

Наиболее полным и качественным кажется подход, рассмотренный в статье “Watch the Neighbors: A Unified K-Nearest Neighbor Contrastive Learning Framework for OOD Intent Discovery”.

Основываясь на подходе, описанном в данной статье, нужно иметь размеченные интенты и неразмеченные OOD. Данный подход в общем более точно выделяет классы запросов. В целом, KCOD использует предварительные знания из размеченных известных интенгов для кластеризации неразмеченных данных OOD. Это позволяет определять, какие данные являются OOD и какие могут быть приняты за известные интенты. Контрастное обучение K-ближайшего соседа и контрастная кластеризация K-ближайших соседей играют ключевую роль в этом процессе, позволяя увеличить различия между классами и уменьшить различия внутри классов, что приводит к более четким границам между классами. Таким образом, с помощью различных KNN подходов, можно будет оценить степень близости кластеров OOD к известным интенгам.

Это даст возможность модели быть близкой к верной классификации запроса. Модель не так сильно огорчит пользователя, задав уточняющий вопрос, близкий к теме запроса пользователя, в сравнении со случаем, когда модель даст такой ответ, как «я не понимаю вашего вопроса».