

Лабораторная работа 6

“Метаобучение”

Моисеенков Илья, М8О-408Б-19

Сиамские нейронные сети

Сиамские нейронные сети - это специальный тип нейронных сетей, который используется для задач сравнения или идентификации объектов. Они были впервые предложены в 1993 году в статье "Signature verification using a 'siamese' time delay neural network"

(https://www.researchgate.net/publication/270283023_Signature_Verification_using_a_Siamese_Time_Delay_Neural_Network)

Основная идея сиамской нейронной сети заключается в том, что она состоит из двух одинаковых подсетей, каждая из которых принимает на вход один объект. И каждая подсеть возвращает выходной вектор, который описывает этот объект. Затем выходные векторы двух подсетей сравниваются между собой, и результатом работы сети является значение, которое показывает, насколько похожи два объекта.

Сиамские нейронные сети могут использоваться для задач распознавания лиц, распознавания рукописного текста, классификации отпечатков пальцев и многих других. Все эти задачи связаны с идентификацией или сравнением объектов на основе их признаков.

Сиамские нейронные сети очень хорошо работают в задачах, где набор данных невелик или когда объекты имеют сложную структуру, например, когда это лица с разными выражениями лица или различные написания одной и той же буквы.

Сиамские нейронные сети относятся - пример непараметрического метаобучения, потому что они обучаются на задаче сравнения объектов, а не на конкретной задаче классификации или регрессии. При этом, сеть должна научиться извлекать общие признаки из двух объектов и на основе этих признаков определять их сходство или различие.

Meta-SGD

Meta-SGD (Meta Stochastic Gradient Descent) - это алгоритм optimization-based метаобучения, который использует градиентный спуск для обучения модели, которая быстро адаптируется к новым задачам.

(<https://arxiv.org/pdf/1707.09835.pdf>)

В Meta-SGD первоначально обучается модель на большом наборе данных, который содержит в себе большое количество задач. Затем эта модель используется для генерации вспомогательных датасетов, которые используются для обучения модели быстрой адаптации. Каждый вспомогательный набор данных соответствует отдельной задаче.

Далее, модель быстрой адаптации обучается на каждом вспомогательном наборе данных, используя градиентный спуск. После этого модель быстрой адаптации проверяется на тестовых данных, которые не использовались при обучении. Если модель быстрой адаптации успешно выполняет задачу, то ее веса используются для обновления весов основной модели.