

ФИНАНСОВЫЕ РЫНКИ

Е. В. Дорохов

канд. экон. наук, преподаватель кафедры статистики и эконометрики Московского государственного университета им. М. В. Ломоносова

О РЕЗУЛЬТАТАХ ПРИМЕНЕНИЯ НЕЙРОСЕТЕВЫХ МЕТОДОВ ДЛЯ АНАЛИЗА РАЗВИТИЯ ФОНДОВОГО РЫНКА РОССИИ

Прогнозирование само по себе еще не выделено в качестве предмета рассмотрения отдельной научной дисциплины, оно является всего лишь имманентной функцией каждой из существующих наук. При наличии объекта исследования и рассмотрении какой-либо проблемы будущего каждая из таких проблем уже является предметом конкретной науки, а не абстрактной «науки о будущем». Эти слова, однако, не следует понимать таким образом, что при прогнозировании в различных областях науки отсутствуют какие бы то ни было общие междисциплинарные положения. В частности, в последние десятилетия был установлен ряд фундаментальных ограничений на предсказуемость сложных систем, к которым, безусловно, можно отнести фондовый рынок. В 1970-е гг. результаты исследований ученых-математиков в области нелинейной динамики и динамического хаоса развеяли иллюзию глобальной прогнозируемости, подтвердив принципиальную невозможность предсказания, начиная с некоторого горизонта прогноза (интервала времени, на который может быть дан прогноз), поведения многих достаточно простых систем (в частности, непериодического маятника). Важнейшую роль в этих теоретических изысканиях сыграло свойство систем, состоящее в чувствительности к начальным данным. Любые долгосрочные прогнозы для подобных систем должны иметь дело только со статистикой. Однако нелинейная динамика позволила увидеть не только принципиальные трудности, но и новые замечательные возможности. Например, несмотря на то что формально сложные системы имеют бесконечно много степеней свободы (величин, которые необходимо задать для определения состояния системы), нелинейная динамика, анализируя системы такого рода, позволяет устанавливать, сколько переменных необходимо для их описания, сколько переменных нужно для прогнозирования, она помогает выяснить, каким должен быть их мониторинг. Оказывается, что в практических задачах нужно не более десятка переменных. Это открывает совершенно новые возможности — у нас есть формально очень сложная система (фондовый рынок), но теоретически оправданно в своем исследовании двигаться по пути выделения из нее самого главного.

Несмотря на то что с момента начала бурного развития прогностических исследований в области теории прогнозирования было разработано достаточно большое количество соответствующих методов (по некоторым оценкам — более 150 (Глуценко, 2000)), к настоящему времени по-прежнему отсутствует их общепринятая классификация. Тем не менее, систематизируя сведения, приводимые

в различных источниках (Арженовский, Молчанов, 2001; Дуброва, 2003; Лугачев, Ляпунцов, 1999), можно условно иметь в виду обобщенный вариант группировки методов прогнозирования, выделяя в порядке уменьшения формальной строгости и, соответственно, расширения области использования следующие основные типы моделей: математические (в том числе статистические), эвристические (в том числе модели на основе нейронных сетей и генетических алгоритмов) и экспертные модели.

Наиболее широко используемыми при прогнозировании моделями из класса эвристических являются *нейросетевые модели*¹. Концепция нейросетевых моделей заключается в моделировании процесса работы человеческого мозга при принятии решений².

Нейронная сеть представляет собой общую схему метода решения вполне определенной типовой задачи. Что касается самой задачи, то в общем виде речь идет о приближенном восстановлении отображения по заданному ряду его значений. Предположим, что имеются наборы чисел (входные векторы) и для каждого из них известно значение функции, которое она принимает на данном наборе (речь идет об исторических данных). Затем предъявляются новые данные, причем в этом случае результат уже неизвестен, и его необходимо (хотя бы приближенно) найти.

Процесс функционирования нейронной сети в этой ситуации можно описать следующим образом. Элементарная операция, которую она производит с данными, такова. Вначале берется взвешенная сумма входных величин. Затем полученная величина преобразуется с помощью нелинейной монотонной функции (функции активации) таким образом, чтобы получившееся в результате значение лежало в интервале от 0 до 1. Описанная конструкция называется искусственным нейроном. Сеть состоит из некоторого числа таких нейронов, причем часть из них непосредственно обрабатывает входные данные (первый слой нейронов), часть — сигналы, полученные на выходе с нейронов предшествующего слоя (скрытые слои нейронов), и, наконец, существует (как правило, единственный) выходной нейрон, который и выдает результат. При этом веса, соответствующие различным нейронам (а иногда и параметры функций активации), могут меняться независимо друг от друга. Обрабатывая исторические (обучающие) данные и меняя при этом веса, сеть стремится наилучшим образом приспособить свой выходной сигнал к известному результату. Этот процесс называется обучением сети³.

¹ Возможность применения нейронных сетей в качестве инструментального средства решения задачи прогнозирования обусловлена их способностью к аппроксимации нелинейных зависимостей. Согласно следствию из теоремы Колмогорова — Арнольда, доказанному Хехт-Нильсенем, произвольная непрерывная функция нескольких переменных может быть аппроксимирована нейронной сетью с любой наперед заданной степенью точности (см. Hecht-Nielsen, 1987, p. 11—13).

² Если в основе классических подходов лежат формализованные каким-либо образом знания человека о предметной области, то для нейронной сети аналитическая форма представления знаний недоступна. Все, что она может, — это запомнить и обобщить предъявленные ей на этапе обучения эмпирические зависимости между входными факторами и результирующими значениями. Иными словами, нейронная сеть строит модель некоего процесса и в дальнейшем воспроизводит его поведение. Собственно говоря, это и дает повод утверждать, что искусственные нейросети моделируют свойственные человеку приемы мышления.

³ Цель обучения заключается в минимизации ошибки оценки объясняемых переменных на основе информации о значениях объясняющих переменных. Для целей прогнозирования наибольшее распространение среди всех архитектур нейронных сетей получил итеративный алгоритм обучения, именуемый алгоритмом обратного распространения ошибки, который впервые был разработан П. Вербосом в 1974 г. (см. Werbos, 1974), а затем в 1986 г. заново открыт и популяризирован Д. Румельхартом, Г. Хинтоном и Р. Вильямсом (см. Rumelhart, Hinton, Williams, 1986, p. 533—536). С начала 1990-х гг. алгоритм обратного распространения стал активно применяться

После того как оно закончено, на вход сети можно подать новые данные, и она выдает свой прогноз¹.

Как видно, здесь действительно есть (в сильно упрощенном виде) некоторая аналогия с работой мозга человека. Иными словами, нейронные сети представляют собой класс аналитических методов, построенных на (гипотетических) принципах обучения мыслящих существ и функционирования мозга и позволяющих прогнозировать значения некоторых переменных в новых наблюдениях по данным других наблюдений (для этих же или других переменных) после прохождения этапа обучения на имеющихся данных, во время которого система прогнозирования в автоматическом режиме осуществляет оценивание параметров модели, минимизируя функцию ошибки.

При применении методов нейронных сетей прежде всего встает вопрос выбора конкретной архитектуры сети². Поскольку на начальном этапе анализа

в прикладных разработках. Суть алгоритма обратного распространения сводится к минимизации функции ошибки, определенной на множестве возможных значений весов сети. Это достигается следующим образом. Сигнал, поступающий на вход нейронной сети, распространяется по всем ее слоям и формирует выходной сигнал. Этот сигнал сравнивается с тестовым, вычисляется ошибка сети, которая распространяется в обратном направлении. При этом сеть корректирует значения весовых коэффициентов нейронов во всех слоях. Затем на первый слой сети подается новый сигнал, и вся процедура проводится заново (пока сети не будут предъявлены все примеры из обучающей выборки). Далее обучающие примеры повторяются до тех пор, пока нейросеть не сможет понять и запомнить их все.

¹ В ходе обучения сети многократно предъявляется один и тот же набор обучающих примеров. Чем дольше продолжается процесс обучения, тем лучше качество аппроксимации, демонстрируемое сетью при оценке значений выходных переменных по обучающей выборке. Однако через определенное число эпох обучения (под эпохой понимается однократное предъявление сети используемого набора обучающих примеров) улучшение качества аппроксимации начинает обеспечиваться не в результате правильной идентификации нелинейной зависимости между объясняющими и объясняемыми переменными, а за счет точности настройки на специфические особенности обучающих примеров. Этот феномен, получивший название переобучение, находит отражение в падении способности сети к обобщению, т. е. к адекватной оценке значений выходных переменных по наблюдениям, не предъявленным в ходе обучения.

Именно для того чтобы разрешить проблему переобучения, массив исходных данных разбивается на обучающую и тестовую выборки. Обучающая выборка используется в процессе работы алгоритма коррекции матрицы весов сети. Тестовая выборка используется для контроля состояния обученности сети. Процесс обучения прекращается, когда значение ошибки оценки значений выходных переменных по тестовой выборке достигает минимума.

² Вообще говоря, на входы нейросетей можно подавать самую разнообразную информацию, как дискретную (различные события), так и непрерывную (цены, индексы, индикаторы). Состав входов обычно является наиболее «сокровенной» тайной разработчика нейросети и держится в секрете. Как показывает опыт нашего исследования, наиболее значимыми входами являются различные фондовые индексы.

Кроме того, большой проблемой при выборе состава входов является определение глубины «погружения» — размера временного окна, данные из которого подаются на вход нейросети. Эта величина определяется только из практического опыта и должна быть адекватной прогнозируемому периоду. Например, по нашему мнению, при дневных прогнозах разумная глубина погружения находится в пределах 5—20 дней.

Необходимо отметить, что многие популярные нейропакеты имеют функцию определения чувствительности по входам, которая рекомендуется разработчиками пакетов в качестве универсального способа отсеивания «лишних» входов. Хотя эта функция и является полезной при определении состава входов, но она не панацея, и ее автоматическое использование может принести больше вреда, чем пользы. Это обусловлено тем, что проблема определения чувствительности по входам для многослойных нейронных сетей не имеет пока строгого математического решения, и решается различными эмпирическими способами и алгоритмами. Поэтому рекомендуется пользоваться этой функцией с осторожностью, а результаты проверять другими косвенными методами, например, традиционными статистическими.

Выбор оптимальной архитектуры сети также в настоящее время не имеет математического решения и производится на основании опыта и знаний. Здесь можно сказать только, что сеть должна иметь разумные размеры, а именно:

природа явления обычно бывает не очень хорошо известна, выбор архитектуры является непростой задачей и часто связан с длительным процессом «проб и ошибок»¹.

Затем построенная сеть подвергается процессу обучения. На этом этапе нейроны сети итеративно обрабатывают входные данные и корректируют свои веса таким образом, чтобы сеть наилучшим образом прогнозировала (в традиционных терминах следовало бы сказать — «осуществляла подгонку») данные, на которых выполняется обучение. После обучения на имеющихся данных сеть готова к работе и может использоваться для построения прогнозов.

Сеть, полученная в результате обучения, выражает закономерности, присутствующие в данных. При таком подходе она оказывается функциональным эквивалентом некоторой модели зависимостей между переменными, подобной тем, которые строятся в традиционном моделировании. Однако, в отличие от традиционных моделей, в случае нейронных сетей эти зависимости не могут быть записаны в явном виде, подобно тому как это делается в других статистических методах².

Одно из главных преимуществ нейронных сетей состоит в том, что они (по крайней мере, теоретически) могут аппроксимировать любую непрерывную функцию, и поэтому исследователю нет необходимости заранее принимать какие-либо гипотезы относительно модели и даже (в ряде случаев) о том, какие переменные действительно важны. Однако существенным недостатком нейронных сетей является то обстоятельство, что окончательное решение зависит от начальных установок сети и, как уже было отмечено, его практически невозможно интерпретировать в традиционных аналитических терминах, которые обычно применяются при построении теории того или иного явления.

Простота лежащего в ее основе принципа не избавляет теорию нейронных сетей от существования ряда значительных сложностей. Прежде всего, следует заметить, что исходная задача сводится к задаче глобальной оптимизации (при этом нелинейной). Несмотря на то что этой задаче уже несколько сотен лет, исчерпывающих методов ее решения до сих пор не найдено. Тем не менее к настоящему времени процедуры этих методов развиты достаточно глубоко и позволяют хорошо решать многие реальные задачи, а производители нейросе-

максимальное число нейронов — 30—40;

максимальное число скрытых слоев — 3—4;

объем входных данных должен превышать в несколько раз количество нейронов (так как в противном случае уменьшаются обобщающие свойства сети — сеть начинает просто запоминать обучающую выборку; такая сеть будет отлично прогнозировать на обучающей выборке, но на новых данных прогноз будет плохим; в частности, для прогнозирования индекса российского рынка ценных бумаг на день вперед минимальный рекомендуемый размер выборки — около 200—250 дней).

¹ Заметим, что нейронные сети — это всего лишь сети, состоящие из связанных между собой простых элементов — формальных нейронов. Ядром используемых представлений является идея о том, что нейроны можно моделировать довольно простыми автоматами, а вся сложность сети, гибкость ее функционирования и другие важнейшие качества определяются связями между нейронами. Каждая связь представляется как совсем простой элемент, служащий для передачи сигнала. Предельным выражением этой точки зрения может служить лозунг «структура связей — все, свойства элементов — ничто».

² Иногда нейронные сети выдают прогноз очень высокого качества. Однако они представляют собой типичный пример нетеоретического подхода к исследованию (иногда это называют «черным ящиком»). При таком подходе мы сосредотачиваемся исключительно на практическом результате (в данном случае — на точности прогнозов и их прикладной ценности), а не на сути механизмов, лежащих в основе явления, или соответствии полученных результатов какой-либо имеющейся теории.

тевых пакетов стараются в полной мере учесть последние достижения теории оптимизации¹.

Что касается использования нейросетевых методов, то применительно к экономической сфере можно выделить три класса таких приложений: прогнозирование временных рядов (курсов акций, обменных курсов), классификация (определение степени устойчивости эмитента, привлекательности объекта для инвестиций) и третий, смешанный, тип, который можно назвать задачей динамической классификации, — анализ и выявление аномалий в поведении объекта во времени (прежде всего, это задача обнаружения злоупотреблений различного рода).

Важно подчеркнуть, что применение нейронных сетей в финансах базируется на одном фундаментальном допущении — замене прогнозирования распознаванием. Нейросеть не предсказывает будущее, она старается «узнать» в текущем состоянии ранее встречавшуюся ситуацию и максимально точно воспроизвести реакцию рынка.

Если говорить о сравнении методов нейронных сетей и традиционных статистических методов, то можно отметить, что четкой границы между ними нет, хотя и существуют некоторые качественные отличия. Используя какой-либо статистический метод, исследователь зачастую навязывает своим данным жесткую структуру некоторой (как правило, линейной) модели, хотя реальные данные могут ей не соответствовать. Нейронные сети в этом смысле представляют собой более гибкое средство, так как они не требуют наложения каких-либо предположений о модели и хорошо находят именно нелинейные зависимости. С другой стороны, по статистической модели вместе с прогнозом можно получить доверительный интервал как для него, так и для параметров самой модели. Нейронная сеть выдает только прогноз и больше ничего, в результате чего ее часто называют «черным ящиком». Поэтому, вероятно, разумнее будет не противопоставлять одни методы другим, а осмысленно пользоваться и теми, и другими.

В данной статье осуществляется проверка гипотезы о существовании нелинейной зависимости между прошлыми значениями индекса российского фондового рынка и ряда других экономических показателей. В качестве инструмента идентификации данной зависимости используются нейронные сети — гибкие непараметрические модели, нашедшие широкое применение в различных финансовых приложениях².

При написании настоящей работы было проведено исследование нейросетевой модели прогнозирования показателей рынка ценных бумаг, целью которого являлось получение краткосрочного прогноза фондового индекса РТС.

Существенным моментом при выборе массива исходных данных является необходимость исследования имеющихся наблюдений на предмет однородности, то есть выяснение вопроса о схожести условий экономической активности на протяжении изучаемого временного интервала, а также о неизменности способов оценивания параметров и инструментов (в противном случае требовалось бы выделение периодов однородности и отдельное моделирование каждого из подобных периодов). Кроме того, следует учитывать тесную связь протяженности

¹ В частности, трактовка процесса приспособления как оптимизационного процесса приводит к идее использования генетических алгоритмов (см. далее) при обучении нейронных сетей. При этом если, например, широко известные градиентные методы гарантируют нахождение локального минимума, то генетические алгоритмы обеспечивают глобальную оптимизацию произвольной функции. Использование таких алгоритмов позволяет обучать большие нейросети, обходя проблему локальных минимумов.

² В случае если характер зависимости между опережающими и прогнозируемым показателем в действительности является нелинейным, то использование нейронной сети вместо множественной регрессии позволяет существенно повысить качество модели.

интервала прогнозирования с временным интервалом, использованным при моделировании. Принято считать, что на каждую прогнозную оценку должно приходиться не менее десяти (в некоторых источниках — трех) исходных наблюдений (Лугачев, Ляпунцов, 1999; Теория статистики, 2002). Исходя из приведенных соображений, при решении задачи прогнозирования развития российского фондового рынка в качестве исходного был взят промежуток времени с января 2003 г. по август 2005 г., а в качестве прогнозного — с сентября по ноябрь 2005 г.¹

Рассматривалось несколько вариантов выбора исходных данных:

1. Динамическая модель, в которой зависимой переменной являлся индекс РТС, а независимыми переменными — лаговые значения этого индекса в предшествующие дни.

Для проверки временного ряда на стационарность можно использовать различные методы. В частности, простейшим из них является визуальный контроль графического представления ряда. В этом случае о вероятной нестационарности будут свидетельствовать очевидные на глаз тренд или периодический компонент, а также возрастание или убывание разброса наблюдений во времени. Другим методом является построение графика выборочной автокорреляционной функции (коррелограммы), которая должна быстро убывать после нескольких первых значений в случае стационарного временного ряда (и, наоборот, убывает достаточно медленно при нестационарности ряда). Кроме автокорреляционной, можно строить график и для частной автокорреляционной функции, которая также должна быстро убывать для стационарного процесса². Наконец, возможно использование формальных тестов на наличие единичного корня (простой и расширенный тесты Дики — Фуллера и др.).

Учитывая, что используемый в задаче прогнозирования исходный временной ряд значений объясняемой переменной (индекса фондового рынка РТС) не является стационарным (нестационарность ряда значений индекса РТС очевидна из рис. 1), был использован ряд преобразований, которые приводят к получению модифицированного ряда с характеристиками, приближенными к характеристикам стационарных рядов. Заметим, что рассмотрены не все возможные способы преобразований, а лишь те, которые приводят к формированию новых рядов, демонстрирующих «хорошие» свойства при их использовании в описанных методах прогнозирования.

Иными словами, в качестве исходной информации для прогнозирования задействовались уровни следующих временных рядов, полученных из значений индекса РТС:

ряд 1 — ряд значений индекса РТС;

ряд 2 — ряда остатков индекса РТС после его аппроксимации алгебраическим полиномом на основе метода последовательных разностей (Айвазян, Мхитарян, 1998);

ряд 3 — ряда остатков индекса РТС после его аппроксимации алгебраическим полиномом и 30 гармониками ряда Фурье;

ряд 4 — ряд разностей значений индекса РТС;

ряд 5 — ряда значений «логарифмической прибыли» (логарифмов цепных темпов роста) индекса РТС;

ряд 6 — ряд значений «логарифмической прибыли» индекса РТС после вычитания выборочного среднего;

¹ Значения индекса российского фондового рынка были получены из архива, хранящегося на FTP-сервере РТС (<ftp://ftp.rtsnet.ru/pub/info/stats/>).

² Содержательно частная автокорреляционная функция представляет собой «чистую корреляцию» между текущим и одним из предыдущих уровнями ряда при исключении влияния всех промежуточных лаговых значений.

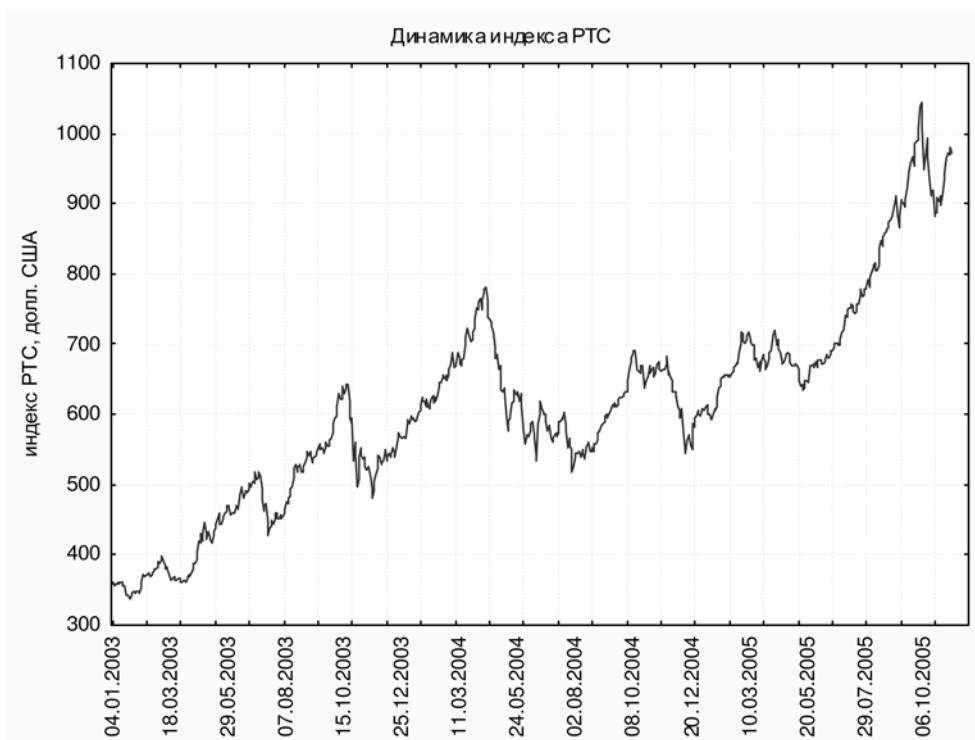


Рис. 1. График значений индекса российского фондового рынка (индекса РТС) на интервале моделирования

ряд 7 — ряд остатков «логарифмической прибыли» индекса РТС после ее аппроксимации алгебраическим полиномом на основе метода последовательных разностей и гармониками ряда Фурье.

2. Многофакторная модель, в которой зависимой переменной являлся индекс РТС, а независимыми переменными — 79 признаков, отражающих системообразующие факторы, оказывающие воздействие на развитие российского фондового рынка:

— фондовые индексы рынков ценных бумаг тех стран, которые способны оказывать влияние на российскую ситуацию (для рассмотрения были отобраны индексы DJI — Dow Jones 30 Industrials, NYA — NYSE Comp, NASD — Nasdaq Comp, XAX — AMEX Composite, FTSE — UK FTSE 100, CAC — France CAC 40, DAX — Germany DAX, NIKKEI — Japan Nikkei 225, TSE — Canada TSE 300 Comp, MIBTEL — Italy MIBTel, AEX — Netherlands AEX General, ASE — Greece General Share, ASX — Australia All Ordinaries, ATX — Austria ATX, BE500 — Bloomberg European 500 Index, BEL20 — Belgium BEL-20, BSE — India BSE 30, BUSE — Brazil Bovespa, BUX — Hungary BUX, CMA — Egypt CMA, CSE — Sri Lanka All Share, HEX — Finland Helsinki General, HSI — Hong Kong Hang Seng, IBC — Venezuela IBC, IGBM — Spain Madrid General, IGRA — Peru Lima General, IPC — Mexico IPC, IPSA — Chile IPSA, JKSE — Indonesia Jakarta Comp, JSE — Johannesburg All Share, KFX — Denmark KFX, KLSE — Malaysia KLSE Comp, KS11 — S. Korea Seoul Comp, KSE — Pakistan Karachi 100, Merval — Argentina MerVal, MID — SnP 400 MidCap, NZSE40 — N. Zealand NZSE 40, OEX — SnP 100 Index, PSE — Philippines PSE Comp, PX50 — Czech Republic, SAX — Slovakia SAX, SET — Thailand SET, SML — SnP 600 SmallCap, SPX — SnP 500 Index, SSEC — China Shanghai Comp, SSMI — Switzerland Swiss Mkt, STI — Singapore Straits

Times, TA100 — Israel TA-100, WIG20 — Warsaw Stock Exchange, XU100 — Turkey ISE National-100);

— объявленные ставки по привлечению (MIBID) и размещению (MIBOR) межбанковских кредитов, в процентах годовых по рублевым вкладам (сроком на 1 день, от 2 до 7 дней, от 8 до 30 дней, от 31 до 90 дней);

— показатели по рынку государственных ценных бумаг — объем и оборот операций (по рыночной стоимости), а также эффективный индикатор рыночного портфеля, отражающий усредненную доходность к погашению по всем выпускам (по ОФЗ-АД, ОФЗ-ФД, ОФЗ-ПД и ГКО — ОФЗ в целом);

— биржевой курс и объем торгов на ММВБ по доллару США и евро;

— котировки драгоценных металлов (золото, серебро, платина, палладий) на LONDON FIX (в долл. США за тройскую унцию);

— ежедневные цены фьючерсного контракта на нефть NYMEX Light Sweet Crude Oil с ближайшей датой поставки (в долл. США за баррель).

Для вычисления значений указанных факторов использовались следующие информационные ресурсы глобальной сети Интернет:

<http://export.rbc.ru> — информация, предоставляемая аналитическим агентством «РосБизнесКонсалтинг»;

<http://www.cbr.ru> — данные, раскрываемые Банком России;

<http://www.kitco.com> — сведения о котировках LONDON FIX из базы данных компании Kitco Inc.;

<http://www.eia.doe.gov> — материалы Министерства энергетики США.

Для многофакторной модели также были рассмотрены как случай использования значений исходных временных рядов, так и случай применения преобразованных различными способами данных (например, по формуле «логарифмической прибыли»). В качестве значений результирующей переменной рассматривались уровни рядов 1—7.

Для того чтобы отразить основные аспекты применения нейронных сетей, ниже приводится описание нейросетевого метода прогнозирования для случая многофакторной модели с исходными данными в виде значений «логарифмической прибыли».

Массив исходных данных за период с января 2003 г. по август 2005 г. был разбит на обучающую, контрольную и тестовую выборки. В состав обучающей выборки было включено 330 наблюдений, в состав контрольной и тестовой выборки — по 165 наблюдений (производился случайный выбор элементов контрольного и тестового множеств).

Необходимость контрольной кросс-проверки объясняется следующим образом. Почти всегда более сложная сеть дает меньшую ошибку, но это может свидетельствовать не о хорошем качестве модели, а о переобучении. Поэтому часть обучающих наблюдений резервируется и не используется при работе алгоритма обратного распространения. Вместо этого по мере работы алгоритма они используются для независимого контроля результата. В самом начале работы ошибка сети на обучающем и контрольном множестве будет примерно одинаковой. По мере того как сеть обучается, ошибка обучения убывает, причем пока обучение уменьшает действительную функцию ошибок, ошибка на контрольном множестве также будет убывать. Если же контрольная ошибка перестала убывать или даже стала расти, это указывает на то, что сеть начала слишком близко аппроксимировать данные, т. е. на то, что обучение следует остановить. При практической работе с нейронными сетями, как правило, приходится экспериментировать с большим числом различных сетей, порой обучая каждую из них несколько раз (чтобы не быть введенным в заблуждение

локальными минимумами) и сравнивая полученные результаты. Главным показателем качества результата является здесь контрольная ошибка. При этом в соответствии с общенаучным принципом, согласно которому при прочих равных показателях следует предпочесть более простую модель, из двух сетей с приблизительно равными ошибками контроля имеет смысл выбрать ту, которая меньше по размеру.

Необходимость многократных экспериментов ведет к тому, что контрольное множество начинает играть ключевую роль в выборе модели, иначе говоря, становится частью процесса обучения. Тем самым ослабляется его роль как независимого критерия качества модели — при большом числе экспериментов есть риск выбрать «удачную» сеть, дающую хороший результат на контрольном множестве. Для того чтобы придать окончательной модели должную надежность, мы поступаем следующим образом. Резервируется еще одно — тестовое — множество наблюдений. Итоговая модель тестируется на данных из этого множества, чтобы убедиться в том, что результаты, достигнутые на обучающем и контрольном множествах реальны, а не являются артефактами процесса обучения. Разумеется, для того чтобы хорошо играть свою роль, тестовое множество должно быть использовано только один раз: если его использовать повторно для корректировки процесса обучения, то оно фактически превратится в контрольное множество.

Поскольку конкретные значения входов принадлежали произвольным диапазонам, то была проведена их нормировка. При этом данные переводились в безразмерную форму, за счет чего все входы нейросети стали сравнимыми по порядку величинами. Аналогичным образом был нормирован и выход нейросети.

Исследовались два различных способа нормировки:

— минимаксная функция (находятся минимальное и максимальное значение переменной по обучающему множеству, после чего выполняется линейное преобразование с применением коэффициентов масштаба и смещения, чтобы значения лежали в нужном диапазоне — на отрезке $[0;1]$);

— масштабирующая функция «среднее/стандартное отклонение» (обучающие данные масштабируются таким образом, чтобы среднее значение переменной равнялось 0, а все другие значения были бы промасштабированы на величину стандартного отклонения).

При проведении исследования было рассмотрено несколько популярных архитектур построения нейронных сетей (Бэстэнс, Ван Ден Берг, Вуд, 1997):

1. Архитектура многослойного персептрона (MLP). Каждый элемент многослойного персептрона выполняет линейное преобразование входного вектора (набора значений, подаваемого на вход элемента) — берется взвешенная сумма входов, и из нее вычитается пороговое значение. Такое преобразование называется линейной PSP (пост-синаптической потенциальной) функцией. Затем результат пропускается через нелинейную функцию активации. Линейность PSP-функции означает, что многослойный персептрон фактически делит пространство входных векторов на части посредством гиперплоскостей (в двумерном пространстве гиперплоскостью является обычная прямая).

2. Архитектура радиальных базисных функций (RBF). В отличие от многослойного персептрона с его линейным подходом, в сети на радиальных базисных функциях используется радиальная PSP-функция. В каждом элементе вычисляется квадрат расстояния от входного вектора до вектора весов. Затем это расстояние умножается на «пороговое значение» (в действительности — на величину отклонения), после чего результат пропускается через функцию активации. Таким образом, сеть на радиальных базисных функциях делит пространст-

во входов посредством гиперсфер (в двумерном пространстве гиперсферой является окружность)¹.

3. Архитектура линейной сети. Главная причина популярности нейронных сетей кроется в их способности моделировать нелинейные задачи, т. е. задачи регрессии, которые нельзя решить, проводя гиперплоскость через данные. Однако часто оказывается, что задача, считавшаяся сложной и нелинейной, на самом деле может быть успешно решена линейными методами. Линейные модели строятся с помощью специального типа нейронных сетей. Линейная сеть имеет всего два слоя: входной и выходной слой, имеющий линейные PSP-функции и функции активации. В практических задачах линейная модель является хорошей точкой отсчета для сравнения эффективности различных более сложных методов.

4. Архитектура обобщенно-регрессионной нейронной сети (GRNN). Стандартные архитектуры сетей, рассчитанные на управляемое обучение (много-слойный перцептрон и радиальные базисные функции), предполагают построение параметрической модели по имеющимся обучающим данным, где в качестве параметров выступают веса. Параметрическая модель (сеть) по объему обычно оказывается гораздо меньше, чем набор обучающих данных, и работает довольно быстро, хотя для обучения сети иногда может требоваться значительное время. К задаче можно подойти иначе и попытаться моделировать искомое отображение более или менее непосредственно по обучающим данным. Преимущество такого подхода состоит в том, что в нем практически не требуется обучения, а недостаток — в том, что в результате может получиться очень громоздкая модель, которая будет занимать много памяти и медленно работать. Обобщенно-регрессионные нейронные сети реализуют методы такого типа, предназначенные для задачи регрессии. Первый промежуточный слой в сетях этих типов состоит из радиальных элементов, каждый из которых хранит в себе центр кластера входных данных (общее число таких элементов бывает примерно того же порядка, что и объем обучающего множества). Выходной сигнал такого радиального элемента представляет собой функцию активации с центром в той точке, которая хранится в данном элементе, а в последующих слоях из этих сигналов составляются оценки для зависимой переменной регрессии.

С целью улучшения качества прогнозирования было выполнено определение статистической значимости входной информации для предсказания выходной, для чего использовался следующий алгоритм анализа чувствительности моделей к входным переменным. Для каждой входной переменной строились две модели, первая из которых получалась обычным образом. При построении второй модели значения анализируемой переменной во всех наблюдениях считались неизвестными, по сути игнорировались. Мерой статистической значимости конкретной входной переменной для исходной нейронной сети выступает отношение ошибки сети для второй модели к ошибке сети для первой модели. Чем более важной для рассматриваемой нейронной сети является анализируемая переменная, тем большим должно быть это отношение. В случае если коэф-

¹ Каждый из двух описанных выше подходов имеет свои достоинства и недостатки. Действие радиальных функций очень локально, в то время как при линейном подходе охватывается все пространство входов. Поэтому, как правило, RBF-сети имеют больше элементов, чем MLP-сети, однако MLP может делать необоснованные обобщения в ситуациях, когда ему попадает набор данных, непохожий ни на какие наборы из обучающего множества, в то время как RBF в таком случае всегда будет выдавать почти нулевой отклик.

В теории отмечается также, что для решения некоторых классов задач с помощью MLP требуется два промежуточных слоя, а иногда для большей эффективности приходится брать даже большее число промежуточных слоев. В отличие от этого, для RBF-сети всегда достаточно одного промежуточного слоя.

коэффициент значимости не превосходит единицы, делается вывод, что исключение переменной не влияет на качество построенной сети, а возможно, даже улучшает его¹. Результирующая информация представлена в табл. 1.

Учитывая, что не имело смысла использовать при прогнозировании те переменные, которые характеризуются невысокой значимостью, проводилась фильтрация входов². Для этого использовался специальный генетический алгоритм

Таблица 1

Статистическая значимость входов нейросети для прогнозирования значения индекса РТС

Переменная	Значимость	Переменная	Значимость	Переменная	Значимость
Биржевой курс долл. США	1,0005	MIBOR30	0,9977	IGRA	1,0009
Объем торгов по долл. США	1,0018	MIBOR90	0,9996	IPC	1,0002
Биржевой курс евро	1,0021	DJI	0,9996	IPSA	1,0001
Объем торгов по евро	0,9994	NYA	1,0000	JKSE	1,0001
Эффективный индикатор ОФЗ-АД	1,0007	NASD	1,0034	JSE	1,0001
Объем ОФЗ-АД	1,0001	XAX	0,9981	KFX	0,9988
Оборот ОФЗ-АД	0,9998	FTSE	0,9996	KLSE	1,0001
Эффективный индикатор ОФЗ-ФД	1,0011	CAC	0,9999	KS11	1,0002
Объем ОФЗ-ФД	1,0001	DAX	0,9980	KSE	0,9948
Оборот ОФЗ-ФД	0,9999	NIKKEI	0,9996	MERVAL	0,9992
Эффективный индикатор ОФЗ-ПД	1,0002	TSE	0,9984	MID	0,9991
Объем ОФЗ-ПД	0,9999	MIBTEL	0,9996	NZSE40	1,0000
Оборот ОФЗ-ПД	0,9997	AEX	0,9998	OEX	1,0013
Эффективный индикатор ГКО — ОФЗ	1,0012	ASE	1,0000	PSE	0,9983
Объем ГКО — ОФЗ	0,9999	ASX	1,0000	PX50	1,0110
Оборот ГКО — ОФЗ	0,9979	ATX	0,9988	SAX	1,0008
Котировки золота	1,0014	BE500	1,0001	SET	1,0001
Котировки серебра	1,0035	BEL20	1,0006	SML	0,9998
Котировки платины	1,0001	BSE	0,9996	SPX	0,9991
Котировки палладия	0,9997	BUSP	1,0008	SSEC	0,9997
Цена на нефть	0,9968	BUX	1,0049	SSMI	0,9992
MIBID1	0,9968	CMA	0,9998	STI	1,0003
MIBID7	0,9989	CSE	1,0015	TA100	0,9999
MIBID30	1,0004	HEX	0,9997	WIG20	1,0003
MIBID90	0,9992	HSI	1,0024	XU100	1,0004
MIBOR1	0,9973	IBC	0,9989	—	—
MIBOR7	0,9957	IGBM	0,9998	—	—

¹ Следует отметить, что указанный коэффициент статистической значимости не является теоретически обоснованной мерой «полезности» переменной, тем более что он совершенно не учитывает возможные взаимозависимости между переменными. Однако в практических исследованиях, сопряженных с большим числом эмпирических допущений, его роль для нахождения способов возможного улучшения анализируемых нейронных сетей очень важна.

² В процессе проведения исследования было отмечено, что уменьшение количества входов существенно сокращает время обучения нейросети, а также дает возможность увеличить ее нелинейные свойства.

отбора входных данных, который позволяет сформировать подходящий набор участвующих в модели переменных. Схема работы генетического алгоритма такова. Берется случайный набор (популяция) битовых строк (отдельный бит, соответствующий каждому входу, показывает, учитывать или нет соответствующую входную переменную) и оценивается степень их пригодности по контрольной ошибке (т. е. качество получаемых решений). Затем плохие строки исключаются из рассмотрения, а из оставшихся порождаются новые строки с помощью искусственных генетических операций мутации и скрещивания (случайных изменений отдельных битов). Таким образом, возникает новая популяция, и весь процесс повторяется, порождая все новые поколения, а в конце отбирается наилучший из рассмотренных экземпляров. Каждое оценивание включает построение GRNN-сети и ее тестирование на контрольном множестве. Сети этой архитектуры выбраны потому, что для них общее время обучения/оценки очень мало, а также потому, что они очень сильно страдают от присутствия ненужных входных переменных и потому являются хорошим средством их обнаружения. Окончательный список использованных в качестве входов нейросети переменных представлен в табл. 2. Проведенные эксперименты показали, что наилучшая точность прогнозирования с точки зрения контрольной ошибки получается в случае применения трехслойного персептрона с девятью нейронами во входном слое (соответствующими девяти объясняющим переменным), одиннадцатью нейронами в единственном скрытом слое и одним нейроном в выходном слое (соответствующим прогнозируемому показателю). Графическая иллюстрация выбранной модели приведена на рис. 2.

Таблица 2

Переменные, использованные в качестве входов нейросети

Переменная	Значимость	Переменная	Значимость	Переменная	Значимость
Объем ГКО — ОФЗ	1,0022	DJI	1,0029	JSE	1,0100
AEX	1,0058	NIKKEI	0,9996	KFX	1,0023
BEL20	1,0017	HSI	1,0096	WIG20	0,9997

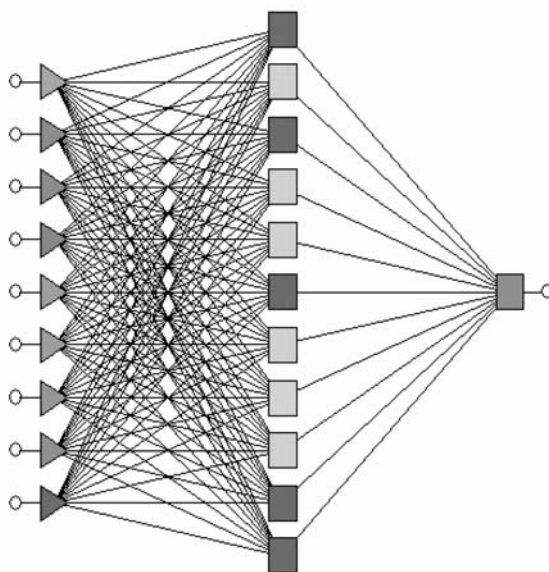


Рис. 2. Графическое представление структуры нейронной сети, используемой для прогнозирования значений «логарифмической прибыли» индекса РТС

В течение примерно 131 эпохи обучения ошибка аппроксимации изменялась как по обучающей, так и по контрольной выборке, достигнув, соответственно, минимальных отметок 0,0924 и 0,0719. В последующем ошибки аппроксимации стабилизировались. Таким образом, оптимальное состояние обученности сети было достигнуто после 131 эпохи обучения (рис. 3).

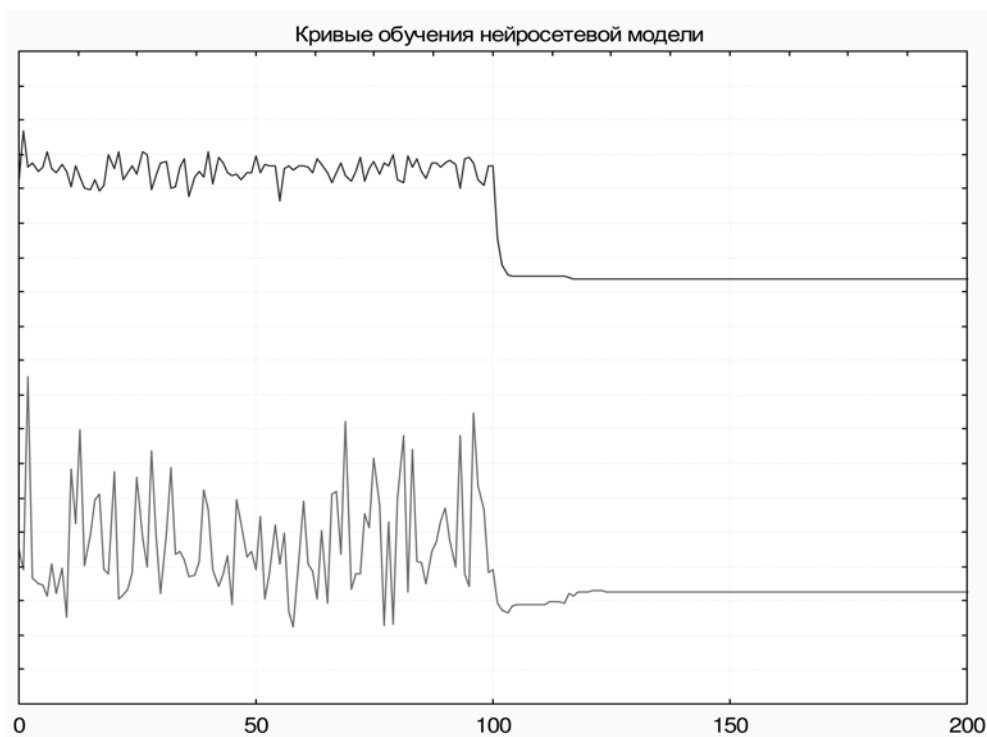


Рис. 3. Кривые обучения нейросетевой модели прогнозирования индекса РТС¹

В качестве итоговых статистик, характеризующих построенную нейросетевую модель, выступают среднее значение и стандартное отклонение, вычисленные для обучающих данных и для ошибки прогноза. В общем случае среднее значение ошибки прогноза будет очень близко к нулю (в конце концов, нулевое среднее для ошибки прогноза можно получить, попросту оценив среднее значение обучающих данных и вовсе не обращаясь к значениям входных переменных). Наиболее важным показателем является стандартное отклонение ошибки прогноза. Если оно не окажется существенно меньше стандартного отклонения обучающих данных, это будет означать, что сеть работает не лучше, чем простая оценка по среднему. Соответственно, важной статистикой является отношение стандартных отклонений, представляющее собой отношение стандартного отклонения ошибки прогноза к стандартному отклонению исходных данных. Если бы у нас вообще не было входных данных, то лучшее, что мы могли бы взять в качестве прогноза для выходной переменной — это ее среднее значение по имеющейся выборке, а ошибка такого прогноза была бы равна стандартному отклонению выборки. Если нейронная сеть работает результативно, мы вправе ожидать, что ее средняя ошибка на имеющихся наблюдениях будет близка к нулю, а стандартное отклонение этой ошибки будет меньше стандартного откло-

¹ Нижняя кривая — ошибка по контрольной выборке, верхняя кривая — ошибка по обучающей выборке.

нения выборочных значений (иначе сеть давала бы результат не лучше, чем простое угадывание). Таким образом, отношение стандартных отклонений значительно меньше единицы (например, ниже 0,1), говорит об эффективности сети. Величина, равная единице минус отношение стандартных отклонений, равна доле объясненной дисперсии модели. В построенной модели отношение стандартных отклонений составило 0,96, т. е. многослойный перцептрон лишь на 4% предсказал остаток ряда. Это означает, что исходный набор данных (значения «логарифмической прибыли» факторов) не имел ощутимой нелинейной структуры.

После завершения процесса обучения рассматриваемой нейронной сети было проведено исследование выборочной дисперсии ошибок формируемого с ее помощью прогноза. Кроме того, указанным способом было выполнено моделирование для всех рассматриваемых рядов (на основе как многофакторной, так и динамической модели). Затем в рамках построенных сетей проводилось прогнозирование на 1, 5 и 50 дней, а также в соответствии с процедурой скользящего экзамена. Суть этой процедуры (известной в зарубежной литературе под названием *jackknife method* — «метод складного ножа») может быть описана следующим образом. Предположим, мы сравниваем q построенных для прогнозирования моделей. Принципиальным признаком этого подхода является требование постоянного обновления модели для определения прогнозной оценки каждой постпрогнозной точки. Будем считать, что выбрано $k = 50$ постпрогнозных точек, находящихся в конце временного ряда: y_{n-k+1}, \dots, y_n . В этом случае процедура скользящего экзамена состоит в следующем.

1. По первым $(n - k)$ наблюдениям исходного ряда оцениваются параметры первой модели в спецификации, определенной для прогнозирования $(n + 1)$ -й точки.

2. Построенная модель используется для оценки постпрогноза \hat{y}_{n-k+1} .

3. Оцениваются параметры модели по первым $(n - k + 1)$ наблюдениям исходного ряда и рассчитывается \hat{y}_{n-k+2} .

4. Расчеты повторяются до определения \hat{y}_n .

5. Оценивается мера качества прогноза

$$\hat{\sigma}^2(1) = \frac{1}{k - p - 1} \sum_{i=n-k+1}^n (y_i - \hat{y}_i)^2.$$

6. Процедура повторяется для всех имеющихся q моделей.

7. Последовательность $\hat{\sigma}^2(1), \dots, \hat{\sigma}^2(q)$ представляет собой значения частного критерия точности прогнозирования для всех рассматриваемых моделей.

Наилучший результат показали:

в случае прогнозирования на 1 день — метод прогнозирования на основе динамической модели и ряда 2;

в случае прогнозирования на 5 дней — метод прогнозирования на основе динамической модели и ряда 4;

в случае прогнозирования на 50 дней — метод прогнозирования на основе динамической модели и ряда 6;

в случае прогнозирования в соответствии с процедурой скользящего экзамена — метод прогнозирования на основе многофакторной модели и ряда 4.

Численные значения прогнозов приведены в табл. 3, графическая иллюстрация — на рис. 4, 5. Для сравнения приведены также результаты прогнозирования индекса РТС с использованием методов множественной регрессии (с 79 перечисленными выше факторами в качестве зависимых переменных) и ARIMA.

Таблица 3

Результаты прогнозирования индекса РТС на 50 дней

Дата	Индекс РТС	Прогноз нейросетевым методом на основе динамической модели и ряда 4	Прогноз нейросетевым методом на основе динамической модели и ряда 6	Прогноз нейросетевым методом на основе многофакторной модели и ряда 4 с помощью процедуры скользящего экзамена	Прогноз методом множественной регрессии с помощью процедуры скользящего экзамена	Прогноз методом ARIMA	Прогноз методом ARIMA с помощью процедуры скользящего экзамена
01.09.05	894,6	882,8	885,0	898,0	894,5	881,1	879,6
02.09.05	909,9	883,8	892,1	902,1	908,6	883,3	899,4
05.09.05	910,1	884,6	893,9	905,6	908,2	887,8	914,4
06.09.05	882,6	885,1	895,8	908,6	909,3	889,6	905,9
07.09.05	866,6	886,5	904,1	881,5	882,6	888,4	876,9
08.09.05	883,1	887,1	908,1	871,6	873,1	886,9	866,3
09.09.05	905,9	888,0	910,0	891,0	891,8	889,4	887,7
12.09.05	903,9	889,2	911,4	905,6	911,6	893,5	906,9
13.09.05	900,2	890,2	912,9	903,3	901,9	895,9	901,7
14.09.05	896,1	891,4	914,7	909,1	903,3	894,3	898,8
15.09.05	906,9	892,1	916,7	906,2	905,8	893,2	898,8
16.09.05	923,0	893,1	918,1	916,0	914,1	895,1	909,4
19.09.05	946,3	894,0	920,1	935,1	932,7	899,7	925,5
20.09.05	945,0	895,1	922,0	949,0	946,3	901,8	948,9
21.09.05	960,5	896,2	923,4	942,1	942,0	900,8	948,5
22.09.05	966,0	897,2	924,9	966,5	960,5	899,1	961,5
23.09.05	965,3	898,3	931,8	955,1	957,9	901,3	967,9
26.09.05	954,8	899,4	935,9	973,2	974,9	905,5	970,3
27.09.05	984,5	900,5	937,8	961,1	960,3	908,2	962,1
28.09.05	988,3	901,6	939,3	988,5	986,2	906,8	986,0
29.09.05	991,1	902,6	940,8	994,2	990,8	905,5	985,6
30.09.05	1007,8	903,7	942,7	992,8	993,5	907,1	991,5
03.10.05	1038,8	904,8	944,3	1015,4	1017,7	911,8	1015,9
04.10.05	1045,0	905,8	947,5	1042,6	1044,1	914,2	1043,8
05.10.05	1009,6	906,9	951,2	1034,8	1036,0	913,3	1040,5
06.10.05	954,6	908,0	953,2	976,2	987,0	911,5	1001,8
07.10.05	949,1	909,1	954,7	952,6	953,9	913,4	955,7
10.10.05	970,3	910,1	956,2	958,3	956,8	917,6	954,5
11.10.05	994,9	911,2	963,4	972,7	972,7	920,6	969,4
12.10.05	966,9	912,3	965,4	983,0	988,3	919,4	984,1
13.10.05	935,2	913,4	967,4	937,4	939,3	918,0	959,5
14.10.05	911,2	914,4	968,9	918,9	918,6	919,3	935,8
17.10.05	915,4	915,5	970,9	920,6	924,3	924,0	913,2
18.10.05	918,2	916,6	972,9	911,7	908,1	926,7	909,2
19.10.05	882,1	917,6	974,5	888,8	896,2	926,1	909,0
20.10.05	887,4	918,7	977,8	881,1	879,9	924,0	877,5
21.10.05	887,9	919,8	982,2	893,0	890,7	925,7	888,4
24.10.05	908,6	920,8	984,2	898,9	892,9	929,9	886,4
25.10.05	904,4	921,9	985,7	908,0	917,7	933,2	904,0
26.10.05	910,4	923,0	987,3	910,4	907,5	932,2	899,8
27.10.05	898,5	924,1	994,7	912,5	900,6	930,7	909,7
28.10.05	910,8	925,1	996,8	897,7	900,2	931,7	898,0
31.10.05	935,0	926,2	998,8	929,9	922,3	936,3	910,7
01.11.05	942,1	927,3	1000,4	941,2	931,9	939,3	935,4
02.11.05	964,4	928,3	1006,0	943,9	952,0	939,0	943,8
03.11.05	972,4	929,4	1008,2	974,1	974,6	936,8	963,4
07.11.05	971,6	930,5	1009,8	976,2	976,1	938,2	971,9
08.11.05	971,2	931,6	1011,3	967,0	967,6	942,3	974,7
09.11.05	980,3	932,6	1017,6	972,7	973,6	946,0	978,1
10.11.05	972,2	933,7	1019,8	978,7	976,7	945,2	982,1
Сумма квадратов ошибок за 50 дней		139 499	131 941	7145	8214	122 255	12 744

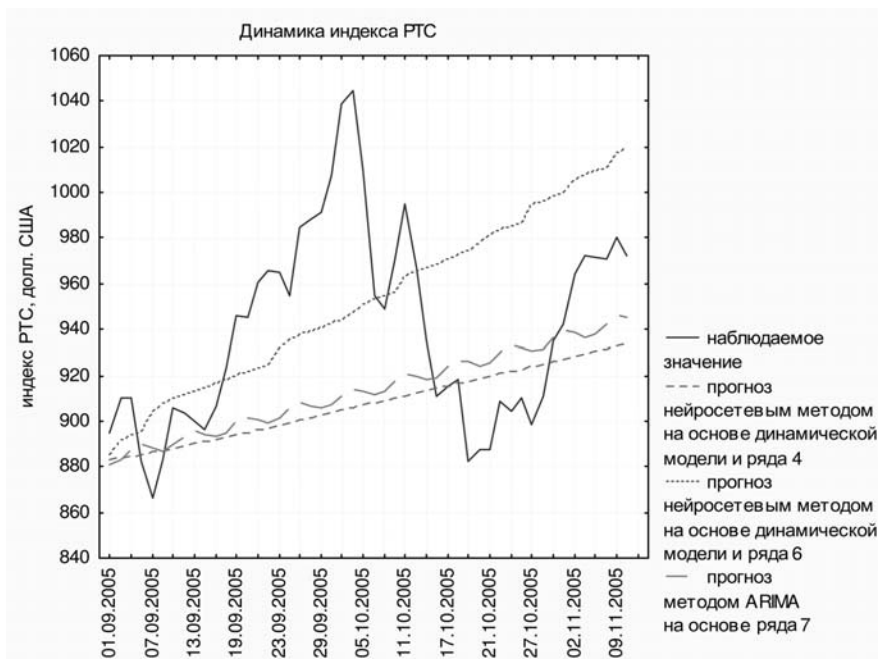


Рис. 4. Прогнозирование индекса РТС на 50 дней

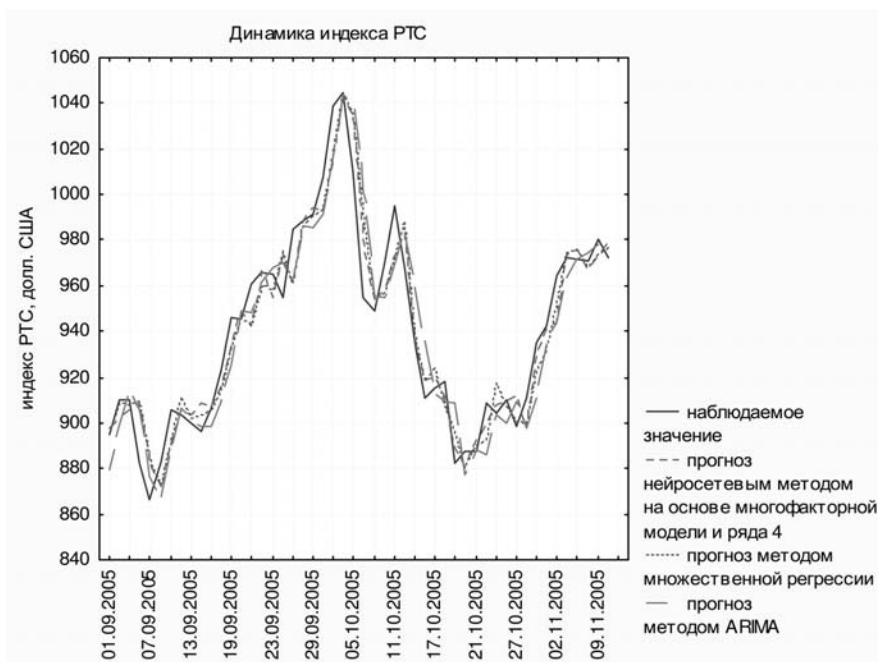


Рис. 5. Прогнозирование индекса РТС на 50 дней с помощью процедуры скользящего экзамена

Таким образом, можно говорить о том, что рассмотренные методы построения прогнозов динамики российского фондового рынка сопоставимы по своей эффективности с точки зрения выбранного критерия сравнения. Некоторое предпочтение может быть отдано системам прогнозирования на базе нейросетевого метода (при прогнозировании с помощью процедуры скользящего экзамена), а также метода ARIMA (для случая единовременного прогнозирования).

Источники

- Айвазян С. А., Мхитарян В. С.* Прикладная статистика и основы эконометрики. М., 1998.
- Арженовский С. В., Молчанов И. Н.* Статистические методы прогнозирования. Учебное пособие для аспирантов. Ростов-на-Дону, 2001.
- Бэстэнс Д. Э., Ван Ден Берг В. М., Вуд Д.* Нейронные сети и финансовые рынки: принятие решений в торговых операциях. М., 1997.
- Глуценко В. В.* Прогнозирование. 3-е изд. М., 2000.
- Дуброва Т. А.* Статистические методы прогнозирования. М., 2003.
- Лугачев М. И., Ляпунов Ю. П.* Методы социально-экономического прогнозирования. М., 1999.
- Теория статистики / под ред. Г. Л. Громыко. М., 2005.
- Hecht-Nielsen R.* Kolmogorov's Mapping Neural Network Existence Theorem // IEEE First Annual Int. Conf. on Neural Networks, San Diego, 1987. Vol. 3.
- Rumelhart D. E., Hinton G. E., Williams R. J.* Learning representations by back-propagating errors // Nature. 1986. N 323.
- Werbos P. J.* Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences. PhD thesis. Harvard University, 1974.