Team Skeletor at Touché 2021: Argument Retrieval and Visualization for Controversial Questions

Notebook for the Touché Lab on Argument Retrieval at CLEF 2021

Kevin Ros[1,2], Carl Edwards[1,2], Heng Ji[1] and ChengXiang Zhai[1]

[1]University of Illinois at Urbana-Champaign, 201 N Goodwin Ave, Urbana, Illinois 61801, U.S.A.
[2]Equal Contribution

Abstract

Arguments are a critical part of education and political discourse in society, especially since more and more information is available online. In order to access this information, argument retrieval is a necessary task. In this work, we leverage the existing techniques of BM25 and BERT-based passage embedding similarity and introduce a new information retrieval technique based on manifold approximation. Evaluation results on the Touché @ CLEF 2021 topics and relevance scores show that the manifold-based approximation helps discover higher-quality arguments. Furthermore, we use these retrieval methods to visualize argument progression for users watching debates. The visualization results show promising directions for future exploration.

Keywords

information retrieval, argument, manifold approximation, visualization

1. Introduction

Arguments are an important part of education and political discourse in society. As the amount of information and social media use grows on the internet, especially surrounding controversial topics, it is critical to improve access to relevant debates, thereby improving public understanding of divisive topics [1, 2]. Furthermore, traditional search engines are often limited in their ability to effectively display and update relevant information during a live debate, especially when the debate topics are constantly changing.

This paper attempts to address these concerns by investigating both argument retrieval and visualization. More specifically, we participate in Touché @ CLEF 2021 [3, 4], which presents two distinct argument retrieval tasks: retrieving arguments for controversial questions and retrieving arguments for comparative questions. We focus on the first task, with the goal of supporting users by retrieving and visualizing relevant arguments and sentences for controversial questions. This argument retrieval task goes beyond traditional information retrieval because the retrieval methods need to capture both relevance and argument strength.

As the basic retrieval models have performed well on this task [5], in addition to the standard baseline BM25 and BERT embedding-based retrieval we explore a new approach in which we

CLEF'21: Conference and Labs of the Evaluation Forum, September 21–24, 2021, Bucharest, Romania
kjros2@illinois.edu (K. Ros); cne2@illinois.edu (C. Edwards); hengji@illinois.edu (H. Ji); czhai@illinois.edu (C. Zhai)

leverage the properties of manifold approximation, which is commonly used for dimensionality reduction [6], as a pseudo-relevance-feedback reranking approach. The manifold-based reranking approach assumes the highest-ranked initially retrieved arguments are relevant to the controversial question, and computes a directed-edge existence probability from each argument to all other arguments in the corpus.

Our hypothesis is that strong, complete, and relevant arguments will have many other arguments "pointing" to it. That is, these arguments should have many high-probability incoming directed edges. Thus, we rerank the arguments based on the aggregation of their incoming edge probabilities. Furthermore, we build on these retrieval approaches to visualize the topics and trajectory of real-time debates as they progress per word with respect to a reference corpus.

Experiments using the args.me corpus [7] and the Touché @ CLEF 2021 [5] topics and relevance scores show that our manifold-based ranking formula improves upon BM25 in argument quality. Additionally, our exploration of visualization techniques using the args.me corpus and a spoken debate shows promise in the direction of debate summarization and augmentation.

2. Related Work

Our retrieval methods are inspired by passage-level evidence, as we treat each argument as a collection of sentences [8]. We follow the general methods described by SBERT's retrieval and re-ranking.[1] Zhao et al. [9] use manifold-based text representations of sentences in the biomedical domain to capture the geometric relationships between sentences. Other work also incorporates manifold learning into text representations [10, 11, 12]. To our knowledge, we are the first to incorporate sentence-level manifold representations into information retrieval.

Regarding conversation augmentation, Lyons et al. investigate leveraging dual-purpose speech, which they define as speech socially appropriate to humans and meaningful to computers [13]. Their software plays the role of an assistant (recording dates, scheduling events) rather than introducing additional knowledge to the conversations, which is what we aim to do. Boyd et al. propose to augment conversations with prosody information to help users with autism detect atypical prosody [14]. We attempt to introduce similar metadata to the debates (however, in the form of conversational topics) as well as introduce additional arguments directly related to the topics being discussed. Popescu-Belis et al. introduce a speech-based just-in-time retrieval system which uses semantic search [15]. That is, they record and transcribe conversations, and provide relevant documents to the participants of the conversation in real-time. Their search methods are based on keywords previously spoken during the conversation using ASR (automatic speech recognition) [16]. A word is considered a keyword if it is in the ASR transcript and is not a stopword, or if it is in a pre-constructed list. Thus, the search queries are limited to what has already been spoken, and high-level dependencies between previously discussed ideas cannot be leveraged. We believe our visualization approaches better address both of these issues.

There has also been much work in the general field of visualizing information retrieval [17], but none of these approaches combine BERT and manifold-based dimensionality reduction to allow for more fine-grained understanding of arguments over time.

[1] https://www.sbert.net/examples/applications/retrieve_rerank/README.html

3. Argument Retrieval

In the following subsections, we describe our argument retrieval methods and results. Each approach retrieves arguments from the args.me corpus (version 2020-04-01), which consists of 387,740 arguments scraped from various online debate portals [7]. For each argument entry in the corpus, we only consider the text in the “premise" field. Our methods are primarily evaluated using the topics and relevance scores from Touché @ CLEF 2021, and we also include the scores of our methods on last year’s iteration of the competition for completeness. The relevance scores from last year consist of -2 (non-argument) or a range from 1 (low relevance, weak argument) to 5 (high relevance, strong argument. This year’s relevance scores use the same range, however, they consist of two separate dimensions: argument relevance and argument quality. There are 50 distinct topics each consisting of a short “title" field and a longer “description" and “narrative" fields. For our queries, we only use the “title" field. Some examples of “title" fields include “Do we need sex education in schools?" and “Should stem cell research be expanded?".

3.1. Methods

3.1.1. BM25

For our baseline approach, we use BM25. BM25 is a bag-of-words ranking formula that relies on keyword matching between a query and a collection of arguments, along with various weighting heuristics. To process, index, and search arguments, we use Pyserini, which is a Python-based information retrieval toolkit built over Anserini and Lucene [18]. All argument premises are processed and indexed using the default Pyserini settings. This includes stopword removal and stemming. All queries are also processed similarly. We use Pyserini’s provided BM25 implementation to search the corpus, only adjusting the k_1 and b parameters. We tune the parameters on last year’s topics and relevance scores.

3.1.2. Semantic Search

Given that BM25 only matches exact terms, we explore the effectiveness of encoder-based k nearest neighbor search to help bridge potential vocabulary gaps. To do this, we first split the premises of each argument by sentence into smaller passages of approximately 200 words each. Then, we encode each passage using the msmarco-distilbert-base-v3 encoder model provided by Sentence Transformers [19]. At a high level, msmarco-distilbert-base-v3 is a BERT-based [20] Siamese sentence encoder fine-tuned for question-answering on the MS MARCO data set [21]. The passage embeddings are stored and indexed using the hnwslib Python library [22], which provides an approximate nearest-neighbor lookup index using hierarchical navigable small world graphs. Each topic title is also encoded using msmarco-distilbert-base-v3, and given the encoded topic, we search for the approximate top k nearest neighbor passages. The top arguments are ordered based on the maximum cosine similarity between the topic and any of its passages. All parameters are again tuned using the previous iteration of the task.

We also investigate combining the scores returned via semantic search with those returned using BM25. To calculate this, we use the following formula:

$$score_{BM25} + \alpha \times score_{semantic}$$

3.1.3. Manifold Approximation

Our third argument retrieval approach attempts to leverage the techniques utilized in UMAP (Uniform Manifold Approximation and Projection) [6]. UMAP is a dimensionality reduction technique that first approximates a uniform manifold for each data point and patches together their local fuzzy simplicial set representations, where a simplicial set is a higher-dimensional generalization of a directed graph. Then, this topoligical representation is used to assess and optimize lower-dimensional representations. A full theoretical description of UMAP is beyond the scope of this paper, so we focus solely on the computational aspects of UMAP's manifold approximation which are relevant to our retrieval approach.

To approximate a uniform manifold for each data point x_i, UMAP first finds the k nearest neighbors to x_i. Then, it defines ρ_i and σ_i, where

$$\rho_i = min\{d(x_i, x_{i_j}) | 1 \leq j \leq k, d(x_i, x_{i_j}) > 0\}, \tag{1}$$

$$\sum_{j=1}^{k} \exp(\frac{-max(0, d(x_i, x_{i_j}) - \rho_i)}{\sigma_i}) = \log_2(k), \tag{2}$$

and $d(x_i, x_{i_j})$ is the distance between x_i and x_{i_j}. Intuitively, ρ_i is the distance to x_i's closest neighbor (in our case, the most similar passage) and σ_i smooths and normalizes the distances to the nearest neighbors. Next, UMAP calculates the following weights between data points:

$$w((x_i, x_j)) = \exp(\frac{-max(0, d(x_i, x_{i_j}) - \rho_i)}{\sigma_i}). \tag{3}$$

Calculating this for every data point x_i results in a k-granularity weighted adjacency matrix between all points in the data. The authors of UMAP note that $w((x_i, x_j))$, or entry i, j of the weighted adjacency matrix, can be interpreted as the probability that a directed edge from x_i to x_j exists.

For the purposes of argument retrieval, our hypothesis is that strong, complete, and relevant arguments will have many other arguments "pointing" to it. That is, these arguments should have many high-probability incoming directed edges. Thus, for a given topic title, we first search using the aforementioned interpolated BM25 and semantic retrieval methods. We encode all of the passages for the top n arguments. Next, for each encoded passage, we find the k nearest neighbors and calculate (1), (2), and (3) as described above. Finally, we score each argument by the sum of all directed edges pointing to the argument.

Note that the sum of these calculated passage weights possess different properties than just the sum of the passage similarities. Most notably, Equation 2 constrains the scaled sum of distances to $\log_2(k)$, where k is the number of nearest neighbors. Our understanding is that this calculation gives importance to points that have fewer highly-similar (closer) neighbors. For example, if we have two points (x) and (y), and the (point,distance) pairs of their three nearest neighbors are

$$(x) : [(a, 0.1), (b, 0.2), (c, 0.9)] \qquad (y) : [(d, 0.1), (e, 0.2), (f, 0.3)]$$

Table 1
Performance on Touché 2021 and 2020 Topics and Relevance Scores

Run Name	Relevance nDCG@5	Quality nDCG@5	2020 nDCG@5
bm25	0.661	0.822	0.6214
semantic	0.570	0.671	0.3475
bm25-0.7semantic	0.667	0.815	0.6347
manifold	0.666	0.827	0.5417
manifold-c10	0.666	0.818	0.5906

then the weight between (x) and (b) will be higher than (y) and (e), even though they have the same relative distances. Here are the resulting weights from the manifold calculation ($\sigma_x = 0.179741, \sigma_y = 0.113319$):

$$(x) : [(a, 1), (b, 0.5733), (c, 0.0117)] \qquad (y) : [(d, 1), (e, 0.4138), (f, 0.1712)]$$

Intuitively, this may help reduce importance of passages that are similar to many other passages, as that passage will contribute lower weights to other passages.

3.2. Results

We submitted five runs to Touché 2021, and the performance measures for these five runs are listed in Table 1. The 2021 runs are judged in two dimensions: argument relevance and argument quality, which correspond to the second and third columns of the table, respectively. We also include the performance of our retrieval models on the topics and relevance scores from Touché 2020 as a reference, see column three. All measures are calculated using normalized discounted cumulative gain at five (nDCG@5).

The first run, “bm25", corresponds to the approach outlined in Section 3.1.1. We tune the parameters using grid search and arrived at $k_1 = 3.2$ and $b = 0.2$ using the 2020 topics and relevance scores. The next row, “semantic", corresponds to Section 3.1.2. We set the number of nearest neighbors $k = 1000$ for each topic. Next, “bm25-0.7semantic" denotes the interpolation of the two aforementioned approaches, with an α value of 0.7. The final two rows correspond to the approach described in Section 3.1.3. For “manifold", we assume the top 3 arguments from “bm25-0.7semantic" are relevant and search for $k = 50$ nearest neighbors for each argument passage. The retrieved passages are completely reranked by aggregating the weights over each argument. For “manifold-c10", we perform the exact same search, but only rerank the top 10 arguments of the “bm25-0.7semantic" run.

For this year’s evaluations, our best-performing run with respect to relevance is “bm25-0.7semantic". However, all of our other runs which utilize BM25 (i.e., excluding “semantic") perform similarly. With respect to quality, our best-performing run is “manifold". Here, it is promising that “manifold" outperformed “manifold-c10", as this implies that the manifold technique is able to increase argument quality by retrieving arguments outside of the top 10 initially-ranked arguments.

It is unclear whether or not our initial hypothesis is supported by the scores listed in Table 1. The evaluation metrics from this year seem to support our hypothesis in the context of our

“manifold" run, but last year’s results show a decrease in performance. This may be because last year’s relevance scores combine many different measures into a single dimension. Furthermore, it is difficult to separate out the effects of BM25 on our manifold-based approaches, since it appears that these approaches perform similarly. This, along with the high scores of our “bm25" run, stresses the importance of well-tuned robust models. Overall, these results are a step in the right direction for our hypothesis, but more analysis is needed to draw firm conclusions.

4. Visualization

While a ranked list of document snippets is often sufficient for ordinary web search, such a list is not necessarily optimal for showing results of argument retrieval to the users because it is common to discuss many topics during a debate and the user may want to see the topical structure. These topics may be discussed at length, briefly mentioned, or revisited as the debate unfolds. Traditional search engines, which require explicit user querying, often display relevant documents and arguments in a ranked list, which makes it difficult to effectively capture and visualize these topic changes. For example, it may be too time consuming for a participant in a debate to constantly search for and read all of the relevant documents. Or, someone may want a high-level summary of the debate at various points. Thus, we explore various visualization techniques to help mitigate these concerns. This is accomplished by minimizing the necessity of constant user input as well as visualizing these structural topic changes. Visualization of search results has been studied before [17, 23, 24, 25]; however, existing visualization methods will not work well for our use case, so we explore new approaches.

For our visualization exploration, we utilize the args.me corpus to help summarize and augment debates in real time. We demonstrate our visualization methods on the publicly-available debate between Bill Nye and Ken Ham on Evolution vs. Creationism.[2] We chose this debate primarily because YouTube provides an accurate transcript of the debate with timestamps, and because of the debate’s diverse topic coverage.

The YouTube transcript timestamps occur approximately every 3 seconds and contain approximately $1 - 8$ words per timestamp. We maintain these groupings for our analysis. The text for the transcript referenced in the analysis is in Table 4. The full text of each referenced argument ID is available on GitHub.[3]

4.1. Visualization Approach with BM25

For any given timestamp t_i, we define a look-back window of size n and collect all the terms that occurred between t_{i-n} and t_i. Then, we search the args.me corpus using our BM25 retrieval approach outlined in Section 3.1.1, with the query being the collected transcript terms. We record the ranks of the top k arguments returned. We choose BM25 because it is well-known to be robust and efficient. Repeating this over a given interval of timestamps results in a smoothed argument-level summary for the interval.

[2] https://www.youtube.com/watch?v=z6kgvhG3AkI

[3] https://github.com/kevinros/toucheRetrievalVisualization/tree/main/arguments

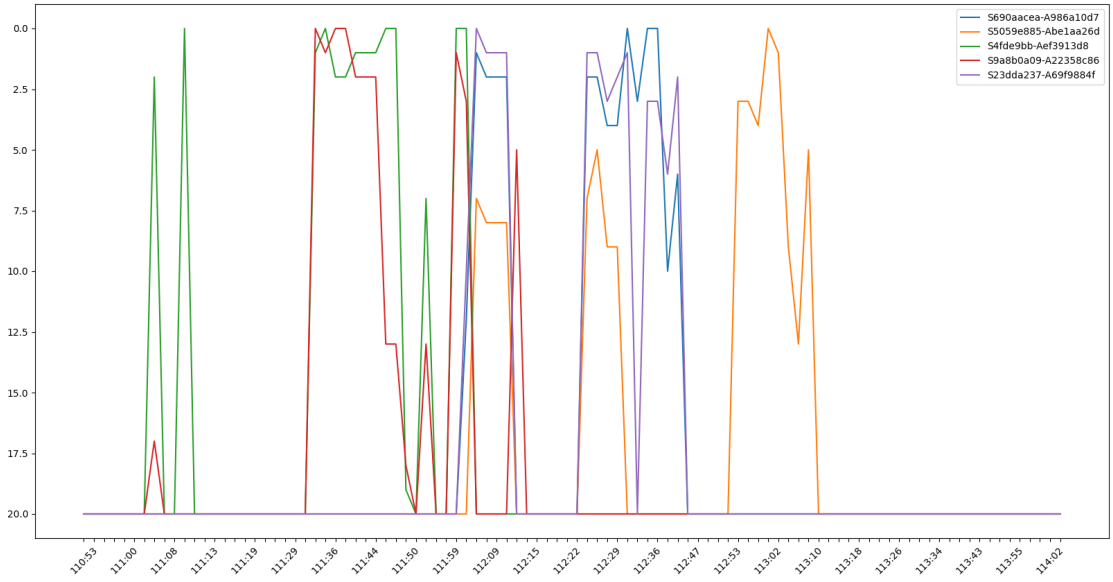


Figure 1: Rankings of the five most frequent arguments over the transcript window 110:53 - 114:04.

Table 2
Arguments from Figure 1

Argument ID	Representative Topic(s)
S4fde9bb-Aef3913d8	description of a bicycle incident
S9a8b0a09-A22358c86	understanding scriptures, the gospel, God
S690aacea-A986a10d7	creator of universe, infinite power, God
S23dda237-A69f9884f	having unlimited power, omnipotence of God
S5059e885-Abe1aa26d	the justness of God

As an example, consider the debate time interval 110:53-114:04. Each timestamp and corresponding text is listed in Table 4. We define a look-back window of size $n = 5$ and retrieve the top $k = 20$ arguments for each timestamp. Then, we collect the number of times an argument is ranked in the top 20 arguments across all timestamps, and consider only the five most frequent arguments. Figure 1 displays the ranks of these five argument at each timestamp, and Table 2 lists a high-level description of each argument. The parameters are manually tuned to demonstrate the benefits and drawbacks of this visualization approach.

Of the five arguments returned, S4fde9bb-Aef3913d8 seems to be topically irrelevant to the transcript text. Interestingly, this argument appears to also be a transcript, and thus it contains many filler words (such as “uh") also present in the debate transcript. It appears to be playing the role of a background language model. The other four arguments seem to be relevant as they discuss topics and themes present in the transcript at different timestamps. From 111:29 to 111:50, argument S9a8b0a09-A22358c86 is one of the highest-ranked, and it discusses “God", “His kingdom", “scripture", and “His actions". From 112:22 to 112:47, we find that arguments

Table 3
Arguments from Figure 2

Argument ID	Representative Topic(s)
S379f0b2-Ab47bd29b	showing the validity of theistic evolution
S56a34f98-A3adb8db7	biblical creationism, unfalsifiable
Scf918055-Af439fe9a	heaven, hell, stars, God
S70cdd68a-A5b15aee9	physics, star formation, modern science
S9ad5951e-A78e904a7	astronomy in the context of the Quran

S690aacea-A986a10d7 and S23dda237-A69f9884f are ranked the highest. Both arguments discuss the powers of the creator of the universe. From 112:53 to 113:10, we observe that argument S5059e885-Abe1aa26d is the highest-ranked, which argues in favor of the justness of God.

One use case for this visualization technique is to help participants of the debate better analyze and justify their stance. For example, the participants can draw on the additional knowledge provided by the retrieved arguments to strengthen their own arguments in real-time. On the other hand, it is also possible that rebuttals to participants' arguments will be retrieved, which could help increase the overall robustness of the debate by exposing counterpoints.

In order to reduce noise and irrelevant arguments, we also explore the possibility of allowing users to specify the search terms or arguments. More specifically, using pre-defined sets of terms, we search the args.me corpus with BM25 to find the most relevant arguments to the provided terms. Then, we display the frequencies of the returned arguments using the methods outlined above, except we consider ranks through 100 rather than 20.

Consider the same debate time interval and the keyword groups "bible god creationism" and "heavens astronomy stars". Figure 2 displays the frequencies of the five most relevant arguments to each keyword group. The first five argument IDs in the legend correspond to the first keyword group, and the second five argument IDs in the legend correspond to the second keyword group. Additionally, high-level descriptions of the arguments that appear in Figure 2 are listed in Table 3. The first two arguments are from "bible god creationism" and the last three arguments correspond to "heavens astronomy stars". From Figure 2, we see that arguments relevant to both keyword groups are highly ranked between 112:10 and 112:36, indicating that the keywords in the retrieved arguments strongly match the keywords from the debate transcript in the time interval.

An important benefit of this visualization technique is that it allows the user to specify specific topics before, during, or after a debate in order to easily track various topic occurrences for further analysis. For example, a user looking to get a high-level summary of a debate can examine the ranking frequencies of known arguments in order to pinpoint the most relevant points in the debate.

As this visualization approach provides a high-level overview of a debate by referencing relevant arguments using keywords, it abstracts away from the actual content of the debate and relevant sentences within arguments. To help address this issue, we explore a more fine-grained visualization approach in the following subsection.

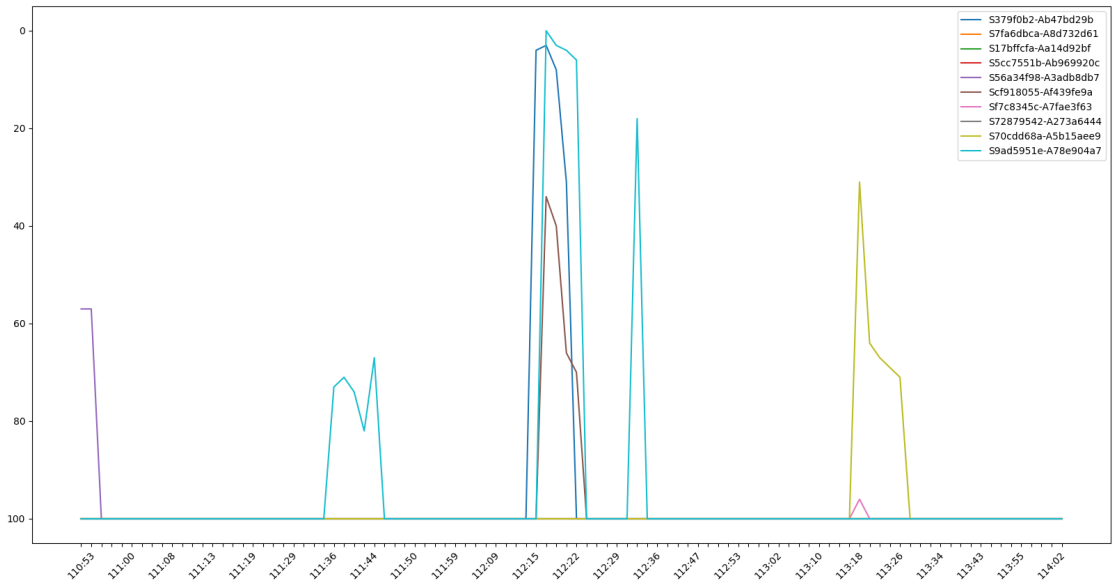


Figure 2: Rankings of the five most relevant arguments to “bible god creationism" and to “heavens astronomy stars" over the transcript window 110:53 - 114:04.

4.2. Visualization Approach with UMAP

The advent of new Transformer-based language models such as BERT [20] have lead to impressive improvement on a variety of NLP tasks. We seek to use BERT’s semantic representation space to better visualize the dynamics of arguments. To do so, we take advantage of UMAP [6]. The goal of UMAP is to visualize high-dimensional embeddings in a low-dimensional space while preserving topological and structural properties. Using the same BERT-based encoder discussed in Section 3.1.2, we combine the encodings of the sentences of relevant arguments and the “caterpillar embeddings" of our debate transcript to visualize how the debate evolves over time. This approach allows us to analyze fine-grained topic changes as they unfold in the debate, as well as their relevance to a reference corpus.

4.2.1. Caterpillar Embeddings

Caterpillar embeddings are used to track the course of the debate over time. They consist of a sequence of encoder representations taken from across the debate. A naïve approach is to slide a window of size n over the sequence of words w in the transcript with stride s. However, this has the downside of both adding and removing information (words) at each step. Instead, we split each step into two: a growth step and a contraction step. Given a window from word w_i to w_{i+n} of the transcript for some i, the next window will grow to be from w_i to w_{i+n+s}. The subsequent window will be a contraction: it will range from w_{i+s} to w_{i+n+s} Hence, this “caterpillar embedding" technique moves along the transcript of the debate like a caterpillar inching along. At step t, the start and end of the window, S and E respectively, are calculated

as follows:

$$S = w_{s\lfloor \frac{t}{2} \rfloor},\ E = w_{s\lfloor \frac{t+1}{2} \rfloor + n}$$

4.2.2. Argument Retrieval-Based Semantic Visualizations

In order to better define the topology of the semantic space, we extract the top k most frequent arguments over the transcript interval 110:53 to 127:01 as described in Section 4.1 from the args.me corpus, split them into sentences, and encode the sentences using the previously-mentioned BERT-based sentence encoder. We combine these argument embeddings with the caterpillar embeddings of the debate transcript and project them into two dimensions using UMAP. This creates a path of the debate as it visits different arguments in the semantic space. We can then use the nearby neighbors of the caterpillar embeddings as relevant arguments to show the user at a given timestamp. The full animation can be found on GitHub.[4]

Regardless of which k value we use, we find that this UMAP projection does not preserve the original space well regarding nearest neighbors. We believe this is because of the large differences between the semantic structures of the conversational YouTube debate and the written structures of the corpus debates. To mitigate this, we use a nearest neighbor search in the original space, and we plot the debate embedding using its m nearest neighbors. Through empirical exploration, we find that $m = 100$ and $k = 100$ yields the clearest results. Additionally, we consider the same window as explored in Section 4.1, namely 110:53 to 114:04. Note that the transcript of the debate in this window is available in Table 4. The resulting path at various timestamps is shown in Figure 3.

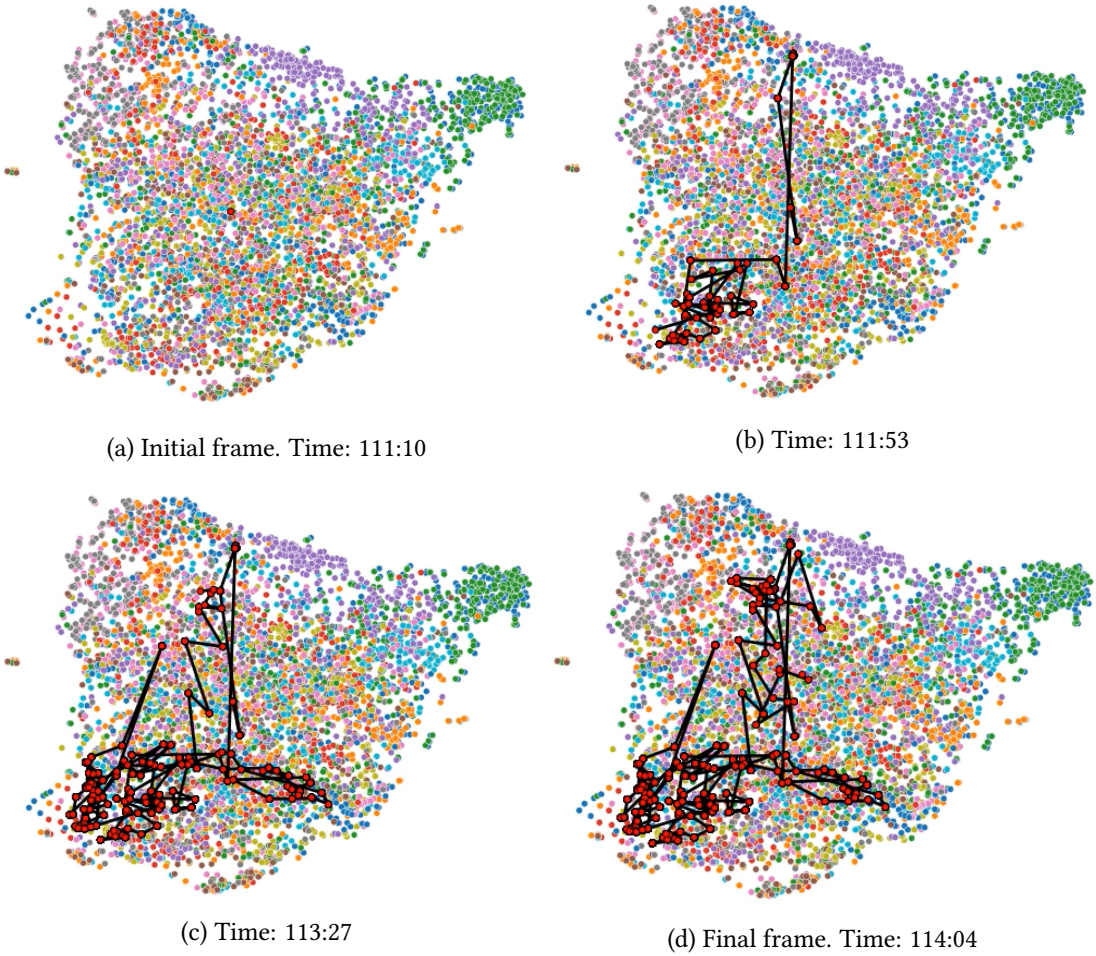
The argument quickly moves to the lower left quadrant, which we find to signify the creation of the universe and heavens, particularly in relation to God. The path briefly moves to the right, when the debate focuses more on the omnipotence and omniscience of God. Finally, the debate moves upward, when the discussion changes to physics, life science, and astronomy. The full video can also be found on GitHub.[5]

In Figure 3, we clearly see groupings of arguments' topics and how they change over time. Interestingly, we can also examine the topic path through the corpus that the YouTube debate took. This could be used to track debate topic progression in a visual manner, and augment live debates with both relevant information at the current point as well as relevant information for future, forecasted points. More work is needed, however, to investigate the effects of parameter selection and the effectiveness in various domains.

5. Conclusion and Future Work

In this work, we apply several techniques to the Touché Argument Retrieval task, such as BM25, semantic search, and manifold-based reranking. Among them, we find that the manifold-based reranking was sometimes more effective in returning high-quality arguments when compared to BM25. In the future, we hope to compute the manifold weights for every argument in the data set as a preprocessing step, and investigate efficient ways to combine these weights with

[4] https://github.com/kevinros/toucheRetrievalVisualization/blob/main/animations/full_anim.mp4
[5] https://github.com/kevinros/toucheRetrievalVisualization/blob/main/animations/100top_mean_anim.mp4



(a) Initial frame. Time: 111:10

(b) Time: 111:53

(c) Time: 113:27

(d) Final frame. Time: 114:04

Figure 3: Visualization of the evolution-creationism debate through the retrieved argument space.

retrieval methods that perform well along the relevance dimension, in order to return the strongest and the most relevant arguments.

To better display search results to users in argument retrieval, we also introduce various visualization techniques based on BM25 keyword matching and UMAP dimensionality reduction, which shows promise in the direction of debate augmentation. Although the benefits of this augmentation are difficult to quantify, we believe it will help improve debate understanding and retention, as well as open up avenues for future work. We also hope to improve the visualization by further testing different parameters, retrieval techniques, and background corpora.

Table 4
Transcript of creationism debate from 110:53 to 114:04

Timestamp	Text	Timestamp	Text
110:53	creationism account for the celestial	112:23	um and just to show us he's an
110:55	bodies	112:25	all-powerful god he's an infinite god so
110:56	planets stars moons moving further and	112:27	i made the stars and he made them to
110:59	further apart	112:29	show us how great he is and he is he's
111:00	and what function does that serve in the	112:31	an
111:02	grand design	112:32	infinite creator god and the more that
111:04	well when it comes to uh looking at the	112:34	you understand what that means that god
111:06	universe of course we believe that in	112:36	is all-powerful infinite you stand back
111:08	the beginning god created the heavens	112:39	in all
111:09	and the earth	112:39	you realize how small we are you realize
111:10	and i believe our uh creationist	112:42	wow that god would consider this planet
111:12	astronomers would say yeah you can	112:44	is is so significant that he created
111:13	observe	112:47	human beings here
111:14	the universe expanding uh why god is	112:48	knowing they would sin and yet stepped
111:16	doing that in fact in the bible it even	112:50	into history to die for us be raised
111:17	says he stretches out	112:52	from the dead
111:19	the heavens and seems to indicate that	112:53	to offer us a free gift to salvation wow
111:22	there is	112:56	what a god and that's what i would say
111:22	an expansion of the universe and so	112:58	when i see
111:26	we would say yeah that you can observe	112:59	the universe as it is mr nye one minute
111:27	that that fits with	113:02	any response
111:29	what we call observational science	113:03	there's a question that troubles us all
111:30	exactly why god did it that way	113:05	from the time we are
111:32	uh i can't answer that question of	113:08	absolutely youngest and first able to
111:34	course uh because you know the bible	113:10	think
111:36	says that uh	113:11	and that is where did we come from where
111:37	god made uh the heavens for for his	113:13	did i come from
111:39	glory and that's why he made	113:15	and this question is so compelling that
111:41	uh the stars that we see out there and	113:18	we've
111:44	it's uh it's to tell us how great he is	113:19	invented the science of astronomy we've
111:46	and how big he is and in fact i think	113:22	invented life science we've invented
111:48	that's the the thing about the universe	113:24	physics we've discovered these natural
111:49	the universe is	113:26	laws
111:50	so large so big out there one of our	113:27	so that we can learn more about our
111:53	planetarium programs	113:29	origin and where we came from
111:54	looks at this we go in and show you uh	113:31	to you when it says he invented the
111:57	how large the universe is	113:34	stars also
111:59	and i think it shows us how great god	113:36	that's satisfying you're done oh good
112:02	is uh how big he is that he's an	113:39	okay to me when i look at the night sky
112:04	all-powerful god he's an infinite god	113:41	i want to know what's out there
112:07	uh an infinite all-knowing god who	113:43	i'm driven i want to know if what's out
112:09	created the universe	113:46	there is any part of me
112:10	to show us his power i mean can you	113:48	and indeed it is the oh by the way
112:13	imagine that and the thing that's	113:51	i find compelling you are satisfied and
112:14	remarkable	113:55	the big thing i want from you
112:15	in the bible for instance says on the	113:56	mr ham is can you come up with something
112:17	fourth day of creation	113:59	that you can predict
112:18	and and oh he made the stars also it's	114:00	do you have a creation model that
112:21	almost like oh by the way i made the	114:02	predicts something that will happen in
112:22	stars	114:04	nature

References

[1] A. Perrin, Social media usage, Pew research center 125 (2015) 52–68.

[2] M. Del Vicario, G. Vivaldo, A. Bessi, F. Zollo, A. Scala, G. Caldarelli, W. Quattrociocchi, Echo chambers: Emotional contagion and group polarization on facebook, Scientific reports 6 (2016) 1–12.

[3] A. Bondarenko, L. Gienapp, M. Fröbe, M. Beloucif, Y. Ajjour, A. Panchenko, C. Biemann, B. Stein, H. Wachsmuth, M. Potthast, M. Hagen, Overview of Touché 2021: Argument Retrieval, in: D. Hiemstra, M.-F. Moens, J. Mothe, R. Perego, M. Potthast, F. Sebastiani (Eds.), Advances in Information Retrieval. 43rd European Conference on IR Research (ECIR 2021), volume 12036 of *Lecture Notes in Computer Science*, Springer, Berlin Heidelberg New York, 2021, pp. 574–582. URL: https://urldefense.com/v3/__https://link.springer.com/chapter/10.1007/978-3-030-72240-1_67__;!!DZ3fjg!qiIStvQ7N0tMq0XWzNrBDwdUszdG_1Cm5f0npcVKkP9lL7BwqrITiN5eveoZNiWt_Q$. doi:10.1007/978-3-030-72240-1_67.

[4] A. Bondarenko, L. Gienapp, M. Fröbe, M. Beloucif, Y. Ajjour, A. Panchenko, C. Biemann, B. Stein, H. Wachsmuth, M. Potthast, M. Hagen, Overview of Touché 2021: Argument Retrieval, in: G. Faggioli, N. Ferro, A. Joly, M. Maistro, F. Piroi (Eds.), Working Notes Papers of the CLEF 2021 Evaluation Labs, CEUR Workshop Proceedings, 2021.

[5] A. Bondarenko, M. Fröbe, M. Beloucif, L. Gienapp, Y. Ajjour, A. Panchenko, C. Biemann, B. Stein, H. Wachsmuth, M. Potthast, M. Hagen, Overview of Touché 2020: Argument Retrieval, in: L. Cappellato, C. Eickhoff, N. Ferro, A. Névéol (Eds.), Working Notes Papers of the CLEF 2020 Evaluation Labs, volume 2696 of *CEUR Workshop Proceedings*, 2020. URL: https://urldefense.com/v3/__http://ceur-ws.org/Vol-2696/__;!!DZ3fjg!qiIStvQ7N0tMq0XWzNrBDwdUszdG_1Cm5f0npcVKkP9lL7BwqrITiN5ever8RPesww$.

[6] L. McInnes, J. Healy, J. Melville, Umap: Uniform manifold approximation and projection for dimension reduction, arXiv preprint arXiv:1802.03426 (2018).

[7] Y. Ajjour, H. Wachsmuth, J. Kiesel, M. Potthast, M. Hagen, B. Stein, Data Acquisition for Argument Search: The args.me corpus, in: C. Benzmüller, H. Stuckenschmidt (Eds.), 42nd German Conference on Artificial Intelligence (KI 2019), Springer, Berlin Heidelberg New York, 2019, pp. 48–59. doi:10.1007/978-3-030-30179-8_4.

[8] J. P. Callan, Passage-level evidence in document retrieval, in: SIGIR'94, Springer, 1994, pp. 302–310.

[9] D. Zhao, J. Wang, H. Lin, Y. Chu, Y. Wang, Y. Zhang, Z. Yang, Sentence representation with manifold learning for biomedical texts, Knowledge-Based Systems 218 (2021) 106869.

[10] T. B. Hashimoto, D. Alvarez-Melis, T. S. Jaakkola, Word embeddings as metric recovery in semantic spaces, Transactions of the Association for Computational Linguistics 4 (2016) 273–286.

[11] S. Hasan, E. Curry, Word re-embedding via manifold dimensionality retention, Association for Computational Linguistics (ACL), 2017.

[12] B. Jiang, Z. Li, H. Chen, A. G. Cohn, Latent topic text representation learning on statistical manifolds, IEEE transactions on neural networks and learning systems 29 (2018) 5643–5654.

[13] K. Lyons, C. Skeels, T. Starner, C. M. Snoeck, B. A. Wong, D. Ashbrook, Augmenting

conversations using dual-purpose speech, in: Proceedings of the 17th annual ACM symposium on User Interface Software and Technology, 2004, pp. 237–246.
[14] L. E. Boyd, A. Rangel, H. Tomimbang, A. Conejo-Toledo, K. Patel, M. Tentori, G. R. Hayes, Saywat: Augmenting face-to-face conversations for adults with autism, in: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, 2016, pp. 4872–4883.
[15] A. Popescu-Belis, M. Yazdani, A. Nanchen, P. N. Garner, A speech-based just-in-time retrieval system using semantic search, Technical Report, Idiap, 2011.
[16] P. N. Garner, J. Dines, T. Hain, A. El Hannani, M. Karafiat, D. Korchagin, M. Lincoln, V. Wan, L. Zhang, Real-time ASR from meetings, Technical Report, Idiap, 2009.
[17] M. Hearst, Search user interfaces, Cambridge university press, 2009.
[18] J. Lin, X. Ma, S.-C. Lin, J.-H. Yang, R. Pradeep, R. Nogueira, Pyserini: An easy-to-use python toolkit to support replicable ir research with sparse and dense representations, arXiv preprint arXiv:2102.10073 (2021).
[19] N. Reimers, I. Gurevych, Sentence-bert: Sentence embeddings using siamese bert-networks, arXiv preprint arXiv:1908.10084 (2019).
[20] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, Bert: Pre-training of deep bidirectional transformers for language understanding, arXiv preprint arXiv:1810.04805 (2018).
[21] T. Nguyen, M. Rosenberg, X. Song, J. Gao, S. Tiwary, R. Majumder, L. Deng, Ms marco: A human generated machine reading comprehension dataset, in: CoCo@ NIPS, 2016.
[22] Y. A. Malkov, D. A. Yashunin, Efficient and robust approximate nearest neighbor search using hierarchical navigable small world graphs, IEEE transactions on pattern analysis and machine intelligence 42 (2018) 824–836.
[23] S. Liu, X. Wang, C. Collins, W. Dou, F. Ouyang, M. El-Assady, L. Jiang, D. A. Keim, Bridging text visualization and mining: A task-driven survey, IEEE transactions on visualization and computer graphics 25 (2018) 2482–2504.
[24] A. M. MacEachren, A. Jaiswal, A. C. Robinson, S. Pezanowski, A. Savelyev, P. Mitra, X. Zhang, J. Blanford, Senseplace2: Geotwitter analytics support for situational awareness, in: 2011 IEEE conference on visual analytics science and technology (VAST), IEEE, 2011, pp. 181–190.
[25] J. Peltonen, K. Belorustceva, T. Ruotsalo, Topic-relevance map: Visualization for improving search result comprehension, in: Proceedings of the 22nd international conference on intelligent user interfaces, 2017, pp. 611–622.