Mistral (8GB) was prompted to classify 9,994 comments into the following 16 categories (relevant/irrelevant). Each topic was categorized separately, resulting in a total of $16 \times 9,994 = 159,904$ prompts. For 4 of the topics, we hypothesized that mostly Republicans or conservative would be triggered. We aim to manually code 10% of the comments of the 9,994 to assess the accuracy of the LLM as a classifier.

	Topics triggering Cons/Reps	Topics triggering Libs/Dems
1	Anti Cons/Reps Rhetoric	Anti Libs/Reps Rhetoic
2	Pro abortion	Anti abortion
3	Anti gun	Pro gun
4	Pro immigration	Anti immigration
5	Pro Hillary Clinton	Anti Hillary Clinton
6	Anti Donald Trump	Pro Donald Trump
7	Drugs	-
8	Suggestive content	-
9	Anti American Rhetoric	-
10	Anti Christianity	-

1 Drugs

In this task, you will be shown a Facebook post and two comments on this post. One comment is from person A and the other comment is from person B.

Your job is to determine if the comment by B is related to the consumption, production, or distribution of drugs. Drugs refer to substances with psychoactive effects.

Your answer must be a single number: 0 if irrelevant, 1 if relevant.

Here are some examples with reasoning. Use the following format for your output:

Example 1: Comment by B: "She is high on drugs of stupidity!" Your reasoning: The comment uses the phrase "high on drugs" metaphorically to describe a state of stupidity, not literal drug use. However, it still references the concept of being under the influence of drugs. Your answer: 1

Example 2: Comment by B: "And who's fault is it that the cartels exist ?? Your people" Reasoning: This comment implies a discussion about drug cartels, which are organizations that produce and distribute drugs. Therefore, it is relevant to the topic of drugs. Your answer: 1

Example 3: Comment by B: "Take her AWAY" Reasoning: This comment does not make any explicit or implicit reference to drugs. Hence, the comment is irrelevant to the topic of drugs. Your answer: 0

2 Suggestive content

In this task, you will be shown a Facebook post and two comments on this post. One comment is from person A and the other comment is from person B.

Your job is to determine if the comment by B is related to suggestive content. Suggestive content refers to any material, either explicit or implicit, that alludes to sexual themes. This can include innuendos, provocative imagery, or language that may be interpreted as sexually charged or erotic.

Your answer must be a single number: 0 if irrelevant, 1 if relevant.

Here are some examples with reasoning. Use the following format for your output:

Example 1: Comment by B: "If those prostitutes have balls." Your reasoning: The comment is involves the term "prostitutes," which refers to sexual activity, and "have balls" could be seen as a crude reference to sexual anatomy. Hence, it is relevant to suggestive content. Your answer: 1

Example 2: Comment by B: "They have hunting… And dicks on their brain!" Reasoning: The term "dicks" is a sexual reference. Thus, this comment involves suggestive content. Your answer: 1

Example 3: Comment by B: "Very painful to watch." Reasoning: This comment conveys discomfort without using explicit or implicit sexual themes. Your answer: 0

3 Negative portrayal of Christianity

In this task, you will be shown a Facebook post and two comments on this post. One comment is from person A and the other comment is from person B.

Your job is to determine if the comment by B is related to the negative characterization of Christianity or Christians. This involves portraying the religion or the religious group in a critical or unfavorable light, highlighting perceived flaws, contradictions, or controversial aspects, often with the intent to diminish its credibility or moral standing.

Your answer must be a single number: 0 if irrelevant, 1 if relevant.

Here are some examples with reasoning. Use the following format for your output:

Example 1: Comment by B: "Another good 'Christian' pedophile!" Your reasoning: This comment portrays Christians unfavorably by implying that there is a connection between being a Christian and engaging in illegal and immoral activities, specifically pedophilia. Your answer: 1

Example 2: Comment by B: "Don't tell Muslim terrors because all Christian terror everyday kill 1000 children in Syria" Reasoning: This comment characterizes Christianity or Christrians negatively by suggesting that they engage in acts of terror, specifically killing children in Syria. Your answer: 1

Example 3: Comment by B: "Take her AWAY" Reasoning: This comment is a general statement without any explicit or implicit reference to Christianity or Christians. Your answer: 0

4 Negative Portrayal of America

In this task, you will be shown a Facebook post and two comments on this post. One comment is from person A and the other comment is from person B.

Your job is to determine if the comment by B is related to the negative characterization of America or Americans. This involves portraying the country or its citizens in an unfavorable light, highlighting perceived flaws, contradictions, or controversial aspects, often with the intent to diminish its credibility or moral standing. Your answer must be a single number: 0 if irrelevant, 1 if relevant.

Here are some examples with reasoning. Use the following format for your output:

Example 1: Comment by B: "You mean unlike the American killers of babies, rapist, or serial killers?" Your reasoning: The comment portrays Americans unfavorably by associating them with negative actions such as being killers of babies, rapists, or serial killers. Your answer: 1

Example 2: Comment by B: "Americans either fail to realize or completely ignore is that America is the #1 drug consumer country in the world !" Reasoning: This comment portrays Americans in an unfavorable light by suggesting that Americans either lack awareness or deliberately ignore the fact that the country is labeled as the top consumer of drugs globally. The comment also implies a negative characterization of American society, alluding to issues of substance abuse and broader societal problems. Your answer: 1

Example 3: Comment by B: "Take her AWAY" Reasoning: This comment is a statement without any explicit or implicit reference to America or Americans. Your answer: 0

5 Anti-Trump

In this task, you will be shown a Facebook post and two comments on this post. One comment is from person A and the other comment is from person B.

Your job is to determine if the comment by B is related to the negative characterization of Donald Trump, the former president of the United States. This involves portraying Trump in an unfavorable light, highlighting perceived flaws, contradictions, or controversial aspects. Your answer must be a single number: 0 if irrelevant, 1 if relevant.

Here are some examples with reasoning. Use the following format for your output:

Example 1: Comment by B: "Trump is a friggen rapist!" Your reasoning: The comment is relevant and falls within the negative characterization of Donald Trump, as it accuses him of a serious crime. Your answer: 1

Example 2: Comment by B: "Can Trump shut up for atleast one week. Every day, he is making a point, where more n more people are hating him" Reasoning: This comment is relevant because it points out that Trump's frequent statements are uncalled for and that people don't like him. Your answer: 1

Example 3: Comment by B: "So tired of this progressive rant. If you come here illegally you need to go home. PERIOD! Get them the hell out of America!" Reasoning: This comment is irrelevant because it is a statement that does not directly refer to Donald Trump. Your answer: 0

6 Pro-Trump

In this task, you will be shown a Facebook post and two comments on this post. One comment is from person A and the other comment is from person B.

Your job is to determine if the comment by B is related to the positive characterization of Donald Trump, the former president of the United States. This involves portraying Trump in an favorable light, highlighting perceived strengths, qualities, and virtues. Your answer must be a single number: 0 if irrelevant, 1 if relevant.

Here are some examples with reasoning. Use the following format for your output:

Example 1: Comment by B: "Vote for Trump! #Trump2016" Your reasoning: The comment by B is relevant because it suggests that Trump is a good presidential candidate, as indicated by the call to vote for him. Your answer: 1

Example 2: Comment by B: "Trump is just what the republican party needs, not a bunch of politians who won't stand up and not do anything but run up the debit and let all these illegals in. Everything he said was right. Trump is a smart man and he wants to make the country better" Reasoning: This comment is relevant because it portrays Trump's actions, statements, and intentions positively. Your answer: 1

Example 3: Comment by B: "If Hillary is elected president, Putin said he will wage nuclear war against the United States, but I guess that's not a big deal…" Reasoning: This comment is irrelevant because it is a statement that does not directly refer to Donald Trump. Your answer: 0

7 Anti-Clinton

In this task, you will be shown a Facebook post and two comments on this post. One comment is from person A and the other comment is from person B.

Your job is to determine if the comment by B is related to the negative characterization of Hillary Clinton, the Democratic Party's 2016 nominee for president of the United States. This involves portraying Hillary Clinton in an unfavorable light, highlighting perceived flaws, contradictions, or controversial aspects. Your answer must be a single number: 0 if irrelevant, 1 if relevant.

Here are some examples with reasoning. Use the following format for your output:

Example 1: Comment by B: "Yes WE NEED TO LOCK HILLARY AND HER FRIENDS UP!" Your reasoning: The comment is relevant because it states that Hillary Clinton and her friends should be incarcerated, suggesting illegal action on their part. Your answer: 1

Example 2: Comment by B: "Hillary is a big fat liar and she will never be president." Reasoning: The comment is relevant because it depicts Hillary Clinton negatively, accusing her of dishonesty and implying she has no shot at the presidency even if she attempts it. Your answer: 1

Example 3: Comment by B: "Trump is just what the republican party needs, not a bunch of politians who won't stand up and not do anything but run up the debit and let all these illegals in." Reasoning: This comment is irrelevant because it is a statement that does not directly refer to Hillary Clinton. Your answer: 0

8 Pro-Hillary

In this task, you will be shown a Facebook post and two comments on this post. One comment is from person A and the other comment is from person B.

Your job is to determine if the comment by B is related to the positive characterization of Hillary Clinton, the Democratic Party's 2016 nominee for president of the United States. This involves portraying Hillary Clinton in an favorable light, highlighting perceived strengths, qualities, and virtues. Your answer must be a single number: 0 if irrelevant, 1 if relevant.

Here are some examples with reasoning. Use the following format for your output:

Example 1: Comment by B: "Recount!! Go girl! Hillary for president!!" Your reasoning: The comment by B is relevant because it suggests that Hillary Clinton is a trustworthy presidential candidate, as indicated by the call to vote for her. Your answer: 1

Example 2: Comment by B: "During the debate season, Hillary Clinton will mop Donald Trump off of the floor and dump him into the sewer with the mop water because she is 20x smarter than he is." Reasoning: This comment is relevant as it presents Hillary Clinton favorably, emphasizing her intelligence and skill in debating. Your answer: 1

Example 3: Comment by B: "Trump is just what the republican party needs, not a bunch of politicians who won't stand up and not do anything but run up the debit and let all these illegals in." Reasoning: This comment is irrelevant because it is a statement that does not directly refer to Hillary Clinton. Your answer: 0

9 Anti-Immigrant

In this task, you will be shown a Facebook post and two comments on this post. One comment is from person A and the other comment is from person B.

Your job is to determine if the comment by B is related to the negative characterization of immigrants, individuals who relocate from their native country to another. This involves portraying the immigrants in an unfavorable light, highlighting perceived flaws, contradictions, or controversial aspects. Your answer must be a single number: 0 if irrelevant, 1 if relevant.

Here are some examples with reasoning. Use the following format for your output:

Example 1: Comment by B: "The immigrants steal our jobs…. even our speeches!" Your reasoning: The comment is relevant because it characterizes immigrants negatively, suggesting that they take away jobs that Americans should have. Your answer: 1

Example 2: Comment by B: "Are those ACTUAL immigrants or illegal trespassing criminals?" Reasoning: This comment is relevant because it associates immigrants as a whole with criminal activity, thereby portraying immigrants in a negative light. Your answer: 1

Example 3: Comment by B: "Take her AWAY" Reasoning: This comment is irrelevant because it does not directly refer to immigrants.

10 Pro-immigrants

In this task, you will be shown a Facebook post and two comments on this post. One comment is from person A and the other comment is from person B.

Your job is to determine if the comment by B is related to the positive characterization of immigrants, individuals who relocate from their native country to another. This involves portraying the immigrants in a favorable light, highlighting perceived strengths, qualities, and virtues. Your answer must be a single number: 0 if irrelevant, 1 if relevant.

Here are some examples with reasoning. Use the following format for your output:

Example 1: Comment by B: "Most immigrants are actually highly educated doctors, lawyers, computer engineers." Your reasoning: The comment is relevant because it portrays immigrants favorably, highlighting that they tend to be highly educated and work in respected professions. Your answer: 1

Example 2: Comment by B: "America was built on the backs of immigrants, the indigenous people were treated and still to this day treated worst than the Buffalo." Reasoning: This comment is relevant because it describes immigrants positively by suggesting that they are hard-working and contributing to the United States. Your answer: 1

Example 3: Comment by B: "Take her AWAY" Reasoning: This comment is irrelevant because it does not directly refer to immigrants.

11 Anti-abortion

In this task, you will be shown a Facebook post and two comments on this post. One comment is from person A and the other comment is from person B.

Your job is to determine if the comment by B is related to the negative characterization of abortion, the termination of a pregnancy either by choice or due to medical reasons. This involves portraying abortion in an unfavorable light, highlighting perceived flaws, contradictions, or controversial aspects. Your answer must be a single number: 0 if irrelevant, 1 if relevant.

Here are some examples with reasoning. Use the following format for your output:

Example 1: Comment by B: "All abortions are murders!" Your reasoning: The comment is relevant because it characterizes abortion negatively, suggesting that it equates to taking a human life. Your answer: 1

Example 2: Comment by B: "Every woman who wants an abortion should see this. Children are truly gifts from God!" Reasoning: This comment is relevant because it states that children are a gift from god, thereby implying a stance against abortion. Your answer: 1

Example 3: Comment by B: "Here in NY, we have a huge population of illegal immigrants from China, Russia, Israel, and the Middle East." Reasoning: This comment is irrelevant because it does not directly refer to abortion. Your answer: 0

12 Pro-abortion

In this task, you will be shown a Facebook post and two comments on this post. One comment is from person A and the other comment is from person B.

Your job is to determine if the comment by B is related to the positive characterization of abortion, the termination of a pregnancy either by choice or due to medical reasons. This involves portraying abortion in a favorable light, highlighting perceived benefits or advantages. Your answer must be a single number: 0 if irrelevant, 1 if relevant.

Here are some examples with reasoning. Use the following format for your output:

Example 1: Comment by B: "Many women are unintentionally impregnated by a RAPIST. Involuntarily having sex is what rape is, if you didn't catch that. A woman should be able to choose if she wants to have an abortion" Your reasoning: The comment is relevant because it characterizes abortion positively, arguing that abortion should be a necessary option for women who become pregnant as a result of rape. Your answer: 1

Example 2: Comment by B: "Some women don't want to have kids at a particular time. Some never want to have kids. Second, most women who have abortions do so because of financial reasons. Third, contraception isn't 100 percent effective. Some types are only 75 percent or so effective." Reasoning: This comment is relevant because it views abortion positively, emphasizing that a woman should be able to decide when and under what circumstances to have kids. Your answer: 1

Example 3: Comment by B: "Here in NY, we have a huge population of illegal immigrants from China, Russia, Israel, and the Middle East." Reasoning: This comment is irrelevant because it does not directly refer to abortion. Your answer: 0

13 Anti-gun

In this task, you will be shown a Facebook post and two comments on this post. One comment is from person A and the other comment is from person B.

Your job is to determine if the comment by B is related to the negative characterization of the possession and use of guns. This involves portraying guns in an unfavorable light, highlighting perceived flaws, contradictions, or controversial aspects. Your answer must be a single number: 0 if irrelevant, 1 if relevant.

Here are some examples with reasoning. Use the following format for your output:

Example 1: Comment by B: "The stats show that you're more likely to shoot a family member than anyone else. Guns don't kill people, but people with guns can and do kill more people than those without. Who would have been at fault if the kids shot out the car window and killed someone?" Your reasoning: The comment is relevant because it

characterizes guns negatively, arguing that they result unintended consequences such as accidental killings. Your answer: 1

Example 2: Comment by B: "We have too many guns, too easily had, by too many people. Guns within easy reach are far more likely to do harm than good." Reasoning: This comment is relevant because it characterizes guns negatively, highlighting their excessive presence and harmful effects. Your answer: 1

Example 3: Comment by B: "Here in NY, we have a huge population of illegal immigrants from China, Russia, Israel, and the Middle East." Reasoning: This comment is irrelevant because it does not directly refer to guns. Your answer: 0

14 Pro-gun

In this task, you will be shown a Facebook post and two comments on this post. One comment is from person A and the other comment is from person B.

Your job is to determine if the comment by B is related to the positive characterization of the possession and use of guns. This involves portraying guns in a favorable light, highlighting perceived benefits or advantages. Your answer must be a single number: 0 if irrelevant, 1 if relevant.

Here are some examples with reasoning. Use the following format for your output:

Example 1: Comment by B: "This is exactly what the government wants, take our guns so they can declare the U.S. a dictatorship!" Your reasoning: The comment is relevant because it portrays guns in a positive light, arguing that they are necessary for preventing government overreach. Your answer: 1

Example 2: Comment by B: "Women are one of the fastest group of gun owners in the country. Many single mothers and single women carry weapons, it is the best defense." Reasoning: This comment is relevant because it views guns positively, emphasizing that they are essential to protection. Your answer: 1

Example 3: Comment by B: "Here in NY, we have a huge population of illegal immigrants from China, Russia, Israel, and the Middle East." Reasoning: This comment is irrelevant because it does not directly refer to guns. Your answer: 0

15 Anti-Conservative or Anti-Republican Rhetoric

In this task, you will be shown a Facebook post and two comments on this post. One comment is from person A and the other comment is from person B.

Your job is to determine if the comment by B is related to the negative characterization of Republicans or conservatives. This involves portraying Republicans and conservatives in an unfavorable light, highlighting perceived flaws, contradictions, or controversial aspects. Your answer must be a single number: 0 if irrelevant, 1 if relevant.

Here are some examples with reasoning. Use the following format for your output:

Example 1: Comment by B: "Conservatives can also be whiny." Your reasoning: The comment is relevant because it portrays conservatives negatively by describing them as complaining, overly critical, or expressing dissatisfaction in a way that may be perceived as annoying or unproductive. Your answer: 1

Example 2: Comment by B: "Republicans want to end abortions so that they will have more population to draw on for their future wars. Think about it: End abortions, brings on many children that the parents can't care for (likely because of finances)." Reasoning: This comment is relevant because it characterizes Republicans negatively by suggesting that

their opposition to abortion is driven by a desire to increase the population for potential future wars, implying a lack of concern for the welfare of children and families. Your answer: 1

Example 3: Comment by B: "Here in NY, we have a huge population of illegal immigrants from China, Russia, Israel, and the Middle East." Reasoning: This comment is irrelevant because it does not directly refer to Republicans or conservatives. Your answer: 0

16 Anti-Liberal or Anti-Democrat Rhetoric

In this task, you will be shown a Facebook post and two comments on this post. One comment is from person A and the other comment is from person B.

Your job is to determine if the comment by B is related to the negative characterization of Democrats or liberals. This involves portraying Democrats and liberals in an unfavorable light, highlighting perceived flaws, contradictions, or controversial aspects. Your answer must be a single number: 0 if irrelevant, 1 if relevant.

Here are some examples with reasoning. Use the following format for your output:

Example 1: Comment by B: "Liberals are actually MURDERING people with their stance. The liberals are the real murderers." Your reasoning: The comment is relevant because it portrays liberals negatively by describing them as murderers. Your answer: 1

Example 2: Comment by B: "Americans are seeing the true colors of Democrats. Most are trash that don't pay taxes. They live off the taxpayers!!" Reasoning: This comment is relevant because it characterizes Democrats negatively by suggesting that they evade taxes. Your answer: 1

Example 3: Comment by B: "Here in NY, we have a huge population of illegal immigrants from China, Russia, Israel, and the Middle East." Reasoning: This comment is irrelevant because it does not directly refer to Democrats or liberals. Your answer: 0