ПРИМЕНЕНИЕ ТРАНСФЕРНОЙ ТЕХНОЛОГИИ В МАШИННОМ ОБУЧЕНИИ

Р.Ф. Долгачев (Университет ИТМО, г. Санкт-Петербург)

Научный руководитель – доцент, Н.А. Осипов (Университет ИТМО, г. Санкт-Петербург)

Введение. Трансферное обучение не является моделью или техникой машинного обучения; это скорее «методология проектирования» в машинном обучении. Другим типом «методологии проектирования» является, например, активное обучение. Общая идея трансферного обучения заключается в том, чтобы использовать знания, извлеченные из задач, для которых много помеченных данных доступно в условиях, когда доступны лишь немного помеченных данных. Создание помеченных данных стоит дорого, поэтому ключевое значение имеет оптимальное использование существующих наборов данных.

Цель работы. Поставленная цель предполагает анализ концепции трансферного обучения и её применения в машинном обучении.

Базовые положения исследования. В традиционной модели машинного обучения основной целью является обобщение данных на основе шаблонов, извлеченных из данных обучения. С помощью трансферного обучения этот процесс обобщения запускается, начиная с шаблонов, которые были изучены для другой задачи. По сути, вместо того чтобы начинать процесс обучения с чистого листа, процесс начинается с шаблонов, которые были изучены для решения другой задачи. В процессе работы рассмотрены текущие варианты применения трансферных технологий в машинном обучении. Были выявлены основные недостатки и преимущества данной технологии.

Промежуточные результаты. В процессе работы была определена общая идея трансферного обучения: использование знаний, извлеченных из задач, для которых имеется большое количество маркированных данных, в задачах, где доступно лишь небольшое количество маркированных данных. Были выявлены причины определяющие, почему хорошее понимание трансферного обучения является критически важным навыком для исследователя данных. Представлены и описаны подходы трансферного обучения, использующиеся в компьютерном зрении:

- Когда доступно достаточное количество изображений (> 1000 изображений на класс);
- Когда доступно только очень небольшое количество изображений (<1000 изображений на класс).

Основной результат. В результате анализа применения трансферной технологии в машинном обучении были выявлены основные преимущества и недостатки, а также описаны причины, которые определяют важность данной технологии.

Автор Долгачев Р.Ф.
Научный руководитель – к.т.н, доцент, Н.А. Осипов
Зав. Кафедрой Программных систем, к.т.н., доцент Т.В. Зудилова