

ИССЛЕДОВАТЕЛЬСКАЯ СТАТЬЯ

<https://doi.org/10.31063/AlterEconomics/2024.21-3.6>

УДК 336.76

JEL C38, C88, D14, D81, D91



## Методы машинного обучения в риск-профилировании инвестора<sup>1</sup>

Сергей В. КУРОЧКИН <sup>1)</sup> , Николай А. ПАВЛОВ <sup>2)</sup>,  
Мария В. ТКАЧЕНКО <sup>3)</sup>  , Елена А. ЯРЕМИЧ <sup>4)</sup>

<sup>1, 3)</sup> Национальный исследовательский университет «Высшая школа экономики», г. Москва, Российская Федерация

<sup>2)</sup> Snke OS GmbH, г. Мюнхен, Германия

<sup>3)</sup> ПАО «Совкомбанк», г. Москва, Российская Федерация

<sup>4)</sup> ПАО «Сбербанк», г. Москва, Российская Федерация

**Для цитирования:** Курочкин, С. В., Павлов, Н. А., Ткаченко, М. В., Яремич, Е. А. (2024). Методы машинного обучения в риск-профилировании инвестора. *AlterEconomics*, 21(3), 527–552.

<https://doi.org/10.31063/AlterEconomics/2024.21-3.6>

**Аннотация.** Риск-профилирование инвестора перед продажей ему финансовых продуктов необходимо для того, чтобы отсеять инструменты, степень риска которых психологически неприемлема для этого инвестора, и избежать мисселлинга. Подавляющее большинство финансовых институтов для определения риск-профиля инвестора просят его заполнить анкету. Однако, многие исследования ставят под сомнение эффективность этого метода — в частности потому, что в реальной жизненной ситуации человек может повести себя совершенно по-другому, нежели предположил, отвечая на вопрос анкеты. Цель исследования заключается в том, чтобы предложить инструмент, основанный не на ответах респондента, а на его поведении в ситуации того или иного изменения рыночных условий. Для достижения поставленной цели был построен веб-интерфейс, через который осуществлялся сбор данных. Для получения выводов использовался регрессионный анализ, реализованный на языке программирования Python. В статье предложена разработанная авторами интерактивная платформа. Ее суть состоит в том, что пользователю предлагаются 10 игровых кейсов, имитирующих реальные торги, в каждом из которых он должен выбрать точку закрытия позиции, зафиксировав убыток или прибыль. Риск-профиль, таким образом, определяется на основании поведения пользователей, а не их ответов на абстрактные вопросы. По итогу эксперимента было показано, что построенная логистическая модель верно предсказывает 65,78 % решений инвесторов, тогда как анкетирование правильно выявило риск-профиль только 53,7 % инвесторов. Полученные результаты могут использоваться в качестве базы для дальнейших разработок инструментов риск-профилирования клиентов, приобретающих инвестиционные продукты.

**Ключевые слова:** риск-профиль, толерантность к риску, машинное обучение, искусственный интеллект, анкетирование

<sup>1</sup> © Курочкин С. В., Павлов Н. А., Ткаченко М. В., Яремич Е. А. Текст. 2024.

# Machine Learning Methods in Investor Risk Profiling

Sergey V. KUROCHKIN <sup>1)</sup> , Nikolay A. PAVLOV <sup>2)</sup>,  
Mariia V. TKACHENKO <sup>3)</sup>  , Elena A. YAREMICH <sup>4)</sup>

<sup>1,3)</sup> National Research University "Higher School of Economics", Moscow, Russian Federation

<sup>2)</sup> Snke OS GmbH, Munich, Germany

<sup>3)</sup> PJSC "Sovcombank", Moscow, Russian Federation

<sup>4)</sup> Sberbank of Russia, Moscow, Russian Federation

**For citation:** Kurochkin, S. V., Pavlov, N. A., Tkachenko, M. V., & Yaremich, E. A. (2024). Machine Learning Methods in Investor Risk Profiling. *AlterEconomics*, 21(3), 527–552.

<https://doi.org/10.31063/AlterEconomics/2024.21-3.6>

**Abstract.** Investment profiling is essential for investors as it differentiates between financial products that align with their risk tolerance and those that are excessively risky. Additionally, investment profiling serves as a tool to prevent misselling. Most financial institutions use risk profiling questionnaires to establish an investor's risk profile. However, the effectiveness of this method is questionable, as actual investor behavior can significantly differ from the responses provided in these questionnaires. This article aims to develop an interactive platform for investment profiling, where data is collected through a web interface. Regression analysis is conducted using Python. On the platform, users engage in a game simulating exchange trading, where they must select a stop loss or take profit in each of ten rounds. This approach allows the investor's risk profile to be determined based on actual user behavior rather than abstract questionnaire responses. The users' actions are then classified using machine learning methods. As a result, it was shown that the constructed model correctly predicts 65.78 % of investors' decisions, whereas the survey correctly identified the risk profile of only 53.7 % of investors. These findings could be used by financial companies for the improvement of their investment profiling process.

**Keywords:** risk profiling, risk tolerance, machine learning, artificial intelligence, questionnaire

## 1. Введение

Определение риск-профиля является важнейшим этапом, предшествующим процессу инвестирования: риск-профиль определяет круг продуктов, риск которых инвестор готов на себя принять. Рост интереса розничных инвесторов к фондовому рынку, усложнение финансовых инструментов, результаты научных исследований в области управления инвестициями, прогресс в законодательной сфере и общей финансовой культуре, — всё это говорит о том, что перед тем, как принять от инвестора денежные средства, необходимо определить его «аппетит к риску» независимо от того, куда именно инвестор желает их вложить: в государственные облигации или же в структурный продукт.

В общем смысле, под риск-профилем понимается небольшой интерпретируемый набор характеристик, на основании которых портфельный управляющий сможет сформировать портфель активов, оптимальный в смысле целей и ограничений данного инвестора. Приведем, с нашими комментариями в скобках, более конкретное определение, данное компанией FinaMetrica<sup>1</sup> — одним из наиболее известных разработчиков инструментария для риск-профилирования: «риск-профилирование — это процесс (не единичный акт) нахождения оптимального уровня инвестиционного риска для вашего клиента, учитывающий одновременно требуемый уровень риска (объем портфеля, которым необходимо рискнуть для достижения целевой доходности), физическую возможность принять риски (размер

<sup>1</sup> FinaMetrica. <https://www.riskprofiling.com> (дата обращения: 29.07.2024).

потерь, которые инвестор физически может себе позволить) и индивидуальную (психологическую) приемлемость риска». Для определения риск-профиля инвестора подавляющее большинство финансовых организаций предлагает ему заполнить анкету.

Можно выделить две связанные с риск-профилированием проблемы, негативным следствием которых стало широкое распространение мисселлинга — продажи инвесторам продуктов, не отвечающих их потребностям и / или с неправильной информацией об их свойствах. Причем такая продажа может быть как умышленной, так и случайной вследствие недостаточной компетентности финансового консультанта.

Во-первых, это граничащее с прямым обманом замалчивание существенных свойств инвестиционного продукта, актуальное для российского рынка из-за низкой финансовой грамотности инвесторов. Данная проблема находится в поле воздействия регулятора, пути ее решения неоднократно комментировались ответственными лицами, однако она едва ли относится к области научного исследования, поэтому в данной работе не затрагивается.

Во-вторых, это нерелевантное профилирование, которое, как правило, связано с форматом анкетирования и может выражаться в том, что:

1) связь между ответом на вопрос и целевой характеристикой, для выявления которой задается этот вопрос, не доказана. Например, практически все анкеты риск-профилирования включают вопрос о возрасте респондента, тогда как лонгитюдное исследование (Van de Venter et al., 2012) показывает, что толерантность к риску, т. е. психологическая готовность принимать его, с возрастом не меняется — измениться может только мощность к риску, т. е. финансовые условия, в которых живет человек;

2) ответ на вопрос предполагает самооценку, однако мнение человека о самом себе субъективно и может расходиться с его действиями: например, в анкете респондент может указать, что для него приемлемо значительное снижение стоимости инвестиционного портфеля, но в реальности испытать панику при потере даже небольшой доли инвестированных средств;

3) ряд вопросов предполагает раскрытие персональных данных — например, уровня дохода и трат. Вследствие некомфортности подобных вопросов ответы могут быть неправдивыми. Помимо этого, присутствие таких вопросов снижает доверие к анкете в целом и ухудшает качество ответов на остальные вопросы.

Сегодняшняя практика риск-профилирования практически полностью представлена анкетированием. В отдельных случаях анкета может включать вопросы с неинтерактивными иллюстрациями: например, клиенту могут предложить выбрать точку на графике, при падении акции до уровня которой он ее продаст. Это улучшает пользовательское восприятие, но не меняет основного принципа: полагаться на анкетирование.

Анкетам как российских, так и зарубежных финансовых компаний присущи два типичных недостатка. Во-первых, в соответствии с требованиями регулятора в анкетах содержится достаточно большое количество вопросов демографического характера, касающихся пола, возраста, семейного положения и т. п. Значимость таких вопросов для конечного результата подвергается серьезному сомнению в большинстве академических исследований — мы подробно осветим этот вопрос в обзоре литературы. Во-вторых, очень малое количество вопросов направлено

на то, чтобы выявить толерантность инвестора к риску, тогда как именно эта характеристика в конечном счете позволяет ответить на вопрос, каким объемом инвестированных средств он готов рискнуть без психологического дискомфорта.

По данным ряда исследований, люди склонны субъективно оценивать склонности своего характера и не могут предсказать, как поведут себя при управлении портфелем активов в ходе реальных торгов. Поэтому представляется научно актуальной и практически важной задача разработки методов и построения платформы для риск-профилирования, которые были бы основаны на иных принципах. Целесообразно отойти от концепции анкетирования и максимально близко воспроизвести условия, в которых инвестору придется принимать решения при реальном управлении портфелем.

В этой статье описан первый опыт разработки и реализации интерактивной платформы для риск-профилирования инвестора. В ходе процесса профилирования инвестору необходимо пройти несколько игровых кейсов, имитирующих торги. В результате сеанса работы в системе возникает набор данных, несущий в несколько раз больше информации, чем возможно получить после классического анкетирования. Затем разработанный алгоритм с применением методов машинного обучения анализирует информацию как по конкретному инвестору, так и по всему пулу. Результатом работы алгоритма является присвоение инвестору того или иного профиля риска.

Платформа, разработанная и представленная авторами статьи — это не готовая программа для риск-профилирования, к которой нечего добавить. Целью работы выступило создание базового инструмента для исследования того, насколько респонденты будут последовательны в своих действиях. В дальнейших исследованиях мы будем дорабатывать платформу, добавляя больше действий в игровую часть и больше вопросов в предшествующую ей анкету, чтобы в конечном счете получить инструмент, готовый к внедрению в финансовых организациях.

Статья состоит из 8 разделов. В разделе 2 проведен обзор академических исследований, а также краткий обзор законодательной базы и применяемых в настоящее время практик риск-профилирования. В разделе 3 описана постановка задачи, требования к платформе и предпосылки, введенные в процессе ее разработки. В разделе 4 описан интерфейс платформы. Данные, на которых проводилось исследование, описаны в разделе 5, применяемая методология — в разделе 6. Результаты исследования представлены в разделе 7. Раздел 8 содержит полученные выводы и вопросы для дальнейшего исследования.

## **2. Обзор литературы, нормативной базы и существующих практик**

### ***2.1. Состояние предметной области в целом***

Как отмечается в статье Дж. Клемента (Klement, 2015), «определение риск-профиля инвестора — важнейшее действие для управления его капиталом». Без знания целей, инвестиционного горизонта, требований по ликвидности и степени неприятия риска инвестором невозможно построить для него подходящую инвестиционную стратегию.

Многие регуляторы финансовых рынков сформировали правила риск-профилирования, в значительной степени защищающие интересы инвесторов. Так, в ЕС вопрос регулируется Статьей 25 директивы MiFID II, в США действует правило 2111

FINRA. В России обязательство профессиональных участников финансового рынка определять риск-профиль клиента впервые было зафиксировано на законодательном уровне Федеральным законом от 20.12.2017 г. N 397-ФЗ (ред. от 17.06.2019) «О внесении изменений в Федеральный закон „О рынке ценных бумаг” и статью 3 Федерального закона „О саморегулируемых организациях в сфере финансового рынка”». В соответствии с ним, риск-профилирование обязательно только в том случае, если клиент хочет воспользоваться услугой инвестиционного консультирования. Закон не требует прохождения риск-профилирования, если клиент желает воспользоваться другой услугой или приобрести какой-либо финансовый инструмент, но на практике финансовые организации просят клиента пройти его, даже если он не собирается использовать услугу инвестиционного консультирования.

Объем информации, которую необходимо запросить у клиента – неквалифицированного инвестора для определения его риск-профиля, регулирует Указание Банка России от 17 декабря 2018 г. N 5014-У «О порядке определения инвестиционного профиля клиента инвестиционного советника, о требованиях к форме предоставления индивидуальной инвестиционной рекомендации и к осуществлению деятельности по инвестиционному консультированию». В соответствии с ним, если клиент является физическим лицом, то он должен сообщить информацию о своем возрасте, образовании, среднемесячных доходах и расходах за срок не менее 1 года, а также о наличии и сумме сбережений. От клиента – юридического лица требуется информация о размере активов и обороте за срок не менее последнего отчетного периода. Если клиент является квалифицированным инвестором, то профессиональный участник финансового рынка имеет право самостоятельно определять вопросы, которые задаст ему.

Применительно к теме нашего исследования представляют непосредственный интерес несколько аспектов риск-профилирования, которые будут рассмотрены в обзоре далее. Это проблемы, связанные с адекватностью анкет, построением и качеством работы робоэдвайзеров, а также применением машинного обучения к риск-профилированию.

## **2.2. Адекватность анкет**

Типичная анкета риск-профилирования включает в себя вопросы про пол, возраст, образование, временной горизонт инвестирования, а также несколько вопросов на гипотетический уровень допустимых потерь. В последние годы все больше исследователей склоняются к мнению, что анкетирование не показывает значимых результатов. Так, Дж. Мурлэнд (Mooreland, 2014) утверждает, что в анкетах люди отвечают не так, как на самом деле ведут себя, а так, как им хотелось бы себя видеть. Кроме того, ответы на вопросы зависят от настроения отвечающего и текущих рыночных условий. Дж. Клемент (Klement, 2015) отмечает, что стандартные современные процедуры риск-профилирования через анкетирование объясняют менее 15 % различий портфелей рискованных активов инвесторов. Согласно Г.Б. Дэвис (Davies, 2017), чем выше у клиента образование и чем ближе его / ее квалификация к экономике или финансам, тем выше его / ее склонность к риску. В исследовании Т. Валентайн (Valentine, 2012) по результатам анализа типовой анкеты риск-профилирования сделан вывод, что такие анкеты могут определить краткосрочную склонность к риску, но не выявляют уровень долгосрочного неприятия риска. Д. Райс (Rice, 2005) проанализировал 131 анкету от разных инвестиционных

компаний и финансовых советников США. Он установил, что 11 % анкет напрямую предлагают инвестору выбрать риск-профиль своего портфеля, т. е. от инвестора требуют, чтобы он сам назвал свой риск-профиль. 35 % анкет не содержали вопросов, касающихся инвестиционного горизонта.

В большинстве анкет содержится вопрос о возрасте инвестора с последующей интерпретацией, что с возрастом человек становится более консервативен как инвестор. Исследования не подтверждают это интуитивно правдоподобное мнение (Van de Venter et al., 2012), хотя утверждают, что толерантность человека к риску зависит от того, пришлось ли ему в молодости пережить трудный в материальном плане период (войны, финансовые кризисы и т. п.).

Многие исследователи стремятся преодолеть ограниченность имеющихся анкет и разрабатывают собственные опросники (De Bortoli et al., 2019; Andreoni et al., 2012; Alsharman, 2019; Alemanni, 2019; Ferreira et al., 2020). Некоторые даже включают в анкету вопросы «биологического» характера (Drichoutis et al., 2015). Поскольку большинство исследований проводится сотрудниками университетов и научных лабораторий, выборка для исследования формируется преимущественно из студентов.

У Дж. Вандербрук и Г. Фортуна (Vandenbroucke & Fortuna, 2019) предпринята попытка на основании существующей анкеты, по результатам которой были риск-профилированы более 1 млн клиентов, выработать более адекватную интерпретацию ее результатов с учетом теории поведенческих финансов. В результате порождается двумерный, а не линейно упорядоченный, как обычно, массив профилей инвесторов с такими интерпретируемыми измерениями, как уровень уверенности (от тревожного до уверенного) и способ действий (от осторожного до импульсивного).

Имеется также большой массив публикаций, где исследуется статистическая значимость влияния отдельных факторов (пол, возраст, образование и т. д.) на риск-профиль безотносительно к тому, были ли получены данные в ходе анкетирования или другим способом. Так, Х. Тэнки и Н. Базер (Thanki & Baser, 2021) показывают, что пол, семейное положение, уровень дохода, место проживания и количество иждивенцев значимо влияют на толерантность к риску. Также они доказали, что чем выше уровень финансовой грамотности инвестора, тем выше его толерантность к риску, хотя значимого влияния формального образования при этом не было обнаружено. Результаты подобных исследований во многом различаются, но в целом большинство современных авторов склоняется к точке зрения, что при оценке толерантности к риску поведенческие аспекты важнее физиологических. Так, например, у В. Хэо и соавторов (Heo et al., 2021) исследовалась склонность людей к риску во время пандемии COVID-19. Авторы заключили, что такой социальный, экономический и экологический шок, как пандемия, снизил склонность к риску у инвесторов. Также они установили, что уровень образования был значимым фактором до пандемии, но перестал быть таковым во время нее, а гораздо более значимыми стали поведенческие факторы — в частности, субъективные финансовые знания, даже не подтвержденные официальным дипломом об образовании.

### **2.3. Рободвайзеры**

В последнее время многие области финансовой индустрии претерпевают серьезные изменения, связанные с автоматизацией процессов, появлением новых

технологий и стремлением заменить человека компьютерными алгоритмами. К портфельному управлению это относится в полной мере: в эту сферу стремительно ворвались робоэдвайзеры — алгоритмы, автоматически формирующие портфель инвестора на основе полученной от него информации. Такие платформы часто показывают результаты, превосходящие результаты деятельности обычных, управляемых людьми фондов и бенчмарков. Так, согласно Р. Тао и соавторам (Тао et al., 2021), за период 2016–2019 гг. в США портфели под управлением робоэдвайзеров показали доходность 18,5 %, тогда как индекс S&P 500 — только 10 %, фонды ценных бумаг — 17 %, фонды денежного рынка — 4,64 %, гибридные фонды — 8,24 %. Коэффициент Шарпа у робоэдвайзеров также был самым высоким — 2,182 против 0,201 у S&P 500, 0,254 — у фондов ценных бумаг, 0,460 — у фондов денежного рынка и 0,094 — у гибридных фондов. У М. Шанмуганатан (Shanmuganathan, 2020) проанализированы четыре ведущих американских робоэдвайзера: Betterment, Wealthfront, FutureAdvisor и Charles Schwab. Согласно выводам статьи, они достаточно эффективны и конкурентоспособны.

Современный робоэдвайзер представляет собой алгоритм, который на выходе выдает структуру портфеля, рекомендуемую данному инвестору, или достаточно конкретные рекомендации по его формированию. С точки зрения целей настоящей статьи робоэдвайзеры заслуживают внимания потому, что этап риск-профилирования в них должен основываться на рациональных, заранее сформулированных моделях, без участия экспертной составляющей на этапе работы с инвестором. Исследования данного аспекта представлены в литературе, но достаточно ограничены. Так, М. Тертильт и П. Шольц (Tertilt & Scholz, 2018) проанализировали 13 робоэдвайзеров (7 немецких и 6 из США и Великобритании), умеющих определять риск-профиль, и пришли к выводу, что их возможность анализировать большие данные не приносит какого-то сверхрезультата, и риск-профилирование робоэдвайзеров не эффективнее, чем риск-профилирование портфельных управляющих. Кроме того, иногда робоэдвайзеры включают в свои анкеты вопросы, вообще не влияющие на оценку риск-профиля: только 6 вопросов из 10 оказали значимое влияние. Также сформированные портфели оказываются слишком консервативными. У С. Тильманс (Tilmans, 2017) проделан анализ аспекта риск-профилирования по большому массиву робоэдвайзеров. Выяснилось, что большинство робоэдвайзеров напрямую задают вопросы, как человек оценивает свою склонность к риску, включают несколько вопросов на эмоциональные склонности инвестора и иногда анализируют текущее состояние его портфеля. Однако этого недостаточно для адекватной оценки риск-профиля, особенно учитывая технический потенциал робоэдвайзеров и их возможность анализировать большие объемы данных. В исследовании М.К.П. Со (So, 2021) 180 вопросов из 20 анкет робоэдвайзеров были проанализированы средствами контент-анализа с целью классификации вопросов и определения их релевантности.

Имеет связь с темой настоящей статьи и представляет интерес публикация Ф. Глейзер и соавторов (Glaser et al., 2019). Это анонс будущей работы, где предполагается спроектировать такого робоэдвайзера, который бы общался с инвестором и адаптировался, наблюдая за его эмоциональными реакциями на со-

бытия в симуляторе и советы робоэдвайзера. Предполагается, что фиксировать реакции будет специальное устройство.

#### 2.4. Применение машинного обучения

По мере доступности массивов данных о финансовых транзакциях в сочетании с персональными данными инвесторов все больше находят применение методы, основанные на данных. Такие методы предполагают, что система прогнозирования, диагностики или принятия решений «самообучается» путем подбора большого числа параметров при оптимизации метрики качества. Среди конкретных приложений здесь безусловным лидером является задача кредитного скоринга. Насколько нам известно, возможности применения машинного обучения для целей риск-профилирования еще недостаточно исследованы: число выявленных публикаций на эту тему невелико. При этом целесообразность использования машинного обучения для анализа поведенческих паттернов в принципе отмечена в публикации Г.Б. Дэвис (Davies, 2017).

В исследовании Д. Мартенса и соавторов (Martens et al., 2016) было предложено применять методы машинного обучения для отслеживания платежей клиентов банка и при помощи полученных данных классифицировать инвесторов по их интересам, уровню трат, скрытым социально-экономическим ограничениям для прогнозирования иных характеристик, например, кредитоспособности.

У П. Ладыжински и соавторов (Ładyżyński et al., 2019) алгоритм «случайный лес» и глубокие нейронные сети применены для того, чтобы, используя данные об исторических транзакциях клиента банка, предсказать вероятность его интереса к новому кредитному продукту.

У А. Ким и соавторов (Kim et al., 2020) с использованием нейросетей решается задача выявления участников рынка, несущих повышенный уровень риска для маркет-мейкеров, с целью определить сделки, которые маркет-мейкеру целесообразно хеджировать.

У Т.Х. Чен и соавторов (Chen et al., 2019) сделана попытка при помощи алгоритмов машинного обучения определить, как личностные качества инвестора влияют на прибыльность инвестиционных стратегий. Применялись методы логистической регрессии, «случайного леса», опорных векторов и другие. Было установлено, что инвесторы с такими личностными качествами, как добросовестность, сговорчивость, экстраверсия и открытость, достигают лучших долгосрочных результатов, чем инвесторы с высоким уровнем невротизма. Главным ограничением своей работы авторы считают необходимость ограничиваться анкетированием.

У Г. Сидху и соавторов (Sidhu et al., 2003) заявлен инструментарий идентификации профиля инвестора из множества predetermined профилей на основе данных о том, как инвестор производил ребалансировку портфеля в ходе симуляции. При этом входными переменными являются две predetermined характеристики инвестора, описывающие его терпимость к риску и инвестиционные цели.

У Дж.Р.Дж. Томпсон (Thompson, 2021) на основе анализа более 50 тыс. счетов 23 тыс. клиентов были сформированы характеристики («features engineering»), позволяющие количественно оценивать поведение инвесторов. Затем с применением методов машинного обучения типа кластеризации инвесторы разбивались на группы сходного поведения. Выяснилось, что информация о клиенте, собран-



ная в результате процедуры KYC («Know Your Customer»), не объясняет поведения инвесторов, в то время как данные о частоте и объеме сделок более информативны.

В исследованиях М.А.С. Брэдбери и соавторов (Bradbury et al., 2015; 2019) ключевой принцип машинного обучения — обучение по данным — был применен не к модели, а к людям. Участникам предлагалось протестировать свои инвестиционные стратегии на выборках из реальных временных рядов активов. Был возможен выбор из линейки продуктов, где крайними позициями служили безрисковый актив и индекс S&P 500, промежуточными — структурные ноты с различным уровнем защиты капитала и коэффициентом участия. В результате прохождения эксперимента большинство участников пересмотрели свои начальные инвестиционные решения.

### **2.5. Обзор существующих практик**

Обзор мировой практики риск-профилирования представлен рассмотрением опыта компании FinaMetrica как общепризнанного лидера в этой области. FinaMetrica разработала инструментальную среду для финансового консультанта, охватывающую различные аспекты его работы по риск-профилированию клиента. Сайт компании<sup>1</sup> содержит большое количество методических материалов, в том числе примеры персонального отчета по клиенту и карту перехода от полученного риск-профиля к рекомендуемой инвестиционной стратегии. Анкета состоит из 25 вопросов, что существенно больше, чем обычно практикуется. Краткое описание вопросов в разбивке на смысловые блоки представлено в таблице 1.

Основная анкета дополнена блоком из 8 вопросов демографического характера. Клиенту предлагается указать свой пол, возраст, образование, собственный доход и доход супруга / супруги, количество иждивенцев и общую стоимость имущества за вычетом долгов. Таким образом, клиент весьма подробно рассматривается в различных измерениях, включая персональные данные. При этом по крайней мере часть вопросов предполагает определенный уровень финансовой грамотности респондента и понимание своей ответственности. Блоку вопросов демографического характера предшествует дисклеймер, что у этих вопросов нет балльной оценки, поэтому ответы на них никак не повлияют на результат тестирования и клиенту необязательно на них отвечать.

По результатам тестирования клиенту присваивается одна из 7 групп риска. Чем больше номер группы, тем выше доля портфеля, которой инвестор готов рискнуть.

Чтобы проиллюстрировать российскую практику риск-профилирования, в таблице 2 мы привели краткое описание анкет, которые используют крупные российские финансовые институты. Мы не будем проводить здесь глубокий анализ каждой из этих анкет на предмет соотношения в них вопросов, «релевантных» и «нерелевантных» главной цели — определению допустимого для респондента уровня риска, поскольку это не является целью нашего исследования. По результатам поверхностного анализа можно заключить, что помимо вопросов, которые обязывает включать в анкету закон, финансовые институты включают в нее вопросы, которые оценивают мощностность к риску, а также вопросы, которые говорят скорее о личных предпочтениях, чем о восприятии риска, например, вопрос о желаемой валюте инвестирования. Вопросы, измеряющие непосредственно толерантность к риску, сведены к минимуму или отсутствуют.

<sup>1</sup> FinaMetrica. <https://www.riskprofiling.com> (дата обращения: 29.07.2024).

Таблица 1

## Описание вопросов из анкеты FinaMetrica

Table 1

## Description of Questions from the FinaMetrica Questionnaire

№ вопроса	Что выявляет вопрос
1–4	Восприятие риска и смысл, который инвестор вкладывает в это понятие
5–10	Финансовое поведение инвестора в условиях риска, его прошлые и предполагаемые будущие финансовые решения
11	Готовность инвестировать заемные средства
12	Степень самоуверенности инвестора
13	Относительную важность прошлого опыта и новой информации при принятии инвестиционных решений
14	Долю всей инвестированной суммы, которую клиент готов потерять вследствие колебаний рынка, не испытывая душевного дискомфорта
15–18, 20–21	Меру выпуклости функции полезности инвестора для последующего конструирования оптимального портфеля
19	Оценку клиента, в какую сторону — риска или надежности — смещаются его инвестиционные предпочтения со временем
22	Отношение к риску изменения законодательства, в т. ч. налогового
23	Отношение к процентному риску текущей стоимости займа: клиенту предлагается сделать выбор между фиксированной и плавающей процентной ставкой
24	Отношение клиента к различным вариантам страхования имущества
25	Клиенту предлагается угадать свое положение на гауссовой кривой, представляющей распределение суммарного балла, набранного за ответы, среди всех инвесторов

Источник: FinaMetrica. <https://www.riskprofiling.com> (дата обращения: 29.07.2024).

Таблица 2

## Описание анкет риск-профилирования российских финансовых институтов

Table 2

## Description of Risk Profiling Questionnaires for Russian Financial Institutions

Компания в обезличенном виде	Краткое описание анкеты риск-профилирования
Компания 1	Анкета состоит из 8 вопросов о возрасте, образовании, среднемесячных доходах и расходах, общей сумме сбережений, финансовых продуктах, которыми пользовался клиент, комфортном соотношении риска и доходности, а также предполагаемой цели инвестирования. По результатам анкеты инвестору присваивается один из пяти риск-профилей: консервативный с ожидаемой доходностью в рублях до 15 % годовых, умеренно-консервативный — 15–20 %, сбалансированный — 20–30 %, умеренно-агрессивный — 25–30 %, агрессивный — свыше 30 %
Компания 2	Анкета состоит из 9 вопросов об опыте инвестирования, ежемесячных доходах и расходах, наличии сбережений и денежных обязательств, цели инвестирования, инвестиционном горизонте, валюте инвестирования, а также возрасте клиента. По результатам клиенту присваивается один из трех риск-профилей: консервативный с допустимым риском до 30 % портфеля, умеренный с риском от 30 % до 60 % портфеля, или агрессивный с возможностью потерять до 100 % инвестиций

Окончание табл. 2 на след. стр.

Окончание табл. 2

Компания в обезличенном виде	Краткое описание анкеты риск-профилирования
Компания 3	Анкета состоит из 13 вопросов об опыте инвестирования и знаниях о финансовом рынке, объеме инвестируемых средств, горизонте и целях инвестирования, валюте инвестирования, действиях при падении стоимости портфеля, предпочтении гарантированного или условного дохода, желании привлечь доверительного управляющего, а также наличии инвестиций в других финансовых организациях
Компания 4	Анкета включает в себя 11 вопросов о валюте инвестирования, возрасте клиента, цели инвестирования и временном горизонте, знаниях и опыте в финансовой сфере, среднемесячных доходах и расходах, наличии сбережений и денежных обязательств. Также у клиента напрямую спрашивается, какое соотношение доходности и риска он считает подходящим для себя. По результатам анкеты клиенту присваивается один из 7 риск-профилей: — консервативный с допустимым риском до 1 % портфеля; — умеренно-консервативный с допустимым риском до 5 % портфеля; — рациональный с допустимым риском до 15 % портфеля; — умеренно-агрессивный с допустимым риском до 25 % портфеля; — агрессивный с допустимым риском до 35 % портфеля; — сверхагрессивный с допустимым риском до 60 % портфеля; — профессиональный с неограниченным допустимым риском
Компания 5	По состоянию на август 2024 г. анкета включает в себя 8 вопросов об инвестиционных целях клиента, ожидаемой доходности инвестиций, знаниях и опыте в финансовой сфере, соотношении доходов и расходов, сумме сбережений и денежных обязательств, а также горизонте инвестирования. По результатам инвестору присваивается один из трех риск-профилей: консервативный, умеренный или агрессивный Примечательно, что предыдущая версия анкеты включала в себя вопросы с использованием интерактивной инфографики. На экране возникал график цены актива, которая пошагово снижалась, и вопрос, что будет делать инвестор на том или ином шаге: продавать, покупать либо держать. Эти вопросы были направлены на определение допустимого риска: авторы стремились дать инвестору возможность ощутить убытки более наглядно. На момент написания этой статьи компания убрала указанные вопросы из анкеты и оставила только 8 вопросов, указанных выше.

Источник: сайты компаний.

### 3. Постановка задачи

Мы сформулировали исчерпывающий набор требований к создаваемой системе, чтобы реализовывать их по мере доступности технических и иных ресурсов. С нашей точки зрения, такими требованиями должны быть:

1) целевая применимость: результаты тестирования потенциального инвестора действительно должны характеризовать его / ее с этой точки зрения и давать финансовому консультанту критерии для формирования портфеля;

2) основанность подхода на данных: система должна использовать современные достижения в области искусственного интеллекта (ИИ), быть свободной от волюнтаристски предопределенных моделей и анкет. В том числе системе не должно быть предписано выдавать риск-профиль только по одной числовой или ординаль-

ной шкале — *a priori* допускаются многомерность. При обучении системы должны использоваться принятые в ИИ методы оценки качества, в т. ч. ошибка на тестовой выборке;

3) анализ действий, а не мнений: входной информацией об инвесторе должно быть не его мнение о себе самом, а его / ее действия в ситуации, приближенной к реальной;

4) индивидуализация: профиль инвестора должен характеризовать именно его / ее. Каждому инвестору соответствует своя модель, «усредненная» модель должна давать результаты хуже, чем индивидуализированные;

5) интерпретируемость: полученные результаты должны переводиться на язык выводов, которые впоследствии будут использоваться при построении оптимального портфеля. Например, что данный инвестор избегает волатильности портфеля даже при отсутствии существенных потерь. Для систем ИИ данное требование не всегда выполнимо, поэтому нужно использовать его по возможности.

Целью данной работы стало создание игровой платформы, которая позволила бы в ходе игры испытать ощущения, подобные испытываемым при реальных торгах на фондовом рынке, но в то же время не перегрузила бы игрока излишними правилами. Игра должна быть простой, интуитивно понятной и показывающей, как человек реагирует на изменение стоимости его инвестиционного портфеля и какие факторы для него важнее при принятии решений.

При разработке платформы были введены следующие предпосылки:

— анкетирование — неэффективный метод риск-профилирования, поскольку человек не способен объективно оценить свои психологические склонности;

— необходима игровая площадка, помещающая инвестора в ситуации, с которыми он потенциально мог бы столкнуться в реальной жизни;

— необходимо полностью отказаться от анализа демографических данных инвесторов, таких, как возраст и пол, а также не задавать вопросов, касающихся персональных данных, например, о ежемесячном доходе. Многие люди предпочитают не раскрывать подобную информацию. Кроме того, такие вопросы могут дать неверный настрой и ухудшить качество и правдивость ответов на другие вопросы;

— вопросов и игровых кейсов не должно быть много, чтобы не утомить инвестора и получить корректные ответы. Инвестор не должен отвечать, не подумав, чтобы побыстрее закончить игру;

— после получения ответов инвестора важно оценить, какие факторы оказались для него наиболее значимыми при принятии решения;

— нет predetermined формата риск-профиля инвестора. Представив по результатам оценки моделей каждого инвестора в виде вектора весов оцениваемых факторов, необходимо выявить структуру совокупности инвесторов в смысле имеющих в ней кластеров, классов или главных компонент и постараться интерпретировать полученные паттерны в данных в терминах риск-профиля.

Полный цикл исследования предполагает два этапа. Первый — сбор и обработка данных для кластеризации или многомерного шкалирования инвесторов по их риск-профилям. Второй — использование полученных кластеров для присвоения риск-профиля новым инвесторам. Таким образом, в терминологии машинного обучения вся задача относится к классу *semi-supervised learning*.

#### 4. Интерфейс системы

В рамках исследования был разработан модуль ввода информации об инвесторе, полученной в ходе прохождения им игры — симуляции торгов, и проведен последующий анализ данных методами машинного обучения. На момент написания статьи результаты доступны только администратору с ролью «аналитик». В будущем персональная диагностика после окончания игры станет доступна любому пользователю.

Модуль ввода информации представляет собой программный комплекс, состоящий из следующих компонент:

1) веб-интерфейс, реализованный на языке JavaScript с помощью фреймворка Angular 10.2.0;

2) модуль аутентификации пользователей (сервис Firebase Authentication<sup>1</sup>) с поддержкой OAuth 2.0, позволяющий использовать как адрес электронной почты, введенный вручную, так и внешние аккаунты, например, аккаунты Google;

3) нереляционная база данных Cloud Firestore, в которой хранятся вопросы, задаваемые пользователю перед игрой, и информация о названии активов для 10 кейсов и динамике их цен, используемая для отображения графиков. Помимо этого, в базе сохраняется информация о пользователе, результаты ответов на вопросы и результаты прохождения кейсов.

По умолчанию пользователи не имеют прав администратора — его можно только назначить. Для администратора в веб-интерфейсе доступен просмотр всех результатов в виде таблицы в формате «пользователь — результаты». Таблица может быть выгружена в виде файла для дальнейшей обработки и анализа.

Исходный код модуля хранится в закрытом репозитории на сайте GitHub<sup>2</sup>, а сам модуль ввода информации размещен на сервере eur3 (europe-west) компании Google посредством сервиса Firebase Hosting<sup>3</sup> и доступен по адресу <https://invest-expert.web.app>. Реализована непрерывная интеграция и доставка (CI/CD) репозитория на GitHub с помощью GitHub Actions.

Веб-интерфейс модуля представляет собой сайт, все содержимое которого доступно только зарегистрированным пользователям. Регистрация бесплатна и не требует ввода персональных данных, а для ее упрощения отсутствует необходимость подтверждать адрес электронной почты. После регистрации пользователю доступна главная страница с приветствием, а также первичный опросник, описание которого приведено далее. Для перехода к игре необходимо дать ответы на три вопроса — изменить их в дальнейшем будет невозможно. После завершения появляется инструкция к прохождению кейсов и кнопка «Перейти к кейсу № 1», по нажатию на которую открывается окно решения кейса. Каждое состояние фиксируется: так, для пользователя, который открыл кейс, но после этого не начал решение и закрыл окно веб-браузера, либо начал, но закрыл окно до завершения показа данных на графике, данный кейс помечается как пропущенный, а при повторном открытии веб-интерфейса пользователю будет предложен следующий кейс при наличии оставшихся нерешенных.

<sup>1</sup> Firebase. <https://firebase.google.com/products/auth>

<sup>2</sup> GitHub. <https://github.com/>

<sup>3</sup> Firebase Hosting. <https://firebase.google.com/products/hosting>

## Вопросы, задаваемые пользователю перед началом игры

Table 3

## Questions Asked to the User before Starting the Game

Вопрос	Варианты ответа	Комментарий
По Вашему мнению, какая имеется связь между будущей ценой финансового актива и ее динамикой в прошлом?	в отсутствие глобальных изменений в экономике поведение цен активов в основном устойчиво; динамика цен носит циклический характер, за падениями следует рост и наоборот; никакой связи нет, будущие изменения определяются новыми событиями	Ответ на данный вопрос позволяет оценить ожидания инвестора от фондового рынка
Предположим, что Вы инвестируете средства в портфель ценных бумаг, торгуемых на фондовом рынке. Как часто Вы будете интересоваться текущей стоимостью Вашего портфеля?	ежедневно; вероятно, раз в неделю; эпизодически, когда вспомню; когда закончится срок инвестирования	Этот вопрос дает понимание того, насколько тревожен инвестор, насколько хорошо он понимает сроки инвестирования, насколько он чувствителен к волатильности цен финансовых активов и готов ли терпеть краткосрочные спады ради выгоды в долгосрочной перспективе
Инвестирование на фондовом рынке может принести убыток. Какую прибыль Вы считаете равноценным возмещением возможного убытка, считая оба события равновероятными?	равную возможному убытку; в полтора раза большую, чем возможный убыток; в два раза большую, чем возможный убыток; я не готов(а) понести убыток	Вопрос вполне традиционен и содержится в большинстве анкет

Источник: составлено авторами.

Анализ результатов и алгоритм кластеризации и присвоения профиля реализован на языке Python. С программным кодом можно ознакомиться на сайте GitHub<sup>1</sup>.

При входе в систему инвестору адресуется приветствие и краткое описание цели платформы. Далее для погружения в тему инвестору предлагается ответить на три вопроса с вариантами ответов (табл. 3). При составлении вопросов преследовалась цель в доступной форме обратить внимание участника на сложный, многомерный характер как рисков, так и возможных целей инвестирования. Кроме того, вопросы были составлены таким образом, чтобы не давать игроку психологических установок перед прохождением игры.

Необходимо отметить, что непосредственная цель данной мини-анкеты — настроить участника на восприятие достаточно сложной информации. В соответствии с общей идеей эксперимента, использовать полученную информацию напрямую для риск-профилирования не предполагалось. Однако, поскольку данные ав-

<sup>1</sup> GitHub. [https://github.com/ShestukhinaElena/VKR\\_risk\\_profiling/blob/main/VKR\\_Shestukhina1.ipynb](https://github.com/ShestukhinaElena/VKR_risk_profiling/blob/main/VKR_Shestukhina1.ipynb)

томатически собирались, они использовались для сопоставления с информацией, полученной в ходе симуляции, и, как описано далее, в целом соответствовали ей.

После прохождения анкеты инвестору предлагается последовательно пройти 10 кейсов. В каждом кейсе пользователь вначале видит отрезок графика исторических цен актива, в который он мог бы инвестировать средства (от  $-10$  до  $0$  по оси абсцисс). Важно отметить, что масштаб графика фиксирован по оси абсцисс, но меняется динамически по оси ординат — это сделано для того, чтобы невозможно было предугадать максимальное или минимальное значение до момента, когда оно появилось на графике. После того, как пользователь нажмет кнопку «Начать построение графика», продолжается построение графика — отображаются значения цены в последующие дни. Время показа одного временного ряда составляет 1 минуту. Если в некоторый момент участник посчитает уровень риска или потерь неприемлемым для себя либо, наоборот, посчитает нужным зафиксировать прибыль и примет решение продать актив, ему нужно нажать кнопку «Закрывать позицию». Предполагается, что пользователь интуитивно понимает, что кейс закончится в тот момент, когда новые появляющиеся точки на графике достигнут его правого края.

При моделировании кейсов использовались только отрезки реальных рядов цен финансовых активов — они будут подробно описаны в разделе «Данные». На графике отображается текущая цена актива в условных единицах. Информацию о том, к какому активу и периоду времени относится данный ряд, инвестор получает сразу после завершения кейса. Также по желанию он может продолжить построение графика уже после закрытия позиции. Это не повлияет на записанный результат, но даст представление о том, как менялась цена актива в реальности.

На этом сеанс работы с пользователем заканчивается и полученные данные собираются в датасет для последующего анализа.

## 5. Данные

В качестве данных пользователю предоставляются отрезки реальных временных рядов цен финансовых активов. Изначально мы предполагали использовать только ряды с общим трендом на понижение, чтобы инвестор закрывал позицию из опасения понести еще больший убыток, а не из желания получить положительную доходность. Однако после ряда экспериментов стало ясно, что такой подход имеет существенный недостаток: игра становится предсказуемой. Пройдя первые 2–3 кейса, игрок начинает закрывать позицию после первого же снижения цены, т. к. уже не ждет ее роста, и рисковать в таких условиях не имеет смысла даже для инвестора с высоким аппетитом к риску. Интерпретировать результат становится сложнее, так как важными оказываются одни и те же факторы, а зависимости имеют один и тот же характер — не столько потому, что люди принимают одни и те же решения, сколько потому, что данные не оставляют им выбора.

В итоге для кейсов мы выбрали ряды цен, разнообразные по трендам и волатильности. Всего было отобрано 10 временных рядов: 9 из них отображали динамику цен акций российских и иностранных эмитентов, один — курс биткоина к доллару США. Каждая точка соответствует одному торговому дню. Выходные и праздничные дни, а также дни, в которые не велись торги, были исключены для большего приближения кейсов к реальности. В таблице 4 представлены использованные ак-

тивы, период наблюдения, волатильность, среднедневная доходность и абсолютная доходность за период наблюдения, страна эмитента и отрасль деятельности.

Результаты действий каждого игрока заносились в таблицу в виде строки из 10 чисел, где для каждого кейса указывался номер шага, соответствующего значению точки по оси абсцисс на графике, в которой была закрыта позиция. Также запоминались ответы на предварительные 3 вопроса.

Следующим этапом было формирование входных переменных («*feature engineering*»). Экспертным путем были выбраны 5 показателей (табл. 5).

Датасет для машинного обучения формировался следующим образом.

1. Для каждого участника, каждого кейса и каждого момента времени формировалась строка входных переменных, как описано выше.

2. Значение выходной переменной для этой строки полагалось равным 1, если инвестор закрывал позицию, и 0, если не предпринимал никаких действий.

3. Выборка делалась сбалансированной для каждого пользователя и кейса путем размножения строк со значением выходной переменной, равным 1.

В эксперименте участвовало 54 чел. Таким образом, в итоге получился массив, состоящий из  $54 \cdot 10 \cdot [\text{среднее время удержания позиции в секундах} \cdot 2] \cdot (5 + 1) \approx 2 \cdot 10^5$  степени чисел.

Таблица 4

## Данные для игровых кейсов

Table 4

## Data for Game Cases

Эмитент акции	Отрасль	Страна	Период	Средняя доходность	Абсолютная доходность	Волатильность
Якутскэнерго	Энергоснабжение	Россия	03.09.20 – 10.12.20	-0,290	-0,112	0,021
Московская Биржа	Биржа		30.07.20 – 05.11.20	-0,064	-0,063	0,010
ГМК «Норильский Никель»	Добыча природных ресурсов		19.05.20 – 27.08.20	0,444	-0,107	0,020
Газпром	Нефтегазовая компания		18.12.20 – 31.03.20	-0,781	0,054	0,015
Polymetal	Горнорудная компания		30.09.20 – 04.01.21	0,093	0,026	0,030
Kroger	Ритейл	США	09.11.20 – 18.02.21	0,034	0,033	0,021
BIT / USD	Курс биткоина к доллару США	Канада	14.11.20 – 22.01.21	0,911	-0,585	0,124
Lowblaw	Ритейл		04.12.20 – 17.03.21	-0,053	0,031	0,011
Фармсинтез	Фармацевтика	Россия	02.12.20 – 15.03.21	-0,357	-0,062	0,013
ОГК-2	Энергоснабжение		08.10.20 – 19.01.21	0,110	0,069	0,016

Источник: составлено авторами.



Таблица 5

## Описание входных переменных

Table 5

## Description of Input Variables

Показатель	Название переменной	Комментарий
Доходность по сравнению с началом игры	return_first	Рассчитывается как отношение текущей цены к цене в точке, в которой началась игра. Показывает, насколько для инвестора важна долгосрочная доходность. Рискованный инвестор будет пережидать падение цен в надежде на более высокую доходность. Напротив, консервативный инвестор не склонен брать на себя такой риск, поэтому с высокой вероятностью закроет позицию, если цена сильно упадет
Дневная доходность	return_last	Рассчитывается как отношение текущей цены к цене в предыдущей точке. Если коэффициент влияния этого показателя значимо положительный, значит инвестор отслеживает изменения на рынке, старается выйти с возможной выгодой для себя. Если значимо отрицательный, то инвестор более консервативный и боится сильного падения цены
Волатильность	volatility	В данном случае рассчитывается не как стандартное отклонение, а как кумулятивная сумма модулей дневных доходностей, деленная на количество прошедших дней. Считается таким образом из предположения, что инвестору сложно визуально оценить стандартное отклонение и он определяет изменчивость доходности интуитивно. Склонный к риску инвестор слабо реагирует на рост волатильности. Консервативный инвестор при росте волатильности скорее будет нервничать и закрывать позицию
Количество дней, когда цена падала, в сравнении с предыдущим значением	number_of_days_when_return_last<1	Рассчитывается как количество дней, в течение которых дневная доходность была отрицательной. Если коэффициент при этом показателе значим и отрицателен, то инвестор консервативен
Количество дней, когда цена актива была ниже, чем текущая	number_of_days_when_price<now	Консервативному инвестору будет некомфортно, если цена пробьет исторический минимум. Если коэффициент при данном показателе значим и отрицателен, то инвестор не склонен к риску

Источник: составлено авторами.

## 6. Методология анализа

Целью анализа было моделирование внутренней структуры риск-профилей инвесторов с последующей разработкой модели классификации инвесторов по риск-профилю. Была принята следующая схема анализа.

1. Все значения признаков из сформированного датасета были стандартизированы на нулевое среднее и единичное стандартное отклонение.

2. Для каждого набора данных, соответствующего отдельному инвестору, решается задача классификации, для чего применяется логистическая регрессия. Входные переменные описаны в таблице 3. Выходная переменная принимает значение 1, если инвестор закрывал позицию, и 0, если не предпринимал никаких действий.

3. Для сравнения была построена единая модель логистической регрессии для всех решений всех инвесторов без разделения на подвыборки для каждого игрока.

4. Полученные на шаге 2 веса отдельных логистических моделей рассматриваются как индивидуальные признаки инвестора, на основании которых можно узнать его риск-профиль. Таким образом, дальнейшие действия связаны с построением модели второго порядка.

5. Для снижения размерности данных и лучшей интерпретируемости целесообразно применить метод главных компонент.

6. Наблюдения, т. е. участники эксперимента, записанные каждый через свою логистическую модель, подвергаются кластеризации на предмет определения внутренней структуры данных.

7. Для предъявления алгоритма присвоения риск-профиля новым клиентам строится модель классификации.

Подробности реализации и результаты представлены в следующем разделе.

## 7. Результаты

На этапе построения модели для каждого отдельного инвестора каждая подвыборка была разбита в соотношении 80:20 на обучающую и тестовую. Были посчитаны коэффициенты моделей и ошибки. В среднем на тестовых подвыборках модель верно определяла 65,78 % решений инвесторов. У 4 испытуемых из 54 точность оказалось экстремально низкой — менее 25 %, что, вероятно, объясняется недобросовестным прохождением игры или непониманием поставленных задач и инструкции. У 8 испытуемых точность превысила 90 %. Полная разбивка выборки по доле пользователей, чьи ответы имели ту или иную точность, показана на рисунке 1.

Построенная для сравнения единая модель без разделения на подвыборки для каждого игрока, показала точность 58,56 %, что ниже среднего показателя по персональным моделям.

Построенная для снижения размерности датасета модель главных компонент со значениями 5 признаков в качестве входных переменных дала результаты, представленные на рисунке 2.

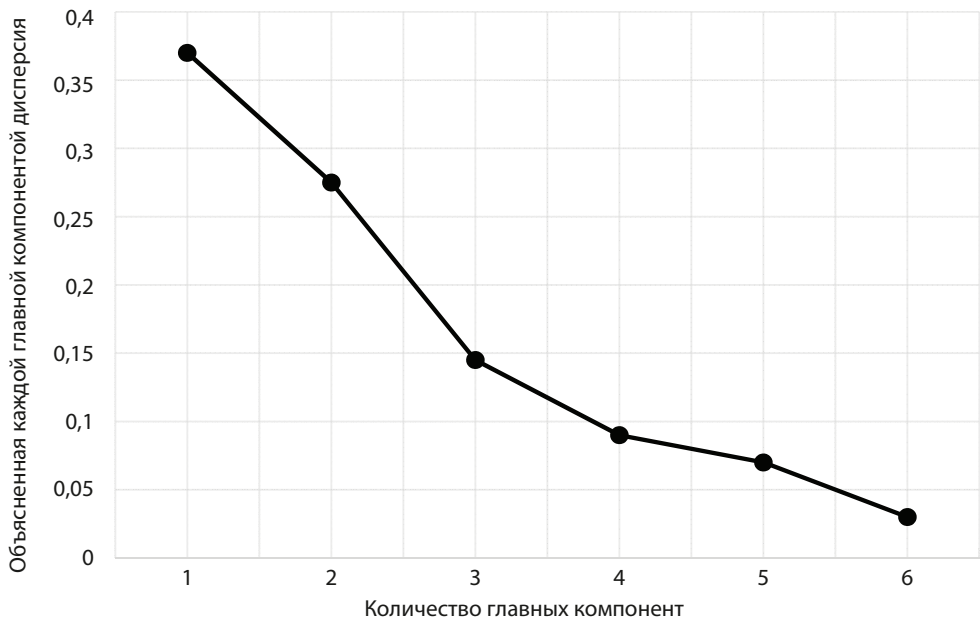
Скорость уменьшения доли объясненной дисперсии снижается в точке, соответствующей трем компонентам, поэтому именно это количество компонент было признано оптимальным. В таблице 6 представлены нагрузки факторов.

В определенной степени компоненты допускают интерпретацию. Отметим, что первичный признак *Intercept* говорит об общем желании выйти из игры. Чем больше его значение, тем выше желание инвестора продать актив. Наоборот, агрессивный инвестор, готовый терпеть последовательное падение цены актива и высокую волатильность, с большой вероятностью будет иметь отрицательный *Intercept*. Анализируя таблицу, можно предположить, что первая главная компонента отвечает за агрессивность поведения, поскольку она обладает:



**Рис. 1.** Распределение выборки по точности модели на тестовых данных  
(источник: расчеты авторов)

**Fig. 1.** Distribution of Sample by Model Accuracy on Test Data



**Рис. 2.** График каменной осыпи для индивидуальных моделей логистической регрессии  
(источник: расчеты авторов)

**Fig. 2.** Screen Plot for Individual Logistic Regression Models

Таблица 6

**Нагрузки компонент массива коэффициентов индивидуальных регрессионных моделей**

Table 6

**Loadings of the Components of the Array of Coefficients of Individual Regression Models**

	PC1	PC2	PC3
Intercept	-0,492293	-0,138510	-0,285159
Доходность	0,535716	0,699149	0,067176
Дневная доходность	-0,038051	-0,111570	0,541924
Количество дней, когда цена падала	-0,470447	0,454450	0,176748
Количество дней, когда цена была ниже текущей	0,476603	-0,381316	-0,357566
Волатильность	-0,144007	0,357255	-0,679262

Источник: составлено авторами.

- высокой положительной нагрузкой на коэффициент влияния доходности — можно предположить, что такие инвесторы гонятся за высокой прибылью;
- высокой отрицательной нагрузкой на коэффициент *Intercept* — такие инвесторы не стремятся выйти из игры;
- высокой положительной нагрузкой на коэффициент влияния количества дней, когда цена была ниже текущей — такие инвесторы не боятся исторических минимумов;
- высокой отрицательной нагрузкой на коэффициент влияния количества дней, когда цена падала — такие инвесторы не переживают из-за падения цен;
- отрицательной нагрузкой при коэффициенте влияния волатильности — такие инвесторы не боятся волатильности.

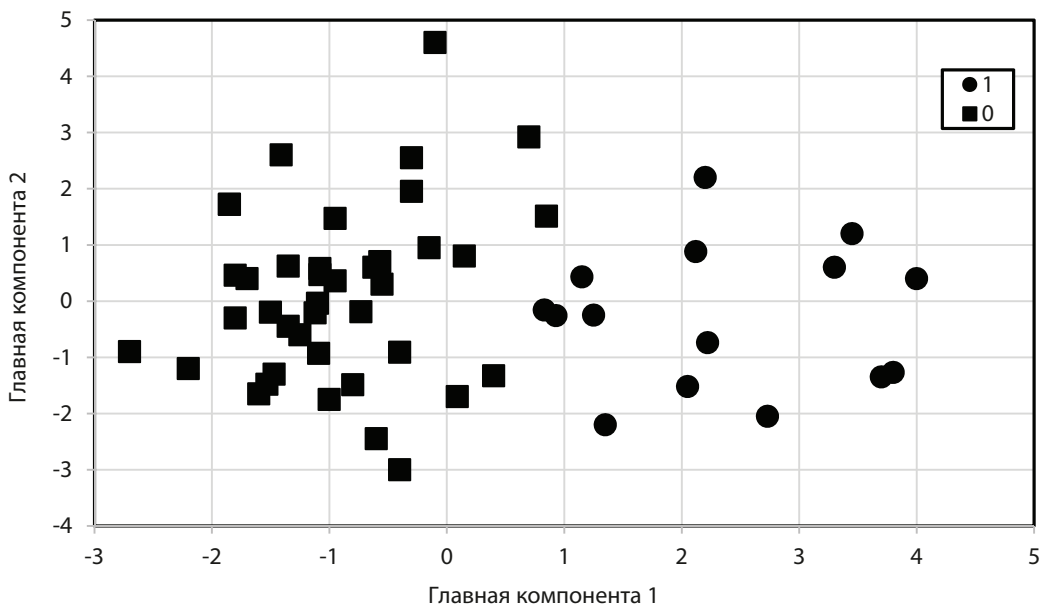
Вторая главная компонента менее интерпретируемая, но можно предположить, что она отвечает за то, что инвестору более важно: изменение доходности или изменение волатильности.

Кластеризация участников эксперимента проводилась различными методами: *k*-средних, иерархической кластеризацией и DBSCAN. Результаты оказались достаточно схожими. На рисунке 3 представлена диаграмма для метода иерархической кластеризации.

Отметим, что:

- 1) задача относится к классу «без учителя», поэтому методы классификации применить невозможно;
- 2) расположение кластеров соответствует значению 1-й компоненты, но общая картина не одномерная.

Для построения модели классификации выборка с 5 показателями и номером кластера в качестве результирующей переменной была разбита в отношении 9:1 на обучающую и тестовую.



**Рис. 3.** Иерархическая кластеризация инвесторов (источник: расчеты авторов)

**Fig. 3.** Hierarchical Clustering of Investors

Были протестированы методы опорных векторов,  $k$ -ближайших соседей, дерева решений и случайного леса. Точность метода опорных векторов составила 1, метода  $k$ -ближайших соседей — 1, метода дерева решений — 0,833333, метода случайного леса — 1. Таким образом, все модели показали хорошую точность, три из них — максимальную. Для классификации новых клиентов был выбран метод опорных векторов.

Затем было протестировано три новых клиента из числа испытуемых, прошедших игру позже остальных. Их ответы и результаты представлены в таблице 7.

Результат игры совпал с результатом предварительного тестирования. Также на этом примере можно продемонстрировать интерпретируемость всей разработанной модели:

Клиент № 1:

— закрывает позицию, когда доходность значительно падает по сравнению с началом игры;

— больше других склонен выходить из игры;

— боится, когда цена часто падает;

— боится волатильности.

Такое поведение явно говорит о принадлежности к консервативному типу.

Клиент № 2:

— закрывает позицию, когда доходность падает по сравнению с началом игры;

— больше других склонен выходить из игры;

— боится, когда цена часто падает;

— закрывает позицию при достижении ценой исторических минимумов.

Такое поведение говорит о принадлежности к консервативному типу.

Клиент № 3:

— закрывает позицию, когда получает прибыль;

— не боится исторических минимумов и падения цен;

— не боится волатильности.

Такое поведение сигнализирует о принадлежности к агрессивному типу.

Таблица 7

**Коэффициенты логистических моделей и результаты классификации для новых клиентов**

Table 7

**Logistic Model Coefficients and Classification Results for New Customers**

Клиент	1	2	3
Intercept	-0,23937	-0,21952	-3,89099
Return_first	-1,25436	-0,51368	3,76673
Return_last	0,48410	-0,20659	-0,15022
number_of_days_ when_return_last < 1	1,04934	0,75503	-1,03466
number_of_days_ when_price < now	0,60541	-0,14394	2,480207
volatility	0,60832	-0,08883	-0,22807
Игра	Консервативный	Консервативный	Агрессивный
Ответ	Консервативный	Консервативный	Агрессивный

Источник: составлено авторами.

Также можно сравнить полученные результаты с профилями, полученными по ответам на вопросы. Последние присваивались традиционным способом — присвоением баллов за ответы и суммированием. Сравнение показало, что ответы на вопросы адекватно выявляют отношение к риск-профилю только у 53,7 % инвесторов, что в очередной раз ставит под сомнение надежность анкет.

## 8. Заключение

В данной работе реализован подход к риск-профилированию инвестора, основанный на данных (*data driven*) и методах машинного обучения. По результатам проведенного исследования мы сделали три ключевых вывода. Во-первых, риск-профиль инвестора статистически значимо выявляется в его действиях при симуляции процесса инвестирования. Во-вторых, риск-профиль инвестора лишь приближенно может быть выражен одним числовым показателем и скорее является многомерной характеристикой. В-третьих, риск-профиль инвестора, определенный из внутренней структуры данных при помощи методов машинного обучения, может быть полезен и конструктивно использоваться при формировании портфеля инвестора.

Однако по результатам исследования многие вопросы остаются открытыми и требуют дальнейшего анализа. Так, нам пока не удалось понять, как оптимально сформировать набор игровых ситуаций, поведение инвестора в которых наиболее достоверно свидетельствовало бы о его / ее действиях и эмоциях в реальном инвестиционном процессе. Также пока нет однозначного ответа на вопрос, как оптимально сформировать набор характеристик инвестора, которые могли бы быть получены по результатам анкетирования или / и симуляции, после чего могли бы быть переведены в общепринятые в инвестиционной отрасли целевые характеристики портфеля.

Мы будем продолжать исследования в данном направлении, в т. ч. с использованием реальных баз клиентских данных в обезличенном виде.

## Список источников

- Alemanni, B., Uberti, P. (2019). What are investors afraid of? Finding the big bad wolf. *International Journal of Financial Studies*, 7(3), 42. <https://doi.org/10.3390/ijfs7030042>
- Alsharman, M. (2019). *The development of a market risk profiling system employing behavioural and emotional finance approaches* (Doctoral dissertation, University of Bath).
- Andreoni, J., Sprenger, C. (2012). Risk preferences are not time preferences. *American Economic Review*, 102(7), 3357–3376. <http://dx.doi.org/10.1257/aer.102.7.3357>
- Bradbury, M. A., Hens, T., & Zeisberger, S. (2019). How persistent are the effects of experience sampling on investor behavior? *Journal of Banking & Finance*, 98, 61–79. <https://doi.org/10.1016/j.jbankfin.2018.10.014>
- Bradbury, M. A. S., Hens, T., Zeisberger, S. (2015). Improving investment decisions with simulated experience. *Review of Finance*, 19(3), 1019–1052. <https://doi.org/10.1093/rof/rfu021>
- Chen, T. H., Ho, R. J., Liu, Y. W. (2019). Investor personality predicts investment performance? A statistics and machine learning model investigation. *Computers in Human Behavior*, 101, 409–416. <https://doi.org/10.1016/j.chb.2018.09.027>
- Davies, G. B. (2017). *New vistas in risk profiling*. CFA Institute Research Foundation.
- Davies, G. B., Brooks, P. (2014). Risk tolerance: Essential, behavioural and misunderstood. *Journal of Risk Management in Financial Institutions*, 7(2), 110–113. <https://doi.org/10.69554/OVZT4879>

De Bortoli, D., da Costa Jr, N., Goulart, M., Campara, J. (2019). Personality traits and investor profile analysis: A behavioral finance study. *PLoS one*, 14(3), e0214062. <https://doi.org/10.1371/journal.pone.0214062>

Drichoutis, A. C., Nayga Jr, R. M. (2015). Do risk and time preferences have biological roots? *Southern Economic Journal*, 82(1), 235–256. <https://doi.org/10.4284/0038-4038-2013.246>

Ferreira, S., Dickason-Koekemoer, Z., McMillan, D. (2020). A structural equation model of financial risk tolerance in South Africa. *Cogent Business & Management*, 7(1). <https://doi.org/10.1080/23311975.2020.1811595>

Glaser, F., Iliewa, Z., Jung, D., Weber, M. (2019). Towards designing robo-advisors for unexperienced investors with experience sampling of time-series data. *Information Systems and Neuroscience: NeuroIS Retreat 2018* (pp. 133–138). Springer International Publishing.

Heo, W., Rabbani, A., Grable, J. E. (2021). An evaluation of the effect of the COVID-19 pandemic on the risk tolerance of financial decision makers. *Finance Research Letters*, 41, 101842. <https://doi.org/10.1016/j.frl.2020.101842>

Kim, A., Yang, Y., Lessmann, S., Ma, T., Sung, M. C., Johnson, J. E. (2020). Can deep learning predict risky retail investors? A case study in financial risk behavior forecasting. *European Journal of Operational Research*, 283(1), 217–234. <https://doi.org/10.1016/j.ejor.2019.11.007>

Klement, J. (2015). *Investor risk profiling: An overview*.

Ładyżyński, P., Żbikowski, K., Gawrysiak, P. (2019). Direct marketing campaigns in retail banking with the use of deep learning and random forests. *Expert Systems with Applications*, 134, 28–35. <https://doi.org/10.1016/j.eswa.2019.05.020>

Martens, D., Provost, F., Clark, J., de Fortuny, E. J. (2016). Mining massive fine-grained behavior data to improve predictive analytics. *MIS quarterly*, 40(4), 869–888. <http://dx.doi.org/10.25300/MISQ/2016/40.4.04>

Mooreland, J. (2014). Obtain a More Accurate Assessment of Your Clients' Risk Profile. *Journal of Financial Planning*, 27(7).

Rice, D. (2005). *Variance in risk tolerance measurement: Toward a uniform solution*. Golden Gate University.

Shanmuganathan, M. (2020). Behavioural finance in an era of artificial intelligence: Longitudinal case study of robo-advisors in investment decisions. *Journal of Behavioral and Experimental Finance*, 27, 100297. <http://dx.doi.org/10.1016/j.jbef.2020.100297>

Sidhu, G., Min, C. (2003). *Method and apparatus for identifying investor profile* (U.S. Patent Application No. 10/128,910). <https://patentimages.storage.googleapis.com/pdfs/US20030055758.pdf> (дата обращения: 29.07.2024).

So, M. K. (2021). Robo-advising risk profiling through content analysis for sustainable development in the Hong Kong financial market. *Sustainability*, 13(3), 1306. <https://doi.org/10.3390/su13031306>

Tao, R., Su, C. W., Xiao, Y., Dai, K., Khalid, F. (2021). Robo advisors, algorithmic trading and investment management: Wonders of fourth industrial revolution in financial markets. *Technological Forecasting and Social Change*, 163, 120421. <https://doi.org/10.1016/j.techfore.2020.120421>

Tertilt, M., Scholz, P. (2018). To advise, or not to advise — how robo-advisors evaluate the risk preferences of private investors. *The Journal of Wealth Management*, 21(2), 70–84. <http://dx.doi.org/10.3905/jwm.2018.21.2.070>

Thanki, H., Baser, N. (2021). Determinants of financial risk tolerance (FRT): An empirical investigation. *The Journal of Wealth Management*, 24(2), 48–64. <http://dx.doi.org/10.3905/jwm.2021.1.144>

Thompson, J. R., Feng, L., Reesor, R. M., Grace, C. (2021). Know Your Clients' behaviours: a cluster analysis of financial transactions. *Journal of Risk and Financial Management*, 14(2), 50. <https://doi.org/10.3390/jrfm14020050>

Tillmans, S. (2017). Robo-advisors: how can automated investment advice change risk profiling practices. <http://theses.ubn.ru.nl/handle/123456789/5205> (дата обращения: 29.07.2024).

Valentine, T. (2012). The usefulness of risk profile questionnaires in financial advising. *JASSA: The Journal of the Securities Institute of Australia*, (3), 54–58.

- Van de Venter, G., Michayluk, D., Davey, G. (2012). A longitudinal study of financial risk tolerance. *Journal of Economic Psychology*, 33(4), 794–800. <http://dx.doi.org/10.1016/j.joep.2012.03.001>
- Vandenbroucke, J., Fortuna, G. (2019). Loss Aversion Implied by a Risk-Based Questionnaire. *The Journal of Wealth Management*, 22(1), 39–48. <http://dx.doi.org/10.3905/jwm.2019.22.1.039>

## References

- Alemanni, B., & Uberti, P. (2019). What are investors afraid of? Finding the big bad wolf. *International Journal of Financial Studies*, 7(3), 42. <https://doi.org/10.3390/ijfs7030042>
- Alsharman, M. (2019). *The development of a market risk profiling system employing behavioural and emotional finance approaches* (Doctoral dissertation, University of Bath).
- Andreoni, J., & Sprenger, C. (2012). Risk preferences are not time preferences. *American Economic Review*, 102(7), 3357–3376. <http://dx.doi.org/10.1257/aer.102.7.3357>
- Bradbury, M. A., Hens, T., & Zeisberger, S. (2019). How persistent are the effects of experience sampling on investor behavior? *Journal of Banking & Finance*, 98, 61–79. <https://doi.org/10.1016/j.jbankfin.2018.10.014>
- Bradbury, M. A. S., Hens, T., & Zeisberger, S. (2015). Improving investment decisions with simulated experience. *Review of Finance*, 19(3), 1019–1052. <https://doi.org/10.1093/rof/rfu021>
- Chen, T. H., Ho, R. J., & Liu, Y. W. (2019). Investor personality predicts investment performance? A statistics and machine learning model investigation. *Computers in Human Behavior*, 101, 409–416. <https://doi.org/10.1016/j.chb.2018.09.027>
- Davies, G. B. (2017). *New vistas in risk profiling*. CFA Institute Research Foundation.
- Davies, G. B., & Brooks, P. (2014). Risk tolerance: Essential, behavioural and misunderstood. *Journal of Risk Management in Financial Institutions*, 7(2), 110–113. <https://doi.org/10.69554/OVZT4879>
- De Bortoli, D., da Costa Jr, N., Goulart, M., & Campara, J. (2019). Personality traits and investor profile analysis: A behavioral finance study. *PLoS one*, 14(3), e0214062. <https://doi.org/10.1371/journal.pone.0214062>
- Drichoutis, A. C., & Nayga Jr, R. M. (2015). Do risk and time preferences have biological roots? *Southern Economic Journal*, 82(1), 235–256. <https://doi.org/10.4284/0038-4038-2013.246>
- Ferreira, S., Dickason-Koekemoer, Z., & McMillan, D. (2020). A structural equation model of financial risk tolerance in South Africa. *Cogent Business & Management*, 7(1). <https://doi.org/10.1080/23311975.2020.1811595>
- Glaser, F., Iliewa, Z., Jung, D., & Weber, M. (2019). Towards designing robo-advisors for unexperienced investors with experience sampling of time-series data. *Information Systems and Neuroscience: NeuroIS Retreat 2018* (pp. 133–138). Springer International Publishing.
- Heo, W., Rabbani, A., & Grable, J. E. (2021). An evaluation of the effect of the COVID-19 pandemic on the risk tolerance of financial decision makers. *Finance Research Letters*, 41, 101842. <https://doi.org/10.1016/j.frl.2020.101842>
- Kim, A., Yang, Y., Lessmann, S., Ma, T., Sung, M. C., & Johnson, J. E. (2020). Can deep learning predict risky retail investors? A case study in financial risk behavior forecasting. *European Journal of Operational Research*, 283(1), 217–234. <https://doi.org/10.1016/j.ejor.2019.11.007>
- Klement, J. (2015). *Investor risk profiling: An overview*.
- Ładyżyński, P., Żbikowski, K., & Gawrysiak, P. (2019). Direct marketing campaigns in retail banking with the use of deep learning and random forests. *Expert Systems with Applications*, 134, 28–35. <https://doi.org/10.1016/j.eswa.2019.05.020>
- Martens, D., Provost, F., Clark, J., & de Fortuny, E. J. (2016). Mining massive fine-grained behavior data to improve predictive analytics. *MIS quarterly*, 40(4), 869–888. <http://dx.doi.org/10.25300/MISQ/2016/40.4.04>
- Mooreland, J. (2014). Obtain a More Accurate Assessment of Your Clients' Risk Profile. *Journal of Financial Planning*, 27(7).
- Rice, D. (2005). *Variance in risk tolerance measurement: Toward a uniform solution*. Golden Gate University.



Shanmuganathan, M. (2020). Behavioural finance in an era of artificial intelligence: Longitudinal case study of robo-advisors in investment decisions. *Journal of Behavioral and Experimental Finance*, 27, 100297. <http://dx.doi.org/10.1016/j.jbef.2020.100297>

Sidhu, G., & Min, C. (2003). *Method and apparatus for identifying investor profile* (U.S. Patent Application No. 10/128,910). <https://patentimages.storage.googleapis.com/pdfs/US20030055758.pdf> (Date of access: 29.07.2024)

So, M. K. (2021). Robo-advising risk profiling through content analysis for sustainable development in the Hong Kong financial market. *Sustainability*, 13(3), 1306. <https://doi.org/10.3390/su13031306>

Tao, R., Su, C. W., Xiao, Y., Dai, K., & Khalid, F. (2021). Robo advisors, algorithmic trading and investment management: Wonders of fourth industrial revolution in financial markets. *Technological Forecasting and Social Change*, 163, 120421. <https://doi.org/10.1016/j.techfore.2020.120421>

Tertilt, M., & Scholz, P. (2018). To advise, or not to advise — how robo-advisors evaluate the risk preferences of private investors. *The Journal of Wealth Management*, 21(2), 70–84. <http://dx.doi.org/10.3905/jwm.2018.21.2.070>

Thanki, H., & Baser, N. (2021). Determinants of financial risk tolerance (FRT): An empirical investigation. *The Journal of Wealth Management*, 24(2), 48–64. <http://dx.doi.org/10.3905/jwm.2021.1.144>

Thompson, J. R., Feng, L., Reesor, R. M., & Grace, C. (2021). Know Your Clients' behaviours: a cluster analysis of financial transactions. *Journal of Risk and Financial Management*, 14(2), 50. <https://doi.org/10.3390/jrfm14020050>

Tillmans, S. (2017). Robo-advisors: how can automated investment advice change risk profiling practices. <http://theses.ubn.ru.nl/handle/123456789/5205> (Date of access: 29.07.2024)

Valentine, T. (2012). The usefulness of risk profile questionnaires in financial advising. *JASSA: The Journal of the Securities Institute of Australia*, (3), 54–58.

Van de Venter, G., Michayluk, D., & Davey, G. (2012). A longitudinal study of financial risk tolerance. *Journal of Economic Psychology*, 33(4), 794–800. <http://dx.doi.org/10.1016/j.joep.2012.03.001>

Vandenbroucke, J., & Fortuna, G. (2019). Loss Aversion Implied by a Risk-Based Questionnaire. *The Journal of Wealth Management*, 22(1), 39–48. <http://dx.doi.org/10.3905/jwm.2019.22.1.039>

### Информация об авторах

**Курочкин Сергей Владимирович** — кандидат физико-математических наук, доцент, Национальный исследовательский университет «Высшая школа экономики»; <https://orcid.org/0000-0001-9484-6012> (Российская Федерация, 109028, г. Москва, Покровский бульвар, д. 11; e-mail: skurochkin@hse.ru).

**Павлов Николай Александрович** — главный проектный инженер, Snke OS GmbH (Германия, г. Мюнхен; e-mail: nikolay.pavlov@snkeos.com).

**Ткаченко Мария Владимировна** — старший преподаватель, Национальный исследовательский университет «Высшая школа экономики»; начальник отдела структурных продуктов, ПАО «Совкомбанк»; <https://orcid.org/0000-0003-4841-0263> (Российская Федерация, г. Москва; e-mail: mtkachenko@hse.ru).

**Яремич Елена Андреевна** — ведущий инженер по разработке, ПАО Сбербанк (Российская Федерация; г. Москва; e-mail: EAnShestukhina@sberbank.ru).

### About the authors

**Sergey V. Kurochkin** — Cand. Sci. (Physics and Mathematics), Associate Professor, National Research University “Higher School of Economics”; <https://orcid.org/0000-0001-9484-6012> (11, Pokrovsky Boulevard, Moscow, 109028, Russian Federation; e-mail: skurochkin@hse.ru).

**Nikolay A. Pavlov** — Chief Design Engineer, Snke OS GmbH (Munich, Germany; e-mail: nikolay.pavlov@snkeos.com).

**Mariia V. Tkachenko** — Senior Lecturer, National Research University “Higher School of Economics”; Head of Structured Products Department, PJSC “Sovcombank”; <https://orcid.org/0000-0003-4841-0263> (Moscow, Russian Federation; e-mail: mtkachenko@hse.ru).

**Elena A. Yaremich** — Lead Development Engineer, Sberbank of Russia (Moscow, Russian Federation; e-mail: EAnShestukhina@sberbank.ru).

*Дата поступления рукописи: 29.07.2024.*

*Прошла рецензирование: 26.08.2024.*

*Принято решение о публикации: 14.09.2024.*

*Received: 29 Jul 2024.*

*Reviewed: 26 Aug 2024.*

*Accepted: 14 Sep 2024.*