

С. А. Горбатков, О. Б. Рашитова, А. М. Солнцев

ИНТЕЛЛЕКТУАЛЬНОЕ МОДЕЛИРОВАНИЕ В ЗАДАЧЕ ПРИНЯТИЯ РЕШЕНИЙ В РАМКАХ НАЛОГОВОГО УПРАВЛЕНИЯ

Рассматривается подход к повышению эффективности принятия решений налогового регулирования за счет использования интеллектуального моделирования. В рамках данного подхода предложен метод селекции факторов оценки финансового состояния предприятий в процедуре их кластеризации на нейронных сетях на основе байесовского подхода. Представлены результаты исследований эффективности предложенного метода на примере оценки кредитоспособности предприятий в различных отраслях народного хозяйства. *Налоговое администрирование; кластеризация; селекция факторов; байесовский подход; нейронные сети*

ВВЕДЕНИЕ

Прикладной аспект рассматриваемой задачи связан с технологиями налогового регулирования (управления). Кластеризация налогооблагаемых предприятий с целью принятия решений по налоговому регулированию повышает объективность принимаемых решений. Лицо, принимающее решение (ЛПР), может отнести то или иное предприятие к одному из получаемых кластеров и принять решение по результатам такого математического моделирования.

Вопросам регуляризации моделей, используемых в экономико-математическом моделировании, уделяется достаточное внимание, поскольку это улучшает их адекватность и повышает качество получаемых решений. В частности, достаточно подробно проработан данный вопрос С. А. Шумским [1] применительно к использованию нейронных сетей для случаев, когда выполняются условия соответствия исходных данных (векторов $\vec{x} = (x_1, x_2, \dots, x_n)$) гауссовым смесям распределения плотности вероятности кластеризируемых объектов.

Однако имеется класс прикладных задач, в которых не выполняется требование возможности представления данных в виде смеси гауссовских распределений в силу сильного искажения исходных данных и их малого объема [2]. В этой связи актуальным является вопрос теоретического обоснования регуляризации нейросетевых кластеризаторов при невыполнении указанных требований.

1. СЕЛЕКЦИЯ ПРИЗНАКОВ ФИНАНСОВОГО СОСТОЯНИЯ ПРЕДПРИЯТИЙ

В вышеуказанных прикладных задачах кластеризации и ранжирования вектор признаков $\vec{x} = (x_1, x_2, \dots, x_n)$ может содержать большое число компонент (до нескольких десятков). Часть из этих компонент имеют высокую информативность в аспекте разделения векторов (объектов) \vec{x} на классы, а часть малоинформативна. В результате качество анализа с использованием моделей систем искусственного интеллекта, например, кластеризации как в нашем случае, может оказаться неудовлетворительным.

В этой связи предлагается метод кластеризации с селекцией признаков и байесовской регуляризацией (КСП и БР), предусматривающий итерационный процесс кластеризации с оценкой ее качества [3].

Селекция признаков при кластеризации представляет собой процедуру выделения из всего множества признаков меньшего подмножества с сохранением информативности [4].

Суть селекции признаков – это выделение признаков, которые приводят к большим расстояниям между классами и к малым расстояниям внутри классов, т. е. к минимизации критерия качества

$$\Theta_q = \frac{\left(\sum_{m=1}^M \sum_{l=1}^{N_m} d^2(\vec{x}_{im}; \vec{x}_{il}) \right)}{\left(\frac{1}{C_M^2} \sqrt{\sum_{l=1}^M \sum_{m=l+1}^M d^2(\vec{x}_{il}; \vec{x}_{im})} \right)}, \quad (1)$$

где $q = 1, 2, \dots, Q$ и Q – соответственно номер гипотезы-нейросети в байесовском ансамбле и их общее количество; N_m – число элементов, по-

павших в m -й кластер; \bar{x}_{Ω_m} – точка центра m -го кластера в n -мерном евклидовом пространстве признаков; $d(\bar{x}_{\Omega_l}; \bar{x}_{\Omega_m})$ – евклидово расстояние от исследуемого объекта \bar{x}_{ij} до центра своего m -го кластера; $d(\bar{x}_{\Omega_l}; \bar{x}_{\Omega_m})$ – расстояние между l -м и m -м кластерами; C_M^2 – число сочетаний из M по 2; M – количество кластеров.

В предлагаемом методе используется скалярная селекция признаков, предусматривающая отдельное независимое рассмотрение используемых признаков.

Суть метода скалярной селекции признаков состоит в оценке дискриминантной способности каждого отдельного признака x_j путем проверки соответствующих статистических гипотез о законах распределения плотности вероятности анализируемого признака в разных классах (кластерах). Если распределения плотности $f(x_j|\Omega_l)$, $l=1, \bar{M}$ совпадают для разных классов ($l \neq m$) при назначенной мере сходства, то признак не различает эти классы Ω_l и Ω_m .

Рассмотрим количественные соотношения для алгоритма скалярной селекции для наглядности на примере двух классов: Ω_1 и Ω_2 , хотя все соотношения, приводимые ниже, справедливы для любого конечного числа классов $\Omega_1, \Omega_2, \dots, \Omega_m, \dots, \Omega_M$.

Примем следующие обозначения: пусть x_j – анализируемый признак. Пусть известны прецеденты для этого признака, т. е. случайные его реализации из класса $\Omega_1 \{x_{ij,1}\} \in \Omega_1; i = \overline{1, N_1}$ и для класса $\Omega_2 \{x_{ij,2}\} \in \Omega_2; i = \overline{1, N_2}$, где N_1, N_2 – числа прецедентов в классах Ω_1 и Ω_2 . Методами математической статистики [5] оцениваются эмпирические законы распределения плотности вероятности для $\{x_{ij,1}\}$ и $\{x_{ij,2}\}$.

Обозначим через H_0 и H_1 две гипотезы: H_1 – средние значения признаков отличаются существенно в классах Ω_1 и Ω_2 ; H_0 – средние значения признаков отличаются несущественно – нуль-гипотеза. Примем соглашение о том, что признаки $\{x_j\}$ при селекции нормализованы стандартным способом [6] и пусть известны дисперсии для нормированных прецедентов \tilde{x}_j в классах. Тогда при селекции основная задача – проверить отличие средних значений μ_1 и μ_2 признаков в двух классах Ω_1 и Ω_2 . Соответствующие гипотезы имеют вид:

$$H_0 : \Delta\mu_j = \mu_{j1} - \mu_{j2} = 0; H_1 : \Delta\mu_j \neq 0. \quad (2)$$

В задаче о значимости различия средних μ_{j1} и μ_{j2} используется критерий Стьюдента [5] с числом степеней свободы $(N_1 + N_2) - 2$:

$$t_j = \frac{|\mu_{j1} - \mu_{j2}|}{S_j \sqrt{\frac{1}{N_1} + \frac{1}{N_2}}}; \quad (2)$$

где S_j – общее среднее квадратическое отклонение анализируемого признака x_j от своих средних в кластерах Ω_1 и Ω_2

$$S_j = \sqrt{\frac{(N_1 - 1)S_{j1}^2 + (N_2 - 1)S_{j2}^2}{(N_1 + N_2) - 2}}.$$

Заметим, что если закон распределения плотности вероятности анализируемого признака x_j в кластерах Ω_1 и Ω_2 отклоняется от нормального, то оценка дискриминантной способности этого признака будет носить приближенный характер. Для уменьшения погрешности такой оценки в предлагаемом методе используется байесовский подход к регуляризации модели кластеризации, подробно описанный ниже.

После вычисления по (2) проверяем нуль-гипотезу:

$$H_0 : t_j \leq t_\alpha(\alpha; v = N_1 + N_2 - 2)? \quad (3)$$

где t_α – табличное значение критерия Стьюдента с принятым уровнем значимости α ; v – число степеней свободы для статистики Стьюдента.

Если неравенство (3) выполнено, то делается вывод о том, что статистически значимого различия в средних μ_1 и μ_2 по классам Ω_1 и Ω_2 нет. Следовательно, анализируемый признак x_j неинформативен в аспекте разделения данных на классы (кластеры) с принятым уровнем значимости α и при имеющихся объемах прецедентов в двух классах N_1 и N_2 . При конечном числе классов M описанная процедура селекции признаков проводится последовательно попарно для классов $\Omega_i, \Omega_m, (m \neq i)$.

Признак $\{x_j\}$ считается неинформативным для выполненного разбиения, если он не различает распределения этого признака более чем в одном из всех сочетаний пар классов $\Omega_l, \Omega_m, (m \neq l); l, m = 1, 2, \dots, M$.

2. БАЙЕСОВСКИЙ ПОДХОД К АНАЛИЗУ ОБЪЕКТОВ НАЛОГООБЛОЖЕНИЯ

Причиной неудовлетворительного качества кластеризации с помощью нейросетевых инструментов является возможная сильная зави-

симось результатов кластеризации от параметров настройки сети. Так, для рассматриваемого примера использования в качестве нейросетевого кластеризатора самонастраивающейся карты Кохонена (SOM) результаты кластеризации сильно зависят от параметров, определяющих динамику скорости ее обучения и величины гауссовой окрестности возбужденного нейрона. Данная неустойчивость нейросетевого кластеризатора по вариации параметров настройки в методе КСП и БР парируется на основе байесовской процедуры регуляризации нейросети.

Главные идеи байесовского подхода [1], используемые в предлагаемом методе, следующие:

- выбор ансамбля априорных гипотез-нейросетей $\{h_q(\bar{x}, W)\}$, где W – множество параметров модели (синаптических весов), осуществляется из фиксированного класса (семейства) H мета-гипотез (сетей Кохонена);

- апостериорная фильтрация обученных гипотез-нейросетей осуществляется по критерию, оценивающему качество кластеризации (1) как по плотности группировки объектов вокруг центров кластеров (числитель отношения (1)), так и по удаленности кластеров друг от друга (знаменатель в (1));

- после фильтрации гипотез-нейросетей осуществляется усреднение критерия качества разбиения векторов (объектов) $\bar{x} \in D$ на кластеры по (1) на отфильтрованном ансамбле гипотез-нейросетей.

В предлагаемом методе КСП и БР нейронной сети формула Байеса непосредственно не используется для апостериорной оценки вероятности $\{P(h_q|D|H)\}$, где $P(\cdot)$ – апостериорная вероятность выбранных гипотез-нейросетей; $\{h_q\}$ – множество априорно выбранных гипотез-нейросетей в ансамбле; H – мета-гипотеза ($\{h_q\} \in H$); D – множество данных, поскольку для оценки указанной вероятности через функцию правдоподобия требуется априорное знание аналитической формы закона распределения кластеризуемых векторов \bar{x} , например в виде гауссовой смеси. Такого знания у нас нет. Поэтому апостериорные вероятности $\{P(h_q|D|H)\}$, несущие информацию о качестве разбиения данных D на кластеры, в предлагаемом методе КСП и БР оцениваются косвенно путем фильтрации гипотез-нейросетей $\{h_q\}$ по критерию (1).

Фильтрация гипотез-нейросетей для случая с большим разбросом качества разбиения

Организуется итерационный процесс пошагового отбора (удаления из ансамбля) гипотез-нейросетей $\{h_q\}$ с низким качеством кластеризации (1), т.е. большим значением $\Theta_q[h_q(\bar{x}, W)|D|H]$:

$$q^* : \Theta^{(q)} \leq \Theta_0; q = 1, 2, \dots, Q, \quad (4)$$

где q^* – номер гипотезы-нейросети, успешно прошедшей процедуру фильтрации; Θ_0 – желаемое значение качества фильтрации, определяемое в предварительных вычислительных экспериментах; $D = \{x_{ij}\}_{i=1}^{i=N}$, ($j = \overline{1, n}$) – совокупность вектор-строк данных.

После фильтрации (4) уточненные значения центров кластеров $\{\bar{x}_{um}\}$ и соответствующего им критерия качества разбиения Θ по (1) найдутся как усредненные на отфильтрованном байесовском ансамбле величины:

$$\bar{x}_{um} = \left[\sum_{q=1}^{Q^*} \bar{x}_{umq^*} \right] / Q^*; \bar{\Theta}^* = \left(\sum_{q=1}^{Q^*} \Theta_{q^*} \right) / \Theta^*. \quad (5)$$

Процесс обучения SOM характеризуется, во-первых, окрестностью взаимодействия k^* -го нейрона с i -м вектором обучающей выборки [11]:

$$h_{k^*i} = \exp\left(-\frac{d_{k^*i}^2}{2\sigma^2}\right), \quad (6)$$

где d_{k^*i} – расстояние взаимодействия по евклидовой метрике; σ – параметр гауссового распределения

$$\sigma(t) = \sigma_0 \exp\left(-\frac{t}{\tau_1}\right), \quad t = 0, 1, 2, \dots, \quad (7)$$

где σ_0 – начальное значение величины σ в алгоритме SOM; τ_1 – некоторый параметр, влияющий на характеристику σ . Во-вторых, процесс обучения SOM характеризуется скоростью изменения (модификации) весов при обучении, характеризуемой параметром $\eta(t)$ экспоненциально изменяющегося от номера повторного прогона обучающей выборки (фактически от времени t):

$$\eta(t) = \eta_0 \exp\left(-\frac{t}{\tau_2}\right), \quad t = 0, 1, 2, \dots, \quad (8)$$

где τ_2 – еще один параметр, влияющий на эффективность работы алгоритма SOM.

При формировании байесовского ансамбля гипотез-нейросетей $\{h_q\}$ варьировались указанные параметры обучения τ_1 и τ_2 .

Организация итераций в алгоритме метода

Шаг 0. $k = 0$ (k – номер итерации), $n^{(k)} = n$. При заданных M , Θ_0 и полной размерности n вектора признаков $\bar{x} \in R^n$ осуществляем начальное разбиение D на кластеры методом самоорганизующихся карт Кохонена [11]. При этом строится байесовский ансамбль из Q нейросетей Кохонена $\{h_q(\bar{x}, W)|D|H\}$ и по (3)–(4) находятся уточненные центры кластеров $\bar{x}_{cm}^{(k)}$ и критерий качества разбиения $\bar{\Theta}^{*(k)}$.

Шаг 1. $k = k + 1$. Производим по каждому признаку x_j расчет критериев Стьюдента t_j по (2) и проверяем соответствующие нуль-гипотезы H_0 по (3) для всех пар образованных кластеров $\Omega_l, \Omega_m, (m \neq l); l, m = 1, 2, \dots, M$.

Шаг 2. Проводим проверку информативности каждого признака, исходя из условия, что он различает распределения этого признака более чем в одном из всех сочетаний пар классов $\Omega_l, \Omega_m, (m \neq l); l, m = 1, 2, \dots, M$.

Шаг 3. Если все рассмотренные признаки информативны, то процедура селекции заканчивается, если нет – осуществляется исключение неинформативных признаков из компонент вектора \bar{x} , т.е. сокращение размерности вектора \bar{x} до $n^{(k)} = n^{(k-1)} - u$, где u – число исключенных на данной итерации признаков.

Шаг 4. Осуществляем кластеризацию с измененным составом признаков и проверяем условие улучшения качества разбиения на k -й итерации селекции признаков:

$$\bar{\Theta}^{*(k)} < \bar{\Theta}^{*(k-1)}? \quad (9)$$

Если неравенство (9) выполняется, то, продолжая процесс отбора, переходим к шагу 1, если нет – «ужесточаем» правила отбора по (3) и (4) и, начиная заново процесс отбора, переходим к шагу 0.

Замечание. Как известно [10], результат кластеризации с помощью сети Кохонена изменяется при повторных запусках нейросети (дублирующих расчетах) за счет случайного характера выбора начальных положений нейронов. В этом случае в обобщенном критерии качества разбиения (5) полезно вместо $\{\Theta_{q^*}\}$ брать значения, осредненные по дублирующим расчетам:

$$\Theta_{q^*} \rightarrow \bar{\Theta}_{q^*} = \frac{1}{R} \sum_{r=1}^R \Theta_{q^*,r}, \quad (10)$$

где $\Theta_{q^*,r}$ – значение критерия качества в r -м повторном расчете; R – количество повторов.

3. КОЛИЧЕСТВЕННЫЕ ОЦЕНКИ

Апробация предлагаемого метода КСП и БР производилась на реальных (закодированных) данных сельскохозяйственных предприятий-налогоплательщиков Республики Башкортостан. В качестве признаков кластеризации использованы показатели, достаточно полно характеризующие состояние предприятий в условиях российской экономики. Они отражают мнения руководителей коммерческих предприятий (по результатам их опроса) и входят в состав методики Федерального управления по делам несостоятельности (банкротства) и моделей Альтмана [8].

В качестве основных используются 16 показателей. Первую группу составляют показатели, характеризующие рентабельность предприятия:

$R1$ – общая рентабельность (отношение балансовой прибыли к сумме выручки от продаж и внереализационных доходов);

$R2$ – рентабельность активов (отношение чистой прибыли к средней балансовой стоимости активов);

$R3$ – рентабельность собственного капитала (отношение чистой прибыли к сумме доходов будущих периодов, капиталов и резервов (за вычетом собственных акций, выкупленных у акционеров) за вычетом целевого финансирования и поступлений);

$R4$ – рентабельность продукции (отношение прибыли от продаж к выручке от продаж);

$R5$ – рентабельность оборотных активов (отношение чистой прибыли к средней стоимости оборотных активов).

Вторую группу составляют показатели, характеризующие ликвидность и платежеспособность предприятия:

$L1$ – быстрый коэффициент ликвидности (отношение вычета запасов, налога на добавленную стоимость по приобретенным ценностям и долгосрочной дебиторской задолженности из оборотных активов к краткосрочным обязательствам, не включая доходы будущих периодов);

$L2$ – коэффициент покрытия запасов (отношение суммы оборотных собственных средств, краткосрочных займов, кредитов и краткосроч-

ной кредиторской задолженности к средней величине запасов);

L3 – текущий коэффициент ликвидности (отношение разности оборотных активов и долгосрочной дебиторской задолженности к краткосрочным обязательствам, не включая доходы будущих периодов).

Третью группу составляют показатели, характеризующие деловую активность:

A2 – оборачиваемость активов (отношение выручки от продажи за вычетом налога на добавленную стоимость, акцизов и др. обязательств к средней стоимости активов);

A4 – оборачиваемость кредиторской задолженности (отношение выручки от продажи без учета коммерческих и управленческих расходов к средней кредиторской задолженности);

A5 – оборачиваемость дебиторской задолженности (отношение выручки от продажи за вычетом налога на добавленную стоимость, акцизов и др. обязательств к разности дебиторской задолженности на конец отчетного периода и задолженности учредителей по вкладам в уставной капитал на конец отчетного периода);

A6 – оборачиваемость запасов (отношение себестоимости к средней величине запасов).

Четвертую группу составляют показатели, характеризующие финансовую устойчивость предприятия:

F1 – коэффициент финансовой зависимости (отношение суммы долгосрочных и краткосрочных обязательства, не включая доходы будущих периодов к сумме доходов будущих периодов, капитала и резервов (за вычетом собственных акций, выкупленных у акционеров) за вычетом целевого финансирования и поступлений);

F2 – коэффициент автономии собственных средств (отношение суммы доходов будущих периодов, капитала и резервов (за вычетом собственных акций, выкупленных у акционеров) за вычетом целевого финансирования и поступлений к сумме внеоборотных и оборотных активов);

F3 – обеспеченность запасов собственными оборотными средствами (отношение собственных оборотных средств к запасам);

F4 – индекс постоянного актива (отношение суммы внеоборотных активов и долгосрочной дебиторской задолженности к сумме доходов будущих периодов, капиталу и резервов (за вычетом собственных акций, выкупленных у акционеров) за вычетом целевого финансирования и поступлений).

Практическое подтверждение предложенного метода селекции признаков осуществлялось для группы из 24 предприятий, для которых были просчитаны вышеназванные 16 показателей в качестве признаков кластеризации. В качестве инструментария кластеризации использовалась нейронная сеть Кохонена, которая успешно применяется в задаче принятия решений при налоговом администрировании [2, 9].

В рассматриваемом примере в гипотезах-нейросетях $\{h_q(\bar{x}, W)|H\}$ варьировались две вышеназванные эвристики – τ_1 и τ_2 .

Параметр $\{\tau_1\}$ варьировался дискретно на трех уровнях: $\tau_1 = \{140; 280; 700\}$.

Параметр $\{\tau_2\}$ варьировался дискретно также на трех уровнях: $\tau_2 = \{125; 250; 625\}$.

Уровни указанных параметров подбирались путем предварительных вычислительных экспериментов. В различных сочетаниях уровней τ_1 и τ_2 было образовано в байесовском ансамбле 9 сетей Кохонена. Для каждой q -й сети проводилось по 3 дублирующих обучения.

Для расчетов использовался программный продукт Deductor Studio 4.4 (демоверсия с ограничением числа записей) для аналитической платформы Deductor Lite.

Проводились предварительные вычислительные эксперименты по выбору параметров адаптивного процесса обучения сети Кохонена с целью определения начального значения ширины σ_0 функции топологической окрестности $h_{k^*,q}(d_{k^*,q})$, начальной скорости обучения η_0 , числа эпох T (итераций) процесса модификации весов: $T = 500$; $\sigma_0 = 4$; $\eta_0 = 0,3$.

Основные результаты процесса кластеризации в представленном ансамбле гипотез-нейросетей заключаются в следующем.

Фильтрация байесовских гипотез $\{h_q(\bar{x}, W)|H\}$ для байесовского метода кластеризации 24 рассмотренных предприятий проведена во всех 27 нейросетевых моделях байесовского ансамбля.

Обобщенный показатель Θ по (1), оценивающий косвенно апостериорную вероятность гипотез-нейросетей $P(h_q(\bar{x}, W)|D|H)$, изменяется на множестве из 27 сетей ансамбля в довольно широких пределах: от $\Theta_q[h_q]_{\min} = 0,43852$ до $\Theta_q[h_q]_{\max} = 0,59826$, т.е. на 36,4%.

Согласно байесовскому методу было использовано предложенное правило фильтрации (4). Уже на первой итерации фильтрации установлено, что все 9 сравниваемых гипотез-

нейросетей $\{h_q\}$ прошли условия отбора и оставлены в апостериорном (отфильтрованном) ансамбле. После выполнения первой итерации процедуры селекции при кластеризации с использованием всех признаков были оставлены в качестве информативных только пять – $L2$, $A6$, $F1$, $F2$, $F3$.

На второй итерации оценка информативности этих признаков не изменилась с сохранением качества кластеризации по (1). Таким образом было сокращено пространство признаков кластеризации для их практического применения. Усредненное по (5) значение критерия качества значительно уменьшилось (рис. 1), что говорит об эффективности процедуры селекции, а окончательно полученное значение $\bar{\Theta}^{*(2)} = 0,505$ – о достаточно высоком качестве разбиения.

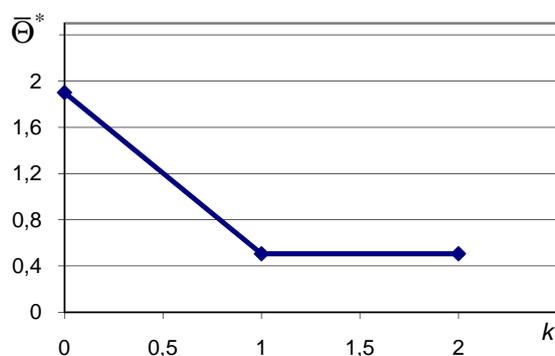


Рис. 1. Динамика критерия качества в процессе селекции

ЗАКЛЮЧЕНИЕ

Показана эффективность предложенного метода селекции признаков кластеризации, выраженная в значительном сокращении их количества, в рассматриваемом примере моделирования принятия решений при налоговом администрировании предприятий.

Применительно к условиям моделирования рассмотренного примера (сильное зашумление данных и их дефицит) проявляется заметное расслоение результатов разбиения на кластеры в зависимости от параметров адаптивного обучения сети Кохонена, которые в широких пределах варьировались в ансамбле сетей. Следовательно, идея о необходимости регуляризации нейросетевого классификатора подтвердилась в вычислительных экспериментах. Идея байесовского метода регуляризации количественно апробирована и подтверждена для рассмотренной практической задачи.

СПИСОК ЛИТЕРАТУРЫ

1. Шумский С. А. Байесова регуляризация обучения // Лекции по нейроинформатике. Ч. 2. М.: МИФИ, 2002. С. 30–93.
2. Нейросетевое математическое моделирование в задачах ранжирования и кластеризации в бюджетно-налоговой системе регионального и муниципального уровней: монография / С. А. Горбатков [и др.]. Уфа: РИЦ БашГУ, 2011. 222 с.
3. Горбатков С. А., Рашитова О. Б. Байесовский итерационный алгоритм кластеризации на основе селекции признаков: Св-во ОФЭРНИО 17538. Рег. 31.10.2011.
4. Местецкий Л. М. Математические методы распознавания образов: курс лекций. М.: МГУ, 2004. 85 с.
5. Айвазян С. А., Мхитарян В. С. Прикладная статистика и основы эконометрики. М.: ЮНИТИ, 1998. 1023с.
6. Ежов А. А., Шумский С. А. Нейрокомпьютеринг и его применение в экономике и бизнесе: учебник. М.: Изд-во МИФИ, 1998. 224 с.
7. Вентцель Е. С. Теория вероятностей. М.: Высшая школа, 1999. 576 с.
8. Давыдова Г. В., Беликов А. Ю. Методика количественной оценки риска банкротства предприятий // Управление риском. 1999. № 3. С. 13–20.
9. Горбатков С. А., Рашитова О. Б. Кластеризация заемщиков для оценки кредитного риска // Проведение научных исследований в области обработки, хранения, передачи и защиты информации: сб. науч. тр: в 4 т. Т4. Ульяновск: УлГТУ, 2009. С. 394–403.
10. Хайкин С. Нейронные сети: полный курс. М.: Изд. дом «Вильямс», 2006. 1104 с.

ОБ АВТОРАХ

Горбатков Станислав Анатольевич, проф. каф. математики и информатики Уфимск. филиала Финансового ун-та при Правительстве РФ. Дипл. инженер-электромеханик (Томск. политехн. ин-т, 1960). Д-р техн. наук по упр-ю в техн. системах (МИЭМ, 1990). Нейросетевое моделирование и управление в технических и экономических системах.

Рашитова Ольга Борисовна, ст. преп. той же каф. Дипл. математик (БГУ, 1999), экономист (ВЗФЭИ, 2007). Иссл. в обл. экон. процессов

Солнцев Андрей Михайлович, нач. налоговой инспекции Межрайон. инспекции Фед. налоговой службы № 39 по РБ. Дипл. экономист (ВЗФЭИ, 1995). Иссл. в обл. процессов налогообложения.