

Рецензия на рукопись «Inductive Bias Meta-Learning with Generative Models»

Авторы: Анна Новокшонова, Олег Бахтеев, Федор Соболевский, Мухаммадшариф Набиев

Рецензент: Владимир Чайкин

Общий комментарий

В данной работе авторами предложен метод извлечения индуктивного смещения моделей машинного обучения с помощью генерации полных синтетических датасетов. В отличие от предыдущего подхода, где генерировались только метки для фиксированного набора входных данных, здесь предлагается генерировать признаки и метки совместно. Проведены эксперименты для сверточных и рекуррентных нейронных сетей, результаты которых согласуются с известными индуктивными смещениями.

Впечатление от работы положительное. Авторы предлагают альтернативный метод, существенно расширяющий и дополняющий существующие способы работы с индуктивным смещением. Далее я дам несколько комментариев по статье, касающиеся в основном легких неточностей или трудных для понимания мест.

Комментарии к тексту

Введение и постановка задачи

1. В разделе 1 авторы отмечают, что нет единого строгого определения индуктивного смещения, не совсем понятно, количественная ли эта метрика.
2. Вводится распределение признаков при условии метки, но на практике генератор это распределение явно не использует. Из описания невозможно понять, каким образом для разных классов получаются разные распределения признаков.

Метод

1. Во втором регуляризаторе R_svd используется порог τ , способ выбора которого не описан. Неясно, является ли он гиперпараметром или выбирается из каких-то соображений.

Эксперименты

1. В таблице 3 для CNN на 4×4 изображениях указана точность около 99 процентов на тестовых данных с перемешанными блоками выглядит очень высокой, но такой результат подразумевается.
2. В RNN-эксперименте точность на перевернутых последовательностях составляет 0 процентов. Такой результат вызывает вопрос, не произошло ли что-то похожее на переобучение или же упрощение правила.
3. В работе отсутствует сравнение с методом Dorrell et al., 2023 в сопоставимых условиях. Было бы неплохо сравнить, даёт ли генерация полных датасетов преимущество перед генерацией только меток.
4. Визуализации полученных датасетов не содержат подписей осей и шкалы значений пикселей, это затрудняет интерпретацию результатов.

Оформление

1. Псевдокод на странице 4 следовало бы оформить как формальный алгоритм, возможно вынести в аппендикс.

Комментарии к коду

Код запускается, README прописан, github максимально приветлив к посетителям