

## Совместное применение модели линейной регрессии и нейронной сети в задаче предсказания тренда котировок криптовалюты Bitcoin

*Е.В. Алымова*

*Российская таможенная академия (Ростовский филиал), Ростов-на-Дону*

**Аннотация:** В статье представлен комбинированный подход с использованием методов машинного обучения для выбора эффективной торговой стратегии на валютной бирже. Представленный подход использует расчет углового коэффициента линейной регрессии по индикаторам логарифмической доходности и определение тренда котировок валютной пары BTC/USD в следующем периоде на основе рассчитанного знака коэффициента. Многослойная нейронная сеть прямого распространения выполняет предсказание значения углового коэффициента в следующем десятиминутном периоде для текущего двадцатиминутного периода. В статье предлагается комбинированный подход к использованию методов машинного обучения для выбора эффективной торговой стратегии на валютной бирже. В исследовании представлены результаты практических экспериментов, оценивающих соотношение эффективных и неэффективных стратегий на основе предсказанных значений коэффициентов линейной регрессии.

**Ключевые слова:** машинное обучение, нейронная сеть, финансовые временные ряды, предсказание поведения котировок валют, коэффициенты линейной регрессии.

Предсказание поведения финансовых временных рядов является необходимым элементом инвестиционной деятельности, а обоснованный выбор стратегии торгов на валютной бирже значительно повышает вероятность прибыльности инвестиций.

Вопросы эффективности применения методов машинного обучения в задаче прогнозирования поведения финансовых временных рядов рассматриваются многими авторами. В работе [1] представлено решение задачи прогнозирования тренда композитного индекса Шанхайской фондовой биржи на момент закрытия торгового периода на основе технологий Deep Learning с использованием многослойной нейронной сети прямого распространения (LSTM). Использованию нейронной сети LSTM для моделирования поведения финансовых временных рядов также посвящены работы [2, 3]. В работе [4] применяется метод опорных векторов для обнаружения сигналов к изменению условий торгов, влияющих на поведение финансового временного ряда. Возможность прогнозирования реализованной волатильности котировок криптовалюты с помощью нейронной сети рассматривается в работе [5]. В работе [6] приводятся результаты сравнения эффективности прогнозирования поведения финансовых временных рядов с помощью нейронной сети LSTM и интегрируемой модели авторегрессии и модели скользящего среднего (ARIMA). Отмечается, что разработка

---

модели поведения временного ряда на основе нейронной сети дает более точный результат прогноза. Вопрос эффективности комбинирования моделей прогнозирования временных рядов поднимается в работе [7] на примере сравнения точности пяти прогнозов будущих трат энергетических ресурсов. Задаче моделирования динамики криптовалют посвящены работы [8, 9].

Несмотря на большое количество проведенных исследований в этом направлении, задача предсказания поведения финансового временного ряда остается актуальной и востребованной в практической деятельности трейдеров валютного рынка.

Представленная работа посвящена исследованию эффективности комбинирования методов машинного обучения в решении задачи прогнозирования поведения котировок криптовалют. Целью исследования является разработка методики обоснованного выбора стратегии ведения торгов на валютной бирже в следующем десятиминутном периоде, основываясь на данных текущего двадцатиминутного периода торгов. Данная работа является продолжением исследования [10], в котором получена модель классификации поведения финансового временного ряда на основе показателей логарифмической доходности валютной пары BTC/USD, реализованная в форме нейронной сети.

В качестве исходной информации используются исторические данные торгов валютной пары BTC/USD в режиме реального времени с интервалом в одну минуту. На основе исходных данных проводятся расчеты показателя логарифмической доходности, как натурального логарифма отношения цены закрытия торгов на следующей минуте к цене закрытия торгов на текущей минуте.

В работе сделано предположение, что показатели логарифмической доходности в текущем двадцатиминутном периоде определяют показатели логарифмической доходности в следующем десятиминутном периоде. Для проверки этого предположения по каждому двадцати значениям логарифмической доходности определяется направление линейного тренда поведения котировок целевой валютной пары в следующие десять минут.

Для обучения нейронной сети используются вектор  $\bar{X} = \{X_1, X_2, \dots, X_m\}$  независимых переменных и вектор  $\bar{Y} = \{Y_1, Y_2, \dots, Y_m\}$  зависимых переменных ( $m$  – количество десятиминутных интервалов в исходных данных, для которых вычислены коэффициенты линии тренда), определенные следующим образом:

$$X_n = \{LOGRET_{n+1}, LOGRET_{n+2}, \dots, LOGRET_{n+20}\};$$

$$Y_n = \{LINREGB_{n+21, n+30}\},$$

где

- $LOGRET_i$  – показатель логарифмической доходности на  $i$ -й минуте ( $i \geq 1$ );
- $LINREGB_{j,k}$  – угловой коэффициент  $B$  уравнения линейной регрессии, построенной по показателям логарифмической доходности  $LOGRET_j, \dots, LOGRET_k$ , ( $k > j$ );
- $n$  – номер наблюдаемого периода ( $n \geq 0$ ).

Предсказание направления тренда котировок наблюдаемой валютной пары в следующем десятиминутном периоде проводится по текущему двадцатиминутному периоду, используя знак углового коэффициента  $B$  уравнения линейной регрессии. Если  $B \geq 0$ , то прогнозируется рост цены на криптовалюту (сигнал к покупке), в противном случае прогнозируется падение цены на криптовалюту (сигнал к продаже).

Таким образом выделяются условия принятия решения о покупке или продаже в следующем десятиминутном периоде:

- стратегия 1: если на конец текущего двадцатиминутного периода коэффициент  $B \geq 0$ , то на первой минуте следующего десятиминутного периода должна быть совершена покупка криптовалюты по текущей цене с тем, чтобы в последующие девять минут совершить продажу по первой более высокой цене;
- стратегия -1: если на конец текущего двадцатиминутного периода коэффициент  $B < 0$ , то на первой минуте следующего десятиминутного периода должна быть совершена продажа криптовалюты по текущей цене с тем, чтобы в последующие девять минут совершить покупку по первой более низкой цене.

Выбранная стратегия может оказаться нерезультативной, если цена на начало следующего десятиминутного периода является самой большой для сигнала к покупке или самой маленькой для сигнала к продаже. Чтобы

проверить результативность стратегий, в работе используются следующие критерии:

- если на первой минуте следующего десятиминутного периода коэффициент  $B \geq 0$ , то цена на момент открытия должна быть меньше всех максимальных цен в следующем десятиминутном периоде. В этом случае стратегия 1 считается результативной;
- если на первой минуте следующего десятиминутного периода коэффициент  $B < 0$ , то цена на момент закрытия должна быть больше всех минимальных цен в следующем десятиминутном периоде. В этом случае стратегия -1 считается результативной.

Проверка сформулированных критериев на исторических данных показывает, что на каждые 5 результативных стратегий приходится одна нерезультативная, то есть в каждом шестом случае возможно принятие ошибочного решения о покупке или продаже криптовалюты.

Для минимизации количества нерезультативных стратегий вводится стратегия 0 ожидания, когда никакие действия по покупке или продаже криптовалюты не предпринимаются. Стратегия 0 выбирается в том случае, если значение коэффициента  $B$  на первой минуте следующего десятиминутного периода является незначительным, то есть коэффициент  $B$  по модулю меньше заданного числа  $h$ . Подбор числа  $h$  в работе произведен эмпирически на исторических данных и равен 0.0003, что позволило повысить число результативных стратегий. При вводе стратегии 0 в каждом восьмом случае возможно принятие ошибочного решения о покупке или продаже криптовалюты.

В данной работе построение и обучение нейронной сети ведется на аналитической платформе Rapid Miner. На рис. 1 представлена цепочка процесса обучения и проверки модели предсказания значения углового коэффициента линейной регрессии. Блок Deep Learning реализует многослойную нейронную сеть прямого распространения, которая обучается с использованием метода стохастического градиентного спуска по обратному распространению ошибки.

В нейронной сети используется четыре скрытых слоя. Каждый нечетный слой содержит 100 нейронов, а каждый четный – 50 нейронов. В каждом скрытом слое используется функция линейного выпрямления (ReLU) активации нейронов. Данные параметры нейронной сети подобраны эмпирически для получения большей точности прогнозирования. При данных параметрах значение среднеквадратической ошибки прогнозирования составляет (RMSE) 0.000086186.

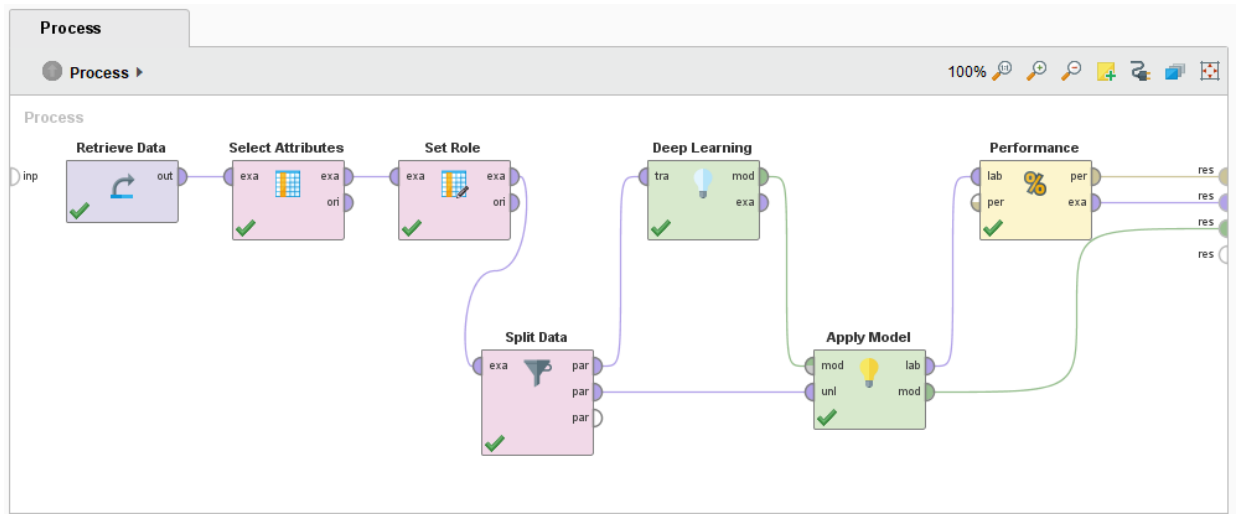


Рис. 1 – Процесс обучения и проверки модели прогнозирования в Rapid Miner

На рис. 2 представлен график сравнения фактических и предсказанных значений углового коэффициента линейной регрессии. Из графика можно сделать вывод, что построенная модель регрессии на основе коэффициентов линии тренда логарифмической надежности в десятиминутных периодах согласуется с реальным процессом.

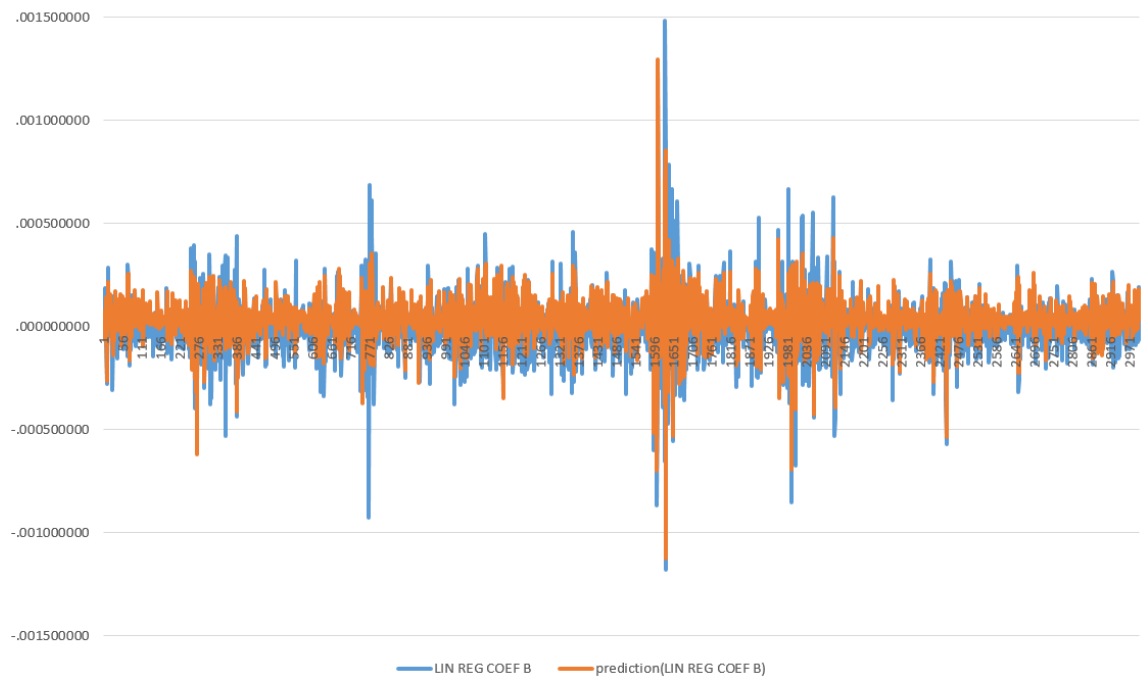


Рис. 2. – Сравнение значений фактических и предсказанных угловых коэффициентов линейной регрессии

Результатом представленного исследования является методика выбора стратегии ведения торгов на валютной бирже в следующем десятиминутном периоде, основываясь на данных текущего двадцатиминутного периода торгов. Выделено три стратегии: покупка в начале следующего периода с целью продажи по первой более высокой цене (стратегия 1), продажа в начале следующего периода с целью покупки по первой более низкой цене (стратегия -1) и бездействие (стратегия 0), если угловой коэффициент линейной регрессии на начало следующего десятиминутного периода является незначительным по выбранному критерию.

На исторических данных проведены расчеты, показывающие, что без введения стратегии 0 каждое шестое решение в выборе стратегии является ошибочным, а при введении стратегии 0 ошибочным является каждое восьмое решение.

Для предсказания значения углового коэффициента построена нейронная сеть прямого распространения, параметры которой подобраны эмпирически с целью повышения точности прогнозирования.

**Благодарность.** Исследование выполнено при финансовой поддержке РФФИ, проект № 18-01-00910А.

### Литература

- 1 Yan H., Ouyang H. Financial time series prediction based on deep learning // Wireless Personal Communications. – 2018. – Т. 102. – №. 2. – pp. 683-700.
- 2 Кондратьева Т. Н. Прогнозирование тенденции финансовых временных рядов с помощью нейронной сети LSTM // Вестник евразийской науки. – 2017. – Т. 9. – №. 4. – С. 61-67.
- 3 Лабусов М.В. Нейронные сети долгой краткосрочной памяти и их использование для моделирования финансовых временных рядов // Инновации и инвестиции. – 2020. – №. 4. – С. 167-171.
- 4 Law T., Shawe-Taylor J. Practical Bayesian support vector regression for financial time series prediction and market condition change detection // Quantitative Finance. – 2017. – Т. 17. – №. 9. – pp. 1403-1416.
- 5 Shintate T., Pichl L. Trend prediction classification for high frequency bitcoin time series with deep learning // Journal of Risk and Financial Management. – 2019. – Т. 12. – №. 1. – pp. 17-33.

- 6 Алжеев А. В., Кочкаров Р. А. Сравнительный анализ прогнозных моделей ARIMA и LSTM на примере акций российских компаний // Финансы: теория и практика. – 2020. – Т. 24. – № 1.– С. 14-23.
- 7 Гребнев В.И., Немцев М.А., Тюков А.П. Оценка эффекта комбинирования моделей прогнозирования временных рядов // Инженерный вестник Дона, 2020, №5. URL: ivdon.ru/ru/magazine/archive/N5y2020/6481.
- 8 Гречко А.С., Кудрявцев О.Е.. Калибровка умеренно устойчивых моделей Леви по данным криптовалют Bitcoin и Ethereum // Инженерный вестник Дона, 2019, №8. URL: ivdon.ru/ru/magazine/archive/N8y2019/6113.
- 9 Сафиуллин М. Р., Абдукаева А. А., Ельшин Л. А. Методические подходы к прогнозированию динамики курса криптовалют с применением инструментов стохастического анализа (на примере биткоина) // Финансы: теория и практика. – 2018. – Т. 22. – № 4.– С. 38-51.
- 10 Alymova E. V., Kudryavtsev O. E. Neural networks usage for financial time series prediction. Abstracts of Talks Given at the 4th International Conference on Stochastic Methods // Theory of Probability & Its Applications, 2020, Vol. 65, № 1, pp. 122-123.

### References

- 1 Yan H., Ouyang H. Wireless Personal Communications. 2018. Vol. 102. № 2. pp. 683-700.
  - 2 Kondratyeva T. N. Vestnik evraziyskoy nauki. 2017. V. 9. № 4. pp. 61-67.
  - 3 Labusov M.V. Innovacii i investicii. 2020. № 4. pp. 167-171.
  - 4 Law T., Shawe-Taylor J. Quantitative Finance. 2017. Vol. 17. № 9. pp. 1403-1416.
  - 5 Shintate T., Pichl L. Journal of Risk and Financial Management. 2019. Vol. 12. № 1. pp. 17-33.
  - 6 Alzheev A. V., Kochkarov R. A. Finansy: teoriya i praktika 2020. Vol. 24. №1. pp. 14-23.
  - 7 Grebnev V.I., Nemtsev M.A., Tyukov A.P. Inzhenernyj vestnik Dona, 2020, №5. URL: ivdon.ru/ru/magazine/archive/N5y2020/6481.
  - 8 Grechko A.S., Kudryavtsev O. E. Inzhenernyj vestnik Dona, 2019, №8. URL: ivdon.ru/ru/magazine/archive/N8y2019/6113.
  - 9 Safiullin M. R., Abdukaeva A. A., El'shin L. A. Finansy: teoriya i praktika. 2018. Vol. 22. № 4. pp. 38-51.
  - 10 Alymova E. V., Kudryavtsev O. E. Theory of Probability & Its Applications, 2020, Vol. 65, № 1, pp. 122-123.
-