A.3 Chain-of-Thoughts Prompting with Cluster Knowledge Incorporation

Prompt Instructions

Please detoxify the provided sentence using the structure below without changing the real meaning of the sentence.
The sentences are **clustered into 3 groups** while each group has its own characterizes.
Cluster 0 is more Offensive, Hostile and Vulgar;
Cluster 1 is more Condescending, Derogatory and Hostile;
Cluster 2 is more Informal, Casual, Dismissive.

For each sentence and cluster that I give, **make the sentence non-toxic by making it Neutral/Informal/Casual without changing the meaning.**
Analysis Structure (do not use ” and [] and "" in your answer and do not suggest improvement!):
{
Sentence: {sentence},
Toxicity level: {Specify here},
Cluster: {cluster},
Fixed sentence: <the non-toxic sentence after making it Neutral/Informal/Casual without changing the meaning>;
},

Example:
{
Sentence: dude should have been taken to api , he would be right at home with all the other knuckleheads there,
Toxicity Level: Medium,
Cluster: 0,
Fixed sentence: It would have been good if he went to api. He would fit in.
}