

## АЛГОРИТМ ОПТИМИЗАЦИИ ПРОСТРАНСТВА ПРИЗНАКОВ ДЛЯ РЕШЕНИЯ ЗАДАЧ КЛАССИФИКАЦИИ

Акинин П.В., Карп В.П.

Московский институт радиотехники, электроники и автоматики,  
ф-т кибернетики, каф. Информационных систем,  
Россия, 119454, г. Москва, пр-т Вернадского 78,  
Тел.: (909)923-22-33, E-mail: pavak@mail.ru

Одной из наиболее актуальных задач, решаемых интеллектуальными информационными системами, является задача классификации. Как правило, из-за отсутствия достаточной априорной информации исследователь рассматривает заведомо избыточный набор признаков, которые несут в себе зашумленную информацию о классах и порождают недостоверные зависимости. В таких случаях появляется необходимость в оптимизации пространства признаков [1].

Идея алгоритма базируется на двух итерационных процедурах: процедуре набора признаков и процедуре исключения признаков. Каждая процедура состоит из нескольких этапов. Процедура набора признаков начинается с упорядочения всех признаков в порядке убывания оценки их качества диагностики ( $Q$ ). Признак, имеющий максимальное  $Q$  становится первым признаком в наборе. Затем к данному признаку добавляется тот признак, который дает вместе с первым лучшую оценку  $Q$ . Далее к этим двум признакам аналогичным путем подбирается тот, который дает в сочетании с ними наилучшее  $Q$  (превышающее качество уже имеющегося набора признаков). Процесс продолжается до тех пор, пока среди не включенных признаков не останется ни одного, включение которого в набор улучшило бы оценку  $Q$ . После этого начинается обратная процедура: алгоритм исключает из получившегося набора те признаки, которые не ухудшают оценку  $Q$  сформированного набора. По завершению данной итерации снова делается попытка подобрать из не включенных в набор признаков такие, которые бы улучшили  $Q$ . По ее окончании начинается процесс исключения признаков. Таким образом, образуется итерационная система, которая нечетной итерацией стремится включить признаки, улучшающие оценку качества диагностики, а четной – исключить признаки, не ухудшающие её. В результате формируется набор признаков, который дает наилучшие результаты диагностики для выбранного метода классификации. Важно отметить, что описанная система оптимизации универсальна, так как не подразумевает использование какого-либо конкретного алгоритма классификации. Апробация алгоритма планируется, в первую очередь, на формировании диагностического пространства признаков методом перебора конъюнкций.

### Литература

1. Карп В.П. Методы и средства контроля и диагностики сложноорганизованных систем. – М.: МИРЭА, 2008. 196 стр.