

# Нейросетевая технология в прогнозировании конъюнктуры фондового рынка\*

А. В. Пекарский

## Общие методические положения

Динамика сложной системы обычно подвержена нелинейным влияниям множества факторов, о большинстве из которых априорные знания отсутствуют. Повысить эффективность ее исследования способен метод моделирования, управляемый данными о таком процессе и имеющий возможность справляться с шумами, присутствующими в этих сведениях, т. е. адаптивный и, таким образом, более гибкий, чем экспертные системы, построенные в соответствии с фиксированными правилами.

Подобным методом является технология нейронных сетей [1, 3]. Такая сеть — это совокупность организованных в устройство параллельных вычислений взаимосвязанных элементов — нейронов. Каждый из них, по сути — простой процессор, который имеет дело только с сигналами, поступающими в него и передаваемыми от него другим нейронам. Вычислительные возможности всякого нейрона позволяют ему, пользуясь определенным правилом, комбинировать значения входных сигналов, затем на основе заданной активационной функции рассчитывать выходной сигнал и распространять его дальше по сети. Это распространение происходит посредством взвешенных связей, в результате чего выходной сигнал может либо усиливаться, либо подавляться.

Схематически структура нейрона представлена на рис. 1. Нейрон принимает значения входного вектора  $x = (x_1, x_2, \dots, x_n)$ , производит синоптическую операцию путем умножения входного вектора на вектор весов  $w = (w_1, w_2, \dots, w_n)$ . Далее нейрон осуществляет операцию агрегирования со смещением, причем часто оно рассматривается как дополнительный входной элемент  $x_0$  с весом  $w_0$  [3]. Результат такого вычисления используется в функции активации, которая и дает выход нейрона:

---

\* Работа выполнена при финансовой поддержке Российского Гуманитарного Научного Фонда (проект 05-02-021800).

$$y = \sigma \left( \sum_{i=0}^n x_i w_i \right), \quad (1)$$

где  $y$  — выход нейрона,  
 $x_i$  — значение  $i$ -го входа нейрона,  
 $w_i$  — вес  $i$ -й связи нейрона,  
 $\sigma(\cdot)$  — функция активации (решающая функция).

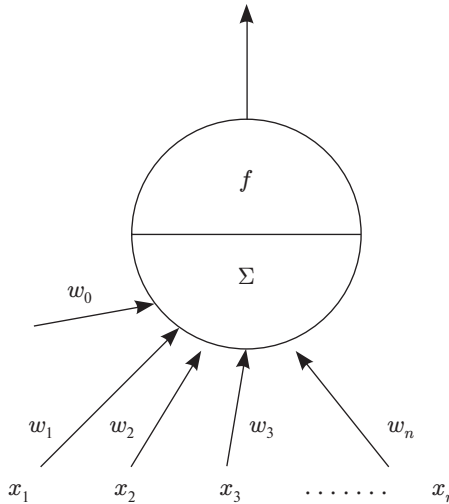


Рис. 1. Нейронные операции

Эта функция производит отображение  $n$ -мерного входного вектора в выходной сигнал. Активность нейрона может представляться либо некоторым непрерывным действительным значением, либо какой-то величиной из определенного дискретного набора. Обычно применяются нелинейные функции активации, но возможны и линейные [1]. Они бывают простыми (при которых выходной сигнал равен взвешенной сумме входов нейрона), с насыщением (когда выход ограничен двумя крайними значениями), ступенчатыми (выход нейрона принимает одно из нескольких последовательных значений). Такие функции чаще всего используются в выходном слое нейронной сети.

Нейроны скрытых слоев сети обычно активизируются монотонно возрастающими непрерывными сигмоидальными функциями [1, 3] типа

$$\sigma(V) = \frac{1}{1 + e^{-bV}}, \quad (2)$$

где  $V$  — взвешенный вход нейрона, коэффициент  $b$  определяет крутизну сигмоида. Значения данной функции находятся на интервале  $(0; 1)$ . Выходные значения от  $-1$  до  $1$  выдает функция

$$\sigma(V) = \frac{e^{-bV} - 1}{e^{-bV} + 1}. \quad (3)$$

Также часто используемые сигмоидальные функции —

$$\sigma(V) = \frac{e^{bV} - e^{-bV}}{e^{bV} + e^{-bV}}, \quad (4)$$

$$\sigma(V) = \frac{1 - e^{-bV}}{1 + e^{-bV}}, \quad (5)$$

$$\sigma(V) = \frac{2}{\pi} \operatorname{arctg} \left( \frac{\pi}{2} V \right), \quad (6)$$

$$\sigma(V) = \frac{V^2}{1 + V^2} \operatorname{sign}(V), \quad (7)$$

где  $\operatorname{sign}(\cdot)$  — знаковая функция.

Конструкция нейронной сети определяется структурой связей в ней, их направлением и весом, а решаемая задача описывается в терминах весовых значений связей между элементами. Структура сети обычно формируется в два этапа: разработчиком, устанавливающим, какие элементы и в каком направлении связаны с помощью межуровневых, внутриуровневых либо рекуррентных соединений, а затем обучением сети выясняются значения соответствующих весовых коэффициентов.

Конструкция применяемой нейронной сети во многом определяется сущностью проблемы, которую предстоит решить. При архитектуре самого общего типа все нейроны сети связаны друг с другом, и в ней отсутствуют обратные связи. Согласно результатам ранее проведенных исследований [1, 3], нейронные сети с прямой связью и сигмоидальными активационными функциями являются универсальным средством аппроксимации различных сложных зависимостей и могут применяться практически везде, где обычно использовались линейные модели либо оценивание при помощи стандартных статистических методов. При этом, однако, не существует правила, позволяющего найти оптимальную топологию сети для конкретной задачи.

Количество входных и выходных элементов сети, как правило, определяется числом признаков и требуемых выходных переменных. Размеры скрытого слоя находятся экспериментально. При этом начинают с одного скрытого слоя размером в 30–50 % числа входных элементов. Нередко

такая конфигурация позволяет достигать хорошей точности в решении проблемы. Далее, если это повышает его эффективность, возможно добавление элементов в данный скрытый слой и новых таких слоев.

Значения весов связей между элементами нейронной сети, в отличие от запрограммированности, присущей самостоятельному определению зависимостей методами математической статистики, определяются в процессе ее обучения на известных данных. Решением множества задач, в каждой из которых верный ответ выясняется с помощью сгруппированного в обучающее множество примеров, посредством изменения весов достигается цель минимизировать функцию ошибок на этом множестве. Зависимость невязки от весов нелинейна, а потому решение такой оптимизационной задачи в аналитической форме невозможно, и минимум находится через итерационный процесс — обучающий алгоритм. В настоящее время существует более сотни различных обучающих алгоритмов, различающихся критериями ошибки и стратегиями оптимизации.

Часто в качестве такого критерия используется средняя квадратичная ошибка (MSE) [1, 3]:

$$E = \frac{1}{P} \sum_{k=1}^P (d_k - y_k)^2, \quad (8)$$

где  $P$  — число примеров в обучающем множестве,  
 $d_k$  — требуемый выход.

Применяется и расстояние Кульбака—Лейблера, связанное с критерием максимума правдоподобия:

$$E = \sum_{k=1}^P \left[ d_k \log \left( \frac{d_k}{y_k} \right) + (1 - d_k) \log \left( \frac{1 - d_k}{1 - y_k} \right) \right]. \quad (9)$$

Ошибка минимизируется посредством градиентных методов. Веса пересчитываются после обработки всего множества обучающих примеров (эпохи) и изменяются в обратном направлении по отношению к наибольшей крутизне для функции стоимости:

$$w_{i+1} = w_i - \varepsilon \frac{\delta E}{\delta w}, \quad (10)$$

где  $\varepsilon$  — коэффициент обучения, или величина градиентного шага.

Наиболее распространенный способ обучения нейронной сети — алгоритм обратного распространения ошибки (BP). Он представляет собой развитие обобщенного дельта-правила [3] и минимизирует суммарную квадратичную ошибку:

$$E = \frac{1}{2} \sum_{k=1}^P \sum_i (d_k^i - y_k^i)^2, \quad (11)$$

где  $i$  проходит все выходы сети.

Чувствительность ошибки сети к изменениям весов изменяется частными производными. Ошибка каждого предыдущего слоя вычисляется через аналогичный показатель последующего и, таким образом, передается в обратном направлении. Весовые значения скрытых слоев корректируются в соответствии с их вкладом в величину ошибки последующего слоя, зависящим от величины ошибки и веса соединения.

Алгоритм обратного распространения используется как в «эпохальном» виде (веса меняются после просчета всего обучающего множества), так и в «стохастическом» (веса пересчитываются всякий раз после прохождения через сеть очередного образца). Значения весов при создании нейронной сети инициализируются малыми случайными величинами. Обучение продолжается, пока изменение среднеквадратичной ошибки при переходе к следующей эпохе не окажется меньше некоторого значения. Для уменьшения вероятности того, что изменения весов приобретут осциллирующий характер, вводится инерционный член, зависящий от предыдущего изменения веса:

$$\Delta w_{t+1} = \mu \Delta w_t - (1-\mu) \varepsilon \frac{\delta E}{\delta w}, \quad (12)$$

где  $\mu$  — задаваемое пользователем число из интервала (0; 1).

Метод инерционного члена весьма чувствителен к способу упорядочения данных в обучающем множестве. При их совпадении несколько раз подряд импульс резко возрастает, а это может привести к нарушению процесса обучения. Поэтому упорядочивать последовательность обучающих примеров необходимо случайным образом, что особенно важно при стохастическом варианте изменения весов.

Сеть, образованная нейронами с радиальными базисными функциями, обучается по иному. В отличие от рассмотренной многослойной сети с обратным распространением ошибки, функционирование нейронов которой основано на оценке степени близости текущего примера тому, что знает обученная сеть, нейроны сети на радиальных базисных функциях определяют степень взаимного различия между нейронным входным вектором и знанием, хранящимся в синаптическом весовом векторе [1]. Подобная сеть обычно имеет три слоя — входной, который распределяет данные образца для первого слоя весов, скрытый и выходной. Отображение входного слоя в скрытый нелинейно, а скрытого в выходной — линейно. Идея сети этого типа в том, что если разместить проблему нелинейным образом в более

многомерном пространстве, то высока вероятность ее превращения в линейно отделимую. Поэтому чаще всего количество скрытых элементов в такой сети больше числа входов. Веса связей ее входных и скрытых элементов определяют центр радиальной функции для каждого скрытого элемента. Комбинированный ввод скрытых элементов равен евклидовой норме:

$$V = \sqrt{\sum_{i=1}^n (x_i - w_i)^2}, \quad (13)$$

где  $n$  — число входных элементов.

Нелинейная активационная функция порождает выход, который является максимальным при нулевой разности (наибольшем сходстве) и уменьшается при больших разностях. Функция активности скрытого элемента может быть, например, гауссова:

$$\varphi(V) = e^{-V^2 / \sigma^2} \quad (14)$$

или

$$\varphi(V) = \sqrt{c^2 + V^2}. \quad (15)$$

Обучение нейронной сети с радиальными базисными функциями требует, чтобы первый слой весов перед началом тренировки был задан случайным выбором центров функций каждого скрытого элемента. Далее такие веса корректируются с помощью неуправляемого обучения [3], причем второй слой — по правилу Видроу—Хоффа:

$$E = d - y. \quad (16)$$

Это делается следующим образом:

$$\Delta w = \epsilon EV, \quad (17)$$

где  $V$  — комбинированный ввод выходного элемента.

Тренируя нейронную сеть, можно столкнуться с положением, когда она достигает высокой степени соответствия аппроксимируемой функции на обучающем множестве, однако при работе с последующими данными показывает крайне низкую степень обобщения. Такая ситуация называется переобучением: сеть просто начинает запоминать данные из обучающего множества, моделируя уже не столько саму функцию, сколько присутствующий в этих данных шум. Проблема переобучения связана с общим вопросом, как отделить сигнал от шума. Ее актуальность тем ниже, чем больше данных в выборке. Их представительность растет, а шум умень-

шается с увеличением числа наблюдений. Но на практике зачастую сложно получить требуемое количество примеров. Поэтому наиболее распространенный способ избежать переобучения — использование подтверждающего множества. Для этого часть примеров из обучающего множества относят к множеству подтверждения и суждение, достигла ли сеть стадии переобучения или нет, основывается на том, сколь последовательно снижается среднеквадратичная ошибка на подтверждающих данных. Если ошибки на обоих типах примеров снижаются, то тренировка сети продолжается до тех пор, пока не начнет расти ошибка в моделировании проверочного множества. Применяя этот метод, приходится отказываться от 20–30 % примеров, которые могли бы служить для тренировки сети.

Обработку временных рядов нейронная сеть с прямой связью осуществляет с помощью разветвленной линии задержки [1]. Суть дела в том, чтобы разбить временной ряд на несколько отрезков и подать образующийся образец на вход этой сети. Имея такую архитектуру, она может моделировать любую конечную временную зависимость вида  $y(t) = F[x(t), x(t-1), \dots, x(t-k)]$ . Кроме того, на вход сети может подаваться не одна, а несколько переменных, также проходящих через линию задержки. Поскольку у сети с данной конфигурацией отсутствуют рекуррентные связи, для ее обучения может использоваться обычный алгоритм обратного распространения.

Прогнозирование временного ряда на несколько шагов вперед требует серии итераций, когда в качестве недостающего входа сети используются ее выходы, полученные на предыдущем этапе. Самый распространенный способ тренировки нейронной сети при этом — алгоритм форсированного управляемого обучения. Он подает на вход сети пример, состоящий из «х» исходных данных и «у» выходов сети, причем  $(x + y)$  равняется числу входов для одного примера ( $z$ ). Обычно тренировка начинается форсированием всех входов из исходных наблюдений (т. е.  $x = z$  и  $y = 0$ ). После каждой эпохи обучения количество форсированных входов уменьшается на некоторую величину, например 1 % ( $x = 0,99 \cdot z$ ,  $y = 0,01 \cdot z$ ). В конце концов, наступает момент, когда число форсированных входов становится равным нулю и все недостающие входы для итеративного прогнозирования сеть берет из своих выходов на предыдущих итерациях.

## **Применения нейронной сети в прогнозировании фондового рынка**

Рассмотренную технологию можно эффективно использовать в прогнозировании конъюнктуры рынка ценных бумаг, повышая достоверность ее многофакторных статистических моделей. Адаптивная, самообучающаяся нейронная сеть позволяет преодолевать трудности их спецификации, выбора значимых для них переменных, поиска адекватных форм этих

зависимостей при недостаточно обширной базе наблюдений и возникновении уникальных событий, которые, сильно влияя на рынок, делают информацию о прошлом малопригодной для построения подобных моделей, которые бы обладали высокой прогностической способностью. Это продемонстрировано в [1], где показано, что, применив вместо регрессии нейронную сеть, удалось при прогнозировании по 33-м переменным индекса курсов акций (на основе теории арбитражного ценообразования) и цен опционов на них добиться шестикратного снижения среднеквадратичной ошибки прогноза.

Нейросети довольно продуктивно применяются для предвидения будущей ценовой динамики фондового рынка и в связке с моделями прогнозирования его конъюнктуры, основанными на постулате о достаточности информации, содержащейся в прошлых значениях курсов ценных бумаг. Этим моделям (графическим и математическим индикаторам технического анализа), ориентированным на неизменность человеческой психологии и повторяемость ситуаций, которые складываются на рынке, по силам подавать сигналы о наличии либо отсутствии на нем господствующей тенденции, ее направлении, стадии развития и вероятном моменте разворота [4, 5]. Но они способны обеспечивать лишь качественное предвидение курсовой динамики, количественно не определяя будущие изменения цен фондовых активов.

Однако, натренировав нейронную сеть на исторических значениях этих цен, можно успешно предвидеть как развитие курсовой тенденции, так и их численные значения. Посредством подобных экспериментов, описанных в [6], было достигнуто ощутимое снижение ошибки прогнозов по сравнению с регрессионной моделью и повышение (в рамках имитационного тестирования их качества) эффективности инвестирования при его ориентации на сигналы выходов соответствующей нейросети.

Пропуская через сеть не оригиналы цен, а технические индикаторы рынка (чтобы содействовать поискам с ее помощью закономерностей в используемых данных за счет их предварительного преобразования, часто рекомендуемого в литературе [1, 3]), удалось удостовериться в следующем. Сеть не улучшает экономические результаты применения для ориентации торговых сделок скользящей средней самих по себе цен, но использование на ее входе ряда разностей цен и значений скользящей средней делает более точным прогноз направления рыночного тренда на один-два периода вперед, а обучение сети на предсказание «Момента» как технического индикатора разворота тенденции и совершение сделок по его сигналам повышает их эффективность.

Представляется возможным, преобразуя ценовые данные при помощи таких нелинейных сглаживающих функций, как сплайн и волновая функция [2], а затем, подавая на вход нейросети полученные при оценке их конфигурации определенные показатели, добиваться большей информа-



тивности ее выходных значений. В [2] сплайн уже использовался для отображения господствующей рыночной тенденции, а краткосрочные движения рынка оценивались с помощью спектрального анализа, причем перед ним из ценовой динамики делением ценового ряда на сплайн и вычитанием среднего удалялся тренд. Там же обоснована целесообразность еще одного сглаживания сплайном для устранения случайных колебаний и предотвращения проблемы «рассеивания» частот. После всех этих сглаживаний получаются кривые, представленные на рис. 2.

## Испытание возможности

Допустим теперь, что на вход нейронной сети фондового рынка подаются показатели, получаемые от совместной оценки динамики цен и упомянутых индикаторов. Источниками информации для такой сети способны быть направления соответствующих кривых, углы их наклона, взаимное расположение цен и сплайна, величина их расхождения в масштабе стандартного отклонения остатков и пр. На выходе сети можно применить как один нейрон с непрерывной, например линейной либо сигмоидальной,<sup>1</sup> или ступенчатой функцией активации, так и несколько бинарных нейронов (по числу классов, которые требуется идентифицировать).

Проверка эффективности этого требует на исторических курсовых данных провести эксперименты, в ходе которых оценить с точки зрения доходности рыночных операций сплайн и волновую функцию, выбрав для

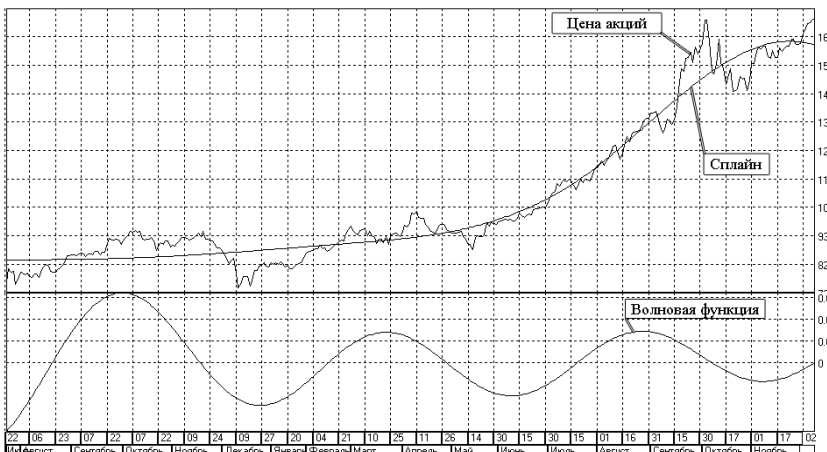


Рис. 2. Графики цен акций ЛУКОЙЛ, сплайна и волновой функции

<sup>1</sup> В таком случае нужно будет сопоставить некоторое значение этой функции определенному классу выходных сигналов.

каждого рассматриваемого выпуска акций наилучшие параметры этих функций. Далее, необходимо, используя их, определить переменные, которые будут подаваться на входы нейронной сети и, таким образом, число этих входов, выбрать некоторую часть данных в качестве обучающего множества и провести на этом отрезке тренировку нейронной сети, нацеливая ее на классификацию ситуации на входе как принадлежащей некоторому из заданных классов, в числе которых возможны как комбинация из «повышение», «понижение», «нейтрально», так и более сложно отделимые множества, отражающие вероятную степень будущего изменения (например «изменение более чем на один процент», «изменение менее процента», «не покроеет транзакционные издержки», и др.). Затем следует проверить работоспособность сети на оставшихся для тестового множества данных и оценить, повысило ли ее применение эффективность торговой системы по сравнению с самостоятельным использованием для ориентации инвестиционных сделок сигналов аналитических функций, действовавших в обучении этой сети.

Такое исследование было выполнено на основе исторических данных о ценах акций, обращавшихся на ММВБ в период с 30 апреля 2002 до 22 марта 2007 г. Первые 900 значений этих цен (по 8 декабря 2005 г.) использовались для поиска и обучения таких нейронных сетей, ориентация на показания которых при осуществлении торговых операций с рассматриваемыми акциями могла бы превзойти по надежности торговые сигналы используемых как входы этих сетей индикаторов. На оставшихся 315 значениях данных оценивалась способность обученных нейронных сетей оставаться эффективными за пределами отрезка, на котором они тренировались. Исследование проводилось среди класса многослойных нейронных сетей с прямыми связями, обучаемых посредством алгоритма обратного распространения ошибки.

Первая задача, которую пришлось решить — выбор входных и выходных переменных нейронной сети. С целью проверить ее способность получать полезную информацию из динамики предложенных сглаживающих функций (сплайновой и волновой) в первом опыте были использованы шесть входных переменных: три сплайна и три волны со случайным выбором параметров. Чтобы помочь сети оценивать динамику и не загружать ее информацией об абсолютных величинах, избыточной в данном случае, значения функций перед входом в сеть преобразовывались с помощью алгоритма насыщения (каждому значению функции в момент  $t$ , большему, чем в момент  $t - 1$ , он ставил в соответствие 1, а в противоположной ситуации  $-1$ ). Аналогично преобразовывалась и выходная переменная, в качестве которой в этом опыте выступала цена ближайшего будущего периода (если завтрашняя цена превышала сегодняшнюю, то на выходе сети показывалась 1, а иначе  $-1$ ). Для соответствия требуемому

выходу за функцию активации выходного нейрона сети была принята функция насыщения (дающая 1, если вход больше нуля, и  $-1$  в противном случае). Функцией активации единственного скрытого нейрона рассматриваемой сети (рис. 3) служил сигмоид.

Сеть обучалась на данных о котировках обыкновенных акций РАО ЕЭС России с помощью алгоритма обратного распространения в его стохастическом варианте (при этом перед каждой эпохой обучения примеры из тренировочного множества перемешивались случайным образом). Норма обучения была установлена на уровне 0,5, а инерционный член равнялся 0,25. После обучения в течение тысячи эпох, сеть показала довольно неважные статистики:  $MSE = 2,064$ , корреляция выходных значений с требуемыми выходами =  $-0,048$ . Но при этом испытание на реальной рыночной информации выявило ее способность вычленять из исходных данных информацию, полезную для повышения эффективности инвестиционных действий на фондовом рынке.

Испытание являлось процессом, при котором последовательно, период за периодом отслеживалась цена акции, рассчитывались значения входных переменных, производился прямой проход нейронной сети и в случае

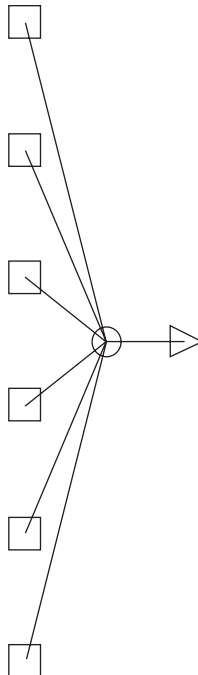


Рис. 3. Нейронная сеть с шестью входами, одним скрытым нейроном и одним выходом

поступления от нее сигнала на покупку (продажу), который в этой конфигурации сети соответствовал выходу равному единице (минус единице) короткие (длинные) позиции, если они были, закрывались, а покупка (продажа) акций осуществлялась по текущей цене. Таким образом, капитал инвестора постоянно изменялся, а отношение разности величин этого капитала в конечном и начальном периодах к его исходному размеру рассматривалось как критерий эффективности торговой системы. Она, основанная на сигналах только выходов нейронной сети, по этому признаку сравнивалась с системой, в которой источниками торговых сигналов выступали индикаторы, являющиеся входами сети, благодаря чему удавалось оценить продуктивность ее применения.

Испытание на данных о тех же акциях, на которых производилось обучение сети, и на том же временном интервале показало, что она способна подавать более точные торговые сигналы, чем индикаторы, выступающие ее входами. Доходность за весь период, полученная от ориентации при совершении сделок на сигналы нейросети, составила 831 %, а лучший результат, показанный индикаторами-входами, оказался равным 401,6 % (это при совершении сделок по сигналам одного из сплайнов — покупка при его росте и продажа при падении, — а совместное использование всех шести индикаторов оказалось чреватым полной потерей начального капитала). Положительный результат дала оценка еще более важной способности сети — оставаться эффективной за пределами обучающего периода: за сопоставимое время доходность операций, ориентированных на показания выходов нейросети, составила 3453,8 % против 1436,2 % как доходности операций, опиравшихся на сигналы самого лучшего сплайна.

Анализ причин этого показал, что в течение временного интервала, на котором проводилось обучение сети, она подавала разные торговые сигналы, а затем открыла длинную позицию и удерживала ее все оставшееся время. Возник вопрос, «сознательно» ли приняла сеть такое решение, правильное на рынке в основном растущем, или просто «не знала», что делать дальше? Поэтому, чтобы определить, сможет ли торговля в соответствии с сигналами нейросети превзойти по доходности пассивную стратегию «купи и держи», в качестве входных индикаторов далее были использованы такие, расчет которых производился на основе параметров, оптимизированных для выпуска акций, использованного в процессе испытаний [2]. А с целью содействовать сети отфильтровывать шум в данных и не обучать ее закрывать выгодную позицию при малейшем движении цены против доминирующей тенденции, за выход были приняты значения цен, сглаженные посредством скользящей средней (также с оптимизированным параметром). Предварительное преобразование и функция активации в выходном нейроне не отличались от предыдущего эксперимента.

Входами первой из рассматриваемых далее сетей (рис. 4) выступали три переменные. Первая — это сплайн, предварительно преобразованный по алгоритму насыщения, вторая — разность сплайна и цены, масштабированная так, чтобы ее значения попадали в область определения сигмоида. Использование абсолютных значений в этой переменной позволяло сети определять, насколько велико расхождение цены и сплайна (чем оно больше, тем выше вероятность их сближения). Третьей переменной служило значение цены, преобразованное по алгоритму насыщения в ряд, состоящий из чисел 1 и  $-1$ , в зависимости от последнего изменения. Три скрытых нейрона активировались функцией насыщения, давая на выход 1 или  $-1$  в зависимости от знака взвешенных входов.

После обучения эта сеть показала MSE 1,05 и корреляцию с требуемыми выходами на уровне 0,46. Имитация торговли по сигналам нейросети на временном интервале, который использовался для обучения, показала очень незначительное ее преимущество перед результатом, показанном при торговле с оглядкой только на используемый на входе сплайн. После включения в тестирование оставшихся 315 данных торговля и по сигналам сети и по направлению сплайна стала приносить убытки, а за весь период по доходности одна превзошла другую лишь на 3 пункта.

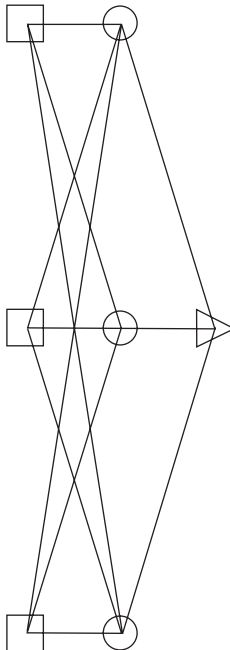
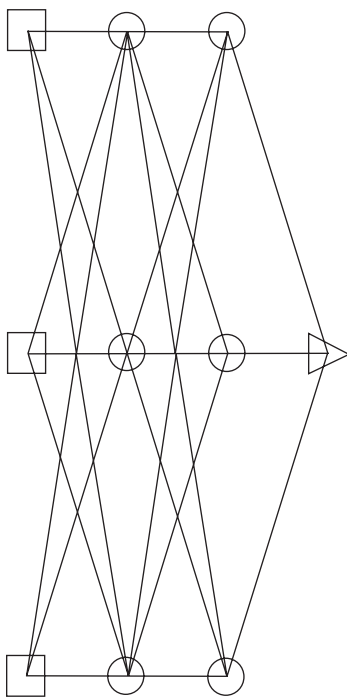


Рис. 4. Нейронная сеть с тремя входами, тремя скрытыми нейронами и одним выходом

Добавление к этой сети еще одного скрытого слоя с тремя нейронами (рис. 5), которые активировались опять функцией насыщения, и проведение дополнительного цикла обучения понизило MSE до 1,04, оставив корреляцию такой же, а имитация показала снижение эффективности сети и на тренировочном отрезке, и на тестовом.

Затем в надежде, что это поможет сети делать более точные прогнозы, к ее предыдущим входным переменным были добавлены еще две предварительно преобразованные масштабированием: соотношение цен текущего и предыдущего периодов, а равно значений в них сплайновой функции. Выходной переменной была принята разность между значением скользящей средней на три периода в будущее и в текущем, снова преобразованная функцией насыщения. В лучшей из сетей с этими входами и выходом оказалось девять скрытых нейронов в трех слоях, активирующихся функцией насыщения. После тренировки MSE оказалась 1,01, а корреляция выходов сети с требуемыми — 0,48. Имитация торговли не выявила способности такой сети превзойти результаты ее входного сплайна ни на тренировочном, ни на тестовом интервале.

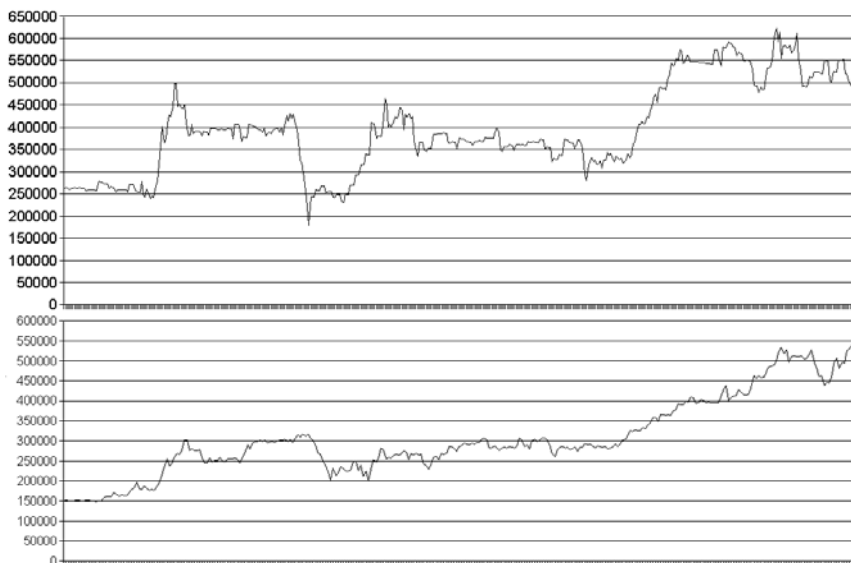


**Рис. 5.** Двухслойная нейронная сеть с тремя входами, шестью скрытыми нейронами и одним выходом

Далее, к трем прежним входам нейросети (приростам сплайн-функции, расхождению цены со сплайном и приростам цен) был добавлен четвертый. Им послужила волновая функция, построенная по остаткам, которые получаются от деления несглаженного ряда цен на сплайн (согласно [2], вторая с точки зрения торговой результативности). Эта волна предварительно преобразовывалась алгоритмом насыщения, чтобы оценивать только характеристики «рост-падение». Выходной переменной было значение скользящей средней на один период в будущее, также преобразованное насыщением. Лучшая из таких сетей содержала два скрытых слоя, по три нейрона в каждом, которые активировались функцией насыщения. После обучения эта сеть показала  $MSE = 1,1$  и корреляцию с требуемыми выходами в размере 0,43.

Испытание данной сети имитацией инвестиционных действий на тренировочном периоде дало за все это время превышение доходности от операций по ее сигналам над доходностью сделок по сигналам волновой функции в 1,79 раз, а над их доходностью по совместным сигналам волны и сплайна в 1,42 раза. При добавлении тестового временного интервала эксперимент засвидетельствовал, что при использовании и сети, и входных индикаторов без нее доходность продолжала расти, но применение только индикаторов приводило к улучшению результатов в конце периода. Анализ динамики капитала на тестовом отрезке данных (рис. 6) позволил выяснить, что исполнение сигналов нейронной сети влечет более резкие изменения его размеров (в случаях как роста, так и падения). При этом почти на всем периоде испытаний объем инвестиционного портфеля в ситуации торговли по сигналам сети превосходит его размеры в других ситуациях, но в конце положение меняется. Возможно, предотвратить это способно периодическое переобучение нейросети после каждого отрезка роста портфеля.

Входы рассмотренных нейронных сетей были также пополнены еще одним — волновой функцией, построенной по частному от деления цен, сглаженных при помощи сплайна с коротким расстоянием между узлами, на сплайн с более широким расстоянием (этот индикатор продемонстрировал наилучшую эффективность при ранее выполненных торговых экспериментах [2]). Предварительно такая входная переменная прошла преобразование посредством алгоритма насыщения. Выход сети не менялся. В ней использовались три скрытых нейрона с их активацией через насыщение. Обучение такой сети привело к  $MSE = 1,48$  и корреляции на уровне 0,21. Имитационный эксперимент на периоде, использованном для обучения, показал превосходство по доходности операций, ориентированных на сигналы нейросети, над сделками по сигналам ее входов в 1,71 раза. После охвата им данных интервала, оставшегося для тестирования, это преимущество достигло 8 с лишним раз, причем на тестовом периоде сеть снова рекомендовала инвестору стратегию «купи и держи».



**Рис. 6.** Динамика капитала при торговом эксперименте на тестовом периоде (вверху — при использовании нейросети, внизу — при ориентации на сигналы волны)

Добавление к такой сети еще одного скрытого слоя с тремя нейронами, которые активировались функциями насыщения, в итоге нового цикла ее обучения привело к снижению MSE до 1,39 и росту корреляции до 0,26. Однако, доходность на тренировочном периоде по сравнению с доходностью в предыдущем случае снизилась в 1,4 раз. То же самое произошло и на тестовом при том, что сеть опять все время сигнализировала инвестору покупать. Следствием еще одного цикла ее обучения в течение тысячи эпох стало снижение MSE до 1,26 и повышение коэффициента корреляции до 0,34. Результаты имитации торговли в период обучения продемонстрировали преимущество по доходности от исполнения сигналов этой сети более чем в 1,5 раза, над «советами» сети, которая использовалась до добавления второго скрытого слоя, и в 2,6 раза над «советами» индикаторов-входов. Сравнение на тестовом периоде динамики капитала при ориентации на сигналы однослойной и двухслойной сетей засвидетельствовало его более быстрый рост в последнем случае на протяжении примерно пятой части этого периода с возникновением затем потерь. Периодическое переобучение такой сети, возможно, позволило бы ей продолжать демонстрировать хорошие результаты.

\*

\*

\*



Проведенное исследование показало, что возможно построить и обучить нейронную сеть, способную превзойти по торговой эффективности те индикаторы, на основе которых она принимает решения. Все сети, показавшие лучшие результаты, содержали в скрытых слоях нейроны, которые активировались функциями насыщения (сигмоиды проявили себя хуже). Высокие значения доходностей при испытании нейросетей на тестовом периоде нельзя считать чистой подгонкой, поскольку последние сто данных из девятистот использовались для множества подтверждений и не принимали непосредственного участия в процессе обучения. А на тестовом интервале, включавшем 2006 г., удачный для российского фондового рынка, нейросети, согласно сигналам которых при имитации торговли совершались в это время сделки, без периодического переобучения не смогли выявить стратегию, по эффективности превосходящую решение удерживать длинную позицию. Чтобы определить, насколько «сознательно» это их решение, необходимы эксперименты на ином тестовом периоде.

Однако, в пользу нейронных сетей в этом аспекте говорит то обстоятельство, что даже сплайн и волновая функция как технические индикаторы, обладающие большим запасом прочности в том, что касается необходимости переоптимизации их параметров [2], не во всех конфигурациях порождения торговых сигналов смогли без нее приспособиться к полугодовому росту в тестовом периоде. В итоге, лучшие результаты показали сети, использовавшие как входы и сплайн, и волны, построенные на основе как однократного, так и двойного сглаживания.

Поиск наилучшей сети затрудняется тем, что при ее нацеливании на предвидение ценовой динамики фондовых активов ошибка такого прогноза после снижения до некоторого уровня начинает осциллировать вокруг него, и не всегда меньшая ошибка означает более весомую способность сети содействовать росту эффективности инвестиционных сделок. Поэтому в процессе обучения нейронных сетей, предназначенных для ориентации реальной торговли ценными бумагами, нужно быть весьма аккуратным, постоянно сопровождая это обучение проверкой посредством описанного имитационного эксперимента. А чтобы можно было рекомендовать инвесторам использовать подобные сети, необходима такая проверка на широком списке акций с выяснением истинной причины отсутствия в тестовом периоде сигналов к сделкам у сетей, показавших лучшие результаты. Кроме того, следует озаботиться ростом надежности процедуры поиска нейронной сети, эффективной в отношении торговых сигналов, возможно применяя при ее обучении алгоритм, изменяющий количество данных во временном ряду, поскольку и сплайн, и волновая функция изменяются на протяжении всей длины при добавлении каждого нового известного значения, а также проверяя, улучшат ли результаты новые классы выходов, например, ограничивающие торговую активность нейронной сети на бестрендовом рынке.

## Литература

1. *Бэстенс Д.-Э., Ван ден Берг В. М., Вуд Д.* Нейронные сети и финансовый рынок. М.: ТВП, 1997.
2. *Завельский М. Г., Пекарский А. В.* Оптимизация торговой системы для деятельности на фондовом рынке // Ежегодник «Системные исследования» 2002. М.: РАН, Эдиториал УРСС, 2004.
3. *Каллан Р.* Основные концепции нейронных сетей. М.: Издательский дом «Вильямс», 2003.
4. *Колби Р. В., Мейерс Т. А.* Энциклопедия технических индикаторов рынка. М.: Альпина, 1998.
5. *Найман Э. Л.* Малая энциклопедия трейдера. Пер. с англ. Киев: Альфа Капитал, Логос, 1997.
6. *Пекарский А. В.* Применение нейросетевой технологии к моделированию фондового рынка // Информационные технологии системного формирования хозяйственной политики государства. М.: РАН, Эдиториал УРСС, 2004.