In our case, ChatGPT performed multitask annotation based on social determinants of mental health, which means that it classified/labeled the text based on multiple predefined categories simultaneously. This made it more efficient for handling complex annotation tasks that involve multiple aspects or classes. ChatGPT based annotation offered us with consistent annotation based on its trained knowledge and language understanding. It eliminated potential human biases or variations that can arise when multiple human annotators are involved. This annotation through ChatGPT helped in processing and analyzing the text quickly, making it suitable for annotation tasks that require fast turnaround times. It had handled a large volume of text and provided annotations in a timely manner.

We have defined the social determinants for mental health from Class1 to Class12 (Table-1) and took help of ChatGPT to get the ChatGPT annotated labels for below Text:

**Text:**I worked with my sister to resolve a family conflict related to my mothers health. She did not do her part and that led to my mother's deteriorating health. I learnt that I can not trust and rely on my sister. I would have found other people to take care of my mom's health and manage the situation. I will identify 2-3 resource people and keep in touch with them to make sure that this doesn't happen again.

Upon analyzing the given Text to ChatGPT, we observed that it belongs to Class 2: Transition of care, while the other labels were not relevant. Subsequently, we generated a response for the same Text, but ChatGPT stated that it doesn't correspond to any of the established classes. Subsequently, we made another attempt to generate a response for the Text, categorizing it under another annotation as Class 3: Barriers to care.

Thus, through our experiences, we can say that the ChatGPT annotated labels are not 100 % reliable. It had miscategorized the labels, even after instructing the name, definition and keywords for each Class labels. Here, ChatGPT changed the Class labels and sometimes did not label at all.

Hence, we can conclude that ChatGPT lacks true human understanding and context. It relies solely on patterns and information from its training data, which means it may not fully grasp the nuanced or subjective aspects of text. This can lead to occasional misinterpretations or inaccurate annotations like we have seen in the above case.