

УДК 519.62



ИСКУССТВЕННЫЕ НЕЙРОННЫЕ СЕТИ КАК МЕТОД ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ДАННЫХ В ФИНАНСОВЫХ СИСТЕМАХ

Г.Г. Четвериков¹, Т.В. Лесовец², О.В. Касала³¹ ХНУРЭ, г. Харьков, Украина, chetvergg@kture.kharkov.ua^{2,3} ХНУРЭ, г. Харьков, Украина, mailolya@ukr.net

В данной работе рассмотрено применение нейронных сетей в решении проблем определения жизнеспособных и склонных к банкротству предприятий и предсказания курса акций. Приведена архитектура искусственных нейронных сетей и алгоритмы их обучения. Показаны результаты работы нейронных сетей на конкретных примерах. В статье также определены направления дальнейших исследований.

НЕЙРОННАЯ СЕТЬ, ВЕКТОР КВАНТОВАНИЯ, РЕГРЕССИЯ, ПРОГНОЗИРОВАНИЕ, КЛАССИФИКАЦИЯ

Введение

Становление теории искусственных нейронных сетей произошло под влиянием исследования функционирования и строения человеческого мозга. Стало возможным создание искусственных нейронных сетей на основе естественных.

Такие результаты были продиктованы как теоретическими, так и прикладными достижениями. Неожиданно открылись возможности использования вычислений в сферах, до этого относящихся лишь к области человеческого интеллекта, возможности создания машин, способность которых учиться и запоминать напоминает мыслительные процессы человека [1].

Лучшее понимание функционирования нейрона и картины его связей позволило исследователям создать математические модели для проверки своих теорий. В первых же работах выяснилось, что эти модели не только повторяют функции мозга, но и способны выполнять функции, имеющие свою собственную ценность. Поэтому возникли и остаются в настоящее время две взаимно обогащающие друг друга цели нейронного моделирования: первая – понять функционирование нервной системы человека на уровне физиологии и психологии и вторая – создать вычислительные системы (искусственные нейронные сети), выполняющие функции, сходные с функциями мозга. Именно эта последняя цель и находится в центре внимания этой статьи [1].

Исследования в области естественного и искусственного интеллекта позволили разработать множество моделей искусственных сетей, способных решать многие задачи. Наиболее используемыми являются: распознавание зрительных и слуховых образов; ассоциативный поиск информации и создание ассоциативных моделей; синтез речи; формирование моделей различных нелинейных и трудно описываемых математических систем, прогнозирование развития этих систем во времени;

системы управления и регулирования с предсказанием; управление работами; принятие решений и диагностика; исключаящий логический вывод [2].

1. Нейронные сети в прогнозировании банкротства

Риск присущ любой форме человеческой деятельности, что связано со множеством условий и факторов, влияющих на положительный исход принимаемых людьми решений. Статистика утверждает, что большая часть вновь созданных компаний разоряется в течение первых трех лет существования. Тем не менее, аналогичное случается и с компаниями, благополучно просуществовавшими и даже удерживавшими лидирующие позиции на протяжении нескольких десятков лет [3].

Современное состояние науки позволяет использовать нейронные сети в решении данной проблемы. Поскольку прогнозирование банкротства является бинарным процессом принятия решений, то целесообразно считать его задачей классификации. Причем, исходя из вышесказанного, все множество организаций разбивается на два класса: банкрот и не банкрот.

Задача классификации представляет собой задачу отнесения образца к одному из нескольких попарно не пересекающихся множеств. При решении задач классификации необходимо отнести имеющиеся статические образцы (в данной задаче это коэффициенты Альтмана) к определенным классам. Возможно несколько способов представления данных. Наиболее распространенным является способ, при котором образец представляется вектором. Компоненты этого вектора представляют собой различные характеристики образца, которые влияют на принятие решения о том, к какому классу можно отнести данный образец [4].

Нейронные сети являются универсальным средством аппроксимации функций, что позволяет использовать их в решении задач классификации. Как правило, нейронные сети оказываются наиболее

лее эффективным способом классификации, потому что генерируют фактически большое число регрессионных моделей (которые используются в решении задач классификации статистическими методами). Нейронная сеть представляет собой совокупность нейроподобных элементов, определенным образом соединенных друг с другом и с внешней средой с помощью связей, определяемых весовыми коэффициентами. В сложных практических задачах обученная нейросеть выступает как эксперт, обладающий большим опытом и способный дать ответ на трудный вопрос. Нейросетевой подход особенно эффективен в задачах экспертной оценки по той причине, что он сочетает в себе способность компьютера к обработке чисел и способность мозга к обобщению и распознаванию.

2. Оптимизированная нейронная сеть векторного квантования

Многие задачи связаны с необходимостью хранения, обработки и передачи больших массивов данных, что требует значительных вычислительных ресурсов и временных затрат. Если массивы данных отличаются незначительно, то существенной экономии как вычислительных средств, так и необходимого для решения задачи времени возможно достичь, используя сжатие данных или специальные методы их кодирования. При этом данные могут быть, например, разгруппированы в некоторые классы, которым присваивается свой код или индекс. Достаточно хорошо разработанные методы сжатия данных требуют для своей реализации наличие статистической информации (плотности распределения) об исследуемых процессах. Применение эффективных методов возможно, если известна частота появления образов. В условиях, когда такая информация отсутствует, наиболее эффективными оказываются искусственные нейронные сети векторного квантования [5].

Под векторным квантованием подразумевается процесс преобразования некоторого вектора X из множества $A \in R^N$ в вектор W из множества $B \in R^N$, где $M < N$.

Векторное квантование осуществляется по методу «ближайшего соседа», причем под «ближайшим» понимается вектор, удовлетворяющий различным требованиям. Если в качестве такового выбирается вектор, находящийся от данного на минимальном евклидовом расстоянии, имеем классификатор, называемый в литературе Voronoi-классификатор. В сети векторного квантования реализован принцип «победитель получает все».

Существуют различные методы обучения сети векторного квантования (Learning Vector Quantization — LVQ): неконтролируемое обучение сети, контролируемое обучение сети (LVQ 1, LVQ 2.1, LVQ 3), оптимизированный OLVQ1. Основным

недостатком метода неконтролируемого обучения сети является то, что если начальные распределения векторов весов и входных образов не являются примерно одинаковыми, то может возникнуть ситуация, когда некоторые из нейронов никогда не станут победителями, то есть их векторы весов не будут изменяться. Основным преимуществом контролируемого обучения сети векторного квантования от рассмотренного выше является использование для каждого входного образа X желаемого соответствующего выходного сигнала [5]. Таким образом, вектор весов нейрона-победителя, который ближе всего расположен к предъявляемому входному вектору, смещается в направлении последнего, если входной вектор относится к одному с ним классу, и удаляется от него в противном случае [5]. Отличие LVQ 2.1 от LVQ 1 состоит в способе коррекции весов: вектор w_i и следующий w_j относятся к разным классам; вектор x принадлежит либо тому классу, к которому относится w_i , либо тому, к которому принадлежит w_j ; вектор x находится в некотором «окне» относительно перпендикуляра между этими двумя классами.

LVQ 3 является модификацией LVQ 1, позволяющей изменить не только веса векторов, но и границу между классами, если оба вектора принадлежат тому же классу, что и x . Метод OLVQ1 улучшает методы контролируемого обучения за счет ускорения скорости сходимости путем адаптации векторов весов и настраивания параметра обучения α .

3. Архитектура нейронной сети

Архитектура оптимизированной нейронной сети векторного квантования (OLVQ1) представлена на рис. 1

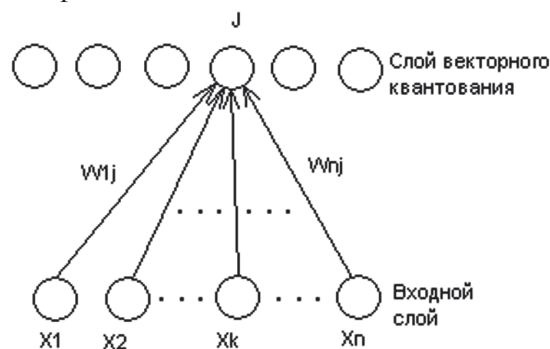


Рис. 1. Топология сети векторного квантования

Сеть состоит из двух слоев: входного слоя и слоя векторного квантования. Каждый нейрон входного слоя соединен со всеми нейронами слоя векторного квантования. Сила связи определяется соответствующим весом.

4. Алгоритм обучения сети

На первом этапе определяется нейрон-победитель путем вычисления евклидова расстояния, как на показано в формуле

$$\|x - w_c\|^2 = \min_j (\|x - w_j\|^2), \quad (1)$$

где x — вектор входов; w_c — вектор весов; w_j — вектор весов j нейрона Кохонена.

Веса выигравшего нейрона изменяются по следующему правилу в представленной ниже формуле [5]:

$$w_c(k+1) = \begin{cases} w_c(k) + \alpha_c(k) * [x(k) - w_c(k)], \\ w_c(k) - \alpha_c(k) * [x(k) - w_c(k)], \end{cases} \quad (2)$$

где $w_c(k+1)$ — следующее значение веса; $w_c(k)$ — предыдущее значение веса; $x(k)$ — значение входа, $\alpha_c(k)$ — коэффициент обучения. Первый вариант используется, если x и w_c принадлежат одному классу, в противном случае используется второй вариант

Параметр обучения $\alpha_c(k)$ вычисляется по формуле:

$$\alpha_c(k) = \frac{\alpha_c(k-1)}{1 + \alpha_c(k-1) * S(k)}, \quad (3)$$

где $\alpha_c(k)$ — следующее значение параметра обучения; $\alpha_c(k-1)$ — предыдущее значение параметра; а $S(k)$ вычисляется по формуле 3:

$$S(k) = \begin{cases} +1, \text{ если } w \text{ и } x \text{ принадлежат одному классу} \\ -1, \text{ в противном случае} \end{cases} \quad (4)$$

5. Применение нейронной сети векторного квантования на примере

На основании данных бухгалтерской отчетности по 5 фирмам решим задачу классификации при помощи оптимизированной нейронной сети векторного квантования.

Представленный пример выступает скорее в роли демонстрационного, поскольку построен на небольшом количестве данных.

Таблица 1
Обучающая выборка

Фирмы	Коэффициенты Альтмана					Классы
	0,22	0,2	0,28	1,49	1,37	
X1	0,22	0,2	0,28	1,49	1,37	1
X2	-0,05	0,14	0,06	0,42	0,74	2
X3	0,00	0,04	0,02	0,54	0,34	2
X4	0,08	0,01	0,08	3,58	0,52	1
X5	0,08	0,03	0,08	1,13	0,4	2

В столбце «Классы» класс 1 определяет отнесенные фирмы к не банкроту, а класс 2 — к банкроту. Мы разделяем поданную выборку на две части: коэффициенты фирмы X1 и X2 используем в качестве начальных весовых коэффициентов. Таким образом, архитектура сети для решения данной задачи представляется в следующем виде: слой входных векторов состоит из трех нейронов (X3, X4, X5) и

слоя векторного квантования, состоящего из двух нейронов, представляющих собой два класса — банкрот и не банкрот. Параметр обучения сети на начальном этапе принимается равным 0,3.

Применив формулы (1)–(5), мы увидели, что сеть обучилась за один такт. Полученные векторы весов позволяют правильно классифицировать фирмы: $w1 = (0,2523; 0,2439; 0,3262; 1,0076; 1,5662)$, $w2 = (-0,0193; 0,095; 0,054; 0,5823; 0,5787)$.

6. Нейронные сети в прогнозировании курса акций

Прогнозирование — одна из самых востребованных, но при этом одна из самых сложных задач интеллектуального анализа данных. Проблемы прогнозирования связаны с недостаточным качеством и количеством исходных данных, изменениями среды, в которой протекает процесс, воздействием субъективных факторов.

Начальным этапом в решении данной проблемы является исследование предметной области, что позволяет выделить значимые информационные ресурсы. Следующим этапом является выбор наиболее подходящего метода прогнозирования курса акций и в конечном итоге его программная реализация.

В данной статье приводится обоснование выбора лучшего из методов — прогнозирование с помощью нейронной сети GRNN. Результаты проведенной работы могут быть использованы для написания специализированного программного продукта с целью дальнейшего его использования на рынках ценных бумаг.

7. Сравнительная характеристика нейронных сетей

Задача прогнозирования курса акций может быть решена при помощи различных нейронных сетей. Наиболее подходящими для этой цели являются такие сети, как сеть обратного распространения (Backpropagation neural network), радиально-базисная сеть (RBFN), обобщенно-регрессионная сеть (GRNN), модульная сеть.

Сеть обратного распространения является самой первой сетью, в которой количество скрытых слоев было больше единицы. Эта сеть способна к успешной аппроксимации любой нелинейной функции преобразования. Она применима к данным с нелинейной зависимостью таким, как: хаотические временные ряды, зашумленные данные, финансовые временные ряды и регрессионные модели. Нежелательно использовать эту сеть при нестационарных данных или для данных, которые скрывают в себе влияние различных факторов, как в случае прогнозирования курса акций. В таком случае лучше применять GRNN, RBFN или модульные сети [6].

Преимуществами радиально-базисной сети является быстрая обучаемость и способность определять границы лучше, чем сеть обратного распространения при решении задачи классификации. К недостаткам можно отнести более медленную реакцию в фазе отклика. В начальной фазе обучения некоторая выборочная информация может быть утеряна. Также сложно определить число нейронов.

Сравнение с сетью обратного распространения показало аналогичную работу модульной сети, но дало лучшие результаты при вычислении ошибки. Принцип работы основан на том, что каждый модуль «участвует» в соревновании, и это позволяет ускорить работу сети. Главным недостатком является сложная архитектура сети, что ведет к возникновению трудностей при программной реализации.

8. Характеристика обобщенно-регрессионной нейронной сети

Обобщенно-регрессионная нейронная сеть (англ. General Regression Neural Network) – разновидность радиально-базисных сетей, предназначена для решения задач регрессии с использованием ядерной аппроксимации. Обобщенно-регрессионная искусственная нейронная сеть в отличие от классического регрессионного анализа не требует задания функциональной зависимости, связывающей входные и выходные переменные. Все входные образы каким-либо образом разбиваются на определенное число кластеров. Если число образов невелико, каждый образ может образовать свой кластер [7].

GRNN состоит из четырех слоев: входные модули, которые образуют полную связь со слоем образов; слой образов; слой суммирования; выходной слой. Функцией активации служит Гауссова функция. ОРНС обучается с учителем.

Архитектура сети представлена на рис. 2 [7].

На входной слой подается обучающая выборка. При этом весовым коэффициентам присваиваются значения выборки и подаются на слой образов, в слое образов весовые коэффициенты проходят через функцию активации, которая имеет следующий вид:

$$\sum_{i=1}^n \exp\left(-\frac{D_i^2}{2\sigma^2}\right), \quad (5)$$

где D_i^2 — квадрат Евклидова расстояния;

$$D_i^2 = (X - X^i)^T (X - X^i); \quad (6)$$

σ — параметр сглаживания, для наиболее результативного значения сети выбирается в промежутке от 2 до 6. Преобразованные на предыдущем шаге значения подаются на слой суммирования, состоящий из двух нейронов: S -нейрон накапливает сумму взвешенных выходов слоя образов.

$$\sum_{i=1}^n Y^i \exp\left(-\frac{D_i^2}{2\sigma^2}\right), \quad (7)$$

D -нейрон вычисляет сумму невзвешенных выходов слоя образов

$$\sum_{i=1}^n \exp\left(-\frac{D_i^2}{2\sigma^2}\right), \quad (8)$$

В выходном слое рассчитывается взвешенное среднее по формуле (9):

$$\hat{Y}(X) = \frac{\sum_{i=1}^n Y^i \exp\left(-\frac{D_i^2}{2\sigma^2}\right)}{\sum_{i=1}^n \exp\left(-\frac{D_i^2}{2\sigma^2}\right)}. \quad (9)$$

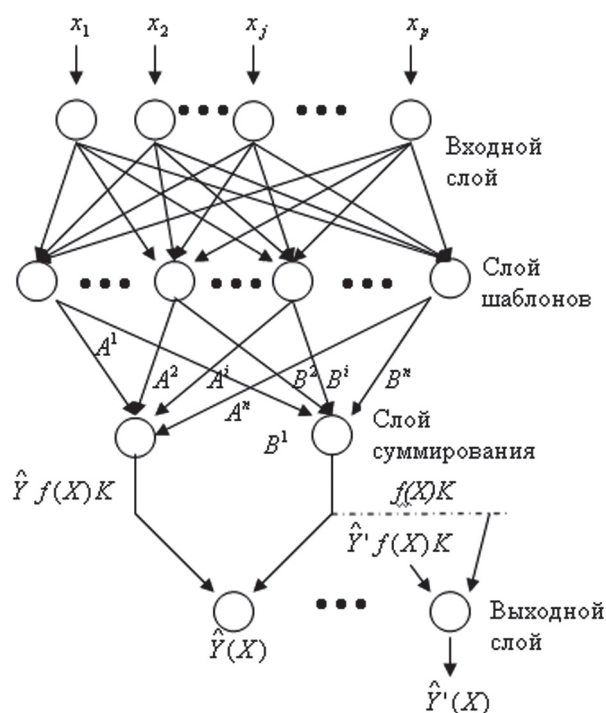


Рис.2. Архитектура сети GRNN

9. Применение обобщенно-регрессионной нейронной сети на примере

Входными данными к сети является выборка курсов акций за предыдущие дни, полученная с finance.yahoo.com. Выходом является прогнозируемый курс на следующий день.

В табл. 2 представлены значения обучающей выборки за 30 дней.

Таблица 2

Входная выборка

108,57	107,24	104,82	102,7	104,26	103,1
101,41	102,59	104,34	104,51	103,24	106,12
102,73	101,89	104,92	105,6	113,56	113,49
115,11	114,42	115,5	114,75	114,5	113,9
113	113,32	114,2	113,78	110,97	113,9

Настраиваемым параметром сети является параметр сглаживания σ . В таблице 3 представлены результаты работы сети при различных значениях параметра сглаживания. Учителем сети является значение 113,9.

Таблица 3

Результаты работы сети

$\sigma = 2$	$\sigma = 3$	$\sigma = 4$	$\sigma = 5$
113,69	113,78	114,08	114,4

Результаты показывают, что при $\sigma = 3$ прогнозируемое значение наиболее приближено к учителю.

Выводы и направление дальнейших исследований

В данной статье было предложено рассмотрение использования OLVQ1 для решения задачи классификации, состоящей в определении жизнеспособных и склонных к банкротству предприятий. Был проведен сравнительный анализ алгоритмов обучения для нейронной сети векторного квантования и выбран наилучший метод — оптимизированная нейронная сеть векторного квантования.

Была изложена общая теория сети векторного квантования, на основании которой был приведен небольшой пример, демонстрирующий возможность практического применения этой сети для решения задач классификации. Представленные результаты дают основания считать сеть векторного квантования достаточно хорошим методом интеллектуального анализа данных.

Также была рассмотрена проблема прогнозирования курса акций и предложено ее решение на основе обобщенно-регрессионной сети.

Направление будущих исследований заключается в изучении самоорганизующихся карт Кохонена, использовании различных моделей конкурентного обучения (правило жесткой конкуренции, справедливой конкуренции, мягкой конкуренции), практическом применении для решения задачи определения жизнеспособных и склонных к банкротству предприятий. Планируется сравнить полученные в ходе эксперимента результаты на примере сети Кохонена и сети вектора квантования. Результаты позволят сравнить две парадигмы обучения нейронных сетей: с учителем для сети OLVQ1 и самоорганизующихся карт Кохонена.

Для задачи прогнозирования курса акций планируется изучение и применение сети радиально-базисных функций с различными функциями активации и сравнение ее работы с работой обобщенно-регрессионной нейронной сети.

Список литературы: 1. *З.Ф. Уоссермен* Нейрокомпьютерная техника: Теория и практика, 1992. — 118 с. 2. *Заенцев И.В.* Нейронные сети основные модели. — Воронеж 1999. — 150 с. 3. *G. Peter Zhang.* Neural networks in business forecasting, October 2004, 287 pages. 4. *Amir F. Atiya.* Bankruptcy prediction for credit risk using neural networks: a survey and new results. // IEEE transactions on neural networks, Vol. 12, No 4, July 2001. 5. *Руденко О. Г., Бодянский Е.В.* Основы искусственных нейронных сетей. — Харьков. — 2003. — 323 с. 6. *Keun-Rong Hsieh, Wen-Tsuen Chen* A Neural Network Model which Combines Unsupervised and Supervised Learning // IEEE Trans. on Neural Networks. — 1993. — Vol.4, №2. 7. *Donald F. Specht* A general regression neural network // IEEE Transactions on neural networks. — November 1991. — Vol. 2, № 6

Поступила в редколлегию 28.04.2008