

Прогнозирование индекса РТС на основе эвристических методов

Дорохов Евгений Владимирович

соискатель

Московский государственный университет им. М.В. Ломоносова, Москва, Россия

E-mail: edorokhov@mail.ru

Наиболее широко используемыми при прогнозировании моделями из класса эвристических являются нейросетевые модели. При проведении данного исследования осуществлялась проверка гипотезы о существовании нелинейной зависимости между прошлыми значениями индекса российского фондового рынка и ряда других экономических показателей. В качестве инструмента идентификации данной зависимости были использованы нейронные сети. 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.

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. Учитывая, что используемый в задаче прогнозирования исходный временной ряд значений объясняемой переменной не является стационарным, был использован ряд преобразований, которые приводят к получению модифицированного ряда с характеристиками, приближенными к характеристикам стационарных рядов.

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; объявленные ставки по привлечению (MIBID) и размещению (MIBOR) межбанковских кредитов; показатели по рынку государственных ценных бумаг; биржевой курс и объем торгов на ММВБ по доллару США и евро; котировки драгоценных металлов; ежедневные цены фьючерсного контракта на нефть.

Массив исходных данных 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).

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.

При проведении исследования было рассмотрено несколько популярных архитектур построения нейронных сетей: архитектура многослойного персептрона (MLP); архитектура радиальных базисных функций (RBF); архитектура линейной сети; архитектура обобщенно-регрессионной нейронной сети (GRNN) [1].

С целью улучшения качества прогнозирования было выполнено определение статистической значимости входной информации для предсказания выходной, для чего использовался специальный алгоритм анализа чувствительности моделей к входным переменным.

Учитывая, что не имело смысла использовать при прогнозировании те переменные, которые характеризуются невысокой значимостью, проводилась фильтрация входов. Для этого использовался специальный генетический алгоритм 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 (то есть 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. Окончательный список использованных в качестве входов нейросети переменных представлен в табл. 1.

Таблица 1. Переменные, использованные в качестве входов нейросети.

Переменная	Значимость	Переменная	Значимость	Переменная	Значимость
объем ГКО – ОФЗ	1.0022	индекс DJI	1.0029	индекс JSE	1.0100
индекс AEX	1.0058	индекс NIKKEI	0.9996	индекс KFX	1.0023
индекс BEL20	1.0017	индекс HSI	1.0096	индекс WIG20	0.9997

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.

В качестве итоговых статистик, характеризующих построенную нейросетевую модель, выступают среднее значение и стандартное отклонение, вычисленные для обучающих данных и для ошибки прогноза. Величина, равная единице минус отношение стандартных отклонений, равна доле объясненной дисперсии модели. В построенной модели отношение стандартных отклонений составило 0.96, то есть многослойный персептрон лишь на 4% предсказал остаток ряда. Это означает, что исходный набор данных не имел ощутимой нелинейной структуры.

Графическая иллюстрация полученного прогноза представлена на рис. 1.

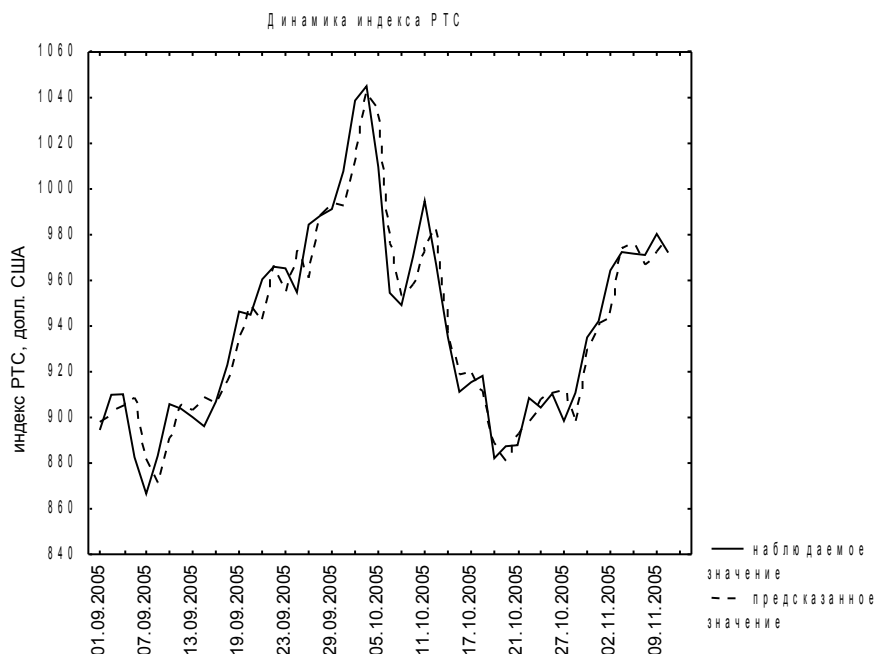


Рис. 1. Прогнозирование индекса РТС нейросетевым методом на основе ряда его значений на 50 дней с помощью процедуры скользящего экзамена.

Литература.

1. Бэстэнс Д.Э., Ван Ден Берг В.М., Вуд Д. Нейронные сети и финансовые рынки: принятие решений в торговых операциях. – М.: ТВП, 1997.