

О МЕТОДАХ ФОРМИРОВАНИЯ НЕЙРОСЕТЕВЫХ АНСАМБЛЕЙ В ЗАДАЧАХ ПРОГНОЗИРОВАНИЯ ФИНАНСОВЫХ ВРЕМЕННЫХ РЯДОВ

И. Л. Каширина

Воронежский государственный университет

Поступила в редакцию 10.10.2009

Аннотация. В статье предлагается несколько подходов к формированию нейросетевых комитетов в задачах прогнозирования финансовых временных рядов. В частности, обсуждается использование для этих целей перспективной методики бустинга.

Ключевые слова: нейронные сети, комитеты сетей, прогнозирование, бустинг

Annotation. In article some approaches to creation the committees of neural networks in tasks of forecasting of financial time rows are offered. In particular, usage for these purposes of a perspective technique boosting is discussed.

Keywords: neural networks, the committees networks, forecasting, boosting

ВВЕДЕНИЕ

Благодаря своей возможности выявлять нелинейные математические закономерности временных рядов, быстро адаптироваться к изменениям рыночных тенденций, нейронные сети являются на данный момент одним из самых перспективных инструментов прогнозирования финансовых временных рядов. Такой подход обладает рядом преимуществ.

Во-первых, нейросетевой анализ, в отличие, например, от технического, не предполагает никаких ограничений на характер входной информации. Это могут быть как индикаторы данного временного ряда, так и сведения о поведении других рыночных инструментов (например, последние изменения цен на нефть, изменения курсов ведущих мировых индексов, объемы продаж на бирже). Во-вторых, нейросетевой анализ не содержит искусственного подгона рыночных взаимосвязей под маску стандартного распределения, а потому представляется более универсальным и перспективным.

Для решения задачи прогнозирования необходимо найти такую нейронную сеть или комитет нейроэкспертов, который бы наилучшим образом строил отображение $F : X \rightarrow y$, обобщающее сформированный на основе ценовой динамики набор примеров $\{x, y\}$. Поиск такой

нейронной сети или комитета нейроэкспертов осуществляется при помощи одного или нескольких алгоритмов обучения.

При этом, однако, исследования в области прогнозирования временных рядов при помощи сетей продолжают и в настоящее время, и никаких стандартных методов здесь пока не выработано. В нейронной сети многочисленные факторы взаимодействуют весьма сложным образом, и успех пока приносит только эвристический подход.

Современные методы обучения многослойных искусственных нейронных сетей подразумевают случайное формирование первоначальных значений весовых (настроечных) коэффициентов. В этой связи предсказания сетей, обученных на одной и той же выборке данных, могут отличаться. Кроме того, как уже отмечалось, нейросетевое моделирование может использовать в качестве входных данных еще и различные финансовые и иные показатели, значения которых влияют на изменение прогнозируемого ряда. Поскольку выявить все такие факторы (и степень их влияния) однозначно, как правило, невозможно, это является аргументом в пользу использования не единственной нейронной сети, а комитета нейронных сетей.

1. КОМИТЕТЫ СЕТЕЙ

Чаще всего в литературе предлагается использовать в качестве итогового значения ко-

митета среднее арифметическое значение прогнозов всех сетей-экспертов [1]. Но эту идею можно существенно развить. Рассмотрим несколько возможных способов организации комитетов прогнозирующих нейросетей.

1. Можно, например, обучить еще одну нейронную сеть («руководителя комитета»), входами которой будут прогнозы всех нейро-экспертов, а выходом — итоговый прогноз комитета (рис. 1).

2. Другим подходом может являться введение понятия «специализации» экспертов. С этой целью предлагается провести предварительную кластеризацию входных образов обучающего множества, то есть разбить исходную выборку на несколько (2—5) групп схожих входных наборов. Например, в некоторые группы данных могут попасть обучающие наборы, характеризующиеся растущим трендом, в другие — убывающим и т.п. Для такой кластеризации может быть использована самоорганизующаяся нейронная сеть, называемая картой Кохонена. Далее, для каждого кластера выделяется, как минимум, два нейроэксперта, которые обучаются только на данных, попавших в этот кластер. Таким образом, создаются подкомитеты нейроэкспертов, специализирующихся на прогнозировании в условиях той или иной сложившейся рыночной ситуации. В режиме функционирования комитета входной образ сначала анализируется картой Кохонена, чтобы определить, к какому из имеющихся кластеров он относится. Затем итоговый прогноз вырабатывается тем подкомитетом, специализацией которого является данный кластер. Нейросетевая карта специализированных экспертов может использоваться не только в процессе прогнозирования, но и для аналитических целей. В частности, на основе карты по мере поступления запросов можно сделать вывод о проблемных областях комитета (например, выявить класте-

ры эксперты которых отличаются худшим качеством прогноза). Это дает новую стратегию обучения и пополнения комитета новыми моделями.

3. При использовании третьего подхода все нейронные сети обучаются на одних и тех же данных, но в комитет изначально включаются сети, отличающиеся принципиально разной архитектурой (многослойные персептроны, рекуррентные сети, RBF-сети и любые другие, хорошо зарекомендовавшие себя при решении задач прогнозирования). Каждая из этих сетей по-своему решает задачу нелинейной аппроксимации отображения $F : X \rightarrow y$, где X — входной вектор сети, y — выходное (прогнозное) значение. Поэтому разница в прогнозах, выдаваемых экспертами, будет следствием не только случайности начальных значений весовых коэффициентов, но и принципиального различия этих сетей. После завершения этапа обучения все входные данные обучающей выборки кластеризуются, как и при использовании предыдущего подхода. А затем для каждого нейроэксперта определяется коэффициент его компетентности на данных каждого кластера (например, в процессе экспериментального тестирования было замечено, что вероятностная сеть обеспечивает более высокое качество прогноза в условиях наличия растущего тренда курса акции, а сеть, обученная по алгоритму обратного распространения, наоборот, убывающего). В процессе функционирования сети коэффициенты компетентности корректируются в зависимости от величины ошибок прогноза нейроэксперта на данных этого кластера. Объединение экспертов в ансамбль при решении задачи итогового прогнозирования производится с весами, соответствующими коэффициентам компетентности нейроэкспертов для того кластера, в который попадает анализируемый входной вектор. То есть в качестве значения комитета берется взве-

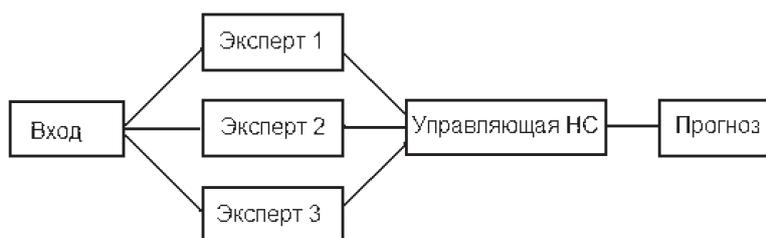


Рис. 1. Архитектура модели прогнозирования комитета нейроэкспертов

шенное среднее арифметическое значение прогнозов всех сетей-экспертов.

4. В четвертом варианте в качестве ответа комитета используется прогноз той из P имеющихся в распоряжении сетей, которая в данный момент t является наилучшей в том смысле, что для нее сумма квадратов ошибок прогнозов за k периодов, предшествующих данному, минимальна:

$$\sum_{s=1}^k (y_i(t-s) - y(t-s))^2 \rightarrow \min_{1 \leq i \leq P}.$$

Здесь $y_i(t-s)$ — прогноз сети с номером i в момент времени $(t-s)$; $y(t-s)$ — истинное значение ряда в этот момент. Величина k задает скорость забывания предыстории и, как показывают численные эксперименты, ее наилучшие значения находятся в промежутке $k \in [3, 5]$.

2. ИСПОЛЬЗОВАНИЕ МЕТОДОВ BOOSTING

Для выработки согласованного решения комитета также предлагается использовать методы бустинга (boosting). Идея бустинга (усиления ансамбля классификаторов) впервые была предложена И. Френдом и Р. Шапиром [2]. Смысл постановки проблемы состоит в следующем.

Пусть имеется некоторый алгоритм классификации объектов по двум классам, точность которого для заданного набора данных лишь незначительно превышает точность угадывания: $p = 1/2 + \varepsilon$. Возможна ли практическая реализация программы обучения этого классификатора, таким образом, что для всяких $0 < \gamma, \delta < 1/2$ с вероятностью более $1 - \gamma$ ошибка классификатора окажется меньше δ ?

В первых вариантах бустинга (BOOST1) рассматривались тройки последовательно обучающихся алгоритмов, достигающих уровня ошибки $\alpha = (1-p) < 1/2$ [3]. Ниже приводится алгоритм применения метода BOOST1 для управления ансамблем, состоящим из трех нейронных сетей.

АЛГОРИТМ

1. Первая нейронная сеть обучается на множестве из m примеров.

2. Вторая сеть также обучается на m примерах, выбираемых так, что первая сеть ровно на половине из них дает точный ответ.

3. Наконец, третья сеть обучается на таких m примерах, на которых ответы первой и второй сетей расходятся.

4. Прогнозом комитета является ответ первых двух сетей, если он одинаков, и ответ третьей сети, если ответы первой и второй сетей разошлись.

Если, как было отмечено, уровень ошибки на тестовом множестве каждой из нейронных сетей равен α , то легко показать, что вероятность ошибки для комитета будет равна $3\alpha^2 - 2\alpha^3 < \alpha$ (при $0 < \alpha < 1/2$), то есть меньше ошибки каждой из сетей.

Отметим, что метод BOOST1 требует, чтобы каждая из сетей комитета осуществляла классификацию обучающего множества исключительно на два класса, что снижает его применимость для решения задач прогнозирования. Однако, например, в системах трейдинга бывает достаточно, чтобы алгоритм прогнозирования выдавал сигнал к покупке (если ожидается рост котировки) или к продаже (если ожидается падение).

Метод AdaBoost, также предложенный Френдом и Шапиром [3], позволяет объединять в комитет произвольное количество сетей, но их обучение и в этом случае должно проводиться последовательно. При использовании метода AdaBoost для управления комитетом все сети могут обучаться на одном и том же множестве, но обучающие векторы приобретают весовые коэффициенты, которые изменяются с течением времени. А именно, для каждой сети комитета, начиная со второй, веса (т.е. значимость) обучающих векторов пересчитываются так, чтобы она точнее настраивалась на тех векторах, на которых чаще ошибались все предыдущие сети.

Как и в методе BOOST1, каждая из сетей комитета должна осуществлять классификацию обучающего множества на два класса. Ответом комитета является взвешенное среднее значение прогнозов всех сетей-экспертов, округляемое в сторону ближайшего допустимого ответа сети. Веса сетей-экспертов вычисляются исходя из числа допущенных ими ошибок.

Обобщающая способность бустинга (применительно к алгоритмам классификации) исследована достаточно хорошо [3]. Во многих случаях качество классификации на тестовой выборке может продолжаться улучшаться даже после достижения безошибочного распознавания обучающей выборки [2]. Бустинг строит выпуклую комбинацию классификаторов, которая проявляет свойство стабильности, то есть

небольшие вариации обучающей выборки приводят к незначительным изменениям получаемого результата.

РЕЗУЛЬТАТЫ И ВЫВОДЫ

1. Как показывают вычислительные эксперименты, использование комитетов нейроэкспертов способно на 20—30 % увеличить точность базовых прогнозов.

2. Наилучшие результаты пока показывают 3-й и 4-й из предложенных в данной статье методов формирования нейросетевых комитетов.

3. В случае, если задачу прогнозирования можно свести к задаче классификации обучающей выборки на два класса (например, периоды, предшествующие падению котировки, и периоды, предшествующие ее росту), представ-

ляется очень перспективным использование методики бустинга для составления ансамблей нейроэкспертов.

СПИСОК ЛИТЕРАТУРЫ

1. *Ежов А. А.* Нейрокомпьютинг и его применения в экономике и бизнесе / А. А. Ежов, С. А. Шумский. — М.: МИФИ, 1998. — 224 с.

2. *Freund Y.* Experiments with a new boosting algorithm / Y. Freund, R. E. Schapire // International Conference on Machine Learning. — 1996. — P. 148—156.

3. *Терехов С. А.* Гениальные комитеты умных машин. IX Всероссийская научно-техническая конференция «Нейроинформатика-2007»: Лекции по нейроинформатике. Часть 2, М., МИФИ, 2007. — С. 11—42.

Каширина Ирина Леонидовна — к. тех. наук, доцент кафедры математических методов исследования операций факультета ПММ Воронежского государственного университета. Тел.: 8-4732-96-33-56, E-mail: kash.irina@mail.ru

Kashirina I. L. — Candidat of Technical Sciences, Associate Professor, the dept. of the Mathematical Methods of Operation Research, Voronezh State University. Tel. 8-4732-96-33-56, E-mail: kash.irina@mail.ru