

УДК 004.852

ОЦЕНКА ЭПИСТЕМИЧЕСКОЙ НЕОПРЕДЕЛЕННОСТИ В ГЛУБОКИХ НЕЙРОННЫХ СЕТЯХ НА ОСНОВЕ БАЙЕСОВСКОГО ВЫВОДА

Каширин М.Ю. (Университет ИТМО)

Научный руководитель – кандидат физико-математических наук, Фильченков А.А.
(Университет ИТМО)

Настоящий доклад рассматривает задачу оценки апостериорного максимума в байесовской нейронной сети, представление вывода предельного распределения правдоподобия как задачи оптимизации. Описываются методы оценки неопределенности в предсказаниях глубоких нейронных сетей с применением методов семплирования Монте-Карло.

Введение. В настоящее время глубокие нейронные сети способны воспринимать сложные нелинейные отношения из наборов обучающих данных, отображать многомерные представления в массив выходных значений. При использовании полученных результатов часто считается, что они являются точными, что, однако, не всегда соответствует действительности. Оценка неопределенности позволяет обнаруживать ошибки в результатах моделей. В научной литературе описано несколько типов неопределенности. Эпистемическая неопределенность выражает неуверенность модели в своих предсказаниях. Оценки этой величины особенно полезны, когда объект, получаемый на вход нейронной сетью не принадлежит исходной выборке, так как предсказания модели не позволяют это обнаружить. Одним из способов оценки неопределенности архитектур глубокого обучения является байесовское глубокое обучение, позволяющее моделям решать сложные задачи, производя обучение иерархических признаков представлений, а также выводить апостериорные распределения данных. Для оценки параметров апостериорного распределения используется вариационный вывод, представляющий вывод предельного распределения правдоподобия как задачу оптимизации.

Основная часть.

Метод отключения нейронов на основе семплирования Монте-Карло можно представить как Байесовскую аппроксимацию Гаусовского процесса (Monte Carlo Dropout). Добавление Dropout и использование его для семплирования модели с разными масками отключения нейронов во время предсказания позволяет получить байесовскую нейронную сеть из любой искусственной нейронной сети.

Выводы. Описанные методы отключения путей по расписанию на основе семплирования Монте-Карло позволяют оценивать эпистемическую неопределенность предсказаний глубоких нейронных сетей.

Каширин М.Ю. (автор)

Подпись

Фильченков А.А. (научный руководитель)

Подпись