



О.П. Солдатова, Д.А. Панкратова, В.Г. Литвинов

## АНАЛИЗ РЕЗУЛЬТАТОВ ПРОГНОЗИРОВАНИЯ КОТИРОВОК АКЦИЙ С ИСПОЛЬЗОВАНИЕМ КОРРЕЛЯЦИОННОЙ ЗАВИСИМОСТИ ОБУЧАЮЩИХ ВЫБОРОК

(Самарский государственный аэрокосмический университет имени академика С.П. Королева (национальный исследовательский университет))

При прогнозировании котировок акций важно выявить наличие корреляционной зависимости между прогнозируемой последовательностью и последовательностями биржевых цен на другие товары.

В качестве прогнозируемых выборок были использованы котировки акций нефтегазовых компаний: ОАО «Лукойл», ОАО «Роснефть» и ОАО «Транснефть». Для прогнозирования котировок были выбраны цены на нефть двух марок – Brent и WTI, а также цены на мазут, что обусловлено спецификой деятельности данных компаний. Для проверки корреляционных зависимостей были построены графики корреляционных функций, рассчитаны временные задержки, которые затем использовались в качестве одного из параметров для прогнозирования и обучения сети. В результате корреляционного анализа выяснилось, что цены на нефть и мазут имеют задержку в 1 день по отношению к котировкам трех компаний.

Для прогнозирования котировок акций была выбрана нейронная сеть Ванга-Менделя. Сеть Ванга-Менделя относится к классу нечётких продукционных нейронных сетей, которые обладают возможностью работы с неопределёнными, неполными или неточными данными.

Для обучения сети Ванга-Менделя использовались два алгоритма: адаптивный и наискорейшего спуска [1]. Второй алгоритм применялся в совокупности с адаптивным алгоритмом (последовательный метод обучения) и с предварительной инициализацией параметров функции Гаусса в первом слое фuzziфикации с помощью алгоритма  $K$ -усреднений.

При обычном обучении с использованием одномерной обучающей выборки входная последовательность данных подается последовательно, при этом происходит формирование обучающей выборки разделением данных на векторы размерностью, равной заданному прогнозному окну.

При использовании корреляционных зависимостей обучающих выборок обучающая выборка включает одно эталонное значение (известной котировки акции) и по одному элементу последовательностей цен на биржевые товары, коррелирующих с прогнозируемыми данными. Таким образом, получается многомерная обучающая выборка, но сохраняется одномерность входной последовательности. В данном случае размер прогнозного окна равен количеству коррелирующих с прогнозируемыми котировками данными.

Для определения корреляционных зависимостей рассчитывалась корреляционная функция, имеющая вид:



$$cor(X, Y, t) = \frac{\sum_{t=1}^N (X_t - MX)(Y_t - MY)}{\sqrt{\sum_{t=1}^N (X_t - MX)^2 \sum_{t=1}^N (Y_t - MY)^2}},$$

$$\text{где } MX = \frac{\sum_{t=1}^N X_t}{N}, \quad MY = \frac{\sum_{t=1}^N Y_t}{N}, \quad N = \min(N_X, N_Y), \quad N = \min(N_X, N_Y),$$

$$-1 \leq cor(X, Y, t) \leq 1.$$

Следует отметить следующие ограничения корреляционного анализа [2]:

1. Наличие достаточного количества наблюдений для изучения. Данное условие в работе выполняется, поскольку используются выборки объемом не менее 250 значений.
2. Совокупность значений всех факторных и результативного признаков должна подчиняться многомерному нормальному распределению. Условие выполняется также по причине достаточно большого объема выборок.
3. Исходная совокупность значений должна быть качественно однородной. Все выборки представляют собой цены на биржевые товары или котировки акций, поэтому условие однородности удовлетворено.
4. Сам по себе факт корреляционной зависимости не даёт основания утверждать, что одна из переменных предшествует или является причиной изменений, или что переменные вообще причинно связаны между собой, а не наблюдается действие третьего фактора. Действительно, нельзя точно утверждать, что именно скачок цены на тот или иной биржевой товар является причиной роста или падения котировок акций. Изменения котировок могут быть вызваны политическими, природными, экономическими и иными факторами, обуславливающими риски при игре на фондовых биржах.

Примем обозначения:

$r$  – максимальный радиус кластера для адаптивного алгоритма;

$e$  – количество эпох;

$rul$  – количество правил (нейронов в скрытом слое);

$d$  – задержка;

$m$  – число коррелирующих выборок.

Для исследования прогнозирующих способностей сети Ванга-Менделя использовались котировки акций трех компаний: «Лукойл», «Роснефть» и «Транснефть».

В работе рассчитывалась средняя относительная погрешность обучения и прогнозирования. Формула расчета данной погрешности для пары «теоретическое значение-рассчитанное значение» выглядит следующим образом:

$$e = \left| \frac{y_i - y_i^{val}}{y_i^{val} \cdot N} \right|,$$

где  $y_i^{val}$  – вычисленное значение выходного сигнала сети;

$y_i$  – эталонное значение выходного сигнала сети,

$N$  – размерность выборки.



В таблице 1 представлены результаты прогнозирования без использования корреляционных зависимостей и прогнозирования с применением корреляций. Во втором случае в качестве коррелирующих выборок использовались цены на нефть марки «Brent», нефть марки «WTI» и мазут.

Таблица 1 – Результаты прогнозирования без использования корреляционных зависимостей и прогнозирования с применением корреляций

Название компании	Название алгоритма обучения	Относительная погрешность			
		Без корреляции		С корреляцией	
		обучения	тестирования	обучения	тестирования
Лукойл	Адаптивный	0,00072	0,01327	0,00511	0,00813
	Наискорейшего спуска	0,01083	0,00995	0,01038	0,00768
	Последовательный	0,06163	0,04315	0,21195	0,00799
Роснефть	Адаптивный	0,00473	0,00604	0,00574	0,00575
	Наискорейшего спуска	0,00931	0,00457	0,01038	0,00499
	Последовательный	0,01337	0,00526	0,01637	0,00535
Транснефть	Адаптивный	0,00294	0,00767	0,00567	0,00849
	Наискорейшего спуска	0,01461	0,00937	0,01054	0,00787
	Последовательный	0,63868	0,54525	0,03954	0,00992

Из таблицы 1 видно, что при использовании корреляционных зависимостей значительно уменьшаются погрешности прогнозирования, особенно, в случае адаптивного алгоритма обучения. Алгоритм наискорейшего спуска не всегда дает лучший результат при использовании корреляции.

Для оценки корреляционных зависимостей между выборками использовались задержки – количество дней от начала отсчета до момента времени, когда корреляционная функция впервые достигла своего максимального значения. В таблице 2 приведены значения погрешности прогнозирования для разного числа коррелирующих выборок. При этом использовались выборки с одинаковыми значениями задержки: нефть марки «Brent», нефть марки «WTI» и мазут – в качестве коррелирующих выборок; котировки акций «Лукойла» – как прогнозируемые.

Как видно из таблицы 2, наилучший результат мы получаем при использовании двух выборок, связанных корреляционной зависимостью с прогнозируемой. Это значит, что наилучший размер прогнозного окна равен 2. Данный факт был замечен при исследовании прогнозирующих способностей сети Ван-га-Менделя при прогнозировании учетных цен на драгоценные металлы [3]. Кроме того, котировки акций «Лукойла» сопоставимы с ценами на золото, что позволяет использовать выводы, полученные в указанном исследовании, в настоящей работе.



Таблица 2 – Значения погрешности прогнозирования и обучения для разного числа коррелирующих выборок

Название алгоритма обучения	Относительная погрешность					
	m=1		m=2		m=3	
	обучения	тестирования	обучения	тестирования	Обучения	тестирования
Адаптивный	0,00607 (e=10, r=7)	0,00783	0,00481 (e=2, r=4)	0,00679	0,00511 (e=10, r=5)	0,00813
Наискорейшего спуска	0,01018(rul=15, e=6)	0,00719	0,00851 (rul=30, e=40)	0,00718	0,00786 (rul=15, e=11)	0,00750
Последовательный	0,01931(e=2, r=10)	0,00960	0,02891 (e=3, r=0,8)	0,00762	0,21195 (e=8, r=0,5)	0,00799

Таким образом, в ходе исследования прогнозирующих способностей сети Ванга-Менделя с использованием корреляционных зависимостей выборок были сделаны следующие выводы:

- использование корреляционных зависимостей улучшает прогнозирующие способности сети Ванга-Менделя, на порядок уменьшая ошибку прогнозирования;
- наилучшие результаты прогнозирования получаются при использовании двух коррелирующих выборок;
- следует использовать коррелирующие выборки со средними задержками (не более 6-7 дней);
- при прогнозировании с использованием корреляций наиболее эффективны адаптивный алгоритм и алгоритм обратного распространения ошибки с инициализацией алгоритмом  $K$ -усреднений.

#### Литература

1 Осовский, С. Нейронные сети для обработки информации [Текст]/ С. Осовский: Пер. с пол. Рудинского И.Д. – М.: Финансы и статистика, 2002. – 344 с.: ил.;

2 Корреляция [Электронный ресурс]. – <https://ru.wikipedia.org/wiki/Корреляция> (дата обращения 09.03.2015 г.);

3 Солдатова О.П., Панкратова Д.А. Анализ результатов прогнозирования учетных цен на драгоценные металлы с помощью нечеткой нейронной продукционной сети Ванга-Менделя // Труды международной научно технической конференции «Перспективные информационные технологии (ПИТ 2014)» (Самара, СГАУ, 30 июня–2 июля 2014 г.). – Самара: Издательство Самарского научного центра РАН, 2014. – с. 165–169;

4 Цены сырьевых товаров [Электронный ресурс]. – <http://www.finanz.ru/birzhevyye-tovary>.