

МЕТОДЫ СИНТЕЗА МОДЕЛЕЙ КОЛИЧЕСТВЕННЫХ ЗАВИСИМОСТЕЙ В БАЗИСЕ ДЕРЕВЬЕВ РЕГРЕССИИ, РЕАЛИЗУЮЩИХ КЛАСТЕР-РЕГРЕССИОННУЮ АППРОКСИМАЦИЮ ПО ПРЕЦЕДЕНТАМ

Субботин С. А. – д-р техн. наук, профессор, заведующий кафедрой программных средств Национального университета «Запорожская политехника», Запорожье, Украина.

АННОТАЦИЯ

Актуальность. Для принятия решений в технических приложениях, как правило, необходимо обладать моделью, позволяющей прогнозировать состояние управляемого объекта или процесса. Объектом исследования является процесс построения моделей зависимостей по прецедентам. Предметом исследования являются методы построения количественных зависимостей по прецедентам на основе кластер-регрессионной аппроксимации.

Цель работы – упрощение моделей кластер-регрессионной аппроксимации путем косвенной реализации кластер-анализа в процессе построения модели.

Метод. Предложен метод древовидной кластер-регрессионной аппроксимации, который для заданной обучающей выборки строит дерево для иерархической кластеризации экземпляров, листовые узлы которого соответствуют кластерам, для каждого кластера строит частную модель зависимости по экземплярам обучающей выборки, попавшим в кластер, стремясь обеспечить наименьшую сложность модели, и использует набор наиболее информативных признаков наименьшей длины. Это позволяет обеспечить приемлемую точность модели, высокие уровни интерпретабельности и обобщения данных, снизить сложность модели, упростить ее реализацию при последовательной организации вычислений.

Результаты. Разработано программное обеспечение, реализующее предложенный метод древовидной кластер-регрессионной аппроксимации. Разработанный метод и реализующее его программное обеспечение исследованы при решении практических задач прогнозирования. Проведенные эксперименты подтвердили работоспособность разработанного математического обеспечения и позволяют рекомендовать его для использования на практике.

Выводы. В отличие от традиционных методов построения регрессионных моделей, строящих модель на основе функции, единой для всего признакового пространства формы, предложенный метод формирует иерархическую комбинацию частных моделей. В отличие от известных методов построения деревьев регрессии, листовые узлы которых содержат усредненные значения выходного признака для кластеров, предложенный метод формирует дерево, состоящее из частных моделей для кластеров, что позволяет обеспечить большую точность модели.

КЛЮЧЕВЫЕ СЛОВА: моделирование, кластер-регрессионная аппроксимация, оценивание, деревья решений, регрессионный анализ.

АББРЕВИАТУРЫ

AIC – Akaike's Information Criterion.

НОМЕНКЛАТУРА

ε – заданное максимально допустимое значение ошибки для всей выборки;

ψ^q – нелинейная функция активации для перцептрона q -й модели;

$\psi^{q,(\eta,i)}$ – функция активации i -го нейрона η -го слоя нейросети q -й модели;

C^q – центр q -го кластера;

C_j^q – координата центра q -го кластера по j -му признаку;

E – функция ошибки модели;

E^q – ошибка частной q -й модели;

F – критерий качества модели;

f – структура модели;

f^q – частная модель-функцию q -го кластера;

I_{IC} – интегральный информационный критерий;

I_Q – интегрированный показатель качества модели;

$I_{interp.}$ – интерпретабельность модели;

$I_{interp.}^{(\eta,i)}$ – интерпретабельность узла;

I_{gen} – коэффициент обобщения модели;

I_{cert} – показатель уверенности модели в принятом решении;

$I^q(x_j, y)$ – индивидуальная информативность j -го признака по отношению к выходному признаку y для экземпляров q -го кластера;

I_{cert}^q – уверенность q -й частной модели в принятом решении;

M – число слоев модели;

$n^{(\eta,i)}$ – сложность функции i -го узла η -го слоя модели;

N – число признаков, характеризующих экземпляры выборки;

N' – число используемых в модели признаков;

N_n – число функциональных узлов модели (для нейросетей – число нейронов, для дерева регрессии – число узлов);

N_w – число настраиваемых параметров модели;

$N_{w=0}$ – число настраиваемых параметров модели, равных нулю;

N_w^{\max} – максимально возможное число настраиваемых параметров модели;

$N^{(\eta,i)}$ – число входов у i -го узла η -го слоя модели;

$N_w^{(\eta,i)}$ – число настраиваемых параметров i -го узла η -го слоя модели;

$N_{q,1}$ – число нейронов в первом слое нейросети q -й модели;

q^s – номер кластера, к которому относится s -й экземпляр;

Q – число кластеров;

$R(x^s, C^q)$ – расстояние от экземпляра x^s до центра q -го кластера;

S – число экземпляров в выборке;

S^q – число экземпляров в q -м кластере (листе дерева);

w – параметры модели;

w_j^q – вес j -го входа частной модели q -го кластера;

$w_j^{q,(\eta,i)}$ – весовой коэффициент j -го входа i -го нейрона η -го слоя нейросети q -й модели;

x – набор экземпляров выборки;

x^s – s -й экземпляр выборки;

x_j^s – значение входного j -го признака s -го экземпляра выборки;

y – набор значений выходного признака экземпляров выборки;

y^s – значение выходного признака, сопоставленное s -му экземпляру выборки;

\bar{y} – набор усредненных значений выходного признака для кластеров;

\bar{y}^q – среднее значение выхода частной модели q -го кластера при подаче на ее входы координат центра q -го кластера;

y_q^{s*} – значение выходного параметра, рассчитанное q -й частной моделью.

ВВЕДЕНИЕ

Для автоматизации управления объектами и процессами разной природы, а также принятия решений в условиях неопределенности, многомерности и нелинейности зависимостей между их свойствами в условиях отсутствия или нехватки экспертных знаний возникает необходимость построения моделей зависимостей по прецедентам – парам точечных наблюдений за моделируемой зависимостью типа «вход-выход».

Объект исследования – процесс построения моделей зависимостей по прецедентам.

Известные методы построения количественных зависимостей такие, как методы регрессионного анализа [1, 2], нейронные сети [3, 4] и метод группового учета аргументов [5, 6] стремятся построить единую модель во всем пространстве признаков путем решения оптимизационной задачи, что требует больших затрат времени и приводит к получению сложной модели. Другие известные методы, такие как деревья

регрессии [7, 8] и нейро-нечеткие сети [9–12], строят комбинацию примитивных частных моделей для областей в пространстве признаков, что позволяет упростить получаемые модели, но существенно снижает их точность. Компромиссом между данными группами методов является метод кластер-регрессионной аппроксимации [13], который позволяет получать более точные, простые и интерпретабельные модели, минимизируя число используемых признаков.

Предмет исследования – методы построения количественных зависимостей по прецедентам на основе кластер-регрессионной аппроксимации.

Цель работы – упрощение моделей кластер-регрессионной аппроксимации путем косвенной реализации кластер-анализа в процессе построения модели.

1 ПОСТАНОВКА ЗАДАЧИ

Пусть задана обучающая выборка S прецедентов $\langle x, y \rangle$, где $x = \{x^s\}$, $y = \{y^s\}$, $s = 1, 2, \dots, S$, $x^s = \{x_j^s\}$, $j = 1, 2, \dots, N$.

Тогда задача построения модели зависимости $y = f(w, x)$ заключается в нахождении такой структуры модели f и значений параметров модели w , при которых будет удовлетворяться критерий качества модели F . Как правило, для задач аппроксимации критерий качества модели определяют как функцию ошибки модели:

$$E = \sum_{s=1}^S \left(y^s - f(w, x^s) \right)^2 \rightarrow 0.$$

2 ОБЗОР ЛИТЕРАТУРЫ

Методы регрессионного анализа [1, 2] позволяют для заданной выборки наблюдений получить полиномиальные модели, коэффициенты которых определяют в общем случае решением оптимизационной задачи минимизации критерия ошибки в пространстве коэффициентов модели. Достоинством методов является их универсальность. Недостатками методов являются неинтерпретабельность получаемых моделей при большом числе признаков, а также проблема выбора оптимальной структуры и сложности модели.

Методы построения прогнозирующих моделей на основе нейронных сетей [3, 4], как правило, так же, как и в случае регрессионного анализа, для построения (обучения) моделей требуют решения оптимизационной задачи минимизации критерия ошибки в пространстве весовых коэффициентов модели. Однако сама модель имеет, как правило, специфическую структуру, которую можно рассматривать как иерархическую комбинацию нелинейных функций и линейных полиномов. Достоинствами метода является его универсальность. Недостатками метода являются неинтерпретабельность и высокая сложность получаемых моделей, проблема выбора оптимальной структуры модели и параметров ее элементов, а также высокая итеративность и значительные затраты времени метода.

Метод группового учета аргументов А.Г. Ивахненко [5, 6] предполагает перебор моделей зависимости, на

основе опорных функций – полиномов различной степени, комбинирующих разные сочетания входных переменных. Для каждой модели определяются коэффициенты методом регрессионного анализа [1, 2]. Среди всех моделей выбираются несколько наилучших. Лучшие из найденных моделей используются на следующей итерации метода как аргументы для формирования более сложных моделей. Качество моделей определяется ошибкой или коэффициентом детерминации, либо коэффициентом парной корреляции выходного признака и входных признаков. Если найдена приемлемая модель, то метод завершает работу. Особенностью метода является преобразуемость полученных моделей в полиномиальные нейронные сети – разновидность глубоких нейронных сетей. Достоинствами метода является его универсальность, пригодность для выборки малого объема, возможность решения задачи отбора информативных признаков в процессе построения модели, структурированность и иерархичность получаемых моделей и, как следствие, их интерпретабельность. Недостатками метода являются существенный рост сложности получаемых моделей с увеличением объема обучающих данных и требований к точности модели, а также увеличением числа используемых входных признаков, высокая итеративность и значительные затраты времени метода.

Общей особенностью всех перечисленных выше методов является то, что они стремятся построить во всем пространстве исходных признаков единую полиномиальную модель, коэффициенты которой подбираются путем решения оптимизационной задачи. Такой подход оказывается весьма затратным при большом числе признаков и экземпляров, а также сопряжен с проблемой выбора начальной точки поиска, задания оптимальной структуры и параметров сети.

Методы построения деревьев регрессии [7, 8] иерархически разбивают исходное пространство на области (кластеры), в которых находят среднее значение выхода для попавших в кластер экземпляров. Это значение присваивается выходному признаку всех экземпляров, которые попадут в данный кластер. Достоинством данной группы методов является простота и интерпретабельность получаемых моделей, а также возможность попутного решения задач кластер-анализа и отбора информативных признаков. Недостатками методов данной группы являются низкая точность получаемых моделей, а также неоднозначность в иерархической комбинации проверок для отнесения экземпляра к кластеру.

Методы построения прогнозирующих моделей на основе нейро-нечетких сетей [9–12] представляют модель зависимости комбинацией линейных регрессионных моделей зависимостей выходного признака от входных признаков для отдельных областей в пространстве признаков (кластеров), определяемых правилами, задаваемыми экспертами или получаемыми в результате кластер-анализа. Достоинством получаемых моделей является их интерпретабельность. Недостатками данных методов является то, что полу-

чаемые модели оказываются недостаточно точными, при большом числе исходных признаков модели получаются крайне сложными, а время на их построение также существенно возрастает. Кроме того, к недостаткам следует отнести зависимость методов данной группы от наличия экспертных знаний или необходимость проведения предварительного кластер-анализа.

Общей для деревьев решений и нейро-нечетких сетей особенностью построения численных моделей является то, что они вместо построения единой для всего пространства признаков модели строят совокупность иерархически получаемых в результате кластер-анализа линейных моделей. Достоинством получаемых моделей является их интерпретабельность. Недостатком получаемых моделей является их низкая точность.

Предложенный в [13] метод кластер-регрессионной аппроксимации комбинирует достоинства и устраняет недостатки рассмотренных выше групп методов. В исходном пространстве признаков выделяются кластеры, для которых метод строит частные модели. При этом в каждом кластере строится набор моделей, использующих наименьший по размеру набор наиболее информативных признаков для построения частной модели, а сами частные модели строятся как линейные или нейросетевые. Данный метод также имеет нейросетевую [13] и нейро-нечеткую интерпретации [14, 15].

Достоинством метода кластер-регрессионной аппроксимации является то, что он позволяет получать более точные модели, чем нейро-нечеткие и деревья решений, но при этом, в отличие от нейронных сетей, строит для каждого кластера отдельно частные модели, стремясь минимизировать число признаков для каждой частной модели, что позволяет получать более простые и интерпретабельные модели. Недостатком метода кластер-регрессионной аппроксимации [13] является то, что он нуждается в реализации кластер-анализа, в процессе которого использует весь набор исходных признаков.

Поэтому актуальной задачей является усовершенствование метода кластер-регрессионной аппроксимации путем косвенной реализации кластер-анализа в процессе построения модели.

3 МАТЕРИАЛЫ И МЕТОДЫ

Рассмотрим базовый метод кластер-регрессионной аппроксимации. Он разбивает выборку в пространстве признаков на компактные группы экземпляров – кластеры, для экземпляров каждого кластера строит частные регрессионные модели, стремясь минимизировать сложность модели и число используемых признаков. Однако полученная модель все же может оказаться избыточной. Поэтому предлагается после построения модели упростить ее посредством контрастирования весов модели. Формально данный метод может быть записан следующим образом.

1. Этап выделения кластеров. На основе экземпляров заданной обучающей выборки $\langle x, y \rangle$ с помощью

заданного метода кластер-анализа [16, 17] выделить кластеры в пространстве признаков, определив координаты их центров $\{C^q\}$, $C^q = \{C_j^q\}$, $q = 1, 2, \dots, Q$, $j = 1, 2, \dots, N$.

2. Этап кластер-анализа обучающей выборки. Определить принадлежность каждого s -го экземпляра x^s выборки $\langle x, y \rangle$ к кластерам, $s = 1, 2, \dots, S$:

– найти расстояния от экземпляра x^s до центра каждого кластера $R(x^s, C^q)$ в метрике соответствующего метода кластер-анализа;

– отнести экземпляр x^s к тому кластеру q^s , расстояние до центра которого от экземпляра является наименьшим:

$$q^s = \arg \min_{q=1,2,\dots,Q} \{R(x^s, C^q)\},$$

где

$$R(x^s, C^q) = \sqrt{\sum_{j=1}^N (x_j^s - C_j^q)^2}.$$

3. Этап построения частных моделей для кластеров. Для экземпляров q -го кластера, $q = 1, 2, \dots, Q$:

– если в q -й кластер попал только один экземпляр, то принять его как синглтон, задав частную модель-функцию кластера как константу:

$$f^q = \{x^s \exists! s : q^s = q\};$$

– если в q -й кластер попало более одного экземпляра, то оценить индивидуальную информативность каждого j -го признака по отношению к выходному признаку y для экземпляров q -го кластера $I^q(x_j, y)$

(например, на основе модуля коэффициента парной корреляции [2]), построить для экземпляров q -го кластера одномерную линейную регрессионную модель зависимости выходного признака y от индивидуально наиболее информативного входного признака

$$f^q = w_1^q x_j^s + w_0^q$$

с помощью метода наименьших квадратов [1, 2], оценить ошибку полученной частной q -й модели E^q , если ошибка E^q является приемлемой (например, если $E^q(S^q/S) \leq \varepsilon$, где ε – заданное максимально допустимое значение ошибки для всей выборки), то перейти к следующему кластеру, в противном случае: построить для экземпляров q -го кластера многомерную линейную регрессионную модель зависимости выходного признака от всего набора входных признаков [1, 2]:

$$f^q = w_0^q + \sum_{j=1}^N w_j^q x_j^s,$$

оценить ошибку полученной частной q -й модели E^q , если ошибка E^q является приемлемой (например, если $E^q(S^q/S) \leq \varepsilon$), то перейти к следующему кластеру, в про-

тивном случае: построить для экземпляров q -го кластера многомерную нелинейную модель зависимости выходного признака от всего набора входных признаков на основе однослойного персептрона [18]:

$$f^q = \psi^q \left(w_0^q + \sum_{j=1}^N w_j^q x_j^s \right)$$

(при этом возможно придется нормировать выходной параметр, отображая его значения на интервал значений функции ψ^q), используя для обучения метод Уидроу-Хоффа [18], оценить ошибку полученной частной q -й модели E^q , если ошибка E^q является приемлемой (например, если $E^q(S^q/S) \leq \varepsilon$), то перейти к следующему кластеру, в противном случае: построить для экземпляров q -го кластера частную многомерную нелинейную модель зависимости выходного признака y от всего набора входных признаков на основе двухслойного персептрона [3, 4]:

$$f^q = \psi^{q,(2,1)} \left(w_0^{q,(2,1)} + \sum_{i=1}^{N_{q,1}} w_i^{q,(2,1)} \psi_i^{q,(2,1)} \left(w_0^{q,(1,i)} + \sum_{j=1}^N w_j^{q,(1,i)} x_j^s \right) \right),$$

обучаемого на основе метода Левенберга-Марквардта [19] и техники обратного распространения ошибки [20], оценить ошибку полученной частной q -й модели E^q , если ошибка E^q является приемлемой (например, если $E^q(S^q/S) \leq \varepsilon$), то перейти к следующему кластеру, в противном случае либо продолжить по аналогии наращивать число слоев в частной модели, либо принять как частную модель q -го кластера наиболее точную из построенных частных моделей.

4. Этап упрощения моделей. Для всех кластеров, для которых построены многомерные регрессионные модели, последовательно перебрать сочетания признаков, удаляя k наименее индивидуально информативных признаков ($k = 1, 2, \dots, N-1$), установив их веса в частной модели q -го кластера равными нулю, до тех пор, пока ошибку соответствующей частной модели остается приемлемой. Для всех кластеров, для которых построены многомерные нейромодели, выполнить контрастирование весов [21], до тех пор, пока ошибки частных нейромоделей являются приемлемыми. Если некоторые из исходных признаков исключены из всех частных моделей, то удалить их из выборки.

5. Этап синтеза модели. На основе полученного набора центров кластеров и построенных частных моделей синтезировать кластер-регрессионную [13], нейросетевую [13] или нейро-нечеткую модель [14, 15], либо дерево регрессии специального вида, описанным ниже методом.

Метод синтеза дерева регрессии на основе построенной кластер-регрессионной аппроксимации предполагает, что каждый лист дерева (узел, не имеющий потомков) считается кластером. При этом, в отличие от известных методов построения деревьев регрессии, где листья содержат лишь среднее значение выходно-

го признака для попавших в них экземпляров, в предлагаемом методе в каждом листе находится функция частной модели. Предлагаемый метод состоит из следующих этапов.

1. Этап формирования псевдовыборки. Сформировать псевдоэкземпляры на основе центров выделенных кластеров $\{C^q\}$, приняв за координаты входных признаков псевдоэкземпляров координаты центров кластеров $C^q = \{C_j^q\}$, а в качестве выхода – среднее значение выхода частной модели соответствующего кластера при подаче на ее входы координат центра кластера:

$$\bar{y}^q = \frac{1}{S^q} \sum_{s=1}^{S^q} \{y^s \mid q^s = q\}.$$

2. Этап построения дерева решения. Для выборки псевдоэкземпляров $\langle C, \bar{y} \rangle$, $C = \{C^q\}$, $\bar{y} = \{\bar{y}^q\}$, построить деревья регрессии на основе известных методов построения деревьев регрессии [7, 8]. Отобрать из них наилучшую в смысле точности модель – дерево, обеспечивающее наименьшую ошибку.

3. Этап задания частных моделей. Относительно построенного наилучшего дерева распознать экземпляры обучающей выборки, определив их попадание в листовые узлы дерева. Для экземпляров каждого соответствующего q -го листа-кластера задать частную модель f^q , выполнив третий этап описанного выше метода кластер-регрессионной аппроксимации.

4. Этап редукции дерева регрессии. Просматривая полученное дерево снизу вверх (от листов к корню):

- если текущий q -й узел является листом, то пропустить его, перейдя к следующему узлу;
- если текущий q -й узел не является листом и один из его потомков также не является листом, то пропустить его, перейдя к следующему узлу;
- если текущий q -й узел не является листом, а все его потомки являются листьями, то для экземпляров узлов-потомков данного узла построить частную регрессионную модель f^q подобно третьему этапу описанного выше метода кластер-регрессионной аппроксимации и оценить ее ошибку E^q для экземпляров данных узлов. Если модель f^q обеспечивает приемлемую ошибку ($E^q(S^q/S) \leq \varepsilon$), то внести ее в текущий q -й узел, а узлы потомки q -го узла удалить и далее считать q -й узел листом.

Построенное дерево можно использовать как самостоятельную модель, а также для построения нейросетевых моделей на основе кластер-регрессионной аппроксимации.

Метод построения нейромодели на основе дерева регрессии можно представить следующим образом.

1. Этап кластеризации выборки. Распознать заданную обучающую выборку $\langle x, y \rangle$ по построенному дереву регрессии и определить принадлежность q^s каждого экземпляра x^s к узлам-листам (кластерам) дерева.

2. Этап определения центров кластеров. Для экземпляров каждого q -го кластера (листа дерева) найти координаты его центра $C^q = \{C_j^q\}$:

$$C_j^q = \frac{1}{S^q} \sum_{s=1}^{S^q} \{x_j^s \mid q^s = q\}, j = 1, 2, \dots, N, q = 1, 2, \dots, Q.$$

3. Этап построения модели. На основе найденных координат центров кластеров $\{C^q\}$ и частных моделей узлов-кластеров $\{f^q\}$ построить кластер-регрессионную модель в виде нейронной [13] или нейро-нечеткой сети [14, 15].

Предложенный комплекс методов позволяет на основе заданной выборки синтезировать кластер-регрессионную модель, на основе которой построить дерево регрессии, которое также может быть преобразовано в кластер-регрессионную модель.

4 ЭКСПЕРИМЕНТЫ

Выбор модели для конкретной задачи определяется наличием у нее свойств, важных для пользователя.

Для сравнения моделей рассмотрим наиболее важные их свойства и определим показатели, количественно их характеризующие.

Ключевым показателем свойств модели является ее ошибка E . Также модель характеризуется числом настраиваемых параметров N_w , и числом используемых признаков N^r .

Известные из литературы [22–24] информационные критерии качества моделей такие, как критерий Ханнан-Квинн (Hannan-Quinn Criterion), информационный критерий Байеса (Bayesian Information Criterion), минимум длины описания (Minimum Description Length), кратчайшее описание данных (Shortest Data Description), информационный критерий Акайке (AIC), скорректированный AIC (Corrected AIC), несмещенный AIC (Unbiased AIC), состоятельный AIC (Consistent AIC), критерий Маллоу (Mallow Criterion), зависят от ошибки модели, размерности обучающей выборки (числа экземпляров и числа признаков), числа настраиваемых параметров модели и предельно допустимого числа параметров модели [25–27].

Поскольку при сравнении моделей, как правило, предполагается одинаковость выборок, то объем выборок целесообразно исключить из критериев сравнения. Поскольку деревья регрессии и нейросети являются графами, то при определении числа настраиваемых параметров таких моделей представляется целесообразным принять во внимание тот факт, что связь считается отсутствующей, если ее вес равен нулю, таким образом, следует из числа настроенных параметров модели исключить веса, равные нулю. В то же время различные синтезированные модели могут использовать не все из представленных в выборке признаков. Поэтому число признаков, используемых в модели, следует рассматривать как важную особенность при их сравнении. На основе данных соображений определим интегральный информационный критерий:

$$I_{IC} = \left(1 - \frac{N^r (N_w - N_{w=0})}{N N_w^{\max}} \right) e^{-E},$$

$$N > 0, N' \leq N, N_w^{\max} > 0, \\ N_w \leq N_w^{\max}, N_{w=0} \leq N_w, E \geq 0.$$

Критерий I_{IC} будет принимать значения в диапазоне от нуля до единицы. Чем меньше будет его значение, тем хуже модель, а чем больше – тем лучше.

Уверенность в принимаемом решении – субъективная оценка моделью принятого решения.

Для кластер-регрессионной аппроксимации уверенность q -й частной модели может характеризоваться расстоянием распознаваемого s -го экземпляра до центра кластера q -й частной модели, а также мгновенной ошибкой (если имеется проверочная выборка):

$$I_{cert}^q(x^s) = \left(e^{-R(x^s, C^q)^2} - e^{-\min_{\substack{p=1,2,\dots,Q \\ p \neq q}} \{R(x^s, C^p)^2\}} \right) \cdot e^{-(y^s - y_q^{s*})^2}.$$

Применительно ко всей модели и всей выборке показатель уверенности модели в принимаемом решении может быть определен следующим образом:

$$I_{cert} = \frac{1}{S} \sum_{s=1}^S \max_{q=1,2,\dots,Q} \{I_{cert}^q(x^s)\} = \\ = \frac{1}{S} \sum_{s=1}^S \max_{q=1,2,\dots,Q} \left\{ \left(e^{-R(x^s, C^q)^2} - e^{-\min_{\substack{p=1,2,\dots,Q \\ p \neq q}} \{R(x^s, C^p)^2\}} \right) \cdot e^{-(y^s - y_q^{s*})^2} \right\}.$$

При отсутствии информации о фактическом значении выходного признака показатель уверенности будет определяться формулой:

$$I_{cert} = \frac{1}{S} \sum_{s=1}^S \max_{q=1,2,\dots,Q} \left\{ \left(e^{-R(x^s, C^q)^2} - e^{-\min_{\substack{p=1,2,\dots,Q \\ p \neq q}} \{R(x^s, C^p)^2\}} \right) \right\}.$$

Показатели субъективной уверенности будут принимать значения в интервале $[0, 1]$: чем выше будет значение показателя, тем ближе по свойствам распознаваемый экземпляр к сформированным сетью эталонам и тем больше сеть уверена в принимаемом решении.

Обобщение – способность модели интегрировать частные данные для определения закономерностей и пролонгации результатов, то есть после обучения на основе обучающей выборки выдавать ответы для экземпляров тестовой выборки, подобных экземплярам обучающей выборки, но не входивших в нее.

Коэффициент обобщения построенной модели определим по формуле:

$$I_{gen} = \frac{NS}{N_w N_n} e^{-E} \text{ либо } I_{gen} = \left\{ \frac{NS}{N_w N_n} \mid E \leq \varepsilon \right\},$$

где $N > 0, N' \leq N, N_w \geq 1, N_n \geq 1, E \geq 0$.

Если коэффициент обобщения будет существенно больше единицы, то модель проявляет большие обобщающие способности, если же коэффициент обобщения существенно меньше единицы, то модель не проявляет обобщающих свойств.

Для удобства сравнения различных моделей определим масштабированный коэффициент обобщения:

$$I_{gen}^* = \frac{e^{-E}}{1 + e^{\frac{NS}{N_w N_n}}}.$$

Масштабированный коэффициент обобщения будет принимать значения в диапазоне от нуля до единицы: чем больше будет его значение, тем более высокие обобщающие свойства проявляет модель, а также обеспечивает меньшую ошибку.

Логическая прозрачность (интерпретабельность) модели – свойство модели быть удобной для восприятия и анализа человеком. Очевидно, что модель, которая содержит как можно меньше функциональных узлов и связей является более простой, но, с другой стороны, при фиксированном числе узлов и связей модель будет тем удобнее для восприятия и анализа, чем более иерархической она является и чем меньше входов у каждого узла.

Интерпретабельность узла будем определять по формуле:

$$I_{interp}^{(\eta,i)} = \begin{cases} \prod_{j=1}^{N^{(\eta,i)}} I_{interp}^{(\eta-1,j)} \\ \left(N^{(\eta,i)} + N_w^{(\eta,i)} \right) \left(n^{(\eta,i)} + 1 \right), \eta > 0; \\ 1, \eta = 0, \\ N^{(\eta,i)} > 0, N_w^{(\eta,i)} \geq 0, n^{(\eta,i)} > 0. \end{cases}$$

где $n^{(\eta,i)}$ для деревьев и кластер-регрессионной аппроксимации определим порядком полинома, реализуемого узлом, а для нейронов – суммой порядков полиномов весовой и активационной функций.

Интерпретабельность модели оценим как интерпретабельность итогового узла модели (для нейронной сети слои будем последовательно нумеровать от входного к выходному, а для дерева – наоборот, от листьев к корню):

$$I_{interp} = I_{interp}^{(M,1)}.$$

Предложенный показатель интерпретабельности будет принимать значения в диапазоне $(0, 1]$: чем больше будет его значение, тем более удобной для восприятия будет модель.

На основе введенных показателей определим интегрированный показатель качества модели как мультипликативный показатель:

$$I_Q = I_{gen}^* I_{cert} I_{interp} I_{IC},$$

либо как аддитивный:

$$I_Q = \frac{I_{gen}^* + I_{cert} + I_{interp} + I_{IC}}{4}.$$

Значение показателя качества модели будет находиться в диапазоне от нуля до единицы. Чем больше будет значение данного показателя, тем лучше будет модель.

Предложенные в работе методы, а также показатели, характеризующие их качество, были программно реализованы и экспериментально исследованы при решении задачи моделирования зависимости коэффициента упрочнения лопаток газотурбинных авиадвигателей [28], для которой имелось 63 экземпляра, характеризовавшихся 9 входными и одним выходным признаками.

5 РЕЗУЛЬТАТЫ

Результаты проведенных экспериментов представлены в табл. 1.

Как видно из табл. 1, предложенные показатели позволяют сравнивать качество регрессионных, нейросетевых и кластер-регрессионных моделей, а также моделей на основе деревьев регрессии.

Наиболее низким уровнем ошибки обладают кластер-регрессионная модель и модель на основе дерева регрессии, синтезированная на основе кластер-регрессионной аппроксимации.

Наиболее высоким уровнем информационного критерия обладает дерево регрессии на основе кластер-регрессионной модели.

Наиболее высоким уровнем интерпретабельности обладают многомерная линейная регрессионная модель и однослойный перцептрон.

Кластер-регрессионная модель и модель дерева регрессии, синтезированная на ее основе, обладают сравнимым уровнем уверенности. Для остальных мо-

делей данный показатель не определен и принят равным нулю.

Наиболее высоким уровнем обобщения данных обладает однослойный перцептрон.

Наиболее высоким уровнем показателя качества обладает дерево регрессии на основе кластер-регрессионной модели.

6 ОБСУЖДЕНИЕ

Результаты проведенных экспериментов подтвердили работоспособность разработанных методов и реализующих их программных средств.

В отличие от традиционных методов построения регрессионных моделей [1, 2], строящих модель на основе функции, единой для всего признакового пространства формы, предложенный метод формирует иерархическую комбинацию частных моделей.

В отличие от известных методов построения деревьев регрессии [7, 8], листовые узлы которых содержат усредненные значения выходного признака для кластеров, предложенный метод формирует дерево, состоящее из частных моделей для кластеров, что позволяет обеспечить большую точность модели.

В отличие от традиционных методов построения нейросетевых моделей на основе слоистых сетей прямого распространения [3, 4], строящих единую модель для всего признакового пространства, предложенный метод формирует иерархическую комбинацию частных моделей.

При этом модели, синтезированные на основе предложенного метода, обеспечивают приемлемую точность решения задачи, высокую интерпретабельность, высокий уровень обобщения данных, высокую уверенность в принимаемом решении, а по совокупности свойств обеспечивают наиболее высокий уровень качества модели.

Это позволяет рекомендовать предложенные методы и реализующее их программное обеспечение для использования на практике при решении задач построения моделей количественных зависимостей по прецедентам.

Таблица 1 – Результаты экспериментов

Вид модели	N'	N_w	E	I_{IC}	I_{cert}	I_{gen}^*	I_{interp}	I_Q
Многомерная линейная регрессия (полином первого порядка)	9	10	0,2730	0,0000	0	0,7611	0,0053	0,1916
Многомерная линейная регрессия (полином второго порядка)	9	55	0,0211	0,0000	0	0,9791	0,0016	0,2452
Многомерная линейная регрессия (полином третьего порядка)	9	220	0,0024	0,7119	0	0,9268	0,0004	0,4098
Дерево регрессии	9	138	0,2894	0,1302	0	0,3994	0,0007	0,1326
Однослойный перцептрон	9	10	0,0704	0,0000	0	0,7703	0,0053	0,1939
Двухслойная нейронная сеть прямого распространения сигнала	9	56	0,0149	0,0000	0	0,8281	0,0015	0,2074
Трехслойная нейронная сеть прямого распространения сигнала	9	65	0,0023	0,0000	0	0,8488	0,0014	0,2125
Кластер-регрессионная модель	7	53	0,0012	0,6720	0,9999	0,8934	0,0021	0,6418
Дерево регрессии на основе кластер-регрессионной модели	7	24	0,0012	0,8508	0,9999	0,9656	0,0040	0,7051

ВЫВОДЫ

В работе рассмотрена задача автоматизации построения моделей количественных зависимостей по прецедентам.

Научная новизна результатов работы состоит в том, что:

– впервые предложен метод древовидной кластер-регрессионной аппроксимации, который для заданной обучающей выборки строит дерево для иерархической кластеризации экземпляров, листовые узлы которого соответствуют кластерам, для каждого кластера строит частную модель зависимости по экземплярам обучающей выборки, попавшим в кластер, стремясь обеспечить наименьшую сложность модели, и использует набор наиболее информативных признаков наименьшей длины. Это позволяет обеспечить приемлемую точность модели, высокие уровни интерпретабельности и обобщения данных, снизить сложность модели, упростить ее реализацию при последовательной организации вычислений;

– предложен комплекс показателей, позволяющих количественно характеризовать такие свойства как уверенность в принимаемом решении, обобщение данных, интерпретабельность, качество для моделей различных видов: нейронных и нейро-нечетких сетей, регрессионных моделей, деревьев регрессии и кластер-регрессионных моделей. Предложенный комплекс позволяет сравнивать различные модели зависимостей, а также формировать на его основе критерии качества моделей, которые можно использовать при обучении и упрощении, а также выборе моделей

Практическая ценность полученных результатов заключается в том, что разработано программное обеспечение, реализующее предложенные метод древовидной кластер-регрессионной аппроксимации и показатели качества моделей. Разработанный метод и реализующее его программное обеспечение исследованы при решении практических задач прогнозирования. Проведенные эксперименты подтвердили работоспособность разработанного математического обеспечения и позволяют рекомендовать его для использования на практике.

Перспективы дальнейших исследований заключаются в том, чтобы опробовать предложенный метод на более широком классе прикладных задач, исследовать зависимости скорости и точности (ошибки) его работы от объема используемых выборок и числа признаков в исходной выборке.

БЛАГОДАРНОСТИ

Работа выполнена на кафедре программных средств Запорожского национального технического университета в рамках научно-исследовательской темы «Разработка и исследование интеллектуальных методов и программных средств диагностирования и неразрушающего контроля качества техники военного и гражданского назначения» (номер гос. регистрации 0119U100360) при частичной поддержке проектов программы «Erasmus+» «Internet of Things: Emerging»

© Субботин С. А., 2019

DOI 10.15588/1607-3274-2019-3-9

«Curriculum for Industry and Human Applications» (ALIOT, Ref. No 573818-EPP-1-2016-1-UK-EPPKA2-CBHE-JP) «Innovative Multidisciplinary Curriculum in Artificial Implants for Bio-Engineering BSc/MSc Degrees» (BIOART, Ref. No 586114-EPP-1-2017-1-ES-EPPKA2-CBHE-JP), финансируемых Европейским Союзом.

ЛИТЕРАТУРА / ЛІТЕРАТУРА

1. Айвазян С. А. Прикладная статистика: Исследование зависимостей / С. А. Айвазян, И. С. Енюков, Л. Д. Мешалкин. – М. : Финансы и статистика, 1985. – 487 с.
2. Афифи А. Статистический анализ: подход с использованием ЭВМ / А. Афифи, С. Эйзен. – М. : Мир, 1982. – 488 с.
3. Computational intelligence: a methodological introduction / [R. Kruse, C. Borgelt, F. Klawonn et. al.]. – London : Springer-Verlag, 2013. – 488 p.
4. Intelligent hybrid systems: fuzzy logic, neural networks, and genetic algorithms / ed. D. Ruan. – Berlin : Springer, 2012. – 374 p.
5. Ивахненко А. Г. Моделирование сложных систем по экспериментальным данным / А. Г. Ивахненко, Ю. П. Юрачковский. – М. : Радио и связь, 1987. – 118 с.
6. Madala H. R. Inductive Learning Algorithms for Complex Systems Modeling / H. R. Madala, A. G. Ivakhnenko. – Boca Raton : CRC Press, 1994. – 384 p.
7. Clarke B. Principles and theory for data mining and machine learning / B. Clarke, E. Fokoue, H. H. Zhang. – New York : Springer, 2009. – 781 p.
8. Classification and regression trees / [L. Breiman, J. H. Friedman, C. J. Stone, R. A. Olshen]. – Boca Raton: Chapman & Hall / CRC, 1984. – 368 p.
9. Rutkowski L. Flexible neuro-fuzzy systems : structures, learning and performance evaluation / L. Rutkowski. – Boston: Kluwer, 2004. – 276 p.
10. Liu P. Fuzzy neural network theory and application / P. Liu, H. Li. – Singapore: World Scientific, 2004. – 376 p. – (Series in Machine Perception and Artificial Intelligence ; vol. 59).
11. Buckleya J. J. Fuzzy neural networks: a survey / J. J. Buckley, Y. Hayashi // Fuzzy sets and systems. – 1994. – Vol. 66, Issue 1. – P. 1–13.
12. Jang J.R. Neuro-fuzzy and soft computing: a computational approach to learning and machine intelligence / J. R. Jang, C.-T. Sun, E. Mizutani. – Upple Saddle River : Prentice-Hall, 1997. – 614 p.
13. Субботин С. О. Алгоритми кластер-регресійної апроксимації та їх нейромеревеві інтерпретації / С. О. Субботин // Радіоелектроніка, інформатика, управління. – 2003. – № 1. – С. 114–121.
14. Субботин С. А. Метод синтеза нейро-нечетких аппроксиматоров / С. А. Субботин // Автоматизация и современные технологии. – 2007. – № 11. – С. 14–18.
15. Субботин С. А. Нейро-нечеткая кластер-регрессионная аппроксимация по обобщенной оси / С. А. Субботин // Нейрокомпьютеры: разработка, применение. – 2009. – № 8. – С. 52–62.
16. Berkhin P. Knowledge discovery: clustering / P. Berkhin, I. S. Dhillon // Encyclopedia of complexity and systems science / ed. R. A. Meyers. – Berlin : Springer, 2009. – P. 5051–5064.

17. Abonyi J. Cluster analysis for data mining and system identification / J. Abonyi, B. Feil. – Basel : Birkhäuser, 2007. – 303 p.
18. Widrow B. 30 years of adaptive neural networks: perceptron, madaline, and backpropagation / B. Widrow, M. A. Lehr // Proceedings of the IEEE. – Vol. 78, Issue 9. – P. 1415–1442. DOI:10.1109/5.58323
19. Ravindran A. Engineering optimization: methods and applications / A. Ravindran, K. M. Ragsdell, G. V. Reklaitis. – New Jersey: John Wiley & Sons, 2006. – 688 p.
20. Rumelhart D. E. Learning representations by backpropagating errors / D. E. Rumelhart, G. E. Hinton, R. J. Williams // Nature. – 1986. – Vol. 323. – P. 533–536. DOI:10.1038/323533a0
21. Gorban A. N. Generation of Explicit Knowledge from Empirical Data through Pruning of Trainable Neural Networks / A. N. Gorban, Eu. M. Mirkes, V. G. Tsaregorodtsev // International Joint Conference on Neural Networks (IJCNN'99), Washington, July 1999 : proceedings. – Los Alamitos: IEEE, 1999. – Vol. 6. – P. 4393–4398.
22. Akaike H. A new look at the statistical model identification / H. Akaike // IEEE transactions on automatic control. – 1974. – Vol. 19, № 6. – P. 716–723.
23. Schwarz G. E. Estimating the dimension of a model / G. E. Schwarz // Annals of statistics. – 1978. – Vol. 6, № 2. – P. 461–464.
24. Hannan E. J. The determination of the order of an autoregression / E. J. Hannan, B. G. Quinn // Journal of the Royal Statistical Society. – 1979. – Ser. B., Vol. 41. – P. 190–195.
25. Субботин С. А. Анализ свойств и критерии сравнения нейросетевых моделей для решения задач диагностики и распознавания образов / С. А. Субботин // Реєстрація, зберігання і обробка даних. – 2009. – Т. 11, № 3. – С. 42–52.
26. Субботин С. А. Модели критериев сравнения нейронных и нейро-нечетких сетей в задачах диагностики и классификации образов / С. А. Субботин // Наукові праці Донецького національного технічного університету. Серія «Інформатика, кібернетика та обчислювальна техніка». – Донецьк : ДНТУ, 2010. – Вип.12 (165). – С.148–151.
27. Субботин С. А. Методика и критерии сравнения моделей и алгоритмов синтеза искусственных нейронных сетей / С. А. Субботин // Радіоелектроніка, інформатика, управління. – 2003. – № 2. – С. 109–114.
28. Прогрессивные технологии моделирования, оптимизации и интеллектуальной автоматизации этапов жизненного цикла авиационных двигателей : монография / [А. В. Богуслаев, Ал. А. Олейник, Ан. А. Олейник, Д. В. Павленко, С. А. Субботин ; под ред. Д. В. Павленко, С. А. Субботина]. – Запорожье : ОАО «Мотор Сич», 2009. – 468 с.

Статья поступила в редакцию 29.03.2019.
После доработки 26.06.2019.

УДК 004.93

МЕТОДИ СИНТЕЗУ МОДЕЛЕЙ КІЛЬКІСНИХ ЗАЛЕЖНОСТЕЙ В БАЗИСІ ДЕРЕВ РЕГРЕСІЇ, ЩО РЕАЛІЗУЮТЬ КЛАСТЕР-РЕГРЕСІЙНУ АПРОКСИМАЦІЮ ЗА ПРЕЦЕДЕНТАМИ

Субботін С. О. – д-р техн. наук, професор, завідувач кафедри програмних засобів Національного університету «Запорізька політехніка», Запоріжжя, Україна.

АНОТАЦІЯ

Актуальність. Для прийняття рішень в технічних застосуваннях, як правило, необхідно мати модель, що дозволяє прогнозувати стан керованого об'єкта або процесу. Об'єктом дослідження є процес побудови моделей залежностей за спостереженнями. Предметом дослідження є методи побудови кількісних залежностей за спостереженнями на основі кластер-регресійної апроксимації.

Мета роботи – спрощення моделей кластер-регресійної апроксимації шляхом непрямої реалізації кластер-аналізу в процесі побудови моделі.

Метод. Запропоновано метод деревовидної кластер-регресійної апроксимації, який для заданої навчальної вибірки буде дерево для ієрархічної кластеризації екземплярів, листові вузли якого відповідають кластерам, для кожного кластера буде часткову модель залежності за екземплярами навчальної вибірки, що потрапили у кластер, прагнучи забезпечити найменшу складність моделі і використовує набір найбільш інформативних ознак найменшої довжини. Це дозволяє забезпечити прийнятну точність моделі, високі рівні інтерпретабельності й узагальнення даних, знизити складність моделі, спростити її реалізацію при послідовній організації обчислень.

Результати. Розроблено програмне забезпечення, що реалізує запропонований метод деревовидної кластер-регресійної апроксимації. Розроблений метод і програмне забезпечення, що його реалізує, досліджені під час вирішення практичних завдань прогнозування. Проведені експерименти підтвердили працездатність розробленого математичного забезпечення і дозволяють рекомендувати його для використання на практиці.

Висновки. На відміну від традиційних методів побудови регресійних моделей, які будують модель на основі функції, єдиної форми для усього простору ознак, запропонований метод формує ієрархічну комбінацію часткових моделей. На відміну від відомих методів побудови дерев регресії, листові вузли яких містять усереднені значення вихідної ознаки для кластерів, запропонований метод формує дерево, що складається з часткових моделей для кластерів, що дозволяє забезпечити більшу точність моделі.

КЛЮЧОВІ СЛОВА: моделювання, кластер-регресійна апроксимація, оцінювання, дерева рішень, регресійний аналіз.

UDC 004.93

METHODS OF SYNTHESIS OF MODELS OF QUANTITATIVE DEPENDENCIES IN THE BASIS OF TREES OF REGRESSION, REALIZING CLUSTER-REGRESSION APPROXIMATION BY PRECEDENTS

Subbotin S. A. – Dr. Sc., Professor, Head of the Department of Software Tools of National University «Zaporizhzhia Polytechnic», Zaporizhzhia, Ukraine.

ABSTRACT

Context. To make decisions in technical applications, it is usually necessary to have a model that allows you to predict the state of a managed object or process. The object of the study is the process of building dependency models by use cases. The subject of the study are the methods for constructing quantitative dependencies based on cluster-regression approximation precedents.

Objective. The aim of the paper is to simplify cluster regression approximation models by indirectly implementing cluster analysis in the process of model building.

Method. A tree-based cluster-regression approximation method is proposed which, for a given training sample, constructs a tree for hierarchical clustering of instances whose leaf nodes correspond to clusters, for each cluster, constructs a particular model of dependence on instances of the training sample that fall into the cluster, in order to provide the least complexity of the model and uses the set the most informative features of the shortest length. This allows to ensure an acceptable accuracy of the model, high levels of interpretation and generalization of data, to reduce the complexity of the model, and to simplify its implementation in the sequential organization of calculations.

Results. The software that implements the proposed method of tree-like cluster-regression approximation is developed. The developed method and the software implementing it are investigated in solving practical problems of prediction. The conducted experiments confirmed the working capacity of the developed software and allow to recommend it for use in practice.

Conclusions. Unlike traditional methods of regression model constructing that build a model based on a function form that is uniform for the entire feature space, the proposed method forms a hierarchical combination of particular models. Unlike the well-known methods of regression tree constructing whose leaf nodes contain averaged values of the output feature for clusters, the proposed method forms a tree consisting of particular models for clusters, which allows to ensure greater accuracy of the model.

KEYWORDS: modeling, cluster-regression approximation, estimation, decision trees, regression analysis.

REFERENCES

1. Ajvazjan S. A., Enjukov I. S., Meshalkin L. D. *Prikladnaja statistika: Issledovanie zavisimostej*. Moscow, Finansy i statistika, 1985, 487 p.
2. Afifi A., Jeizen S. *Statisticheskij analiz: podhod s ispol'zovaniem EVM*. Moscow, Mir, 1982, 488 p.
3. Kruse R., Borgelt C., Klawonn F. et. al. *Computational intelligence: a methodological introduction*. London, Springer-Verlag, 2013, 488 p.
4. *Intelligent hybrid systems: fuzzy logic, neural networks, and genetic algorithms*, ed. D. Ruan. Berlin, Springer, 2012, 374 p.
5. Ivahnenko A. G., Jurachkovskij Ju. P. *Modelirovanie slozhnyh sistem po jeksperimental'nym dannym*. Moscow, Radio i svjaz', 1987, 118 s.
6. Madala H. R., Ivakhnenko A. G. *Inductive Learning Algorithms for Complex Systems Modeling*. Boca Raton, CRC Press, 1994, 384 p.
7. Clarke B., Fokoue E., Zhang H. H. *Principles and theory for data mining and machine learning*. New York, Springer, 2009, 781 p.
8. Breiman L., Friedman J. H., Stone C. J., Olshen R. A. *Classification and regression trees*. Boca Raton, Chapman & Hall, CRC, 1984, 368 p.
9. Rutkowski L. *Flexible neuro-fuzzy systems : structures, learning and performance evaluation*. Boston, Kluwer, 2004, 276 p.
10. Liu P., Li H. *Fuzzy neural network theory and application*. Singapore, World Scientific, 2004, 376 p. (Series in Machine Perception and Artificial Intelligence ; vol. 59).
11. Buckley J. J., Hayashi Y. *Fuzzy neural networks: a survey*, *Fuzzy sets and systems*, 1994, Vol. 66, Issue 1, pp. 1–13.
12. Jang J. R., Sun C.-T., Mizutani E. *Neuro-fuzzy and soft computing: a computational approach to learning and machine intelligence*. Upple Saddle River, Prentice-Hall, 1997, 614 p.
13. Subbotin S. O. *Algoritmy klaster-regresijnoi aproksimacii ta jih nejromerezhevi interpretacii*, *Radio Electronics, Computer Science, Control*, 2003, No. 1, pp. 114–121.
14. Subbotin S. A. *Metod sinteza nejro-nechjotkih approksimirov, Avtomatizacija i sovremennye tehnologii*, 2007, No. 11, pp. 14–18.
15. Subbotin S. A. *Nejro-nechjotkaja klaster-regressionnaja approksimacija po obobshhjonnoj osi*, *Nejrokomputery: razrabotka, primenenie*, 2009, No. 8, pp. 52–62.
16. Berkhin P., Dhillon I. S. ed. *R. A. Meyers Knowledge discovery: clustering*, *Encyclopedia of complexity and systems science*. Berlin, Springer, 2009, pp. 5051–5064.
17. Abonyi J., Feil B. *Cluster analysis for data mining and system identification*. Basel, Birkhäuser, 2007, 303 p.
18. Widrow B., Lehr M. A. 30 years of adaptive neural networks: perceptron, madaline, and backpropagation, *Proceedings of the IEEE*, Vol. 78, Issue 9, pp. 1415–1442. DOI:10.1109/5.58323
19. Ravindran A., Ragsdell K. M., Reklaitis G. V. *Engineering optimization: methods and applications*. New Jersey, John Wiley & Sons, 2006, 688 p.
20. Rumelhart D. E., Hinton G. E., Williams R. J. Learning representations by back-propagating errors, *Nature*, 1986, Vol. 323, P. 533–536. DOI:10.1038/323533a0
21. Gorban A. N., Mirkes Eu. M., Tsaregorodtsev V. G. Generation of Explicit Knowledge from Empirical Data through Pruning of Trainable Neural Networks, *International Joint Conference on Neural Networks (IJCNN'99)*, Washington, July 1999 : proceedings. Los Alamitos, IEEE, 1999, Vol. 6. pp. 4393–4398.
22. Akaike H. A new look at the statistical model identification, *IEEE transactions on automatic control*, 1974, Vol. 19, № 6, pp. 716–723.
23. Schwarz G. E. Estimating the dimension of a model, *Annals of statistics*, 1978, Vol. 6, No. 2, pp. 461–464.
24. Hannan E. J., Quinn B. G. The determination of the order of an autoregression, *Journal of the Royal Statistical Society*, 1979, Ser. B, Vol. 41, pp. 190–195.
25. Subbotin S. A. *Analiz svojstv i kriterii sravnenija nejrosetevyh modelej dlja reshenija zadach diagnostiki i raspoznavanija obrazov*, *Reestracija, zberigannja i obrobka danih*, 2009, Vol. 11, No. 3, pp. 42–52.
26. Subbotin S. A. *Modeli kriteriev sravnenija nejronnyh i nejro-nechjotkih setej v zadachah diagnostiki i klassifikacii obrazov*, *Naukovi praci Donec'kogo nacional'nogo tehničnogo universytetu. Serija «Informatyka, kibernetyka ta obchysljuval'na tehnika»*. Donec'k, DNTU, 2010, Vyp. 12 (165), pp. 148–151.
27. Subbotin S. A. *Metodika i kriterii sravnenija modelej i algoritmov sinteza iskusstvennyh nejronnyh setej*, *Radio Electronics, Computer Science, Control*, 2003, No. 2, pp. 109–114.
28. Boguslaev A. V., Olejnik Al. A., Olejnik An. A., Pavlenko D. V., Subbotin S. A.; pod red. Pavlenko D. V., Subbotina S. A. *Progressivnye tehnologii modelirovaniya, optimizacii i intellektual'noj avtomatizacii etapov zhiznennogo cikla aviacionnyh dvigatelej : monografiya*. Zaporozh'e, OAO «Motor Sich», 2009, 468 p.