

METHODS OF REDUCING ATTRIBUTES IN MACHINE LEARNING WHEN WORKING WITH APPROXIMATE SETS

Usov A.Ye.¹, Varlamov A.A.², Babkin O.V.³, Dos E.V.⁴, Mostovshchikov D.N.⁵
(United States of America) Email: Usov561@scientifictext.ru

¹Usov Aleksey Yevgenyevich – Lead Systems Architect;
²Varlamov Aleksandr Aleksandrovich – Senior Solution Architect;
³Babkin Oleg Vyacheslavovich – Senior System Architect;
⁴Dos Evgeniy Vladimirovich – System Architect;
⁵Mostovshchikov Dmitriy Nikolayevich – Senior System Architect,
IT INTEGRATOR «LI9 TECHNOLOGY SOLUTIONS»,
RALEIGH, UNITED STATES OF AMERICA

Abstract: the article provides a detailed analysis of known approaches to the generalization of concepts in terms of their application to real data arrays based on the use of attributes in machine learning. In order to visualize the main stages of machine learning, Figure 1 is shown. It becomes clear that in order to build the necessary qualification rules for the task, it is necessary to work with attributes that meet the requirements of the task.

The analysis of scientific papers on the subject of this article suggests that the use of the theory of approximate methods provides an opportunity to successfully solve tasks related to contradictory and incomplete information. A review of current methods in the theory of approximate sets is carried out, as well as the listed main contradictions with which the researcher can combine to extract and systematize the knowledge contained in the information arrays. A practical implementation is described using the theory of lower and upper approximations, since they allow an assessment of the belonging of the element under study to a set that has fuzzy boundaries.

Successful practical implementation of the method of reducing attributes in machine learning requires digitized data, since it is the data that will determine the quality and completeness of the solution of the problem, based on this, the main theoretical aspects of the generalization algorithm construction using the theory of approximate sets.

Keywords: machine learning, approximate set, attributes, model, information system, discretization, approximation, boundary region, array.

МЕТОДИКА СНИЖЕНИЯ ТРИБУТОВ В МАШИННОМ ОБУЧЕНИИ ПРИ РАБОТЕ С ПРИБЛИЖЕННЫМИ МНОЖЕСТВАМИ

Усов А.Е.¹, Варламов А.А.², Бабкин О.В.³, Дос Е.В.⁴, Мостовщиков Д.Н.⁵
(Соединенные Штаты Америки)

¹Усов Алексей Евгеньевич – ведущий архитектор;
²Варламов Александр Александрович – старший архитектор;
³Бабкин Олег Вячеславович – старший архитектор;
⁴Дос Евгений Владимирович – архитектор;
⁵Мостовщиков Дмитрий Николаевич – старший архитектор,
системный интегратор «Li9 Technology Solutions»,
г. Райли, Соединенные Штаты Америки

Аннотация: в статье осуществляется детальный анализ известных подходов к обобщению понятий в плане их применения к реальным массивам данных на основе использования атрибутов в машинном обучении. С целью визуализации основных этапов машинного обучения приведен рисунок 1, где становится ясно, что для того, чтобы построить необходимые к поставленной задаче квалификационные правила, необходимо работать с атрибутами, которые соответствуют требованиям задачи.

Проведенный анализ научных работ по тематике данной статьи позволяет утверждать, что использование теории приближенных методов предоставляет возможность успешно решать задания, связанные с противоречивой и неполной информацией.

Проводится обзор актуальных методов в теории приближенных множеств, а также перечисленные основные противоречия, с которыми может столкнуться исследователь при извлечении и систематизации знаний, которые содержатся в информационных массивах. Описана практическая реализация с использованием теории нижней и верхней аппроксимации, поскольку они позволяют провести оценку принадлежности исследуемого элемента к множеству, которое имеет нечеткие границы.

Успешная практическая реализация методики снижения атрибутов в машинном обучении требует оцифрованных данных, поскольку именно от данных будет зависеть качество и полнота решения поставленной задачи, исходя из этого в статье изложены основные теоретические аспекты построения алгоритма обобщения с использованием теории приближенных множеств.

Ключевые слова: машинное обучение, приближенное множество, атрибуты, модель, информационная система, дискретизация, аппроксимация, граничная область, массив.

УДК 331.225.3

Введение: В современном мире стремительного развития информационных технологий, одним из приоритетных направлений ученых является разработка систем для принятия решений, где одним из главных вопросов, возникает проблема построения общих моделей, которые будут учитывать результаты обработки огромного количества массивов экспериментальных данных [2]. Указаны массивы, которые представляют информационные потоки данных, имеют место в большинстве фундаментальных наук, таких как: математика, медицина, биология, банковская деятельность и многие другие [1, 4].

Учитывая то, что в силу своих индивидуальных свойств данные имеют разные представления, общим для них, является, то, что они могут содержать огромное количество скрытых закономерностей, которые прежде всего важны при принятии стратегических решений. Определение различных закономерностей в информационных данных не под силу человеческим возможностям, в связи с постоянным стремительным возрастанием объема данных [3].

Для решения описанной проблемы, а именно, для установления закономерностей, исследователи используют методы обобщения и с учетом полученных результатов разрабатывают основные направления машинного обучения, где компьютерные системы способны успешно реализовать поставленные перед ними задачи [1-3].

Анализ последних исследований и публикаций. Изучая научные работы исследователей нашего времени, отметим, что теория приближенных множеств в большинстве работ используется как теоретически основанная база и в то же время как аппарат методов для решения практических задач при извлечении знаний из баз данных (Data Mining) [4-6].

Рассмотрим известный пакет "RoughSet", который имеет в своем составе две части: Теория приближенных множеств (Rough Set Theory (RST)) и Теория нечетких приближенных множеств (Fuzzy Rough Set Theory (FRST)). Если говорить о появлении теории приближенных множеств, то у ее истоков стоит Здислав Павлак, данная теория имеет в своем составе сложный математический аппарат для того чтобы моделировать и проводить детальный анализ информационных систем, которые за своей природой могут иметь неоднородности и неточности. Описанная система использует отношения неразличимости между элементами, и, кроме того не требует каких либо других параметров при необходимости извлечения информации. При практической реализации, используется теория нижней и верхней аппроксимации, которые предоставляют возможность для проведения оценки принадлежности элемента к множеству, которое имеет нечеткие границы [1].

Приблизительные множества в научных работах рассматриваются как множества с неопределенными границами, то есть такие которые невозможно предельно точно описать известными науке наборами атрибутов (признаков) [2, 3].

Теория нечетких приближенных множеств (FRST) можно назвать неким расширением RST, ее фундаментами являются – D. Dubois и H. Prade, которые положили в основу понятия неопределенности и неразличимости, данные понятия и выражены в нечетких множествах, которые предложил L.A. Zadeh. Описанная концепция предоставляет возможность проводить глубокие анализы непрерывных атрибутов (переменных) и при этом не нужно дискретизировать данные. Описанные концепции были использованы учеными в данной области, как результат были предложены методы, которые возможно применить в различных направлениях [5].

Перечислим ряд основных алгоритмов, использование которых позволяет решать задачу по обобщению объектов, представленных определенными наборами атрибутов, к таким алгоритмам относятся:

- 1) фокусирование представлено в работах – Т. Митчелла, Б. Смита;
- 2) индукция решающих деревьев представлена в работах – Р. Куинлана, Р. Кохави, Дж. Шлиммера, Л. Бримана;
- 3) нейронные сети используются в работах Д. Румельхарта;
- 4) привлечение теории приближенных множеств представлено в работах – З. Павлака, Я. Комовски, С. Нгуена [1, 4-6].

Наличие перечисленных алгоритмов для работы с большими массивами данных, которые, как правило можно охарактеризовать в плане размера, полноты, противоречивости, зашумленности собранной информации, являются или вовсе непригодными или же не всегда полученные результаты являются удовлетворительными, в связи с этим возникает вопрос о исследовании методов обобщения с

применением теории приближенных множеств, которая представляет возможность решения проблемы связанной с неполнотой и противоречивостью информации [5].

Формулирование целей статьи (постановка задачи). Провести исследования основных методов для снижения атрибутов с использованием теории приближенных множеств для успешной их реализации в машинном обучении.

Изложение основного материала исследования. Существующие, на данный момент, методы в теории приближенных множеств "RoughSet" можно классифицировать следующим образом:

1) фундаментальные понятия RST и FRST, которые представляют четыре разноплановые задачи – отношение неразличимости; нижнее и верхнее приближение; положительный регион, а также матрица отличий;

2) дискретизация, которая способствует преобразованию вещественных данных в номинальные, если анализировать данный метод с позиции RST, то возможна поддержка различимости между объектами;

3) выбор предикторов, представляет собой некий процесс, который способствует нахождению подмножества предикторов, кроме того, главной целью является выбор существенных особенностей и ликвидации их зависимости, поскольку данная процедура является необходимой, когда пользователь сталкивается с наборами данных, которые в своем составе содержат огромное количество всяческих признаков;

4) выбор экземпляров главной целью представляет удаление шумных, лишних или же неоднозначных экземпляров из рабочих наборов данных, но отметим, что производится сохранение согласованных экземпляров, исходя из описанного принципа работы данного метода, следует отметить, что становится возможным становление отменной точности классификации за счет удаление тех экземпляров, которые не представляют интереса для будущей работы;

5) индукция правил, позволяет осуществлять генерацию правил, которые представляют знания таблицы решений, в научной среде, данный процесс принято называть фазой обучения в машинном обучении;

6) Прогнозирование решений (классификация) используется с целью предсказания значения переменной от сформированного набора данных в разноплановой подачи информации [1-2].

Если характеризовать основные принципы работы перечисленных методов по поводу выбора предикторов и экземпляров, то прежде всего необходимо сформировать наборы данных входного и выходного характера, при необходимости следует их преобразовать в необходимый для работы класс "DecisionTable", поскольку речь идет об использовании номинальных данных, то имея дело с числовыми непрерывными данными необходимо преобразить их в номинальные, для этого следует воспользоваться представленной в данном пакете специализированной функцией дискретизации.

При выборе важных переменных, пользователь не увидит данные, которые не были разделены, поскольку они будут выброшены из набора. Отметим, что если проводить процесс дискретизации текстового набора, то данные необходимо преобразовать в соответствии с осуществленной редукцией. После этого пользователь имеет возможность извлечь набор правил, которые объединяют предикторы и целевую функцию, используя для этого возможности пакета, а именно «индукция правил».

Если говорить о выборе функции, то это специальный тип уменьшения размерности, где набор построенных функций должен выступать подмножеством исходных функций, это значит, что функции не были кардинально изменены, не учитывая включения и исключения.

Процесс уменьшения размерности состоит в том, чтобы сначала модифицировать исходные функции в некое новое представление, это можно сделать, используя преобразование координат к необходимому размеру, или, как вариант, спроектировать данные в пространство, которое имеет высшую размерность, для совершения описанного действия необходимо использовать функцию ядра, на следующем этапе нужно завлечь некую меру информативности с целью отсекация функции в новом пространстве. При уменьшении размерности, имеет место, использование более простого преобразования, которое состоит из усреднении вместе нескольких компонентов вектора, при условии наличия признаков коллинеарности [4].

Отметим, что упрощенный способ преобразования, по-прежнему отличается от выбора функции, тем что новая рассматриваемая функция не будет подмножеством оригинала. Исходя из этого, необходимо учитывать и понимать, следующее: процесс выбора функции не влечет за собой кардинальных изменений, кроме таких, как способность не учитывать функции исходного ввода, которые являются менее информативными, что касается остальных функций, то они сохраняются без изменений.

Необходимость и результативность использования приближенных методов можем найти в работах Я. Базана, С. Нгуена, Я. Степанюка, они предоставили доказательства того, что в алгоритмах обобщения имеет место использование теории приближенных множеств с целью повышения точности при построении классификации объектов.

Изучив перечисленные выше работы, можно сформировать основные этапы при работе алгоритма обобщения с использованием теории приближенных множеств: на первом этапе необходимо

осуществить дискретизацию непрерывных областей значений атрибутов, на следующем этапе – отделить значимые атрибуты иными словами необходимо искать срезы, а также выводить формулировку конечных правил [3].

Решение задачи относительно дискретизации и проведения поиска минимального среза NP-сложны, что определяет необходимость разработки алгоритмов эвристического характера для их успешной практической реализации. Таким образом, исходя из изложенного материала, видим, что вопрос исследования машинного обучения с использованием методики снижения атрибутов с целью обработки огромных выборок за определенное время, а также разработка и практическая реализация алгоритмов обобщения, является приоритетной задачей среди большинства исследователей в современной науке.

В научных работах, принято задачу обобщения информации представлять как одной из фундаментальных задач искусственного интеллекта. Результатом успешной реализации алгоритмов обобщения является формирование основных правил классификации, которые эффективно определяют нужные объекты.

При исследовании методов обобщения с применением теории приближенных множеств, будем под обобщением подразумевать некий переход от рассмотрения локального единичного объекта x или даже некоторого множества объектов X к исследованию такого множества объектов W такого, что $x \in W$ или $X \subset W$. Отметим, что при практической реализации описанного алгоритма в роли объекта может быть реальный физический объект, как следствие некоторого процесса или явления. Для успешного выполнения алгоритмов необходимо, также, припустить, что все объекты находящиеся в множестве W имеют некоторые общие свойства [1, 5].

Для того чтобы детализировать необходимые объекты, принято использовать некие признаки – a_1, a_2, \dots, a_k , которые у многих научных работах называют еще атрибутами. Атрибуты предоставляют возможность для каждого объекта $x \in X$ построить конкретный набор значений его признаков (атрибутов): $x = \{v_1, v_2, \dots, v_k\}$, где v_i – значение i -го признака. Данное описание свойств объектов принято называть признаковым описанием, в роли признаков могут выступать различные свойства в зависимости от природы происхождения рассматриваемых объектов: цена, прибыль, размер, возраст, цвет и др.

Главной целью процесса обобщения является, то, чтобы сравнить между собой описания начальных объектов, которые были заданные совокупностью определенных значений атрибутов, а также выделить наиболее значимые фрагменты этих описаний. В литературе процесс обобщения получил еще одно название – индуктивное формирование понятий.

Множество всех существующих объектов W , которому соответствует определенное понятие, будем называть объемом понятия, при этом, будем различать положительный и отрицательный объект для понятия, если он входит или не входит в объем этого понятия.

Постановка задачи обобщения понятий по атрибутам (задача индуктивного формирования понятий) состоит в том, чтобы, построить понятие, используя при этом основы анализа обучающей выборки U , которое будет способствовать с помощью определенного правила распознавания успешно отобрать все положительные и отрицательные объекты множества обучающей выборки U .

Припустим необходимые условия для решения задачи обобщения, а именно, что: 1) $U = \{x_1, x_2, \dots, x_n\}$ – непустое конечное множество объектов, данное множество является обучающим (универсумом); 2) $A = \{a_1, a_2, \dots, a_k\}$ – непустое конечное множество атрибутов, при этом также известно множество значений атрибута $a - V_a$. Определенное значение атрибута a для некоторого объекта $x \in U$ будем обозначать – $a(x)$.

В процессе решения задачи обобщения может возникнуть необходимость в получении описания понятия, которое вводится значением одного из атрибутов, как правило, его называют решением или решающим атрибутом и обозначим его d . Если атрибуты будут входить в A , то будем их называть условными атрибутами. В зависимости от исходной задачи, исследователь может определить ранг решения $r(d)$ (количество возможных значений для решающего атрибута d), а множество всех значений решения будем обозначать

$$V_d = \{V_1^d, V_2^d, \dots, V_{r(d)}^d\}.$$

Если рассматривать понятия, которые возникли на базе обучающего множества U , то в общем случае данное понятие будет приближением к понятию множества X , и степень близости этих понятий будет зависит от представительности обучающей выборки, в плане полноты отражения особенностей множества X .

Отметим, что для различных методов обобщения знаний многое зависит от способа представления знаний, прежде всего, речь идет о таких способах: исчисление предикатов первого порядка; продукционные правила; решающие деревья; семантические сети.

Следует учитывать, то что при извлечении и обобщении знаний, на практике, может возникнуть перечень проблем таких как: данные имеют различное происхождения; на практике базы данных достаточно велики и по этому не всегда реально использовать алгоритмы экспоненциальной сложности;

массивы данных могут содержать информацию которая будет неправдоподобной, и кроме этого значения ряда атрибутов могут отсутствовать, чтобы это избежать необходимо брать во внимание только существенные атрибуты, тогда будет вполне реально построить классификационные правила; интерпретация результатов должна быть доступной и понятной в информационном плане для конкретного человека. Путем изучения индивидуально каждой из возникших задач, необходимо искать способы для ее успешного решения.

Анализ литературных источников позволяет утверждать, что известные подходы к решению задачи о индуктивном формированию понятий не все соответствуют описанным выше требованиям, по этому следует отметить, что наиболее успешным является подход связанный с использованием теории приближенных множеств, поскольку данный подход позволяет устранить проблемы связанные с противоречивостью, полнотой массивов информационных данных, в результате исследователь имеет комплекс решающих правил, которые помогают ему охарактеризовать скрытые закономерности [1, 4,6].

Наряду с преимуществами теории приближенных множеств существуют и некоторые недостатки с которыми каждый в зависимости от поставленной задачи справляется индивидуально, прежде всего недостатки связаны с затруднением при проведении вычислительных процессов.

Анализируя теорию «приближенного множества», следует отметить, что исследователь работает из «несовершенством данных», которые относятся к «гранулярности» информации. За своей природой возникновения «теория приближенных множеств» является топологической и, как показано на практике, успешно дополняет известные подходы для работы с неполной информацией, речь идет о таких подходах как: нечеткие множества, методы Байеса, нейронные сети, статистические методы для анализа данных, эволюционные алгоритмы.

Тем не менее учитывая все преимущества и недостатки теории приближенных множеств её активно используют для осуществления классификации объектов. Для того чтобы описать методику снижения атрибутов при работе с приближенными множествами рассмотрим некую информационную систему которая представляется у виде пары – $S=(U,A)$, а также решающую таблицу (систему) вида –

$$S = (U, A \cup \{d\}), \quad (1)$$

где $d \notin A$. Пусть $B \subseteq A$, исходя из этого определим отношения неразличимости по B : $IND(B) = \{(x, y) \in U \times U: \forall a \in B (a(x) = a(y))\}$. Множество классов эквивалентности отношения $IND(B)$ обозначим как $\{X_1^B, X_2^B, \dots, X_m^B\}$.

Одну из главных ролей в информационной системе играет то, каким способом осуществляется представления знаний. Если говорить о системах индуктивного формирования понятий то наиболее использованными способами являются: продукционные правила, решающие деревья, исчисление предикатов и семантические сети.

Таким образом, исходя из введенных определений возникает возможность приближенно определять произвольные классы объектов $X \subseteq U$ через построение нижнего и верхнего приближений для X , соответственно:

$$\underline{BX} = \bigcup_{X_i^B \subset X} X_i^B \quad \overline{BX} = \bigcup_{X_i^B \cap X \neq \emptyset} X_i^B.$$

На практике, под нижним приближением для X принимают объединение классов эквивалентности отношения неразличимости, которые принадлежат X , а верхнее приближение множества X понимают как объединение классов эквивалентности, где часть объектов принадлежит X . Таким образом, построенное приближенное множество состоит из пары $\langle \underline{BX}, \overline{BX} \rangle$, при этом разницу между нижним и верхним приближением составляет граничную (недостовверную) область множества $X - BN_B(X) = \overline{BX} \setminus \underline{BX}$ кроме этого, множество $U \setminus \overline{BX}$ имеет в своем составе отрицательные объекты.

Во время извлечения и систематизации знаний, которые расположенные в информационных массивах, пользователи могут сойтутся из такими сложностями:

- 1) необходимые данные имеют разную природу возникновения в плане количественных, качественных и других структурных характеристик;
- 2) базы данных имеют весьма масштабные размеры, а это значит что во время извлечения из них знаний, нужные для этого алгоритмы экспоненциальной сложности не всегда будут работать;
- 3) необходимых значений для ряда атрибутов может вовсе не быть и кроме этого, присутствующая информация в массивах данных может быть деформированная в плане полноты, правдоподобности и лишена других не менее важных свойств информации [3].

Для правильного восприятия использования методики снижения атрибутов при работе с приближенными множествами необходимо понимать, что в понятия нижней аппроксимации исследователи припускают все элементы, которые точно есть в составе X , а у верхнюю аппроксимацию относят элементы, которые возможно принадлежат X . Если говорить о граничной области рассматриваемого множества X то выбирают разницу между верхней и нижней аппроксимацией,

поскольку граничную область составляют элементы из множества X , которые находятся у верхней аппроксимации и не принадлежат нижней аппроксимации.

Отметим, что практическая реализация приближенных множеств имеет широкий спектр применения, а особенно активно приближенные множества используются для классификации объектов, перечислим их: 1) описание $\underline{BC} \rightarrow C$; 2) описание $U \setminus \overline{BC} \rightarrow \neg C$; 3) описание $\overline{BC} \setminus \underline{BC} \rightarrow$ возможно C . Среди перечисленных классификаций описание множества имеет набор характерных атрибутов, где C – некоторый класс решения, подмножество множества V_d . Следует отметить, что для того чтобы построить правильные во всех смыслах классификационные правила необходимо использовать атрибуты, которые соответствуют всем существенным требованиям.

Принято выделять положительную область решающей системы S , у форме некоторого множества $POS_A(d) = \underline{A}C_1 \cup \dots \cup \underline{A}C_{r(d)}$, которое включает объекты, те, что обязательно входят в один из классов решения. Для объекта x , за обобщенное решение следует принимать множество решений объектов, которые входят в один и тот самый класс эквивалентности отношения $IND(B)$ что и сам объект x

$$\partial_B(x) = \{v \in V_d : \exists x' \in U ((x', x) \in IND(B) \wedge d(x') = v)\}.$$

При решении задач могут возникнуть случаи, когда в таблице (1) будут несущественные условные атрибуты или наоборот – одни условные атрибуты, которые будут зависеть од других. При определении среза для решающей системы (1), необходимо найти минимальное подмножество атрибутов $B \subseteq A$, которое предоставит возможность сохранить обобщенное решение для всевозможных объектов обучающего множества, иными словами для которых выполняется условие – $\partial_B(x) = \partial_A(x) \forall x \in U$.

Использование машинного обучения предполагает использование оцифрованных данных, поскольку от того с какими данными работает тот или иной алгоритм зависит качество и полнота решения. Рассмотрим основные этапы машинного обучения в виде блок-схемы на рис. 1 [1].



Рис. 1. Основные этапы машинного обучения

Отметим, что данная иллюстрация основных этапов, на практике, очень редко после обучения сразу получают успешно работающую модель, тогда исследователи между вторым и третьим этапом совершают оценку качества модели с целью устранения ошибок, и только за результатами оценки осуществляется переход на следующий этап или же возврат к предыдущему этапу.

Исследование основных подходов использования атрибутов в машинном обучении, предоставляет возможность утверждать, что для каждой конкретной задачи необходимо выделять N -мерное пространство атрибутов, где N – число, которое показывает измеряемые характеристики объектов,

поскольку для того чтобы задать объекты в пространстве атрибутов используют N-мерные векторы, где каждая компонента которых задает собой значение конкретной характеристики [3].

Таким образом, теорию приближенных множеств можно назвать универсальной поскольку, концепция которая, заложена в ее основу успешно реализовывается как в исследованиях теоретического направления – алгебре, топологии так и в прикладных исследованиях – в плане задач об искусственном интеллекте, теории принятия решений, обработке изображений и многих других.

Выводы. В статье исследованы основные подходы для снижения атрибутов, которые используются в машинном обучении, а именно, популярный на практике метод, в основе которого используется теория приближенных множеств. Для лучшего понимания основных этапов машинного обучения приведен рис 1.

В работе осуществляется обзор основных подходов к обобщению и сравниваются возможности их применения к реальным массивам данных. Исследованы новые возможности, представленные в работах современных ученых, по поводу оценки предикторов, их визуализации, выбору наиболее значимых, следует отметить, что искомым результатом зависит от различных уровней важности, зависимости предикторов и их влияния на отклики.

Приводится алгоритм, который за основу использует теорию приближенных множеств, основными этапами которого являются:

- 1) определение классов эквивалентности отношения неразличимости;
- 2) установление верхнего и нижнего приближений;
- 3) поиск среза для решающей системы;
- 4) формирование набора решающих правил.

Необходимо отметить, что при исследовании атрибутов, которые имеют непрерывные границы решений, следует применять дополнительные методы дискретной математики, что влечет за собой предельно высокую вычислительную сложность этапов дискретизации при выборе существенных атрибутов и построению решающих правил.

Детальное изучение теории приближенных множеств позволяет утверждать что данная теория является универсальной и основная её концепция находит своё применение, как в теоретических так и практических исследованиях, кроме этого теория приближенных методов успешно используется при решении задач об искусственном интеллекте и многих других отраслях современной науки, что представляет интерес для проведения новых исследований в рамках тематики данной статьи.

Список литературы / References

1. Воронина В.В., Мошкин В.С. Разработка приложений для анализа слабоструктурированных информационных ресурсов: учебное пособие / Воронина В.В., Мошкин В.С. Ульяновск: УлГТУ, 2015. 162 с.
2. Клячкин В.Н. и др. Применение методов машинного обучения при решении задач технической диагностики // Научный вестник УВАУ ГА (И), 2016. Т. 8. С. 158–161.
3. Hinton G., Osindero S., Teh Y. A Fast learning Algorithm for Deep Belief Nets. *Neural Computation*, 2006. 18 (7): 1527-1554.
4. Lan A.S. et al. Mathematical language processing: Automatic grading and feedback for open response mathematical questions // Proceedings of the Second (2015) ACM Conference on Learning@ Scale. ACM, 2015. С. 167–176.
5. Rasmussen C.E. The infinite Gaussian mixture model. In *Advances in Neural Information Processing Systems*, Vol. 12, 2000.
6. Wu X. et al. Top 10 algorithms in data mining // *Knowledge and information systems*, 2008. Т. 14. № 1. С. 1–37.