

УДК 658.6
DOI: 10.25559/SITITO.16.202002.510-517

Применение глубокого обучения с подкреплением в алгоритмической торговле

П. В. Никитин*, Р. И. Горохова, С. А. Корчагин, В. С. Красников

ФГБОУ ВО «Финансовый университет при правительстве Российской Федерации», г. Москва, Россия

125993, Россия, ГСП-3, г. Москва, Ленинградский пр., д. 49

* pvnikitin@fa.ru

Аннотация

На текущий момент существует большой объем литературы по торговле на бирже. Очевидно, что с каждым годом математическая база работ усложняется вместе с увеличением вычислительной мощности, машины за единицу времени могут обрабатывать больше метрик от года к году и выдавать более точные решения. Применение глубокого обучения уже хорошо зарекомендовало себя, так как применение такого подхода дало качественный скачок в алгоритмической торговле. В статье представлен алгоритм торговли длинными контрактами одним активом на финансовом рынке на языке программирования Python с применением нейронной сети LSTM при помощи библиотеки Keras, который используется в качестве демонстрационного примера по дисциплине «Обучение с подкреплением». Формализовано модель LSTM решает проблему исчезающих градиентов, которая может удерживать градиент целевой функции относительно сигнала состояния. В применении к нашей задаче такое улучшение модели позволяет собирать данные о тех или иных паттернах изменения цены, то есть при предсказании цены следующего шага опираемся не только на данные предыдущего шага, но и на более ранние данные, когда было схожее состояние среды. Коэффициент Шарпа используется для определения оптимальной стратегии и принятия решения в каждый момент времени применения. Определены оптимальный минимальный временной промежуток для работы модели; задержка передачи сигнала от момента изменения ситуации на рынке до приема сигнала моделью, которая будет бесконечно малой, а вычислительную мощность будем считать бесконечно большой. Эти допущения дают право говорить: при изменении ситуации на рынке модель мгновенно готова среагировать и принять решение о продаже, покупке или удержании актива.

Ключевые слова: алгоритмическая торговля, глубокое обучение, обучение с подкреплением, рекуррентные нейронные сети, модель LSTM, демонстрационный пример.

Для цитирования: Никитин, П. В. Применение глубокого обучения с подкреплением в алгоритмической торговле / П. В. Никитин, Р. И. Горохова, С. А. Корчагин, В. С. Красников. – DOI 10.25559/SITITO.16.202002.510-517 // Современные информационные технологии и ИТ-образование. – 2020. – Т. 16, № 2. – С. 510-517.

© Никитин П. В., Горохова Р. И., Корчагин С. А., Красников В. С., 2020



Контент доступен под лицензией Creative Commons Attribution 4.0 License.
The content is available under Creative Commons Attribution 4.0 License.



Applying Deep Reinforcement Learning to Algorithmic Trading

P. V. Nikitin*, R. I. Gorokhova, S. A. Korchagin, V. S. Krasnikov

Financial University under the Government of the Russian Federation, Moscow, Russia
49 Leningradskiy Ave., Moscow 125993, GSP-3, Russia

* pvnikitin@fa.ru

Abstract

At the moment, there is a large volume of literature on exchange trading. Obviously, every year the mathematical base of work is becoming more complicated along with an increase in computing power, machines can process more metrics from year to year and produce more accurate solutions per unit of time. The use of deep learning has already proven itself well, as the application of this approach has given a quantum leap in algorithmic trading. The article presents an algorithm for trading long contracts with one asset in the financial market in the Python programming language using the LSTM neural network using the Keras library, which is used as a demo example in the Reinforcement Learning discipline. The formalized LSTM model solves the vanishing gradient problem, which can hold the gradient of the objective function relative to the state signal. As applied to our problem, such an improvement in the model allows us to collect data on certain patterns of price changes, that is, when predicting the price of the next step, we rely not only on the data of the previous step, but also on earlier data, when there was a similar state of the environment. Sharpe Ratio is used to determine the optimal strategy and make decisions at each time of application. The optimal minimum time period for the model operation has been determined; the signal transmission delay from the moment the market situation changes until the signal is received by the model, which will be infinitely small, and the computing power will be considered infinitely large. These assumptions give the right to say: when the market situation changes, the model is instantly ready to react and make a decision to sell, buy or hold an asset.

Keywords: algorithmic trading, deep learning, reinforcement learning, recurrent neural networks, LSTM model, demo example.

For citation: Nikitin P.V., Gorokhova R.I., Korchagin S.A., Krasnikov V.S. Applying Deep Reinforcement Learning to Algorithmic Trading. *Sovremennye informacionnye tehnologii i IT-obrazovanie* = Modern Information Technologies and IT-Education. 2020; 16(2):510-517. DOI: <https://doi.org/10.25559/SITITO.16.202002.510-517>



Введение

Торговля на бирже в настоящее время набирает все большие обороты и становится все актуальнее. На текущий момент существует большой объем литературы по торговле на бирже. Данные вопросы рассматривали как отечественные, так и зарубежные исследователи. Проблемы и перспективы алгоритмической торговли на финансовых рынках рассмотрены в работах [1-7]. Внедрение вычислительных систем и увеличение вычислительных мощностей делает необходимым применение глубокого обучения. Вопросы непосредственного рассмотрения глубокого обучения, его применения к моделированию в алгоритмической торговле, обработки разнородной информации ценообразования, коммерческой ценности и другие нашли свое отражение в целом ряде исследований [8-17]. Вместе с этим происходит усложнение математической базы выполняемых работ, которая и влечет увеличение вычислительной мощности, благодаря чему машины за единицу времени могут обрабатывать больше метрик от года к году и выдавать более точные решения [18]. Очевидно, что с каждым годом математическая база работ усложняется вместе с увеличением вычислительной мощности, машины за единицу времени могут обрабатывать больше метрик от года к году и выдавать более точные решения [19-22]. Применение глубокого обучения уже хорошо зарекомендовало себя на практике, так как применение такого подхода дало качественный скачок в алгоритмической торговле [23, 24].

В определенной степени алгоритмическая торговля может быть применена в той или иной степени на каждом финансовом рынке, где данные не разрежены [25, 26, 27]. Такому критерию отвечают валютный, индексный рынки, голубые фишки, рынок казначейских облигаций. В рамках данного исследования остановимся на рынке голубых фишек и индексов. Рынок казначейских облигаций слишком неповоротлив, а для получения более-менее достойных результатов на валютном рынке необходимо рассмотреть большое количество значимых факторов.

Инвестиционные банки, хедж-фонды и другие участники финансового рынка с определенной периодичностью реструктурируют портфель ценных бумаг, обращающихся на рынке. При принятии решения о включении актива или его производных в портфель исследуются: волатильность (риск) и ожидаемая доходность (математическое ожидание) актива [16, 21, 24]. Для этого важно понимать долгосрочный и среднесрочный тренд и целевую цену с определенной точностью, исходя из этих данных можем формировать портфель.

При торговле одним активом у обучаемого агента есть 3 варианта действий: купить, продать, держать. Вообще говоря, действий может быть намного больше: занять короткую позицию, выйти в наличные, взять опцион и так далее, но в рамках работы рассмотрим конкретную задачу. Будем опираться на то, что в любой момент времени можем оценить стоимость активов в соответствии с актуальной ценой на рынке, оценить риск в соответствии с историческими данными, математическое ожидание стоимости актива на следующем шаге, также не забудем учесть комиссию при совершении транзакции, иначе задача будет оторвана от реальности. Также в случае рассмотрения средне- и долгосрочной торговли необходимо учитывать ставку дисконтирования по базовой валюте.

При решении задачи можем опираться на спрос-предложение

по активу. Эти данные говорят об общем настрое рынка в текущий момент на рассматриваемый актив. В случае, если разница между спросом и предложением значительная, то есть спрос меньше предложения по цене, то рынок находится в равновесии, агент ожидает бокового тренда (Таблица 1).

Таблица 1. Рынок находится в равновесии

Table 1. The market is in equilibrium

Сторона	Объем, млн. у.е.	Цена, у.е.
Предложение	3000	107
Предложение	2000	106
Предложение	1000	105
Спрос	1000	103
Спрос	1800	102
Спрос	2700	101

В случае, если рынок разрежен в части спроса, тогда эта ситуация является драйвером к продаже, так как незначительные колебания на рынке могут спровоцировать срабатывание стоп-лоссов и резкое снижение стоимости актива (Таблица 2).

Таблица 2. Рынок разрежен в части спроса

Table 2. The market is thin on the demand side

Сторона	Объем, млн. у.е.	Цена, у.е.
Предложение	3000	107
Предложение	2000	106
Предложение	1000	105
Спрос	100	103
Спрос	250	102
Спрос	330	101

Полностью противоположная ситуация в случае с разреженностью рынка в части предложения (Таблица 3).

Таблица 3. Рынок разрежен в части предложения

Table 3. The market is thin on the supply side

Сторона	Объем, млн. у.е.	Цена, у.е.
Предложение	3000	107
Предложение	2000	106
Предложение	1000	105
Спрос	1000	103
Спрос	1800	102
Спрос	2700	101

Цель исследования

Описать и реализовать алгоритм торговли длинными контрактами одним активом на финансовом рынке на языке программирования Python с применением нейронной сети LSTM при помощи библиотеки Keras, который можно будет использовать в качестве демонстрационного примера при обучении студентов по дисциплине «Обучение с подкреплением».



Основная часть

В применении обучения с подкреплением к торговле на финансовом рынке наблюдением будут являться цена актива, объем и цена спроса и предложения, базовая ставка валюты, в котором торгуется актив.

Действием является покупка/продажа/удержание актива. Вознаграждением является изменение показателя Шарпа.

Политика обучаемого агента заключается в максимизации показателя Шарпа. В математических терминах:

$$\pi^* = \underset{\pi}{\operatorname{argmax}} E[Sh|\pi]$$

$$Sh = \sum_{t=0}^{\infty} \gamma^t sh_t$$

$$sh_t = E[R_t - R_0] / \sqrt{\operatorname{var}[R_t - R_0]},$$

где R_t – доходность актива в месяц t ,

R_0 – доходность безрискового актива за месяц t .

γ – показатель дисконтирования, ($\gamma \in [0,1]$). Параметр определяет важность будущим вознаграждений и мотивирует агента к действию. Чем выше данный показатель – тем более модель ориентирована на долгосрочный результат.

Принцип работы LSTM модели заключается в следующем. В процессе использования рекуррентных нейронных сетей возникла следующая проблема: при обучении градиент «затухает», другими словами, модель «забывает» входные данные из предыдущего набора. Для решения этой проблемы была изобретена модель LSTM, которая определяет каким образом будут учитываться текущие данные на следующем шаге. Формализовано модель LSTM решает проблему исчезающих градиентов, которая может удерживать градиент целевой функции относительно сигнала состояния. В применении к нашей задаче такое улучшение модели позволяет собирать данные о тех или иных паттернах изменения цены, то есть при предсказании цены следующего шага опираемся не только на данные предыдущего шага, но и на более ранние данные, когда было схожее состояние среды.

Для решения проблемы давайте разделим данные предыдущего и текущего состояния на 2 части:

$$\vec{s}[n+1] = \vec{F}_s(\vec{s}[n-1]) + \vec{F}_u(\vec{r}[n-1], \vec{x}[n])$$

$$\vec{r}[n-1] = G_d \vec{s}[n], \text{ где}$$

$G_d \vec{s}$ – гиперболический тангенс активации нейрона.

Первая часть уравнения для $\vec{F}_s(\vec{s}[n-1])$ – принимает на вход данные предыдущего состояния.

Вторая часть: $\vec{F}_u(\vec{r}[n-1], \vec{x}[n])$ комбинирует информацию с предыдущего результата и текущего состояния и объединяет данные в определенной пропорции, управлять пропорцией можем с помощью соответствующих коэффициентов:

$$\vec{s}[n] = \vec{g}_{cs}[n] \odot \vec{F}_s(\vec{s}[n-1]) + \vec{g}_{cu}[n] \odot \vec{F}_u(\vec{r}[n-1], \vec{x}[n])$$

$$\vec{0} \leq \vec{g}_{cs}[n], \vec{g}_{cu}[n] \leq \vec{1}$$

Кроме этого, модель имеет «фильтр забывания», этот слой позволяет исключить незначимые значения предыдущего шага, подстраиваться под изменяющуюся среду и быть более гибкой.

$$\vec{F}_s(\vec{v}[n-1], \vec{x}[n]) = W_s \vec{v}[n-1] + \vec{g}_{cx}[n] \odot W_x \vec{x}[n] + \vec{\theta}_s$$

На основе данных, поступающих от обоих блоков, экстраполируется значение следующего состояния, модель выдает ответ.

В текущей конфигурации модель с каждым шагом будет переобучаться, то есть значения параметров будут подгоняться

под обучающие данные. Для предотвращения переобучения модели необходимо случайным образом исключать связь некоторых нейронов на каждом шаге обучения с фиксированной вероятностью.

Так как мы применяем метод обучения с учителем, то входные временные ряды преобразуются в выборку с одной меткой, длина временного ряда – T , размер окна – W , вход выборки: $s(x_t, x_{t+1}, \dots, x_t + W)$. Как функцию активации будем использовать ReLU.

В качестве функции потерь используем MSE.

В качестве функции оптимизации параметров используем ADAM - Adaptive Moment Estimation.

Для определения корректировок параметров будем определять среднюю скользящую градиентов:

$$m_p = \beta_1 m_{p-1} + (1 - \beta_1) \cdot \nabla L(w_{p-1})$$

m_p – значение скользящего среднего градиента на шаге p

β_1 – корректировочный коэффициент

∇L – штрафная функция

w_{p-1} – веса, в нашем случае, нейронной сети на шаге $p - 1$

Также будем определять разброс значений градиента:

$$v_p = \beta_2 v_{p-1} + (1 - \beta_2) \cdot (\nabla L(w_{p-1}))^2$$

v_p – разброс значений градиента на шаге p

β_2 – корректирующий коэффициент

Скользящая средняя и уровень разброса значений приближают первый и второй момент градиентов. Уровень разброса значений градиента v_p используется для изменения скорости обучения, а m_p как поправку для изменения значения параметров (весов) модели.

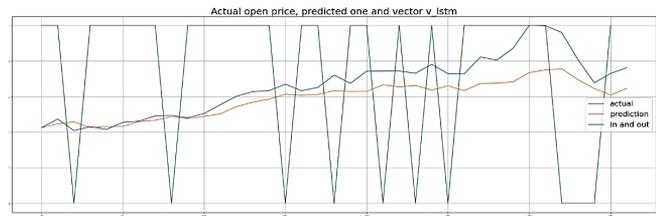
Формула изменения параметров модели записывается следующим образом:

$$w_p = w_{p-1} - \frac{\eta \cdot \widehat{m}_p}{\sqrt{\widehat{v}_p} + \epsilon}$$

С помощью коэффициентов η и ϵ можем корректировать скорость подбора значений параметров модели.

Полученные результаты исследования и их обсуждение

Для апробации модели использовались данные тикера AAPL с 14.03.2000 по 13.03.2019, как безрисковая ставка была использована ставка Федеральной резервной системой США за аналогичный период.



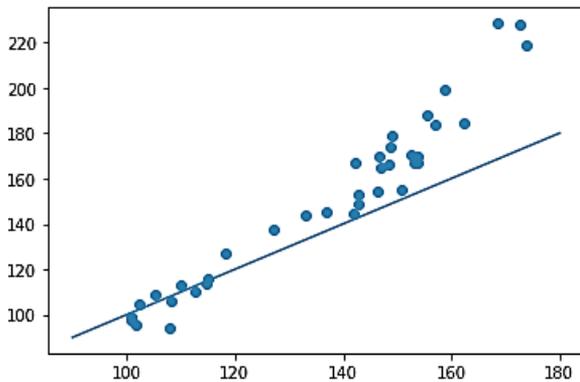
Р и с. 1. График модели за 36 месяцев

F i g. 1. Model graph for 36 months

На графике, представленном на рисунке 1, можем наблюдать, что за 36 месяцев модель советует занимать 9 длинных позиций разной длительности, из которых лишь одна с 0 валовой доходностью. Данный эксперимент говорит о том, что алгоритм имеет достаточно высокую точность для планирования управления наличными. Мы можем понять, сколько времени в течение нескольких лет деньги будут находиться в активе, а сколько бездействовать.



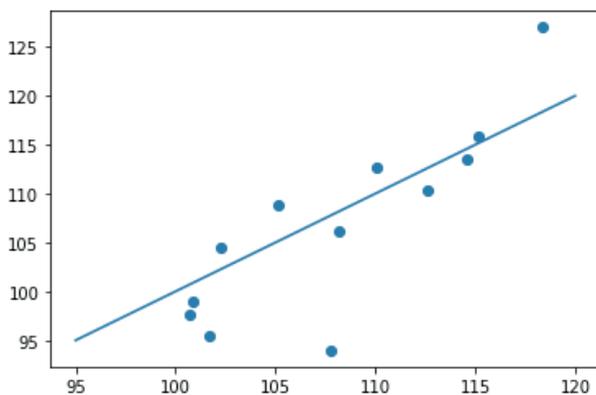
На рисунке 2 представлен график отклонения предсказанных цен от реальных на следующие 3 года после обучающей выборки. На оси ox отложены предсказанные цены, на оси oy отложены реальные цены.



Р и с. 2. График отклонения предсказанных цен
Fig. 2. Price forecast deviation chart

Интерпретация графика: график $x=y$ – идеальное предсказание, чем дальше точка находится от графика $x=y$ – тем хуже модель предсказала значение цены. В случае, если точка лежит ниже графика ($x < y$), модель недооценила актив, то есть мы могли не войти в позицию и упустили бы возможность заработать. В случае, если точка лежит выше графика ($x > y$), модель переоценила актив, мы могли ошибочно войти в позицию и потерять деньги.

Средняя ошибка предсказания составила 15,99%, при этом цена актива выросла на 42%, следовательно ошибка составила 38% от изменения цены актива. Это значит, что с помощью модели на трехлетнем горизонте можем предсказать общий тренд движения и приблизительную скорость роста актива. На рисунке 3 представлены отклонения предсказанной цены от реальной при годовом прогнозировании.



Р и с. 3. График отклонения предсказанной цены от реальной
Fig. 3. The graph of price forecast deviation from real

Средняя ошибка предсказания составила 5% при росте цены актива на 29%, следовательно, ошибка составила 16% от изменения цены актива. Это значит, что на годовом горизонте модель также предсказывает тренд и более точно скорость роста актива.

Заключение

Разработанная модель позволяет проводить ежемесячный тренд, на основании технического анализа предсказать тренд на 3 года, прогнозировать скорость роста актива, волатильность актива.

Исходя из этих данных можем планировать портфель в долгосрочной перспективе, как часто можем переключаться из одного актива в другой.

Для улучшения качества предсказания имеет смысл во входные данные включить годовую и квартальную отчетность компании, выпустившей ценные бумаги. С помощью этих данных сможем рассчитать такие показатели как уровень долга по отношению к выручке, уровень операционной прибыли, свободный денежный поток и динамику этих показателей во времени. Наиболее значимые показатели включить в политику агента, тем самым улучшив модель. Большой объем финансовых данных о компании предоставят модели более обширное пространство наблюдений. Во-вторых, необходимо глубже проработать методику вознаграждений модели, чтобы цель алгоритма была ближе к цели трейдера – извлечь из торговли как можно большую прибыль. С другой стороны, интересным направлением для исследования является рассмотрение распределения, в соответствии с которым определяется цена актива, что поможет управлять неопределенностью.

Список использованных источников

- [1] Батюк, Б. В. Проблемы и перспективы алгоритмической торговли на финансовых рынках / Б. В. Батюк. – DOI 10.24182/2073-9885-2020-13-2-9-16 // Путеводитель предпринимателя. – 2020. – Т. 13, № 2. – С. 9-16. – URL: <https://www.elibrary.ru/item.asp?id=42763702> (дата обращения: 11.05.2020). – Рез. англ.
- [2] Пономарев, Е. С. Обучение с подкреплением в задаче алгоритмической торговли / Е. С. Пономарев, И. В. Оселедец, А. С. Чихоцкий // Информационные процессы. – 2019. – Т. 19, № 2. – С. 122-131. – URL: <https://www.elibrary.ru/item.asp?id=38538082> (дата обращения: 11.05.2020). – Рез. англ.
- [3] Касьянов, Р. А. MIFID II: регулирование алгоритмической торговли и высокочастотной алгоритмической торговли / Р. А. Касьянов // Финансовое право. – 2018. – № 6. – С. 3-7. – URL: <https://www.elibrary.ru/item.asp?id=35133781> (дата обращения: 11.05.2020). – Рез. англ.
- [4] Красников, В. С. Алгоритм выставления стоп-ордеров при торговле длинными контрактами / В. С. Красников // Научные записки молодых исследователей. – 2018. – № 4. – С. 21-30. – URL: <https://www.elibrary.ru/item.asp?id=35588843> (дата обращения: 11.05.2020). – Рез. англ.
- [5] Шаипова, С. А. Высокочастотный трейдинг и его развитие на торговом рынке / С. А. Шаипова // Вестник Академии права и управления. – 2018. – № 3. – С. 146-149. – URL: <https://www.elibrary.ru/item.asp?id=36319246> (дата обращения: 11.05.2020). – Рез. англ.
- [6] Xie, M. Blockchain financial investment based on deep learning network algorithm / M. Xie, H. Li, Yu. Zhao. – DOI 10.1016/j.cam.2020.112723 // Journal of Computational



- and Applied Mathematics. – 2020. – Vol. 372. – Pp. 112723. – URL: <https://www.sciencedirect.com/science/article/pii/S0377042720300145> (дата обращения: 11.05.2020).
- [7] Borges, T. A. Ensemble of machine learning algorithms for cryptocurrency investment with different data resampling methods / T. A. Borges, R. F. Neves. – DOI 10.1016/j.asoc.2020.106187 // Applied Soft Computing. – 2020. – Vol. 90. – Pp. 106187. – URL: <https://www.sciencedirect.com/science/article/pii/S1568494620301277> (дата обращения: 11.05.2020).
- [8] Пальмов, С. В. Глубокое обучение: определение и отличительные особенности / С. В. Пальмов, Е. С. Артюшкина // Форум молодых ученых. – 2020. – № 3. – С. 311-316. – URL: <https://www.elibrary.ru/item.asp?id=42857736> (дата обращения: 11.05.2020). – Рез. англ.
- [9] Чистова, Е. В. Применение глубокого обучения к моделированию диалога на естественном языке / Е. В. Чистова, А. О. Шелманов, И. В. Смирнов. – DOI 10.14357/20790279190110 // Труды Института системного анализа Российской академии наук. – 2019. – Т. 69, № 1. – С. 105-115. – URL: <https://www.elibrary.ru/item.asp?id=37313176> (дата обращения: 11.05.2020). – Рез. англ.
- [10] Sineglazov, V. M. Deep learning fuzzy classifier / V. M. Sineglazov, R. S. Koniushenko. – DOI 10.18372/1990-5548.60.13813 // Electronics and Control Systems. – 2019. – Vol. 2, No. 60. – Pp. 35-42. – URL: [http://jrnл.nau.edu.ua/index.php/ESU/article/view/13813/0](http://jrnل.nau.edu.ua/index.php/ESU/article/view/13813/0) (дата обращения: 11.05.2020).
- [11] Потемкин, А. В. Обработка разнородной информации с помощью глубокого обучения нейронных сетей / А. В. Потемкин // Мягкие измерения и вычисления. – 2019. – № 9. – С. 44-48. – URL: <https://www.elibrary.ru/item.asp?id=41688926> (дата обращения: 11.05.2020). – Рез. англ.
- [12] Ефремцев, В. Г. Применение методов глубокого обучения для оценки степени коммерческой ценности визуальных объектов / В. Г. Ефремцев, Н. Г. Ефремцев, Е. П. Тетерин, П. Е. Тетерин, В. В. Гансовский. – DOI 10.18287/2412-6179-CO-515 // Компьютерная оптика. – 2020. – Т. 44, № 1. – С. 127-132. – URL: <https://www.elibrary.ru/item.asp?id=42709516> (дата обращения: 11.05.2020). – Рез. англ.
- [13] Liu, F. State representation modeling for deep reinforcement learning based recommendation / F. Liu, R. Tang, X. Li [et al.]. – DOI 10.1016/j.knosys.2020.106170 // Knowledge-Based Systems. – 2020. – Vol. 205. – Pp. 106170. – URL: <https://www.sciencedirect.com/science/article/pii/S095070512030407X> (дата обращения: 11.05.2020).
- [14] Pandey, V. Deep reinforcement learning algorithm for dynamic pricing of express lanes with multiple access locations / V. Pandey, E. Wang, D. Boyles. – DOI 10.1016/j.trc.2020.102715 // Transportation Research Part C: Emerging Technologies. – 2020. – Vol. 119. – Pp. 102715. – URL: <https://www.sciencedirect.com/science/article/pii/S0968090X20306306> (дата обращения: 11.05.2020).
- [15] Liu, F. Top-aware reinforcement learning based recommendation / F. Liu, R. Tang [et al.]. – DOI 10.1016/j.neucom.2020.07.057 // Neurocomputing. – 2020. – Vol. 417. – Pp. 255-269. – URL: <https://www.sciencedirect.com/science/article/pii/S0925231220311656> (дата обращения: 11.05.2020).
- [16] Park, H. An intelligent financial portfolio trading strategy using deep Q-learning / H. Park, M. K. Sim, D. Gu Choi. – DOI 10.1016/j.eswa.2020.113573 // Expert Systems with Applications. – 2020. – Vol. 158. – Pp. 113573. – URL: <https://doi.org/10.1016/j.eswa.2020.113573> (дата обращения: 11.05.2020).
- [17] Lu, R. A Dynamic pricing demand response algorithm for smart grid: Reinforcement learning approach / R. Lu, S. H. Hong, X. Zhang. – DOI 10.1016/j.apenergy.2018.03.072 // Applied Energy. – 2018. – Vol. 220. – Pp. 220-230. – URL: <https://www.sciencedirect.com/science/article/pii/S0306261918304112> (дата обращения: 11.05.2020).
- [18] Tymchenko, B. Segmentation of Cloud Organization Patterns from Satellite Images Using Deep Neural Networks / B. Tymchenko, Ph. Marchenko, D. Spodarets. – DOI 10.15276/hait01.2020.2 // Herald of Advanced Information Technology. – 2020. – Vol. 3, No. 1. – Pp. 352-361. – URL: <https://www.elibrary.ru/item.asp?id=43091153> (дата обращения: 11.05.2020).
- [19] Горшенин, А. К. Анализ конфигураций LSTM-сетей для построения среднесрочных векторных прогнозов / А. К. Горшенин, В. Ю. Кузьмин. – DOI 10.14357/19922264200102 // Информатика и ее применения. – 2020. – Т. 14, № 1. – С. 10-16. – URL: <https://www.elibrary.ru/item.asp?id=43146928> (дата обращения: 11.05.2020). – Рез. англ.
- [20] Chakole, J. B. A Q-learning agent for automated trading in equity stock markets / J. B. Chakole [et al.]. – DOI 10.1016/j.eswa.2020.113761 // Expert Systems with Applications. – 2021. – Vol. 163. – Pp. 113761. – URL: <https://www.sciencedirect.com/science/article/pii/S0957417420305856> (дата обращения: 11.05.2020).
- [21] Betancourt, C. Deep reinforcement learning for portfolio management of markets with a dynamic number of assets / C. Betancourt, W.-H. Chen. – DOI 10.1016/j.eswa.2020.114002 // Expert Systems with Applications. – 2021. – Vol. 164. – Pp. 114002. – URL: <https://www.sciencedirect.com/science/article/pii/S0957417420307776> (дата обращения: 11.05.2020).
- [22] Lei, K. Time-driven feature-aware jointly deep reinforcement learning for financial signal representation and algorithmic trading / K. Lei, B. Zhang, Y. Li, M. Yang, Y. Shen. – DOI 10.1016/j.eswa.2019.112872 // Expert Systems with Applications. – 2020. – Vol. 140. – Pp. 112872. – URL: <https://www.sciencedirect.com/science/article/pii/S0957417419305822> (дата обращения: 11.05.2020).
- [23] Nguyen, T. T. A multi-objective deep reinforcement learning framework / T. T. Nguyen [et al.]. – DOI 10.1016/j.engappai.2020.103915 // Engineering Applications of Artificial Intelligence. – 2020. – Vol. 96. – Pp. 103915. – URL: <https://www.sciencedirect.com/science/article/pii/S0952197620302475> (дата обращения: 11.05.2020).
- [24] Riesener, M. Applying Supervised and Reinforcement Learning to Design Product Portfolios in Accordance with Corporate Goals / M. Riesener, C. Dölle, C. Dierkes, M.-H. Jank. – DOI 10.1016/j.procir.2020.02.157 // Procedia CIRP. – 2020. – Vol. 91. – Pp. 127-133. – URL: <https://www.sciencedirect.com/science/article/pii/S2212827120307952>



(дата обращения: 11.05.2020).

- [25] Jeong, G. Improving financial trading decisions using deep Q-learning: Predicting the number of shares, action strategies, and transfer learning / G. Jeong, H. Y. Kim. – DOI 10.1016/j.eswa.2018.09.036 // *Expert Systems with Applications*. – 2019. – Vol. 117. – Pp. 125-138. – URL: <https://www.sciencedirect.com/science/article/pii/S0957417418306134> (дата обращения: 11.05.2020).
- [26] Jeong, G. Improving financial trading decisions using deep Q-learning: Predicting the number of shares, action strategies, and transfer learning / G. Jeong, H. Y. Kim. – DOI 10.1016/j.eswa.2018.09.036 // *Expert Systems with Applications*. – 2019. – Vol. 117. – Pp. 125-138. – URL: <https://www.sciencedirect.com/science/article/pii/S0957417418306134> (дата обращения: 11.05.2020).
- [27] Gao, W. Analysis on block chain financial transaction under artificial neural network of deep learning / W. Gao, Ch. Su. – DOI 10.1016/j.cam.2020.112991 // *Journal of Computational and Applied Mathematics*. – 2020. – Vol. 380. – Pp. 112991. – URL: <https://www.sciencedirect.com/science/article/pii/S037704272030282X> (дата обращения: 11.05.2020).

Поступила 11.05.2020; принята к публикации 23.07.2020;
опубликована онлайн 30.09.2020.

Об авторах:

Никитин Петр Владимирович, доцент департамента анализа данных и машинного обучения, ФГОБУ ВО «Финансовый университет при правительстве Российской Федерации» (125993, Россия, ГСП-3, г. Москва, Ленинградский пр., д. 49), кандидат педагогических наук, доцент, ORCID: <http://orcid.org/0000-0001-8866-5610>, pvnikitin@fa.ru

Горохова Римма Ивановна, доцент департамента анализа данных и машинного обучения, ФГОБУ ВО «Финансовый университет при правительстве Российской Федерации» (125993, Россия, ГСП-3, г. Москва, Ленинградский пр., д. 49), кандидат педагогических наук, доцент, ORCID: <http://orcid.org/0000-0001-7818-8013>, rigorokhova@fa.ru

Корчагин Сергей Алексеевич, доцент департамента анализа данных и машинного обучения, ФГОБУ ВО «Финансовый университет при правительстве Российской Федерации» (125993, Россия, ГСП-3, г. Москва, Ленинградский пр., д. 49), кандидат физико-математических наук, ORCID: <http://orcid.org/0000-0001-8042-4089>, sakorchagin@fa.ru

Красников Владимир Сергеевич, магистрант департамента анализа данных и машинного обучения, ФГОБУ ВО «Финансовый университет при правительстве Российской Федерации» (125993, Россия, ГСП-3, г. Москва, Ленинградский пр., д. 49), ORCID: <http://orcid.org/0000-0003-2201-5425>

Все авторы прочитали и одобрили окончательный вариант рукописи.

References

- [1] Batiuk B.V. Problems and Prospects of Algorithmic Trade in Financial Markets. *Entrepreneur's Guide*. 2020; 13(2):9-16. (In Russ., abstract in Eng.) DOI: <https://doi.org/10.24182/2073-9885-2020-13-2-9-16>
- [2] Ponomarev E.S., Oseledets E.V., Cichocki A.S. Algorithmic Trading with Reinforcement Learning. *Information Processes*. 2019; 19(2):122-131. Available at: <https://www.elibrary.ru/item.asp?id=38538082> (accessed 02.09.2020). (In Russ., abstract in Eng.)
- [3] Kasyanov R.A. MIFID II The Regulation of Algorithmic Trading and High-frequency Algorithmic Trading. *Financial Law*. 2018; (6):3-7. Available at: <https://www.elibrary.ru/item.asp?id=35133781> (accessed 02.09.2020). (In Russ., abstract in Eng.)
- [4] Krasnikov V.S. Algorithm for Placing Stop Orders when Trading Long Contracts. *Scientific notes of young scientists*. 2018; (4):21-30. Available at: <https://www.elibrary.ru/item.asp?id=35588843> (accessed 02.09.2020). (In Russ., abstract in Eng.)
- [5] Shaipova S.A. High-Frequency Trading and its Development in the Trade Market. *Bulletin of the Academy of Law and Management*. 2018; (3):146-149. Available at: <https://www.elibrary.ru/item.asp?id=36319246> (accessed 02.09.2020). (In Russ., abstract in Eng.)
- [6] Xie M., Li H., Zhao Yu. Blockchain financial investment based on deep learning network algorithm. *Journal of Computational and Applied Mathematics*. 2020; 372:112723. (In Eng.) DOI: <https://doi.org/10.1016/j.cam.2020.112723>
- [7] Borges T.A., Neves R.F. Ensemble of machine learning algorithms for cryptocurrency investment with different data resampling methods. *Applied Soft Computing*. 2020; 90:106187. (In Eng.) DOI: <https://doi.org/10.1016/j.asoc.2020.106187>
- [8] Palmov S.V., Artyushkina E.S. Deep Learning: Definition and Distinctive Features. *Forum molodyh uchenyh*. 2020; (3):311-316. Available at: <https://www.elibrary.ru/item.asp?id=42857736> (accessed 02.09.2020). (In Russ., abstract in Eng.)
- [9] Chistova E.V., Shelmanov A.O., Smirnov I.V. Natural Language Dialogue Modelling with Deep Learning. *Proceedings of the Institute for Systems Analysis Russian Academy of Sciences*. 2019; 69(1):105-115. (In Russ., abstract in Eng.) DOI: <https://doi.org/10.14357/20790279190110>
- [10] Sineglazov V.M., Koniushenko R.S. Deep learning fuzzy classifier. *Electronics and Control Systems*. 2019; 2(60):35-42. (In Eng.) DOI: <https://doi.org/10.18372/1990-5548.60.13813>
- [11] Potemkin A.V. Processing Different Information Using the Deep Learning of Neural Networks. *Soft Measurements and Computing*. 2019; (9):44-48. Available at: <https://www.elibrary.ru/item.asp?id=41688926> (accessed 02.09.2020). (In Russ., abstract in Eng.)
- [12] Efremtsev V.G., Ejtemtsev N.G., Teterin E.P., Teterin P.E., Gantsovsky V.V. Deep Learning Application for Box-office Evaluation of Images. *Computer Optics*. 2020; 44(1):127-132. (In Russ., abstract in Eng.) DOI: <https://doi.org/10.18287/2412-6179-CO-515>
- [13] Liu F, et al. State representation modeling for deep reinforcement learning based recommendation. *Knowl-*



- edge-Based Systems*. 2020; 205:106170. (In Eng.) DOI: <https://doi.org/10.1016/j.knosys.2020.106170>
- [14] Pandey V., Wang E., Boyles S.D. Deep reinforcement learning algorithm for dynamic pricing of express lanes with multiple access locations. *Transportation Research Part C: Emerging Technologies*. 2020; 119:102715. (In Eng.) DOI: <https://doi.org/10.1016/j.trc.2020.102715>
- [15] Liu F., Tang R., et al. Top-aware reinforcement learning based recommendation. *Neurocomputing*. 2020; 417:255-269. (In Eng.) DOI: <https://doi.org/10.1016/j.neucom.2020.07.057>
- [16] Park H., Sim M.K., Choi D.G. An intelligent financial portfolio trading strategy using deep Q-learning. *Expert Systems with Applications*. 2020; 158:113573. (In Eng.) DOI: <https://doi.org/10.1016/j.eswa.2020.113573>
- [17] Lu R., Hong S.H., Zhang X. A Dynamic pricing demand response algorithm for smart grid: Reinforcement learning approach. *Applied Energy*. 2018; 220:220-230. (In Eng.) DOI: <https://doi.org/10.1016/j.apenergy.2018.03.072>
- [18] Tymchenko B., Marchenko Ph., Spodarets D. Segmentation of Cloud Organization Patterns from Satellite Images Using Deep Neural Networks. *Herald of Advanced Information Technology*. 2020; 3(1):352-361. (In Eng.) DOI: <https://doi.org/10.15276/hait01.2020.2>
- [19] Gorshenin A.K., Kuzmin V.Yu. Analysis of Configurations of LSTM Networks for Medium-Term Vector Forecasting. *Informatics and Applications*. 2020; 14(1):10-16. (In Eng.) DOI: <https://doi.org/10.14357/19922264200102>
- [20] Chakole J. B., et al. A Q-learning agent for automated trading in equity stock markets. *Expert Systems with Applications*. 2021; 163:113761. (In Eng.) DOI: <https://doi.org/10.1016/j.eswa.2020.113761>
- [21] Betancourt C., Chen W.-H. Deep reinforcement learning for portfolio management of markets with a dynamic number of assets. *Expert Systems with Applications*. 2021; 164:114002. (In Eng.) DOI: <https://doi.org/10.1016/j.eswa.2020.114002>
- [22] Lei K., Zhang B., Li Y., Yang M., Shen Y. Time-driven feature-aware jointly deep reinforcement learning for financial signal representation and algorithmic trading. *Expert Systems with Applications*. 2020; 140:112872. (In Eng.) DOI: <https://doi.org/10.1016/j.eswa.2019.112872>
- [23] Nguyen T.T., et al. A multi-objective deep reinforcement learning framework. *Engineering Applications of Artificial Intelligence*. 2020; 96:103915. (In Eng.) DOI: <https://doi.org/10.1016/j.engappai.2020.103915>
- [24] Riesener M., Dölle C., Dierkes C., Jank M.-H. Applying Supervised and Reinforcement Learning to Design Product Portfolios in Accordance with Corporate Goals. *Procedia CIRP*. 2020; 91:127-133. (In Eng.) DOI: <https://doi.org/10.1016/j.procir.2020.02.157>
- [25] Jeong G., Kim H.Y. Improving financial trading decisions using deep Q-learning: Predicting the number of shares, action strategies, and transfer learning. *Expert Systems with Applications*. 2019; 117:125-138. (In Eng.) DOI: <https://doi.org/10.1016/j.eswa.2018.09.036>
- [26] Jeong G., Kim H.Y. Improving financial trading decisions using deep Q-learning: Predicting the number of shares, action strategies, and transfer learning. *Expert Systems with Applications*. 2019; 117:125-138. (In Eng.) DOI: <https://doi.org/10.1016/j.eswa.2018.09.036>
- [27] Gao W., Su Ch. Analysis on block chain financial transaction under artificial neural network of deep learning. *Journal of Computational and Applied Mathematics*. 2020; 380:112991. (In Eng.) DOI: <https://doi.org/10.1016/j.cam.2020.112991>

Submitted 11.05.2020; revised 23.07.2020;
published online 30.09.2020.

About the authors:

Petr V. Nikitin, Associate Professor of the Department of Data Analysis and Machine Learning, Financial University under the Government of the Russian Federation (49 Leningradskiy Ave., Moscow 125993, GSP-3, Russia), Ph.D. (Pedagogy), Associate Professor, ORCID: <http://orcid.org/0000-0001-8866-5610>, pvnikitin@fa.ru

Rimma I. Gorokhova, Associate Professor of the Department of Data Analysis and Machine Learning, Financial University under the Government of the Russian Federation (49 Leningradskiy Ave., Moscow 125993, GSP-3, Russia), Ph.D. (Pedagogy), Associate Professor, ORCID: <http://orcid.org/0000-0001-7818-8013>, rigorokhova@fa.ru

Sergey A. Korchagin, Associate Professor of the Department of Data Analysis and Machine Learning, Financial University under the Government of the Russian Federation (49 Leningradskiy Ave., Moscow 125993, GSP-3, Russia), Ph.D. (Phys.-Math.), ORCID: <http://orcid.org/0000-0001-8042-4089>, sakorchagin@fa.ru

Vladimir S. Krasnikov, Undergraduate Student of the Department of Data Analysis and Machine Learning, Financial University under the Government of the Russian Federation (49 Leningradskiy Ave., Moscow 125993, GSP-3, Russia), ORCID: <http://orcid.org/0000-0003-2201-5425>

All authors have read and approved the final manuscript.

