

А. Ю. Минц,

ГВУЗ «Приазовский государственный
технический университет»

МЕТОДЫ ОТБОРА ВХОДНЫХ ДАННЫХ ДЛЯ НЕЙРОСЕТЕВОГО МОДЕЛИРОВАНИЯ

АНОТАЦІЯ. У статті здійснено постановку і вирішення актуальної проблеми відбору вхідних даних для нейромережевого моделювання в умовах малого об'єму навчальної вибірки. Запропоновано критерії визначення завершення ітеративного процесу зменшення розмірності вхідного вектора даних. Запропоновано метод рішення поставленої задачі за допомогою алгоритму автоматичної побудови дерева рішень С4.5. На прикладі задачі класифікації банківських позичальників розглянуто його ефективність у порівнянні з методом кореляційного аналізу.

КЛЮЧОВІ СЛОВА. Нейронна мережа, дерево рішень, вхідні дані, значущість, алгоритм С4.5.

АННОТАЦИЯ. В статье осуществлена постановка и решение актуальной проблемы отбора входных данных для нейросетевого моделирования в условиях малого объема обучающей выборки. Предложены критерии для определения окончания итеративного процесса уменьшения размерности входного вектора данных. Предложен метод решения поставленной задачи с помощью алгоритма автоматического построения деревьев решений С4.5. На примере задачи классификации заемщиков рассмотрена эффективность предложенного метода по сравнению с методом корреляционного анализа.

КЛЮЧЕВЫЕ СЛОВА. Нейронная сеть, дерево решений, входные данные, значимость, алгоритм С4.5.

ANNOTATION. The article contain the formulation and solution of actual problem of selecting input data for neural network modeling in case of small amount of input data. Author propose a set of criteria for determining the end of the iterative process of reducing the dimension of input vector. Also proposed and compared with method of correlation analysis a new method of solving this problem by using the algorithm of automatic decision trees construction С4.5.

KEYWORDS. Neural network, decision tree, the input data, the significance of the algorithm С4.5.

Искусственные нейронные сети (ИНС) уже длительное время являются одним из самых распространенных инструментов поиска сложных зависимостей при анализе экономических данных. Однако ряд задач, непосредственно связанных с практическим применением ИНС, до сих пор нельзя назвать полностью решенными. К ним, в частности, относится проблема формализации ме-

тодов отбора входных данных для нейросетевого моделирования.

Пусть T — обучающая выборка, представляющая собой множество данных, состоящее из векторов X_i и ожидаемого отклика сети $f(X_i)$:

$$T = \{X_i, f(X_i)\}_{i=1}^I,$$

где I — количество примеров в выборке, а X_i — вектор, имеющий вид:

$$X_i = (x_{i,1}, x_{i,2}, \dots, x_{i,M}),$$

где M — количество параметров, составляющих один пример.

Пусть NS — структура нейронной сети, которая в общем виде может быть описана таким образом:

$$NS = \langle N, L, \psi \rangle,$$

где N — множество нейронов, составляющих ИНС;

L — множество связей между нейронами;

ψ — отношение инцидентности, ставящее в соответствие каждой связи из множества L два нейрона из множества N .

Для полносвязных многослойных нейронных сетей прямого распространения (именно этот тип ИНС находит наиболее широкое применение при решении экономических задач) при описании структуры можно ограничиться указанием количества нейронов в каждом слое. В этом случае структуру (точнее — архитектуру) сети можно записать формулой вида

$$N^i - N^h - \dots - N^o,$$

где N^i — количество нейронов во входном слое, совпадающее с M — количеством параметров в обучающем примере;

N^h — количество нейронов в каждом из скрытых слоев;

N^o — количество нейронов в выходном слое, совпадающее с количеством выходов ИНС.

Так, формула 5-6-3-1 описывает архитектуру ИНС, имеющей 5 входов, 6 нейронов в первом скрытом слое, 3 нейрона во втором скрытом слое и 1 выход.

В общем виде проблема отбора входных данных сводится к нахождению такой выборки $\bar{T} \subset T$, чтобы при заданном параметре объема выборки I обеспечивалось достаточно хорошая аппроксимация функции $f(X_i)$ и соответственно решение иссле-

дуемой задачи. При этом на свойства выборки \bar{T} могут накладываться ограничения, в зависимости от условий задачи.

Существование проблемы отбора входных данных обусловлено известной зависимостью между сложностью структуры нейронной сети и количеством примеров, которые необходимы для её обучения. Теоретически достаточное количество примеров в обучающей выборке можно определить, отталкиваясь от концепции *VC-измерения* (измерения Вапника—Червоненкиса), введенного в [1] и отражающего «вычислительную мощность семейства функций классификации, реализованных обучаемыми машинами» [2]. *VC-измерение* также может быть интерпретировано как число образов, достаточное для адекватного обучения искусственной нейронной сети заданной сложности. Хотя точное аналитическое определение этого параметра для конкретной архитектуры ИНС в большинстве случаев невозможно, оценки, сделанные в [3], показывают, что для наиболее распространенного типа нейронной сети, то есть сети прямого распространения, состоящей из нейронов с сигмоидальной активационной функцией, *VC-измерение* имеет порядок W^2 , где W — количество степеней свободы (в некоторых источниках — свободных параметров) ИНС. Поскольку к свободным параметрам сети относятся весовые коэффициенты связей и пороговые значения активации нейронов, то:

$$W = |L| + |N|. \quad (1)$$

При этом, если архитектуру ИНС обозначить формулой вида $N_1 - N_2 - \dots - N_{lay}$, где N_1 — входной слой, N_{lay} — выходной, а остальные — скрытые слои, то:

$$|L| = \sum_{i=1}^{lay-1} N_i \cdot N_{i+1}, \quad (2)$$

$$|N| = \sum_{i=2}^{lay-1} N_i. \quad (3)$$

Рассмотрим требования к количеству примеров во входной выборке на примере сравнительно небольшой полносвязной нейронной сети прямого распространения, с архитектурой 5-5-1. В такой сети $lay = 3$, следовательно из (1)-(3) $|L| = 25 + 5 = 30$, $|N| = 5 \Rightarrow W = 35$. Следовательно для обучения такой сети необходима выборка, в которой количество независимых примеров имеет тот же порядок, что и 35^2 .

Другие методы оценки (например, предложенный в [4]) позволяют задавать допустимый уровень ошибки сети, что дает возможность получить несколько более оптимистические значения размеров обучающей выборки при приемлемом уровне ошибки, однако требования к её величине всё равно остаются достаточно высокими.

Из (2) следует, что в полносвязной нейронной сети количество свободных параметров непосредственно зависит от количества нейронов во входном и выходном слоях, а следовательно от размерности вектора входных данных и его представления в ИНС. При этом нередко встречается ситуация, когда размерность велика, а количество примеров в обучающей выборке напротив — мало. Так, например задача предсказания банкротств в банковской системе Украины связана с анализом более чем 30 параметров, характеризующих активы, пассивы и финансовые результаты банков, что обуславливает необходимость наличия выборки данных, состоящей из примерно 25 000 примеров (расчеты сделаны для нейронной сети с архитектурой 30-5-1). В то же время выборка, сделанная на основе банковской статистики за период с сопоставимыми условиями ведения банковской деятельности имеет объем на порядок меньше требуемого. Аналогична ситуация возникает и во многих других случаях, когда собрать объем данных, необходимый для нормального обучения нейронной сети не представляется возможным.

Проблема отбора входных данных тесно связана с задачей нахождения такой структуры сети NS , которая бы обеспечивала минимум ошибки при решении задачи. В совокупности эти две задачи не имеют и скорее всего не могут иметь аналитических методов решения. Не существует также и достаточно надежных эмпирических способов их решения, поэтому распространенным способом решения является простой перебор возможных вариантов.

Рассмотрим эффективность различных методов отбора данных для нейросетевого моделирования на примере решения некоторой стандартной задачи. В качестве программной базы для исследования выбран аналитический пакет *Deductor* компании *BaseGroup*. Преимуществами этого пакета является большое количество поддерживаемых методов анализа данных, развитые средства визуализации, интерактивный интерфейс, наличие функциональной бесплатной академической версии. К недостаткам нейросетевой части пакета можно отне-

сти работу только с полносвязными сетями, то есть отсутствуют возможности прореживания связей и создания нестандартных архитектур. Кроме того, пакет поддерживает всего два алгоритма обучения ИНС.

В качестве стандартной задачи возьмем задачу анализа кредитоспособности заемщика на основании анкетных данных. Эта задача имеет большое значение в банковской сфере, поскольку от эффективности её решения напрямую зависят возможные потери банков по кредитным операциям. Для использования скоринговых методов анализа, которые являются традиционными при решении этой задачи [5], необходима большая выборка данных, полученных в сопоставимых условиях. Выполнение этого условия влечет для банков большие затраты, которые связаны либо с накоплением такой выборки (при решении задачи своими силами), либо с покупкой готового решения у сторонней компании — разработчика. При этом в случае резких изменений в экономическом окружении, как это произошло, например, с наступлением экономического кризиса, существующие аналитические процедуры утрачивают свою актуальность и требуются новые инвестиции в их разработку. Именно поэтому решение задачи анализа кредитоспособности заемщиков при небольшом объеме данных предыстории имеет самостоятельную практическую ценность для банковской системы.

В рамках данного исследования анализ проводился на выборке данных о кредитозаемщиках, включающей 149 обучающих примеров, в том числе 114 позитивных и 35 негативных. Для описания заемщика используется 25 параметров. По приведенным выше формулам нетрудно подсчитать, что при пяти нейронах в скрытом слое требуемый объем выборки должен иметь тот же порядок, что и 161^2 , то есть выборка должна содержать несколько десятков тысяч примеров.

Рассмотрим теперь задачу с обратной стороны, то есть определим, какое количество степеней свободы может иметь нейронная сеть при данном объеме выборки. Если значение VC -измерения приравнять к имеющемуся количеству примеров, то соответствующее количество свободных параметров $W = \sqrt{VC \dim(\bullet)}$ $= \sqrt{149} = 12,2$. Примерно столько имеет нейронная сеть конфигурации 3-2-1. Следовательно, задачей отбора данных является существенное уменьшение размерности входного вектора. Поскольку этот процесс носит итеративный характер, необходимо сформулировать некоторый набор критериев, выполнение кото-

рых позволяет определить окончание процесса. К таким критериям можно отнести следующие:

1. достижение незначительной ошибки при распознавании обучающего множества;
2. достижение незначительной ошибки при распознавании тестового множества;
3. достижение хорошей сопряженности результатов;
4. достижение адекватной динамики процесса обучения.

Последние два критерия следует пояснить подробнее.

Матрица сопряженности отражает количество правильно и неправильно классифицированных образцов и позволяет контролировать результаты обучения в случае асимметрии цены ошибок классификации первого и второго рода [6]. В рассматриваемом случае к ошибкам первого рода относится классификация «плохих» заемщиков как «хороших», а к ошибкам второго рода — классификация «хороших» заемщиков как «плохих». В этом случае стоимость ошибок первого рода значительно выше, чем ошибок второго рода.

Анализ динамики обучения нейронной сети необходимо проводить с целью избежать переобучения ИНС, что является основной опасностью работы с малыми выборками данных. На рис. 1 показан классический вид графика изменения средней ошибки ИНС на обучающем и тестовом множествах.

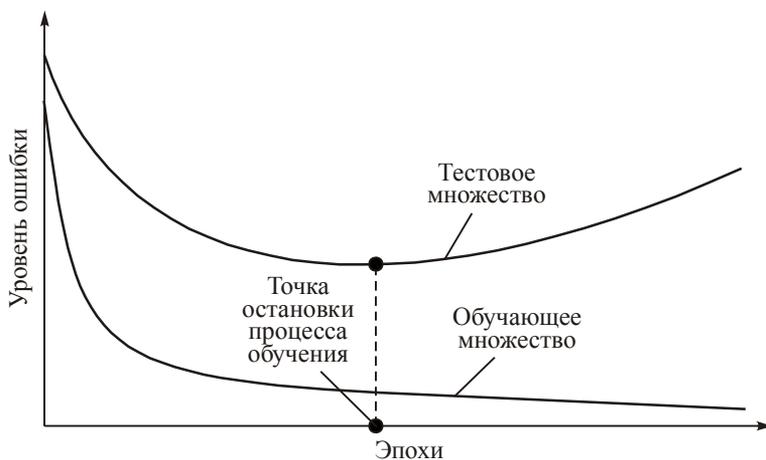


Рис. 1. Динамика процесса обучения ИНС

Из рис. 1 видно, что если средняя ошибка на обучающем множестве монотонно уменьшается, то средняя ошибка ИНС на тестовом множестве проходит через некоторый экстремум, после чего начинает увеличиваться. Точка этого экстремума и есть момент, в который процесс обучения необходимо остановить. В противном случае ИНС вместо поиска зависимостей входных и выходных данных фактически начинает «запоминать» примеры из обучающего множества. Согласно [2, с. 293], перекрестная проверка с использованием тестового множества и ранняя остановка процесса обучения целесообразна при выполнении условия

$$N < 30 W, \quad (4)$$

где W — количество свободных параметров нейронной сети; N — количество примеров в обучающей выборке. То есть, при достаточном малых объемах обучающей выборки перекрестная проверка необходима.

График, показанный на рис. 1, также может служить образцом при анализе процесса обучения, поскольку при явном несоответствии характеристик обучающей выборки и ИНС его вид будет существенно отличаться от приведенного.

До рассмотрения различных методов отбора входных данных построим и проанализируем поведение ИНС, на вход которой подается максимально возможное количество значений из входной выборки. Перед построением такой ИНС целесообразно провести *предварительный (вербальный) анализ* данных, с целью отсечения информации, которая не несет информации о характеристике клиента, либо в принципе не может быть использована в нейросетевом анализе. Так, из имеющейся выборки были исключены поля «Порядковый номер», «№ паспорта», «ФИО», «Адрес», «Название предприятия».

Оставшиеся 20 полей сформировали входной вектор данных ИНС. Поскольку для представления строковых типов данных использовалась нормализация в виде битовой маски, общее количество нейронов во входном слое ИНС составило 32. Количество нейронов в скрытом слое ИНС варьировалось от 2 до 10. Входная выборка данных была случайным образом разделена на обучающую (90 %) и тестовую (10 %). Результаты обучения ИНС алгоритмом *Back Propagation* при различном количестве нейронов в скрытом слое приведены в табл. 1. Сопряженность результатов показана в табл. 2, а график процесса обучения для конфигурации 32-5-1 изображен на рис. 2.

Таблица 1

РЕЗУЛЬТАТЫ РАБОТЫ ИНС КОНФИГУРАЦИИ 32-Н-1

Количество нейронов в скрытом слое (Н)	Распознавание обучающей выборки, %	Распознавание проверочной выборки, %
2	99	66
3	100	73
4	100	67
5	100	73
7	100	73
10	100	60

Таблица 2

СОПРЯЖЕННОСТЬ РЕЗУЛЬТАТОВ КЛАССИФИКАЦИИ ЗАЕМЩИКОВ В КОНФИГУРАЦИИ 32-5-1

Фактически	Классифицировано		
	Плохой	Хороший	Итого
Плохой	34	1	35
Хороший	3	111	114
Всего	37	112	149

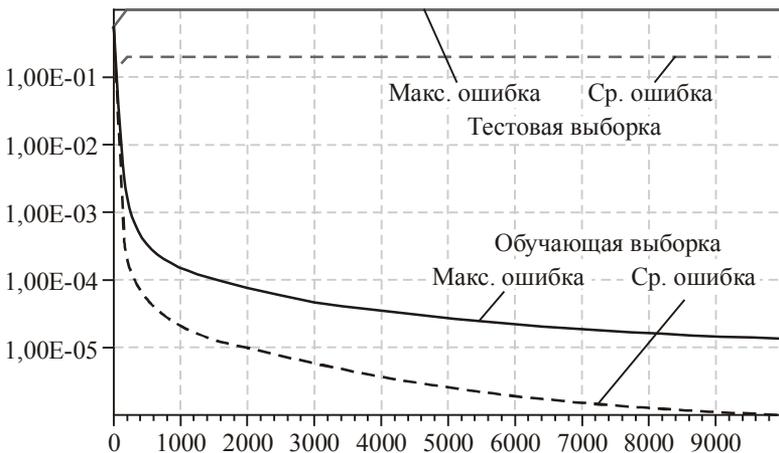


Рис. 2. Динамика процесса обучения ИНС конфигурации 32-5-1

Как видно из анализа приведенных данных, результаты, полученные на обучающем и тестовом множествах, существенно различаются. Кроме того вид графика процесса обучения отличается от приведенного на рис. 1. Всё это свидетельствует о переобучении нейронной сети, что делает невозможным её практическое использование.

Целью сокращения размерности вектора входных данных и соответственно степеней свободы ИНС является исключение параметров, слабо влияющих на конечный результат. Традиционно для этого применяют методы математической статистики, в частности корреляционный анализ, однако ниже будет показана возможность использования и других приемов.

Недостатком корреляционного анализа является возможность обнаруживать зависимости между изменением только тех параметров, которые имеют числовое выражение. Нечисловые параметры могут быть проанализированы лишь в том случае, если они могут быть ранжированы по принципу «лучше»-«хуже», что не всегда возможно.

Результаты корреляционного анализа имеющихся данных в программе *Deductor* приведены на рис. 3.

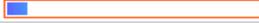
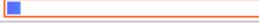
Входные поля		Корреляция с выходными полями	
№	Поле	Давать кредит	▲
15	Обеспеченность займа		0,336
5	Наличие недвижимости		0,327
7	Наличие банковского счета		0,232
14	Срок проживания в данной местности, лет		0,202
9	Срок работы на данном предприятии, лет		0,201
10	Срок работы на данном направлении, лет		0,193
6	Наличие автотранспорта		0,167
13	Количество иждивенцев		0,095
12	Количество лет		0,094
8	Наличие страховки		0,080
4	Среднемесячный расход, грн		-0,073
2	Срок ссуды, мес		0,053
1	Размер ссуды, грн		0,046
3	Среднемесячный доход, грн		-0,024
11	Семейное положение		0,003

Рис. 3. Корреляционный анализ параметров кредитоспособности заемщика

В дополнение к анализу корреляции между входными и выходными параметрами целесообразно также провести аналогич-

ный анализ для входных данных. Так, в рассматриваемом случае коэффициент корреляции между параметрами «Срок работы на данном предприятии» и «Срок работы на данном направлении» оказался равным 0,951, что позволяет исключить один из них из дальнейшей обработки.

Для определения оптимального набора показателей было построено несколько сетей конфигурации I-4-1, то есть с четырьмя нейронами в скрытом слое и различным их количеством на входе. Результаты приведены в табл. 3

Таблица 3

**РЕЗУЛЬТАТЫ РАБОТЫ ИНС
С ФОРМИРОВАНИЕМ НАБОРА ВХОДНЫХ ДАННЫХ
НА ОСНОВАНИИ РАСЧЕТА КОЭФФИЦИЕНТОВ КОРРЕЛЯЦИИ**

№ ИНС	Количество входных параметров	Распознавание обучающей выборки, %	Распознавание проверочной выборки, %
1	2	70,15	73,33
2	3	75,37	73,33
3	4	80,6	73,33
4	5	84,33	73,33
5	6	88,06	73,33
6	7	87,31	66,67
7	8	94,78	73,33
8	9	97,76	80
9	10	96,27	80

Из анализа табл 3 видно, что с ростом количества анализируемых параметров, эффективность распознавания входной выборки сетью в основной улучшается. Однако добавление параметра «Количество иждивенцев» ухудшило распознавание как обучающей, так и проверочной выборки. В связи с этим была проверена также ИНС, аналогичная №7 (по табл. 3), но без этого параметра. В результате обучающая выборка была распознана на 98,51, а тестовая на 86,67%, что в целом лучше, чем результаты, полученные выше с помощью ИНС конфигурации 32-5-1.

График процесса обучения нейронной сети показан на рис. 4. При его анализе можно увидеть, что динамика процесса обучения также гораздо лучше соответствует эталонной, чем в случае ИНС с полным набором входов.

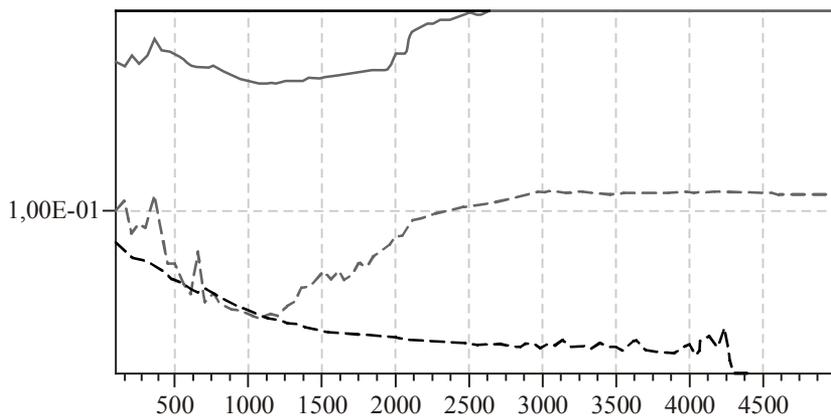


Рис. 4. Динамика процесса обучения ИНС конфигурации 7-4-1

Ранняя остановка обучения в точке минимума средней ошибки тестовой выборки (эпоха 1200) позволила получить по распознаванию обучающей выборки результат 94,03 %, а по тестовой 93,33 %. Это свидетельствует о практически полной идентичности качества распознавания обучающей и тестовой выборок. Анализ сопряженности результатов может быть сделан на основании табл. 4.

Таблица 4

**СОПРЯЖЕННОСТЬ РЕЗУЛЬТАТОВ КЛАССИФИКАЦИИ
ЗАЕМЩИКОВ В КОНФИГУРАЦИИ 7-4-1**

Фактически	Классифицировано		
	Плохой	Хороший	Итого
Плохой	33	2	35
Хороший	2	112	114
Всего	35	114	149

Таким образом, ИНС сделала по две ошибки первого и второго рода, что является приемлемым результатом для практического применения.

Как уже отмечалось, при корреляционном анализе за пределами рассмотрения остаются нечисловые факторы, либо факторы, кото-

рые не могут быть приведены к числовому виду. В то же время фактически для отбора данных подходит любой метод, обеспечивающий ранжирование параметров входного вектора по степени их влияния на конечный результат, даже если такое ранжирование не является его основной функцией. Такие методы анализа отнесем к косвенным. К ним можно отнести алгоритм автоматического построения деревьев решений *C4.5*, предложенный в 1993 году [7].

Для работы алгоритма *C4.5* необходимо соблюдение следующих условий:

- каждый элемент входного набора данных должен соответствовать одному из предопределенных классов;
- каждый пример должен однозначно относиться только к одному из классов;
- количество классов должно быть значительно меньше количества записей во входной выборке.

Легко заметить, что данные условия в целом справедливы и для любой задачи, решаемой с помощью перцептронных нейронных сетей. Рассмотрим возможности данного метода на нашем примере.

Поскольку построение деревьев решений относится к категории методов интеллектуального анализа данных (*Data mining*), в пакете *Deductor* предусмотрена возможность разделения исходной выборки данных на обучающую и проверочную. Однако в рассматриваемом случае проверочная выборка не нужна, поэтому все данные используются только для обучения. Для самого алгоритма задаются следующие параметры (в скобках приведены значения, заданные в ходе исследования):

- минимальное количество примеров, при котором будет создан новый узел (1);
- повышение достоверности дерева, за счет снижения компактности (включено);
- уровень доверия для отсечения узлов дерева (20).

Для уровня доверия оставлено значение по умолчанию. Чем выше это значение, тем более точным, но в то же время громоздким будет дерево решений, поскольку при этом возникает эффект, родственному переобучению нейронных сетей.

В результате работы алгоритма создано дерево решений, правильно классифицирующее 146 из 149 примеров обучающей выборки. Ввиду того, что проверочная выборка в рассматриваемом случае отсутствовала, практическую ценность этого результата рассматривать не будем. В рамках данного исследования больший интерес представляет анализ значимости атрибутов, представленной на рис. 5.

Целевой атрибут: Давать кредит			
N ²	Номер	Атрибут	Значимость, %
1	7	Наличие недвижимости	15,028
2	12	Срок работы на данном предприят...	12,749
3	19	Срок проживания в данной местно...	12,497
4	20	Обеспеченность займа	11,444
5	9	Наличие банковского счета	11,357
6	8	Наличие автотранспорта	8,973
7	5	Среднемесячный расход, грн	8,365
8	4	Среднемесячный доход, грн	6,006
9	17	Количество лет	5,018
10	16	Семейное положение	4,626
11	10	Наличие страховки	3,936
12	18	Количество иждивенцев	0,000
13	3	Цель ссуды	0,000
14	1	Размер ссуды, грн	0,000
15	2	Срок ссуды, мес	0,000
16	6	Основное направление расходов	0,000
17	13	Направление деятельности заем...	0,000
18	11	Отраслевая принадлежность пред...	0,000
19	15	Пол	0,000
20	14	Срок работы на данном направлен...	0,000

Рис. 5. Значимость атрибутов входной выборки по алгоритму C4.5

Как можно заметить, рейтинг показателей, полученный с помощью алгоритма C 4.5, значительно отличается от рейтинга, полученного в результате корреляционного анализа. Рассмотрим эффективность использования данного набора показателей для нейросетевого анализа. Для эксперимента воспользуемся тем же методом, что и выше, то есть обучая несколько нейронных сетей с различным количеством параметров на входе. Результаты экспериментов приведены в табл. 5.

Таблица 5

**РЕЗУЛЬТАТЫ РАБОТЫ ИНС С ФОРМИРОВАНИЕМ НАБОРА
ВХОДНЫХ ДАННЫХ НА ОСНОВАНИИ АНАЛИЗА ДЕРЕВА РЕШЕНИЙ**

№ ИНС	Количество входных параметров	Распознавание обучающей выборки, %	Распознавание проверочной выборки, %
1	2	54,48	80
2	3	85,07	73,33
3	4	78,36	80
4	5	86,57	86,67
5	6	94,03	93,33
6	7	97,76	93,33
7	8	94,03	100
8	9	97,01	93,33
9	10	98,51	100
10	11	97,01	93,33

Рассмотрим более подробнее работу ИНС № 9, показавшей лучшие результаты из рассмотренных. График динамики процесса обучения этой сети показан на рис. 6. Из анализа этого графика можно сделать вывод, что в целом процесс обучения близок к эталонному, но для улучшения качества прогноза целесообразно применить раннюю остановку обучения на уровне эпохи 2500.

Сопряженность результатов полученной сети приведена в табл. 6. Из этой таблицы видно, что ИНС достоверно распознала всех «плохих» заемщиков, то есть вообще не допустила дорогостоящих ошибок первого рода. Два случая неправильной классификации заемщиков относятся к ошибкам второго рода, то есть менее дорогостоящим для банка. Это увеличивает практическую ценность нейросетевой модели.

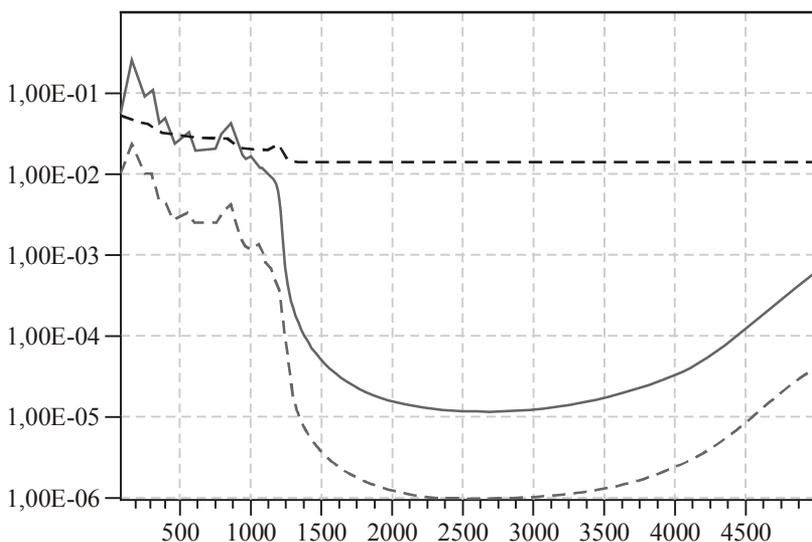


Рис. 6. Динамика процесса обучения ИНС конфигурации 10-4-1

Таблица 6

**СОПРЯЖЕННОСТЬ РЕЗУЛЬТАТОВ КЛАССИФИКАЦИИ
ЗАЕМЩИКОВ В КОНФИГУРАЦИИ 10-4-1**

Фактически	Классифицировано		
	Плохой	Хороший	Итого
Плохой	35	0	35
Хороший	2	112	114
Всего	37	112	149

В целом, на основании данного исследования можно сделать вывод, что косвенный анализ значимости показателей при отборе входных данных для нейросетевого моделирования позволяет в некоторых случаях получить результаты сопоставимые, или лучшие, нежели при наиболее часто используемом в этих целях корреляционном анализе. Использование косвенных методов, в частности анализ значимости показателей с помощью алгоритмов построения деревьев решений, позволяет расширить инструментарий разработки нейросетевых моделей в условиях ограниченности объемов входных данных.

Литература

1. *Vapnik V. N. and A. Ya. Chervonenkis.* On the uniform convergence of relative frequencies of events to their probabilities. — *Theoretical Probability and Its Applications*, 1971, vol. 17, p. 264—280.

2. *Хайкин С.* Нейронные сети: Полный курс / Хайкин С. — [2-е изд.]. — М.: Вильямс, 2006. — 1104 с.

3. *Koiran P. and E. D. Sontag.* Neural networks with quadratic VC dimension. — *Advances in Neural Information Processing Systems*, 1996, vol. 8, p. 197—203, Cambridge, MA: MIT Press.

4. *Baum E. B. and D. Haussler.* What size net gives valid generalization? — *Neural Computation*, 1989, vol. 1, p. 151—160.

5. *Ишина И. В., Сазонова М. Н.* Скоринг-модель оценки кредитного риска// *Аудит и финансовый анализ.* — № 4. — 2007 — 297—304.

6. *Уланов С. В.* Оценка качества и сравнение скоринговых карт.// *Экономические науки*, № 9(58), 2009. С. 330—335.

7. *J. Ross Quinlan.* C4.5: Programs for Machine learning. Morgan Kaufmann Publishers 1993. — 302 p.

Стаття надійшла до редакції 29.06.2011 р.