

УДК 004.89

ПРИМЕНЕНИЕ ФОРМАЛЬНОГО КОНТЕКСТА ДЛЯ АУГМЕНТАЦИИ ДАННЫХ В ЗАДАЧЕ АВТОМАТИЧЕСКОГО ЛОГИЧЕСКОГО ВЫВОДА

Чирковский А.Д.

(Университет ИТМО)

Научный руководитель – д. т. н., доцент, Бессмертный И. А.

(Университет ИТМО)

Недостаток специализированных данных затрудняет процесс обучения моделей для задач логического вывода. В работе рассматривается метод применения формального контекста целевой предметной области для увеличения объёма данных и дообучения модели.

Исследования выполнены за счет финансирования университета ИТМО в рамках НИР № 619423 «Разработка виртуального диалогового помощника для поддержки проведения дистанционного экзамена на основе аргументационного подхода и глубокого машинного обучения».

Логический вывод (Natural Language Inference) остаётся нерешённой задачей и области обработки естественного языка (Natural Language Processing). Продолжают появляться новые наборы данных, посвящённые данной проблеме, как отечественные (LRWC), так и зарубежные (SNLI, MNLI), и новые методы. Одним из наиболее успешных современных методов решения этой проблемы являются нейросетевые модели на основе архитектуры Transformer, однако такие подходы имеют ряд недостатков: требовательная к вычислительным ресурсам и объёму данных процедура предобучения модели, а также необходимость последующего дообучения модели на данных целевой предметной области. Кроме того, подобные модели основываются на распределении слов в исходном корпусе текстов, на основе которых часто невозможно построить семантические отношения всех слов ввиду их разреженности. В связи с этим при обучении модели предлагается использовать источник знаний о мире в виде формального контекста.

Формальный контекст – структура, описывающая некую предметную область, состоящая из множества объектов, атрибутов и бинарного отношения между атрибутами и объектами. Механизм анализа формальных понятий предлагает множество методов для манипулирования формальными контекстами – извлечения концептов (совокупности объектов и соответствующих им атрибутов, которым зачастую соответствует какое-то понятие предметной области) и определения их отношений, таких как импликация, независимость и несовместность. В каноничном подходе задача логического вывода представляется как задача классификации, где для каждой пары высказываний следует указать, следует ли логически второе высказывание из первого, есть ли логическое противоречие, или независимы ли два высказывания между собой. Таким образом, для дообучения модели по имеющемуся формальному контексту случайным образом извлекалась пара концептов, устанавливалось их отношение, и затем данная пара преобразовывалась в пример для обучения сети на естественном языке посредством шаблонных конструкций и проведения аналогии между типами связей концептов и типами меток в задаче логического вывода. Полученные данные использовались для дообучения сети, а затем проверялось качество работы на концептах и понятиях, которых не было в обучающей выборке.

Эксперименты показали, что предложенный метод позволяет значительно увеличить точность классификации для объектов исходного контекста, которые не участвовали при обучении модели, и при этом качество работы на исходном наборе данных упало незначительно.