

# UniPoll: A Unified Social Media Poll Generation Framework via Multi-Objective Optimization

## Supplementary Material

### I. CASE STUDY

We have conducted quantitative evaluations of UniPoll in the main paper. To further demonstrate its effectiveness, we present two case studies as qualitative examples, with the results illustrated in Figure 2. Specifically, we include cases from *ChatGPT* (March-23 version) and *NewBing*, which could not be included in the main comparisons due to their reliance on case-sensitive prompts for guiding generation. For reproducibility, the prompts used in these case studies are provided in Figure 1.

In Case 1, UniPoll demonstrates a high-quality poll whose generated question adeptly captures the input post’s specific concerns (about exam fairness) and well reflects that point in the question. Moreover, all three answer choices show high consistency with the poll question. On the contrary, T5 suffers from inconsistency between the question (“Do you...”) and answers (“Can or Cannot”) because the QA relations were not carefully exploited. ChatGPT, without learning from social media data, managed to generate a fine poll yet somehow looks too serious, lacking the potential to engage social media engagements. NewBing’s results look good with sufficient information, whereas it fails to generate a poll without comments even though we tried multiple times with varying prompts.

Case 2 exemplifies a scenario with implicit context-question-answer relationships and noisy social media data, where comments discuss topics unrelated to the post and provide useless contexts. Despite the challenging conditions, UniPoll successfully captures the point about “movie re-releasing” from the post and generates a poll closely related to it. In contrast, T5 again suffers from QA inconsistency problems and lacks details in the output.

In summary of the case study results, we observe UniPoll’s outputs are usually specifically related to the context’s key points, whereas others’ outputs are relatively more general. It is attributed to UniPoll’s efforts to strengthen learning in social media context-question-answer relations. For this reason, its output polls are also more engaging, which is helpful in drawing other users’ attention to participating in discussions.

**[Case 1]:** 要求: 根据以下微博内容和微博评论, 生成一个简短的投票标题和几个选项, 要求选项的字数不超过10个字, 标题字数不超过20个字。微博内容: 线上复试是否能保障公平高考延期惹的祸考研线上复试那还能保证公平吗。微博评论: 这个世界上本来就没有绝对的公平你可以说一个倒数第一考了第一但考上了他也还是啥都不会也可以说他利用一切机会达到目的反正结果就是人家考的好你还找不出来证据线上考试平时考倒数的人进了年级前十好平时考试有水分线上之后那不就是在在水里考。

**[Case 1 w.o. comments]:** 要求: 根据以下微博内容, 生成一个简短的投票标题和几个选项, 要求选项的字数不超过10个字, 标题字数不超过20个字。微博内容: 线上复试是否能保障公平高考延期惹的祸考研线上复试那还能保证公平吗。

**[Case 2]:** 要求: 根据以下微博内容和微博评论, 生成一个简短的投票标题和几个选项, 要求选项的字数不超过10个字, 标题字数不超过20个字。微博内容: 哪吒大鱼海棠重映动画电影哪吒之魔童降世大鱼海棠以及雷佳音佟丽娅主演的超时空同居确定将重映据最新数据显示DIGIT月DIGIT日全国复工影院DIGIT家复工率DIGIT单日票房DIGIT万元。微博评论: 我在人间贩卖黄昏只为收集世间温柔去见你谢谢你的分享来看看你我的微博随时恭候你的到来一个今天胜过两个明天一个评论胜过两个今天中午好我来啦支持不变微博在于互动所以我来了眼泪不是答案拼搏才是选择只有回不了过去没有到不了的明天看电影可以找我啊绝对精彩你要热门了老铁没有人能一路单纯到底但别忘了最初的自己好久不见啦我自带小板凳来的。

**[Case 2 w.o. comments]:** 要求: 根据以下微博内容, 生成一个简短的投票标题和几个选项, 要求选项的字数不超过10个字, 标题字数不超过20个字。微博内容: 哪吒大鱼海棠重映动画电影哪吒之魔童降世大鱼海棠以及雷佳音佟丽娅主演的超时空同居确定将重映据最新数据显示DIGIT月DIGIT日全国复工影院DIGIT家复工率DIGIT单日票房DIGIT万元。

Fig. 1. Prompt Template for ChatGPT and NewBing.

<p><b>[Post]:</b> #线上复试是否能保障公平# 高考延期惹的祸，考研线上复试，那还能保证公平吗？ (#Can online retesting ensure fairness# The disruptions caused by the postponement of the entrance exams have resulted in postgraduate retests being conducted online. Can this method still ensure fairness?)</p> <p><b>[Comments]:</b> 这个世界上本来就没有绝对的公平。你可以说一个倒数第一考了第一，但考上了他也还是啥都不会。也可以说他会利用一切机会达到目的，反正结果就是人家考的好，你还找不出来证据。线上考试，平时考倒数的人进了年级前十。平时考试有水分，线上之后，那不就是在水里考？ (Absolute fairness doesn't exist in this world. Even if someone who ranks last comes first in their exams, it doesn't mean they possess the necessary knowledge when they enter university. They will seize every opportunity to achieve their goals, resulting in good exam performance, without any tangible evidence against them. In the case of online exams, individuals who typically rank last may end up in the top ten of their grade. Regular exams can be subject to manipulation, and when conducted online, there may be even more room for unforeseen circumstances.)</p> <hr/> <p><b>[Gold]:</b> 你觉得线上复试是否公平 (Do you think the online retest is fair)      &gt; 公平 (Fair); 不公平 (Unfair); 其他 (Others)</p> <p><b>[T5]:</b> 你支持线上复试吗 (Do you support online retesting)      &gt; 能 (Can); 不能 (Cannot)</p> <p><b>[UniPoll]:</b> 你认为线上复试公平吗 (Do you think the online retest is fair)      &gt; 公平 (Fair); 不公平 (Unfair); 评论区补充 (Please comment if you have more to say.)</p> <p><b>[ChatGPT]:</b> 线上考试公平性投票 (Online exam fairness poll)      &gt; 保障公平 (Guaranteeing fairness); 存在瑕疵 (Deficiencies exist); 无法确定 (Not sure)</p> <p><b>[NewBing]:</b> 线上复试是否能保障公平 (Can online retesting guarantee fairness)      &gt; 可以 (Yes); 不可以 (No); 不确定 (Not sure)</p> <hr/> <p><b>[T5 w.o. comments]:</b> 你觉得线上复试能保障公平吗 (Do you think online retesting guarantees fairness)      &gt; 能 (Yes); 不能 (No); 看情况 (Depends)</p> <p><b>[UniPoll w.o. comments]:</b> 你认为线上复试公平吗 (Do you think the online retest is fair)      &gt; 公平 (Fair); 不公平 (Unfair); 看情况 (Depends)</p> <p><b>[ChatGPT w.o. comments]:</b> 考研线上复试公平性投票 (Online retest fairness poll)      &gt; 能保证公平 (Fairness guaranteed); 不能保证公平 (Fairness can't be guaranteed); 取消线上复试 (Eliminate online retesting)</p> <p><b>[NewBing w.o. comments]:</b> Failed to generate      &gt; Failed to generate</p>	<p><b>[Post]:</b> #哪吒，大鱼海棠重映# 动画电影《哪吒之魔童降世》、《大鱼海棠》，以及雷佳音、佟丽娅主演的《超时空同居》确定将重映。据最新数据显示，3月24日全国复工影院495家，复工率4.36%，单日票房2.7万元。 (#Nezha, Big Fish and Begonia to be Re-released# The animated films "Nezha: Birth of the Demon Child" and "Big Fish and Begonia," along with "How Long Will I Love U," starring Lei Jiayin and Tong Liya, have been confirmed for re-release. According to the latest data, on March 24, a total of 495 cinemas nationwide resumed operations, with a resumption rate of 4.36% and a single-day box office of 27,000 yuan.)</p> <p><b>[Comments]:</b> 我在人间贩卖黄昏，只为收集世间温柔，去见你。谢谢你的分享，来看看你。我的微博，随时恭候你的到来... (I am selling Twilight on Earth, just to collect the tenderness of the world to meet you. Thank you for sharing and coming to see you. My Weibo feeds, waiting for you anytime...)</p> <hr/> <p><b>[Gold]:</b> 你最期待哪部重新上映 (Which one are you most looking forward to re-releasing)      &gt; 哪吒之魔童降世 (Nezha: Birth of the Demon Child); 大鱼海棠 (Big Fish and Begonia); 超时空同居 (How Long Will I Love U)</p> <p><b>[T5]:</b> 你会去看哪部 (Which one would you go to watch)      &gt; 期待 (Expectations); 不期待 (No expectations)</p> <p><b>[UniPoll]:</b> 你最期待哪部电影重映 (Which film are you most looking forward to re-releasing)      &gt; 哪吒之魔童降世 (Nezha: Birth of the Demon Child); 超时空同居 (How Long Will I Love U); 都期待 (Both); 都不期待 (None)</p> <p><b>[ChatGPT]:</b> 你更想看哪部电影 (Which film would you prefer to watch)      &gt; 哪吒之魔童降世 (Nezha: Birth of the Demon Child); 大鱼海棠 (Big Fish and Begonia); 超时空同居 (How Long Will I Love U)</p> <p><b>[NewBing]:</b> 哪吒大鱼海棠重映，你会选择哪部电影 (Which film would you choose in a re-release of Nezha or Big Fish and Begonia?)      &gt; 哪吒之魔童降世 (Nezha: Birth of the Demon Child); 大鱼海棠 (Big Fish and Begonia); 都想看 (Both)</p> <hr/> <p><b>[T5 w.o. comments]:</b> 你觉得哪部电影最值得一看 (Which film do you think is most worth watching)      &gt; 期待 (Expectations); 不期待 (No expectations)</p> <p><b>[UniPoll w.o. comments]:</b> 你还会去看重映的电影吗 (Do you still go to re-releases)      &gt; 会 (Yes); 不会 (No); 看情况 (Depends);</p> <p><b>[ChatGPT w.o. comments]:</b> 选择你要观看的电影吧 (Select the film you want to watch)      &gt; 哪吒之魔童降世 (Nezha: Birth of the Demon Child); 大鱼海棠 (Big Fish and Begonia); 超时空同居 (How Long Will I Love U)</p> <p><b>[NewBing w.o. comments]:</b> Failed to generate      &gt; Failed to generate</p>
--	---

(a) Case 1

(b) Case 2

Fig. 2. Case studies about online retesting for the national postgraduate entrance exam (Case 1), and the first post-COVID-19 movie to watch (Case 2). Each case is displayed in three parts. The top presents the input to the model, including the [Post] and [Comments]. In the middle, we show the model-generated polls, [T5] and [UniPoll], with the author-written gold standard, [Gold]. Besides, we include evaluations of the latest models, [ChatGPT][NewBing]. At the bottom, we show the model results without taking comments for training (e.g. [UniPoll w.o. comments]) to examine the qualitative effects of comments.