

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

Доклад посвящен применению методов формальных понятий для построения оптимальной архитектуры нейронной сети на примере задачи Breast Cancer Wisconsin (Diagnostic) Data Set из репозитория UCI. Стоит отметить, что данный датасет состоит всего из 569 объектов и 32 признаков, поэтому он далек от большинства реальных задач машинного обучения. В целом была проведена достаточно большая исследовательская работа, поэтому дальше будут сформулировано лишь небольшое количество замечаний, чтобы помочь автору в дальнейших выступлениях.

Во вступлении были введены основные понятия метода анализа формальных понятий (формальный контекст, предикат, операторы Галуа и замыкания, формальное понятие, ...), а также приведен наглядный пример построения решетки формальных понятий для заданного формального контекста (на слайдах, насколько я понял, представлен лишь фрагмент решетки).

В основной части были приведены алгоритмы шкалирования и порождения формальных понятий. Так как обрабатывать все возможные формальные понятия вычислительно трудно, было предложено решение порождать лишь понятия, состоящие не более чем из 4 признаков. Также был введен *Score*, по которому отбирались лишь значимые формальные понятия из порожденных. *Score* был введен как взвешенная комбинация *Confidence* и *Support* формального понятия, причем эти понятия эквивалентны точности и полноте модели при бинарной классификации. Кажется, если ввести *Score* как F_1 -меру (именно этот функционал автор использует для сравнения итоговых моделей), то итоговое качество модели по F_1 -мере может возрасти.

В завершающей части была представлена архитектура нейронной сети, построенной лишь по отобранным формальным понятиям, а также приведены результаты работы обученных моделей на выбранном датасете. Как отмечалось непосредственно на семинаре, не очень понятно, почему количество принятых гипотез не монотонно по числу верхних уровней решетки, а также почему результаты модели FCA не зависят от числа верхних уровней решетки, если это число превосходит 1. Очень не хватило хотя бы какой-нибудь интуиции по этим вопросам. Также было бы интересно увидеть сравнение с другими моделями на этих же данных, источников таких результатов достаточно много, например [1].

В заключение отмечу, что для меня осталось не совсем понятным, что же является основной целью исследования автора. Если формальные понятия действительно позволяют строить «оптимальные» в некотором смысле нейронные сети, то было бы интересно сравнить такие модели с обычными на большем датасете; если же основной целью является построение интерпретируемой модели, то на это стоит сделать более явный акцент в презентации.

[1] You, H. *Comparative Study of Classification Techniques on Breast Cancer FNA Biopsy Data*, 2010.