Step 1: Prompt for Persona\_creation

Given a text: ##input\_text

classification\_label: {Hatespeech, Not Hatespeech}

Your task is to create a pool of 6 to 10 debate agents, each of whom will analyse the given text from a distinct perspective. Each agent should represent a unique viewpoint whether it caste based hate or not that is relevant to the input text.

For each agent, assign a generic caste persona description in one sentence, along with a corresponding claim that focuses on a specific angle to analyse the input text. Ensure that each agent's claim and response is distinct and relevant to the caste based hate speech input text. To foster diversity and fairness, the agents should reflect a variety of Indian caste communities and perspectives.

Please format your persona descriptions as follows, with each line being a JSON object:

{"agent\_id": 0, "description": the\_description\_of\_Agent\_0, " classification\_label ": the\_ classification\_label \_of\_Agent\_0}

{"agent\_id": 1, "description": the\_description\_of\_Agent\_1, " classification\_label ": the\_ classification\_label \_of\_Agent\_1}

{"agent\_id": 2, "description": the\_description\_of\_Agent\_2, " classification\_label ": the\_ classification\_label \_of\_Agent\_2}

...

Step 2: Prompt for Personna pool selection

Given a text: ##input\_text

classification\_label: {Hatespeech, Not Hatespeech}

You need to build a team of three agents, to work together and collaboratively formulate a persuasive classification\_label that analyse the given text.

Now given the following candidates, where each candidate has a unique persona offering a different perspective relevant to the topic at hand. You need to select three agents that you think can together form a strong team to achieve the task.

When making your selections, consider the importance of diversity to ensure a balanced and fair discussion. For each selection, give the reason why you select the candidate.

## Candidate list:

###candidate\_list

Please select three candidates and add a reason. Each line of output should be a JSON object as follows:

{"agent\_id": 0, "description": the\_description\_of\_Agent\_0, " classification\_label ": the\_ classification\_label \_of\_Agent\_0, "reason": the\_reason\_of\_selection}

{"agent\_id": 1, "description": the\_description\_of\_Agent\_1, " classification\_label ": the\_ classification\_label \_of\_Agent\_1, "reason": the\_reason\_of\_selection}

...

Step 3: Prompt for Debate\_discussion\_plan

## Goal: Modeling a debate process to analyze a given text on a caste based topic, and formulate a well-structured classification label plan to analyse the text based on the debate discussion.

## Participants and Roles

1. A Main Team of three members: Agent A, Agent B, and Agent C

- Stance: Analyse the Given text;

- Goal: Discuss together to propose a persuasive classification\_label plan outlining the overall logical flow to analyse text.

- Specific Personas and classification\_label of the team members:

- Agent A: persona\_a; classification\_label;

- Agent B: persona\_b; classification\_label;

- Agent C: persona\_c; classification\_label;

2. A Critic

- Stance: Analyse the Given text;

- Goal: You Disagree with the Main Team. Identify and challenge weaknesses in the Main Team's discussion, and debate with the Main team.

## Additional Guidelines

- The discussion should be conducted for multiple rounds until the Main Team members are satisfied with their classification\_label plan and Critic is persuaded.

- The discussion should provide a rigorous reasoning so that the logic flow is persuasive and coherent.

- The discussion does not need to follow a strict order of participants but should prioritize the sequence of topics and sub-ideas to maintain a coherent progression of ideas.

Text: {input\_text}

Now, simulate this discussion process for the given the text.

Step 4: Prompt for classification\_label

## Text: {text}

Output format: {classification\_label}

## Task:

Give the classification\_label as a response. Do not add any other things.

classification\_label: