

АКТИВНОЕ ОБУЧЕНИЯ ДЛЯ МУЛЬТИМОДАЛЬНОЙ СЕТИ

Крылов Д.П. Университет ИТМО

Научный руководитель - к.т.н., Сметанников И.Б. Университет ИТМО

Аннотация:

Активное обучение занимается решением вопроса выбора новых данных, зная разметку которых можно ускорить обучение моделей машинного обучения. В данной работе рассматриваются мультимодальные нейронные сети, особенность которых заключается в том, что им на вход могут подаваться данные принципиально разных типов, например, картинка и текст. Целью данной работы является проведение исследования подходов к активному обучению мультимодальных сетей.

Введение. Одним из ключевых факторов, влияющих на качество получаемых машинным обучением моделей является количество данных, на которых она обучена, так в общем случае чем больше данных, тем лучше можно обучить модель. Как правило при решении задачи изначально имеется корпус размеченных данных, например, врачебные документы, в которых уже есть и симптомы и диагнозы.

Однако, в ряде задач не имеется изначально размеченных данных, например такой задачей является, категоризация постов в социальной сети. В этом случае требуется нанимать экспертов, которые смогут классифицировать данные вручную. В этом случае актуальной становится проблема уменьшения количества данных, которые отдаются на разметку эксперту при сохранении или улучшении качества получаемых моделей, что является одной из возможных постановок задачи активного обучения.

Основная часть. В активном обучении есть часть методов, основанных на расстояниях между объектами. Специфика мультимодальных данных в том, что не сразу понятно, как эти расстояния находить. Хорошо находить расстояния получается только для данных одной модальности (между двумя картинками или двумя текстами). В данной работе исследуется, как применять стратегии активного обучения для мультимодальных сетей.

Одним из возможных вариантов решения является использование методов, не зависящих от расстояний между объектами, например, Uncertainty sampling или Query-By-Committee.

Другой возможный подход - определение наиболее важной модальности и применение методов, учитывающих расстояния только по одной модальности.

Также возможно посчитать метрики важности для объекта отдельно по каждой модальности и научиться их сочетать.

Выводы.

Данное исследование применимо при обработке данных социальных сетей, где одна сущность (запись на странице) может включать в себя данные разных модальностей (текст и картинка). Активное обучение важно для таких исследований, так как записи социальных сетей размечаются людьми, чье рабочее время это ограниченный ресурс.

Крылов Д. П. (автор)

Подпись

Сметанников И. Б. (научный руководитель)

Подпись