

ARGUMENT GATE

ECE1786 - CREATIVE APPLICATIONS OF NATURAL LANGUAGE PROCESSING - FINAL REPORT.

> UMARPREET SINGH - 1008632945 DHIRAJ TAWANI - 1007504871

WORD COUNT (Without References, Figures, etc.) - 1985 WORD COUNT (With References, Figures, etc.) - 2559 Penalty - 0%

Permissions:

Umarpreet Singh:

Permission to Post Video: Yes Permission to Post Final Report: Yes Permission to Post Source Code: Yes

Dhiraj Tawani:

Permission to Post Video: Yes Permission to Post Final Report: Yes Permission to Post Source Code: Yes

Introduction:

If natural language processing (NLP) is meant to utilize computational techniques to emulate natural human dialogue, then it stands to reason that there should be paradigms for extracting and constructing a method for validating an assessment for perhaps what humans do most with their language - argue [1]. This field is referred to as computational argumentation which focuses on the analysis and synthesis of arguments that can be used in many applications such as virtual assistants [2], search engines [3], and writing.

The goal of the project is to classify an argument with respect to a given statement, judging the argument on multiple factors into four classes (changed view, excellent argument, adequate argument, inadequate argument). This would indeed help writers, debaters, journalists, and bloggers to assess their written work with feedback that target individual aspects of writing an argument. We assume an argument to be excellent since it has influenced a lot of people positively.

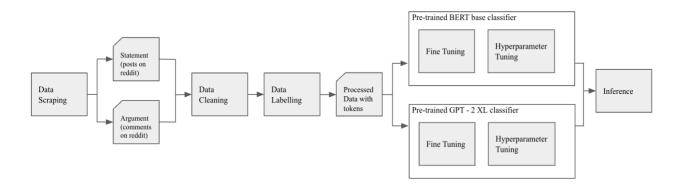


Figure 1: Working of Argument Gate.

Background & Related Work:

Researchers have been developing approaches to extract arguments from natural language text. These approaches belong to the task of Argument Mining, which not only aims to extract argumentative discourse units [4] but also tries to find different sub-structures within arguments [5].

Wachsmuth et al., 2017 [6] present a completely novel work on computational argumentation quality in natural language. It takes into consideration the diverse existing theories and approaches to assess logical, rhetorical, and dialectical quality dimensions and derives a systematic taxonomy from these. This paper defines a platform for the automatic assessment of argumentation quality in natural language.

Lauscher et al., 2020 [7] advance theory-based argument quality research by performing a thorough analysis covering three various domains of online argumentative writing: Q&A forums, debate forums, and review forums. They demonstrated that the relations between theory-based argument quality dimensions can be manipulated to yield performance improvements and demonstrates the utility of theory-based argument quality predictions with respect to the practical Argument Quality (AQ) assessment view. The paper suggests a novel approach of computational theory-based Argument Quality models and establishes that mutually predicting Argument Quality scores can improve the performance of the models (RQ3) and that generally, models benefit from involving out-of-domain training data (RQ4).

Swanson et al., [8] raised a litany of hand-curated features that they claim as indicators of argument presence and quality: Sentence Length (SLEN), Word Length (WLEN), Speciteller (SPTL), Kullback-Leibler Divergence (KLDiv), Discourse (DIS), Part-Of-Speech N-Grams (PNG), Syntactic and (SYN).

Data and Data Processing:

Data Source:

The data is scraped from the **ChangeMyView** subreddit. People post their views or opinions about different topics on this subreddit to have a healthy and open-minded argument. Users who have opposite views, comment on a logical and explained argument in response to the original statement.

Q 🋞 n	🔾 🛞 r/changemyview 🛞 Search Reddit							
	☆ 1.3k ♂	Posted by u/fuckusnowman 2 days ago ③ ⑤ Delta(s) from OP CMV: there's nothing racist about asking an ethnic minority where they're from	1 - R					
		This came up in the news today after some comments made by a Buckingham Palace aide. I know that this is generally considered as a very racist thing to ask but I just can't see it. This is why:						
		 The question is clearly intended to ask about someone's heritage. This is something that many people are proud of, to the extent that they will describe themselves as a hyphenated nationality - Italian-American and so on. Someone my age, in my country (the UK) who is an ethnic minority is demographically likely to be a first or second generation immigrant. I don't understand why effectively asking where someone's parents are from is racist. People ask me where I grew up all the time and I don't regard that as offensive. I enjoy telling people about my background. How could it be offensive? I hand concrete for a number of upper and uppe action of upper a class of the offensive? 1.2k Comments Share Save 	A State of the sta					
		Posted by u/Impossible_Active271 2 days ago						
	187 	Delta(s) from OP CMV: most of the time we should relocate victims of famine instead of giving them food						
		If you would like a humorous explanation of my view, here is a <u>Sam Kinison joke</u> . Now, more seriously : I know that there are famines around the world that are happening because of natural disasters (like floods for example) or because of war, or because it was deliberately forced upon a population. But there is a kind of famine that we have been seeing in every international charity ads and school geography books for 50 years. Yes, I am talking about that decimated African kid with flies all over the face, and his 11 brothers and sisters. From all the documentaries and pictures on the subject, we can clearly see that nothing can grow where they live. The soil is dry and have been so for the last 500 years. Why do they keep living in deserts? In western civilisation, you can find inhabited deserts like Las Vegas, but those places were created knowing it was possible to bring food and water there. If, for some political reason, Las Vegas became a place where you can neither grow food/get water, nor have it brought there, I am						
		□ 215 Comments → Share □ Save …						

Figure 2: ChangeMyView SubReddit.

The main motivation to use this subreddit is due to the rich nature of the subreddit with respect to users, posts, comments, and methodologies to appreciate good content and not suffer from mass spam.

Data Explanation - Posts:

Users create posts with a statement and an explanation of a topic and their opinions. These posts are created with the intention for vivid users to comment and discuss different opinions and in the process, maybe, change the view of a user.



🕼 r/changemyview - Posted by u/Creepy-Soil2698 4 days ago

Δ

^{1.2k} CMV: I don't think any information is trustworthy because no one knows what information is influenced by propaganda and what isn't.

Delta(s) from OP

There's alot of propaganda, manipulation and radicalization going on the internet. There's alot of groups radicalizing people into hateful ideologies. Could it be that alot of the information out there is also influenced by propaganda?

I often feel very confused where to stand with certain political issues because every side tends to be so strongly opinionated and have their own studies and statistics that they use to support their views. Take the trans debate for example. (Note that I'm fully pro-trans) The pro-trans side have their own strong points, opinions and studies that validate their side and prove that the other side is the wrong and bigoted one, the side that I belong to. The anti-trans side also have their own points, opinions and studies that the other side is the wrong and mentally ill one, the side that I strongly believe is wrong.

Although, I'm all for trans people but recently, after "What is a woman?" and the amount of support it got really made me think. How do *I* know which side is spreading an agenda and which side isn't? How do *I* know which side is cherry picking scientific studies, using flawed data etcetera to pust an agenda and which side isn't. I feel the same way about abortion (I'm pro choice), race/gender issues or any political debate for that matter. How do I know which side is more right which side isn't? I don't know if that's even possible.



Figure 3: Example of a post in ChangeMyView Subreddit.

Data Explanation - Comments:

Individuals who wish to provide their opinion or argument on a topic can provide so with the mechanism of a comment. Comments serve as the argument to the problem statement i.e., the post to our model. If an argument changed someone's perspective, they comment a delta (represent change) on that to show appreciation. Moreover, they upvote an excellent argument and downvote an inadequate argument. We scrape these metrics (delta and votes) provide us with information on whether arguments are best or worst for a given statement.

Figure 4: Example of a comment in ChangeMyView Subreddit.

Data Scrapping:

Data is scrapped using the PRAW library. We scrapped all the posts from November 2013 till November 2022. Then all the corresponding comments along with metadata are fetched such as the number of upvotes, and depth in the comment tree. We have scrapped 107717 posts and 1425075 comments in total.

The logs of delta reward given to a comment are present on Delta Log Reddit. In total, we scrapped 56121 delta comments. This metadata is then merged with the main scrapped dataset to build a combined dataset. We also added a data_id column through which can uniquely identify our data point.

data id	delta	statement_id	statement	statement author	arguement_id	argument	argument_author	upvotes	depth
1088814	4.0	orxd63	CMV: Covid Vaccination is NOT going to end covid	TruerInfo	h6kuc06	>Malta has vaccinated the entire eligible popu	Runiat	13.0	0.0
1088815	4.0	orxd63	CMV: Covid Vaccination is NOT going to end covid	TruerInfo	h6kvaje	>Isn't it better to not have ANY covid in your	edwardlleandre	9.0	0.0
1088819	4.0	orxd63	CMV: Covid Vaccination is NOT going to end covid	TruerInfo	h6ku8gs	>Malta has vaccinated the entire eligible popu	barbodelli	5.0	0.0
762913	3.0	gattcf	cmv: The concept of cultural appropriation is	Jamo-duroo	fp2ndod	I think there needs to be a distinction betwee	moon_truthr	41.0	0.0
1085413	3.0	ouloho	[deleted by user]	0	h6z8r01	>Our government however is much more likely to	dublea	10.0	0.0
1240015	0.0	7hqddz	CMV: 'The Future is Female' movement should r	FluffyRadcliffe	dqtav9s	Their goal is not "equality" though. Their goa	cdb03b	-170.0	0.0
115349	0.0	b0iwy1	CMV: Transgender athletes shouldn't compete in	Hey-I-Read-It	eif15m4	Your post is extremely insensitive and I advis	Misspelt	-186.0	0.0
919286	0.0	anivfz	CMV: The controversy surrounding Liam Neeson's	OddlySpecificReferen	eftvkhm	[deleted]	0	-206.0	0.0
919285	0.0	anivfz	CMV: The controversy surrounding Liam Neeson's	OddlySpecificReferen	eftod8v	Reasons why people may be mad: \n\n1. He is tr	Helpfulcioning	-246.0	0.0
1088751	0.0	orzj99	CMV: The US should not re-impose lockdowns/res	0	h6lcgs2	>the vaccines have been proven effective with	roofied_elephant	-267.0	0.0
	1088814 1088815 1088815 1088819 762913 1085413 1240015 115349 919286 919285	1088814 4.0 1088815 4.0 1088815 4.0 1088819 4.0 762913 3.0 1085413 3.0 1240015 0.0 919286 0.0 919285 0.0	I088814 4.0 orxd63 1088815 4.0 orxd63 1088819 4.0 orxd63 1088819 4.0 orxd63 1088819 3.0 gattcf 1085413 3.0 ouloho 1240015 0.0 7hqddz 115349 0.0 b0iwy1 919286 0.0 anivfz	1088814 4.0 orxd63 CMV: Covid Vaccination is NOT going to end covid 1088815 4.0 orxd63 CMV: Covid Vaccination is NOT going to end covid 1088815 4.0 orxd63 CMV: Covid Vaccination is NOT going to end covid 1088819 4.0 orxd63 CMV: Covid Vaccination is NOT going to end covid 1088819 4.0 orxd63 CMV: Covid Vaccination is NOT going to end covid 762913 3.0 gattef cmv: The concept of cultural appropriation is 1085413 3.0 ouloho [deleted by user] 1240015 0.0 7hqddz Female' movement should r 115349 0.0 b0iwy1 CMV: The controversy surrounding Liam Neeson's 919286 0.0 anivfz CMV: The controversy surrounding Liam Neeson's 919285 0.0 anivfz CMV: The US should not re-impose	1088814 4.0 orxd63 CMV: Covid Vaccination is NOT going to end covid TruerInfo 1088815 4.0 orxd63 CMV: Covid Vaccination is NOT going to end covid TruerInfo 1088815 4.0 orxd63 CMV: Covid Vaccination is NOT going to end covid TruerInfo 1088819 4.0 orxd63 CMV: Covid Vaccination is NOT going to end covid TruerInfo 1088819 4.0 orxd63 CMV: Covid Vaccination is NOT going to end covid TruerInfo 762913 3.0 gattef cmv: The concept of cultural appropriation is Jamo-duroo 1085413 3.0 ou1oho [deleted by user] 0 1240015 0.0 7hqddz Female' movement should r FluffyRadcliffe 115349 0.0 b0iwy1 CMV: The controversy surrounding Liam Neeson's OddlySpecificReferen 919286 0.0 anivfz CMV: The controversy surrounding Liam Neeson's OddlySpecificReferen 1088751 0.0 orzj99 CMV: The US should not re-impose 0 <td>10888144.0orxd63CMV: Covid Vaccination is NOT going to end covidTruerInfoh6kuc0610888154.0orxd63CMV: Covid Vaccination is NOT going to end covidTruerInfoh6kvaje10888154.0orxd63CMV: Covid Vaccination is NOT going to end covidTruerInfoh6kvaje10888194.0orxd63CMV: Covid Vaccination is NOT going to end covidTruerInfoh6kvage10888194.0orxd63CMV: Covid Vaccination is NOT going to end covidTruerInfoh6kvage7629133.0gattefcmv: The concept of cultural appropriation isJamo-duroofp2ndod10854133.0ouloho[deleted by user]0h6z8r0112400150.07hqddzFemale' movement should rFluffyRadcliffedqtav9s1153490.0b0iwy1CMV: The controversy surrounding Liam Neeson'sOddlySpecificRefereneftvkhm9192850.0anivfzCMV: The controversy surrounding Liam Neeson'sOddlySpecificRefereneftod8v10887510.0orzj99CMV: The US should not re-impose0h6icgs2</td> <td>1088814 4.0 orxd63 CMV: Covid Vaccination is NOT going to end covid TruerInfo h6kuc06 >Malta has vaccinated the entire eligible popu 1088815 4.0 orxd63 CMV: Covid Vaccination is NOT going to end covid TruerInfo h6kuc06 >Int it better to not have ANY covid in your 1088819 4.0 orxd63 CMV: Covid Vaccination is NOT going to end covid TruerInfo h6kus9s >Malta has vaccinated the entire eligible popu 1088819 4.0 orxd63 CMV: Covid Vaccination is NOT going to end covid TruerInfo h6kus9s >Malta has vaccinated the entire eligible popu 1088819 4.0 orxd63 CMV: Covid Vaccination is NOT going to end covid TruerInfo h6kus9s >Malta has vaccinated the entire eligible popu 1088413 3.0 gattcf cmv: The concept of cultural appropriation is Jamo-duroo fp2ndod I think there needs to be a distinction betwee 1240015 0.0 7hqdz Female' movement should r FluftyRadcliffe dqtav9s "Cur post is extremely insensitive and i advis 115349 0.0 anivtz CMV: The controversy surrounding Liam Neeson's OddlySpecificReferen Neeson's eftod8v Reasons wh</td> <td>1088814 4.0 orxd63 CMV: Covid Vaccination is NOT going to end covid TruerInfo h6kuc06 >Malta has vaccinated the entire eligible popu Runiat 1088815 4.0 orxd63 CMV: Covid Vaccination is NOT going to end covid TruerInfo h6kuc06 >Malta has vaccinated the entire eligible popu edwardlleandre 1088819 4.0 orxd63 CMV: Covid Vaccination is NOT going to end covid TruerInfo h6kusgs >Malta has vaccinated the entire eligible popu barbodelli 762913 3.0 gattcf cmv: The concept of cultural appropriation is Jamo-duroo fp2ndod Ithink there needs to be a distinction be a distinction moon_truthr 1085413 3.0 ou1oho [deleted by user] 0 h6z8r01 Norever is much more likely to dublea 1240015 0.0 7hqddz Female' movement should athletes shouldn't competion FluffyRadcliffe dqtav9s "Qui post is extremely insensitive and i advis Misspelt 115349 0.0 anivtz CMV: The controversy surrounding Liam Neeson's OddySpecificReferen Neeson's efwofsv Reasons why people may be mad: Initi Helpfulcioning is tr 919286 0.0<td>1088814 4.0 oxd63 CMV: Covid Vaccination is NOT going to end covid Truerinfo h6kuc06 >Maita has vaccinated the entire eligible popu Runiat 13.0 1088815 4.0 orxd63 CMV: Covid Vaccination is NOT going to end covid Truerinfo h6kuc06 >Int it bitter to not have ANY covid in your edwardleandre 9.0 1088819 4.0 orxd63 CMV: Covid Vaccination is NOT going to end covid Truerinfo h6ku8gs >Maita has vaccinated the entire eligible popu edwardleandre 9.0 1088819 4.0 orxd63 CMV: Covid Vaccination is NOT going to end covid Truerinfo h6ku8gs >Maita has vaccinated the entire eligible popu barbodelli 5.0 762913 3.0 gattef cmv: The concept of cultural appropriation is Jamo-duroo fp2ndod Ithink there needs to be a distinction be a distinction betwee moon_truthr 41.0 1085413 3.0 ou1oho [deleted by user] 0 h6z8r01 however is much more likely to moon_truthr 41.0 1240015 0.0 7hqddz Female' movement should r FluftyRadcliffe dqta9s "equality'thought Misspet <td< td=""></td<></td></td>	10888144.0orxd63CMV: Covid Vaccination is NOT going to end covidTruerInfoh6kuc0610888154.0orxd63CMV: Covid Vaccination is NOT going to end covidTruerInfoh6kvaje10888154.0orxd63CMV: Covid Vaccination is NOT going to end covidTruerInfoh6kvaje10888194.0orxd63CMV: Covid Vaccination is NOT going to end covidTruerInfoh6kvage10888194.0orxd63CMV: Covid Vaccination is NOT going to end covidTruerInfoh6kvage7629133.0gattefcmv: The concept of cultural appropriation isJamo-duroofp2ndod10854133.0ouloho[deleted by user]0h6z8r0112400150.07hqddzFemale' movement should rFluffyRadcliffedqtav9s1153490.0b0iwy1CMV: The controversy surrounding Liam Neeson'sOddlySpecificRefereneftvkhm9192850.0anivfzCMV: The controversy surrounding Liam Neeson'sOddlySpecificRefereneftod8v10887510.0orzj99CMV: The US should not re-impose0h6icgs2	1088814 4.0 orxd63 CMV: Covid Vaccination is NOT going to end covid TruerInfo h6kuc06 >Malta has vaccinated the entire eligible popu 1088815 4.0 orxd63 CMV: Covid Vaccination is NOT going to end covid TruerInfo h6kuc06 >Int it better to not have ANY covid in your 1088819 4.0 orxd63 CMV: Covid Vaccination is NOT going to end covid TruerInfo h6kus9s >Malta has vaccinated the entire eligible popu 1088819 4.0 orxd63 CMV: Covid Vaccination is NOT going to end covid TruerInfo h6kus9s >Malta has vaccinated the entire eligible popu 1088819 4.0 orxd63 CMV: Covid Vaccination is NOT going to end covid TruerInfo h6kus9s >Malta has vaccinated the entire eligible popu 1088413 3.0 gattcf cmv: The concept of cultural appropriation is Jamo-duroo fp2ndod I think there needs to be a distinction betwee 1240015 0.0 7hqdz Female' movement should r FluftyRadcliffe dqtav9s "Cur post is extremely insensitive and i advis 115349 0.0 anivtz CMV: The controversy surrounding Liam Neeson's OddlySpecificReferen Neeson's eftod8v Reasons wh	1088814 4.0 orxd63 CMV: Covid Vaccination is NOT going to end covid TruerInfo h6kuc06 >Malta has vaccinated the entire eligible popu Runiat 1088815 4.0 orxd63 CMV: Covid Vaccination is NOT going to end covid TruerInfo h6kuc06 >Malta has vaccinated the entire eligible popu edwardlleandre 1088819 4.0 orxd63 CMV: Covid Vaccination is NOT going to end covid TruerInfo h6kusgs >Malta has vaccinated the entire eligible popu barbodelli 762913 3.0 gattcf cmv: The concept of cultural appropriation is Jamo-duroo fp2ndod Ithink there needs to be a distinction be a distinction moon_truthr 1085413 3.0 ou1oho [deleted by user] 0 h6z8r01 Norever is much more likely to dublea 1240015 0.0 7hqddz Female' movement should athletes shouldn't competion FluffyRadcliffe dqtav9s "Qui post is extremely insensitive and i advis Misspelt 115349 0.0 anivtz CMV: The controversy surrounding Liam Neeson's OddySpecificReferen Neeson's efwofsv Reasons why people may be mad: Initi Helpfulcioning is tr 919286 0.0 <td>1088814 4.0 oxd63 CMV: Covid Vaccination is NOT going to end covid Truerinfo h6kuc06 >Maita has vaccinated the entire eligible popu Runiat 13.0 1088815 4.0 orxd63 CMV: Covid Vaccination is NOT going to end covid Truerinfo h6kuc06 >Int it bitter to not have ANY covid in your edwardleandre 9.0 1088819 4.0 orxd63 CMV: Covid Vaccination is NOT going to end covid Truerinfo h6ku8gs >Maita has vaccinated the entire eligible popu edwardleandre 9.0 1088819 4.0 orxd63 CMV: Covid Vaccination is NOT going to end covid Truerinfo h6ku8gs >Maita has vaccinated the entire eligible popu barbodelli 5.0 762913 3.0 gattef cmv: The concept of cultural appropriation is Jamo-duroo fp2ndod Ithink there needs to be a distinction be a distinction betwee moon_truthr 41.0 1085413 3.0 ou1oho [deleted by user] 0 h6z8r01 however is much more likely to moon_truthr 41.0 1240015 0.0 7hqddz Female' movement should r FluftyRadcliffe dqta9s "equality'thought Misspet <td< td=""></td<></td>	1088814 4.0 oxd63 CMV: Covid Vaccination is NOT going to end covid Truerinfo h6kuc06 >Maita has vaccinated the entire eligible popu Runiat 13.0 1088815 4.0 orxd63 CMV: Covid Vaccination is NOT going to end covid Truerinfo h6kuc06 >Int it bitter to not have ANY covid in your edwardleandre 9.0 1088819 4.0 orxd63 CMV: Covid Vaccination is NOT going to end covid Truerinfo h6ku8gs >Maita has vaccinated the entire eligible popu edwardleandre 9.0 1088819 4.0 orxd63 CMV: Covid Vaccination is NOT going to end covid Truerinfo h6ku8gs >Maita has vaccinated the entire eligible popu barbodelli 5.0 762913 3.0 gattef cmv: The concept of cultural appropriation is Jamo-duroo fp2ndod Ithink there needs to be a distinction be a distinction betwee moon_truthr 41.0 1085413 3.0 ou1oho [deleted by user] 0 h6z8r01 however is much more likely to moon_truthr 41.0 1240015 0.0 7hqddz Female' movement should r FluftyRadcliffe dqta9s "equality'thought Misspet <td< td=""></td<>

Figure 5: A screenshot of scrapped data along with meta data

Data Cleaning:

After scrapping, the first cleaning step was to remove all the rows which have upvotes less than 1. These comments with negative upvotes were mostly derogatory or misinformed and does not contribute to productive arguments. As a result, we end up with 1313022 data points after this step.

Next, there were comments and posts that had multiple keywords such as [deleted], "your submission has been deleted" and [removed]. So, we removed all the data points which had any of these keywords in the text and very short comments.

Next, we used the NLTK library along with regular expressions to clean the text. Some comments had external links, references, and pictures. We removed all this non-textual information from both statements and comments. After all the cleaning, we had 1208959 data points.

Statement : CMV: Covid Vaccination is NOT going to end covid Argument : >Malta has vaccinated the entire eligible population at 73.5%. However they had a recent rise. If cova x was so effective, this should not have happened.

So there is a R naught which states how many people a typical infected person will infect. The higher the R naught the more people need to be vaccinated in order to achieve herd immunity. The fact that places with high vaccine ra tes still experience spikes only tells us that they have not achieved herd immunity. Nothing more nothing less. Th e fact that Coronavirus has an unusually high R naught is hardly new information.

>Covid is highly mutating. Vaccines against highly mutating viruses does not work. 2019 Flu vaccine was only 39% e ffective for example.

So far the vaccines are holding firm against variants.

>The ONLY solution is to drive covid to ZERO like New Zealand has. These are seen as draconian measures that most countries such as USA would not adopt. Greenzone strategy proposed by https://zerocovid.us/ is probably the only real way to eliminate covid.

>Isn't it better to not have ANY covid in your system, neither from the virus nor the covax?

One important aspect of the Covid vaccine is that IT DOES NOT PREVENT INFECTIONS. You can still get sick and test positive for coronavirus even if you are completely vaccinated. What it does do is make the likely hood of you hav ing complications due to Coronavirus significantly less likely. This is a big problem with the common understandin g of what vaccines actually do. They prepare your system to fight the virus. They don't necessarily make it imposs ible for the virus to infect you.

Figure 6: Example of Unprocessed Statement and Argument

Statement Start Token

Statement End Token

Input : <S_START> Covid Vaccination is NOT going to end covid <S_END <A_START> >Malta has vaccinated the entire eligible population at 73.5%. However they had a recent rise. If covax was so effective, this should not have happ ened.

So there is a R naught which states how many people a typical infected person will infect. The higher the R naught the more people need to be vaccinated in order to achieve herd immunity. The fact that places with high vaccine ra tes still experience spikes only tells us that they have not achieved herd immunity. Nothing more nothing less. Th e fact that Coronavirus has an unusually high R naught is hardly new information.

>Covid is highly mutating. Vaccines against highly mutating viruses does not work. 2019 Flu vaccine was only 39% e ffective for example.

So far the vaccines are holding firm against variants.

>The ONLY solution is to drive covid to ZERO like New Zealand has. These are seen as draconian measures that most countries such as USA would not adopt. Greenzone strategy proposed by [is probably the only real way to eliminate covid.

>Isn't it better to not have ANY covid in your system, neither from the virus nor the covax?

One important aspect of the Covid vaccine is that IT DOES NOT PREVENT INFECTIONS. You can still get sick and test positive for coronavirus even if you are completely vaccinated. What it does do is make the likely hood of you hav ing complications due to Coronavirus significantly less likely. This is a big problem with the common understandin g of what vaccines actually do. They prepare your system to fight the virus. They don't necessarily make it imposs ible for the virus to infect you. <A_END>

Argument End Token -

Argument Start Token -

Figure 7: A cleaned and processed Input and Label for the above example.

Lastly, the argument and statement are combined using start and end tokens which will be our final input to the model.

Data Labelling:

We gave deltas more importance because they drive the change. We sorted all the data points based on the number of deltas they received. Then the result was again sorted on the number of upvotes. This ordered data is divided equally into 4 quality groups and each group is labeled from 3 - 0, (3 to the best).

Statement	Argument	Delta	UpVote	Label	Meaning
XYZ	ABC	>=1	-	3	Changed an opinion
XYZ	DEF	0	>=30	2	Excellent argument
XYZ	GHI	0	>15 & <=10	1	Adequate
XYZ	JKL	0	==1	0	Inadequate

 Table 1: Data Labeling Examples.

Architecture:

The Language model that we are going for this to use is GPT-2. The main reason for our choice is their remarkable success in complex language tasks. A pre-trained GPT-2 Large model will be used which consists of 48 layers and 25 heads. This version of GPT-2 has more than 1.5 billion parameters which are 5-10 times more as compared to our baseline BERT models.

A classification head will be used for output and will be mapped into 4 different classes. A SoftMax will be applied to these classes to get probabilities and the highest probability class will be our final prediction. A Cross Entropy loss will be used on these probabilities during training which will be minimized by Adam Optimizer.

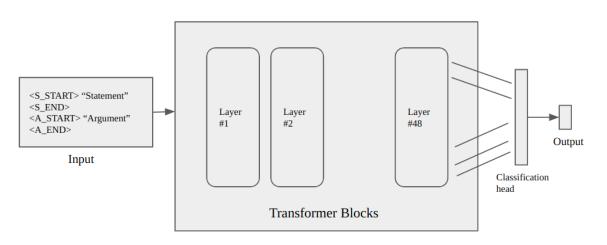


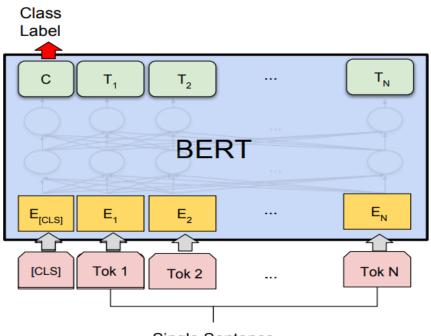
Figure 8: End-to-End Architecture Diagram

Hyperparameters

- Learning Rate 1e-6 with linear decay
- Weight Decay 0.1
- Dropout 0.5
- Batch Size 256
- Training Size 113363
- Number of Epochs 10 (Early stopping at 9 to prevent overfitting)
- Optimizer AdamW

Baseline Model:

We employ a BERT-based [9] model as our baseline model. BERT's model architecture is a multilayer bidirectional Transformer encoder based on the original implementation described in [10] and released in the tensor2tensor library. We report results on one model size: BERTBASE (L=12, H=768, A=12, Total Parameters=110M). We intend to use pre-trained BERT and make use of the classification head to fine-tune it.



Single Sentence

Figure 9: Bert Architecture Diagram

Hyperparameters:

- Learning Rate 1e-6
- Weight Decay 0.05
- Dropout 0.5
- Batch Size 512
- Training size 113363
- Number of Epochs 30
- Optimizer AdamW

Quantitative Results:

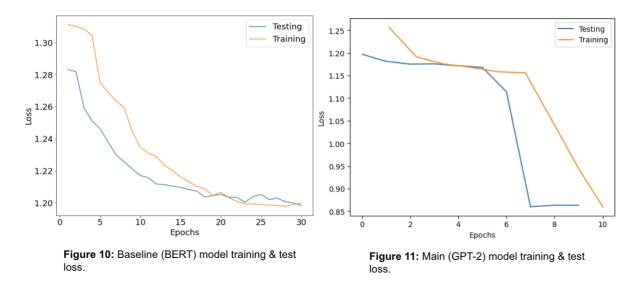
The metrics that we use to measure results are as follows:

- Accuracy The fraction of predictions a model got right.
- **Precision** Quantifies the number of positive class predictions that belong to the positive class.
- **Recall** Quantifies the number of positive class predictions made of all positive examples in the dataset.
- **F1** Harmonic mean of precision and recall.
- **Training Loss** Metric used to assess how a model fits the training data.
- Test Loss Metric used to assess the performance of a model on the Test set.

Model	Train Size	Test Size	Accuracy	Precision	Recall	F1	Training Loss	Test Loss
Baseline (BERT)	113363	28341	0.44	0.44	0.43	0.44	1.19	1.19
Main (GPT-2)	113363	28341	0.63	0.63	0.66	0.65	0.86	0.86

Table 2: Quantitative Results for both the models.

The results rightly indicate that the GPT-2 model outperforms the BERT language model with fewer epochs on the same training data and validation data.



The loss curves do also represent the same phenomena uncovered with the various classification metrics that the GPT-2 model performs much better than the BERT base language model. The GPT-2 model shows convergence right after some epochs which is not the case with BERT where the learning slows down after a certain number of epochs.

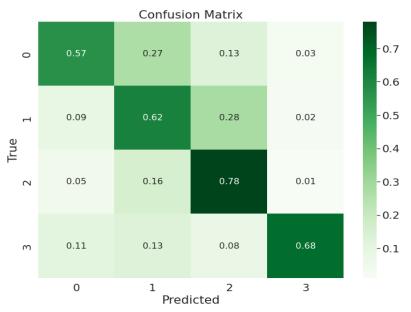


Figure 12: Confusion Matrix for main model (GPT - 2).

The confusion matrix confirms that the model is classifying well on the different balances of the dataset and not on any one class. There are a few instances where the model is predicting a label into the neighboring class. For example, the model is predicting label 0 to be in label 1, 27% of the time, and label 1 to be in label 2, 28% of the time. This is quite acceptable due to the nature of our problem and does not possess an immediate concern to be reconsidered.

Qualitative Results:

statement	output	
Not believing in science makes sense.	CHANGER score: 0.53	
argument		
Medical scientists don't sell or market drugs. They develop they use scientific method to create drugs that work exact not morally good or bad thing. It's just a tool. Tool that can drugs or other "evil" things. But it is a tool that works and t hungry bosses) that make morally questionable choices	ly as they were ordered. Scientific method works but it is be used to make weapons of mass destruction, addictive	
Clear	Submit	

Figure 13: Example of a statement that is classified as "CHANGER".

statement	output
Not believing in science makes sense.	INADAQUATE score: 0.916
argument	
Science doesn't care if you believe in it or not. It's there and it's happening no matter what you be	elieve.
Clear	

Figure 14: Example of a statement that is classified as "INADAQUATE".

In Figure 13, the argument is very well explained with proper example from the medical science. The model has good confidence in predicting it as **"changed a view"** argument. On the other hand, the second example (Figure 14) is bad. The model predicts it is as **"Inadequate"** with very high confidence.

statement	output	
People who sneeze extremely loudly do it on purpose.	EXCELLENT score: 0.647	
argument		
It's more due to the size of your diaphragm which is below y out your body. People with diaphragm that move more spar [Diaphragmatic breathing](you are better able to control ho knowledge and has to be taught it usually not in the user's o loudly can cause extra pain to the user and it causes more v		
Clear	Submit	

Figure 15: Example of a statement that is classified as "EXCELLENT".

statement	output	
People who sneeze extremely loudly do it on purpose.	EXCELLENT score: 0.687	
argument		
It's more due to the size of your diaphragm which is below your body. If you have diaphragm that move more span [Diaphragmatic breathing](you are better able to control h knowledge and has to be taught it usually not in the user's if they have some condition and it can be more annoying to		
Clear	Submit	

Figure 16: Rephrased example of the statement in Figure 15 that is classified as "EXCELLENT".

statement		output
People who sneeze extremely loudly do it on purpose.	INADAQUATE score: 0.409	
argument		
Unfortunately, there are times when I am unable to politely to my ears and sinuses from when I was a kid, and if it's a pa one of my ears or cause a nosebleed. It's not fun, and in fact however, try my best to cover my mouth in the crook of my- body finds it necessary, which usually involves a loud noise don't have time to react accordingly. Can't really give any pr info as to why at least some people can't sneeze quietly or in	rticularly big sneeze, if I hold it in I can easily blow out , even when I do let them out it can be quite painful. I do, elbow, but sometimes I just can help but let it out as my Especially if I'm hit with a sudden need to sneeze and oof about this, but I hope that this can at least give some	
Clear	Submit	

Figure 17: Example of a statement that is classified as "INADAQUATE".

statement	ou	itput					
People who sneeze extremely loudly do it on purpose.		INADAQUATE score: 0.938					
argument	argument						
I don't know if sneeze volume has to do with size but with so noticed my sneezes got a lot louder. I blame it on my lungs g							
Clear							

Figure 18: Example of a statement that is classified as "INADAQUATE".

In Figure 15, the example is well framed and gives proper explanation using science. The model is able to predict it properly with good confidence. In Figure 16, we did rephrasing of some sentences from the first example. But despite of that, model is very accurately able to identify it as an excellent argument.

Figure 17 is an example of a wrong prediction as its actual label is adequate and the model is predicting it as inadequate. This result is still not too bad because confidence score for this prediction is low.

In Figure 18, the argument is very vague and convincing. The model is able to predict it as inadequate with very high confidence.

Discussions and Learnings:

The baseline model as intended does not match the capabilities offered by the main model even though the main model ran for fewer epochs (~10) than the baseline model (~30). The results produced by the main model are remarkable in terms of the fewer number of epochs and small training size. Our prediction right now is that the model would not adapt to the entire dataset due to the biased way of the data labeling process. The results are similar with respect to the balance of the dataset where each class is well represented and well predicted by the main model.

The results are matching our expectations and did not hit us with any surprises. With enough computational power and time, we would very much like to run the models on more training data and epochs to maximize the accuracy as high as possible. Our learnings are very clear with the fact in mind that the data handling process is much more sensitive, and important than any other part of the project. With correct data, an individual can very easily tune any model to adapt to their dataset and with the wrong data, even the strongest model would fail to perform. Another important takeaway for us was to understand the importance of overfitting prevention techniques and their usage in huge language models.

Individual Contributions:

	Task	Contributions
1	More Thorough research of existing literature	Both
2	Data Scraping	Umarpreet
3	Data Cleaning & Preprocessing	Both
4	Baseline Model Implementation & Training	Dhiraj
5	Progress report	Both
6	GPT & Training Loop implementation	Both
7	Fine Tuning	Both
8	Testing on different examples	Umarpreet
9	UI Setup	Dhiraj
10	Deployment on Server	Umarpreet
11	User Testing	Dhiraj
12	Final Presentation	Both
13	Final Report	Both

References:

[1] Applying Transformers and NLP Computational Techniques to America in One Room Stanford CS224N Custom Project Haroun Ahmed Department of Computer Science Stanford University haroun@stanford.edu

[2] Rinott, Ruty, Lena Dankin, Carlos Alzate Perez, et al. (2015). "Show Me Your Evidence - an Automatic Method for Context Dependent Evidence Detection". In: Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing. Lisbon, Portugal: Association for Computational Linguistics, pp. 440–450.

[3] Stab, Christian Matthias Edwin, Johannes Daxenberger, Chris Stahlhut, et al. (2018). " ArgumenText: Searching for Arguments in Heterogeneous Sources". In: Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations. New Orleans, Louisiana: Association for Computational Linguistics, pp. 21–25.

[4] Ajjour, Yamen, Wei-Fan Chen, Johannes Kiesel, Henning Wachsmuth, and Benno Stein (2017). "Unit segmentation of argumentative texts". In: Proceedings of the 4th Workshop on Argument Mining, pp. 118–128.

[5] Stab, Christian Matthias Edwin (2017). "Argumentative writing support by means of natural language processing". PhD thesis. Technische Universität Darmstadt

[6] Wachsmuth, Henning, et al. "Computational argumentation quality assessment in natural language." Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers. 2017.

[7] Lauscher, Anne, et al. "Rhetoric, logic, and dialectic: Advancing theory-based argument quality assessment in natural language processing." arXiv preprint arXiv:2006.00843 (2020).

[8] Swanson, Reid, et al. "Argument Mining: Extracting Arguments from Online Dialogue." ACL Anthology, Association for Computational Linguistics, 2015, <u>www.aclweb.org/anthology/W15-4631</u>.

[9] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pretraining of deep bidirectional transformers for language understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), pages 4171–4186, Minneapolis, Minnesota, June. Association for Computational Linguistics.

[10] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In Proceedings of NeurIPS, pages 5998–6008.