



МОСКОВСКИЙ АВИАЦИОННЫЙ ИНСТИТУТ
(НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ)

Лабораторная работа №7

Выполнила: Прудникова А. А.

Группа: М8О-408Б-20

Обучение с подкреплением является областью машинного обучения. Оно связано с выбором подходящего действия для максимизации вознаграждения в конкретной ситуации. Оно используется различными программами и машинами для нахождения наилучшего возможного поведения или пути, который они должны предпринять в определенной ситуации. Обучение с подкреплением отличается от обучения с учителем тем, что в обучении с учителем обучающие данные содержат ключ с ответом, поэтому модель обучается на самом правильном ответе, тогда как в обучении с подкреплением нет ответа, но агент подкрепления решает, что делать для выполнения заданной задачи. При отсутствии обучающего набора данных он вынужден учиться на своем опыте.

Обучение с подкреплением (RL) - это наука принятия решений. Речь идет об изучении оптимального поведения в среде для получения максимального вознаграждения. В RL данные накапливаются из систем машинного обучения, использующих метод проб и ошибок. Данные не являются частью входных данных, которые мы можем найти в обучении с учителем или без учителя.

Обучение с подкреплением использует алгоритмы, которые учатся на результатах и решают, какое действие предпринять дальше. После каждого действия алгоритм получает обратную связь, которая помогает ему определить, было ли сделанный им выбор правильным, нейтральным или неправильным. Это хороший метод для использования в автоматизированных системах, которые должны принимать множество мелких решений без участия человека.

Обучение с подкреплением - это автономная самообучающаяся система, которая по сути учится методом проб и ошибок. Она выполняет действия с целью максимизировать вознаграждения или, другими словами, учится путем практики для достижения наилучших результатов.

Одним из методов, который можно считать методом метаобучения, является Model-Agnostic Meta-Learning (MAML). Он относится к оптимизационным подходам в метаобучении. MAML - это алгоритм, который обучает модель

так, чтобы она могла быстро адаптироваться к новым задачам путем мелкой настройки с небольшим количеством примеров. Он использует градиенты для обновления параметров модели таким образом, чтобы максимизировать производительность на новых задачах после нескольких шагов обновления.

Другим методом метаобучения является Prototypical Networks, который относится к методам адаптации черного ящика. Prototypical Networks - это непараметрический подход, использующий принципы немногократного обучения. Он извлекает представления данных (прототипы) из обучающих примеров и классифицирует новые примеры на основе их сходства с этими прототипами.

В контексте методов метаобучения, таких как MAML и Prototypical Networks, θ и ϕ могут быть использованы для обозначения различных параметров модели. Например:

- θ может представлять исходные параметры модели до адаптации к новой задаче.
- ϕ может обозначать параметры модели после адаптации к новой задаче путем мелкой настройки или обновления параметров.

Таким образом, алгоритмы метаобучения часто нацелены на оптимизацию параметров θ для быстрой адаптации к новым задачам, получая ϕ после нескольких шагов обновления или настройки.

Статья: [Reinforcement learning](#)