

Венцов Николай Николаевич

E-mail: vencov@list.ru

Chernyshev Jury Olegovich

Don State Technical University.

E-mail: vencov@list.ru

1, Strana Sovetov street, Rostov-on-Don, 344023, Russia.

Phone: 88632589136.

Vencov Nikolay Nikolaevich

E-mail: vencov@list.ru.

УДК 681.3

Ю.А. Кравченко

МЕТОД ПОСТРОЕНИЯ ИМИТАЦИОННЫХ МОДЕЛЕЙ ПРИНЯТИЯ РЕШЕНИЙ НА ОСНОВЕ МНОГОАГЕНТНЫХ ТЕХНОЛОГИЙ*

Описывается новый метод использования интеллектуальных агентов для построения имитационных моделей принятия решений. В основе метода лежит использование двух типов агентов: агент обучения классификатора и агент комбинирования классификатора. Каждый агент обучения классификатора отвечает за чтение вертикального среза выборок и обучение локального классификатора, в то время как агент комбинирования классификатора разработан для комбинации результатов классификации агентов обучения классификатора. Ключевым моментом метода является то, что каждое из подмножеств для агентов обучения классификатора образуется путем разделения признаков, а не серии выборок среды распределения.

Агент обучения классификатора; агент комбинирования классификатора; искусственный интеллект; интеллектуальные агенты; индивидуальные траектории; референтные группы.

Y.A. Kravchenko

THE METHOD OF SIMULATION MODELS CREATION BASED ON THE THEORY OF AGENTS FOR SUPPORT DECISIONS

The object of this paper attention is a new method of intellectual agents application for support decision simulation models creation. The method based on the usage of two types of agents: the classifier training agent and the classifier combining agent. Each classifier training agent is responsible to read a vertical slice of the samples and train the local classifier, while the classifier combining agent is designed to combine the classification results of all the classifier training agents. The key of our method is that the sub-datasets for the classifier training agents are obtained by dividing the features rather than by dividing the sample set in distribution environment.

Classifier training agent; classifier combining agent; artificial intelligence; intelligent agents; individual trajectories; reference groups.

Введение. В данной работе для реализации комбинации гиперплоскостных классификаторов (Hyper Surface Classifier – HSC) предлагается использовать технологию на базе многоагентных систем. Агенты могут имитировать принятие решений в референтной группе обучаемых для разрешения проблем или выполнения учебных проектов, подобно групповым консультациям экспертов. Известно, что модель гиперплоскости получают во время процесса обучения, затем она непо-

* Работа выполнена при поддержке: РФФИ (гранты № 08-01-00473, № 10-07-00538), г/б № 2.1.2.1652.

средственно используется для классификации большой базы данных в зависимости от четности или нечетности числа витков, основанного на теореме Жордана о кривой [1]. Эксперименты показывают, что метод HSC способен эффективно и точно классифицировать массивы данных большого объема в двухмерном и трехмерном пространстве. Хотя метод HSC может теоретически классифицировать данные с более высокой размерностью в соответствии с теоремой Жордана о кривой, его реализация в более чем трехмерном пространстве не так проста. Потребуется алгоритм, который сможет обрабатывать данные не только большого объема, но и высокой размерности. В последнее время внимание специалистов, занимающихся интеллектуальным анализом данных, привлек распределенный интеллектуальный анализ данных (Distributed Data Mining – DDM) [2-5]. Под DDM понимается анализ существенно распределенных массивов данных, направленный на создание глобальных шаблонов из объединения множеств локальных данных. Однако проблема обеспечения безопасности локальных массивов данных и огромная стоимость затрат на перемещение информации не позволяет сделать данные общедоступными. Таким образом, алгоритмы DDM, зачастую, принимают вычислительную парадигму обработки локальных данных и глобального синтеза. Это означает, что процесс интеллектуального анализа данных имеет место сначала на локальном уровне, а затем на глобальном уровне, где результаты локального поиска данных объединяются для получения глобальных результатов [1-3].

Рассматривая проблему распределенной вычислительной среды и то, что агенты могут имитировать принятие решения группой людей для разрешения проблем подобно групповым консультациям экспертов, введем методы многоагентных систем для понимания комбинации гиперплоскостных классификаторов. Комбинация классификаторов – ряд классификаторов, отдельные решения классификации которых объединены определенным образом, как правило, взвешенным или равным голосованием, чтобы классифицировать новые примеры. Существует два способа комбинации: горизонтальная и вертикальная комбинации. В данной работе будем рассматривать вертикальную комбинацию. Комбинация HSC для данных с высокими размерностями строится делением размерности, а не ее сокращением. Будем использовать два типа агентов: агент обучения классификатора и агент объединения классификатора [1]. Каждый агент обучения классификатора выполняет чтение вертикальной части выборок и обучает локальный классификатор, в то время, как агент объединения классификаторов предназначен для объединения результатов классификации всех агентов обучения классификатора.

1. Обзор методов классификаций, основанных на гиперплоскости. HSC –

это универсальный метод классификации, основанный на топологической теореме Жордана о кривой. Основным отличием метода от хорошо известного алгоритма «Метод опорных векторов» является то, что HSC позволяет непосредственно решать нелинейную задачу классификации в исходном пространстве без необходимости отображения данных в пространство с большей размерностью, и, соответственно, без использования функции ядра.

Теорема Жордана о кривой. Пусть X – ограниченное множество в n -мерном пространстве R^n . Если X гомеоморфно сфере в $(n-1)$ -мерном пространстве, тогда его дополнение $R^n \setminus X$ имеет две связанных компоненты, одна из которых называется внутренней, а другая – внешней [1].

Теорема о классификации. Для любой точки $x \in R^n \setminus X$, если x лежит внутри множества X , тогда число кручения, т.е. количество пересечений между любым радиусом от x до X является четным, и, если x не находится в X , тогда количество пересечений между любым радиусом от x до X является четным [1].

Важной задачей является построение разделяющей гиперплоскости. Основываясь на теореме Жордана о кривой, можно выдвинуть следующий метод классификации HSC [1]:

1. Распределить данные выборки внутри прямоугольной области.
2. Преобразовать область к единичной.
3. Разбить область на более мелкие равные части. Если в каких-либо участках будут содержаться выборки, принадлежащие двум или более классам, то разбить их на несколько меньших частей до тех пор, пока в каждой части будет максимальное количество выборок одного и того же класса.
4. Пометить каждую область в соответствии с классом находящихся внутри выборок. Тогда граничные векторы и вектор класса сформируют строку для каждой области.
5. Объединить смежные области одного класса и получить секущую гиперплоскость, сохранить ее в виде строки.
6. Ввести новую выборку и вычислить количество пересечений выборки с секущей гиперплоскостью. Этого можно добиться, нарисовав окружность от выборки. Тогда класс выборки рассчитывается в зависимости от четности или нечетности количества пересечений радиальной окружности и секущей гиперплоскости.

2. Комбинация гиперплоскостных классификаторов. В качестве реализации комбинации гиперплоскостных классификаторов будем использовать много-агентную технологию. Агенты могут имитировать принятие решений референтной группой в процессе учебного проектирования, мозгового штурма или иной ситуации. Каждый классификатор разрабатывается в качестве агента с заданными условными атрибутами. Используются два типа агентов: агент обучения классификатора (АОК) и агент комбинирования (объединения) классификатора (АКК) [1]. После определения соответствия агентом-координатором (АК) между моделями обучаемых, референтных групп и процессов обучения, формируются модели обучения и классификации. Агент обучения классификатора изучает модель классификации с помощью заданного алгоритма на массиве данных и предсказывает возникновение экземпляров с использованием модели обучения. Агент комбинирования классификатора реализует объединение результатов предсказаний множественных агентов обучения классификатора и находит непомяченные экземпляры выборки, позволяющие внести коррективы в исследуемые модели обучения и классификации (рис. 1).

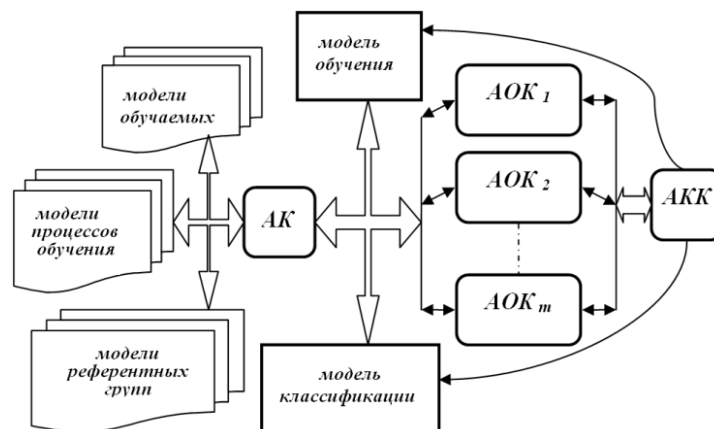


Рис. 1. Имитационная модель принятия решений на основе комбинации агентов обучения и объединения классификаторов

Когда агенты совместно выполняют задачу по классификации, каждый агент обучения классификатора, прежде всего, должен закончить свое обучение самостоятельно и создать независимую модель классификации с заданными различными условными признаками. Затем, когда задача предсказания классификации выполнена, агент объединения классификатора отправляет экземпляры, которые будут предсказаны, каждому агенту обучения классификатора. Агенты обучения классификатора используют модель классификации, чтобы предсказать метку экземпляра, одновременно оценивая его свойства и посылая полные результаты оценки с меткой класса агенту комбинирования классификатора. Агент комбинирования классификатора разрешает окончательную метку класса, основанную на результатах всех агентов обучения классификатора с точки зрения логики управления, таких как голосование или взвешенное голосование. Таким образом, агент объединения классификатора, возможно, превосходит пределы способности одиночных классификаторов.

Следовательно, основанная на агентах система гетерогенных классификаторов HSC может достигнуть высокой точности классификации без дополнительного обучения группы классификаторов. Существуют два метода комбинирования: горизонтальная и вертикальная комбинация. Наиболее существенным отличием является получение подмножеств делением признаков, и не делением исходного множества выборок, поэтому, в случае отсутствия несовместности, размер каждого подмножества равен размеру исходного множества выборок, занимая лишь немногим больше места в памяти, чем исходное множество. Поэтому предпочтение отдается вертикальной комбинации.

Стремясь к решению проблемы нехватки памяти HSC на массивах данных с высокими размерностями, предлагается идея комбинирования. Придавая то же самое значение каждому признаку, во-первых, группируются множественные признаки данных, согласно определенным правилам формируются некоторые подмножества данных, затем начинается процесс обучения и создается классификатор для каждого подмножества данных, и в заключении, окончательное решение определяется объединением последовательности результатов классификации соответствующим образом. Представим описанные действия более детально.

Алгоритм работы агента обучения.

1. Получить размерность условных атрибутов d из массива данных обучающих выборок.

2. Разделить признаки на подмножества $[d/3]$, где $[d/3]$ – наименьшее целое число, большее, чем $[d/3]$. Причем, i -ое подмножество охватывает признаки $3i-2$, $3i-1$, $3i$, $i = 1, 2, 3, \dots, [d/3]$ и атрибут решения. Если d не делится нацело на 3, тогда один или два признака в $(d-1)$ -ом подмножестве добавляются к последнему подмножеству. После этого обучающая выборка разделяется на $[d/3]$ подмножества, каждое из которых имеет одинаковый размер, с тремя условными атрибутами и одним атрибутом решения.

3. Для каждого подмножества, избавиться от несовместности, которая могла возникнуть на Шаге 2. Для каждой выборки в данном подмножестве, если существуют какие-либо другие выборки, в которых значения условных атрибутов совпадают с таковыми в данной выборке, а различие состоит только в атрибуте принятия решения, тогда можно говорить о несовместности. В этом случае необходимо просто удалить данные выборки из этого подмножества.

4. Для каждого подмножества после шага 3 запустить независимый процесс обучения HSC и сохранить результат обучения в виде модели. В итоге получим $[d/3]$ классификаторов HSC, которые и являются комбинацией классификаторов.

Алгоритм работы агента классификации.

1. Получить размерность условных атрибутов d из массива данных обучающих выборок.

2. Разделить признаки на подмножества $[d/3]$, i -ое подмножество охватывает признаки $3i-2, 3i-1, 3i, i = 1, 2, 3, \dots, [d/3]$. Если d не делится нацело на 3, тогда один или два признака в $(d-1)$ -ом подмножестве добавляются к последнему подмножеству. После этого обучающая выборка разделяется на $[d/3]$ подмножества, каждое из которых имеет три атрибута решения.

3. Для каждого подмножества запустить соответствующий классификатор HSC и сохранить результаты классификации. В итоге получим $[d/3]$ результатов классификации для каждой проверочной выборки в исходном множестве данных.

4. Окончательное решение для каждой тестовой выборки в исходном множестве принимается путем голосования. За основу принимается схема голосования множества, в которой коллективное решение – результат классификации, достигнутый большим количеством классификаторов, чем любой другой. При этом дается одинаковое значение всем классификаторам. Таким образом, в случае одинакового количества голосов за два или больше результата классификации, можно случайным образом выбрать один из них.

Комбинация HSC классифицирует множества данных с высокими размерностями путем попытки анализа множественных сечений обучающих и проверочных выборок. Кроме того, комбинация классификаторов намного менее подвержена ошибкам, чем единичный классификатор HSC [1].

Среди методов комбинации для агентов HSC решение большинством голосов является самым простым для реализации, так как оно не требует никакого предшествующего обучения [6]. Использование этого метода является особенно подходящим в ситуациях, когда другие количественные формы выходных данных не могут быть легко получены отдельным агентом классификатора, или когда использование других точных методов комбинирования может быть слишком сложным. Очевидные примеры – некоторые структурные классификаторы. Этот метод комбинации, как было установлено, оказался очень эффективен [6-8]. Путем процесса голосования простым большинством, когда решение каждого классификатора имеет равный вес, могут выполняться различные обработки данных. Можно назначать различные веса для каждого агента классификатора при оптимизации работы объединенного агента классификатора на обучающей выборке. Для первичной обработки веса значения могут быть присвоены при помощи генетического алгоритма и выставлены на голосование каждого классификатора для определения оптимального значения для целевой функции. Эта функция может включать условия распознавания и частоты возникновения ошибок. При этом применяется метод голосования большинством голосов. Агент объединения классификатора принимает решение о принадлежности тестовой выборки к искомому классу C_j , если число классификаторов, которые поддерживают это, значительно больше, чем число классификаторов, которые поддерживают любой другой класс.

Многоагентные среды очень важны для систем комбинации гиперплоскостных классификаторов (Combination HyperSurface Classifiers System – CHCS). Причина, по которой за основу была принята многоагентная технология MAS, состоит в том, что она хорошо подходит под особенности и требования подобных исследований, т.к. естественная распределенная вычислительная среда позволяет используемым интеллектуальным агентам одновременно достигать высокого уровня эффективности. Более того, многоагентная среда является открытой системой. Структура открытой системы способна к динамическому изменению. Ее компо-

ненты могут быть заранее не определенными и изменяться в течение долгого времени, также они могут быть гетерогенными по своей сути. При обучении и тестировании сложного процесса результаты предыдущих вычислений являются входными данными для последующих [9]. Для функционирования подобных систем необходимо разработать массу агентов, специализирующихся на решении разбиения задач автономными способами. Такие агенты будут координироваться через стандартизированный язык взаимодействия агентов, поддержка которого в много-агентной среде осуществляется с помощью протоколов коммуникации.

Заключение. В работе рассмотрена возможность интеллектуальных агентов имитировать принятие решения группой людей для разрешения проблем в различных предметных областях, подобно групповым консультациям экспертов. Описаны два типа агентов: агент обучения классификатора и агент комбинирования классификации HSC. Каждый агент обучения классификатора разрабатывается в виде агента с заданными условными признаками вместо общих, что важно для считывания вертикального среза выборок и обучения локального классификатора. Агент комбинирования (объединения) классификатора предназначен для объединения результатов классификации всех агентов обучения классификатора посредством голосования. Наиболее важное различие между комбинацией HSC и традиционной комбинацией – это то, что подмножества данных получают, разделяя признаки, а не собственно множество выборок, и таким образом, в случае несогласованности, размер каждого подмножества данных равен исходному множеству выборок, а размер занимаемой памяти не намного превышает объем, необходимый для хранения исходных выборок. Временная сложность алгоритма комбинации HSC составляет $o((nd + n^2d)/3)$, где n – размер множества выборок; d – число размерностей.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. He Q. Combination methodologies of multi-agent hyper surface classifiers: design and implementation issues / Qing He, Xiu-Rong Zhao, Ping Luo, Zhong-Zhi Shi // Second international workshop, AIS-ADM 2007, Proceedings. – Springer Berlin Heidelberg, 2007. – P. 100-113.
2. Кравченко Ю.А. Метод определения познавательных стилей на основе теории агентов // Известия ЮФУ. Технические науки. – 2009. – № 12 (101). – С. 120-128.
3. Бова В.В. Технологии интеллектуального анализа и извлечения данных на основе принципов эволюционного моделирования / В.В. Бова, Л.А. Гладков, Ю.А. Кравченко, В.В. Курейчик, В.М. Курейчик, С.Н. Щеглов // Таганрог: Изд-во ТТИ ЮФУ, 2009. – 124 с.
4. Cannataro, M., Congiusta, A., Pugliese, A., Talia, D., Trunfio, P. Distributed data mining on grids: Services, tools, and applications // IEEE Transactions on Systems, Man and Cybernetics. – 2004. – № 34 (6). – P. 2451-2465.
5. Fu Y. Distributed data mining: An overview // IEEE TCDP newsletter. – 2001.
6. Shi Z. Z., H. Zhang Y. Cheng Y. Jiang Q. Sheng and Z. Zhao MAGE: An Agent-Oriented Programming Environment // In Proceedings of the IEEE International Conference on Cognitive Informatics. – New York, 2004. – P. 250-257.
7. Курейчик В.М. Писаренко В.И., Кравченко Ю.А. Инновационные образовательные технологии в построении систем поддержки принятия групповых решений // Известия ЮФУ. Технические науки. – 2008. – № 4 (81). – С. 216-221.
8. Курейчик В.В. Концепция эволюционных вычислений, инспирированных природными системами // Известия ЮФУ. Технические науки. – 2009. – № 4 (93). – С. 16-25.
9. Литвиненко В.А., Ховансков С.А., Норкин О.Р. Оптимизации мультиагентной системы распределенных вычислений // Известия ЮФУ. Технические науки. – 2009. – № 4 (93). – С. 226-235.

Кравченко Юрий Алексеевич

Технологический институт федерального государственного автономного образовательного учреждения высшего профессионального образования «Южный федеральный университет» в г. Таганроге.

E-mail: krav-jura@yandex.ru.

347928, г. Таганрог, пер. Некрасовский, 44.

Тел.: 88634371651.

Kravchenko Yury Alekseevich

Taganrog Institute of Technology – Federal State-Owned Educational Establishment of Higher Vocational Education “Southern Federal University”.

E-mail: krav-jura@yandex.ru.

44, Nekrasovskiy, Taganrog, 347928, Russia.

Phone: +78634371651.

УДК 330.43

Р.Р. Рзаев, Р.А. Кулиев

**ОБ ОДНОМ ПОДХОДЕ К ИДЕНТИФИКАЦИИ ПОТРЕБИТЕЛЬСКОЙ
ФУНКЦИИ ПОЛЕЗНОСТИ**

В целях моделирования потребительского спроса и последующей оптимизации потребительского поведения на рынке товаров и услуг предлагается нейро-нечёткий подход для построения функции полезности, основанный на применении метода нечёткого вывода. Для выборочных наборов товаров и услуг, удовлетворяющих заданным ограничениям по доходам и рыночным ценам, вычисляются соответствующие значения функции полезности. На основе полученной выборки вида «набор товаров – полезность» проведена нейронная идентификация функции полезности.

Потребительская полезность; функция полезности; пространство товаров; нечёткое множество; рекуррентная нейронная сеть.

R.R. Rzaev, R.A. Kuliev

**ABOUT ONE APPROACH TO IDENTIFICATION OF THE CONSUMER
UTILITY FUNCTION**

With a view of modelling a consumer demand and the subsequent optimization of consumer behaviour on a commodity market and services it is offered the fuzzy approach for construction of the utility function based on application of a fuzzy logic. For sample sets of the goods and the services satisfying given set of constraints on incomes and market prices, corresponding values of utility function are calculated. On the basis of the received sample “set of goods – utility” it is carried out neuron identification of utility function.

Consumer utility; utility function; space of the goods; fuzzy set; fuzzy rules; feedforward neural network.

Введение. Традиционные модели прогнозирования спроса не учитывают неметризуемые факторы, обуславливаемые, прежде всего, поведением потребителей. Существенное влияние этих факторов на потребительский спрос является неоспоримым. По существу, их наличие не позволяет с большим основанием считать приемлемой модель временных рядов. Поэтому, чтобы понять основы спроса и получить его адекватную кривую, необходимо исследовать одну из самых важных, трудно формализуемых, характеристик потребительского поведения, каковой является полезность (степень удовлетворённости) от потребляемого набора товаров и услуг.