

ПРИМЕНЕНИЕ АЛГОРИТМОВ АВТОМАТИЗИРОВАННОГО ФОРМИРОВАНИЯ НЕЧЕТКИХ КЛАССИФИКАТОРОВ В ЗАДАЧАХ АНАЛИЗА РЕЗУЛЬТАТОВ ДИСТАНЦИОННОГО ЗОНДИРОВАНИЯ ЗЕМЛИ

Сергиенко Р. Б.

Сибирский государственный аэрокосмический университет им. академика М. Ф. Решетнева,
г. Красноярск, Россия

Нечеткий классификатор [1] – алгоритм классификации, основанный на извлечении нечетких правил из массивов данных. Преимуществом данного подхода является возможность явной интерпретации причинно-следственных закономерностей, приводящих к отнесению объекта классификации к различным классам.

В [2] разработан и исследован новый подход к формированию нечетких классификаторов, использующий самонастраивающиеся коэволюционные алгоритмы и гибридизирующий основные подходы к формированию нечетких систем генетическими алгоритмами – Питтсбургский и Мичиганский. Процедура состоит из следующих основных этапов:

1) *Формирование начальной популяции для Мичиганского этапа.* Данная операция очень важна, т. к. случайное генерирование правил для начального заполнения популяции неприемлемо – при значительном числе информативных признаков в задаче классификации вероятность случайной генерации правила, которому соответствовал хотя бы один элемент из обучающей выборки, становится крайне малой. Эта проблема становится существенной уже при размерности четыре и выше. Поэтому необходимо использовать априорную информацию из обучающей выборки.

2) *Мичиганский этап генерирования нечеткого классификатора.* Индивиды представляют собой отдельные нечеткие правила. Длина хромосомы равна числу информативных признаков, каждый ген – число от 1 до 6, соответствующее нечеткому числу. Функция пригодности индивидов – доверительный уровень правила, вычисляемая по обучающей выборке. Используется коэволюционный генетический алгоритм безусловной оптимизации. Популяция с наибольшей точностью классификации используется на следующей стадии генерирования нечеткого классификатора.

3) *Питтсбургский этап генерирования нечеткого классификатора.* Индивиды представляют собой базу нечетких правил целиком. Длина хромосомы равна числу правил, найденных на Мичиганском этапе. Хромосомы бинарные, бит «1» означает использование соответствующего нечеткого правила, найденного на предыдущем этапе, бит «0» – исключение правила из базы. Пригодность – точность классификации базы правил. Вводится ограничение на максимально допустимое число правил, используемых в базе. Используется коэволюционный генетический алгоритм условной оптимизации [3].

Следует отметить, что в основу работы метода положены стохастические алгоритмы оптимизации. Несмотря на статистическую устойчивость метода, разброс в показателях эффективности получаемых нечетких классификаторов при увеличении сложности решаемых задач классификации (увеличение числа классов и/или числа признаков) и ограниченности вычислительных ресурсов возрастает. Показательной является ситуация, когда в обучающей выборке некоторые классы представлены ограниченным числом элементов. При различных запусках автоматизированной процедуры формирования нечетких классификаторов могут получаться базы правил примерно одного уровня точности классификации в целом, но при этом в одной базе правил имеются характерные правила для одних «редких» классов и отсутствуют для других, в другой же ситуация аналогична, но для иных классов. Интуитивно понятно, что подобные нечеткие классификаторы могли бы «взаимно дополнять» друг друга, существенно повышая точность классификации в целом. Поэтому возникла идея разработки метода коллективов нечетких классификаторов.

Здесь можно рассмотреть несколько подходов. Тривиальным является объединение нескольких баз правил в одну большую базу. Однако при этом теряется свойство компактности нечеткого классификатора, важное для простоты интерпретируемости алгоритма экспертами в соответствующей проблемной области. Кроме того, исследования в [2] показали, что увеличение числа используемых правил может приводить к существенному снижению точности классификации. Разнообразные методы голосования в коллективах решающих правил также не лишены указанных недостатков.

Поэтому предлагается на основе нескольких сгенерированных баз нечетких правил формировать новую базу нечетких правил ограниченного объема путём отбора определенных правил из исходных баз правил. По сути, в модифицированном виде повторно реализуется Питтсбургский этап формирования нечеткого классификатора. Аналогом множества правил, полученного на Мичиганском этапе, является множество правил из всех исходных нечетких классификаторов, сгруппированных в единый массив. Целевая функция – точность классификации базой нечетких правил, вводится ограничение на число используемых правил, используется коэволюционный алгоритм условной оптимизации. Данный метод должен позволить генерировать компактные базы правил «повышенной точности», обладающие преимуществами нескольких исходных нечетких классификаторов.

Планируется апробация данного метода при решении задач обработки данных дистанционного зондирования Земли, в частности, в рамках проекта «Малые космические аппараты СибГАУ». В декабре 2011 г. планируется запуск МКА «Юбилейный-2», оснащенного оборудованием для ДЗЗ. Алгоритм обработки данных ДЗЗ на основе метода коллектива нечетких классификаторов может стать важным программным инструментом для использования Студенческим центром управления полётами СибГАУ.

ЛИТАРАТУРА

1. H. Ishibuchi, T. Nakashima, and T. Murata. Performance Evaluation of Fuzzy Classifier Systems for Multidimensional Pattern Classification Problems // *IEEE Trans. on Systems, Man, and Cybernetics*, vol.29, May 1999, pp. 601–618.
2. R. B. Sergienko, E. S. Semenkin, and V. V. Bukhtoyarov. Michigan and Pittsburgh Methods Combining for Fuzzy Classifier Generating with Coevolutionary Algorithm for Strategy Adaptation // *2011 IEEE Congress on Evolutionary Computation*, June 5–8, 2011, New Orleans, LA, USA.
3. R. B. Sergienko and E. S. Semenkin. Competitive Cooperation for Strategy Adaptation in Coevolutionary Genetic Algorithm for Constrained Optimization // *WCCI 2010 IEEE World Congress on Computational Intelligence*. – CCIB, Barcelona, Spain, July, 18–23, 2010. – pp. 1626–1631.