

УДК 004.8

**АЛГОРИТМ АВТОМАТИЧЕСКОГО ВЫБОРА ОТСЕКАЮЩИХ ПРАВИЛ ДЛЯ
ФИЛЬТРУЮЩИХ АЛГОРИТМОВ ВЫБОРА ПРИЗНАКОВ**

Куриленко В.Ю. (Университет ИТМО)

Научный руководитель – кандидат технических наук, Сметанников И.Б.
(Университет ИТМО)

В данном докладе рассматривается алгоритм поиска и автоматического выбора отсекающих правил для фильтрующих алгоритмов выбора признаков.

На данный момент существует множество фильтрующих алгоритмов (filter algorithms) для выбора признаков, которые предполагают ранжирование признаков с помощью какой-либо метрики также называемой важностью (importance) и последующего отсека ненужных признаков с помощью отсекающих правил (cutting rules). В то время как выбор метрики играет не настолько большую роль, очень важно правильно определить отсекающее правило для конкретного набора данных и конкретной метрики, чтобы оставшийся набор признаков был оптимален в каком-то смысле этого слова (например, согласно какой-либо метрике качества). Так как фильтрующие алгоритмы примечательны своей скоростью работы, разработка алгоритма для подбора нужного отсекающего правила позволит избежать трудоёмкого перебора множества правил и обучения модели на получившемся наборе признаков.

Предлагаемое решение использует методы мета-обучения для выявления зависимостей между мета-признаками наборов данных, метриками и отсекающими правилами. На основании этих зависимостей появляется возможность предсказать отсекающее правило или набор таких правил, которые при их применении позволят получить хорошие практические результаты.

Описанный метод позволяет существенно сократить время работы фильтрующих алгоритмов за счет отсутствия необходимости в переборе отсекающих правил. Алгоритм был протестирован как на синтетических наборах данных, сгенерированных специально для задач по отбору признаков, так и на реальных наборах.

Куриленко В.Ю. (автор)

Подпись

Сметанников И.Б. (научный руководитель)

Подпись