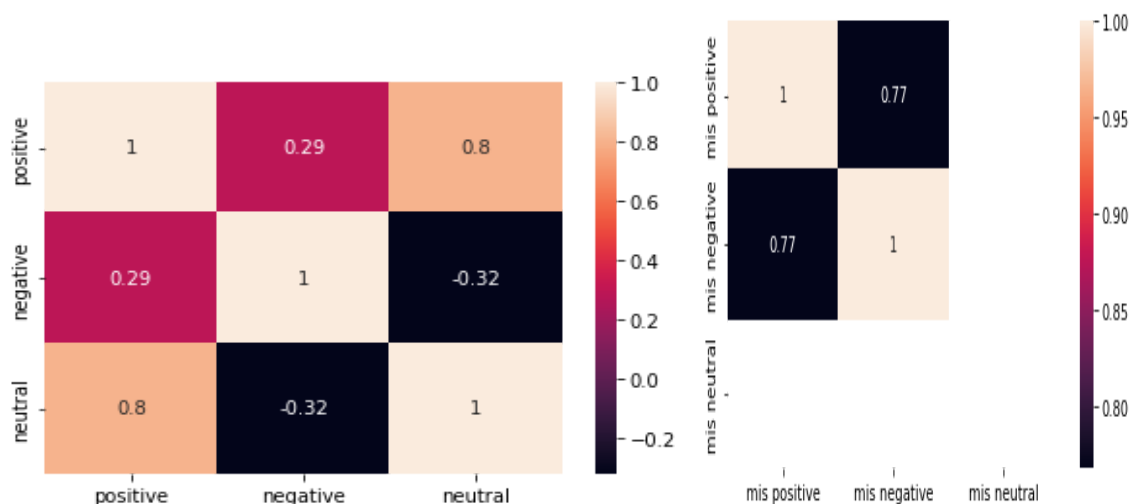


Figure 4. True information matrix

Figure 5. Misinformation matrix

The correlation graph is made on the various sentiments (positive,negative and neutral) of the users. We use the **Pearson's coefficient** as a metric of correlation. It uses covariance and gives the relationship between the variables. Pearson's Point Binomial Correlation Coefficient is used when a variable is dichotomous and this is a special case of the normal Pearson's coefficient.

- We see that a positive correlation indicates that information is **more likely to be retweeted**. From our graph, we see that positive misinformation and positive neutral emotions tweets tend to be more diffused x. Similarly, the misinformation for negative and positive is well diffused.
- A negative correlation is less retweeted.
- This analysis gives the relation between **veracity** and network diffusion characteristics. We see that there is no neutral misinformation in our dataset.
- The closer the relationship, the **stronger the veracity and network diffusion** of the information.

Diffusion Network Structure

We used retweet relationships in our data to analyze diffusion of information in social networks. Through the number of cascades in a tweet we get an estimate of the number of users engaged in diffusion of that information. We then use this network to estimate the relationship between Information veracity and various attributes which might contribute to its diffusion.

We used the Fruchterman Reingold Layout for our diffusion network graph as it shows flow of information between equidistant nodes and topology of forwarding more clearly.

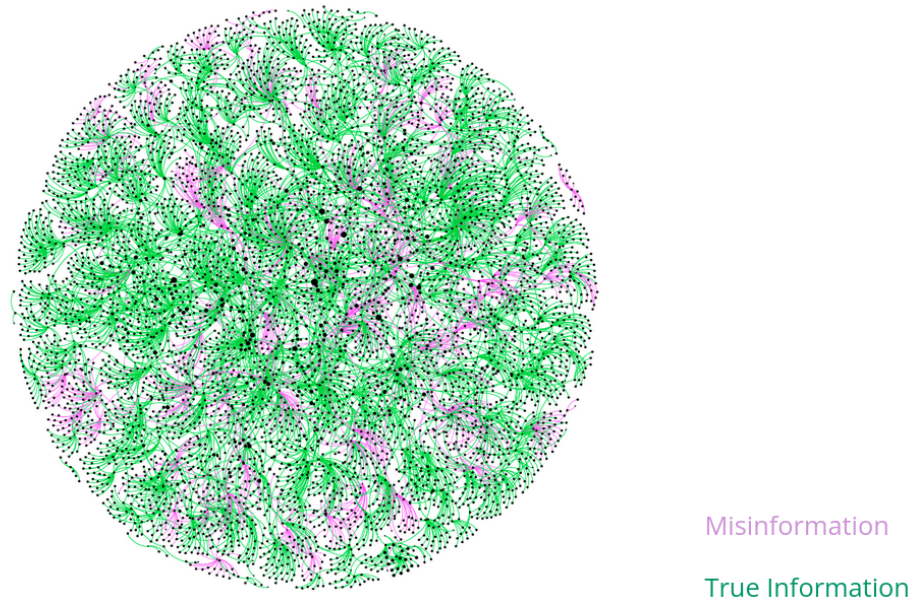


Figure 6. Misinformation vs True Information diffusion network

What we discovered is that truthful information was more widely disseminated and accepted by social media users than misinformation. True information performed better in terms of retweet network range and structural virality, thus disseminated more deeply and extensively than misinformation, according to all of the indexes, and reached a greater audience on social media.

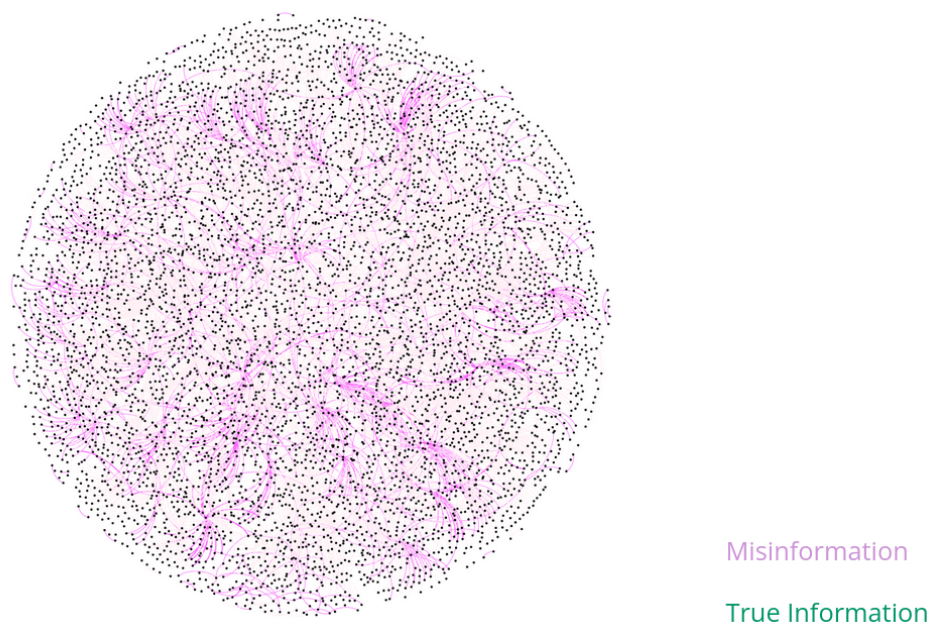


Figure 7. Misinformation diffusion network - lots of cascades

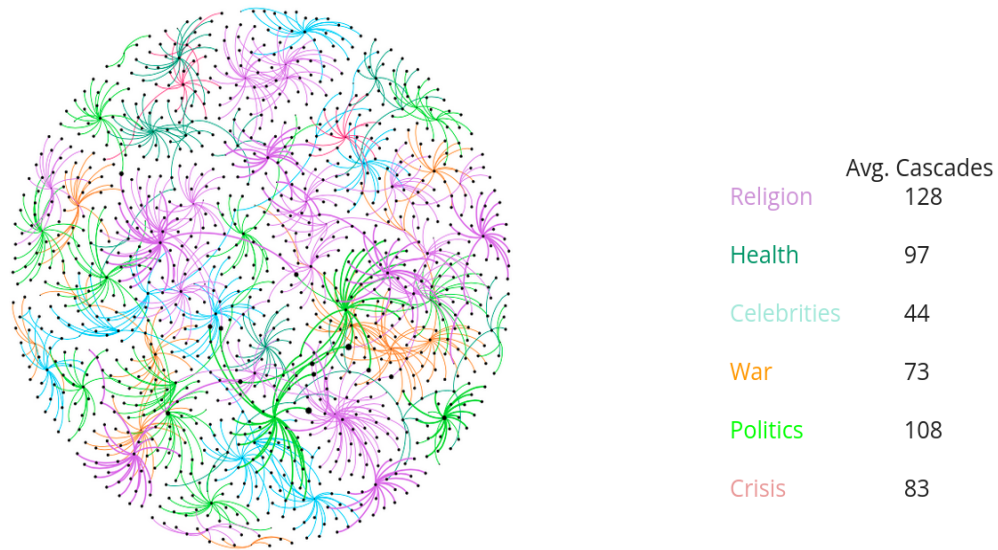


Figure 8. Misinformation spread by topic

However, as compared to correct information, misinformation has an unusual property of having greater cascades. This shows that, while genuine information reaches a larger audience, not all users actively participate in its diffusion, whereas popular misinformation has a larger number of users contributing to its dissemination.

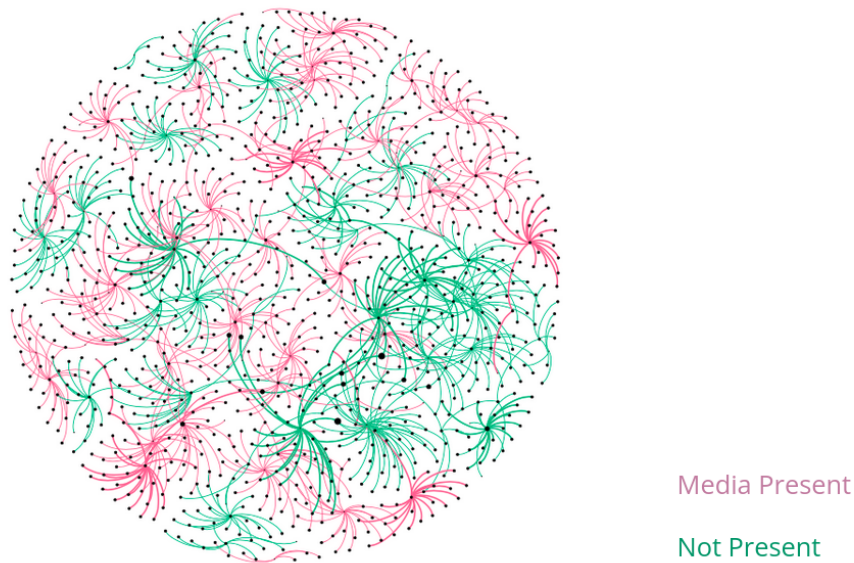


Figure 9. Misinformation spread based on presence of media (video or image)

When it comes to misinformation classification based on media presence, we can see that media presence has a significant impact on tweet popularity. This finding is similar in case of true information as well, suggesting that tweets containing media/hyperlinks are far more likely

to be retweeted by users. Also, we find that misleading tweets containing pictures/videos are relatively well received.

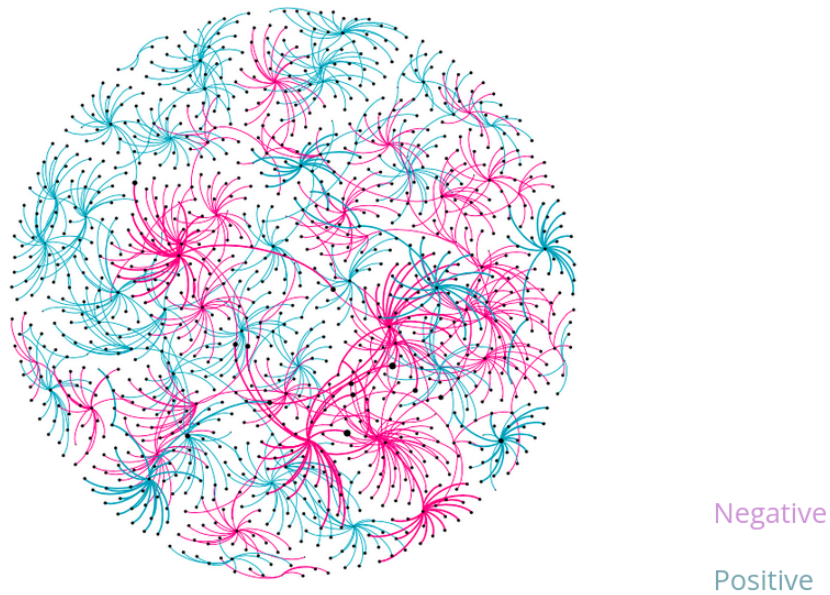


Figure 10. Misinformation spread based on sentiment

We discovered that positive sentiment misinformation is more likely to spread than negative sentiment misinformation after analyzing the attitudes of the tweets. Some examples of positive sentiment include tweets like "Home-remedy for COVID" , "Indian anthem rated by UNESCO" etc. Users appear to not only respond better to these tweets, but are also more likely to share them.

Diffusion Characteristics Based on Veracity

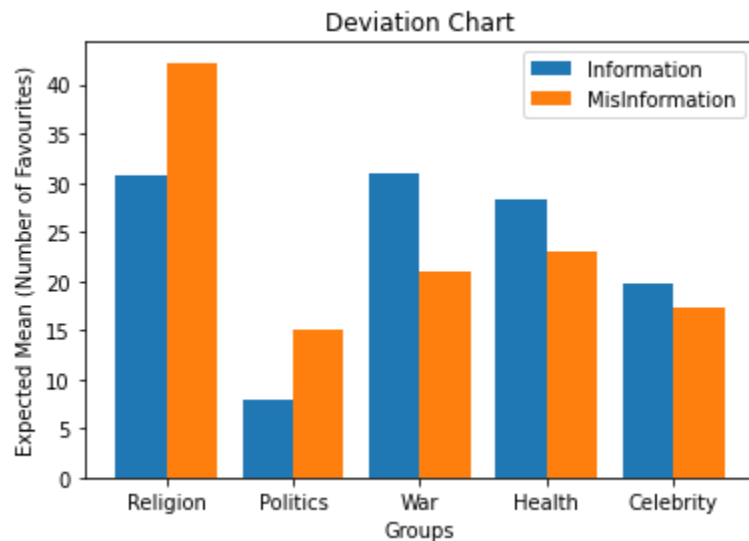


Figure 11. Number of favorites

The GLM model helps perform statistical tests on the data. The above graph tells us that a tweet classified as belonging to religion, based on expected number of favorites is more likely to be spreading misinformation than information. On the other hand, a tweet classified as belonging to health, based on number of favorites is more likely to be spreading information rather than misinformation. In other words the above graph shows that for a particular number of favorites whether a tweet belonging to one topic is spreading information or misinformation. However, for the number of retweets and comments, the diffusion spread is varied and means are at different levels for each category.

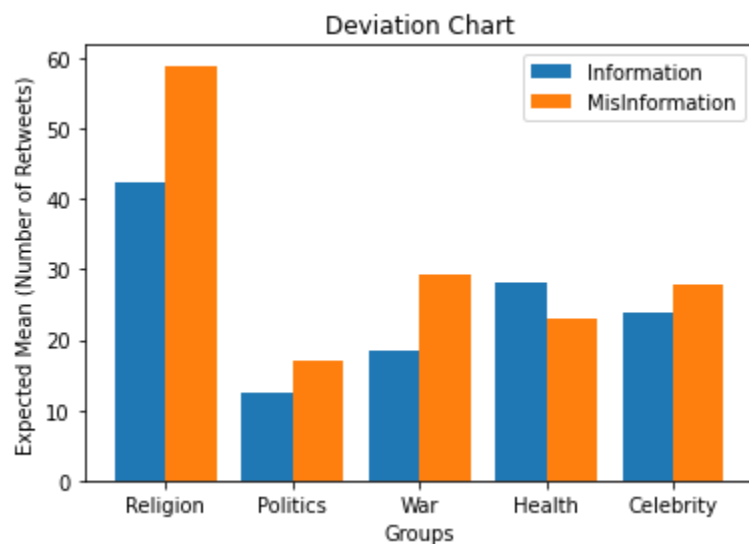


Figure 12. Deviation Chart

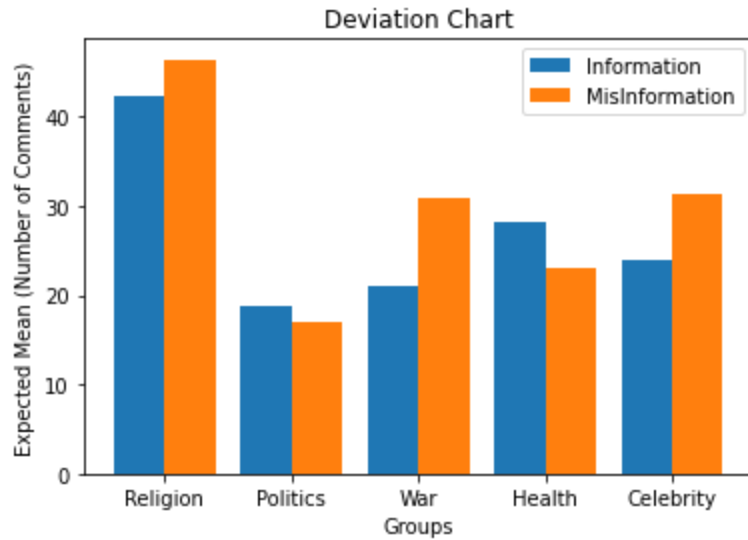


Figure 13. Number of comments

Conclusion

As vital as social media is in our lives, it has played a significant role in the dissemination of disinformation, and Twitter has been chastised for this. Related study has attempted to investigate distinct dissemination features of misinformation on Twitter, but has been severely hampered by the Twitter API's direct retweet capability. With this research, we wanted to gain a better understanding of the diffusion network by looking at how people contribute to the tweet's popularity. With our study, we've discovered that the authenticity of information has an impact on how it spreads. With content and topic analysis, we find that politics, war and religion categories are significantly more prone to misinformation, also religion shows a high number of cascades indicating that a lot more people contribute to its spread. Also, if there is media presence like photo or video, the tweet is more likely to be well received and circulated. Surprisingly, with sentiment analysis, we observe that users are much more likely to share positive misinformation. This may not seem like an issue at first, but can prove to be a matter of concern when it comes to categories like health.

References

- 1 . Fernández-Luque L, Bau T. Health and social media: perfect storm of information. *Healthc Inform Res* 2015 Apr;21(2):67-73
2. Søe SO. Algorithmic detection of misinformation and disinformation: Gricean perspectives. *J Doc* 2018 Mar 12;74(2):309-332.
3. <https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/>
4. Oh O, Agrawal M, Rao HR. Community intelligence and social media services: a rumor theoretic analysis of tweets during social crises. *MIS Q* 2013 Feb 2;37(2):407-426.
5. De Domenic M, Lima A, Mougél P, Musolesi M. The anatomy of a scientific rumor. *Sci Rep* 2013 Oct 18;3:2980
- 6.H. Allcott and M. Gentzkow, "Social media and fake news in the 2016 election," *The Journal of Economic Perspectives*, vol. 31, no. 2, pp. 211–236, 2017.