

Evaluation of Voting Rules on LLM Generated Ballots

Predicting the Popular Vote Using Artificial Voting on
Generated Statements

Ishan Hitesh Mistry

OR/CSC-591 Computational Social Choice

Instructor : Adolfo Escobedo

Motivation

- **Structured Data** (e.g., preference matrix) is key
- **Scalability**: There's only so much you can tell from the data
- **More than meets the eye** : Traditional “candidates” may not fully capture people's opinions.
- **Can we do better?**: Including non-traditional candidates could better represent individual preferences.
- **LLMs superpower** : Unstructured data to Structured data
- **Meeting in the middle?** : LLMs + CSC

Problem Statement

Objective: Develop a model to simulate and predict the popular vote outcome based on artificial voting mechanisms.

Context: Inspired by computational social choice principles to study voting rules and their applications.

Problem Statement:

- How accurately can we predict the popular vote using generated statements as a proxy for voter preferences?
- Can artificial voting methods emulate real-world voting decision making processes?
- Is it possible to **treat comments as votes** to reflect people's preferences or stances?

Concepts from the Paper

- Justified Representation
- Extended Justified Representation

Additional concepts

- Validity of Statements
- Agentic Voting mechanisms

The Playground



r/AmltheAsshole · 8 hr. ago
throwaway3747434



AITA for leaving in the middle of Thanksgiving dinner because of pumpkin pie?

My (32f) Mother (60f) hosts Thanksgiving dinner at her house every year. It's a small event, with my parents, me, my brothers family and my SILs family attending. We avoid family quarrels by implementing a strict "no politics" rule and trying our best to be civil. I should probably mention that we are not a particularly close-knit family. We rarely see each other beyond these events since my Brother lives in South Africa and I travel a lot due to my work. Thanksgiving is important to my mom since it's one of the rare times we're all together.



jphistory · 3d ago ·

Asshole Enthusiast [5]

NTA. Your mom is the woorst, like seriously the worst. God forbid anyone eat and then compliment your cheesecake! You were not crying for some reason, you were crying for valid reasons. You were not behaving irrationally, you were reacting with perfect rationality given your circumstances. Maybe take off the rest of the holidays and give yourself some grace? Or spend them with a friend or family member that doesn't treat you like an inconvenience? From one black sheep to another: you deserve way better.



6K



Reply



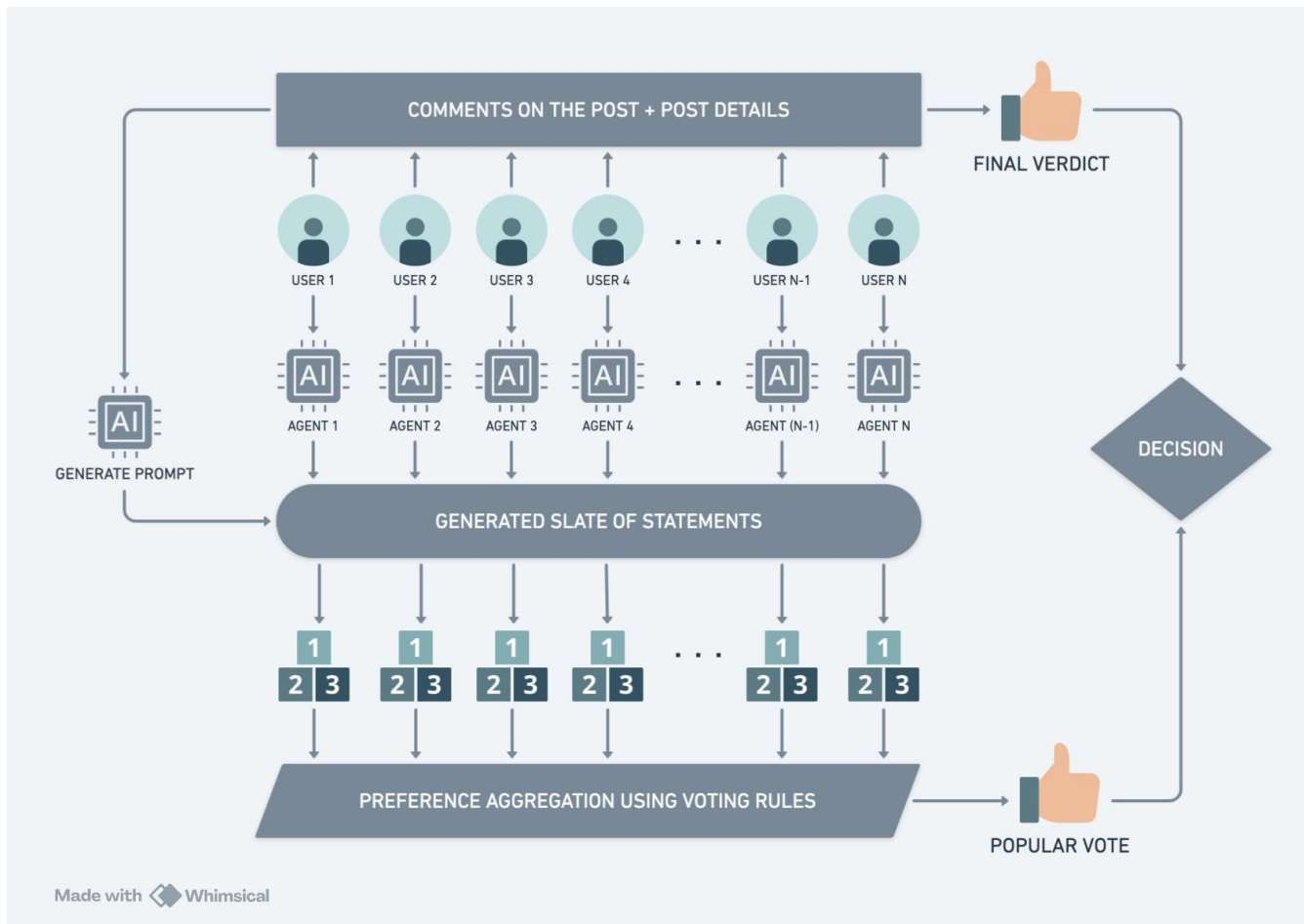
1



Share



Workflow



PROMPT for Slate Generation

"""

TASK :

Summarize the comments as CONCISELY as possible—no need for formal language or nice phrasing.

Ignore all the verdicts that are present in the comments. There should not be any verdicts in the slate of comments.

DATA : Comments = {comments}

INSTRUCTIONS :

1. Determine the OPTIMAL(not too many and not too few) number of UNIQUE and SUCCINCT SUMMARIZED viewpoints you can find and give that many statements.
2. OUTPUT should in a JSON format, where the keys are strings of numbers and values are the summarized statements. For example = dict('1' : 'Statement_1','2' : 'Statement_2')
3. ONLY return the output as a valid JSON(no preamble)
4. Usage of similar ideas and words in your summaries is a good indicator of duplicate ideas and such ideas should be avoided. Beware not every comment has a unique viewpoint.
5. DO NOT use your knowledge about the world; stick to what the participants said. Do not put quotes around your response.

"""

EXAMPLE : POST

Post Title

AITA for snapping at my boyfriend after he targeted me in a game?

Post Description

"I (F) was playing games with my boyfriend, and his friends. We played a game where the goal is to make a controversial prompt that the players of the game will be split on. The more split it is, the more points. Some more background about me, this is important. I was adopted by other relatives and have lived with them since I was 1. My bio mother is dead and has been for my entire life. I don't remember anything about her. I don't fully remember what the prompt was, it was something about getting money by having to delete the only picture of something you have off your phone. My boyfriend filled in the blank and wrote that the only picture you'd have to delete would be of your birth mom. I immediately said I felt targeted, and he told me I was. I was not happy and I asked him why he did that in our DMs. He told me he did it because he wanted to win and wanted the split. He said sorry, but, he was insistent that he thought it wouldn't be a big deal and I wouldn't be so hurt. He told me he felt awful afterwards and like I had gotten way too upset. AITA?"

Comment

'YTA, honestly. It sounds like you're overreacting a bit. It's just a game, and it was meant to be controversial. He probably didn't think it would hit so close to home. If you can't handle a little banter in a game, maybe you shouldn't play. Plus, you said yourself you don't remember your bio mom, so it seems like you're making a bigger deal out of it than necessary. Just let it go; it's not like he did it to hurt you on purpose.'

EXAMPLE : Slate of Statements

{1: 'Boyfriend crossed a line by using a personal and sensitive topic to win a game.',

2: 'He should have known better than to use a dead mom joke, even in a game meant to be offensive.',

3: "The game's goal is to be controversial, but that doesn't mean it's okay to hurt someone's feelings.",

4: "Boyfriend's apology was sincere, and OP was overreacting.",

5: "OP should set boundaries and communicate what is and isn't okay to joke about.",

6: "Boyfriend's actions show he's willing to hurt OP for personal gain, which is a red flag in the relationship.",

7: "OP should consider whether boyfriend's behavior is a consistent pattern and if it's a sign of deeper issues.",

8: "The game is meant to be uncomfortable, but that doesn't mean OP should just accept being hurt.",

9: "Boyfriend's actions were not just about winning the game, but also about targeting OP specifically.",

10: "OP's feelings are valid, and boyfriend should have been more understanding and empathetic."}

PROMPT : Generate Preferences

""Given, the description and the comment as your stance, think like this person and try to order ALL the statements from the slate in an order that goes from MOST LIKED to LEAST LIKED.

A random example would be as follows :, if the example post description == "AITA for getting a rescue cat and giving it away later on because it was not my cup of tea?".

Your example comment to this == "YTA, it was your duty to check and educate yourself about the responsibilities that come with adopting a pet cat and how sometimes they can be more difficult because they have been through abuse".

The example slate has 5 statements for example,

1 : You are not the asshole here because you were doing a good deed but did not prepare for it.

2 : Well if it did not work out then it did not, no point in blaming people now.

3 : You are in the wrong here because you not only gave into your impulses but also were irresponsible with the duties.

4 : Cats look easy to maintain is not a good enough reason to get a stray animal as your pet. You are not prepared to get one

5 : You need to first interact and see if you are a cat person at all. Reach out to people or friends who have pet cats.

Given all this information you would have the following order of the statements and the reasoning for the order.

ORDER = [3,1,4,5,2]

REASONING = Based on the comment you disagree and the poster is in the wrong. Logically, you agree with 1,3 but then also agree with 4,5 a little because that would mean being responsible. You do not agree with 2 so it goes last.

Now that the example is done. Here are the inputs for the task that you have to perform. You are a commenter on a post. The post is from r/AITA the with the following description:

```
----  
DESCRIPTION = {description}  
----  
You have put the following comment on the post as a reflection of what you think :  
----  
COMMENT = {comment}  
----  
These are the summarized statements/opinions across all the other comments(including yours) that you have  
to use for the ordering and give reasoning for that order.  
----  
SLATE = {slate}  
----
```

RULES TO STRICTLY FOLLOW :

1. DO NOT SKIP any statements.You NEED TO RANK all of the statements.
2. ONLY respond with ONE COMPLETE LIST OF ORDERING and the REASONING(based on your comment and slate) justifying that ordering.
3. FOLLOW the format from the example. You might have a different number of statements(10 0r 15) so ADAPT accordingly.
4. Each RANK should be in the output EXACTLY ONCE.
5. Given above is JUST AN EXAMPLE for you. Do not use the same context, but try to understand how the task is to be done from this example.
6. DO NOT USE THE SAME SLATE or THE SAME ORDERING from the EXAMPLE, use the slate statements, description and comment given to you after the example for the task
7. DO NOT RETURN ANY EXTRA ITEMS except for the VALID ORDER and REASONING """)

BAD LLM! :|

LLM Response

ORDER : [3,1,4,5,2]

REASONING : Based on the situation, I can agree with 4 and 5 a little bit. This is because they are saying you should be responsible if you bring a pet into the house. To be responsible while raising one. However, I disagree also because it's a game if it's not your responsibility to raise a pet or not. I can also agree and say that this isn't the right game to have used that strategy to win. I do not agree at all that his intent was to gain points. It was his intention to make the game a bit more fun. I am a bit disappointed with your explanation of your friends, be careful next time or not taking it to seriously or just play fair. Next time, you should consider friends before using someone's personal life for a game, especially if you haven't considered your friend's feelings. Bottom line, he could have been more careful in the first place but maybe he's too competitive. It's evident that you should have known the emotional cost of using a sensitive topic during a game. You should have considered that the prompt was sensitive in nature as humor is subjective, so using such a sensitive topic for winning a point, you need to be more aware of the potential emotional hammering that sensitive topics in the next time. Also, be cautious and aware of how it might make someone feel: you could have been more specific and not just one directional."

This means you should reconsider using sensitive topics to win a point, be more aware of the emotional impact, and be more specific and not just one-sided.

Still bad :(

LLM Output

```
ORDER = [5, 4, 1, 3, 2, 3, 6, 7, 9, 1, 8, 10, 2, 6, 10, 7]
```

ORDERING AND REASONING :

1: This is MOST BEHAVING THERE'S a comment in the post that this is not even a real level of concern considering the person being upset, it is reasonable to feel concerned about a topic that deeply affects the person, but it is better to understand everyone's feelings are different and sometimes it is okay to avoid uncomfortable or controversial topics.

2: There is no way to force others to play the game. It is a game made to be a controversial game, but it seems like you've shown a picture of the person's mother during a game. Also, you should know your partner's taste for such games.

3: The game isn't just for the winner- just game for the NPC. It's necessary to have a heart before hurting someone.

4: Boyfriend crossed a line though the last one isn't that great.

5: The winner is the one who shows care for the game , trying to hurt someone that way.

6: It is easy to hurt someone the game; it's easy to play a game with someone and it's not your mom so it's not about being insensitive enough if you play such a game.

7: If it bothers here, it bothers others also.

8: If it bothers others, it bothers you too just because you had an uncomfortable or controversial topic.

9: It was not worth hurting her mom or mom. Don't make fun of someone's pain.

Tears Incoming!

LLM Response

```
[1,2,8,9,4,3,5,7,6,10.]
```

LLM Response

Here are the rankings and reasoning:

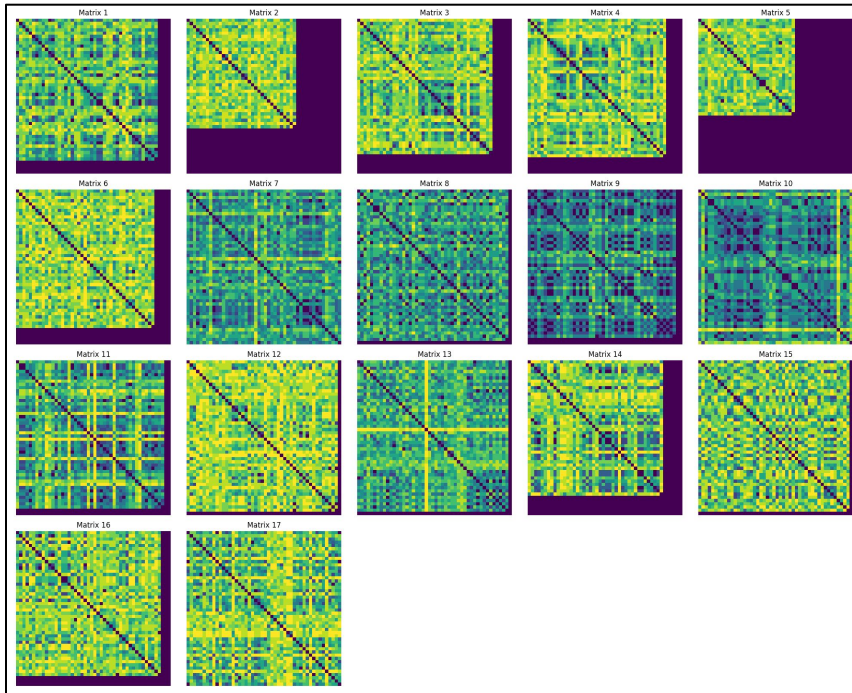
1. ****YTA**** - (Based on the example rule 2. You may not like what he did. This could be due to personal feelings/issues but don't make it a personal attack. Explaining the rule you are following) The only reason this shows up 1st is that it is the most objective and least emotional, but this is an example to follow. Therefore, it shows first and sets the tone, then follow the rest based on that tone.**

A Few Prompts
Later!

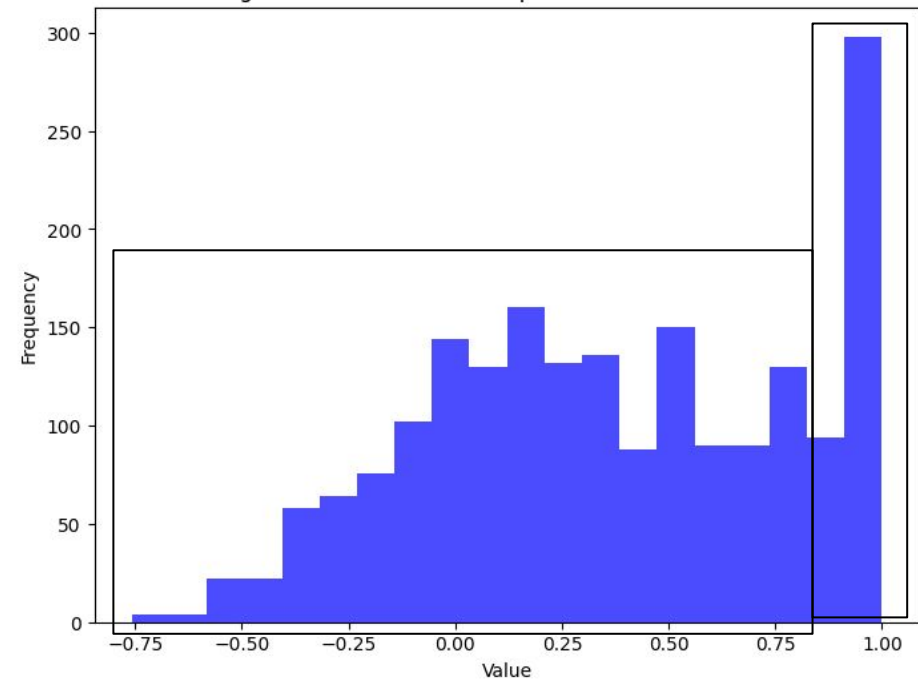
Sanity checks on the generated ballots

ORDER = [2, 3, 4, 9, 1, 10, 6, 7, 5, 8]

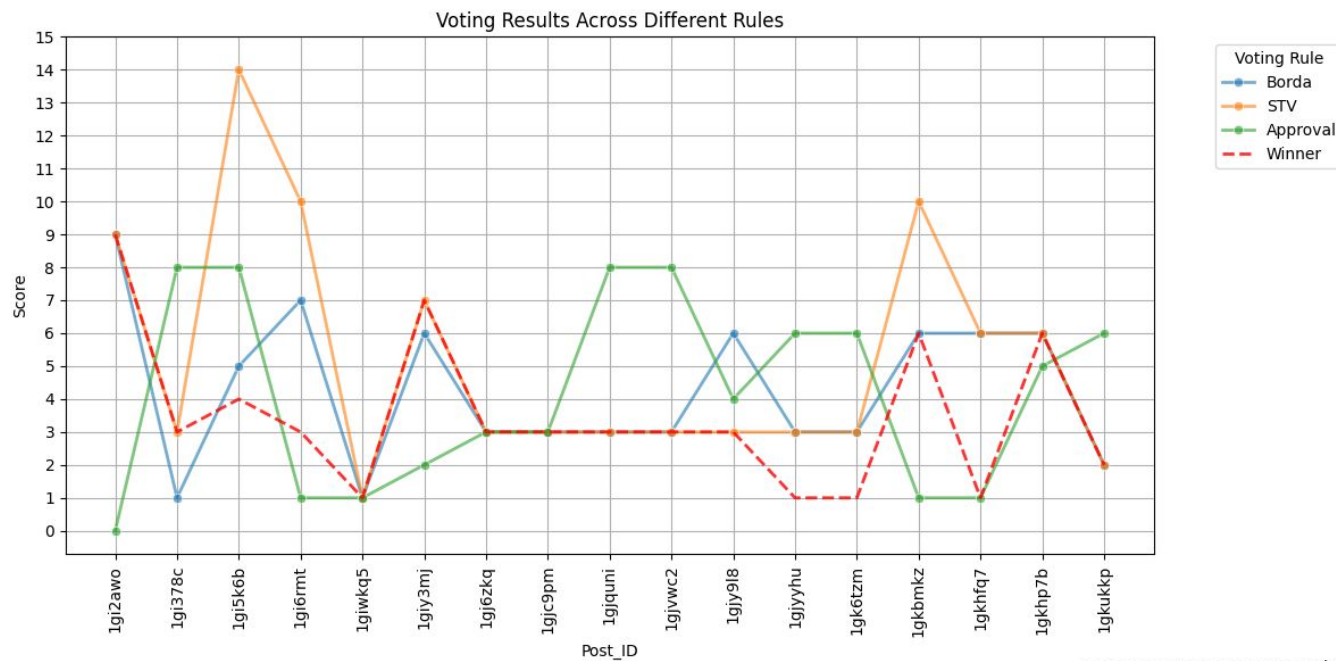
REASONING: Based on my comment, I agree that going behind someone's back to do something they explicitly said no to is not okay (2). I also think that forcing someone to accept something they don't want is violating their boundaries (3). Sneaking money into someone's wallet is not a good way to show appreciation (4), and not respecting someone's wishes can make them feel uncomfortable or even stressed (9). I agree that a simple 'thank you' or acknowledgement can be enough, without needing to force payment or compensation (10). I agree that everyone involved is trying to do the right thing, but some methods are better than others (6), and it's better to be transparent and direct when showing appreciation, rather than sneaky or manipulative (7). Instead of paying for services, consider doing something nice for the person, like taking them out for a meal (5) and making a charitable donation in someone's honor instead of forcing them to accept payment (8). I think respecting people's wishes and boundaries, even if it's a kind gesture (1) is less important than the other statements because it's important to consider the other person's feelings and boundaries.



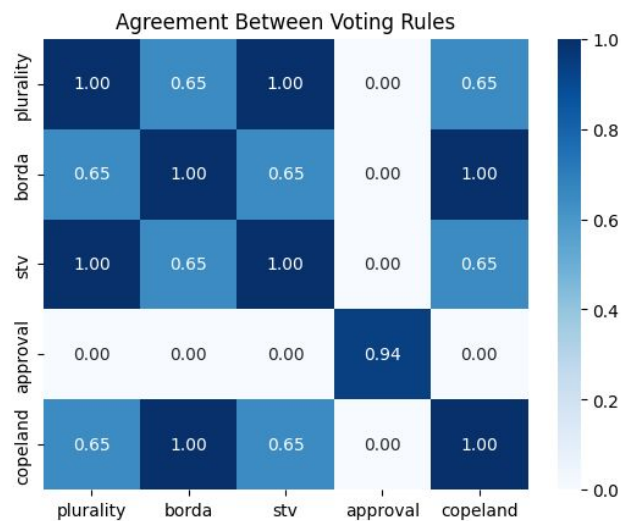
Histogram of Values from All Spearman distances Matrices



Results



Voting Rule	Accuracy
Plurality	0.6471
Borda	0.5294
STV	0.6471
Approval	0.2353
Copeland	0.5294



Future Work

- **Iterative Refinement:** Optimize configurations to match popular vote outcomes.
- **Diverse Voting Rules:** Experiment with alternative voting mechanisms.
- **Expanded Context:** Increase context windows for richer input data.
- **Fine-Tuning LLMs:** Adapt models with domain-specific examples (challenging but impactful?).
- **Advanced Prompting:** Explore innovative prompting techniques for better results.
- **Validation Goal:** Conduct surveys to validate findings with real participants.