

Рецензия на доклад Махажанова Нуртаса на тему:
«Применение методов формальных понятий к построению оптимальной архитектуры нейронной сети»

Рецензент: Лобачева Екатерина

Краткое содержание

В первую очередь в данном докладе кратко освещаются две области: методы анализа формальных понятий и нейронные сети. После чего описывается идея применения анализа формальных понятий для выбора структуры нейронной сети. Рассматривается нейронная сеть с двумя скрытыми слоями, структура которых выбирается следующим образом:

1. С помощью решетки формальных понятий выбираются положительные и отрицательные гипотезы. При этом используется не вся решетка формальных понятий, а только ее нижняя часть (то есть в гипотезах ограничено число возможных входящих признаков).
2. Из выбранных гипотез отбираются наилучшие путем отсеивания гипотез с низкой достоверностью или низким коэффициентом поддержки.
3. Каждой гипотезе, отобранной на втором шаге, ставится в соответствие нейрон на первом скрытом слое. Этот нейрон соединяется с нейронами входного слоя, соответствующими входящим в гипотезу признакам.
4. На втором скрытом слое берутся два нейрона: один из них соединяется со всеми нейронами, соответствующими положительным гипотезам, а второй – со всеми нейронами, соответствующими отрицательным гипотезам.

Описанная выше процедура построения нейронной сети и процедура ее обучения были запрограммированы, после чего был проведен эксперимент на одном наборе данных.

Замечания

1. Презентацию в текущем виде довольно сложно воспринимать: на каждом слайде слишком много текста. Стоит оставить на слайдах только основные понятия и формулы, а весь текст говорить словами. Так рассказ будет значительно проще понимать.
2. В презентации стоит выделить отдельный слайд для описания набора данных, на котором проводились эксперименты. Я долго не могла найти, какого же объема были эти данные.
3. Код в текущем виде достаточно долго работает. Стоит поискать более эффективные реализации поиска гипотез путем полного перебора (возможно не на питоне, а на С) или поискать методы, которые делают только приближенный поиск, но более эффективны по времени.
4. В экспериментах нет baseline-решения на основе простой нейронной сети. Сравнение разных вариантов нового метода без их сравнения со стандартным baseline-решением не дает нам понять, дает ли предложенная идея какую-то прибавку в качестве.

Общее впечатление и предложения

Предложенные в докладе идеи показались мне интересными и стоящими исследования, однако в текущем виде эксперименты очень сырые и нужно их доводить до ума. Интересно было бы:

1. Посмотреть, на каких объемах данных предложенный алгоритм при условии эффективной реализации может работать за приемлемое время.
2. Сравнить структуру сети, получаемую с помощью предложенного метода, и сети, которая получается в результате обучения стандартной полносвязной сети. Возможно, при достаточной регуляризации весов стандартной нейронной сети, в ней как раз появятся нейроны, которые будут значительно зависеть только от подмножества признаков. Было бы интересно, если бы эти нейроны были похожи на те, что выдает предложенный метод.
3. Исследовать возможность выбора структуры сети с большим числом скрытых слоев, возможно можно слои как-то привязать к слоям в решетке понятий.