

УДК 330.46

Е.В. Синицын¹*Уральский федеральный университет
имени первого Президента России Б.Н.Ельцина,
г. Екатеринбург, Россия***А.В. Толмачев²***ООО «Датател-Урал»,
г. Екатеринбург, Россия*

МОДЕЛЬ СИСТЕМЫ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ НА ФИНАНСОВЫХ РЫНКАХ ДЛЯ ПРЕДПРИЯТИЙ НА ОСНОВЕ ВЕРОЯТНОСТНОГО АНАЛИЗА И МАШИННОГО ОБУЧЕНИЯ

Аннотация. Расширение портфеля инструментов для управления финансами предприятия с целью повышения доходности вложений является актуальной задачей. В статье мы обсудили модель системы поддержки принятия торговых решений на финансовых рынках на основе вероятностного анализа и машинного обучения, которая может быть использована для ее решения. Целью работы является разработка и апробация модели системы поддержки принятия решений при совершении торговых операций с биржевыми финансовыми инструментами в рамках процесса управления финансами предприятия. Модель основана на технологиях машинного обучения, обеспечивающих получение больших объемов исходных данных, их первичную обработку, формирование многомерного пространства векторов признаков и его трансформацию. Метод прогнозирования рассмотрен на основе правила Байеса. Полученные байесовские вероятности собраны в гиперкуб, который используется для определения правил принятия торговых решений. Разработанная модель протестирована на исторических данных срочного рынка Московской биржи на примере фьючерса на индекс РТС в качестве основного инструмента для выполнения операций и фьючерса на курс доллара США к рублю в качестве вспомогательного инструмента, используемого для анализа. Для оценки результатов тестирования разработаны количественные метрики, включающие в себя количество и объем прибыльных и убыточных сделок, среднюю/средний прибыль/убыток в расчете на одну сделку. С их помощью проанализирована эффективность и границы применимости разработанной модели. Модель может быть реализована в виде программного HFT-робота, способного обеспечить вероятность получения прибыли, превышающую вероятность потерь. В качестве дальнейших шагов по развитию данной темы могут быть предприняты исследования механизмов формирования векторов признаков с помощью методов интеллектуального анализа данных.

Ключевые слова: модель; система поддержки принятия решений; машинное обучение; прогнозирование временных рядов; вероятностный анализ; метод Байеса.

Актуальность темы исследования

Задача срочного размещения временно свободных денежных средств возникает у многих предприятий и организаций в процессе их деятельности. Она может быть решена с помощью банковских депозитов, а также государственных и корпоративных облигаций [1]. Однако такие вложения имеют минимальный риск и низкую

доходность. Для увеличения общей доходности часть денежных средств может быть направлена на спекулятивные торговые операции с ценными бумагами и производными финансовыми инструментами, например фьючерсами, опционами [2]. Такие операции имеют более потенциально более высокую доходность, однако сопряжены с повышенными рисками. Для их успешно-

го выполнения представители финансовых служб предприятий, для которых деятельность на финансовых рынках не является основной (они не являются профессиональными трейдерами, совершающими большое количество сделок на регулярной основе), нуждаются в работоспособных системах поддержки принятия решений. Такие модели базируются на прогнозах ценовых изменений на финансовых рынках и выдают человеку необходимую для принятия решений информацию либо предлагают ему непосредственные рекомендации по купле-продаже определенных финансовых инструментов. Возможен также вариант, когда сделки на основании рекомендаций заключаются с помощью торгового робота [3], реализующего модель в автоматическом режиме без участия человека.

Степень изученности и проработанности проблемы

Задача прогнозирования ценовых изменений на финансовых рынках существует уже давно и относится к категории задач анализа и прогнозирования временных рядов. Для ее решения часто применяются методы математической статистики, использующие широкий набор моделей, например AR, MA, ARMA, ARIMA, ARCH и др. Изучение временных рядов в контексте

экономических задач началось еще в 20-х годах прошлого века. В качестве примера можно привести работы [4–5], хотя это далеко не исчерпывающий список. К настоящему времени статистических модели достаточно хорошо проработаны и нашли свое отражение в литературе [6–8]. К данной группе моделей относятся инструменты технического анализа [9], которые с различной степенью успешности широко применяются на практике. Также для решения задач анализа и прогнозирования временных рядов активно используются методы статистической физики [10–12], использующие модели детерминированного и стохастического хаоса. При этом исследование в данной области с целью улучшения существующих моделей продолжаются. Как правило, такие исследования направлены на решение каких-либо локальных задач с определенными инструментами на конкретных рынках, например валютных [13] или срочных (фьючерсных) [14] рынков. Высокая активность в исследованиях временных рядов и связанных с ними процессов в экономических системах обусловлена сложностью этих процессов с точки зрения математического моделирования, ограниченной применимостью существующих моделей, а также финансовой мотивированностью участников финансовых рынков на поиск эффективных моделей и их применение на практике для получения прибыли. Активное развитие информационных технологий в последние десятилетия привело к появлению доступных высокоскоростных каналов связи и больших вычислительных мощностей, что стимулировало развитие технологий машинного обучения [15] и интеллектуального анализа данных [16], которые нашли свое применение и на финансовых рынках [17]. Также появилось новое направление – высокочастотный трейдинг (англ. HFT: High-Frequency Trading), который представляет собой форму скоростной

¹ Синицын Евгений Валентинович – доктор физико-математических наук, профессор кафедры анализа систем и принятия решений Высшей школы экономики и менеджмента Уральского федерального университета имени первого Президента России Б.Н. Ельцина, г. Екатеринбург, Россия (620002, г. Екатеринбург, ул. Мира, 19); e-mail: sinitsyn_ev@mail.ru.

² Толмачев Александр Владимирович – генеральный директор ООО «Датател-Урал», г. Екатеринбург, Россия (620100, г. Екатеринбург, Сибирский тракт, 12б, офис 311); старший преподаватель кафедры анализа систем и принятия решений Высшей школы экономики и менеджмента Уральского федерального университета имени первого Президента России Б.Н. Ельцина, г. Екатеринбург, Россия (620002, г. Екатеринбург, ул. Мира, 19); e-mail: at@idtu.ru.

торговли на финансовых рынках [18] и также опирается на математические модели. В работе будет рассмотрена разработка модели машинного обучения для прогнозирования и поддержки принятия торговых решения на основе метода Байеса на примере фьючерсных контрактов Московской биржи [19].

Методика исследования

В качестве исходных данных будем рассматривать потоки обезличенных сделок для выбранных финансовых инструментов (deals), представляющие собой последовательности записей такого типа (далее ITS или Initial Time Series):

$$X_{Instr}^{ITS} = \{t_i, x_i, v_i\}_{i=1, \dots, N}^{Instr}, \quad (1)$$

где t_i – момент времени заключения сделки (момент фиксации сделки в системе биржи);

x_i – цена, по которой заключена сделка;

v_i – объем сделки (количество лотов финансового инструмента в сделке);

N – количество членов временного ряда (количество сделок).

Временные ряды ITS являются дискретными и неравномерными. Если рассматривать несколько различных финансовых инструментов, то их ITS будут не синхронны друг с другом. Это затрудняет их совместный анализ и прогнозирование.

Под прогнозированием будем понимать вероятностную оценку изменения цены финансового инструмента по величине ($|\Delta x| \geq x_{\min \Delta}$ или $\Delta x = x_{\text{prediction}}$) и направлению (Up: $\Delta x > 0$; Down: $\Delta x < 0$; Flat: $\Delta x = 0$) на ближайший, заранее определенный интервал времени Δt_{fix} . Прогнозировать точную величину изменения цены не имеет смысла. С практической точки зрения достаточно, чтобы изменение было больше некоторого порогового значения и имело нужное направление.

Для целей анализа и прогнозирования целесообразно использовать не исходные, а модифицированные временные ряды, далее MTS (Modified Time Series), которые также являются дискретными, но отличаются от ITS равномерностью, синхронностью и равным количеством членов. Для перехода к ним разобьем временную ось на равные интервалы времени (тайм-фреймы или bars). Тайм-фрейм выбирается в соответствии с интервалом прогнозирования Δt_{fix} . В настоящей работе будут рассмотрены тайм-фреймы 1s, 2s, 3s, 4s, 5s, 10s, 60s, и проанализирована зависимость эффективности модели от величины тайм-фрейма. В работе будем обозначать тайм-фрейм моментом времени его начала; время приведем к POSIX формату; при формировании MTS ограничимся дневной торговой сессией с максимальным количеством участников и ликвидностью. В качестве цены, характеризующей тайм-фрейм в MTS, будем использовать цену закрытия, то есть цену последней сделки в этом интервале времени. Если в какой-либо интервал времени не было сделок, то для него будем брать цену закрытия предыдущего интервала.

В качестве примера анализ и прогнозирование будем выполнять для фьючерса на индекс РТС на Московской бирже³. Для машинного обучения определим пространство признаков. Возьмем два исходных временных ряда:

X_{RTS}^{ITS} – цены на МБ на фьючерс индекса РТС (основной инструмент);

X_{SI}^{ITS} – цены на МБ на фьючерс USD – RUB (вспомогательный инструмент).

Используя описанную выше методику, перейдем от ITS к MTS с выбранными для изучения тайм-фреймами 1s, 2s, 3s, 4s, 5s, 10s, 60s.

³ Сайт Московской биржи. Деривативы. <http://moex.com/ru/derivatives/>

Цены находятся в постоянном движении и на каждом тайм-фрейме могут изменяться вверх – U (Up), вниз – D ($Down$) или оставаться без изменения – F ($Flat$). MTS в этом случае может быть схематически представлен в виде последовательностей: $D-U-F-U-F-U-U-D$, $U-D-D-F-D-U$. Последовательности тайм-фреймов, когда цена изменялась в одну сторону, например $U-U-U$ или $D-D-D-D$, будем называть цепочками ($Chains$), а ситуации типа $U-D$, $U-F$, $D-U$, $D-F$, когда направленное движение цены закончилось, – обрывом цепочки.

Для применения в модели метода Байеса [19–20] сформулируем гипотезу, справедливость которой будем проверять и на основании чего будем принимать решения, а также набор событий, о реализации которых мы располагаем информацией. Сформулируем две гипотезы, которые покрывают все возможные случаи изменения цены:

H_1 – обрыв цепочки (ситуации типа $U-D$, $U-F$, $D-U$, $D-F$, $F-F$);

H_2 – продолжение цепочки (ситуации строго $U-U$ или $D-D$).

Для количественного описания цепочек определим 4-компонентный вектор признаков ($attribute$), который получается с помощью оператора $\hat{F}(\Delta t)$:

$$\begin{aligned} \hat{F}: X_{RTS}^{MTS}, X_{Si}^{MTS} &\rightarrow Y_{H_m/\Delta x}^{atr} = \\ &= \{L_t^{RTS}, L_x^{RTS}, L_{dx}^{RTS}, L_x^{Si}\}_{H_m/\Delta x}, \end{aligned} \quad (2)$$

где L_t^{RTS} – длина цепочки в единицах времени ($L_t^{RTS} = t_{end} - t_{start}$);

L_x^{RTS} – длина цепочки в единицах цены ($L_x^{RTS} = x_{t_{end}}^{RTS} - x_{t_{start}}^{RTS}$);

L_{dx}^{RTS} – относительная производная ($L_{dx}^{RTS} = \Delta x_{last}^{RTS} / L_x^{RTS}$, $\Delta x_{last}^{RTS} = x_{t_{end}}^{RTS} - x_{t_{end}-\Delta t}^{RTS}$);

L_x^{Si} – изменение цены вспомогательного инструмента ($L_x^{Si} = x_{t_{end}}^{Si} - x_{t_{start}}^{Si}$);

H_m , $m = 1, 2$ – одна из двух гипотез об обрыве или продолжении цепочки;

Δx – изменение цены на следующем шаге.

Механизм определения элементов вектора признаков основан на факторном анализе и здесь рассматриваться не будет. Размерность пространства признаков принципиально не ограничена 4 и может быть без проблем увеличена за счет включения в число элементов каких-либо других индикаторов, например индикаторов технического анализа [9].

Имея пространство признаков на исторических данных, сформируем 4-мерный гиперкуб Байесовских вероятностей, с помощью которого по вектору признаков на тестовых данных, не задействованных при обучении, будем определять вероятность продолжения (гипотеза H_2) или обрыва цепочки (гипотеза H_1), делать прогноз на следующий тайм-фрейм, на основании которого принимать решение о заключении сделки.

Для подсчет Байесовских вероятностей для 4-мерного гиперкуба воспользуемся:

$$\begin{aligned} P(H_m | Y^{atr}) &= \frac{P(H_m) \cdot P(Y^{atr} | H_m)}{\sum_m P(H_m) \cdot P(Y^{atr} | H_m)}, \\ m &= 1, 2. \end{aligned} \quad (3)$$

Вероятности $P(H_1)$, $P(H_2)$, $P(Y^{atr} | H_1)$, $P(Y^{atr} | H_2)$ вычисляются на основе статистической обработки векторов пространства признаков Y^{atr} . При подсчете Байесовских вероятностей выполним трансформацию векторов признаков, которая будет заключаться в нормализации и фильтрации векторов. В рамках процедуры нормализации векторов признаков мы приведем значения их компонентов к нужной сетке дискретизации, а фильтрацией исключим ненужные и несущественные данные (статистический «мусор»). Пример алгоритма нормализации, используемого в работе, приведен в табл. 1.

Такая процедура формирования гиперкуба является одним из видов машинного обучения, поэтому будем ее также называть

обучением системы. Полная схема обучения показана в рис. 1.

С помощью полученного гиперкуба можно делать прогнозы изменения цены по направлению и величине на следующем тайм-фрейме. Алгоритм прогнозирования, когда мы имеем гиперкуб, обученный на исторических данных, и временной ряд тестовых данных, не пересекающийся с использованными при обучении данными, показан рис. 2.

Правила принятия решения на основе сделанного прогноза могут быть, например, такими, как приведены в табл. 2 (работа только по гипотезе H_1). Такие правила используются в настоящей работе. Решение «BUY» означает покупку финансового инструмента, «SELL» – его продажу, «OUT» – отсутствие каких-либо действий. Короткие позиции, когда продажа предшествует покупке, разрешены.

Таблица 1

Пример алгоритма нормализации

Компоненты вектора Y^{attr}	ДО нормализации	ПОСЛЕ нормализации	Фильтрация
L_t^{RTS}	в секундах	в шагах тайм-фреймов	$L_t^{RTS} \geq 2$
L_t^{RTS}	в ед. цены с шагом 10	остается как есть	$ L_x^{RTS} \geq 10$ $ \Delta_{x(next)}^{RTS} \geq 10$
L_{dx}^{RTS}	действительное число из отрезка [0, 1]	действительное число из отрезка [0, 1], кратное 0,1	$L_{dx}^{RTS} \leq 0,95$
L_x^{Si}	в единицах цены с шагом 1	в единицах цены с шагом 2	не фильтруется

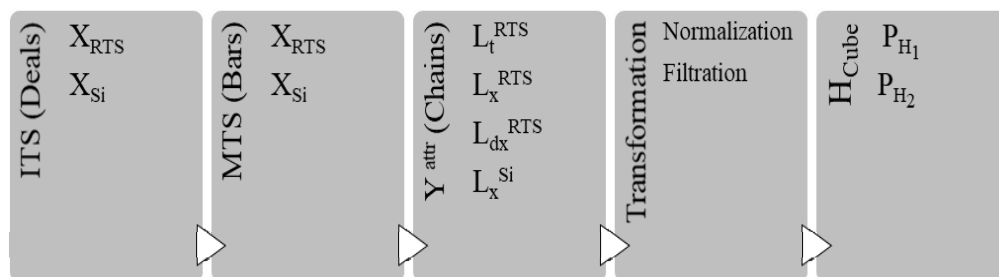


Рис. 1. Схема машинного обучения



Рис. 2. Алгоритм выделения вектора признаков и прогнозирования на его основе
помощью гиперкуба с байесовскими вероятностями

Таблица 2

Пример правил принятия решений на основе гиперкуба

Направление цепочки до t_i	Вероятности гипотез H_1 и H_2	Прогноз на $[t_p, t_{i+D}]$	Сделка в t_i	Сделка в t_{i+D}
U (Up)	$P(H_1) > P(H_2)$	D (Down)	SELL	BUY
	$P(H_1) < P(H_2)$	U (Up)	OUT	OUT
	$P(H_1) = P(H_2)$ или не определена	–	OUT	OUT
D (Down)	$P(H_1) > P(H_2)$	U (Up)	BUY	SELL
	$P(H_1) < P(H_2)$	D (Down)	OUT	OUT
	$P(H_1) = P(H_2)$ или не определена	–	OUT	OUT
F (Flat)	$\square P(H_1), P(H_2)$	–	OUT	OUT

Анализ полученных результатов

Для тестирования модели возьмем исходные временные ряды в виде таблиц обезличенных сделок на Московской бирже для инструментов RTS-9.17 и Si-9.17 за 58 дней в период с 03.07.2017 по 20.09.2017. Из них будем брать данные для обучения и тестирования. Например, обучение проводится по 10 дням, тестирование на 11-й день. Таким образом множества данных для обучения и тестирования никогда не пересекаются.

Для оценки эффективности разработанной модели будем использовать количественные характеристики, которые позволят оценить качество прогнозов как по направлению изменения цены, так и по величине:

- Количество правильных и ошибочных решений.
- Суммы прибылей и убытков по заключенным сделкам.
- Производные характеристики, например, средняя прибыль на 1 сделку.
- Нас интересует зависимость эффективности модели от:
- Величины тайм-фрейма. Рассмотрим 1s, 2s, 3s, 4s, 5s, 10s, 60s.
- Глубины обучения гиперкуба. Рассмотрим диапазон от 1 до 30 дней.

Для каждой комбинации тайм-фрейма и глубины обучения будем делать тесты на 25 последовательных днях. Результаты получены следующие.

На рис. 3 показаны результаты тестирования модели в виде средней прибыли на одну сделку для разных тайм-фреймов и глубины обучения. Этот показатель является самым важным в оценке результативности модели с практической точки зрения. Если его величина положительна и выше суммарных расходов, нормированных на одну сделку (разного рода комиссий и иных расходов), то результат положительный и модель можно использовать на практике. Результативность модели уменьшается при увеличении тайм-фрейма. Для тайм-фреймов 1s...10s результативность снижается, но остается в положительной области. При тайм-фрейме 60s результативность имеет математическое ожидание

близкое к нулю и высокое стандартное отклонение. Это говорит о применимости модели на коротких тайм-фреймах, измеряемых в секундах в области HFT [21]. Результативность имеет минимумы на минимальных и максимальных глубинах обучения и максимум между ними, то есть существует оптимальная глубина обучения. Снижение результативности модели при большой глубине обучения объясняется тем, что гиперкуб заполняется противоречивыми прогнозами, отражающими изменение общей ситуации на рынке. Пониженная результативность на малых глубинах обучения связана с недостаточным объемом статистики при подсчете гиперкуба. Вид графиков на средних

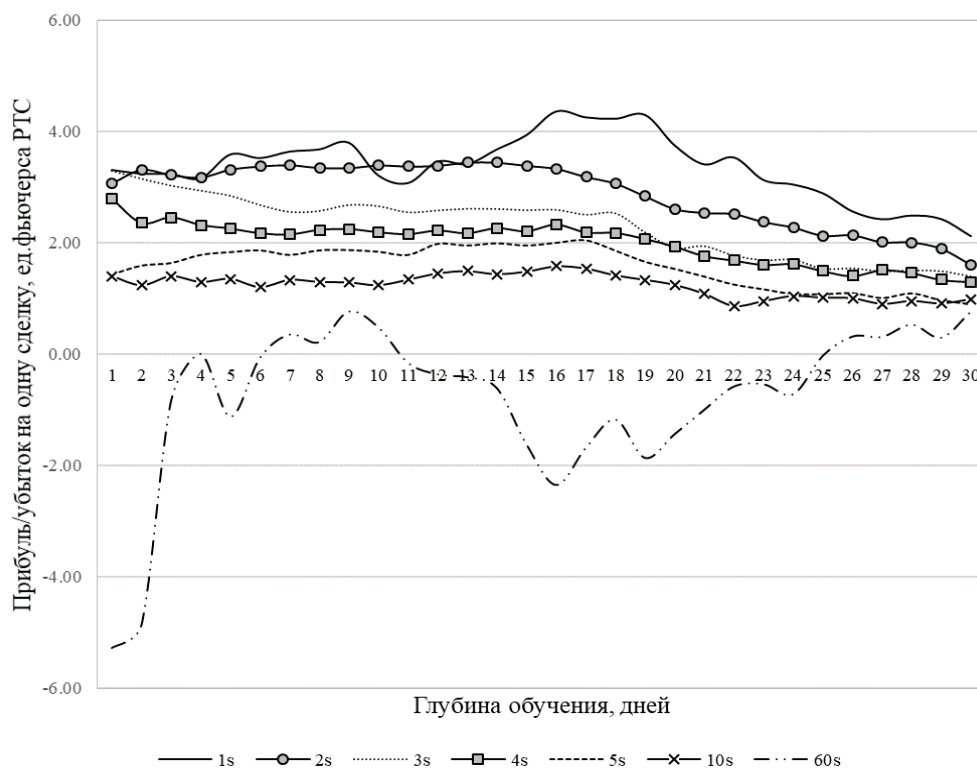


Рис. 3. Зависимость от глубины обучения средней прибыли/убытка на 1 сделку для разных тайм-фреймов

тайм-фреймах 2s...5s говорит о стабильности модели при разной глубине обучения.

На рис. 4, 5 показаны разности числа (в штуках) и сумм (в денежных единицах) правильных и ошибочных сделок в зависимости от глубины обучения и тайм-фрейма. Ситуация аналогичная – результативность падает с увеличением тайм-фрейма, имеет оптимальную глубину обучения. На средних тайм-фреймах 2s...5s модель стабильна при изменении глубины обучения.

На рисунках 6–8 приведены кривые средней прибыли на одну сделку и заполняемости гиперкуба для разных тайм-фреймов и глубины обучения. Заполняемость гиперкуба определена как отноше-

ние общего числа цепочек, взятых в расчет гиперкуба, к числу заполненных ячеек гиперкуба. На средних тайм-фреймах до 10s видно, что существует критическая величина заполняемости гиперкуба (~10), после которой начинается резкий спад результативности. На больших тайм-фреймах 60s такого не наблюдается.

На рис. 9–11 показана суммарная прибыль/убыток нарастающим итогом при использовании модели. На коротких тайм-фреймах (например, 2s) наблюдается стабильный рост прибыли практически без просадок (наличия дней с отрицательным результатом). По мере роста тайм-фрейма просадка растет и для тайм-фрейма 60s ре-

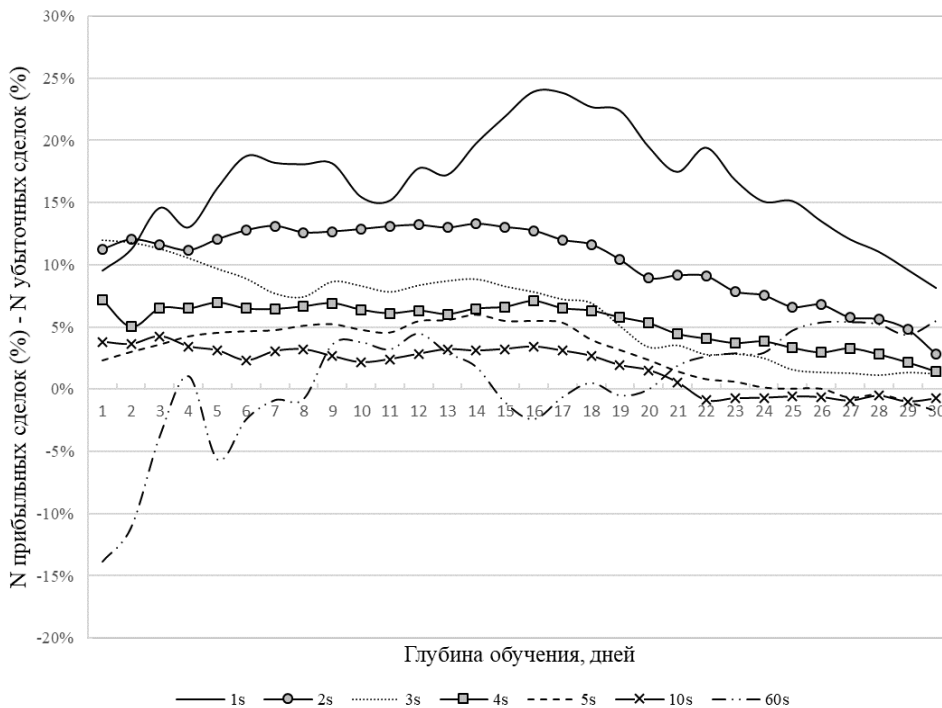


Рис 4. Зависимость от глубины обучения разницы числа правильных и ошибочных сделок в % для разных тайм-фреймов

Синицын Е.В., Толмачев А.В.

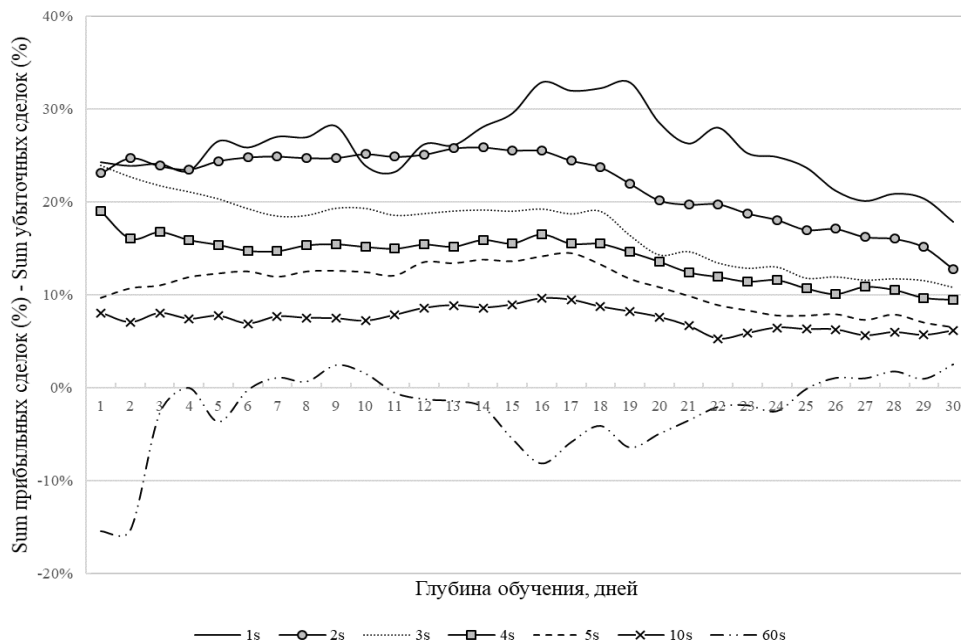


Рис. 5. Зависимость от глубины обучения разницы сумм правильных и ошибочных сделок в % для разных тайм-фреймов

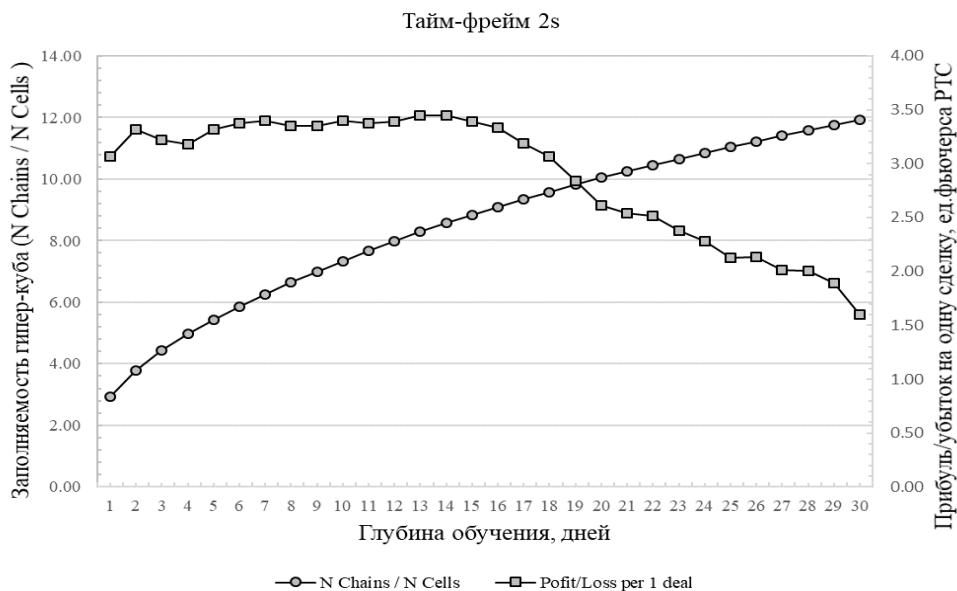


Рис. 6. Средняя прибыль на 1 сделку и заполняемость куба для тайм-фрейма 2s

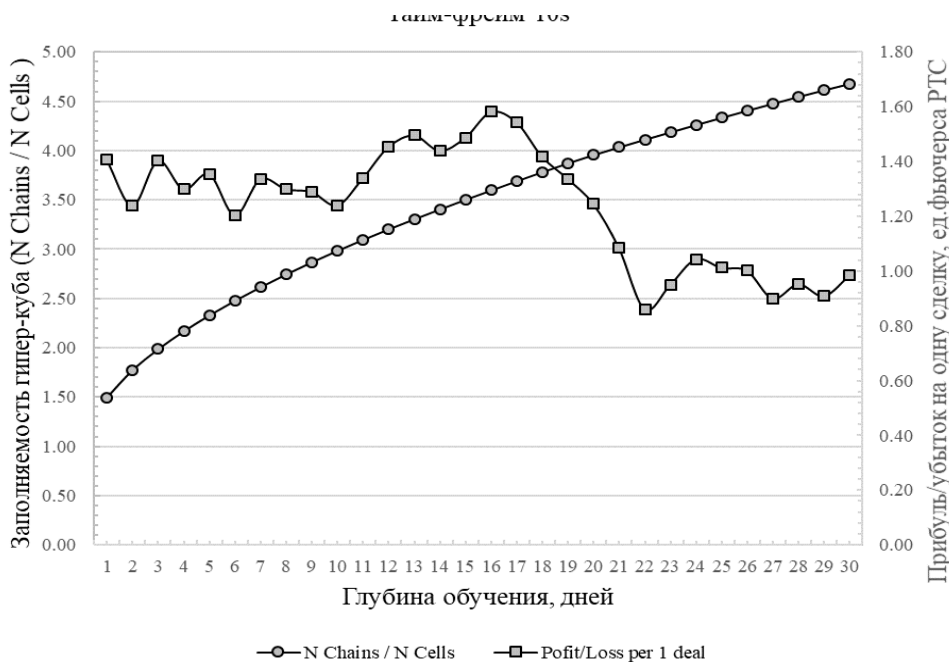


Рис. 7. Средняя прибыль на 1 сделку и заполняемость куба для тайм-фрейма 10s

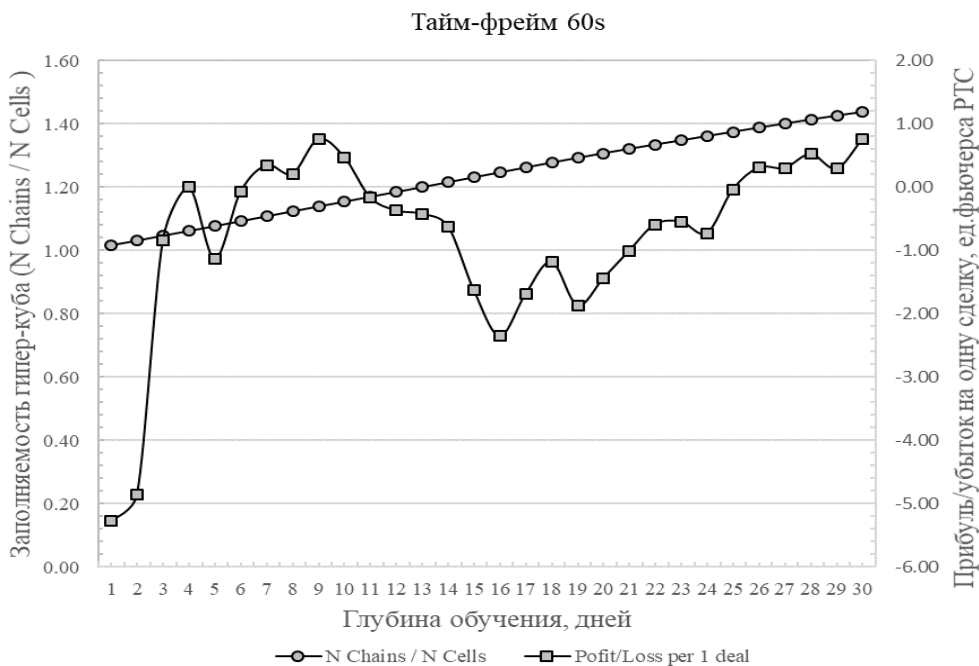


Рис. 8. Средняя прибыль на 1 сделку и заполняемость куба для тайм-фрейма 60s

зультат становится в большинстве случаев отрицательным с сильными просадками.

Основные выводы

Предложенная модель имеет стабильный положительный результат на коротких тайм-фреймах 1-2s, относящихся к области HFT, и если позволяет быстродействие серверов и ПО, скорость сетей передачи данных, то может быть использо-

вана на практике для реализации системы управления финансами предприятия, направленных на спекулятивные торговые операции с ценными бумагами и производными финансовыми инструментами. Реализация системы – программный торговый робот, осуществляющий анализ рынка, принятие торговых решений и заключение на их основе сделок в автоматическом режиме.

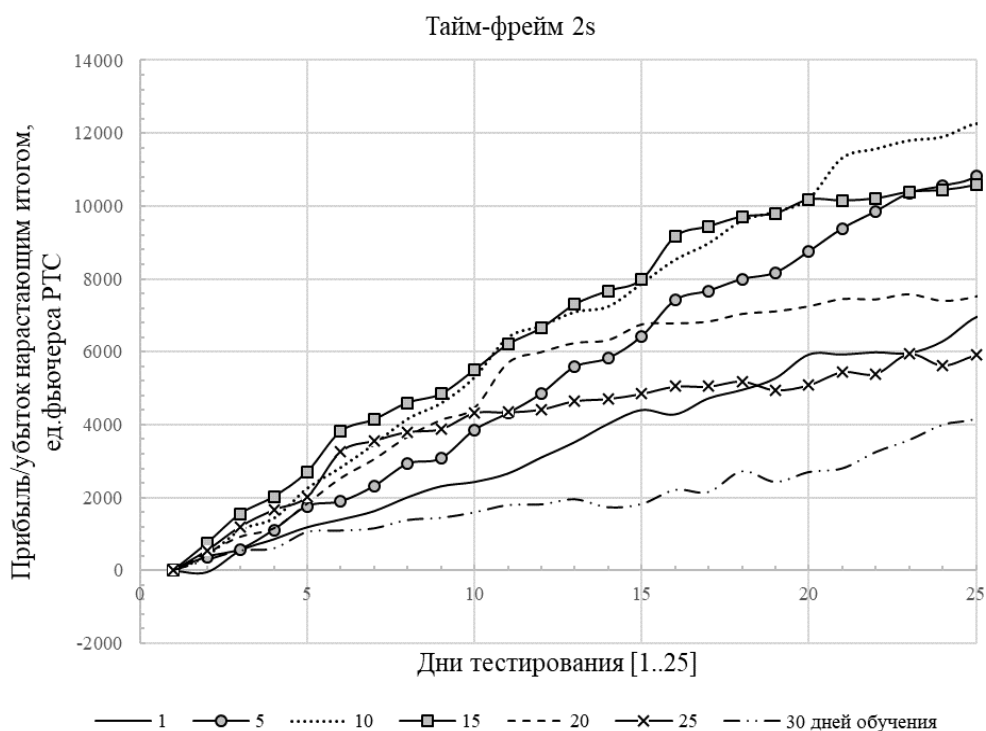


Рис. 9. Прибыль/убыток нарастающим итогом за 25 дней в зависимости от глубины обучения для тайм-фрейма 2s

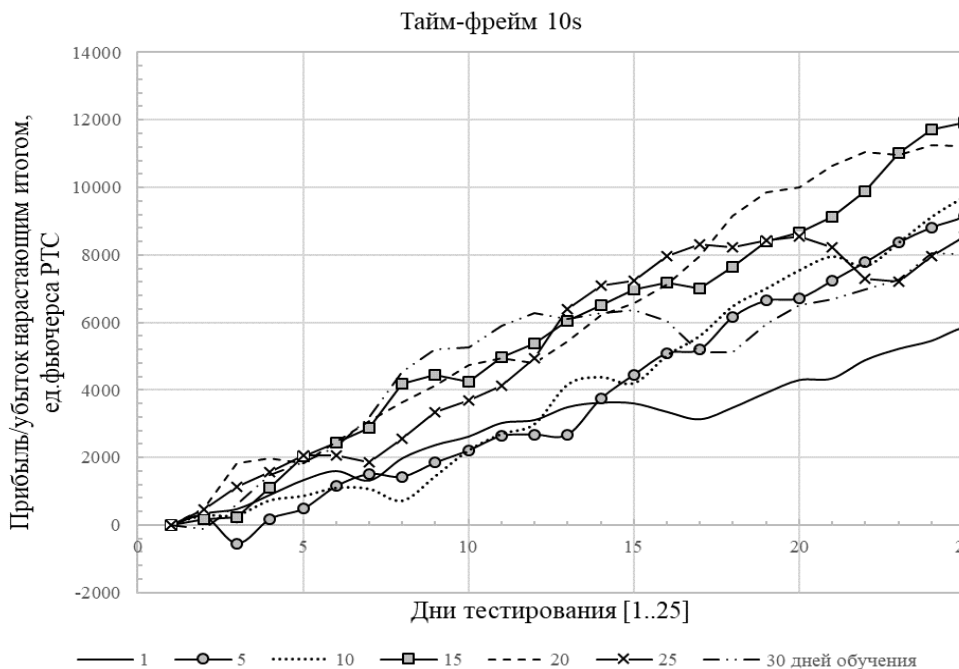


Рис. 10. Прибыль/убыток нарастающим итогом за 25 дней в зависимости от глубины обучения для тайм-фрейма 10s

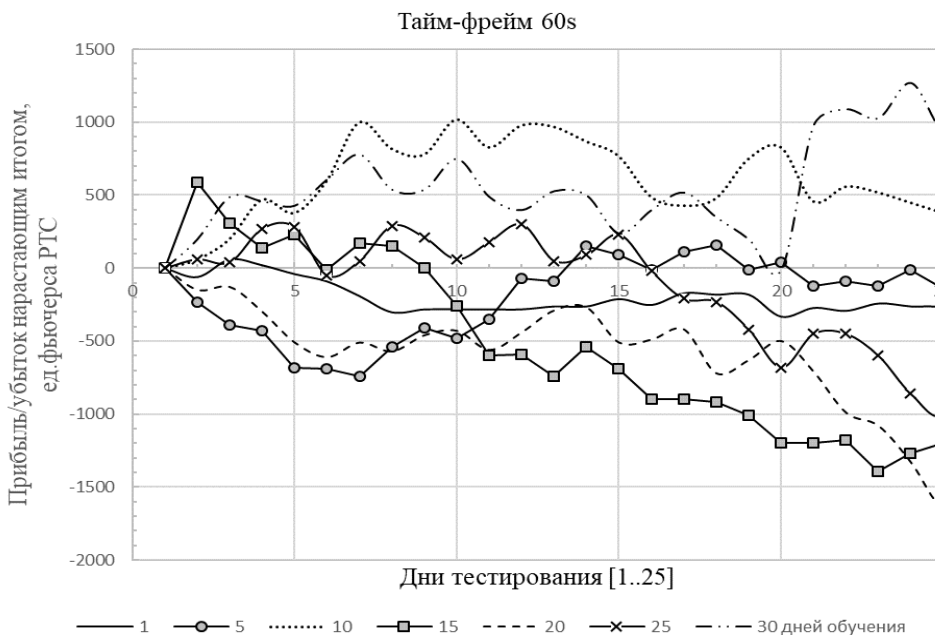


Рис. 11. Прибыль/убыток нарастающим итогом за 25 дней в зависимости от глубины обучения для тайм-фрейма 60s

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Брейли Р., Майерс С. Принципы корпоративных финансов / пер. с англ. Н. Барышниковой. М. : ЗАО «Олимп-Бизнес», 2008. 1008 с.
2. Шимко П.Д. Международный финансовый менеджмент : учебник и практикум для бакалавриата и магистратуры. 2-е изд., перераб. и доп. М. : Издательство «Юрайт», 2014. 494 с.
3. Чебогарев Ю.А. Торговые роботы на российском фондовом рынке. 2-е изд., перераб. и доп. М. : Смартбук, 2011. 160 с.
4. Yule G. On the method of investigating periodicities in disturbed series // *Philosophical Transaction*. 1927. Vol. 226A. P. 267–298.
5. Slutsky E. The summation of random causes as the source of cyclical processes // *Econometrica*. 1925. Vol. 56. P. 105–146.
6. Арженовский С.В., Молчанов И.Н. Статистические методы прогнозирования : учебное пособие. Ростов-н/Д. : Рост. гос. экон. ун-т «РИНХ», 2001. 73 с.
7. Уотшем Т.Дж., Паррамоу К. Количественные методы в финансах / пер. с англ. под ред. М.Р. Ефимовой. М. : Финансы, ЮНИТИ, 1999. 527 с.
8. Vox G., Jenkins G.M., Reinsel G. *Time Series Analysis: Forecasting and Control*. 3rd edition. New Jersey : Prentice Hall, 1994. 614 p.
9. Murphy J. *Technical Analysis of the Financial Markets: A Comprehensive Guide to Trading Methods and Applications*. New York : Institute of Finance, 1999. 576 p.
10. Мантенья Р.Н., Стенли Х.Ю. Введение в экономфизику. Корреляции и сложность в финансах / пер. с англ. В.И. Гусева, С.В. Малахова, А.И. Митуса под ред. В.Я. Габескирия. М. : ЛИБРОКОМ, 2009. 187 с.
11. Шустер Г. Детерминированный хаос. Введение / пер. с нем. М. : Мир, 1988. 253 с.
12. Безручко Б.П., Смирнов Д.А. Математическое моделирование и хаотические временные ряды. Саратов : ГосУНЦ «Колледж», 2005. 320 с.
13. Балонишников А.М., Балонишникова В.А., Копыльцов А.В. Прогнозирование временных рядов методами Фармера-Сидоровича и Бокса-Дженкинса // *Известия Российского государственного педагогического университета им. А.И. Герцена*. 2011. № 141. С. 7–16.
14. Григорьев В.П., Козловских А.В., Марьясов Д.А. Исследование математической модели фьючерсных рынков // *Рынок ценных бумаг*. 2005. № 9(288). С. 38–42.
15. Bishop C. *Pattern Recognition and Machine Learning*. New York : Springer, 2006. 738 p.
16. Leskovets Y., Radjaraman A., Ulman J. *Mining of Massive Datasets*. Cambridge University Press, 2011. 457 p.
17. Shen S., Jiang H., Zhang T. *Stock Market Forecasting Using Machine Learning Algorithms* [Электронный ресурс]. Режим доступа: <http://cs229.stanford.edu/proj2012/ShenJiangZhang-StockMarketForecastingusingMachineLearningAlgorithms.pdf>.
18. Miller R.S., Shorter G. *High Frequency Trading : Overview of Recent Developments*. Washington, DC Congressional Research Service, 2016. 19 p.
19. Вентцель Е.С. Теория вероятностей. М. : Наука, 1969. 366 с.
20. Феллер В. Введение в теорию вероятностей и ее приложения. Том 1. М. : Мир, 1964. 484 с.
21. Толмачев А.В., Синицын Е.В. Вероятностный анализ в высокочастотном трейдинге // *Российские регионы в фокусе перемен: сборник докладов XI Международной конференции*. Екатеринбург : Издательство УМЦ УПИ, 2016. Ч. 1. С. 91–97.

Sinitsyn E.V.*Ural Federal University**named after the first President of Russia B.N. Yeltsin,**Ekaterinburg, Russia***Tolmachev A.V.***Datatel-Ural LLC,**Ekaterinburg, Russia*

MODEL OF THE DECISION SUPPORT SYSTEM ON THE FINANCIAL MARKETS FOR ENTERPRISES BASED ON PROBABILITY ANALYSIS AND MACHINE LEARNING

Abstract. Expanding the portfolio of instruments for finance management of an enterprise in order in order to increase the return on investments is a task of current interest. In the article, we discuss a model of the trading decision support system in financial markets based on probability analysis and machine learning, which can be used to solve the above problem. The aim of this work is to develop and test a model of the decision support system for trading operations with stock financial instruments as a part of the enterprise financial management process. The model is based on machine learning technologies that provide acquisition of large amounts of input data, its primary processing, the formation of a multi-dimensional space of feature vectors, and its transformation. The forecasting method is based on the Bayesian rule. The obtained Bayesian probabilities are stored in a hypercube which is used to determine the rules for trading decisions making. The developed model was tested on historical data of the futures market of the Moscow Exchange in the case of the RTS Index Futures as the main instrument for transactions and the USD-RUB Futures as an auxiliary instrument used for analysis. To evaluate the results of testing, quantitative metrics have been developed, which include the number and volume of profitable and unprofitable transactions, the average profit/loss per transaction. These metrics were used for analysis of effectiveness and limits of applicability for the developed model. The model can be implemented as a software HFT robot that can provide the probability to get profit greater than the probability of losses. As a further step in the development of this topic, research can be undertaken on the mechanisms for the formation of feature vectors using data mining methods.

Keywords: model; decision support system; machine learning; time series forecasting; probabilistic analysis; the Bayes method.

References

1. Brealey, R., Myers, S. (2008). *Principles of Corporate Finance*. Irwin, McGraw-Hill.
2. Shimko, P.D. (2014). *Mezhdunarodnyi finansovyi menedzhment [International Financial Management]*. Moscow, Iurait.
3. Chebotarev, Iu.A. (2011). *Torgovye roboty na rossiiskom fondovom rynke [Trading Robots in the Russian Stock Market]*. Moscow, Smartbuk.
4. Yule, G. (1927). On the method of investigating periodicities in disturbed

- series. *Philosophical Transaction*, Vol. 226A, 267–298.
5. Slutsky, E. (1925). The summation of random causes as the source of cyclical processes. *Econometrica*, Vol. 56, 105–146.
 6. Arzhenovsky, S.V., Molchanov, I.N. (2001). *Statisticheskie metody prognozirovaniia [Statistical Forecasting Methods]*. Rostov-on-Don. Rostov State University of Economics RINKh.
 7. Watsham, T., Parramore, K. (1996). *Quantitative Methods for Finance*. Cengage Learning EMEA.
 8. Box, G., Jenkins, G.M., Reinsel, G. (1994). *Time Series Analysis: Forecasting and Control*. 3rd edition. New Jersey, Prentice Hall, 614 p.
 9. Murphy, J. (1999). *Technical Analysis of the Financial Markets: A Comprehensive Guide to Trading Methods and Applications*. New York, Institute of Finance, 576 p.
 10. Mantegna, R., Stanley, H.E. (2000). *An Introduction to Econophysics*. Cambridge University Press.
 11. Schuster, H.G. (1984). *Deterministic Chaos*. Weinheim, Wiley-VCH.
 12. Bezruchko, B.P., Smirnov, D.A. (2005). *Matematicheskoe modelirovanie i khaoticheskie vremennye riady [Mathematical modeling and Chaotic Time Series]*. Saratov, GosUNTs Kollodzh.
 13. Balonishnikov, A.M., Balonishnikova, V.A., Kopyltsov, A.V. (2011). Prognozirovanie vremennykh riadov metodami Farmera-Sidorovicha i Boksa-Dzhenkinsa (Prediction of Time Series by Box-Jenkins and Farmer-Sidorovich Methods). *Izvestiia Rossiiskogo gosudarstvennogo pedagogicheskogo universiteta im. A.I. Gertsena (Izvestia: Herzen University Journal of Humanities & Science)*, No. 141, 7–16.
 14. Grigoryev, V.P., Kozlovskikh, A.V., Maryasov, D.A. (2005). Issledovanie matematicheskoi modeli f'uchersnykh rynkov [A Study of the Mathematical Model Futures Market]. *Rynok tsennykh bumag [Stock Market]*, No. 9(288), 38–42.
 15. Bishop, C. (2006). *Pattern Recognition and Machine Learning*. New York, Springer, 738 p.
 16. Leskovets, Y., Radjaraman, A., Ulman, J. (2011). *Mining of Massive Datasets*. Cambridge University Press, 457 p.
 17. Shen, S., Jiang, H., Zhang, T. *Stock Market Forecasting Using Machine Learning Algorithms*. Available at: <http://cs229.stanford.edu/proj2012/ShenJiangZhang-StockMarketForecastingUsingMachineLearningAlgorithms.pdf>.
 18. Miller, R.S., Shorter, G. (2016). *High Frequency Trading: Overview of Recent Developments*. Washington, DC, Congressional Research Service, 19 p.
 19. Ventsel, E.S. (1969). *Teoriia veroiatnostei [Theory of Probabilities]*. Moscow, Nauka.
 20. Feller, W. (1957). *An Introduction to Probability Theory and its Applications*. Vol. 1. John Wiley and Sons Inc.
 21. Tolmachev, A.V., Sinitsyn, E.V. (2016). Veroiatnostnyi analiz v vysokochastotnom treidinge [Probability Analysis in High Frequency Trading]. *Proceedings of 11th international conference "Russian Regions in the Focus of Transformations"*. UMTs UPI, Part 1, 91–97.

Information about the authors

Sinitsyn Evgeny Valentinovich – Doctor of Physical and Mathematical Sciences, Professor, Department of Systems Analysis, Graduate School of Economics and Management, Ural Federal University named after the first President of Russia B.N. Yeltsin, Ekaterinburg, Russia (620002, Ekaterinburg, Mira street, 19); e-mail: sinitsyn_ev@mail.ru.

Tolmachev Alexander Vladimirovich – CEO, Datatel-Ural LLC, Ekaterinburg, Russia (620100, Ekaterinburg, Sibirskij trakt, 12B, office 311); Senior Lecturer, Department of Systems Analysis, Ural Federal University named after the first President of Russia B.N. Yeltsin, Ekaterinburg, Russia (620002, Ekaterinburg, Mira street, 19); e-mail: at@idtu.ru.

Для цитирования: Синицын Е.В., Толмачев А.В. Модель системы поддержки принятия решений на финансовых рынках для предприятий на основе вероятностного анализа и машинного обучения // Вестник УрФУ. Серия экономика и управление. 2019. Т. 18, № 3. С. 378–393. DOI: 10.15826/vestnik.2019.18.3.019.

For Citation: Sinitsyn E.V., Tolmachev A.V. Model of the Decision Support System on the Financial Markets for Enterprises Based on Probability Analysis and Machine Learning. *Bulletin of Ural Federal University. Series Economics and Management*, 2019, Vol. 18, No. 3, 378–393. DOI: 10.15826/vestnik.2019.18.3.019.

Информация о статье: дата поступления 2 апреля 2019 г.; дата принятия к печати 26 апреля 2019 г.

Article Info: Received April 2, 2019; Accepted April 26, 2019.