

Categorization and Prediction of Persuasive Evidences on Reddit

Ram Rattan Goyal, Dr. Dana McKay

June 2022

1 Introduction

There has been years spent on research trying to define what exactly convinces people into believing someone. This research talks about the influence of the involvement of different types of data like images, videos, news, articles, research papers etc. in convincing the other person to believe someone's point. I collected two datasets containing thousands of different posts from the platform R/CHANGEMYVIEW and split the dataset based on the posts that received a 'delta', which is a marker provided by the user to whoever they think have successfully convinced them.

A non-profit movement named as 'The Buddy Project' was started in April 2013 which aimed at creating awareness about the increases number of suicide cases held various campaigns to interact with people through various activities like poster making, slogan writing, script writing and short films making to prevent the increase in number of these cases. They noted that we need different mediums like images(poster making), videos(short film making) and textual information(slogan writing), as each of these have a different impact of convincing people to what they trying to convey.

As we might notice, with various types of data like images, videos, articles, we have different influential impacts in terms of convincing the other person in believing us. As Benjamin Proudfit said, "Your opinion matters especially if it's supported by facts and observations.", [15] also conclude that even though winning an argument requires several factors, the amount of evidence provided plays a crucial role in convincing the other person.[1] suggests that a photograph convinces the most of a story's truth than other media.

Based on these, I have defined my research questions in the following section which will indicate what my research will be focused on in this project. By answering these questions, I want to understand if different types of data can have different convincing power or do they work the same as an evidence.

2 Research Questions

2.1 A quick overview of the dataset

R/ChangeMyView is a platform where people post their opinions and the other members of the platform comment on the post in an effort to change the opinion of the user. Upon receiving a convincing comment the user marks the comment as a **delta** comment.

Redundant comments or posts that are against the strict rules of the platform are removed by the active moderators of the community and the platform is maintained to represent positive arguments and credibility.

Post	Comment
I'm talking about making the human race smarter, forever.Intelligence is at least partially genetic and therefore passed down by parents,yes? http://www.the-scientist.com/?articles.view/articleNo/40459/title/Inherited-Intelligence/So, what if, instead of killing off the less-intelligent people,we just limit offspring?	In addition to all of the other ways in which your view has already changed... I would like to point out that whatever we're doing <i>*already*</i> is having the effect that you would like, and without the negative downsides that any actual eugenics program implemented by actual humans would have.up on the (Flynn Effect) http://en.wikipedia.org/wiki/Flynn_effect . Basically, we've increased our IQ by more than a standard deviation in less than 100 years.

Table 1: Sample of the dataset

2.2 Research Questions

1. How many types of data like images, videos, articles are used by people as evidences?

People try to convince the author of the original post by citing evidence of different types like an image, video, news articles, research papers etc. in the form of hyperlinks like <http://example.com/evidence.mp4> as seen above in table 1. Under this research question, I will identify how many and which are the different type of data that are being used by the commentators to cite their evidence. Before, we conclude which type of evidence is the most influential or predict the use of a particular type of

evidence, we first need to know how many types are present. There are several thousand people who are active on this platform and structure their opinions differently. Someone may use an image to show a proof of an incident, someone may cite a research paper under a post that argues about a scientific topic. Therefore, I will identify and categorize all the types of data into images, videos, news, articles, blogs, research papers, etc.

2. Which one of the types of data found above, is used the most frequently in the comments which are awarded as being successful in changing the author's mind ?

There are two types of comments that are posted on the platform that is being used to conduct the research. First, are the comments that are successful in convincing the author of the post to change their mind in believing what is written in the comment. Second, are those which are not able to convince the author of the post. The successful comments are awarded a **delta** by the author marking them to be the persuasive ones.

After the extraction of the comments which are awarded a delta, I will visualize the use of the various types of data in the delta comments. This will indicate how different types of evidences like images, videos are used by the people to prove their point to the author of the original post. Also, I will be able to identify the most used evidence amongst the comments that are able to successfully convince which will give an idea of which is the most influential type of data amongst all.

3. How to predict which evidence will be used by people to spread information ?

Using the help of machine learning models, I will predict the amount of comments that will receive a delta i.e. will be convincing to the author by citing an evidence using the type of data that we receive from research question 2.

In the era where misinformation is rife, if we know which medium like images, videos are more influential, we will be able to target that particular type of data to look for misinformation. Also, in the case of holding campaigns for public welfare like the promotion of the vaccines against a pandemic, rather than spending time and money over various modes of communication like short videos, textual posts, we can predict which type of data can be the most persuasive and use that to spread awareness against the social issue.

3 Literature Review

3.1 Defining Persuasion

In social psychology, the effects of social influence on an individual’s opinion or choice have been an active area of study. ([3]). I measure influence using the r/ChangeMyView dataset focusing my attention on the role of distinct types of digital content in the act of persuading individuals. Following [10], convincing is understood then as “an attempt or intention of one participant to change the behavior, feelings, intentions, or viewpoint of another by [linguistic] communicative means.” Persuasion or the act of convincing has been defined as containing three aspects: logos (usage of arguments), pathos (emotional involvement of the audience) and ethos (credibility) as mentioned by [4]. The first one tells about the change in performance of the statements by using two or more alternative versions of the same statement to mark the possible effects on the recipients, the second one deals with the semantics of the statement, the field on which the research of [16] was based, and the last one consists of various components including the effect of evidence or proof to prove what one is trying to convey to generate a sense of credibility amongst the audience.

3.2 Work based on evidence.

The persuasiveness of the debate being held between two sides may be influenced by aspects such as the common backgrounds or ethnicity, ([11], [3], [14]) but prior work suggests that people use evidences to back up their beliefs. However, the issues which are very close to our identities such as moral or political beliefs are believed by us based on the thoughts of who we are [2] and we thus are highly resistant to being changed.[8]. Given this situation, this resistance to evidence may pose extreme challenges in the society we live in ([13]). ([9],[12]).

As suggested by [13] that correcting misconceptions about vaccines by providing corrective information by the possible side effects and other harms of the vaccine could have unexpected results and may even backfire due to the concerns of the people who are taking the vaccine. On the contrary, [6] found that generating awareness about emphasizing the risk and consequences of failing to provide vaccination to the people could positively influence the involvement of more population towards getting vaccinated.

These are two studies that indicate contradicting opinions of people in the same situation where evidence is provided. But, the methods used in the previous studies required the careful selection of the people being chosen for the experiment. These methods were carried out in the restricted environment of laboratories and hence they might not apply at a larger scale and not represent the real-world implementation.

3.3 Studies on R/ChangeMyView

People on the other hand, found a platform which appears to be driven by crowdsourcing information who were demotivated by a lack of good face-to-face conversations on some issues. The inadequacy of other online platforms and the active community of the platform R/ChangeMyView suited them better([7]). They also found that people who comment on the post find their main aim to be earning deltas. This made R/ChangeMyView, a choice for many researchers as the data is present in a more naturalistic and raw form which could represent the real world in a better way.

One of the first to gain insights from this dataset present on the platform of <https://www.reddit.com/r/changemyview/> was [18]. They examined the data posted in the first two years of the platform’s existence. According to their research, the earlier a commentator joins a topic, the more likely they are to win a delta. They also discovered a link between the number of individuals who commented on the post and the likelihood that the original author’s perspective would be modified. By using natural language processing techniques such as n-grams, they discovered that effective arguments are less similar to the original post by words and more similar by stopwords such as the, of, and so on. According to them, longer replies also had a better chance of success than shorter comments. Although, the findings from [18] are undoubtedly intriguing, these leave ample space for further research into how a sentence as a whole can infer a different meaning rather than going for the similarity between two sentences only by considering the common words like of, the, are, etc.

A different method of going with this experiment was provided by [20]. Rather than targeting separate words, they looked at a sentence as a whole. They used a machine learning based classifier based on the features provided by [17] to classify sentences as argumentative or non-argumentative which simply means if the sentence represents a debatable nature or not. They further approached their problem using term-frequency metric with the help of natural language toolkit to determine the similarity between the comments and the original opinions posted by the author. They discovered that text properties such as the word count, POS tags, punctuation are insignificant, while argumentation-related qualities produce the greatest results. This was a new discovery in addition to the work of [18] however they did not take into consideration some external factors which do not define if the sentence is argumentative or not but can still influence the opinion of the original author. With some of these factors like the number of likes/upvotes on the comments, the number of comments on the post and the presence of evidences in the form of citations, the research by [20] leave some space for further research.

This platform was further used by [15] based on the shortcomings of the work done above. They were interested in the evidence and statistical language used

in these discussions and understand how were people convinced based on the presence of these evidences and statistics. The authors define evidence as a citation or a reference to a particular document/image/video using a URL provided by the individual who is commenting as a proof to their point or a piece of information to support against the motion provided by the original author. These researchers looked at evidence as a whole and restricted their research on socio-moral topics or topics on social issues that affect the people as one. Due to these restrictions, their research was not extended to more topics as R/ChangeMyView is a house to topics of several fields, and they overlooked the fact that different evidences can have different impacts on persuasiveness.

My research is a further extension on the research of [15] where I emphasize the difference in the influence of using various types of data like images, videos, articles etc. being used as evidences to convince the authors at the platform known as R/ChangeMyView.

4 Research Methodologies

My proposed study has 6 sequential steps: 1) Data Retrieval 2) Categorization into Delta and Non-delta comments. 3) Detection of comments that cite evidence. 4) Identification of different type of evidences. 5) Inferring insights from the number of evidences used graphically. 6) Predicting the number of different evidences.

4.1 Data Retrieval

The dataset is composed of the format '.jsonlist' in two files. These files are downloaded from the internet given under the link mentioned in the evaluation section below and is retrieved through a python script. A jsonlist file is a type of file that contains a list of JSON type objects. Each object is iterated and converted into compatible formats using the python's JSON API. These JSON objects are appended into an array through numpy and with the help of the pandas framework, the array is finally converted into a dataframe. This allows us to visualize the dataset in a more understandable using the various functions present in the pandas library with ease than the complex format of JSON.

4.2 Categorization into Delta and Non-delta comments

The raw data that was captured into the dataframe will be traversed and categorized into posts and comments. These will represent the post that was originally written by the author and the comments that were written under the original post. Afterwards, the comments will be categorized into positive and negative i.e. the comments which have been awarded a delta by the original author and the comments which have not been awarded with any delta respectively into separate columns of the dataframe. These two columns will further be used to carry out the research.

4.3 Detection of comments that cite evidence

In addition to collecting both types of comments: the delta awarded and the non-delta awarded comments, I was required to make some inferences as to what defines an evidence in this platform. With the help of the visualization in the first step, I calculated the evidences by the number of hyperlinks to external websites/documents that are cited by people in their comments. To collect these hyperlink references, the text inside the comments of the people will be searched for typical website and extension identifiers like 'http' using natural language processing tools like regular expressions.

4.4 Identification of different type of evidences

After the extraction of the comments that cite evidence, regular expressions will be used to figure out which type of evidences are used. Each of the hyperlinks received above will be matched using regular expressions against an array of extensions like 'pdf', 'jpeg', 'mp4', '/news', '/blog', '3gp', 'webm', 'mov', 'gif' and so on. Further, a python dictionary with key as the extension and the value as the number of comments falling under that extension will be stored.

4.5 Inferring insights from the number of evidences used graphically

In this step, I will create pie graphs and bar graphs with the various types of extensions that we get from above and the amounts of comments that use that type. To capture a better image of how many types of data are used as evidences, and to capture interesting facts like the most-used type of evidence, the least-used type of evidence and to infer more interesting facts that may be present in the data, I will use the plotting libraries of python like plotly, seaborn and matplotlib to plot these graphs.

4.6 Predicting the number of different evidences

Finally, we will use the multivariate negative binomial regression model to create a machine learning model to predict the number of comments that will use the different types of evidences. This model will be used to predict the number of images/videos/articles etc. in the delta awarded comments or the comments that were able to convince someone, and then we will predict the amount of these evidences in the non-delta comments or the comments which were not able to convince the original author. This will the predict the type of data that can be used to communicate our information which will have the most convincing power.

5 Evaluation

5.1 Dataset

On the internet forum site Reddit, 'ChangeMyView' is a channel, or 'subreddit' which has become a key hub for forming friendships and communities. The ChangeMyView site on Reddit is a platform where you may "share a view-point you recognise may be erroneous, in an effort to understand various other perspectives on the topic" <https://www.reddit.com/r/changemyview/>. An author who wants to participate in a conversation would usually make a new post outlining their point of view and encouraging other users to change what they believe by mentioning their views using the comment section below. The dataset that I used was collected by [19] from this platform. A sample of the dataset is given below.

Users add compelling comments to the initial post, sometimes just deliberative, sometimes with connections to other sources and figures, in an attempt to persuade the author ([15]). If the author considers a comment persuasive, they will reward it with a delta.

Human moderators also ensure that authors expound on their points of view and that the conversation stays respectful and clear. The policy of the platform is very strict and does not allow random conversations and citations that are not credible. If someone uploads any evidence or citation, they can be downgraded by the people and the moderators will remove the comment. The strict rules and followed by both the authors and the commenters, and the active human moderators are responsible for the credibility of the data that is captured from this dataset. The policy can be found [here](#).

5.2 Metrics

Two measures of the machine learning model that we will be using to predict the amount of comments with specific evidences will help us to define if our model is optimal or not and then define the correctness of the model.

1. **Pearson-Chi Squared:** This is a metric that we receive from the report generated by the negative binomial regression model. It is compared with the another metric which is the chi-squared value that is calculated using a standard chi-squared table and the value of df_residuals that we obtain from the same report.

By comparing the Pearson-Chi Squared with the calculated Chi-Squared value, we can conclude if our model is optimal or not. The nearer the value of the Pearson-Chi Squared metric is to the calculated value, the more efficient our model can be considered, but if the absolute difference between the two is large, then our model can be optimized further.

2. **Standard-Error:** This metric is used to identify the how well our model

Posted by author 23 hours ago	1.Number of up-votes(likes) by other users.
CMV: The average homeowner does not benefit from constantly rising house prices	2.Concise statement of the original poster’s opinion
I often hear that consistently inflation beating rises in house prices are A Good Thing. People who own houses seem very happy that their house has increased in monetary value, despite the fact that the utility they get from it has not increased at all. [...]	3. Detailed statement of the original poster’s opinion
B 21 hours ago	4.Persuasive comment by user B
All property does not increase at the same rate. [...]	
OP 23 hours ago This approach only works for people who are free to move easily and whose income is not linked to where they live, or for investors who do not live in the property. [...]	5. OP states they are not convinced by B’s comment
C 20 hours ago You are not taking into account leverage (i.e. taking out a loan to buy a house) which almost 100% of home buyers do. [...]	6.Persuasive comment by user C.
OP 19 hours ago That is a good illustration of why falling into negative equity is a problem. Maybe never having been in that situation I underestimate the severity of the issue. [...] !delta for highlighting the importance of avoiding negative equity.	7. OP states their opinion has been changed by user C. 8. OP state they would like to award a delta to C’s comment

Table 2: Example dataset taken from [5]

is performing. This simply returns the error that is generated after the prediction of the amount of comments with different types of data. Standard Deviation defines a similar function as standard error which is frequently confused but standard deviation reflects variability within a sample, while the standard error estimates the variability across samples

of a population. This means that using standard error, we will be able to compare the predicted values with the original test values taken from the sample of our dataset.

References

- [1] The photograph as evidence. <https://www.photopedagogy.com/evidence.html>.
- [2] Dana R. Carney, John T. Jost, Samuel D. Gosling, and Jeff Potter. The secret lives of liberals and conservatives: Personality profiles, interaction styles, and the things they leave behind. *Political Psychology*, 29(6):807–840, 2008.
- [3] Robert Cialdini, Wilhelmina Wosinska, Daniel Barrett, Jonathan Butner, and Malgorzata Górnik-Durose. Compliance with a request in two cultures: The differential influence of social proof and commitment/consistency on collectivists and individualists. *Personality and Social Psychology Bulletin*, 25:1242–1253, 10 1999.
- [4] Robert Cockcroft, Susan Cockcroft, Craig Hamilton, and Laura Hidalgo-Downing. *Persuading People: An Introduction to Rhetoric*, 3rd edition. 01 2014.
- [5] Daria Dayter and Thomas Messerli. Persuasive language and features of formality on the r/changemyview subreddit. 5(1):165–195, January 2022.
- [6] Zachary Horne, Derek Powell, John E. Hummel, and Keith J. Holyoak. Countering antivaccination attitudes. *Proceedings of the National Academy of Sciences*, 112(33):10321–10324, 2015.
- [7] Shagun Jhaver, Pranil Vora, and Amy Bruckman. Designing for civil conversations: Lessons learned from changemyview. Technical report, Georgia Institute of Technology, 2017.
- [8] Dan Kahan, Ellen Peters, Maggie Wittlin, Paul Slovic, Lisa Ouellette, Donald Braman, and Gregory Mandel. The polarizing impact of science literacy and numeracy on perceived climate change risks. *Nature Climate Change*, 2, 12 2012.
- [9] Joshua Klayman. Varieties of confirmation bias. volume 32 of *Psychology of Learning and Motivation*, pages 385–418. Academic Press, 1995.
- [10] R. Lakoff. Persuasive discourse and ordinary conversation, with examples from advertising. In D. Tannen, editor, *Analyzing Discourse: Text and Talk*, pages 25–42. Georgetown University Press, Washington DC, 1982.
- [11] Tanushree Mitra and Eric Gilbert. The language that gets people to give: Phrases that predict success on kickstarter. pages 49–61, 02 2014.

- [12] Raymond Nickerson. Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, 2:175–220, 06 1998.
- [13] Brendan Nyhan and Jason Reifler. Does correcting myths about the flu vaccine work? an experimental evaluation of the effects of corrective information. *Vaccine*, 33, 01 2015.
- [14] Richard Petty, Duane Wegener, and Leandre Fabrigar. Attitudes and attitude change. *Annual review of psychology*, 48:609–47, 02 1997.
- [15] Hunter Priniski and Zach Horne. Attitude change on reddit’s change my view. 01 2018.
- [16] Marie-Thérèse Rohr. *Persuasion in smoking cessation online: an interpersonal pragmatics perspective*. PhD thesis, 01 2018.
- [17] Christian Stab and Iryna Gurevych. Identifying argumentative discourse structures in persuasive essays. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 46–56, Doha, Qatar, October 2014. Association for Computational Linguistics.
- [18] Chenhao Tan, Vlad Niculae, Cristian Danescu-Niculescu-Mizil, and Lillian Lee. Winning arguments: Interaction dynamics and persuasion strategies in good-faith online discussions. *WWW ’16: Proceedings of the 25th International Conference on World Wide Web*, 02 2016.
- [19] Chenhao Tan, Vlad Niculae, Cristian Danescu-Niculescu-Mizil, and Lillian Lee. Winning arguments: Interaction dynamics and persuasion strategies in good-faith online discussions. In *Proceedings of WWW*, 2016.
- [20] Zhongyu Wei, Yang Liu, and Yi Li. Is this post persuasive? ranking argumentative comments in online forum. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 195–200, Berlin, Germany, August 2016. Association for Computational Linguistics.