Information Cascades in Discussions

Teddy Heidmann
Cornell University
ath55@cornell.edu

Andrew Wang
Cornell University
azw7@cornell.edu

Abstract

Herding behavior in the form of information cascades has been known to occur in lab experiments and on social media. In this paper we attempt to shine new light on cascades outside of their typical social context, investigating their occurrence in objective discussions. Studying Wikipedia "Article for Deletion" discussions, we analyze the series of votes cast by users, along with their rationale for their votes, finding evidence that information cascades occur. We investigate a number of language-centric features to understand the qualities of positive and negative cascades, and to see which posts have more influence over future posts. We use our findings to build a classifier that predicts the extent to which a cascade will grow in a discussion. We find not only that information cascades exist in these discussions, but that cascades in discussions are reasonably predictable, which is somewhat surprising given their subtlety.

1 Introduction

Collective decision-making in a community can be challenging to orchestrate correctly and fairly, especially if done through discussion. Editors of Wikipedia, the popular online knowledge base, participate in a particularly interesting group deliberation process in the form of Article for Deletion (AfD) discussions. If an editor believes an article is out of place or should not exist, he can nominate it for deletion by creating an AfD discussion for the article and giving his initial reasoning for this nomination. For example, articles can be nominated due to copyright infringement or lack of credible sources, or because they are hoaxes or duplicates of existing articles (Wikipedia, 2017). A particularly divisive issue is notability: due to differing perspectives on the mission of an online encyclopedia, some members are sympathetic towards articles about less notable topics, while others favor deleting them.

Once an AfD discussion is initiated, users can visit the page, contribute to the discussion and "vote" by recommending the article be kept, deleted, or redirected to or merged with another. Once the community has come to a consensus (as judged by an impartial observer), a moderator performs the agreed-upon action, and the discussion is closed. In this paper, we investigate information cascades within these discussions, specifically the effect of previous users' votes and comments on subsequent users' decisions to keep or delete an article.

1.1 Information Cascades

In an AfD discussion, users can view the entire thread of events chronologically, including every user's comment and decision. If many users in a row vote the same way, this could potentially create a cascade effect, yielding more of the same vote. Information cascades begin when a person's decision is influenced by the decisions of others, leading him to make a decision that aligns with the choice of the group (Easley and Kleinberg, 2010). An information cascade is the chain reaction that occurs when future decision-makers are put in the same situation, continuing the bandwagoning trend.

Cascades can be quite powerful, and exist in everyday life as well as on almost every form of social media. One example occurs when deciding where to eat in a new town — should you should go to the restaurant you heard about from a friend, or the one next door with many more diners? Another example is on the once-popular social media application, Yik Yak, where a recent study showed that "upvoting" or "downvoting" a post a number of times in the first few minutes influenced more people to vote the same way (Liedell et al., 2015).

1.2 Motivation

Why study cascades in AfD discussions?

Easley and Kleinberg (2010) show that cascades can occur even when the agents involved are "perfectly rational": an agent may judge that a body of outside evidence (previous people's decisions) carries more evidential weight than her private information, then make the same decision as the herd, which places subsequent decision-makers in the same situation and forces them to follow suit. These "epistemically-motivated" cascades have indeed been observed in people in lab experiments (Anderson and Holt, 1997) and can have deep consequences for discussions in which members aim to make a collective correct decision — if these cascades do occur in discussions, it means that despite people's best interests, people with potentially valuable information may get discouraged from sharing it, and that the final decision reached is "unstable" in that it could depend merely on the order in which members share their views. If we want to move towards better discussions, it's an important goal to be able to identify early signs of herding behavior in discussions, and to be able to predict herding behavior before it grows out of hand. Thus, we aim to explore both analysis and prediction of cascades in discussions.

2 Related Work

To our knowledge, this is the first in-depth exploration of cascades in textual discussions, and the first to attempt to understand how "epistemically-motivated" cascades arise from participants' use of language. In studying discussions, we draw from a large body of work that implicitly emphasizes cascades as a social (as opposed to epistemically-motivated) phenomenon, in the sense that the cascades studied are intuitively socially motivated, and depend crucially on the way members of a community are connected together in a social network. Cheng et al. (2014) finds that photo-sharing cascades in a large social network (Facebook) are predictable, but much of their model's predictive power comes from rich structural features that aren't applicable to the barebones network structure of a discussion. Indeed, content features performed by far the worst in their task, and textual features based on photo captions did no better than random guessing. Tsur and Rappoport (2012) and Ma et al. (2013) look at a more textual task, tracking the spread of hashtags on Twitter, and find similarly that hashtag cascades are predictable given content and network features, but that network features are far more important. They also explore mostly shallow (orthographic, lexicon) features that don't dig into the information content of posts.

In contrast to this line of work, we aim to study cascades in a setting in which social motivation is minimized, and in which we can abstract from the effects of network topology on cascade growth, focusing as much as possible on the epistemic phenomenon. Thus, we study AfD discussions, which have these desired properties: users try to make the correct decision, discussions follow a long tradition of commitment to objectivity and the structure of the network is constant across discussions (it can be thought of as a sequence of nodes corresponding to votes, each node influenced by all preceding nodes.) Helpfully, we also have explicit labels of users' positions in the form of votes.

The Wikimedia Foundation provides some initial analysis of AfD discussions: Taraborelli and Ciampaglia (2010) find evidence of herding behavior, in that the distribution of votes in any prefix of a discussion influences the distribution of votes in the rest of the discussion. We extend their work on herding to show that cascades of the type we're interested in also occur — specifically, we give evidence that the mere order in which votes are cast can influence subsequent votes in a discussion. We also further the analysis by considering language, not just the votes cast, which is crucial to understanding epistemically-motivated cascades.

3 Dataset

We scraped Wikipedia AfD discussions from 2005 to 2016, resulting in 360,211 total discussions. We extracted usernames of participants, times posted, discussion text (taking note of strikethrough text, which indicates changed opinion) and final decision made. We generated part of speech tags and sentiment scores for each post using the NLTK libraries `tag` (Loper and Bird, 2017) and `vaderSentiment` (Hutto et al., 2017). We find that discussions have a mean length of 6.1 posts and 4.8 votes, and the longest discussion consists of 335 posts and 316 votes. 32% of votes are for "keep," 60% are for "delete" and the rest are for "redirect" or "merge."

4 Method

These hypotheses map our course:

1. Cascades occur in AfD discussions.
2. AfD cascades have distinguishing features.
3. AfD cascades are predictable.

We also look into what best explains AfD cascades: Do cascades occur because the first post in a cascade introduces game-changing, convincing new information, or do they occur *in spite of* a lack of compelling rhetoric? To this end, we also look into the differences between "keep" and "delete" cascades, and their potentially different causes.

Define a *chain* of type $t \in \{\text{keep}, \text{delete}\}$ and length k as a sequence of k votes for t (with no dissenting votes in between). We take cue from Cheng et al. (2014) and frame our exploration in terms of this question: Given that we currently observe a chain of type t and length k, will it grow past length $f_t(k)$, where $f_t(k)$ is the median length over all chains of type t and length $\geq k$? This formulation is practical because it implies a binary decision problem with approximately equal class sizes, and a classifier that predicts the result can be used in practice to track the growth of a cascade with finer detail over time, as more votes are observed. Furthermore, it allows us to conveniently compare statistics between the two classes, moving beyond pure vote counts (since the votes observed so far in the current chain are the same between the two classes) and helping us answer the question: Given a set of ongoing discussions that each end with the same sequence of votes, what aspects of language or the users involved account for the fact that some of these cascades will grow bigger than others?

5 Analysis

5.1 Cascades Occur

Figure 1 shows that the larger a chain is, the longer it is expected to keep growing, supporting the claim that cascades occur. It's worth noting that the graphs tend to dip towards the end, but that this isn't too exciting, and is merely a consequence of the fact that discussions are finite — so even the longest chain must eventually end, bringing the mean down with it.

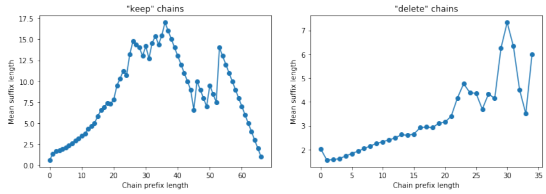

Figure 1: Mean chain suffix length vs prefix length — given a chain currently of length k, how much longer can we expect it to grow? The slope of the regression line is positive for both keep and delete chains ($p < 0.01$).

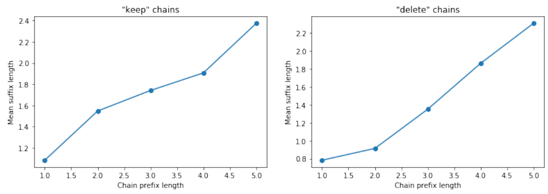

Figure 2: Mean chain suffix length vs prefix length, but only among discussions that have currently received 5 keep votes, 5 delete votes and 0 others (in any order) at the time the current chain is observed. The slope of the regression line is positive for both keep and delete chains ($p < 0.01$).

There is still the concern that the graphs in Figure 1 can be explained by an alternate hypothesis: Perhaps some discussions are intrinsically more keep-worthy than others; if we observe more keep votes in the beginning of a discussion, we're bound to see more in the rest of the discussion simply because the article is particularly worth keeping, and not due to herding effects. We try to control for this effect in two ways.

The first is to take a "God's-eye view" and only consider discussions that have received some fixed proportion of keep to delete votes, say $50 \pm 1\%$ (with zero redirect or merge votes) — this gives us some grasp of the "objective keep-worthiness" of the article under discussion. Among these discussions, we find that we get graphs similar to those in Figure 1, supporting the claim that the order of the votes itself is responsible for cascades (rather than the mere fact that seeing more "keep" votes is evidence that the discussion has more "keep" votes overall in the first place.)

But the closing of a discussion (and thus, indirectly, the final distribution of votes in the discussion) is determined by a very human process (an editor's judgment of consensus) which could introduce any number of confounding factors, so our second approach is to consider an ongoing discussion from a participant's perspective: given that x keep votes and y delete votes have been cast so far in any order (for fixed x and y), how much will the next voter be swayed by the most recent chain? We take this approach in Figure 2, and the resulting trend suggests that even in these cases, the length of the most recent chain can influence users' votes. This suggests that even among discussions where users' overall attitude toward the article so far is approximately fixed, the most recent attitude has a stronger influence.

To be sure, this could be due to new developments in the situation, such as editors improving the article in question. However, we note that there is less reason to suspect this in "delete" chains, where the trend also occurs. In the case of "delete" chains, due to the format of the discussion, most of the objective information in favor of deletion (e.g. specific rules or guidelines) has been presented at the beginning of the discussion by the nominator, and new posts merely discuss differences in priorities and opinions. Furthermore, new developments rarely make an article worse (since damaging edits to the article are quickly reverted by moderators), so it's unlikely that the trend can be explained by the objective quality of the article changing over time. Still, we'll continue trying to tease apart these alternate hypotheses: Do cascades occur because users bring up genuinely new information, or because of some more subtle phenomenon, perhaps some feature of language that affects how knowledgeable users seem to others?

Having established that some form of herding behavior occurs, we'll use the term "cascade" in place of "chain," to match terminology with (Cheng et al., 2014). We try to explore the above questions in the next section by comparing cascades.

5.2 Cascades Have Distinguishing Features

After confirming that cascades occur in the AfD discussions, we turn to our next question: How does content affect cascade growth? We now explore the effects of word count, sentiment, and others on cascades. Note that by definition, a cascade starts either at the beginning of a discussion or immediately after a cascade of the opposite decision,

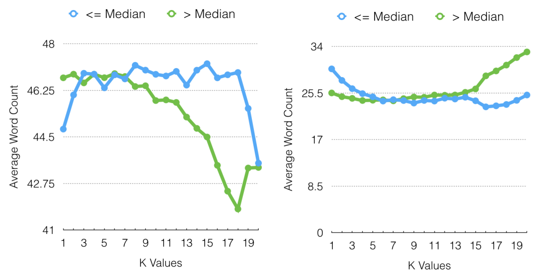

Figure 3: Average word lengths of first posts of "keep" (left) and "delete" (right) cascades, versus k. Lengths are significantly different ($p < 0.05$) for "keep" cascades with $k \in [1, 2]$ and for "delete" cascades with $k \in [1-5, 9, 16-20]$.

which will be important in our analysis of the first posts of cascades.

5.2.1 Word Count

How does the length of the first post of a cascade influence the cascade's growth? One hypothesis is that if the first post of a cascade contains more information, it could be more persuasive, leading to a longer cascade. If this is true, it confirms the "new information" hypothesis that people at the beginning of a cascade bring up genuinely new information, changing more users' minds.

To control for any effect that the position in the discussion might have on post length, we look only at cascades that begin at the beginning of the discussion. For various values of k, we take all cascades that reached a length of at least k, then split them into two groups based on whether they reached the median cascade length for that k value or not. We then graph the average number of words in each post versus k. The graph for average word count for both "keep" and "delete" cascades can be seen in Figure 3.

In the "delete" graph, we see that cascades that do not exceed the median length tend to start with posts with more words than cascades that exceeded the median length — but as cascade length becomes very large, this trend flips. In the "keep" graph, we see the opposite trend. One possible interpretation is that in taking the "default" position (supporting "delete"), a brief post may seem more convincing and self-assured in the short run, but that in the long run, when these short posts get drowned out by many other voices, it takes a more substantive argument to influence people far down the cascade. On the other hand, fighting for "keep" is an uphill battle, so more rhetoric must go into

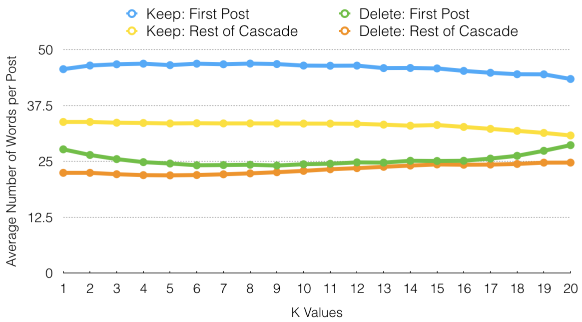

Figure 4: Average word length for posts in "keep" and "delete" cascades. Split between first post of a cascade and the remainder of the posts, for various values of the final cascade length K.

initially convincing others to vote "keep"; whereas the influence of the first post also diminishes as more voices join in the cascade. Interestingly, when we compare the number and frequency of content words used between the two classes (as a measure of the number of ideas introduced), following Niculae and Danescu-Niculescu-Mizil (2016), we find that the first k posts of a cascade that exceeds the median consistently tend to introduce more ideas than the first k posts of a cascade that doesn't reach the median, across all values of k and for both "keep" and "delete" cascades. This is very intuitive, suggesting that people are more convinced by arguments with more substance, more ideas.

When we look at the average word length for both types of cascades, split between the first post of the cascade and the remainder of posts in that cascade, we see some interesting results (Figure 4). First, we note that the "keep" cascades have a much higher word count than the "delete" cascades. A potential explanation is that since an article is nominated for deletion, then someone has already moved to delete the post; thus, users voting to prevent the post from being deleted must work harder to support their ideas, writing more words on average. The first post of a "keep" cascade is generally the longest (first post is significantly longer than the average of the rest, t-test, $p < 0.05$ for $K \in [11, 20]$, $p < 0.001$ for $K \in [1, 10]$), which we speculate is because the poster is the first person to oppose the initial proposal to delete, and must build a case for a new position from scratch. This difference persists for larger values of K, suggesting the first post in a "keep" cascade has longstanding influence.

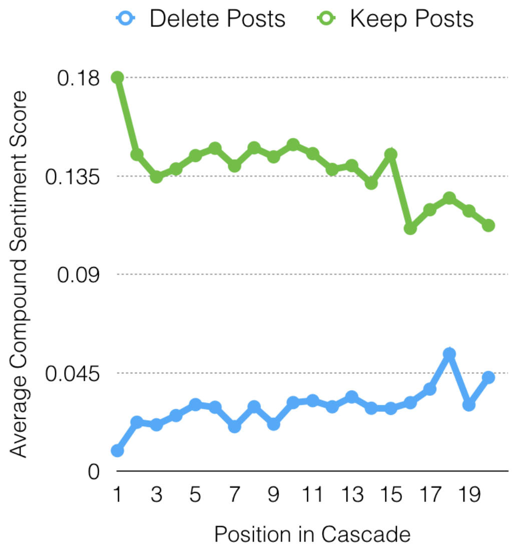

Figure 5: Average sentiment score for "keep" and "delete" cascades at $k = 10$ based on position in the cascade.

In contrast, the first post of a "delete" cascade tends to not be much longer than posts in the rest of the cascade. This again suggests that "keep" cascades have to work harder to persuade users, whereas a "delete" cascade has an easier time coasting on momentum.

5.2.2 Sentiment

We also investigate how differences in sentiment differentiate cascades. We compared differences in sentiment between "keep" and "delete" cascades, as well as the difference between the first post in a cascade and the remainder of the posts in the cascade. The average sentiment score based on position in the cascade can be seen in Figure 5.

As shown in the graph, the first post of a "keep" cascade has the highest average sentiment in the cascade, while the first post of a "delete" cascade has the lowest — thus, the first post in a cascade tends to be the most polarized. The remainder of the cascade tends to be more neutral. Expanding this to various values of k (Figure 6), we can see the average sentiment remains significantly different between all lines ($p < 0.05$ for all points with $k \in [1 - 20]$). This trend continues as k increases.

We also looked at the other sentiment scores output by vaderSentiment: positive, negative and neutral. The compound score that we looked at above is the weighted sum of each of these valence scores, normalized from -1 to 1. Since the compound score is simply a weighted sum of the

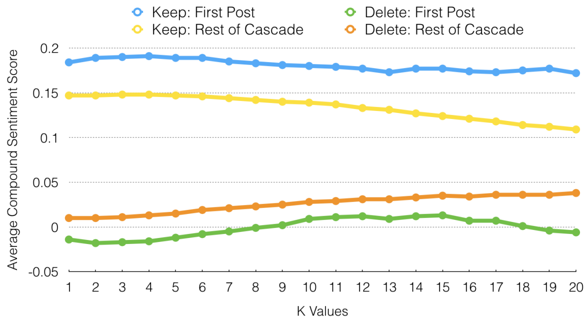

Figure 6: Average sentiment score for "keep" and "delete" cascades. Split between first post of a cascade and the remainder of the posts, for various final cascade lengths K.

other measures, we got similar results when analyzing the individual measures. Neutral sentiment scores were similar between "keep" and "delete" posts. We find that the graphs for positive and negative sentiment are analogous to the graphs for compound sentiment.

We also took the difference between the sentiment of the first post of a cascade and the average sentiment of the remainder of that cascade, and compared this statistic between "keep" and "delete" cascades. We find that the first post of a "keep" cascade is significantly more polarized, relative to the rest of the cascade, than that of a "delete" cascade (t-test, $p < 0.05$ for $K \in [11, 20]$, $p < 0.001$ for $K \in [1, 10]$). This is in line with the theme that "keep" cascades must work harder to take root, while it is more likely that a "typical" "delete" post initiate a cascade.

Lastly, we again split the data into "keep" and "delete" cascades that either reach or don't reach the median length. We found that there was no significant difference between the two classes for either of the two types of cascades, so we did not continue in the analysis of splitting by median length.

5.2.3 Other features

We tried a collection of other features and compared their values between cascades of at least length 5 that exceed the median length and those that don't. Call these first 5 observed posts the "prefix." We compared:

- Average pairwise cosine/Jaccard similarity between posts in the prefix (measures cohesiveness of the paricipants' ideas)
- Cosine/Jaccard similarity between first post in prefix, and the rest of the posts concatenated together
- Percentage of voters in the prefix who vote only once in the entire dataset (first-time voters)
- How far into the discussion the cascade begins (absolute, and normalized by length of discussion)
- What percent of votes earlier in the discussion were for "keep" versus "delete"
- How many votes in the prefix are marked "speedy"
- How often the word "agree" is used in the prefix
- Average historical keep rate of users in the prefix (i.e. how much were users swayed from their typical voting patterns)
- Part-of-speech frequencies

The results are shown in Figures 7 and 8. Some takeaways:

We notice that features fall into two groups. For some features X, more of X in the prefix entails a longer "keep" cascade, while less of X entails a longer "delete" cascade. Specifically, we see this in the word count (as seen earlier), speedy votes, historical keep rate (since we consider 1 minus this number in delete cascades) and adverb usage statistics. While these features are useful markers of each particular cascade type, and will thus be useful in prediction, we should be wary of concluding that they influence cascades in general — an alternate explanation, for example, is that these features are indicative of users tending toward some position in particular, rather than toward the position of the current voter. However, these features do give further evidence of difference between "keep" and "delete" cascades: cascade-generating "delete" posts are brief, aren't as polarized sentiment-wise as in "keep" cascades, trigger less strong sentiment in terms of speedy votes and don't seem characterized by their ability to sway users who typically vote the other way. This gives further evidence that the two cascade types occur in fundamentally different ways: since the discussion is premised upon deleting the article, "keep" supporters are the "underdogs" and must rally support to build a case against the default position, while "delete" cascades are potentially more insidious — they are a result of users "falling into" the default conclusion without being checked by a dissenting voice, potentially more in line with the "knowledge-perception" explanation of cascades (in which dissent is underexpressed.)

The other group of features can be thought of as more generally indicative of cascades (i.e. more of X means a bigger cascade, no matter what type.) This includes the content words introduced (as seen earlier), agreement, post similarity, start position and initial vote distribution features. These features suggest that there are still positive aspects to cascades: more ideas are introduced in longer cascades, meaning participants have considered more reasons for supporting a position; posts are more similar in the prefixes of longer cascades, suggesting members are more united in support of common ideas; and in longer cascades, there are actually *less* votes in agreement with the cascade earlier in the discussion, suggesting that the participants of the cascade have reason to believe in their position beyond blind support of earlier members' positions.

5.3 Cascades are Predictable

We used LIBLINEAR (Fan et al., 2008) to train a Linear SVM on a performant subset of the above features as well as term frequencies of the 186 non-stopword unigram and bigrams with a document frequency of at least 0.01. We again tackle the binary decision task: Given a cascade of type t and length k, will it grow past length $f_t(k)$? We split our dataset into 90% training and 10% testing, and developed two classifiers for each value of k, one for "keep" and one for "delete" cascades. We found that the features that generalize best are the vocabulary features, previous votes in the discussion, historical voting record of users and first-time voter features. Figure 9 plots performance under the AUC metric for various values of k. We find that "delete" cascades are much more predictable but become harder to predict as k increases, while "keep" cascades remain relatively consistent in terms of predictability. We compared our classifiers to random guessing by comparing our model against a weighted random classifier under 10 permutations of training-test splits, finding that both "keep" and "delete" classifiers outperform this baseline for $k = 5$ (Mann-Whitney U

"Keep" Cascades (k=5)		
Feature	> median	≤ median
average cosine similarity	0.904 (7856)	0.9 (10928)
first post vs rest (cosine)	0.447 (7652)	0.437 (10673)
first post vs rest (Jaccard)	0.766 (7652)	0.759 (10673)
agreement	0.455 (7955)	0.287 (11033)
single-time voters	0.088 (7955)	0.076 (11033)
start position	7.541 (7955)	5.759 (11033)
start position (norm.)	0.243 (7955)	0.283 (11033)
initial keep votes (norm.)	1.199 (7955)	1.229 (11033)
speedy votes	0.239 (7955)	0.203 (11033)
keep rate	0.503 (7955)	0.517 (11033)
adverb usage	0.046 (7955)	0.044 (11033)

Figure 7: Other features compared for "keep" cascades. All features significant ($p < 0.01$) except for start position (norm.), for which $p = 0.05$ (Welch's t-test). Bin sizes listed in parentheses.

"Delete" Cascades (k=5)		
Feature	> median	≤ median
first post vs rest (cosine)	0.372 (15921)	0.369 (23962)
first post vs rest (Jaccard)	0.804 (15921)	0.795 (23962)
agreement	0.201 (16132)	0.148 (24230)
start position	2.228 (16132)	2.067 (24230)
start position (norm.)	0.101 (16132)	0.127 (24230)
initial delete votes (norm.)	0.76 (16132)	0.792 (24230)
initial keep votes (norm.)	0.067 (16132)	0.07 (24230)
speedy votes	0.227 (16132)	0.262 (24230)
keep rate	0.225 (16132)	0.23 (24230)
proper noun usage	0.179 (16132)	0.171 (24230)
adverb usage	0.045 (16132)	0.047 (24230)
period usage	0.108 (16132)	0.103 (24230)

Figure 8: Same but for "delete" cascades. Listed features are significant ($p < 0.01$) except for initial keep votes (norm.), for which $p < 0.05$ (Welch's t-test). Bin sizes listed in parentheses.

test, $p < 0.01$).

5.3.1 Classifier Analysis

A Linear SVM is conveniently interpretable: we can look at the magnitude of the coefficients it learns to get a sense of which features are more important (Guyon et al., 2002). Figure 10 charts the relative importance of term features for each post in the beginning of the cascade ($k = 5$)[1]. Note that we use term frequencies instead of counts since the total length of a post (as seen earlier) is a confound. We again see a dramatic difference between the "keep" and "delete" classifiers, suggesting the two kinds of cascades have different causes.

To the "keep" classifier, the first post of the cascade is most important (average coefficient magnitude is greater than the average over the next 4 posts, Mann-Whitney U test, $p < 0.05$), supporting the claim that here, a stronger initial idea indicates the start of a larger cascade.

To the "delete" classifier, the first post of the cascade is not clearly more important — indeed, we find an almost opposite trend. For each feature, consider the difference between its magnitude in the first post and its average magnitude in the next 4 posts. If we compare this statistic for all features between the "keep" and "delete" classifiers, we see that features in the first post of a "keep" cascade have on average 0.02 greater magnitude than corresponding features in the next 4 posts, while features in the first post of a "delete" cascade have on average 0.02 *less* magnitude than corresponding features in the next 4 posts. In other words, the first post of a "keep" cascade is generally more important than any of the next 4 posts, while the first post of a "delete" cascade is less important than its succeeding posts. Furthermore, this delta is significantly different between the "keep" and "delete" classifiers (Mann-Whitney U test, $p < 0.05$).

This lack of strong first posts in "delete" cascades adds evidence to the "epistemic-motivation" hypothesis: if cascade growth depends on the extent to which a user brings up important information contrary to what was previously said (since by definition a cascade begins after a decision of the opposite type), we should see the first post in the cascade being most predictive of the eventual support it garners. However, we find that the opposite seems to be the case. This raises the concern that even though subsequent users contribute more supporting points and agree more in the beginning of the cascade (as found earlier), they are supporting a position "merely because" of the tendency to agree, rather than because the cascade-initiating post used language that distinguished it from a less-influential post.

[1]We get similar results on a classifier trained *only* on unigram-bigram features, with no meta features.

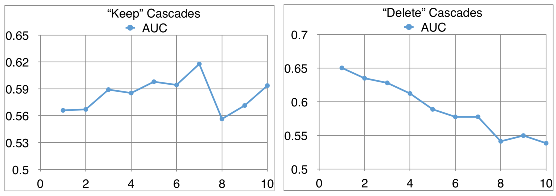

Figure 9: Performance of the "keep" and "delete" classifiers respectively, for various values of k. Random guessing would score 0.5.

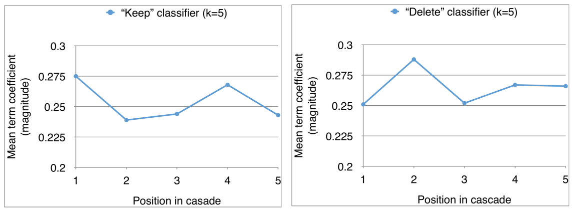

Figure 10: Mean magnitude of term features for each of the first five posts of the cascade ($k = 5$), for "keep" and "delete" respectively, as a proxy for the importance of each post.

Figures 11 and 12 list the terms with the highest-magnitude coefficients. It's worth noting that terms like "speedy delete" in "keep" posts occur because people refer back to previous suggestions, although we did find a few mislabelings (often due to users changing their votes in a subthread of their original posts.) The fact that successful "keep" movements refer back to the other side's points is in line with the findings of Zhang et al. (2016) with respect to Oxford-style debates. WP refers to specific editing guidelines; for example, WP:BIO sets notability criteria for articles about people, while WP:GNG refers to the "general notability guideline." It's interesting that "WP:BIO" and "WP:GNG" are both positive while "bio" and "gng" are negative, suggesting that more careful or official-sounding language is more credible.

6 Conclusions

We have presented a number of methods for analyzing the content of Wikipedia AfD discussions to understand information cascades. We've

Most positive features	
First post	speedy delete, hoax, useful, simply, meet
Last post	fails wp, vanity, wp bio, speedy delete, far
Most negative features	
First post	bio, possible, looks, content, redirect
Last post	fails, original, exist, comment, meets

Figure 11: Features with the most positive and most negative coefficients for the "keep" classifier ($k = 5$).

Most positive features	
First post	original research, going, verifiable, wp gng, added
Last post	fails wp, vanity, wp bio, speedy delete, far
Most negative features	
First post	gng, content, current, multiple, establish
Last post	original research, section, pages, note, little nom

Figure 12: Features with the most positive and most negative coefficients for the "delete" classifier ($k = 5$).

shown that cascades occur, that cascades have distinguishing features and that cascades are predictable. We've also given evidence to support the claim that AfD cascades can grow in spite of a lack of strong impetus at the beginning of the cascade, particularly in "delete" cascades. Given the possibility of discussions being quietly susceptible to herding effects, it becomes valuable to be able to track the eventual growth of a cascade from its inception. Our classifiers make first steps in this direction. We hope these findings can be applied to improving the format of future discussions, furthering the objective and balanced exchange of ideas.

7 Future Work

It would be interesting to make more use of temporal features. It would also be interesting to cluster cascades to see if we can further unravel their distinct causes. We also did not consider "merge" or "redirect" votes and the way in which these could interact with cascades. Finally, there are interesting ways in which the AfD discussions deviate from the clean sequential model explored: users sometimes reply directly to each other's posts, occasionally leading people to change their votes. It would be valuable to look more into the asynchronous aspects of the discussion — effects of side conversations and decision changes on cascades. Exploring these directions could reveal interesting new insights about the nature of cascades in discussions.

8 Division of Work

Teddy focused on feature analysis, particularly on word count and sentiment. Andrew looked at other features and focused on the classifier. We worked together on figuring out which directions to take the project.

References

Lisa R Anderson and Charles A Holt. 1997. Information cascades in the laboratory. *The American economic review*, pages 847–862.

Justin Cheng, Lada Adamic, P Alex Dow, Jon Michael Kleinberg, and Jure Leskovec. 2014. Can cascades be predicted? In *Proceedings of the 23rd international conference on World wide web*, pages 925–936. ACM.

David Easley and Jon Kleinberg. 2010. *Networks, crowds, and markets: Reasoning about a highly connected world.* Cambridge University Press.

Rong-En Fan, Kai-Wei Chang, Cho-Jui Hsieh, Xiang-Rui Wang, and Chih-Jen Lin. 2008. LIBLINEAR: A library for large linear classification. *Journal of Machine Learning Research*, 9:1871–1874.

Isabelle Guyon, Jason Weston, Stephen Barnhill, and Vladimir Vapnik. 2002. Gene selection for cancer classification using support vector machines. *Machine learning*, 46(1):389–422.

C.J. Hutto, Ewan Klein, Pierpaolo Pantone, and George Berry. 2017. Natural language toolkit: Vader.

David Liedell, David Castillo, Albert Heidmann, and Glenn Mathews. 2015. Yik yak and information cascades: A study of viral yaks.

Edward Loper and Steven Bird. 2017. Natural language toolkit: Tag.

Zongyang Ma, Aixin Sun, and Gao Cong. 2013. On predicting the popularity of newly emerging hashtags in twitter. *Journal of the American Society for Information Science and Technology*, 64(7):1399–1410.

Vlad Niculae and Cristian Danescu-Niculescu-Mizil. 2016. Conversational markers of constructive discussions. In *Proceedings of NAACL-HLT (to appear)*, June.

Dario Taraborelli and Giovanni Luca Ciampaglia. 2010. Beyond notability. collective deliberation on content inclusion in wikipedia. In *Self-Adaptive and Self-Organizing Systems Workshop (SASOW), 2010 Fourth IEEE International Conference on*, pages 122–125. IEEE.

Oren Tsur and Ari Rappoport. 2012. What's in a hashtag?: content based prediction of the spread of ideas in microblogging communities. In *Proceedings of the fifth ACM international conference on Web search and data mining*, pages 643–652. ACM.

Wikipedia. 2017. Deletion policy, May.

Justine Zhang, Ravi Kumar, Sujith Ravi, and Cristian Danescu-Niculescu-Mizil. 2016. Conversational flow in oxford-style debates. *arXiv preprint arXiv:1604.03114*.