## Cross evaluation - Used smaller LLM for notion extraction model davinci2 0.6 davinci3 chatgpt prompt zeroshot 0.5 fewshot accuracy zeroshotCOT 0.3 0.2 0.825 0.830 0.835 0.840 0.845

score