УДК 336.763.4

ПОЛОВНИКОВ ДМИТРИЙ СЕРГЕЕВИЧ
*студент 4-го курса факультета прикладной математики и механики,
Пермский Национальный Исследовательский Политехнический
Университет, г. Пермь, Россия
e-mail: polovnikov.161@mail.ru*

ВЛАДИМИРОВА ДАРЬЯ БОРИСОВНА
*к.физ.-мат.н., доцент, Пермский Национальный Исследовательский
Политехнический Университет г. Пермь, Россия,
e-mail: da0807@mail.ru*

ПЕРВАДЧУК ВЛАДИМИР ПАВЛОВИЧ
*д.т.н., профессор, заведующий кафедрой Прикладной математики,
Пермский Национальный Исследовательский Политехнический
Университет, г. Пермь, Россия
e-mail: pervadchuk@mail.ru*

DOI:10.26726/1812-7096-2021-6-155-166

О ПРИМЕНИМОСТИ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ НА РЫНКЕ ДЕРИВАТИВОВ

Аннотация. *В данной статье рассматривается возможность объединения научных методов анализа финансового рынка вторичных ценных бумаг двух подходов: стохастического и принципа стохастического детерминизма. Данные подходы были использованы для исследования возможности принятия оптимального решения о выборе покупки опциона «call» или «put» европейского типа для получения наилучшего финансового результата. В качестве механизма принятия решения были применены классические методы машинного обучения, позволяющие решить задачу классификации. В результате проведенной работы было показано, что применение алгоритмов машинного обучения способно значительно повысить ожидаемый финансовый результат от приобретения опционных контрактов, причем он достигается за счет объединения 2 рассматриваемых в работе подходов. Полученные в статье выводы могут быть использованы региональными финансовыми организациями: банками, инвестиционными фондами — для достижения большей эффективности в их экономической деятельности.*

Ключевые слова: *рынок ценных бумаг, стохастический подход, принцип стохастического детерминизма, опционы, машинное обучение, инвестиции.*

POLOVNIKOV DMITRY SERGEEVICH
*4th year student of the Faculty of Applied Mathematics and Mechanics,
Perm National Research Polytechnic University, Perm, Russia,
e-mail: polovnikov.161@mail.ru*

VLADIMIROVA DARYA BORISOVNA
*Ph. D. in Physics and Mathematics Associate Professor,
Perm National Research Polytechnic University, Perm, Russia,
e-mail: da0807@mail.ru*

PERVADCHUK VLADIMIR PAVLOVICH
*Dr.Sc of Technics, Professor, Head of the Department of Applied Mathematics,
Perm National Research Polytechnic University, Perm, Russia,
e-mail: pervadchuk@mail.ru*

ABOUT THE APPLICABILITY OF MACHINE LEARNING METHODS TRAINING IN THE DERIVATIVES MARKET

Abstract. *This article considers the possibility of combining scientific methods of analyzing the financial market of secondary securities of two approaches: stochastic and the principle of stochastic determinism. These approaches were used to study the possibility of making an optimal decision on the choice of buying a European-type "call" or "put" option to obtain the best financial result. As a decision-making mechanism, classical machine learning methods were used to solve the classification problem. As a result of the work carried out, it was shown that the use of machine learning algorithms can significantly increase the expected financial result from the purchase of option contracts, and it is achieved by combining the 2 approaches considered in the work. The conclusions obtained in the article can be used by regional financial organizations: banks, investment funds-to achieve greater efficiency in their economic activities.*
Keywords: *securities market, stochastic approach, the principle of stochastic determinism, options, machine learning, investment.*

Введение. Вопросы выявления и обоснования закономерностей в поведении финансовых рынков изучаются достаточно давно. Согласно теории эффективного рынка [1] Юджина Фама, зародившейся еще в XX в., логарифмы ценовых величин, описывающих колебания валютных пар, процентных ставок и прочих активов, ведут себя как случайное блуждание:

$$S_t = S_0 e^{H_n}, \tag{1}$$

где $S = S(t) = S_t$ — стоимость актива в момент времени $t, t \in [0,T]$; S_0 – цена актива в начальный момент времени; H_n — сумма независимых случайных величин $h_1, \dots, h_n$.

В рамках этой теории в 1965 г. П. Самуэльсон предложил модель геометрического броуновского движения для цен активов на фондовом рынке:

$$S_t = S_0 e^{\left(\mu_t - \frac{\sigma_t^2}{2}\right)t + \sigma W_t}, t \geq 0, \tag{2}$$

где $\mu = \mu(t) = \mu_t$ — детерминированный тренд; $\sigma = \sigma(t) = \sigma_t$ — стохастическая волатильность; W_t — стандартное броуновское движение.

При более детальном изучении [2] можно видеть, что в краткосрочном периоде изменение цены актива определяется исключительно волатильностью σ_t и броуновским движением W_t, т. е. будущая цена актива $S_T, T > t$ является случайной величиной и не может быть спрогнозирована. Модель П. Самуэльсона легла в основу модели Блэка-Шоулза-Мертона, в рамках которой была выведена в 1973 г. формула для рациональной стоимости стандартного колл-опциона европейского типа:

$$C_T = S_0 \Phi\left(\frac{ln\frac{S_0}{K} + T\left(r + \frac{\sigma^2}{2}\right)}{\sigma\sqrt{T}}\right) - Ke^{-rT}\Phi\left(\frac{ln\frac{S_0}{K} + T\left(r - \frac{\sigma^2}{2}\right)}{\sigma\sqrt{T}}\right), \tag{3}$$

где *K* – цена страйк опциона; *T* – время экспирации опциона; *r* – безрисковая процентная ставка; *Φ*(.) – кумулятивная функция стандартного нормального распределения.

Можно видеть, что справедливая стоимость опциона рассчитывается, исходя из вероятностных характеристик броуновского движения, и не предусматривает детерминированного движения цены (тренда), но учитывает изменение цены с учетом безрисковой процентной ставки.

В теории фрактального рынка предполагается, что будущие значения цен зависят от их прошлых изменений, т. е. процесс ценообразования на рынках глобально детерминирован. Фрактальный анализ, получивший развитие благодаря работе Б. Мандельброта и Р. Хадсона, заимствует свойства фракталов для получения прогнозов [3]. В частности, используется фрактальная размерность множества, которая связана с экспонентой Херста, предложенной Г. Херстом в качестве статистики для выявления случайных и неслучайных процессов, трендов и циклов. Данная величина входит в безразмерное отношение, называемое нормированным размахом:

$$E\left[R_T/\sqrt{Var[S_T]}\right] = cT^H, \tag{4}$$

где c – некоторая константа; T – величина промежутка времени; $R_T = \max_{t\in[0,T]} S_t - \min_{t\in[0,T]} S_t$ — размах цены на промежутке T; $Var[S_T]$ — дисперсия цены актива на данном промежутке времени; H – показатель Херста. Подробно методология R/S-анализа (здесь S – стандартное отклонение) была рассмотрена в работе [7].

Величина $H \in [0,1]$, так что:

$H < 0.5$ — характеризует антиперсистентный ряд. Для него характерен возврат к среднему и сильные колебания, что говорит о смене тренда.

$H > 0.5$ — характеризует персистентный ряд. Для него характерен устойчивый тренд, который продолжится в будущем.

$H = 0.5$ — характеризует белый шум, случайное блуждание.

Предметом исследования данной работы является выявление потенциальной возможности применения характеристики Херста для идентификации устойчивого тренда и приобретения соответствующего колл- или пут-опциона с учетом прогноза направления движения цены базового актива.

Основная часть.

Постановка и алгоритм решения задачи. Оставаясь в рамках теории эффективного рынка, авторами осуществляется предположение, что вся информация уже отражена в цене так, что нет необходимости анализировать информацию, поступающую извне, и использовать только исторические данные движения цены. В качестве исследуемых характеристик рассматривается волатильность $\sigma_t\left(=\sqrt{Var[S_t]}\right)$, тренд $\mu_t\left(=\ln\frac{S_t}{S_0}\right)$ и экспонента Херста $H_t\left(=\frac{\ln\frac{R_t}{\sigma_t}}{\ln\frac{t}{2}}\right)$.

Учитывая, что наибольший интерес представляет только направление движения цены (ее положение относительно текущей цены или цены исполнения опциона), авторами формулируется задача классификации: Дано множество объектов $X \in \mathrm{X}$, описываемых числовыми признаками $x = (\mu, \sigma, H) \in \mathbb{R}^3$, которые можно разделить на два непересекающихся класса $Y = \{-1(put), +1(call)\}$. Необходимо найти модель алгоритма принятия решения (разделения на классы):

$$a(x) = sign\big(f(x)\big), \tag{5}$$

где $f(x): X \to \mathbb{R}$.

Фактически цель заключается в получении алгоритма, который для подаваемых на вход параметров исследуемого актива будет выдавать решение о приобретении call- или put-опциона.

В качестве такого алгоритма мы рассмотрим несколько классических моделей машинного обучения: *Support Vector Machine (SVM), Decision Tree, Random Forest, Light Gradient Boosting Machine (LightGBM), k Nearest Neighbors (kNN), Logistic Regression* [4].

Учитывая стохастическую природу волатильности, принято решение об исследовании совокупности движений цен активов, полученной на основе ретроспективных данных, предполагая, что были совершены сделки по приобретению опционов так, что известны их финансовые результаты. В качестве параметров для оценки качества моделей будем использовать медианное и среднее значения доходностей от приобретенных опционов.

Для исследования были использованы временные ряды – котировки всех торгуемых на бирже NYSE активов (акций, индексов, фондов) за период от 01.06.2019 до 15.05.2021 [5]. Каждый временной ряд был разбит и представлен набором временных рядов длиной в 3 месяца со сдвигом на 1 месяц вперед. В качестве уровня временно́го ряда взяты цены закрытия часовых торговых сессий. С учетом выходных и праздничных дней, а также отсутствующих данных, были отобраны для анализа ряды, для которых число наблюдений не меньше 250. Среднее число наблюдений 440, максимальное – 462. Всего исследовано 24186 временных рядов (табл. 1).

По каждому временно́му ряду длиной в 3 месяца рассчитаны: волатильность, тренд, экспонента Херста. На основе этих данных был рассчитан результат гипотетического приобретения call- или put-опциона сроком на 3 месяца вперед (т.е. использовались данные по активу за 6 месяцев: 3 – для оценки параметров, 6-й – для фиксации конечной цены актива) и цены исполнения, равной текущей цене актива. В качестве результата сделки использована величина:

$$B_C = \frac{max\{S_T - K, 0\} - C_T}{C_T}, \tag{6}$$

$$B_P = \frac{max\{K - S_T, 0\} - P_T}{P_T}, \tag{7}$$

где C_T, P_T — стоимость опционов call и put, рассчитанных по модели BSM для вычисленных параметров волатильности (использованная процентная ставка $r = 0.0025$ — ставка ФРС США); B_C, B_P — доходность приобретения опционов call и put соответственно. Очевидно, что $B \in [-1, M_T]$, где -1 означает, что цена исполнения опциона хуже рыночной и опцион не исполняется, что эквивалентно потере всей стоимости опциона. M_T – некоторая максимально возможная доходность опциона для промежутка времени T.

В качестве базового результата от модельных сделок рассмотрим результат, полученный от сделок, где выбор типа опциона определялся следующим образом:

Показатель	Тренд	Волатильность	Пок. Херста	Фин. результат
Количество наблюдений	24186	24186	24186	24186
Среднее арифметическое	0.070078	0.335072	0.577282	0.324151
Стандартное отклонение	0.60687	0.205936	0.056909	2.68133
Минимальное значение	-4.27663	0.017708	0.273775	-1
Максимальное значение	4.657632	0.999369	0.827781	56.04935

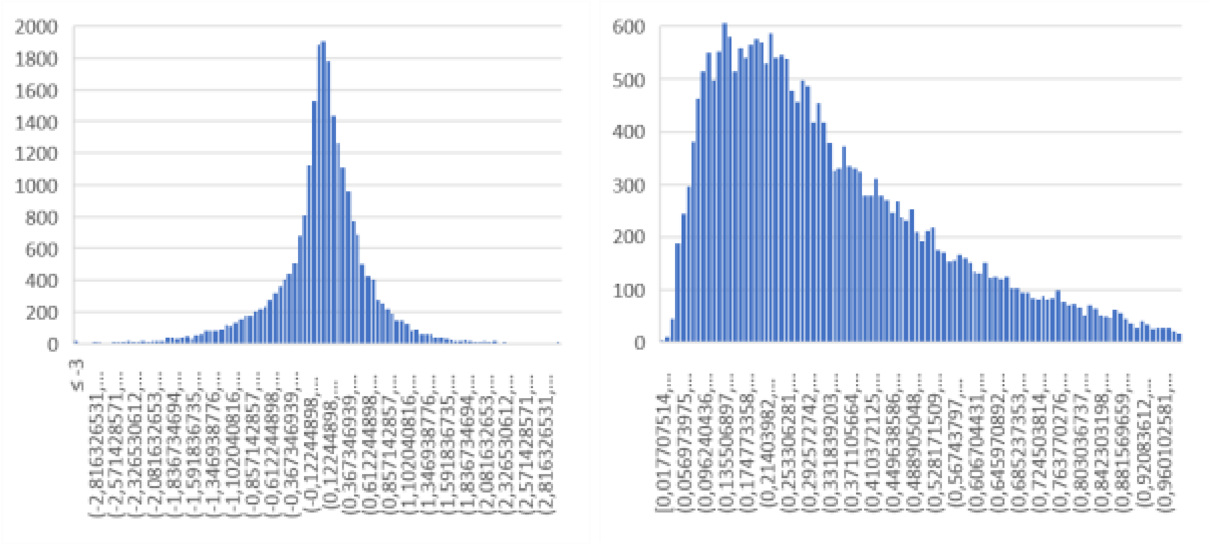

Рис. 1. *Гистограммы частот тренда (слева) и волатильности (справа)*

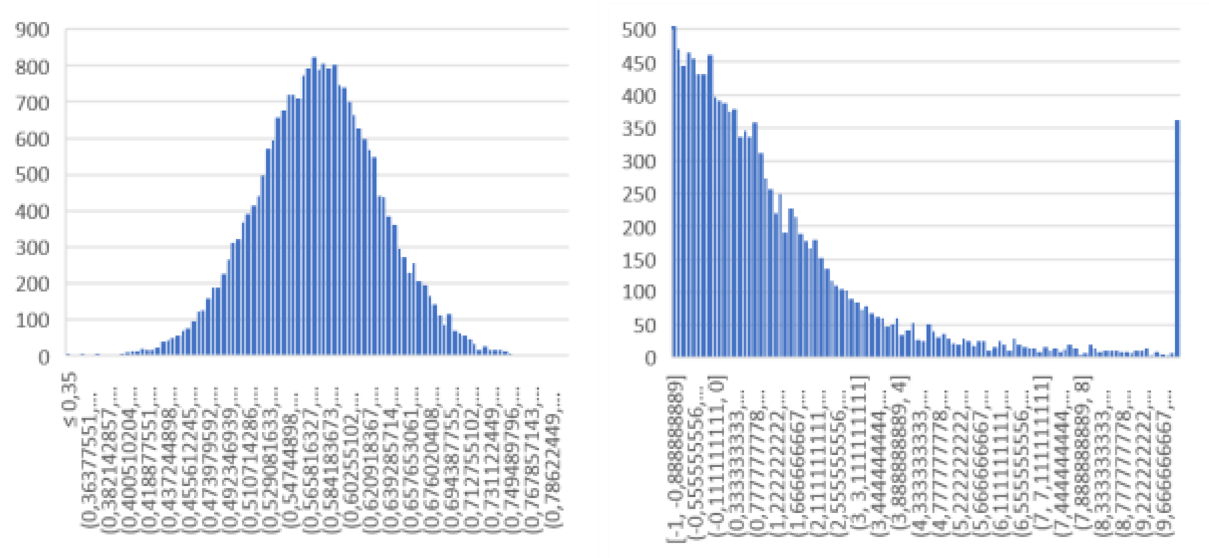

Рис. 2. *Гистограммы частот экспоненты Херста (слева) и доходности сделок (справа)*

Обратим внимание, что все три рассматриваемые характеристики подчиняются разным законам распределения: экспонента Херста – нормальное, волатильность – логнормальное, тренд, который фактически является доходностью актива на рассматриваемом промежутке времени, подчиняется усеченному распределению Леви. Кроме того, отметим тот факт, что число опционных контрактов для покупки, выбранных на основе предложенного алгоритма (8), обеспечивающих доходность выше -1, составило 12210 или 50,48% от числа всех контрактов, что эквивалентно случайному выбору типа контракта (подбрасывание монеты).

С учетом различия законов распределения данных имеет смысл провести преобразования над исходными данными и сравнить результаты. Отметим, что нормирование волатильности визуально делает данные более различимыми (рис. 3).

Мы не будем останавливаться на алгоритмах машинного обучения, подробно описанных в работах [4, 6]. Алгоритмы обучались на разных выборках, случайно отобранных данных из всего массива, с различными настройками гиперпараметров так, чтобы получить в среднем наилучшую точность – долю угаданных контрактов опционов, обеспечивающих доходность выше -1. Далее алгоритмы с оптимальными настройками обучались с помощью кросс-валидации – по полученным результатам рассчитывались статистические данные.

Ниже представлены графики обучения некоторых алгоритмов: с изменением гиперпара-

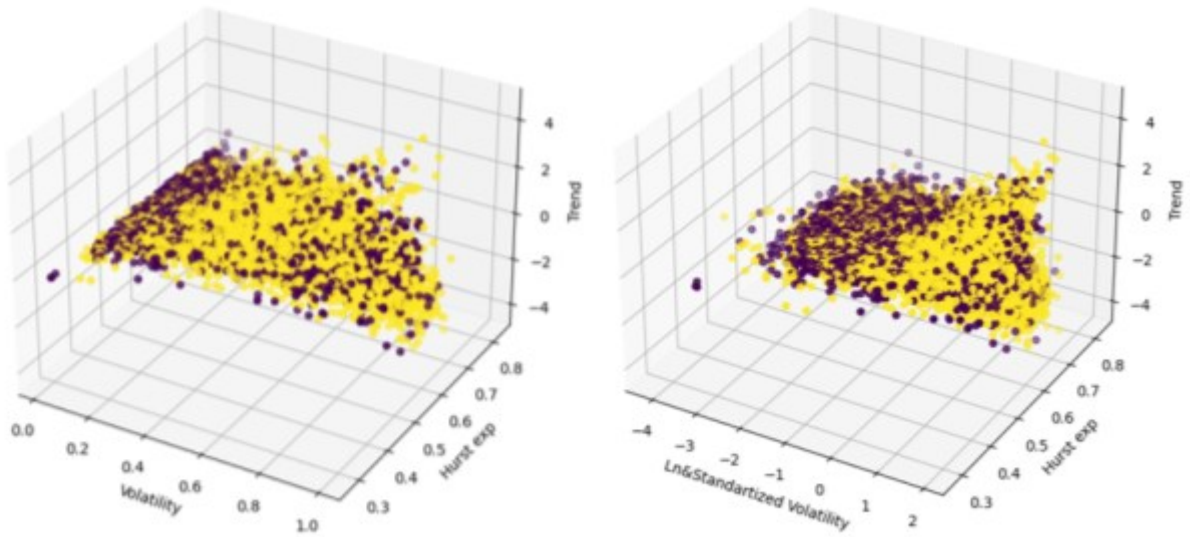

Рис. 3. *Визуализация данных. Положительные результаты от приобретения опционов в зависимости от тренда, волатильности, экспоненты*

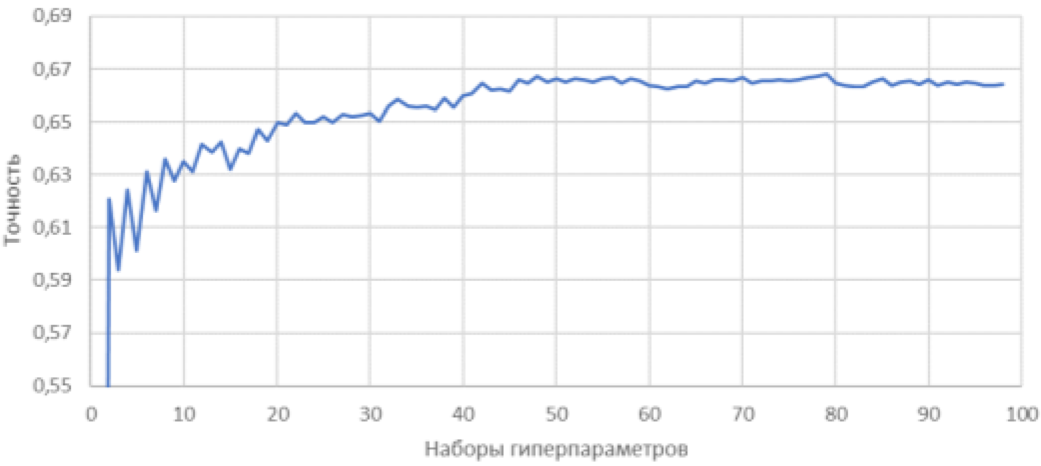

Рис. 4. *Точность алгоритма k Nearest Neighbors для разных наборов гиперпараметров*

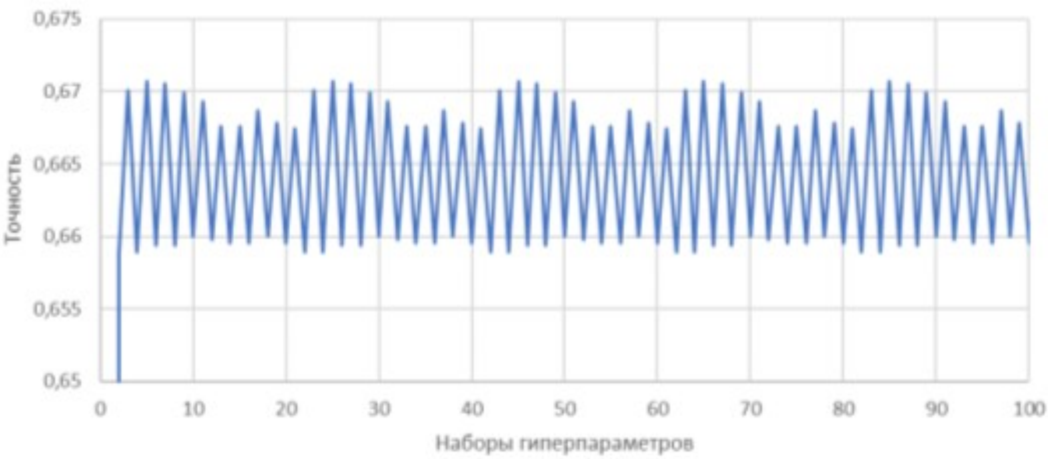

Рис. 5. *Точность алгоритма LGBM для разных наборов гиперпараметров*

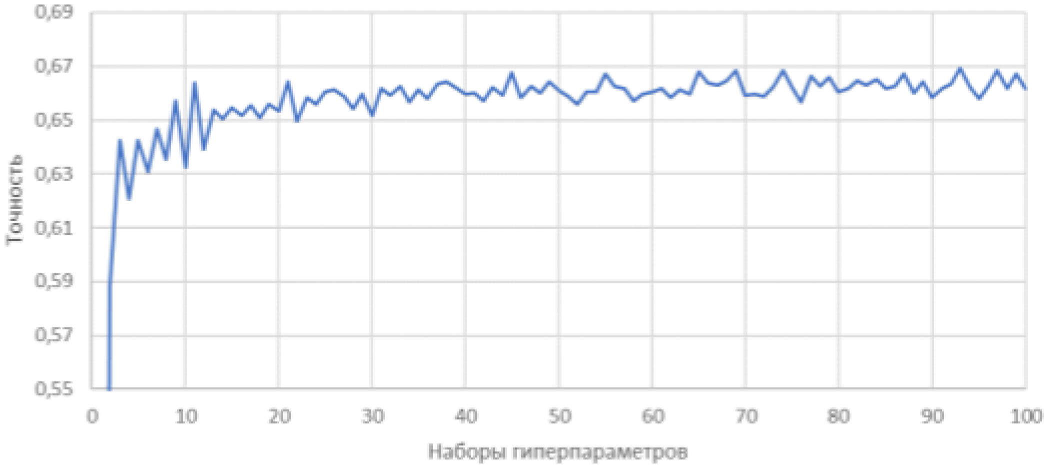

Рис. 6. *Точность алгоритма Random Forest для разных наборов гиперпараметров*

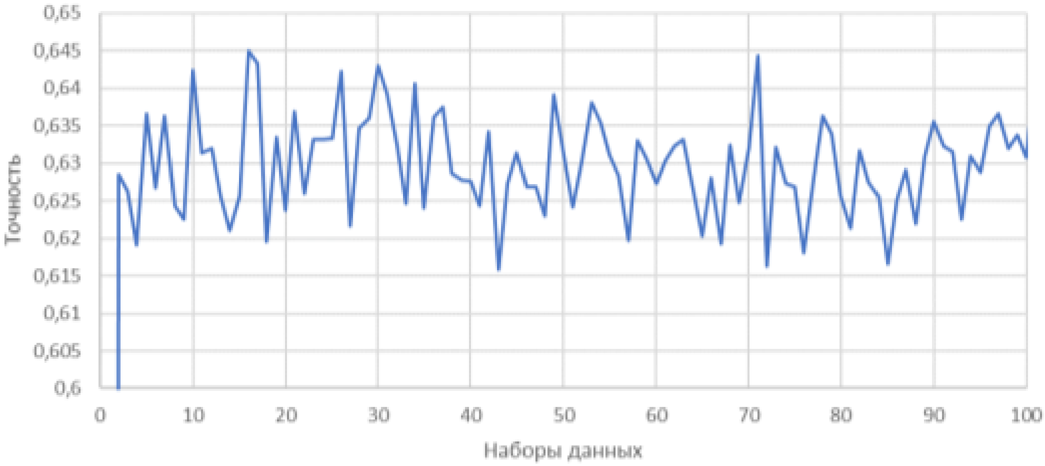

Рис. 7. *Колебания точности алгоритма на разных выборках*

Кросс-валидация для алгоритмов машинного обучения не имеет практической пользы, поскольку колебания точности уже обученного алгоритма на разных выборках представляют собой белый шум. Однако стоит учитывать, что в данном случае полученная информация может использоваться в качестве статистических данных при анализе эффективности алгоритмов на реальных данных.

Результат от модельных сделок приобретения опционных контрактов представлен ниже (табл. 3–10). Можно видеть, что в среднем опционные контракты обеспечивают доходность на уровне 32,4%, обеспечиваемую единичными сделками с многократной доходностью. Медианное значение лучше описывает реальную доходность: половина всех контрактов обеспечивает почти полную потерю всей стоимости опциона (табл. 2).

Средние характеристики медианной и средней ожидаемой доходности рассчитывались на 1000 выборках объемом 120 контрактов, отобранных случайным образом из всего набора данных (24186).

Таблица 2

Результат базового алгоритма принятия решений	
Медиана прибыли	Среднее значение
-0.97022	0.324151

Сначала будут представлены результаты для алгоритмов, обученных только с использованием параметра Херста, затем тренда и волатильности и, наконец, всех 3 параметров, объединяющих стохастический подход и стохастический детерминизм.

Таблица 3

Среднее и медианное значения доходностей сделок, обеспеченных методами машинного

Параметры	Алгоритм Random Forest		Алгоритм k-ближайших соседей		Алгоритм LightGBM	
	Медиана	Среднее	Медиана	Среднее	Медиана	Среднее
Среднее	-0.31948	0.786555	-0.28994	0.324906	-0.3271	0.333046
Ст. откл.	0.225177	0.294293	0.212556	0.165048	0.210426	0.172799
Мин. знач.	-0.98626	-0.05699	-0.96601	-0.18841	-1	-0.17334
Макс. знач.	0.460259	1.951127	0.431084	0.997774	0.407213	1.026552
Параметры	Алгоритм логистической регрессии		Алгоритм опорных векторов		Алгоритм решающих деревьев	
	Медиана	Среднее	Медиана	Среднее	Медиана	Среднее
Среднее	-0.2685	0.296264	-0.28994	0.324906	-0.41789	0.39229
Ст. откл.	0.213282	0.152295	0.212556	0.165048	0.215032	0.200492
Мин. знач.	-0.98724	-0.1919	-0.96601	-0.18841	-1	-0.14979
Макс. знач.	0.431084	0.934698	0.431084	0.997774	0.253733	1.178496

Таблица 4

Среднее и медианное значения доходностей сделок, обеспеченных методами машинного обучения для стандартизованного набора

Параметры	Алгоритм Random Forest		Алгоритм k-ближайших соседей		Алгоритм LightGBM	
	Медиана	Среднее	Медиана	Медиана	Среднее	Медиана
Среднее	-0.31738	0.789064	-0.28994	0.324906	-0.29457	0.313708
Ст. откл.	0.226992	0.297029	0.212556	0.165048	0.210367	0.161554
Мин. знач.	-0.98626	-0.05699	-0.96601	-0.18841	-1	-0.13543
Макс. знач.	0.460259	1.928083	0.431084	0.997774	0.385941	0.998703
Параметры	Алгоритм логистической регрессии		Алгоритм опорных векторов		Алгоритм решающих деревьев	
	Медиана	Среднее	Медиана	Среднее	Медиана	Среднее
Среднее	-0.2685	0.296264	-0.2685	0.296264	-0.41789	0.39229
Ст. откл.	0.213282	0.152295	0.213282	0.152295	0.215032	0.200492
Мин. знач.	-0.98724	-0.1919	-0.98724	-0.1919	-1	-0.14979
Макс. знач.	0.431084	0.934698	0.431084	0.934698	0.253733	1.178496

Как можно видеть по табл. 3, 4, обученные алгоритмы позволяют немного увеличить точность прогнозов по сравнению с базовым алгоритмом принятия решения. Нормализация показателя Херста позволяет несущественно увеличить качество исходной модели.

Далее будет рассмотрен стохастический подход – в качестве параметров модели будут использованы параметры тренд и волатильность (табл. 5–7).

Таблица 5

Среднее и медианное значения доходностей сделок, обеспеченных методами машинного

Параметры	Алгоритм Random Forest		Алгоритм k-ближайших соседей		Алгоритм LightGBM	
	Медиана	Среднее	Медиана	Среднее	Медиана	Среднее
Среднее	-0.07351	1.000662	-0.1805	0.813079	-0.15777	0.844134
Ст. откл.	0.217663	0.311572	0.223211	0.283101	0.224978	0.29483
Мин. знач.	-1	0.067788	-0.90699	0.093017	-0.96971	0.054618
Макс. знач.	0.824688	2.109471	0.585742	1.959512	0.585742	2.2079
Параметры	Алгоритм логистической регрессии		Алгоритм опорных векторов		Алгоритм решающих деревьев	
	Медиана	Среднее	Медиана	Среднее	Медиана	Среднее
Среднее	-0.2685	0.296264	-0.2685	0.296264	-0.25169	0.864522
Ст. откл.	0.213282	0.152295	0.213282	0.152295	0.229728	0.3064
Мин. знач.	-0.98724	-0.1919	-0.98724	-0.1919	-0.99461	0.030966
Макс. знач.	0.431084	0.934698	0.431084	0.934698	0.478668	2.036569

Таблица 6

Среднее и медианное значения доходностей сделок, обеспеченных методами машинного

Параметры	Алгоритм Random Forest		Алгоритм k-ближайших соседей		Алгоритм LightGBM	
	Медиана	Среднее	Медиана	Среднее	Медиана	Среднее
Среднее	-0.06708	0.999922	-0.19707	0.801048	-0.15777	0.844134
Ст. откл.	0.216968	0.309485	0.222787	0.282264	0.224978	0.29483
Мин. знач.	-0.81561	0.031154	-0.97043	0.052331	-0.96971	0.054618
Макс. знач.	0.636098	2.201314	0.568152	1.989719	0.585742	2.2079
Параметры	Алгоритм логистической регрессии		Алгоритм опорных векторов		Алгоритм решающих деревьев	
	Медиана	Среднее	Медиана	Среднее	Медиана	Среднее
Среднее	-0.24457	0.604982	-0.2685	0.296264	-0.25169	0.864522
Ст. откл.	0.220763	0.256406	0.213282	0.152295	0.229728	0.3064
Мин. знач.	-1	-0.03135	-0.98724	-0.1919	-0.99461	0.030966
Макс. знач.	0.493257	1.776516	0.431084	0.934698	0.478668	2.036569

Таблица 7

Среднее и медианное значения доходностей сделок, обеспеченных методами машинного

Параметры	Алгоритм Random Forest		Алгоритм k-ближайших соседей		Алгоритм LightGBM	
	Медиана	Среднее	Медиана	Среднее	Медиана	Среднее
Среднее	-0.01057	1.056867	-0.17956	0.832947	-0.14589	0.910221
Ст. откл.	0.217807	0.318432	0.226319	0.290291	0.220484	0.298737
Мин. знач.	-0.74012	0.165842	-0.92174	0.111021	-0.74957	0.165742
Макс. знач.	0.800498	2.223631	0.516735	1.888386	0.537282	1.88513
Параметры	Алгоритм логистической регрессии		Алгоритм опорных векторов		Алгоритм решающих деревьев	
	Медиана	Среднее	Медиана	Среднее	Медиана	Среднее
Среднее	-0.24148	0.604642	-0.26796	0.296106	-0.24838	0.786773
Ст. откл.	0.220923	0.261719	0.212658	0.152739	0.221935	0.287635
Мин. знач.	-0.92766	-0.03135	-0.98724	-0.1919	-1	0.020537
Макс. знач.	0.493257	1.776516	0.431084	0.934698	0.446037	2.133309

Использование в качестве параметров моделей тренда и волатильности позволяет уже в значительной мере повысить доходность опционных контрактов. Наилучший результат дает модель Random Forest для стандартизованных данных. Использование этого алгоритма позволяет увеличить медианное значение от -0,97 до -0,01, а среднее – от 0,324 до 1,057.

Далее представлены результаты для алгоритмов, обученных на всех 3 параметрах: тренде, волатильности и показателе Херста (табл. 8–10).

Таблица 8

Среднее и медианное значения доходностей сделок, обеспеченных методами машинного обучения для исходного набора данных с использованием всех 3

Параметры	Алгоритм Random Forest		Алгоритм k-ближайших соседей		Алгоритм LightGBM	
	Медиана	Среднее	Медиана	Среднее	Медиана	Среднее
Среднее	-0.00606	1.067242	-0.16116	0.822927	-0.14106	0.883382
Ст. откл.	0.214367	0.317475	0.223358	0.283575	0.218128	0.292947
Мин. знач.	-0.72615	0.080676	-0.8617	0.022561	-0.78721	0.104616
Макс. знач.	0.653166	2.267343	0.513287	2.01327	0.585742	1.959421
Параметры	Алгоритм логистической регрессии		Алгоритм опорных векторов		Алгоритм решающих деревьев	
	Медиана	Среднее	Медиана	Среднее	Медиана	Среднее
Среднее	-0.26593	0.29666	-0.26692	0.296447	-0.24199	0.791091
Ст. откл.	0.211157	0.154872	0.21047	0.153386	0.221389	0.28822
Мин. знач.	-0.94089	-0.1919	-0.94089	-0.1919	-1	0.020537
Макс. знач.	0.431084	0.934698	0.431084	0.934698	0.446037	2.133309

Таблица 9

Среднее и медианное значения доходностей сделок, обеспеченных методами машинного обучения для набора данных с трендом, линеаризованной

Параметры	Алгоритм Random Forest		Алгоритм k-ближайших соседей		Алгоритм LightGBM	
	Медиана	Среднее	Медиана	Среднее	Медиана	Среднее
Среднее	-0.00798	1.045489	-0.16375	0.843746	-0.14426	0.898248
Ст. откл.	0.219179	0.311541	0.22933	0.294552	0.220542	0.303339
Мин. знач.	-0.62953	0.181332	-0.98611	0.044849	-0.83732	0.071173
Макс. знач.	0.738362	2.150598	0.791463	1.911632	0.52696	1.918275
Параметры	Алгоритм логистической регрессии		Алгоритм опорных векторов		Алгоритм решающих деревьев	
	Медиана	Среднее	Медиана	Среднее	Медиана	Среднее
Среднее	-0.24515	0.604338	-0.26539	0.296681	-0.24661	0.786644
Ст. откл.	0.220715	0.256878	0.210438	0.154346	0.221246	0.286955
Мин. знач.	-1	-0.03135	-0.94089	-0.1919	-1	0.020537
Макс. знач.	0.493257	1.776516	0.431084	0.934698	0.446037	2.133309

Наилучшую медианную доходность обеспечивает алгоритм Random Forest на исходных данных: -0.00606 против -0.97022. Почти половина опционных контрактов, выбранных с применением алгоритма Random Forest, обеспечит положительный финансовый результат, а средний ожидаемый доход составит +106,7%. Кроме того, этот же алгоритм обеспечивает наименьшее стандартное отклонение для медианного значения.

Можно видеть, что добавление к стохастическим параметрам, тренду и волатильности третьего параметра – фрактальной размерности – показателя Херста позволяет добиться наилучшего качества моделей и получить алгоритм принятия решений о покупке опционов, обеспечивающий наибольшую доходность. Так, наилучшая медианная доходность для моделей с 2 параметрами составляет -0,01, а с 3 – 0,006. При этом достигается большая средняя доходность при меньшем стандартном отклонении финансового результата. Ниже представлены графики медианной и средней доходности для разных алгоритмов с разными параметрами. (рис. 6, 7).

Таблица 10

Среднее и медианное значения доходностей сделок, обеспеченных методами машинного

Параметры	Алгоритм Random Forest		Алгоритм k-ближайших соседей		Алгоритм LightGBM	
	Медиана	Среднее	Медиана	Среднее	Медиана	Среднее
Среднее	-0.01057	1.056867	-0.17956	0.832947	-0.14589	0.910221
Ст. откл.	0.217807	0.318432	0.226319	0.290291	0.220484	0.298737
Мин. знач.	-0.74012	0.165842	-0.92174	0.111021	-0.74957	0.165742
Макс. знач.	0.800498	2.223631	0.516735	1.888386	0.537282	1.88513
Параметры	Алгоритм логистической регрессии		Алгоритм опорных векторов		Алгоритм решающих деревьев	
	Медиана	Среднее	Медиана	Среднее	Медиана	Среднее
Среднее	-0.24148	0.604642	-0.26796	0.296106	-0.24838	0.786773
Ст. откл.	0.220923	0.261719	0.212658	0.152739	0.221935	0.287635
Мин. знач.	-0.92766	-0.03135	-0.98724	-0.1919	-1	0.020537
Макс. знач.	0.493257	1.776516	0.431084	0.934698	0.446037	2.133309

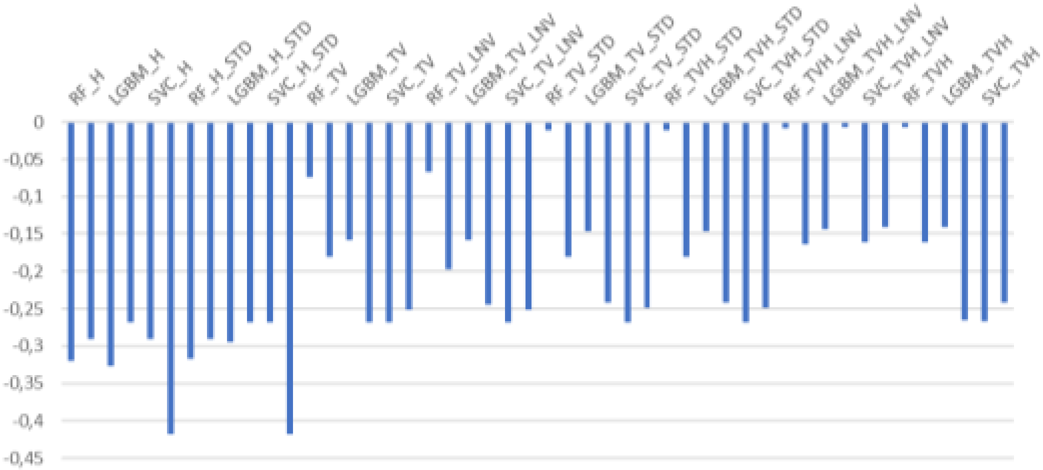

Рис. 6. *Медианное значение доходности для разных параметров и алгоритмов*

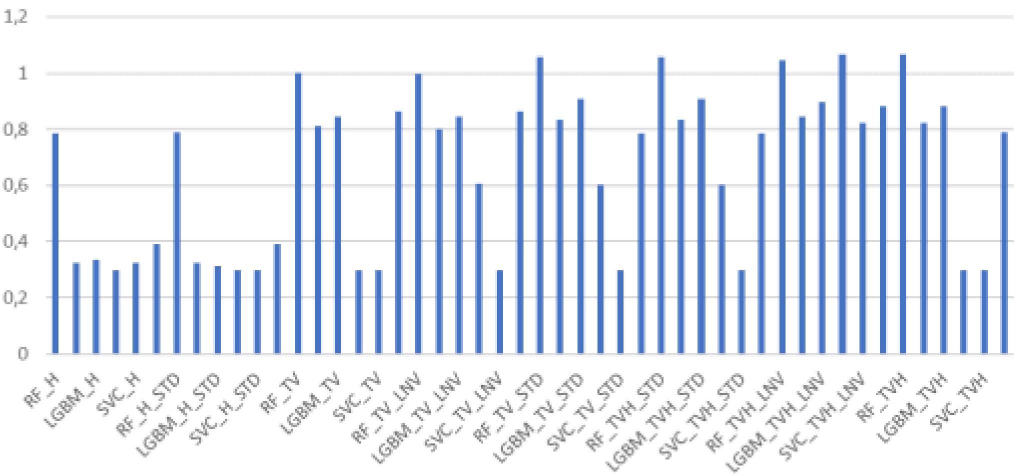

Рис. 7. *Среднее значение доходности для разных параметров и алгоритмов*

Таким образом, применение методов машинного обучения позволяет незначительно повысить долю опционных контрактов с положительной доходностью (в среднем от 50% до 65%), но обеспечить значительное увеличение медианного и среднего значения доходности заключаемых контрактов (от полной потери стоимости опциона до нулевой доходности для медианы и от +30% до +100% для среднего значения).

Выводы. В данной работе авторы попытались совместить 2 подхода к изучению финансовых рынков: стохастический и принцип стохастического детерминизма. В качестве механизма такого объединения выступили классические методы классификации, применяемые в машинном обучении, для создания алгоритма принятия решения – выбора типа опциона. Предложенная методика опробована на реальных данных, и было показано, что эффект от учета фрактальной характеристики данных – показателя Херста – приводит к большей эффективности рассматриваемых алгоритмов по сравнению с использованием методов, рассматриваемых отдельно в рамках каждого из указанных подходов.

Использование методов машинного обучения, как было показано, позволяет значительно повысить ожидаемую прибыль от совершаемых на рынке производных ценных бумаг сделок [7, 8], что, в свою очередь, может способствовать более быстрому и эффективному экономиче-

Литература

1. Boyle, P., McDougall, J., Trading and Pricing Financial Derivatives. – Berlin : Walter de Gruyter Inc., 2019.
2. Avramidi, I. G. Heat Kernel Method and its Application. – Springer Cham Heidelberg Publishing Switzerland, 2015.
3. Половников, Д. С., Колпаков, И. Ю. Показатель Херста в техническом анализе фондовых рынков // Тенденции развития науки и образования. 2019. № 49. С. 5–11. DOI: https://doi.org/10.18411/lj-04-2019-170.
4. Вандерплас, Д. Python для сложных задач : наука о данных и машинное обучение. – СПб. : Питер, 2018.
5. The New York Stock Exchange NYSE [Электронный ресурс]. – URL : https://www.nyse.com/index (дата обращения 25.05.2021).
6. Geron, A. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow. – Second Edition. Concepts, Tools, and Techniques to Build Intelligent Systems. – O'Reilly Media Inc., 2019.
7. Kolanovic, M. Big Data and AI Strategies. Machine Learning and Alternative Data Approach to Investing. – JPMorgan Chase & Co, 2017.
8. Pricope, T.-V. Deep Reinforcement Learning in Quantitative Algorithmic Trading : A review. 2021. – URL : https://arxiv.org/abs/2106.00123.
9. Безсмертная, Е. Р. Место и роль деривативов в эволюции денежной системы // Экономика. Налоги. Право. 2020. № 5. С. 62–70. DOI: https://doi.org/10.26794/1999-849X□2020-13-5-62-70.
10. Россиина, Н. С. Производные финансовые инструменты, их риски и возможность использования в региональной экономике // Социально-политические исследования. 2020. № 4 (9). С. 106–120. DOI: https://doi.org/10.20323/2658-428X-2020-4-9-106-120.

References:
1. Boyle, P., McDougall, J., Trading and Pricing Financial Derivatives. – Berlin : Walter de Gruyter Inc., 2019.
2. Avramidi, I. G. Heat Kernel Method and its Application. – Springer Cham Heidelberg Publishing Switzerland, 2015.
3. Polovnikov, D. S., Kolpakov, I. YU. Pokazatel' Hersta v tekhnicheskom analize fondovyh rynkov // Tendencii razvitiya nauki i obrazovaniya. 2019. № 49. S. 5–11. DOI: https://doi.org/10.18411/lj-04-2019-170.
4. Vanderplas, D. Python dlya slozhnyh zadach : nauka o dannyh i mashinnoe obuchenie. – SPb. : Piter, 2018.
5. The New York Stock Exchange NYSE [Elektronnyj resurs]. – URL : https://www.nyse.com/index (data obrashcheniya 25.05.2021).
6. Geron, A. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow. – Second Edition. Concepts, Tools, and Techniques to Build Intelligent Systems. – O'Reilly Media Inc., 2019.
7. Kolanovic, M. Big Data and AI Strategies. Machine Learning and Alternative Data Approach to Investing. – JPMorgan Chase & Co, 2017.
8. Pricope, T.-V. Deep Reinforcement Learning in Quantitative Algorithmic Trading : A review. 2021. – URL : https://arxiv.org/abs/2106.00123.
9. Bezsmertnaya, E. R. Mesto i rol' derivativov v evolyucii denezhnoj sistemy // Ekonomika. Nalogi. Pravo. 2020. № 5. S. 62–70. DOI: https://doi.org/10.26794/1999-849X□2020-13-5-62-70.
10. Rossiina, N. S. Proizvodnye finansovye instrumenty, ih riski i vozmozhnost' ispol'zovaniya v regional'noj ekonomike // Social'no-politicheskie issledovaniya. 2020. № 4 (9). S. 106–120. DOI: https://doi.org/10.20323/2658-428X-2020-4-9-106-120.